



**Pacific
Northwest**
NATIONAL LABORATORY

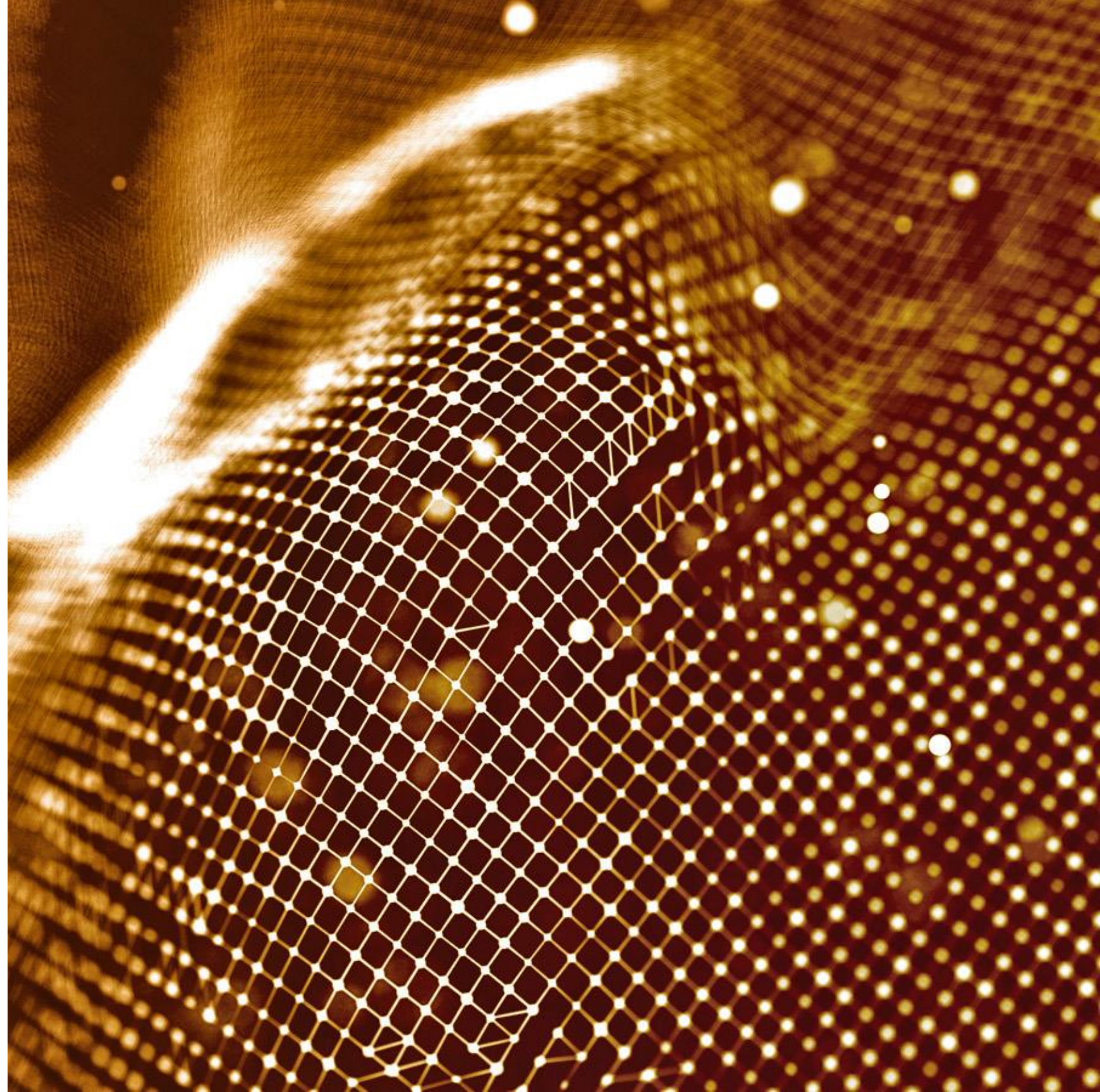
ChemReasoner: A Generative AI approach for Discovering Novel Catalysts

Sutanay Choudhury

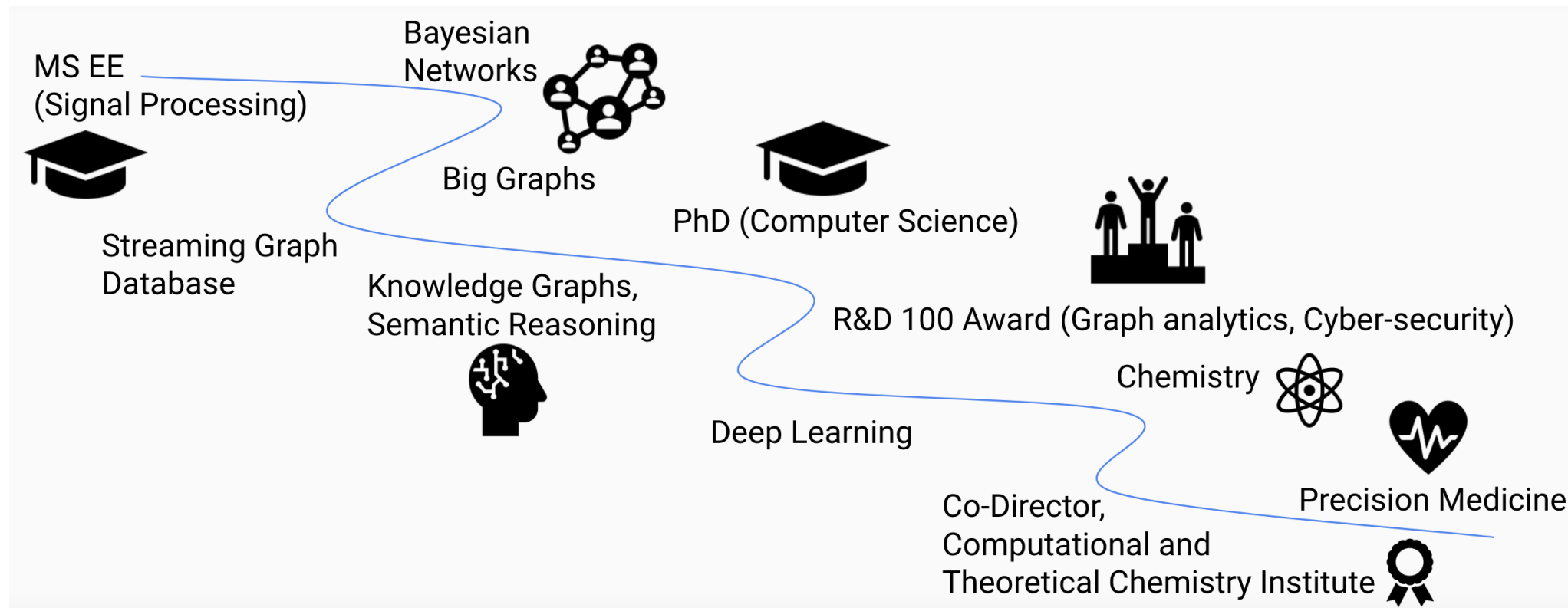
Chief Scientist
Advanced Computing, Mathematics and
Data Division
Co-Director
Computational and Theoretical Chemistry
Institute

U.S. DEPARTMENT OF
ENERGY **BATTELLE**

PNNL is operated by Battelle for the U.S. Department of Energy



My Journey: From Reasoning to Science



You do not need any background beyond high-school chemistry for this talk. Let's just think step-by-step together.

Why should we care about Catalysis?

How can Generative AI advance this field of Science?

How can hardware innovations advance Scientific discovery?

Driving application: Synthetic Biofuels



HOME » NEWS & MEDIA

August 19, 2021 | News Release

Faster and Cheaper Ethanol-to-Jet-Fuel on the Horizon

New catalyst and microchannel reactors improve efficiency and cost

[Karyn Hede, PNNL](#)



SHARE: [f](#) [t](#) [in](#) [✉](#)

Related Content

DECEMBER 12, 2022 |
Staff Accomplishment

[Devaraj on npj Materials Degradation Editorial Board](#)

RICHLAND, Wash. — A patented process for converting alcohol sourced from renewable or industrial waste gases into jet or diesel fuel is being scaled up at the U.S. Department of Energy's [Pacific Northwest National Laboratory](#) with the help of partners at Oregon State University and the carbon-recycling experts at [LanzaTech](#).

DECEMBER 12, 2022 |

We need better catalysts to lower cost and improve efficiency for industrial production

Jet Fuel from Carbon Waste Powers Virgin Atlantic Flight

BY ALAN, ON OCTOBER 3RD, 2018



Virgin Atlantic chairman Richard Branson flags-in flight 16 at London's Gatwick Airport. Flight 16 was the company's first flight to use fuel from recycled industrial carbon waste. (Virgin Atlantic, Twitter)

3 Oct. 2018. A Virgin Atlantic flight arriving earlier today from Orlando, Florida to London's Gatwick airport is the first to use a biofuel made from fermented carbon-rich industrial waste gases. The recycled jet fuel for the Boeing 747 aircraft is a project of industrial recycled carbon company [LanzaTech](#), based in Chicago, with [Pacific Northwest National Laboratory](#) or PNNL in Richland, Washington.

Advancing Catalysis with AI

One of our laboratory's
signature strength

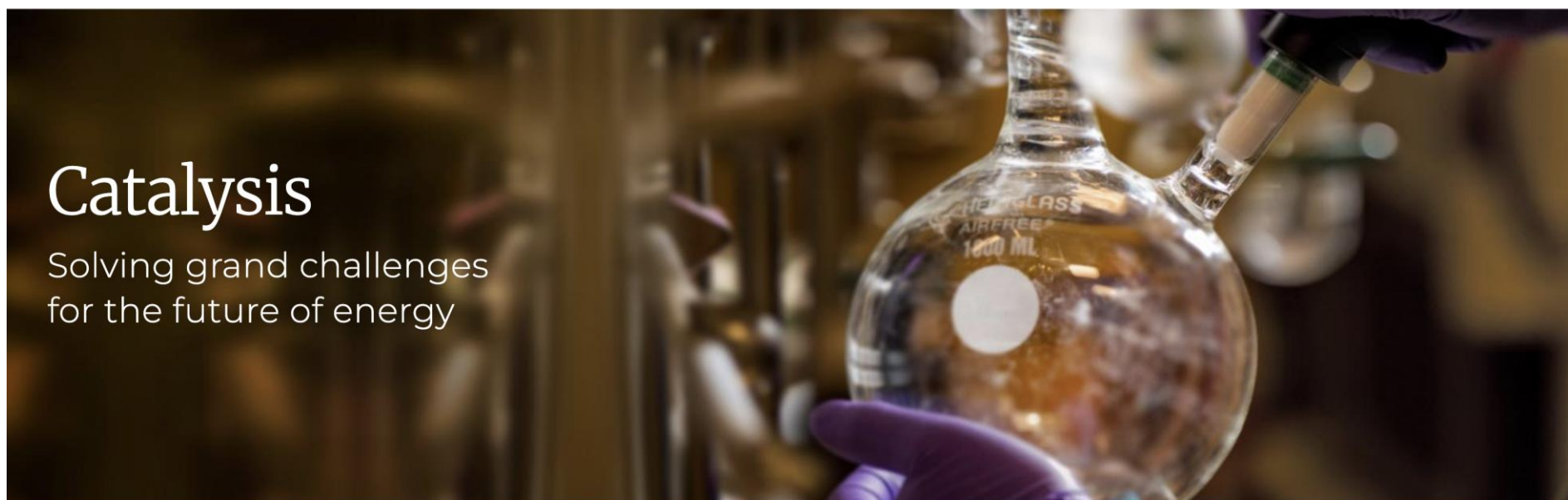


**FEBRUARY 23, 2023 |
News Release**

**Plastic Upcycling to Close
the Carbon Cycle**



APRIL 4, 2023 | Feature
**The Flightpath from a
Groundbreaking Catalyst
to Jets that Soar on
Renewable Fuel from
Waste**



Catalysis

Solving grand challenges
for the future of energy

Need: Recommendation of Novel Catalysts

- **Reaction**
 - $\text{CO}_2 + \text{H}_2 \rightarrow \text{CO} + \text{H}_2\text{O}$
- **Candidate questions:**
 - 1. Recommend a catalyst for this reaction**
 1. Answer: Platinum. and Rhodium
 2. Explanation: They are single atom catalyst
 - 2. Recommend a cheap catalyst for this reaction**
 - 3. Recommend an alternate for Platinum for this reaction**
 1. Answer: Cobalt
 2. Explanation:

Generative AI approach:

Integrate logical reasoning from scientific literature and structure driven reasoning from simulations

How does an out-of-box LLM perform?



which metal is a good alternative to Pt for hydrogenation?



There are several metals that can be used as alternatives to platinum for hydrogenation reactions, including nickel, palladium, and ruthenium. These metals have similar properties to platinum, including a high affinity for hydrogen, and can be used in many of the same applications. However, they may not be as effective in some cases, so it's important to carefully consider the specific requirements of the reaction before choosing a metal to use.



provide more details

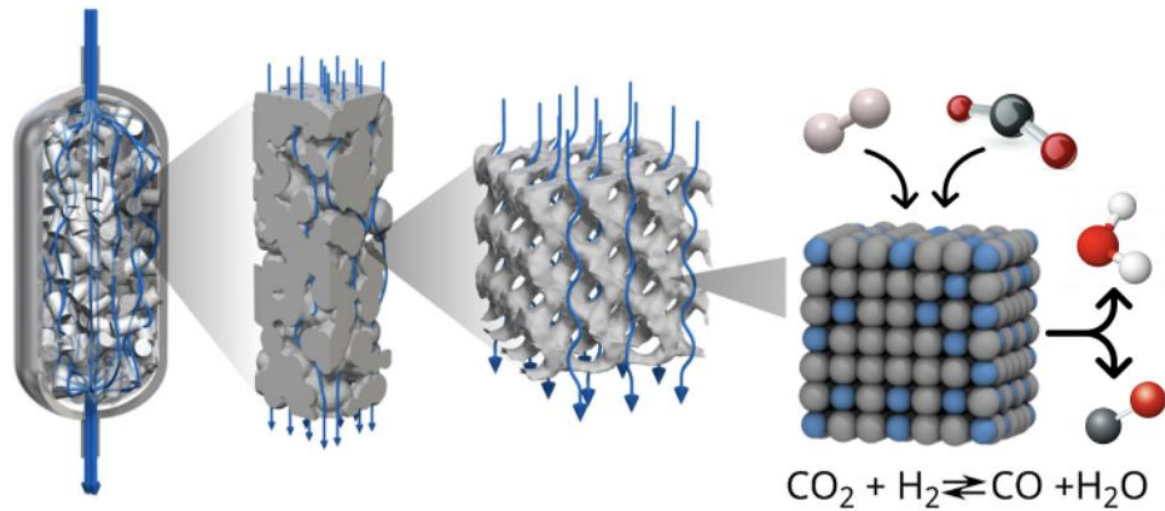


Hydrogenation is a chemical reaction in which hydrogen is added to a molecule. This reaction is often used in the production of a variety of chemicals, including oils, plastics, and pharmaceuticals. Platinum is a commonly used catalyst for hydrogenation reactions because it has a high affinity for hydrogen and can speed up the reaction without being consumed by it. However, platinum is also a very expensive metal, so there is a need for alternative catalysts that are less expensive and more readily available.

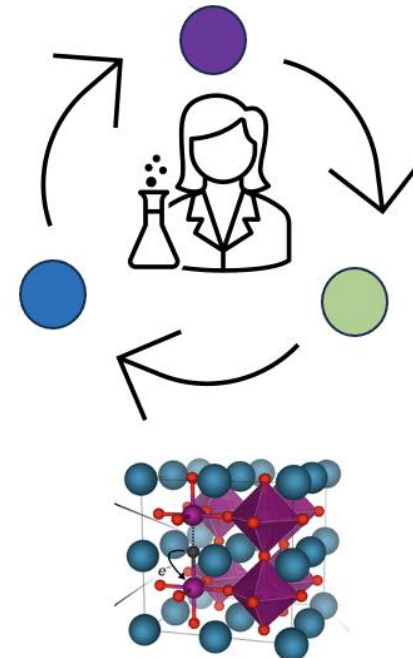


Nickel, palladium, and ruthenium are three common alternatives to platinum for hydrogenation reactions. These metals are less expensive and more readily available than platinum, and they have similar properties that make them effective catalysts for hydrogenation reactions. For example, they all have a high affinity for hydrogen and can speed up the reaction without being consumed. However, they may not be as effective as platinum in some cases, so it's important to carefully consider the specific requirements of the reaction before choosing a metal to use.

The Gaps – beyond Hallucination

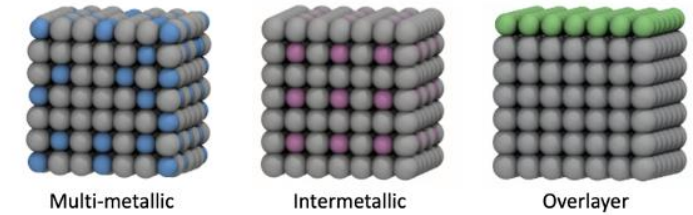


Query: What are the top-3 catalysts that are cheap, perform reverse water gas shift reaction at lower temperature (<200 C) and demonstrate higher adsorption energy for both CO_2 and H_2 ?



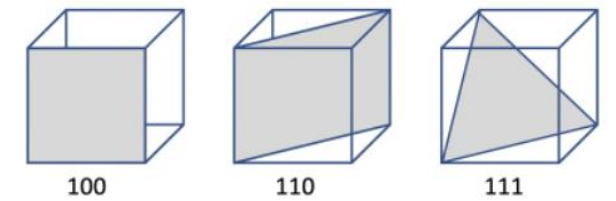
Reason-via-electronic structure

Q + “Let’s think in terms of metal-support interaction”



Reason-via-catalyst composition

Q + “Let’s think in terms of bulk structural properties”



Reason-via-target properties

Q + “Let’s think in terms of adsorption and crystal planes”

- Answers from LLM often lack the specificity experts seek
- Non-deterministic nature of the querying is another issue

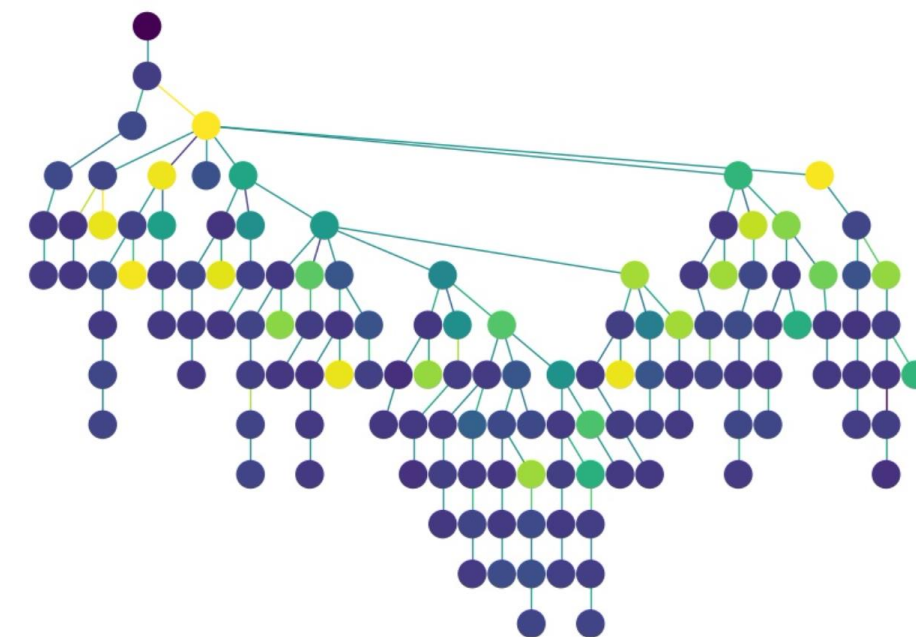
Developed a new reasoning dataset focused on Catalysis

Questions	Answers	Reasoning criteria
What are the top catalysts with higher adsorption energy for both CO ₂ and H ₂ (or facilitates both CO ₂ and H ₂ adsorption)	<p>Noble metal catalysts such as Pt, Rh, Pd, Ru supported on reversible metal oxide i.e., CeO₂ (cerium oxide), TiO₂ (Titanium dioxide)</p> <p>While noble metals are active for hydrogen adsorption, reversible metal oxide facilitates the CO₂ adsorption. The oxygen vacancy present in the reversible metal oxide facilitates C-O bond cleavage of CO₂. Generally, interface sites are coined as the active sites. Higher metal-support interaction is key for their high activity.</p>	<p>Adsorption energy</p> <p>Electronic structures</p> <p>Metal-support interaction</p>
Identify the top catalysts that exhibit weak adsorption energy for CO (product)	Metal catalysts such as Au, Ag, Cu, Zn demonstrate weak adsorption energy corresponding to CO	Adsorption energy
What are the top catalysts that perform RWGS reaction at lower temperature (<200 °C)	Atomically dispersed Pt, Rh, Pd and Ru catalysts on CeO ₂ or TiO ₂ . Atomically dispersed metal sites bind CO very weakly due to their unique electronic structure consequently exhibits high selectivity to CO	Electronic structures
Is RWGS reaction structure sensitive?	Yes. Particles that are typically <2 nm are more active for the RWGS reaction. smaller particle size enables higher metal-support interaction which facilitates the CO ₂ activation and facilitates the reaction. In general step and kink sites i.e., coordinatively unsaturated sites (such as 110 planes) exhibit weaker CO binding energy.	<p>Electronic structures</p> <p>Adsorption on specific crystal planes</p>
What are the crystal planes that are most active for the adsorption of CO ₂ for noble metal catalysts?	Open structure such as (100) or (110) planes of metal catalyst are more active towards CO ₂ activation. Binding energy of CO ₂ is higher on these crystal surfaces	<p>Surface structure</p> <p>Adsorption on specific crystal planes</p>

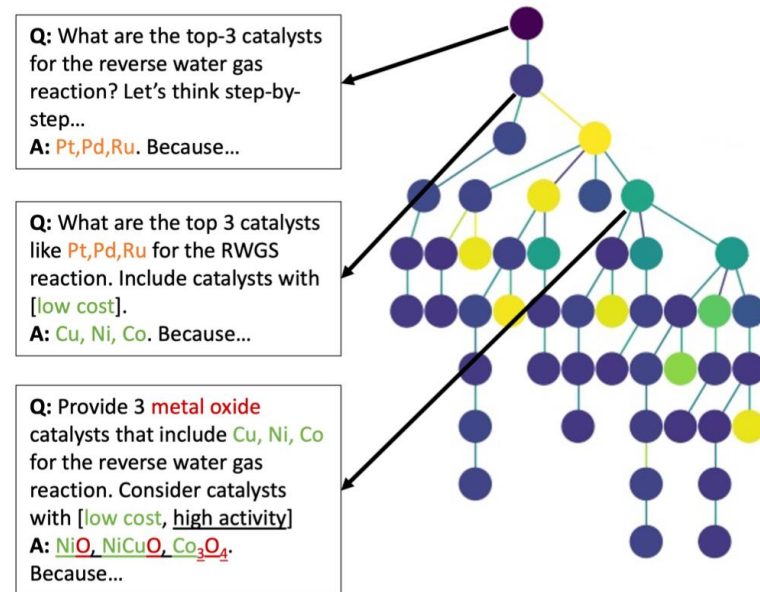
Results

- Superior results over ChatGPT
- Superior performance comes with high-inference cost

Method	OpenCatalysis			BioFuelQR		
	Reward	N_P	d_{max}	Reward	N_P	d_{max}
CoT	2.04	1	N/A	2.27	1	N/A
CoT w/ Self-consistency	4.04	10	N/A	6.38	10	N/A
ToT (breadth-first-search)	9.91	253	5	13.8	253	5
MCR (ours)	12.47	301	9.33	15.6	301	9.5



Qualitative Analysis



Analyze how search progresses through the tree from root to leaf

Question: What are the top-3 catalysts that perform the RWGS reaction at a lower temperature (<200 C) and demonstrate higher adsorption energy for both CO₂ and H₂ (or facilitates both CO₂ and H₂ adsorption)?. Include candidate catalysts with the following properties: low cost. Provide scientific explanations and return a list of top 3 answers and their explanations as a list of pairs. Let's think step-by-step.

Bimetallic PtRe (platinum-rhenium Alloy) – This alloy has been found to be highly effective for both CO₂ and H₂ adsorption, and it catalyzes the RWGS reaction at temperatures lower than 200 C using a lower amount of energy. The PtRe alloy also has a lower cost compared to other catalysts making it a popular choice.



“The answer is correct and consistent with literature. Integrating with Rhenium does bring the cost down. I will surely use it.”

Ni-Mn-Cu: Ni-Mn-Cu catalysts demonstrate one of the highest known rate constants for the RWGS reaction, and also have higher adsorption energy for both CO₂ and H₂ compared to PtRe, Ni-Cu, and Cu-Zn.



“I am really impressed by the specificity of the suggestion and use of its rate constant based justification. If another colleague told this, I would be really intrigued by their insight. I am inclined to validate this in a lab experiment.”

Ni-Al-Co-Mo; Ni-Al-Co-Mo is a low-cost metal oxide catalyst that is effective for both CO₂ and H₂ adsorption. The catalyst is a mixture of nickel (Ni), aluminum (Al), cobalt (Co) and molybdenum (Mo). In addition to being low-cost, Ni-Al-Co-Mo is a highly active and selective catalyst due to the presence of active metal sites and it possesses strong acid sites, which improves the activity of the catalyst.



“The answer is logically coherent and not incorrect. However, it ignores the practical challenge of combining four metals. This recommendation is not actionable.”

Qualitative Analysis

Expert-based review of GPT 3.5 and MCR output

1) Quality: How did the AI methods matched your answer?

Answer: One answer from GPT-3.5 can be considered partially correct (transition metal) while the Monte Carlo Reasoner partially matched my answers and reasoning for the noble metal catalysts' RWGS activity. Both models were not able to address the requirement of catalyst activity of less than $<200\text{C}$. The Monte Carlo Reasoner identified noble metals, Platinum and Ruthenium. Hafnium was something that I would not have considered. For catalysts that have high adsorption energies for CO_2 and H_2 , my answers were three Pt-based catalysts (PtRe/SiO₂, Pt/CeO₂ and Na-doped Pt/ZrO₂). I also identified Ni-based (Ni/La-dopedCeO₂, NiCu, Ni/Ce-Zr-O) and Cu-based (4Cu-Al₂O₃) catalysts from my research. My reasoning is that catalysts that would be expected to demonstrate higher adsorption energies for both CO_2 and H_2 would contain noble and base metals such as Pt, Ru and Ni supported on oxides with a high level of oxygen vacancies to facilitate high adsorption energies for both CO_2 and H_2 . From the manuscripts that I reviewed that have tested RWGS at 200C , none resulted in any significant CO_2 conversion ($>5\%$). Lastly, calculated equilibrium constants from another paper reported 0.0043 at 200C and 0.0830 at 400C .

2) Specificity: Which AI method matched the specificity of your explanation?

Answer: Both methods didn't completely match the specificity of my explanation, but I would choose the Monte Carlo Reasoner because it identified Pt, even as part of a bimetallic catalyst. However, even Pt catalysts do not have much activity ($>5\%$ CO_2 conversion) for RWGS at $<200\text{C}$.

3) Reasoning: Which AI methods used similar chemical descriptors as yours to reason about?

Answer: In part, the Monte Carlo Reasoner. It correctly identified strong adsorption properties for both CO_2 and H_2 for the noble metal catalysts.

4) Did the AI method return any wrong answer?

Answer: Yes, they both did. GPT-3.5's claim that the ionic liquid and zeolite were good catalysts for RWGS was incorrect. They were not identified as RWGS catalysts in my search. Transition metal catalysts, like Ni, Cu, and their alloys, were identified as potential RWGS catalysts but they are not active at $<200\text{C}$. The Monte Carlo Reasoner incorrectly identified Hafnium as a potential RWGS catalyst. However, I conducted a follow-on search because I am not very familiar with its chemistry. Hafnium seems to be able to activate CO_2 but whether it can produce CO selectively through RWGS was not conclusive.

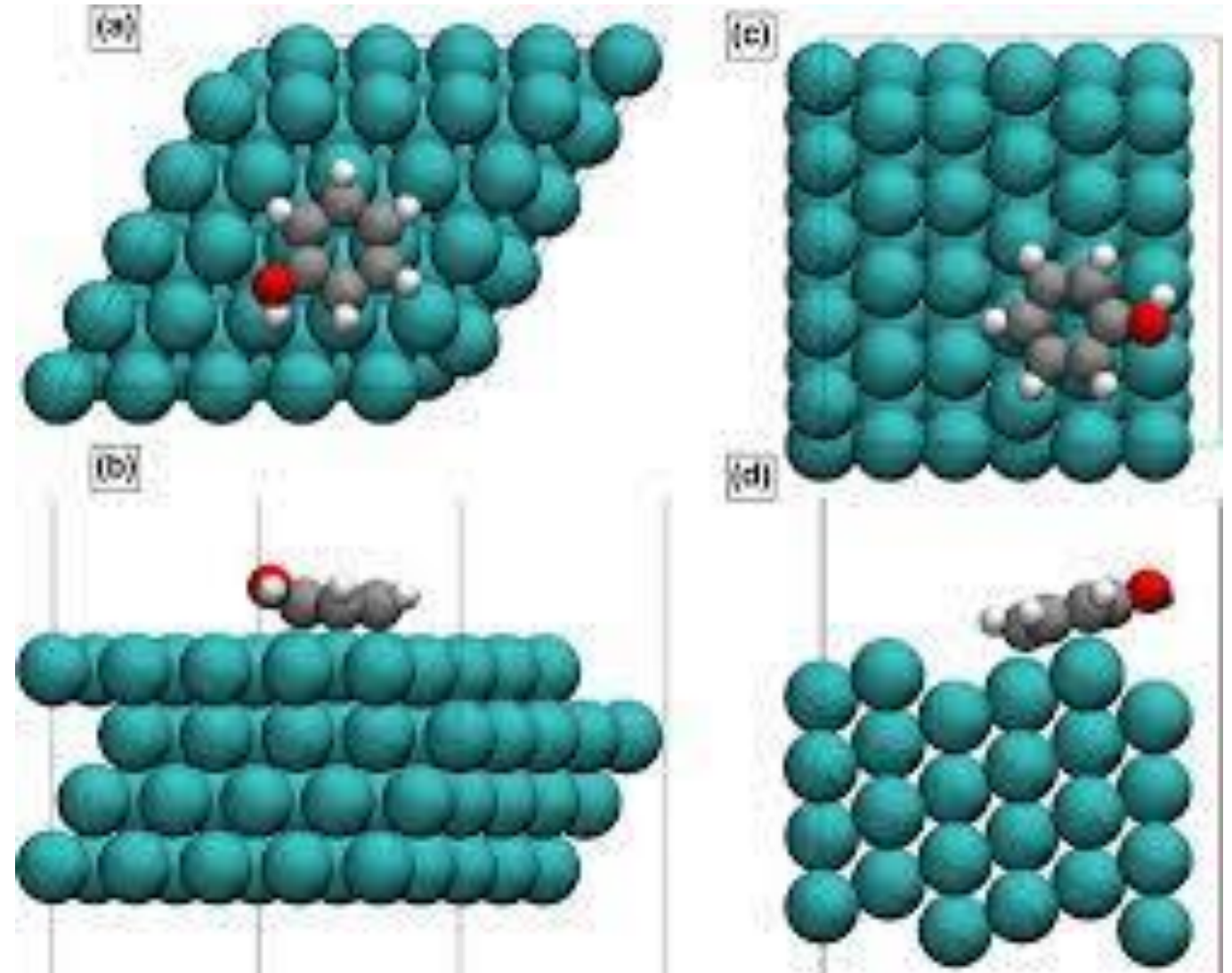
4) Are any of the AI-generated answers novel/superior to the human expert answer?

Answer: The Hafnium suggestion was novel for me, but it was not superior to the human expert answer.

Where can we use help from this community?

Augmenting LLMs via Structure-Driven Reasoning

- Given a string from the LLM output such as “**Platinum**”, we want to infer the 3D structure
- We represent the 3D structure as a “**3D-Atom Graph**” – such a representation considers relative positions and orientations
- This configuration will be passed to a DFT simulation or a trained 3DGNN



Increasing throughput of Hypothesis Generation

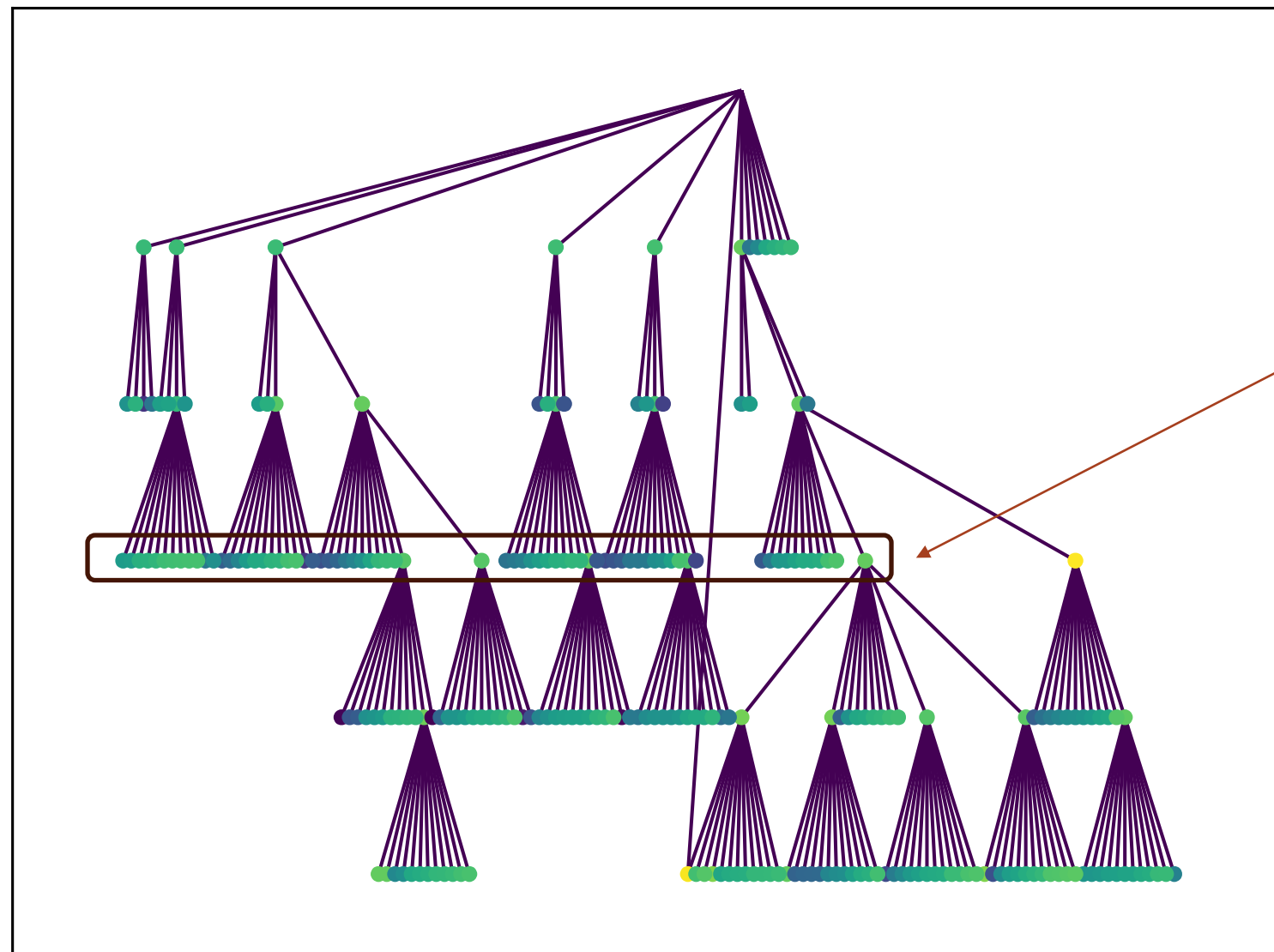


Illustration of the Search Space

- All states in a given level can be evaluated in parallel
- Each state evaluation is a sequence of 3 tasks:
 - *2-3 LLM inferences*
 - *16-32 GNN inferences*
 - *Potentially, execute a DFT simulation*

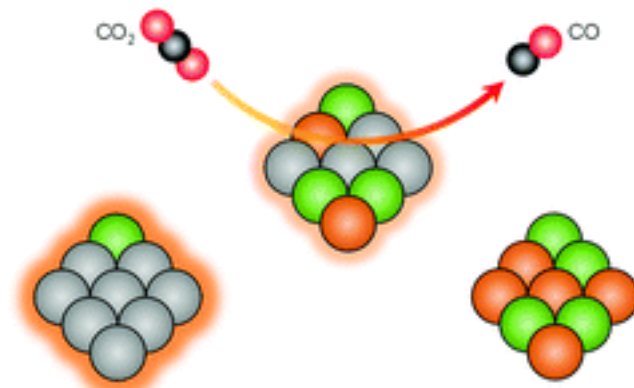
Let's expand that **one** green dot in the next slide

A Single Cycle of Hypothesis Generation and Testing



Hypothesis generation
(LLM)

3D Atom
Graph Generator



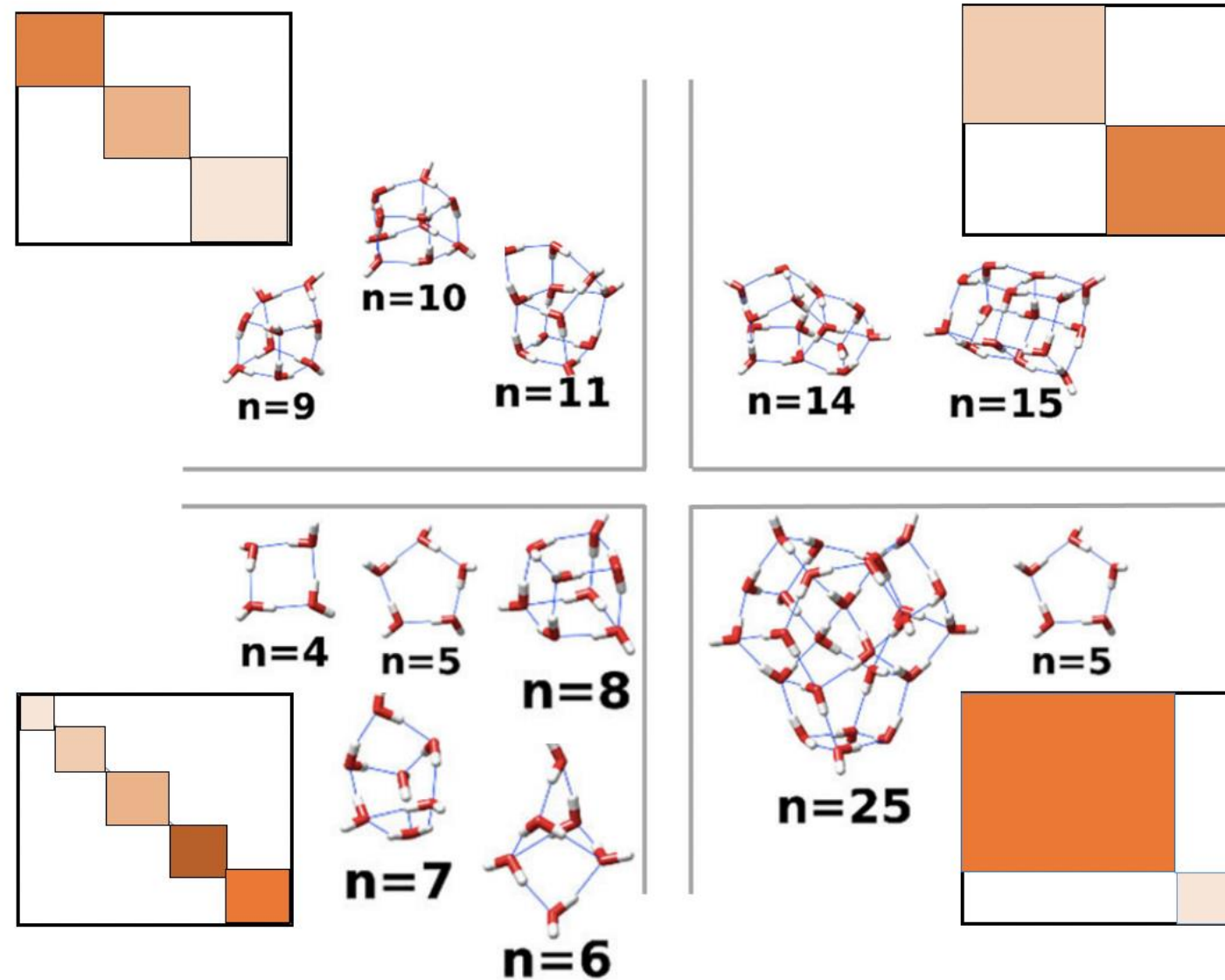
Property prediction
with UQ (GNN)

Active Learning

High uncertainty

DFT Simulation

Training and Inference on Molecular GNNs require processing many small and sparse graphs



See [<https://sites.google.com/view/ai4hydronet/home>] for details

Vision for Future: Computational Pipeline for AI-driven Scientific Discovery

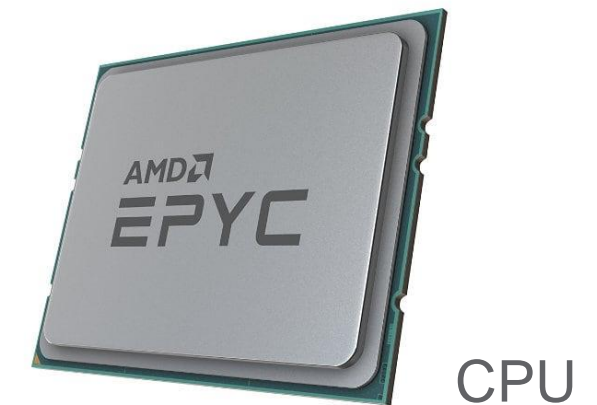
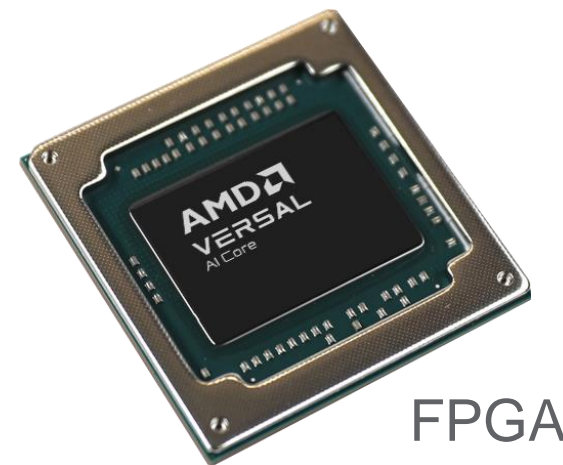
Hypothesis generation
(LLM)

3D Atom Graph
Generator

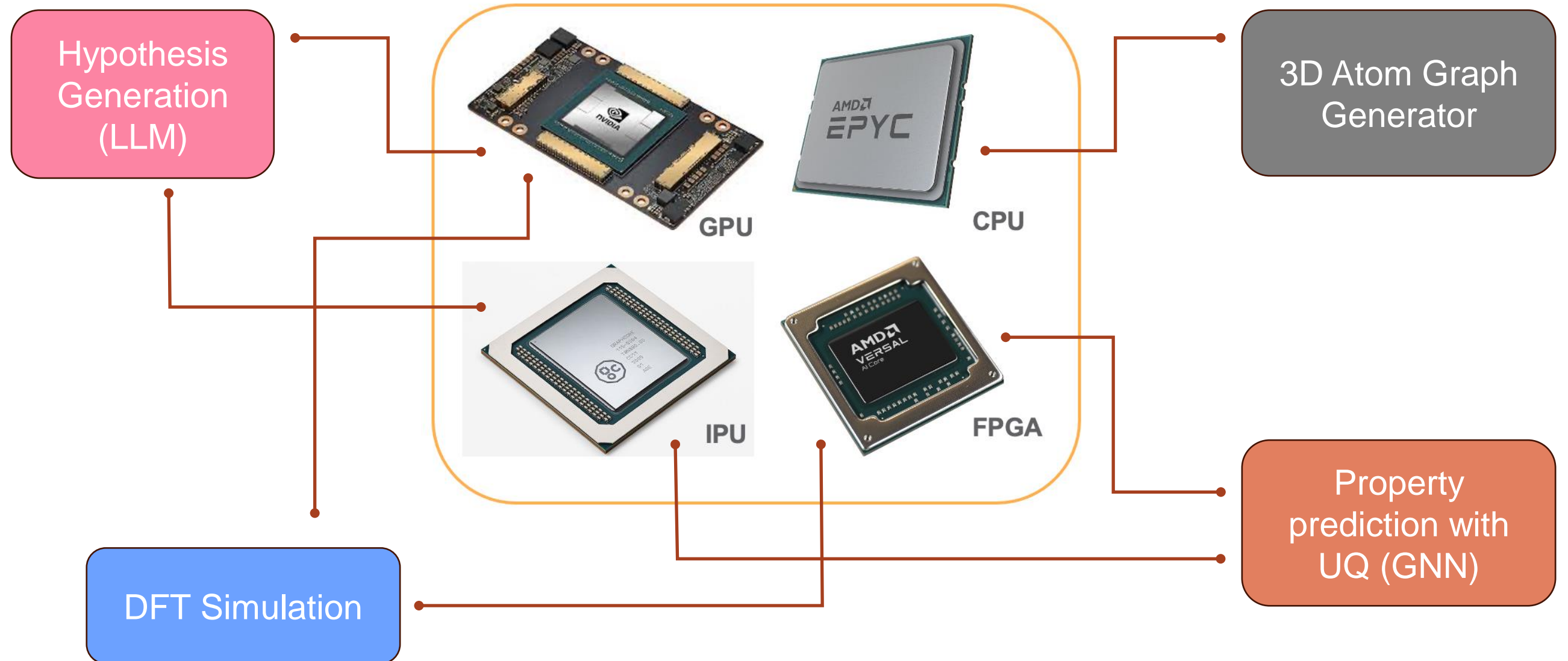
Property prediction with
UQ (GNN)

Active Learning

DFT Simulation



Vision for Future: Computational Pipeline for AI-driven Scientific Discovery





Project Team



<https://github.com/pnnl/chemreasoner/>

Thank you