# X2Sim: Rapid digital twin creation from text and videos for natural hazard modeling

Jonathan Gaucin<sup>1</sup>, Cristian Moran<sup>2</sup>, Cheng-Hsi Hsiao<sup>3</sup>, Justin Bonus Ph.D.<sup>4</sup>, Krishna Kumar Ph.D.<sup>3</sup>

<sup>1</sup>University of Houston, <sup>2</sup>Texas A&M University-San Antonio, <sup>3</sup>University of Texas at Austin, <sup>4</sup>University of California, Berkeley

X2Sim transforms real-world footage into precise digital twins for natural hazard modeling. Our platform combines advanced LLM processing and multi-modal simulation, providing a text-and-video-to-simulation workflow for researchers and practitioners in disaster



Scan QR code to









# preparedness.

#### access website

# Introduction

- Traditional natural hazard modeling methods are computationally intensive and time-consuming (Figure 1).
- Our X2Sim framework utilizes an agentic LLM to generate digital twins from video using natural language, integrating two distinct 3D object generation methods.



**Fig. 1.** TACC Galaxy and MPM rendering of Oso Landslide [1].

### Methods

• An Large Language Model (LLM)



#### Results



- agent generates a 3D object, fluid geometry and simulation from a given prompt (Figure 2).
- Text-to-3D objects are generated through DALL-E [2] image generation and TRELLIS [3] 3D asset generation.
- Video-to-3D object generation is realized through COLMAP [4] (Figure 3).
- Simulation is realized through the Material Point Method (MPM) or the learned Graph Network-based Simulator (GNS) [5].

**Fig. 2.** X2Sim framework: LLM agent with text- and video-to-3D tools.



**Fig. 6.** Rendering of a text-to-simulation output of a house impacted by a tsunami generated using MPM.

## References

[1] Minority Report: A Graph Network Oracle for In Situ Visualization. Scientific Figure on ResearchGate.
[2] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, and M. Chen, "Hierarchical text-conditional image generation with CLIP latents," arXiv preprint, Apr. 13, 2022. Available: https://doi.org/10.48550/arXiv.2204.06125.
[3] J. Xiang, Z. Lv, S. Xu, Y. Deng, R. Wang, B. Zhang, D. Chen, X. Tong, and J. Yang, "Structured 3D latents for scalable and versatile 3D generation," arXiv preprint, Dec. 2, 2024. Available: https://doi.org/10.48550/arXiv.2412.01506.
[4] Schönberger and J.-M. Frahm, "Structure-from-Motion Revisited." Available: <u>https://demuc.de/papers/schoenberger2016sfm.pdf</u>
[5] Choi, Y., and K., Kumar. 2024. "Graph neural network-based surrogate model for granular flows." Comput. Geotech. 166, 106015. https://doi.org/10.1016/j.compgeo.2023.









