Michael Hind – IBM Research

ML vs traditional SW evaluations

- Correctness?
- Testing corner cases w/o knowing the corners
- Focus on risks
  - Bias, adversarial robustness, explainability, uncertainty, privacy, generalizability, data quality, etc.

  ➔ Transparency

Experiences with enterprises

- Large variation
- AI Governance emerging

Most useful evaluation outputs

- Evaluation is an abstraction of efficacy
- Know your concerns, consumers
- Don’t measure what is easy, measure what matters

My passion

- AI FactSheets: aifs360.mybluemix.net

  –Templates, methodology, automatic collection

- Open source tools
  - aif360.mybluemix.net (Fairness)
  - aix360.mybluemix.net (Explainability)
  - art360.mybluemix.net (Adversarial Robustness)
  - uq360.mybluemix.net (Uncertainty Quantification)