

MACHINE LEARNING in SCIENCE

BRIDGING DATA-DRIVEN & MECHANISTIC MODELLING APPROACHES



SNAPSHOTS
of AI in
SCIENCE



BUILDING
EFFECTIVE
SIMULATIONS

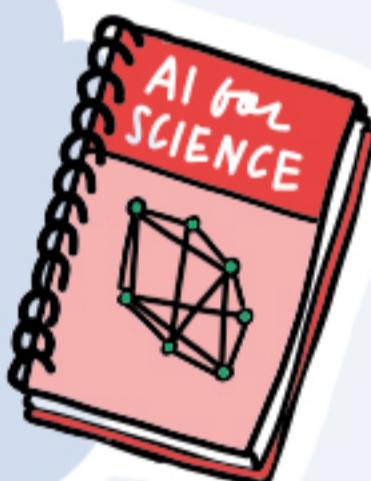


CONNECTING
DATA to
CAUSALITY



ENCODING
DOMAIN
KNOWLEDGE

an
EMERGING
RESEARCH
AGENDA



the ML for
SCIENCE
ECOSYSTEM



THIS CONFERENCE
WAS HELD at
SCHLOSS DAGSTUHL
19 - 23 SEPTEMBER,
2022.



AI for SCIENCE
ROADMAP

PARTICIPANTS' TALKS

LEARNING from
RECENT EXPERIENCES

CONNECTING
DATA to CAUSALITY

BUILDING EFFECTIVE
SIMULATIONS

ENCODING DOMAIN
KNOWLEDGE



CLICK EACH TITLE
to go to PAGE



UNIVERSITY OF
CAMBRIDGE

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



NYU



WISCONSIN
UNIVERSITY OF WISCONSIN-MADISON

REAL WORLD APPLICATIONS

EXAMPLES of MACHINE LEARNING APPLICATIONS in SCIENCE can be FOUND ACROSS a WIDE SPECTRUM of SCIENTIFIC STUDY - from ATOMIC to ASTRONOMICAL!

IN EARTH SCIENCES...

INTEGRATING OBSERVATIONAL DATA w/ PHYSICS-INFORMED MODELLING - ACROSS SCALES



THIS ABILITY can HELP us BETTER UNDERSTAND the IMPLICATIONS of CLIMATE CHANGE!

TOOL-KITS to MAKE ICE-MELT MODELS more ACCURATE

HYBRID MODELS can HELP FORECAST the IMPACT of CLIMATE CHANGE on DIFFERENT LANDSCAPES

DIFFERENT DOMAINS have SPECIFIC NEEDS from ML TOOLS!

IN PHYSICAL SCIENCES...

NEURAL NETWORKS can HELP w/ COMPLEX PREDICTIONS

ML TOOLS help FIND NEW APPROACHES to FIND APPROXIMATE NUMERICAL SOL. in DIFFUSION MODELING

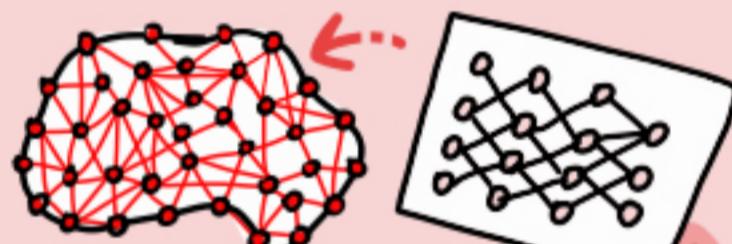


THIS is USEFUL ACROSS MULTIPLE DOMAINS!

IN BIOLOGICAL SCIENCES...

COMBINING STATISTICAL & MECHANISTIC APPROACHES to RECONSTRUCT GENE DYNAMICS

ALLOWS RESEARCHERS to PREDICT CELL SPECIALIZATION & ASSOC. GENETIC CHANGES!



STREAMLINING the MODEL DEFINITION PROCESS to RAPIDLY DEVELOP SIMULATIONS of COMPLEX STRUCTURES like BRAINS or NERVOUS SYSTEMS

IN ENVIRONMENTAL SCIENCES...

REQUIRES CAREFUL CALIBRATION of INPUT SOURCES



COLLABORATING ACROSS DOMAINS & FROM FARMERS to BUILD a ROBUST SYSTEM for ANALYZING POULTRY FECAL SAMPLES



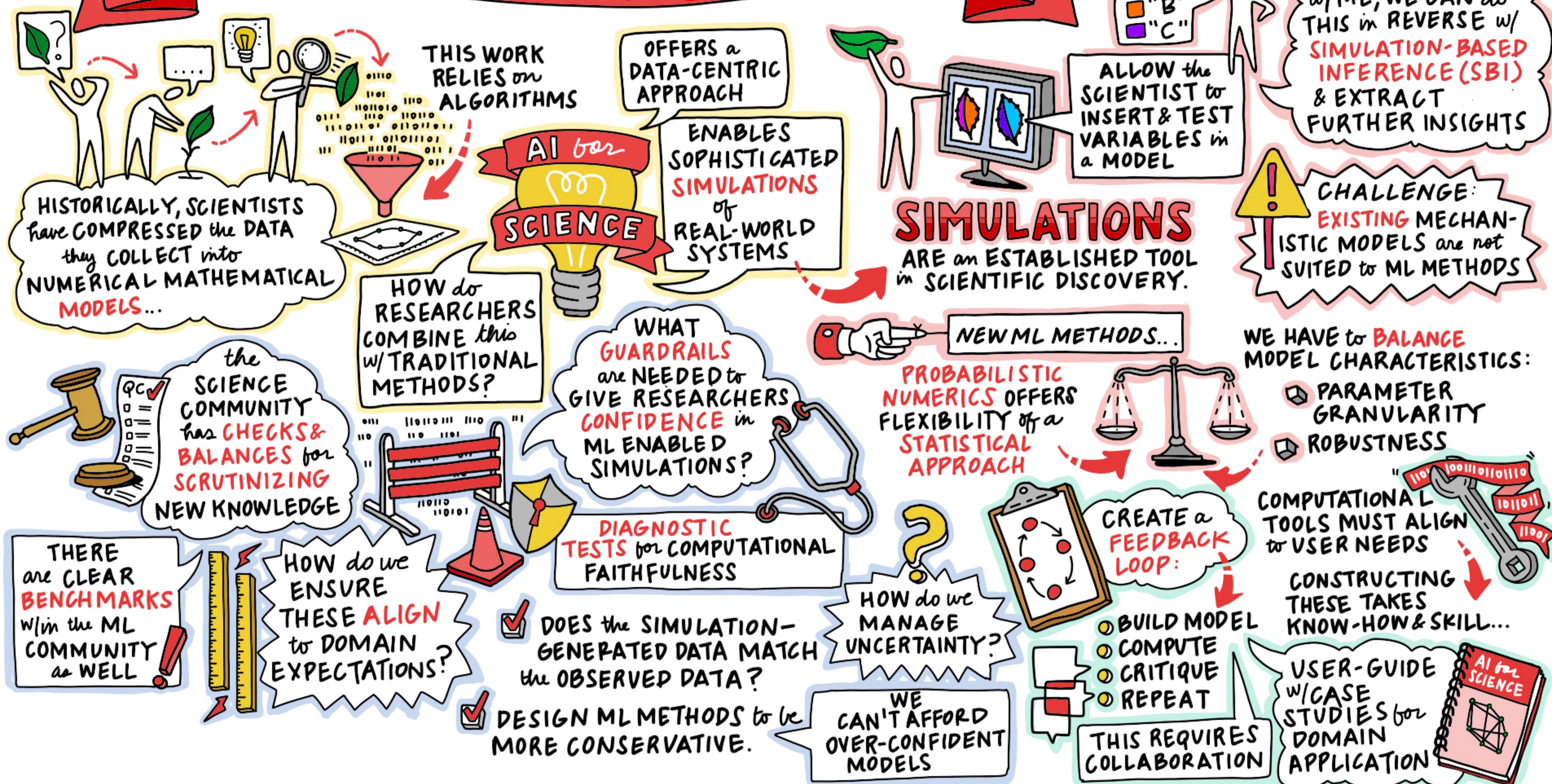
APPLYING ML to ANALYZE MULTIPLE INPUTS to ESTIMATE # & BIOMASS of TREES

USING ML INSIGHTS from SATELLITES to MANAGE VECTOR BORNE ILLNESSES

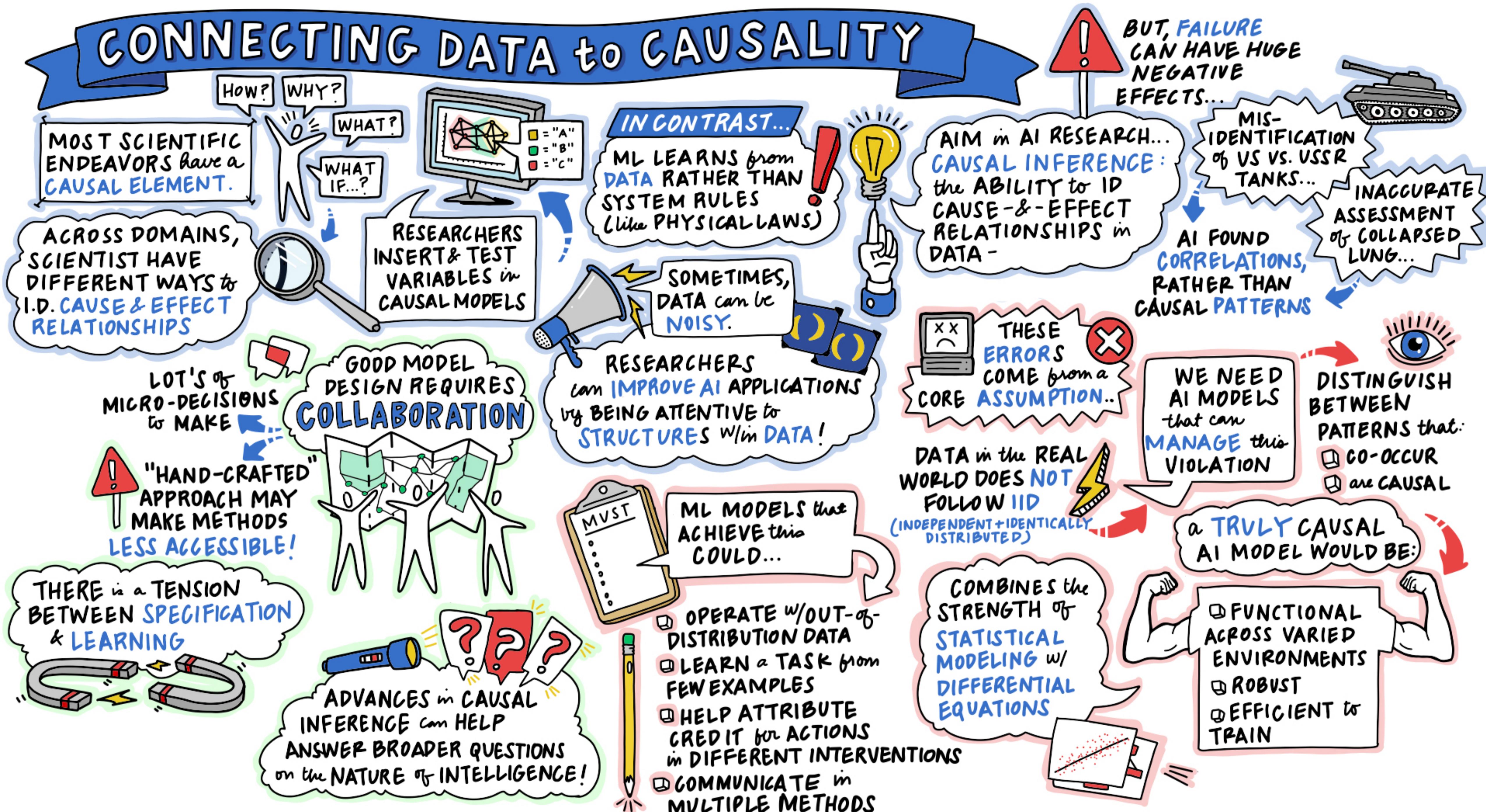
MODELS have INHERENT ASSUMPTIONS, WHICH SHOULD be made TRANSPARENT to USER GROUPS



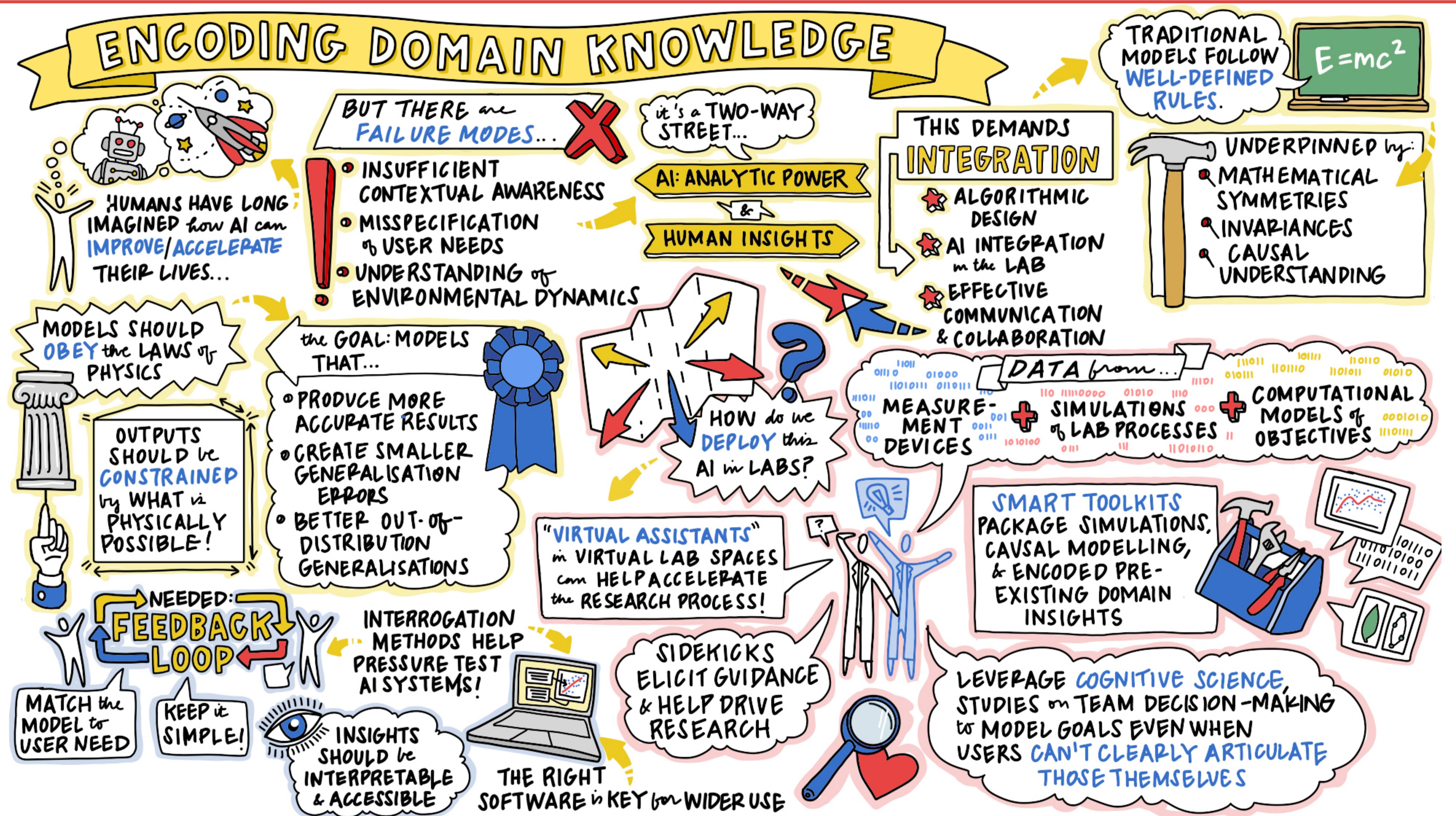
BUILDING EFFECTIVE SIMULATIONS



CONNECTING DATA to CAUSALITY



ENCODING DOMAIN KNOWLEDGE



A RESEARCH AGENDA in AI for SCIENCE



COMBINING HUMAN & MACHINE INTELLIGENCE

WHICH INSIGHTS are RELEVANT, how to BEST COMMUNICATE them, & the CULTURAL ENVIRONMENT of the SCIENCE CONDUCTED!

AREAS of PROGRESS..

- ★ DESIGNING PRECISE INTERFACES
- ★ BUILDING MECHANISMS to INTERROGATE AI SYSTEMS
- ★ ACCELERATING PACE of KNOWLEDGE CREATION & USE



UNIVERSITY OF
CAMBRIDGE

HOME



EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



NYU



WISCONSIN
UNIVERSITY OF WISCONSIN-MADISON

BUILDING AI SYSTEMS for SCIENCE

THERE are GAPS in our CURRENT CAPABILITIES

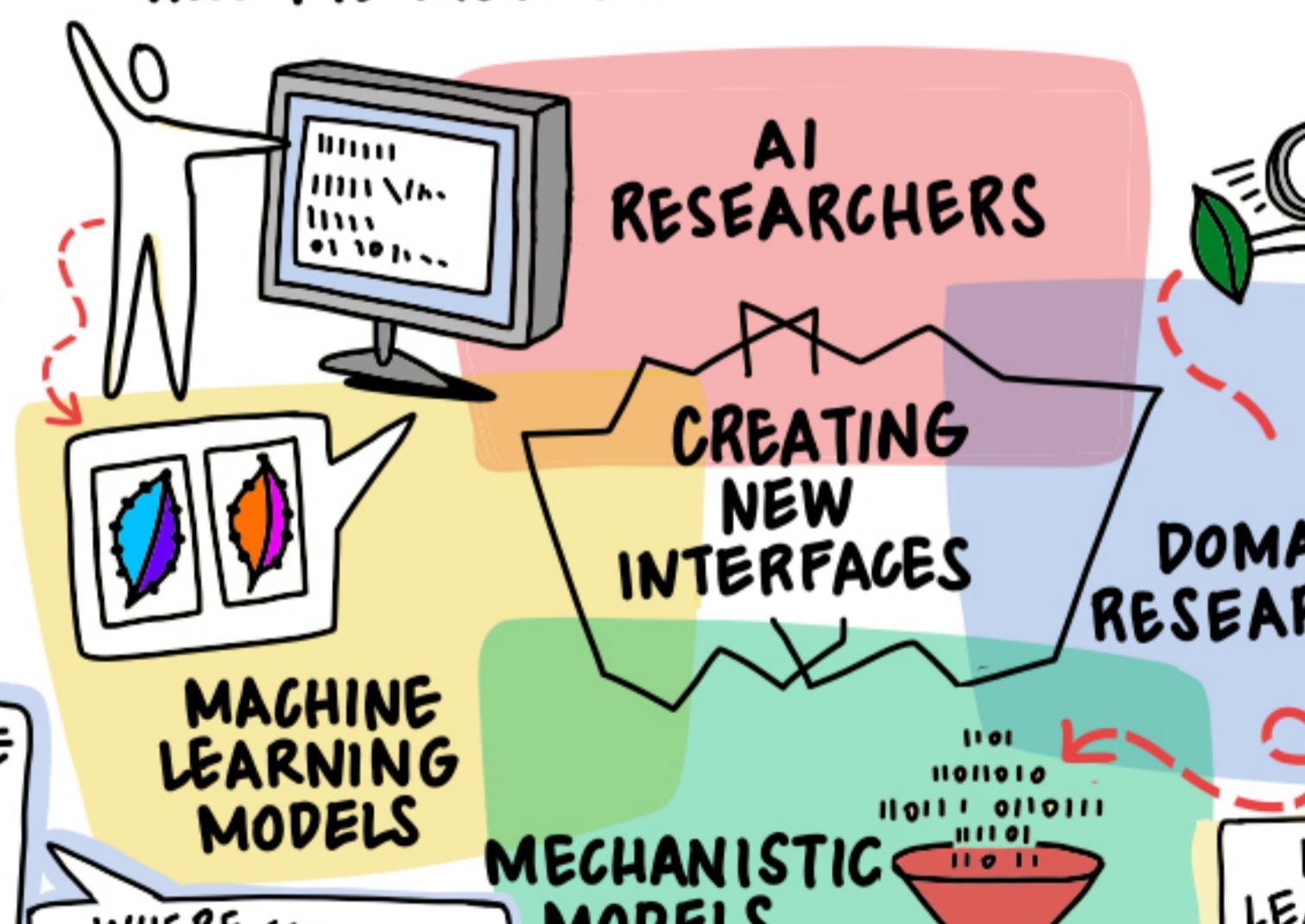
& w/a CORE SHARED INTEREST:



LEARNING from DATA!

AREAS of PROGRESS..

- ★ ADVANCING METHODS & TOOLKITS
- ★ INTEGRATE SCIENTIFIC LAWS into ML MODELS
- ★ ADVANCE CAUSAL MACHINE LEARNING



- HOW can AI CAPTURE the QUALITATIVE UNDERSTANDING that RESEARCHERS have in THEIR DOMAIN to APPLY in a SYSTEM?
- WHERE can AUTOMATION SUPPORT RESEARCH PROGRESS?
- WHAT MIX of AI DESIGN, ENGINEERING, SOCIAL INTERACTION & EDUCATION can MAKE EFFECTIVE INTERFACES for RESEARCHERS & AI SCIENTISTS?

KEY QUESTIONS to CONSIDER...

WHAT is the COMPUTATIONAL COST of COMPLEXITY, & WHAT METHODS can HELP MANAGE THIS?

HOW can 'DIGITAL SIBLINGS' be USED to EXPLORE the IMPACT of DIFFERENT INTERVENTIONS on COMPLEX SYSTEMS?

HOW can AI RESEARCHERS BUILD MEANINGFUL MODELS from DATA to ACCURATELY REPRESENT CAUSAL MECH. in a SYSTEM of STUDY?

WHAT DOES it MEAN to UNDERSTAND a MODEL? HOW can RESEARCHERS COMBINE EXPLAINABILITY w/ COMPLEXITY

HOW can ADVANCES in SIMULATION METHODS be APPLIED in DOMAINS WHERE the SYSTEM at HAND is LESS EASILY DESCRIBED by EQUATIONS?

WE MUST PROMOTE UPTAKE & PROGRESS for this WORK!

INFLUENCING PRACTICE & ADOPTION

AREAS of PROGRESS..

- ★ PROMOTE INTER-DISCIPLINARY COLLABORATIONS
- ★ DEVELOP TOOLKITS & USER GUIDES to HELP RESEARCHER APPLY APPROPRIATE AI TOOLS

- HOW do we BRING AI TOOLS out of the AI COMMUNITY & INTO "the LAB"?
- WHICH AI TOOLS are SUITABLE for WHICH PURPOSES, DISCIPLINES, or EXPERIMENTAL DESIGNS? IS IT POSSIBLE to CREATE a TAXONOMY for SCIENCE?



ML for SCIENCE ECOSYSTEM

UNIVERSITY of CAMBRIDGE:
the ACCELERATE PROGRAMME
for SCIENTIFIC DISCOVERY is
BUILDING BRIDGES ACROSS
DISCIPLINES, CREATING a
COMMUNITY PASSIONATE ABOUT
OPPORTUNITIES for AI in SCIENCE.

UNIVERSITY of COPENHAGEN:
SCIENCE AI CENTRE
PROVIDES a FOCAL POINT AI
RESEARCH & EDUCATION in its
FACULTY for SCIENCE.

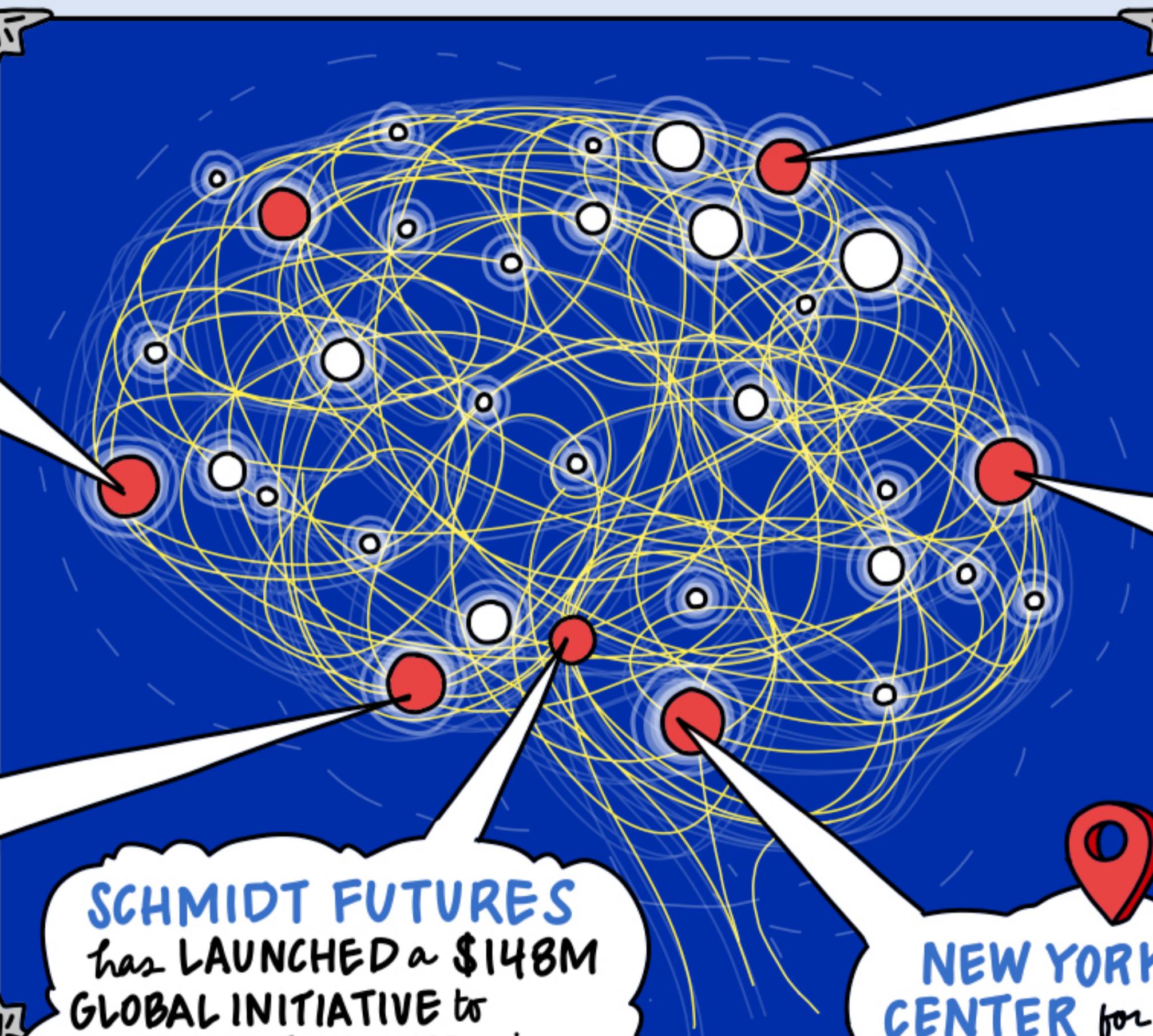
THIS MAP is INTERACTIVE.
CLICK EACH BUBBLE to VISIT
EXTERNAL LINKS of INTEREST

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



NYU

WISCONSIN
UNIVERSITY OF WISCONSIN-MADISON



SCHMIDT FUTURES
has LAUNCHED a \$148M
GLOBAL INITIATIVE to
ACCELERATE AI USE in
POSTDOCTORAL RESEARCH.
CLICK to LEARN MORE about
the 1st COHORT of SCHOOLS.

NEW YORK UNIVERSITY:
CENTER for DATA SCIENCE
HOSTS INTERDISCIPLINARY
FACULTY PURSUING INNOVATIVE
RESEARCH & EDUCATION.



UNIVERSITY of TÜBINGEN:
the ML LEARNING for SCIENCE
CLUSTER is LEVERAGING LOCAL
STRENGTHS in AI to DRIVE WIDER
PROGRESS in RESEARCH &
INNOVATION.



UNIVERSITY of WISCONSIN-
MADISON: AMERICAN FAMILY
INSURANCE DATA SCIENCE
INSTITUTE is DEVELOPING
STRATEGIC PARTNERSHIPS to
ACCELERATE the USE of DATA
SCIENCE in RESEARCH.

AI for SCIENCE ROADMAP

INVEST in TOOLS & TOOLKITS

MAKE KNOW-HOW ACCESSIBLE!

WE MUST CHART a PATH BETWEEN the STATISTICIAN & the MATHEMATICIAN



INVEST in TOOLKITS that SUPPORT GENERALISATION & EFFECTIVE APPROACHES

"OFF-the-SHELF" SOLUTIONS
CLEAN, EASY-to-USE INTERFACE for USER!

DEVELOP PIPELINES for DATA MGMT

- ★ BEST PRACTICES
- ★ FRAMEWORKS

DEVELOP BEST PRACTICES in SOFTWARE ENGINEERING & USER-GUIDES for NON-EXPERTS



BUILD CAPABILITY across DISCIPLINES

NEEDED

RESEARCHERS,
ENGINEERS, & CONVENERS
w/a PASSION for ADVANCING
SCIENCES THROUGH AI.



PROVIDE ACCESS to LEARNING & DEVELOPMENT ACTIVITIES in ML

TRAINING

TUTORIALS

USER GUIDES

CREATE INSTITUTIONAL INCENTIVES
RECOGNIZE & REWARD SUCCESSES

BUILD CAPABILITIES in COMMUNICATION, ORGANISATION, & CONVENING.

AI for SCIENCE is a RENDEZVOUS POINT

IT'S NEXT WAVE of DEVELOPMENT will COME from the STRENGTH of its DIVERSE COMMUNITY!

GROW COMMUNITIES of RESEARCH & PRACTICE

COMMUNITIES of RESEARCH & PRACTICE are the BACKDROP of this EFFORT!



CREATE NEW COMMUNICATION CHANNELS



FORUMS
NEWSLETTERS
RESEARCH SYMPOSIA

DEVELOP "AI for SCIENCE" JOURNALS



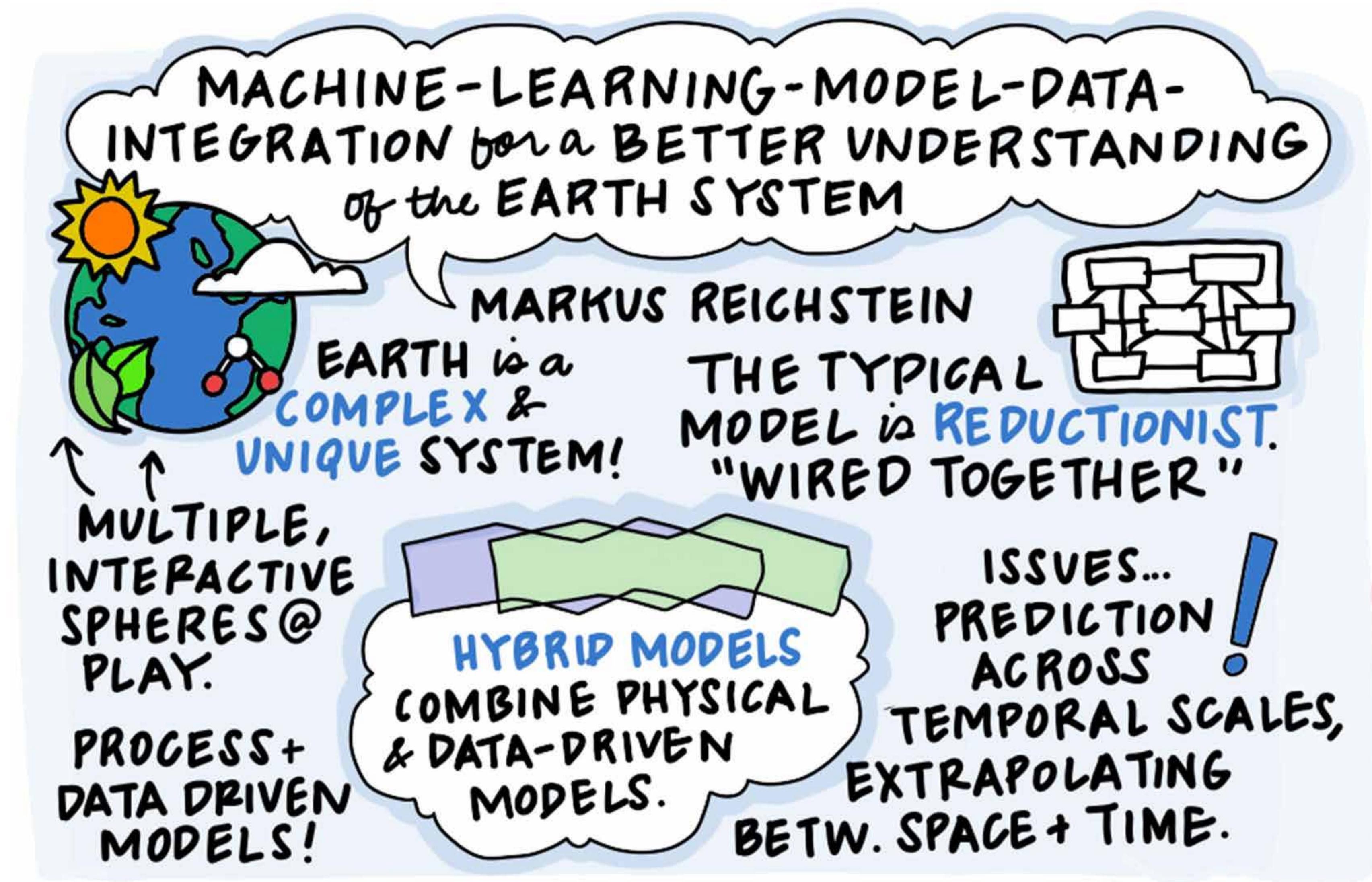
ESTABLISH MEMBERSHIP ORGANISATIONS

to FOSTER COMMUNITY & SUPPORT CAREER DEV.



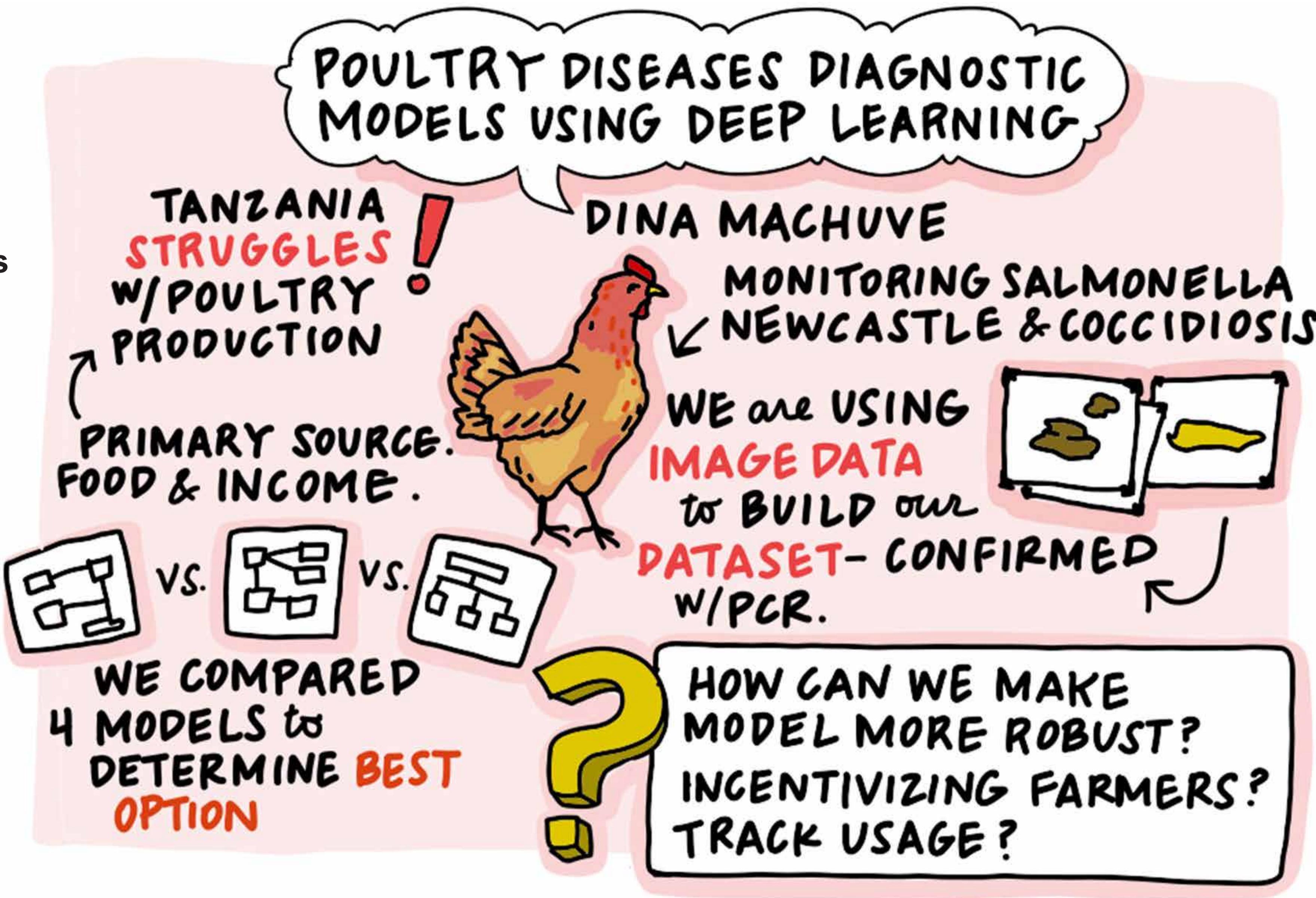
Machine - Learning - Model - Data - Integration for a Better Understanding of the Earth's System

Markus Reichstein
MPI für Biogeochemistry
Max Planck Institute
Jena, DE



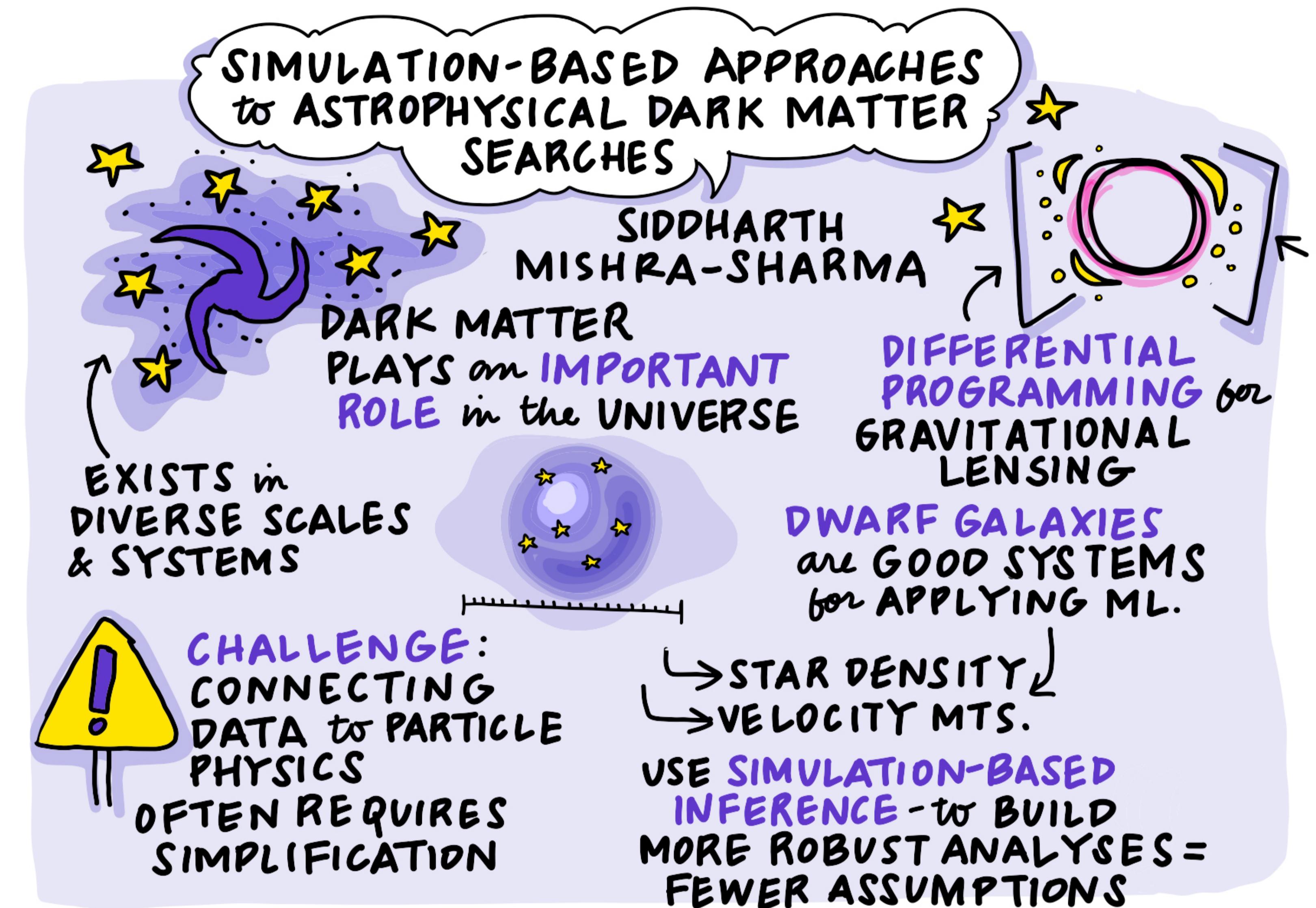
Poultry Diseases Diagnostics Models using Deep Learning

Dina Machuve
DevData Analytics
Arusha, TZ



Finding new physics: Simulation-based Approaches to Astrophysical Dark Matter Searches

Siddharth Mishra-Sharma
MIT
Cambridge, US



Single-cell Transcriptomics

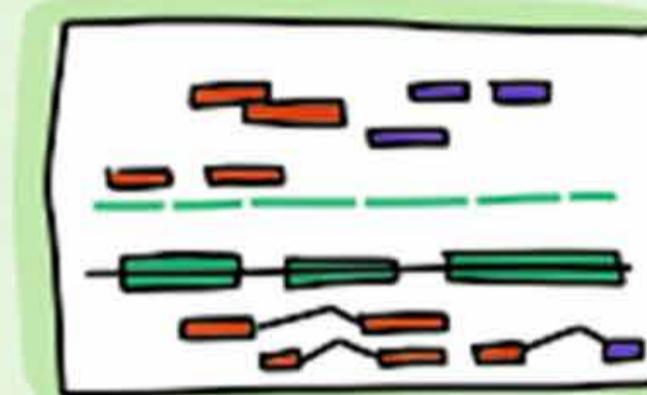
Maren Büttner

Helmholtz Zentrum München &
Universität Bonn
München, Bonn, DE

SINGLE-CELL TRANSCRIPTOMICS

MAREN BÜTTNER

CELLS are
INCREDIBLY
COMPLICATED
SYSTEMS



WE ANALYZE &
VISUALIZE THIS
DATA, & MAKE
a VARIETY of
INFERENCES



MULTIPLE AREAS
to APPLY ML:

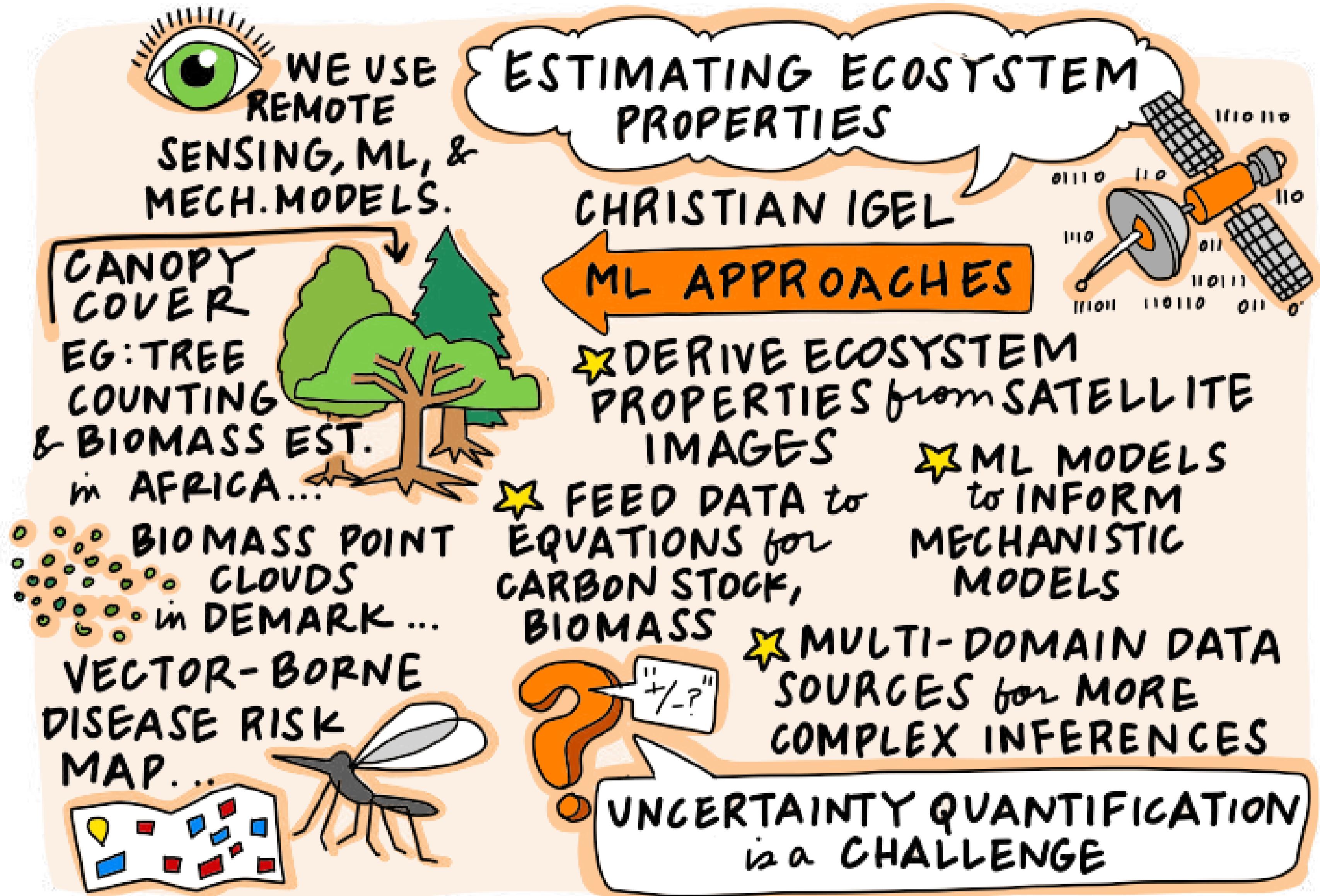
- DATA CORRECTION
in BATCHES + CLUSTERING
- QC of DATA INTEGRATION
- RECONSTRUCT CELL
IMAGES, LOCATION,
EVOLUTION, FATE



Estimating Ecosystem Properties

Christian Igel

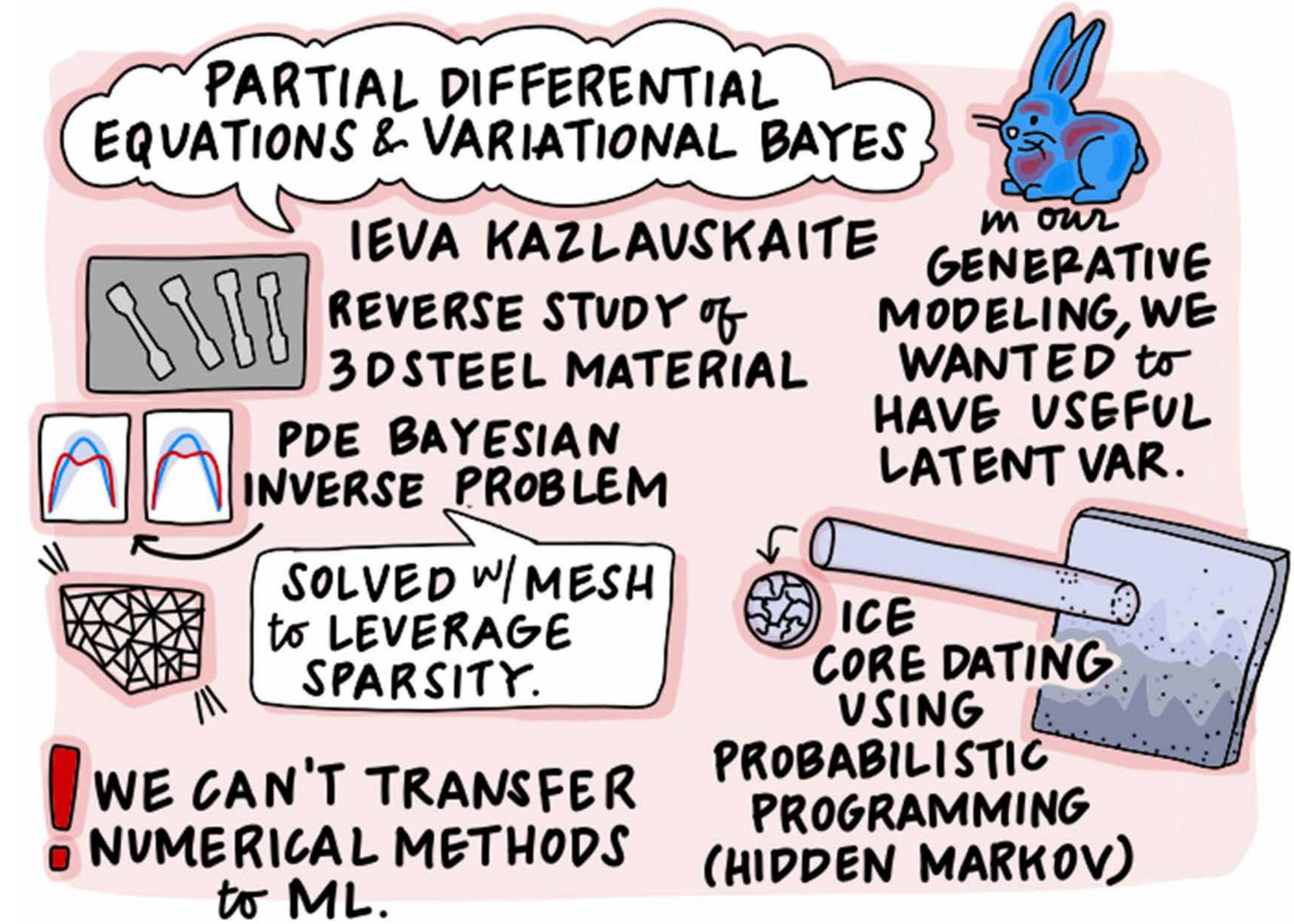
University of Copenhagen
Copenhagen, DK



Partial Differential Equations and Variational Bayes

Ieva Kazlauskaitė

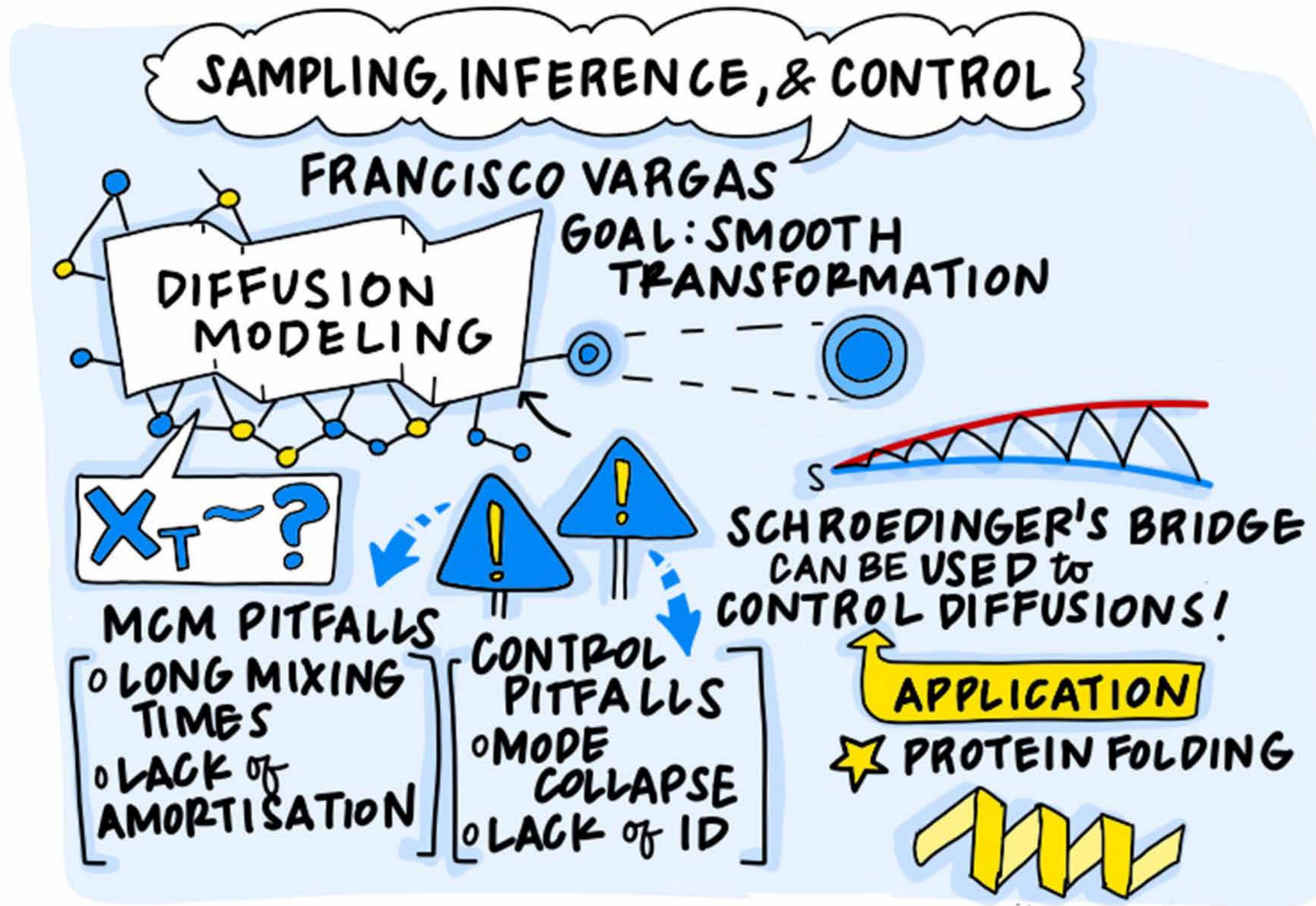
University of Cambridge
Cambridge, GB



The Schrödinger Bridge Problem

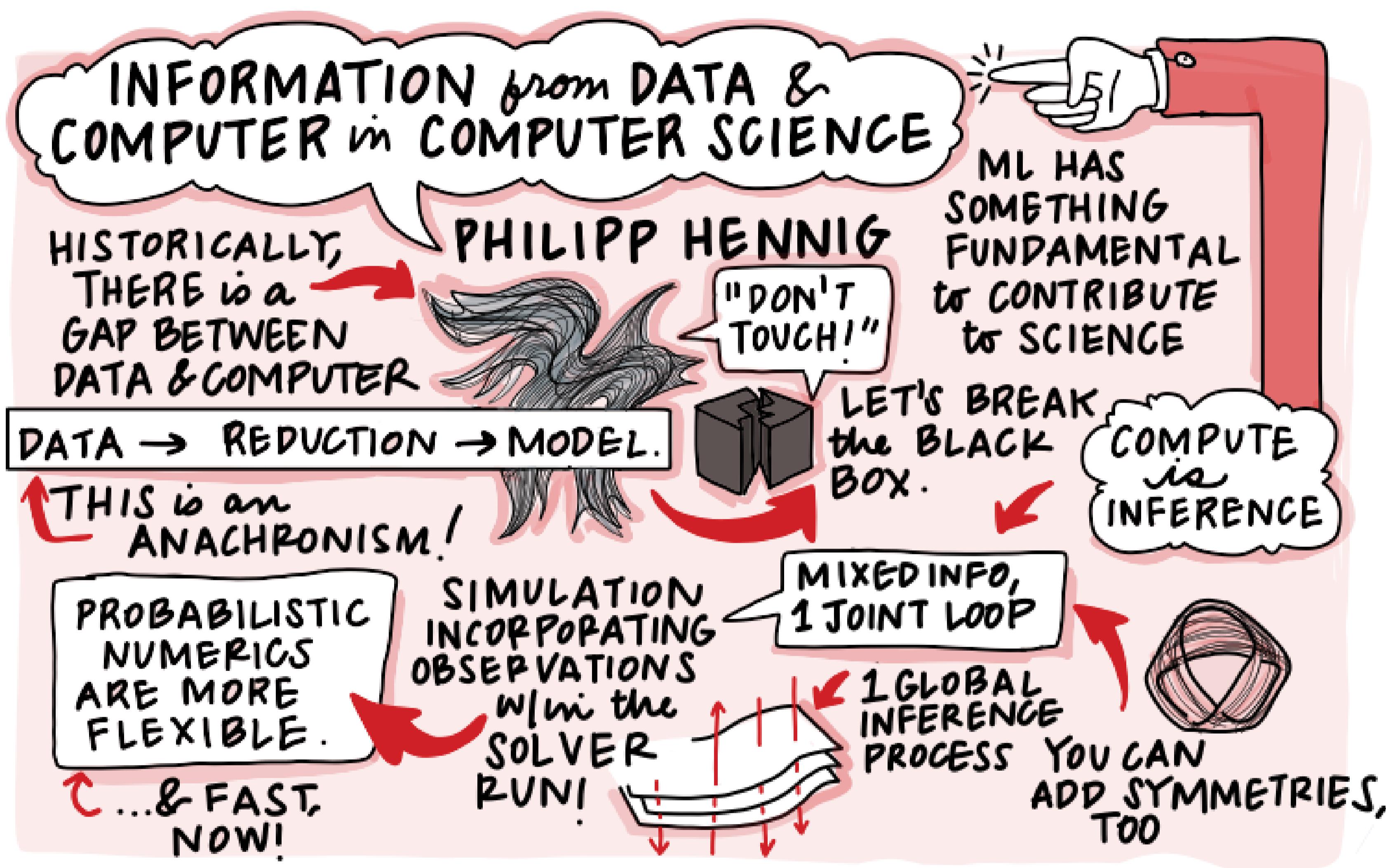
Francisco Vargas

University of Cambridge Cambridge,
GB



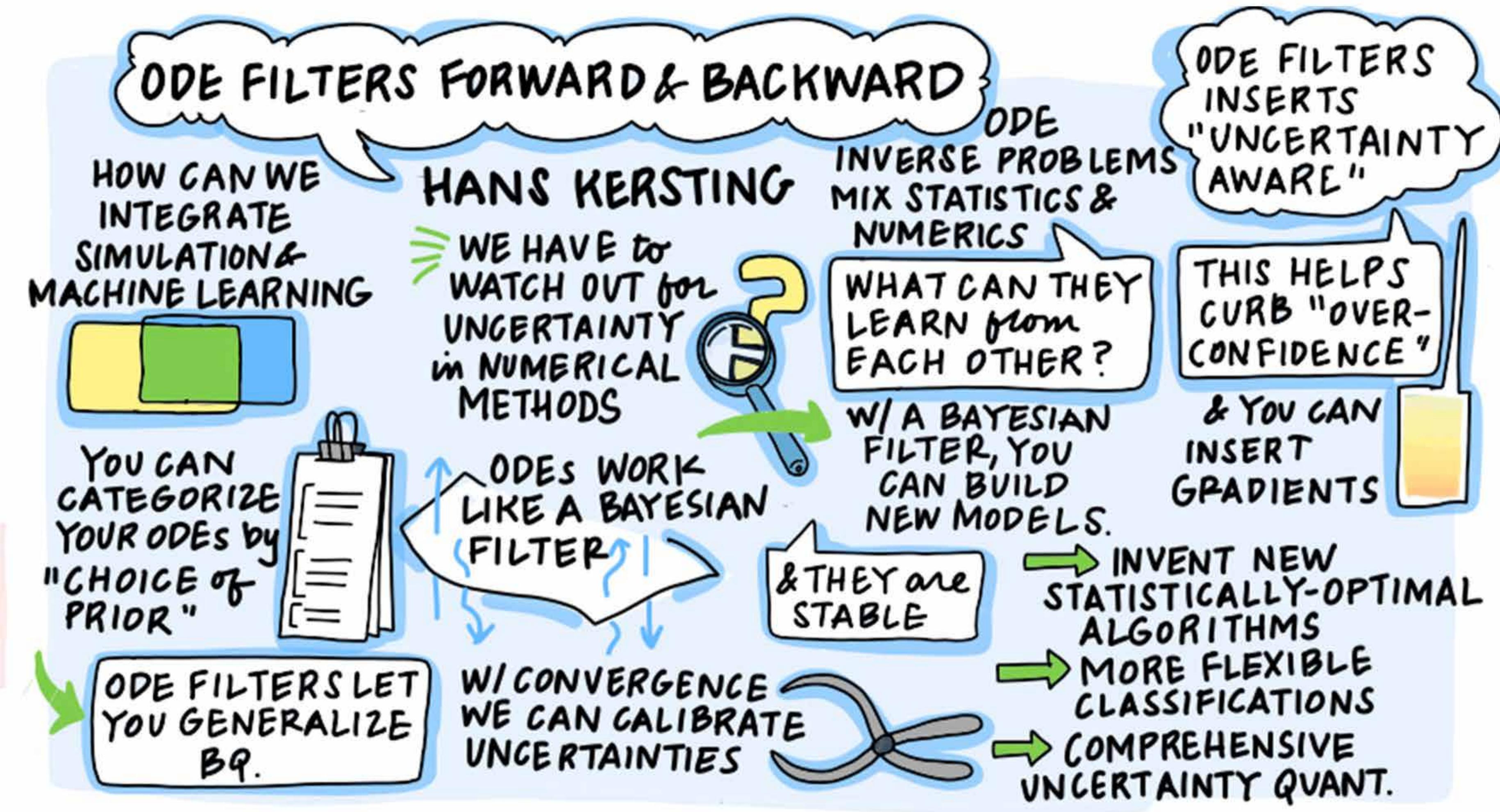
Simulation and Scientific Computing

Philipp Hennig
Universität Tübingen
Tübingen, DE



ODE Filters and Smoothers: Probabilistic Numerics for Mechanistic Modelling

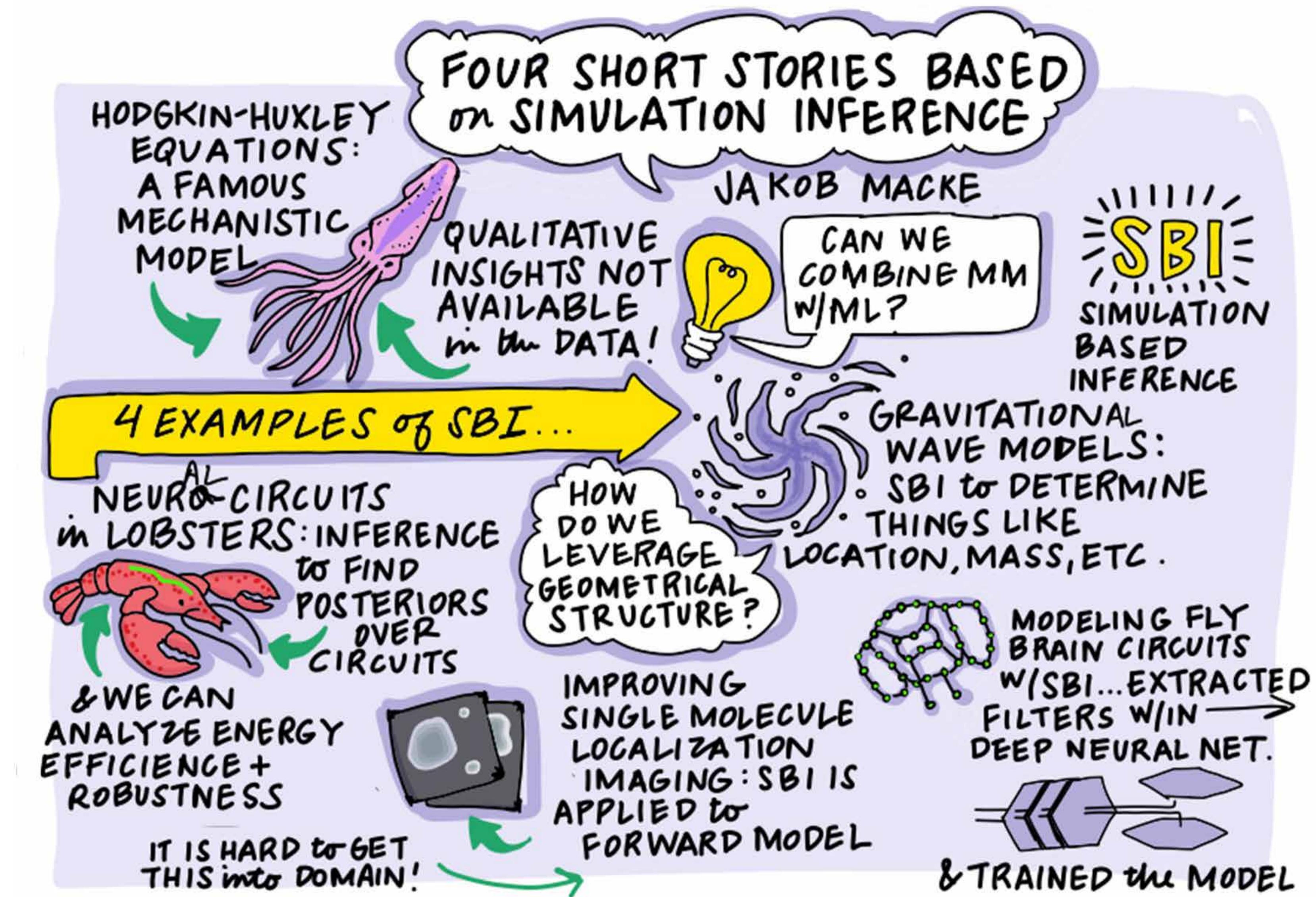
Hans Kersting
INRIA
Paris, FR



Simulation Based Inference for Scientific Discovery: Opportunities & Challenges

Jakob Macke

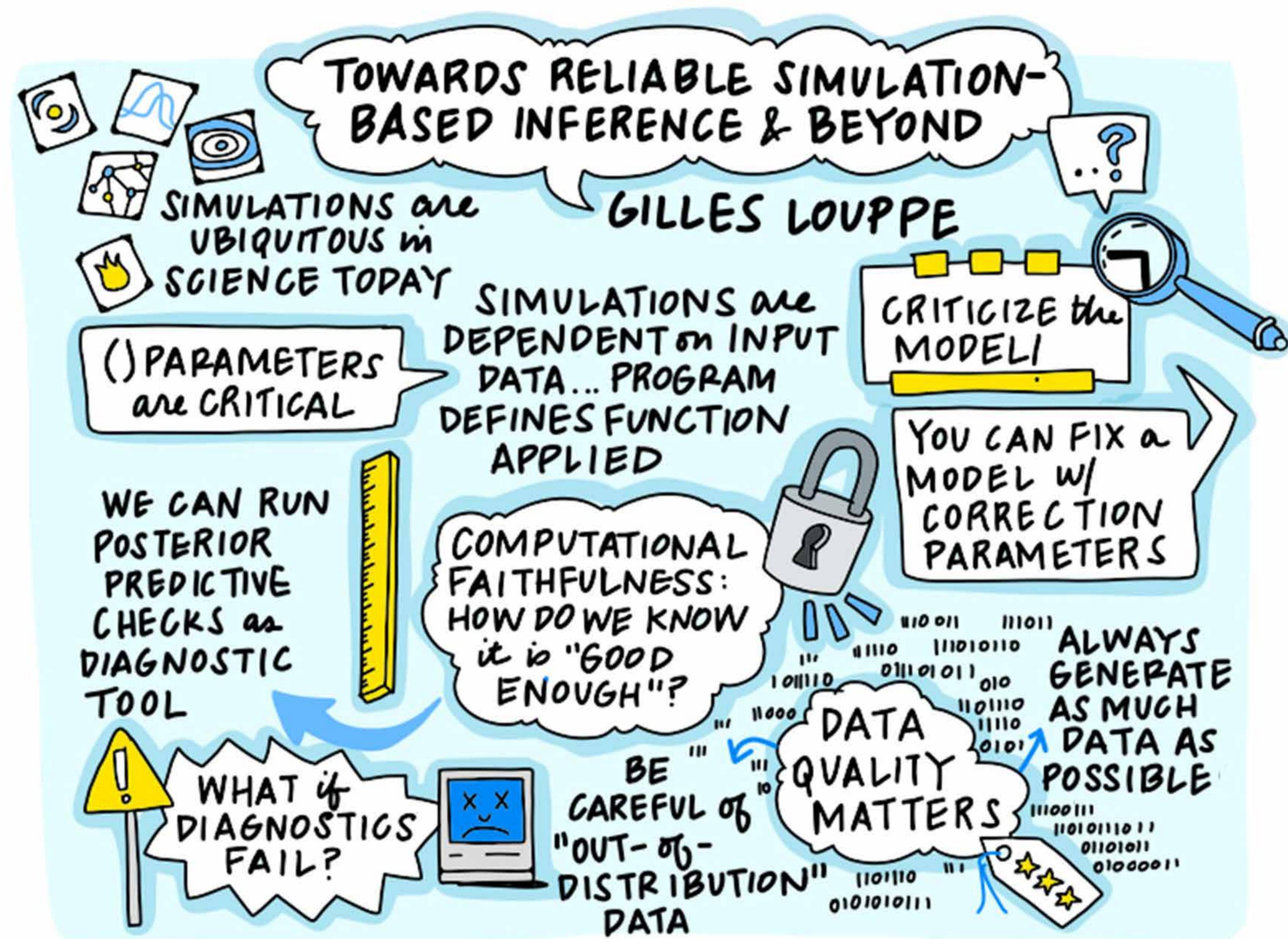
Universität Tübingen
Tübingen, DE



Towards Reliable Simulation-based Inference

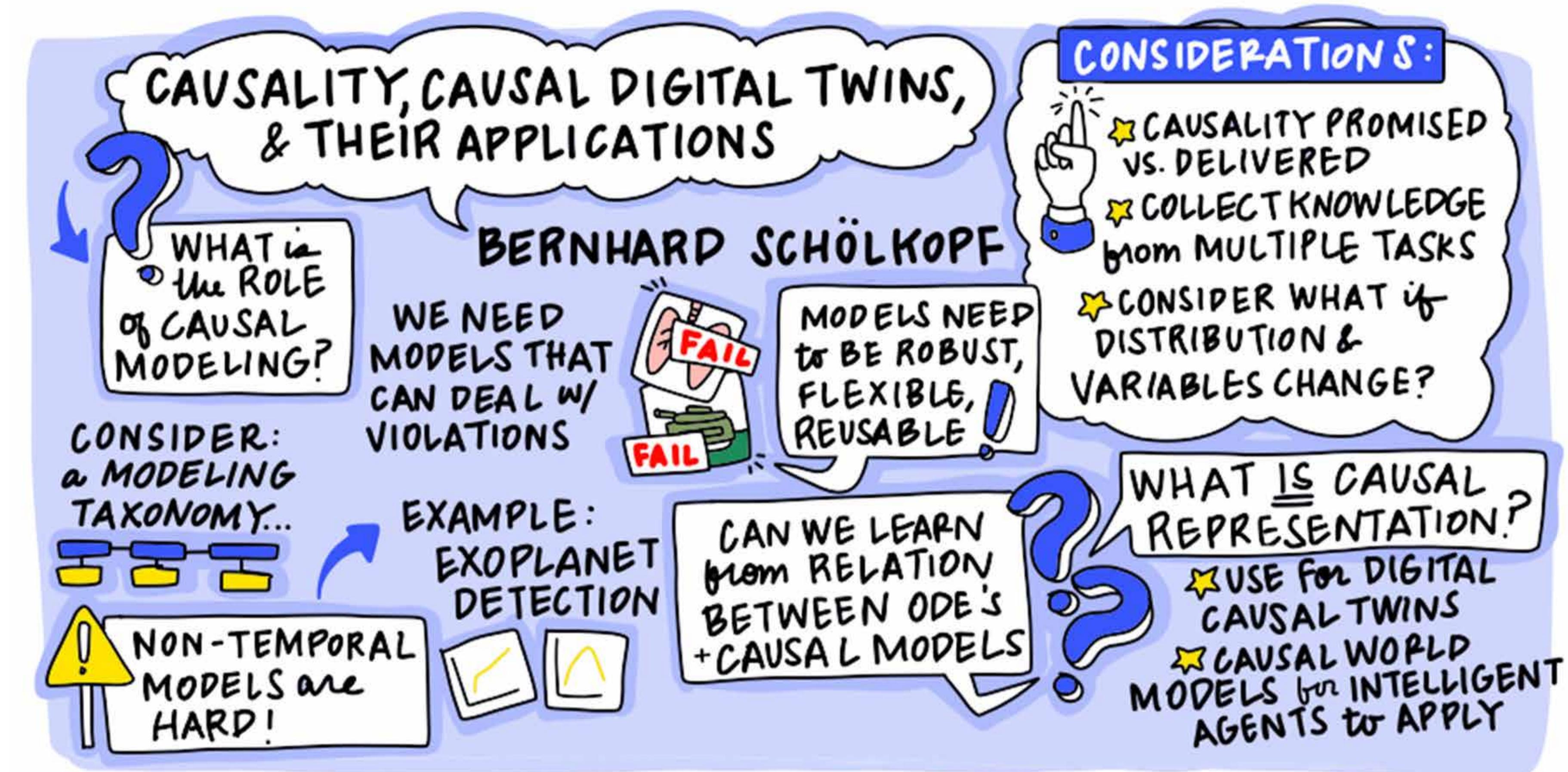
Gilles Louppe

University of Liège, BE



Causality, Causal Digital Twins, and their Applications

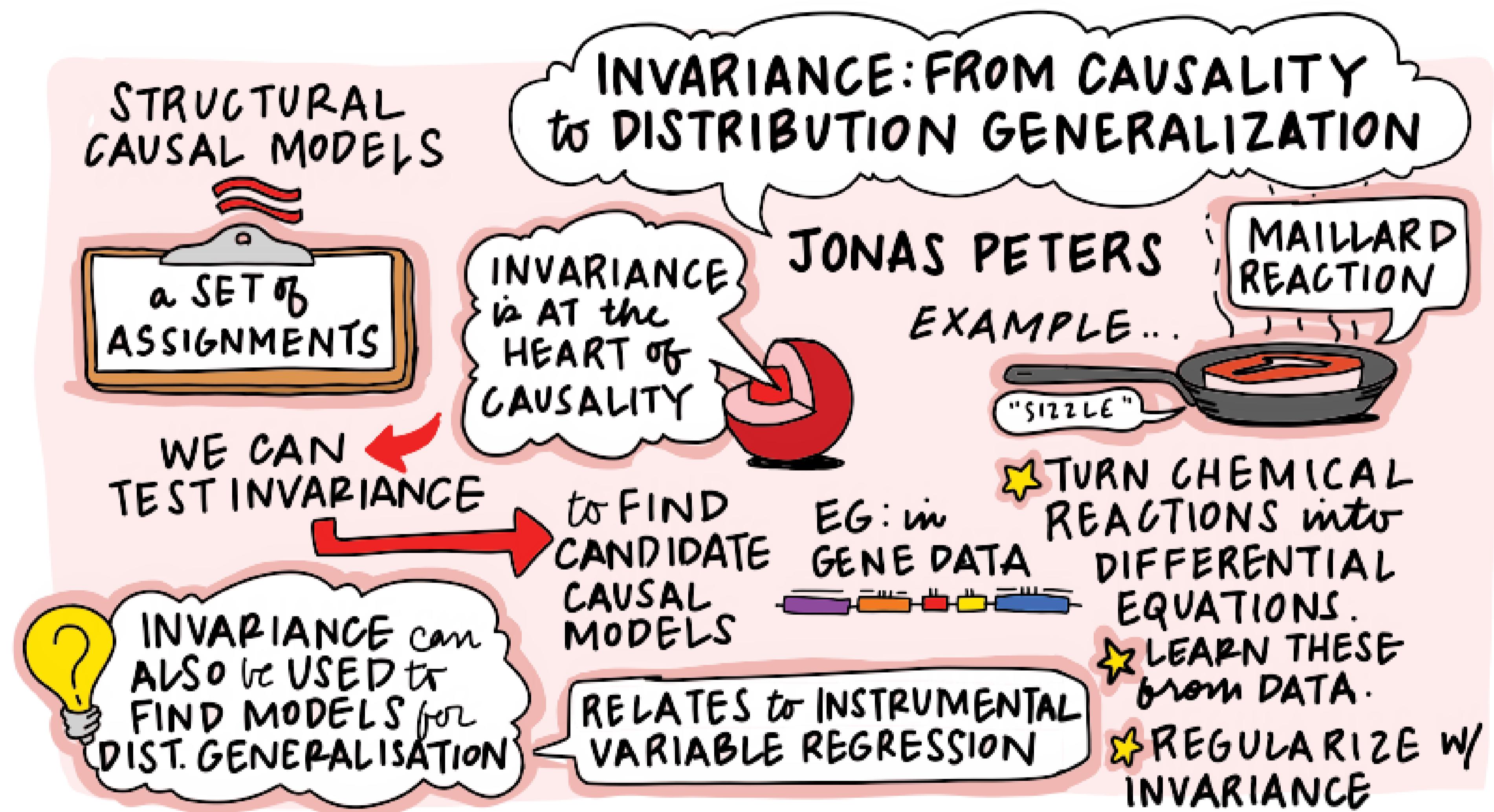
Bernhard Schölkopf
MPI für Intelligente Systeme
Universität Tübingen
Tübingen, DE



Invariance: From Causality to Distribution Generalization

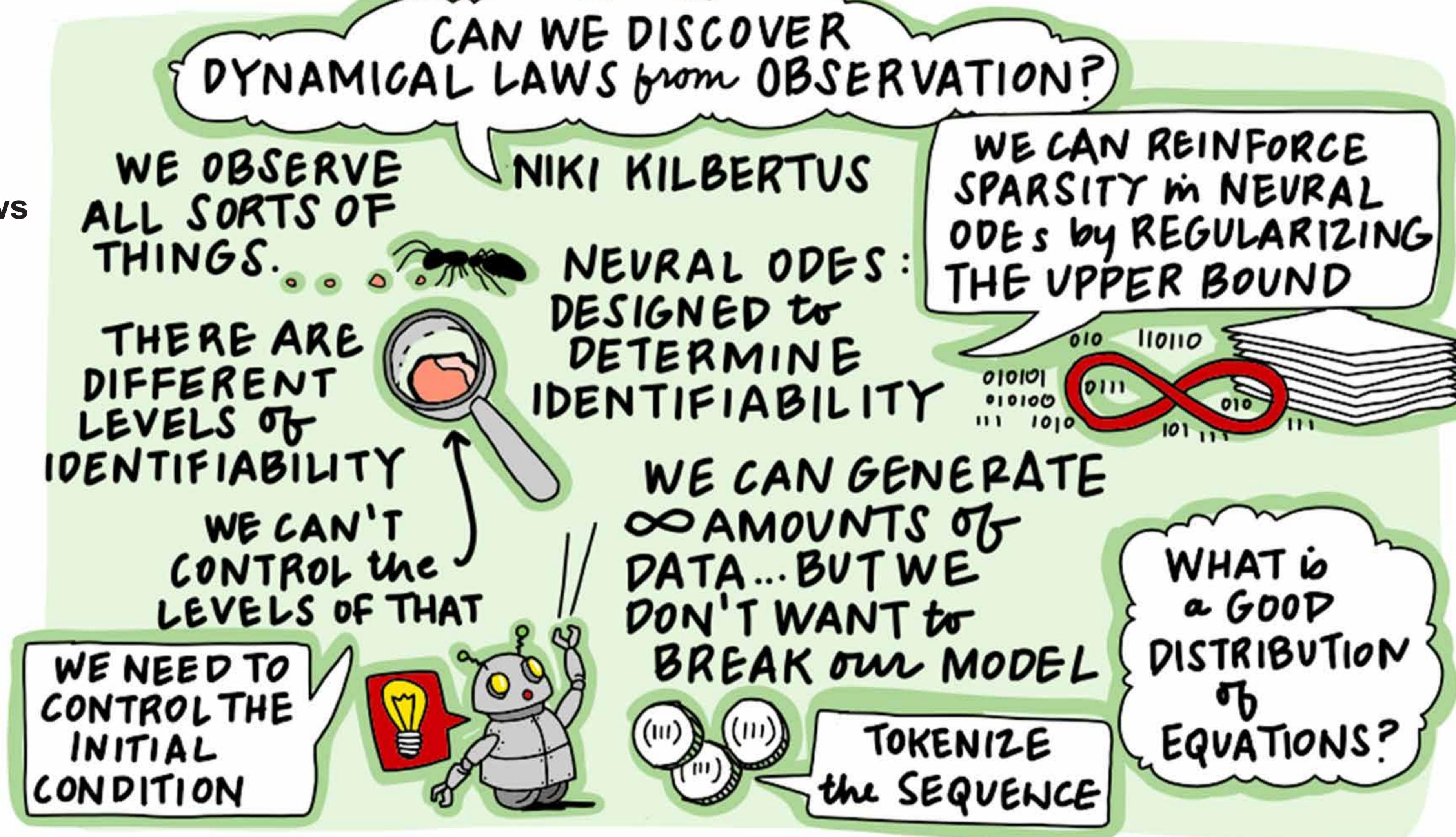
Jonas Peters

University of Copenhagen
Copenhagen, DK



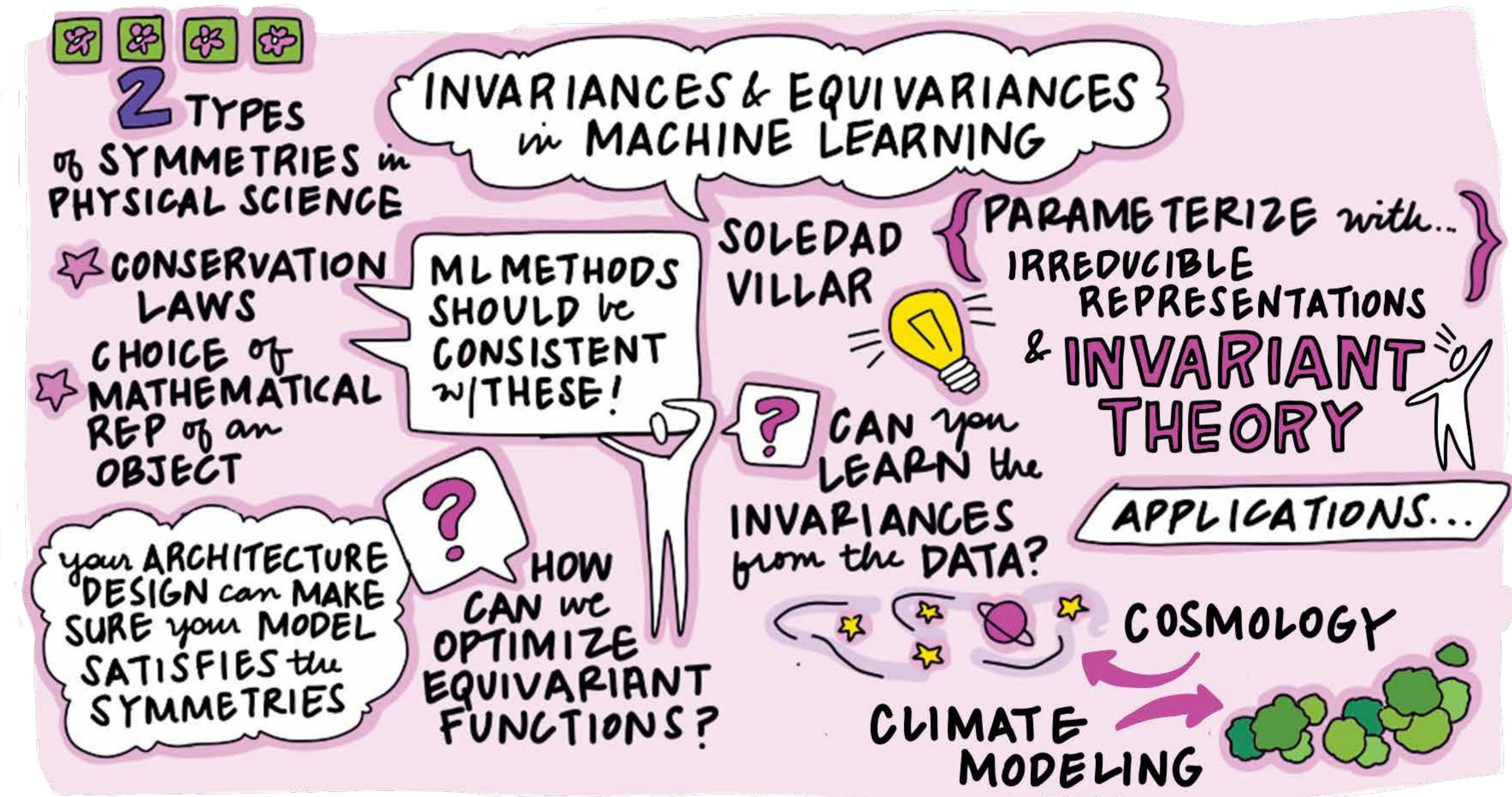
Can we discover Dynamical Laws from Observation?

Niki Kilbertus
TU & Helmholtz AI
München, DE



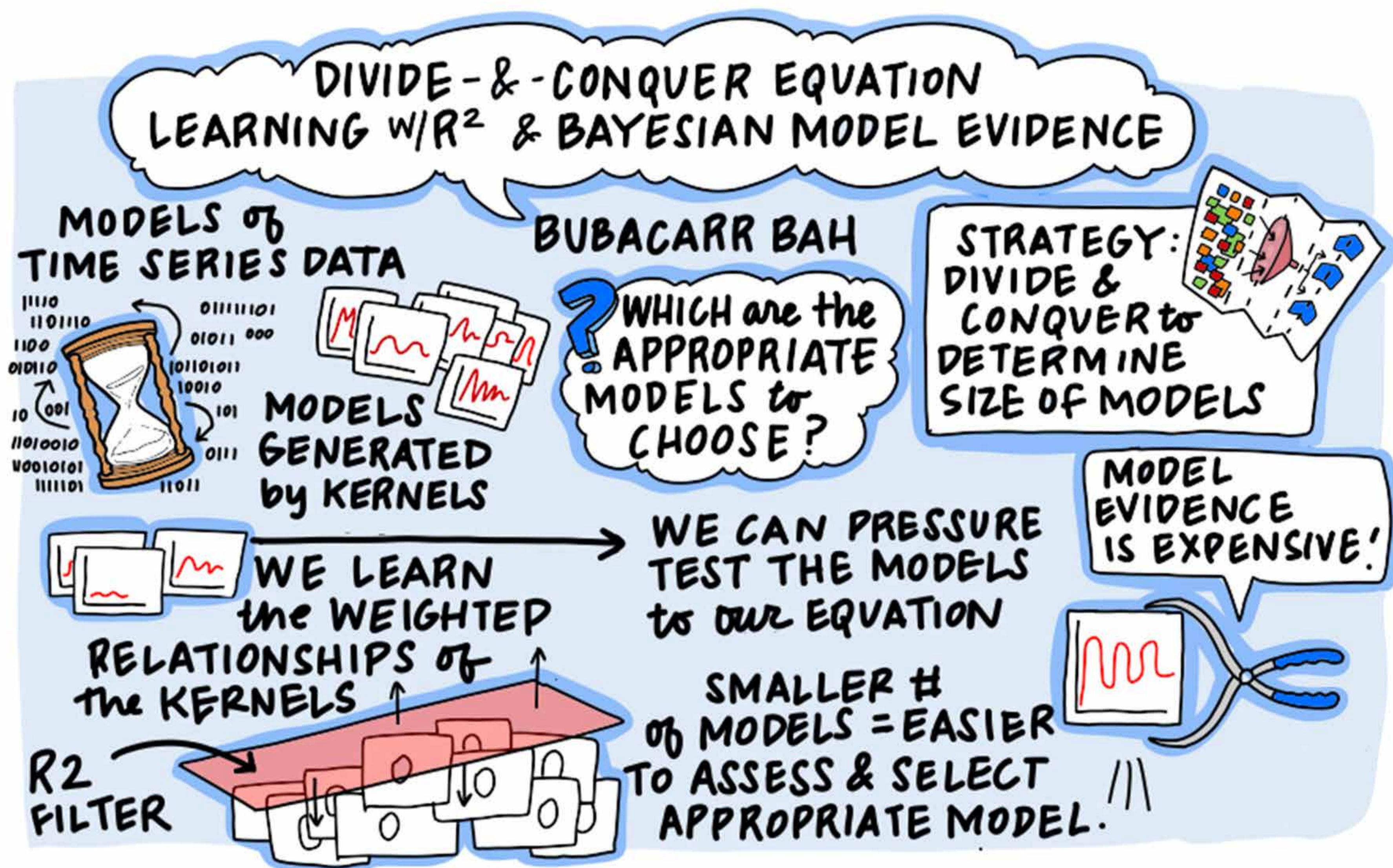
Invariances and Equivariances in Machine Learning

Soledad Villar
Johns Hopkins University
Baltimore, US



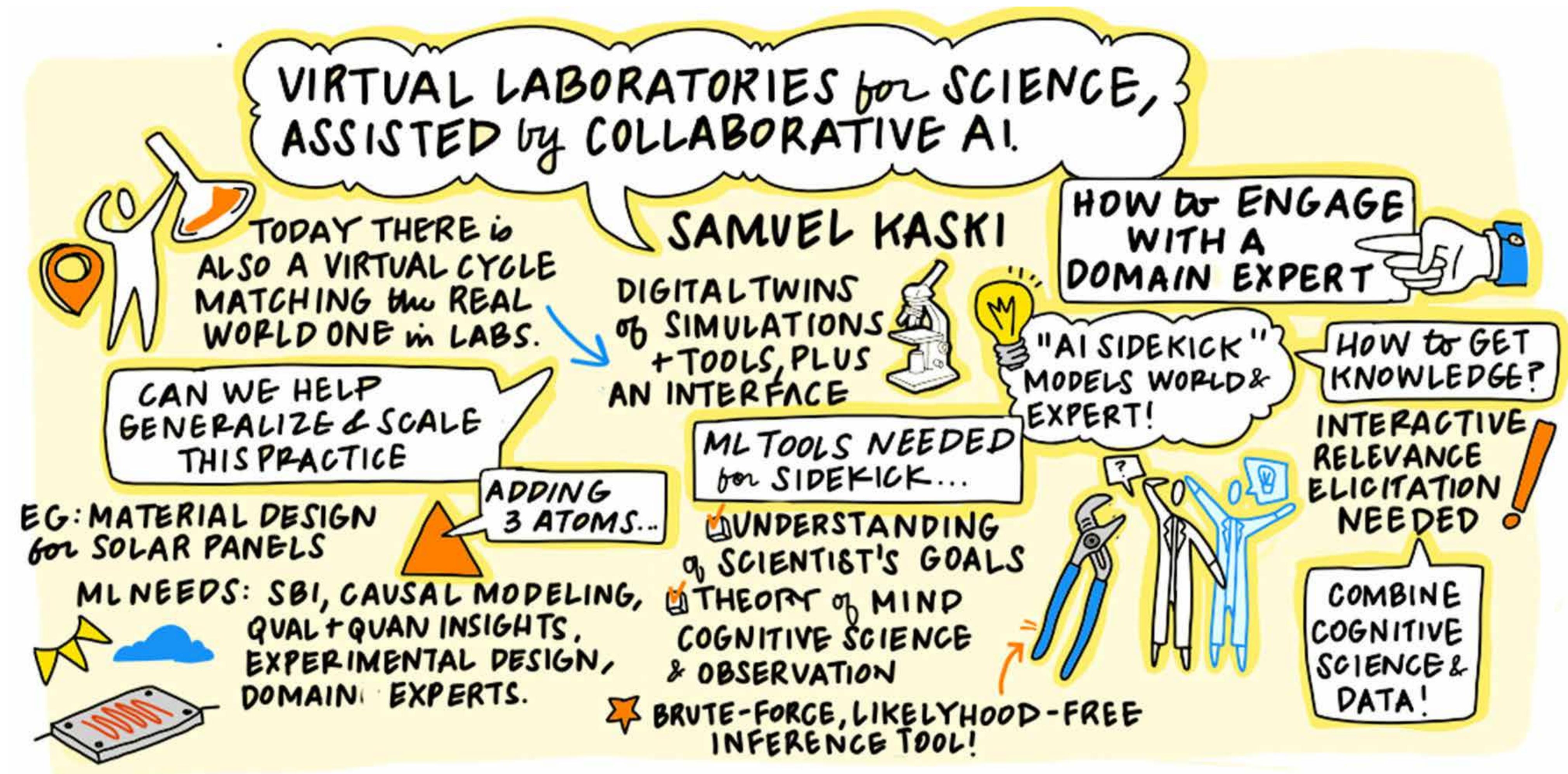
Divide-and-Conquer Equation Learning with R² and Bayesian Model Evidence

Bubacarr Bah
AIMS South Africa
Cape Town, ZA



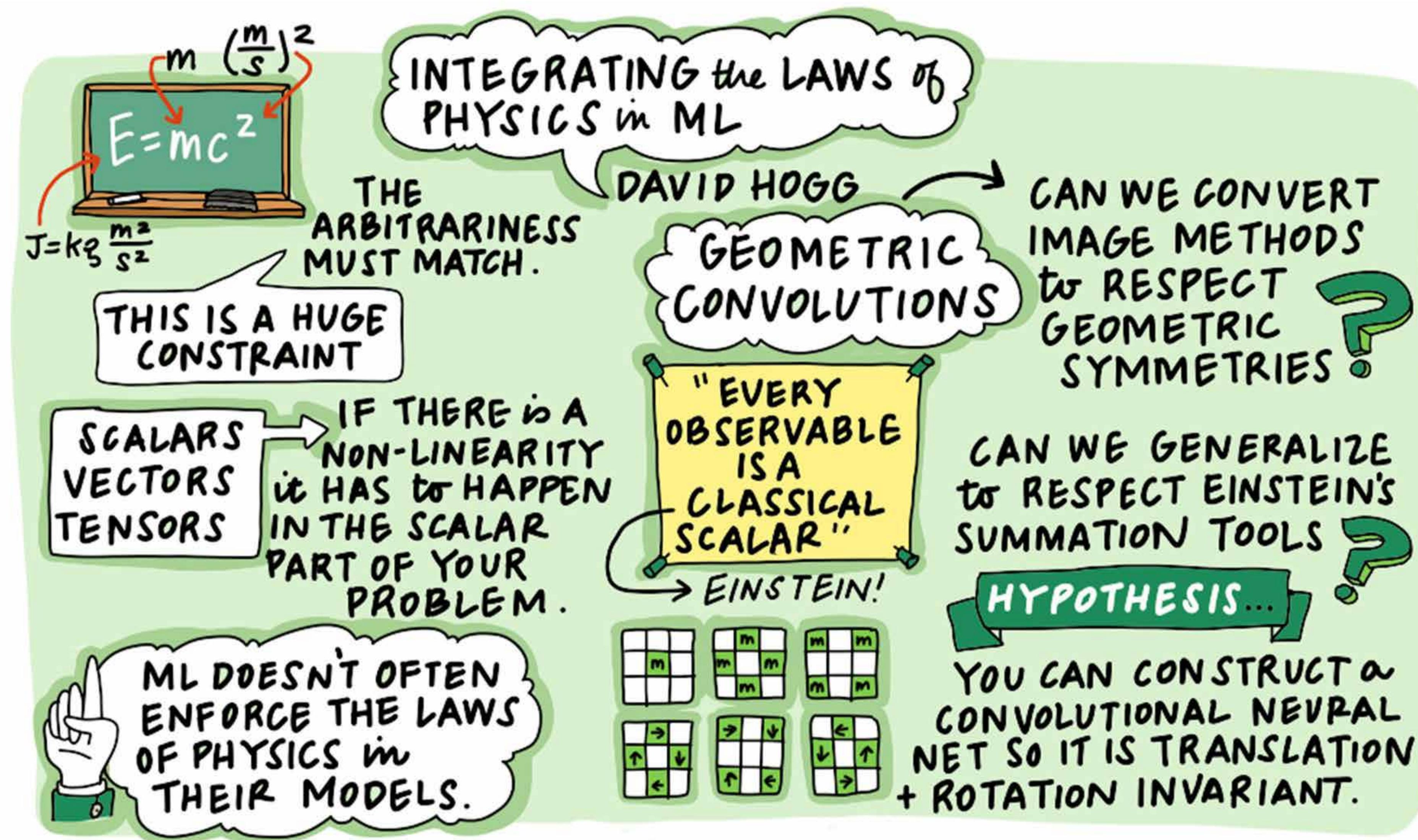
Virtual Laboratories for Science, Assisted by Collaborative AI

Samuel Kaski
Aalto University
Helsinki, FI
University of Manchester
Manchester UK



Integrating the laws of Physics in ML

David W. Hogg
New York University
NYC, US



Translating Mechanistic Understandings to Stochastic Models

Carl Henrik Ek
University of Cambridge
Cambridge, GB

