Preventing Skynet
(or at least understanding why it happened)

Ben Zorn
Microsoft Research
Why understanding data and models matters

• An existential threat?
  • “…If its [function] is just something like getting rid of e-mail spam and it determines the best way of getting rid of spam is getting rid of humans …” Elon Musk, CNBC Interview, (Research Priorities for Robust and Beneficial Artificial Intelligence: an Open Letter)

• The problem: complexity, data, and human reason...
  • Errors affect: policy (Reinhart & Rogoff error), money (Spreadsheet Mistake Costs Tibco Shareholders $100 Million, WSJ 10/16/2014), health (Google Flu Trends errors, Tim Havorford “Disruptions: Data Without Context Tells a Misleading Story”, NYT 2/24/2013)

• The solution: testing, data analysis, verification of data and ML models
  • Insight: Data quality analysis cannot detach data from the logic that processes it
  • Insight: Historically, malicious attacks on code, now attacks on data/models
  • Insight: Programming Languages is the discipline to attack this problem
    • We’ve been doing it at scale in a principled way for decades
  • Examples: CheckCell (Barowy, Gochev, Berger, OOPSLA 2014)