ChemReasoner: Bridging Generative AI and Computational Chemistry

Sutanay Choudhury

Chief Scientist
Data Science and Machine Intelligence

Co-Director
Computational and Theoretical Chemistry Institute

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

Pacific Northwest NATIONAL LABORATORY

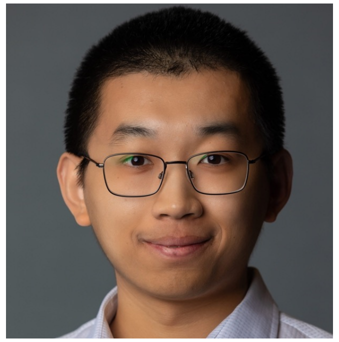
Pacific Northwest NATIONAL LABORATORY

Heterogeneous catalysis, biofuel development, catalysis synthesis and characterization

Large language models, graph neural networks, neural symbolic reasoning

Machine learning for computational chemistry

Microsoft



Computational Chemistry, Quantum Computing, Condensed Matter Physics

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

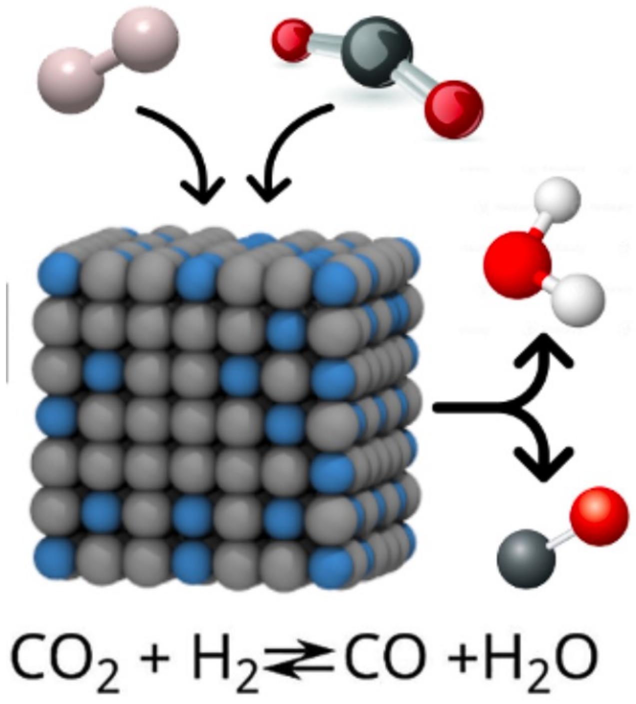
Natural Language Processing, Large Language Models

Outline

- Scientific driver
- Motivation and limitation for LLMs
- Multi-Modal/Compound AI
- Quantitative and Qualitative Analysis
- Scaling needs

What is ChemReasoner?

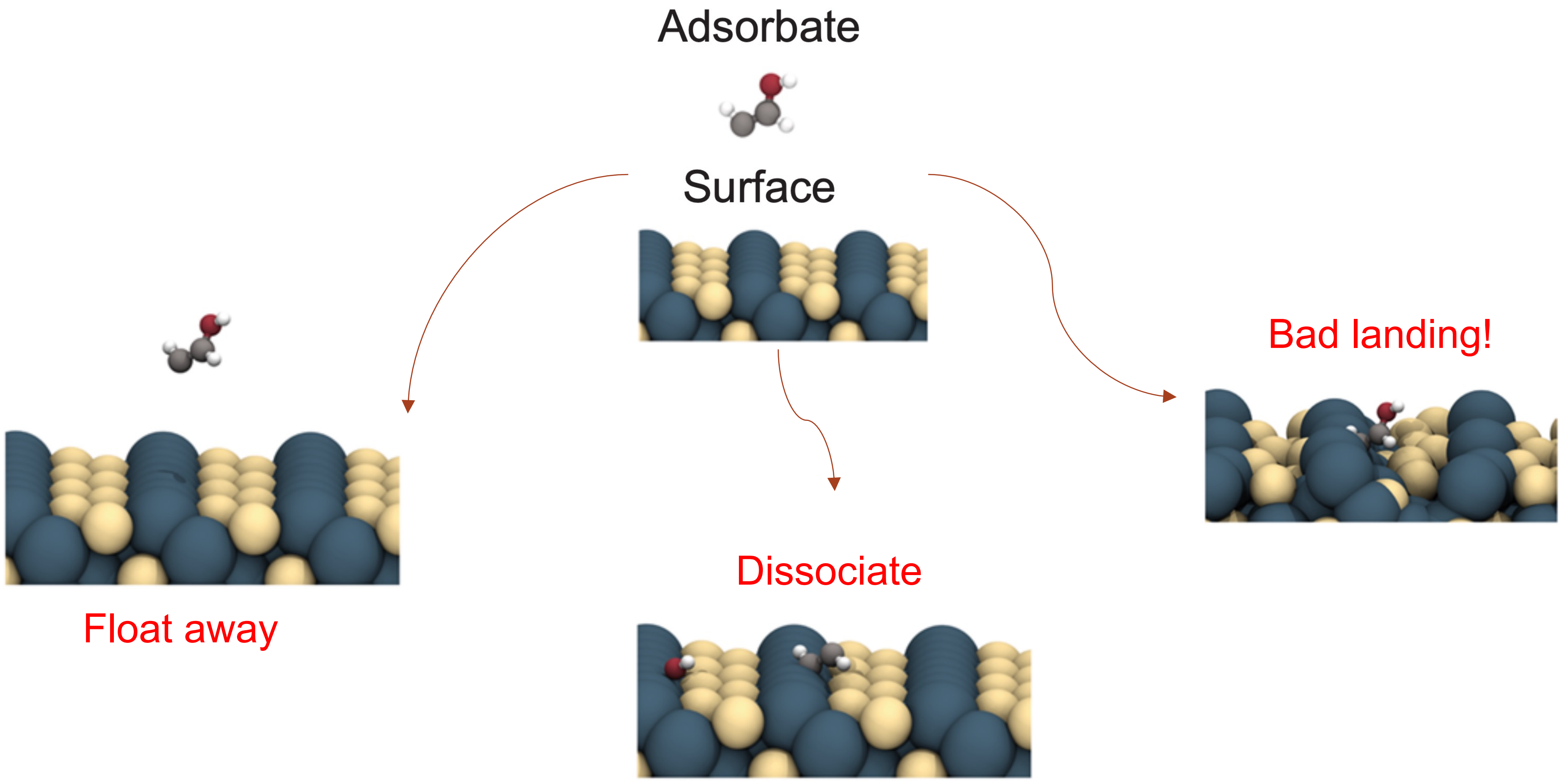
A new AI system designed for Chemistry



Let us begin with our focus: Catalysis

Catalysis – Need and Challenges

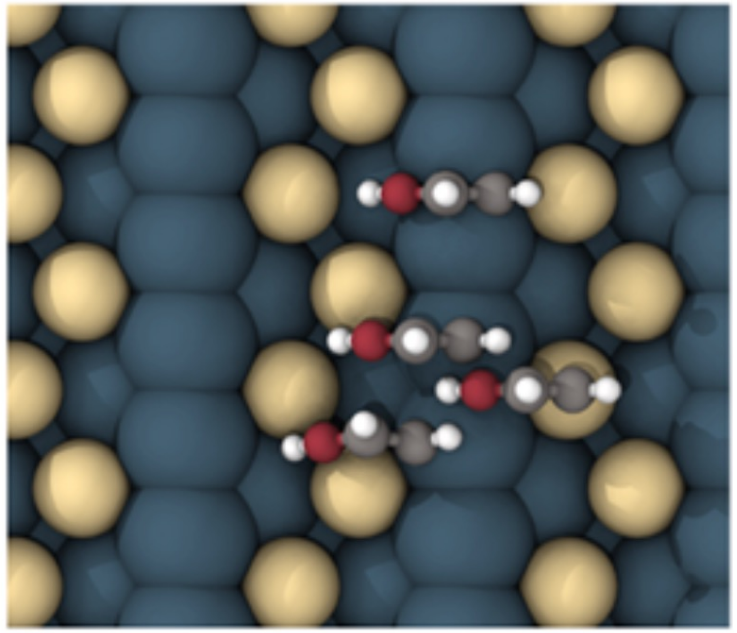
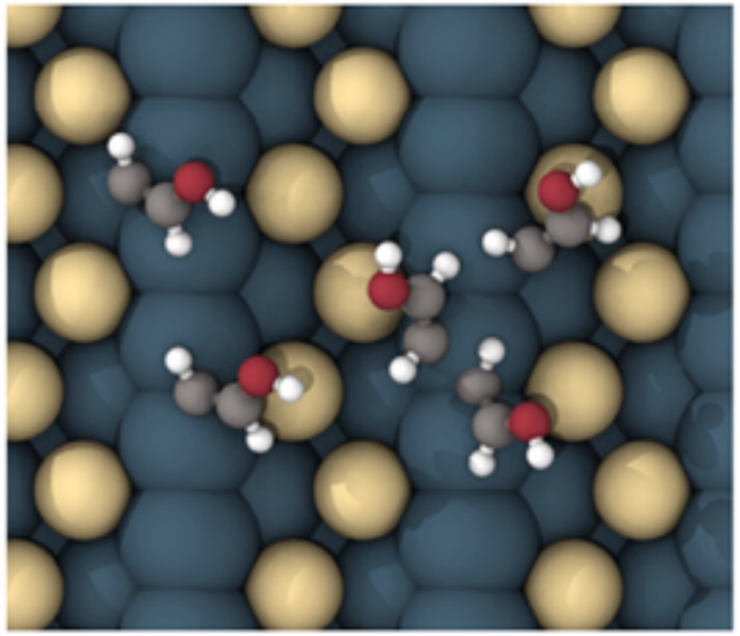
Why Catalysis is Hard?



The process unfolding at microscopic scale can go wrong in many ways

Image courtesy: Lan, J., Palizhati, et al., 2023. AdsorbML: a leap in efficiency for adsorption energy calculations using generalizable machine learning potentials. npj Computational Materials, 9(1), p.172.

Good catalyst – "works ~~every~~ most of the time"

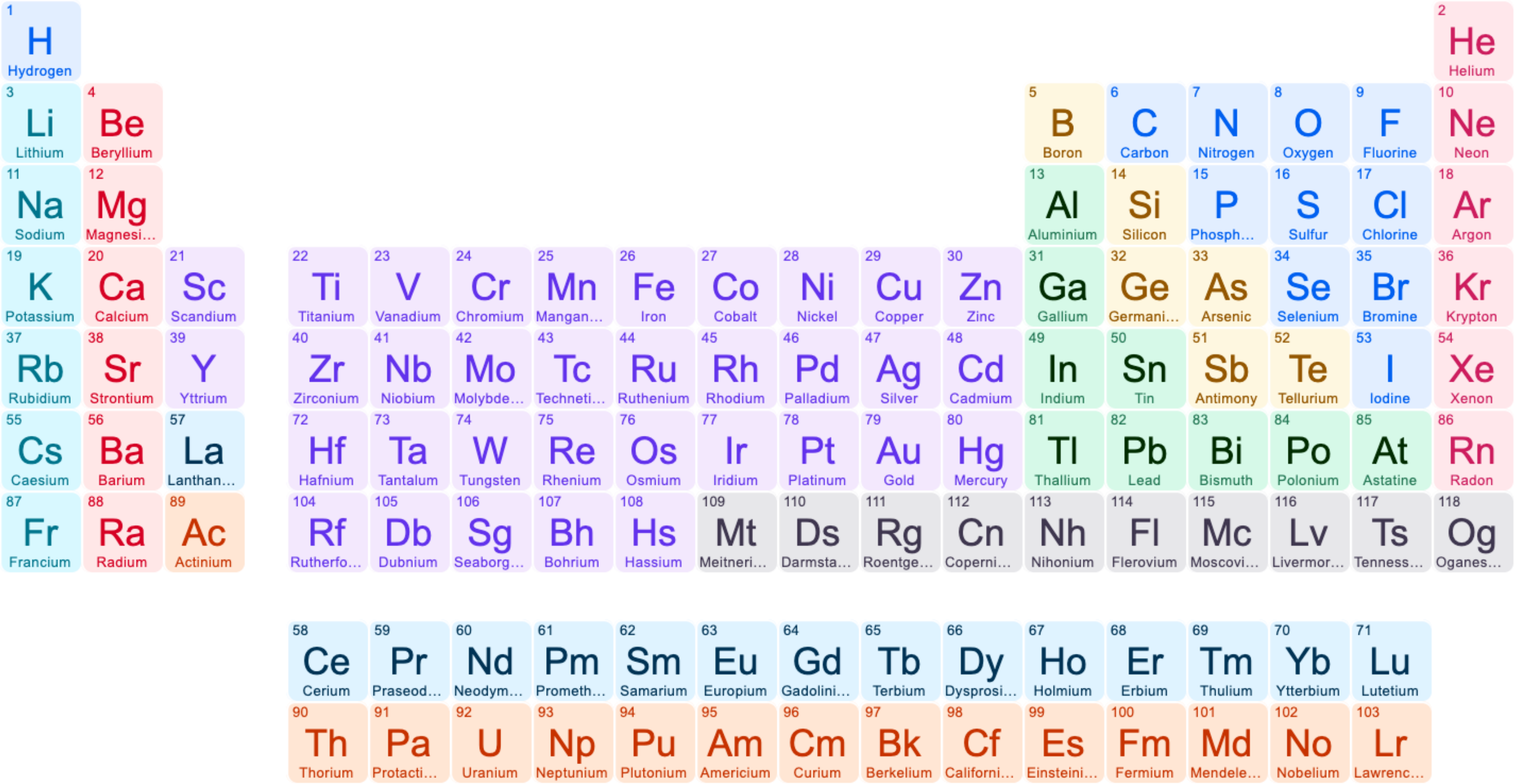



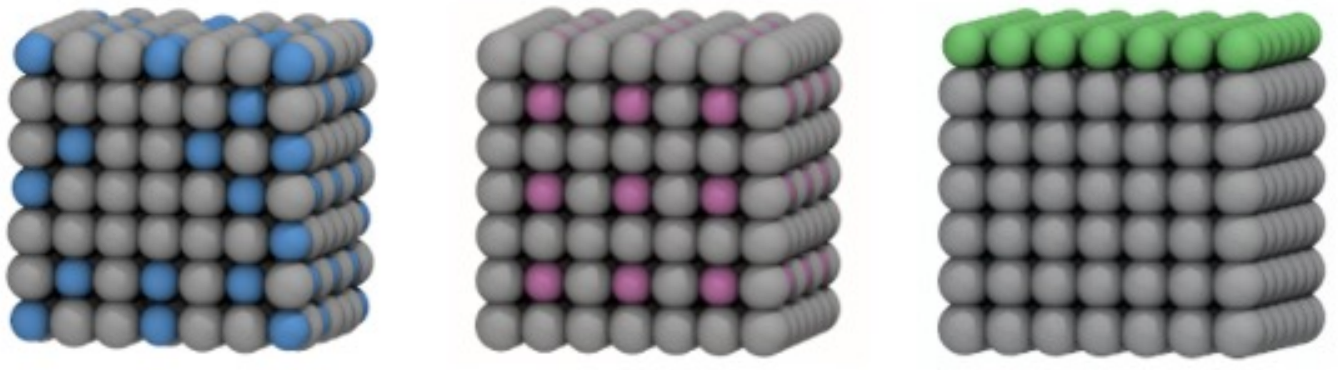
Imagine facilitating this controlled dance of molecules at large scale

Image courtesy: Lan, J., Palizhati, et al., 2023. AdsorbML: a leap in efficiency for adsorption energy calculations using generalizable machine learning potentials. npj Computational Materials, 9(1), p.172.

Pacific Northwest
NATIONAL LABORATORY

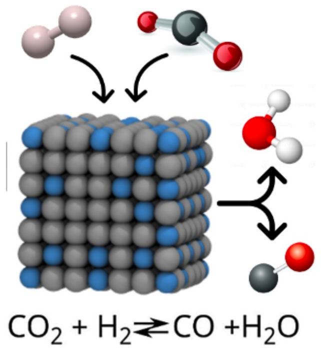
The curse of Combinatorics





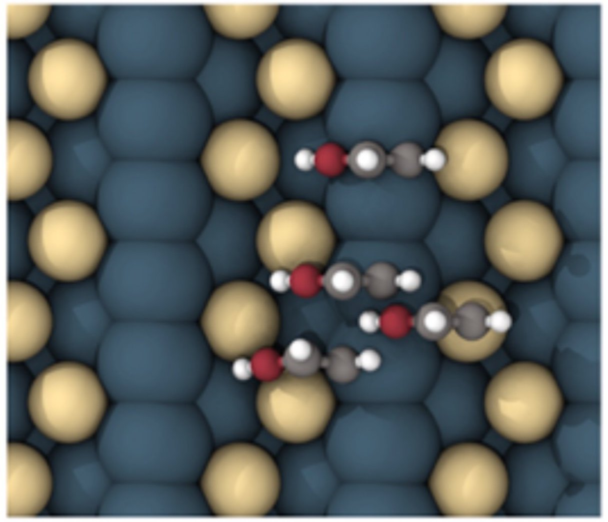
Multi-metallic | Intermetallic | Overlayer

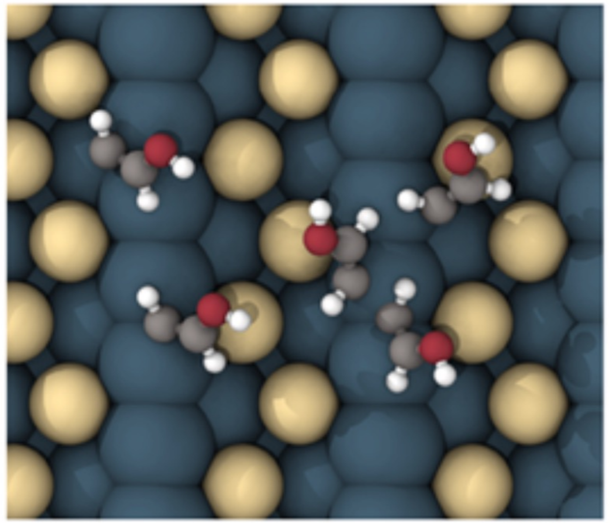
Number of possible catalysts involving **3** metals drawn from a set of **50** is … **19600**



Multiply that by the number of reactants

Economic and Societal Impact





It takes energy for everything

- To break or form a bond
- To hold molecules together in a stable way

A good catalyst is one

- Who will get the job done with less extrinsic energy

Less extrinsic energy → Lower carbon footprint

Inspiration from Literature

Inspiration

- Generate a sequence of questions/answers that mimic human reasoning
- Similar to how we justify in a scientific publication

The catalytic conversion of synthesis gas (also referred to as syngas, $CO + H_2$) to higher oxygenates (C_{2+} oxy) including alcohols, aldehydes, acetates, etc., offers a promising alternative to the production of higher-value fuels and chemicals.[1] Given the intrinsic selectivity of rhodium (Rh) toward C_{2+} oxy, catalysts based on this metal are by far the most studied systems for higher alcohol synthesis (HAS).[2,3] However, due to the inherent reaction kinetics of HAS, which shift the selectivities away from the desired higher alcohols (HA) and C_{2+} oxy, no commercial catalyst exists to date with practically relevant activity and selectivity.[4] Depending on how the CO binds to the catalyst surface, methanol and hydrocarbon (HC) synthesis pathways compete with C_{2+} oxy formation.[5] Specifically, transition metals that facilitate molecular CO adsorption, e.g., Cu and In, are selective to methanol, while those that promote CO dissociation, e.g., Fe and Co, are selective toward HCs.[4,5] The most favorable mechanism for C_{2+} oxy formation on Rh is through the insertion of CO/CHO into CH_x (x = 1−3) species, which requires simultaneous molecular and dissociative chemisorption of CO.[6] While monometallic Rh is primarily selective toward methane and acetaldehyde, it hardly suffices the bifunctional requirement for HAS. For this purpose, various alkali- and transition-metal promoters are added, increasing alcohol formation and carbon chain growth capabilities.[4,5] This significantly improves

ACS Catalysis

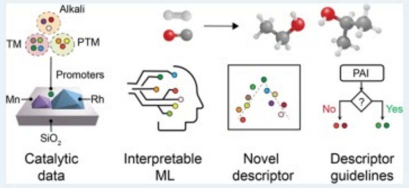
pubs.acs.org/acscatalysis Research Article

Identifying Descriptors for Promoted Rhodium-Based Catalysts for Higher Alcohol Synthesis *via* Machine Learning

Manu Suvarna, Phil Preikschas, and Javier Pérez-Ramírez*

Cite This: *ACS Catal.* 2022, 12, 15373−15385 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Rhodium-based catalysts offer remarkable selectivities toward higher alcohols, specifically ethanol, *via* syngas conversion. However, the addition of metal promoters is required to increase reactivity, augmenting the complexity of the system. Herein, we present an interpretable machine learning (ML) approach to predict and rationalize the performance of $Rh\text{-}Mn\text{-}P/SiO_2$ catalysts (P = 19 promoters) using the open-source dataset on Rh-catalyzed higher alcohol synthesis (HAS) from Pacific Northwest National Laboratory (PNNL). A random forest model trained on this dataset comprising 19 alkali, transition, post-transition metals, and metalloid promoters, using catalytic descriptors and reaction conditions, predicts the higher alcohols space-time yield (STY_{HA}) with an accuracy of R^2 = 0.76. The promoter's cohesive energy and alloy formation energy with Rh are revealed as significant descriptors during posterior feature-importance analysis. Their interplay is captured as a dimensionless property, coined promoter affinity index (PAI), which exhibits volcano correlations for space-time yield. Based on this descriptor, we develop guidelines for the rational selection of promoters in designing improved $Rh\text{-}Mn\text{-}P/SiO_2$ catalysts. This study highlights ML as a tool for computational screening and performance prediction of unseen catalysts and simultaneously draws insights into the property−performance relations of complex catalytic systems.

KEYWORDS: *syngas, computational screening, feature engineering, alloy formation energy*

INTRODUCTION

The catalytic conversion of synthesis gas (also referred to as syngas, $CO + H_2$) to higher oxygenates (C_{2+} oxy) including alcohols, aldehydes, acetates, etc., offers a promising alternative to the production of higher-value fuels and chemicals.[1] Given the intrinsic selectivity of rhodium (Rh) toward C_{2+} oxy, catalysts based on this metal are by far the most studied systems for higher alcohol synthesis (HAS).[2,3] However, due to the inherent reaction kinetics of HAS, which shift the selectivities away from the desired higher alcohols (HA) and C_{2+} oxy, no commercial catalyst exists to date with practically relevant activity and selectivity.[4] Depending on how the CO binds to the catalyst surface, methanol and hydrocarbon (HC) synthesis pathways compete with C_{2+} oxy formation.[5] Specifically, transition metals that facilitate molecular CO adsorption, e.g., Cu and In, are selective to methanol, while those that promote CO dissociation, e.g., Fe and Co, are selective toward HCs.[4,5] The most favorable mechanism for C_{2+} oxy formation on Rh is through the insertion of CO/CHO into CH_x (x = 1−3) species, which requires simultaneous molecular and dissociative chemisorption of CO.[6] While monometallic Rh is primarily selective toward methane and acetaldehyde, it hardly suffices the bifunctional requirement for HAS. For this purpose, various alkali- and transition-metal promoters are added, increasing alcohol formation and carbon chain growth capabilities.[4,5] This significantly improves selectivity and/or activity toward C_{2+} oxy, specifically ethanol synthesis from syngas.

Fe and Mn are among the most commonly investigated promoters to improve Rh's catalytic performance, either as binary or ternary systems.[4,7] Bimetallic Rh−Fe catalysts are known to improve ethanol selectivity while decreasing methane selectivity due to the formation of $Rh\text{-}FeO_x$ interfacial sites.[8] However, due to their low stability under reaction conditions and *in situ* Rh−Fe alloy formation induced by FeO_x reduction *via* hydrogen spillover, the addition of a third metal is required for stable catalytic behavior. Likewise, the impact of Mn promotion on selectivity toward higher alcohols has been well reported.[9−11] Mn is known to remain in an oxidized state under harsh reaction conditions (523−593 K, 3−8 MPa)[1,5] typical of HAS. The modification of Rh with MnO_x has mainly been described as (i) stabilization of isolated Rh^+ sites, (ii) formation of $Rh\text{-}MnO_x$ interfacial sites,[12] (iii) increase of Rh dispersion, or (iv) a combination of these effects.[13] Although the role of Mn as a promoter is not fully elucidated yet, it has been proposed that the formation of bimetallic $Rh\text{-}MnO_x$ sites

Received: September 3, 2022
Revised: October 28, 2022

https://doi.org/10.1021/acscatal.2c04349
ACS Catal. 2022, 12, 15373−15385

Hypothesis for catalyst design – Chain of thought

What is good catalyst

Shows higher activity and selectivity

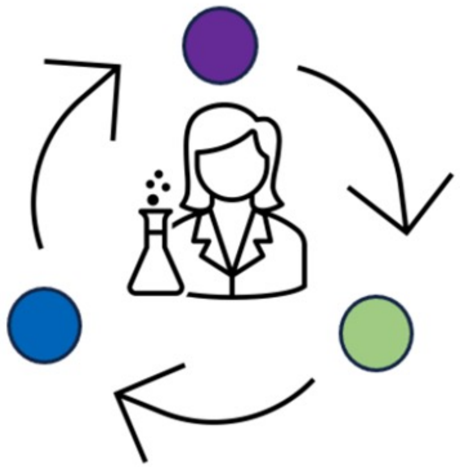
What is the descriptor for catalytic activity

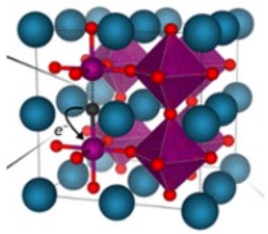
Adsorption energy of reactants and intermediates on catalyst surface

How do I tune adsorption energy??

Let's think step by step…..

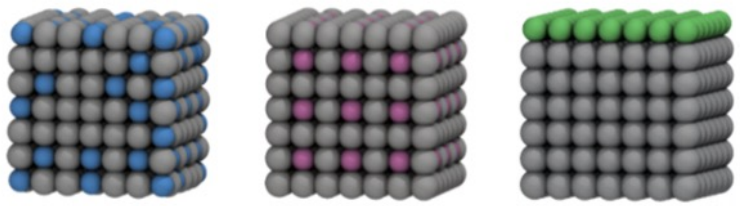
Reasoning/thinking via catalyst descriptor





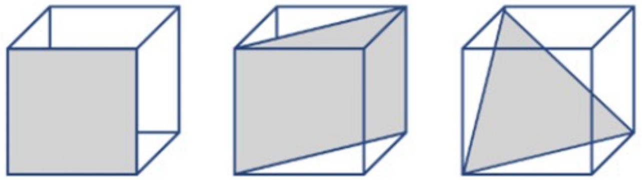
Reason-via-electronic structure

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



Reason-via-catalyst composition

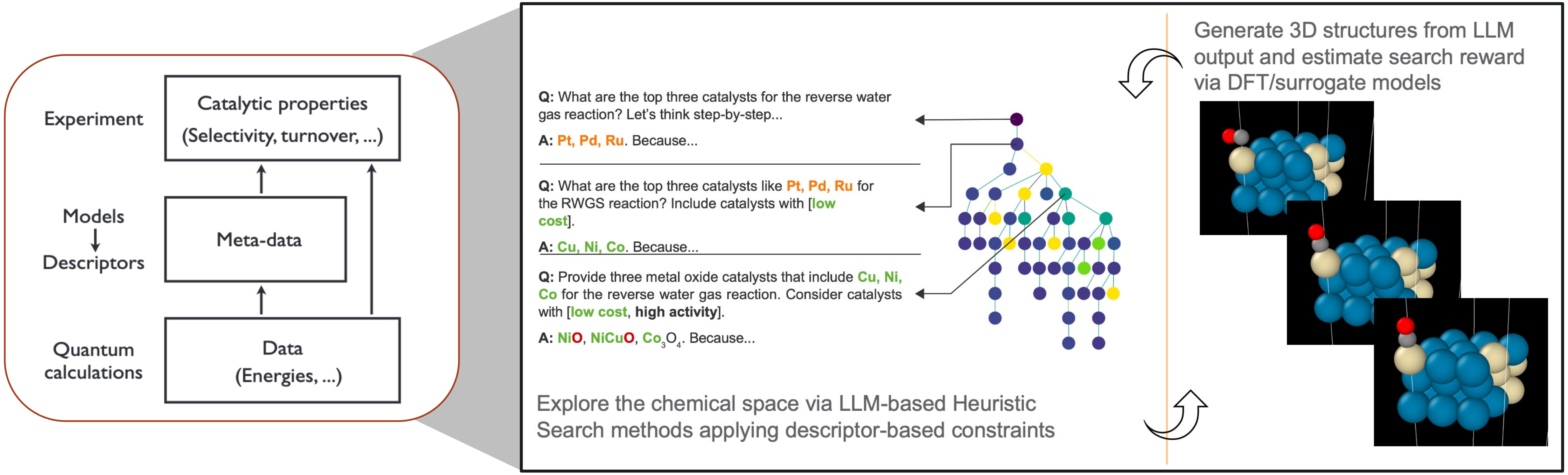
Q + “Let’s think in terms of bulk structural properties“



Reason-via-target properties

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

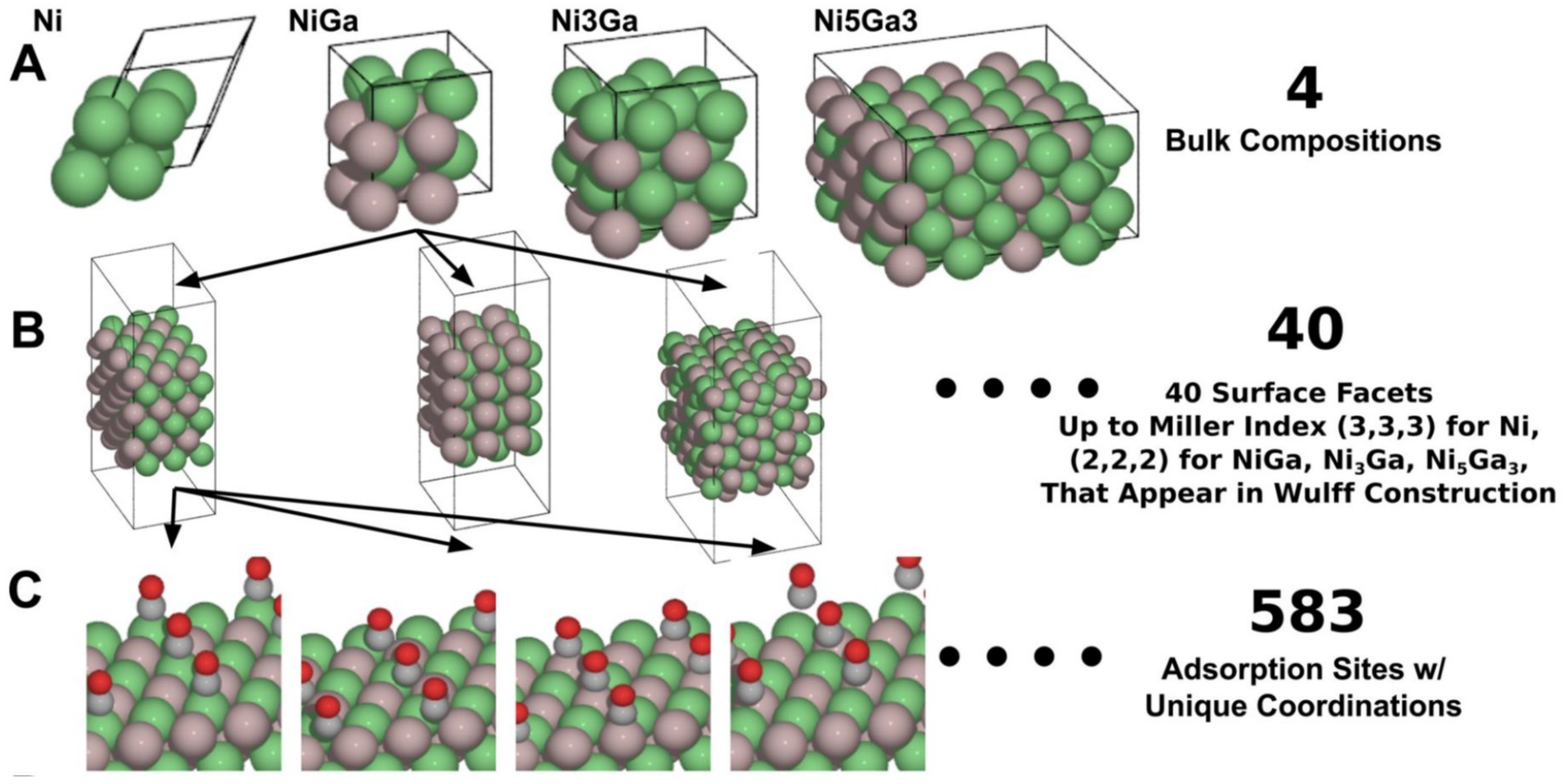
The vision from Catalysis community



The original paper [1] expressed the vision for searching over a "hypothetical database" that stored performance metrics for any group of chemical descriptors. Promising candidates returned by such a search would be further screened by computational chemistry simulations and experiments in the laboratory.

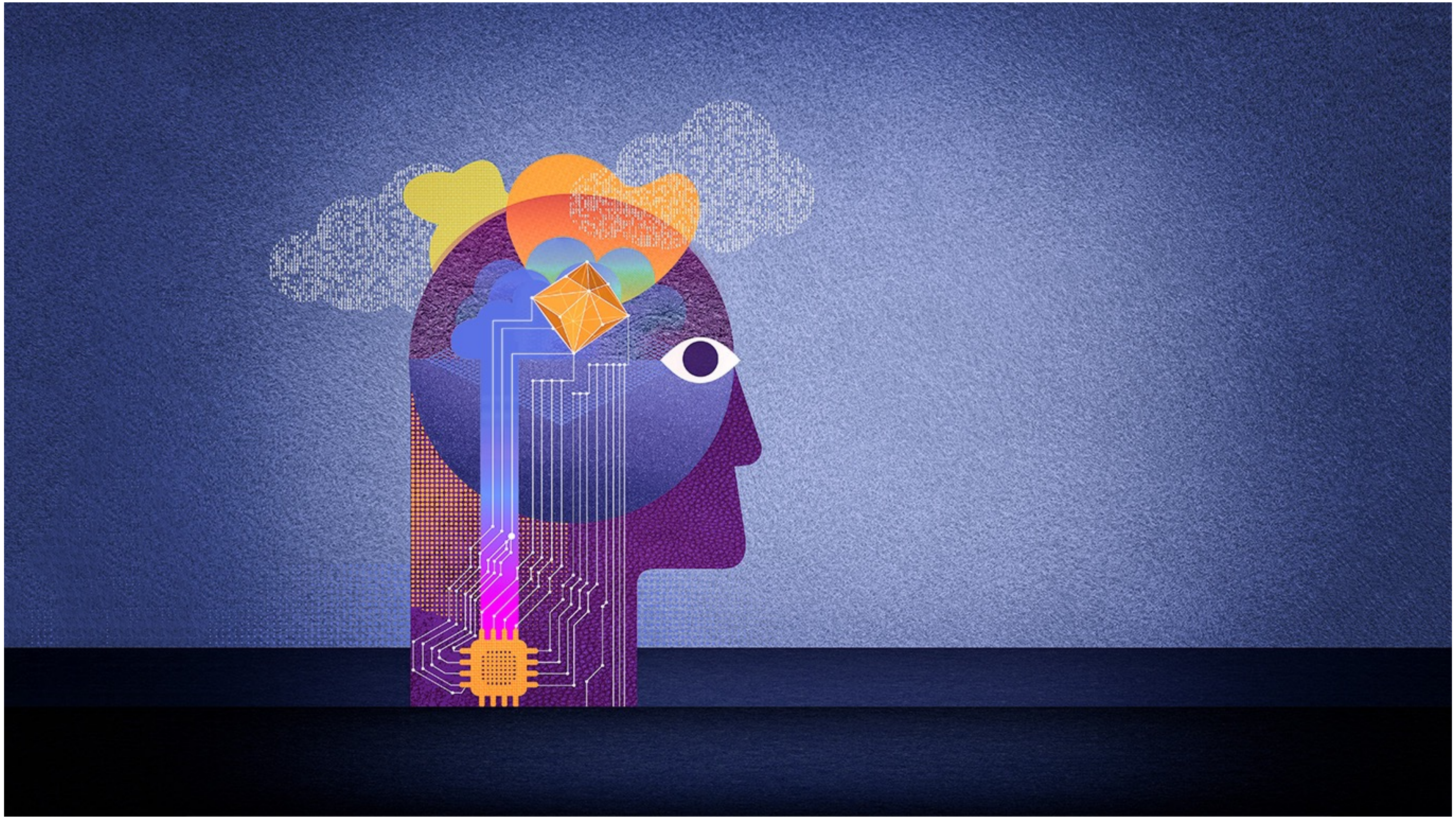
[1] Nørskov, J.K., Abild-Pedersen, F., Studt, F. and Bligaard, T., 2011. Density functional theory in surface chemistry and catalysis. *Proceedings of the National Academy of Sciences*, *108*(3)

Examples of Automated Screening from Literature



Reproduced from: Ulissi, Z. W.; Tang, M. T.; Xiao, J.; Liu, X.; Torelli, D. A.; Karamad, M.; Cummins, K.; Hahn, C.; Lewis, N. S.; Jaramillo, T. F.; Chan, K.; Nørskov, J. K. **Machine-Learning Methods Enable Exhaustive Searches for Active Bimetallic Facets and Reveal Active Site Motifs for CO2 Reduction**. ACS Catal. 2017, 7, 6600−6608.

Came along … ChatGPT



Pacific Northwest
NATIONAL LABORATORY

How did ChatGPT do?

Chemistry experts in the laboratory said its answer looks like taken out of a Science Encyclopedia

Lacked specificity and **explanations grounded in deep scientific principles** chemistry experts demand

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.

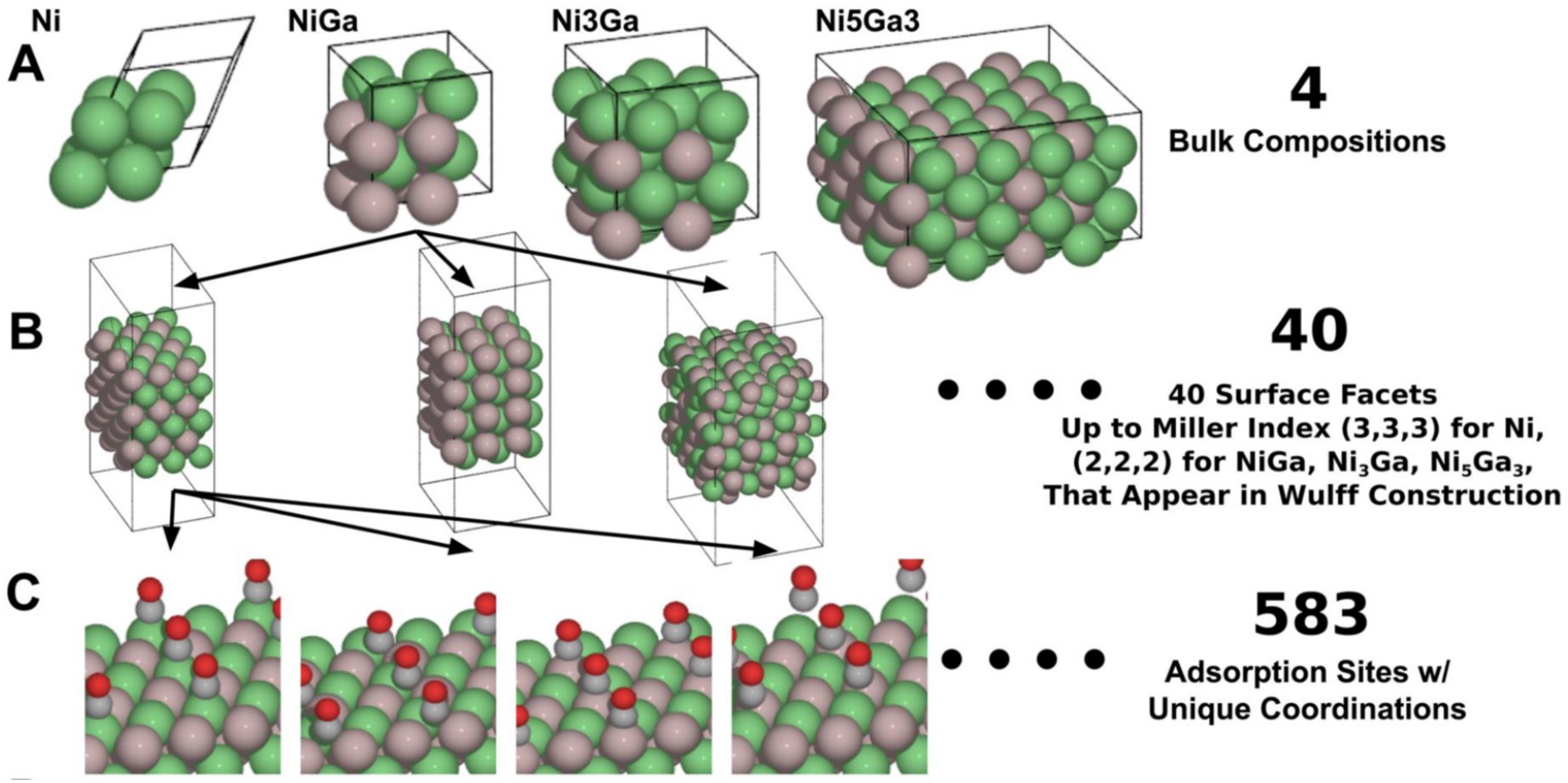
provide more details

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.

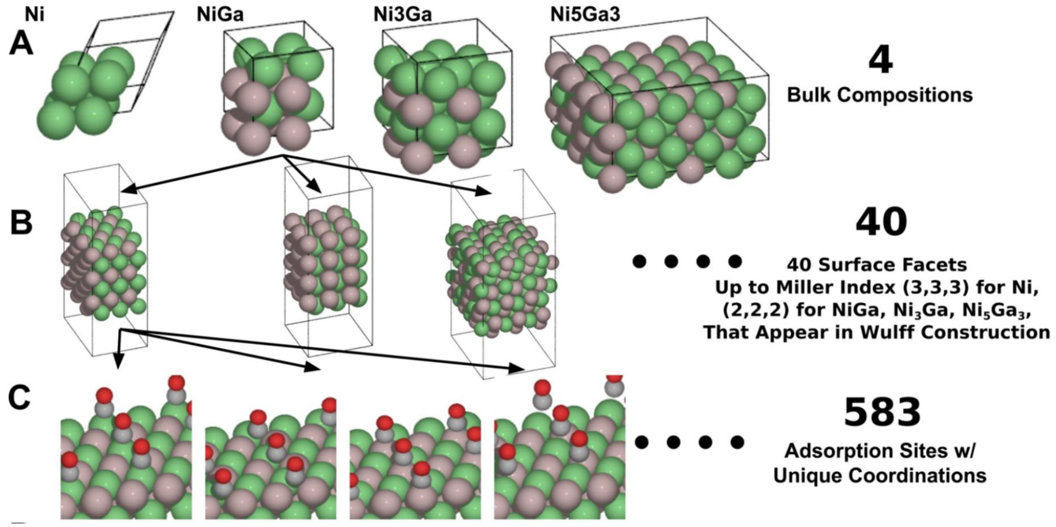
LLM-driven Computational Screening

Revisiting the idea for Computational Screening



Reproduced from: Ulissi, Z. W.; Tang, M. T.; Xiao, J.; Liu, X.; Torelli, D. A.; Karamad, M.; Cummins, K.; Hahn, C.; Lewis, N. S.; Jaramillo, T. F.; Chan, K.; Nørskov, J. K. **Machine-Learning Methods Enable Exhaustive Searches for Active Bimetallic Facets and Reveal Active Site Motifs for CO2 Reduction**. ACS Catal. 2017, 7, 6600−6608.

Our Idea: Expanding the Search by considering a wide range of Chemical Descriptors



Why limit ourselves to a few options?

Action Type	Possible Values	# possible
Add Include Prop.	high activity, high selectivity, high stability, novelty, low cost, low toxicity, high surface area, high porosity, crystal facet, availability	11
Add Exclude Prop.	low activity, low selectivity, low stability, high cost, high toxicity, low dispersion, low porosity, high scarcity	9
Change Catalyst Type	unary catalyst, binary catalyst, trinary catalyst, catalyst	4
Toggle Oxide	on/off	1
Change Relation to Prev. Answer	including elements that are different from, including elements similar to, introducing new elements to, including elements from	4

Let's explore a wider chemical space using a larger set of descriptors!

Our novelty: Generative AI receiving feedback from Quantum Chemistry

Q: What are the top three catalysts for the reverse water gas reaction? Let's think step-by-step...

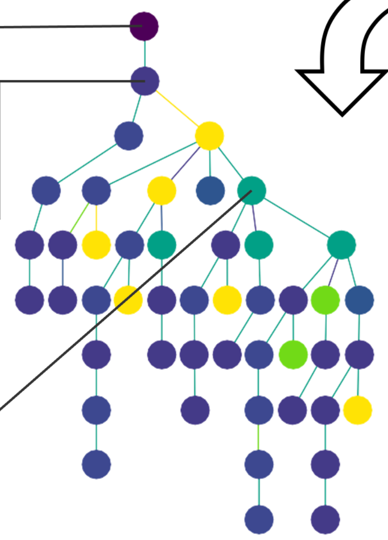
A: **Pt, Pd, Ru**. Because...

Q: What are the top three catalysts like **Pt, Pd, Ru** for the RWGS reaction? Include catalysts with [**low cost**].

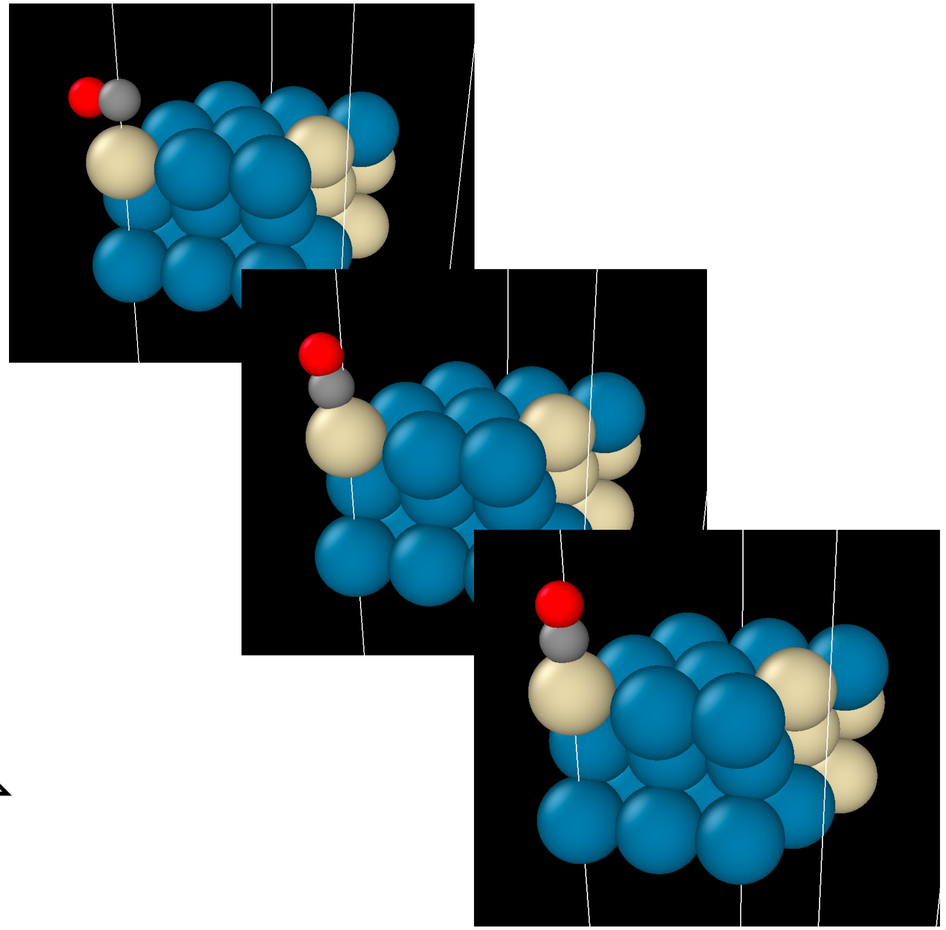
A: **Cu, Ni, Co**. Because...

Q: Provide three metal oxide catalysts that include **Cu, Ni, Co** for the reverse water gas reaction. Consider catalysts with [**low cost**, **high activity**].

A: **NiO**, **NiCuO**, **Co_3O_4**. Because...



Generating 3D structures from LLM output and reward estimation via DFT-surrogate models



Explore the chemical space via LLM-based Heuristic Search methods

ChemReasoner's Exploration of Chemical Space

Instruction: Generate a list of top-5 metallic catalysts for the adsorption of *CO

["Platinum (Pt)", "Palladium (Pd)", "Gold (Au)", "Nickel (Ni)", "Copper (Cu)]

Instruction: **search within** transition metals

Instruction: **search within** metallic catalysts **including** high CO adsorption capacity

Instruction: **search within** metallic catalysts **excluding** poor CO selectivity

['Rhodium (Rh)', 'Ruthenium (Ru)', 'Iron (Fe)', 'Silver (Ag)', 'Iridium (Ir)']

Instruction: **filter candidates with** low stability

Instruction: **constrain with** high resistance to CO poisoning

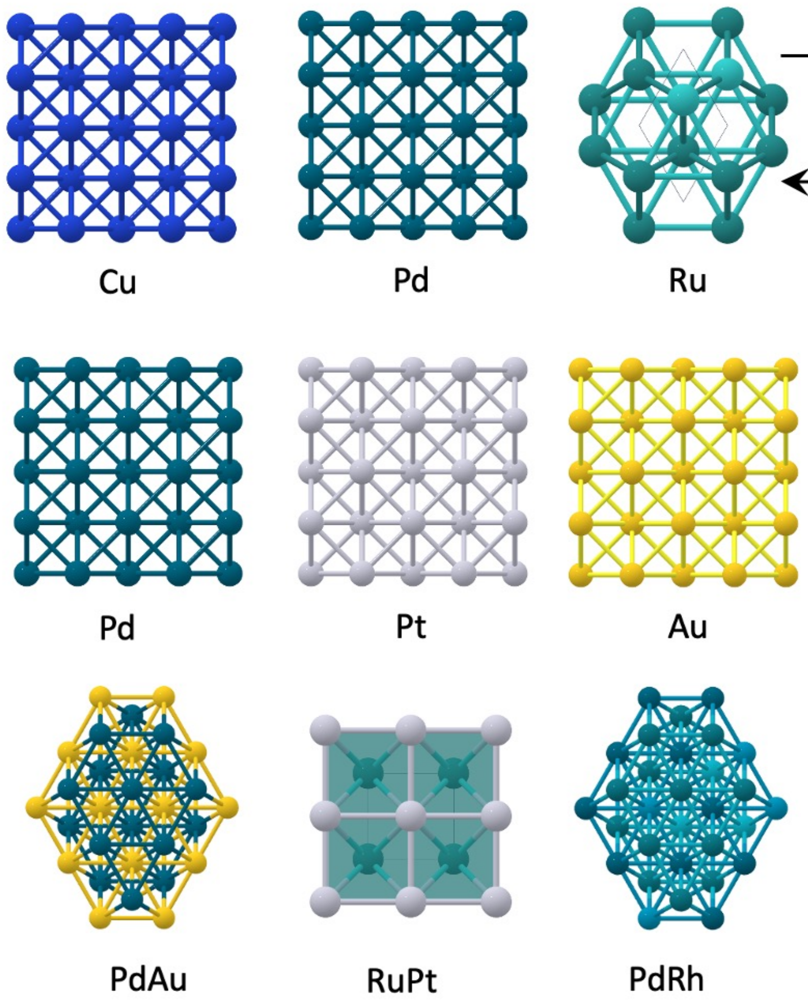
Instruction: **filter candidates with** low stability, weak interaction with CO

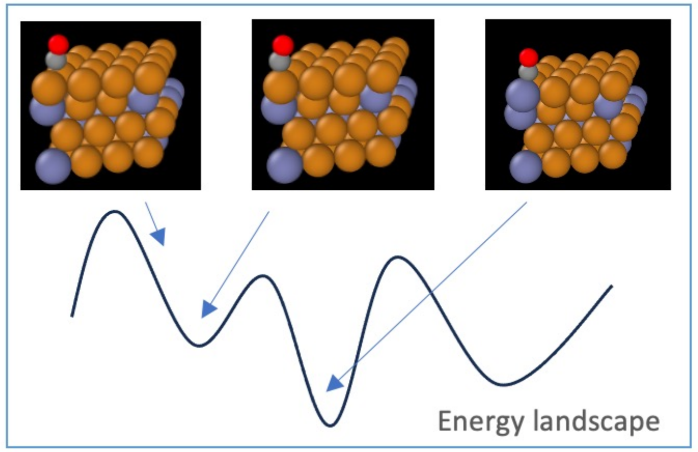
How about we build a new AI system?

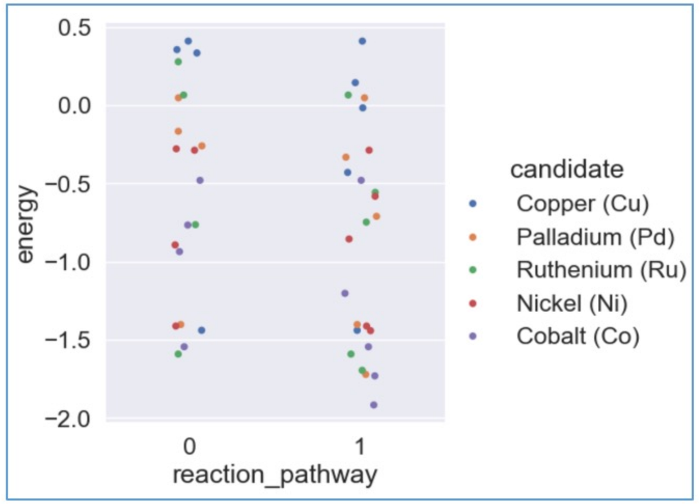
```
{
  Catalyst Type:  ["metallic catalysts"]
  Inclusion Criteria: ["commercial viability",
       "high resistance to CO poisoning"]
  Exclusion Criteria: []
}
```

```
{
  Catalyst Type:  ["metallic catalysts"]
  Inclusion Criteria: ["commercial viability",
       "high resistance to CO poisoning"]
  Exclusion Criteria: [“low catalytic activity for the
       CO2 to methanol conversion reaction”]
}
```

```
{
  Catalyst Type:  [”noble metallic catalysts"]
  Inclusion Criteria: ["commercial viability",
       "high resistance to CO poisoning"",
       "ability to dissociate CO2 into CO and O"]
  Exclusion Criteria: [“low catalytic activity for the
       CO2 to methanol conversion reaction”]
}
```







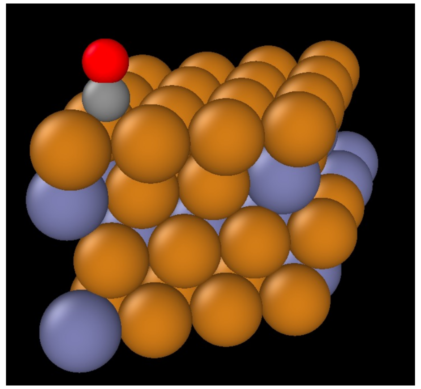
That brings the best of ChatGPT and Computational Chemistry?

3D Atomistic Structure-driven Reward Computation and Screening

Validation of Candidates – Concept of Reward Function

- The language model driven search returns strings in English
- We need a numeric measure (reward) to determine which candidates to prune and expand promising ones

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

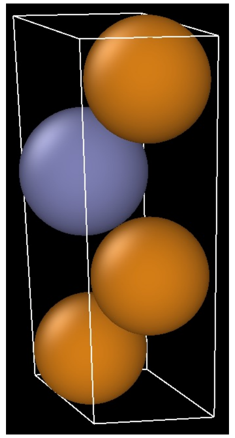


Adsorption Energy

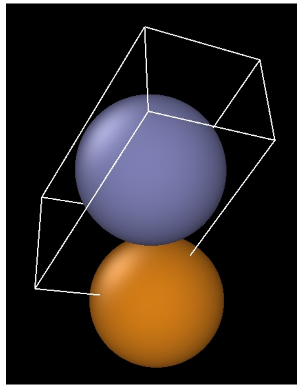
Reward

More sophisticated reward functions are explored later

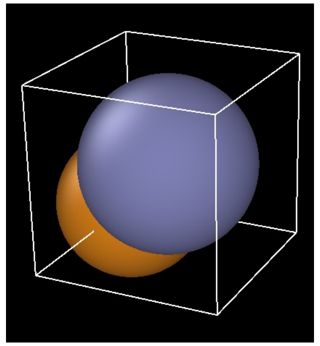
Choosing the bulk structure



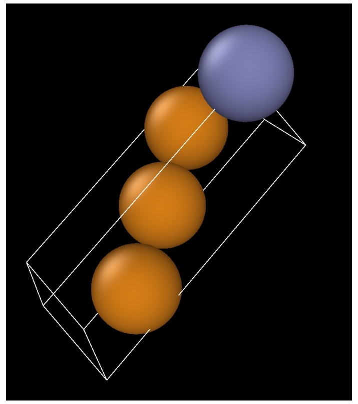
Cu_3Zn



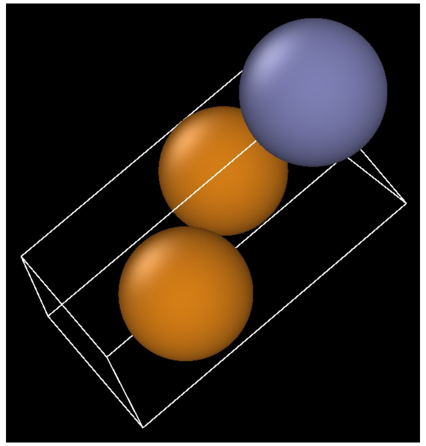
Cu_3Zn



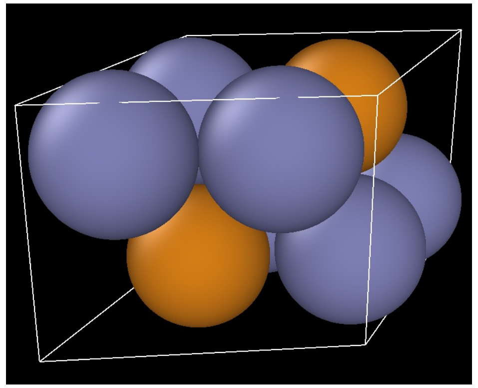
Cu_3Zn



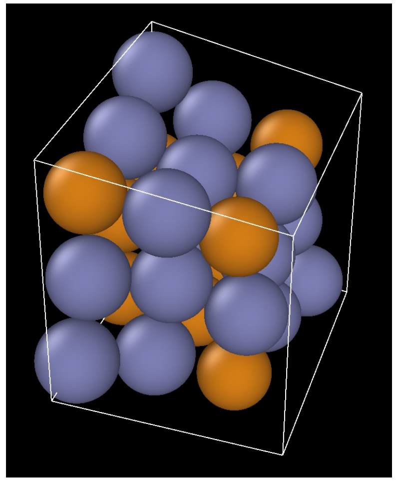
Cu_3Zn



Cu_2Zn



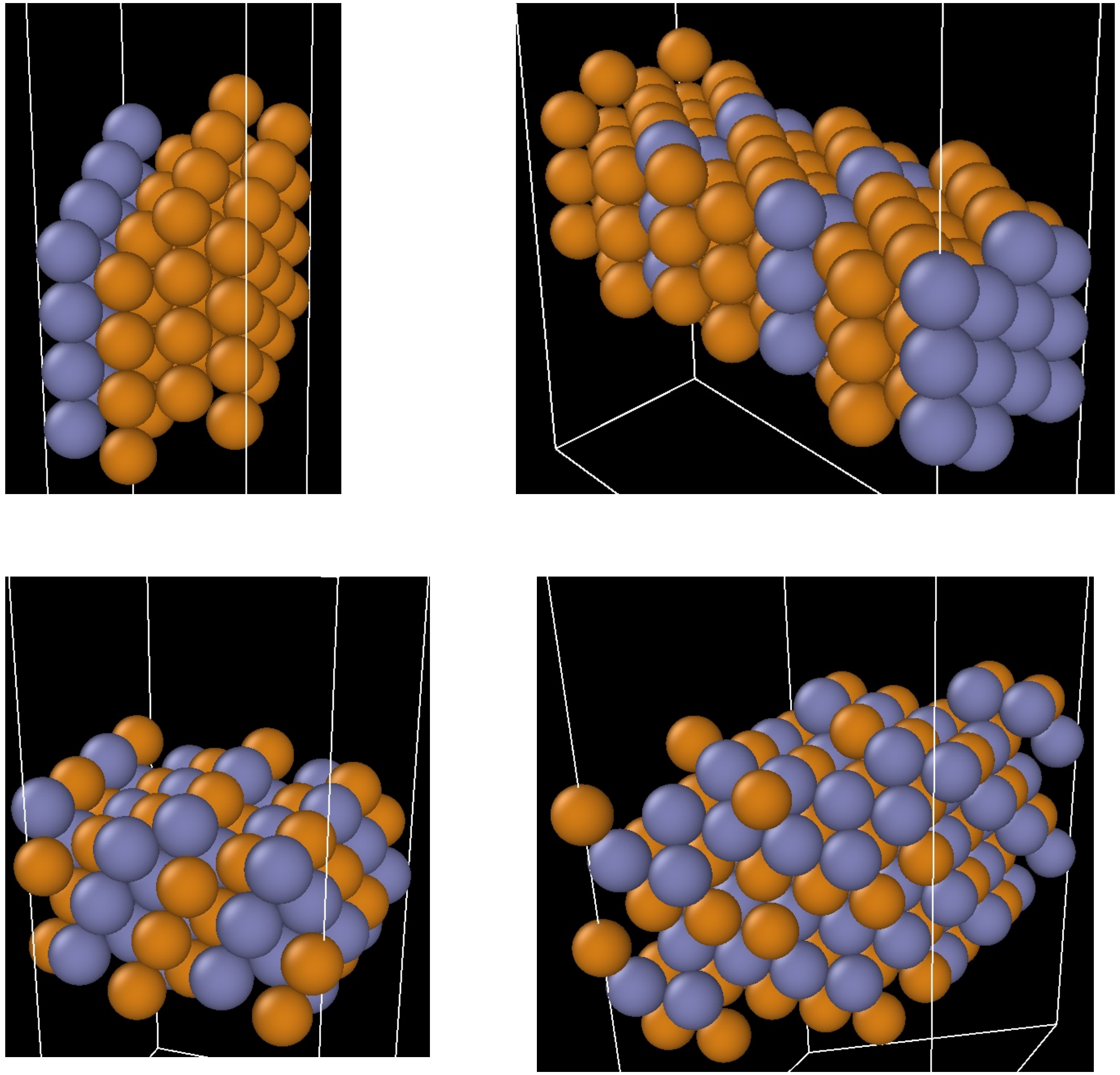
Cu_2Zn_6



$Cu_{10}Zn_{16}$

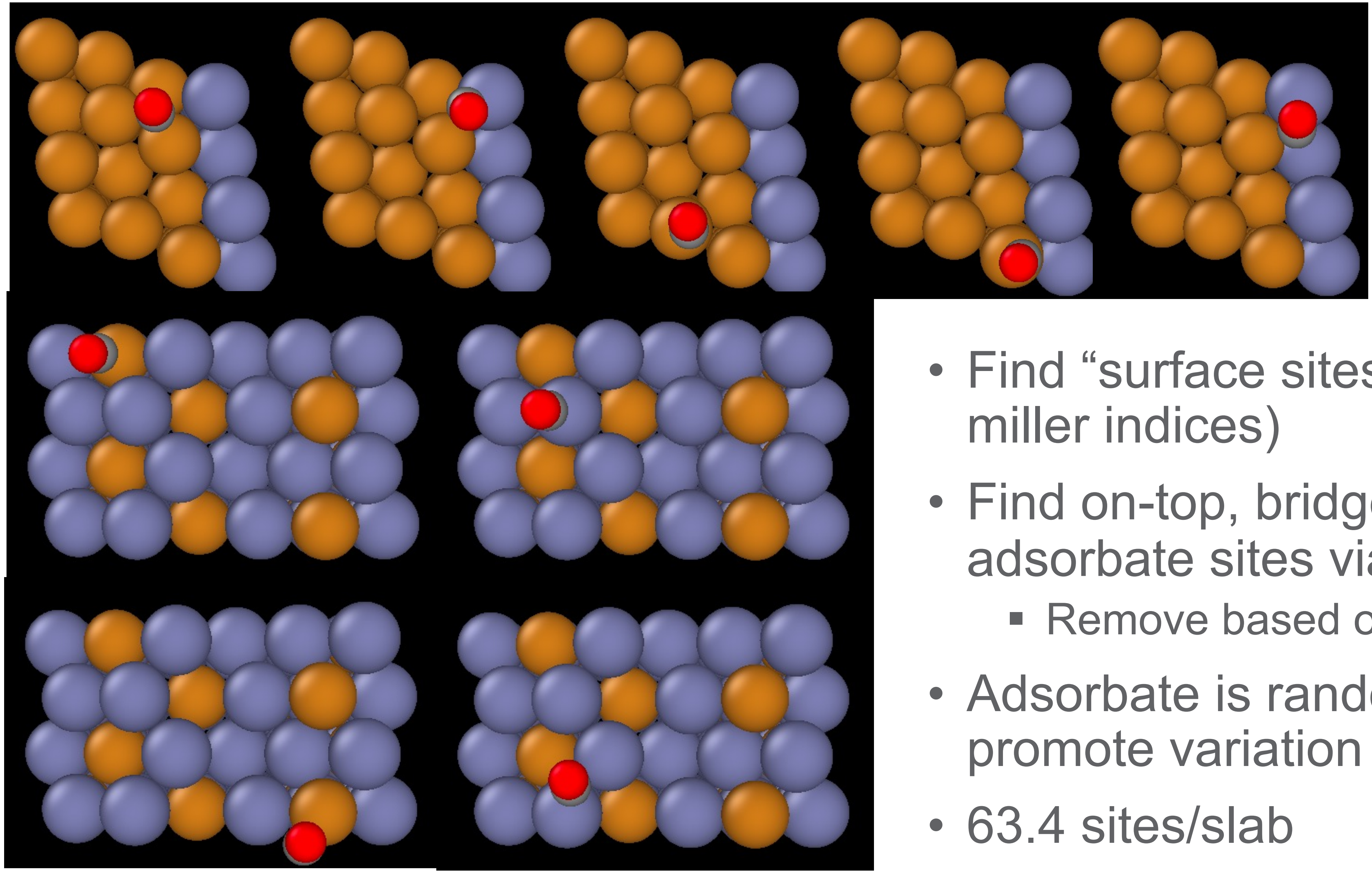
7 Different Compositions and structures of Cu-Zn
Taken from Materials Project Database

Sampling slab configurations



- Sample slab configurations for each bulk:
  - Miller indices from 0 to 2
  - Unique Cell shifts
  - 36.7 slab/bulk
- Following Methods of Open Catalyst Project https://github.com/Open-Catalyst-Project/ocp

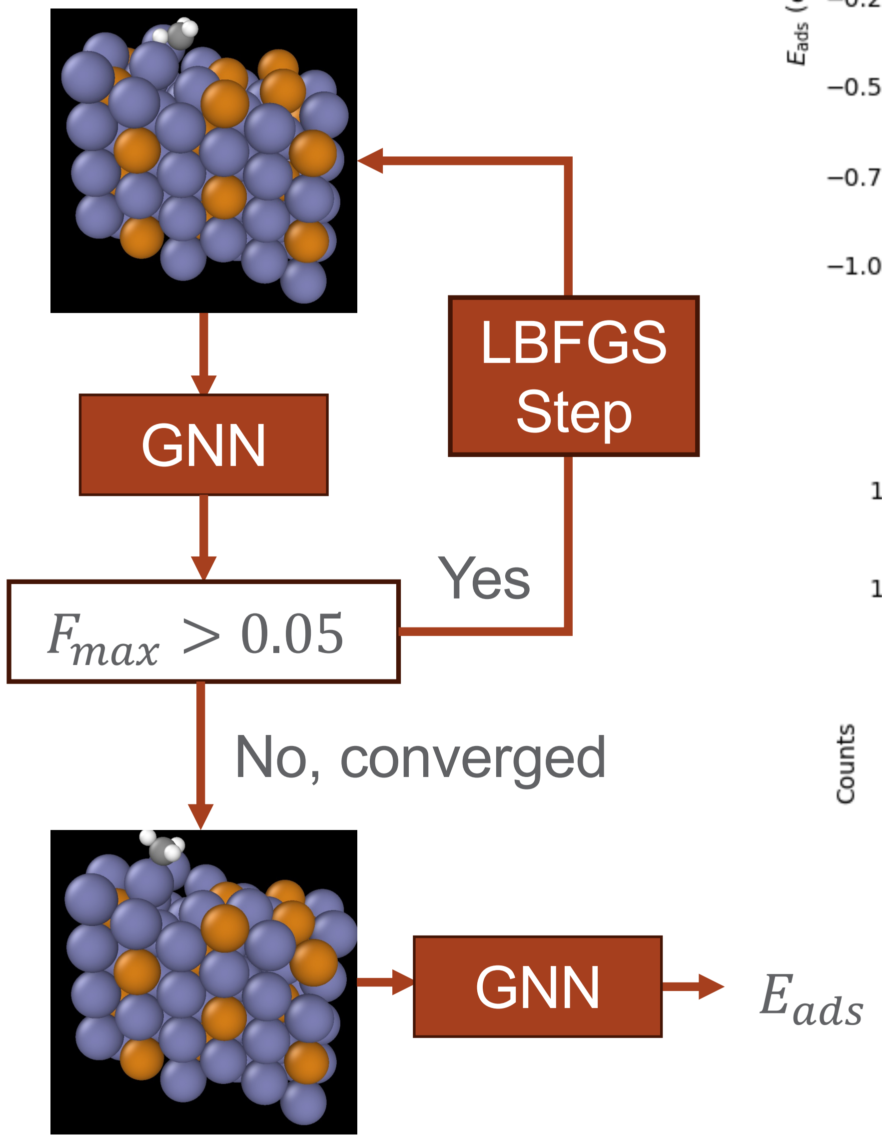
Heuristic placement of *CO (pymatgen)

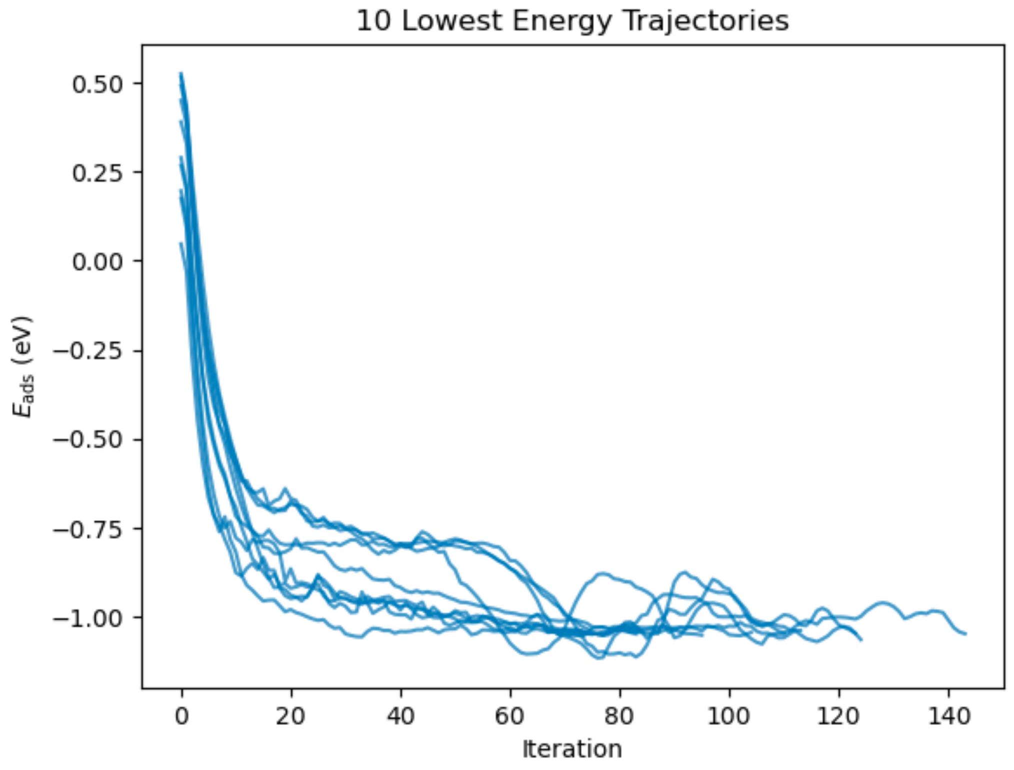


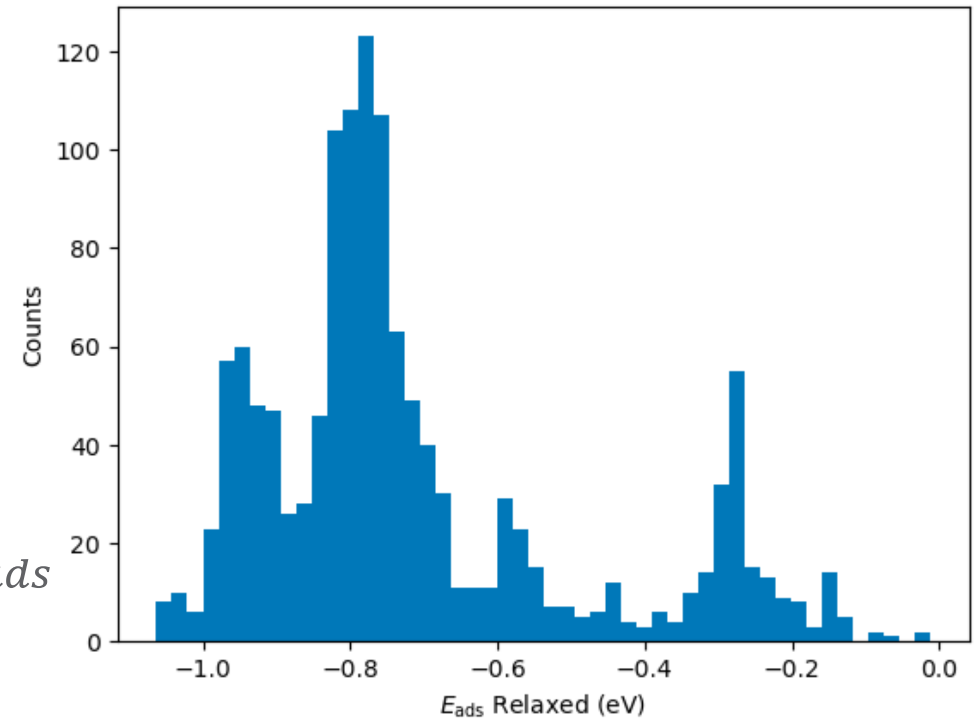
- Find "surface sites" (sites close to miller indices)
- Find on-top, bridge, hollow adsorbate sites via triangulation
  - Remove based on symmetry
- Adsorbate is randomly rotated to promote variation
- 63.4 sites/slab

Relaxation of *CO w/ GNN

- Given an initial structure, relax atoms according to a GNN trained to predict DFT adsorption energies of catalysts
  - GemNet-dT
- Global minimum of all initial configurations is the adsorption energy for CO/CuZn
- 300 steps, fmax 0.05, batch size 20
- ~2 minutes/batch/GPU
  - V100

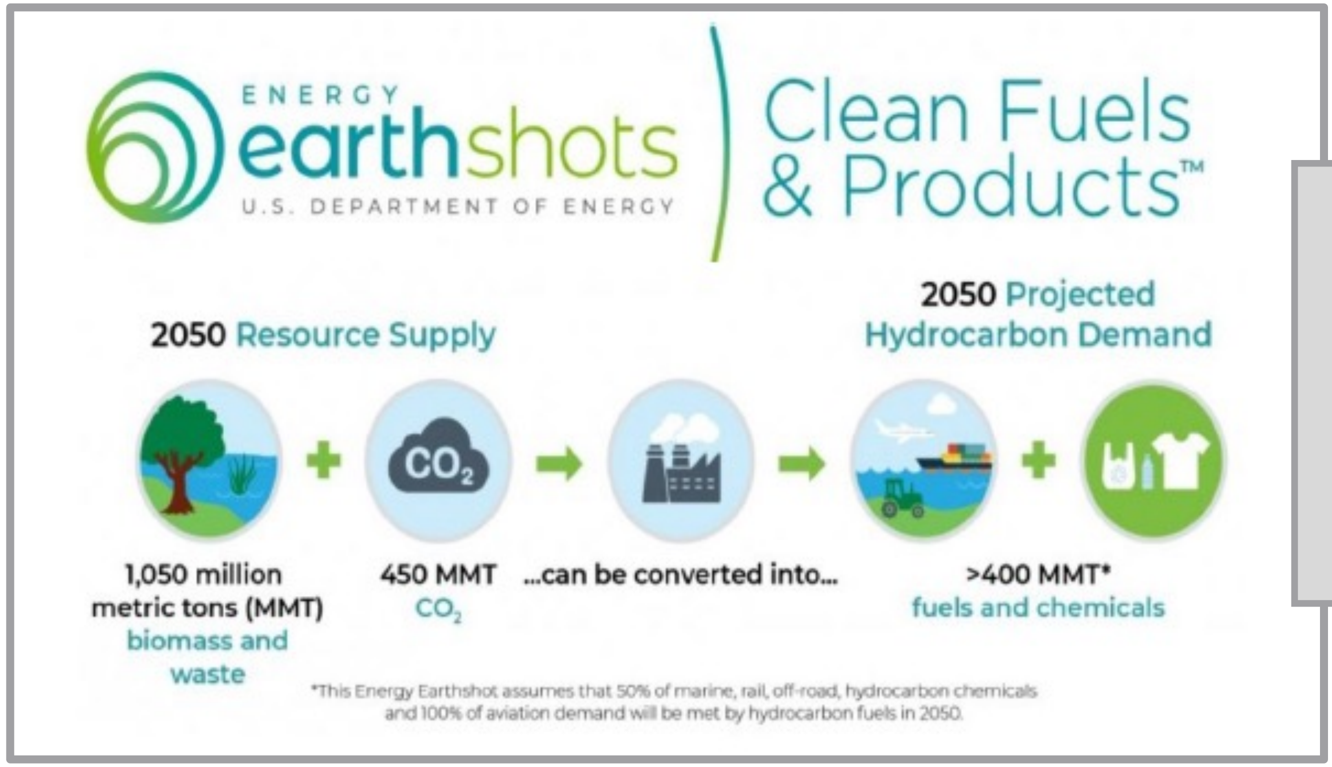






Exploration of Complex Reward Functions

Design of catalyst and catalytic processes are key to achieve net-zero target

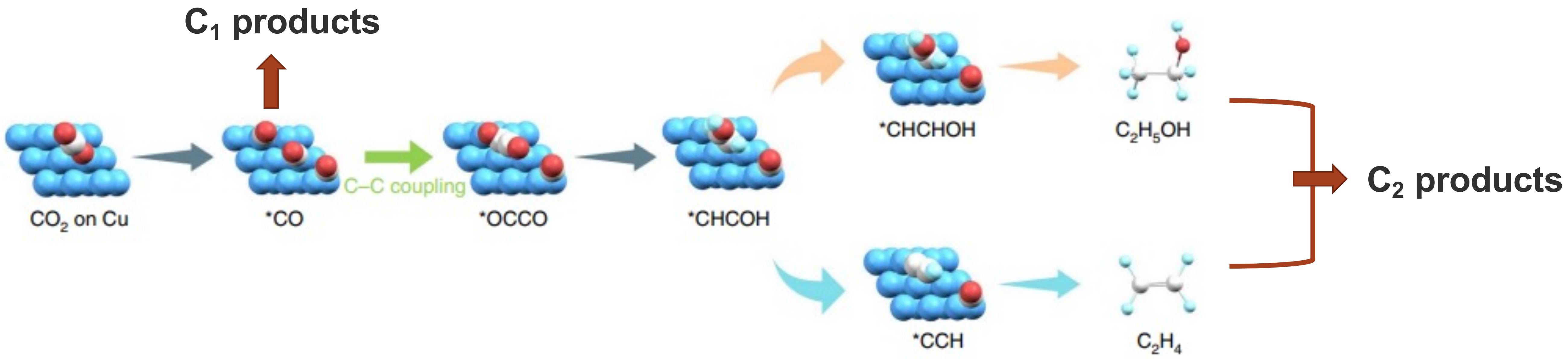


Multifunctional catalyst

Catalyze complex feedstocks with multiple functional groups, through the coordination between different active sites

- Active sites
- Overall reaction pathway
- Intermediate binding energy
- Activation barrier

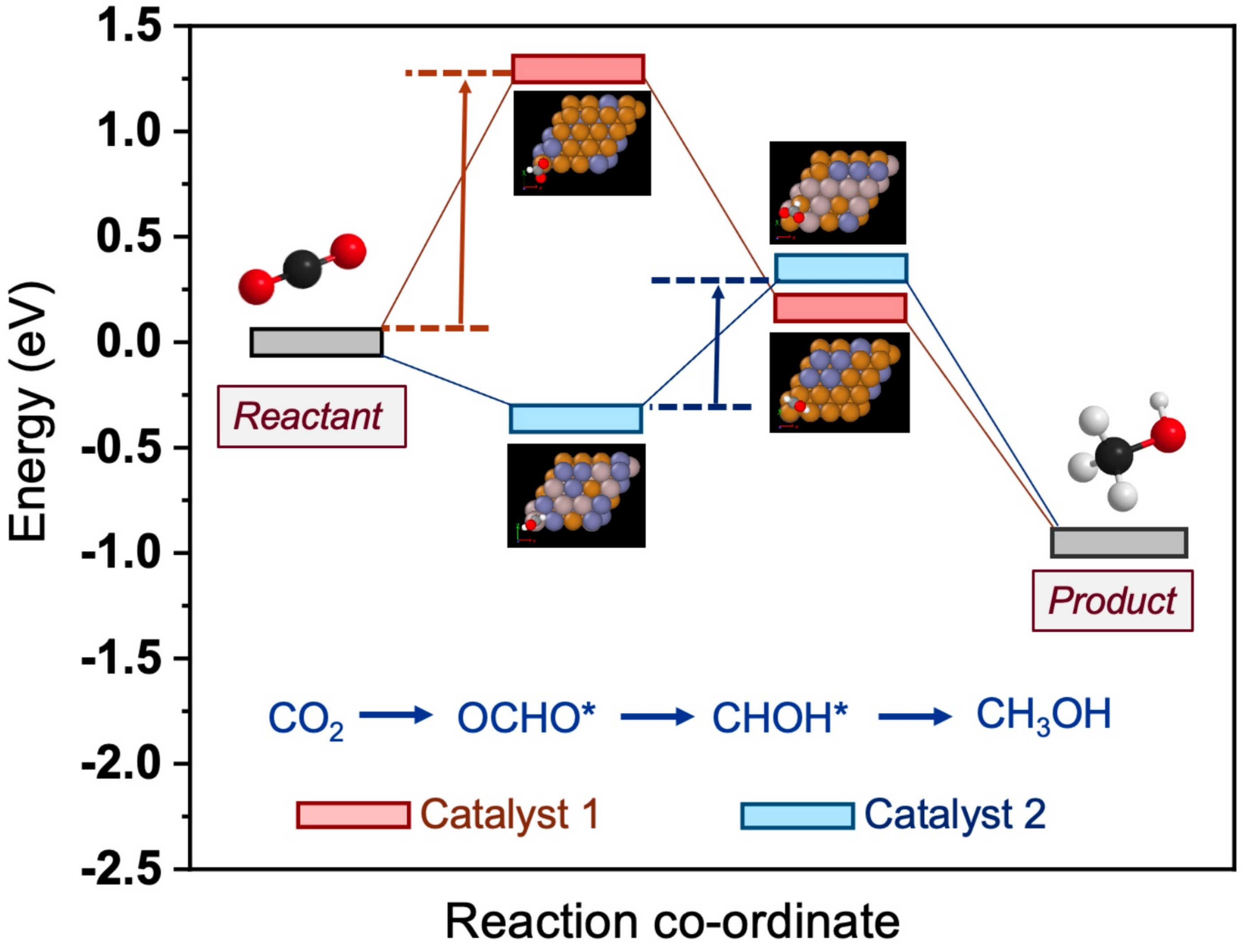
Controlling surface intermediate is key to enhance selectivity



Nature catalysis 2020, 3, 75

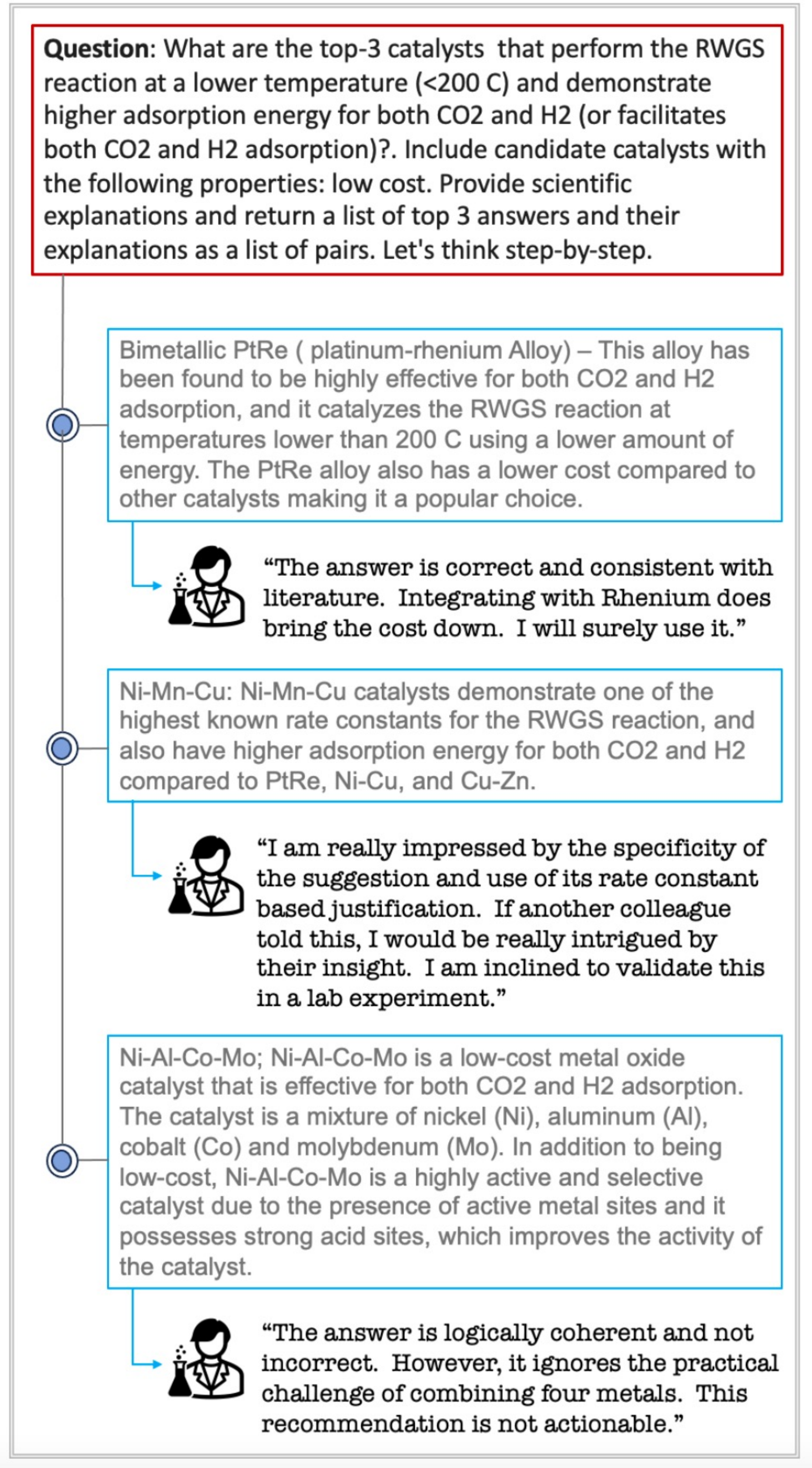
- **Major algorithmic contribution**:
  - Computing the reward function by reasoning in terms of intermediate steps
  - We query the LLM to consider different possible reaction pathways (such as above) and corresponding surface intermediates
  - Then compute a single score based on all the possibilities and their activation barrier

Formulation of Reactivity-focused Reward Functions



$$r(c) = -\min_{p \in \text{Paths}} \left(\max_{\text{ads}_t \in p} \left(E_{\text{ads}_t} - E_{\text{ads}_{t-1}} \right) \right)$$

Quantitative and Qualitative Validation



We had catalysis researchers at PNNL compare ChemReasoner outputs with ChatGPT

ChemReasoner was better.

Recent studies shows it also outperformed GPT-4.

This is expected because modern AI systems such as GPT-4 are not designed to reason at atomistic scale

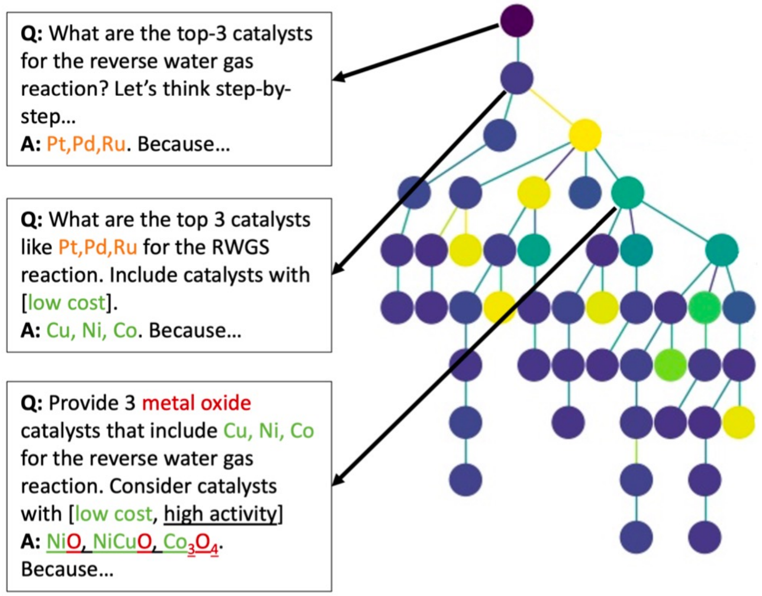
If anyone (OpenAI, Google etc.) developed a better Large Language Model, ChemReasoner should get even better. **It complements and enhances them – does not complete.**

Pacific Northwest
NATIONAL LABORATORY

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)	Noble metal catalysts such as Pt, Rh, Pd, Ru supported on reversible metal oxide i.e., CeO2 (cerium oxide), TiO2 (Titanium dioxide) While noble metals are active for hydrogen adsorption, reversible metal oxide facilitates the CO2 adsorption. The oxygen vacancy present in the reversable 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.	**Adsorption energy** **Electronic structures** **Metal-support interaction**
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 CeO2 or TiO2. 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 CO2 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.	**Electronic structures** **Adsorption on specific crystal planes**
What are the crystal planes that are most active for the adsorption of CO2 for noble metal catalysts?	Open structure such as (100) or (110) planes of metal catalyst are more active towards CO2 activation. Binding energy of CO2 is higher on these crystal surfaces	**Surface structure** **Adsorption on specific crystal planes**

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."

Pacific Northwest
NATIONAL LABORATORY

Qualitative Analysis

Expert-based review of GPT 3.5 and our 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 <200C. 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/SiO_2, Pt/CeO_2 and Na-doped Pt/ZrO_2). I also identified Ni-based (Ni/La-dopedCeO_2, NiCu, Ni/Ce-Zr-O) and Cu-based (4Cu-Al_2O_3) 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.

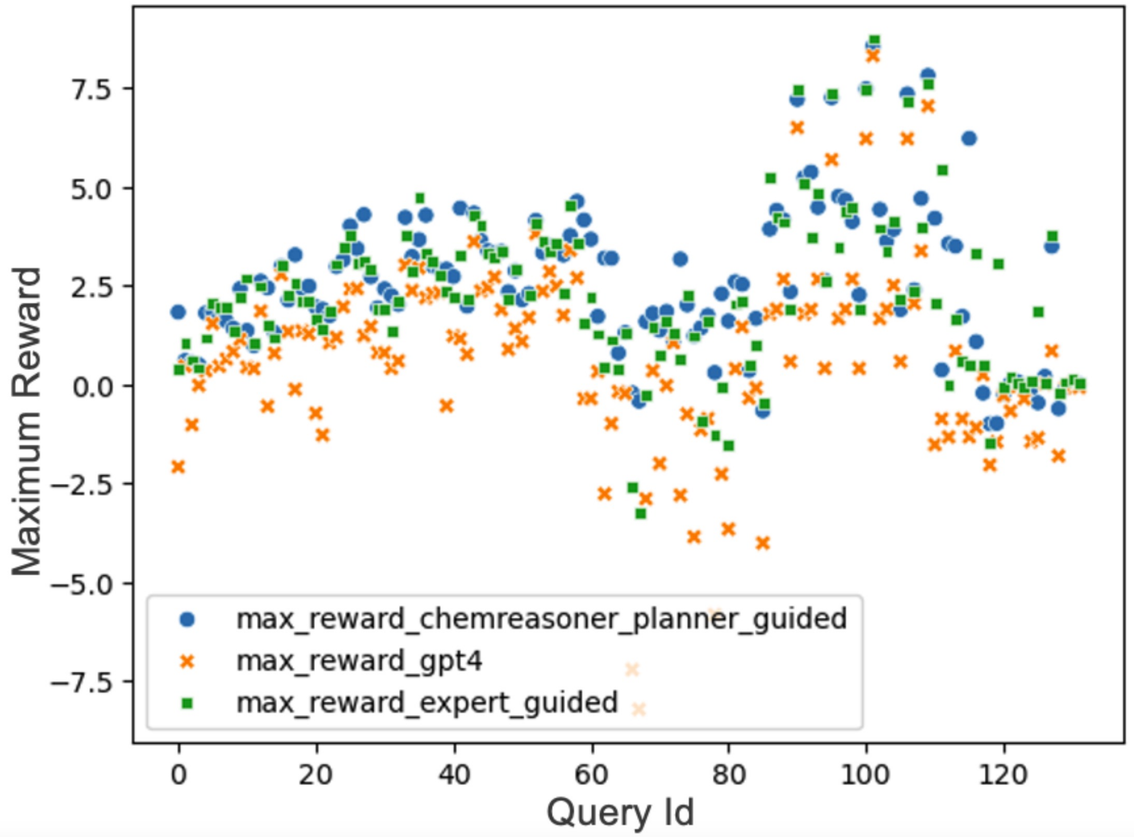
2) Specificity: Which AI method matched the specificity of your explanation?
Answer: Both methods didn't completely match the specificity of my explanation, but I would choose the Monte Carlo Reasoner because it identified Pt, even as part of a bimetallic catalyst. However, even Pt catalysts do not have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

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 <200C. 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.

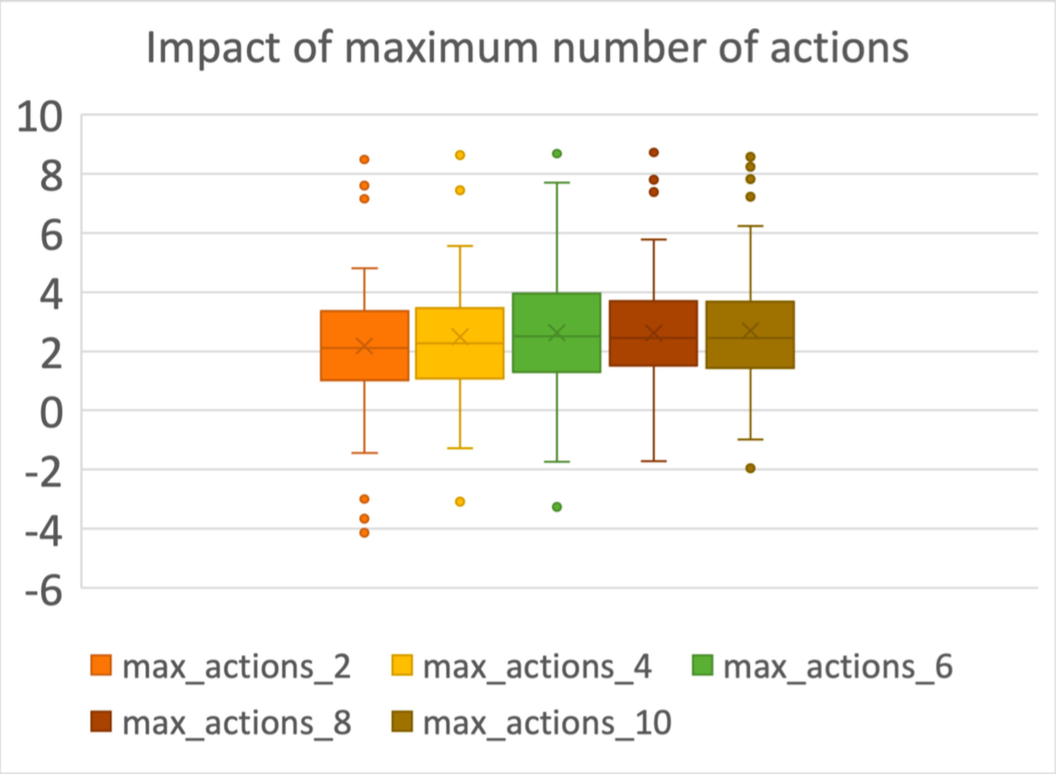
4) Are any of the AI-generated answers novel/superior to the human expert answer?
Answer: The Hafnium suggestion was novel for me, but it was not superior to the human expert answer.

Quantitative Measure



	OpenCatalyst		BioFuels		CO_2-Conversion	
	GPT-4	GPT-3.5	GPT-4	GPT-3.5	GPT-4	GPT-3.5
Chain-of-Thought	0.37	0.66	2.08	2.10	-0.62	-0.54
Self Consistency	0.73	0.76	2.08	2.12	-0.54	-0.36
ChemReasoner-Expert	1.90	2.11	3.90	3.79	0.45	**0.78**
ChemReasoner-Planner	**2.36**	2.16	**4.15**	3.29	0.01	0.49

Modeling Search Efficiency



Method	Category	GPT-3.5-turbo	GPT-4	Search depth reduction (%)
CHEMREASONER-Expert	OpenCatalyst	3.87	3.45	10.97
CHEMREASONER-Expert	BioFuels	3.71	3.50	5.76
CHEMREASONER-Expert	CO_2-Fuel	4.30	3.30	23.25
CHEMREASONER-Planner	OpenCatalyst	3.79	3.67	3.08
CHEMREASONER-Planner	BioFuels	3.5	3.21	8.16
CHEMREASONER-Planner	CO_2-Fuel	4.25	3.55	16.47

- What are the optimal number of actions in the search tree?
- What are the typical search depth for best results?

Inner Monologue: Inspecting the LLM's Rationale

Level 1: Internal monologue of Planner

To address the root question, we need to modify the search state to find metallic catalysts that are effective in the CO2 to methanol conversion reaction. Given the existing state, we need to add inclusion criteria that align with this requirement.

Let's consider the following factors:

1) Catalyst Type: The root prompt asks for metallic catalysts, so we don't need to change the type of catalyst. However, to narrow down our search, we could specify a subclass of metallic catalysts that are known to be effective in this reaction. For example, transition metal catalysts such as copper, zinc, and palladium are known to be active in CO2 to methanol conversion.

2) Inclusion Criteria: Since we're looking for catalysts that exhibit high activity for the CO2 to methanol conversion reaction, we should include this as a criteria. We might also want to include catalysts that are stable under reaction conditions and have a high selectivity towards methanol.

3) Exclusion Criteria: Since we're looking for metallic catalysts, we can exclude non-metallic catalysts. Additionally, we might want to exclude catalysts that degrade quickly or are toxic.

4) Relationship to Candidate List: We're looking for catalysts that are similar to those in our candidate list. This means they should have similar properties and performance in the CO2 to methanol conversion reaction.

Here are my suggestions:

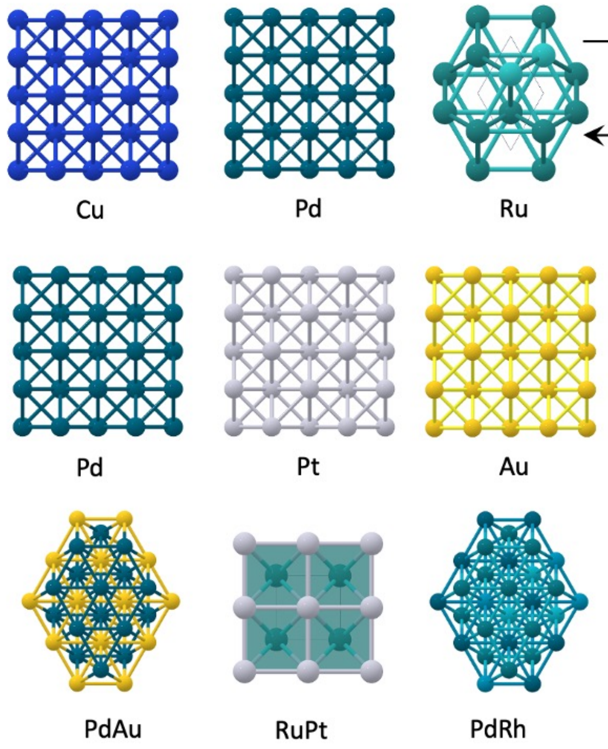
- Catalyst Type: ["transition metal catalysts"] - Inclusion Criteria: ["high activity in CO2 to methanol conversion", "stability under reaction conditions", "high selectivity towards methanol"] - Exclusion Criteria: ["non-metallic catalysts", "quickly degrading catalysts", "toxic catalysts"] - Relationship to Candidate List: ["similar to"]

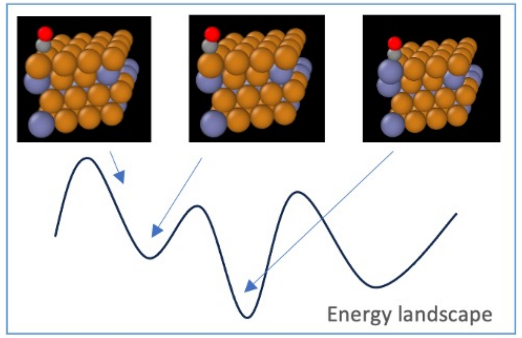
This way, we can find catalysts that are not only similar to those in our candidate list, but also meet the specific requirements of the root question.

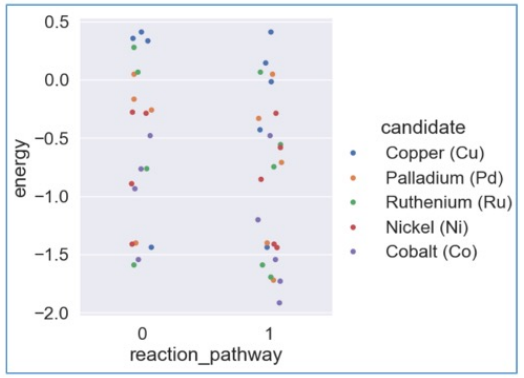
```
{
  Catalyst Type: ["metallic catalysts"]
  Inclusion Criteria: ["commercial viability",
       "high resistance to CO poisoning"]
  Exclusion Criteria: []
}
```

```
{
  Catalyst Type: ["metallic catalysts"]
  Inclusion Criteria: ["commercial viability",
       "high resistance to CO poisoning"]
  Exclusion Criteria: ["low catalytic activity for the
       CO2 to methanol conversion reaction"]
}
```

```
{
  Catalyst Type: ["noble metallic catalysts"]
  Inclusion Criteria: ["commercial viability",
       "high resistance to CO poisoning"",
       "ability to dissociate CO2 into CO and O"]
  Exclusion Criteria: ["low catalytic activity for the
       CO2 to methanol conversion reaction"]
}
```







Scalability

ChemReasoner is Inference-Intensive

Q: What are the top three catalysts for the reverse water gas reaction? Let's think step-by-step...

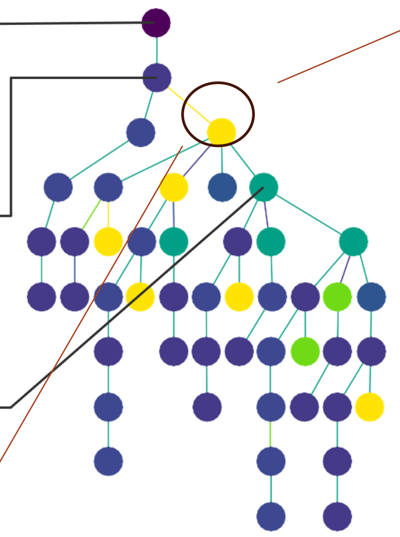
A: Pt, Pd, Ru. Because...

Q: What are the top three catalysts like **Pt, Pd, Ru** for the RWGS reaction? Include catalysts with [**low cost**].

A: Cu, Ni, Co. Because...

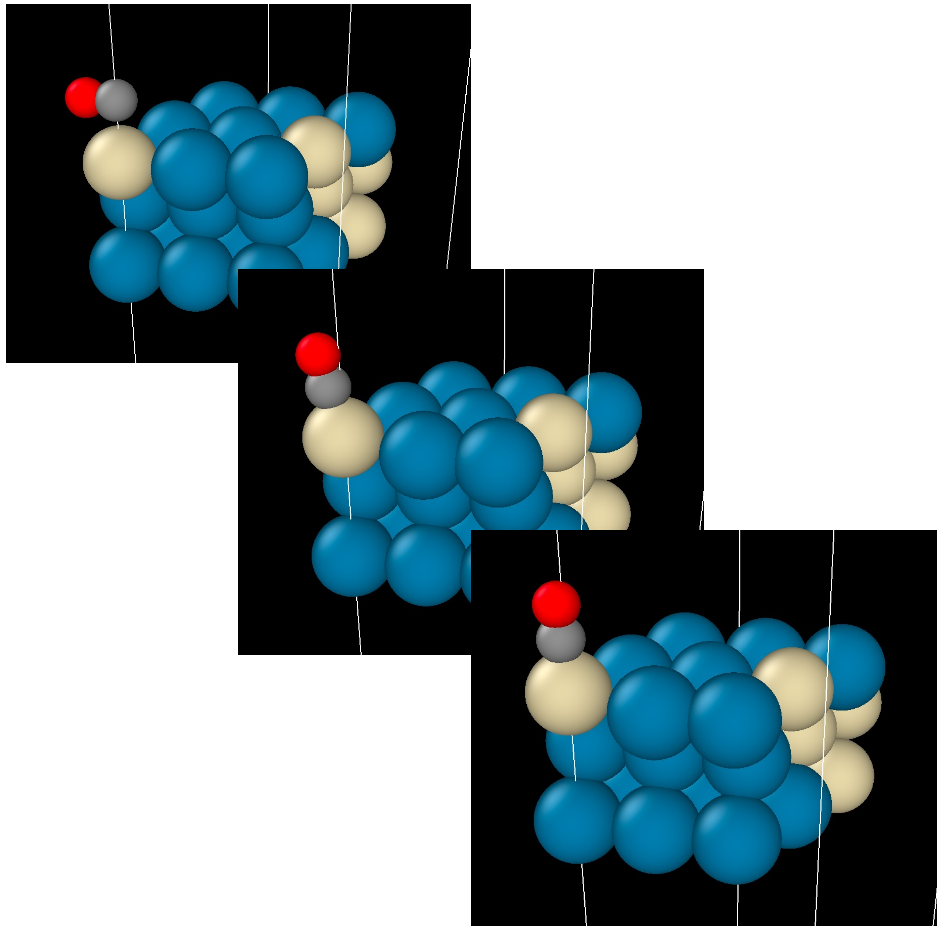
Q: Provide three metal oxide catalysts that include **Cu, Ni, Co** for the reverse water gas reaction. Consider catalysts with [**low cost**, **high activity**].

A: NiO, **NiCuO**, **Co_3O_4**. Because...

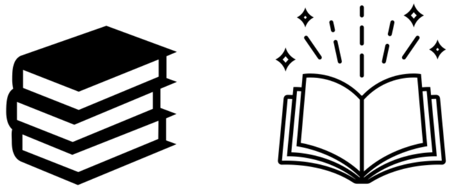


- Each node in the search tree executes 2-3 LLM inferences
- 400-600 inferences per search tree
- We use parallel beam search with asynchronous LLM calls

- Each node in the search tree executes 2400-3200 GNN inferences
- We use caching to avoid duplicate calls

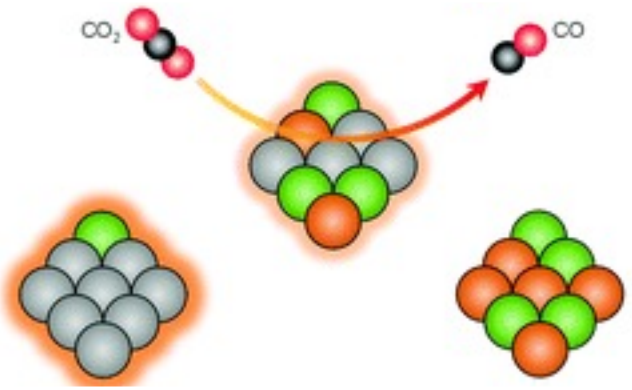


A Single Cycle of Hypothesis Generation and Testing



Hypothesis generation (LLM)

3D Atom Graph Generator



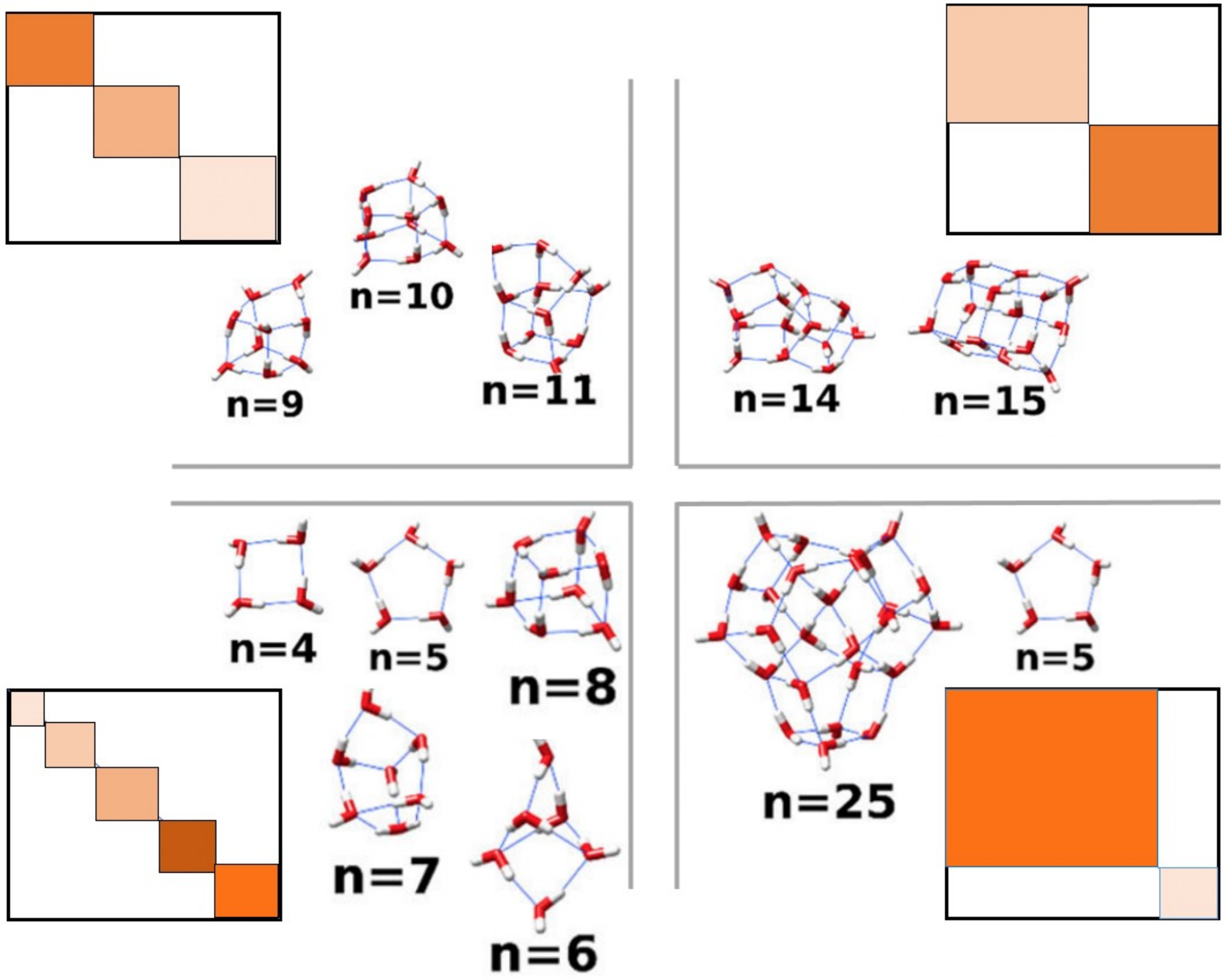
Property prediction with UQ (GNN)

Active Learning

High uncertainty

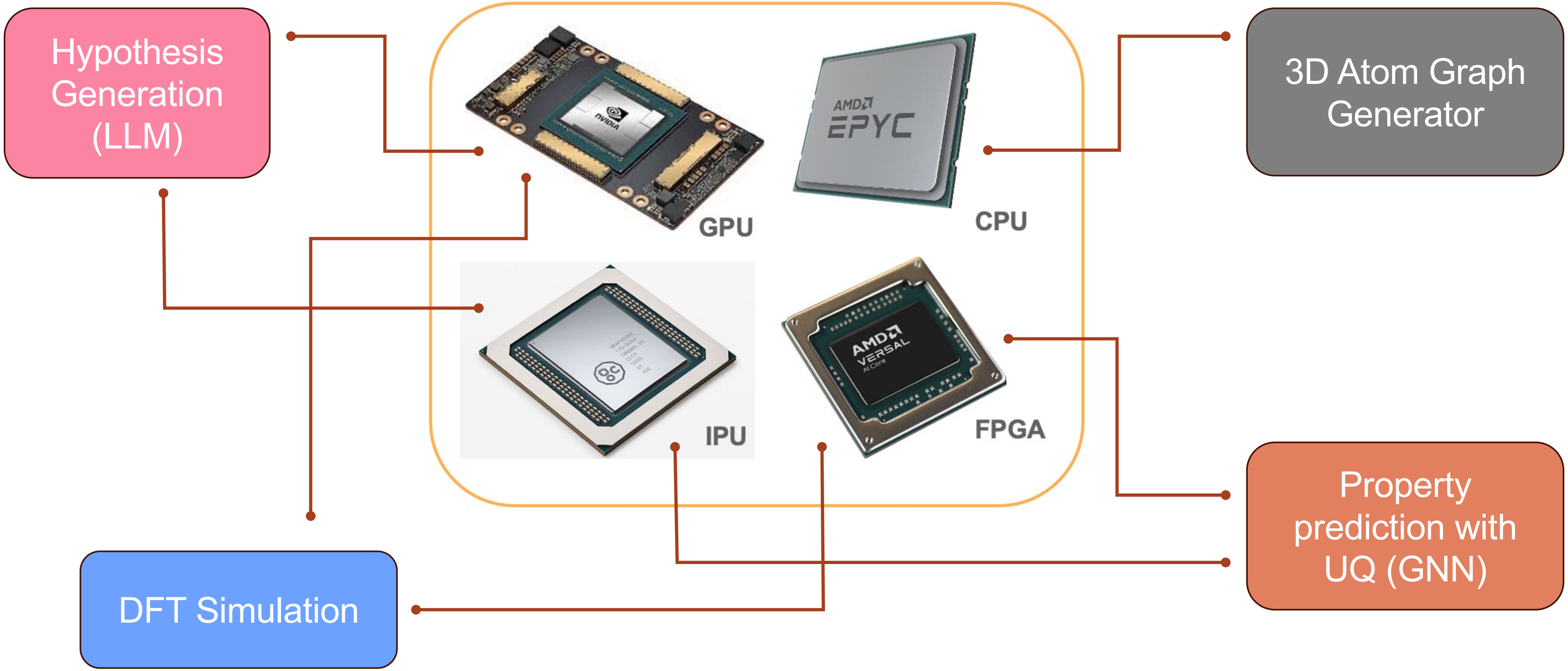
DFT Simulation

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: Heterogeneous Computing-driven Computational Pipeline



Conclusions

- ChemReasoner represents an emerging class of AI systems (Compound AI) that are promising to build on the foundation of Large Language Models
- Initial studies demonstrate that the integration of Generative AI and Computational Chemistry can outperform pure LLM (such as GPT-3.5 and GPT-4) based approaches
- Such systems involve heterogeneous computing workload with distinct scaling characteristics

Pacific Northwest NATIONAL LABORATORY

Acknowledgements

- Accelerating Foundation Models Research award, Microsoft
- Seed LDRD Program, Pacific Northwest National Laboratory
- Generative AI for Science LDRD Program, Pacific Northwest National Laboratory
- https://opencatalystproject.org

Pacific Northwest NATIONAL LABORATORY

- Scaling Atomistic Neural Network Potentials: https://sites.google.com/view/ai4hydronet/home
- ChemReasoner: https://github.com/pnnl/chemreasoner/

- Helal H., J.S. Firoz, J.A. Bilbrey, H.W. Sprueill, K.M. Herman, M.M. Krell, and T. Murray, et al. 2024. "Acceleration of Graph Neural Network-based Prediction Models in Chemistry via Co-design Optimization on Intelligence Processing Units." Journal of Chemical Information and Modeling.
- Sprueill H.W., C. Edwards, U. Sanyal, M.V. Olarte, H. Ji, and S. Choudhury. 2023. "Monte Carlo Thought Search: Large Language Model Querying for Complex Scientific Reasoning in Catalyst Design." In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP2023) Findings.
- Sprueill H.W., C. Edwards, K. Agarwal, M.V. Olarte, U. Sanyal, C. Johnston, and H. Liu, et al. 2024. "CHEMREASONER: Heuristic Search over a Large Language Model’s Knowledge Space using Quantum-Chemical Feedback." https://arxiv.org/abs/2402.10980

Questions?