SCIENTIFIC DATA

OPEN

DATA DESCRIPTOR

Received: 1 February 2019
Accepted: 17 April 2019
Published online: 28 May 2019

High-throughput calculations of catalytic properties of bimetallic alloy surfaces

Osman Mamun[1,2], Kirsten T. Winther[1,2], Jacob R. Boes[1,2] & Thomas Bligaard[1]

A comprehensive database of chemical properties on a vast set of transition metal surfaces has the potential to accelerate the discovery of novel catalytic materials for energy and industrial applications. In this data descriptor, we present such an extensive study of chemisorption properties of important adsorbates - e.g., C, O, N, H, S, CH_x, OH, NH, and SH - on 2,035 bimetallic alloy surfaces in 5 different stoichiometric ratios, i.e., 0%, 25%, 50%, 75%, and 100%. To our knowledge, it is the first systematic study to compile the adsorption properties of such a well-defined, large chemical space of catalytic interest. We propose that a collection of catalytic properties of this magnitude can assist with the development of machine learning enabled surrogate models in theoretical catalysis research to design robust catalysts with high activity for challenging chemical transformations. This database is made publicly available through the platform www.Catalysis-hub.org for easy retrieval of the data for further scientific analysis.

Background & Summary

Electronic structure calculations from Density functional theory (DFT)[1,2] is a well established approach for predicting a large range of material properties[3]. In the field of heterogeneous catalysis and electrocatalysis, DFT has provided a deeper understanding of catalytic activity and reaction mechanisms[4], and has guided the exploration of new catalytic materials. Importantly, the adsorption energy of chemical species to the surface obtained from DFT, has been found to be a strong descriptor for catalytic activity of surfaces[5,6].

Chemical reactions of interest for sustainable energy applications, including the conversion of CO_2 and syngas to carbon-based fuels[7,8], fuel-cell operation[9,10], and electrochemical water splitting[11], noble metals such as Pt, Ru, Ag, Ir and Cu are the most active materials. However, a key challenge for large-scale sustainable energy technologies is to identify catalytic materials that are also of high abundance and low cost. In this search, it is instructive to investigate metal alloys, which span a vast set of materials, with the potential to mimic the catalytic properties of the highly active pure metals. Several bimetallic alloys with high catalytic activity have already been identified, including CoMo[12], BiPt[13] and Pt-lanthanide alloys such as Pt_3Y[14].

Here, we present a large-scale DFT study of chemical adsorption and hydrogenation on 1,998 bimetallic alloy and 37 pure metal surfaces. Consisting of more than 90,000 systematic DFT calculations, this dataset is intended for machine learning model generation. The alloys were chosen by combining 37 selected metals and transition metals (outlined in the periodic table in Fig. 1) to form alloys in the $L1_2$ and $L1_0$ Strukturbericht designation, which corresponds to face-centered cubic crystal structures with A_3B and AB stoichiometries, respectively. The 37 pure metals in the A1 (FCC) structure were included in addition to the 1,998 bimetallic alloys resulting from all possible combinations, such that stoichiometric A:B ratios of 0%, 25%, 50%, 75% and 100% are sampled. The metal surfaces were modeled by cleaving three-layer slabs with a (111) termination for A1 and $L1_2$ and a (101) termination for $L1_0$, although this termination is also referred to as the (111) miller index[15,16] when cleaved from the cubic bulk unit cell which is not the standard conventional form.

Atomic adsorption of H, C, N, O, and S was studied for all 2,035 surfaces. In order to systematically sample the adsorption energies, all unique adsorption sites were considered. The unique sites for each of the surface structures are shown in Fig. 1 where the number of sites are 4, 9, and 10 for the A_1, $L1_2$, and $L1_0$ surfaces respectively.

[1]SUNCAT Center for Interface Science and Catalysis, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, California, 94025, United States. [2]SUNCAT Center for Interface Science and Catalysis, Department of Chemical Engineering, Stanford University, Stanford, California, 94305, United States. Correspondence and requests for materials should be addressed to T.B. (email: bligaard@stanford.edu)

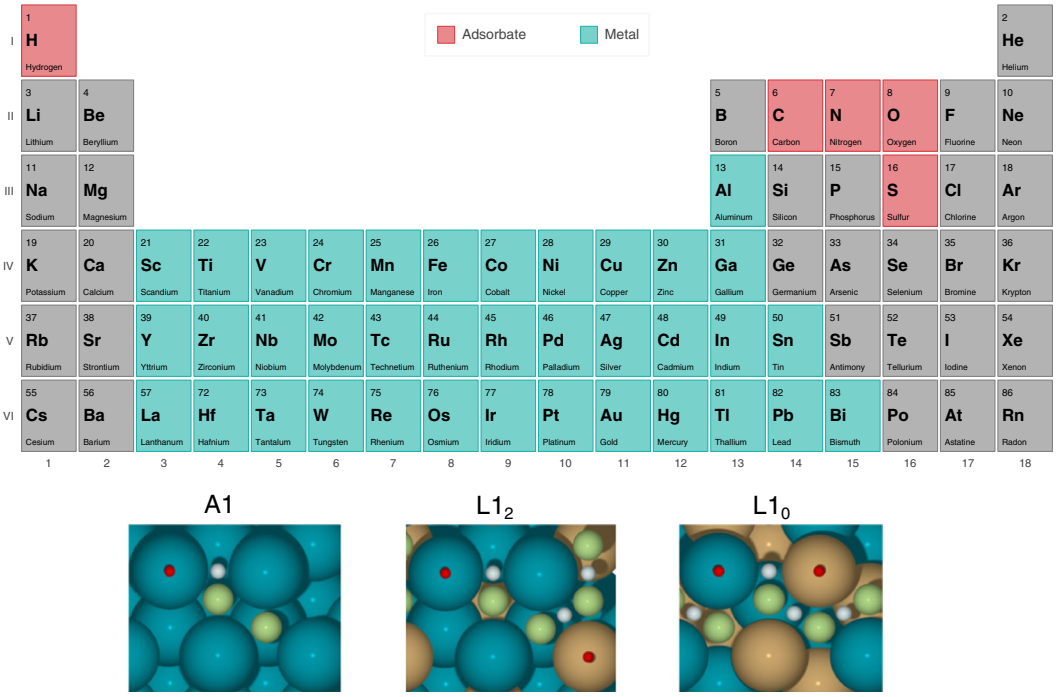

Fig. 1 The periodic table outlining five adsorbate elements and the 37 metals included in the dataset. This includes six metals from group 13–15, 17 transition metals, and Lanthanum. Surface geometry and enumerated adsorption sites for the three structures are provided in the lower panel, where top, bridge, and hollow sites are shown in red, white, and green, respectively.

This gives a total of 96,015 unique surfaces, adsorbate and site combinations (including the empty slabs), where roughly 65,000 calculations are completed. Also, the adsorption of the hydrogenated species CH, NH, CH_2, CH_3, SH, OH and H_2O has been studied for a smaller subset of alloy surfaces, where alloys formed from 16 metals of particular interest for catalysis have been chosen, with approximately 25,000 calculations completed. We note that due to reorientation of adsorbates during structure relaxation, the number of unique surface structures are lower than the number of initially sampled configurations. More than 90,000 adsorption and reaction energies have been parsed from the dataset, where approximately 30,000 adsorption energies stems from the monoatomic adsorbates (H, C, N, O and S), and 10,000 adsorption and reaction energies stems from the hydrogenated adsorbates. The remaining reaction energies are generated by decomposing a set of gas phase molecules of interest for catalysic applications, such as CH_4(g), NH_3(g), CO_2(g), CH_2CH_2(g), CH_3OH(g), H_2O_2(g), CH_3COOH(g), into their atomic constituents on the surfaces.

The dataset is made available from the open repository Catalysis-Hub.org[17], where reaction energies and barriers from more than 50 publications and datasets can be accessed.

Examples of calculated adsorption energies are given in Fig. 2, showing the results for the most stable adsorption sites for atomic carbon (C), oxygen (O), and nitrogen (N). In Fig. 2(a,b) the adsorption energies are plotted as a function of metal A and B, that are arranged on an improved Pettifor scale[18,19], with small adjustments for magnetic elements, which ensures a smooth variation of the adsorption energies with composition. Grey areas in the figure can be seen for structures where converged adsorption energies could not be obtained due to surface reconstruction, mismatch in the magnetic structure of the slab and the adsorbate-slab structure or convergence problems for the electronic structure calculation.

Another approach for visualizing adsorption energy trends is to plot the adsorption energy of two adsorbates against each other, which often gives rise to linear scaling relationships for similar surface geometries. Utilizing scaling relationships is a well established approach in theoretical catalysis to model and understand catalytic activity and selectivity[6,20]. In the lower panel of Fig. 2 the correlation between the adsorption of carbon with (c) oxygen and (d) nitrogen is shown. Metals containing a partially filled *d*-band versus a filled or empty *d*-band is labeled as *d*- and *non-d* metals respectively. All alloys containing a *non-d* metal are labeled as *non-d* alloys. While a close to linear relationship between the adsorption of C and O can be seen for the *d* and *non-d* pure metals separately, the correlation between the atomic adsorption energies on the alloys are more complicated, emphasizing the need for more sophisticated methods for modelling these systems, such as data-driven approaches. A link to the script used to plot Fig. 2 by fetching the data directly with the Catalysis-Hub Python API is provided in the *Methods* section.

Methods

Adsorption energies were calculated with DFT and obtained from the equation:

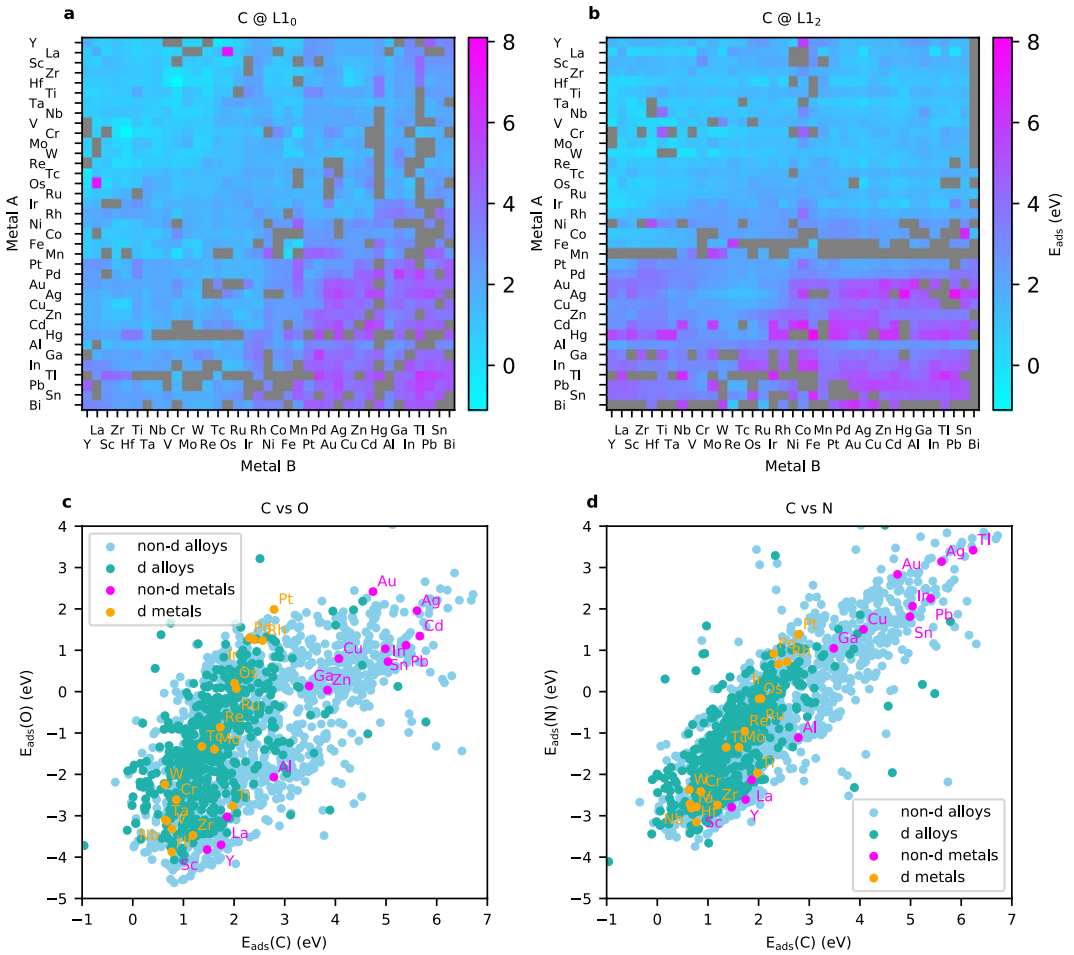

Fig. 2 Adsorption energies of selected atomic species. In (**a**,**b**) the C adsorption energy is shown for the 666 $L1_0$ (AB) and 1332 $L1_2$ (A_3B) alloy surfaces respectively. Results for the 37 pure metals are shown along the diagonal. Adsorption energies of atomic (**c**) O and (**d**) N are plotted with respect to the C adsorption energy for all materials. References are taken with respect to the reactions: $CH_4(g) - 2\,H_2(g) + * \rightarrow C*$, $0.5\,N_2(g) * \rightarrow N*$ and $H_2O(g) - H_2(g) + * \rightarrow O*$ with all species adsorbed to their lowest energy site.

$$E_{\text{ads}} = E^{\text{DFT}}_{\text{slab+adsorbate}} - E^{\text{DFT}}_{\text{slab}} - \sum_i \alpha_i E^{\text{DFT}}_{\text{i,gas}}, \tag{1}$$

where the gas phase species are chosen among $CH_4(g)$, $H_2(g)$, $N_2(g)$, $H_2O(g)$ and $H_2S(g)$. A full list of studied adsorbates and references used are given in Table 1.

The Catalysis Kit (CatKit)[21] software was used to generate the slab structures from optimized bulk systems and to systematically enumerate all high-symmetry adsorption sites. The generated structures were stored and handled with the ASE database[22].

DFT calculations were performed in the Quantum Espresso (QE) electronic structure code[23], using the BEEF-vdw exchange correlation functional[24], a 500 eV plane-wave cutoff, and a 5,000 eV density cutoff. Monkhorst-Pack *k*-point grids of (12, 12, 12) for bulk and (6, 6, 1) for slab calculations were used, with a 0.15 eV Fermi smearing. Spin-polarized calculations were performed only for alloys containing Fe, Ni, Co, and Mn, while allowing the magnetic moments to converge during the electronic structure optimization. Initial magnetic moments of 3, 3, 2, 1 μ_B was chosen for Fe, Mn, Co and Ni respectively, and set to zero for all other elements. For the A_1 and $L1_2$ structures, lattice constants were obtained from bulk alloy calculations with an equation of state combined with an energy minimization in QE. For $L1_0$ structures we used a variable cell optimization in QE with a high plane wave cutoff (800 eV) and then used the resulting lattice constants as initial guess for the final energy minimization with respect to lattice constant parameters - i.e., 'a' and 'c' - using the Scipy fmin optimizer[25]. Slab geometries were optimized by fixing the two bottom layers and allowing the top layer and adsorbates to relax. Due to the large number of calculations, job submissions were handled with FireWorks[26] and the CatFlow submodule of CatKit, that provides a FireWorks interface to QE and other electronic structures calculators supported by ASE.

Upon relaxation we found that reconstructions of the metal surfaces, e.g. horizontal sliding or dissociation of the top layer from the slab, are quite common. Also, we found that the adsorbates often reorient into other sites. The relaxed geometries were therefore post-processed with a tailored classification method to label reconstructed

Adsorbate	Gas phase references
H*	$0.5\ H_2(g)$
N*	$0.5\ N_2(g)$
C*	$CH_4(g) - 2\ H_2(g)$
O*	$H_2O(g) - H_2(g)$
S*	$H_2S(g) - H_2(g)$
*CH	$CH_4(g) - 1.5\ H_2(g)$
*CH_2	$CH_4(g) - H_2(g)$
*CH_3	$CH_4(g) - 0.5\ H_2(g)$
*NH	$0.5\ N_2(g) + 0.5\ H_2(g)$
*OH	$H_2O(g) - 0.5\ H_2(g)$
H_2O*	$H_2O(g)$
*SH	$H_2S(g) - 0.5\ H_2(g)$

Table 1. Studied adsorbates listed together with the choice for gas phase reference used to calculate the adsorption energy, Eq. 1. The '*' symbol in the adsorbate formula specifies which atom of the molecule binds to the surface.

	Computed lattice constant [a]	Experimental lattice constant [a]
Pb	4.89	4.91[29]
Ni	3.53	3.51[29]
Pt	3.95	3.91[29]
Ir	3.88	3.84[30]
Al	4.02	4.02[29]
Cu	3.67	3.60[29]
Pd	3.99	3.88[29]
Au	4.21	4.06[29]
Ag	4.22	4.06[29]

Table 2. Equilibrium lattice constant of pure FCC metals, in Å.

	Computed lattice constant [a]	Reported lattice constant [a]
Pd_3Co	3.88	3.93[31]
Pd_3Ni	3.89	3.93[31]
Fe_3Pt	3.72	3.74[16]
Ni_3Pt	3.67	3.66[16]
Zr_3Al	4.38	4.37[32]
Sc_3Al	4.41	4.42[32]

Table 3. Equilibrium lattice constant of $L1_2$ metals, in Å.

surfaces and reclassify the adsorption sites. Only non-reconstructed surfaces have been used to generate adsorption energies, although, as the reconstructed structures could be of interest for model generation, the atomic structures are still made available via the web and python APIs discussed in the *Usage Notes* section.

Data Records

All the DFT calculations stems from one dataset, generated by O. Mamun, K. Winther, and J. Boes, in the group of Thomas Bligaard at the SUNCAT-Center for Interface Science and Catalysis. The data are made available from two platforms, where the open electronic structure database Catalysis-Hub.org[17], is the main resource for easy access to parsed adsorption energies. The dataset has been assigned its own permanent link at https://www.catalysis-hub.org/publications/MamunHighT2019. The Catalysis-Hub web interface enables in-browser search for reactions and chemical compositions, with a visualization of atomic geometries, that can be downloaded in several formats including CIF, JSON, VASP POSCAR and Quantum ESPRESSO input. A description on how to download reaction energies and atomic structures with the Catalysis-Hub (CatHub) Python API, available from the Zenodo repository[27], is provided in the *Usage Notes* section.

	Computed lattice constant [a]	Computed $\left[\frac{c}{a}\right]$	Reported lattice constant [a]	Reported $\left[\frac{c}{a}\right]$
PdCo	3.80	0.98	3.83[31]	0.98[31]
PdNi	3.87	0.94	3.83[31]	0.95[31]
PtFe	3.88	0.97	3.86[16]	0.97[16]
PtNi	3.86	0.94	3.85[16]	0.95[16]
CoPt	3.83	0.97	3.83[33]	0.97[33]

Table 4. Equilibrium lattice constant of $L1_0$ metals, in Å.

Reaction	Metal	Computed adsorption energy [eV]	Reported adsorption energy [eV]
$0.5\,H_2(g) + * \rightarrow H*$	*Pt*	−0.31	−0.24[8]
$0.5\,H_2(g) + * \rightarrow H*$	*Pd*	−0.29	−0.28[34]
$0.5\,H_2(g) + * \rightarrow H*$	*Re*	−0.51	−0.60[34]
$0.5\,H_2(g) + * \rightarrow H*$	*Rh*	−0.28	−0.34[35]
$0.5\,H_2(g) + * \rightarrow H*$	*Ir*	−0.16	−0.19[8]
$0.5\,H_2(g) + * \rightarrow H*$	*Ag*	0.34	0.44[35]
$0.5\,H_2(g) + * \rightarrow H*$	*Cu*	−0.05	0.03[35]
$O_2(g) + * \rightarrow 2O*$	*Pt*	−1.90	−1.96[35]
$H_2O(g) + * \rightarrow OH* + 0.5\,H_2(g)$	*Pd*	0.66	0.60[8]
$H_2O(g) + * \rightarrow OH* + 0.5\,H_2(g)$	*Co*	−0.24	−0.31[8]
$H_2O(g) + * \rightarrow OH* + 0.5\,H_2(g)$	*Ag*	0.72	0.63[8]
$H_2O(g) + * \rightarrow OH* + 0.5\,H_2(g)$	*Cu*	0.19	0.28[8]
$H_2O(g) + 2* \rightarrow OH* + H*$	*Ag*	1.02	1.13[8]
$H_2O(g) + 2* \rightarrow OH* + H*$	*Cu*	0.28	0.36[8]
$CH_3* + * \rightarrow CH_2* + H*$	*Ag*	1.87	2.11/1.89[8,36]
$CH_3* + * \rightarrow CH_2* + H*$	*Cu*	0.92	1.15/0.94[8,36]
$CH_3* + * \rightarrow CH_2* + H*$	*Au*	0.71	0.77[36]
$0.5\,N_2(g) + * \rightarrow N*$	*Pd*	1.06	0.94[34]
$0.5\,N_2(g) + * \rightarrow N*$	*Re*	−1.33	−1.46[34]
$0.5\,N_2(g) + * \rightarrow N*$	*Pt*	0.78	0.91[34]
$0.5\,N_2(g) + * \rightarrow N*$	*Ir*	0.34	0.30[34]
$0.5\,N_2(g) + * \rightarrow N*$	*Rh*	0.05	0.06[34]
$0.5\,N_2(g) + * \rightarrow N*$	*Au*	2.80	2.92[34]
$0.5\,N_2(g) + * \rightarrow N*$	*Ru*	−0.65	−0.57[34]
$0.5\,N_2(g) + * \rightarrow N*$	*Ag*	3.11	3.19[34]
$0.5\,N_2(g) + * \rightarrow N*$	*Cu*	1.50	1.53[34]

Table 5. Adsorption energies of various adsorbates on pure metals in eV. The raw DFT reaction energy without energy corrections is reported for all values in the table. (Note in ref.[8] a +0.1 eV correction was applied to $H_2(g)$).

In addition, all the raw text output files from the Quantum Espresso calculations have been uploaded to the Materials Cloud archive[28]. The output files can be downloaded and inspected with any text editor, or opened with ASE[22] to create Atoms objects containing the atomic structures and the results of the calculations.

Technical Validation

To ensure the quality of the adsorption properties reported herein, the convergence with respect to all calculation parameters have been carefully checked. Adsorption and reaction energies have only been included for surface structures that do not undergo reconstruction upon relaxation. In the case of magnetic surface structures, we have only parsed adsorption and reaction energies if the discrepancy in total magnetization between the empty surface and the surface with the adsorbate is less than 4 in atomic units.

To illustrate the validity of the data, we compare the lattice constant reported in reputed journal articles to the computed lattice constant. We found excellent agreement between our results and previously computed lattice constants which are presented in Tables 2, 3 and 4. In Table 5, we also show a comparison between previously computed adsorption energies to those reported in this article. We also see good agreement between the reported and the computed adsorption energies with slight deviation which may be an artifact of different calculation setup and/or system size, i.e., pseudopotential, smearing scheme, number of layers and lateral size of the slab. For

A1	$L1_2$	$L1_0$
top|A	top|A	top|A
	top|B	top|B
bridge|A_A|A	bridge|A_A|A	bridge|A_A|A*
		bridge|B_B|B*
	bridge|A_A|B	bridge|A_A|B
		bridge|B_B|A
	bridge|A_B|A	bridge|A_B|A
		bridge|A_B|B
hollow|A_A_A|HCP	hollow|A_A_A|HCP	
hollow|A_A_A|FCC	hollow|A_A_A|FCC	
	hollow|A_A_B|HCP	hollow|A_A_B|HCP
	hollow|A_A_B|FCC	hollow|A_A_B|FCC
		hollow|A_B_B|HCP
		hollow|A_B_B|FCC
subsurface*	subsurface*	subsurface*
		4fold|A_A_B_B*

Table 6. Names of sampled adsorption sites, where A and B refers to the choice of metals. The sites marked with '*' have not been sampled with the initial configuration shown in Fig. 1, but stems from deviation from the hexagonal surface structure for some of the $L1_0$ alloys or reorientation of the adsorbate into a subsurface site.

example, the differences in adsorption energies between this work and ref.[8], which is also available at https://www.catalysis-hub.org/publications/SchumannSelectivity2018, can be attributed to the use of a 4 layer (3×3) repeated surface slab model, compared to the 3 layer (2×2) slab used in this study.

Usage Notes

The CatHub software module, which is available from the Zenodo repository[27], provides a Python API which is better suited for fetching a large amount of data from the Catalysis-Hub repository. A small script for obtaining pre-parsed adsorption energies in Python is provided below:

```
from cathub.query import get_reactions
get_reactions(pubId='MamunHighT2019',
              n_results=2,
              surfaceComposition='Mo+Ru',
              reactants="CH4gas+H2",
              sites="~hollow",
              products='C')
```

which returns a JSON dictionary on the form:

```
{'reactions':
  {'edges': [x'
    {'node':
         {'Equation': 'CH4(g) - 2.0H2(g)+* ->C*',
         'activationEnergy': None,
         'chemicalComposition': 'Mo3Ru9',
         'coverages': {'C': 0.25},
         'dftCode': 'Quantum ESPRESSO 5.1',
         'dftFunctional': 'BEEF-vdW',
         'facet': '111',
         'products': {'Cstar': 1},
         'pubId': 'MamunHighT2019',
         'reactants': {'star': 1, 'H2gas': -2.0,
                       'CH4gas': 1.0}',
         'reactionEnergy': 1.2068607934897955,
         'sites': {'C': 'hollow|A_A_A|HCP'},
         'surfaceComposition': 'Ru3Mo'}
    },
    {'node':…}
    ],
   'totalCount': 10}
  }
}.
```

Note that each data entry is given as a 'node' in a list of 'edges', utilizing the graph-theory based query language GraphQL (https://graphql.org/). Since the Catalysis-Hub repository contains several datasets from different

Key	Description	Datatype
id	Local database id	int
unique_id	Globally unique hexadecimal id	str
ctime	Creation time	float
mtime	Modification time	float
user	User name	str
numbers	Atomic numbers	int
pbc	Periodic boundary condition flags	bool
cell	Unit cell	float
positions	Atomic positions	float
initial_magmoms	Initial atomic magnetic moments	float
initial_charges	Initial atomic charges	float
masses	Atomic masses	float
tags	Tags	int
momenta	Atomic momenta	float
constraints	Constraints	list of dict
energy	Total energy	float
forces	Atomic forces	float
stress	Stress tensor	float
dipole	Electrical dipole	float
charges	Atomic charges	float
magmom	Magnetic moment	float
magmoms	Atomic magnetic moments	float
calculator	Calculator name	str
calculator_parameters	Calculator parameters	dict
metalA	Name of A metal (see Fig. 1)	str
metalB	Name of B metal (see Fig. 1)	str
slab_name	Chemical stiochemitry of slab	str
SB_symbol	Structurbericht designation	str
adsorbate	Adsorbate (see Table 1)	str
fw_id	User specific Fireworks job id	int
geometry	Unique structure descriptor	str
pub_id	Dataset id (MamunHighT2019)	str
reconstructed	Surface reconstruction	bool (0 or 1)
relaxed	Relaxed structure	bool (0 or 1)
site	Primary site descriptor	str
site_type	Secondary site descriptor	str
state	State (bulk, molecule or slab)	str

Table 7. Data structure for storing atomic structures with the ASE database. The upper panel show the native ASE database columns and the lower panel dataset specific key value pairs.

publications, the "pubId = 'MamunHighT2019'" argument must be assigned in the script above in order to query only this dataset. The script above queries for entries with hollow-site adsorption of C (product) with respect to the relevant gas phase species (reactants), on surfaces containing Mo as well as Ru. The reaction and product entries must be chosen (and matched) among the adsorbates and gas phase references in Table 1. A more specific query for adsorption site can be made by using the site names specified in Table 6.

Furthermore, easy access to all the atomic structures, calculation results and parameters in the study, can be obtained with the ASE database interface[22], where the CatHub module features a convenient wrapper around the ASE db command line interface (CLI), used directly from a terminal. For example, the query:

```
cathub ase 'pub_id=MamunHighT2019, relaxed=1'
```

will return a list with the first 20 results (out of approximately 90,000) for the relaxed configurations in the study. The initial geometries can be queried by setting 'relaxed = 0'. The atomic structures are labeled with an several key-value-pair metadata, that can be used to query the dataset. For example:

```
cathub ase 'Pt,pub_id=MamunHighT2019,relaxed=1,
  reconstructed=0,SB_symbol=L10
  -c formula,energy,adsorbate,site,site_type -L 100'
  --gui
```

will return the 100 first relaxed and non-reconstructed structures containing Pt in the $L1_0$ structure, as a table containing the chemical formula, the DFT total energy, the name of the adsorbate, and site information as well as opening all the matching structures in the ASE gui visualizer. A description of ASE native columns as well as special key value pairs assigned in this study is given in Table 7. We refer to ref.[17] for a detailed description of the Catalysis-Hub database structure.

Code Availability

The CatHub python API and the CatKit software packages are available open-source from the GitHub repository at https://github.com/SUNCAT-Center/. In addition, the latest stable version of the CatHub module is available from the Zenodo repository[27]. The Python scripts used for plotting the data shown in Fig. 2 is made available as a tutorials at https://github.com/SUNCAT-Center/CatHub/tree/master/tutorials/1_bimetallic_alloys/. The code used to classify the adsorption sites is made available at https://github.com/SUNCAT-Center/CatHub/tree/master/cathub/classification.py.

References

1. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Physical Review* **136**, B864 (1964).
2. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Physical Review* **140**, A1133 (1965).
3. Jones, R. O. Density functional theory: Its origins, rise to prominence, and future. *Reviews of Modern Physics* **87**, 897 (2015).
4. Nørskov, J. K., Abild-Pedersen, F., Studt, F. & Bligaard, T. Density functional theory in surface chemistry and catalysis. *Proceedings of the National Academy of Sciences* **108**, 937–943 (2011).
5. Nørskov, J. K., Bligaard, T., Rossmeisl, J. & Christensen, C. H. Towards the computational design of solid catalysts. *Nature Chemistry* **1**, 37 (2009).
6. Medford, A. J. *et al*. From the sabatier principle to a predictive theory of transition-metal heterogeneous catalysis. *Journal of Catalysis* **328**, 36–42 (2015).
7. Subramani, V. & Gangwal, S. K. A review of recent literature to search for an efficient catalytic process for the conversion of syngas to ethanol. *Energy & Fuels* **22**, 814–839 (2008).
8. Schumann, J. *et al*. Selectivity of synthesis gas conversion to C2+ oxygenates on fcc(111) transition-metal surfaces. *ACS Catalysis* **8**, 3447–3453 (2018).
9. Shao, M., Chang, Q., Dodelet, J.-P. & Chenitz, R. Recent advances in electrocatalysts for oxygen reduction reaction. *Chemical Reviews* **116**, 3594–3657 (2016).
10. Debe, M. K. Electrocatalyst approaches and challenges for automotive fuel cells. *Nature* **486**, 43 (2012).
11. Wang, J. *et al*. Recent progress in cobalt-based heterogeneous catalysts for electrochemical water splitting. *Advanced Materials* **28**, 215–230 (2016).
12. Jacobsen, C. J. *et al*. Catalyst design by interpolation in the periodic table: Bimetallic ammonia synthesis catalysts. *Journal of the American Chemical Society* **123**, 8404–8405 (2001).
13. Greeley, J., Jaramillo, T. F., Bonde, J., Chorkendorff, I. & Nørskov, J. K. Computational high-throughput screening of electrocatalytic materials for hydrogen evolution. *Nature Materials* **5**, 909–913 (2006).
14. Escudero-Escribano, M. *et al*. Tuning the activity of Pt alloy electrocatalysts by means of the lanthanide contraction. *Science* **352**, 73–76 (2016).
15. Chen, Z.-X., Neyman, K. M., Lim, K. H. & Rösch, N. CH3O decomposition on PdZn (111), Pd (111), and Cu (111). A theoretical study. *Langmuir* **20**, 8068–8077 (2004).
16. Hensley, A. J. R., Schneider, S., Wang, Y. & McEwen, J.-S. Adsorption of aromatics on the (111) surface of PtM and PtM_3 (M = Fe, Ni) alloys. *RSC Advances* **5**, 85705–85719 (2015).
17. Winther, K. T., Hoffmann, M. J., Mamun, O., Boes, J. R. & Bligaard, T. Catalysis-Hub.org, an open electronic structure database for surface reactions. *Scientific Data* **6**, 81 (2019).
18. Pettifor, D. G. A chemical scale for crystal-structure maps. *Solid State Communications* **51**, 31–34 (1984).
19. Glawe, H., Sanna, A., Gross, E. & Marques, M. A. The optimal one dimensional periodic table: A modified pettifor chemical scale from data mining. *New Journal of Physics* **18**, 093011 (2016).
20. Abild-Pedersen, F. *et al*. Scaling properties of adsorption energies for hydrogen-containing molecules on transition-metal surfaces. *Physical Review Letters* **99**, 016105 (2007).
21. Boes, J. R., Mamun, O., Winther, K. & Bligaard, T. Graph theory approach to high-throughput surface adsorption structure generation. *The Journal of Physical Chemistry A* **123**, 2281–2285 (2019).
22. Larsen, A. H. *et al*. The atomic simulation environment—a Python library for working with atoms. *Journal of Physics: Condensed Matter* **29**, 273002 (2017).
23. Giannozzi, P. *et al*. Advanced capabilities for materials modelling with QUANTUM ESPRESSO. *Journal of Physics: Condensed Matter* **29**, 465901 (2017).
24. Wellendorff, J. *et al*. Density functionals for surface science: Exchange-correlation model development with bayesian error estimation. *Physical Review B* **85**, 235149 (2012).
25. Nelder, J. A. & Mead, R. A simplex method for function minimization. *The Computer Journal* **7**, 308–313 (1965).
26. Jain, A. *et al*. Fireworks: a dynamic workflow system designed for high-throughput applications. *Concurrency and Computation: Practice and Experience* **27**, 5037–5059 (2015).
27. Winther, K. T. *et al*. CatHub: A Python API for the Surface Reactions Database on Catalysis-Hub.org. *Zenodo*. https://doi.org/10.5281/zenodo.2600391 (2019).
28. Mamun, O., Winther, K. T., Boes, J. R. & Bligaard, T. High-throughput calculations of catalytic properties of bimetallic alloy surfaces. *Materials Cloud Archive*. https://doi.org/10.24435/materialscloud:2019.0015/v1 (2019).
29. Haas, P., Tran, F. & Blaha, P. Calculation of the lattice constant of solids with semilocal functionals. *Physical Review B* **79**, 085104 (2009).
30. Donohue, J. *Structures of the Elements* (John Wiley and Sons, 1974).
31. Li, B., Greeley, J. & Prakash, J. Understanding the oxygen reduction reaction on Pd-based alloys (Pd-M, M = Ni, Co) surfaces using density functional theory calculations. *ECS Transactions* **19**, 109–116 (2009).
32. Arıkan, N. The first-principles study on Zr3Al and Sc3Al in L12 structure. *Journal of Physics and Chemistry of Solids* **74**, 794–798 (2013).
33. Liu, Z., Lei, Y. & Wang, G. First-principles computation of surface segregation in L1 CoPt magnetic nanoparticles. *Journal of Physics: Condensed Matter* **28**, 266002 (2016).
34. Montoya, J. H., Tsai, C., Vojvodic, A. & Nørskov, J. K. The challenge of electrochemical ammonia synthesis: A new perspective on the role of nitrogen scaling relations. *ChemSusChem* **8**, 2180–2186 (2015).
35. Wellendorff, J. *et al*. A benchmark database for adsorption bond energies to transition metal surfaces and comparison to selected DFT functionals. *Surface Science* **640**, 36–44 (2015).
36. Hansen, M. H., Nørskov, J. K. & Bligaard, T. First principles micro-kinetic model of catalytic non-oxidative dehydrogenation of ethane over close-packed metallic facets. *Journal of Catalysis* **374**, 161–170 (2019).

Acknowledgements
This work was supported by the U.S. Department of Energy, Chemical Sciences, Geosciences, and Biosciences (CSGB) Division of the Office of Basic Energy Sciences, via Grant DE-AC02-76SF00515 to the SUNCAT Center for Interface Science and Catalysis.

Author Contributions
The dataset was generated by O. Mamun in collaboration with J. Boes and K. Winther. K. winther was responsible for parsing and uploading data to the database platforms, and surface geometries were generated by J. Boes. T. Bligaard formed the vision and scope of the work. The manuscript was written by K. Winther and O. Mamun, and has been revised and approved by all authors.

Additional Information
Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.