International Workshop, The Schneefernerhaus, Zugspitze, Germany, Sept 11-13, 2023

Information Theory as a Bridge Across the Geosciences & Modeling Sciences

Help Desk

Ask workshop related questions

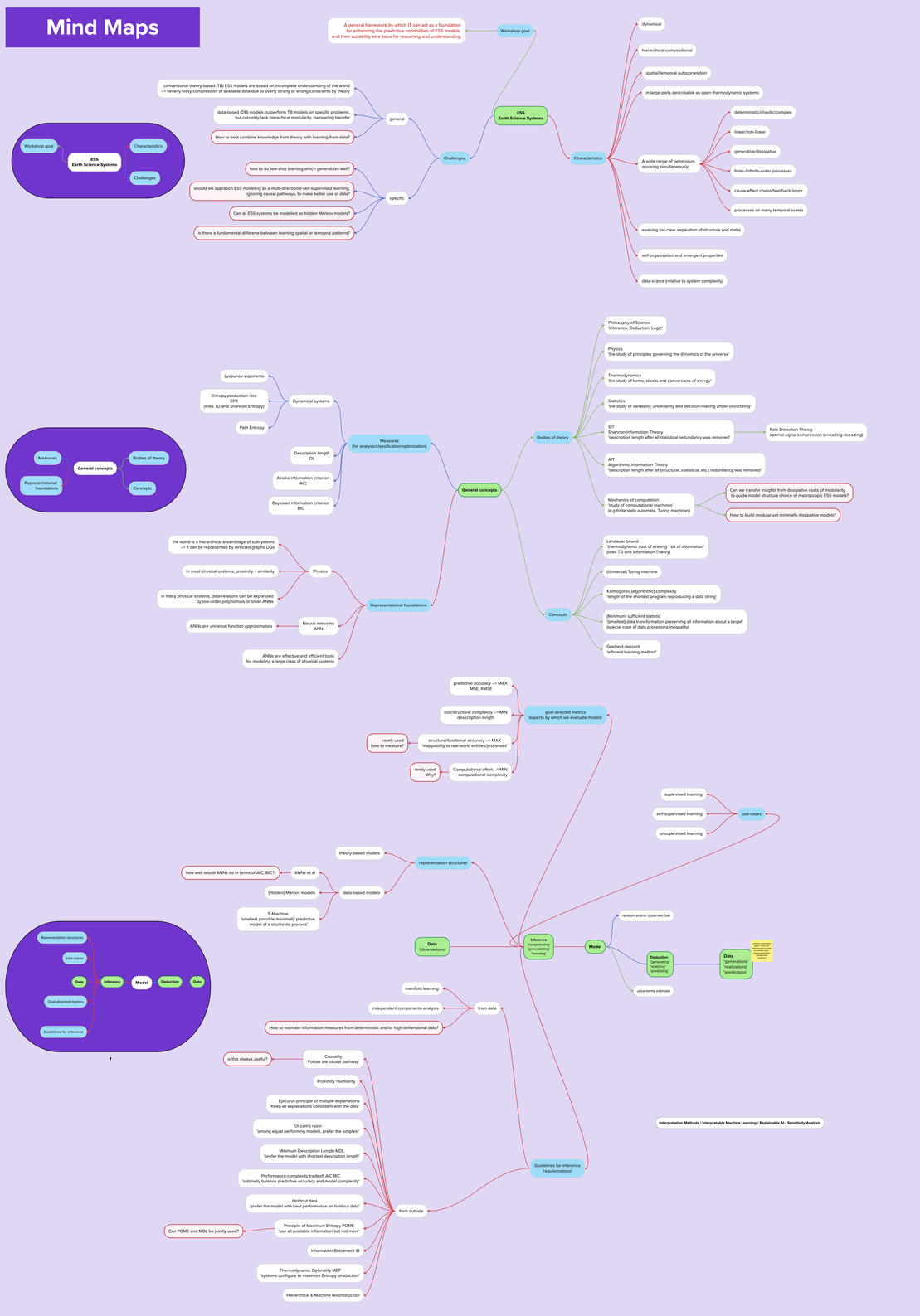
Answers can benefit all

Stationary

Copy/paste sticky notes to add your comments or icons to emphasize existing notes

Type your paragraph...

Mind Maps



0. Pre-Workshop

0.0 WELCOME !

We will use this space to share and organize resources, questions and insights throughout the workshop

Mural How To

Logistical

Schedule — Logistics — Participants — Hiking Trip

Learning Resources

- July Theory in the Geosciences Journal Club
- **YouTube Archive**
- A Brief History of Information Theory — **Ehsan Arabzadeh Talk Summary**
- Introduction to Information Theory Crash Course — **Recording**
- Introduction to Information Theory Crash Course — **Slides**

Access all Resources

Schneefernerhaus_IT_workshop_2023 - Google Drive

Contact organizers:
uwe.ehret@kit.edu, hoshin@arizona.edu

0.1 Participants

Tell us who you are by positioning yourself in or between the workshop disciplines and tag yourself with one or multiple domains of interests/expertise using examples from the list below.

Information Theory
IT - Information Theory
AIT - Algorithmic Info Theory
RIT - Representational Info Theory
IB - Information Bottleneck

Modelling Science
CS - Computer Science
ODE - Ordinary Differential Equations
ML - Machine Learning / Deep Learning
STAT - Probability and Statistics
CD - Causal Discovery
PB - Physically-Based Modeling

Domain Science
ESS - Earth Science
HYDRO - Hydrology/Hydrometeorology
HYD - Hydrological Science
ECO - Ecological Science
GEO - Geology

0.2 Preliminary Reading

Read literature suggested by our invited speakers and organizing committee
Contribute your impressions and questions in the discussion notes

Additional reading suggestions from participants?

0.3 Preliminary Reflections

Share your reflections about the greatest opportunities and obstacles in different domains with respect to the overarching workshop goal

or/and add an icon to emphasize existing notes

(please, sign your notes)

Overarching Workshop Goal

Development of a general framework by which Information Theory can act as a foundation for enhancing the predictive capabilities of Geoscientific Models, and their suitability as a basis for Reasoning and Understanding

Opportunities & advances

Obstacles & challenges

Information Theory — Modelling Sciences — Domain Sciences — Bridge

1. Invited Talks and Plenaries (Days 1 & 2)

Information Theory as a bridge (Day 1)

- **Hoshin Gupta and Uwe Ehret** Introductory talk — Recording
- **Alec Boyd** Thermodynamics Overfitting: Limits on Complexity in Thermodynamic Learning — Recording
- **Alexandra Jurgens** Information Theory, Relativeness, Structure and Dynamics: Predicting Complex Systems — Recording

Data-Based Learning and Modeling (Day 1)

- **Ravi Shwartz-Ziv** Information Bottleneck in Self-Supervised Representation Learning: Optimal Representations and Semantic Alignment — Recording
- **Andrew Schwiegler** Probabilities are probably not enough — Recording
- **Peter Grünwald** Minimum Description Length, E-Values and Evidence: a brief introduction — Recording

Modeling in the Geosciences (Day 2)

- **Grey Nearing** Data Based Modeling at Scale — Recording
- **Laura Condon** Groundwater-surface water mechanistic modeling in hydrology: successes and ongoing challenges — Recording
- **Holger Metzler** Information theory in ecological system modelling — Recording

2. Posters (Days 1 & 2)

Deposit and view pdfs of your posters on google drive

Please use a sticky note to contribute to the discussion. Use colors, title and emojis to make it easy to follow the color-coded discussion threads.

- **Manuel Bravo Chiara, Uwe Ehret, Hoshin Gupta & Arnold Gutiérrez** — Toward an Information Theoretic Diagnostic Evaluation Framework for Hybrid Models
- **Eduardo Acuna, Rolf Lutz, Marcel Geneiro, Nicola Bulaevskaya and Uwe Ehret,** Information theory and Markov chains: a perfect marriage for modelling?
- **Cristóbal Sandoval, Julio Biscarrisa, Noelia Otero, Guillermo Gallego, Ana Ramos, Esteban Hernández** — Bridging the long memory Maximum Caliber and Information Theory approaches
- **Rolf Lutz and Uwe Ehret** We — Learnable Learned Tuple Tracking in Discrete Dynamical Systems: Preliminary Results
- **Da Gerundo & Nicola Otero** — Minimizing Network Optimization
- **Timothy Hodson, Thomas Omey, Tyler Smith, and Lucy Marshall** — Test your objectives per formance metrics: Loss functions for data assimilation using information theory
- **Uwe Ehret and Perdrig Dey** — A scheme, & method to analyse, classify and compare dynamical systems by uncertainty and complexity

3. Breakouts (Day 2)

Relative entropy
scale dependent concepts
surprise and scale length-being related to the concept of information
need of reference measures for it to be meaningful
multivariate definition of entropy
IT as a language for science

Information theory and other theory notes

- Heritage theories (related to Brian's slides) self-perception 101
- Different "information theories" (not just Shannon's)
  - IB (compress per relevance)
  - Kolmogorov complexity (compress per algorithm)
  - Algorithmic information theory (compress per Turing machine)
  - Information theory of complexity (compress per scale-by-scale lens)
  - Information theory of causality (compress per intervention)
  - Information theory of the optimal (compress per utility)
  - Complexity is a context-dependent concept

Data Issue
- 14% of Users in Data Rich Sciences have access to no scientists
- Data Sharing Incentives Misaligned: data share stats across physical disciplines

Foundational modeling: what is could-look like and what are the steps to get from current state to that? Ideas: Use organizer gestalts similar to LLM (pretrained, fine-tunable) - Requirements: - LLM with ability to self-organize to predict and generate - reasoning, data, and modeling capacity

Breakout Topic Suggestions

- Use AI for new and large problems
- Advancing IT (and other) theory
- Theoretical vs. practical
- uncertainty
- foundation models
- Advancing AI theory
- data foresight

4. Speaker Reflections (Day 3)

Alec Boyd

Thermodynamics and physics tells us about the reason for certain types of learning

- Time is a directional channel - can stockpile information obtained in the past inform physical processes
- Thermodynamics of learning: how dynamical technique/strategies in ML feel as physically/may systems (overfitting/charging)

connecting Thermodynamics + IT and ML

Thermodynamics + IT might help to understand why training etc.works so well (stoch gradient descent...)

Physicals justifications

Grey: information flow through ML models

explore links between AI and IT (see breakout groups)

opportunities for large-scale models, starting from small-scale models, generating synthetic datasets to begin with and make proof-of-concept

Laura: investigate emergent behaviour of systems, using diverse datasets to assess what do we have and what is lacking

Andrew Schwiegler

- Certain assumptions for choosing right model/ loss function
- Use theory to tell us what is the right loss/eval metric, and what is the right model class
- ITA and and and efficient approaches now to inform IT advances by practical needs
- Using ITA tools and IT-like models/ approaches for integrating LLMs
- IT as a unifying lens into ITS

Andrew: parallels that between multiple objectives - IT is navigator on IT

- entropy as loss property to distributions
- Entropy is a loss function, but not always (ex IB), but may be
- let's take sign out of models (ML are better theory)
- LLMs in similar as next: big thing that should be focused on pursuers Foundation modeling

Round table interests:

Ripping the role of information
physical phenomena and self similarity for foundational models
characterizing latent space in minimum representation of our system
Is information a fundamental property of physics?
info as a fundamental quantity in physics to observe phenomena
self organization and theory of information
Fundamental
IT to explain systems that are comprehensible to the human mind
how to go from theory to practical
Causality
categorize and identify data and making them talk to each other
proximity in space vs time
Information in systems and spatial phenomena
loss IT and AI
foundational models
Information relationship and resources to the macroscale
capturing scale emergent and evolutionary laws with ML through the lens of information as the driver of these patterns
Fundamental
IT for better communications in applications
IT for better foundational models
information theory beyond Shannon
concrete experiments for concrete applications
self organization and using information
self-organization
how is information expressed through representations
Information theory and its relationship to the world with models
making data talk through self organization

5. Synthesis Discussion (Day 3)

6. Next Avenues (Day 3

Concluding Session
- **Hoshin Gupta - Continue reading related literature from related domains (contact Lalo Hernandez Rodriguez** <hlhernandezrodriguez@kit.edu>**) - Info Theory Reading Group**
- **Invited speakers**
- **Uwe Ehret - Manuscript resulting from Workshop - target journal?** (HRR, Entropy Journal, Geoscientific Model Development (perspectives format), Jaines? Broader than hydrology journal, Nature Reviews, Nature Geoscience..)
- **Uwe Ehret - Workshop Report** (2 weeks?)
- **Papers/Initiatives**
- **Jonathan (@hamelyntee.com), Laura - Foundation Models for Geosciences - Workshop Proposal** (Jonathan to share draft)
- **Qingyun, Praveen, Laura, Lalo, Marcel) - Quality of Physics and Information**
- **Timothy, Juliana, Cris, Manuel, Lalo - A General Framework for Evaluation of ES Models**
- **Timothy, Daniel - Obj Fn Experiments with LSTM**
- **Marcys - Science Fiction Story / metaphors on Lens of IT and Modeling (Nature?)**
- **Qingyun, Lalo - Data Sharing Incentives**
- **Florian, Uwe - Complexity Analysis & Info in Spatial patterns**
- **Jonathan - Synchronization & Emergence in Geoscience Systems --> tied to FDIN workshop**
- **Hoshin, Florian - How info is encoded in/expressed through representational entities/ness**
- **Timothy - Broad critique of ES modeling practices**
- **Uwe - How to continue WORKSHOP(discussion) format & SUMMER SCHOOL**
- **Uwe - Workshop Report - Lessons Learned - Workshop Report - Lessons Learned** (design, conversation etc. **USA, CANADA (Waterloo Institute), CHILE, in person non-hybrid (not Synchronized-Connected on 2 continents)**
- **Uwe - Lessons Learned - Workshop Report - Lessons Learned** - Notes etc.
- **Lalo/Ben** - Update of publications list, proper subscription method
- **Praveen** - Out of the Box - write position papers for **funding** agencies to pick up on. Tap into the funding ecosystem and influence it towards
- **XXX** - Sessions at AGU, EGU etc (on your own)

7. Post-Workshop Reflections & Sharing

Random notes & thoughts

Main take-aways: opportunities to leverage...

Main take-aways: Challenges to solve...

I'm really excited about...

Meeting all of you

Being in the Alps

I'm surprised that...

I'm confused about ...

I'm inspired to ...

I want to share this cool link/resource/paper

Access all Resources

Schneefernerhaus_IT_workshop_2023 - Google Drive

Talk Recordings

Workshop Videos

View/share photos

Share your photos here :)