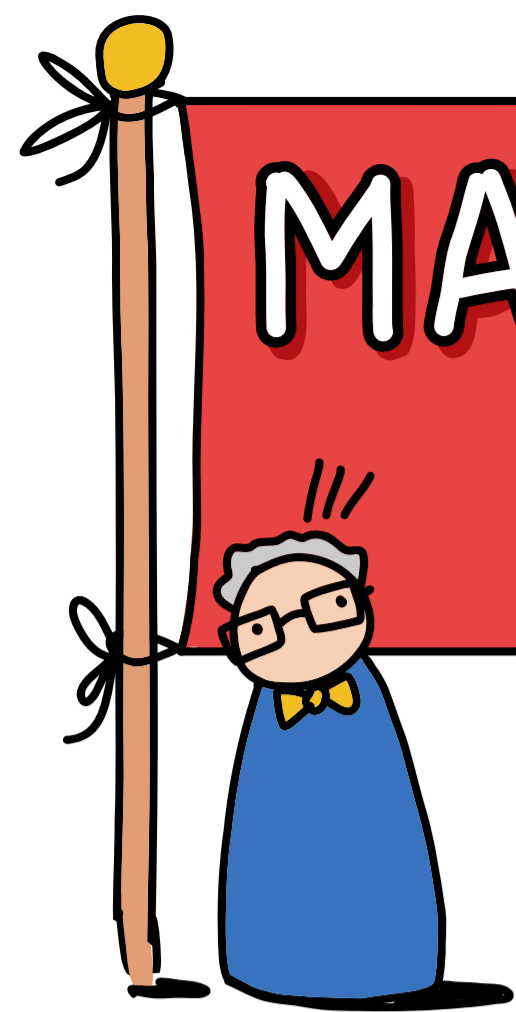
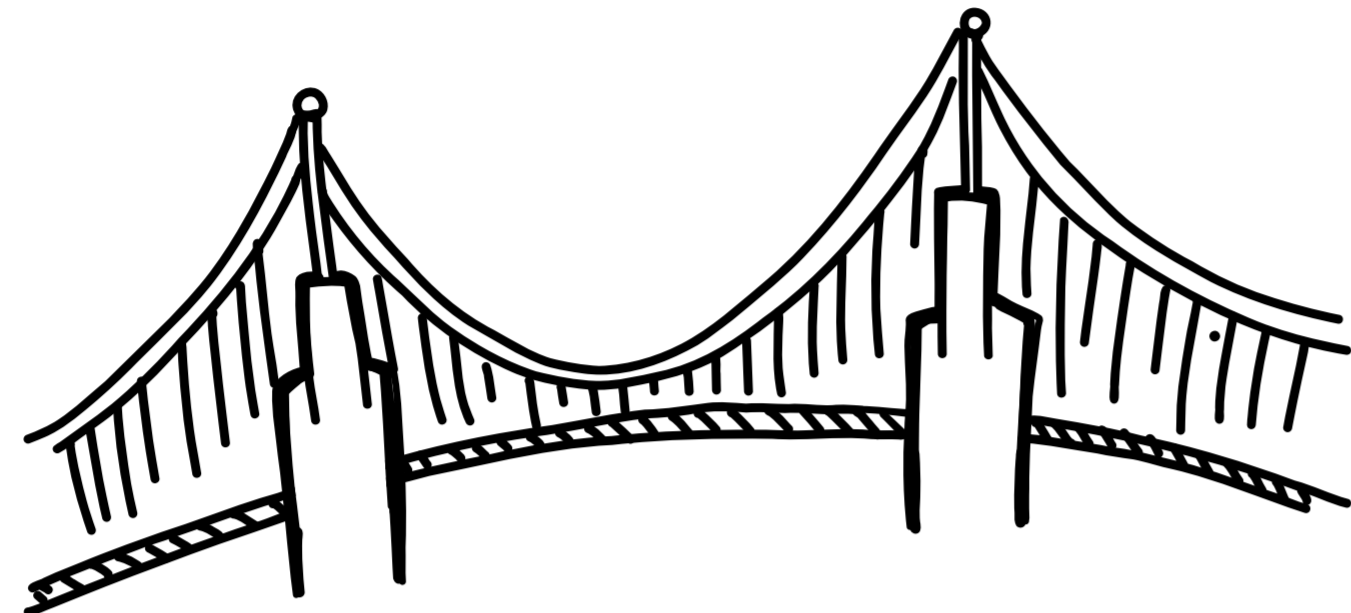


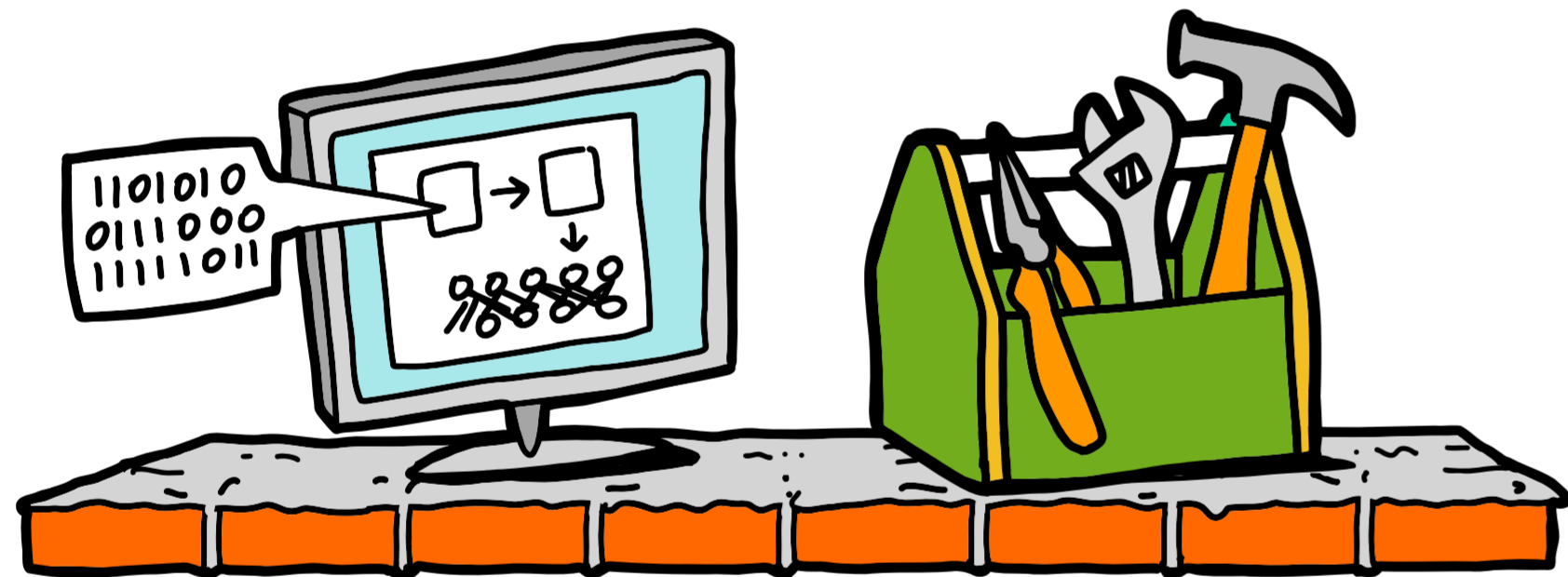
MACHINE LEARNING for SCIENCE



MATHEMATICS at the INTERFACE of DATA-DRIVEN & MECHANISTIC MODELLING



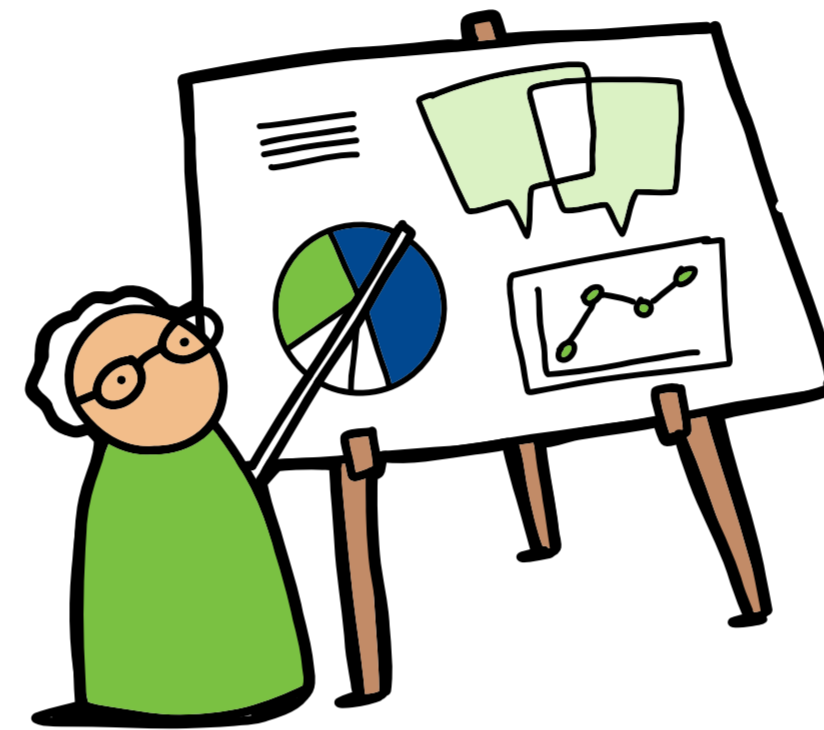
BRIDGING ML & SCIENCE:
DIRECTIONS in MATHEMATICS



FOUNDATIONAL CONCEPTS
& EMERGING METHODS

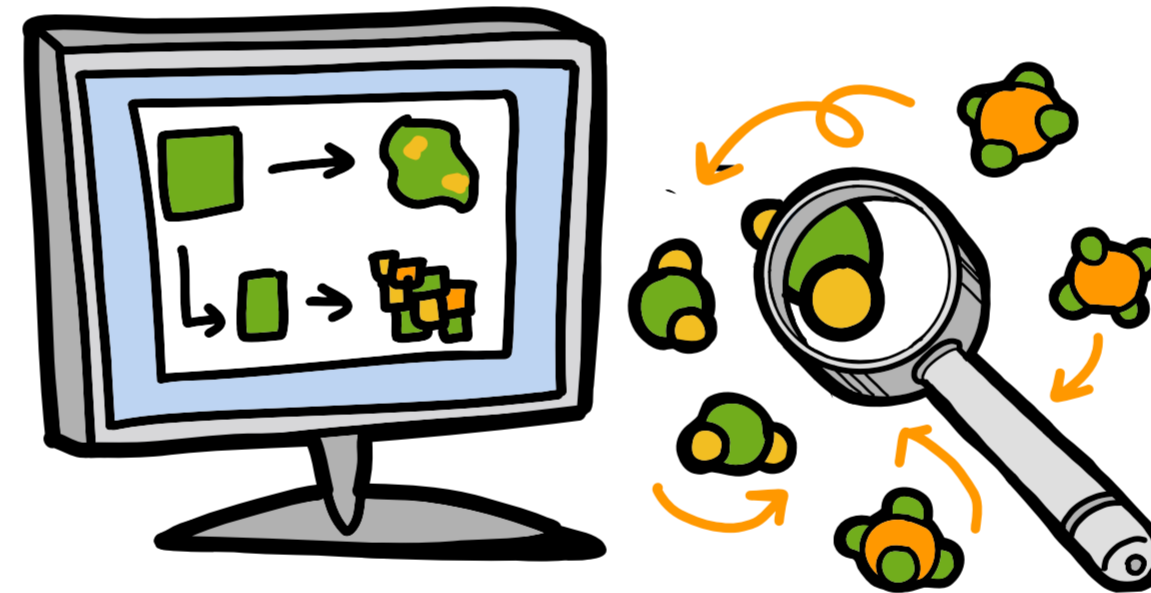
- ROBERT C. WILLIAMSON
- PHILIPP HENNIG
- CARL-HENRIK EK
- RICHARD WILKINSON
- BUBACARR BAH

- BERT KAPPEN
- WILSON GREGORY
- SAMUEL KASKI
- JAKOB MACKE
- NIKI KILBERTUS



LESSONS from the APPLICATION
of MACHINE LEARNING in SCIENCE

- MAREN BÜTTNER
- CHRISTIAN IGEL
- DINA MACHUVE
- CHALLENGER MISHRA
- KLAUS-ROBERT MÜLLER



MACHINE LEARNING for
EARTH & CLIMATE SCIENCES

- VERONIKA EYRING
- MARKUS REICHSTEIN
- GUSTAV CAMPS-VALLS
- CHRISTIAN REINERS,
ALEXANDER WINKLER

THIS CONFERENCE
was HELD at the
MATHEMATISCHES
FORSCHUNGSINSTITUT
OBERWOLFACH on
JULY 12-16, 2023

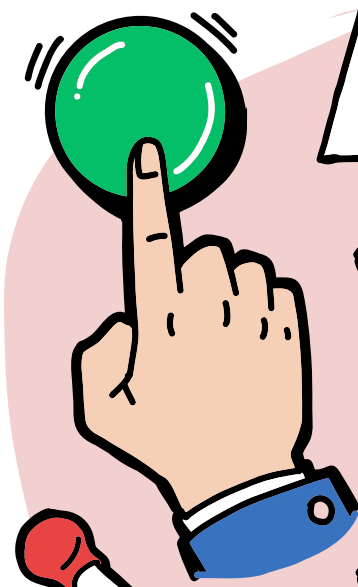
CLICK the
SECTION TITLES/
SCIENTISTS' NAMES
to GO to PAGE

ORGANISED by NEIL LAWRENCE,
JESSICA MONTGOMERY,
& BERNHARD SCHÖLKOPF



BRIDGING ML & SCIENCE: DIRECTIONS in MATHEMATICS

TODAY'S ML is USEFUL for...



✓ STITCHING TOGETHER DIFFERENT TYPES OF DATA for a NUANCED VIEW.

✓ EXTRACTING INSIGHTS from DATA for MORE ACCURATE UNDERSTANDING of the PROPERTIES of a SYSTEM

✓ IDENTIFYING AREAS for EXPERIMENTATION & THEORIZING

✓ SPEEDING UP ANALYSIS

✓ SIMULATING COMPLEX SYSTEMS



WHAT does "MATHEMATICAL FOUNDATIONS for ML" REALLY MEAN?



... ML'S FOUNDATIONAL INTERFACE with the WORLD?

MATHS HELPS US DESCRIBE RELATIONSHIPS

...to REPRESENT DOMAIN KNOWLEDGE to be EMBEDDED in ML SYSTEMS

...to EXPRESS CONCEPTS like FAIRNESS or INTERPRETABILITY

...to CREATE MODELS w/ UNCERTAINTY AWARENESS

...to HANDLE HIGH-DIMENSIONAL & SPARSE DATA

...to SUPPORT VALIDATION & BENCHMARKING

...to MODEL USERS

...to PRODUCE EFFICIENT METHODS

FOUNDATIONAL CONCEPTS & EMERGING MATHS can ACT as BRIDGES!



LIKE...

- ... STATISTICS & PROBABILITY
- ... RIEMANNIAN GEOMETRY
- ... LINEAR ALGEBRA
- ... OPTIMISM TECHNIQUES & CALCULUS



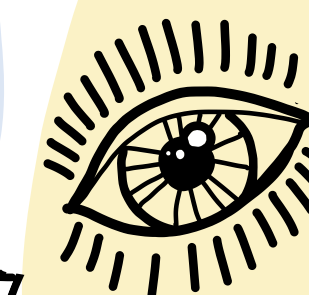
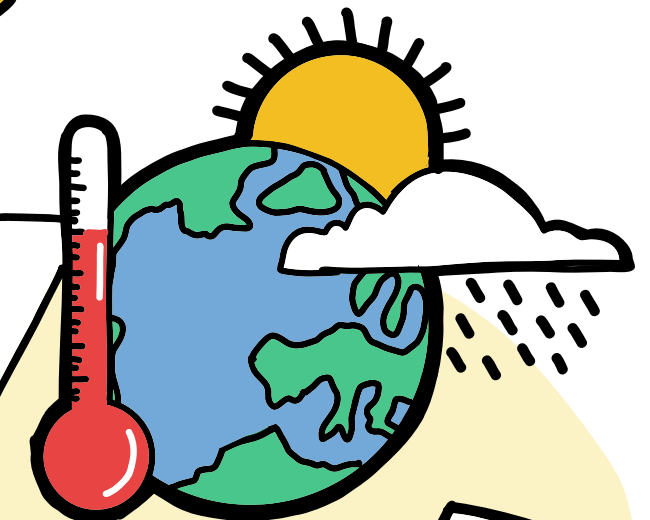
& ALSO...

- ... PROBABILISTIC NUMERICS
- ... IMPRECISE PROBABILITIES
- ... FINSLERIAN GEOMETRY
- ... ADJOINT LATENT FORCE MODELS
- ... EQUIVARIANCES

COULD we CREATE a SUITE of these AI TOOLS to DEPLOY ACROSS FIELDS?



RECENT CLIMATE CHANGE is UNPRECEDENTED



MODELS HELP us SIMULATE & ANALYSE the CLIMATE SYSTEM.



MODELS are VITAL, we NEED THEM to be TRUST-WORTHY

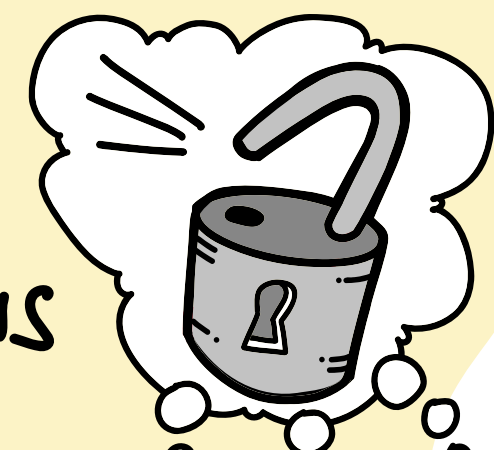
- ✓ STABLE
- ✓ PHYSICALLY-PLAUSIBLE
- ✓ BASED ON CAUSAL DRIVERS

CURRENT ML APPLICATIONS in CLIMATE SCIENCE HELP US...

- ★ UNDERSTAND SYSTEMS for CO₂ EXCHANGE & CARBON DYNAMICS
- ★ PREDICT LANDSCAPE & ECOSYSTEM CHANGES UNDER CHANGING CLIMATE CONDITIONS
- ★ FORECAST EXTREME WEATHER EVENTS
- ★ I.D. CAUSAL RELATIONSHIP THAT DRIVE THESE CHANGES

OPEN CHALLENGES...

- ... PARAMETER ESTIMATION & BAYESIAN INFERENCE
- ... EXPLAINABLE AI
- ... IMPROVED GENERALISATION ACROSS CHANGING DISTRIBUTIONS
- ... UNCERTAINTY QUANTIFICATION
- ... MANAGING HYBRID DATA-DRIVEN & MECHANISTIC MODELS



We NEED a COMMUNITY DEDICATED to DEVELOPING this WORK!



BUT...



! HARD to ENGAGE BOTH ML & DOMAIN EXPERTISE TOGETHER

! TONS of DATA... BUT NO CLEAR WAY to TRACK SUCCESS

! we STILL LACK SOPHISTICATED TECHNIQUES for ESTABLISHING CAUSAL RELATIONSHIPS

! SYSTEMS NEED to be DEPLOYABLE & EASY to USE

LOWERING BARRIERS...



→ PROVIDE ACCESSIBLE BENCHMARKS that REFLECT DOMAIN INTERESTS & EXISTING KNOWLEDGE

→ SIMPLIFY USER INTERFACES

→ ADDRESS a CLEAR NEED

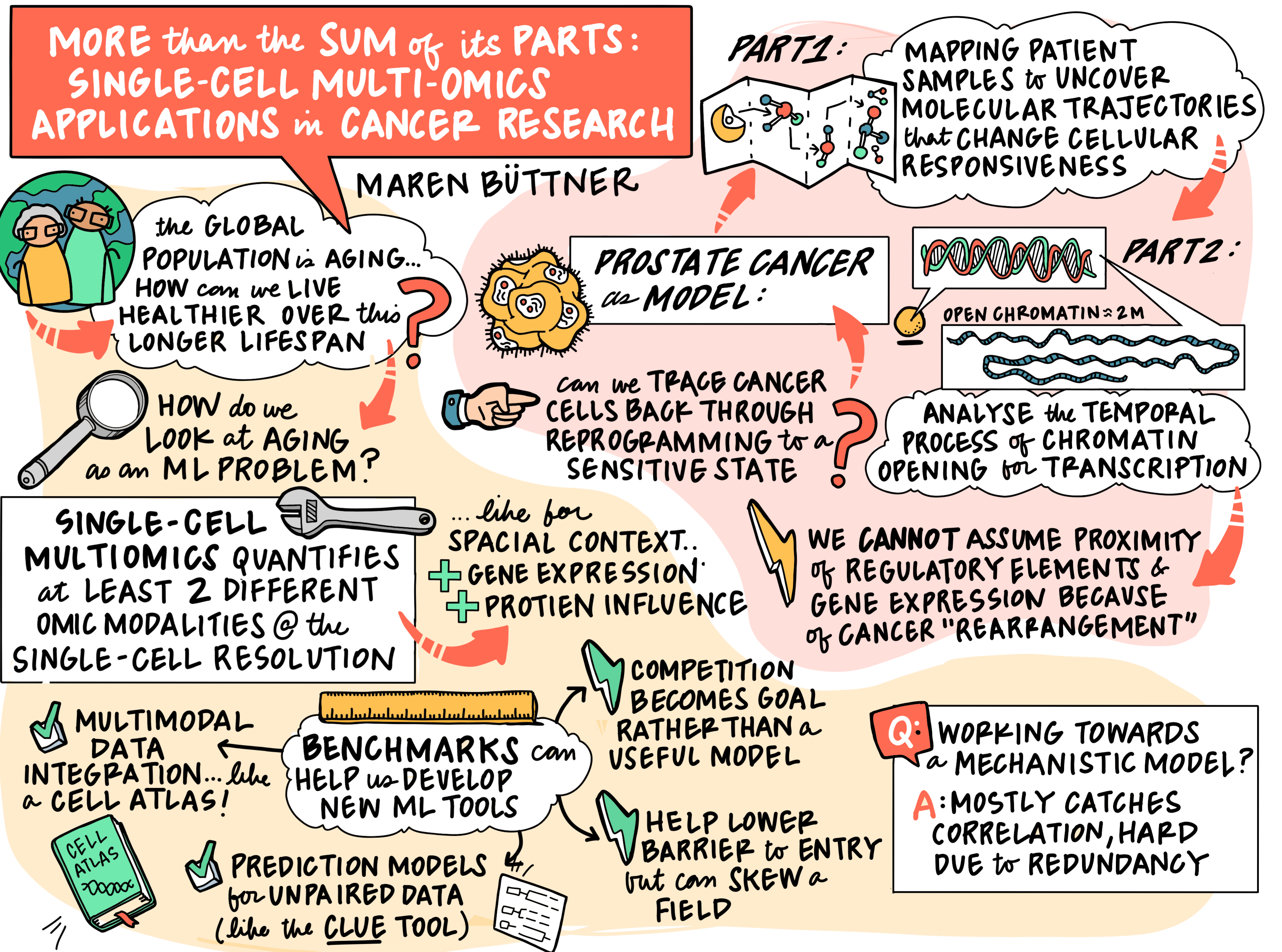
→ THEORIES of MIND for HUMAN INTERACTION

→ NARRATIVES for SCIENTIFIC RESULTS



More than the Sum of its Parts: Single-Cell Multi-omics Applications in Cancer Research

Maren Büttner
Calico Life Sciences
San Francisco, CA

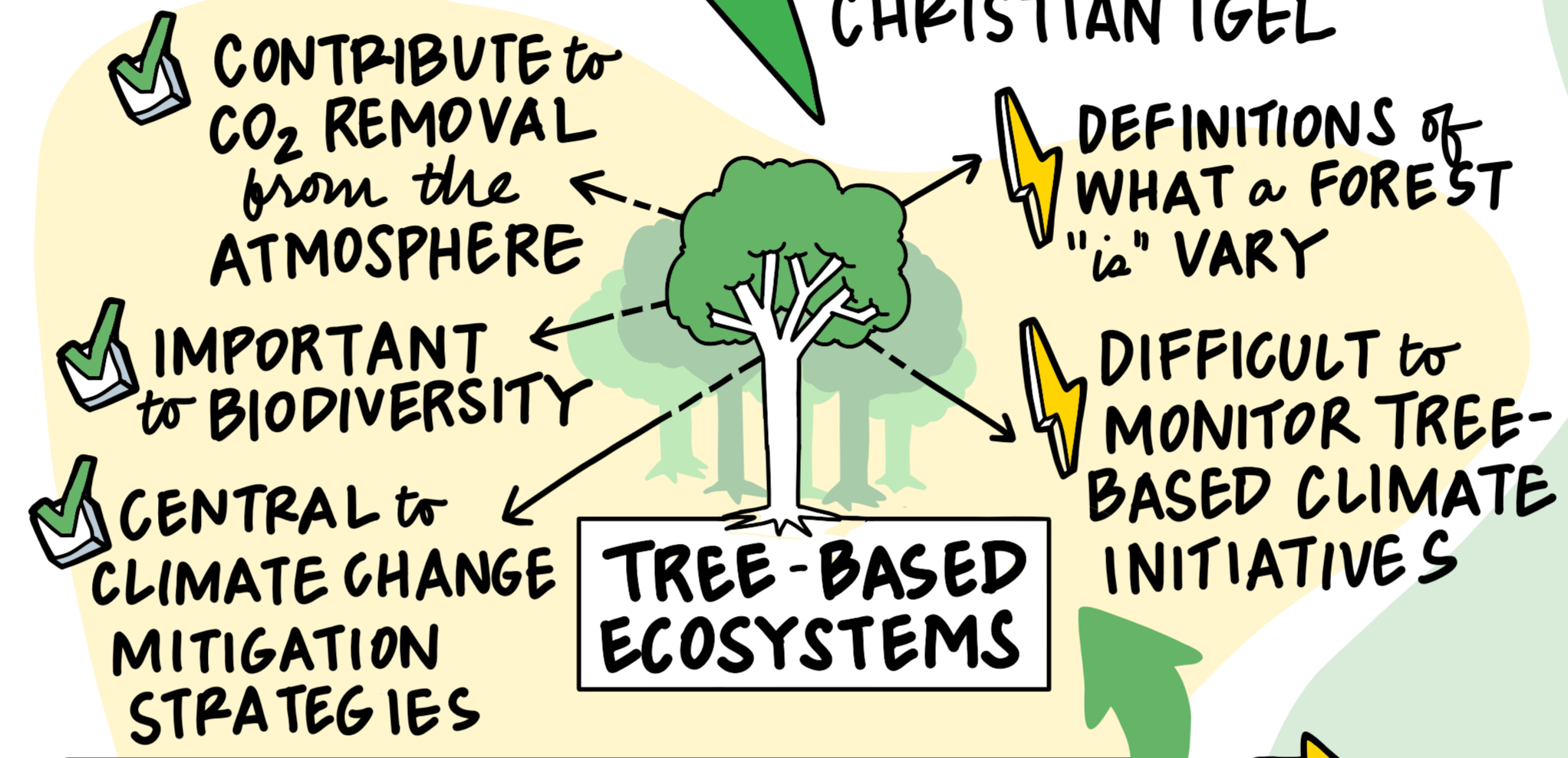


Regression of Ecosystem Properties: Bias, Monotonicity, & Uncertainty

Christian Igel
University of Copenhagen
Copenhagen, DN

REGRESSION of ECOSYSTEM PROPERTIES: BIAS, MONOTONICITY, & UNCERTAINTY

CHRISTIAN IGEL



we CAN USE ML to MORE ACCURATELY QUANTIFY TREE CARBON SEQUESTRATION!

... we can APPLY ML to SATELLITE IMAGES to SEGMENT & COUNT TREES.

... we can ESTIMATE TREE BIO-MASS to UNDERSTAND HOW MUCH CARBON THEY HOLD

★ LARGE-SCALE TREE ECOSYSTEM MONITORING w/ UNPRECEDENTED ACCURACY

1 FITTING ALLOMETRIC EQUATIONS

we NEED a GOOD FIT to DATA w/ LOW BIAS, & WANT TIGHT UNCERTAINTY ESTIMATES

CHALLENGE: NON-SYMMETRIC, HETEROSKEDASTIC ERROR DISTRIBUTION

2 MONOTONIC NEURAL NETWORKS for BIOMASS PREDICTION

SMOOTH MONOTONIC APPROXIMATIONS of MIN/MAX = NO SILENT NEURONS

SMOOTH MODULE can be USED in a LARGER DL SYS. & TRAINED E-to-E.

BIAS in DEEP LEARNING for REGRESSIONS

ERRORS can ACCUMULATE!

SATELLITE IMAGES ALONE AREN'T ENOUGH... COMBINE w/ LIDAR POINT CLOUDS for MORE ACCURATE BIOMASS EST.

I.D. SPECIES!

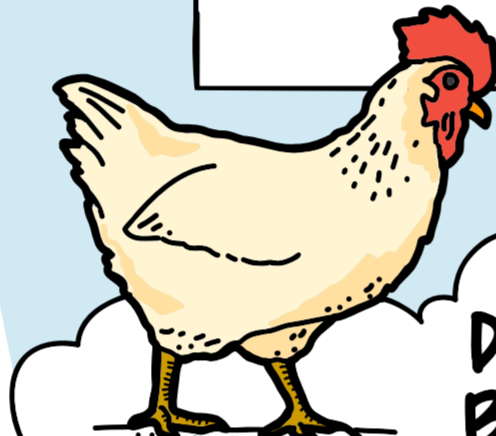
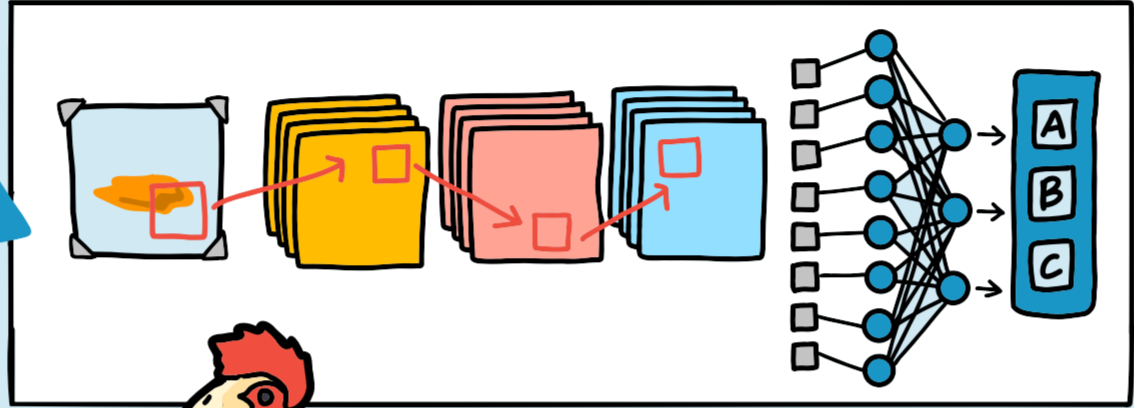


DATA-DRIVEN SOLUTION for POULTRY DISEASE DIAGNOSTICS

Data-Driven Solution for Poultry Disease Diagnostics

Dina Machuve
DevData Analytics
Arusha, TZ

we BUILT a CNN MODEL to I.D. POULTRY DISEASE from FECAL IMAGES



the GOAL: →

DELIVER a SMARTPHONE-BASED MODEL to TANZANIAN FARMERS to HELP I.D. SICK POULTRY from FECAL IMAGES.

MOTIVATION:

- LOWER BARRIERS to ACCESS for LAB WORK (COST, TIME, LITERACY, LANGUAGE)
- TAKE ADVANTAGE of UBIQUITY of MOBILE PHONES to INCREASE ACCESS to DIAGNOSTIC TOOLS.

UPCOMING WORK: SEMANTIC SEGMENTATION of IMAGES

DINA MACHUVE

- LOW RATIO of EXTENSION OFFICERS to FARMERS
- RADIO BROADCASTS TOO SHORT to ADDRESS SPECIFIC NEEDS
- BARRIERS to INTERNET ACCESS

WHAT ABOUT a DIGITAL ASSISTANT?

HOW?

1. FARMER ASKS Q.
2. Q ANALYSED by STT.
3. CHATGPT GENERATES RESPONSE.
4. ANSWER CONVERTED to SPEECH by IVR.

we NEED to be ABLE to PROBE the SYSTEM, & ACCOMODATE HOW HUMANS EXPRESS UNCERTAINTY.

...HOW do we INCENTIVISE FARMERS to USE the TOOL?

...HOW do we DELIVER the TOOL to FARMERS?

DEPLOYMENT CHALLENGES

CONVERSATIONAL IVR TECHNOLOGY: ...a SWAHILLI VOICE BOT! HABARI!

PROBLEM: FARMERS HAVE LOW-END PHONES

OPEN MATHEMATICAL QUESTIONS

- ... MODELS to USE to REPRESENT the CONCEPT
- ... to ENABLE INTERPRETABILITY
- ... to REPRESENT FAIRNESS
- ... to HANDLE UNCERTAINTY

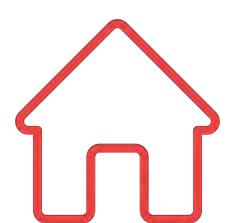
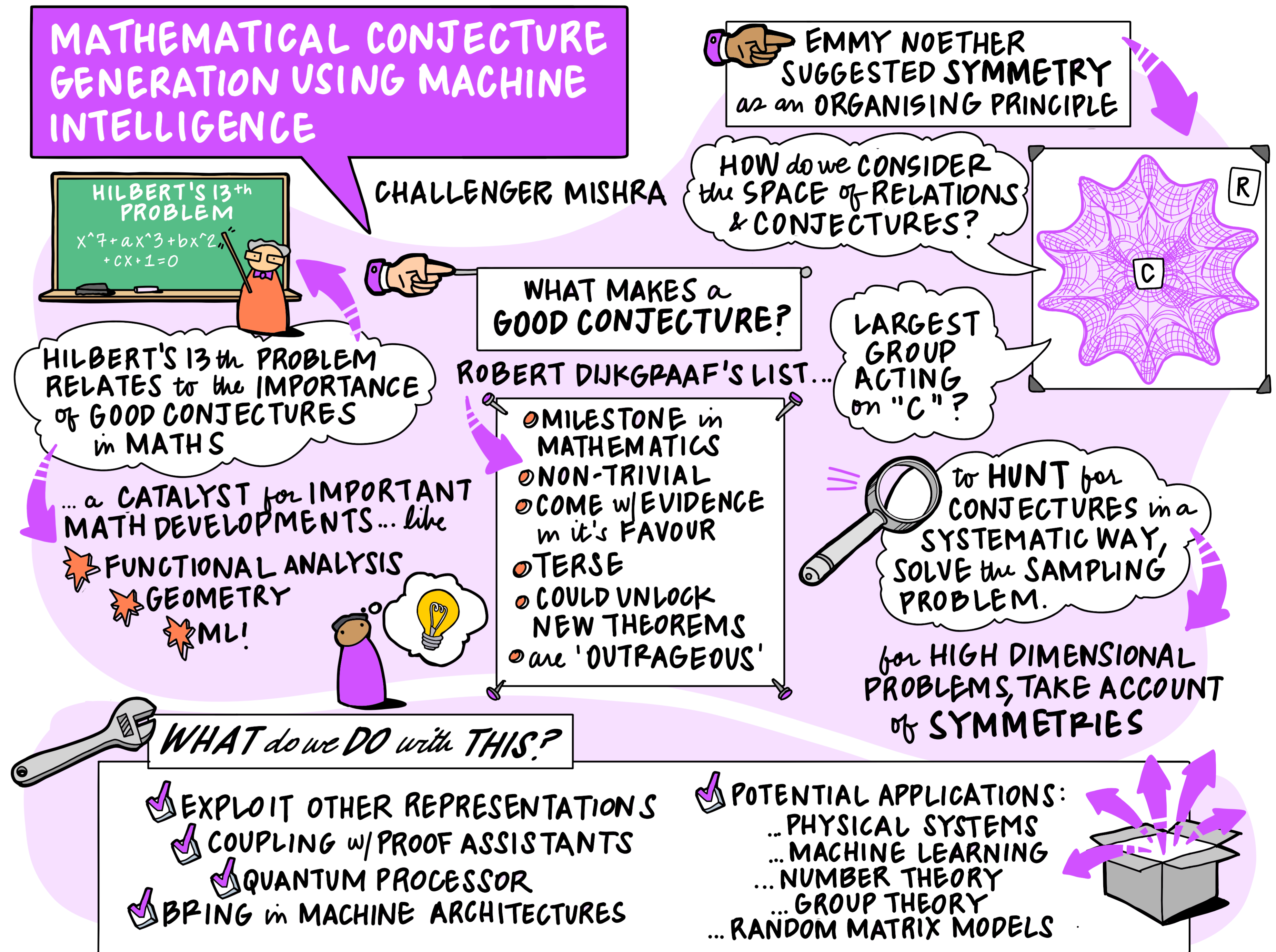
POLITICS... CLIMATE... CULTURE... LANGUAGE...

THESE SYSTEMS are COMPLEX & MASSIVE in SCALE!



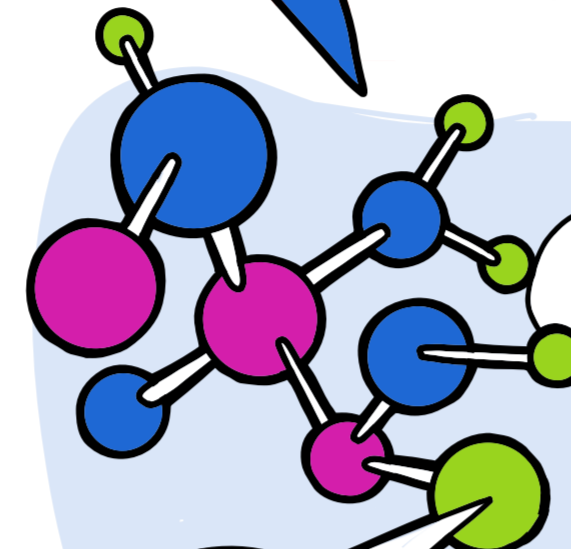
Mathematical Conjecture Generation using Machine Intelligence

Challenger Mishra
University of Cambridge
Cambridge, UK



ML for QUANTUM CHEMISTRY

KLAUS-ROBERT MÜLLER



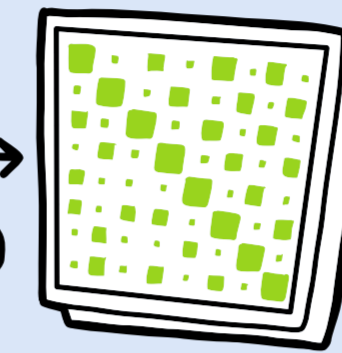
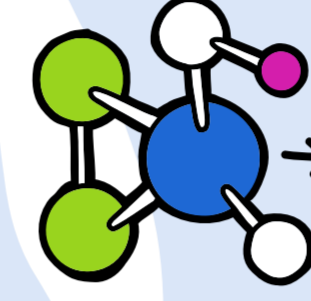
We NEED EFFICIENT WAYS to UNDERSTAND the PROPERTIES of MOLECULES



... like an ATOM in its CHEMICAL CONTEXT - we SHOULD be ABLE to PREDICT ENERGY & INTERACTIONS.



Can we SHORTCUT the SCHRODINGER EQUATION & USE ML BASED on REPRESENTATIONS of MOLECULES?



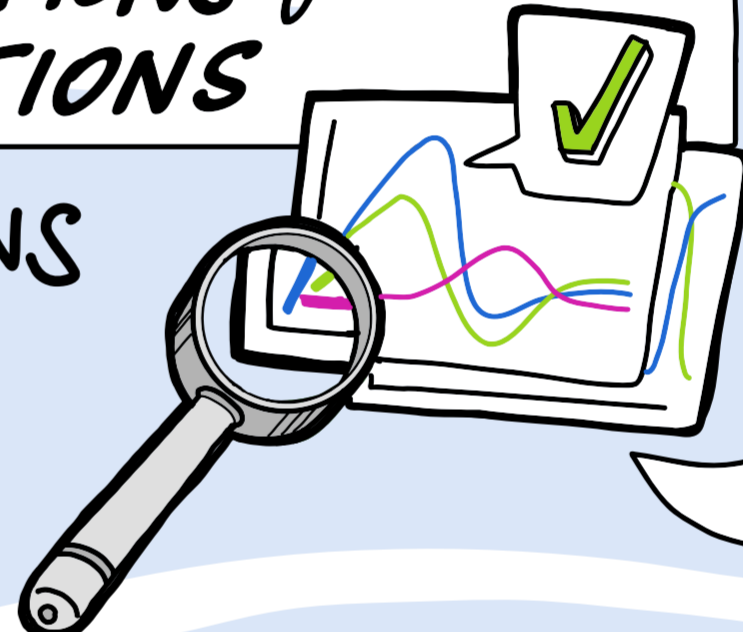
- COULOMB REPRESENTATION of EACH MOLECULE
- ML to COMPARE MATRICES

MULTIPLE TOOLS for THIS...

SchNet for QUANTUM INTERACTIONS from CONVOLUTIONS



CONVOLUTIONS ACCOUNT for NON-LINEAR INTERACTIONS



BUT, we MUST CHECK MODEL to KNOW if it is LEARNING ARTEFACTS, or CHEMISTRY.

Can ML SAMPLE this SPACE EFFICIENTLY to be ABLE to DO THIS REGRESSION PROBLEM?

10 ATOMS in 3 POSITIONS

the LAW of CONSERVATION HOLDS.

USE THIS as a WAY to RESTRICT to ONLY PLAUSIBLE RESULTS

- EXPLICITLY CODE THIS in w/ GRADIENT ML
- ADD MORE KNOWLEDGE
- ADD SYMMETRIES

... & for LARGER MOLECULES?

- CURRENT SYSTEMS aren't STRONG ENOUGH, COMPUTATIONALLY..
- FRAGMENT, DEV. LOCAL MODELS, STITCH w/ML?

NEXT CHALLENGE...

STITCHING TOGETHER LOCAL & GLOBAL INFO!

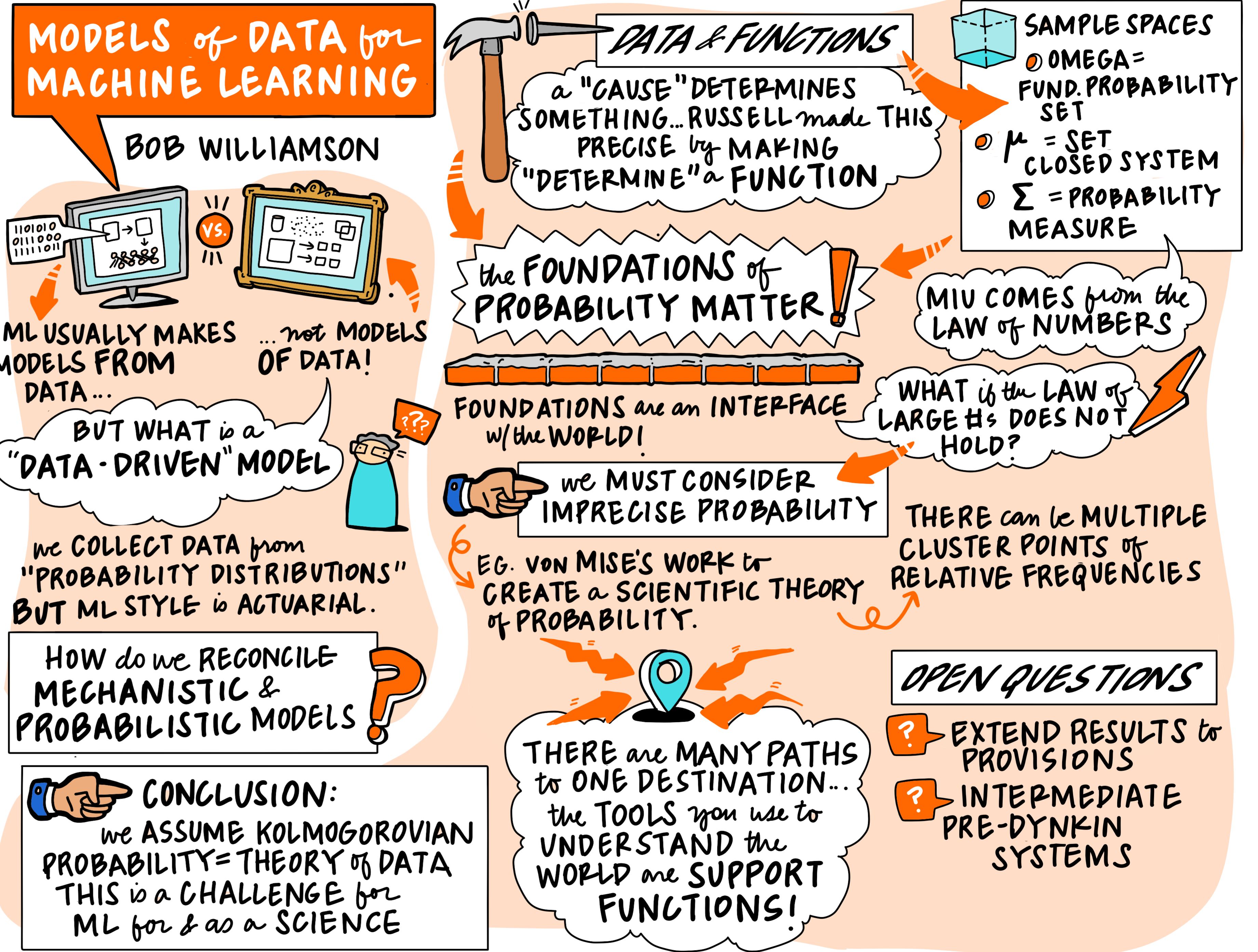
ML for Quantum Chemistry

Klaus-Robert Müller
Technical University Berlin
Berlin, DE



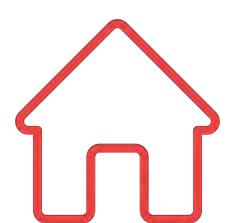
Models of Data for Machine Learning

Robert C. Williamson
University of Tübingen
Tübingen, DE



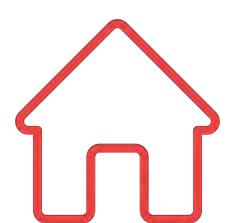
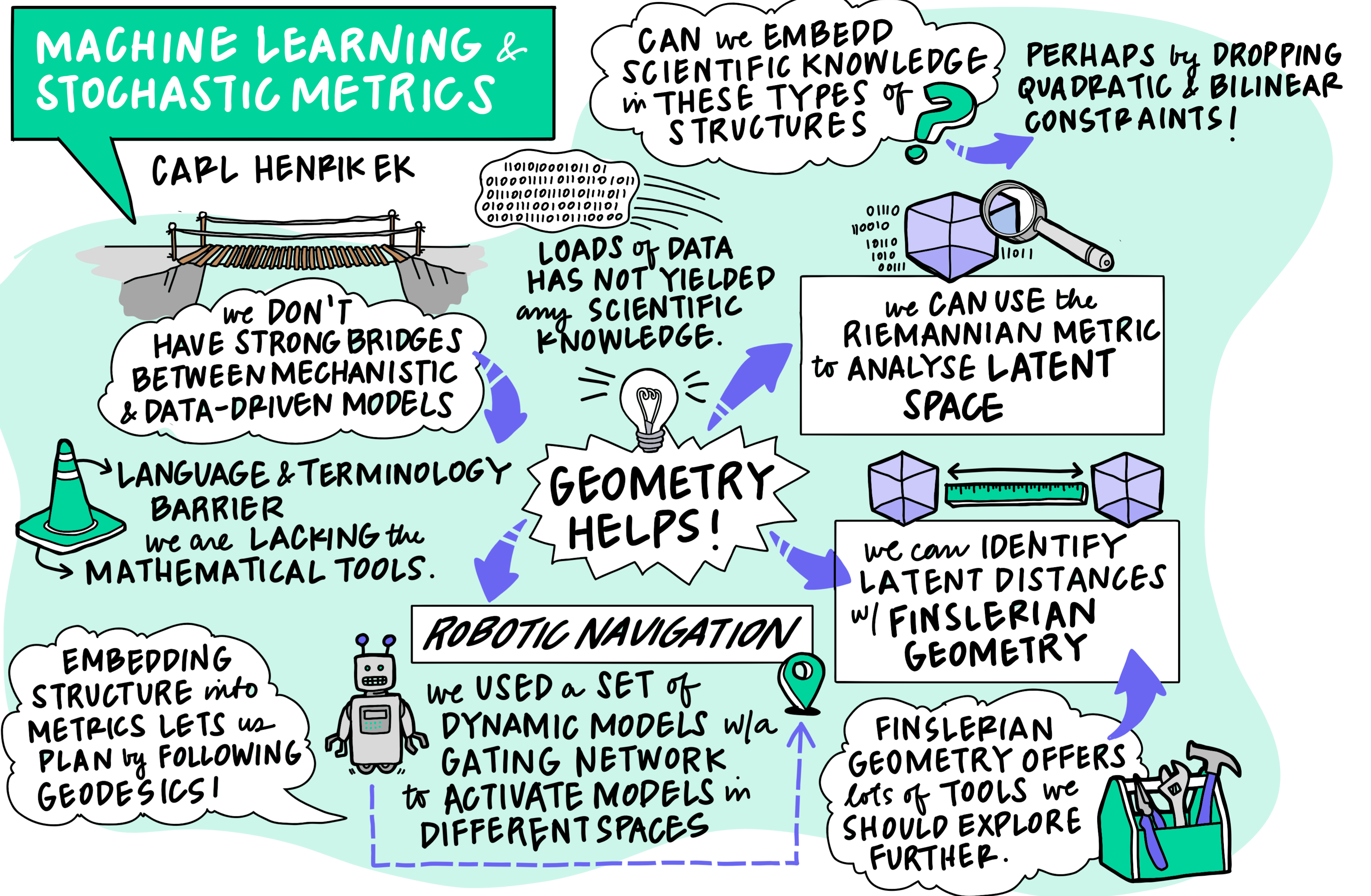
Deep Learning only works if its Bayesian, & Bayesian Deep Learning is Easy

Philipp Hennig
University of Tübingen,
Max Planck Institute for Intelligent
Systems
Tübingen, DE



Machine Learning and Stochastic Models

Carl Henrik Ek
University of Cambridge,
Cambridge, UK



Adjoint-added Inference for Latent Force Models

Richard Wilkinson
University of Nottingham
Nottingham, UK

ADJOINT-ADDED INFERENCE for LATENT FORCE MODELS

RICHARD WILKINSON

HOW to do INFERENCE for LINEAR SYSTEMS in NOISY ENVIRONMENTS ?

WHAT is it EVERYWHERE?

WHERE will it be TOMORROW?

SAMPLE PROJECT: CAUSES of AIR POLLUTION in KAMPALA, UGANDA.



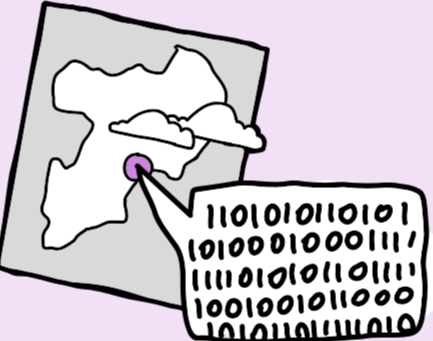
BODA BODA DRIVERS TRAVERSE CITY w/SENSORS to MEASURE PARTICULATES & CALIBRATE AGAINST FIXED SENSORS.

we WANT to MODEL the POLLUTION THROUGH SPACE & TIME - we NEED DATA MODELS that KNOW PHYSICS.

INCLUDING MECHANISTIC BEHAVIORS will HELP us to...

- ✓ INFER SOURCES
- ✓ PLAN INTERVENTIONS
- ✓ PREDICT BETTER

...to BUILD a DIGITALTWIN



ADJOINTS of LINEAR SYSTEMS

- ✓ can be AUTOMATED
- ✓ INSENSITIVE to # of BASIS FUNCTIONS USED
- ✓ GIVE NUMERICALLY STABLE DERIVATIVES of the COST FUNCTION \approx to OTHER PARAMETERS

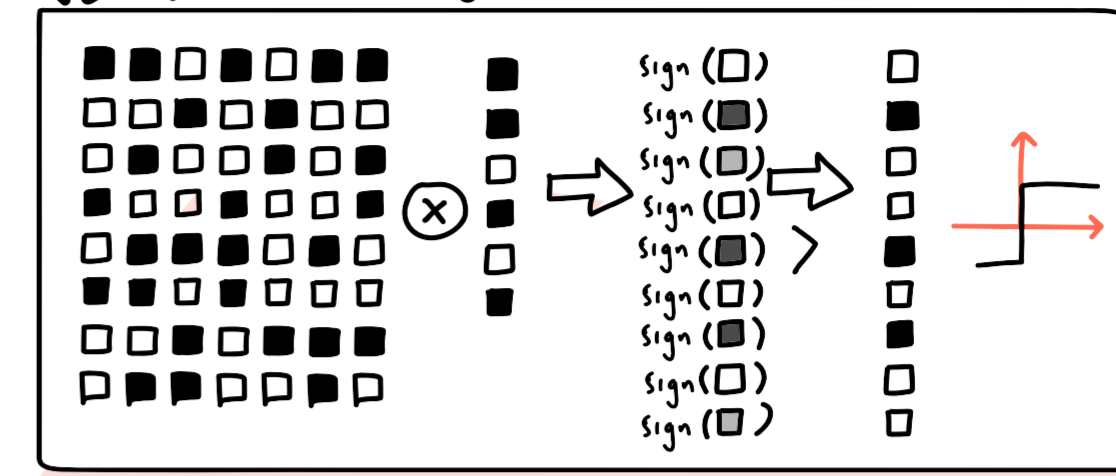
EFFICIENTLY INFER FORCING FUNCTIONS MODELLED as GAUSSIAN PROCESSES!

LINEAR PDEs MODEL the DISTRIBUTION & SPREAD of PARTICULATE AIR POLLUTION



EFFICIENT & ROBUST OPTIMIZATION MODELS FOR TRAINED BINARIZED DEEP NEURAL NETWORKS

(EXAMPLE OF BDNN)



Efficient & Robust Optimization Models for Trained Binarized Deep Neural Networks

Bubacarr Bah

Medical Council Unit The Gambia,
London School of Hygiene & Tropical
Medicine
Serrekunda, GM

BUBACARR BAH

DEEP LEARNING
REVOLUTIONISED
MACHINE LEARNING

BUT IT ALSO BRINGS MANY
CHALLENGES.

- ⚡ ABSENCE OF MATHEMATICS IT WORKS, BUT HOW?
- ⚡ HIGH DIMENSIONALITY
- ⚡ HETEROGENEITY & INCOMPLETENESS
- ⚡ SCALE-MASSIVE DATASETS FOR STATISTICIANS TO HANDLE
- ⚡ EXPLAINABILITY: DL ALGORITHMS ARE BLACK BOXES
- ⚡ ROBUSTNESS: DL ALGORITHMS NOT ROBUST TO PERTURBATIONS OR NOISE

BINARIZED NEURAL NETWORKS (BDNN)

TESTED ON WISCONSIN BREAST CANCER DATASET & A SYNTHETIC DATASET OF RANDOM POINTS W/ DIFFERENT DIMENSIONS.

- 👁️ BDNN HEURISTIC ALGORITHM ACCURACY WAS HIGHER THAN OTHER APPROACHES ↓ 50 DIMENSIONS OF POINTS
- 👁️ ↑ 50 DIMENSIONS OF POINTS, NN PERFORMED BETTER

RESULTS:

- ✓ MIXED INTEGER PROGRAMMING FORMULATIONS TO TRAIN BDNN
- ✓ HEURISTIC VARIANTS HAVE HIGH ACCURACY ON REAL DATASETS
- ✓ ROBUST OPTIMISATION MODELS CAN BE USED TO ENFORCE ROBUSTNESS DURING TRAINING

CONSUME LESS MEMORY
MORE ROBUST AGAINST NOISE/ADVERSARIAL ATTACKS

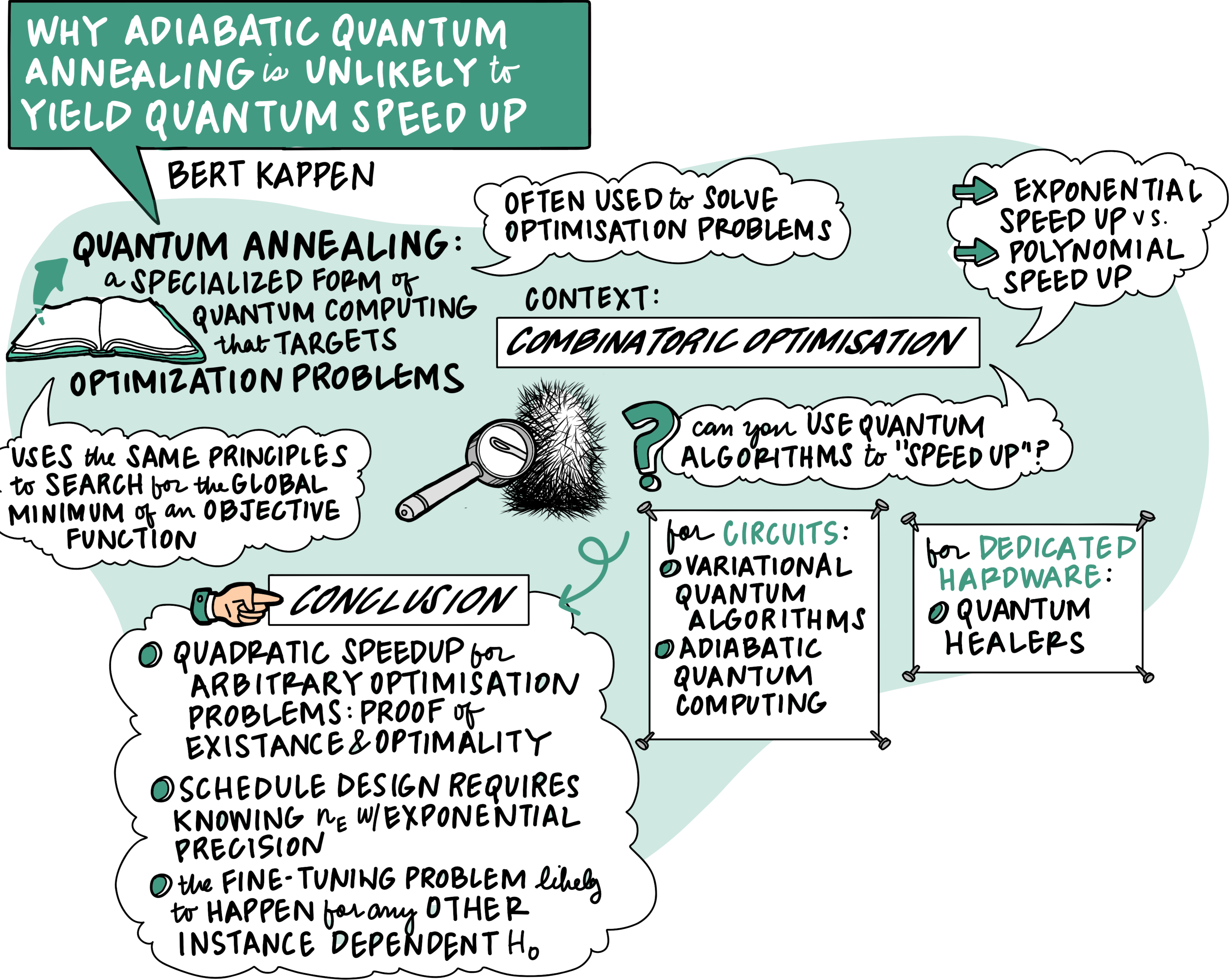
OPEN PROBLEMS:

- DERIVATIVES OF MORE TRACTABLE REFORMATIONS
- USE MORE GENERAL DISCREET ACTIVATION FUNCTIONS TO INCREASE THE COMPLEXITY OF THE NETWORK



Why Adiabatic Quantum Annealing is Unlikely to Yield Quantum Speed Up

Bert Kappen
Radboud University
Nijmegen, NL



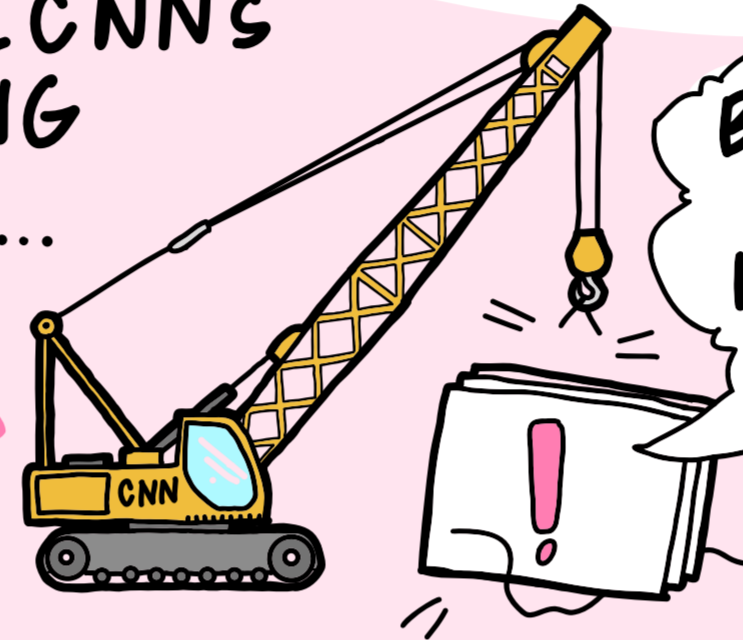
Equivariant Convolutions with Tensor Images

Wilson Gregory
Johns Hopkins University
Baltimore, US

EQUIVARIANT CONVOLUTIONS W/TENSOR IMAGES

WILSON GREGORY

we HAVE
POWERFUL CNNs
for WORKING
w/IMAGES...

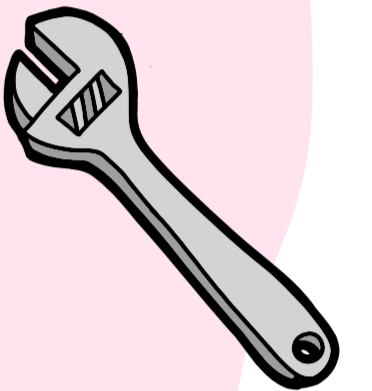


BUT PHYSICS IMAGES
are VECTORS &
MUST FOLLOW the
GEOMETRIC
PRINCIPLE

the LAWS of PHYSICS MUST be
EXPRESSIBLE as GEOMETRIC
RELATIONSHIPS betw. GEOMETRIC
OBJECTS that REPRESENT
PHYSICAL ENTITIES.



our GOAL: EXTEND CNNs
to ACCOMODATE IMAGES of
GEOMETRIC OBJECTS, &
BUILD NETWORKS to
RESPECT COORDINATE
FREEDOM



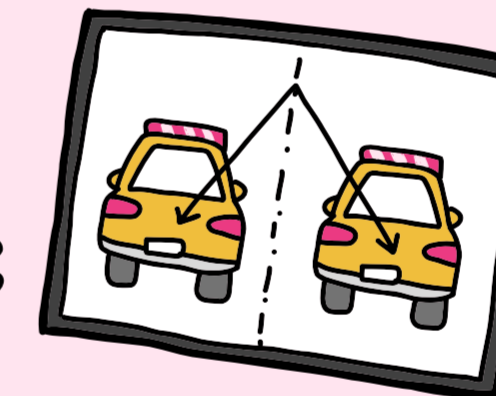
MATH INVOLVED

- **TENSORS** to DESCRIBE
MULTILINEAR RELATION-
SHIPS betw. ALGEBRAIC
OBJECTS in a VECTOR
SPACE...
- **CONVOLUTIONS** to
UNDERSTAND OVERLAPS
of FUNCTIONS..
- **EQUIVARIANCES**
to SHOW SYMMETRIES
betw. FUNCTIONS

our MODEL:

GEOMETRICIMAGENET

- ... MULTIPLE INPUT
CHANNELS
- ... SPECIFY TENSOR
ORDER of FILTERS
- ... CAPTURE LAYER
INTERACTIONS



EQUIVARIANCES
SHOULD PROVIDE
BIG GAINS!

the APPLICATIONS

LEARNING GRAVITATIONAL
FIELDS are a GOOD
WAY to FIND MATCHES
BETWEEN GROUND
TRUTHS & PREDICTIONS



COLLABORATIVE ML for SCIENCE

SAMUEL KASKI

EMPIRICAL RESEARCH FIELDS all FOLLOW the "TEST-ANALYZE-DESIGN-MAKE" PROCESS

for EXAMPLE...

CAN we BUILD a SINGLE BAYESIAN OPTIMISATION TOOL that COULD be USED ACROSS FIELDS?

a COOPERATIVE AI ASSISTANT to HELP INVESTIGATE more EFFECTIVELY!

EG DRUG DESIGN

ALREADY RUNNING a VIRTUAL CYCLE. (SIMULATIONS)

WHAT KIND of SOFTWARE do we NEED for VIRTUAL LABS?

DIGITAL TWINS of INSTRUMENTS & PROCESSES w/ MODELS of INTENTIONS

ENGAGING w/ DOMAIN EXPERTS

(1) PROJECTIVE PREFERENTIAL BAYESIAN OPTIMISATION USER HELPS ID LOCATIONS for SIMULATIONS

(2) ACTIVE LEARNING FEEDBACK for a REINFORCEMENT LEARNING ENGINE

WHAT ABOUT TASKS WHERE 2 PEOPLE NEED to COLLABORATE?

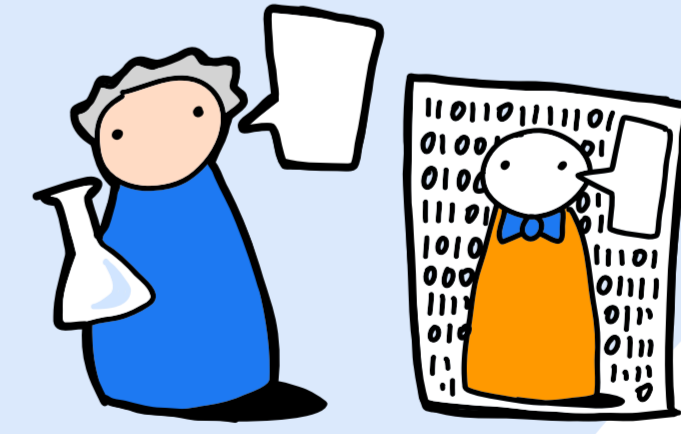
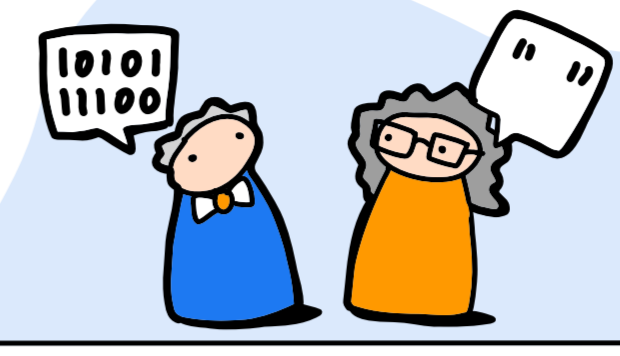
CHALLENGE: PEOPLE DON'T JUST PASSIVELY GIVE the FUNCTION

USER MODELS can HELP AI LEARN to BEST ASSIST

We can USE COGNITIVE SCIENCE for PRIOR KNOWLEDGE on USER BEHAVIORS/INTERACTIONS

EVEN w/ TACIT, UNKNOWN, or EVOLVING DESIRES!

HOW to AVOID BURDEN on SCIENTIST?



Collaborative ML for Science

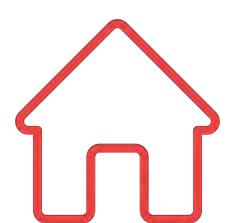
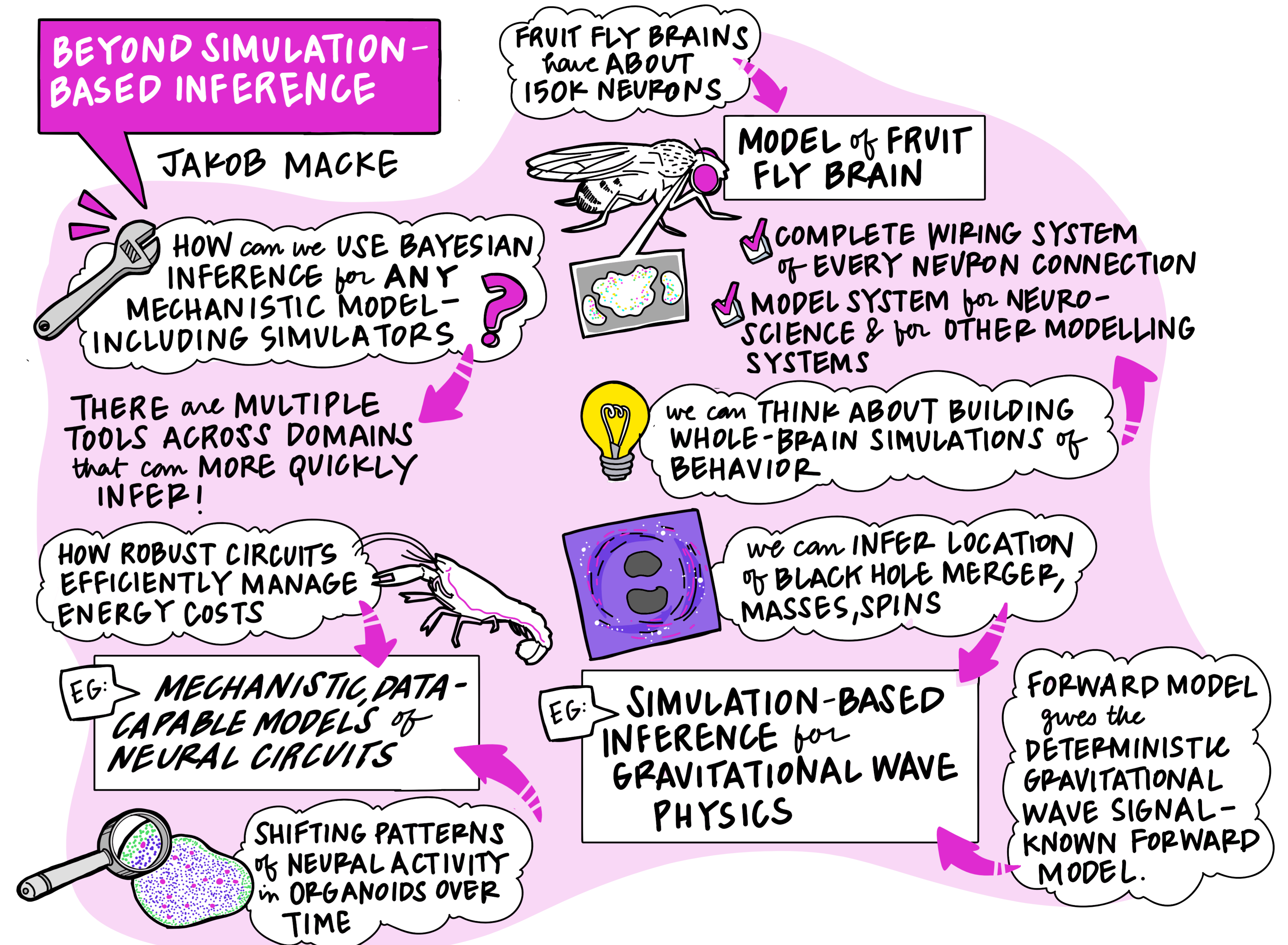
Samuel Kaski

University of Manchester,
Aalto University
Manchester, UK,
Espoo, FI



Beyond Simulation- Based Inference

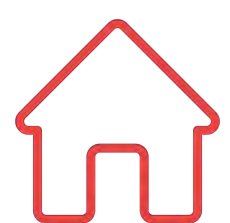
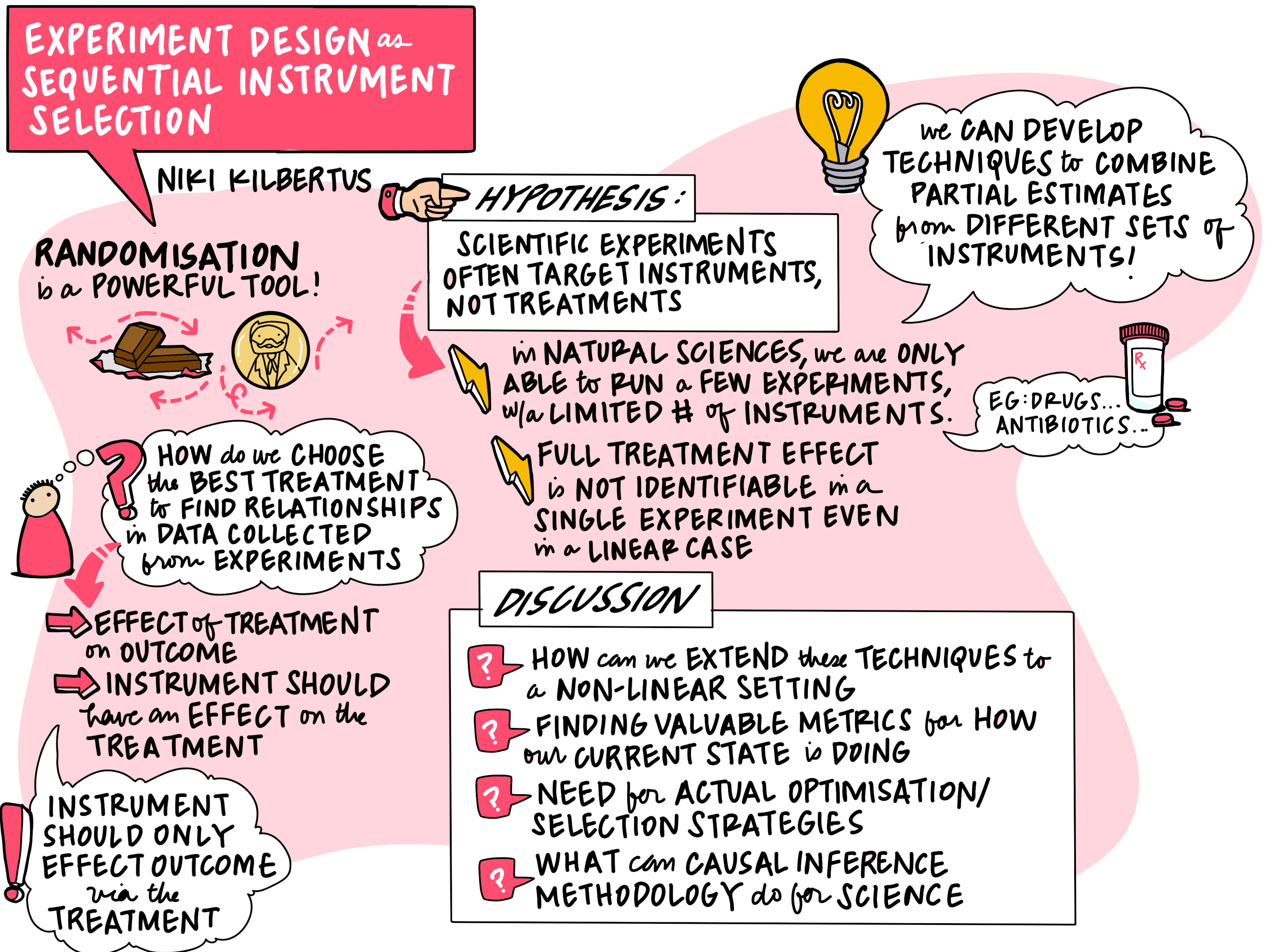
Jakob Macke
University of Tübingen
Tübingen, DE



Experiment Design as Sequential Instrument Selection

Niki Kilbertus

Technical University of Munich
Munich, DE



Machine Learning for Improved Understanding & Projections of Climate Change

Veronika Eyring

German Aerospace Center (DLR) Institute of Atmospheric Physics,
University of Bremen
Bremen, DE

MACHINE LEARNING for IMPROVED UNDERSTANDING & PROJECTIONS of CLIMATE CHANGE

VERONIKA EYRING

USMILE PROJECT



DATA-DRIVEN, PHYSICS-AWARE APPROACHES

- ➔ ML OBSERVATION PRODUCTS for ADVANCED MODEL EVALUATION
- ➔ PATTERNS of CLIMATE VARIABILITY & EXTREMES
- ➔ CONSTRAINTS for FEEDBACKS & PROJECTIONS
- ➔ ML BASED PHYSICS-AWARE PARAMETERISATIONS & HYBRID MODELS

ACTION ITEMS

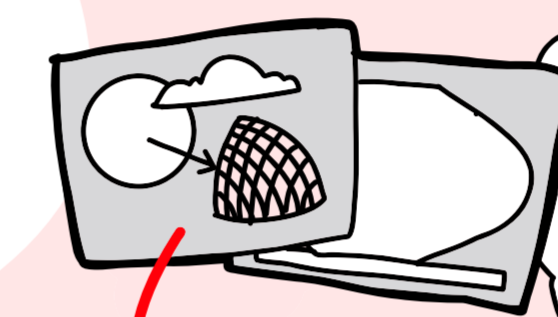
- ML SIDE ○ PHYSICS-AWARE ML & GENERALIZATIONS
- DATA-MODEL INTEGRATION
- GENERATION of ENSEMBLES w/PROBABILISTIC UNCERTAINTIES
- BENCHMARKS
- REQ. TRAINING of NEW GEN. of SCIENTISTS @INTERFACE ↔ CLIMATE SCIENCE & ML

CO₂ is 50% ABOVE PRE-INDUSTRIAL LEVELS

EFFECTING EVERY INHABITED REGION of the WORLD



RECENT CLIMATE CHANGE is UNPRECEDENTED



MODELS are our WAY of SIMULATING the CLIMATE SYSTEMS we OBSERVE

BASED on the FUNDAMENTAL LAWS of PHYSICS

VISUALISATIONS HELP us COMMUNICATE FINDINGS to POLICYMAKERS

UNCERTAINTIES in PROJECTIONS REMAIN



- ! SUBGRID PARAMETERISATION?
- ! HOW can we TRAIN MODELS to GENERALISE ACROSS LATITUDES? (eg. CLOUD COVER)
- ! HOW to IMPROVE TRUST, INTERPRETABILITY, GENERALISATION to EASILY IMPLEMENT MODELS?
- ! DEVELOPMENT of METHODS that ALLOW CAUSAL INFERENCE



PRACTICAL AI CHALLENGES for EARTH & SUSTAINABILITY

CHALLENGES to LOOK at IN the SCIENTIFIC COMMUNITY

Practical AI Challenges for Earth and Sustainability

Markus Reichstein
Max Planck Institute for Biogeochemistry
Jena, DE

MARKUS REICHSTEIN

the EARTH SYSTEM is **UNIQUE**

HIGHLY COMPLEX - COMBINING BIOLOGY, CHEMISTRY, & PHYSICS

we MUST RELY on OBSERVATIONS RATHER than REPLICATIONS

NET CO₂ EXCHANGE
DIFFERENT ECOSYSTEMS have DIFFERENT CHARACTERISTICS

FLUXNET: OBSERVATIONAL ML BASED MODEL of "BREATHING of the BIOSPHERE"

we are TRYING to get to **HIGH SPATIO-TEMPORAL RESOLUTION.**

this is a **DATA CHALLENGE...** CURRENTLY can REACH 5km RESOLUTION

DO we HAVE a **CONCEPTUAL MODEL** of ECOSYSTEM-ATMOSPHERE INTERACTIONS?

SPATIAL PREDICTION: CLASSIC CROSS-VALIDATION STRATEGIES DON'T WORK WELL

HOW can we do THIS at **LANDSCAPE SCALE?**

THERE are **MULTIPLE DRIVERS** of ECOSYSTEM BEHAVIORS.

EARTHNET is LOOKING at **LANDSCAPE EVOLUTION**

HYBRID APPROACHES can HELP ADDRESS DISTRIBUTION SHIFT CONCERNS

VITUS BENSON

ATMOSPHERIC TRANSPORTATION

INTEGRATING a PDE in 3D

TRANSPORT MODELS can SIMULATE CIRCULATION of TRACE GAS GLOBALLY

EMULATORS for WEATHER FORECASTING PREDICT CHANGES in CO₂ CONCENTRATION



LIVING in the PHYSICS-ML INTERPLAY for the EARTH SCIENCES

GUSTAV CAMPS-VALLS



we MEASURE the EARTH w/a VARIETY of SYSTEMS & DEVICES

"MODELS w/o DATA are FANTASIES; DATA w/o MODELS are CHAOS"
- PATRICK GALL

HYBRID MODELS



PHYSICS-AWARE ML MEANS you can EXPLOIT DATA & some PHYSICAL KNOWLEDGE & CREATE MODELS that MINIMISE ERRORS/BREAK PHYSICAL LAWS

- ... we CAN PREDICT CROP YIELDS from SPACE!
- ... we CAN LOOK@ the STATUS of OCEANS & COASTLINES!
- ... we CAN ESTIMATE the ATMOSPHERE & AIR QUALITY!

ENSURING PHYSICAL CONSISTENCY is SIMILAR to ENSURING FAIRNESS & VICEVERSA

BUT THERE are ML CHALLENGES as WELL...

- ! can we MAKE ML RESPECT the LAWS of PHYSICS?
- ! DOES INCLUDING DOMAIN KNOWLEDGE HELP in EXTRAPOLATION?
- ! DO HYBRID MODELS get CAUSAL RELATIONS?
- ! PHYSICS is INVARIANT - can INVARIANCES HELP in LEARNING PHYSICAL PROCESSES?

KAI-HENDRIK COHRS...
DOUBLE ML for CAUSAL HYBRID MODELING
CAUSAL EFFECT ESTIMATION for INTERPRETABLE PHYS. PARAMETERS
Q10 TEMP SENSITIVITY in ECOSYS. RESPIRATION
LIGHT-USE EFFICIENCY in NET ECOSYS. EXCH.

EMILIANO DIAZ...
EARTH SYSTEM SCIENCE DATA is HETEROGENOUS & NOISY
USE HETEROGENEITY to FIND CAUSAL DRIVERS w/ GUARANTEES
USE NNETS & ICP to LEARN a CAUSAL REPRESENTATION & ID. SOURCES of HETEROGENEITY

Living in the Physics-ML Interplay for the Earth Sciences

Gustau Camps-Valls
Universitat de València
València, ES



Machine Learning for Uncovering Hidden Relationships & Improving Predictions in the Coupled Earth System Models

Christian Reimers & Alexander Winkler

Max Planck Institute for Biogeochemistry
Aachen University
Jena, DE
Aachen, DE

MACHINE LEARNING for UNCOVERING HIDDEN RELATIONSHIPS & IMPROVING PREDICTIONS in the COUPLED EARTH SYSTEM MODELS

ALEXANDER WINKLER & CHRISTIAN REIMERS



FOCUS: I.D. FEEDBACKS & CAUSAL LINKS in the EXCHANGE of CARBON, WATER, ENERGY FLUXES betw. the TERRESTRIAL BIOSPHERE & ATMOSPHERE



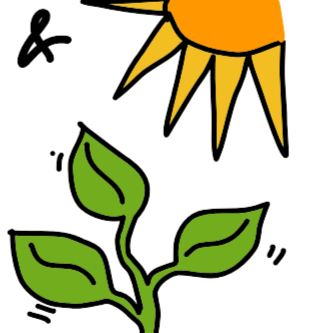
our TOOLS

- MECHANISTIC MODELS of VARYING COMPLEXITY
- CONCEPTUAL MODELS to FULL COUPLED EARTH SYSTEM MODELS
- STATISTICAL ANALYSIS

MODELS HELP us UNDERSTAND LAND FLUX to ANALYZE CARBON UPTAKE



we USED FLUXNET for REAL DATA to LOOK@ CO₂ & TRENDS in PHOTOSYNTHESIS



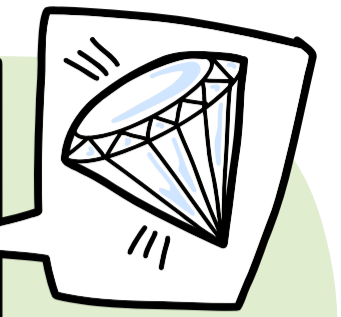
we CONVENED a HACKATHON on CAUSAL CONSISTENCY in ML for EXTRAPOLATION



we TRAINED a STATISTICAL MODEL to LOOK at DRIVERS to CHANGE in PLANT PHENOLOGY



ML is a VALUABLE ASSET BECAUSE...



USE ML & OBSERVATIONAL DATA to IMPROVE our PROCESS UNDERSTANDING.

DRIVER ATTRIBUTION is TRICKY- ML HELPS to DISENTANGLE the SYSTEM

ML HELPS REDUCE UNCERTAINTIES for EXTRAPOLATION

USE ML to INCORPORATE OBSERVATION-INFORMED DATA PARAMETERISATIONS to COUPLED MODELS



SCIENTIFIC QUESTIONS

- can we DETECT & QUANTIFY CO₂ FERTILISATION EFFECT & REDUCE UNCERTAINTY?
- UNDERLYING DRIVERS CAUSING CHANGES in PHENOLOGY?
- ID DIRECT & INDIRECT EFFECT of SOIL MOISTURE VARIABILITY in CONTROLLING CARBON INTAKE?

