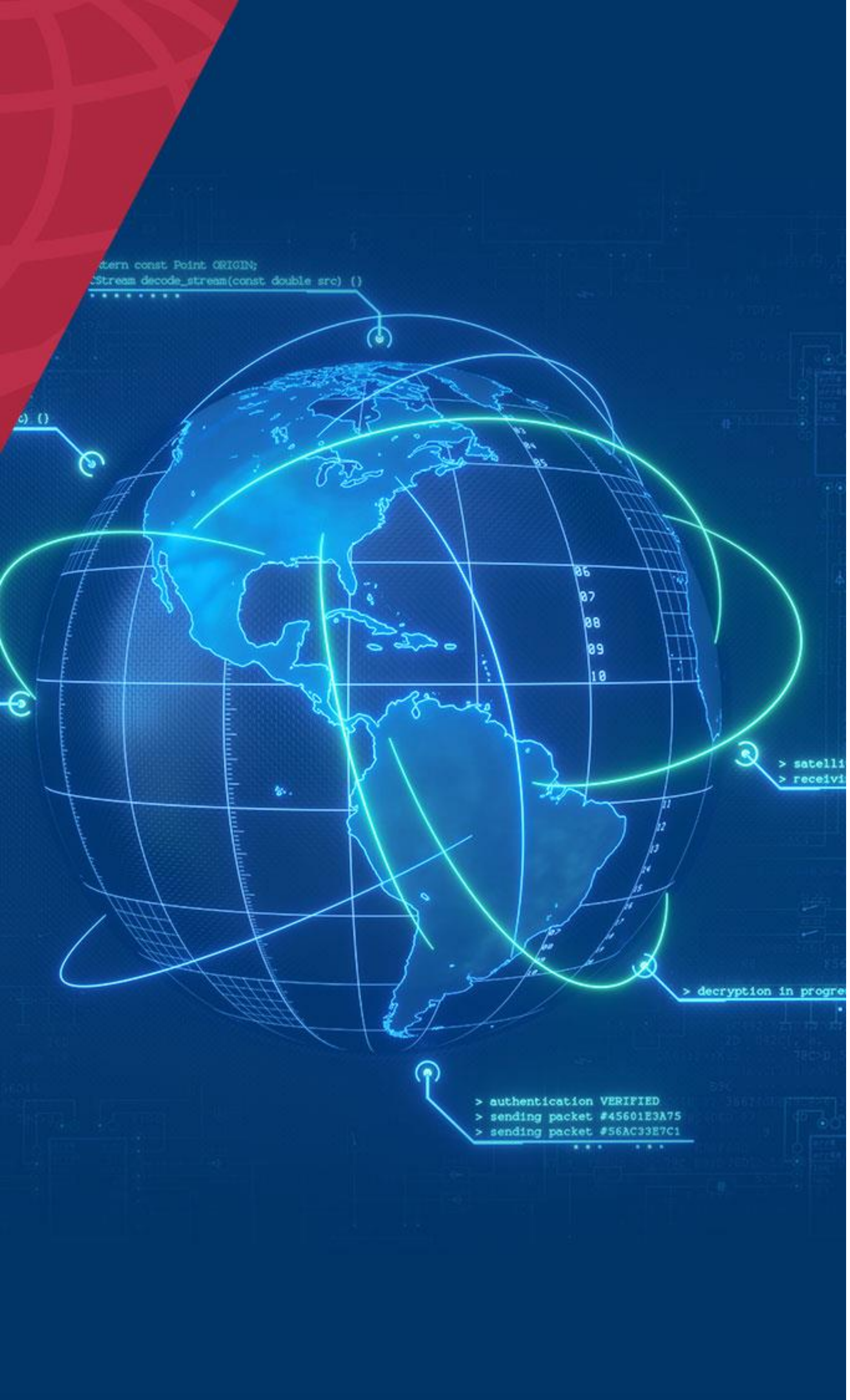
Joint IT & SW Cost Forum

Intersections of AI and Cost Estimating: Explainability

Presented by:
Jon Kilgore, PRICE Systems

Estimate with Confidence™

Contents



Introduction to Artificial Intelligence and Machine Learning

Definitions and Concepts

Artificial Intelligence (AI)

"…any program is an AI system simply by the fact that it does something that we would normally think of as intelligent in humans. How it does so is not the issue; just that it is able to do it at all is the key point."

- Strong AI: work aimed at genuinely simulating human reasoning; explain how humans think
- Weak AI: systems that can behave like humans, the results tell us nothing about how humans think
- Broad AI: systems designed around the ability to reason in general
- Narrow AI: systems designed around specific tasks

Practical Artificial Intelligence For Dummies®, Narrative Science Edition. John Wiley & Sons, Inc. p. 6, 8. ISBN 978-1-119-14984-2

Machine Learning

"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E."

Tom Michael, Carnegie Mellon University (CMU); Mitchell, T. (1997). Machine Learning. McGraw Hill. p. 2. ISBN 978-0-07-042807-2

The Technical Core of AI Systems

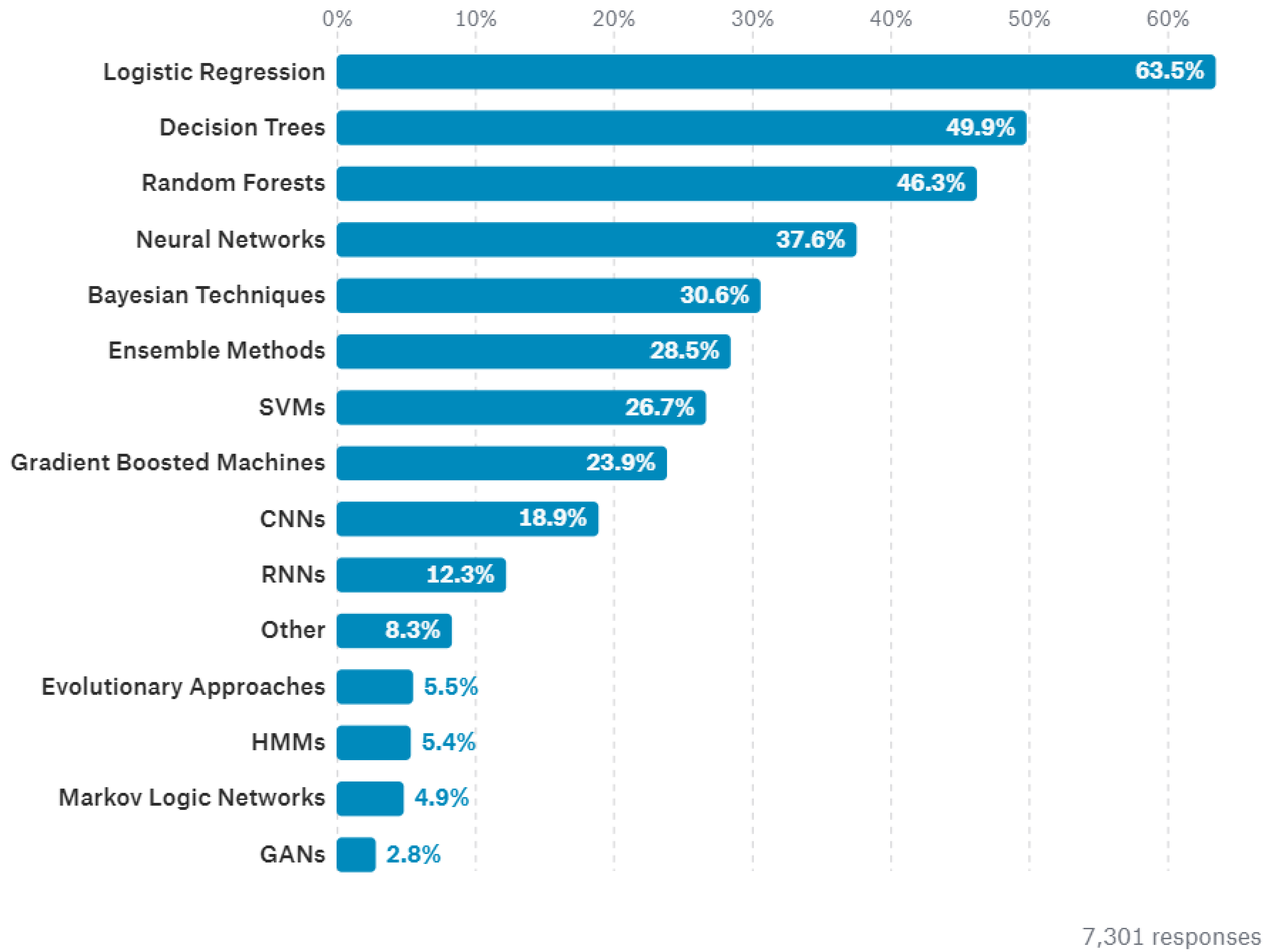
"AI Stack"

Technology Blocks	Description
Talent	Drives the creation and management of all other components; most valuable resource
Data	Enables machine learning to create new applications and improve performance of existing AI applications
Hardware	Provides the computing power to analyze ever-growing data pools and run applications
Algorithms	Mathematical operations that tell system how to navigate the data to provide answers to specific questions
Applications	Make the answers useful for specific tasks
Integration	Integrate existing data flows, **<u>decision pipelines</u>**, legacy equipment, testing designs, etc. • Critical to fielding successful end-to-end system • Requires significant engineering talent and investment • Can be daunting and **<u>historically has been underestimated</u>**

National Security Commission on Artificial Intelligence. Interim Report November 2019. p. 8.

The State of Data Science and Machine Learning

What data science methods are used at work?



kaggle: https://www.kaggle.com/surveys/2017

Why Should AI Matter to Us?

- "The Commission's attempts to predict AI's impact on national security is like Americans in the late 19th century pondering the impact of electricity on war and society." National Security Commission on AI
- Impacts of AI on National Security
  - Will change the way we defend, detect, identify, safeguard, respond in ways humans simply can't
  - Speed and accuracy; processing huge amounts of data
  - Semi-autonomous and autonomous combat systems; "algorithmic warfare"
- Impacts of AI on Cost Estimating and Predictive Analytics
  - Many of us will support programs with AI capabilities; estimate the "AI Stack"
  - The cost and effort of weapons systems of almost every type are increasingly driven by software; AI will also primarily be implemented via software in warfighter systems
  - New, powerful analytical techniques will be introduced to our community
  - Explaining our models and convincing decision makers why they should trust them is likely to become increasingly challenging

Will DoD AI Implementation be more Complex?

- Military applications are frequently distinct from commercial applications due to:
  - Noisy, incomplete, uncertain and erroneous data inputs during operations
  - Limited access to real data to train AI/ML
  - Rapidly changing situations
  - Peer adversaries that employ deceptive techniques to defeat algorithms
    - *AI hacking, data poisoning, adversarial machine learning*
    - *Examples: Inserting a few tactically inserted pixels (for a computer vision algorithm) or some innocuous looking typos (for a natural language processing model)*
    - *Difficult to detect using traditional methods*

A Tesla Model S thought this was 85 mile-an-hour speed limit sign, according to researchers at McAfee

U.S. Army CCDC Army Research Laboratory Public Affairs. https://www.army.mil/article/235531.
Hacking AI: Exposing Vulnerabilities in Machine Learning. https://www.datanami.com/2020/07/28/hacking-ai-exposing-vulnerabilities-in-machine-learning/?utm_source=rss&utm_medium=rss&utm_campaign=hacking-ai-exposing-vulnerabilities-in-machine-learning.

Ethics Debate Regarding AI for Military Applications

- Pro Weaponization of AI
  - U.S. needs to accelerate the fielding of AI-enabled weapons systems; competition
  - AI-enabled systems could save American lives and reduce civilian casualties
  - Enable military operations in communications-degraded or -denied environments
- Con Weaponization of AI
  - U.S. should slow down or forswear adoption of AI-enabled weapons systems
  - Possible catastrophic incident, crisis instability, weaponization of AI is immoral
- Debates on-going within companies regarding whether and how to engage DoD regarding AI technologies
- Economic and intellectual gains of sharing AI, even with countries posing strategic threat, may outweigh national security risks

> "Creating the principles is the easy part. It's the implementation part that is the hard part." - Lt. Gen. Jack Shanahan, JAIC Director

National Security Commission on Artificial Intelligence. Interim Report November 2019. p. 13-14.
Ethical Principles for Artificial Intelligence, Joint Artificial Intelligence Center, https://www.ai.mil/docs/Ethical_Principles_for_Artificial_Intelligence.pdf

Explainability in AI Challenges

Descriptions of the Explainability Challenge

"Last Mile" Communication Failures – Harvard Business Review

"…many companies aren't getting the value they could from data science. Even well-run operations that generate strong analysis fail to capitalize on their insights. Efforts fall short in the last mile, when it comes time to explain the stuff to decision makers."

Harvard Business Review: Data Science and the Art of Persuasion, https://hbr.org/2019/01/data-science-and-the-art-of-persuasion?

Lack of Traceability - Forbes

"Many of the algorithms used for machine learning are not able to be examined after the fact to understand specifically how and why a decision has been made. This is especially true of the most popular algorithms currently in use – specifically, deep learning neural network approaches."

Forbes: Understanding Explainable AI, https://www.forbes.com/sites/cognitiveworld/2019/07/23/understanding-explainable-ai/#7d9cac687c9e

Descriptions of the Explainability Challenge

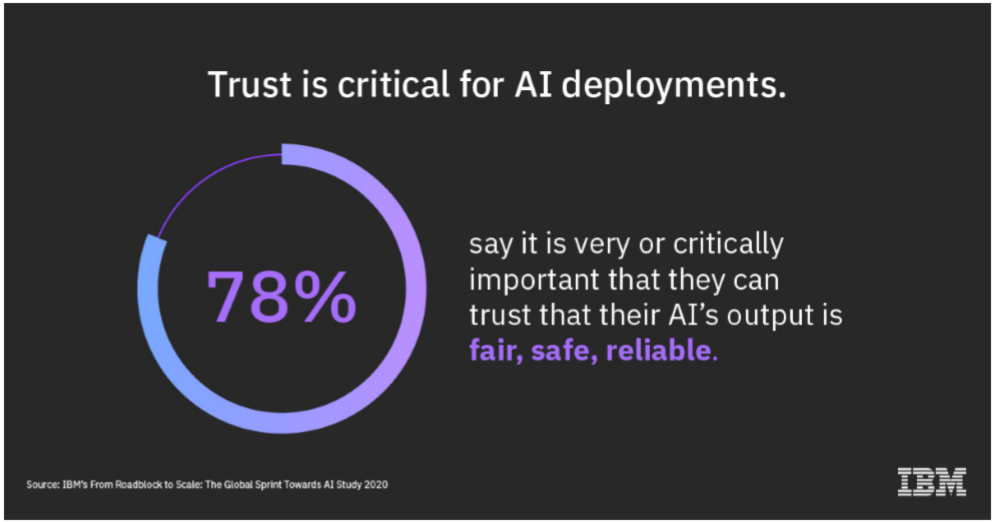
"Black Box" AI Causes Distrust - Cognilytica

"Right now, too much of what AI systems do are a "black box". We have little visibility into how decisions are being made, conclusions drawn, ... AI black boxes are scary … these AI black boxes lack transparency, making their decisions too secretive causing issues of trust and accountability. As the consequences of mistakes and certain decisions become more significant, it becomes more important to have visibility into the inner workings of AI decision-making, or in other words, Explainable AI (XAI)."

Cognilytica Podcast: AI Today Podcast #016: Explainable AI (XAI), https://www.cognilytica.com/2017/12/20/ai-today-podcast-016-explainable-ai-xai/

Trust is Critical for AI Deployments

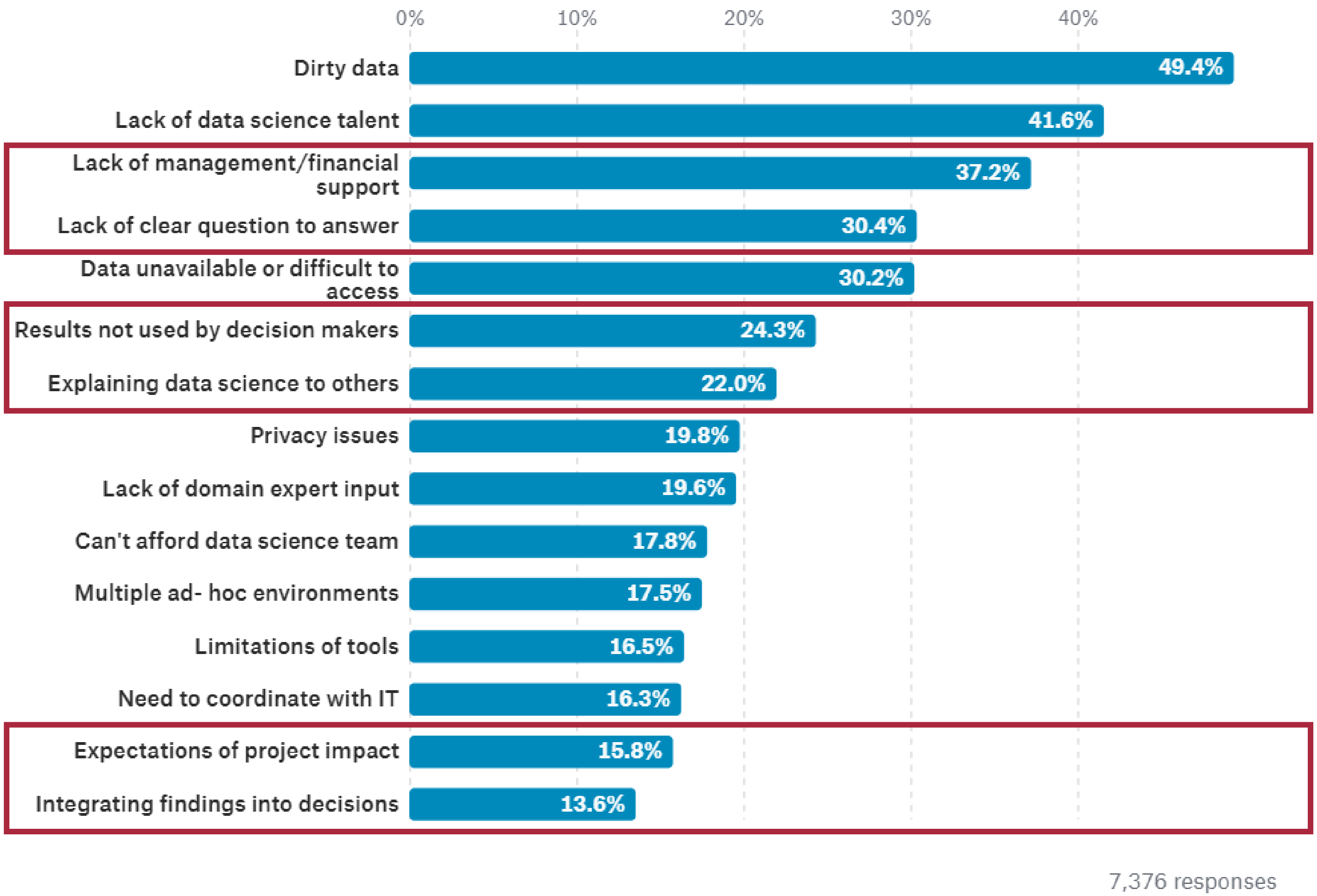
- 2019 IBM survey across an international sample of 4,514 senior business decision-makers
- Globally, 78% of respondents say it is very or critically important that they can trust that their AI's output is fair, safe, and reliable
- Being able to explain how AI arrived at a decision is universally important (83% global respondents), but it is particularly important to those currently deploying AI (92%)



From Roadblock to Scale: The Global Sprint Towards AI, New research commissioned by IBM in partnership with Morning Consult
http://filecache.mediaroom.com/mr5mr_ibmnews/183710/Roadblock-to-Scale-exec-summary.pdf

The State of Data Science and Machine Learning

What barriers are faced at work?



kaggle: https://www.kaggle.com/surveys/2017

Definitions and Concepts

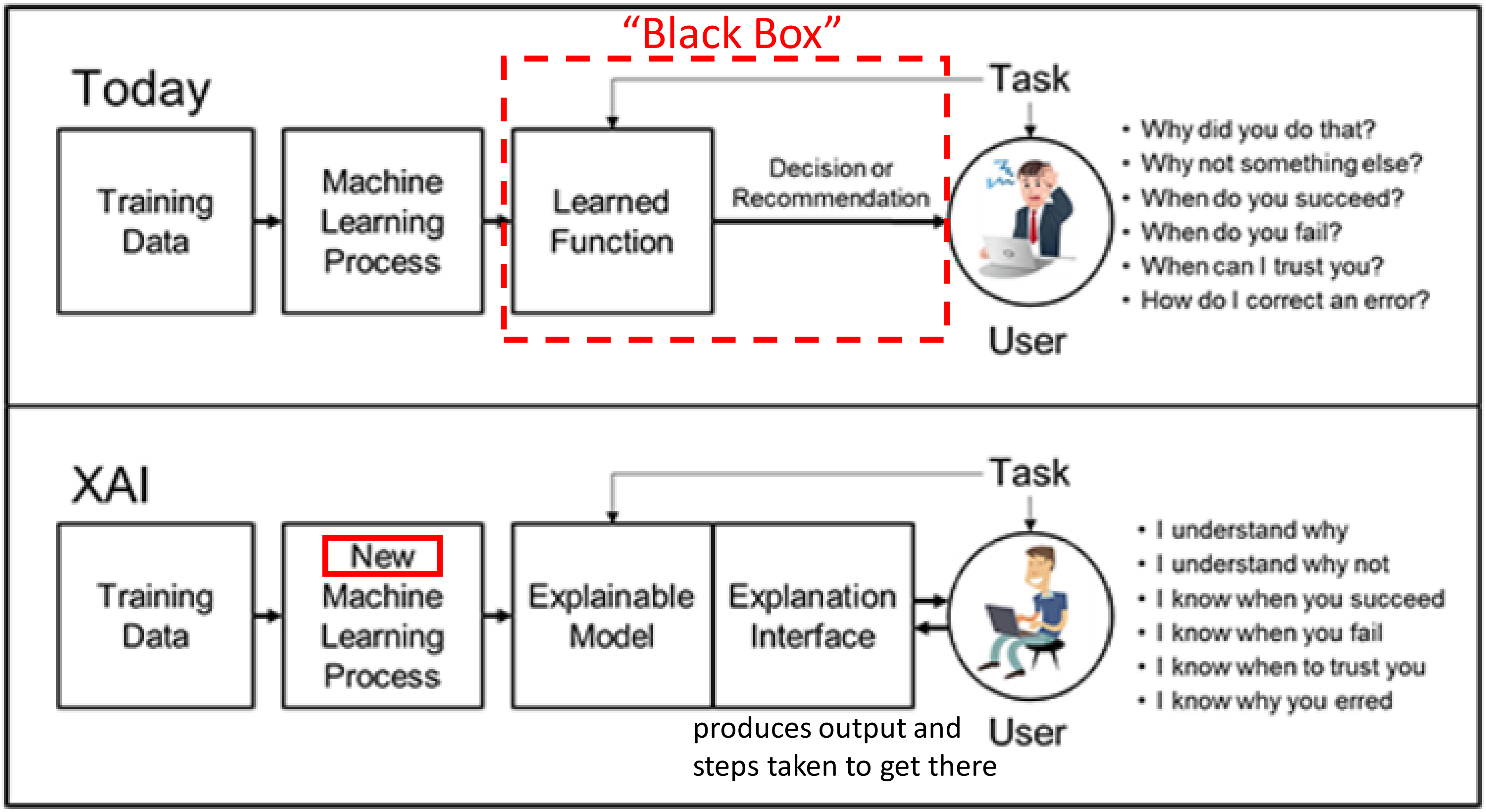
Explainable Artificial Intelligence (XAI)

- Emerging field within AI
- Focused on helping decision makers to understand and trust underlying machine learning logic and algorithms
    - *Inherent in many AI algorithms is a tradeoff between explainability and potential predictive power*
- XAI tries to answer these questions that we don't currently have answers to:
    - *Why did the AI system to that?*
    - *Why didn't the AI system do something else?*
    - *When did the AI system succeed and when did it fail?*
    - *When does the AI system give enough confidence in the decision that you can trust it?*
    - *How can the AI system correct an error?*

Cognilytica Podcast: AI Today Podcast #016: Explainable AI (XAI), https://www.cognilytica.com/2017/12/20/ai-today-podcast-016-explainable-ai-xai/

Future State of Machine Learning and XAI

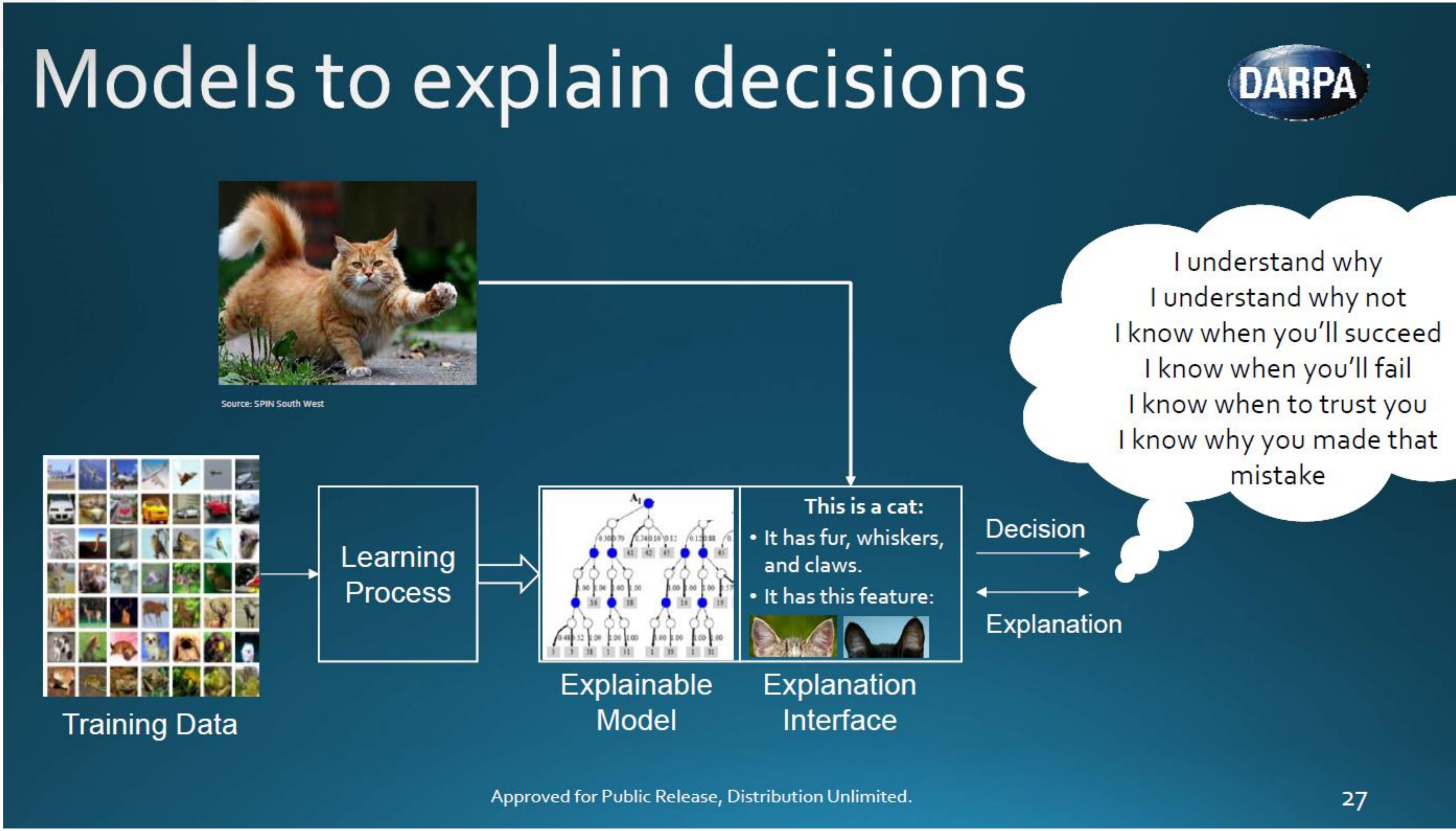
XAI is one of several DARPA initiatives to enable "third-wave AI systems"



DARPA: Explainable Artificial Intelligence (XAI), https://www.darpa.mil/program/explainable-artificial-intelligence

"Our strategy is to pursue a variety of techniques in order to generate a portfolio of methods … covering the performance-versus-explainability trade space."

Future State of Machine Learning and XAI Example



DARPA: A DARPA Perspective on AI, https://www.darpa.mil/about-us/darpa-perspective-on-ai

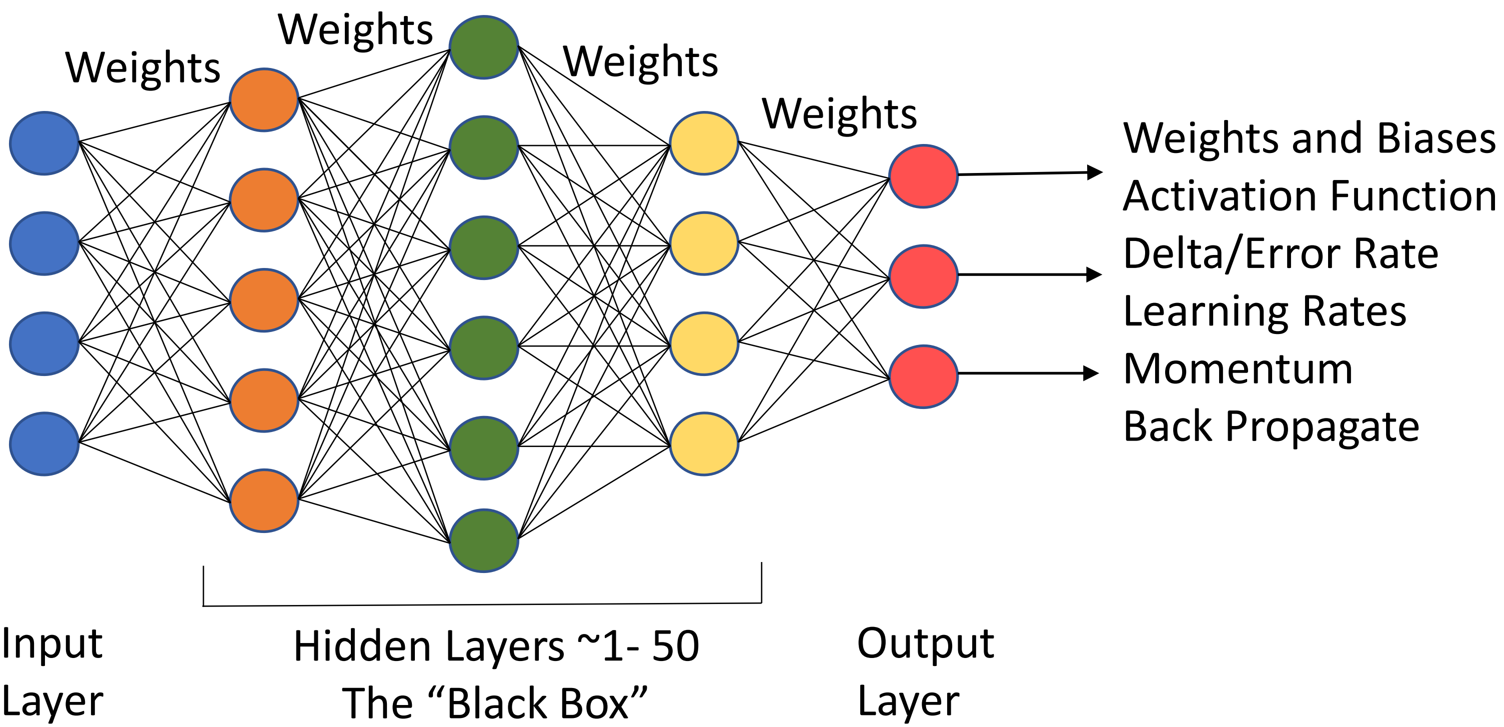
Some Approaches to AI and Machine Learning are More Transparent and Explainable than Others

- Inherent in many AI algorithms is a tradeoff between explainability and potential predictive power
- Less complex techniques such as Bayesian and influence diagrams are explainable and transparent
- More powerful and complex techniques, such as neural networks and deep learning, are much less so
- Most approaches have been known for some time, but today's computing power makes them practical (thanks video games!)
  - First proposed Neural Networks in 1944
  - "There's this idea that ideas in science are a bit like epidemics of viruses," Tomaso Poggio, MIT

Deep Learning and Neural Networks

- "Currently, deep learning is responsible for the best-performing systems in almost every area of artificial-intelligence research", MIT News
- Consists of thousands or millions of densely interconnected processing nodes

Opacity and lack of transparency still troubles many practitioners and decision makers



MIT News, Explained: Neural networks, http://news.mit.edu/2017/explained-neural-networks-deep-learning-0414
But what is a Neural Network? | Deep learning, chapter 1, https://www.youtube.com/watch?v=aircAruvnKk

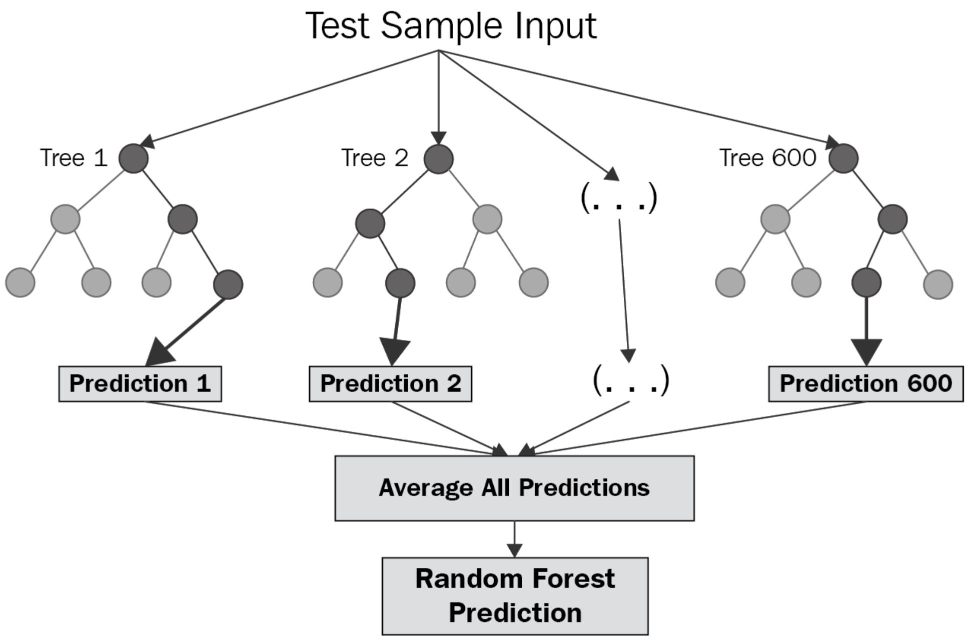
Bayes' Theorem and Machine Learning

- Probabilistic approach where the model can be fully exposed
- P(A|B) = P(B|A) * P(A) / P(B), where:
  - P(A|B): Posterior probability
  - P(B|A): Likelihood
  - P(A): Prior probability
  - P(B): Evidence
- Bayes' Theorem for Modeling Hypotheses
  - A model can be thought of as a hypothesis about the relationships in the data
  - Bayes' Theorem provides a probabilistic model to describe the relationship between data (D) and a hypothesis (h), or in this case a model; for example:
    - *P(h|D) = P(D|h) * P(h) / P(D)*
  - Testing different models on a dataset can be thought of as estimating the probability of each hypothesis (or model) being true given the observed data

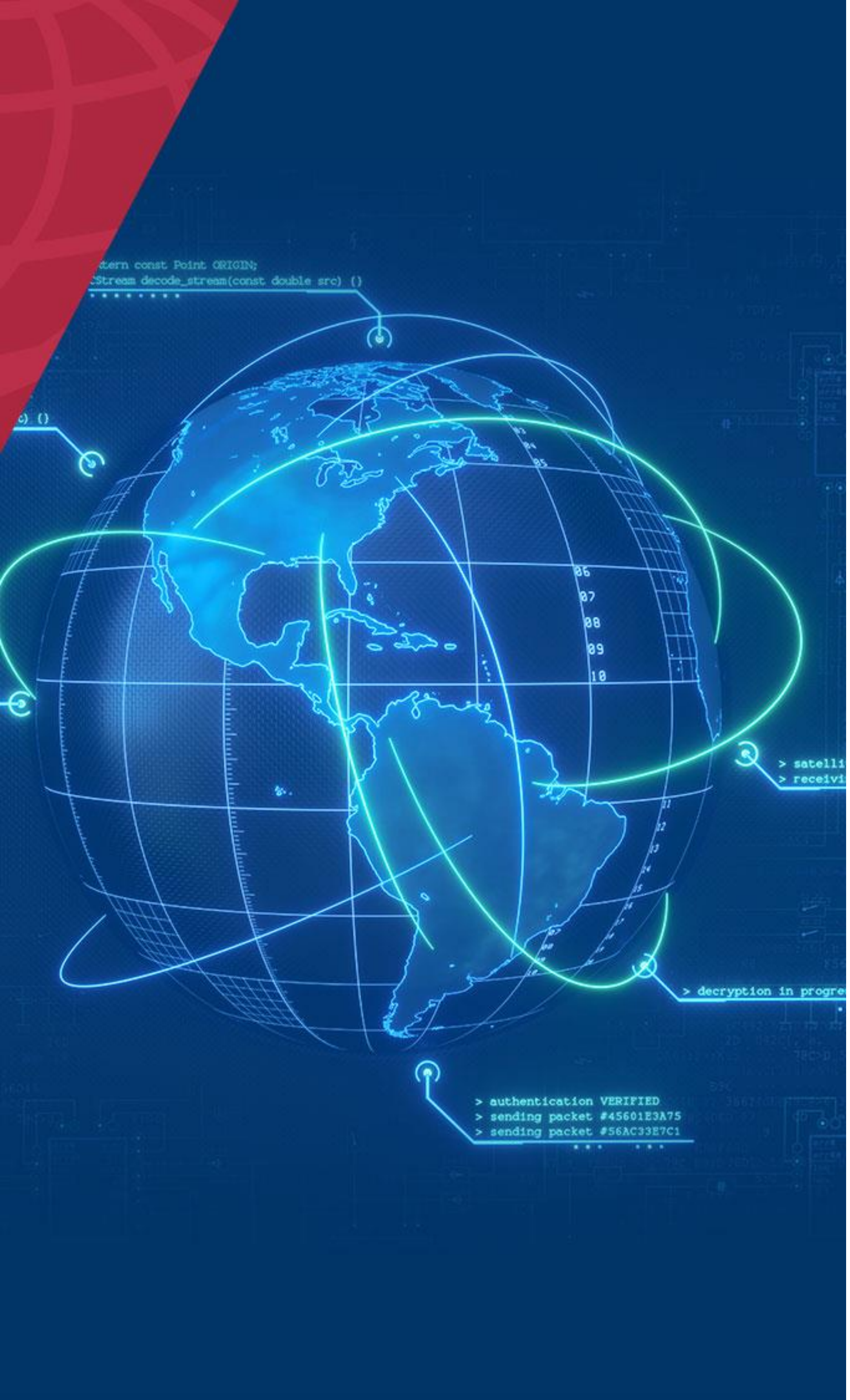
Machine Learning Mastery, https://machinelearningmastery.com/bayes-theorem-for-machine-learning/.

Actual Recent Example: Random Forest Regression

- Performed random forest regression using sklearn on a large software dataset for a client to develop a predictive model
- Technique was of interest because it is considered a powerful ML technique, while also addressing some of the weaknesses of other decision tree methods
  - However, the resulting prediction model was not clearly traceable and transparent as it was the aggregation of hundred of individual decision trees
  - Will the results of such techniques be accepted by government decision makers?



https://towardsdatascience.com/random-forest-and-its-implementation-71824ced454f



Explainability in Predictive Analytics and Cost Analysis

Existing Cost Community Efforts Towards Explainability

- How do we address explainability in the cost analysis and predictive analytics community?
- What are some of the standards and guidelines we rely upon?
- GAO Cost Estimating and Assessment Guide (March 2020)
  - Chapter 3: The Characteristics of Credible Cost Estimates and a Reliable Process for Creating Them
  - Chapter 13: Document the Estimate
  - Chapter 14: Present the Estimate to Management
  - Chapter 16: Auditing and Validating the Cost Estimate ***(New)***
- ICEAA CEBoK
  - Module 1, Sections: Cost Estimate Qualities, Estimate Documentation, Estimate Development and Validation
  - Module 2, Section: Documentation
  - Module 6, Section: Data Validation

GAO: The Four Characteristics of a Reliable Cost Estimate

- Emphasis on elements related to Explainability:

Characteristic	Summary Definition
Comprehensive	Completely define the program; assumptions and exclusions are reasonable, **clearly identified, explained, and documented**.
Well-documented	Easily be repeated or updated and **can be traced to original sources through auditing**. Thorough documentation **explicitly identifies** the primary methods, calculations, results, rationales or assumptions, and sources of the data.
Accurate	Developed using the best methodology from the data collected.
Credible	**Discuss and document any limitations**; includes sensitivity, uncertainty and risk analysis; cross-checked with alternative estimating methodologies.

GAO Cost Estimating and Assessment Guide. Chapter 3: The Characteristics of Credible Cost Estimates and a Reliable Process for Creating Them. pp. 31-32.

GAO: Best Practices Related to Explainability

- A comprehensive cost estimate:
  - Based on a WBS that is product-oriented, traceable to the statement of work
  - Documents all cost-influencing ground rules and assumptions
- A well-documented cost estimate:
  - Shows the source data used, the reliability of the data, and the estimating methodology
  - Describes how the estimate was developed so that a cost analyst unfamiliar with the program could understand what was done and replicate it
  - Discusses the technical baseline description and the data in the technical baseline are consistent with the cost estimate
- An accurate estimate:
  - Documents, explains, and reviews variances between planned and actual costs
  - Based on a historical record of cost estimating and actual experiences from other comparable programs

GAO Cost Estimating and Assessment Guide. Chapter 3: The Characteristics of Credible Cost Estimates and a Reliable Process for Creating Them. pp. 32-33.

GAO Best Practices for Communicating Estimates

- "A best practice is to present the cost estimate in a consistent format that facilitates management's understanding of the completeness and the quality of the cost estimate."
- "Communicating results simply and clearly engenders management confidence…"

Items that should be included in the information presented to decision makers:

- ✓ Date and intended audience
- ✓ A top-level outline
- ✓ The estimate's purpose, including why it was developed and what approval is needed
- ✓ A brief program overview, including scope, physical and performance characteristics, and acquisition strategy to enable management to understand the program's technical foundation and objectives
- ✓ Estimating ground rules and assumptions
- ✓ Life cycle cost estimate, including time-phased costs and constant year dollars
- ✓ Changes from any previous estimates
- ✓ A discussion of WBS elements, including:
  (1) a breakout of element costs and their percentage of the total cost estimate to help identify key cost drivers;
  (2) the estimating method for each WBS element; and
  (3) data sources and historical data
- ✓ Sensitivity analysis, including an interpretation of cost drivers and results
- ✓ Discussion of risk and uncertainty analysis, including:
  (1) cost drivers and top risk areas;
  (2) the corresponding S curve, the level of confidence in the point estimate, and contingency associated with select confidence levels; and
  (3) how risk and uncertainty distributions were defined
- ✓ Comparison to an independent cost estimate with a discussion of differences and the results of reconciliation
- ✓ A comparison of the life cycle cost estimate to the program budget, expressed in budget year dollars, including contingency based on the risk and uncertainty analysis and any budget shortfall and its effect
- ✓ Concerns or challenges with the estimate
- ✓ Conclusions and recommendations

GAO Cost Estimating and Assessment Guide (2020). Chapter 14: Step 11: Present the Estimate to Management. pp. 176- 177.

Relevant ICEAA Guidelines and Standards

- Characteristics of high-quality cost estimates:

Characteristic	Summary Definition
Accuracy	Techniques based on sound analysis. Unbiased.
Comprehensiveness	**Appropriate WBS. GR&A documented.**
Replicability and Auditability	**Someone else can reproduce the same answer.**
Traceability	**Data traceable to source. WBS element tasks traceable.**
Credibility	Combination of prior characteristics.
Timeliness	Provides decision makers insight when needed.

ICEAA CEBoK. Module 1 Cost Estimating Basics. Slide 13.

- Data Validation (some elements of an Explainable Model)
  - Scatter Plots
  - Descriptive Statistics
  - Outlier Identification
  - Compare to History
  - Charting and Histograms

ICEAA CEBoK. Module 6 Data Analysis. Slides 19-25.

Relevant ICEAA Guidelines and Standards

- CEBoK Technique Documentation:
  - Within reason, more information is better than less
  - Any information that is used in the analysis must be included in the documentation
  - Documentation should be adequate for another cost analyst to replicate your technique
  - Like they used to tell you in math class....

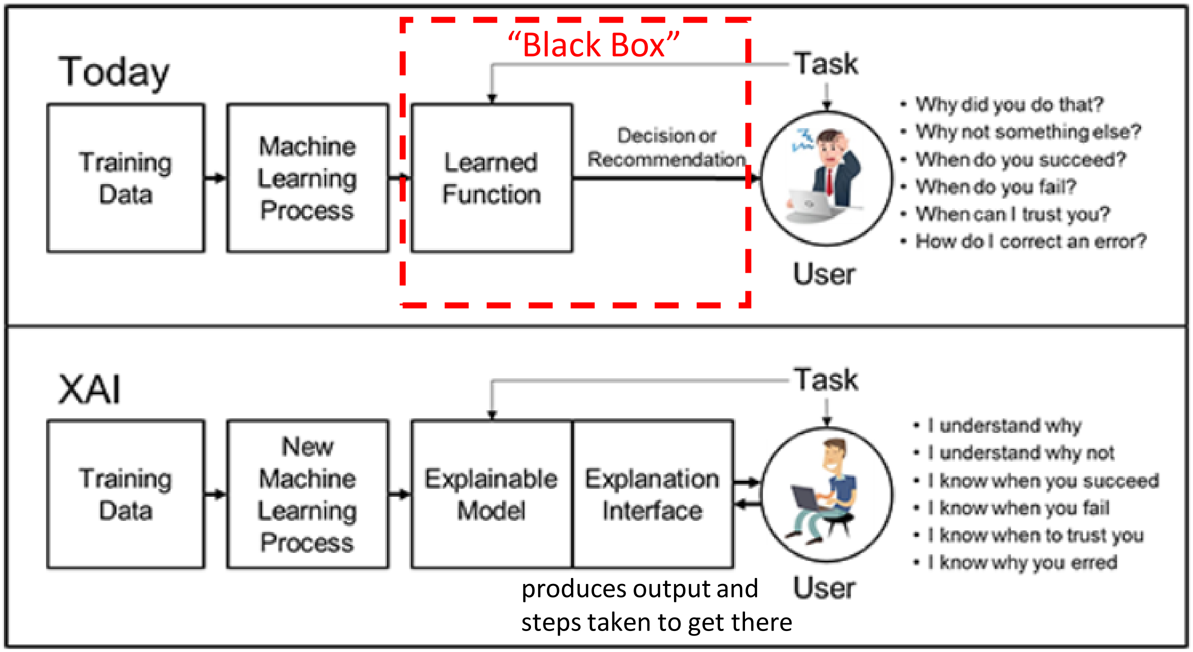
If You Don’t Show Your Work,
You Don’t Get Any Credit!

ICEAA CEBoK. Module 2 Costing Techniques. Slide 25.

Which of these Models does the cost estimating community most resemble today?

Our guidelines and best practices align towards the XAI approach (below); however experience suggests that reality is more like the AI and machine learning challenges of today (above)

<u>Key Difference:</u> Today our constraints regarding Explainability are generally not due to the techniques and models themselves

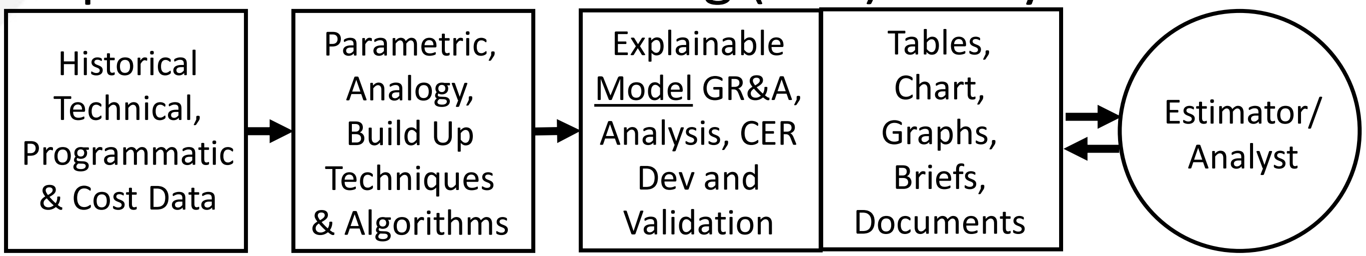


DARPA: Explainable Artificial Intelligence (XAI), https://www.darpa.mil/program/explainable-artificial-intelligence

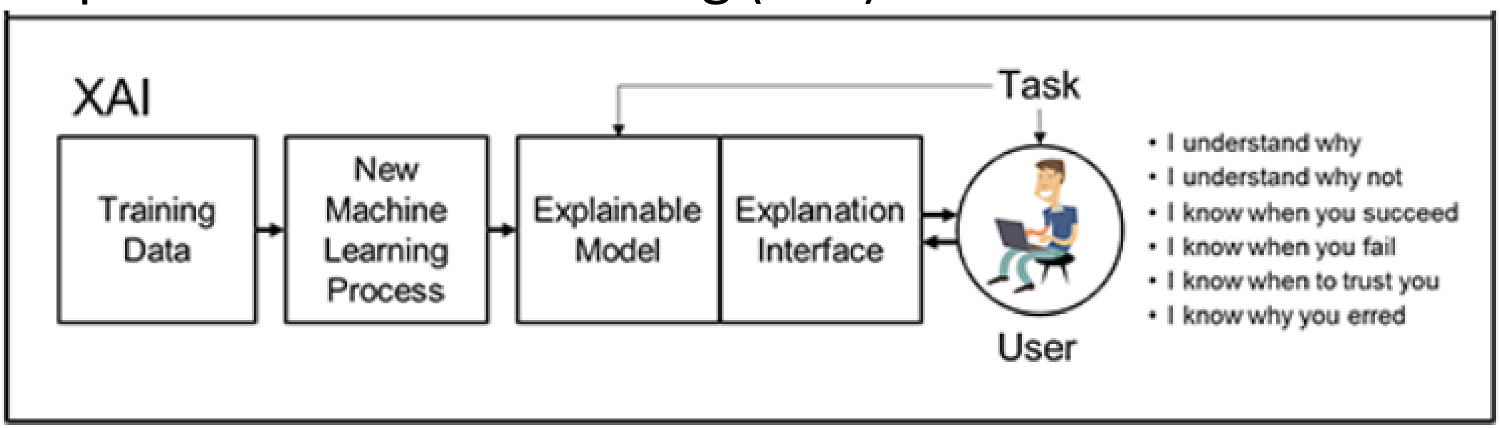
Can/Should We Mimic the XAI Explainability Model?

Over time, it's likely that they will converge as the cost estimating community evolves to include AI and machine learning techniques into our solutions

Explainable Cost Estimating (XCE) Today

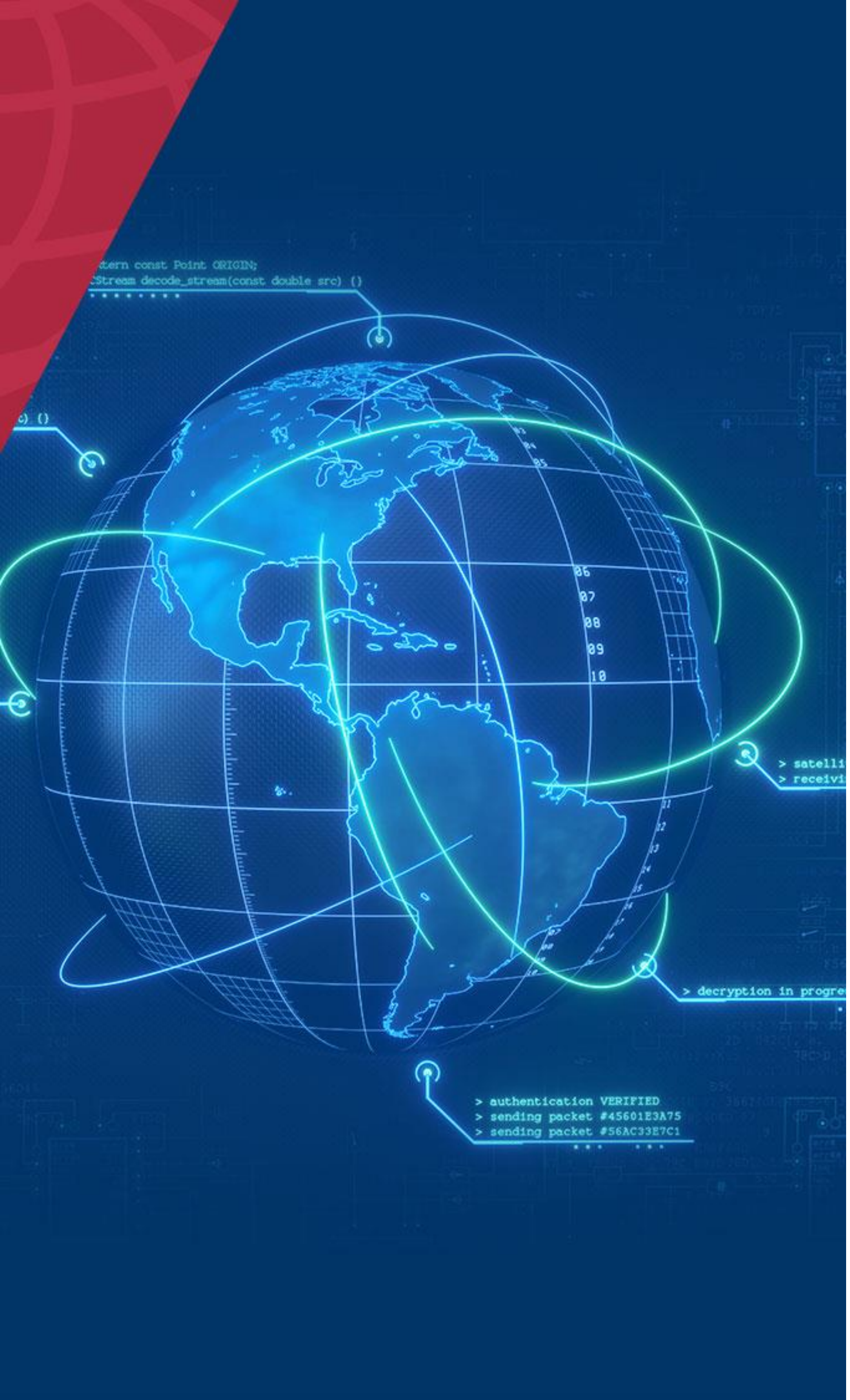


Explainable Cost Estimating (XCE) of the Future?



Will human-computer interface techniques capable of translating models into understandable explanation dialogues for the end user be necessary for us in the future?

DARPA: Explainable Artificial Intelligence (XAI), https://www.darpa.mil/program/explainable-artificial-intelligence



Final Thoughts

AI and Machine Learning offer both Opportunities and Challenges

- As our environment continues to become more data rich (e.g., CADE, ERP systems, RCE, parametric bidding), the relevance of powerful, new AI/ML techniques will increase
- Learning to effectively explain and communicate AI/ML based analyses will be critical to the success of our future efforts
  - Early adopters in other industries have struggled to realize the benefits
  - Historically, explainability has already been a challenge for the cost community
  - Can we avoid the same pitfalls of early adopters of AI/ML techniques?
    - *End consumer focus*
    - *Skills & knowledge*

What Role Might I Have in an AI Infused World?

There is room for many in the data science field

"But in the rush to grab in-demand data scientists, organizations have been hiring the most technically oriented people they can find, ignoring their ability or desire (or lack thereof) to communicate with a lay audience."

A possible solution?

"…proposes a way for those that aren't getting the most out of their operations to free data scientists from unreasonable expectations and introduce new types of workers to the mix. It relies on cross-disciplinary teams composed of members with varying talents who work in close proximity."

Harvard Business Review: Data Science and the Art of Persuasion, https://hbr.org/2019/01/data-science-and-the-art-of-persuasion?