Advances in Chemistry with Artificial Intelligence (AI)

This project explores how AI models are now driving modern chemistry in directions that range from molecular design and retrosynthesis to protein folding and reaction prediction. The integration of AI into chemistry makes the leap from conventional trial-and-error experimentation toward data-driven chemical discovery.

I'll give an overview of how AI models have evolved, how they work, and how specialized large language models (LLMs) like ChemDFM are changing the game with respect to chemical synthesis and molecular understanding. This presentation would demonstrate practical applications of AI in retrosynthesis planning, protein modeling by examples such as AlphaFold, and automated decision systems for chemical design.

Key Research Questions

1. What is an AI model, and what types of models are used in chemistry today?
2. How have AI models evolved, and what does the general term “AI” mean in modern chemistry?
3. What is a system prompt, and how is it used in large language models for chemical reasoning?
4. How do LLM-augmented synthesis design programs help automate and optimize reaction planning?
5. What makes ChemDFM a breakthrough compared to earlier models like AlphaFold?
6. What is retrosynthesis, and how can AI enhance it through atom-level reasoning?

Starting Sources

1. LLM-Augmented Chemical Synthesis and Design Decision Programs (arXiv: 2505.07027) – focus on pp. 16–18 (System Prompts).
2. Developing ChemDFM as a Large Language Foundation Model for Chemistry (arXiv: 2401.14818) and ChemDFM on HuggingFace: https://huggingface.co/OpenDFM/ChemDFM-v1.0-13B
3. Deep Learning of Proteins with Local and Global Regions of Disorder (arXiv: 2502.11326) – compares protein understanding to AlphaFold.
4. Atom-Anchored LLMs Speak Chemistry: A Retrosynthesis Demonstration (arXiv: 2510.16590).

Presentation Format

Format: Interactive Webpage (AI-style simulation).

The presentation will mimic Google's Gemini interface called "Chemini", it will be an application permanently published on the web. When opened, the webpage begins an auto-running presentation as if the AI itself were explaining the research interactively. Viewers can type questions, and an AI version of me will explain it. I will also run the LLM that was trained on everything related to Chemistry, ChemDFM.

Visual and interactive elements:

- Background animation resembling a live AI terminal. AI generated me.
- Sections presented in a "chat-like" format for immersion.
- Embedded code blocks and molecule visualizations (if possible).
- Summaries narrated in real-time text.

Chemistry Units Linked

- Structures and Bonding: Chemical structure prediction via neural networks.
- Rates of Reaction: AI-assisted reaction kinetics simulations.
- Organic Chemistry: Retrosynthesis and molecular pathway prediction.
- Equilibrium: LLMs predicting stable reaction states.

Timeline

- Nov 7: Proposal and rubric submission.
- Nov 8–30: Build application layout, integrate research.
- Dec 1–15: Script interactive flow and design elements.
- Dec 18: Submit final interactive webpage.

RUBRIC (Total: 30 Marks)

Criteria	Level 1 (1-5)	Level 2 (6-10)	Level 3 (11-15)	Level 4 (16-20)	Marks
Research Depth	Minimal sources, unclear focus.	Basic understanding of AI in chemistry.	Clear explanations and integration of examples.	Deep, accurate research with strong source use.	/10
Chemistry Connection	Weak chemistry relevance.	Links to one unit vaguely.	Connects clearly to multiple units.	Thorough explanation linking several units coherently.	/6
Creativity and Presentation	Plain or non-interactive design.	Some interactivity.	Strong visuals and structure.	Highly creative, engaging AI-style simulation.	/6
Clarity and Organization	Unclear or incomplete sections.	Some logical flow.	Well-organized content and transitions.	Seamless flow and polished presentation.	/4
Technical and Aesthetic Quality	Low-quality visuals or bugs.	Basic webpage with errors.	Stable and functional.	High-quality layout resembling Gemini's UI.	/4