

CAUSAL FAULT LOCALISATION IN DATAFLOW SYSTEMS

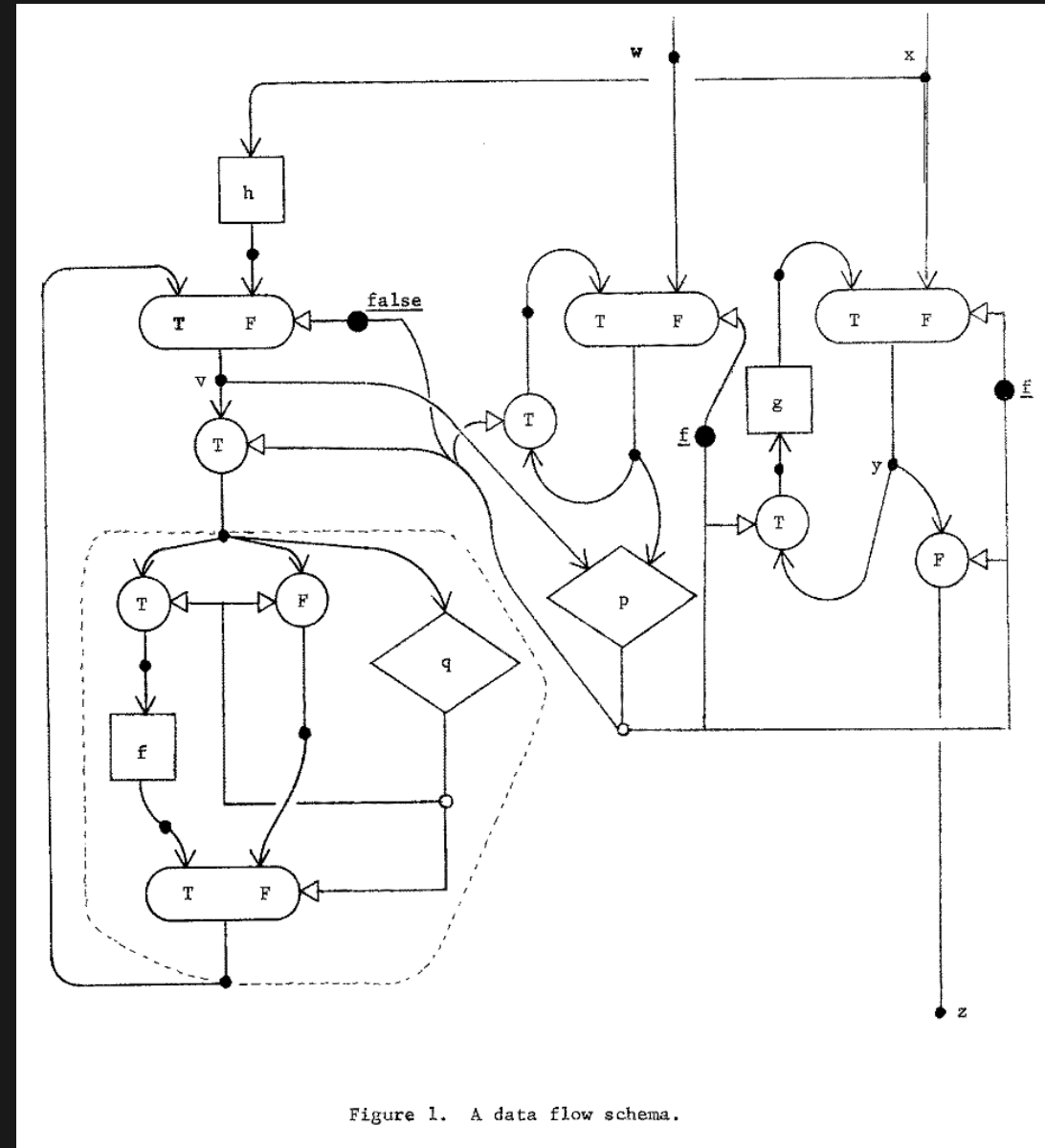
Andrei Paleyes and Neil Lawrence

University of Cambridge

EuroMLSys

May 2023

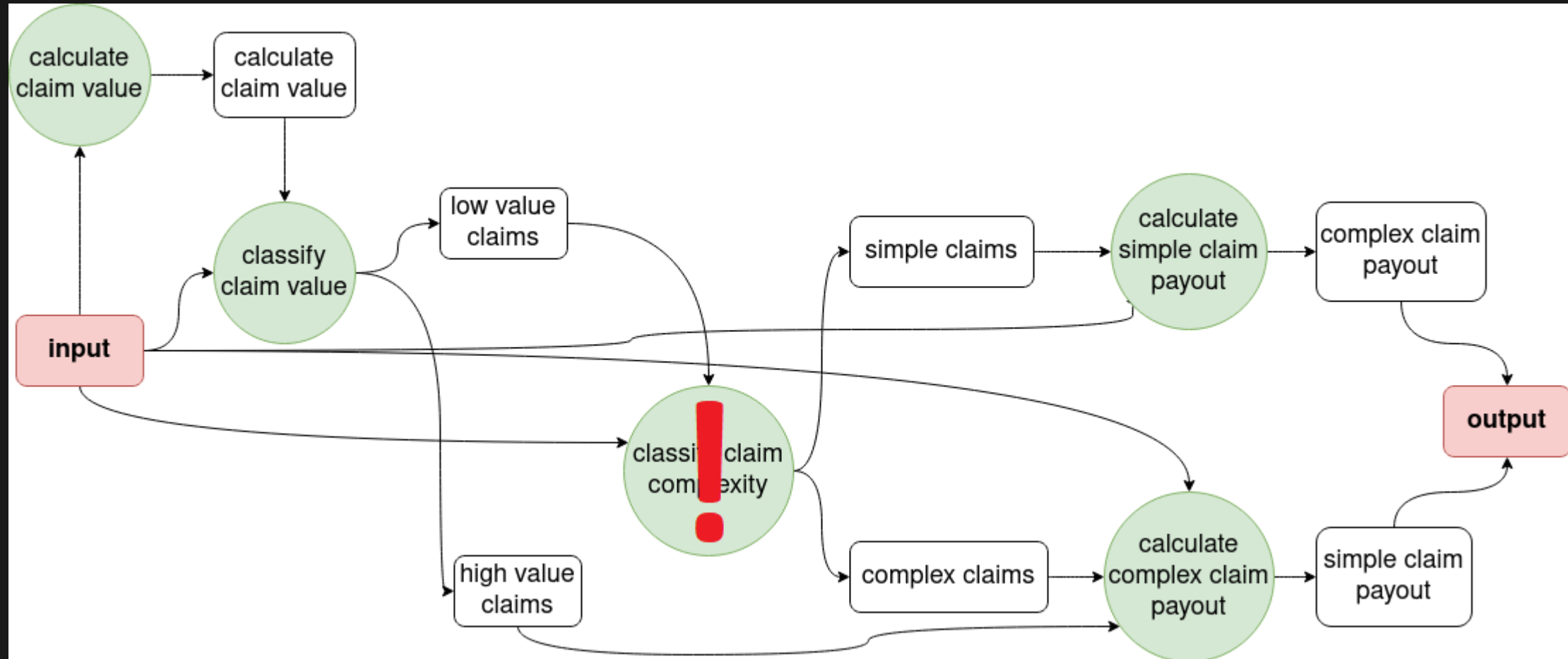
WHAT IS DATAFLOW ARCHITECTURE?



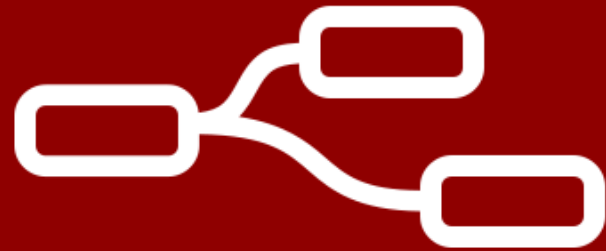
*“Dataflow graph can be used as complete causal graph,
with implications for debugging, business analysis and
experimentation.”*

"Dataflow graphs as complete causal graphs", Paleyes et al., CAIN 2023

OUR GOAL



DEMOS AND EXPERIMENTS



Node-RED

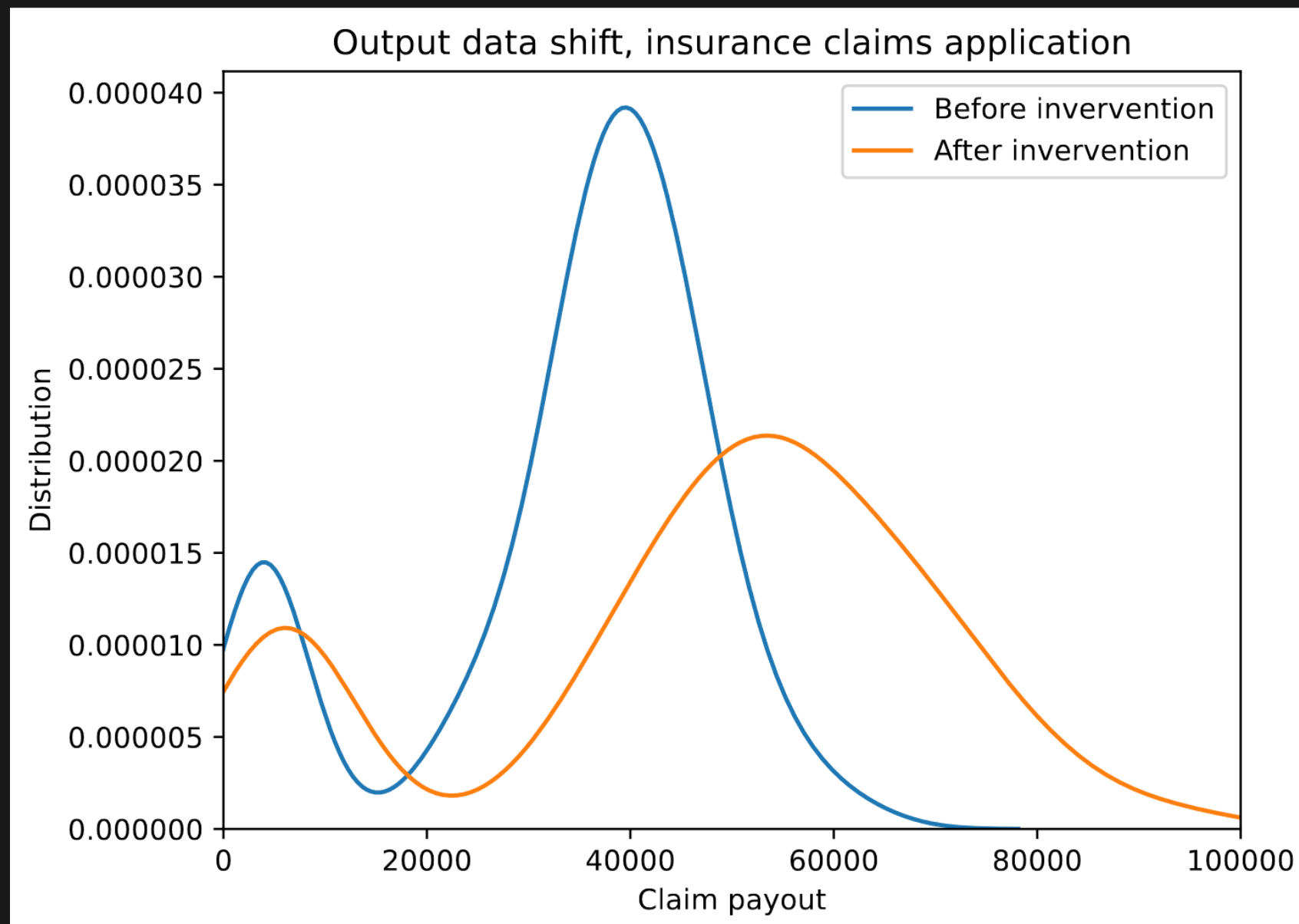


SciPipe

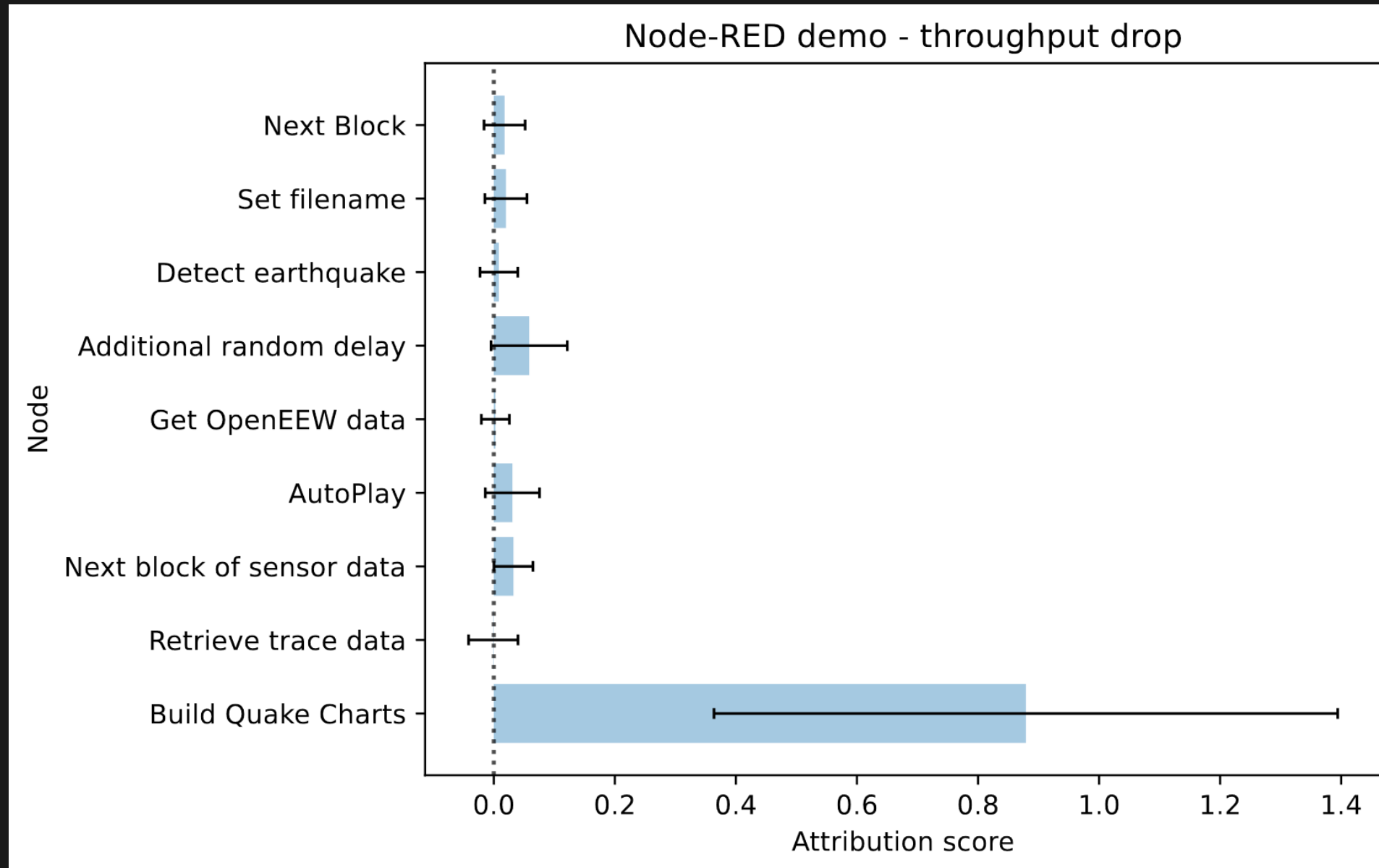


SELDON
CORE

DEMOS AND EXPERIMENTS



DEMOS AND EXPERIMENTS



FUTURE WORK

- Scale up graph size
- Estimate data storage costs
- Develop fully automated plugins
- Demos for business analysis and experimentation

THANK YOU!

Paper: <https://arxiv.org/abs/2304.11987>

Code: <https://github.com/apaleyes/causality-fbp>

Questions?

<https://mlatcl.github.io>

✉ ap2169@cam.ac.uk

