

Synthetic Geology: Structural Geology Meets Deep Learning

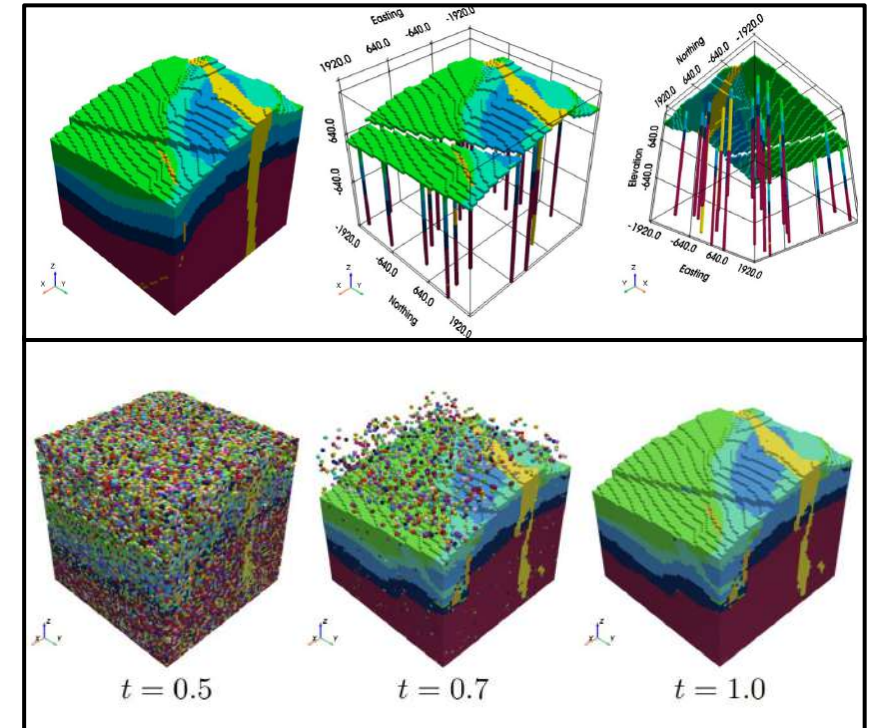
Presenter:
Simon Ghyselincks

Authors:
Simon Ghyselincks¹; V. Okhmak³, S. Zampini³,
G. Turkiyyah³, D. Keyes³, E. Haber².

¹ Dept. of Computer Science, UBC.

² Dept. of Earth, Ocean and Atmospheric Sciences, UBC.

³ CEMSE Division, KAUST.



THE UNIVERSITY OF BRITISH COLUMBIA



A Hidden Earth



Data Type

Boreholes and Direct Sampling

Strength

True direct observations of geology.

Weakness

Sparse, often incomplete picture, expensive to acquire



Potential Fields and Remote Sensing

Continuous coverage over large areas.

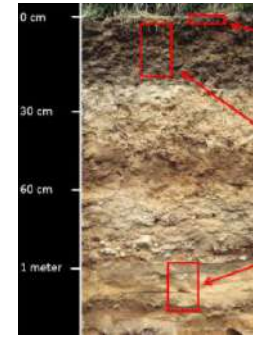
Ill-posed problem with many plausible solutions. Dimension reduction from 3D to 2D.



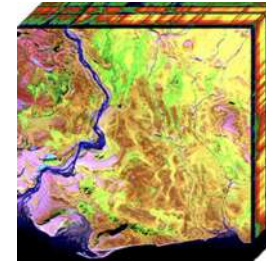
Structural Mapping

Lithological pooling of observations, measurable geometries.

Heavily biased toward surface outcrops, ill-conditioned depth continuation.



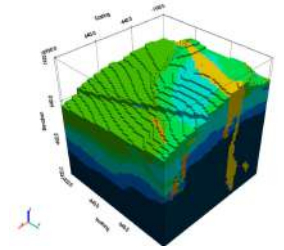
USGS Survey



Hyperspectral Data, NASA

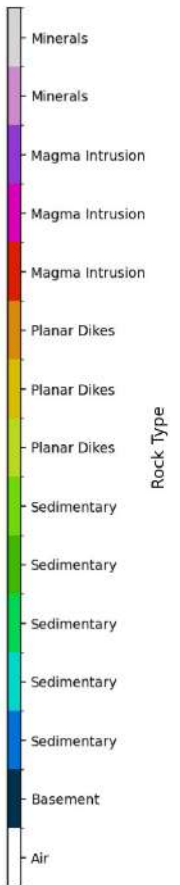
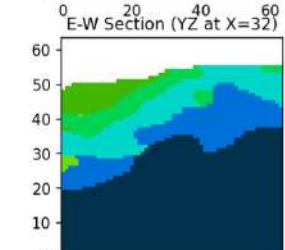
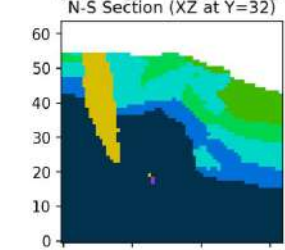
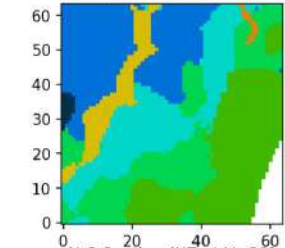


GTK Finland from Nironen et al. (2016)



Central Orthogonal Slices

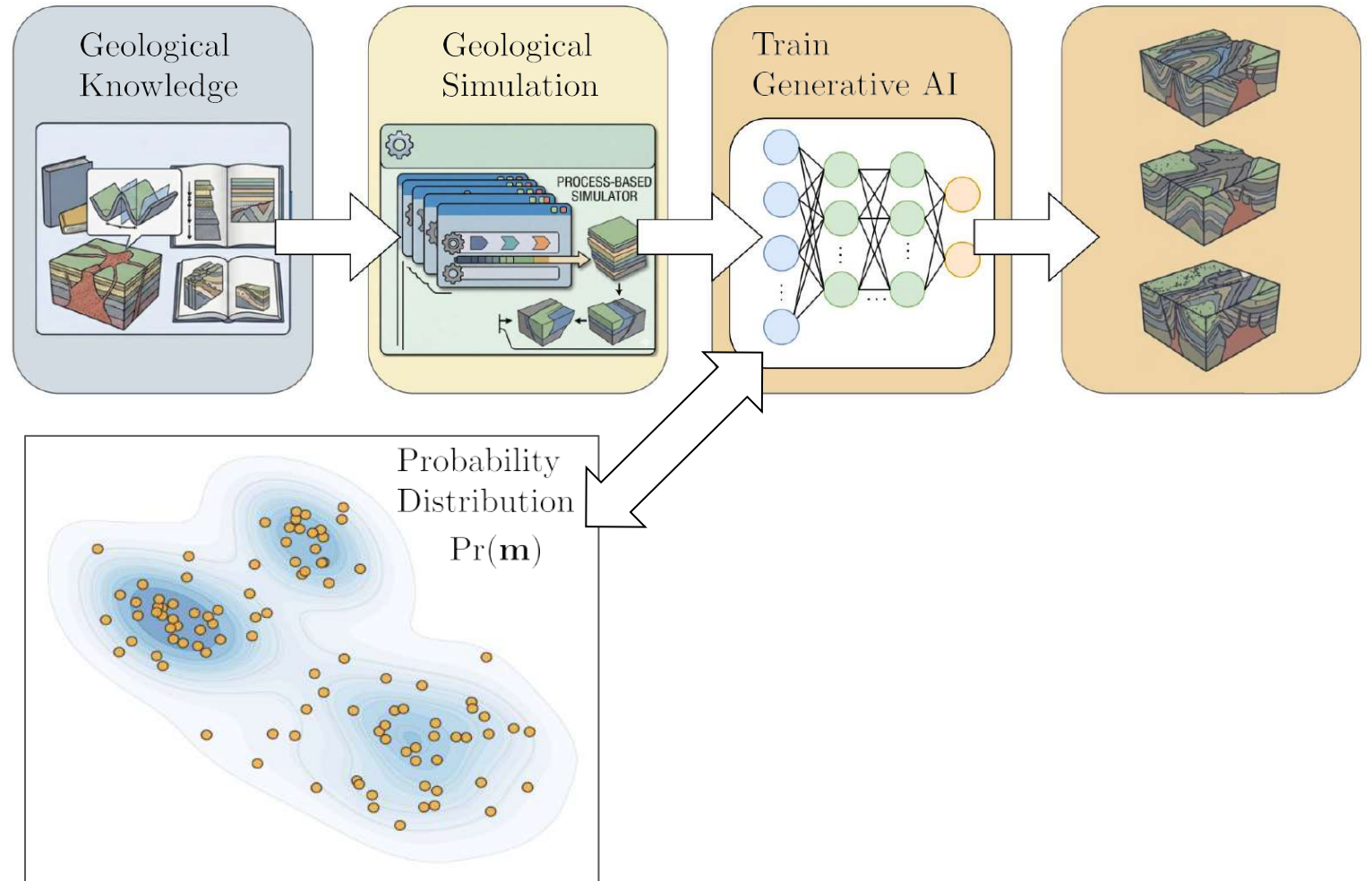
Map View (XY at Z=42)

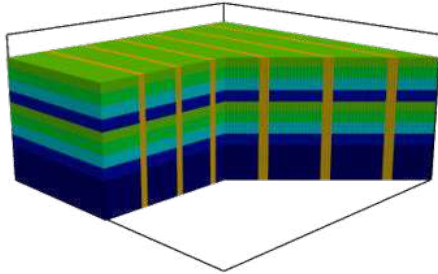


Rock Type

Geological Generative AI

- Find a way to encode abstract geological concepts
- Simulate the 3D lithology
- Train a probabilistic model
- Condition on forward observables (measured data)

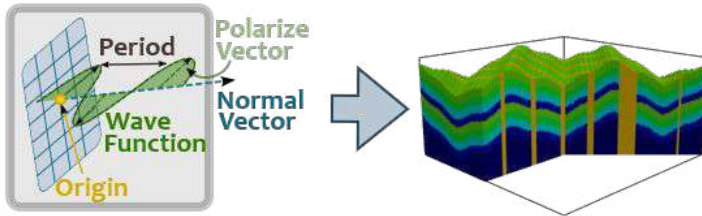




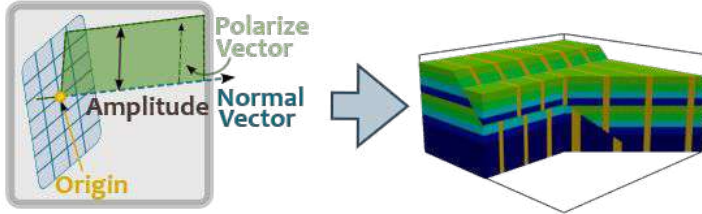
Process-Based Simulation

- Modular framework
- Fast compute (point cloud)

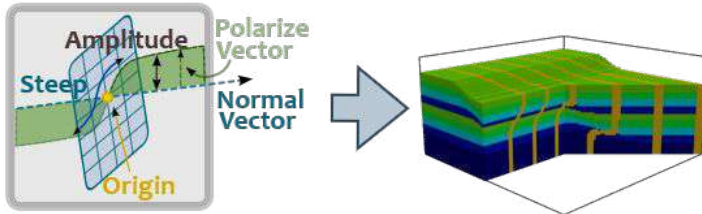
Fold



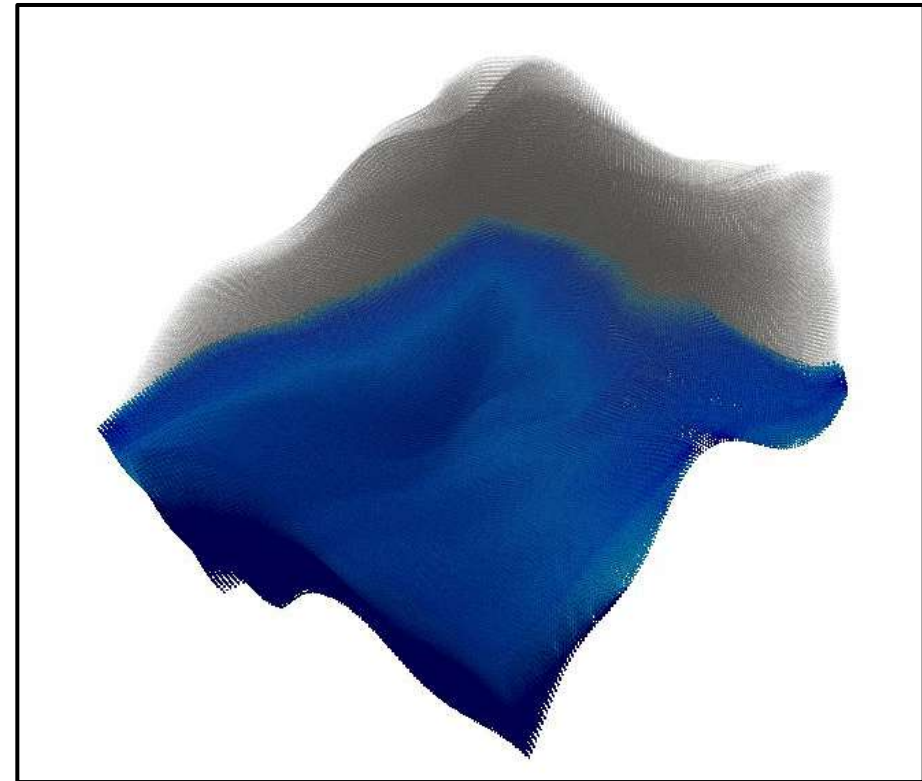
Fault



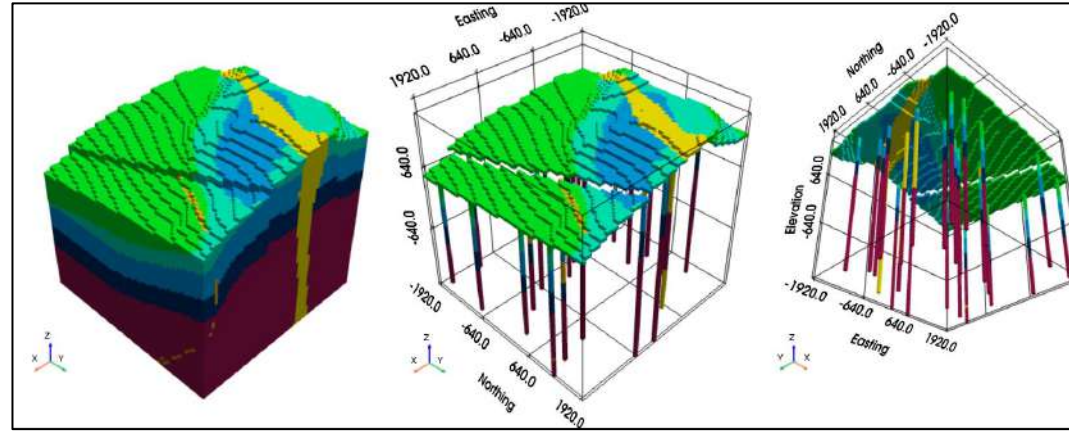
Shear



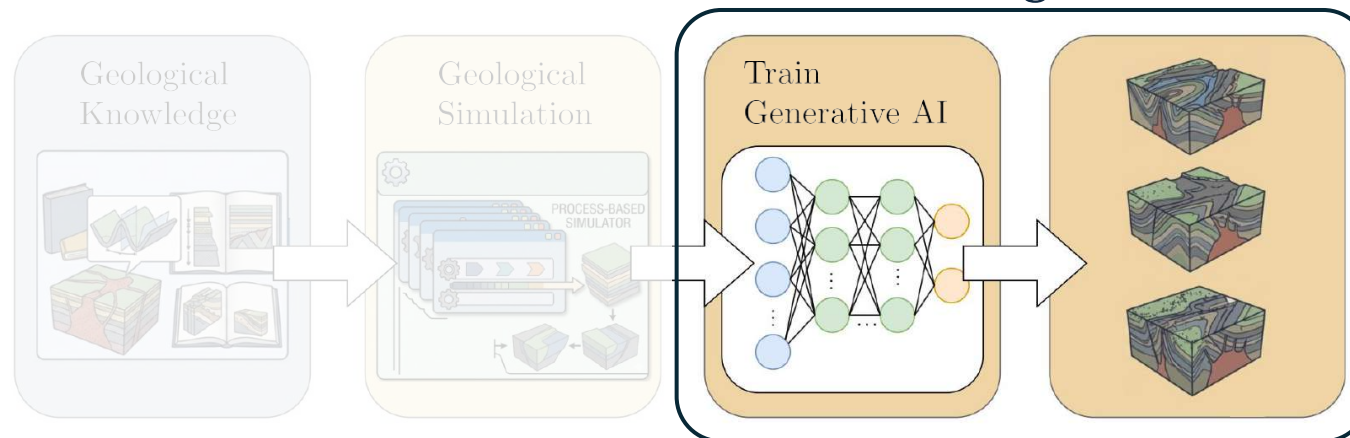
Example of tectonic geological events



Our Research: Generation



Part 2: Training a Generative AI



Flow Matching

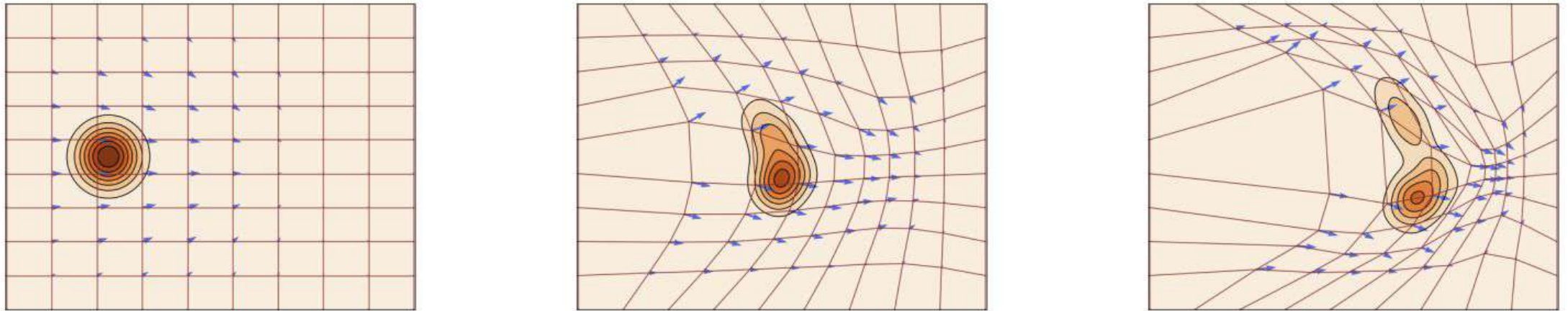
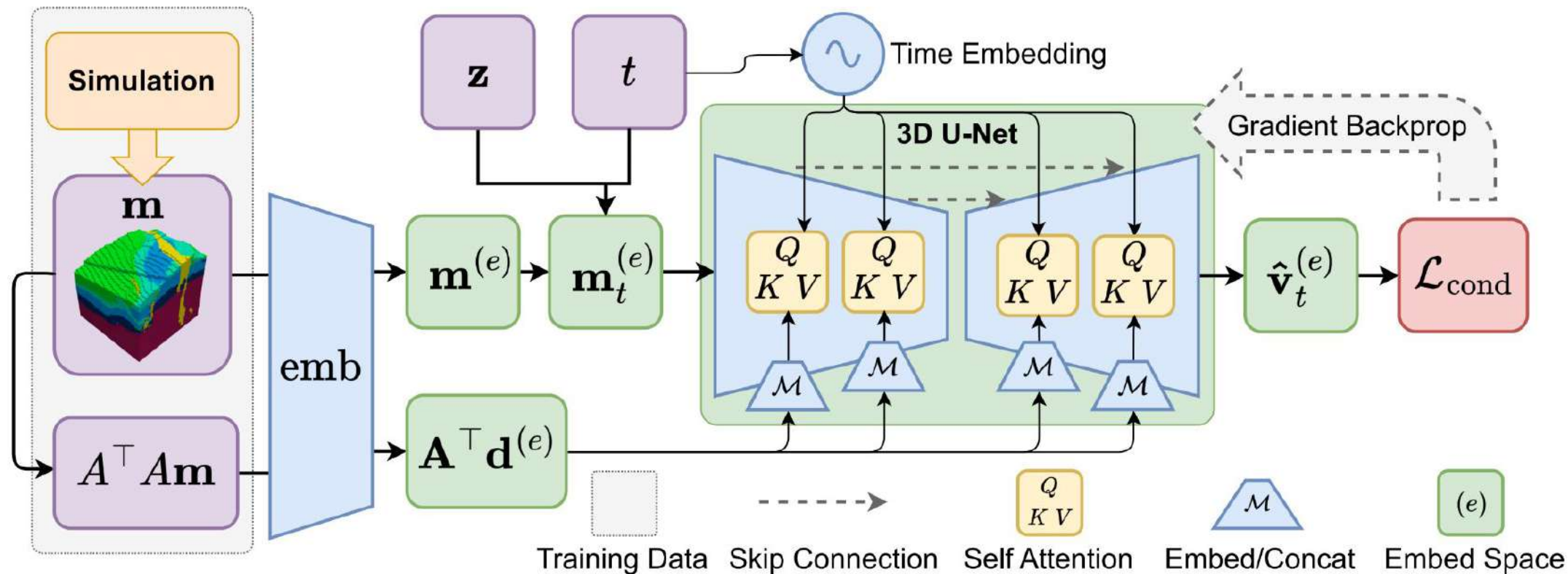


Figure 7 A velocity field u_t (in blue) *generates* a probability path p_t (PDFs shown as contours) if the flow defined by u_t (square grid) reshapes p (left) to p_t at all times $t \in [0, 1)$.

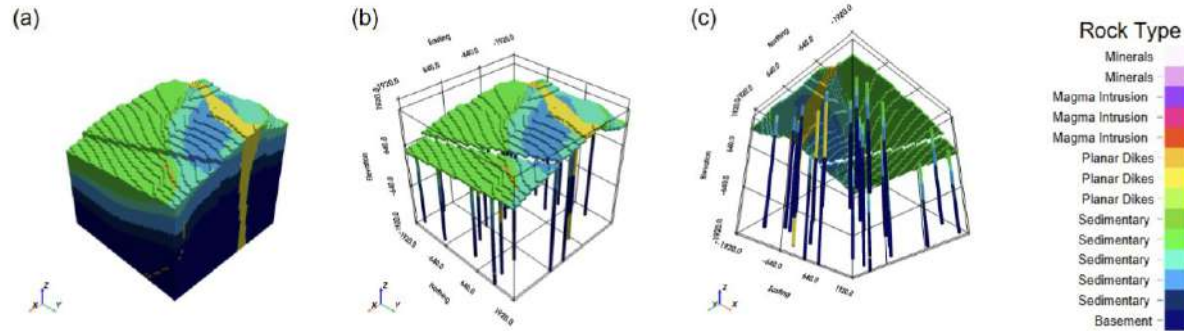
(Lipman et. al, 2024)

Incorporating Time and Conditional Data

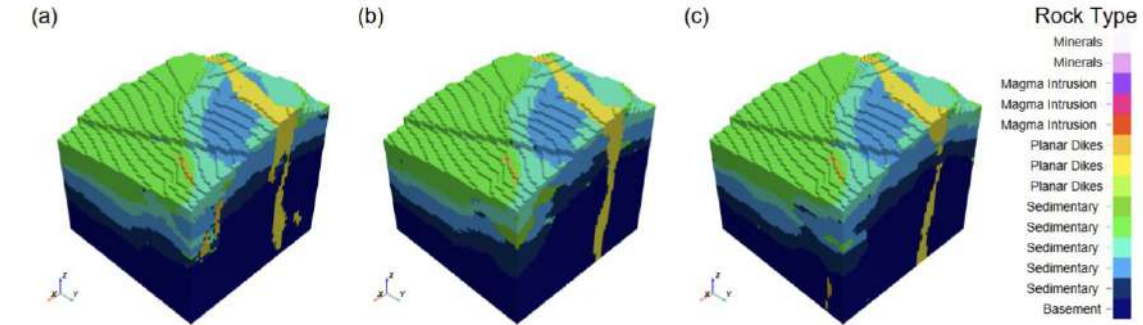


$$\mathcal{L}_{\text{cond}} = \|\hat{\mathbf{v}}(\mathbf{m}_t, \mathbf{d}, t; \xi) - \mathbf{v}_t\|_2^2 + \lambda \cdot t \cdot \|\hat{\mathbf{d}} - \mathbf{d}\|_2^2.$$

Results

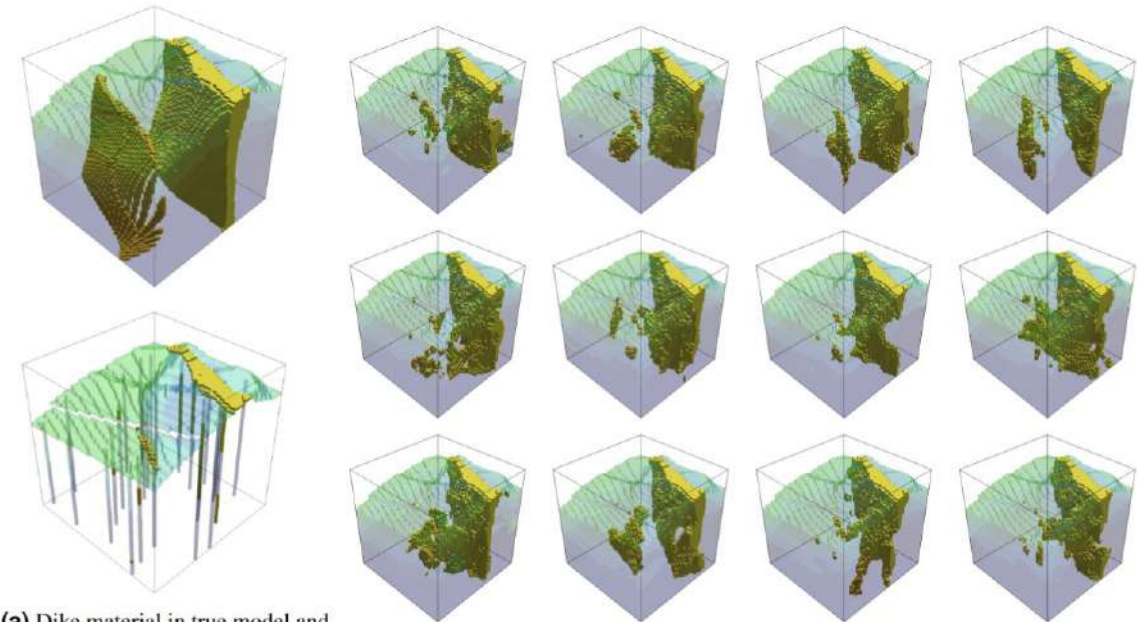


(a) Ground truth; (b) and (c) borehole and surface data



(a)-(c) Three conditional realizations

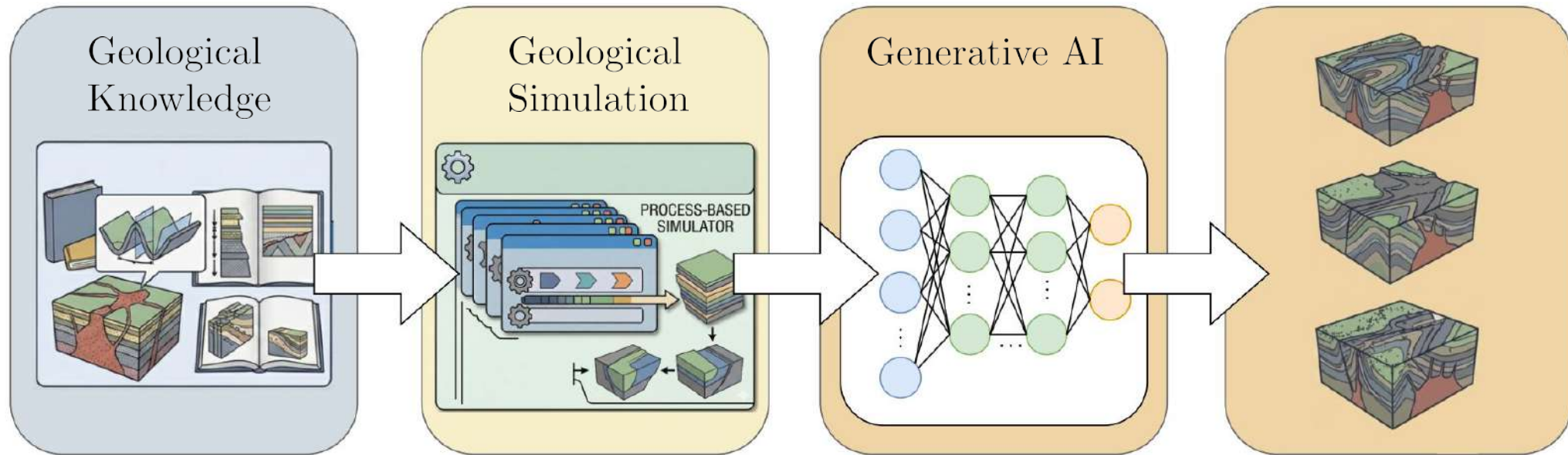
- Multiple reconstructions
- Ensemble analysis



(a) Dike material in true model and conditional data.

(b) 12 different dike realizations conditioned on surface and borehole data.

Human-centered AI: A probabilistic system that informs geologists of plausible scenarios, integrating expert interpretation with automated pipelines to better understand the Earth's subsurface.





References

Image by wirestock on Freepik

References

Albergo, M. S., Boffi, N. M., & Vanden-Eijnden, E. (2023). Stochastic Interpolants: A Unifying Framework for Flows and Diffusions. arXiv:2303.08797.

Bergen, K. J., Johnson, P. A., De Hoop, M. V., & Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. *Science*, 363(6433), eaau0323.

Jessell, M., Guo, J., Li, Y., Lindsay, M., Scalzo, R., Giraud, J., Pirot, G., Cripps, E., & Ogarko, V. (2022). Into the Noddyverse: A massive data store of 3D geological models for machine learning and inversion applications. *Earth System Science Data*, 14(1), 381–392.

Ruthotto, L., & Haber, E. (2021). An Introduction to Deep Generative Modeling. *GAMM-Mitteilungen*, 44(2), e202100008.

de la Varga, M., Schaaf, A., and Wellmann, F. (2019). GemPy 1.0: open-source stochastic geological modeling and inversion, *Geosci. Model Dev.*, 12, 1-32.

Lipman, Y., Havasi, M., Holderrieth, P., Shaul, N., Le, M., Karrer, B., Chen, R. T. Q., Lopez-Paz, D., Ben-Hamu, H., and Gat, I. (2024). Flow Matching Guide and Code, arXiv preprint arXiv:2412.06264.

Synthetic Geology for Deep Learning

A probabilistic system that informs geologists of plausible scenarios, integrating expert interpretation with automated pipelines to better understand the Earth's subsurface.

