Galvanize

Machine learning for governance

How governance, risk, & compliance professionals can start using this evolving technology

Table of contents

Machine learning 101

You might have recently binge-watched the latest season of a TV show on Netflix and noticed that after the final episode, Netflix automatically served up the trailer for another similar series.

Netflix collects and parses large amounts of data around your viewing habits (e.g., preferred genres, ratings, and featured actors), and then uses algorithms to serve up other series that you might enjoy. The more data you give the system—by selecting your own content, or clicking on recommendations—the more accurate these predictions become.

This is machine learning in a nutshell—the idea that computers can find patterns, make decisions, and learn from data. And conversations about machine learning have made their way into the boardroom. As we continue to find new applications for this technology, entire sectors, industries, and individual roles will evolve.

No industry or job will be untouched by machine learning. Hospitals are already using it to better diagnose cancer. Musicians use it to create formulas that more often result in hit records. And retailers put machine learning to use in profiling and upselling customers.

So, are you ready for this machine learning revolution?

> “Robot-led automation has the potential to transform today’s workplace as dramatically as the machines of the Industrial Revolution changed the factory floor.”
>
> » ***Deloitte, 2017, Automate this: The business leader’s guide to robotic and intelligent automation***

The beginnings of machine learning

To join the revolution, it's important to understand where it all began.

If we really wanted to get geeky, we could go back to 1600 BC and talk about the origin of algorithms in Babylon. Or to 1763, when the Bayes theorem was introduced, which describes the probability of an event based on prior knowledge or conditions related to that event. But that's not our main goal here.

Instead, let's flash forward to 1952. Arthur Samuel joined IBM's Poughkeepsie Laboratory and started working on a program that would revolutionize the world. He designed and tested two machine learning procedures that enabled a computer to learn to play checkers better than the person who wrote the program. This was a breakthrough moment. Samuel realized that instead of having to instruct a computer, he could instead provide certain parameters, and allow the computer to learn through trial and error.

From that moment on, the technological innovations and discoveries of machine learning came more frequently: the Stanford Cart in 1979; explanation-based learning in 1981; the application of back propagation in 1986; and the victory of IBM's Deep Blue over then-worldwide chess champion, Garry Kasparov, in 1997.

The sheer amount of research and dedication spent on exploring the increasing functionality of machine learning has brought us to where we are today: An interconnected "smart" world that feeds off the constant firehose of data we put into it.

As the internet was adopted widely in the '90s and 2000s, all of that data started to become a commodity. And now, Domo reports that over 2.5 quintillion (1030) bytes of data are created every single day.[1] All of our transactions, likes, comments, and favoriting of cat videos will mean that by 2020, 1.7MB of data will be created every second for every person on earth. (That's a lot of cat videos!)

Add to that the non-stop advances in computer technology, like increased memory and more powerful processing capabilities, and there's no stopping this technology.

[1] **DOMO,** *2017, Data never sleeps 6.0*

Are machine learning and AI the same?

It's a common misconception that they are ...

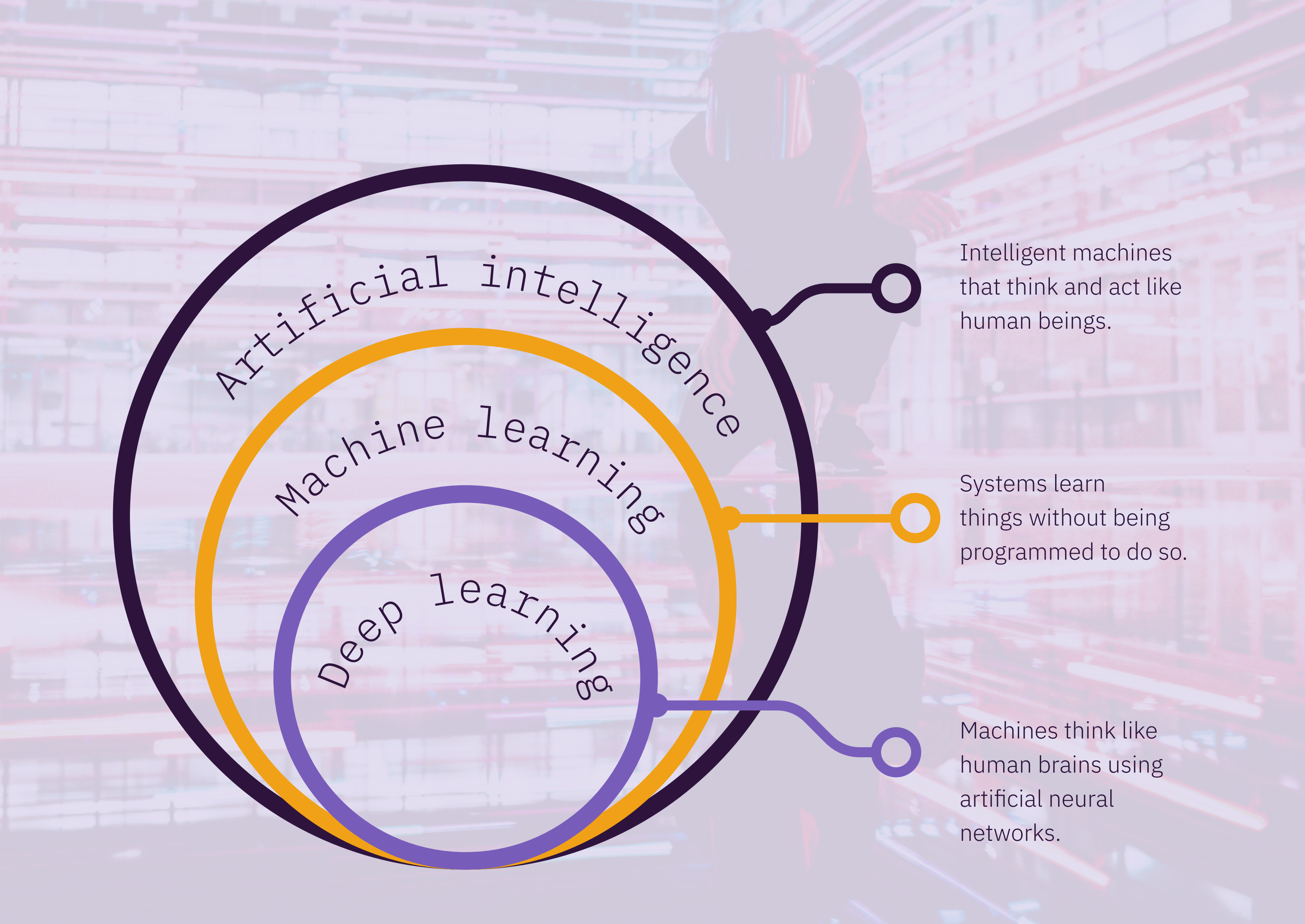
It's common for the terms machine learning and AI to be used interchangeably—but they're not the same thing. Machine learning is a subfield of artificial intelligence (AI). And while all machine learning counts as AI, not all AI counts as machine learning. Confusing? Let's break this down.

Advances in computing power have increased the ability of neural networks to process more and more data, meaning that a computer can now learn beyond what a programmer has told it.

We briefly called out IBM's Deep Blue, the first machine to beat a reigning world chess champion in a six-game match. This was deemed to be a major milestone in the world of AI. But was it machine learning?

Deep Blue relied on state space search, the number of legal game positions reachable from the initial position of the game. Researchers input data about the game of chess and any given board position, and the supercomputer found the best possible board position that would win. Using a wise optimization strategy and a fast computer, the system was able to execute millions of calculations per second to win the chess game—all based on the data input by the researcher.

Researchers took the same technology from Deep Blue and tried to apply it to solving Go, a far more complex game than chess (there are 10^100 times more possibilities than in chess). This near-incomprehensible number of possibilities made it hard for a computer to play, and left AI researchers perplexed.



Intelligent machines that think and act like human beings.

Systems learn things without being programmed to do so.

Machines think like human brains using artificial neural networks.

FIGURE 1: INTEGRATED RISK & CONTROL MODEL

They had to take a different approach. By using a combination of advanced search trees (the way the computer would locate key data within a dataset) with deep neural networks (modeled on the human brain and nervous systems), AlphaGo was able to beat Lee Sedol, one of the best Go players in the world.

So in the case of Deep Blue, the computer required detailed input of all gameplay possibilities, and would calculate ideal outcomes based on moves. AlphaGo used its policy network to select a move, and its value network to predict a winner. It was exposed to numerous amateur games to understand how people play, and played against different versions of itself thousands of times to learn from its mistakes and become better at decision-making.

By using a common risk language across departments and with individuals in all three lines of defense, an auditor can truly evaluate the effectiveness of a cybersecurity program and get an accurate picture of where the organization stands.

A risk-based approach also lets internal audit meet expectations set by the board and identify major tactical and strategic gaps in cybersecurity governance.

Supervised & unsupervised learning

Machines can learn in one of two ways: supervised or unsupervised.

SUPERVISED LEARNING

The majority of practical machine learning uses this approach. With supervised learning, a trainer gives the system an input (X) and desired output (Y); Y = f(X).

Feedback is given to the machine until it learns to consistently give the correct output and get more accurate with its predictions.

Everyday examples of supervised learning:

- Training your email inbox to classify mail as spam or safe.
- A real estate app predicting the price of a house in a specific city.
- Bank systems tracking your spending patterns and flagging fraudulent activity.

UNSUPERVISED LEARNING

With unsupervised learning, there is no trainer—no feedback saying “you got it correct, or you got it wrong”—the machine just uses the data it has.

Everyday examples of unsupervised learning:

- Identifying values in large data groups that are distinctly different from other values.
- Identifying differences and grouping like data values into logical groups.

Common machine learning applications

Now for the real reason we're all here—practical applications of this amazing technology in the world of governance. Because as regulations change, emerging technologies multiply cybersecurity threats, and fraudsters become increasingly sophisticated, there's no better time to embrace machine learning than the present.

So, how do organizations, who are creating and collecting massive amounts of data, apply machine learning to help streamline workflows, protect the organization, and meet their objectives? Let's take a look at three areas of governance, risk, and compliance (GRC) where machine learning can make a big impact.

01 Fraud detection

PwC's Global Economic Crime and Fraud Survey 2018 found that 49% of the 7,200 companies surveyed had some kind of fraud.[2] It's actually not a surprising figure when you consider that fraud transactions:

- Are usually so minor, they're often easily overlooked until they become huge losses.
- Are still often manually reviewed. According to the CyberSource Fraud Benchmark Report,[3] 83% of North American businesses conduct manual reviews.

And machine learning has proven to be great at detecting and preventing fraud in banking, finance, commercial businesses, and even state and local governments.

It makes large data sources more manageable by clustering like data points together (e.g., like when you've got 100,000 P-Card transactions to go through). Based on criteria that you input, your software will assign a rating and group the transactions accordingly.

DETECTING & ELIMINATING FALSE POSITIVES

In the graph on the right, let's say the green dots represent the low- or no-risk transactions, blue the medium risk, and red high risk. Your software would take the red transactions and place them into automated workflows for remediation. The system will "watch" for patterns in the remediation to see which of the red dots turn out to be true fraud. If instead they're false positives, the software will learn from that information. This reduces the false positives over time, and helps surface only those transactions that are genuinely fraud.

Identifying high-risk P-Card transactions is a clearly defined, relatively easy-to-execute, and measurable task—ideal for machine learning.

[2] **PwC,** *2018, Pulling fraud out of the shadows: Global Economic Crime and Fraud Survey 2018*

[3] **CyberSource,** *2016, Annual fraud benchmark report: A balancing act*

Case study

Visa Advanced Authorization uses machine learning to detect fraud in real time

Visa has been using machine learning to combat fraud since 1993. And that's a good thing, because its global network, VisaNet, processed more than 127 billion transactions in 2018 alone!

PROBLEM

Prior to adopting machine learning, Visa relied on cashiers to manually search through a massive book of stolen cardholder account numbers, or pick up the phone to get a verbal authorization from a call center. This was an extremely slow and inaccurate process.

SOLUTION

"Visa was the first payments network to apply neural network-based AI in 1993 to analyze the riskiness of transactions in real time, and the impact on fraud was immediate," said Melissa McSherry, SVP and global head of credit and data products, Visa.[4]

The Visa Advanced Authorization starts when a transaction is initiated. The model reviews and analyzes the data, looking for more than 500 specific risk attributes (e.g., the amount, location, PIN security).

A score is then instantly generated between one and 99. That score determines how likely it is that the transaction is fraudulent (one is low risk, 99 is high risk). The score is then sent to the account holder's financial institution to either approve or reject the transaction. All of this happens in seconds or less.

By using machine learning to score and route transactions, Visa announced that it was able to prevent an estimated $25 billion in annual fraud.[5]

[4] **PaymentsJournal,** *2019, Using artificial intelligence, Visa is combating fraud at nearly the speed of light*

[5] **Visa,** *2019, Visa prevents approximately $25 billion in fraud using artificial intelligence*

02 Compliance management

Since the 2008 global financial crisis, the responsibilities of compliance teams have increased. If you have anything to do with compliance, you know that keeping up with changing standards and regulations is a never-ending struggle. And when you have to demonstrate compliance to multiple regulations like ISO or GDPR, each with different requirements, creating and modifying controls as these requirements change can leave you constantly scrambling.

Supervised machine learning algorithms can be used to automatically compare data about rules. Using this data comparison, compliance managers can map controls to regulations more efficiently, saving huge amounts of time and preventing major headaches—not to mention potential financial penalties.

But machine learning doesn't just help you keep on top of regulatory updates. It also helps analyze huge quantities of both structured and unstructured data to perform comparisons and make decisions. This is a huge benefit for compliance professionals, who find themselves drowning in an increasing sea of data.

WAYS THAT THESE MACHINE LEARNING ALGORITHMS CAN HELP

+ Clustering regulation requirements into similar groups and suggesting controls that satisfy multiple requirements.
+ Automatically identifying contracts that are affected by regulatory changes (e.g., vendors under GDPR requirements).
+ Automating how data is classified for retention or disposal.
+ Although human oversight is still very much required in compliance management, the speed and accuracy of machine learning makes it possible for compliance teams to more efficiently meet regulatory requirements and interpret data.

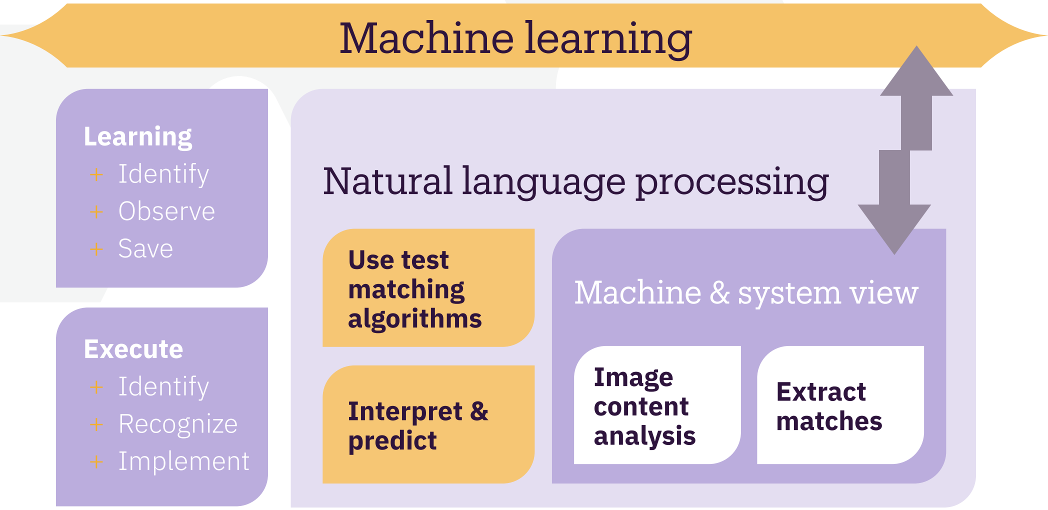


FIGURE 2: MACHINE LEARNING IN COMPLIANCE MANAGEMENT

03 Risk assessments

By analyzing large datasets in short times, machine learning is changing the way risks are assessed. The following are just some examples of how machine learning can be used in risk management.

DETERMINING CREDITWORTHINESS

Machine learning can help lenders determine the creditworthiness of potential borrowers by examining datasets like their digital footprint. This has become more common for evaluating borrowers with little or no credit history—like young adults, or 500 million people in China.

Startups like Lenddo, UpStart, and ZestFinance use machine learning in their systems to examine alternative data like social media usage, search engine browsing history, SAT scores, and GPAs to help predict whether a borrower is likely to pay back a loan. These companies use this data to generate a more accurate credit score that they can share with financial institutions. A study at MIT found that these applications of machine learning could reduce bank losses due to delinquent customers by up to 25%.[6]

IDENTIFYING OPERATIONAL RISKS

Operational risk is present in every organization, from a small business to a global corporation. Here are a few ways machine learning can help with operational risk.

CYBERSECURITY THREATS

Humans can't possibly sort through millions of files to identify potential cybersecurity risks. Machine learning can use statistical analysis and algorithms to stop threats before they cause damage. Proofpoint's MLX technology uses machine learning to protect itself against spammers by analyzing the language in millions of messages to detect potential threats.

MONEY LAUNDERING ATTEMPTS

The cost of anti-money laundering (AML) compliance is estimated at $23.5 billion per year in the US, while European banks spend 18.3 billion euros annually.[7] By using machine learning clustering techniques that can classify transactions based on how suspicious they are, or even find people with similar behaviors working together to commit a crime, money laundering attempts can also be uncovered.

[6] **Journal of Banking & Finance,** *2010, Consumer credit-risk models via machine-learning algorithms*

[7] **Business Recorder,** *2019, AI brings down AML cost*

CUSTOMER COMPLAINTS

Chatbots can recognize why a customer is messaging them based on previous contact patterns, find similar cases to recommend a resolution based on what has worked in the past, and escalate issues to the right person at the right time.

ALLOCATING RESOURCES

By using past data to project transactions from one period to the next, risk managers can help determine where to direct resources. Forget about the labor-intensive work of manually collecting and reconciling data—machine learning allows risk managers to automatically predict which branch locations are likely to fail an audit, and which are likely to pass, and only focus efforts on locations that need more attention.

MODELING SCENARIOS

Once a machine learning model is set up for a dataset, risk managers can alter input data to find out what impact it might have on predicted outcomes (e.g., how it might increase or decrease risk scores). Machine learning can explore a multitude of models, allowing GRC professionals to make predictions and continue to repeat and refine them, based on results.

REMOVING SUBJECTIVE RISK SCORING

A major benefit machine learning brings to risk management teams is the ability to remove the subjective scoring of risk. By feeding data into systems—and using a model to determine data-driven risk scores—you can avoid the manual and human process of risk scoring, which is often inaccurate.

Getting started with machine learning

So, ready to jump into the world of automation and machine learning? Before you do, there are a few things you should keep in mind …

DATA ACCURACY

Obviously, the accuracy of your data is essential in any machine learning project; outliers, noise, and missing values could render your results meaningless. Regularly testing and validating the model is a best practice that your organization needs to adopt.

DATA BIAS

Do you have suitable data? Machine learning models are only as good as the data that you feed them. So, if your data is skewed, you won't get the most from your efforts, or you could face legal issues (e.g., if you're a bank and you use data points like race, gender, or religion to refuse or approve credit).

CHOOSE THE RIGHT MODEL

From linear, to regression, to random forest, you have some decisions to make when it comes to machine learning algorithms. Your decisions depend on several factors, including your business goals, the scalability of the model, and the complexity of your data.

CLEARLY DEFINE GOALS & OBJECTIVES

What problems are you trying to solve? Before you implement machine learning in your organization, evaluate which processes require it—not all automated processes need machine learning. Your company should have specific use cases in mind for machine learning to ensure it provides value.

> "Essentially, all models are wrong, but some are useful."
>
> » ***George Cox, father of modern statistics***

HighBond's machine learning capabilities

Galvanize's HighBond platform includes easy-to-use, built-in machine learning capabilities—so you don't need to be a data scientist to uncover hidden patterns or anomalies.

That said, if you happen to be a data scientist and want those advanced capabilities, our data automation tool, ACL Robotics, fully integrates with R and Python to make that possible.

The machine learning commands in ACL Robotics help you uncover the blind spots that rules-based analytics can't:

- Take advantage of AI with machine learning algorithms using unsupervised learning.
- Perform enhanced analytic capabilities to find patterns, anomalies, and predict categories in your data that you didn't know to look for.
- Watch as our clustering command identifies patterns within your dataset without having to tell the analytic what to look for.
- Demonstrate results easily.
- Create a threshold where if exceptions are likely to be a false positive, they are automatically eliminated from the analytic workflow.

Summary

Machine learning isn't a new concept. A lot of time, effort, and research has gone into developing and building it, and it's adapted over time to be primed for our current digital environment.

Around the world, data continues to grow, making now the ideal time to turn to machine learning to help manage and mine it.

Machine learning elevates existing processes and systems. Automation streamlines work, improves your resource allocation, and frees up your staff to focus on the things that require real human attention.

Remember: It doesn't have to be about humans versus machines; both can work together for the most effective and efficient outcome.

Further learning & resources

There are endless resources on machine learning, AI, and how they relate to GRC functions. Here are some we thought you might like to dig into.

MACHINE LEARNING FOR FRAUD DETECTION

Everything you need to know about models, neural networks, risk scores, thresholds and adding human insight.

https://www.ravelin.com/insights/machine-learning-for-fraud-detection

ROBOTIC PROCESS AUTOMATION (RPA) VS. AI, EXPLAINED

What is the difference between RPA and AI? How do RPA and AI work together? How does machine learning fit in? What are some RPA and AI use cases and best practices?

https://enterprisersproject.com/article/2019/8/rpa-robotic-process-automation-vs-ai-explained

MACHINE LEARNING & AI FOR RISK MANAGEMENT

Explore how machine learning and AI solutions are transforming risk management.

https://link.springer.com/chapter/10.1007/978-3-030-02330-0_3#Fn1

HOW AI SHAPES THE NEXT GENERATION OF DATA & COMPLIANCE PART 1: SECURING SENSITIVE DATA

https://blog.netapp.com/how-artificial-intelligence-shapes-the-next-generation-of-data-and-compliance-part-1-of-3

AI & MACHINE LEARNING IN FINANCIAL SERVICES COMPLIANCE MANAGEMENT: USE CASES FOR FINANCIAL INSTITUTIONS

https://www.finextra.com/blogposting/16050/ai-and-ml-in-financial-services-compliance-management-use-cases-for-fis

Let us help you take advantage of the power of machine learning.

↳

To find out how Galvanize can help your organization automate critical processes, deliver the answers that drive strategic change, and improve your bottom line, call 1-888-669-4225, email info@wegalvanize.com, or visit **wegalvanize.com**.

Galvanize

ABOUT GALVANIZE

Galvanize delivers enterprise governance SaaS solutions that help governments and the world's largest companies quantify risk, stamp out fraud, and optimize performance.

Our integrated family of products—including our cloud-based governance, risk management, and compliance (GRC) solution and flagship data analytics products—are used at all levels of the enterprise to help maximize growth opportunities by identifying and mitigating risk, protecting profits, and accelerating performance.

wegalvanize.com

©2020 ACL Services Ltd. ACL, Galvanize, the Galvanize logo, HighBond, and the HighBond logo are trademarks or registered trademarks of ACL Services Ltd. dba Galvanize. All other trademarks are the property of their respective owners.