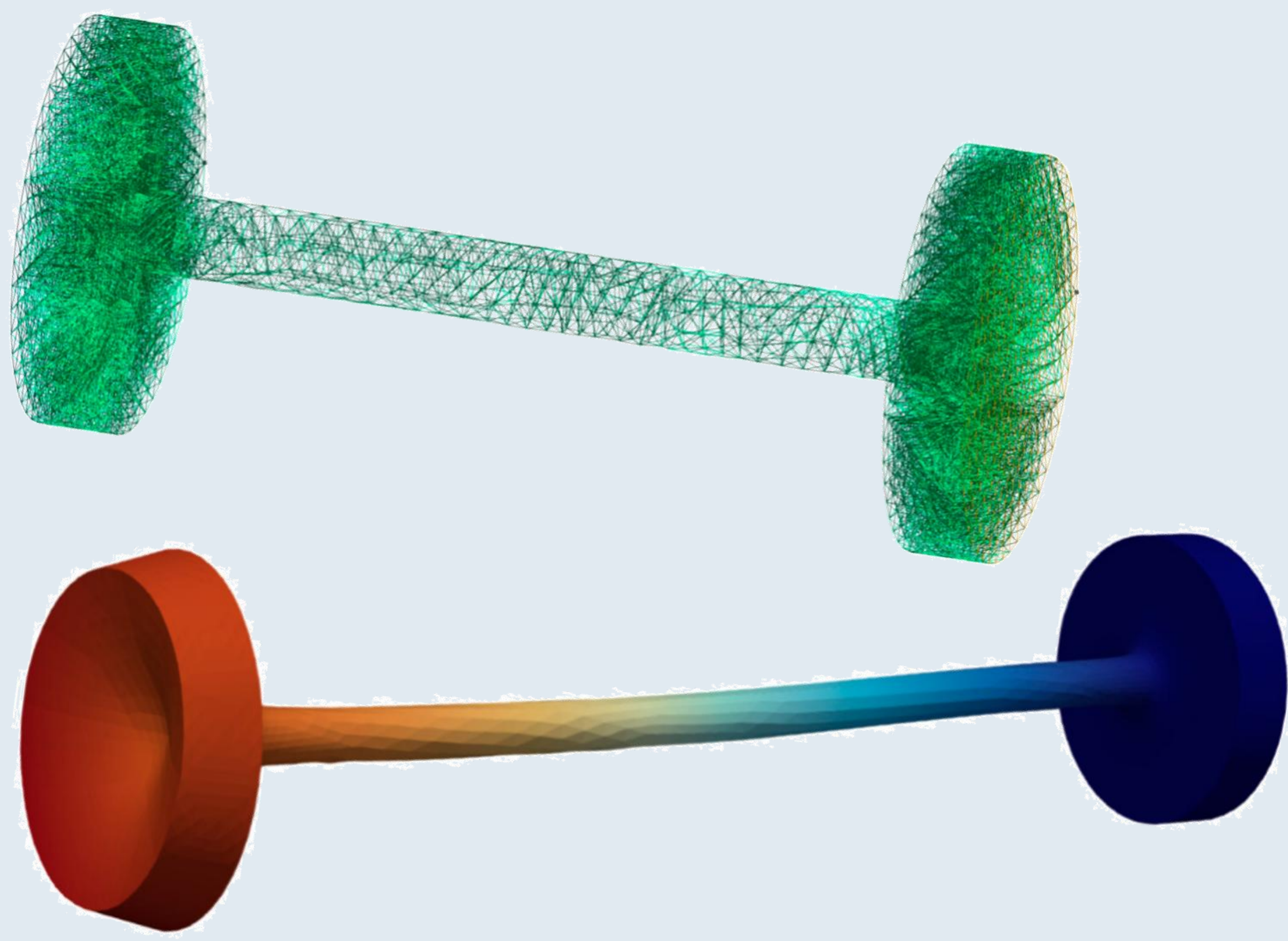




From Text to Tech: Sketch- and Language-Driven Physics-Based Simulation with LLMs

This project explores how large language models can turn a plain-language description, or even a rough sketch, into a simulation-ready geometry for COMSOL Multiphysics®. The aim is to lower the expertise barrier to physics-based simulation, letting engineers focus on the problem rather than the software.

Shafiq, O.¹, Alexiadis, A.², & Ghiassi, B.¹

¹School of Engineering, University of Birmingham, UK

²School of Chemical Engineering, University of Birmingham, UK



A pair of eyes into the world of geometry – visualising shape, structure, and simulation through AI-powered lenses.

Abstract

COMSOL Multiphysics® and similar platforms give engineers powerful tools to model and optimise processes, and COMSOL has begun building LLM assistance directly into that workflow; its Chatbot window generates and debugs COMSOL API code from natural-language prompts, and in recent versions consults the COMSOL documentation to ground its answers. Even so, setting up a simulation remains complex: you need to create the geometry, generate a mesh, define boundary conditions, and configure the solver, each step requiring expertise in specialised software. Our goal is to make this process easier. Instead of manually creating input files, engineers describe their setup in plain language, and a Large Language Model (LLM) generates the necessary files for the simulation software. This shifts the focus from learning complex tools to defining the problem, whether it's a reactor, a pipeline, or a mixing vessel, while the LLM handles the technical details. To test this approach, we evaluated LLMs on their ability to generate input files for both simple and complex geometries (Ref. 1). Larger models performed well out of the box, but smaller models, when fine-tuned with domain-specific data such as geometry templates, produced comparable or even better results (Ref. 2). We also explored using additional inputs, such as engineering drawings or partial sketches, to further improve accuracy through two routes: fine-tuning a vision-language model to interpret a drawing directly, and a fixed base model paired with a CLIP model that retrieves the closest-matching geometry templates as context (retrieval-augmented generation, conceptually parallel to COMSOL's documentation-aware Chatbot, but applied to sketches and geometry). Together, this points toward a path from a plain-language description or a rough sketch to a simulation-ready geometry, and suggests how sketch-driven, retrieval-augmented generation could extend the kind of LLM assistance now emerging inside COMSOL Multiphysics®.

Methodology

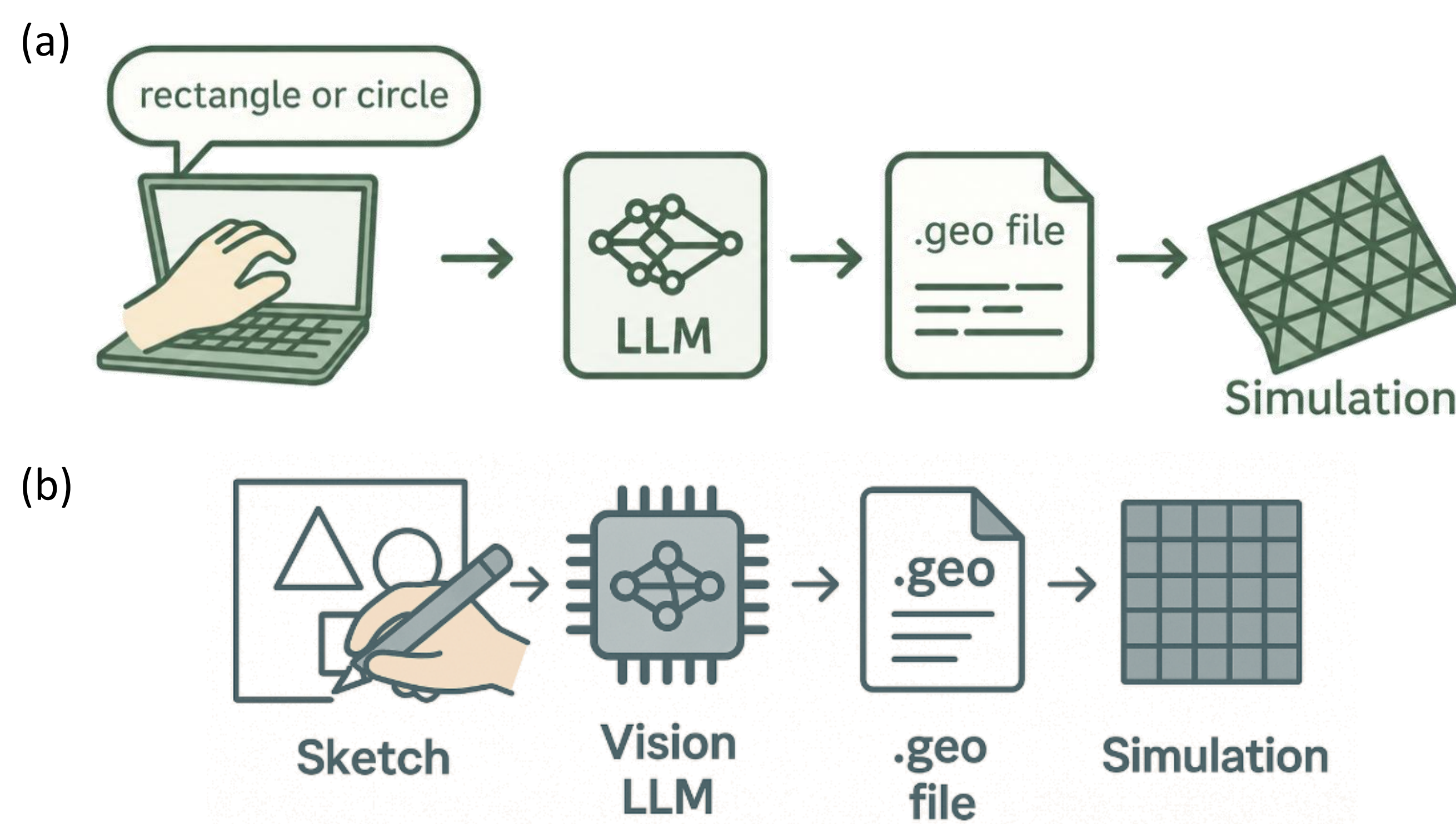
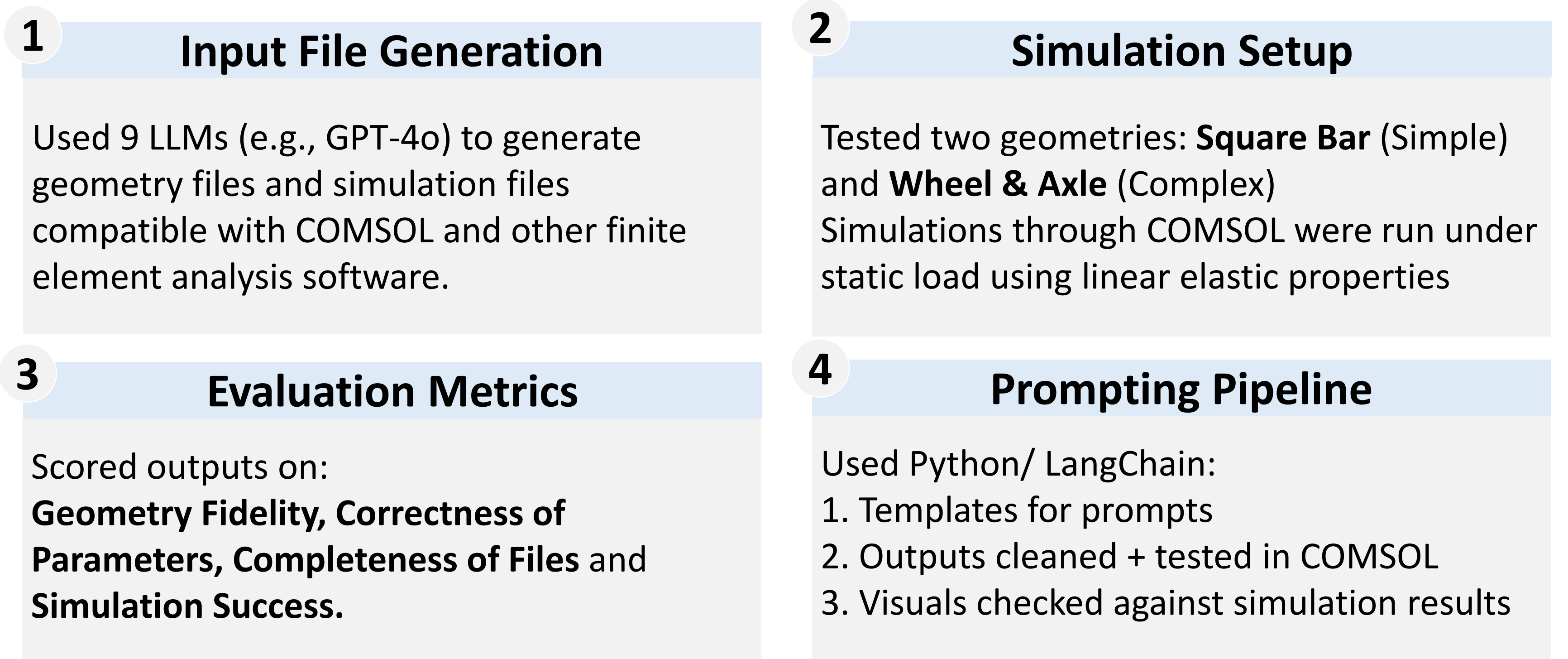
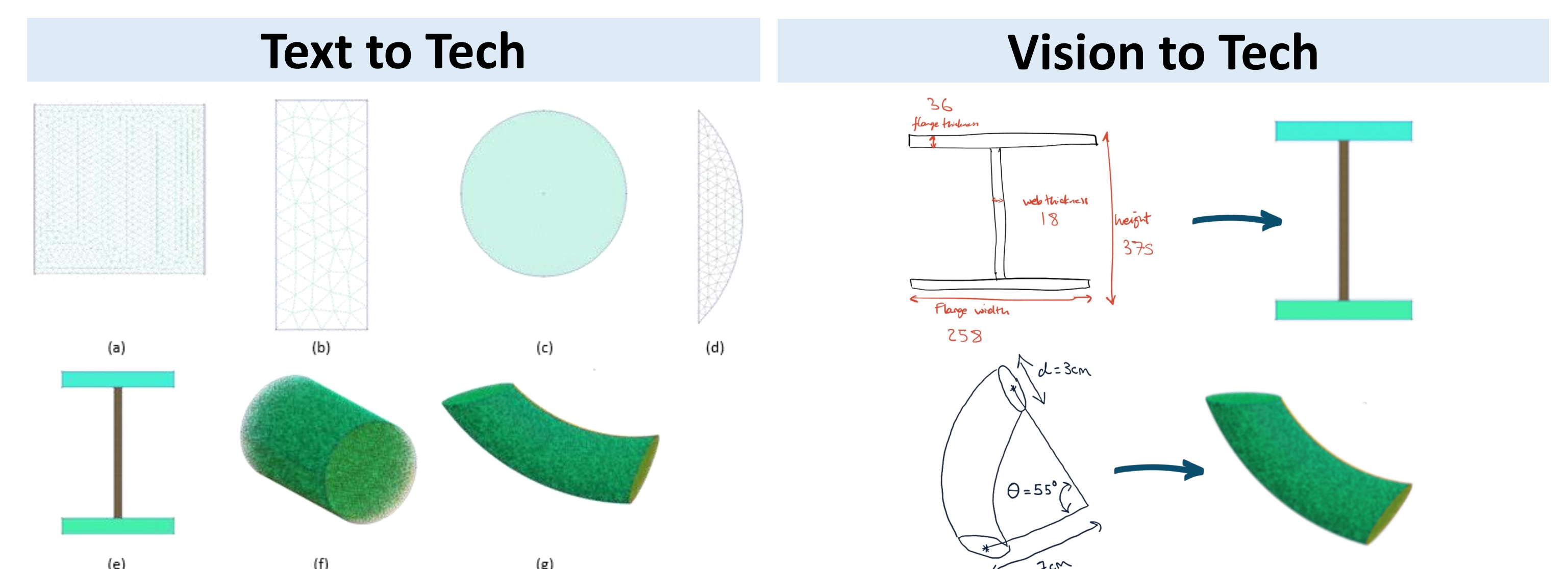


FIGURE 1.(a) Text to Tech pipeline, (b) Vision to Tech pipeline

Results

This project demonstrates how large language models (LLMs) can transform simulation workflows in civil and chemical engineering that utilise COMSOL Multiphysics® by automating the generation of geometry allowing practitioners to focus on simulation efforts. We began by evaluating the capabilities of general-purpose and fine-tuned LLMs in converting text prompts into simulation-ready .geo files. Our **Text to Tech** pipeline showed that even lightweight models, when fine-tuned, can achieve near-perfect results in geometry fidelity, dimensional accuracy, and simulation completeness rivaling larger models like GPT-4o. To extend automation to conceptual design stages, we developed a **Vision to Tech** pipeline that accepts hand-drawn sketches as inputs. Using the UnsLoth FastVisionModel and LoRA adapters, we fine-tuned vision-language models on annotated image-geometry pairs. In our comparative study, Gemma 3 4B significantly outperformed other vision models (e.g., Qwen 2.5 VL 3B & 7B), accurately translating sketches into valid .geo files across a range of geometries. Together, these pipelines demonstrate that LLMs, both in text-only and multimodal forms can reduce technical barriers, speed up simulation setup, and improve accessibility for engineers working on real-world physical systems.



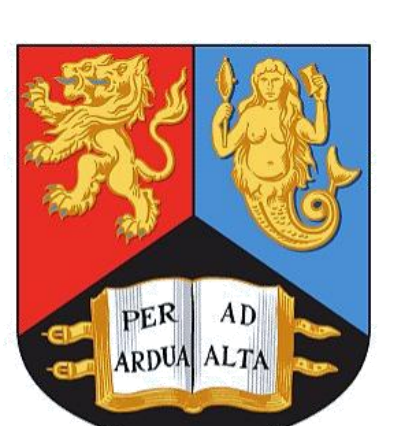
Before fine-tuning, small models were unable to produce executable or working geometry. After fine-tuning, they excelled.

Small models were successfully fine-tuned to convert sketches into accurate .geo scripts. Handled complex shapes (I-beams, bent pipes) with high precision.

FIGURE 2. Outputs of Text to Tech and Vision to Tech pipelines.

REFERENCES

1. Shafiq, O., Rahmat, A., Alexiadis, A. and Ghiassi, B., 2025. Evaluating the performance of large language models for geometry and simulation file generation in physics-based simulations. *Applied Sciences*, 15(22), p.12114.
2. Shafiq, O., Alexiadis, A. and Ghiassi, B., 2025, July. Automating Engineering Simulation Setup with AI-Driven Generative Models. In *International Conference of Steel and Composite for Engineering Structures* (pp. 691-699). Cham: Springer Nature Switzerland.



UNIVERSITY OF BIRMINGHAM