

# Hallucinations and Phantoms: Correcting AI Attempts at Chemistry

Author: Dorian Canelas

Affiliation: Duke University

Target Audience: Educators in general; science teachers; STEM professors; High school students; college students

Estimated Engagement Time: 5 minutes (video slideshow set to music)

## Description

This 5 minute video set to music mixes real information about the problems with hallucination in large language models (LLM's) with a humorous attempt to generate a correct image with a single specific chemical structure. It describes early efforts in teaching LLM's to reason about molecular structure in organic chemistry, a widely taken undergraduate chemistry course. A new concept, "phantom molecules," is introduced and explored by setting the presentation to a musical improvisation of *Phantom of the Opera*, performed by the contributing author's late father. The slides feature custom-generated illustrations.

The work explores the limitations of pattern-matching AI systems when mechanistic reasoning and symbolic representations are required. Highlights include examples of the key contributors to these domain-specific hallucinations, such as tokenization failures, data gaps, and biased training sets. This educational resource is designed to build awareness and generate discussion between the key stakeholders in this area: educators, scientists, students, and AI-developers.