THE NATIONAL ACADEMIES BOARD ON HUMAN SYSTEM INTEGRATION (BOHSI)

EXPLAINABLE AI, SYSTEM TRANSPARENCY, AND HUMAN MACHINE TEAMING

Session Chair: Emilie Roth, Roth Cognitive Engineering

Session Co-Chair: Toby Warden, National Academies

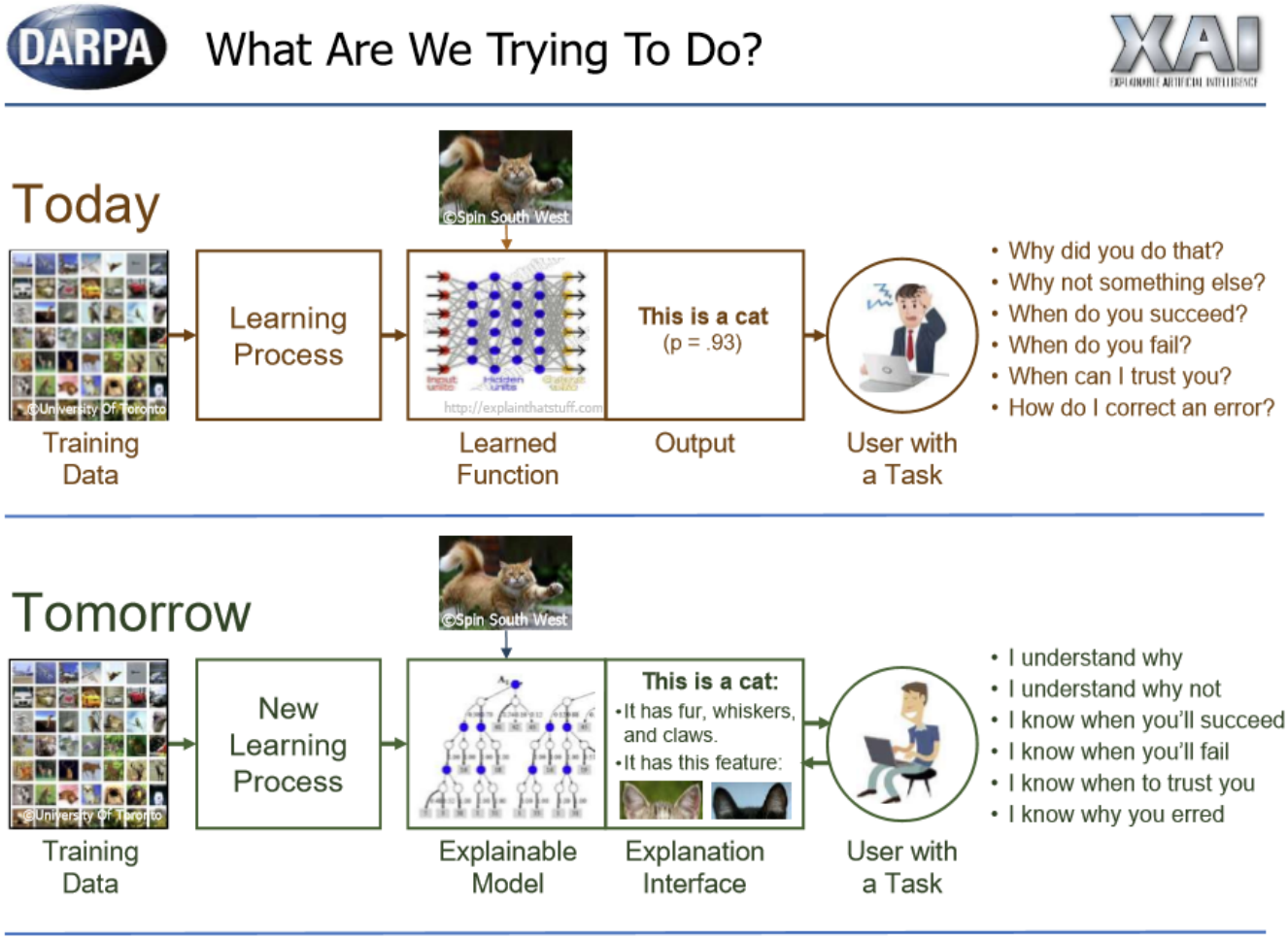
(Promise of) Application of AI is Ubiquitous

However…

Challenges Remain

- Effectiveness of AI agents cannot be judged on 'stand-alone' performance
  - It is the joint performance of Humans and Machine agents (Human Machine Teaming) that matters.
- This requires:
  - Common ground
  - A basis for developing **Appropriate** Trust

Emerging Research



Goals of Panel

- Bring together AI and Human Factors Researchers tackling these challenges
- Discuss current state of the art, challenges and shortfalls and ways forward in developing systems that engender 'appropriate' trust.
- Goal is to:
  - Explore points of consensus
  - Draw insights from unique and provocative perspectives

Panelists

Jessie Chen	Senior Research Scientist for Soldier Performance US Army Research Laboratory
William Clancey	Senior Research Scientist, Florida Institute of Human & Machine Cognition
Robert Hoffman	Senior Research Scientist, Florida Institute for Human & Machine Cognition
Mica Endsley	SA Technologies
Marc L. Steinberg	Office of Naval Research

Follow-up Questions

- Does an AI system need to be 'explainable' to be useful and usable and/or to promote appropriate trust?
- How does tempo of operations affect needs for 'transparency'/'explainability' – e.g., in a moving car or a plane – If the AI system says 'change lane' or descend … vs. after action review when there is time to be more reflective?
- Some of us can remember earlier 'waves of AI /automation' that eventually failed to live up to the hype – Are there lessons learned from those experiences that are relevant today?