EVOLVING SOLUTIONS

AI FOR MODERN OPERATIONS

While AI won’t address every challenge you have today, it has started the journey to providing incremental benefits across business disciplines, including Modern Operations.

CONTENTS

AI FOR MODERN OPERATIONS

If you're following the media and vendor hype, you've heard that AI is the next shiny object that will solve all your IT Operations challenges. At this point in time, however, this promise is more of a vision than a reality. While AI tools provide useful insights and automation, it's still in the beginning stages for general business use and IT Operations.

Those working in the IT world know that AI isn't new. It's been at work for years, quietly working in the background for retail, marketing, and other applications to offer more personalized websites and services. More recently, IBM and other big tech companies have been incrementally enhancing their products with AI to automate processes to reduce the complexity of hybrid cloud environments.

While AI won't address every challenge you have today, it has started the journey to providing incremental benefits across business disciplines, including Modern Operations.

IT OPERATIONS IS A LOGICAL APPLICATION FOR AI

Operations is a logical place for AI to flourish and take hold. Organizations have plenty of structured and semi-structured data available in predictable formats that can be fed to an AI to perform inference to make recommendations, deliver insights, or ease burdens for IT operations teams.

The visionary goal of AI for IT operations is fully autonomous operations. Theoretically, you could deploy an application, and your AI could put it on the right infrastructure and environment to optimize performance and cost, while always complying with security standards. But we have a long way to go to get there.

In the meantime, operations professionals have many AI tools available to tame growing complexity and help them work smarter, not harder. **In this e-book, we'll highlight some of the practical applications for AI in Modern Operations today and hint at what we can expect in the near future.**

IBM IS LEADING THE WAY IN AI

IBM watsonx is a data and AI platform built for business applications. Its predecessor, IBM Deep Blue, is famous for beating Garry Kasparov at chess in 1996. As Watson, the platform won Jeopardy in 2011. Since then, IBM has had a consistent strategy for building out the platform as a practical business tool based on decades of development. **Given this rich AI legacy, IBM is well ahead of the curve in operationalizing AI for business, making watsonx more mature than emerging platforms.**

As an ecosystem of tools designed to accelerate and scale the use of AI across your organization, watsonx core components include watsonx.ai to build foundational models and generative AI (GenAI), watsonx.data for collecting, storing, querying, and analyzing enterprise data, and watsonx.governance to direct and monitor AI activities.

IBM has already built several products on the watsonx platform, including IBM watsonx Assistant to build chatbots, IBM Turbonomic Management to optimize cost and performance for IT operations, IBM Instana Observability for monitoring applications, services, and infrastructure, and watsonx Code Assistant to increase developer productivity.

As with any AI platform, organizations naturally have concerns over data sources, ownership, and privacy. IBM handles these issues particularly well, partly because it doesn't have a business model around making money from your data. Your data isn't used by other IBM customers, and your trade secrets and regulated information aren't at risk. This foundational governance gives organizations a head start in developing governance to address risk. One of the most significant issues for AI in business today is that the opportunities are getting ahead of the governance.

AI FOR IT OPERATIONS

A primary goal of AI for general business applications is to get actionable insights from data quickly within stringent performance requirements. For AI to work for business, organizations must have IT Operations that can support high volumes of quality data to train models, low latency, and fast infrastructure.

KEY CONSIDERATIONS IN SUPPORTING AI WORKLOADS

AI's Impact on Storage

As AI gradually becomes a bigger part of day-to-day business operations, it will dramatically change how organizations use storage. For example, machine learning requires large data sets for training, including object storage. The bigger the training set, the better. Organizations that plan to train their own models will need adequate storage resources.

Good Data to Train Models

An important activity of machine learning and AI is to recursively determine what data is useful so that the AI tool gets smarter, more accurate, and more refined over time. So, the quality and integrity of data for AI is key to getting consistently accurate outcomes. For example, a security analyst wants to separate the noise from data that delivers real intelligence about security events. Access to accurate, reliable, and complete data, including metadata to make it easily searchable, is a requirement for successful AI projects. Missing the mark on any of these factors will result in unsatisfactory outcomes.

Low Latency

For many applications, speed is of the essence to make AI practical. Today, retailers use AI for heat maps and digital signage to engage shoppers with personalized offers while they're browsing. For this type of inference to be practical, it can't rely on cloud-based inference because it takes too long. In cases like this, inference must happen where the data is being gathered, requiring local compute and storage.

Workloads need the right data at the right location with the right connectivity to meet the low latency requirements for AI workloads. Some data may be best stored in a file system, while other data may be best stored in an object repository. Either way, latency between storage and compute must be low, so the workload must be located as close to the data as possible.

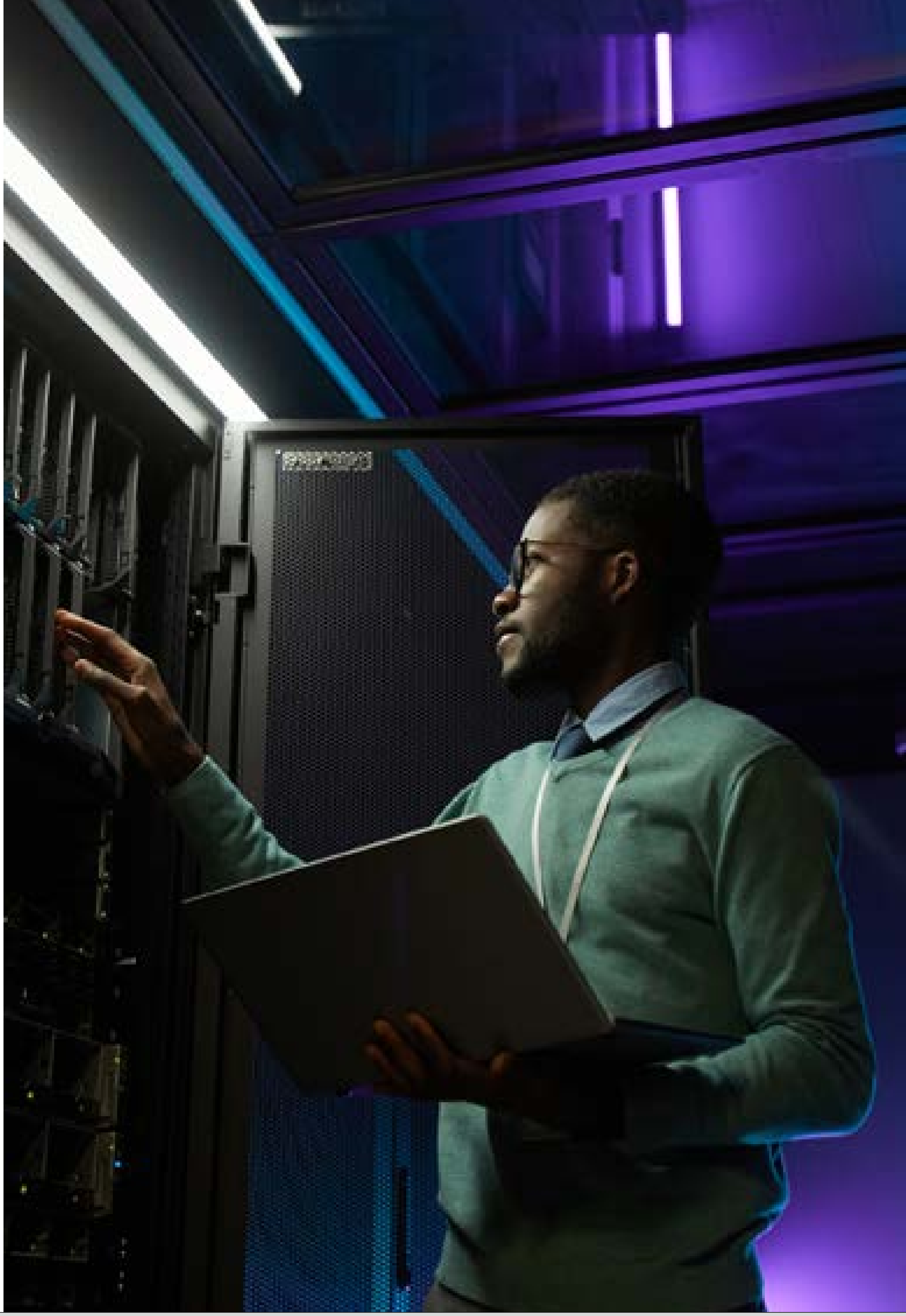
Fast Infrastructure

Low latency can be supported by fast infrastructure, such as AI-optimized chips and ultra-fast storage. Connections between the data and the workload must be fast, and compute must also be fast. Graphics processing units (GPUs) are often used for AI inference as they're optimized for training and running AI models. Given these requirements, operations personnel will, at some point, need to support new deployment models, such as running Kubernetes at the edge.

Backup and Recovery

When relying on AI to run the business, protecting the raw data becomes important. The methods for backing up AI workloads differ from traditional backup strategies, where everything gets copied to a large image. For instance, if you're running containers at the edge, you want to protect the starting points: the container templates, the workflows, and the configurations. You also need to protect the insights gathered at the end. In many cases, what happens in the middle doesn't need to be protected.

Traditional methods like backing up to tape won't work in this environment because the backup window is too small.



THE FUTURE OF AI FOR OPERATIONS

Failure Analytics

As AI for operations matures, it can be used for failure analytics. By sensing patterns, an AI can alert you to a failing network link, disk, or degrading application.

Dynamic Thresholding

Setting thresholds has been more of an art than a science for operations professionals. As AI matures, monitoring tools will be able to look at all the possible parameters to learn what normal behavior is for an application and use that information to set thresholds dynamically.

Code Suggestions

AI is already at work in IT operations to co-generate code or make code suggestions for infrastructure as code to give developers a starting place to, for instance, deploy a new cluster with an AI-generated Helm chart or a deployment file to get the job done based on your performance standards.

GenAI Assistants

GenAI holds a lot of promise for IT operations. While it's still somewhat new, its importance for IT operations can't be understated. Security operations, network management, performance and diagnostic troubleshooting, and application mapping are already being supported by GenAI.

As GenAI for operations matures, it will be able to help you start or review a configuration or build a new Kubernetes cluster based on your standards. But of course, the AI must be trained on the data in your environment using a large language model pre-trained specifically for operations tasks. We're not there yet.

As GenAI and AI for Modern Operations mature, it will be built into product offerings. As a result, you can expect to see dramatic improvements in areas such as storage consumption.

AI FOR SECURITY OPERATIONS

As AI for business comes of age, it will literally be everywhere. Your organization will likely invest in developing or customizing AI tools to help run marketing, finance, business operations, and other areas. Because AI has vulnerabilities, security analysts should keep a watchful eye on any new AI technology your organization wants to implement.

AI FOR THREAT DETECTION AND RESPONSE

Perhaps the best use case today for AI in security operations is the support of threat detection and response operations. When a significant security event occurs, time is always of the essence. Organizations want to respond to a security threat as it happens.

Security is already generating a lot of activity for IT. Still, it's not always obvious what activity is anomalous and if anomalies are associated with a threat actor or failing infrastructure.

Platforms like IBM Security QRadar Suite can correlate events across systems and identify anomalies in near real-time, which can take hours for people to do manually. AI-based security platforms can prioritize risks and reduce or eliminate false positives and other noisy data. This makes incident response teams much more effective in coordinating response efforts because they're focused on solving real problems rather than spending precious hours manually sifting through data. QRadar makes it easy for security analysts to get started by automating the process of connecting to data sources. Some security platforms can even automate remediation, which we'll likely see more of in the near future.

The ability to automate security tasks with AI gives security analysts the freedom to apply human skills to an event. When an AI hands an event off to human security analysts, they have all the information they need to make decisions or take action.

Where AI for security operations is going to excel and be extremely valuable in security operations centers is its ability to adapt to the latest threats, including zero-day threats.

RISKS OF AI IN SECURITY OPERATIONS

Like any other technology, AI has its vulnerabilities. You want to ensure you're getting accurate results when relying on AI to perform critical functions. Erroneous results can be a result of bad data or an immature model. These are vulnerabilities that can be exploited.

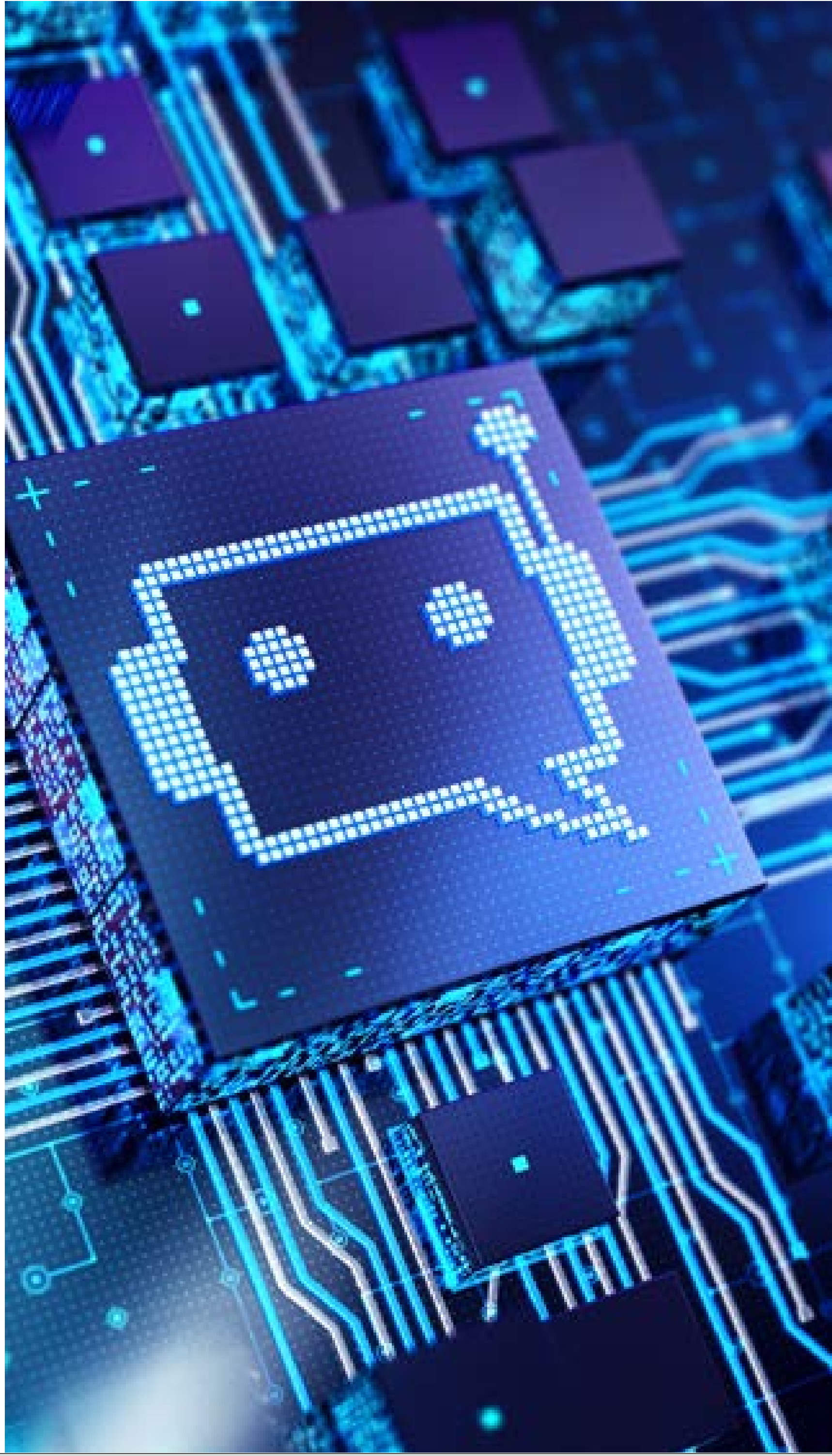
Perhaps the most considerable risk is for security teams not to know what an AI is doing for them. AI has flaws. Even when you're getting consistently reliable feedback from an AI-enhanced security platform, you need to be ready for those times when it doesn't deliver on what it's designed to do. The AI can glitch, fail, or miss something, so you can't blindly rely on it.

RISKS OF AI IN THE ORGANIZATION

Security teams need to understand AI's role in the organization and ensure standards around how to use it are in place. It's important that all employees, and security teams in particular, are trained on how to use AI safely and effectively within the organization.

Using a third-party GenAI platform such as ChatGPT has risks. Some early adopters have fed company information into third-party GenAI platforms, which aren't owned, controlled, or governed by their organization. That confidential information is now "in the wild" and can be exposed to anyone using the platform, by design or by accident.

Getting this right is vitally important because hackers have already begun attacking AI platforms with unusual prompts to trick an AI platform into divulging private information that may have been uploaded to it.



PRIVACY AND COMPLIANCE CONSIDERATIONS

Personal data used to train and run an AI needs to be protected in the same way that it must be protected when centralized in a data lake for analysis. For instance, if your company tracks users' movements and uses AI to predict where an individual might shop next, that data must be protected. If a bad actor gets access to this data, they can also track your users and predict events.

OPERATIONALIZING AI FOR SECURITY

To avoid struggles in operationalizing AI to automate security tasks, you need clearly defined processes and a good idea of where AI can fit into them. The idea is to clearly delineate which processes can be automated and which will continue to require higher levels of human involvement.

THE FUTURE OF AI FOR SECURITY OPERATIONS

As AI for security operations becomes more mature, providers of security tools and platforms will develop more purpose-built models to detect threats and predict what bad actors will do next. For example, we may see the ability to expand correlations beyond our established security platforms to incorporate endpoints and user behavior. This will be especially useful as threat actors start leveraging AI to help them access sensitive data, systems, and devices.

AI TO ENHANCE OBSERVABILITY AND AUTOMATION

There are several use cases for AI today in Modern Operations. AI is currently helping organizations optimize operations and build resilience using platforms such as IBM Instana Observability and its AI counterpart, IBM Turbonomic Management, to help manage cost and performance through automation.

Consolidate Observability Data for AI Analysis

The key to using AI for IT operations is to consolidate observability data for ingestion into an AI platform for analysis and feedback. Unless you're working with legacy platforms and software, you likely already have access to plenty of observability data that AI can use to automate observability tasks and optimize cost and performance.

By connecting an AI observability platform such as IBM Instana to your data sources, you can start automating processes with an AI platform such as IBM Turbonomic.

Root Cause Analysis

AI can analyze observability data to get to the root cause of an issue quickly. For example, if an incident creates hundreds of service tickets, AI can sift through the tickets to identify a root cause. Some AI tools can also suggest a fix or be used to automate remediation.

Monitoring

AI can find logs in external systems that may be useful for certain monitoring tasks and can help you determine how to instrument that dependency to access the data. Since you can't observe everything all the time, AI can help you expand or contract observability as needed based on incoming incidents.

Right-size Observability in Real Time

Using a tool like IBM Instana for connecting to data sources, AI can differentiate what data is needed or not needed to meet observability needs based on system activity. The idea isn't to observe everything all the time but to have the data available when needed and have the AI decide what should or shouldn't be observed at any given moment.

Workflow Automation

AI-powered observability tools can help automate IT workflows to streamline operations work. For example, incident remediation may involve running command lines remotely, but that can be difficult to automate using traditional methods. AI can help tame this level of complexity and make automation like this possible.

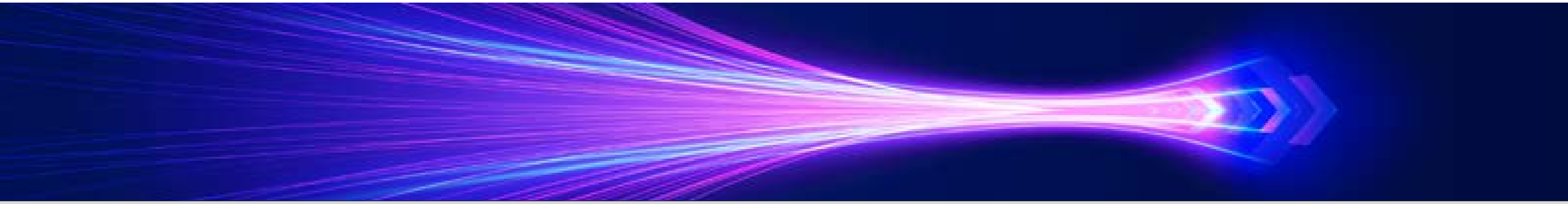
The Future of AI for Observability

The next wave of AI for IT operations will center around GenAI assistants, which will be able to summarize long service tickets, draw on online documentation to help review, create, or optimize a process, or use telemetry data to provide prescriptive advice customized to your environment.

AI AND NETWORKING

AI is making inroads in networking, but at this time, most AI tools are vendor-specific. That's not much help if your organization uses networking gear from multiple providers, which is common. However, these vendor-specific capabilities hint at the near future of AI for networking, including the ability to automate event correlation and proactively detect a root cause.

One promise of AI for networking is to enable intent-based networking. Historically, networking relies on complex protocols that require advanced skills to build complex configurations. AI will enable general practitioners to configure networks based on desired outcomes without the need for advanced knowledge where AI handles the configuration based on a desired outcome. A general practitioner can outline the requirements based on a desired outcome and let AI handle the configuration details.



ABOUT EVOLVING SOLUTIONS

Evolving Solutions helps clients modernize and automate mission-critical applications and infrastructure to support business transformation. We provide consulting services and technical solutions to enable Modern Operations in a hybrid cloud world.

www.evolvingsol.com

EVOLVING SOLUTIONS | IBM Platinum Partner

Let us help you understand your AI options.

Let’s get to work.