pmsymposium.umd.edu

NEXT SESSION

MANAGING AI: RISKS AND OPPORTUNITIES

Donald Polaski & Marissa Brienza

Director of AI and ML, Booz Allen Hamilton

Chief Data Scientist, Booz Allen Hamilton

This session will be recorded.

University of Maryland
Project Management
Symposium

PROJECT MANAGEMENT
CENTER FOR EXCELLENCE
A.J. CLARK SCHOOL OF ENGINEERING
Civil & Environmental Engineering Department

Managing AI: Risks & Opportunities

Donald Polaski and Marissa Brienza

2023 Project Management Symposium

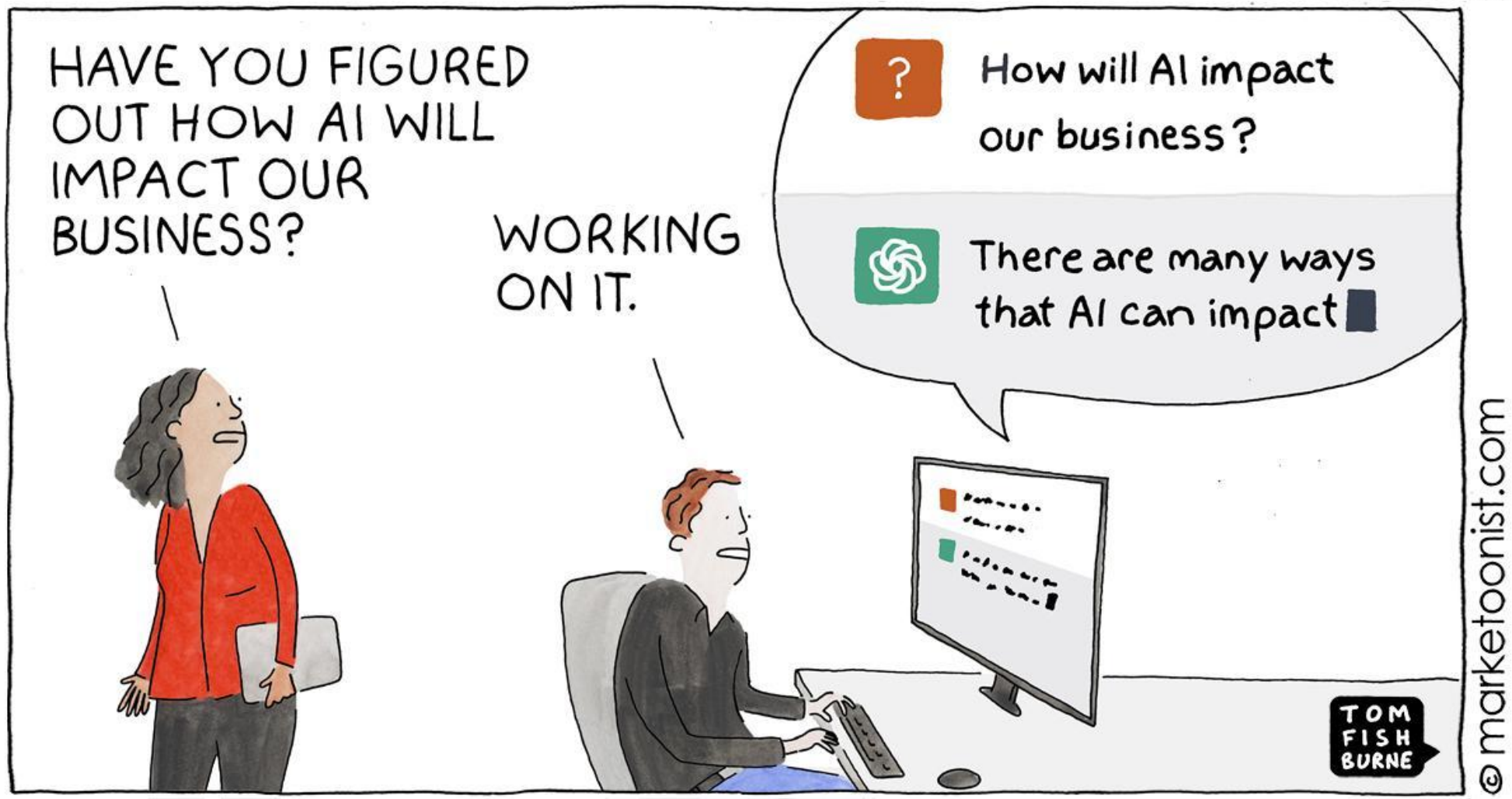
Speaker Introductions

Donald Polaski

AI Director – Booz Allen Hamilton

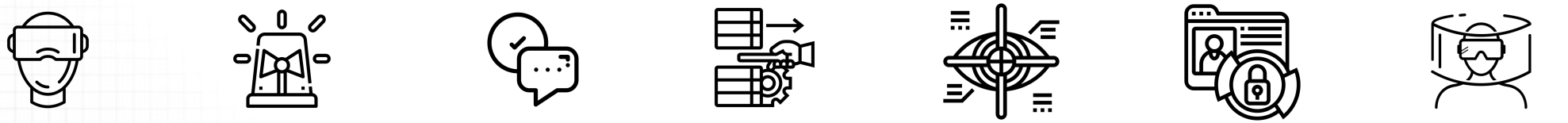
Marissa Brienza

Chief Data Scientist – Booz Allen Hamilton


HAVE YOU FIGURED OUT HOW AI WILL IMPACT OUR BUSINESS?
WORKING ON IT.
How will AI impact our business?
There are many ways that AI can impact
TOM FISH BURNE
©marketoonist.com

What is Artificial Intelligence (AI)?

AI is a broad field of study focused on using computers to do things that require human-level intelligence.



PERCEPTION | NOTIFICATION | SUGGESTION | AUTOMATION | PREDICTION | PREVENTION | SITUATIONAL AWARENESS

Hype Around AI

We are in the 4th Industrial Revolution - marked by the convergence of technologies such as AI, robotics, IoT and blurring the lines between the physical, digital, and biological spheres.

Generated by AI for
Coca-Cola Art Competition

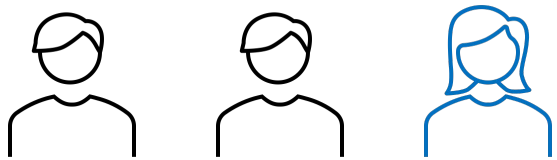
"Created by humans, AI should be overseen by humans. But in our time, one of AI's challenges is that the skills and resources required to create it are not inevitably paired with the philosophical perspective to understand its broader implications."

- Henry A Kissinger; Eric Schmidt; Daniel Huttenlocher in The Age of AI

Case Study: Bias in AI

Amazon's experimental hiring tool used AI to give job candidates a score, ranging from one to five stars.

In 2015, they realized the new system was not rating candidates in a gender-neutral way.



About a third of employees in tech occupations in the U.S. are women.

Case Study: Legal Implications in AI

The current legal cases against generative AI are just the beginning

AI that can generate art, text and more is in for a reckoning

Kyle Wiggers @kyle_l_wiggers / 11:30 AM EST • January 27, 2023 Comment

Yoda and Harry Potter Chatbots Could Be the Next Big Legal Battle

By Brad Stone

March 20, 2023 at 7:00 AM EDT

'Robot' Lawyer DoNotPay Sued For Unlicensed Practice Of Law: It's Giving 'Poor Legal Advice'

Cameron Burke

Mon, March 20, 2023 at 2:34 PM EDT · 4 min read

Legal Doomsday For Generative AI ChatGPT If Caught Plagiarizing Or Infringing, Warns AI Ethics And AI Law

Lance Eliot Contributor

Dr. Lance B. Eliot is a world-renowned expert on Artificial Intelligence (AI) and Machine Learning...

Follow

0 Feb 26, 2023, 08:00am EST

MARCH 22, 2023

From tort law to cheating, what is ChatGPT's future in higher education?

by Jason Pohl, University of California - Berkeley

Call to Action

As AI Project Managers, we have a responsibility to continuously learn and evolve with the field to ensure we **promote responsible AI uses and practices** across our teams and provide customers with **AI solutions they can trust and rely on**.

We are responsible for the solutions we create.

AI is a complex integration of people, processes, and technology and although may seem like magic, it comes with **considerable risks**.

- Bias
- Transparency
- Concept Drift
- Privacy
- Legal

Risk: Bias

Bias occurs when an AI system produces outputs that lead to discrimination against specific groups or individuals (e.g., race, gender, biological sex, nationality, age).

Example: Bias

Machine Learning algorithms are being designed to detect skin cancer but run the risk of being less accurate for people with dark skin.

A study examined available datasets:

- 21 datasets with 106,950 images
- 2,436 images had data on skin type, 10 listed brown, dark brown, or black skin
- 1,585 images had data on ethnicity, 0 listed African, African-Caribbean or South Asian

So, where do we go from here?

Questions to Ask Yourself: Bias

- What is the source of my training and testing data?
- How was the data collected?
- Could there be biases or inaccuracies in the data?
- How can we ensure proper demographic representation in the data?
- What tools can we use to monitor model performance and alert to potential bias?

Mitigations: Bias

- Host regular sessions to discuss potential bias and ethical risks
- Conduct proper testing on data used for training and evaluation
- Leverage an AI ethical assessment tool or capability
- Continuously evaluate models for bias after deployment
- Be proactive in identifying potential bias and communicating those issues

Risk: Transparency

AI systems are often referred to as "black boxes“ – they can be difficult for humans to interpret. Knowing why a model makes predictions the way it does is critical.

Example: Transparency

Dialpad, an AI customer intelligence platform, offers a Sentiment Analysis product enabling teams to track coachable moments on a customer calls. A common user complaint was around the lack of understanding on why a statement was tagged positive or negative.

- Performed poorly on sarcasm and swear words
- User perspectives between what is positive and negative varies
- Now, highlights most influential words in determining the correct label

Questions to Ask Yourself: Transparency

- What features are most important in the model?
- What technologies and techniques can we use to interrogate our model?
- Are there certain types of data that the model performs poorly on?
- Are there tradeoffs we can make to make our models more transparent?
- How do we measure accuracy and performance of our AI system?

Mitigations: Transparency

- Determine tradeoff between explainability and algorithm selection and performance
- Leverage available model explainability tools (e.g., LIME)
- Ensure training data set is well-labeled and robust
- Perform consistent Quality Assurance tests for early detection of deviations from the norm/expected results

Risk: Concept Drift

AI models begin to deteriorate soon after they are deployed.

- **Model Drift** - predictive power of predictors declines
- **Data Drift-** your data can change over time

Example: Concept Drift 

Zillow Offers used AI to assess a property's value and enabled homeowners to sell their homes without a real estate agent. They lost $304M in 2021 because their model couldn't keep pace with market conditions.

- Overestimated the value of the homes
- Employed sound development, testing, and deployment procedures
- Lacked a monitoring solution

Questions to Ask Yourself: Concept Drift

- How are we monitoring model performance over time?
- How can we build automated alerts for performance issues?
- How often do we retrain our models?
- How are we doing data quality control?
- How can we build automated alerts for changes in data (e.g., means, distributions)?

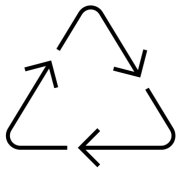
Mitigations: Concept Drift

- Leverage rigor of software development and configuration management principles (e.g., MLOps)
- Employ automated and continuous monitoring tools (e.g., Amazon SageMaker Model Monitor)
- Construct model metrics dashboards for stakeholders
- Create a retraining schedule
- Build a new model if necessary

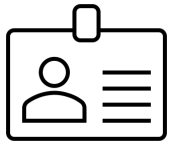
Risk: Privacy

Training AI models requires large volumes of data – we must safeguard data and ensure data storage meets regulatory standards and protects the privacy of the users.

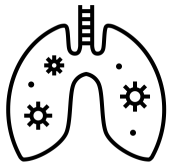
Data Security



Data Repurposing



Data Re-Identification

Example: Privacy 

In 2017, Google cancelled a project with the National Institutes of Health days before they were set to make more than 100,000 images of human chest X-rays publicly available.

- Data could be used to train models to identify lung disease
- Concerns that the images could be used to identify patients
- Privacy violations come with legal consequences

Questions to Ask Yourself: Privacy

- What regulations exist for using this data?
- How do we meet these regulations throughout the development lifecycle?
- How will we perform data governance?
- What is our data management lifecycle?
- How will we secure our data (e.g., encryption, access control)?
- How might nefarious actors try to use our data?
- What mechanisms are in place to prevent, detect, and address security breaches?

Mitigations: Privacy

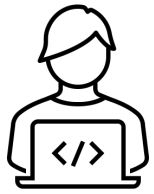
- Conduct a privacy impact assessment
- Understand the regulations for data safeguarding and governance for your organization, industry, state, and country
- Use accurate, fair, and representative data sets
- Obtain consent to use individual's data
- Leverage data minimizing methods (e.g., de-identification, aggregation)
- Audit privacy measures regularly
- Monitor for data leaks or data breaches

Risk: Legal

The AI field is evolving faster than legislation. It is still our responsibility to 1) ensure AI is used responsibly and 2) to protect our teams and organizations from potential legal issues.

- Discrimination
- Data Privacy
- Antitrust and Competition
- Intellectual Property
- Liability for Harm or Damage

Example: Legal



GitHub Copilot uses generative AI to suggest new code and entire functions in software development tools. A class action lawsuit was filed in late 2022 against Copilot creators – GitHub, Microsoft, and OpenAI.

- Ingested billions of lines of computer code from GitHub public repositories
- Contends that the legal rights of a vast number of creators were violated
- Users can generate blocks of copyrighted code, with no attribution and without attaching the required license

Questions to Ask Yourself: Legal

- What is our risk tolerance for incorporating Generative AI?
- What data was used for training?
- Who owns the generated outcomes (e.g., code, art, text, etc.)?
- What are the potential licensing, copyright, or patent infringements?
- How do I include/apply the licenses and attributions needed?

Mitigations: Legal

- Engage with your organization's lawyers early
- Evaluate risk of using Generative AI
- Understand the terms of use for Generative AI tools
- Ensure data used in Generative AI or a custom model is legally obtained with appropriate licenses in place
- Establish guidelines for the use of AI-generated content to protect your IP

Resources Available to AI Project Managers

- Nvidia's Deep Learning Institute
- NIST Artificial Intelligence Risk Management Framework
- NIST AI Risk Management Framework Playbook

Closing

As *Project Managers*, it is our responsibility to:

- Understand the challenges and risks with developing and operationalizing AI solutions
- Manage those risks to maximize impact
- Stay up to date with industry standards and news
- Shepherd our teams, customers, and clients on their AI journeys

Evaluate Session