Karlsruhe Institute of Technology

Information Theory as a Bridge Across the Geosciences and Modeling Sciences

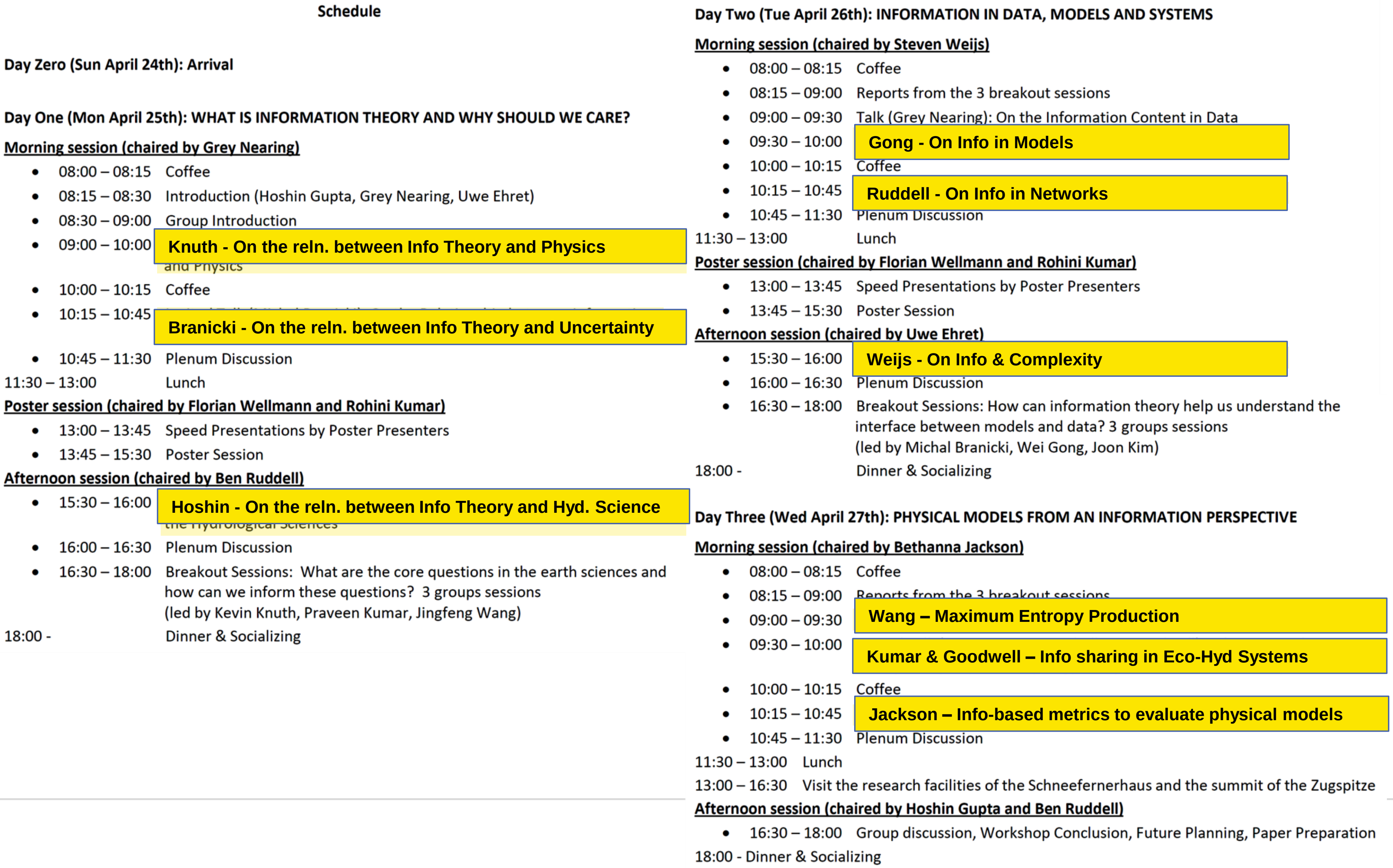
Uwe Ehret (KIT) & Hoshin Gupta (UofA)

Introduction To The Workshop

Sept 11, 2023

Remember 2016 …

Remember 2016 …



Schedule

Day Zero (Sun April 24th): Arrival

Day One (Mon April 25th): WHAT IS INFORMATION THEORY AND WHY SHOULD WE CARE?

Morning session (chaired by Grey Nearing)

- 08:00 – 08:15 Coffee
- 08:15 – 08:30 Introduction (Hoshin Gupta, Grey Nearing, Uwe Ehret)
- 08:30 – 09:00 Group Introduction
- 09:00 – 10:00 **Knuth - On the reln. between Info Theory and Physics**
  and Physics
- 10:00 – 10:15 Coffee
- 10:15 – 10:45 **Branicki - On the reln. between Info Theory and Uncertainty**
- 10:45 – 11:30 Plenum Discussion

11:30 – 13:00 Lunch

Poster session (chaired by Florian Wellmann and Rohini Kumar)

- 13:00 – 13:45 Speed Presentations by Poster Presenters
- 13:45 – 15:30 Poster Session

Afternoon session (chaired by Ben Ruddell)

- 15:30 – 16:00 **Hoshin - On the reln. between Info Theory and Hyd. Science**
  the Hydrological Sciences
- 16:00 – 16:30 Plenum Discussion
- 16:30 – 18:00 Breakout Sessions: What are the core questions in the earth sciences and how can we inform these questions? 3 groups sessions (led by Kevin Knuth, Praveen Kumar, Jingfeng Wang)

18:00 - Dinner & Socializing

Day Two (Tue April 26th): INFORMATION IN DATA, MODELS AND SYSTEMS

Morning session (chaired by Steven Weijs)

- 08:00 – 08:15 Coffee
- 08:15 – 09:00 Reports from the 3 breakout sessions
- 09:00 – 09:30 Talk (Grey Nearing): On the Information Content in Data
- 09:30 – 10:00 **Gong - On Info in Models**
- 10:00 – 10:15 Coffee
- 10:15 – 10:45 **Ruddell - On Info in Networks**
- 10:45 – 11:30 Plenum Discussion

11:30 – 13:00 Lunch

Poster session (chaired by Florian Wellmann and Rohini Kumar)

- 13:00 – 13:45 Speed Presentations by Poster Presenters
- 13:45 – 15:30 Poster Session

Afternoon session (chaired by Uwe Ehret)

- 15:30 – 16:00 **Weijs - On Info & Complexity**
- 16:00 – 16:30 Plenum Discussion
- 16:30 – 18:00 Breakout Sessions: How can information theory help us understand the interface between models and data? 3 groups sessions (led by Michal Branicki, Wei Gong, Joon Kim)

18:00 - Dinner & Socializing

Day Three (Wed April 27th): PHYSICAL MODELS FROM AN INFORMATION PERSPECTIVE

Morning session (chaired by Bethanna Jackson)

- 08:00 – 08:15 Coffee
- 08:15 – 09:00 Reports from the 3 breakout sessions
- 09:00 – 09:30 **Wang – Maximum Entropy Production**
- 09:30 – 10:00 **Kumar & Goodwell – Info sharing in Eco-Hyd Systems**
- 10:00 – 10:15 Coffee
- 10:15 – 10:45 **Jackson – Info-based metrics to evaluate physical models**
- 10:45 – 11:30 Plenum Discussion

11:30 – 13:00 Lunch

13:00 – 16:30 Visit the research facilities of the Schneefernerhaus and the summit of the Zugspitze

Afternoon session (chaired by Hoshin Gupta and Ben Ruddell)

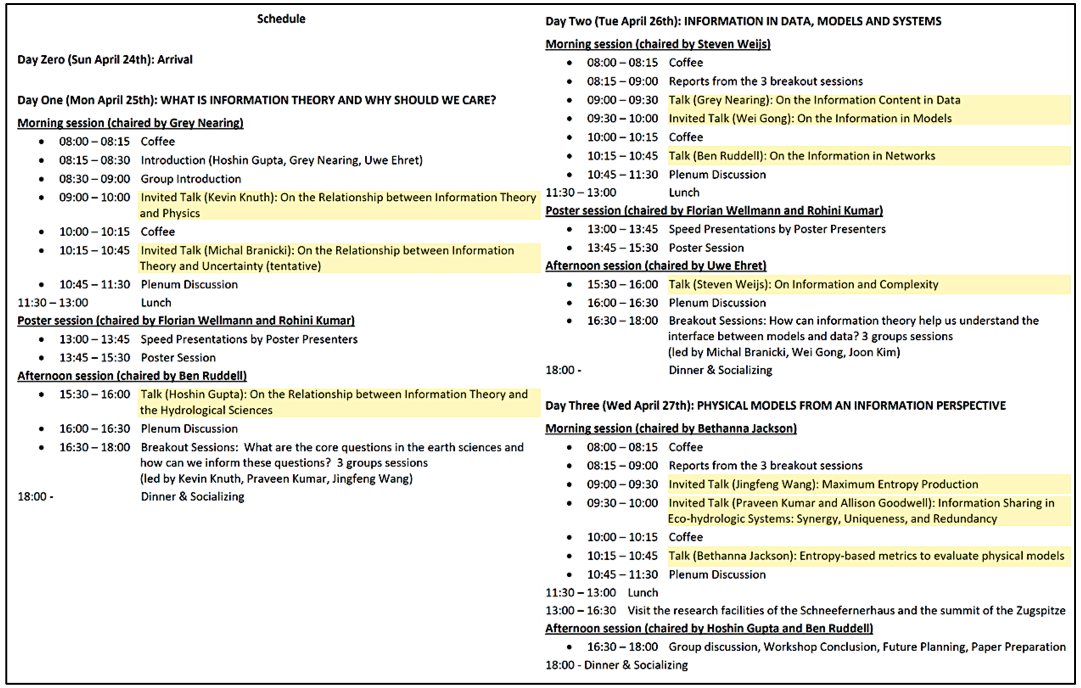
- 16:30 – 18:00 Group discussion, Workshop Conclusion, Future Planning, Paper Preparation

18:00 - Dinner & Socializing

Remember 2016 … Fussball Champions

Since 2016 …

- Moved beyond "*Info ≡ Shannon Info*“
- Progress on *Causality* and *Transfer Entropy*
- No limits to applicability discovered (yet :-)
- *Information Bottleneck* as an inferential guideline
- The rise of ML
  - A broad *Representational Framework*
  - Sobering moments for *Theory-based modeling*
  - Better awareness of the tight integration of all inferential components



Schedule

Day Zero (Sun April 24th): Arrival

Day One (Mon April 25th): WHAT IS INFORMATION THEORY AND WHY SHOULD WE CARE?

Morning session (chaired by Grey Nearing)

- 08:00 – 08:15 Coffee
- 08:15 – 08:30 Introduction (Hoshin Gupta, Grey Nearing, Uwe Ehret)
- 08:30 – 09:00 Group Introduction
- 09:00 – 10:00 Invited Talk (Kevin Knuth): On the Relationship between Information Theory and Physics
- 10:00 – 10:15 Coffee
- 10:15 – 10:45 Invited Talk (Michal Branicki): On the Relationship between Information Theory and Uncertainty (tentative)
- 10:45 – 11:30 Plenum Discussion

11:30 – 13:00 Lunch

Poster session (chaired by Florian Wellmann and Rohini Kumar)

- 13:00 – 13:45 Speed Presentations by Poster Presenters
- 13:45 – 15:30 Poster Session

Afternoon session (chaired by Ben Ruddell)

- 15:30 – 16:00 Talk (Hoshin Gupta): On the Relationship between Information Theory and the Hydrological Sciences
- 16:00 – 16:30 Plenum Discussion
- 16:30 – 18:00 Breakout Sessions: What are the core questions in the earth sciences and how can we inform these questions? 3 groups sessions
  (led by Kevin Knuth, Praveen Kumar, Jingfeng Wang)

18:00 - Dinner & Socializing

Day Two (Tue April 26th): INFORMATION IN DATA, MODELS AND SYSTEMS

Morning session (chaired by Steven Weijs)

- 08:00 – 08:15 Coffee
- 08:15 – 09:00 Reports from the 3 breakout sessions
- 09:00 – 09:30 Talk (Grey Nearing): On the Information Content in Data
- 09:30 – 10:00 Invited Talk (Wei Gong): On the Information in Models
- 10:00 – 10:15 Coffee
- 10:15 – 10:45 Talk (Ben Ruddell): On the Information in Networks
- 10:45 – 11:30 Plenum Discussion

11:30 – 13:00 Lunch

Poster session (chaired by Florian Wellmann and Rohini Kumar)

- 13:00 – 13:45 Speed Presentations by Poster Presenters
- 13:45 – 15:30 Poster Session

Afternoon session (chaired by Uwe Ehret)

- 15:30 – 16:00 Talk (Steven Weijs): On Information and Complexity
- 16:00 – 16:30 Plenum Discussion
- 16:30 – 18:00 Breakout Sessions: How can information theory help us understand the interface between models and data? 3 groups sessions
  (led by Michal Branicki, Wei Gong, Joon Kim)

18:00 - Dinner & Socializing

Day Three (Wed April 27th): PHYSICAL MODELS FROM AN INFORMATION PERSPECTIVE

Morning session (chaired by Bethanna Jackson)

- 08:00 – 08:15 Coffee
- 08:15 – 09:00 Reports from the 3 breakout sessions
- 09:00 – 09:30 Invited Talk (Jingfeng Wang): Maximum Entropy Production
- 09:30 – 10:00 Invited Talk (Praveen Kumar and Allison Goodwell): Information Sharing in Eco-hydrologic Systems: Synergy, Uniqueness, and Redundancy
- 10:00 – 10:15 Coffee
- 10:15 – 10:45 Talk (Bethanna Jackson): Entropy-based metrics to evaluate physical models
- 10:45 – 11:30 Plenum Discussion

11:30 – 13:00 Lunch

13:00 – 16:30 Visit the research facilities of the Schneefernerhaus and the summit of the Zugspitze

Afternoon session (chaired by Hoshin Gupta and Ben Ruddell)

- 16:30 – 18:00 Group discussion, Workshop Conclusion, Future Planning, Paper Preparation

18:00 - Dinner & Socializing

Of Course… Many Problems Still Remain to be Solved …

- *Theory-based (TB) models* are based on incomplete understanding of the world
  - (severely) *Lossy Compression* of *Data* due to overly strong (or wrong) constraints imposed by theory

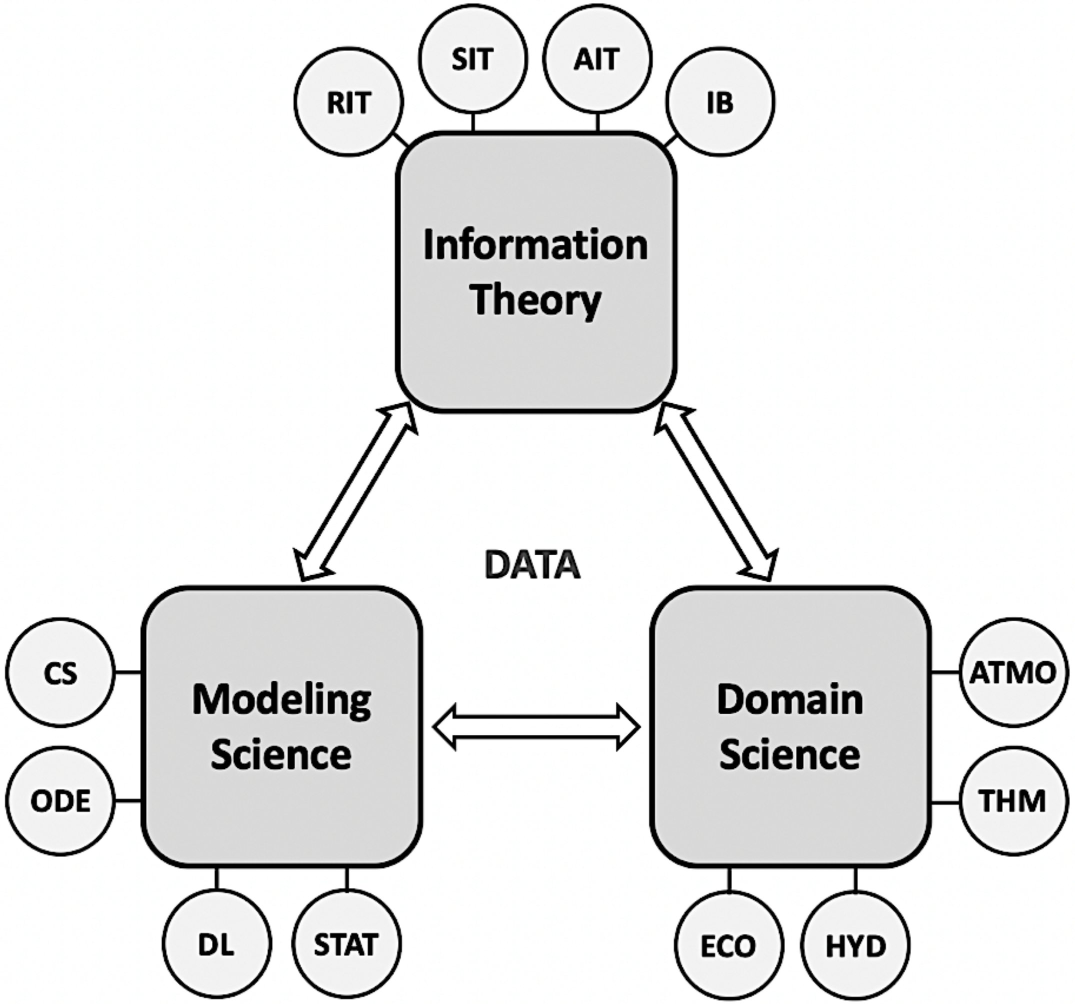
- *Data-based (DB) models* outperform TB models on specific problems, but typically lack the hierarchical modularity of TB representations
  - This hampers *Interpretation, Reasoning, Transfer* (generalization across domains)

Workshop Focus

- **To Explore the Nexus of:**
  - *Information Theory (IT)*
  - *Modeling Sciences (MS)*
  - *Domain Relevant Theory (DRT)*

- **Goal**
  - To enhance the predictive capabilities of ESS models, and their suitability for *Reasoning* and *Understanding*

- **Approach**
  - Closer integration of the *Modeling* and *Domain Sciences*
  - A general framework with IT as a *conceptual* and *linguistic* foundation
  - Expanded understanding of the richness of how "*Information*" is expressed by *Models*



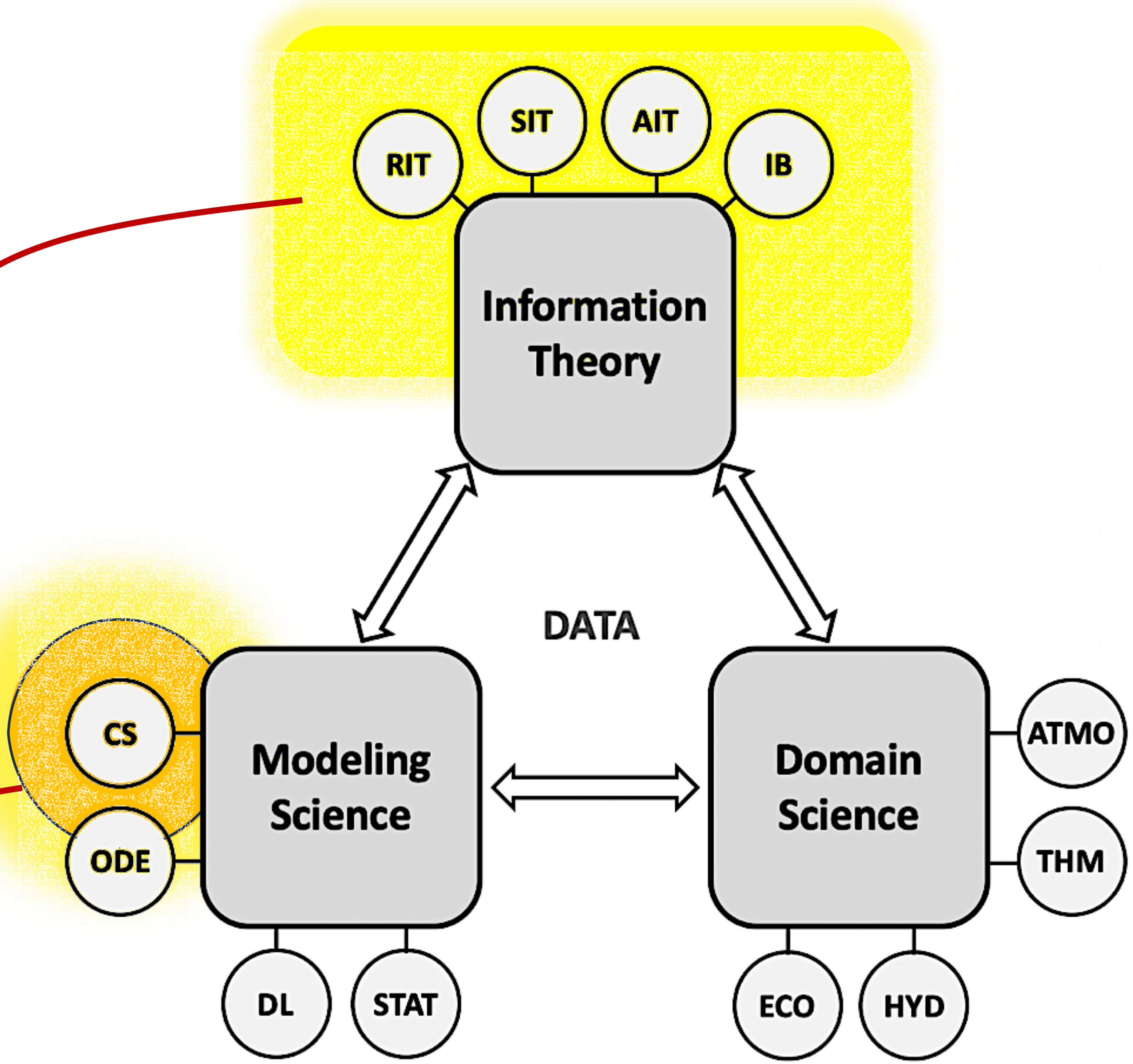
Workshop Focus

- **To Explore the Nexus of:**
  - *Information Theory (IT)*
  - *Modeling Sciences (MS)*
  - *Domain Relevant Theory (DRT)*

- ***Shannon (Statistical) Information***
- ***Algorithmic Information***
- ***Representational Information***

Theory of Computation

(Finite State Automata, Turing Machines, Chomsky Grammars)



Shannon **versus** ***Algorithmic*** **Information**

Shannon (Statistical) Information:

$$I_S(x) = \log_2\left(\frac{1}{p(x)}\right)$$ *Code length of a Probabilistic Description*

- *Relates to* ***repeated events****/objects*
- *Characterizes the (expected/average) "surprise" associated with encountering such events*
- ***Description/Code Length after removing all Statistical Redundancy*** *(statistical compression)*

Algorithmic Information:

$$I_K(x) = K(x) = \log_2\left(\frac{1}{2^{-K(x)}}\right)$$ *Code length of an Algorithmic Description*

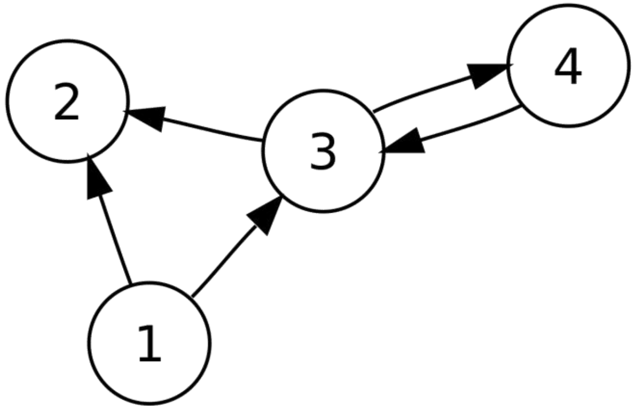
- *Relates to* ***individual events****/objects*
- *Characterizes the (expected/average) "surprise" associated with encountering such events*
- ***Description/Code Length after removing ALL Redundancy (statistical & non-statistical)*** *(structural compression)*

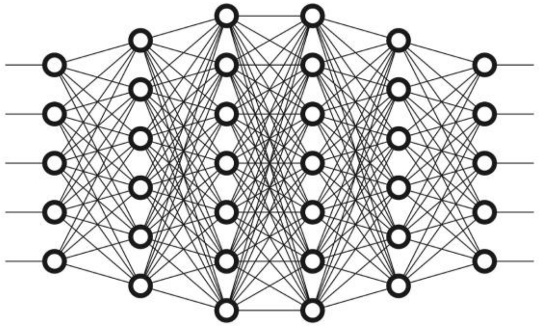
We are interested in
"Minimal" Description Lengths

But there is also Information bound up in ***Representations***

Representational Information for Building Dynamical Systems Models:

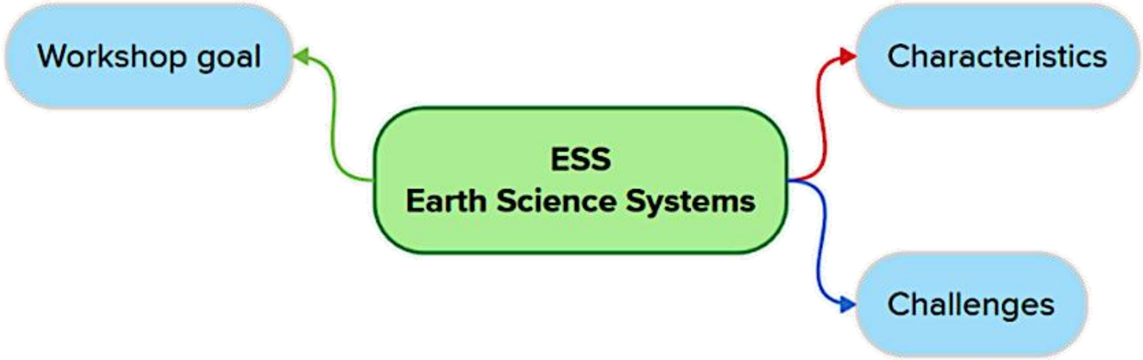
- *Symbols (Alphabet)*
- *Types of Objects (Dictionary … Features, States, Parameters, Mass Energy and Info Flows)*
- *Directed Graphs Structures (Nodes & Links, Associative Relationships, Short- & Long-term Memory)*

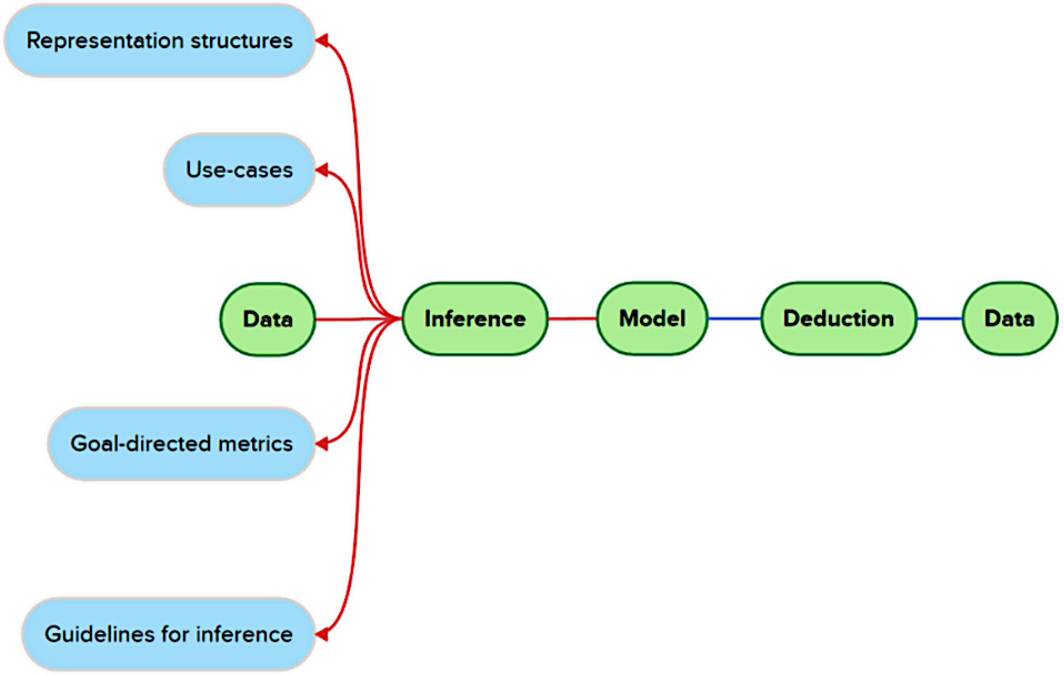


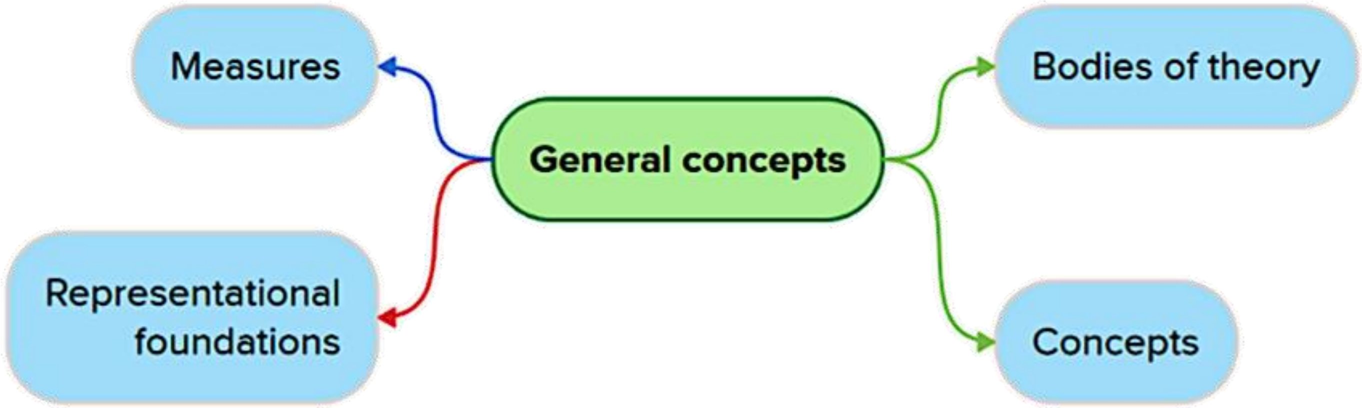


- *Relates to **Holoarchic Structural Organization of Systems***
- *Characterizes the (hypothesized) "generative/relational structure" underlying the generation of events*
- ***Description/Code Length after removing all Representational Redundancy** (representational compression)*
  - ***But also more than simply code length … the structure of the description is important***

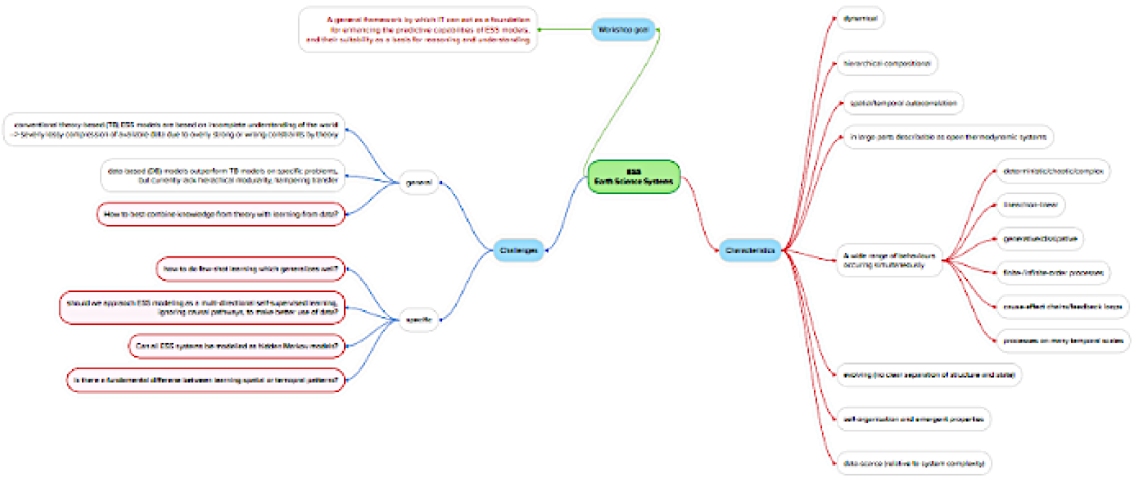
Possible Conceptual Framework for our Discussions

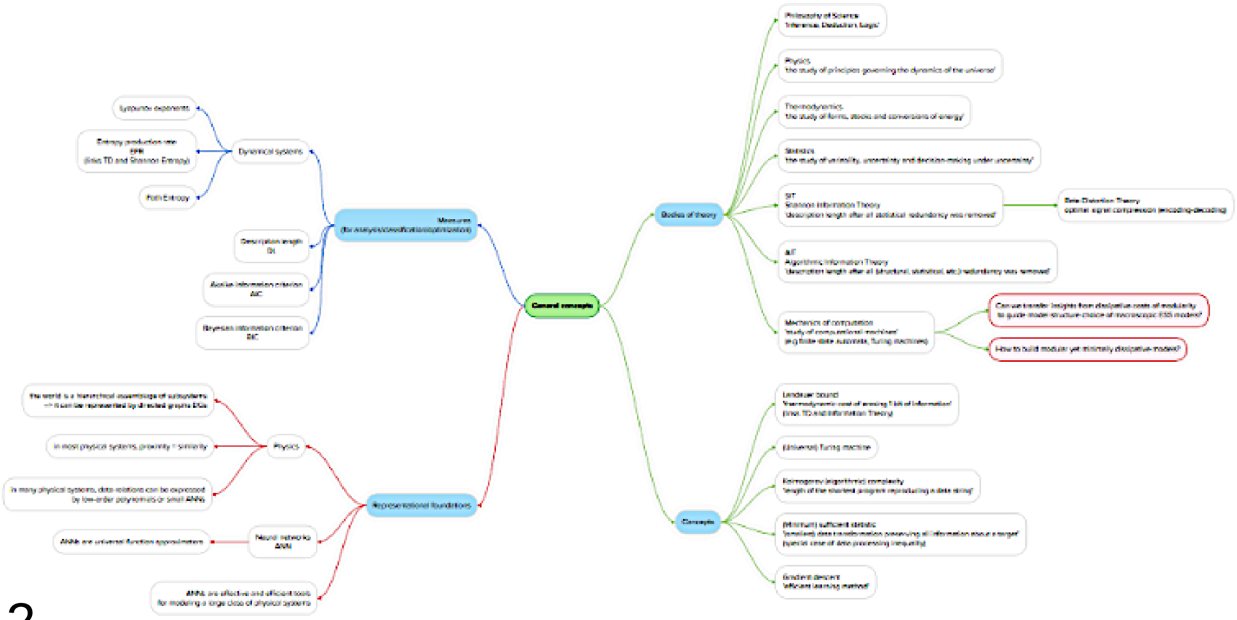


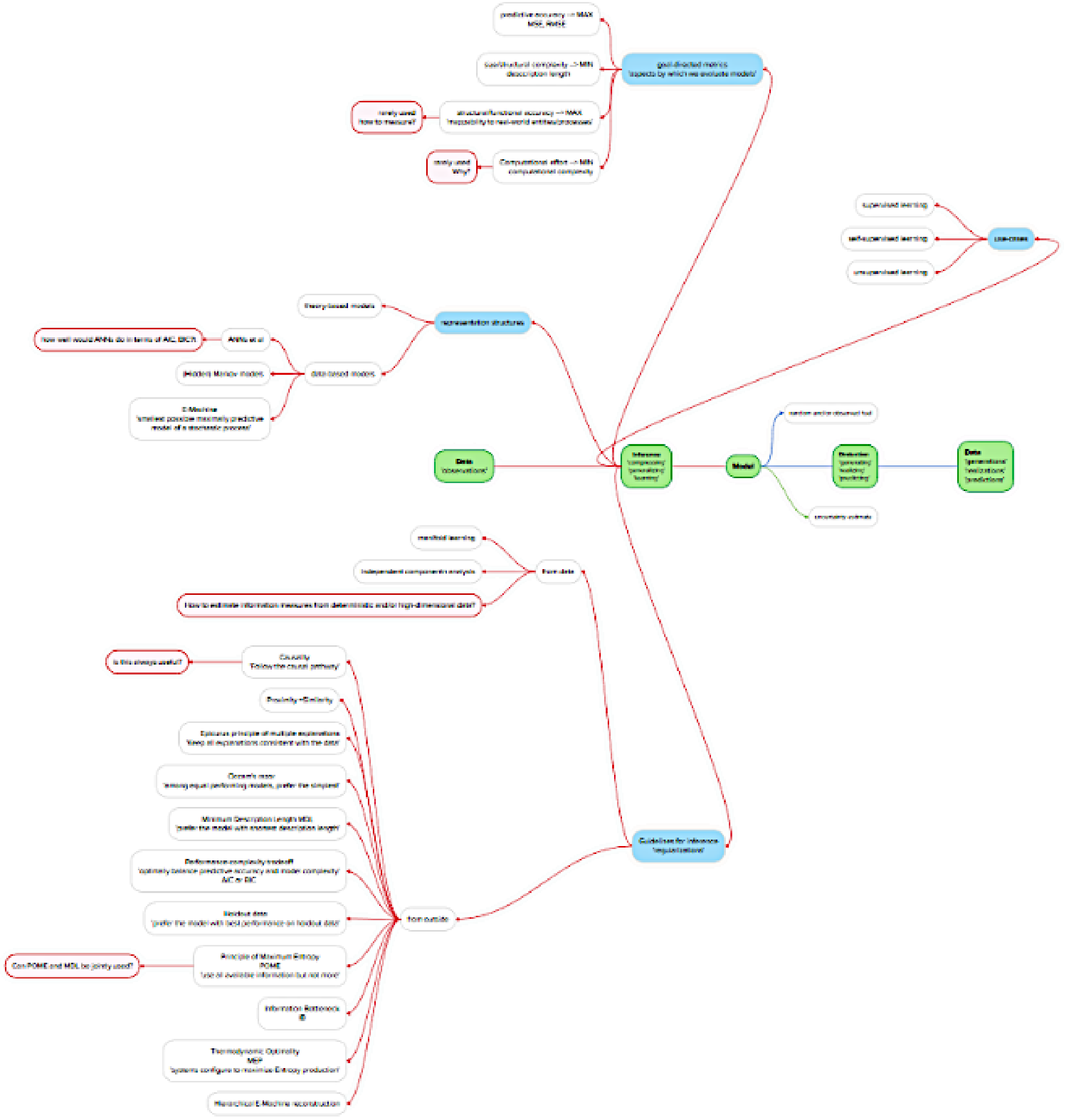




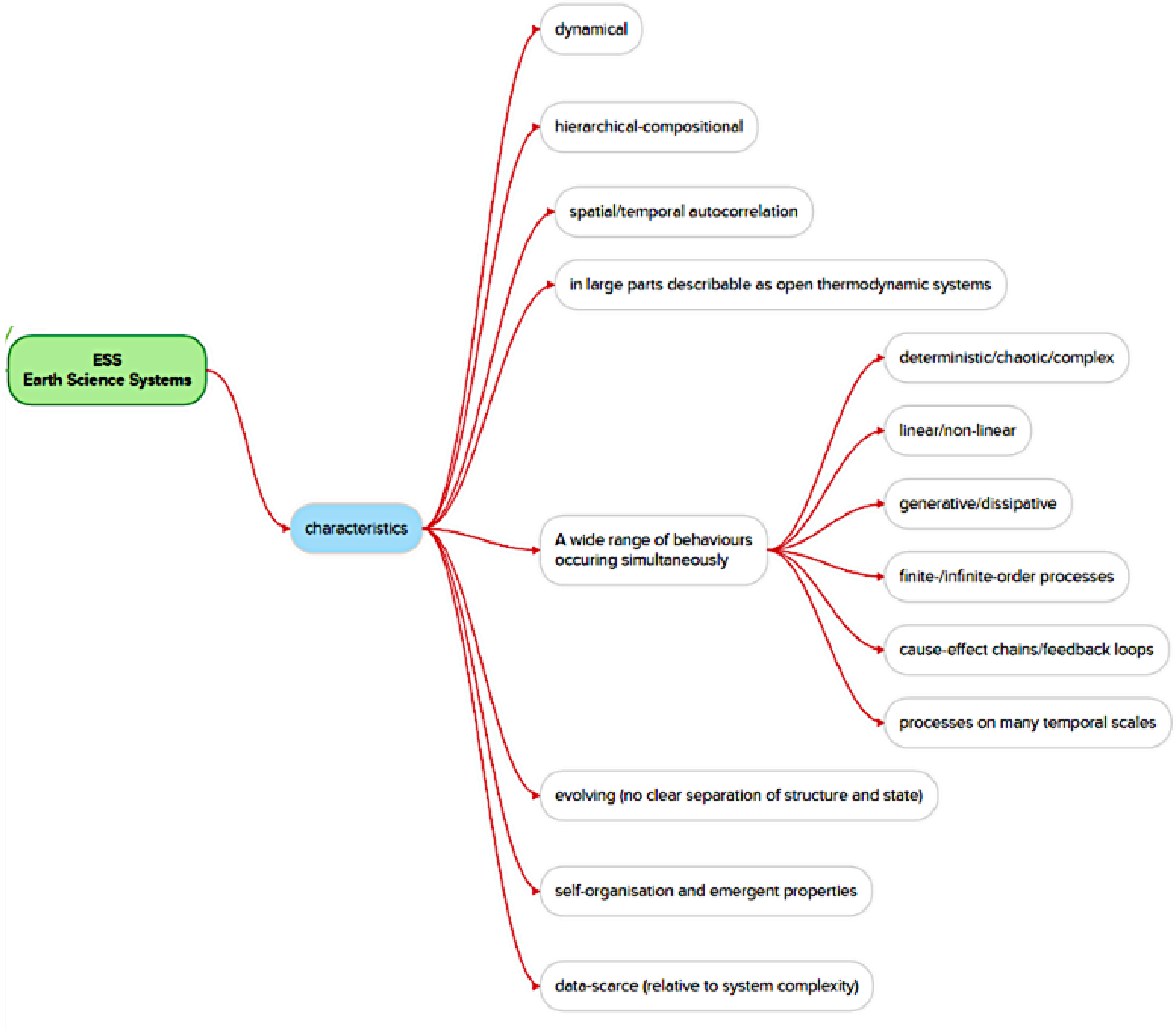
Possible Conceptual Framework for our Discussions


ESS
Earth Science Systems
Workshop goal
A general framework by which IT can act as a foundation for enhancing the predictive capabilities of ESS models, and their suitability as a basis for reasoning and understanding
Challenges
general
conventional theory based (TB) ESS models are based on incomplete understanding of the world -> severly lossy compression of available data due to overly strong or wrong constraints by theory
data based (DB) models outperform TB models on specific problems, but currently lack hierarchical modularity, hampering transfer
How to best combine knowledge from theory with learning from data?
specific
How to do few-shot learning which generalizes well?
Should we approach ESS modeling as a multi-directional self-supervised learning, ignoring causal pathways, to make better use of data?
Can all ESS systems be modeled as Hidden Markov models?
Is there a fundamental difference between learning spatial or temporal patterns?
Characteristics
dynamical
hierarchical compositional
spatial/temporal autocorrelation
in large parts describable as open thermodynamic systems
A wide range of behaviours occuring simultaneously
deterministic/chaotic/complex
linear/non-linear
generative/dissipative
finite-/infinite-order processes
cause-effect chains/feedback loops
processes on many temporal scales
evolving (no clear separation of structure and state)
self-organisation and emergent properties
data scarce (relative to system complexity)

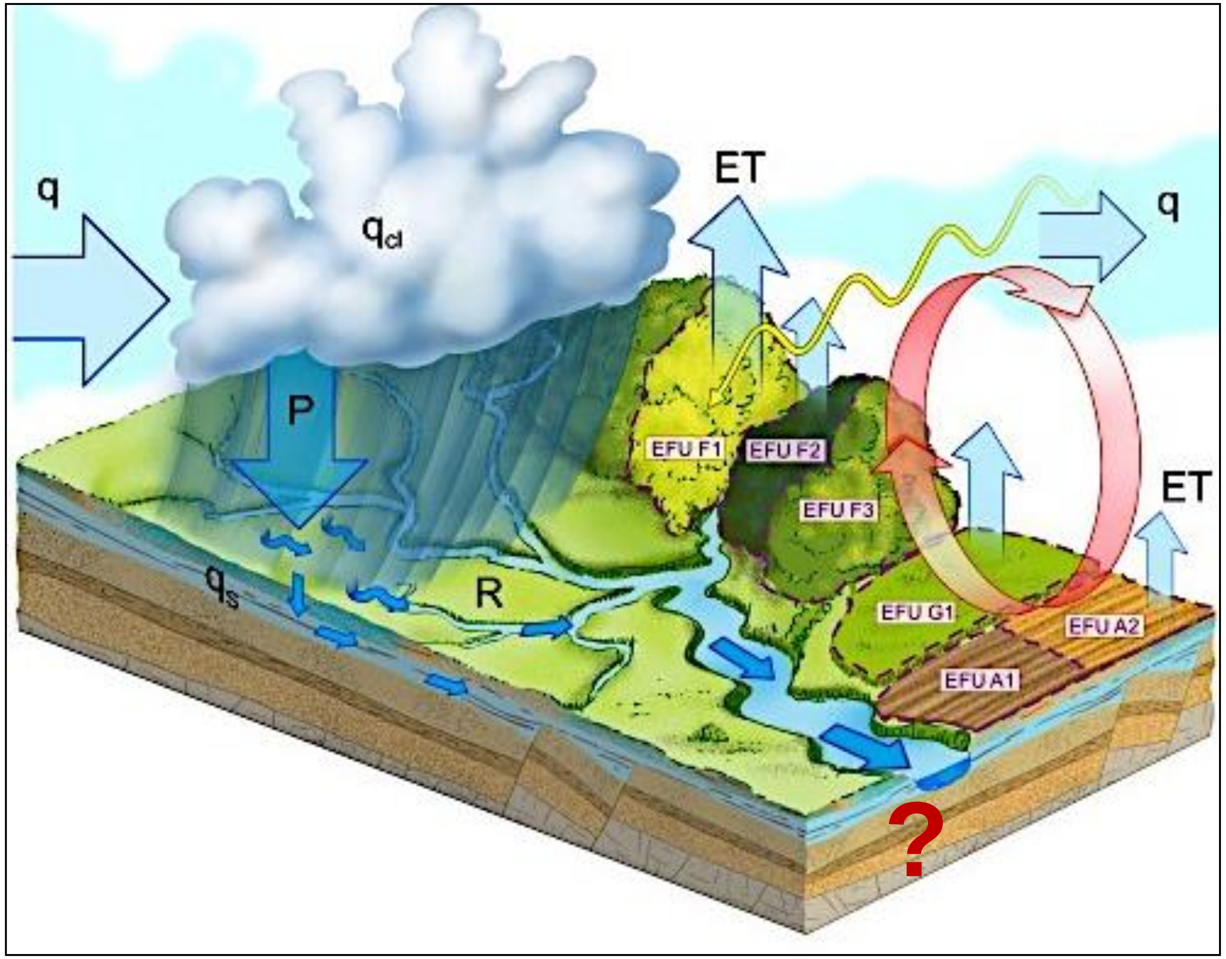

General concepts
Measures
(for analysis/classification/optimization)
Dynamical systems
Lyapunov exponents
Entropy production rate
EPR
(links TD and Shannon Entropy)
Path Entropy
Description length
DL
Akaike information criterion
AIC
Bayesian information criterion
BIC
Representational foundations
Physics
the world is a hierarchical assemblage of subsystems
--> it can be represented by directed graphs DGs
in most physical systems, proximity = similarity
in many physical systems, data relations can be expressed by low-order polynomials or small ANNs
Neural networks
ANN
ANNs are universal function approximators
ANNs are effective and efficient tools for modeling a large class of physical systems
Bodies of theory
Philosophy of Science
'Inference, Deduction, Logic'
Physics
'the study of principles governing the dynamics of the universe'
Thermodynamics
'the study of forms, stocks and conversions of energy'
Statistics
'the study of variability, uncertainty and decision-making under uncertainty'
SIT
Shannon Information Theory
'description length after all statistical redundancy was removed'
Rate Distortion Theory
optimal signal compression (encoding-decoding)
AIT
Algorithmic Information Theory
'description length after all (structural, statistical, etc.) redundancy was removed'
Mechanics of computation
'study of computational machines'
(e.g finite-state automata, Turing machines)
Can we transfer insights from dissipative costs of modularity to guide model structure choice of macroscopic ESS models?
How to build modular yet minimally dissipative models?
Concepts
Landauer bound
'thermodynamic cost of erasing 1 bit of information'
(links TD and Information Theory)
(Universal) Turing machine
Kolmogorov (algorithmic) complexity
'length of the shortest program reproducing a data string'
(Minimum) sufficient statistic
'(smallest) data transformation preserving all information about a target'
(special case of data processing inequality)
Gradient descent
'efficient learning method'


predictive accuracy --> MAX
MSE, RMSE
size/structural complexity --> MIN
description length
rarely used
How to measure?
structural/functional accuracy --> MAX
'mappability to real-world entities/processes'
rarely used
Why?
Computational effort --> MIN
computational complexity
goal-directed metrics
'aspects by which we evaluate models'
supervised learning
self-supervised learning
unsupervised learning
use-cases
theory-based models
representation structures
How well would ANNs do in terms of AIC, BIC?!
ANNs et al
data-based models
(Hidden) Markov models
E-Machine
'smallest possible maximally predictive model of a stochastic process'
Data
'observations'
Inference
'compressing'
'generalizing'
'learning'
Model
random and/or observed bat
Deduction
'generating'
'realizing'
'predicting'
Data
'generations'
'realizations'
'predictions'
uncertainty estimate
manifold learning
independent components analysis
from data
How to estimate information measures from deterministic and/or high-dimensional data?
Is this always useful?
Causality
'Follow the causal pathway'
Proximity =Similarity
Epicurus principle of multiple explanations
'keep all explanations consistent with the data'
Occam's razor
'among equal performing models, prefer the simplest'
Minimum Description Length MDL
'prefer the model with shortest description length'
Performance-complexity tradeoff
'optimally balance predictive accuracy and model complexity'
AIC or BIC
Holdout data
'prefer the model with best performance on holdout data'
Can POME and MDL be jointly used?
Principle of Maximum Entropy
POME
'use all available information but not more'
Information Bottleneck
IB
Thermodynamic Optimality
MEP
'systems configure to maximize Entropy production'
Hierarchical E-Machine reconstruction
from outside
Guidelines for inference
'regularizations'

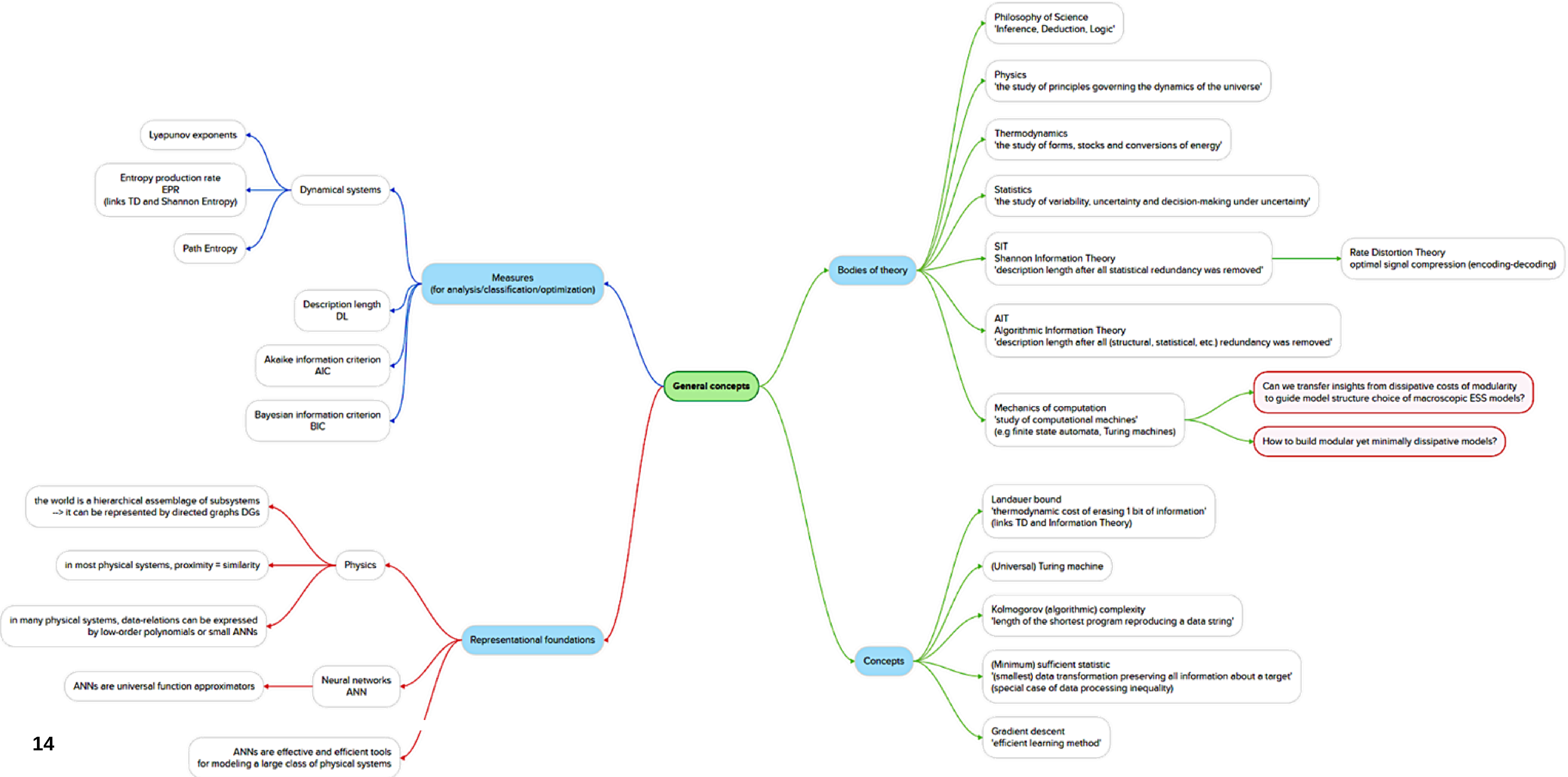
1. Earth Science Systems



- ESS Earth Science Systems
  - characteristics
    - dynamical
    - hierarchical-compositional
    - spatial/temporal autocorrelation
    - in large parts describable as open thermodynamic systems
    - A wide range of behaviours occuring simultaneously
      - deterministic/chaotic/complex
      - linear/non-linear
      - generative/dissipative
      - finite-/infinite-order processes
      - cause-effect chains/feedback loops
      - processes on many temporal scales
    - evolving (no clear separation of structure and state)
    - self-organisation and emergent properties
    - data-scarce (relative to system complexity)



2. General Concepts



General concepts

Measures
(for analysis/classification/optimization)

Dynamical systems

Lyapunov exponents

Entropy production rate
EPR
(links TD and Shannon Entropy)

Path Entropy

Description length
DL

Akaike information criterion
AIC

Bayesian information criterion
BIC

Representational foundations

Physics

the world is a hierarchical assemblage of subsystems
--> it can be represented by directed graphs DGs

in most physical systems, proximity = similarity

in many physical systems, data-relations can be expressed
by low-order polynomials or small ANNs

Neural networks
ANN

ANNs are universal function approximators

ANNs are effective and efficient tools
for modeling a large class of physical systems

Bodies of theory

Philosophy of Science
'Inference, Deduction, Logic'

Physics
'the study of principles governing the dynamics of the universe'

Thermodynamics
'the study of forms, stocks and conversions of energy'

Statistics
'the study of variability, uncertainty and decision-making under uncertainty'

SIT
Shannon Information Theory
'description length after all statistical redundancy was removed'

Rate Distortion Theory
optimal signal compression (encoding-decoding)

AIT
Algorithmic Information Theory
'description length after all (structural, statistical, etc.) redundancy was removed'

Mechanics of computation
'study of computational machines'
(e.g finite state automata, Turing machines)

Can we transfer insights from dissipative costs of modularity
to guide model structure choice of macroscopic ESS models?

How to build modular yet minimally dissipative models?

Concepts

Landauer bound
'thermodynamic cost of erasing 1 bit of information'
(links TD and Information Theory)

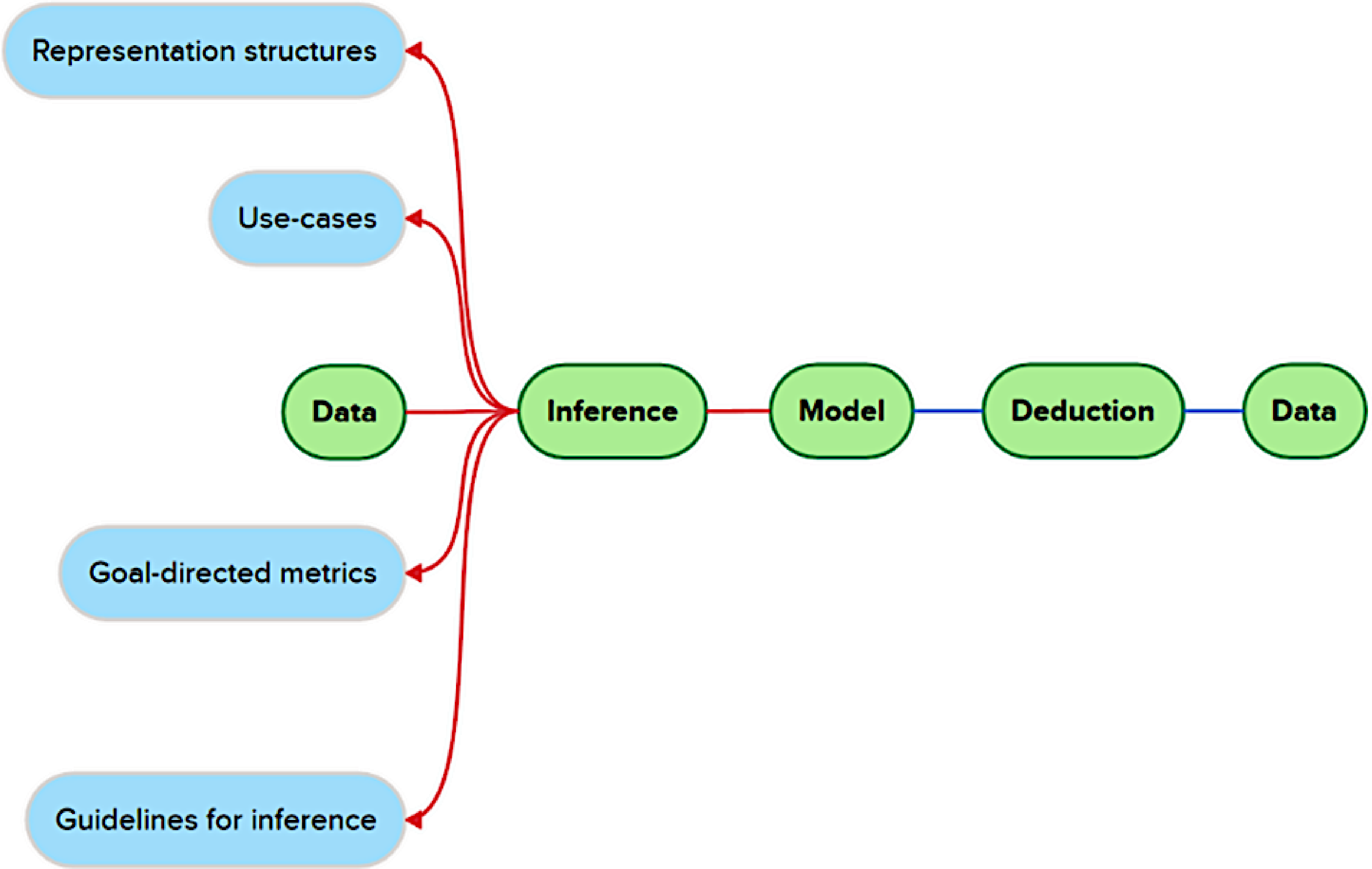
(Universal) Turing machine

Kolmogorov (algorithmic) complexity
'length of the shortest program reproducing a data string'

(Minimum) sufficient statistic
'(smallest) data transformation preserving all information about a target'
(special case of data processing inequality)

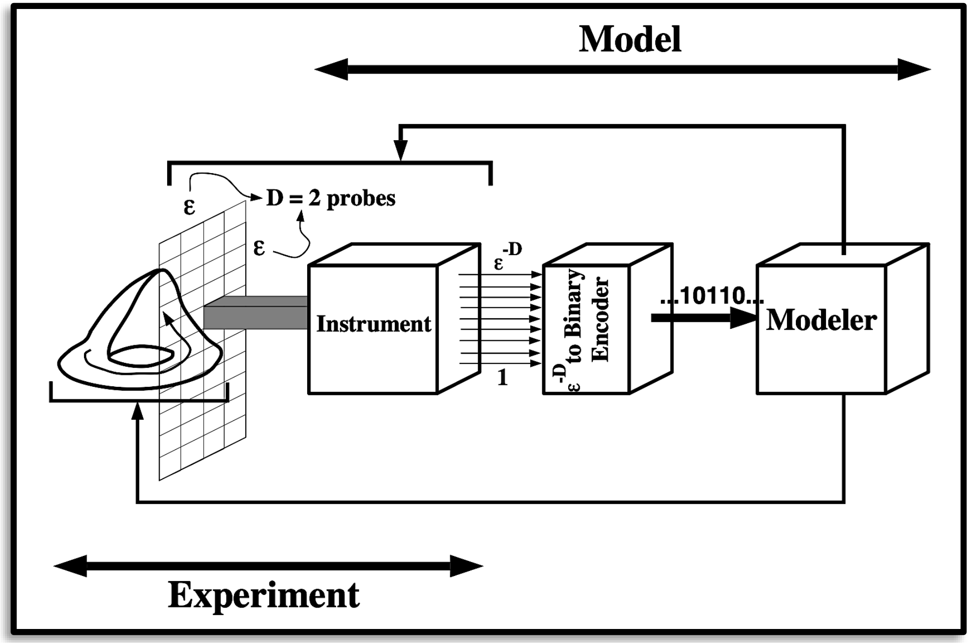
Gradient descent
'efficient learning method'

3. The Operational Pathway (Data → Inference → Deduction)



The Central Problem of Inference

Given an ***Instrument***, some number of ***Measurements***, and fixed ***Finite Inference Resources*** … ***how much Computational Structure in the underlying process can be extracted***?



On what sort of structure in the data stream should the models be based … given that:

- *Individual measurements are only indirect representations of the state*
- *The instrument may not supply data of quality sufficient to discover the true states*

So how can the process's "effective" states be accessed?

JP Crutchfield (1992) Knowledge and meaning… chaos and complexity

The "Info Theory" Persective on Model Development

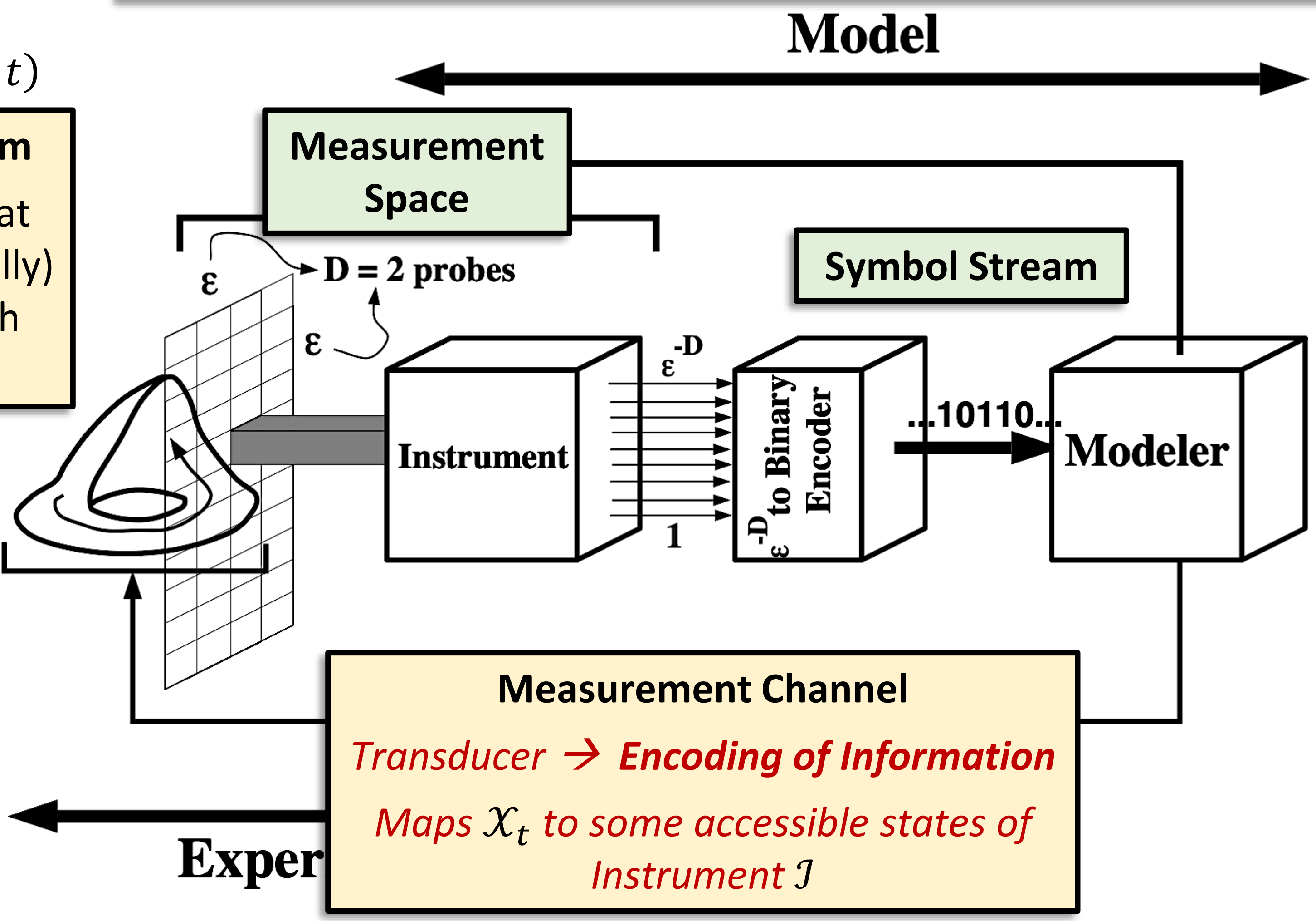
Shannon’s Coding Theorems

The *Capacity* (in bits) of the *Communication Channel* (instrument) between the *Process* and the *Modeler* larger than the *Entropy Generation Rate* of the Process.

$\mathcal{X}_{t+\Delta t} = F(\mathcal{X}_t, \zeta_t, t)$

Real-World System

Configurations that evolve (stochastically) through time with some *Entropy*



Measurement Space

Symbol Stream

Measurement Channel

*Transducer → **Encoding of Information***

Maps $\mathcal{X}_t$ to some accessible states of Instrument $\mathcal{I}$

$X_{t+\Delta t} = f(X_t, \eta_t, t)$

Model/Representation

Effective States that evolve through *State Transitions*

Defined in terms of Available Resources

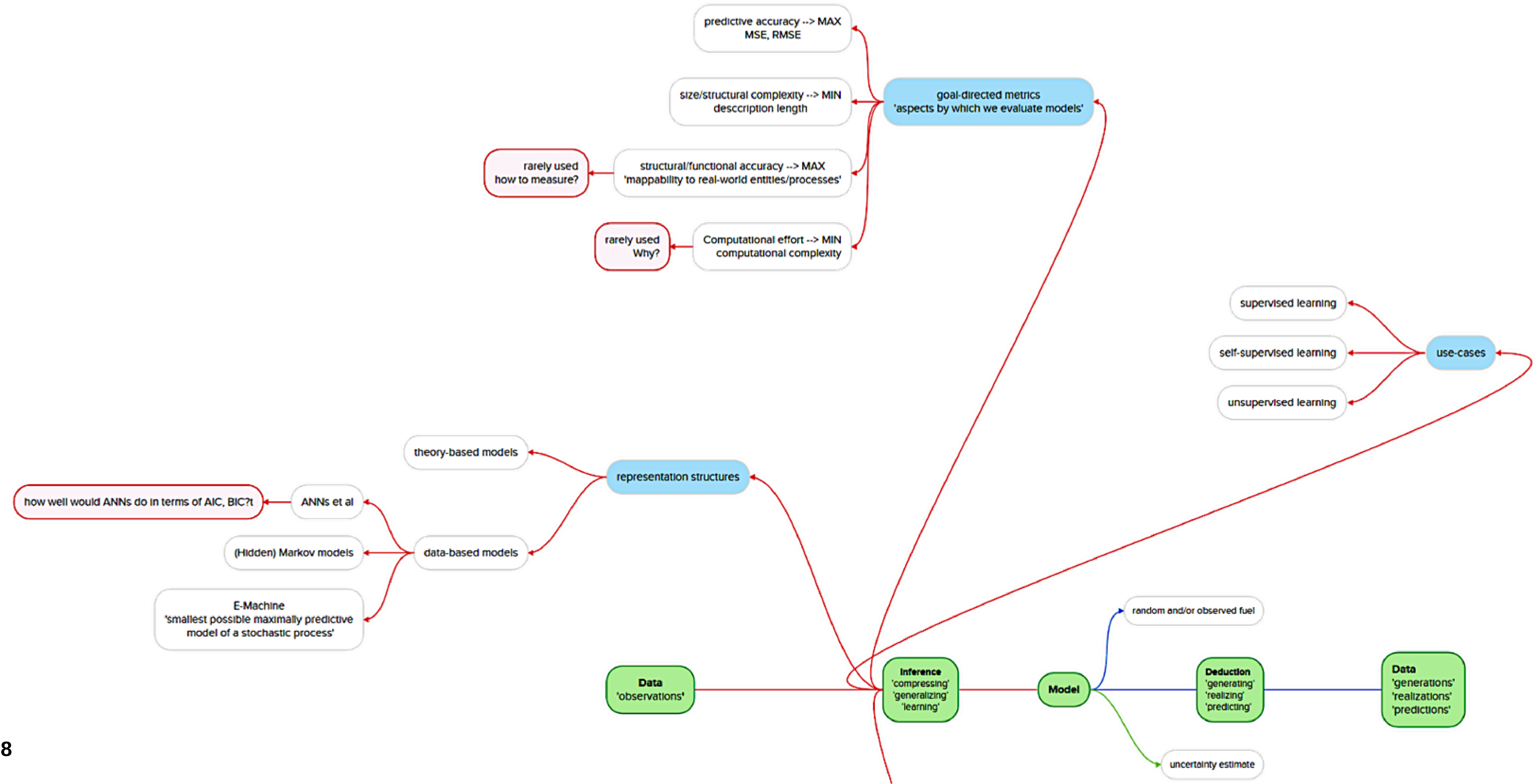
storage capacity, computational power, inferential efficacy, etc.

Effective Noise Level

Determined by the amount of data “unexplained” by the inferred model

JP Crutchfield (1992) Knowledge and meaning… chaos and complexity

3. The Operational Pathway (Data → Inference → Deduction)



predictive accuracy --> MAX
MSE, RMSE

size/structural complexity --> MIN
desccription length

rarely used
how to measure?

structural/functional accuracy --> MAX
'mappability to real-world entities/processes'

rarely used
Why?

Computational effort --> MIN
computational complexity

goal-directed metrics
'aspects by which we evaluate models'

supervised learning

self-supervised learning

unsupervised learning

use-cases

theory-based models

representation structures

how well would ANNs do in terms of AIC, BIC?t

ANNs et al

(Hidden) Markov models

data-based models

E-Machine
'smallest possible maximally predictive model of a stochastic process'

random and/or observed fuel

Data
'observations'

Inference
'compressing'
'generalizing'
'learning'

Model

Deduction
'generating'
'realizing'
'predicting'

Data
'generations'
'realizations'
'predictions'

uncertainty estimate

The Tension between Engineering & Science Perspectives

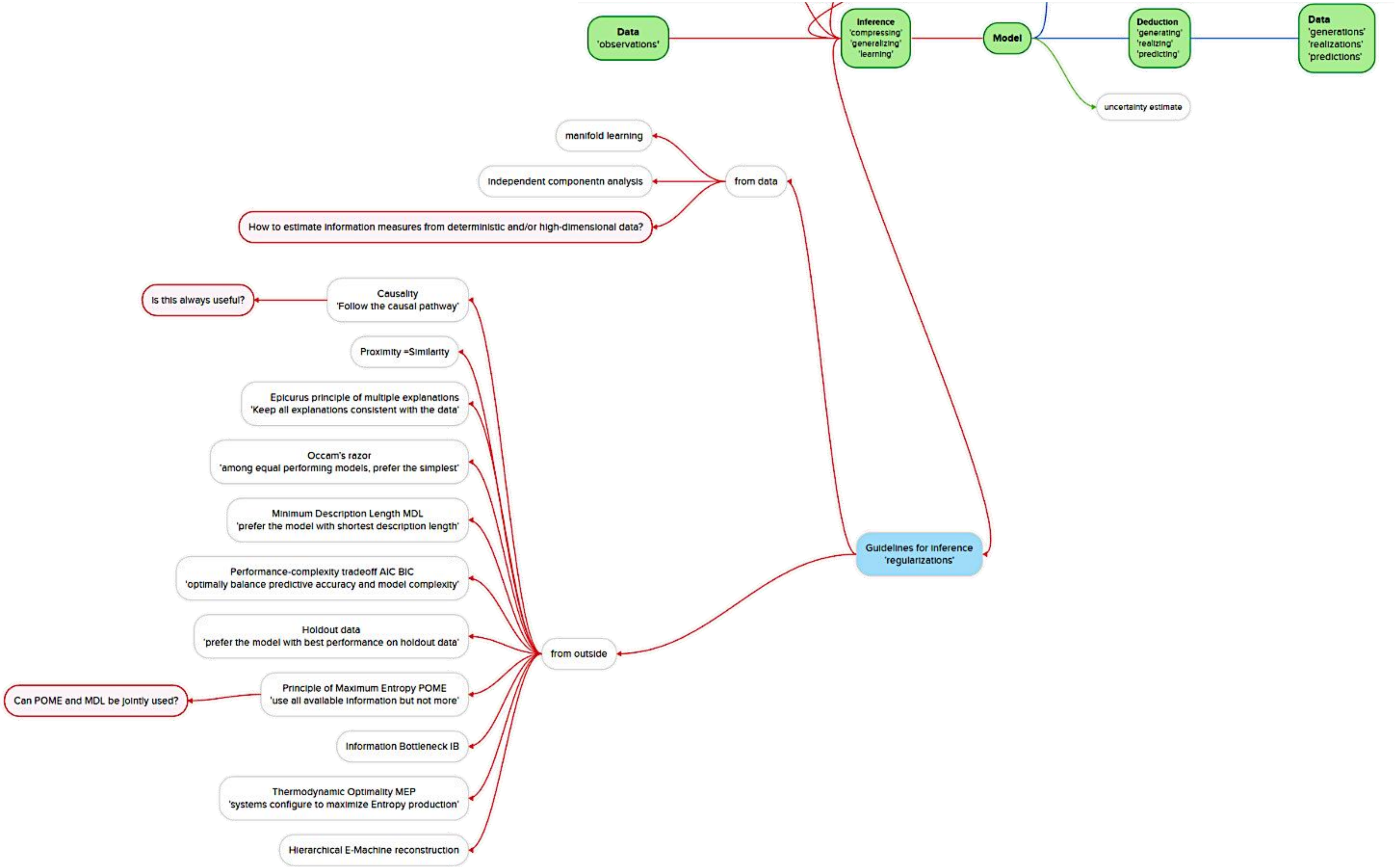
"The *Engineering* view of *Science* is that it is mere *Data Compression …* [but] *Scientists* seem to be *motivated by more than this*."

ENGINEERING: If a *representation is task-effective*, ***the engineer does not*** [necessarily] ***care what it implies about the underlying mechanisms*** *(although certainly concerned with minimizing implementation cost … representation size, compute time, storage, etc.).*

SCIENCE: ***To the scientist the implication makes all the difference in the world*** – the scientist presumes to be focused on *what the model means vis á vis natural laws*.

JP Crutchfield (1991) "Reconstructing language hierarchies," in Information Dynamics (HA Atmanspracher and H Scheingraber, eds.), (New York), p. 45, Plenum.

3. The Operational Pathway (Data → Inference → Deduction)

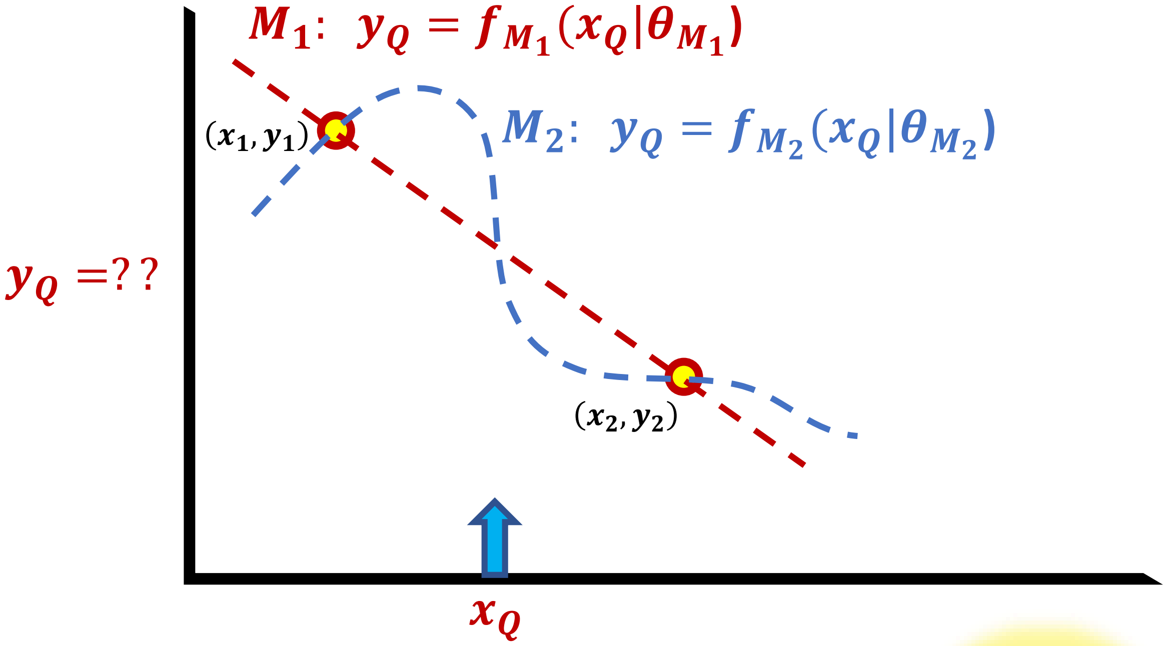


Data 'observations'
Inference 'compressing' 'generalizing' 'learning'
Model
Deduction 'generating' 'realizing' 'predicting'
Data 'generations' 'realizations' 'predictions'
uncertainty estimate
manifold learning
independent componentn analysis
from data
How to estimate information measures from deterministic and/or high-dimensional data?
Is this always useful?
Causality 'Follow the causal pathway'
Proximity =Similarity
Epicurus principle of multiple explanations 'Keep all explanations consistent with the data'
Occam's razor 'among equal performing models, prefer the simplest'
Minimum Description Length MDL 'prefer the model with shortest description length'
Performance-complexity tradeoff AIC BIC 'optimally balance predictive accuracy and model complexity'
Holdout data 'prefer the model with best performance on holdout data'
from outside
Guidelines for inference 'regularizations'
Can POME and MDL be jointly used?
Principle of Maximum Entropy POME 'use all available information but not more'
Information Bottleneck IB
Thermodynamic Optimality MEP 'systems configure to maximize Entropy production'
Hierarchical E-Machine reconstruction

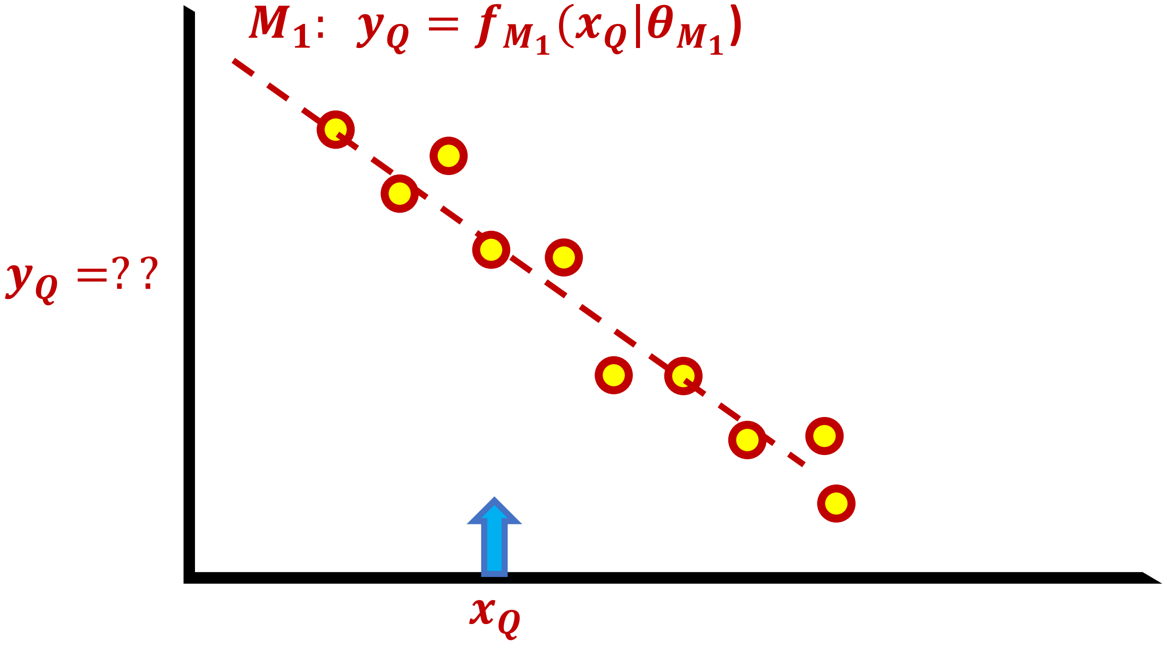
The Epistemological Problem of Inference

"Have we *discovered something in our Data …*

… or *<u>have we projected the new-found structure</u>* onto it?"



M_1: $y_Q = f_{M_1}(x_Q|\theta_{M_1})$
M_2: $y_Q = f_{M_2}(x_Q|\theta_{M_2})$
(x_1, y_1)
(x_2, y_2)
$y_Q = ??$
x_Q



$$\mathbf{I}(y_Q|x_Q, Model, Data) \neq \mathbf{I}(y_Q|x_Q, Data) + \mathbf{I}(y_Q|x_Q, Model)$$

The Epistemological Problem of Inference

"Have we *discovered something in our Data …*
… or *have we projected the new-found structure* onto it?"

This was the main lesson of attempting to reconstruct equations of motion from a time series:

When it works, it works; when it doesn't, you don't know what to do; and in both cases it is ambiguous what you have learned.

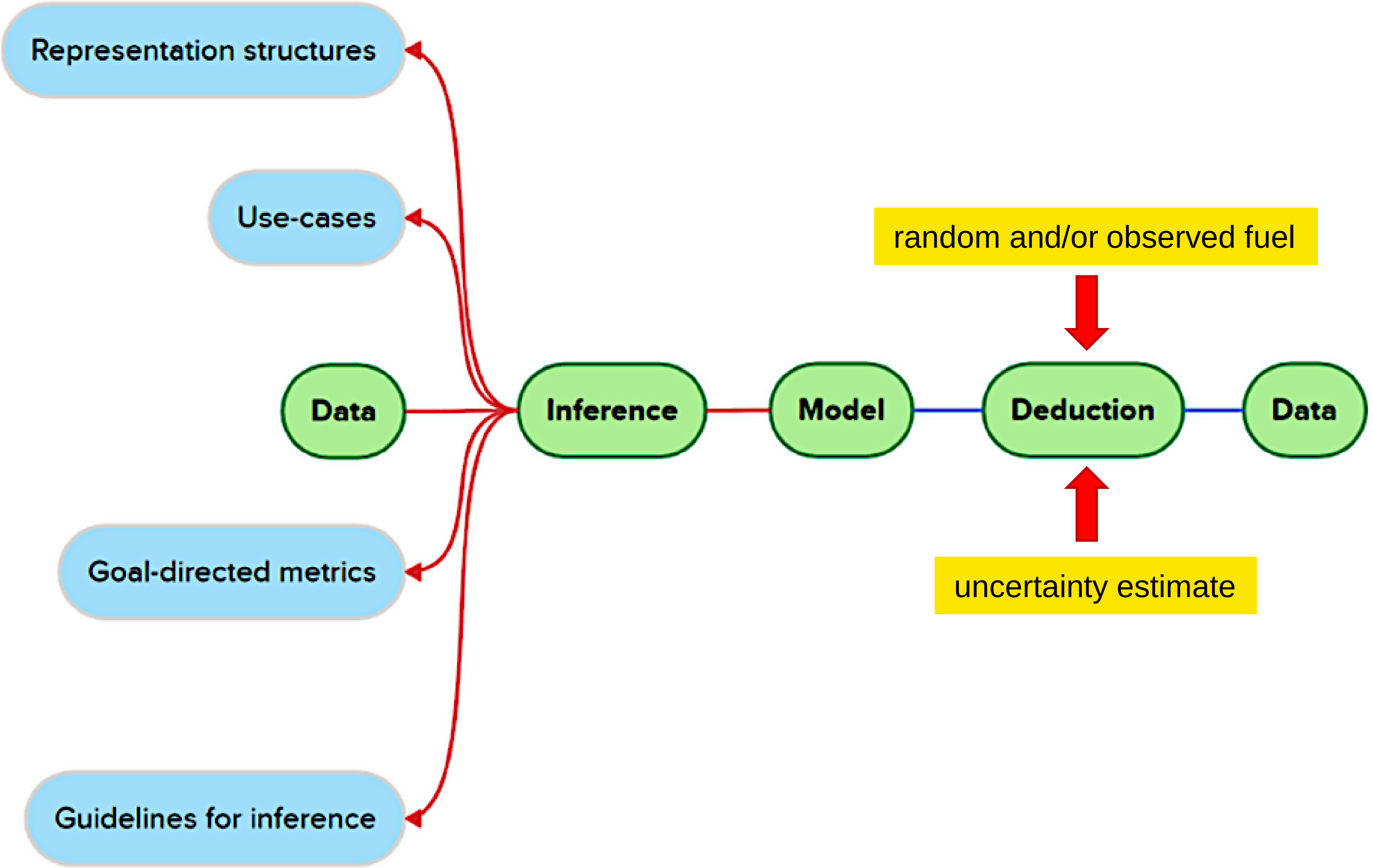
Even though data was generated by well-behaved, smooth dynamical systems, there was an ***extreme sensitivity to the assumed model class*** that completely swamped "model order estimation".

Worse still there was ***no a priori way to select the class appropriate to the process*** (*In AI this is referred to as the "**representation problem**"*.

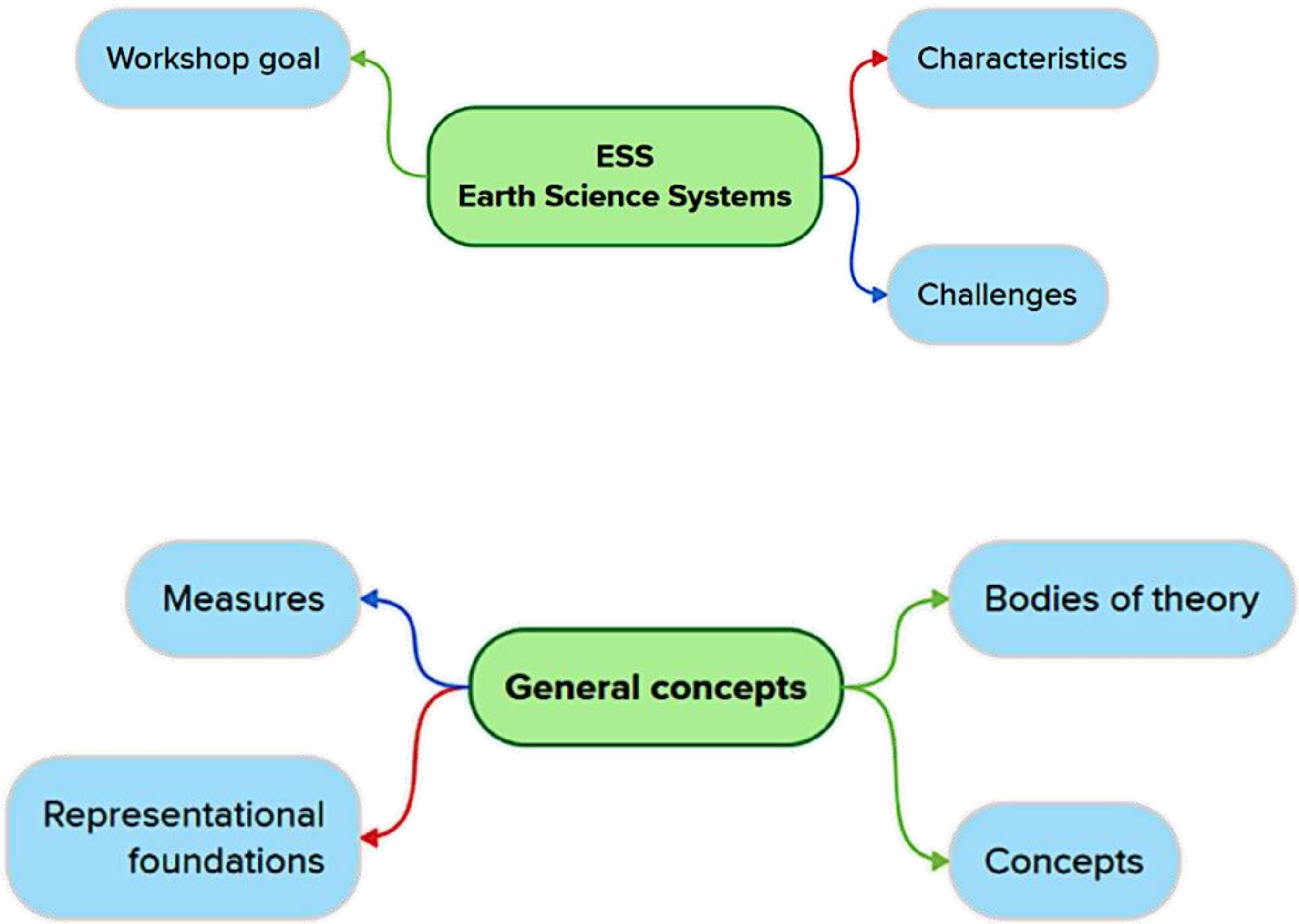
Despite representations to the contrary, … ***"model order estimation" does not address issues of class inappropriateness and what to do when confronted with failure***.

JP Crutchfield and BS McNamara (1987) Equations of motion from a data series," Complex Systems, vol. 1.

3. The Operational Pathway (Data → Inference → Deduction)



Please Contribute to the Mural



ESS
Earth Science Systems

Workshop goal

Characteristics

Challenges

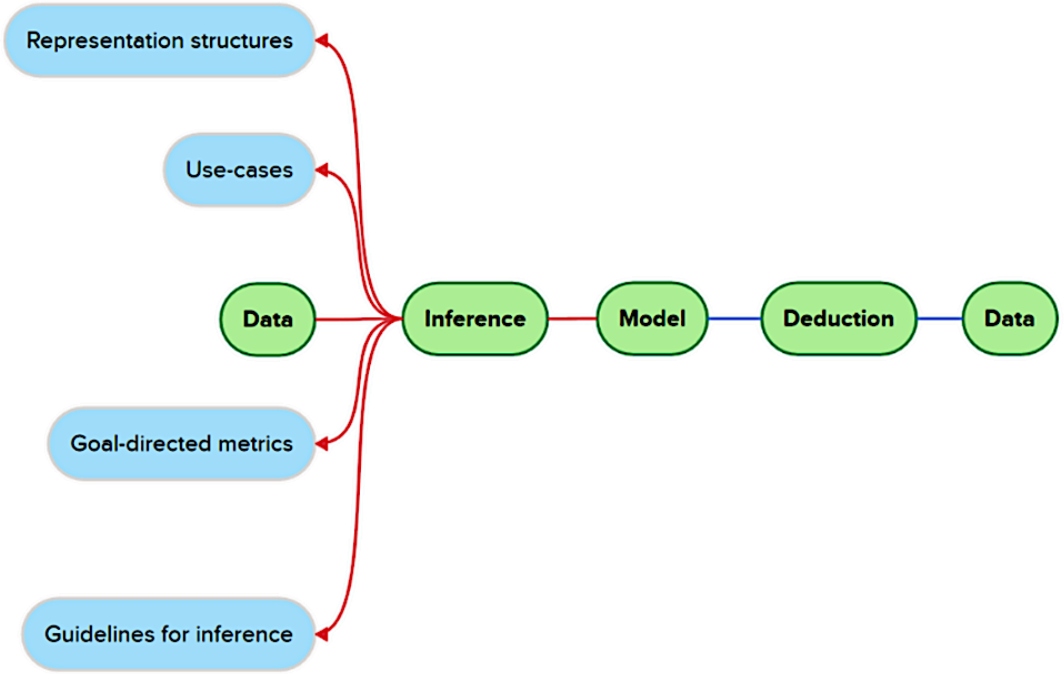
General concepts

Measures

Representational foundations

Bodies of theory

Concepts



Data

Inference

Model

Deduction

Data

Representation structures

Use-cases

Goal-directed metrics

Guidelines for inference

Schedule – Day One

Topical block I: Information Theory as a bridge

9:30-10:30 Invited talk 1 (45+15) A. Boyd
"Thermodynamic Overfitting: Limits on Complexity in Thermodynamic Learning"

10:30-11:30 Invited talk 2 (45+15) A. Jurgens
"Epsilon Machines and Randomness, Structure and Complexity:
Predicting Complex Systems"

11:30-11:45 --- Break --- (15)

11:45-12:30 Plenum discussion (45)

12:30-14:00 --- Lunch --- (90)

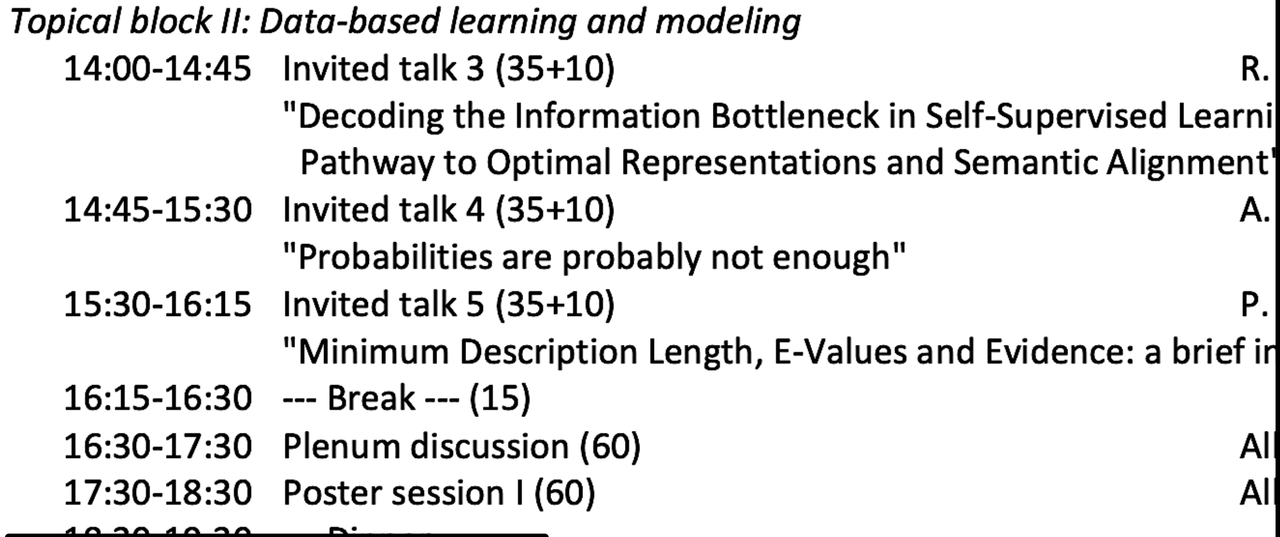
Alexandra Jurgens

Epsilon Machines.
Intrinsic Randomness.
Structural Complexity.
Finite-state generators.
Statistical complexity dimension.
Minimal memory resources.

Alec Boyd

Max-Work corresponds to Thermodynamic Learning.
Requisite Complexity.
Information Engines.
Predictive Hidden Markov Models.
Overfitting.

Schedule – Day One



Topical block II: Data-based learning and modeling

14:00-14:45 Invited talk 3 (35+10) R.
"Decoding the Information Bottleneck in Self-Supervised Learni
Pathway to Optimal Representations and Semantic Alignment"

14:45-15:30 Invited talk 4 (35+10) A.
"Probabilities are probably not enough"

15:30-16:15 Invited talk 5 (35+10) P.
"Minimum Description Length, E-Values and Evidence: a brief in

16:15-16:30 --- Break --- (15)

16:30-17:30 Plenum discussion (60) Al

17:30-18:30 Poster session I (60) Al

18:30-19:30 --- Dinner

Peter Grünwald

Minimum Description Length
E-Values and Evidence.
Compression.
Optimal codes.
Combining data from different sources.

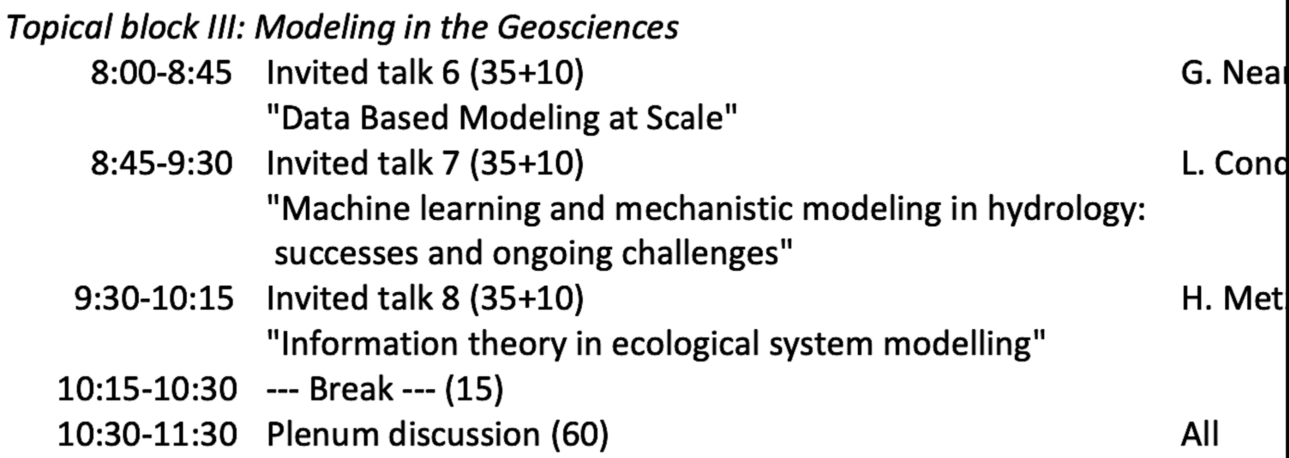
Ravid Schwartz-Ziv

Compression.
Self-Supervised IT Learning.
Information Bottleneck.
Minimal IT Statistics.
Information Plane.

Andreas Scheidegger

Good Predictions – versus --
Good Decisions.
Natural Systems.
Causal Inference.

Schedule – Day Two



Topical block III: Modeling in the Geosciences

8:00-8:45	Invited talk 6 (35+10) "Data Based Modeling at Scale"	G. Near
8:45-9:30	Invited talk 7 (35+10) "Machine learning and mechanistic modeling in hydrology: successes and ongoing challenges"	L. Cond
9:30-10:15	Invited talk 8 (35+10) "Information theory in ecological system modelling"	H. Met
10:15-10:30	--- Break --- (15)	
10:30-11:30	Plenum discussion (60)	All

Holger Metzler

Ecological System Modelling.
Conserving (compartmental) systems.
Complexity measures for Dissipative systems.
Maximum Entropy principle.

Grey Nearing

Data Based Geoscientific Modeling at Scale.
Challenges in operational ML.
Relationship between academia and industry.

Laura Condon

Hydrological Systems.
Bridging ML & Physical Modeling.
ML for Model Emulation.
Accelerating simulation-based inference.

Schedule – Day Two

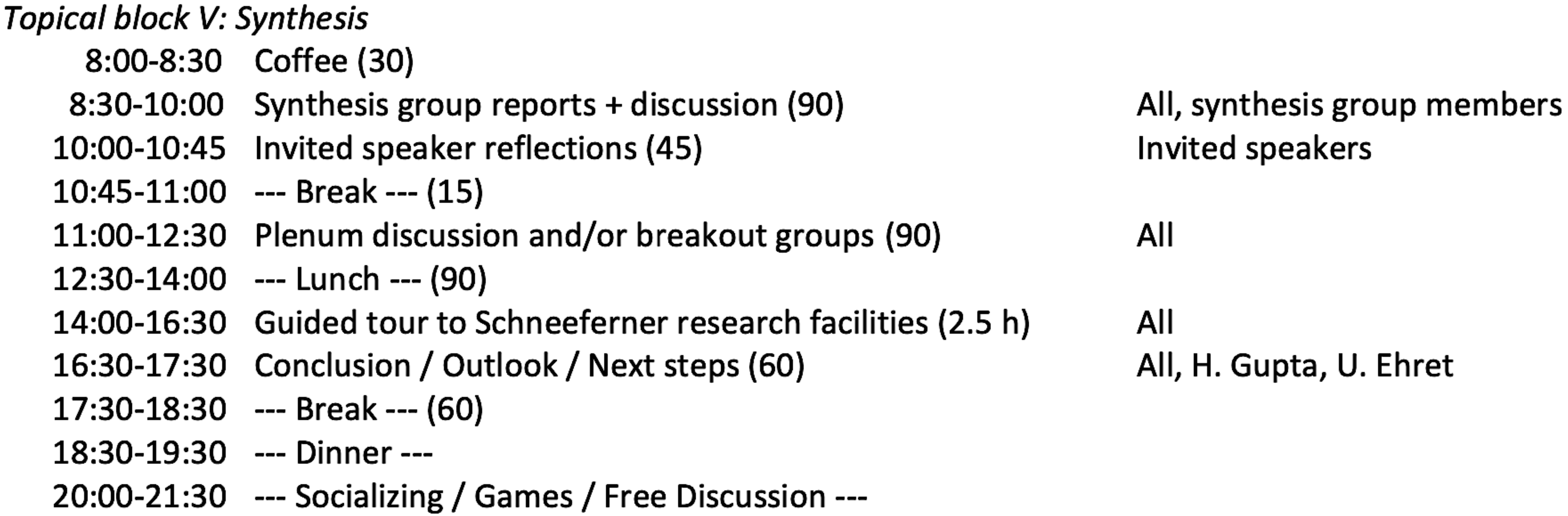
Topical block III: Modeling in the Geosciences

8:00-8:45	Invited talk 6 (35+10) "Data Based Modeling at Scale"	G. Nearing
8:45-9:30	Invited talk 7 (35+10) "Machine learning and mechanistic modeling in hydrology: successes and ongoing challenges"	L. Condon
9:30-10:15	Invited talk 8 (35+10) "Information theory in ecological system modelling"	H. Metzler
10:15-10:30	--- Break --- (15)	
10:30-11:30	Plenum discussion (60)	All

Topical block IV: Ways forward

11:30-12:30	Promises and challenges revisited (60) (identify forward-looking ideas for breakout groups)	All, M. Bassiouni, M. Höge
12:30-14:00	--- Lunch --- (90)	
14:00-15:30	Breakout groups	All
15:30-16:30	Poster session II (60)	All
16:30-18:00	Breakout group reports + discussion (90)	All
18:00-18:30	--- Break --- (30)	
18:30-19:30	--- Dinner ---	
20:00-21:30	--- Socializing / Games / Free Discussion ---	

Schedule – Day Three



Topical block V: Synthesis

8:00-8:30	Coffee (30)	
8:30-10:00	Synthesis group reports + discussion (90)	All, synthesis group members
10:00-10:45	Invited speaker reflections (45)	Invited speakers
10:45-11:00	--- Break --- (15)	
11:00-12:30	Plenum discussion and/or breakout groups (90)	All
12:30-14:00	--- Lunch --- (90)	
14:00-16:30	Guided tour to Schneeferner research facilities (2.5 h)	All
16:30-17:30	Conclusion / Outlook / Next steps (60)	All, H. Gupta, U. Ehret
17:30-18:30	--- Break --- (60)	
18:30-19:30	--- Dinner ---	
20:00-21:30	--- Socializing / Games / Free Discussion ---	

Potential Outcomes

- Enhanced *Dialogue* and *Collaboration*
- Community Progress towards *Development of a General Framework* (*rooted in the marriage of Information Theory, Modeling Science, and Domain Science*) for:
  - *Constructing Task-Relevant Models that can Learn from Data*
  - *While maintaining Interpretable Representational Structures consistent with physical understanding*
- Joint/Collaborative *Papers* (or series)
- Other …

We Invite Creative Suggestions

Discussion

100011101001111111000000101111001110010101000001111111111000000011010101010010