RSAConference2020

San Francisco | February 24 – 28 | Moscone Center

HUMAN ELEMENT

SESSION ID: **MLAI2-F01**

Does Artificial Intelligence Need a General Counsel?

Alan Brill

Senior Managing Director
Kroll Cyber Risk

Stacy Scott

Managing Director
Kroll Cyber Risk

Agenda

- Where Is Artificial Intelligence (AI) Used In Business?
- AI's Popularity
- What Could & Has Gone Wrong?
- How To Course Correct
- Key Takeaways

Where Is AI Used in Business?

- MOBILE APPLICATIONS
- CUSTOMER EXPERIENCE
- SUPPLY CHAIN
- HR
- FRAUD DETECTION
- RESEARCH & DEVELOPMENT
- RISK MANAGEMENT & ANALYTICS
- PRICING & PROMOTION
- OPERATIONS MANAGEMENT

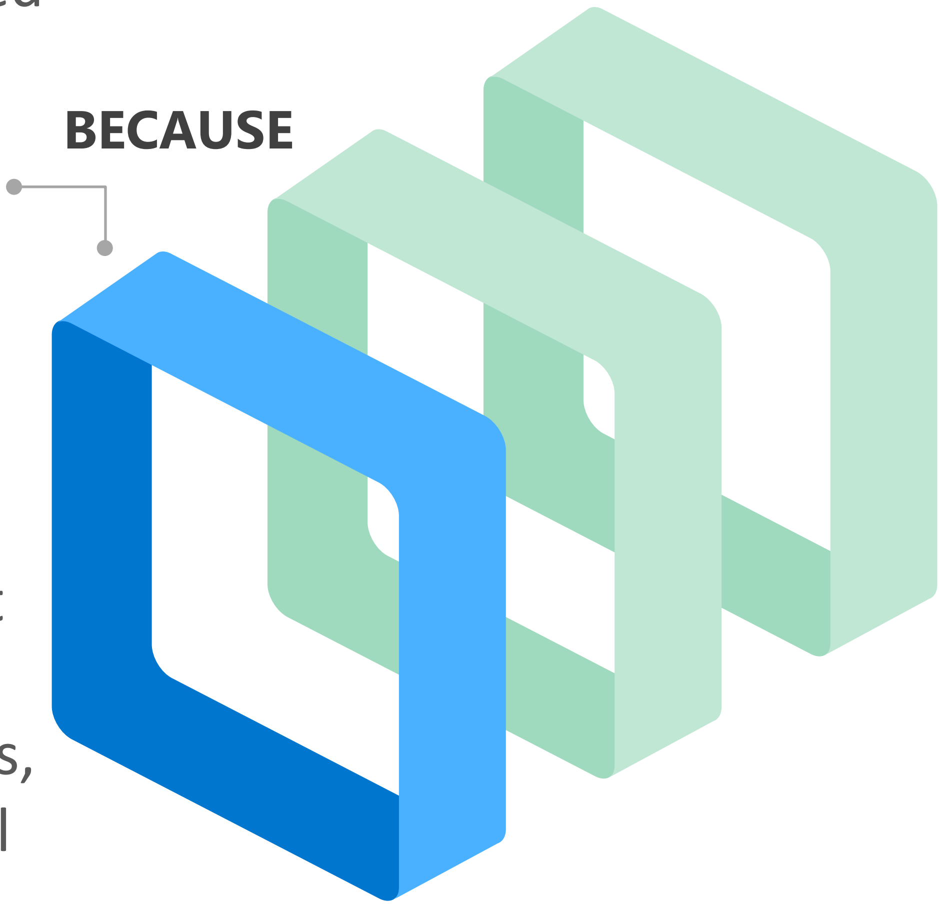
And Its Use is Only Getting Bigger…

- By 2030, the average simulation shows that some **70% of companies** might have **adopted** at least one type of **AI technology**.[1]
- The **AI market** will grow to a **$190 billion** industry by 2025.[2]
- **83% of businesses** say **AI is a strategic priorit**y for their businesses today.[3]

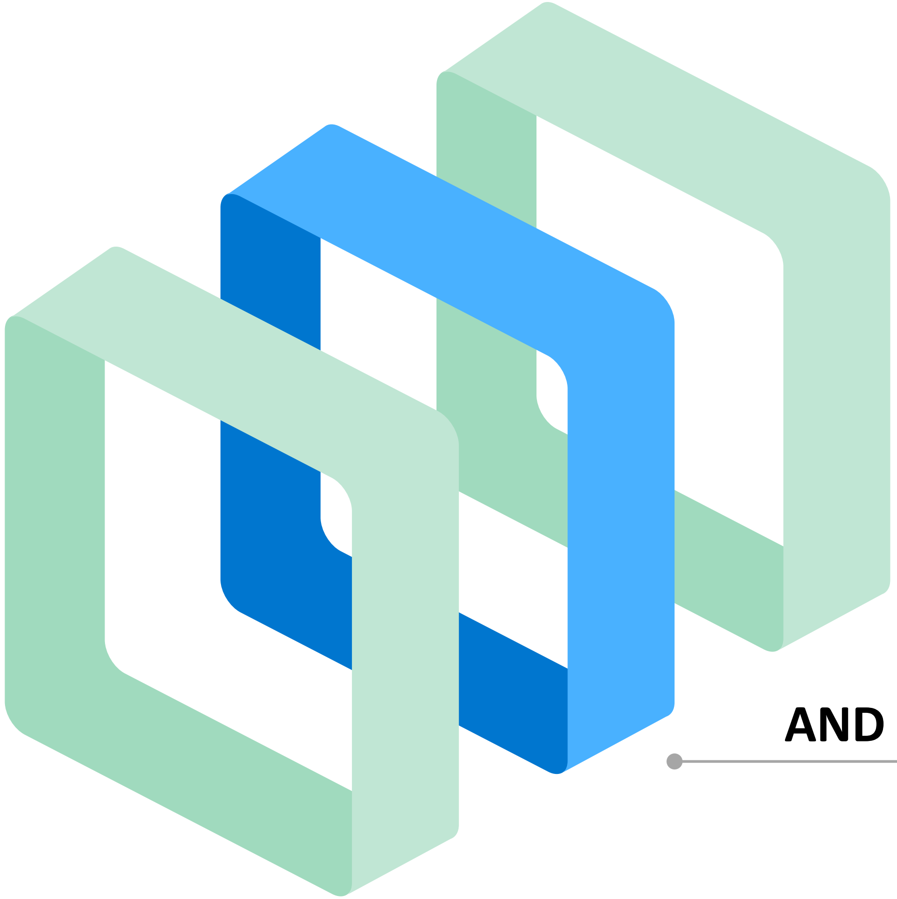
[1] mckinsey.com [2] marketsandmarkets.com [3] forbes.com

A Starting Point…Why Build AI?

- Traditional computer programming follow defined rules (i.e. it does what it does today and will do the same tomorrow)
- Basic Tenet of Compliance and testing
- Corporate counsel believe they are tech savvy but acknowledge their comfort level and confidence with technology have limitations, specifically around artificial intelligence (AI).



A Starting Point…Why Build AI?



AND

- “Everything we love about civilization is a product of intelligence, so amplifying our human intelligence with artificial intelligence has the potential of helping civilization flourish like never before – as long as we manage to keep the technology beneficial.“
- - Max Tegmark

A Starting Point…Why Build AI?

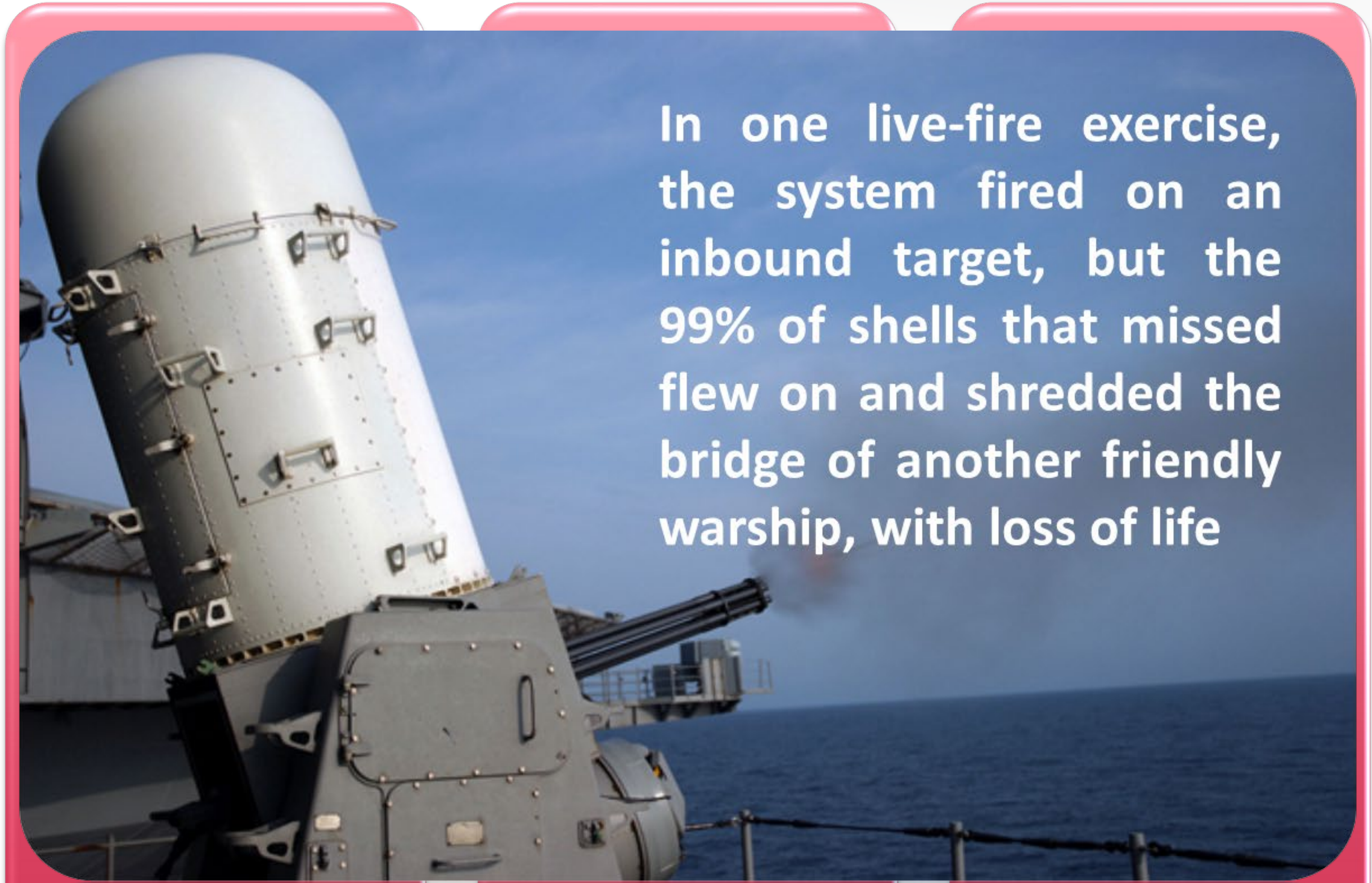
One of the Goals of AI & Deep Learning (DL) = Avoid Avoidable Problems

THEN

- AI rules will evolve and rules will change
- We should NOT build AI to be able to say we built it, or we have them
- The competition has one, so we need one too – and right now schedule the AI/DL timeline
- But rather, because we want to avoid avoidable problems and AI technology should be fair

What Could & Has Gone Wrong?

Autonomous weapons: AI programmed to do something dangerous



In one live-fire exercise, the system fired on an inbound target, but the 99% of shells that missed flew on and shredded the bridge of another friendly warship, with loss of life

Phalanx CIWS

Discrimination

⚠ What Could & Has Gone Wrong?

OPERATIONS MANAGEMENT

Consider a U.S. AI/Deep Learning System Used by a Financial Services Company to Make Decisions About Loans….

- AI System “training” provided 250,000 records of prior loans
- AI software analyzes factors related to the likelihood of successful repayment

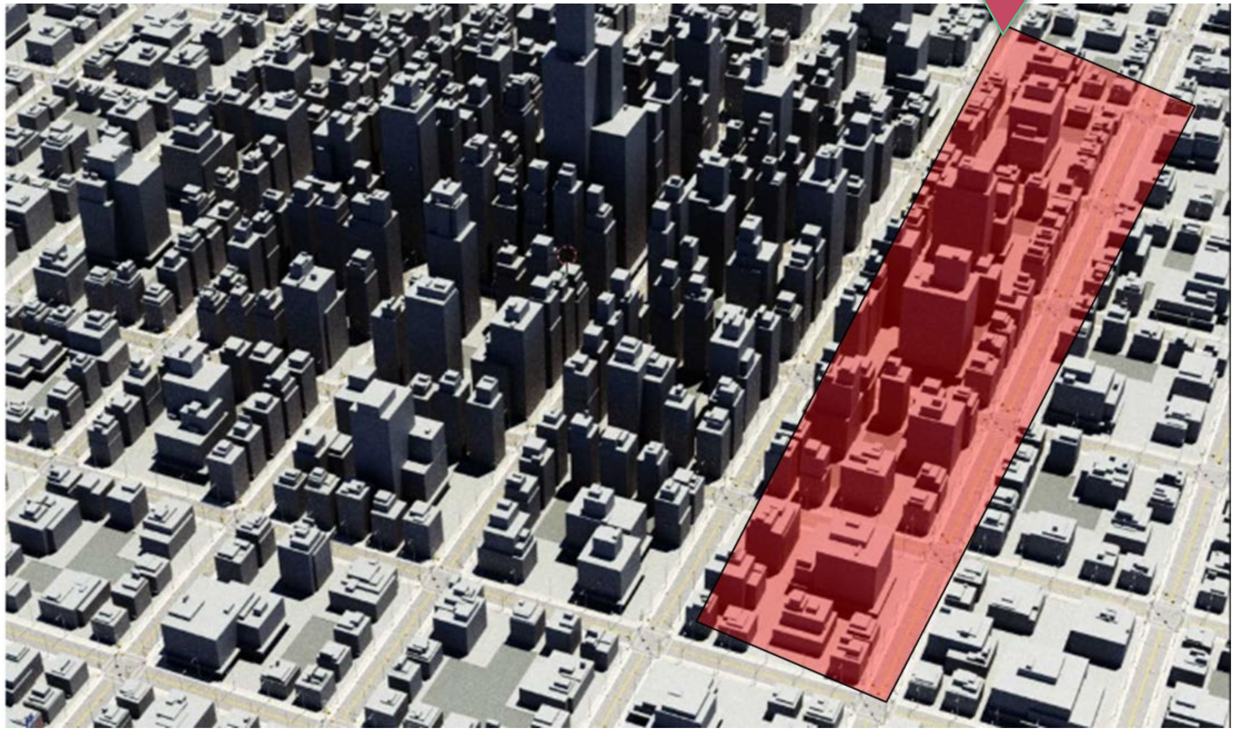
What Could & Has Gone Wrong?

OPERATIONS MANAGEMENT

Consider a U.S. AI/Deep Learning System Used by a Financial Services Company to Make Decisions About Loans....

POSTAL CODE OUTLINED IN RED STATISTICALLY LESS LIKELY TO REPAY A LOAN TO THE BANK

- One learned approval/denial factor is postal code
- AI now places greater weight in loan approval/denial based on where the borrower lives

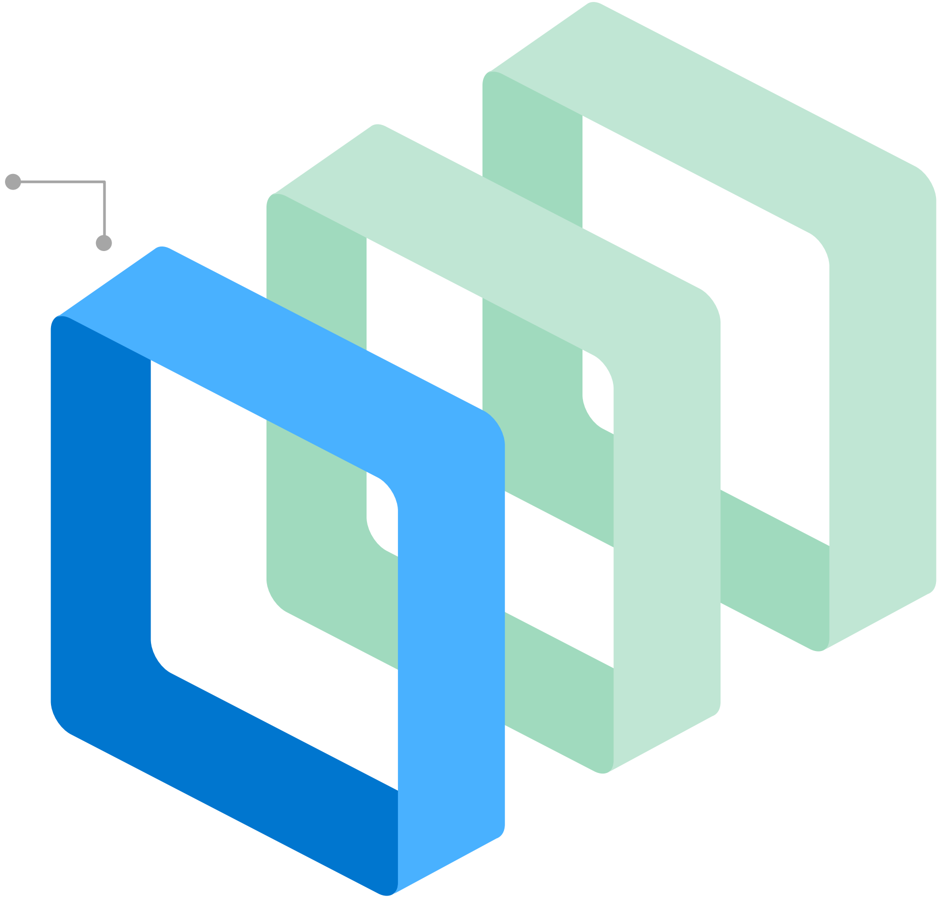


What Could & Has Gone Wrong?

The *Community Reinvestment Act* of 1977 made it illegal to base lending decisions on the neighborhood where a person lives.

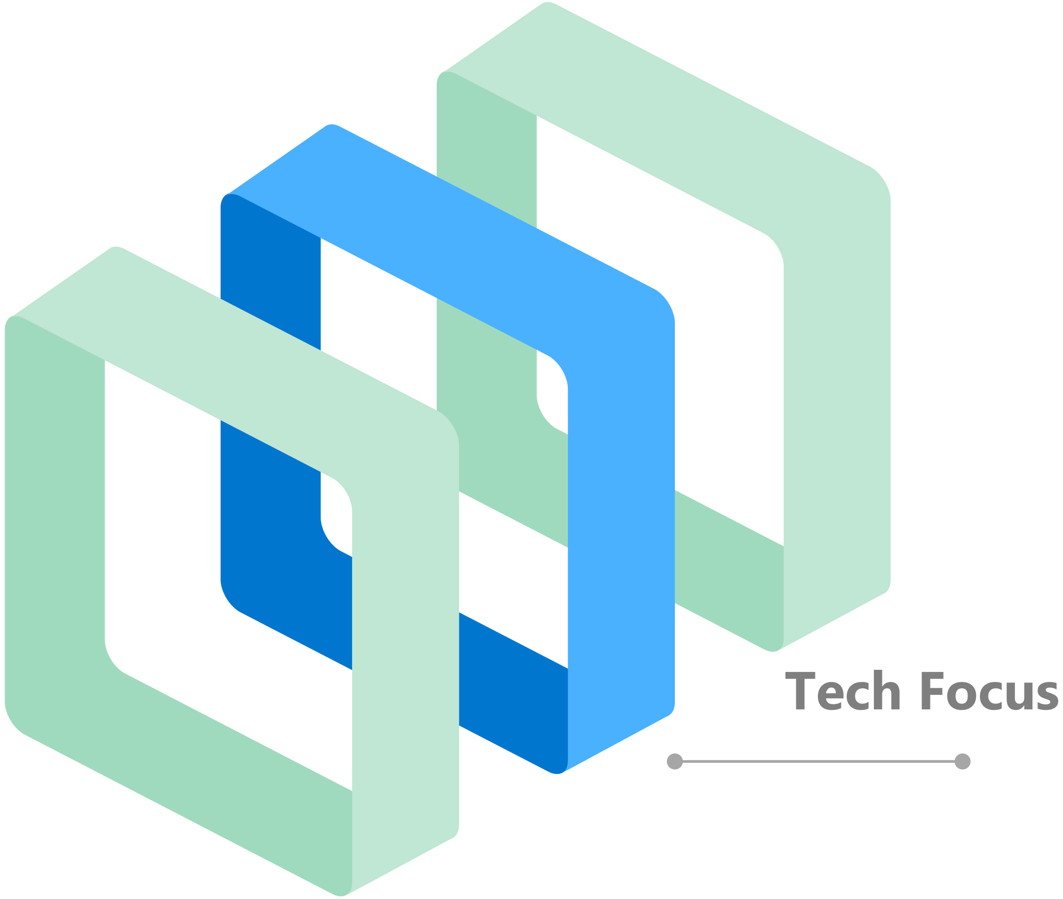
Without it, financial institutions literally drew red lines on maps around minority communities and denied them access to services.

⚠ Why does this happen?



Organization may hire AI experts ***but not necessarily subject matter experts***.

⚠ Why does this happen?



Focused on building (e.g., make loan decisions) & **never think about legal or regulatory issues**

⚠ Why does this happen?

Who’s Going To Tell Them?

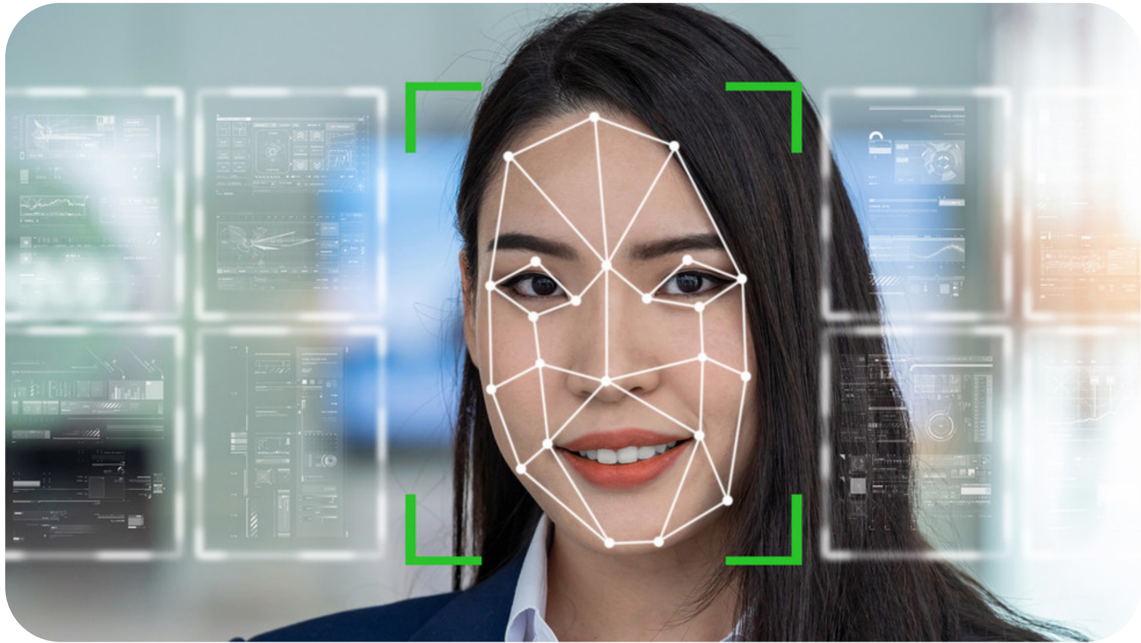
- Can you **afford to hope** that **they figure this out** or research it themselves?

But Wait…There's Another Danger

> **Bias can creep in at many stages of the deep learning process, and the standard practices in computer science aren't designed to detect it.**
>
> MIT Technology Review in February 2019

⚠ An example of implicit bias…

- A **facial recognition system used by police** was **99% accurate** with **white males.** It was **substantially less accurate in** identifying **women or people of color**.
- This turned out to be because the enormous **training set used** for the system **consisted mostly of photographs of white males**.



This Photo by Unknown Author is licensed under CC BY

⚠ Another example of implicit bias…

This Photo by Unknown Author is licensed under CC BY-SA-NC

- According to Reuters, **Amazon stopped** using a new **AI** system that **reviewed applicant's resumes** looking for top-level talent. But the system was **trained on resumes submitted for 10 years**, and **most of those were from men**.
- The AI system **penalized resumes** that **included** the word **"women's"**. It also **downgraded graduates of two all-women's colleges**.

How Do We Course Correct on AI Systems?

1. Get These Experts Involved in Defining & Building AI

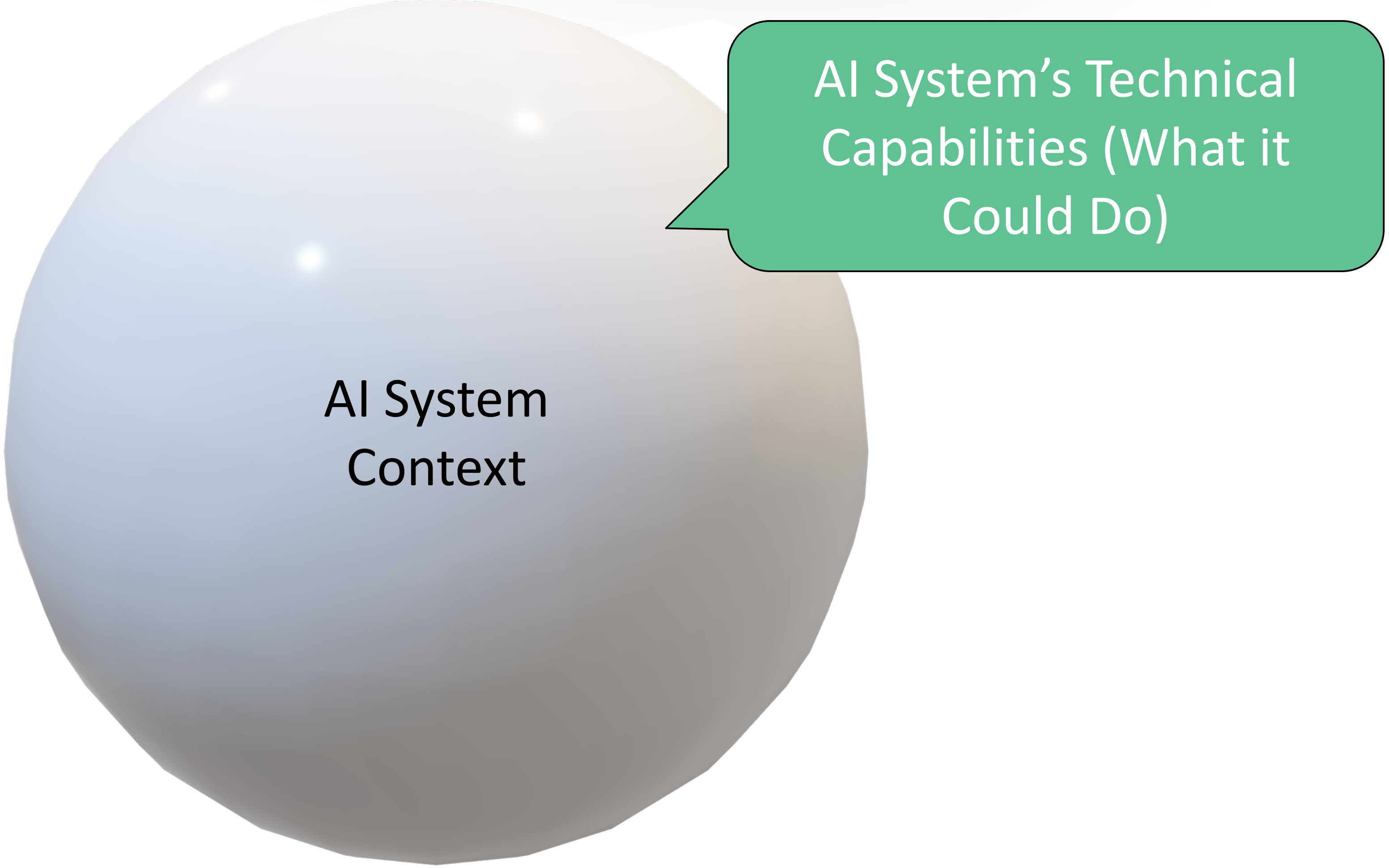
LEGAL

COMPLIANCE
&
REGULATORY

2. Collect larger & more diverse data sets to use for AI training

How Do We Course Correct on AI Systems?

Get Legal, Compliance, & Regulatory Experts Involved in Defining & Building AI



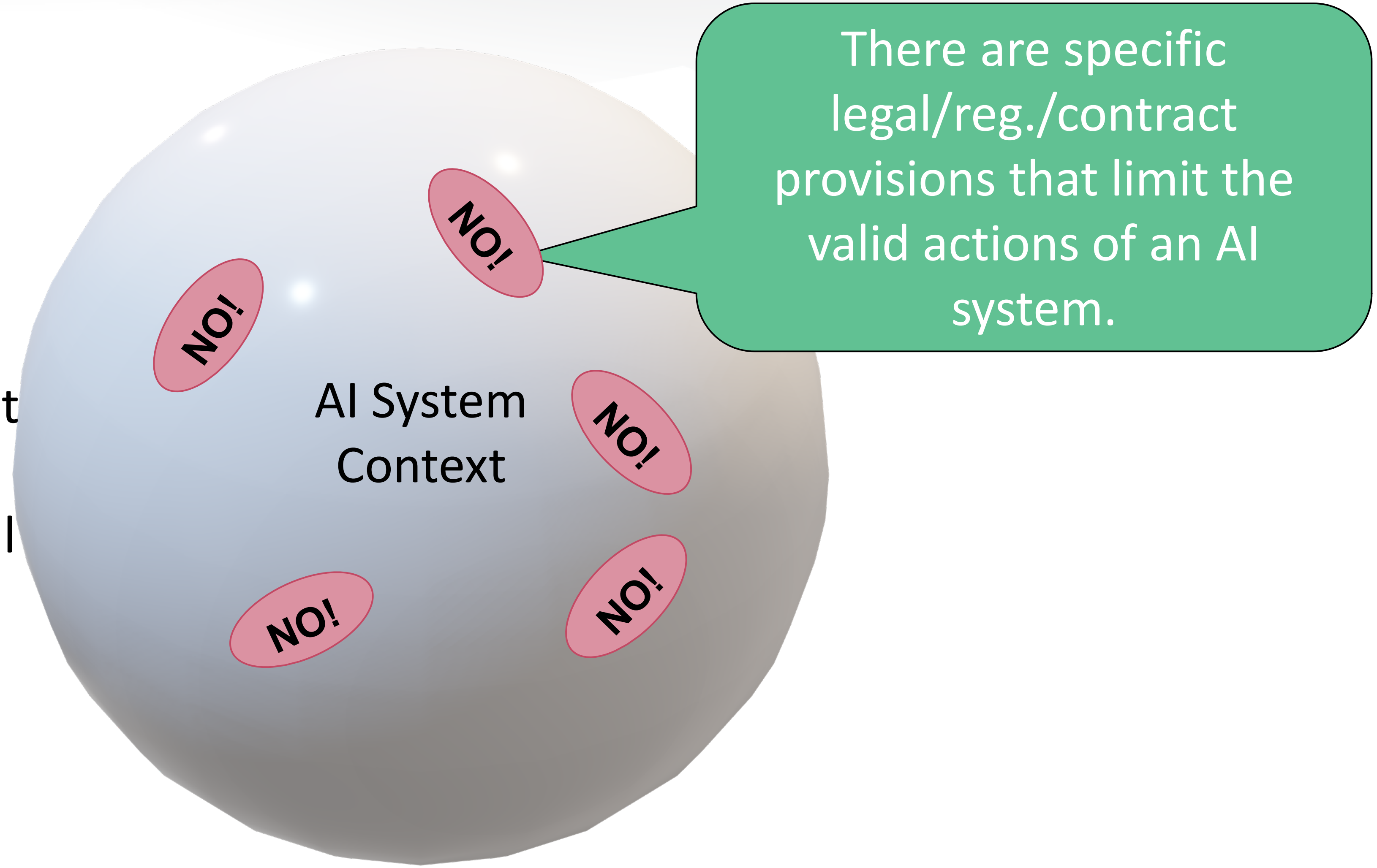
AI System Context

AI System’s Technical Capabilities (What it Could Do)

How Do We Course Correct on AI Systems?

Get Legal, Compliance, & Regulatory Experts Involved in Defining & Building AI

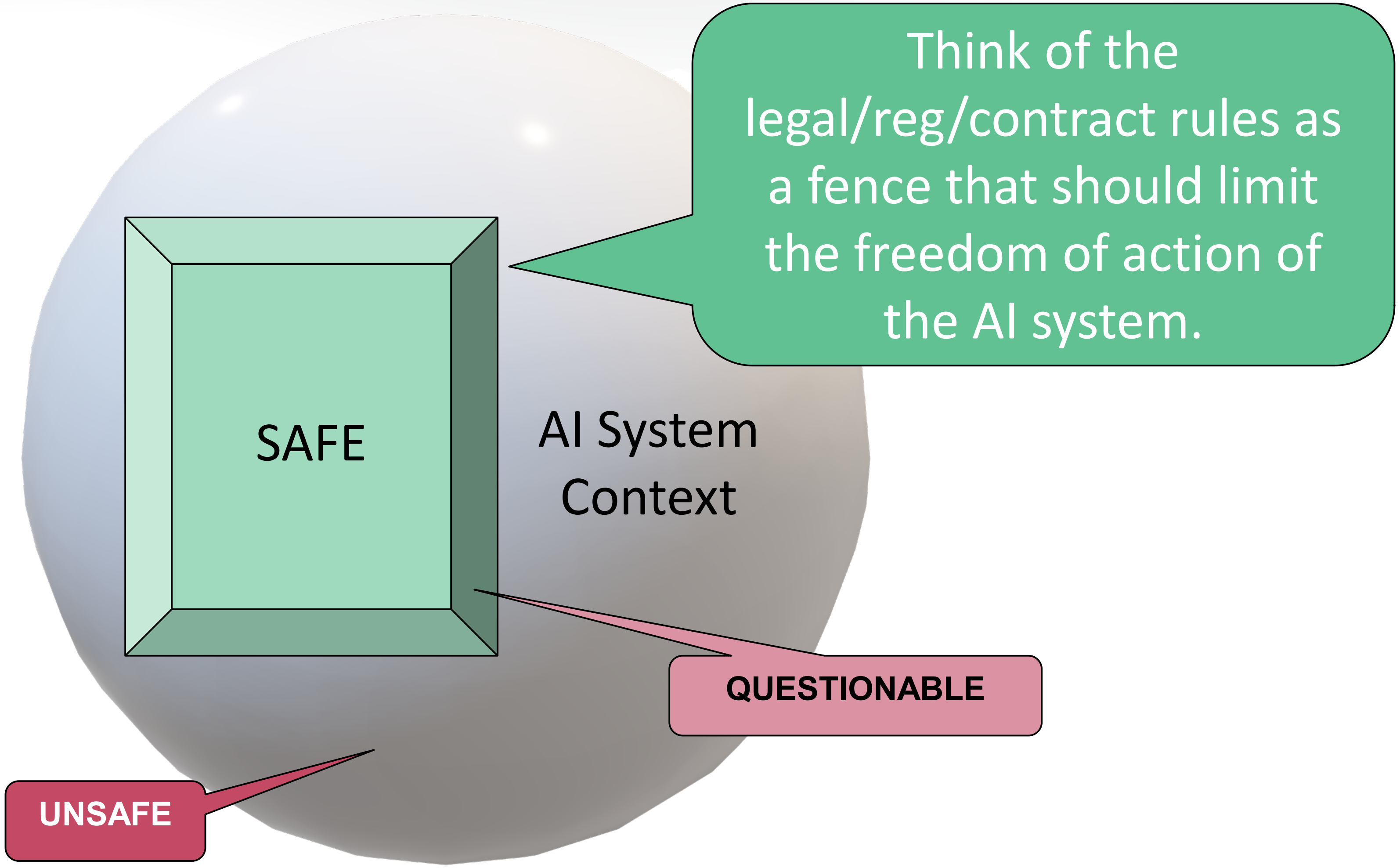
Part of an AI System's Context is the Legal/Regulatory/Contractual Framework Within Which it Operates



There are specific legal/reg./contract provisions that limit the valid actions of an AI system.

How Do We Course Correct on AI Systems?

Get Legal, Compliance, & Regulatory Experts Involved in Defining & Building AI



Think of the legal/reg/contract rules as a fence that should limit the freedom of action of the AI system.

How Do We Course Correct AI Systems?

Get Legal, Compliance, & Regulatory Experts Involved in Defining & Building AI

Collect larger & more diverse data sets to use for AI training

IMPLICIT BIAS

First, if you don't know about it, you're unlikely to prevent it.

PREVENTING IT REQUIRES THINKING ABOUT

- **How you frame the problem (what constitutes "success"?)**
- **How you collect data (does the data represent reality? Does it reflect existing/past prejudices?**
- **How was the data prepared? What attributes are provided to the AI system to use?**

How Do We Course Correct AI Systems?

Dealing with implicit bias can be harder than you think…

Unknown unknowns

- Amazon tried to fix its software by eliminating penalties for the word "women's".
- But there were implicitly gender-related words that appear more on male applicant's resumes than females (e.g. "executed" or "captured". **Those were still used**.

Process problems

- One standard test divides the potential training material into a group to be used to train, and one to be used to validate the training.
- If both have the same implicit bias content, **YOU'LL NEVER NOTICE IT.**

Social Context

- An algorithm designed for one purpose might be used by computer scientists in a different application.
- "Fairness" may vary depending on the subject of the system, and bias factors have to be considered in light of what the system does.

Key Takeaways

- AI/DL systems being built in private & public sector systems

AI IS GROWING

- AI built by Computer/AI/Data Science Specialist with potential for no knowledge of legal & social science areas

BRING IN LEGAL/ COMPLIANCE EXPERTS

Key Takeaways

- Identifying a problem after the system is launched may be too late to avoid consequences.

IDENTIFY PROBLEMS DURING DESIGN & BUILD

- AI training data sets can include bias due to being uncomprehensive, incomplete, and too small.

USE LARGER, MORE DIVERSE, & MULTIPLIE TRAINING DATA SETS

Thank you for inviting us. If we can help, please contact us.

- abrill@kroll.com
- stacy.scott@kroll.com
- If you would like to receive our cyberdefense or intel threat newsletters or our reports on fraud or other subjects, just let us know.

Kroll | A Division of DUFF & PHELPS

The MONITOR

VOLUME 2

Issue 4	Understanding and Fighting Against Banking Trojans	Page 2
Issue 5	Point-of-Sale (POS) Compromise and MID Refund Frauds	Page 5
Issue 6	Web Application Compromise and E-commerce Exploits	Page 8