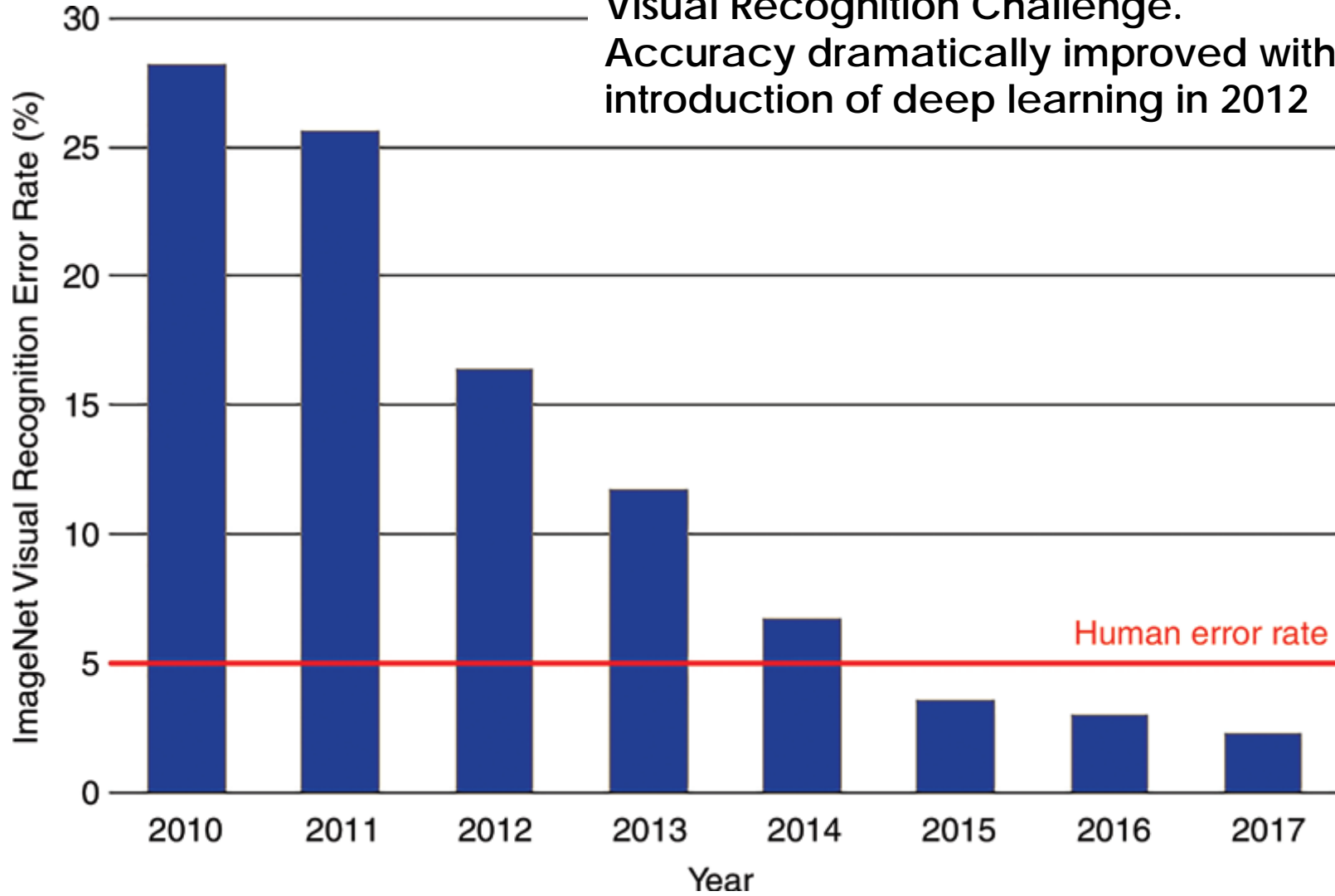# Explainable AI and Regulation in Medical Devices

David Ritscher
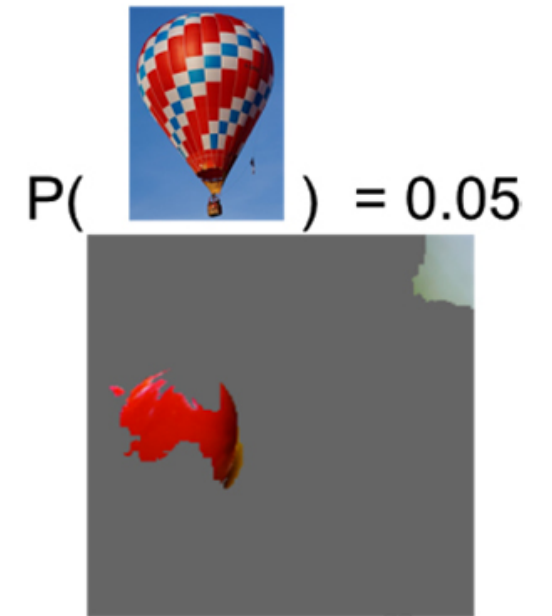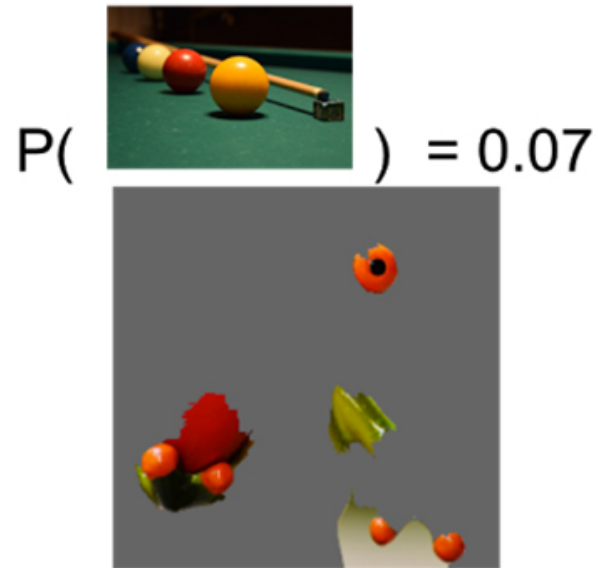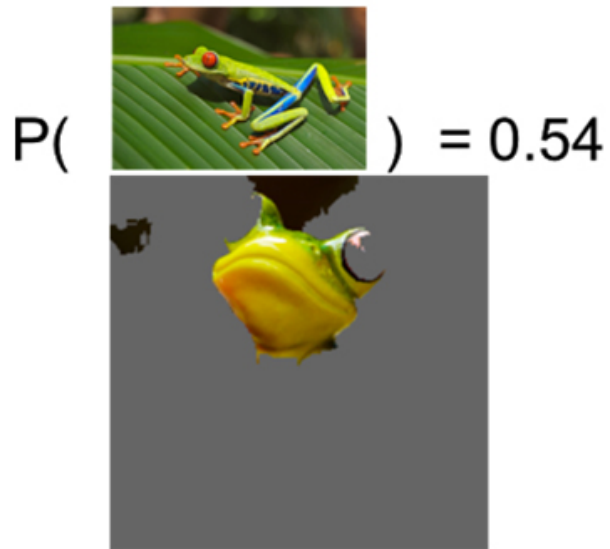
Senior Consultant

Cambridge Consultants

david.ritscher@cambridgeconsultants.com

Error rates on the ImageNet Large-Scale Visual Recognition Challenge. Accuracy dramatically improved with the introduction of deep learning in 2012

Radiology: Volume 291: Number 3—June 2019; *radiology.rsna.org*
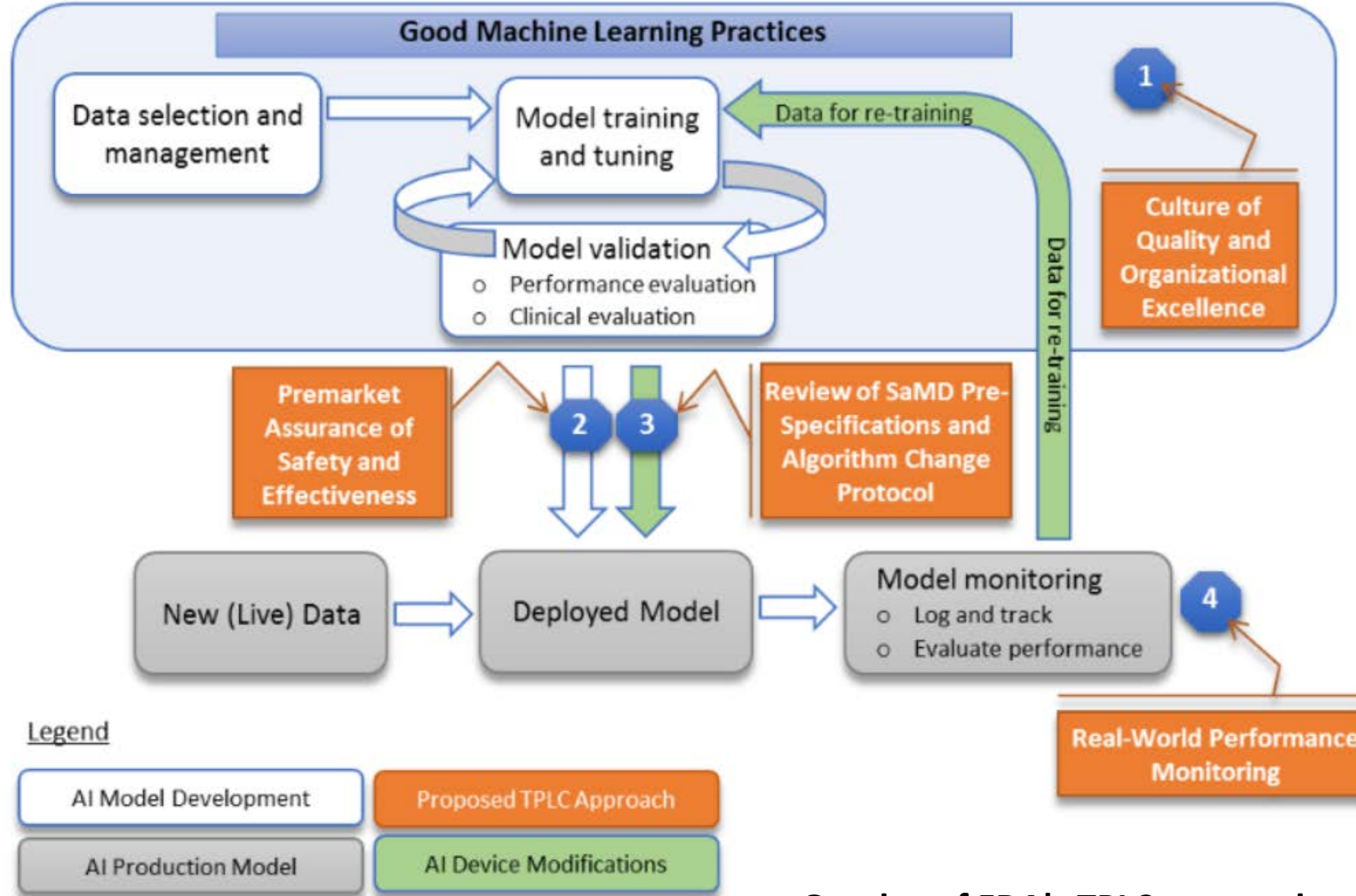
The Value of XAI (Explainable AI)

- Much of the value of AI/ML is that it can outperform rule-based algorithms
- AI is typically not explainable – it's a black box
- The incremental benefit of AI is the non-explainable part
- What does it mean to be explainable?

- XAI doesn't explain
- It provides hints
- It helps generate insights
- It allows probing of behaviors and debugging of the process

Explanation for a prediction from Inception. The top three predicted classes are "tree frog," "pool table," and "balloon."
Local Interpretable Model-Agnostic Explanations (LIME): An Introduction, Marco Tulio Ribeiro, Sameer Singh, Carlos Guestrin, Aug12, 2016, www.oreilly.com/learning/introduction-to-local-interpretable-model-agnostic-explanations-lime

Where does XAI fit in?

Overlay of FDA's TPLC approach on AI/ML workflow

XAI can provide

XAI Can provide hints on

- What is 'of interest' to the AI
- Hints about weak points of the AI

XAI does not

- Explain the AI
- Replace GMLP (Good Machine Learning Practices), which are the key to assuring safety and efficacy