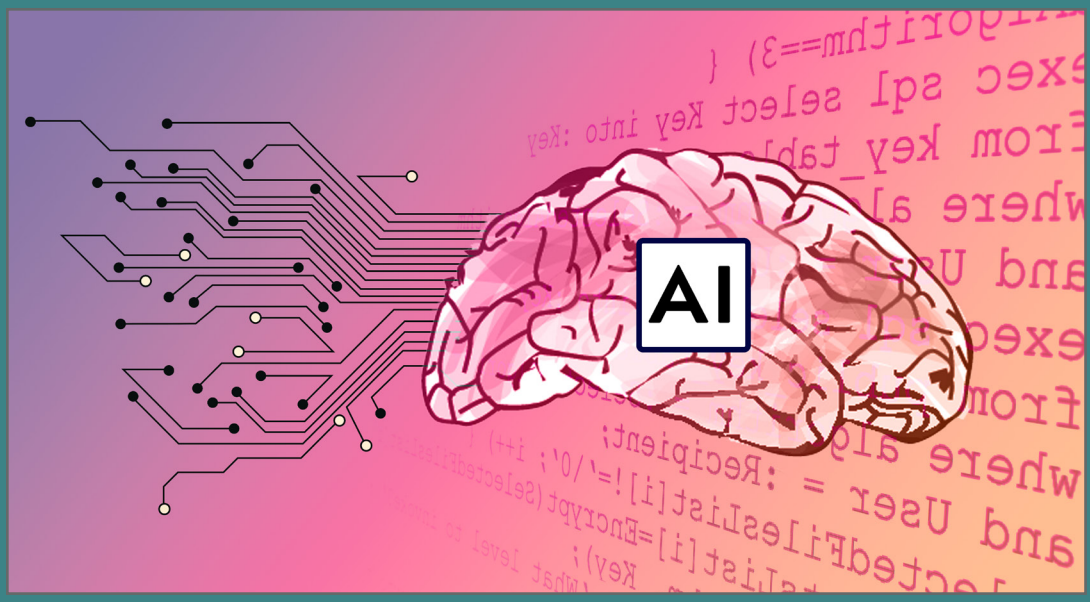

HARNESS THE POWER OF AI

WHILE AVOIDING ITS PITFALLS

Artificial intelligence is changing how organizations work – and it's changing people's lives. It's bringing about major advances in healthcare, helping people recover faster and live longer. It's making our lives safer, preventing accidents on the road and in the factory. And because it handles essential but arduous tasks, it gives us the freedom to focus on other things.

But artificial intelligence has its potential dark side as well. If it isn't designed and deployed properly and responsibly, it can make problems worse instead of resolving them. If biases are allowed to slip into the system, entire segments of people, patients, or customers can be negatively impacted. Outcomes can suffer, and real people can feel real-world consequences.

How can your organization harness the power of AI while avoiding its pitfalls? By practicing responsible AI. In this white paper, we'll explore:

- What responsible AI is, and the role data provenance plays in it
- The four pillars that form the foundation of responsible AI
- The data biases you should avoid when designing and deploying AI
- How organizational data governance can help you most effectively use AI

WHAT IS RESPONSIBLE ARTIFICIAL INTELLIGENCE (AI)?

Artificial intelligence (AI) creates an unparalleled opportunity for businesses and organizations, allowing them to process and make decisions based on massive amounts of data – with an eye on improving outcomes in the process. But as the use of AI has risen, so too have concerns. What information is being captured? How is it being used? How can we be sure we use it responsibly?

Dr. Kendall Giles, Bradley Department of Electrical and Computer Engineering Collegiate Assistant Professor at Virginia Tech, outlines the problem of responsible AI as follows:

> *"Increasingly, we are seeing more artificial intelligence systems fail because of 'black-box mechanisms, social discrimination, security vulnerabilities, privacy harms, and the decay of system quality over time.' This is troubling, because machine learning, artificial intelligence, and automation are becoming so pervasive all around us. At a time when our society is becoming increasingly dependent on these systems, their potential for causing harm in society is likewise increasing."*

RESPONSIBLE AI

So what exactly is responsible AI, and how can organizations practice it? Computer science researcher Virginia Dignum defines responsible artificial intelligence as" Human responsibility for the development of intelligent systems along fundamental human principles and values, to ensure human flourishing and wellbeing in a sustainable world." And putting it into practice, says Dr. Giles, is about "designing systems that are Fair, Accountable, Transparent, and Explainable (FATE)." We'll explore this concept more momentarily.

THE ROLE OF DATA PROVENANCE

The proliferation of tools and technologies that create and collect data – from smart speakers to social media – have only muddied the data quality waters, jeopardizing the ability of companies to deploy AI responsibly. As the number of data channels has increased, data quality has not experienced a corresponding rise. Today's organizations are beginning to realize that the problem isn't the amount of data gathered but the relevancy and quality of that data and its ability to help them make informed decisions.

That's where data provenance comes in. Data provenance provides a constantly-updated record showing where the data originated and how it's been processed. Because the data's origin and lifecycle are clear and easily accessible, organizations can feel more confident making decisions based on that data. By prioritizing data provenance, you'll not only be practicing responsible AI. Still, you will be improving the fairness, accountability, transparency, and explainability (or FATE) of the insights your AI shares, boosting trust and confidence.

THE FOUR CHARACTERISTICS RESPONSIBLE ARTIFICIAL INTELLIGENCE BUILDS FROM

"To evaluate the quality of and trust in data, we need not only to understand the data context but also to trace and record the origins and processing of data. In other words, we need to know where data comes from, how it is collected and how it can best be used. This task ensures the reliability and quality of data and is called data provenance."

–Dr. Nektaria Tryfona, Bradley Department of Electrical and Computer Engineering Collegiate Associate Professor, Virginia Tech

According to "Establishing Data Provenance for Responsible Artificial Intelligence Systems," four characteristics form the foundation of responsible artificial intelligence.

Fairness: AI systems are programmed by humans, meaning they are not immune to bias. For example, feeding AI training data that only includes information about Caucasian patients would lead to an algorithm that discriminates against non-white minorities.

Accountability: Because AI gathers data from numerous sources and has many moving parts, it's often hard for end users to determine who ultimately holds responsibility for the results. For example, training a self-driving car's AI system on data from India would lead to potentially disastrous recommendations for self-driving vehicles in the U.S. – and determining accountability is nearly impossible.

Transparency: By their very nature, AI systems are often highly opaque – the "magic" happens behind an impenetrable curtain. But organizations need transparency to feel confident in the system's recommendations. Remember how your fifth-grade math teacher always insisted you "show your work" so she could see how you arrived at a conclusion? Your AI system should do the same.

Explainability: For AI to be truly effective, users must understand how and why the system makes specific predictions and recommendations. If users cannot understand how the AI came to a particular conclusion, they are less likely to accept and trust it.

HOW A FATE ASSESSMENT HELPS TO EFFECTIVELY DEPLOY AI

> *"To effectively build and deploy AI systems, we need to 'train' Machine Learning (ML) algorithms using various data sets, i.e., initially, ML algorithms understand data entities and their relationships and then they identify patterns, explore clusters and provide insights."*
>
> *–Dr. Nektaria Tryfona, Bradley Department of Electrical and Computer Engineering collegiate assistant professor, Virginia Tech*

As we've found, FATE forms the foundation of deploying AI responsibly. But it does more than that – performing a FATE assessment helps your organization pinpoint and eliminate data biases, reaping all of the benefits AI offers.

THE EFFECT ON DATA BIASES WHEN DEPLOYING RESPONSIBLE AI

Artificial intelligence and machine learning offer businesses and organizations the ability to process massive amounts of data – much more than they ever could in the past. But AI isn't immune to bias. Both programming and data sources can potentially introduce bias into the algorithm, leading to real-world impacts on humans – from being denied a mortgage loan to being discharged from the hospital.

To deploy AI responsibly at your organization, ensure that you avoid inadvertently introducing bias into your process. Specifically, be mindful of these common data biases:

POPULATION DATA

A population data bias occurs when you survey the wrong people or gather data from the wrong audience. An example of this would be developing an algorithm using consumer data from one country, then using that algorithm to predict shopping behavior in another country. Attempting to use data created for one context in another context without retraining the system is likely to lead to poor outcomes.

MEASUREMENT ERROR

Measurement errors occur when reported data differs from the actual data. One example of a measurement error may be data gathered from a survey that asks patients to identify how many 1:1 interactions they've had with their care coordinator in the past year; the number they report may or may not reflect reality. Training an AI system with data that suffers from measurement errors will deliver imprecise results and recommendations.

DATA QUALITY CHASM

A chasm in data quality can create challenges for AI, ultimately impacting the quality of the recommendations it delivers. For example, training an AI algorithm using images taken using the latest iPhone camera, then retraining it using photos taken with an inferior camera would likely lead to inaccurate results and recommendations.

DATA REPURPOSING

AI systems regularly use data originally captured for a different purpose – for example, using the results of a patient's heart enzyme test to identify other potential health concerns. But because it wasn't explicitly collected for that purpose, the data may be missing valuable information necessary to make an informed decision.

DATA AUGMENTATION

Artificial intelligence craves data; the more, the better. So when your available data is too small for your AI to process, you may opt to augment that data by slightly modifying your existing data or using synthetic data. However, data augmentation tends to exacerbate any existing biases, making it more difficult for the AI algorithm to deliver fair, accountable recommendations.

REASONS FOR ESTABLISHING ORGANIZATIONAL DATA GOVERNANCE

What is organizational data governance? Simply stated, it's a clearly-defined, standardized approach to how the organization will strategically manage data. With solid data governance in place, organizations can better understand and even improve their data, increasing its value and decision-making power. And data provenance plays a key role; after all, it's imperative to understand the inputs and processes that impact data, providing context.

There are many reasons for establishing data governance at your organization:

Provide transparency: There's a reason police track the chain of custody regarding evidence handling – it proves that the evidence hasn't been inadvertently or intentionally tampered with at any point in the process. Similarly, organizational data governance allows you to track and view data at every stage of the process, allowing you to review and understand the recommendations your analytics or AI makes. Transparent data is also much easier to audit.

Enhance data reliability: If you can't pinpoint where the data came from, how can you trust that it's accurate? Making decisions based on data with an unknown provenance is a risky move at best. Capturing and timestamping data sources is a great way to address data provenance and ensure the accuracy of your data.

Ensure you're using the right data for the job: Remember the old edict, "garbage in, garbage out"? It holds true when discussing data collection and artificial intelligence. Your analytics and AI can only be effective if the inputs they're receiving are relevant, accurate, and up-to-date. Organizational data governance ensures you use the right data for your particular situation.

Capture the whole story, not the "averaged" version: Just because your data is accurate doesn't mean it's appropriate for your needs. In some cases, you may be dealing with averaged data, which can wreak havoc on AI and analytics insights. Organizational data governance allows you to avoid the "flaw of averages" and make truly informed decisions.

POSITION YOUR ORG FOR AI SUCCESS

Artificial intelligence has the power to significantly alter our lives for the better – if we do it right and responsibly. To capture all of AI's benefits while avoiding its pitfalls, design AI systems that are fair, accountable, transparent, and explainable. Prioritize and practice data governance across your organization. And avoid inadvertently adding biases into your AI system. By following these best practices, you'll be setting your organization up for AI success.

References:

[1] Establishing Data Provenance for Responsible Artificial Intelligence Systems, K. Werder, B. Ramesh, R. Zhang, ACM Transactions on Management Information Systems, Vol. 13, Issue 2, June 2022, Article No.: 22, doi. org/10.1145/3503488.

[2] The Seattle Report on Database Research, D. Abadi et. al., Communications of the ACM, August 2022, Vol. 65, No 8. DOI:10.1145/3524284.