An open-source suite of causal AI tools and libraries

Emre Kıcıman

emrek@microsoft.com

Microsoft Research

together with many others

Causal Questions in Practice

- What is the impact of a marketing campaign?
- What are the drivers of customer churn?

- How do farmer practices effect soil carbon sequestration?
- Identifying retail demand factors for inventory planning
- Prioritizing maintenance to mitigate climate change effects

- Who doesn’t benefit from a drug?
- What are long-term effects of a disease?
- How can AI models generalize across hospitals?

Observational studies at large scale

Multiple kinds of scale in practice

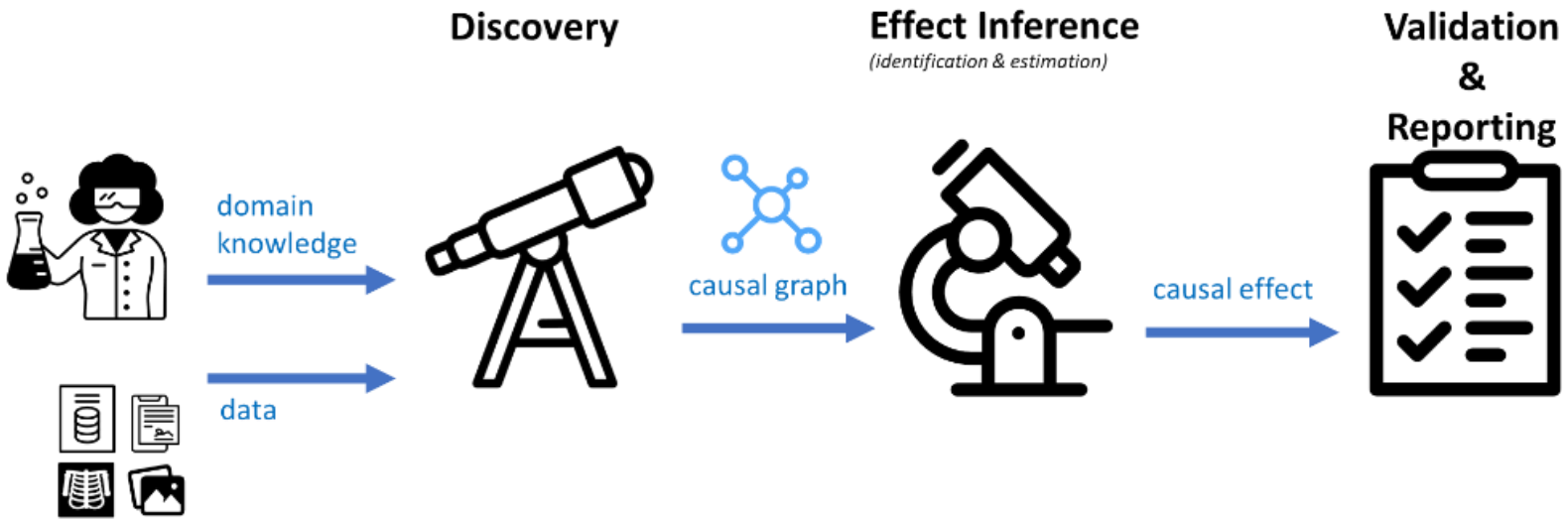
- Large data scale
- Large number of analyses in parallel
- Low latency

- Many people designing and running analyses independently

Design Principles

- Scaffolding causal analysis process for practitioners
  - Especially causal assumptions
- Separation of causal questions from algorithms and implementations
- Bridge across causal frameworks with common abstractions

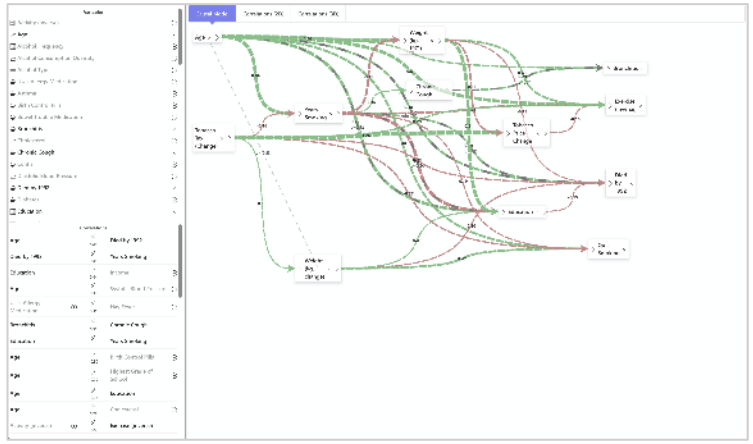
Causal AI Suite



- **DoWhy:** Scaffolds end-to-end causal analysis and best practices
- **EconML:** Latest advances in causal machine learning for effect inference
- **Causica:** Deep learning for causal discovery and end-to-end inference
- **ShowWhy:** Interactive tools for causal decision making

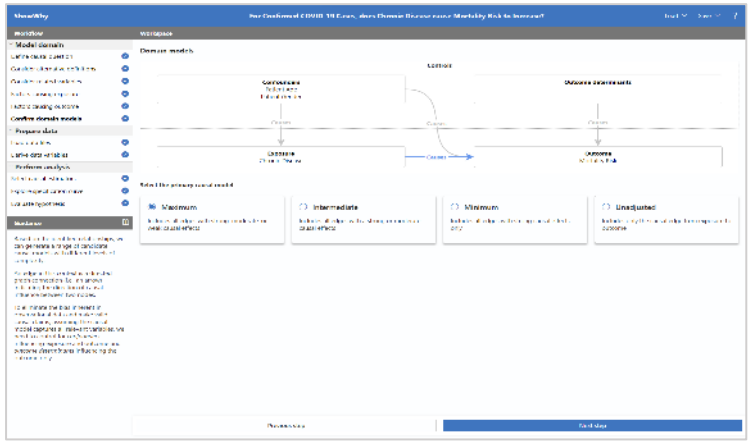
ShowWhy: Interactive tools for causal decision making

Causal discovery



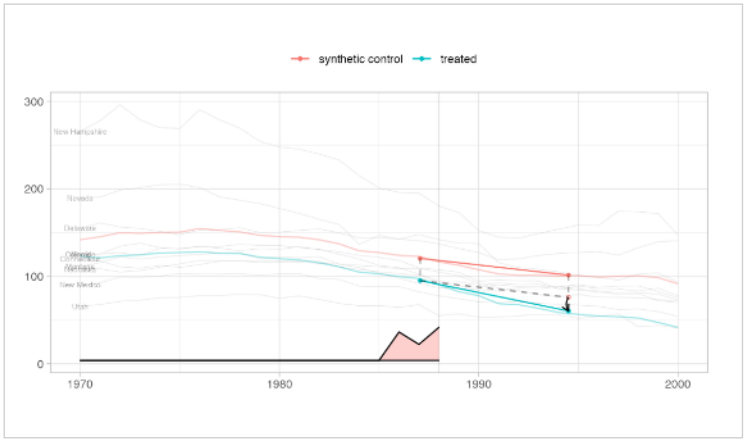
End-to-end causal discovery + inference,
Bayesian/deep learning foundation

Effect inference via IV and conditioning methods



Emulated randomized controlled trial,
Do-calculus/machine learning foundation

Synthetic controls and diff-in-diff



Synthetic difference-in-differences,
Econometrics/regression foundation

Discovering candidate interventions, deciding which interventions to make, evaluating impact of interventions over time

Research needs inspired from practical use

- Improved elicitation of domain knowledge
- Improved validation, refutation, and sensitivity analyses
- Improved support for unstructured and high-dimensional text and image data

Questions?

✉ emrek@microsoft.com @emrek@hci.social @emrek

Links:

- DoWhy/PyWhy: https://pywhy.org/
- EconML: https://github.com/microsoft/econml
- Causica: https://github.com/microsoft/causica
- ShowWhy: https://github.com/microsoft/showwhy

Team:

- Eleanor Dillon, Darren Edge, Adam Foster, Joel Jennings, Chao Ma, Robert Ness, Nick Pawlowski, Amit Sharma, Cheng Zhang, Keith Battocchi, Mónica Carvajal, Denise Chen, Nathan Evans, Andrés Morales Esquivel, Friederike Niedtner, Dayenne de Souza, Christopher Trevino, Ha Trinh, Robert King, Ahmed Mostafa, and Thomas Kapler
- Our collaborators in the PyWhy organization, Peter Götz, Patrick Blöbaum, Kailash Budhathoki, and all our open-source contributors