Integrating Technology, Driving Growth

TechTalk

GALAXY Integrating Technology | Driving Growth

GALAXY OFFICE AUTOMATION PVT. LTD. Issue No. 132, June 2023

We are proud to announce that our team at Galaxy has been honoured with the award

"INNOVATOR OF THE YEAR in APJ Region"

at Dell Tech World 2023 Las Vegas.

This recognition is a testament to our unwavering commitment to innovation, exceptional service, and delivering cutting-edge IT solutions to our valued clients.

MD SPEAKS

Anoop Pai Dhungat
Chairman & Managing Director

Dear Readers,

I am extremely happy to inform you that Galaxy has been recognised by Dell as the "Innovator of the Year in the Asia Pacific Japan Region". This was announced at a glittering ceremony at Dell Tech World 2023 at Las Vegas. Serving our customer by innovating solutions using best of breed technologies is at the very heart of our business philosophy so getting this award reinforces our belief that we are on the right track. I thank all our stakeholders and especially Dell of this recognition.

We are continuously working closely with our technology partners to architect the best solutions for helping our customers achieve their business outcomes in a rapid, cost effective and secure manner. We have recently added some solutions that can reduce your cloud usage costs substantially. We can also help you in cloud-native application modernization, which includes transforming legacy programs into scalable, robust, and flexible cloud-native solutions that leverage the power of hybrid and multi cloud to drive efficiencies. Please contact our experts to know more about this

Happy Reading

Future Is Now

Google is killing off the password forever. Here's what could replace it

Google's new passkey software is a biometric replacement for old-fashioned password. Can we finally forget about having to remember what all our passwords are? Sounds good. But what is a passkey and how will it make you and your devices more secure?

What's wrong with passwords?

The very first digital passwords were invented by an MIT professor in the mid-1960s who needed to give multiple users private access to the same giant computer. Passwords soon became ubiquitous in our computers and it's easy to see why – a simple, memorable word is quick and easy to input when you want to gain access to your computer.

But that's also the problem with passwords. A simple, memorable word such as 'password' or '123456' is very easy to guess, and when hackers ask their computers to guess millions of passwords a second, even quite complex words and codes can be broken instantly.

Nevertheless, it's recommended that you use a different password for every new app – it's not possible for us to remember hundreds of different codes.

So passwords are a weak spot. Doesn't two-factor authentication solve that?

To some extent, yes. But two-factor (or multi-factor) authentication (2FA/MFA) still relies on you remembering the relevant password.

MFA-enabled devices work by asking you for your password before they use another method of identifying you – sending a text or email, or asking for a response via a dedicated app. The theory is that even if hackers have your password, they'd still be unable to gain access because they'd need your phone or computer.

But 2FA is still vulnerable to hackers through various methods. For example, simply resetting a password can sometimes bypass the 2FA, or hackers could 'SIM-jack' your SIM card so that texts go to their device instead of yours.

How do biometric passkeys work?

When your device knows it's really you, then it has to send that approval securely to the application demanding authentication. Passkeys provide that mechanism. They use cryptographic security – the same kind of system used for Secure Socket Layer (SSL) websites to ensure that data transferred between sender and recipient cannot be intercepted and deciphered.

Your phone maintains a private cryptographic key stored on the device and releases a public key to the application. This enables your phone to send a private message to the application that can only be read by that application saying:"the biometric test has been passed".

All you needed to do was look at the phone or put your finger on the fingerprint reader.

And passkeys are better because…

Once we have biometrics and passkeys, we no longer need passwords. And this looks like the next stage in the evolution of computer security. Google recently announced that it's switching from passwords to passkeys, turning off passwords and 2FA altogether for those users who wish to switch.

It's a better solution for everyone: no more passwords to remember, no codes sent to your phone that you have to type in. And should your phone be lost or stolen, it's no problem: the authentication requires your face or your fingerprint. So it won't work for anyone else.

Like all changes, this may take some getting used to – some of us have been using the (same) passwords for a very long time! But adoption is likely to be offered as a choice and given the alternatives, this is a considerable improvement. If you're offered the option of a passkey with biometric authentication, it's worth a try.

https://tinyurl.com/drz6wyux

Synthetic data is about to transform artificial intelligence

Synthetic data is information that's been generated on a computer to augment or replace real data to improve AI models, protect sensitive data, and mitigate bias.

Aim a firehose of data at a human, and you get information overload. But if you do the same to a computer, you get machine-learning models that can learn to complete sentences as you type or detect tumours in medical scans that are often too subtle for a human eye to see.

Data is the raw material fueling much of today's progress in artificial intelligence, producing fresh insights, new discoveries, and decisions backed by more evidence. Data is now so essential to the modern economy that demand for real, high-quality data has grown exponentially. At the same time, stricter data privacy rules and ever larger AI models have made gathering and labeling real data increasingly difficult or impractical.

Synthetic data is cheap to produce, comes automatically labeled, and sidesteps many of the logistical, ethical, and privacy issues that come with training deep learning models on real-world examples. The research firm Gartner estimates that, by 2030, synthetic data will overtake actual data in training AI models.

An unlimited supply of annotated data

The beauty of synthesizing data on a computer is that it can be procured on-demand, customized to your exact specifications, and produced in nearly limitless quantities. Computer simulations are one popular way of creating synthetic datasets. With the help of a graphics engine, you can churn out an endless supply of realistic images and video created in a virtual world.

A second way of creating artificial data is with AI itself, using generative models to create realistic text, images, tables, and other data types. Model architectures that fall under the generative AI umbrella include transformer-based foundation models, diffusion models, and GANs that learn representations of the underlying data to generate versions in a similar style. DALL-E is one of the best known models for generating images, and GPT for text.

One of synthetic data's key advantages is that it comes pre-labeled. Gathering real data and annotating it by hand is time-consuming, expensive, and often humanly impossible. The benefit of having a machine churn out digital facsimiles is that it already understands the data, eliminating the need for humans to painstakingly describe each image, sentence, or audio file.

Keeping sensitive data safe

Another advantage of synthetic data is that it allows companies to sidestep some of the regulatory issues involved in handling personal data. Healthcare records, financial data, and content on the web, are all protected by privacy and copyright laws that make them difficult for companies to analyze at scale.

Financial services often rely on sensitive customer data for internal work like testing software, detecting fraud, and predicting stock market trends. To keep this information safe, companies follow strict internal procedures for handling the data. As a result, it can take months for employees to gain access to the anonymized data. Errors can also get introduced through anonymization that severely compromise the quality of the final product or prediction.

The challenge, then, is to create synthetic financial datasets that can't be traced to individuals but preserves the statistical properties of the original data. "We want to clone the data almost exactly so that it's as useful as the real data but contains none of the sensitive private information," said IBM's Kate Soule, a senior manager of Exploratory AI Research who co-leads, with Akash Srivastava, Project Synderella, a privacy-preserving synthetic data product.

Reducing vulnerability and bias

Synthetic data is also commonly used to test AI models for security flaws and biases. AI models that do well on benchmarks are often easy to trick with adversarial examples — images and text that have been subtly altered to trigger mistakes.

Using publicly available data, IBM researchers recently built a tool to fabricate quote tweets on Twitter to test the robustness of stock prediction models that trawl social media for tips. After ingesting the fake tweet, an AI stock picker that might have predicted that a stock price was falling, and suggested that investors sell, might reverse its decision, and instead nudge investors to buy.

Large models almost always contain hidden biases, too, picked up from the articles and images they have ingested. IBM researchers recently created a tool that finds these flaws and creates fake text to undo the model's discriminatory assumptions. It works by generating a counterfactual conditioned on the class you want to test — a topic, tense, or sentiment — to flip the model's decision.

Take the statement: "my boss is a man." The tool generates a hypothetical statement with the gender reversed: "my boss is a woman." Such a minor change shouldn't cause a classifier to change its "positive" sentiment-rating to "negative," but in this case it does. To mitigate the bias, the model could be retrained on a dataset augmented with counterfactuals, so that it learns that the statements are equivalent and should be classified similarly.

"Real world data is rarely problem-free," said IBM's Inkit Padhi. "Synthetic data allow us to find and fix problems in AI models to make them more fair, robust, and transferrable to other tasks."

Training AI models faster

Training a billion-parameter foundation model takes time and money. Replacing even a fraction of real-world training data with synthetic data can make it faster and cheaper to train and deploy AI models of all sizes.

Synthetic images can be created in multiple ways. IBM researchers have used the ThreeDWorld simulator and related Task2Sim platform to simulate images of realistic scenes and objects for pretraining image classifiers. Not only do the fakes reduce the amount of real training data needed, they can be as effective as real images in pretraining a model to do things like detect cancer in a medical scan.

Synthetic images can be cranked out even faster using generative AI. MIT and IBM researchers recently combined thousands of small image-generating programs to crank out fake images with simple colors and textures. A classifier pretrained on these basic images performed more accurately than models trained on more detailed synthetic data, they found.

Offsetting real data with more synthetic data can also reduce the chances that a model pretrained on raw data scraped from the internet will go off on a racist or sexist tangent. Artificial data made-to-order comes pre-vetted with fewer biases.

Injecting more variety into datasets

The self-driving car industry embraced synthetic data early on. Collecting samples of all potential scenarios on the road, including rare, so-called edge cases, would be impractical to impossible. Synthetic data makes it possible to create customized data to fill the gaps.

Customer-care chatbots also see variation — in the accents, rhythm, and style of how people speak. It could take a chatbot years to learn the nuances of every customer request and how to respond effectively. As a result, synthetic data has become crucial to improving chatbot performance.

An algorithm developed by IBM Research, called LAMBADA, generates fake sentences aimed at filling a chatbot's knowledge gaps. LAMBADA generates the sentences with GPT then vets them for accuracy. "You need to be very creative to imagine all of the edge cases," said IBM's Ateret Anaby-Tavor, an expert in natural language processing. "Instead, you can use a machine that with a push of a button gives you thousands of sentences. You just need to evaluate and filter them."

Sometimes, though, there isn't enough data to create a fake sentence. This is true for thousands of languages spoken worldwide by relatively few people. To train AI models on these so-called low resource languages, IBM researchers have tried pretraining language models on image-grounded gibberish.

They recently showed that a model pretrained on complete nonsense performed nearly as well on a fill-in-the-blank fluency test as a model pretrained on Spanish. No matter what language we speak, said IBM researcher Chuang Gan, our visual world varies very little, creating a common foundation for natural language.

"Teaching the model an emergent language first can make it easier to learn non-Indo-European languages, while avoiding some of the cultural biases that come with pretraining on a Western language," he said.

http://surl.li/htckp

Special Focus

User and Entity Behaviour Analytics (UEBA)

In today's world all cybersecurity tools are generating vast amount of data, which in turn is increasingly becoming tough to uncover and conclude that it is truly a potential information of a real attack. Analytics tools help make sense of the vast amount of data that SIEM, IDS/IPS, system logs, and other tools gather. The challenge here is from determined and persistent threat actors who purposely stretch out their activity across weeks or even months, especially when most SIEM and XDR solutions are incapable of piecing together events across time.

Even worse, is that these solutions primarily use rule-based Machine Learning, which is essentially pattern matching. This makes them ineffective in detecting new attacks and/or variants, which are highly successful in breaching organizations. Sophisticated cyberattackers will find a way to enter a system in some way, and detection even of the seemingly smallest anomaly is crucial now.

Galaxy is now associated with Gurucul in this space of User and Entity Behaviour Analytics (UEBA) solutions.

User and entity behaviour analytics (UEBA) is a cybersecurity solution that uses algorithms and machine learning to detect anomalies in the behaviour of not only the users in a corporate network but also the routers, servers, and endpoints in that network. These UEBA systems produce more data and provide more complex reporting options to our customers. Allowing customers to identify risky, suspicious and malicious behaviour's for accelerated detection of external and internal threats. The analysis includes entity activities that take place but that are not necessarily directly linked or tied to a user's specific actions but that can still correlate to a vulnerability, reconnaissance, intrusion breach or exploit occurrence.

Gurucul User & Entity Behaviour Analytics (UEBA) detects and responds quickly to threats based on an understanding of normal activity that continuously learns and adjust to characterize suspicious and anomalous activity. Helping security teams quickly distinguish malicious activity from false positives. Its risk engine combines all telemetry, analytics and behavioural modelling into a unified risk score that helps security teams prioritize investigation and response actions.

Let us take a closer look at the benefits of UEBA and why companies need to consider adopting it.

- It Identify high-risk profiles with risk-based analytics, data mining, anomaly, and behaviour detection. Help security teams by creating a baseline using profiling attributes from HR records, events, access repository, log management solutions and more.

- Detect advanced persistent threat (APT) attacks and attack vectors and predict data exfiltration by performing entity-centric anomaly detection. Correlate a wide range of parameters including endpoint security alerts, vulnerability scan results, risk levels of users and accounts used, targets accessed, packet level inspection of the requested payloads, and more.

- Detect attacks using ML algorithms tuned to inspect various parameters like timestamp, location, IP address, device, transaction patterns, high-risk event codes and network packets. Identify any deviation from the normal behaviour that may be indicative of a threat.

- Identify as threat actors attempt to traverse the network in search of finding better vantage points to download additional malware, communicate to external servers, and eventually find the location of sensitive data. By detecting unusual activity and suspicious access, Gurucul UEBA can detect this coming technique used by threat actors as part of an attack campaign.

Social engineering and phishing are also on the rise. These strategies do not attack an organization's hardware but rather its people, convincing employees to click on links, download software, and send passwords. Infecting one computer is only the start of a potentially large-scale cyberattack. UEBA seeks to detect even the tiniest of unusual behaviours and prevent a small phishing scheme from escalating into a massive data breach.

Galaxy now offers UEBA solution to its customer's for detecting fraudulent activities, monitoring employees & trusted hosts. Achieving a holistic visibility across the organization's environments, users, and devices, maximizes the SOC team's efficiencies thus delivering cost savings.

Don't let networking and security complexity delay your journey! Galaxy can help your organization extend a consistent solution. To talk to our experts, email us at **marketing@goapl.com**

Tech News

Report: 73% of Indian firms now report being a ransomware victim

The rate of ransomware attacks has increased in India, with 73 per cent of organisations reporting they were a victim of ransomware in 2023, up from 57 per cent the previous year, a new report said on Wednesday. According to cybersecurity company Sophos, about 77 per cent of ransomware attacks against surveyed organisations succeeded in encrypting data, but only 44 per cent paid the ransom to recover their data -- a significant drop from 78 per cent last year.

On a global scale, the report showed that when organisations paid a ransom to decrypt their data, they ended up additionally doubling their recovery costs ($7,50,000 in recovery costs versus $3,75,000 for organisations that used backups to get their data back).

"Although dipping slightly from the previous year, the rate of encryption remains high at 77 per cent, which is certainly concerning. Ransomware crews have been refining their methodologies of attack and accelerating their attacks to reduce the time for defenders to disrupt their schemes," said Chester Wisniewski, field CTO, Sophos. When the root cause of ransomware attacks was examined, the most common was an exploited vulnerability (in 35 per cent of cases), followed by compromised credentials (in 33 per cent of cases). Moreover, the report mentioned that in 30 per cent of cases where data was encrypted, data was also stolen, suggesting this "double dip" method (data encryption and data exfiltration) is becoming commonplace.

Globally, the education sector reported the highest level of ransomware attacks, with 79 per cent of higher education organisations and 80 percent of lower education organisations reporting that they were victims of ransomware. "Human-led threat hunting is very effective at stopping these criminals in their tracks, but alerts must be investigated, and criminals evicted from systems in hours and days, not weeks and months," said Wisniewski.

https://rb.gy/zql0x

India public cloud services market revenue to reach $17.8 billion by 2027: IDC

The revenue of India's public cloud services market totaled $6.2 billion in 2022, and it is expected to reach $17.8 billion by 2027 growing at a CAGR of 23.4% during their period, according to the latest IDC report released Thursday.

The India public cloud services market includes infrastructure-as-a-service (IaaS), platform as-a-service (PaaS) solutions, and software-as-a-service (SaaS). SaaS accounted for the largest share of the overall public cloud services market, followed by IaaS and PaaS in 2022, said IDC.

According to the research firm, the top two cloud service providers held more than 40% of the overall market. "Post the COVID period, the pace of digital transformation has been steady in India as organizations are driving innovation, introducing digital products and services, and automating processes," said Rajiv Ranjan, Associate Research Director, Cloud and Artificial Intelligence, IDC India.

Ranjan said that migrating from legacy infrastructure to the cloud has been the most preferred way to modernise IT infrastructure which is, in turn, driving public cloud services growth.

In the coming years, there may be growing adoption of AI technologies, containerised applications, edge computing, serverless computing, and kubernetes technologies to further enhance efficiency and agility of infrastructure and applications on the cloud.

According to the research firm, enterprises increasingly availed of compute and storage services as part of their IT infrastructure modernisation initiatives. It added that there was also an increased demand for cloud-based collaborative applications, ERM, CRM, and security software.

https://rb.gy/3fn0b