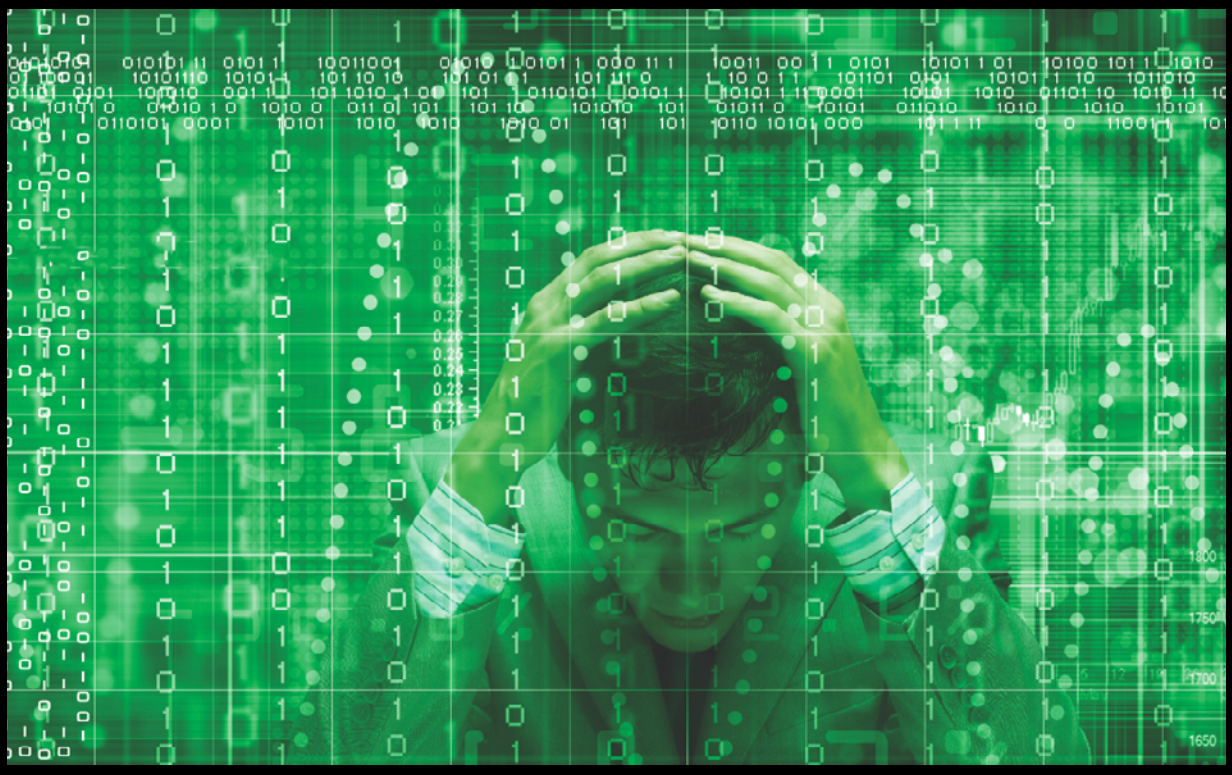
idm.
information & data manager
JUNE-JULY 2024

Building a privacy layer for AI



ICT Commitment-Creep & Organisational Schizophrenia

Knowledge Graphs Redefining Data Management for the Modern Enterprise

Managing information in the GenAI era

Mature IM Strategies are Critical to Realising AI Benefits: Report

Blind Spots in AI Overconfidence: Report

Nearly half (44%) of IT leaders surveyed believe their organizations are fully set up to realize the benefits of AI, according to a global research report commissioned by Hewlett Packard Enterprise. However, the report reveals critical gaps in their strategies, such as lack of alignment between processes and metrics, resulting in consequential fragmentation in approach, which will further exacerbate delivery issues.

The report, 'Architect an AI Advantage', which surveyed more than 2,000 IT leaders from 14 countries, found that while global commitment to AI shows growing investments, businesses are overlooking key areas that will have a bearing on their ability to deliver successful AI outcomes – including low data maturity levels, possible deficiencies in their networking and compute provisioning, and vital ethics and compliance considerations. The report also uncovered significant disconnects in both strategy and understanding that could adversely affect future return on investment (ROI).

Sylvia Hooks, VP, HPE Aruba Networking, said, "These findings clearly demonstrate the appetite for AI, but they also highlight very real blind spots that could see progress stagnate if a more holistic approach is not followed. Misalignment on strategy and department involvement – for example – can impede organizations from leveraging critical areas of expertise, making effective and efficient decisions, and ensuring a holistic AI roadmap benefits all areas of the business congruently."

Ricoh adds capture software firm natif.ai

Ricoh has announced the acquisition of natif.ai, a German software startup, offering artificial intelligence (AI)-enabled Intelligent Capture, advanced image recognition and optical character recognition (OCR) technologies. The acquisition was made by Ricoh Group company DocuWare, also headquartered in Germany, which Ricoh acquired in 2019.

The company says the acquisition enhances the data extraction function of Ricoh's Process Automation portfolio from various documents, including paper and handwritten documents, enabling Ricoh to offer customers automation and sophistication in a wide range of business processes.

Ricoh states the Process Automation business is one its growth areas, beginning with data capture and processing with its portfolio of edge devices (including multifunction printers (MFP) and scanners), as well as the management of document and workflow management applications, outsourcing services, and intelligence process automation.

natif.ai is a software startup with 46 employees founded in Germany in 2019. In addition to its strength in Intelligent Capture, leading image recognition and OCR technologies that utilize AI, natif.ai provides a service platform for document classification and data extraction, as well as research and development of high-performance AI models and advanced OCR technology through machine learning.

These technologies enable improved accuracy in reading unstructured data and automatic data extraction from a variety of non-standard business documents such as invoices, order forms, and contracts.

By combining DocuWare and natif.ai's Intelligent Indexing technologies, Ricoh claims a new level of AI powered capture and automation will be available, which can be applied to a wider range of business areas.

UNSW adopts Kapish CM Cloud

The University of New South Wales is migrating its Records & Archives Management System (RAMS) to the Kapish Content Manager Cloud.

UNSW joins CSIRO, University of Newcastle, Victoria University, ACT Education Directorate and Bureau of Meteorology on Content Manager Cloud.

UNSW supports over 60,000 students and over 7,000 staff within 47 schools across 4 campus locations (Kensington, Paddington, Sydney CBD and the Australian Defence Force Academy – Canberra).

Kapish Content Manager Cloud will provide a Cloud eDRMS Platform, based on OpenText Content Manager, for over 2,000 RAMS users and 5TB of documents across all business units.

RAMS contains information classified at a PROTECTED level, that must be available to users securely 24/7 from any device.

Kapish Content Manager Cloud is a zero-footprint solution including ISO27001 (Information Security Management) Certification and IRAP PROTECTED.

For further information, contact Kapish.

idm. information & data manager

Publisher/Editor: Bill Dawes
Email: bill@idm.net.au

Web Development & Maintenance: Cordelta

Advertising Phone: 02 90432943
Email: idm@idm.net.au

Published by Transmit Media Pty Ltd
PO Box 392, Paddington NSW 2021, Australia

Managing information in the GenAI era

In 2024, data management looks a little different than it has in the past. This is largely because generative AI has brought with it new possibilities. However, as with most new opportunities, it has come with new challenges. Two data security experts to share their thoughts on the most important things every organisation needs to know when it comes to getting your data in order in the age of generative AI.

James Greenwood is Regional Vice President of Technical Account Management at Tanium.

"Data has become an organisation's most valuable asset in the eyes of hackers. When we talk about records and information management, identifying all the endpoints an organisation's data is stored on needs to be at the top of the list," said Greenwood.

"In 2024, data is everywhere—from the server room to an employee's smartwatch. This has made it increasingly difficult to protect data because so many endpoints are unknown. Tanium research found that up to 20% of endpoints are unknown in 94% of organisations, with many only auditing a small number of devices connected to their network at any given time.

"Implementing effective information management strategies without visibility across all endpoints is impossible. If you can't see it, you can't secure it. But with traditional manual processes, that's much easier said than done. In fact, it's near impossible.

"That's why, when it comes to an information management strategy, organisations need autonomous real-time visibility to ensure continuous monitoring against threats. This enables an organisation to identify, and in some cases, remove sensitive information living on endpoints either unnecessarily or without protection.

James Greenwood, Tanium

"Autonomous Endpoint Management can also improve process efficiency and ensure that all endpoints are patched in a timely manner to protect sensitive data.

"When starting out on your information management cleanup, it's important to keep in mind four essential steps. Firstly, build a comprehensive inventory of your data storage locations by identifying all endpoints across your network and the sensitive data stored on them. Secondly, prioritise encrypting your data to safeguard it from unauthorised access.

"Thirdly, automate the real-time monitoring of each endpoint, ensuring rapid detection and response to any potential threats. And finally, streamline your data storage practices to minimise unnecessary accumulation. These measures not only enhance efficiency and cost-effectiveness but also help to mitigate risks, particularly in the event of a cyberattack."

Alyssa Blackburn, Director, Information Management at AvePoint, points to a recent report conducted in partnership with the Association for Intelligent Information Management (AIIM) and Center for Information Policy Leadership (CIPL), which found that 77% of organisations acknowledge the need to implement new information management measures to keep pace with the increasing integration of AI into their operations.

Alyssa Blackburn, AvePoint

"This has become particularly evident since the introduction of Microsoft's Co-Pilot AI tools. We know though that AI is only as effective as the data and information it is provided. And old or sensitive content that has not been managed properly is going to become a major headache for organisations leveraging AI-driven technologies.

"To be truly effective, an information management strategy must establish control over data and information throughout its entire lifecycle, from creation to eventual disposal or archival. Firstly, organisations need to understand the state of their data and information in order to make informed decisions. Once it has been thoroughly analysed, critical business data and information can be identified, managed, and protected accordingly.

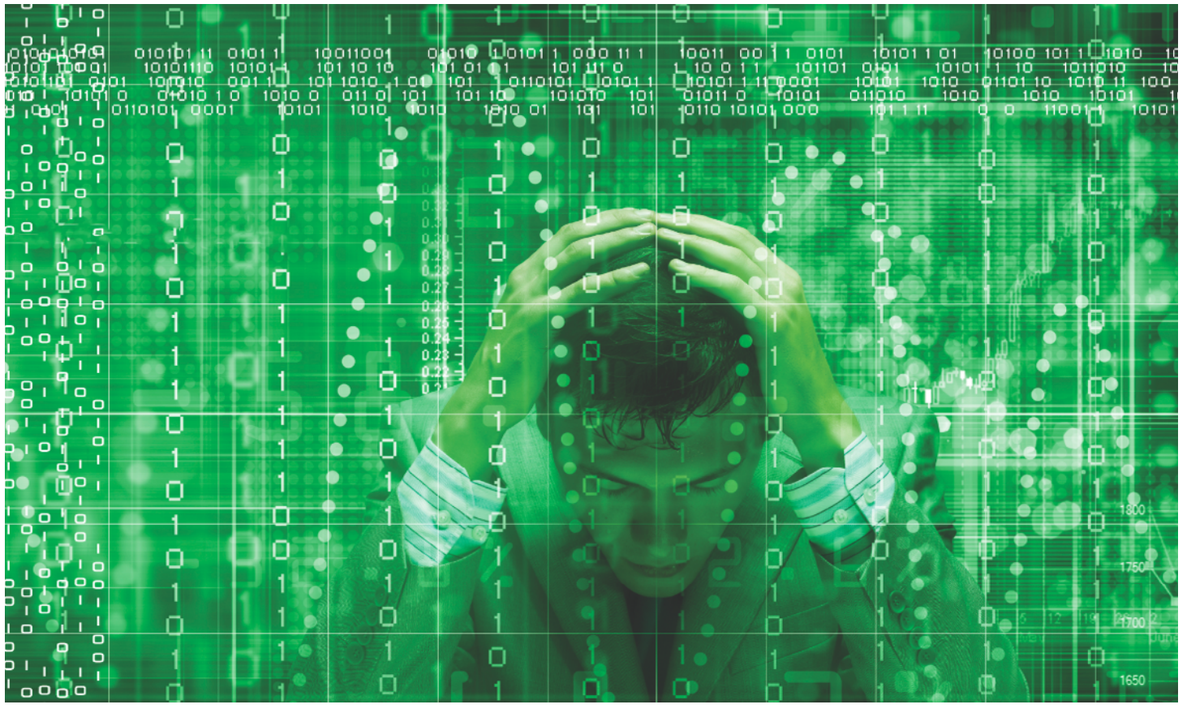
"This process will also uncover Redundant, Obsolete, or Trivial (ROT) content that can be disposed of or stored in cheaper storage tiers to lower costs and reduce regulatory risk. It also helps to improve efficiencies by making sure information that is required on a regular basis can be found more easily. Further to this, it helps to ensure that the quality of the data and information we're running AI tools across is as high as possible.

"Secondly, organisations need to implement ongoing lifecycle management to reduce risks, particularly in relation to data breaches, regulatory non-compliance, and legal liabilities. Lifecycle management also supports better decision-making, improved system performance, and an overall more efficient information management environment.

"It involves setting rules and policies for different types of data based on their value, sensitivity, and compliance requirements and should be regularly reviewed and updated to adapt to changing business needs, evolving regulatory landscapes, and technological advancements. Again, an essential step when preparing to implement AI tools.

"Lastly comes automation. With the explosion of data sets, manually tracking and classifying content has become an impossible task. This is when automating information lifecycle management and policies becomes your best friend.

"For organisations looking to get their data in order, AI is going to present new data challenges whilst also creating significant opportunities. Organisations that embrace AI to help clean up their information management habits will reduce risk, gain valuable insights into their data, make more informed decisions, and maintain a competitive edge," concludes Blackburn.



ICT Commitment-Creep & Organisational Schizophrenia

by Martin Erasmuson

The last couple of years has seen many organisations around the world establishing executive-level Chief Data Officers (CDO). In making those appointments, perhaps those organisations are realising (finally) that their decades-long preoccupation with technology, at the expense of their data, has left them with significant 'data-debt'? And further, that data-debt is having a drag-effect, a negative impact on their plans to join the digital transformation wave sweeping the planet.

The bad-data-trifecta consisting of - 1) prolonged and ineffective data strategy; 2) poor data culture and literacy; leading to 3) data-debt - is the perfect storm to manifest organisational schizophrenia!

Schizophrenia sufferers typically have a 'sensory gating deficit' i.e., impaired neurological processes for filtering out redundant or unnecessary stimuli.

For schizophrenics, everything hits them at once, at the same volume which overwhelms their cognitive ability to figure out what they should be paying attention to; and what to ignore. Does that sound familiar?

Many organisations suffer the same symptoms i.e. their information infrastructure which does, or should consist of the various organisational, strategic, policy, process and financial arrangements necessary to plan for and provide ready discovery and access to relevant information, are inadequate in responding to the perfect storm of an ever-changing business environment and a tsunami of information.

And so, like the schizophrenia sufferer, organisations get overwhelmed as they struggle to figure out what they should be paying attention to; and what to ignore.

Above I introduced the bad-data-trifecta. Add to that a decades-long technology commitment-creep.

What is commitment creep?

We'll all have friends or acquaintances in a long-term de facto relationship, possibly with kids and a mortgage. If you roll back the clock, that relationship likely started with a chance meeting at a bar or social gathering; then a date; a second date; sleeping together; a toothbrush at each other's house; then a microwave; buying a sofa together; then a bassinette!!#%!; and finally, a mortgage.....that's 'commitment-creep'.

You end up committed, but there was never a time when you made a conscious decision to commit.

Commitment-creep is an appropriate description of many organisations de facto-relationship relationship with their information and communications technology infrastructure.

The was no overarching strategy, it just crept up on them, and before they knew it, they have dozens (hundreds) of systems and applications which they find themselves committed to.

And it all started a bit like that initial bar meeting with the couple I mention above. Let's quickly unpack how that happened with a quick 50-year ICT recap.

Back in the 1960s there was the data store, magnetic tapes with files stored on a mainframe and the hard-copy 'print-out'.

(Continued over)

ICT Commitment-Creep & Organisational Schizophrenia

(From previous page

There was no real distinction between data and information because data and the process to extract it were the same application.

In the 1970s the first relational data models appeared meaning for the first time you could separate different information from the source data.

With structured query language (SQL), different applications could access the same data. Into the 1970s and 1980s, relational databases on a mainframe become best practice. So far so good!

Into the 80s, data management began to go awry with the release of applications like 'Vulcan' (later re-released as dBASE II), followed by Microsoft 'Omega' (later becoming Access).

This pivotal turning point in data management seems to have happened without much thought of the consequences as users with only a passing technical ability began creating their own personalised databases filled with all manner of new, copied or modified corporate data.

Into the 90s ICT architectures focused almost exclusively on desktop PCs. As more applications moved to the desktop, the data followed. As silos of systems and information proliferated on every server and every desktop, issues of data management were forgotten.

While IT departments continued to manage the main databases, no one had a clue who had made copies; who had created new data; how, when or if it was being updated. As network servers fill up with a myriad of data, the typical response was to keep installing cheap disk space to keep pace.

Into the 00s, 10s and 20s, virtual servers and the cloud - with every conceivable "(any letter of the alphabet here) as a Service" (XaaS) offering - didn't help, they just created yet another data store for corporate data.

Add to that the global explosion of information and content, and we find every organisation drowning in information; with a poor understanding of the data they need to support their critical organisational objectives; or where and how it is managed.

Time to tune into 'RealityFM'?

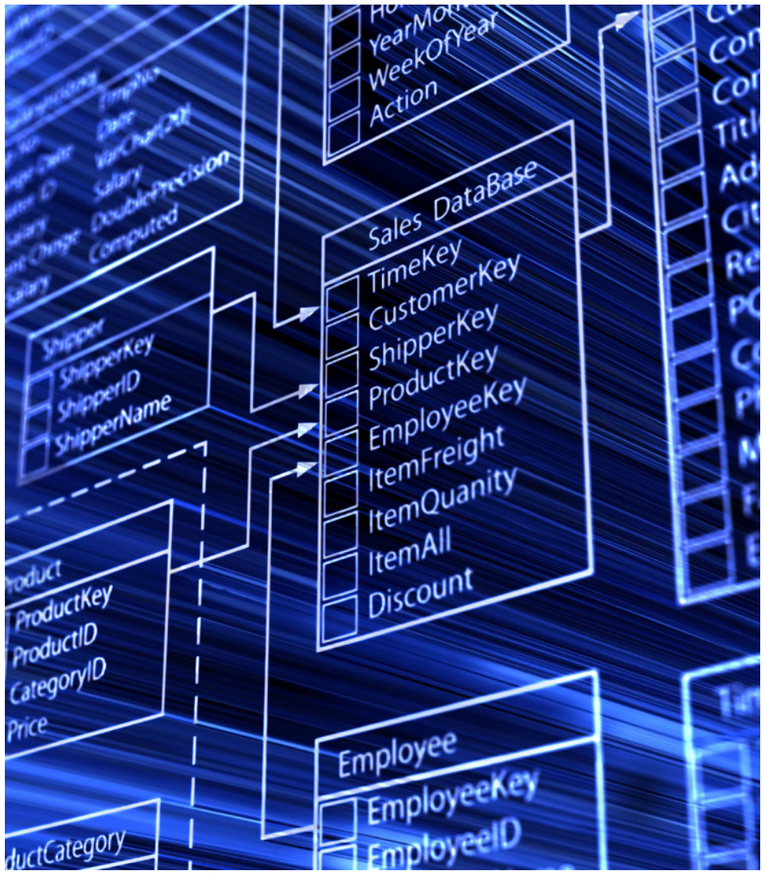
High-performing economies employ high-performing economists i.e. people who can 'speak economics', to negotiate a course through troubled, ever-changing waters.

As organisations come to realise that data is their lifeblood, that realisation should go hand in hand with a need for senior, high-performing 'data-fluent' people. And so, introducing again, the Chief Data Officer (CDO) – they do, or should, speak data.

What makes a good CDO? Here in New Zealand, a large proportion of folk with the job title 'Chief Data Officer' come from an ICT background.

Indeed, on reviewing their LinkedIn profiles, you will find most recording 'IT Manager' (or similar) appearing near the top of their 'Experience', while the word 'data' or 'information' are mostly absent. Why is that a problem?

In answering that, it is not my intention to denigrate what are likely dedicated, hard-working ICT professionals.



That said, this article has sought to make clear that a decades-long infatuation with technology is the key reason we're in this mess, i.e. if you take a problem to a person with a hammer, the solution will be a nail.

In 2023, '*Technology Magazine*' ran a story 'Top 10: Chief Data Officers'. All but one of the ten globally-recognised Chief Data Officers in that story has a background in Data Science, Business Intelligence, Analytics or a related data field.

That tells me how important the Executive and Boards of those organisations consider 'data', and having a senior data-fluent-speaker at the helm.

Most schizophrenia sufferers rely on medication to make it through the day. Organisational-schizophrenia will require an awakening and acknowledgement of data is a strategic asset, and perhaps rehab for our addition to the next shiny piece of kit.

Technology is just plumbing, and the folk managing it, plumbers. We have fleeting, transient flirtations with technology, but data; that's a long-term relationship.

Martin Erasmuson is an Information Architect.

TAFE adopts secure CM Cloud

iCognition has announced the addition of Wodonga TAFE to its growing list of OpenText Content Manager cloud clients. Moving Content Manager to the iCognition cloud will further strengthen the TAFE's commitment to quality and secure information management.

"This cloud service is not only Government IRAP assessed to the level of Protected, but also upgraded to the new ISO27001:2022 Information Security, Cybersecurity and Privacy Protection certification, making it the most secure Content Manager cloud service available, as far as we can tell," said iCognition CEO, Joe Mammoliti.

"iCognition is very excited about this, as Wodonga TAFE will join National Capital Authority, Tasmanian Dept. of Natural Resources and the Environment, Australian Digital Health Agency, Perpetual Limited, and other clients in our secure OpenText Content Manager-as-a-Service cloud service", Mammoliti said.

"Clients are increasingly turning to cloud services not just to reduce costs but also to embrace future innovation. The cloud-based approach facilitates regular updates and new enhancements, providing the TAFE with the opportunity to extend a standard Content Manager solution to iCognition's Ingress Content Services Platform," said Nigel Carruthers-Taylor, Principal at iCognition.

Ingress extends a standard Content Manager system with technologies such as enterprise search, artificial intelligence, business process automation, and content functions that can be accessed directly from corporate applications and Microsoft 365.

"iCognition is proud to partner with the TAFE in their journey towards efficient document and records management. By leveraging our industry-leading solutions, the TAFE will experience greater efficiency, enhanced compliance, and a seamless path to future innovation and digital transformation," said Carruthers-Taylor.

DDOS Attacks Surge in Financial Services

The Financial services sector is now the number one target for DDoS attacks, with a 154% increase in DDoS attacks between 2022 to 2023 according to a new report by global financial service cyber-intelligence sharing group FS-ISAC.In APAC, financial services ranked as the third-most attacked sector, representing 11% of DDoS attacks.

The report, *DDoS: Here to Stay*, revealed that more than one-third (35%) of all DDoS attacks in 2023 were aimed at the financial services industry, which has surpassed the gaming sector as the most-attacked vertical. It blames the dramatic surge on the power of botnets and hacktivism motivated by the Russia-Ukraine War.

The report explains how nation-states, ransomware attackers, criminal groups, and hacktivists have all leveraged DDoS as a part of point attacks or campaigns, often using low-cost DDoS-for-hire services that are available on underground markets.

It also examines how organizations can mitigate the impacts of these attacks through thoughtful cyber hygiene policies, including regularly assessing their networks, applications, and security measures.

Other major findings of the report include:

- DDoS attacks are quickly becoming one of the most prevalent types of cyberthreats, experiencing rapid growth in both number and volume over the past year, with significant jumps in the number of attacks during the second and third quarters of 2023.
- Larger firms and banks with strong brand recognition are more likely to be targeted, as attackers aim to create the appearance of widespread disruption and disinformation. However, they are also the most likely enterprises to have strong mitigations in place.
- Hacktivists and DDoS attacks can disrupt business operations, leading to a loss of credibility, customer trust, and financial damage. Moreover, DDoS attacks may serve as a smoke screen for other malicious activities, such as data theft or cyber espionage.

"While DDoS is an age-old problem, there is a renewed focus driven by heightened geopolitical tensions as nation-states and hacktivists seek to disrupt operations and break trust in the global financial system," said Teresa Walsh, Chief Intelligence Officer and Managing Director, EMEA, at FS-ISAC.

"These DDoS campaigns are becoming more persistent and increasingly multi-vector as they target all areas of the financial sector, including wealth management, banking, credit cards, digital payments, and insurance."

Download the report here.

nib completes migration to cloud

Top-100 ASX-listed company nib holdings has revealed it has completed migrating more than 95% of its on-premise workloads across all business lines into Amazon Web Services (AWS). The cloud-first strategy completed in February led to the closure of the last of its seven data centres, five in Australia and two New Zealand, in March 2024.

"AWS allows nib to expand its current use and future exploration of artificial intelligence (AI) and the ways that AI might drive nib's consumer-facing businesses as they grow to support Australians managing their everyday health," said nib Group Chief Information Officer, Brendan Mills.

Mr Mills said cloud-based computing provides nib with an expansive environment; one that can be dialled-up or down on a needs-basis, is cost efficient and allows better member data analytics.

In 2019, nib worked closely with the regulator, the Australian Prudential Regulation Authority (APRA), and has again worked to meet APRA's Cloud Adoption Standards and Techniques (CAST) requirements.

APRA defines a system of record as one that maintains information essential for an institution to determine its obligations to customers and counterparties.

nib has moved multiple Extreme Inherent Risk (EIR) systems in the final stages of the migration, which aligns with industry best practice.

"By transferring these EIR systems, we've further fortified our technology backbone, ensuring that critical operations are running on advanced and resilient platforms," Mr Mills said.

Building a privacy layer for AI

By Vivek Vaidya

As the democratization of artificial intelligence continues, a new reality is dawning for brands. Data governance is not only a hot topic — it is a potentially lethal pitfall. Brands that want to survive and thrive in this new environment need to bring their data governance policies up to speed. And fast.

How data practitioners think about the confluence of AI and data — how AI models are trained, how data quality is ensured, how permissions are managed — can seem overwhelming. On the ground, many data engineers, business leaders, and sales and marketing teams are looking for an AI data governance north star. What does data governance mean in the realm of AI, practically, as it pertains to privacy and AI?

Regulations are catching up, as seen in 2023 with both the EU AI Act and, in the U.S., the Biden-Harris administration›s Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. However, beyond the realm of regulation, there is a pressing need for practical guidance at the grassroots level. So, what steps should teams take to build a comprehensive AI governance framework? The answer lies in proactive engagement and the development of an enterprise privacy layer for AI, which can be broken down into several actionable components.

Defining AI governance

AI governance is a dual pillar system that includes data governance and model governance.

Data governance is relatively familiar terrain for many brands. It involves ensuring data collection, processing and utilization align with privacy policies, user commitments and regulatory requirements. Remember, this extends to employee interactions with AI tools and their usage must fit within a company's data governance framework. Model governance, on the other hand, focuses on ensuring that both the input, the data used for training models, and output, the data produced by models, adhere to relevant regulations and organizational commitments.

Data governance: A closer look

Data governance can be further broken down into two subcomponents: data cataloging and AI-specific data governance.

Many organizations already engage in data cataloging, maintaining comprehensive data maps and lists of systems where data is stored and processed. This includes tracking data purposes and regulatory compliance, as well as performing data processing impact assessments.

AI-specific data governance involves additional considerations unique to AI applications. Brands must adapt data governance strategies to account for AI's specific needs and ethical implications.

Purpose-built AI governance capabilities

If a specific dataset is used to train an AI model, it's crucial for AI practitioners to determine whether the data purposes are disclosed to the end user. Is their consent obtained for building AI models from their data?

The first pillar of AI governance, specifically, is purpose and permission. This involves clearly defining the reasons for using data, disclosing this usage and always obtaining user consent.

Access control in AI

Access control in AI governance focuses on the data used to build models. For instance, if a model is trained on user data containing unique identifiers, like email addresses, it is worth considering whether such sensitive information is necessary for the model. Often, tokenization or pseudonymization techniques, like hashing with salting, can be employed to enhance privacy. Why use personal identifiable information, such as an email address, when a unique identifier, like simply a number, could suffice? More advanced techniques like differential privacy can also be used to add noise to training data while still allowing models to learn, make accurate predictions and/or make relevant business decisions.

We are already learning from real-world examples that if a user prompts a large language model with enough innocuous information, it can spit out the data used to train itself. If a model inputs personal identifiable information for whatever reason, that data at risk of exposure. Brands must consider why they are inputting data in the first place.

Data quality

Beyond cataloguing, data governance for AI also includes ensuring data quality. This involves checking for bias, representation, labels and feature quality in the data before it is used in models.

Model governance

When building out model AI governance within your organization, consider these three areas:

- **Purpose and permission:** Similar to data governance, this ensures model training complies with your privacy policies and user consent.
- **Model-data lineage:** This is crucial. Model-data lineage is the process of documenting the connection between datasets and models. When deployed effectively, model-data lineage enables brands to look holistically at what datasets are used to train their models and how they compare and relate to each other. This means brands can be confident the data used to train their models is representative of the data that will be used when the model is conducting inference.
- **Model evaluation and assessments:** Finally, it is time to wrap a bow around all this work. Using evaluation and assessments, brands can actively demonstrate that they are doing exactly what they said they would. With the right model evaluations and assessments, they can determine whether a model is performing as intended, what datasets were used to evaluate the model, the properties of those datasets and whether the system contains any bias. Outputs must be fair to the constituents the model serves.

Going one step further, to really build a responsible privacy layer, disturb existing datasets or generate synthetic datasets that have bias, then evaluate models against those sets to verify the model outputs are resilient to bias when the models are used for inference after training.

All this information can be used to prove to the leadership team, business and even regulators, that the company is doing the right thing. Using data governance and model governance for AI is responsible AI governance — from the ground up.

As AI continues to redefine the landscape of data and technology, the onus is on practitioners, leaders and innovators to pave the way for an AI-driven future that upholds the highest standards of privacy, security and ethical practice.

Start with building out data governance and model governance for AI. Think of this as a call to action to embrace the role as a steward of responsible data practices in the AI era.

Let's champion a future where AI benefits all, guided by the principles of responsibility and trust.

Vivek Vaidya is a contributor to the International Association of Privacy Professionals (IAPP)

Originally published HERE

Cybercrime Index ranks countries by threat

Following three years of intensive research, an international team of researchers have compiled what they claim is the first ever 'World Cybercrime Index', which identifies the globe's key cybercrime hotspots by ranking the most significant sources of cybercrime at a national level.

The Index shows that a relatively small number of countries house the greatest cybercriminal threat. Russia tops the list, followed by Ukraine, China, the USA, Nigeria, and Romania. Australia comes in at number 34.

Co-author of the study, Dr Miranda Bruce from UNSW Canberra and the University of Oxford said the study will enable the public and private sectors to focus their resources on key cybercrime hubs and spend less time and funds on cybercrime countermeasures in countries where the problem is not as significant.

"The research that underpins the Index will help remove the veil of anonymity around cybercriminal offenders, and we hope that it will aid the fight against the growing threat of profit-driven cybercrime," Dr Bruce said.

"We now have a deeper understanding of the geography of cybercrime, and how different countries specialise in different types of cybercrime.

"By continuing to collect this data, we'll be able to monitor the emergence of any new hotspots and it's possible early interventions could be made in at-risk countries before a serious cybercrime problem even develops.

"For the first time we have reliable data on the location of cybercriminals, and we also have a way to measure their impact. Government agencies and private enterprises tasked with tackling cybercrime now have a much better understanding of the scale of the problem in their own backyard.

"Up until now, you had to be an experienced cybercrime investigator to know where cybercriminals actually live, but now we can share that information with the public, governments and businesses. It means that we now have a much clearer picture of the extent of the problem and can target our efforts in the right direction."

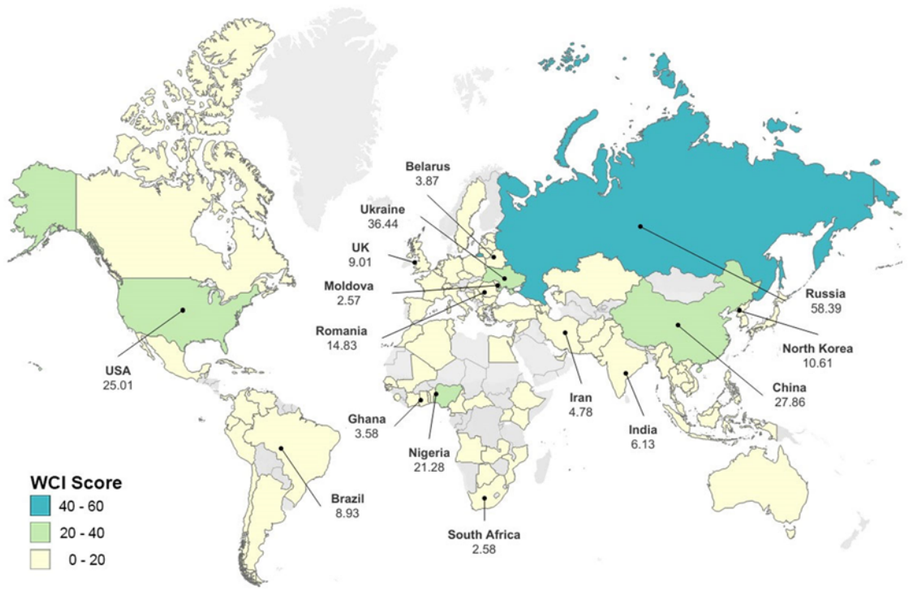
The data that underpins the Index was gathered through a survey of 92 leading cybercrime experts from around the world who are involved in cybercrime intelligence gathering and investigations. The survey asked the experts to consider five major categories of cybercrime, nominate the countries that they consider to be the most significant sources of each of these types of cybercrime, and then rank each country according to the impact, professionalism, and technical skill of its cybercriminals.

Another co-author of the study, Associate Professor Jonathan Lusthaus from the University of Oxford in the UK, said cybercrime has largely been an invisible phenomenon because offenders often mask their physical locations by hiding behind fake profiles and highly technical protections.

"Due to the illicit and anonymous nature of their activities, cybercriminals cannot be easily accessed or reliably surveyed. They are actively hiding. If you try to use technical data to map their location, you will also fail, as cybercriminals bounce their attacks around internet infrastructure across the world. The best means we have to draw a picture of where these offenders are actually located is to survey those whose job it is to track these people," Dr Lusthaus said.

Joint author of the study, Professor Federico Varese from Sciences Po in France, said the World Cybercrime Index is the first step in a broader aim to understand the local dimensions of cybercrime production across the world.

"We are hoping to expand the study so that we can determine whether national characteristics like educational attainment, Internet penetration, GDP or levels of corruption are associated with cybercrime. Many people think that cybercrime is global and fluid, but this study supports the view that, much like forms of organised crime, it is embedded within particular contexts," Professor Varese said.



The World Cybercrime Index has been developed as a joint partnership between the University of Oxford and UNSW and has also been funded by CRIMGOV, a European Union-supported project based at the University of Oxford and Sciences Po. The other co-authors of the study include Professor Ridhi Kashyap, from the University of Oxford and Professor Nigel Phair from Monash University.

'Mapping the global geography of cybercrime with the World Cybercrime Index' is available online at: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0297312

Rising Cyber Threats Heighten Financial Stability Risks, IMF Warns

The International Monetary Fund (IMF) has sounded the alarm on the escalating threat of cyberattacks to global financial stability. The IMF's warning comes amidst a surge in cyber incidents, more than doubling since the onset of the pandemic, with dire implications for the financial sector.

According to the newly released IMF April 2024 Global Financial Stability Report, Financial firms have reported significant direct losses, totalling almost $US12 billion since 2004 and $US2.5 billion since 2020, with indirect impacts such as reputational damage exacerbating the financial toll.

"The financial sector is uniquely exposed to cyber risk. Financial firms—given the large amounts of sensitive data and transactions they handle—are often targeted by criminals seeking to steal money or disrupt economic activity, it revealed in a blog post.

"Attacks on financial firms account for nearly one-fifth of the total, of which banks are the most exposed."

The IMF warns of the potential for cyber incidents to trigger widespread economic turmoil. Citing examples such as the disruption of the national payment system in Lesotho, it emphasizes the significant impact such attacks can have on critical services and economic activity.

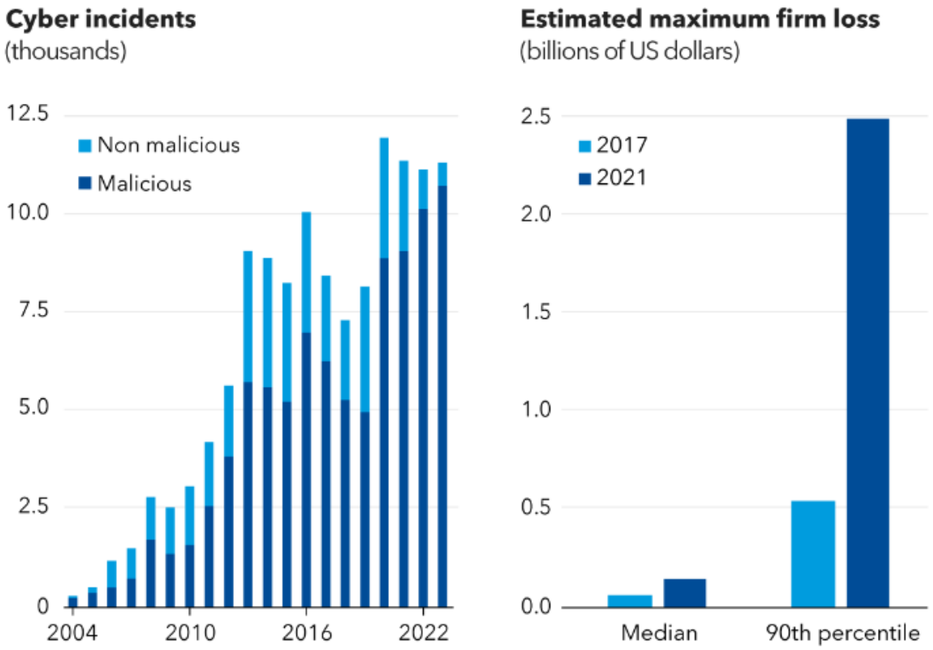
The IMF observed that a variety of reasons are involved in the increase in cyber incidents.

It stated that they include the COVID-19 pandemic's acceleration of the fast-expanding digital connectivity as well as the rising reliance on technology and financial innovation.

The article went on to say that given the spike in cyberattacks following Russia's invasion of Ukraine in February 2022, geopolitical tensions could also be a cause.

Greater threat

The risk of suffering a cyberattack and extreme losses from it has increased.

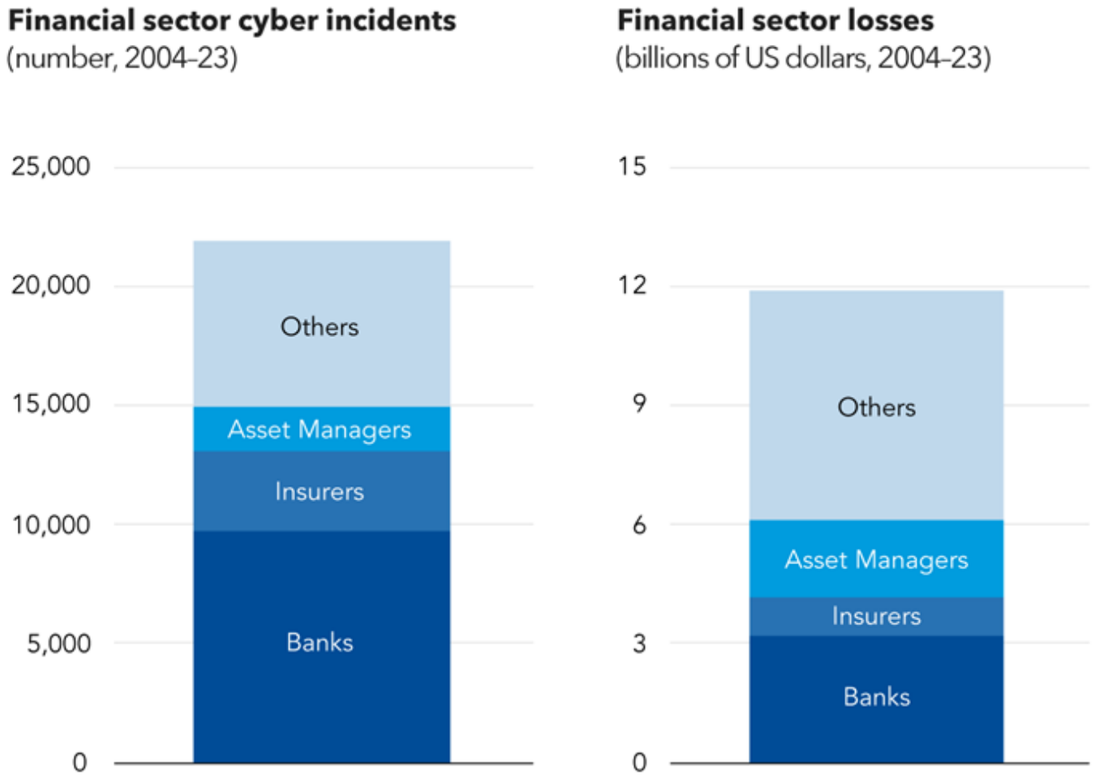


Sources: Advisen Cyber Loss Data; Capital IQ; and IMF staff calculations.
Note: Panel 1 cyber events are classified according to Advisen. Delayed reporting may lead to the underestimation of cyber events in more recent periods. Panel 2 is based on the estimated posterior density function of the highest loss of all firms within a year.

IMF

Attractive target

The financial sector has suffered more than 20,000 cyberattacks, causing $12 billion in losses, over the past 20 years.



Source: Advisen cyber loss data and IMF staff calculations.

IMF

"A cyber incident at a financial institution or a country's critical infrastructure could generate macro-financial stability risks through three key channels: loss of confidence, lack of substitutes for the services rendered, and interconnectedness. While cyber incidents thus far have not been systemic, ongoing rapid digital transformation and technological innovation such as artificial intelligence and heightened global geopolitical tensions exacerbate the risk.

"Recent significant cyber incidents—such as the ransomware attack on the US arm of China's largest bank, the Industrial and Commercial Bank of China, on November 8, 2023, which temporarily disrupted trades in the US Treasury market—further underscore that cyber incidents at major financial institutions could threaten financial stability," it said.

One of the key concerns raised by the IMF, is the reliance of financial firms on third-party IT service providers, which can amplify systemic shocks. An incident in 2023 involving a cloud IT service provider led to simultaneous outages at 60 US credit unions, illustrating the interconnected nature of cyber risks in the financial sector.

The IMF underscores the need for robust cybersecurity policies and governance frameworks to mitigate these risks. Quoting from the post, it stresses the importance of public intervention, particularly in developing economies where cybersecurity regulations are often lacking.

To enhance resilience in the financial sector, the IMF proposes a multifaceted approach, including regular assessments of cybersecurity risks, promoting cyber maturity among firms, improving cyber hygiene, and fostering information sharing among financial institutions.

Highlighting the imperative of international cooperation, the IMF underscores the need for coordinated efforts to address cyber risks effectively. As cyber incidents transcend national borders, collaborative measures are deemed essential to combatting this growing threat.

In conclusion, it urges financial firms to bolster their capacity to deliver critical services during cyber disruptions through robust response and recovery procedures.

The IMF reaffirms its commitment to assisting member countries in strengthening their cybersecurity frameworks through policy advice and capacity-building initiatives.

NSW State and Local Governments on Automation Overdrive, Report Reveals

The use of automated decision-making (ADM) systems by state and local governments in NSW is widespread, varied in function and technology, and actively expanding according to a report from the NSW Ombudsman.

A majority of state government departments and agencies that responded to a survey (46 of 77) reported using or planning to use ADM systems.

Of NSW's 128 local councils, 35 responded, with 14 reporting a total of 77 ADM systems performing a range of purposes, with 23 of the reported systems planned, in development or being piloted.

Use of ADM systems was more likely in metropolitan and city councils, with no concrete ADM systems reported in rural councils.

At both the state government and local government level, ADM systems are being used for a range of purposes, although the pattern of use at each level is different. Local councils primarily reported using ADM systems for public service delivery, user interaction, resource allocation and planning, whereas use in state government was more diverse, with a strong emphasis on compliance.

"We found ADM systems involved across government services, from low to high stakes contexts", said Professor Kimberlee Weatherall, Professor of Law, The University of Sydney Law School and Chief Investigator, ARC Centre of Excellence on Automated Decision-Making and Society (ADM+S).

Key findings included:

- The NSW government sector use of ADM systems is widespread and increasing;
- NSW government organisations are interested in AI, but simpler forms of automation and data linkage and matching are widespread;
- There is widespread use of sensors, computer vision and analysis, including use by local councils;
- Humans are mostly 'in the loop' for now, but further automation is a short step away; and
- There may be a need for wider expertise and testing at the development stage of ADM systems.

"We found that a mapping of this kind is challenging for a whole range of reasons, and so we also provide insights, learned through the process of conducting this mapping, about how to identify, and record ADM system use in government.

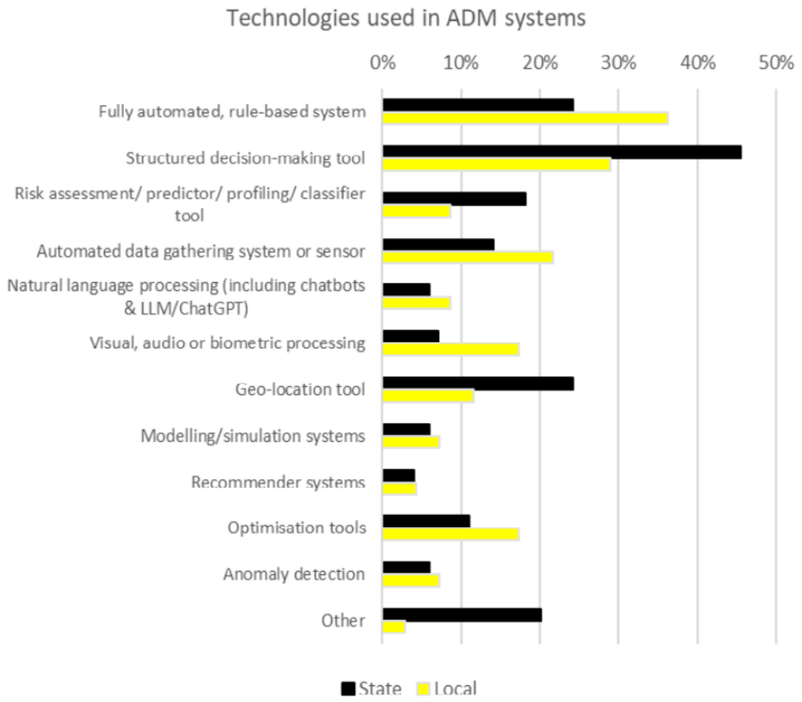
"We believe this will be useful both for researchers, and for governments seeking to be transparent and accountable for their use of technology".

This research was undertaken as a partnership between ADM+S and the NSW Ombudsman seeking better visibility on when and how ADM systems are being used to support or replace the work of NSW public servants in making decisions that affect the public in NSW.

NSW Ombudsman, Paul Miller PSM said, "we hope that all departments, agencies and local councils that have contributed to this research will find the analysis and insights in the report of value, and useful as they continue to consider and pursue their own current and future ADM projects".

The project follows from a ground-breaking report on the use of technology in government decision-making published by the NSW Ombudsman in 2021.

The full report is available HERE



Mature IM Strategies are Critical to Realising AI Benefits: Report

A new global survey has found that data quality remains the top challenge with AI. However, organisations with more mature information management (IM) strategies are 1.5x more likely to realise benefits from AI than those with less mature strategies.

The *AI and Information Management Report* surveyed over 750 digital workplace leaders across the world, of which 20% were from APAC. It was undertaken by the Association for Intelligent Information Management (AIIM) and the Centre for Information Policy Leadership (CIPL).with AvePoint.

It found in addition:

- Before implementing AI, 71% of organisations were concerned about data privacy and security, while 61% were concerned about quality and categorisation of internal data.

- Fewer than half have an AI Acceptable Use Policy, despite widespread use of publicly available generative AI tools (65% of organisations use ChatGPT and 40% use Google Gemini today).

- When implementing AI, 45% of organisations encountered unintended data exposure, meaning one of their biggest concerns became reality.

"Unsurprisingly, data privacy and security were among the top concerns for organisations before implementing AI," said Dana Simberkoff, Chief Risk, Privacy and Information Security Officer, AvePoint.

"The reality is that not enough organisations have the proper policies in place today, which exposes them to risks that could be mitigated, if they better protect and govern their data and educate their employees on the safe usage of this technology."

The survey exposed contradictions in organisations' perception of readiness for AI compared to their reality. Many companies experience gaps in data readiness and information management that have already or will pose significant obstacles to safe and successful AI implementation.

- 88% of organisations report they have an IM strategy in place, but 44% lack basic measures such as archiving and retention policies and lifecycle management solutions.

- Just 29% of organisations use automation in most aspects of their IM strategy. But data volume is growing, with 64% of organisations managing at least 1 petabyte of data and 41% managing at least 500 petabytes of data.

- When implementing AI, 52% of organisations faced challenges with internal data quality.

"The amount of data companies are generating and must manage is growing rapidly, and this will only accelerate as more organisations utilise AI technology," said Alyssa Blackburn, Director of Information Management, AvePoint.

"If organisations don't establish or adapt their information management strategies, the challenges they are already facing will be exacerbated by AI. The good news is, 77% of organisations recognise they must implement additional strategies to keep pace with AI, which is a promising step in the right direction."

Additionally, effective IM strategies can lead to a meaningful return on AI investments, according to the survey. Organisations with mature IM strategies are 1.5x more likely to realise benefits from AI than those with less mature strategies.

Despite this correlation, not enough organisations acknowledge the value of information management in AI success, with only 17% recognising a robust IM strategy as the most effective way to ensure ROI on their AI investments.

This year, organisations are significantly increasing their investments in AI, with 83% planning to increase their AI spending and 79% investing in licensed AI such as Copilot for Microsoft 365.

In addition, 60% of organisations plan to allocate at least a quarter of their technology budget to AI in the next 5 years. However, less than half of organisations (46%) offer AI-specific training, hindering their employees from safely using and optimising this technology.

To read the full AvePoint AI and Information Management Report, visit avpt.co/AI-report-2024

Aussie Early Adopters of Copilot for Security

AustralianSuper, Powerlink Queensland and TAL among the first organisations exploring Microsoft's Copilot for Security to enhance defence capabilities.

The service builds on large language models (LLMs) and harnesses Microsoft's security expertise and global threat intelligence – including 78 trillion daily threat signals – to help security defenders outpace their adversaries and move and respond at machine speed and scale.

TAL, a leading Australian life insurer, and one of the first organisations to deploy Copilot for Microsoft 365 has also signed up to the Copilot for Security EAP.

The company says the investment supports its commitment to technology innovation, with security-focused AI tools key to continuously improving employee and customer experiences – faster, safely.

TAL Chief Information Officer, Hinesh Chauhan said the collaboration with Microsoft has been instrumental in integrating AI technology responsibly and securely.

"Microsoft's Copilot program has allowed us to experiment with and learn from AI technology in a safe way. AI will play an increasingly significant role in tackling cyber security threats at scale, and at speed. We believe our involvement in the Copilot for Security EAP will give our specialist cyber teams valuable insights on what this technology can do to boost our threat detection and response capabilities, and strengthen their knowledge and expertise."

Accelerating detection and response

Speed is paramount in security, with recent data from Microsoft's Digital Defence Report indicating that it takes just 72 minutes for an attacker to access an organisation's private data once they gain access via a phishing email, and 1 hour and 42 minutes for an attacker to begin moving laterally within the corporate network.

Using generative AI in security means that defenders can respond to security incidents in a fraction of the time when every minute matters.

Mick Dunne, Chief Information Security Officer at AustralianSuper spoke of how the technology is helping the Fund's cyber defence team accelerate detection and response.

"We are using Copilot for Security to augment our cyber defence capability. First through the EAP, and now through GA, we are able to leverage AI to speed up event correlation across the suite of Microsoft tools.

"This helps us find productivity benefits and time savings in generating event and incident reporting, while providing an opportunity to enhance our team's effectiveness by summarising security incidents with recommendations and additional context.

"It has streamlined log analysis by using natural language questions, instead of creating complex KQL queries, which accelerates productivity and response for detection and response.

"Another benefit is detecting abnormal indicators of compromise due to the cross-platform integration with Microsoft's security products, where events can be more easily correlated".

With AustCyber predicting an urgent need for an additional 7,000 skilled cyber security specialists in Australia over the next two years, the industry is looking to AI to augment and scale team skills and capacity.

For Powerlink Queensland, joining the Copilot for Security EAP program allowed them to explore generative AI to enable the security team to streamline processes, while fast tracking skills development for their analysts.

"Cybersecurity talent remains one of the biggest challenges when it comes to effectively defending against cyber threats," said Mark Pozdena, General Manager Business IT & Digital Delivery at Powerlink Queensland.

> "We are able to leverage AI to speed up event correlation across the suite of Microsoft tools." - Mick Dunne, Chief Information Security Officer at AustralianSuper

"Equipping our talent with the right tools is critical to optimising our resources and being part of the EAP meant we could be among the first to access the latest AI advancements in security."

"By using natural language prompts, our analysts can initiate and perform tasks more accurately and immediately than what was historically possible and as an added bonus – assist in the development of their skills in the long-term."

Copilot for Security has been offered to select organisations through the Microsoft's Early Access Program (EAP) which kicked off in December 2023. Earlier this month it became available for general purchase to organisations across Australia on a consumption-based pay-as-you-go model.

"With recent reports exposing early moves of threat actors using AI to research and refine their attacks, we are at a pivotal moment where we need to bring new capabilities to ensure we can defend at machine speed," said Evan Williams, Business Group Director for Security at Microsoft ANZ.

"Given the global skills shortage we are facing in cybersecurity, there is an immense opportunity for generative AI to help up-skill security teams, save them time when it matters most, and free them to focus on more complex or strategic work."

A recent economic study examined Copilot for Security users against a control group and found experienced security analysts were 22 per cent faster with Copilot, 7 per cent more accurate across all tasks, and most notably, 97 per cent said they want to use Copilot the next time they do the same task.

An earlier study that focused on new-in-career analysts, found that participants using Copilot were 44 per cent more accurate and 26 per cent faster on foundational tasks such as investigation and response, threat hunting and threat intelligence assessments.

NSW Automated Debt Recovery System Failed Legal Standards: Report

In echoes of the Robodebt report, the NSW Ombudsman has criticised the use of automation to recover fines and debts from defaulters' bank accounts as practised by Revenue NSW until 2022.

Revenue NSW's garnishee order (GO) process system allows it to recover overdue debts directly from the bank accounts of defaulters, a practice it has had the power to conduct since 1998 under the Fines Act 1996 and since 2018 under the State Debt Recovery Act 2018 (SDR Act).

The number of garnishee orders issued annually by Revenue NSW increased dramatically over the years. In the 2010-11 financial year, 6,905 garnishee orders were issued.

This number rose to 1,517,748 in the 2017-18 financial year.

"There are lessons for all public sector agencies from our investigation of Revenue NSW," said NSW Ombudsman, Paul Miller.

"In particular, when agencies are thinking about introducing automation technology to assist in the performance of their functions, whether that involves Artificial Intelligence or otherwise, it is imperative that they seek relevant legal advice and involve legal experts in the design and implementation process," said Mr Miller.

The report Revenue NSW – The lawfulness of its garnishee order process found that Revenue NSW's conduct in using the garnishee order system was contrary to law until March 2019, and wrong until March 2022.

During the investigation, Revenue NSW obtained advice from the NSW Solicitor General. That advice has confirmed that the version of the garnishee order system that is now in place (and has been in place since May 2022) for the recovery of overdue fines is compliant with the Fines Act 1996.

Revenue NSW has advised that it ceased using the garnishee order to recover any other State debts in 2020.

The garnishee order process has evolved over the years, from a manual system to one assisted by technology, and since 2016, an automated system capable of issuing high volumes of garnishee orders daily.

This automated system, referred to as the GO system, has varied in its degree of automation.

From January 2016 to March 2019 the GO system period was found to be contrary to law as it did not comply with the Fines Act.

The automated system issued garnishee orders without an authorized decision-maker forming the required evaluative judgment or making a decision to issue the order.

From March 2019 to March 2022, while the system was modified to include a human review step, it remained defective.

Decision-makers lacked a clear and complete basis for decisions, and the system failed to fully record the decision-making process, making it wrong within the meaning of the Ombudsman Act

The Ombudsman's investigation highlighted the significant hardship caused by the garnishee orders, particularly on vulnerable individuals. Revenue NSW had implemented several changes over the years, such as:

- Introducing a minimum protected amount left in accounts post-garnishment.
- Ceasing the practice of imposing multiple garnishee order costs for each order issued.
- Developing a machine learning model to identify and exclude vulnerable persons from the GO system.
- Implementing a documented hardship policy and criteria for assessing refund requests.

NSW Finance minister Courtney Houssos said the recommendations of the report will be fully implemented.

"While garnishee orders can be an important tool of last resort to recover fines and state debts, it is crucial they are implemented in a fair and equitable way," she said in a statement.

She said Revenue NSW has already put changes in place and will further modify the system to better protect people on welfare and increase transparency.

"In line with the NSW Ombudsman's recommendations, Revenue NSW will also be publishing further information about garnishee orders on its Web site, with a view to better educating the public on how they work and on what grounds they can be challenged."

Knowledge Graphs Redefining Data Management for the Modern Enterprise

By Gergana Petkova

In the current data management landscape, enterprises have to deal with diverse and dispersed data at unimaginable volumes. Among this complexity of siloed data and content, valuable business insights and opportunities get lost.

Not surprisingly, the last decade has witnessed a paradigm shift in enterprise data management, leading to a rise in leveraging knowledge graphs. Providing unified information access, flexible data integration and automation of data management tasks, knowledge graphs have a huge impact on many systems and processes across various industries.

The value proposition of knowledge graphs

One of the key advantages of knowledge graphs is their ability to integrate and unify data from diverse sources. In this way, data is no longer fragmented or potentially lost across different systems or departments of an organization. Instead, it can be viewed, explored and analyzed from a single access point.

Knowledge graphs can also act as a central hub that brings together not only the actual data, but also metadata. This enables enterprises to have a holistic view of all information and better understand the relationships between its different pieces. This is a core component of most data fabric-based implementations.

Knowledge graphs also ensure that data is always represented consistently. Regardless of its original format or source, data is transformed and unified in a way that provides all users with a common framework for understanding and working with it.

Using semantic modelling techniques (such as ontologies and controlled vocabularies) allows knowledge graphs to define precise meanings and relationships between the data. This solves ambiguity issues that plague traditional data management systems.

Another unique ability of knowledge graphs is that they can enhance their proprietary information by leveraging global knowledge as context for interpretation. This means that in addition to the data contained within the graph, external knowledge sources can be integrated to provide a richer and more comprehensive understanding of the data.

Rising above the challenges

The value of data depends on our ability to use it effectively. The vast amounts of data enterprises have today need to be processed, understood and leveraged in realtime. Traditional approaches to data management are no longer sufficient to handle current demands. More and more enterprises are realizing the importance of managing their data differently to reduce costs, improve maintenance and unlock potential revenue and gain competitive advantage.

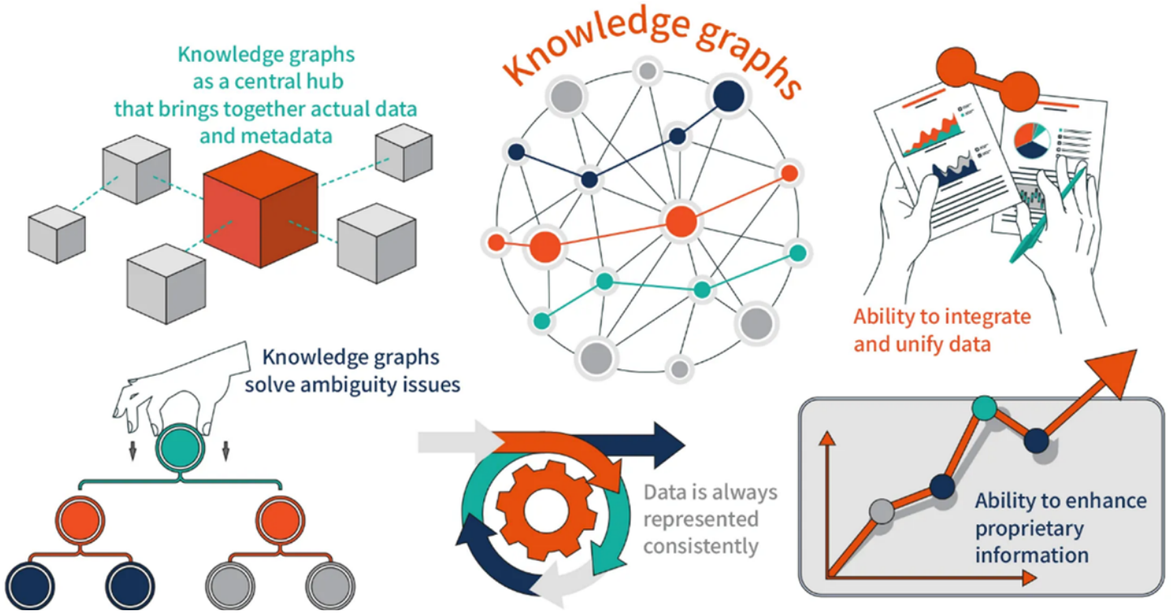
Knowledge graphs provide a viable solution to many data management challenges. They are suitable for every organization, regardless of its size, and can easily handle the diversity and lack of centralized control inherent in modern data ecosystems.

Let's have a look at some of the challenges enterprise data management faces today and how knowledge graphs address them.

Multiplicity of data sources and types

Traditionally, data representation is tightly coupled to specific formats, which dictate how the data is organized and stored. As a result, information and communication technology infrastructures are required to support a wide range of data formats and types spread across various systems. This also includes legacy systems that are used in ways beyond their original purposes.

Knowledge graphs address this problem by providing a



higher-level abstraction for representing data. This allows the format of the data to be decoupled from its purpose or intended use. It's a more flexible approach towards data modelling, where concepts are represented as nodes and relationships between concepts are represented as edges. Since the graph structure is not tied to any specific format, it becomes easier to integrate data from different sources and systems, which enhances data interoperability.

Also, as the purpose or use of data changes over time, the underlying model can be modified without the need to restructure the data completely. This flexibility enables enterprises to better respond to changing business needs or data requirements.

Disconnect between data and the real-world

Relational databases have significant limitations when it comes to organizing and accessing information in an intuitive manner. "Relational databases" are an oxymoron. Relationships are not first-class citizens in such databases. Their rigid tabular structure fails to capture the inherent complexities and rich interconnectedness of real-life data, which often leads to compromises, simplifications and impedance mismatch.

As a result, traditional data management solutions often rely on human efforts to organize real-world information in a way that conforms to the available software. This requires extensive upfront planning and schema design to determine how the data would be stored, connected and queried.

In contrast, knowledge graphs organize information intuitively. By enabling an expressive and flexible representation of data, graph structures capture the relationships between different pieces of information and provide richer context. This empowers enterprises to work with data in a way that closely mirrors their understanding of the domain. It also leads to better insights, decision-making and utilization of data resources.

According to McKinsey, data professionals in most organizations spend 25-30% of their time finding and searching for relevant data. Knowledge graphs help semantically discover, find and explore relevant data, using Linked Data principles. By leveraging shared vocabularies and standards, data can be semantically linked across different sources making it possible to traverse across the connections to find the information they need.

Lack of meaning

Many enterprises fail to make proper use of their data because it's represented in a way that obscures its meaning and the underlying modelling assumptions. This restricts the usability of the data to systems where these assumptions are explicitly hardcoded. It also makes it difficult to integrate data with other systems or leverage it in different contexts.

The representation of data in a knowledge graph conforms to a formal meaning that can be interpreted unambiguously by both humans and machines. This enables a clear understanding of its content. This formalism of semantics powers automated reasoning, which leads to new data-driven insights as well as identifying hidden patterns or relationships.

By leveraging the semantic richness of knowledge graphs, enterprises can better perform important tasks such as answering complex queries, making predictions or generating recommendations. They can track data as it flows through the enterprise, monitor its quality, discover errors, and trace them to the source, reducing bad data quality and data duplication. Knowledge graphs provide high-quality data based on enriched and linked metadata.

Rigid and brittle schemas

The traditional approach in data management calls for defining an exhaustive data schema upfront, based on a presumed understanding of all requirements. However, it's next to impossible to fully grasp the intricacies of a business case and to capture all possible attributes and relationships in a single schema from the outset.

In addition, the dynamic nature of businesses and evolving data needs make it inconceivable to expect a single schema to remain applicable and valid for long periods. The constantly changing technologies, market trends and business requirements demand adaptable data management approaches.

The good news is that knowledge graphs have the ability to model data in a flexible and extensible manner. As new

(Continued over)

(Continued)

data requirements arise, they allow adding or modifying schema elements, without a complete overhaul of the existing structure. This enables enterprises to keep pace with evolving business needs and future demands.

Data silos

Data silos are one of the biggest hurdles to using data efficiently and enterprises often revert to using point-to-point data integration as a quick workaround. However, each integration significantly increases the development time, complexity, resources and effort, slowing down crucial business processes that rely on timely and accurate information.

Most importantly, this approach often fails to address the fundamental problem. Data silos occur when data is isolated and confined within specific systems or departments. Such integrations only establish direct connections between selected systems, leading to a fragmented integration landscape.

The interconnected nature of knowledge graphs, on the other hand, allows enterprises to reuse their existing data assets more effectively and efficiently as well as to easily incorporate external third-party data sources. This empowers them to leverage their proprietary knowledge into the context of global information and enhance business analysis, decision-making and innovation.

In addition, the use of open standards when building knowledge graphs not only ensures interoperability and facilitates the integration of data from multiple domains, but also avoids proprietary formats and vendor lock-in. They connect entities between records in the different datasets to get a 360-degree view. Knowledge graphs excel in providing an entity-centric view that encompasses data across heterogeneous data sources

Expensive data management

As we already discussed, dealing with various data formats and types, reliance on manual efforts, and never-ending data integration projects puts a strain on the overall cost and timelines of enterprise data management.

Knowledge graph approaches offer many cost-effective benefits. For example, the ability to reuse data allows enterprises to better utilize their existing data assets across different applications, projects and teams. Bootstrapping systems with Linked Open Data reduces the costs of data acquisition and maintenance. Flexible graph models eliminate the need for exhaustive (and expensive!), future-proof schema designs and redesigns. The list can go on and on.

Limited time and resources

For most enterprises across all industries, the majority of data remains untapped, invisible, inaccessible and only a small fraction of it is being actively used. This disparity between the vast amount of data and the limited capacity to process it hinders their ability to extract meaningful information and derive actionable insights.

By capturing the rich semantic relationships between concepts, knowledge graphs enable inference capabilities, complex analysis and discovery of hidden patterns. They form a robust backbone for every artificial intelligence and analytics platform and empower users to uncover insights locked in the data and do it in real time.

To wrap it up

The abundance of data requires a data model that is aligned with our complex understanding of information, domain and context. To make data smart, we need to abandon inflexible data schemas and choose data models that can represent the real world with its rich and intricate relationships. When this is done in a machine-readable format with formal semantics, it enables automated reasoning, which complements and facilitates human expertise and decision-making.

Semantic knowledge graphs fulfill these requirements and find applications in many data and information-intensive services across various industries. By adopting them as a foundational component of their data management strategies, enterprises can navigate the complexities of the modern data landscape and make their data and decisioning smarter, faster and data-driven.

Gergana Petkova is Content Manager at Ontotext. Originally published Here

What AI should be doing with documents

By Duff Johnson

In the 1990s, optical character recognition (OCR) got a lot of attention, enabling search engines to work on scanned documents, providing near-instant access to relevant content without laborious manual indexing.

In the mid-to-late 1990s, I owned an imaging service bureau. We delivered OCR output in addition to scanning services. In selling these services it was very important to be able to help customers understand vendors' accuracy statistics.

OCR's power was real, but it wasn't a panacea. Recognition errors led to misses or false positives. These problems were eventually alleviated, in part, by additional software such as language identification and dictionary lookups that post-processed OCR results to improve accuracy. But one fundamental problem was (and remains) harder to solve - complacency. As academic research on the subject has demonstrated, the magic of OCR and search engines inspired far more trust in these systems than the results warranted, leading in some cases to costly mistakes.

Trust in extraordinary technology should not be blindly given, but a key lesson of the information age is that the easier the technology is to use the more readily it gets trusted, even if that trust isn't earned.

AI's promise is great, risks are likewise great

Although Artificial Intelligence (AI) is beginning to dramatically transform how users interact with and process documents, weaknesses remain. Even when the training data is tightly quality controlled, AI's results are not necessarily trustworthy.

Although AI is already assisting in authoring, data extraction, and content management, AI tools can only be as good as their inputs and training. If input data is biased, AI models are helpless to resolve - or even detect - the problem, a major source of AI hallucinations. But, as Air Canada, Michael Cohen, and many others have already realized, relying on AI for critical tasks should prompt far more due diligence than their ease of use implies. AI is a near-miraculous assistant, but trust must be earned, not given.

Competent processing of inputs is a sine qua non for competent, trustworthy AI. So far, however, we see AI developers relying on data volume and (apparently) ignoring data quality.

It's not the AI's fault - it's what they are fed

While the vast majority of content available for training might be unstructured or poorly semantically structured, this doesn't mean that the semantic information it includes should be ignored. When proper semantics are present in a document, such as with WAI-ARIA or Tagged PDF, this information becomes a far richer source of trusted knowledge for AI, whereas attempting to retrospectively guess can result in the familiar garbage-in-garbage-out (GIGO) problem.

If AI assistants helped creators to make richly structured documents - and if consuming AIs were equipped to leverage such enriched inputs, including associated machine-readable source data and provenance information, results would evolve to become more trustworthy.

However, document authoring tends to be oriented towards visual consumption, with unstructured or unreliably structured content. Yes, today's AI assistants can help recognize headings and paragraphs and apply these simple semantics for novice users, but the richer semantics of quotes, referencing, indexing, math, illustrations, etc. do not get the same level of attention. Even structured source information for tabular data (say, an Excel file) is commonly published in unstructured ways while the associated source data is rarely published at all.

This problem is less acute with HTML, as semantic structures are commonly integrated into the content, offloading the complexities of determining appearance to the browser. PDF is necessarily a more complex format than HTML because it's self-contained. That's why good-quality PDF files and a full PDF parser is essential to extracting anything useful from PDF.

The choice of parser matters … a lot. There are literally thousands of PDF parsers covering a vast array of applications and cases …. but only a relative few are truly competent at ingesting PDF content in all its variety. If your parser doesn't support all versions of PDF, uses old Unicode or outdated CMaps, can't understand Tagged PDF, ignores annotations, doesn't process language markers, has no idea about right-to-left or vertically typeset languages, … then inputs into AI will be biased accordingly.

How should AI be integrated with documents?

For creators, AI integration should focus on helping content creators to not only draft and refine their content but also to richly structure and contextualize it. Beyond recognizing tables and lists and offering to structure them, authoring applications should:

- identify quotations, and include the source (either visibly or as metadata);

- when pasting content, include the source (either visibly or as metadata);

- recognize math, and include MathML, even when equation editors are not used;

- expose the AI's confidence in AI-generated alt text of images, and ask the author to review;

- recognize abbreviations and acronyms (does "Dr." mean Doctor or Drive)?

- suggest improvements to the document's structure and ensure appropriate document navigation;

- ensure metadata, referencing and cross-referencing all remain meaningful to the content as it is edited;

- recognize the intentions behind character formatting (are bold words emphasis or defined terms, etc.) and embed those semantics;

- ensure hyphenation and whitespace is semantically indicated;

- retain the semantics of common tools such as org-charts, flow-charts, and drawing tools (these are not just lines and words, but represent semantics!)

- in slide decks and document templates, ensure that page "chrome" is semantically identified;

- ensure that generated PDF files include all this information via Tagged PDF, embedded and associated files, links to structured sources, document and object metadata, ARIA roles and C2PA and digital signatures for provenance and authentication. At the bare minimum, born-digital documents created using AI assistants, regardless of format, should include the full range of accessibility features such as meaningful alternate text for images, logically arranged headings, and MathML for mathematical equations, to name a few. They should also provide open data to support any graphs and charts. We already see some of these features today in various modern office suites, but the author-side AIs aren't (yet) necessarily assisting or prompting the authors to make all these enhancements.

As it happens, the exact same features necessary to support accessibility for users with disabilities can also greatly improve results in AI reuse and extraction scenarios. Authoring applications that provide AI writing support but do not thereafter generate Tagged PDF have only implemented half a solution (and are impeding all AIs that consume these documents in the future)!

On the consumption side

AI systems have a massive appetite for both data and computing resources and can consume vast troves of content from many types of systems across many different formats.

With an "estimated 3 trillion PDF documents on the planet", PDFs are a very attractive source of data. Even so, AI developers often seem to lack "situational awareness" in terms of recognizing when the data fed to their AI is correct and meaningful, and whether the ingestion tools they choose are simply up to the task. The choice of ingestion tool(s) drives the quality of results from content. Is your HTML parser understanding ARIA roles? Does your PDF parser process the full richness of PDF?

As a crude example, there are demonstrable live AI systems today that have "learned" mojibake and "understand" complete gibberish as a Slavic or Asian language! The root cause of such problems often lies in the use of inadequate and outdated technologies that do not support proper text and content extraction from PDF files.

Another approach seen too often is to "dumb everything down" as part of an attempt to support input from pure images of documents (TIFFs, JPEGs, etc). In this case, otherwise rich documents (including, but not limited to PDF) are simply rendered to pixels and then OCR-ed to achieve a form of "consistency" for consumption by the AI engine. Not only is this computationally expensive, but all existing rich semantics and metadata is ignored and replaced by a guess from the OCR process.

Yet another very serious issue making bad headlines for AI systems arises from their unconsented consumption of Personally Identifiable Information (PII) and copyrighted content. This article doesn't address the ethical or legal issues except to point out that PDF (and other formats) support various means to identify and protect content including encryption, digital signatures, and well-defined metadata (such as Dublin Core and C2PA).

Today, most AIs are fed with lousy and/or dumbed-down data that contributes to bias and untrustworthy results- and this is all before the problem of malice. To become truly reliable, AIs need schemes for preserving rich semantics and data when they encounter it. Let the author's AI help the author to provide this richness, and let consumer AIs leverage it when it's provided.

Duff Johnson is CEO of the PDF Association. This article originally published Here.

ISACA Digital Trust Ecosystem launched

ISACA has launched a framework to help organisations achieve digital trustworthiness, and reap the significant benefits that come with high levels of trust. The new Digital Trust Ecosystem Framework (DTEF) has been four years in development and incorporates input from experts across the globe.

It is claimed to be the first single body of knowledge that thoroughly addresses the scenarios, risk and controls required to operate in a digitally trusted environment.

Upcoming ISACA research indicates more than half of organisations (55 percent) believe it is extremely/very important for an organisation to have a digital trust framework. The DTEF is designed to help organisations focus on trust holistically by leveraging technology securely, increasing collaboration, reducing reaction times to unforeseen events, focusing on brand management and improving financial performance through enhanced trust.

Jo Stewart-Rattray, ISACA's Oceania Ambassador said the new framework is progressive, and will provide organisations of all sizes with a roadmap to achieving optimal trust.

"We know the importance of digital trust ranks high among business leaders in Australia," said Stewart-Rattray. "Over two years of gruelling cyber-attacks in this region has taken its toll and the importance of digital trust and digital reputation has been escalated significantly in the eyes of company leaders.

"Most organisations in Australia have the intention to be trustworthy, but sometimes there is a disconnect between intent and action, largely driven by uncertainty around how to achieve trust across all areas of the business. ISACA's new DTEF outlines the practices and steps to achieve trust and the enormous benefits that come with this, including customer loyalty, investor confidence and increased inovation."

The Digital Trust Ecosystem Framework (DTEF) and portfolio of resources helps organisations understand the practices they can undertake to impact their trustworthiness and reputation by addressing the key components of digital trust: integrity, security, privacy, resilience, quality, reliability and confidence.

According to ISACA's tate of Digital Trust research, the benefits of high levels of digital trust include enhancing positive reputation (67 percent), having more reliable data for decision-making (57 percent), and experiencing fewer privacy breaches and fewer cybersecurity incidents (56 percent each). Digital trust, defined as confidence in the integrity of the relationships, interactions and transactions within an associated digital ecosystem, can also strengthen an organisation's bottom line.

"Strengthening digital trust is not just a one-time exercise, but a continuous practice of proactive relationship building between enterprises and stakeholders that is both responsible and profitable while also addressing underlying ethical questions," says Rolf von Roessing, member of the ISACA Digital Trust Working Group and lead developer of the framework, and partner and CEO at FORFA Consulting AG.

"Having a framework to guide organisations as they integrate trustworthiness into their technology and business operations ensures that they are covering all ground, minimising risk and maximising impact."

ISACA is also publishing an interactive guide to help organisations effectively use the framework in a format that is innovative and new to the organisation. The DTEF Interactive Guide allows users to explore the content of the framework with 3D modeling, use cases, and a visual experience. A DTEF Implementation Guide and the Introduction to the Digital Trust Ecosystem Framework Short Course are also available. In addition, ISACA has issued the Digital Trust Board Briefing, available exclusively to members for 60 days, followed by a wider public release.

To access the Digital Trust Ecosystem Framework and resources, visit https://www.isaca.org/digital-trust.

Parliament tackles Digital Transformation

The rapid uptake of automated decision making and machine learning techniques in the modern workplace will b e a major focus of a new Parliamentary inquiry initiated by Australia's Minister for Employment and Workplace Relations, Tony Burke. The House Standing Committee on Employment, Education and Training commenced its inquiry on April 9, 2024. It has invited submissions by Friday 21 June 2024.

The Committee stated it would like to hear directly from workers, employers, software developers and providers, academics, employer groups and trade unions about the digital transformation of workplaces that is currently taking place across Australia.

Chair of the Committee Lisa Chesters MP said, 'following the Committee's recent work on generative artificial intelligence in the Australian education system, the Committee is now turning its attention to examining the intersection between advanced digital technology and the labour market'.

'Australian workplaces are changing the way they operate,' Ms Chesters said. 'The Committee wants to understand what these changes mean for employees and employers, our workplaces and the way we regulate and govern our employment practices'.

Of particular interest to the Committee are the benefits and risks of automated decision making and machine learning in the context of work, the role of business software and regulatory technology companies, and how to ensure the safe and responsible use of these technologies.

The Committee welcomes submissions from interested individuals and organisations by Friday 21 June 2024. More information on the inquiry, including the full terms of reference and details on making a submission, can be found on the inquiry website.

EzeScan is one of Australia's most popular production capture applications and software of choice for many Records and Information Managers. This award winning technology has been developed by Outback Imaging, an Australian Research and Development company operating since 2002. Solutions range from centralised records capture, highly automated forms and invoice processing to decentralised enterprise digitisation platforms which uniquely align business processes with digitisation standards, compliance and governance requirements. With advanced indexing functionality and native integration with many ECM/EDRMS, EzeScan delivers a fast, cost effective method to transform your manual business processes into intelligent digital workflows. EzeScan benefits include: initiate intelligent automated processes; accelerate document delivery; minimise manual document handling; capture critical information on-the-fly; and ensure standards compliance.

www.ezescan.com.au | info@ezescan.com.au | 1300 393 722

EncompaaS is a global software company specialising in information management, powered by next-gen AI. Leading corporations, government departments and statutory authorities trust EncompaaS to govern and optimise information that resides within on-premises and multi-cloud environments. Organisations are empowered to solve information complexity, proactively address compliance and privacy risk, and make better use of data to act strategically at pace. EncompaaS is distinguished in the way the platform utilises AI to build a foundation of unparalleled data quality from structured, unstructured and semi-structured data to de-risk every asset. From this foundation of data quality, EncompaaS harnesses AI upstream to unlock knowledge and business value that resides within information. EncompaaS maintains a robust partner ecosystem, including global consulting and advisory firms, technology partners, and resellers to meet the diverse needs of highly regulated organisations.

encompaas.cloud | enquiries@encompaas.cloud | 1300 474 288

Hyland is a leader in providing software solutions for managing content, processes and cases for organisations across the globe. For 30 years, Hyland has enabled more than 16,000 organisations to digitise their workplaces and fundamentally transform their operations. Hyland has been a leader in the Gartner Magic Quadrant for Content Services for the past 12 years and named one of Fortune's Best Companies to Work For® since 2014, Hyland is widely known as both a great company to work for and a great company to do business with. Our solutions are intuitive to use so organisations can focus on what they do best. Managing information doesn't have to be complicated. At Hyland, our mission is to empower efficiency and agility so our customers can grow and innovate with confidence. We help organisations handle their most critical content and processes with flexible, configurable software solutions.

www.hyland.com/en/| info-onbase@onbase.com| 02 9060 6405

InoTec provides the Australia and New Zealand market a wide range of digitisation and workflow solutions for standard and wide format paper, microfilm, microfiche, book and 3D scanning. As the only factory subsidiary of InoTec Organisationssysteme GmbH (IOG) in Germany, for over 30 years it has supplied and supported SCAMAX production scanners, designed for use in all applications in which large volumes of documents must be scanned very fast and with maximum image quality. SCAMAX customers use their scanners for very diverse applications and the expectations they place on professional scanners are just as clearly defined: high throughput, outstanding image quality, simple and intuitive operation, genuine 24/7 staying power as well as minimum maintenance are obligatory. SCAMAX also adds long service life, low overall operating costs and upgrade options that protect the investment.

https://inotec.com.au| info@inotec.com.au| 1300 447 553

Newgen offers a unified digital transformation platform that includes native process automation, content services, and communication management capabilities. Globally, many successful enterprises across various industries rely on the NewgenONE digital transformation platform—a comprehensive and unified cloud-based platform with low code capability for rapid development of content-driven, customer-engaging business applications. The platform can transform and simplify complex business processes. Equipped with cutting-edge technologies, including mobility, social listening/sensing, analytics, cloud, artificial intelligence (AI), machine learning (ML), and robotic process automation (RPA), the NewgenONE platform helps enterprises stay ahead of the curve. From grass-root citizen experience management, dynamic case management to electronic documents and records management, lending to underwriting, the platform solves multiple use cases across various industries, including government, banking, insurance, and others. Furthermore, Newgen has a robust partner ecosystem, including global system integrators, consulting and advisory partners, value-added resellers, and technology partners.

newgensoft.com/home-anz/ | info@newgensoft.com | 02 80466880

Information Management and Governance (IMG) specialist, iCognition Pty Ltd, helps our clients to maximise the value of their information assets, while minimising cost and risk. We use an integrated Information Management and Governance approach that combines the disciplines of data, records, and information management to value, manage, control and harness information across the enterprise. iCognition's Electronic Document and Records Management System-as-a-Service (EDRMSaaS) represents 20 years of iCognition experience. It is a proven, secure and trusted Software-as-a-Service offering for Content Manager. It can also include iCognition's award-winning RM Workspace for secure web-based end-user access and collaboration, Office365RMBot for fast and easy information governance of Office 365 information, RM Workflow to deliver easy-to-use Content Manager workflows, and RM Public View for publishing and sharing to non-Content Manager users.

www.icognition.com.au | info@icognition.com.au| 1300 00 4264

Kodak Alaris is a leading provider of information capture solutions that simplify business processes. We make it easy to transform documents and data into valuable business information and is where digital transformation starts. Kodak Alaris delivers intelligent document processing and information capture solutions that make sense. We exist to help the world make sense of information with smart, connected solutions powered by decades of image science innovation. Unlock the power of your information with our award-winning range of scanners, software and professional services available worldwide, and through our network of channel partners.

www.alarisworld.com/en-au | AskMe@kodakalaris.com| 1300 252 747

Kapish is a member of the Citadel Group (ASX:CGL).Citadel solve complex problems and lower risk to our clients through our tailored advisory, implementation and managed services capabilities. With over 250 staff nationwide and an ability to 'reach back' and draw on the expertise of over 1,500 people, we are specialists at integrating knowhow, systems and people to provide information securely on an anywhere-anytime-any device basis. Servicing both large and small, public and private sector organisations across all industries, our team of highly qualified staff have global experience working with all versions of Micro Focus Content Manager (CM). It is this experience coupled with our extensive range of software solutions that enable our customers and their projects to be delivered faster, more cost-effectively and with more success. At Kapish we are passionate about all things Content Manager. As a Tier 1, Micro Focus Platinum Business Partner, we aim to provide our customers with the best software, services and support for all versions of the Electronic Document and Records Management System, Content Manager. Quite simply, our products for CM make record-keeping a breeze.

kapish.com.au | info@kapish.com.au | 03 9017 4943

OPEX® Corporation is the industry leader in document and mail automation, providing innovative, unique solutions that help streamline processes, and set the standard for operational efficiency. This includes seamless mail opening and sorting as well as document imaging (scanning), which increases throughput, maximises efficiency, saves time and money, and provides better output. Since 1975, the family-owned and operated company has served as a trusted partner to clients around the world, with more than 1,500 employees continuously reimagining automation technology that solves the most significant business challenges of today and in the future. OPEX provides advanced document and mail automation solutions across numerous industries, including service bureaus, law firms, banks, medical and health organisations, forms processing and archival agencies, and government institutions. OPEX is headquartered in Moorestown, NJ, with facilities in Pennsauken, NJ; Plano, TX; France; Germany; Switzerland; the United Kingdom; and Australia.

https://opex.com | info@opex.com

INFORMOTION is an innovative professional services organisation specialising in the design and implementation of modern information management, collaboration and governance solutions – on-premises, in the cloud or hybrid. INFORMOTION's workflow tools, custom user interfaces and utilities seamlessly combine to deliver compliance, collaboration, capture and automation solutions that provide greater business value and security for all stakeholders. We can help you map and successfully execute your digital transformation strategy. Boasting the largest specialist IM&G consulting teams in Australia with experience that spans over twenty years, INFORMOTION consultants have a deep understanding of business and government processes and the regulatory frameworks that constrain major enterprises. Our compliance experience is second-to-none. INFORMOTION is a certified Micro Focus Platinum Partner and global Content Manager implementation leader. We are also an accredited Microsoft Enterprise Business Partner, Ephesoft Platinum Partner and EncompaaS Diamond Partner.

informotion.com.au | info@informotion.com.au | 1300 474 288

Collaborate with confidence. AvePoint is the largest Microsoft 365 data management solutions provider, offering a full suite of SaaS solutions to migrate, manage and protect data. More than 8 million cloud users rely on our solutions to make their organisations more productive, compliant and secure. Founded in 2001, AvePoint is a five-time Global Microsoft Partner of the Year and headquartered in Jersey City, New Jersey. AvePoint Cloud Records is a SaaS based, IRAP certified and VERS compliant solution used to manage the information lifecycle including content classification; retention and disposal; comprehensive auditing; reporting; and physical records. The Public Office Record of Victoria (PROV) has certified that government agencies and enterprise customers alike can leverage AvePoint Cloud Records to overcome physical and electronic records management challenges around authenticity, reliability, and ensuring content is maintained in a compliant format long-term.

www.avepoint.com| sales@avepoint.com | (03) 8535 3200

Google Gemini powers AI workflow

Automation Anywhere has announced it is utilizing Google Cloud's Gemini models and Vertex AI platform to launch new generative-AI powered solutions developed on top of leading large language models and trained with anonymized metadata from more than 150 million automation processes across thousands of leading enterprise applications.

The company says the models have a deep understanding of how work gets done across enterprise applications and give companies the ability to develop and run complex processes and workflows across enterprise applications automatically, first by understanding the work request via natural language, converting natural language inputs into steps, and then dynamically creating new process workflows. Companies have the power now able to automate 40 to 80 percent of their enterprise processes and tasks.

Using Google Cloud's Gemini models and the advanced capabilities of Vertex AI, Automation Anywhere's platform now can integrate different types of information from text to code to audio to image and video. This empowers employees to use natural language to request automations across systems, generate personalized content, summarize dense documents and automate workflows without leaving their preferred application.

Some use cases include:

Transforming underwriting processes: In commercial lending and insurance, generative AI can review video footage, images, and text from hundreds of pages of documents, leveraging the generative AI-powered assistant to review, summarize, and submit necessary information to help accelerate the underwriting process.

Simplifying patient communications: Generative-AI can generate discharge after-visit summaries in audio format in a patient's native language, assisting doctors and clinicians in how they relay vital healthcare information to their patients to ensure a better understanding of the details of their care.

Automating data extraction with accuracy: Generative-AI powered assistants can extract text and images automatically including retail barcodes, review the output, provide product support, and authorize returns. In pharmaceuticals, generative AI validates package image data including size, colours, logo, usage instructions, expiration dates, and more to ensure quality standards.

Combat Money Laundering: Automation Anywhere's Intelligent Automation platform further enables financial services organizations to improve the efficiency and effectiveness of their anti-money laundering (AML) compliance programs, seamlessly integrating the results of Google Cloud's AML AI product into the investigation workflow and generation of suspicious activity reports.

"Partnering with Google Cloud gave us access to an ecosystem of best-in-class machine learning models, like MedLM for example, a pre-built solution that eases the burden on doctors and healthcare professionals by automating after-visit summaries, reducing manual work and burnout," said Adi Kuruganti, Chief Product Officer, Automation Anywhere.

http://www.automationanywhere.com/

Securing GenAI Data in the Cloud

Baffle has announced new capabilities that allow organizations to secure structured and unstructured data when moving to and across the cloud. As organizations increasingly adopt GenAI, more and more data is being moved across the cloud, resulting in organizations needing to ensure privacy and review security practices to ensure that all data complies with relevant privacy regulations across regions.

"The ability to move files has been around for quite some time," said Ameesh Divatia, founder and CEO of Baffle.

"The explosion of data as GenAI tools have infiltrated all aspects of life has created the need to move data more efficiently and securely. Securing unstructured data at the file or field level creates a unique challenge for organizations trying to get the most out of their data especially with the complexity of various file formats across multiple locations.

"Protecting that data is paramount, and with our latest offering, data can be moved to and within the cloud while meeting all compliance requirements globally for analytics and AI."

Deloitte estimates that by 2025, global data volume will reach 175 zettabytes. With this increase in data volume, organizations have modernized their file transfer process in the cloud, and security teams are reevaluating security practices to protect data in one location and ensure compliance with relevant regulations as that data moves.

With the new capabilities announced today, Baffle is able to protect sensitive data at all times while maintaining compliance requirements and increasing data utility.

Baffle provides organizations with the ability to move data to the cloud that is:

Easy - Baffle integrates seamlessly into existing processes, and no changes to operational processes are required.

Secure - Baffle protects information at the field level for unstructured, semi-structured, and structured files. Advanced access control actively enforces compliance with privacy regulations at the endpoint of the receiving system.

Flexible - Baffle can be customized to protect information in non-standard file formats including PDF, CSV, Parquet, JSON objects and more.

Baffle delivers an enterprise-class data security platform that secures data stores for applications and GenAI with "no code" changes. The solution supports masking, tokenization, and encryption with role-based access control at the logical database, column-, row-, or field level.

https://baffle.io/

Data Detection for Employee Offboarding

Concentric AI has announced a new employee offboarding risk monitoring and reporting module that delivers data detection and response capabilities to identify true risk to data and secure sensitive information when employees leave a company.

Concentric AI's new release also expands its supported structured and unstructured data repositories, including new support for Salesforce, Azure Blob Storage, Azure SQL, and BigQuery, broadening the company's ability to discovery, monitor, classify, and protect organizations' structured and unstructured data.

As organizations experience employee churn, ensuring that data is protected and employees don't leave with sensitive IP is critical. Employees can have access to massive amounts of sensitive information, including customer databases, strategic plans, software, software design, NDAs, contracts, and trade secrets. Employee offboarding presents unique challenges for data security, especially with a workforce that's heavily cloud-driven and more remote than ever before.

Concentric AI enables deep contextual understanding of data sensitivity and permissions management to enable security teams to easily determine which employee data is personal and which is private corporate data, independent of the folder the file may be in, so they can be prescriptive on which data they allow to leave the organization.

The new employee offboarding risk monitoring and reporting module also identifies by category any sensitive data downloaded and/or modified; and any sensitive data shared with personal emails, sent to anyone outside the organization, and emailed or shared across messaging channels.

"Comprehensive DSPM functionalities are essential for identifying the private IP and sensitive company information that needs to remain in the hands of the organization when an employee leaves," said Karthik Krishnan, Concentric AI CEO.

"This new functionality is a result of our customers' requests as they deal with offboarding employees and potential data loss. This industry-first functionality further distances Concentric AI from the competitive field which does not support these capabilities, and also best equips our partners to deliver the industry's leading DSPM capabilities to their customers."

Added data repository support enables Concentric AI to leverage its market-leading data discovery, risk monitoring and remediation capabilities across a broader variety of popular data repositories. With data sprawled across all types of repositories, all organizations are challenged with the inability to clearly identify where all sensitive data resides.

The distributed nature of data repositories in the cloud posed risks of mismanagement and security breaches. Sifting through log files in SIEMs and endpoint folders is time consuming, while ineffective rules-based content segregation and folder-level permissions lead to manual oversight burdens and potential audit and compliance issues.

"Our additional connector support brings the most accurate AI-enabled data classification and robust remediation capabilities to important and popular new data repositories," added Krishnan.

"By expanding support for both structured and unstructured data, organizations can gain a complete understanding of their data without using multiple tools. This release is the latest demonstration of Concentric AI's ability to innovate as the data landscape expands into new areas."

Concentric AI secures data-centric work using AI to protect business-critical information hidden in the numerous applications, data stores and databases used by today's distributed workforce. The company's deep learning solution uses natural language to autonomously and accurately identifies sensitive data, assesses risk, and remediates security issues. It automates unstructured and structured data security both on premises and in cloud environments using deep learning to categorize data, uncover business criticality and reduce risk in enterprises' most popular databases, data repositories, email and messaging applications, and enterprise Gen AI tools and AI assistants.

Concentric AI's Semantic Intelligence DSPM solution scans organizations' data, detects sensitive or business critical content, identifies the most appropriate classification category, and automatically tags the data.

https://concentric.ai/product-overview/

DynaFile enhances HR DocManagement

An electronic filing solution for HR, DynaFile, has been updated with improved user experience, streamlining document handling, and increasing compliance and security.

The new intuitive interface offers features like customizable column display, ordering, and sorting to allow users to focus on the information that is most important to them. Users can also now have multiple search terms selected to view, allowing for cross-correlating information across documents or even employees.

In addition, users can now connect to their Microsoft 365 or Google email accounts, allowing for direct contact lookup and sending emails from users' accounts. The advanced customization of index fields reduces time spent on document handling tasks, making a more streamlined document registration process.

Adding new documents to the system has been simplified with enhanced drag-and-drop features, allowing users to add documents directly to employee's folders. Innovative pre-population features that understand context and recognize unique values have made adding new documents and registration faster and more user-friendly.

With this release, DynaFile has also implemented a fully web-based intra-document document page management interface that allows users to reorganize pages within multiple documents at once. Finally, barcode processing has been enhanced to utilize 2D barcodes that persist index information and allow for future adjustments to document indexing without the need to regenerate barcode cover pages.

Enhanced document-sharing features have been added that allow for one-time access, password/passcode protection, remembering of link expiration dates, and link access notifications. Features like the "share current search" allow users to copy a link to the exact search they are viewing to share with other users, enhancing cloud collaboration.

Other additional document-sharing features, such as sharing specific document pages as attachments and the ability to convert to PDF and send as attachments, increase team collaboration further. In addition, the ability to recover deleted files gives admins a safety net and the ability to recover from mistakes.

Increased compliance and security have been added with advanced audit trails with new user event logs that provide a history of the logged-in user's actions and upgraded report building.

The refactored open API allows for increased interoperability through seamless integrations with systems such as HCM, HRIS, LMS, Payroll, and ATS.

https://www.dynafile.com

Contract drafting with Copilot integration

The Henchman contract drafting platform has announced its integration with Microsoft Copilot, allowing legal professionals to harness the power of collective knowledge stored in their Document Management System to streamline drafting processes.

With this integration, Henchman users can seamlessly access suggestions directly from their Document Management System (eg. iManage, NetDocuments, SharePoint…) when looking for answers through Microsoft Copilot.

The integration ensures that users benefit from the security protocols provided by Microsoft – who announced in July 2023 its support for new voluntary commitments crafted by the Biden-Harris administration to help ensure that advanced AI systems are safe, secure, and trustworthy.

With the Henchman Microsoft Copilot integration:

■ Quickly surface clauses and definitions stored in your Document Management System, along with their contextual data, when prompting in Microsoft Copilot;

■ Easily determine the best precedent for your negotiation with quick access to the source document; and

■ Knowledge managers can ensure the right practice groups have access to the right sources, with the help of mirrored user permissions from your Document Management System.

Similar to Microsoft Copilot's availability for the general audience, Henchman's Microsoft Copilot integration is now available via early access for early adopters of Microsoft Copilot in Microsoft Teams. As Microsoft is gradually rolling out Copilot across its suite of products, Henchman will follow this same rollout cadence. The integration is available free of charge to all Henchman customers.

https://henchman.io/

Cloud Data Access Management

Informatica is launching Cloud Data Access Management (CDAM), a solution based on Informatica's 2023 acquisition of Privitar and its data access management products.

Now integrated into the Intelligent Data Management Cloud (IDMC), Informatica's flagship data management platform, this AI-powered solution is a key component of data access governance, leveraging IDMC's common metadata foundation.

CDAM promises to help organizations manage, share and use their data by ensuring compliance with policies and fostering customer trust.

It integrates seamlessly with IDMC to offer secure and automated AI and analytics use cases.

At the heart of CDAM's offering is the CLAIRE AI engine, which ensures automated sensitive data classifications at enterprise scale.

With CDAM, data teams have universal, automated controls at their disposal to govern data access, ensuring data is not only secure and private but also readily available.

Key Features Include:

■ Intuitive policy authoring to simplify enterprise-wide data access management;

■ Leverages CLAIRE AI Engine to automate sensitive data classifications to identify personal records, personal identifiable information, financial records and more across large datasets;

■ Fully Integrated with Cloud Data Integration and Cloud Data Marketplace for automated contextual controls on data use and sharing; and

■ Streamlines policy enforcement and auditing across diverse environments to enhance security, privacy and data management;

Key Benefits Include:

■ Accelerate access to trusted data for analytics and AI;

■ Reduce the cost of compliance and reduce the risk of data misuse;

■ Simplify controls on data across complex, hybrid data operations; and

■ Automate self-service access to data from hundreds of data sources

"We've heard directly from data leaders, and 40% of them have stated data privacy, protection and compliance as their top challenges when managing their data strategy priorities.

"To address these needs, we are excited to launch our Cloud Data Access Management as part of a holistic data access and governance solution on IDMC," said Jitesh S Ghai, Chief Product Officer at Informatica.

"CDAM combines data access control with classification, discovery and cataloging tools, using metadata to automate processes.

"This makes operations in data-driven organizations more efficient and easier to scale, simplifying and speeding up tasks."

https://www.informatica.com/

Eigen IDP on Azure Marketplace

The Eigen Technologies (Eigen) intelligent document processing (IDP) solution is now available on the Microsoft Azure Marketplace, enabling streamlined deployment and management of Eigen's technology for Microsoft customers.

Eigen's no-code platform gives business users the ability to transform their documents into usable, structured data to automate processes, connect systems and increase efficiencies at reduced costs.

With Eigen, non-technical users can easily train and re-train robust machine learning models to extract data out of structured and unstructured sources.

Users can easily extract information out of complex documents, set confidence thresholds for extracted data, safely leverage the power of large language models within Eigen to uncover additional insights and pass extracted data into other business systems.

"Eigen and the Azure Marketplace are a winning combination for Microsoft customers looking to enable enterprise grade IDP," said Brian Troesch, Eigen's CRO.

"Through our listing, Eigen's intelligent automation platform is now available to more customers worldwide, which we're very excited about."

To learn more about Eigen's technology, visit Eigen's website. To view Eigen's Microsoft Azure Marketplace listing, click here.

Appian unites AI & Process Automation

Appian has announced the latest version of the Appian Platform, introducing Process HQ, a combination of process mining and enterprise AI unified with the Appian data fabric.

Process HQ provides visibility into business operations to enable data-driven decisions and process improvements. The latest version of Appian also extends the practical value of generative AI through enhancements to Appian AI Copilot and the prompt builder AI skill.

By combining the latest technologies in data fabric, process mining, machine learning, and generative AI, Process HQ helps monitor and improve every business process built on Appian.

Process HQ includes:

■ Process insights, which lets business users without a background in process mining or data science uncover insights and explore their business processes through an AI-powered analysis of their workflows.

■ Data fabric insights, allowing business users to explore enterprise data and build custom reports and dashboards. When partnered with Appian AI Copilot, users can gain new insights even faster. Data fabric insights makes report creation possible for business users without any Appian development knowledge and also empowers them to answer common business questions faster, without needing to rely on a data expert or developer to build a report.

Additional generative AI enhancements in Appian's latest release include:

■ Eleven New AI Skills Using Generative AI. These use low-code design to enable users to leverage an LLM to easily tune AI prompts for specific use cases, including: document summarization, PII identification, unstructured document and email extraction, text generation and more. By presenting a curated list of common use cases, the AI Skills simplify incorporating generative AI in Appian applications, enabling users to start from a contextually relevant prompt and efficiently generate reliable AI responses.

■ Appian AI Copilot automates tedious development tasks by generating sample data and application unit tests. Simply provide context using natural language and the amount of data, and let the AI copilot handle the rest, generating data for individual records and for complex sets of related records. AI generated sample data is ideal for user acceptance testing and stakeholder demos, and accelerates the development lifecycle realistic data for any application.

"Process Mining has been held back by two problems - expensive and often subjective manual data prep because even 'standard' systems have been customized, and then limited tools to resolve the process bottlenecks you uncover," says Michael Beckley, CTO & Founder, Appian.

"Process HQ integrates Appian Data Fabric to reduce manual data prep and is unified with Appian process automation so getting from insight to action has never been easier."

https://appian.com/

Mindee Does Away with Data Prep

Mindee has announced the release of docTI (Document Tailored Intelligence), an AI-powered optical character recognition (OCR) tool designed to revolutionize the way SaaS products process documents. The company claims it is the first intelligent document processing (IDP) tool on the market to allow the processing of any document type, in any language, without the requirement of data model training.

Much of data scientists' time is dedicated to data preparation, including tasks like cleansing and annotating data before it can be used to train a model. This not only underscores the extensive efforts behind data model training, but also highlights a significant pain point for product managers: the challenge of finding solutions that can both quickly integrate into their systems and meet their product's unique needs.

docTI directly addresses this challenge by enabling product managers to create custom document processing APIs, test them, and integrate them in minutes rather than months.

"Mindee's core vision is to deliver solutions that align perfectly with the specific needs of our clients," said Jonathan Grandperrin, CEO, Mindee.

"With docTI, we're bringing a one-of-a-kind expertise of deep learning, computer vision, and large language models (LLMs) to provide the most adaptable, dynamic, and high-performing document processing tool on the market."

Building on Mindee's initial offering - an extensive API catalogue for processing common documents like invoices, identity documents, and more - docTI extends this capability to any document type, making it an essential tool for fintech, HRIS, legal systems, and more.

Mindee helps developers automate application workflows by standardizing the document processing layer. Mindee's API helps companies avoid manual data entry and can be used in expense management, accounts payable automation, procurement, accounting, insurance, user and employee onboarding, loan applications, underwriting, and more.

https://www.mindee.com/

Graph Technology enhances GenAI

Neo4j, a Graph Database and Analytics developer, has announced a collaboration with Microsoft to deliver a unified data offering that will address critical data needs for Generative AI (GenAI).

Specifically, the collaboration will see Neo4j's graph capabilities natively integrated into Microsoft Fabric and Microsoft Azure OpenAI Service to combine structured and unstructured data and enable users to uncover hidden patterns and relationships within their data.

"By 2025, graph technologies will be used in 80% of data and analytics innovations - up from 10% in 2021 - facilitating rapid decision-making across the enterprise," predicts Gartner in its *Emerging Tech Impact Radar: Data and Analytics* November 20, 2023 report.

Gartner also notes, "The ability to discover and document data use cases and help build knowledge graphs out of data uses is becoming a vital capability. It is the first step to resolving fragmented data management issues by enabling a GenAI-augmented data fabric," in its January 23, 2024 report entitled *Innovation Insight: How Generative AI is Transforming Data Management Solutions.*

Azure OpenAI Service enables businesses to use advanced AI models and tools that unlock the full potential of their data. Microsoft Fabric is an AI-powered analytics platform that enables everything from data movement to data science, realtime analytics, and business intelligence.

Developers can use OpenAI Service to process unstructured data, structure it, and load it into a knowledge graph. Once in a knowledge graph, users extract insights leveraging Neo4j data visualization and query tools like Bloom or use Neo4j connector with Power BI for business intelligence (BI).

With Neo4j's GenAI functions, Azure OpenAI Service can be used for fully integrated GraphRAG applications, whereby LLM queries can be used against enterprise data in knowledge graphs.

GraphRAG is an enhanced form of Retrieval Augmentation Generation (RAG) whose results demonstrate intelligence or mastery that outperforms other approaches previously applied to private datasets. Enterprises can also use Gen AI orchestration platforms like LangChain and LlamaIndex to build intelligent GenAI applications.

Neo4j's Azure OpenAI Service integration is generally available now. The integration of Neo4j's graph database into Microsoft Fabric will be generally available later this year.

Arun Ulag, CVP, Azure Data, Microsoft, said, ""Microsoft is committed to empowering organizations with the tools and technologies they need to thrive in today's data-driven world.

"Our collaboration with Neo4j represents a significant step forward in delivering innovative data solutions that will enable businesses to unlock new opportunities and drive digital transformation in the GenAI era."

https://neo4j.com/

LLM Data extraction from Planet AI

Planet AI, a developer of cognitive software solutions, has announced the release of IDA 5.2. The latest version offers improvements for document capture, classification, and extracting data from unstructured documents.

This supports businesses and public administrations extensively in integrating AI into their document processes. IDA is a software suite for document processing that can be deployed on-premises in compliance with data protection regulations.

At the core of IDA 5.2 is the LLM-based (Large Language Model) data extraction, which analyzes unstructured data.

"The LLM-based data extraction feature, now in beta with IDA 5.2, is a revolution in document processing," said Jesper Kleinjohann, CEO of Planet AI.

"We use advanced language models, which have become well-known to everyone since the introduction of ChatGPT. However, our technology is specialized in analytical tasks, thus precisely extracting existing data from documents.

"It identifies relevant information through a deep understanding of context – a capability that goes far beyond conventional approaches, providing our partners and customers with a decisive added value in document processing."

IDA's handwriting and machine-print capture is now available as a plug-and-play SDK (software development kit).

New functions for image binarization and cropping dramatically improve the readability and storage efficiency of documents, while the redaction of data fields ensures the protection and security of sensitive information.

https://planet-ai.com/

FSI Communications Archiving Solution

A cloud-based, all communications-in-one archiving and records management solution for financial services firms has been announced by NICE Actimize, a division of NICE, a technology company that specializes in customer relations management software.

NICE Actimize's ARCHIVE-X provides buy-side, sell-side, and online trading platforms, as well as insurance and wealth management firms, with one solution that supports the regulatory requirements for communications archiving and records management.

With ARCHIVE-X, firms can achieve more robust records lifecycle management and reduce compliance risk by archiving all communications securely in the NICE Compliance Cloud.

Because ARCHIVE-X is open and agnostic, it can capture any communications and related data the financial institution needs to store for regulatory or other purposes. It can ingest captured communications from NICE's NTR-X compliance recording solution and other data sources.

With the ability to store and manage more than one hundred different regulated employee communication types, ARCHIVE-X eliminates data silos and supports many existing communication platforms. ARCHIVE-X works directly with NICE Actimize's SURVEIL-X surveillance solution or integrates with third-party solutions.

Among the communication types it may archive are email (Outlook 365) eComms (chat, IM,), voice, video, screenshare, SMS, social, document sharing, CRM, and more. Whether communications originate from the front or back office (via turrets or desktop phones), unified communications platforms (Microsoft Teams, Webex, Symphony, Zoom, RingCentral), chat platforms (Bloomberg, Reuters), or mobile phones, ARCHIVE-X manages all of them in one place.

"Communication archiving and retention remains a serious challenge. Global regulators have issued massive fines at record levels for failures by financial firms and their employees to maintain and preserve required communication records," said Chris Wooten, Executive Vice President, NICE. "ARCHIVE-X can assist in managing this data securely and effectively while protecting both the institution and investors from recordkeeping lapses."

NICE Actimize's ARCHIVE-X also provides these additional benefits:

- ARCHIVE-X provides high levels of security with advanced data encryption, WORM-compliant archiving, SOC-audited data centers, 24/7 monitoring for security threats, and more.

- As a cloud-based SaaS (software as a service) solution, ARCHIVE-X offers flexibility, scalability, and cost-effectiveness. With an availability SLA of 99.99%, ARCHIVE-X is highly resilient and can scale to support archiving 6+ terabytes of data daily.

- ARCHIVE-X makes data governance more manageable and failsafe because it controls retention limits for different communication types in one system, virtually eliminating the need to micromanage data retention in different systems.

Together with NICE Actimize's compliance solutions, NTR-X Compliance Recording and SURVEIL-X Holistic Conduct Surveillance, ARCHIVE-X completes the compliance trifecta to form Compliancentral, a holistic end-to-end communication and trade compliance solution platform. Compliancentral unifies communications capture, archiving, and surveillance into a robust cloud compliance platform, eliminating costly integrations and ensuring interoperability.

https://www.niceactimize.com/

Nominal to Disrupt ERP Market

Nominal is a US startup harnessing generative AI to bridge the gap between what it claims are outdated, costly ERP systems and the financial management needs of modern mid-market, multi-entity businesses.

The company launched with $US9.2 million in seed funding from Bling Capital and Hyperwise Ventures, with participation from Vela Partners, Incubate Fund and executives from Bill.com, Salesforce, Justworks, ServiceNow and Intel.

The $US44 billion ERP market (Gartner 2023), originally revolutionized by the cloud, now grapples with outdated systems requiring costly engineering to meet modern companies' needs.

Nominal believes that generative AI is a pivotal technology, especially in light of the decline in CPA candidates and the surge of financial and accounting point solutions. Compared to existing retrieval-based assistants such as SAP Joule and Sage Copilot, Nominal takes a proactive approach, transforming business logic into automated accounting workflows.

The company plans to use the funds to accelerate its product offering, expand its market reach and increase sales and support resources in the US. The company primarily focuses on mid-market, multi-entity companies such as holding companies, real estate, energy and multinational technology companies in its initial go-to-market plans.

The current ERP landscape is dominated by outdated technologies, costly system integrators and lengthy customization projects. Nominal surveyed over 120 financial professionals and found that for every dollar spent on software licenses, an additional $3 is spent on system integrators. The problem is further exacerbated in multi-entity companies. The complexity of managing multiple ledgers, currencies and regulations has amplified the need for expensive customizations and convoluted spreadsheets.

Nominal says it is bridging this gap with an entirely new, generative AI-powered and cloud-based platform. Nominal's shadow ledger seamlessly extends existing ERPs and general ledgers, requiring no migration and carrying no risk to current operations.

Nominal's key innovation is its generative subledgers that transform free text, documents and spreadsheets into automated workflows. The platform also offers custom workflows, enabling businesses to tailor their financial operations to their specific needs. The generative workflows can operate in auto-pilot mode or under human supervision to ensure financial accuracy.

"Today's dominant ERP software companies were started around 50 years ago and built for large enterprises. These solutions are expensive and difficult to implement, particularly for smaller companies with limited time and resources," said Kyle Lui, general partner at Bling Capital.

https://www.nominal.so/

Monitor Content in Microsoft Teams

Reveille Software, developer of management and monitoring solutions for Enterprise Content Management (ECM) systems, has announced the launch of Reveille Sentry for Microsoft Teams.

This solution offers visibility into Microsoft Teams user activity, content sharing, and configuration changes. Reveille Sentry addresses the complexities of managing multiple ECM platforms - including SharePoint and closely related collaboration solutions that share any ECM repository content.

With the native capability to monitor over 40 Microsoft Teams transactions, Reveille Sentry for Microsoft Teams can operate securely behind a firewall. This approach ensures that sensitive Microsoft 365 access information is not exposed to third-party solution providers, bolstering internal security measures and reducing the operating risk profile.

Key Features include:

Comprehensive Monitoring: Reveille Sentry for Microsoft Teams gathers and observes user activity within Teams, providing out-of-the-box (OOTB) dashboard reports and insights to enhance decision-making and operational efficiency.

Security and Response: The new solution detects unusual Teams activities and integrates seamlessly with Microsoft Defender, enabling rapid containment of any content access breaches, thereby safeguarding sensitive information.

Competitive Edge: Unlike general-purpose Microsoft 365 management tools focused on administration and reporting for only SharePoint, Reveille Sentry offers a specialized view across different ECM platforms. This consolidated perspective is crucial for organizations relying on multiple ECM solutions to manage their collaborative ecosystems effectively. A typical organization has at least 4 ECM platforms per the Association for Information and Image Management (AIIM).

https://www.reveillesoftware.com/

Rossum IDP on Coupa Marketplace

Rossum is offering its next-generation AI for document understanding and end-to-end automation in the Coupa App Marketplace, connecting businesses with certified, pre-built solutions.

Coupa helps teams collaborate to build more agile and sustainable operations, delivering intelligent and responsible spending strategies to meet their companies' purpose.

Rossum's platform leverages advanced AI to automate transactional documents end-to-end and deliver business value beyond productivity. For Coupa customers, this certified integration is an opportunity to gain time and accuracy by automating the ingestion and processing of all invoices, from PDFs to paper, straight into their downstream ERP system without the need for templates.

With a language-agnostic approach and master data validation, Rossum's AI-first cloud-native platform ensures seamless and accurate invoice processing, enabling global businesses to streamline their procurement and financial processes efficiently, no matter the region or document format.

As a certified Coupa App Marketplace solution, it meets the requirements established by Coupa through its Partner Program and is available in the Coupa App Marketplace.

The App Marketplace Partner Program enables software partners to build complementary solutions that easily connect to the Coupa platform. Customers benefit by discovering and connecting solutions to optimize their business spend and mitigate business risk while reducing the cost of third-party software integration.

https://marketplace.coupa.com/en-US/home

Enhancing Purview data governance

Microsoft has named Solidatus as its preferred data lineage integration partner as it launches a reimagined data governance experience in Microsoft Purview.

This new software-as-a-service (SaaS) platform provides integration across data sources, AI-driven efficiency, and actionable insights for data governance processes.

Solidatus claims data lineage forms the foundation of modern data governance. Without it, complexity results in silos, information mismatches, and limited visibility into data movement across the organization.

This partnership will allow Microsoft Purview's global customer base to harness Solidatus' data lineage and its capabilities in fine grain lineage, visualization and version control, boosting their data governance efforts and fostering stronger, more resilient businesses.

"Whether applying AI, driving digital transformation, or ensuring regulatory compliance, our best-in-class data lineage enables users to build a living twin of their enterprise metadata, laying the foundation for informed decision-making," said Philip Dutton, Solidatus CEO.

"Purview customers will gain a unified and actionable view of their governance workflows, empowering them to tackle their most pressing data challenges."

https://www.solidatus.com/

UiPath Unveils New Family of LLMs

UiPath has announced several new generative AI (GenAI) features in its platform designed to help enterprises realize the full potential of AI with automation.

The company states are four key factors that business leaders seeking to embed AI in their automation program must keep top of mind: business context, AI model flexibility, actionability, and trust. The new AI features of the UiPath Platform address these areas to enhance the performance and accuracy of GenAI models and tools and more easily tackle diverse business challenges with AI and automation.

Graham Sheldon, Chief Product Officer at UiPath, said: "Businesses need an assortment of AI models, the best in class for every task, to achieve their full potential. Our new family of UiPath LLMs, along with Context Grounding to optimize GenAI models with business specific data, provide accuracy, consistency, predictability, time to value, and empower customers to transform their business environments with the latest GenAI capabilities on the market.

"These new features ensure that AI has the integrations, data, context, and ability to take action in the enterprise with automation to meet our customers' unique needs."

The family includes:

■ **Generative Large Language Models (LLMs)** - The new LLMs, DocPATH and CommPATH, give businesses LLMs that are extensively trained for their specific tasks, document processing and communications. General-purpose GenAI models like GPT-4 struggle to match the performance and accuracy of models specially trained for a specific task. Instead of relying on imprecise and time-consuming prompt engineering, DocPATH and CommPATH provide businesses with extensive tools to customize AI models to their exact requirements, allowing them to understand any document and a huge variety of message types.

■ **Context Grounding to augment GenAI models with business specific data** - Businesses need a safe, reliable, low touch way to use their business data with AI models. To address this need, UiPath is introducing Context Grounding, a new feature within the UiPath AI Trust Layer that will be entering private preview in April. UiPath Context Grounding helps businesses improve the accuracy of GenAI models by providing prompts a foundation of business context through retrieval augmented generation. This system extracts information from company specific datasets, like a knowledge base or internal policies and procedures to create more accurate and insightful responses. Context Grounding makes business data LLM-ready by converting it to an optimized format that can easily be indexed, searched, and injected into prompts to improve GenAI predictions. Context Grounding will enhance all UiPath Gen AI experiences in UiPath Autopilots, GenAI Activities, and intelligent document processing (IDP) products like Document Understanding.

■ **Autopilot for Developers and Testers** - UiPath Autopilot is a suite of GenAI-powered experiences across the platform that make automation builders and users more productive. Autopilot experiences for Developers and Testers are now available in preview with a targeted general availability in June. Over 1,500 organizations are using UiPath Autopilot resulting in over 7,000 generations and over 5500 expressions generated per week. Autopilot for Testers transforms the testing lifecycle, from planning to analysis, reducing the burden of manual testing and allowing enterprise testing teams to test more applications faster. Autopilot for Testers empowers testing teams to rapidly generate step-by-step test cases from requirements and any other source documents, generate automations from test steps, and surface insights from test results, allowing testers to identify root cause of issues in minutes, not hours or days.

■ **Prebuilt GenAI Activities for faster time to value** - New prebuilt GenAI Activities utilize the UiPath AI Trust Layer and are easy to access, develop with, and leverage high quality AI predictions in automation workflows that deliver faster time to value. GenAI Activities provide access to a growing collection of GenAI use cases, such as text completion for emails, categorization, image detection, language translation, and the ability to filter out personally identifiable information (PII) enabling enterprises to do more with GenAI. With GenAI Activities, enterprises can reduce time to build and achieve a competitive edge using GenAI to help customize the customer experience, optimize supply chains, forecast demands, and make informed decisions.

https://www.uipath.com/

OCR Image Analyzer eyes the cloud

Symmetry Systems, a data security company, has announced the launch of its advanced Optical Character Recognition (OCR) image analyzer which can uncover and secure business-critical images across a hybrid cloud including all unstructured data object stores, including OneDrive, Amazon S3, Google Cloud Storage, Azure Blob Storage, and Snowflake.

It also has the ability to identify sensitive information hidden within various image formats as well as PDFs, and supports over 100 languages.

"As regulatory requirements for document management are becoming increasingly stringent, having an OCR image analyzer is a critical data security capability for all businesses," said Dr. Mohit Tiwari, CEO of Symmetry Systems.

Where other solutions rely solely on existing metadata, Symmetry Systems works closely with clients to strategically deploy OCR where it provides the most value, focusing on reducing data risks and enhancing classification based on existing metadata and context.

Additional features include:

Enhanced Data Discovery and Visibility: Symmetry's OCR capability provides businesses with enhanced visibility into their data stores, revealing critical images and information that were previously challenging to access.

Improved Data Security and Compliance: With the ability to identify sensitive information within images and PDFs, Symmetry's OCR technology allows businesses to identify potential data security and compliance issues.

Efficient Data Management: By automating the evaluation, tagging, and metadata enforcement of images, Symmetry's OCR technology streamlines data retrieval and processing. This efficiency not only saves time but also significantly reduces the risk associated with improperly secured or non-compliant data.

https://www.symmetry-systems.com/

Data Stewardship for Healthcare

Verato has announced the launch of Smart Steward, a generative AI-based assistant for healthcare data stewardship teams. The solution enables healthcare organizations to combine Verato's patented approach to referential matching (the Verato Referential Matching technology) with new generative AI features leveraging Google Cloud's Vertex AI, to drive operational efficiencies and help improve healthcare data integrity.

Smart Steward is the latest addition to the Verato hMDM platform for healthcare identity data management, purpose-built to combine the benefits of EMPI with trusted identity verification, enterprise master data management, and data enrichment.

Key capabilities of the Verato Smart Steward application include:

AI-Powered Recommendations: Verato Smart Steward provides recommendations on how tasks should be resolved. Data stewards can choose to automatically resolve tasks based on AI-powered recommendations to improve their productivity.

Match Transparency: Verato Smart Steward clarifies match decisions and recommendations in easy-to-understand language--allowing data governance teams to make informed decisions on match results.

Conversational Match Explainer: Verato Smart Steward uses conversational AI to explain its recommendations, including why a task was created or why two identities are a match. Data stewards can ask follow-up questions and receive replies in a chatbot-like manner.

The net effect at an organization using Smart Steward is that only the most challenging tasks that require review get raised to the data stewardship team, reducing the team's workload by 50% to 65%. Moreover, the assistant application helps the customer understand the overall match logic in a conversational interface, allowing customers to gain trust and understanding of the logic applied.

Smart Steward has increased automated task resolution by 50% to 70% across the initial set of customers and will gain further accuracy as Verato starts to offer this technology to its other 100 customers.

Visit the Verato hMDM listing on Google Cloud Marketplace at https://console.cloud.google.com/marketplace/product/verato-marketplace/verato-hmdm

INTRODUCING RIGHT-SPEED™ SCANNING

Traditional high-speed scanning requires extensive prep and lots of labour, especially as jobs get messier and messier. High-speed scanners sometimes require multiple operators to keep them in continuous operation. This leads to additional labour hours driving up cost per image and driving down profitability.

The OPEX® Gemini® scanner is designed for maximum versatility and configurability and handles documents at the right speed while requiring minimal prep and controlling costs.

DM DOCUMENT MANAGER AWARDS 2023
WINNER
Imaging Product of the Year: High Volume
OPEX
OPEX Gemini

Visit opex.com to learn more or contact info@opex.com to schedule a demo today.

OPEX®