



Citrine Informatics

The data analytics platform for the physical world

Citrination Tutorial

Eric Lundberg

Dec. 2017

Outline

Objective: Familiarize new users with the key functionality of the Citrination platform

- Search
- Upload Data
- Create and Apply ML Models:
 - Create a Data View
 - Plot data
 - Predict properties of unknown materials
 - Design materials to meet parameters
 - Assess model quality

NOTE: Your experience may vary because each Citrination site is different, and we update our platform regularly

- All example images are from the base Citrination.com website. Each private Citrination site may look different
- We add functionality, fix bugs, and update the user interface frequently. Revisit the Citrination website for an updated version of these instructions
- If the site doesn't seem to work or you see an error message, closely check that you followed these instructions and contact us at training@citrine.io

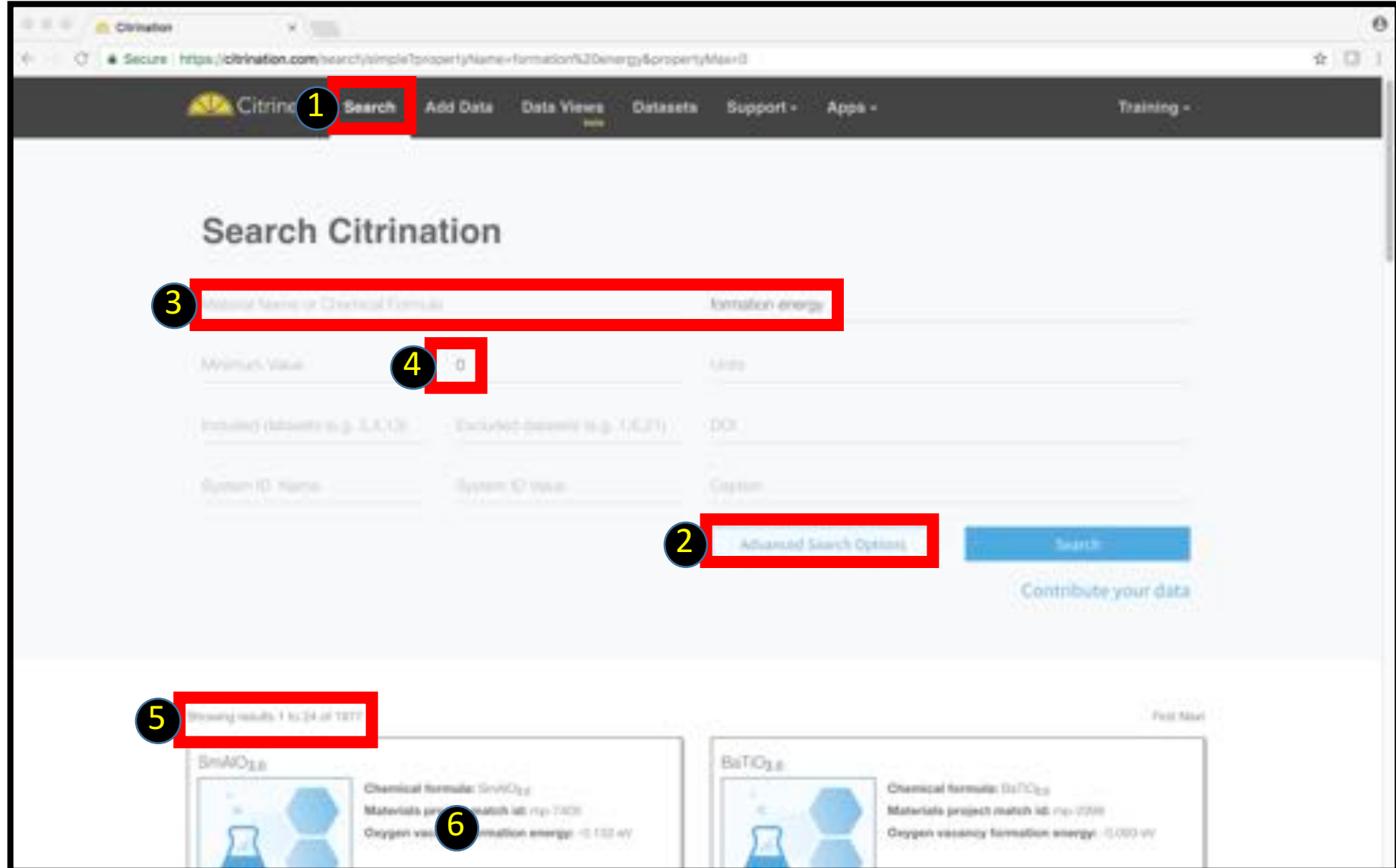
Search Function

- **What is it?** Explore the world's largest database of materials and chemicals information, including: polymers, alloys, semiconductors, and many more. This database is broken up into material cards, where the properties of a known material are consolidated into a single view.
- **How does this help you?** Want to know the properties of a material? Want to find materials with certain properties? Use Advanced Search to apply more filters to narrow down your options.
- **How it works:** You can search in 2 ways: all datasets or a specific dataset. Your search will return a number of materials cards that meet the requirements you set. Click on one to explore more!

Search – All Datasets (1/3)

Explore materials data across all available data sets with the Search tab.

1. Click Citrination's Search tab
2. Click "Advanced Search Options"
3. Type material OR property of interest
4. Set constraints
5. See number of results returned
6. Click on an material card under "results"

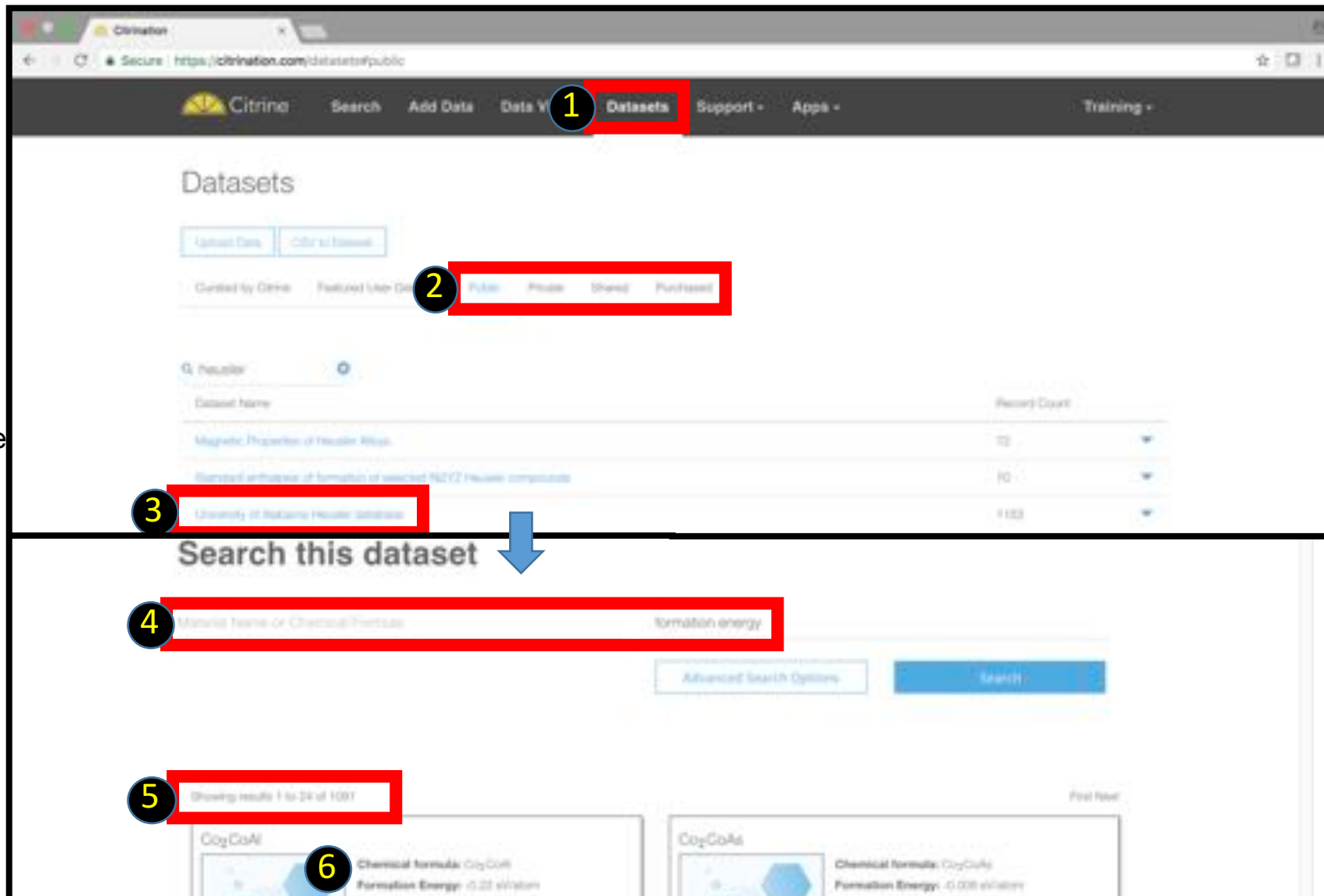


The screenshot shows the Citrination web application interface. The browser address bar displays the URL: <https://citrination.com/search/simple?propertyName=formation%20energy&propertyMap=0>. The navigation bar includes the Citrine logo, a "Search" button (annotated with a red box and number 1), and links for "Add Data", "Data Views", "Datasets", "Support", and "Apps". The main heading is "Search Citrination". Below this is a search form with a text input field (annotated with a red box and number 3) containing "formation energy". To the right of the input field is a "Units" dropdown menu (annotated with a red box and number 4) set to "eV". Below the input field are several filter options: "Included datasets (e.g. 3A, 12)", "Excluded datasets (e.g. 1, 6, 21)", "DOI", "System ID Name", "System ID Value", and "Caption". A button labeled "Advanced Search Options" (annotated with a red box and number 2) is located to the right of these filters. A blue "Search" button is also present. Below the search form, a status bar (annotated with a red box and number 5) indicates "Showing results 1 to 24 of 1877". The results section displays two material cards. The first card is for $\text{SrTiO}_{3,6}$ and the second is for $\text{BaTiO}_{3,6}$. Both cards show the chemical formula, a materials project match ID, and the oxygen vacancy formation energy. The first card's energy is -1.132 eV (annotated with a red box and number 6), and the second card's energy is -1.000 eV. A "Full View" link is visible in the top right corner of the results section.

Search – Specific Dataset (2/3)

Learn more about a specific research project's collection of records with the Search-Dataset tab.

1. Click Datasets tab
2. Choose type of dataset
 - Public: Visible to everyone on the site
 - Private: Uploaded by you
 - Shared: Shared with you by others
 - Purchased: Purchased by your organization
3. Click on dataset
4. Type material name or property of interest
5. See results
6. Click on a material card



The screenshot shows the Citrine Informatics website interface. The 'Datasets' tab is highlighted in the top navigation bar (1). Below it, the 'Public' dataset type is selected (2). A dataset titled 'University of Indiana Heusler database' is chosen (3). A search bar is used to enter 'formation energy' (4). The results show 'Showing results 1 to 24 of 1001' (5). A material card for 'Co₂CoAl' is highlighted (6), showing its chemical formula and formation energy.

Search – Material Card

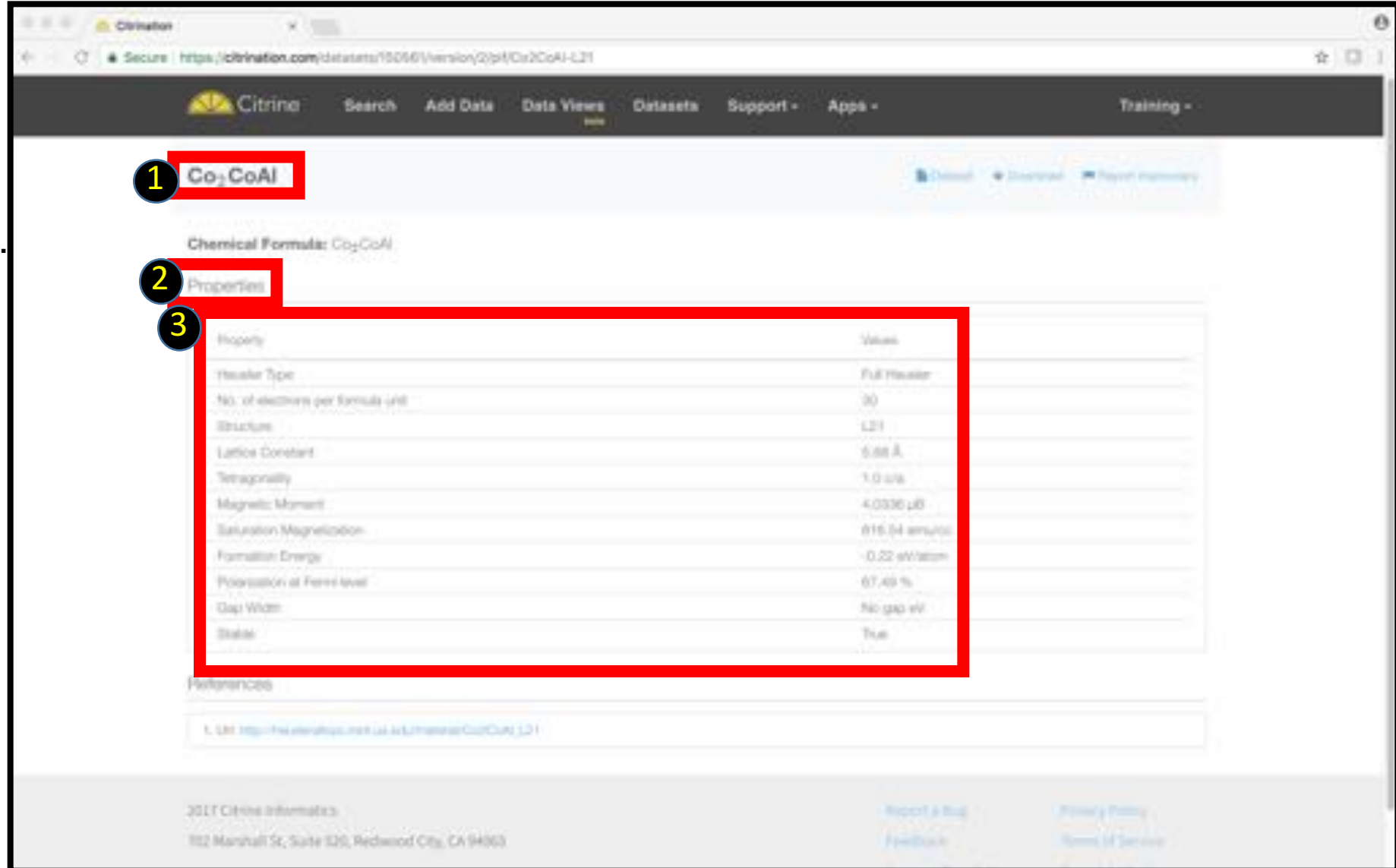
(3/3)

Once you click on a search result you can see this card. It displays the properties of a known material

1. Chemical formula/ composition
2. Type of data (e.g. property, composition, preparation, method, or references)

See <http://help.citrine.com/> for more information what can go into a material card or what types of data Citrine recognizes

3. Data



1 **Co₂CoAl**

Chemical Formula: Co₂CoAl

2 **Properties**

3

Property	Value
Heuser Type	Full Heuser
No. of electrons per formula unit	30
Structure	L21
Lattice Constant	5.68 Å
Tetragonality	1.0 u/a
Magnetic Moment	4.0336 μB
Saturation Magnetization	616.34 emu/mg
Formation Energy	-0.22 eV/atom
Polarization at Fermi level	67.49 %
Gap Width	No gap eV
Stable	True

References

1. URL http://www.earthchem.mcgill.ca/earthchem/Co2CoAl_L21/

2017 Citrine Informatics
112 Marshall St, Suite 120, Redwood City, CA 94063

[Report a Bug](#) [Privacy Policy](#)
[Feedback](#) [Terms of Service](#)

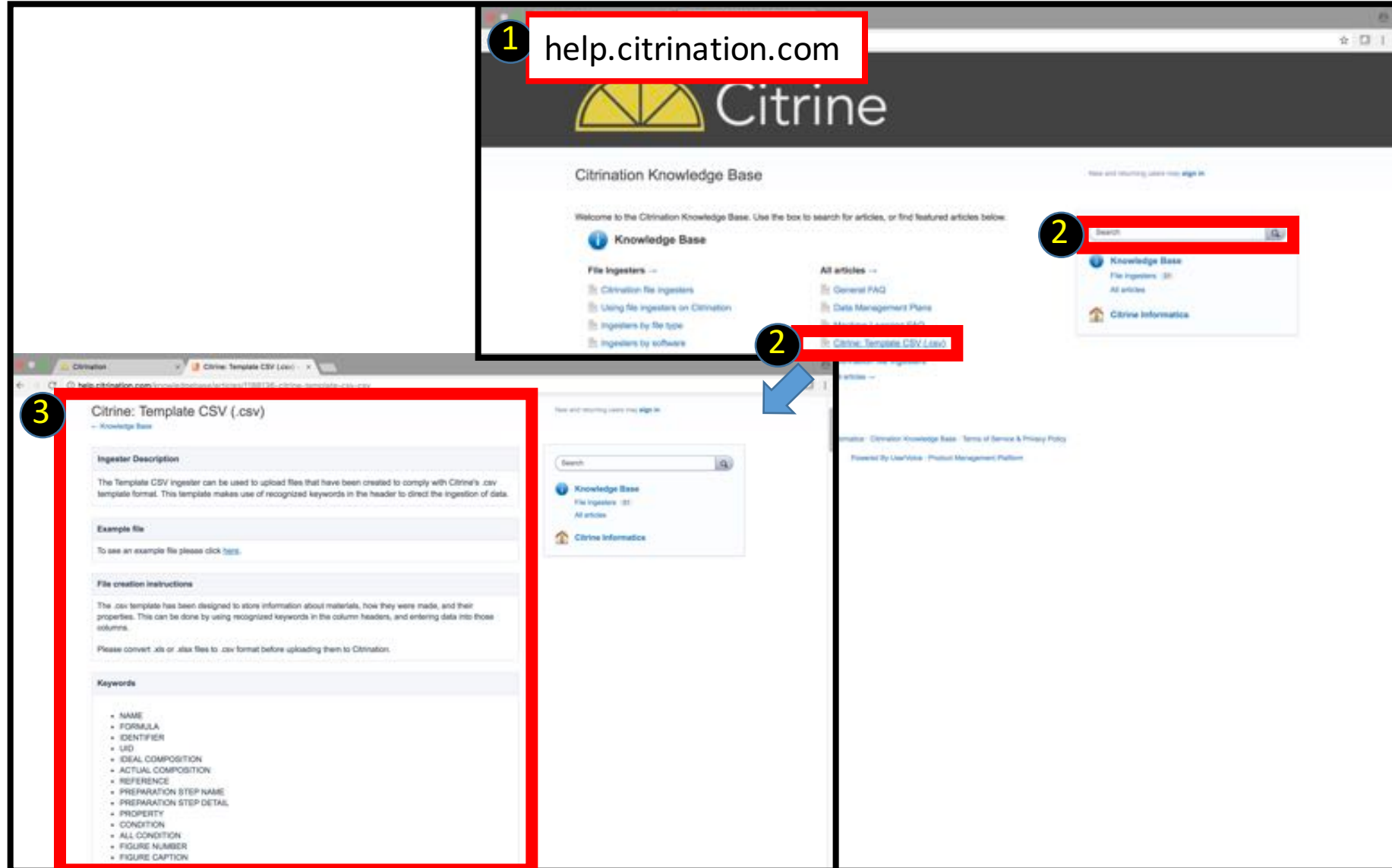
Upload Data

- **What is it?** This feature allows you to upload data (or any type of file) to the Citrination platform, organize it, and keep it private or share it with the rest of the users on the site. If you have a private Citrination site, that information is kept secure.
- **How does this help you?** Uploading your data to our site makes it searchable, shareable, and accessible for machine learning (ML)
- **How it works:** The Citrination platform will turn your data into a series of materials cards. We have dozens of ingesters to upload structured data files (e.g. CSV files, XRD files, and VASP DFT files)

Help on Ingesters

Learn how to upload your data so Citrine can process it. One common data format is the CSV Template. See the help.citrine.com page for more details

1. Open Citrine help site
2. Browse or search key words for a concept
3. Template CSV page



The screenshot illustrates the steps to find the CSV Template page on the Citrine Informatics Knowledge Base. The website header includes the Citrine logo and the title 'Citrine Knowledge Base'. A search bar is located in the top right corner. The main content area is divided into 'File Ingesters' and 'All articles' sections. The 'File Ingesters' section lists various ingestion methods, including 'Citrine File Ingesters', 'Using file ingesters on Citrine', 'Ingesters by file type', and 'Ingesters by software'. The 'All articles' section lists various articles, including 'General FAQ', 'Data Management Plans', 'Machine Learning FAQ', and 'Citrine Template CSV (.csv)'. A red box highlights the 'Citrine Template CSV (.csv)' link in the 'All articles' section. A blue arrow points from this link to the 'Citrine: Template CSV (.csv)' page, which is also highlighted with a red box. The page content includes an 'Ingestor Description', an 'Example file' link, 'File creation instructions', and a list of 'Keywords' such as NAME, FORMULA, IDENTIFIER, UID, IDEAL COMPOSITION, ACTUAL COMPOSITION, REFERENCE, PREPARATION STEP NAME, PREPARATION STEP DETAIL, PROPERTY, CONDITION, ALL COMPOSITION, FIGURE NUMBER, and FIGURE CAPTION.

1 help.citrine.com

2

3

Add Data

(1/2)

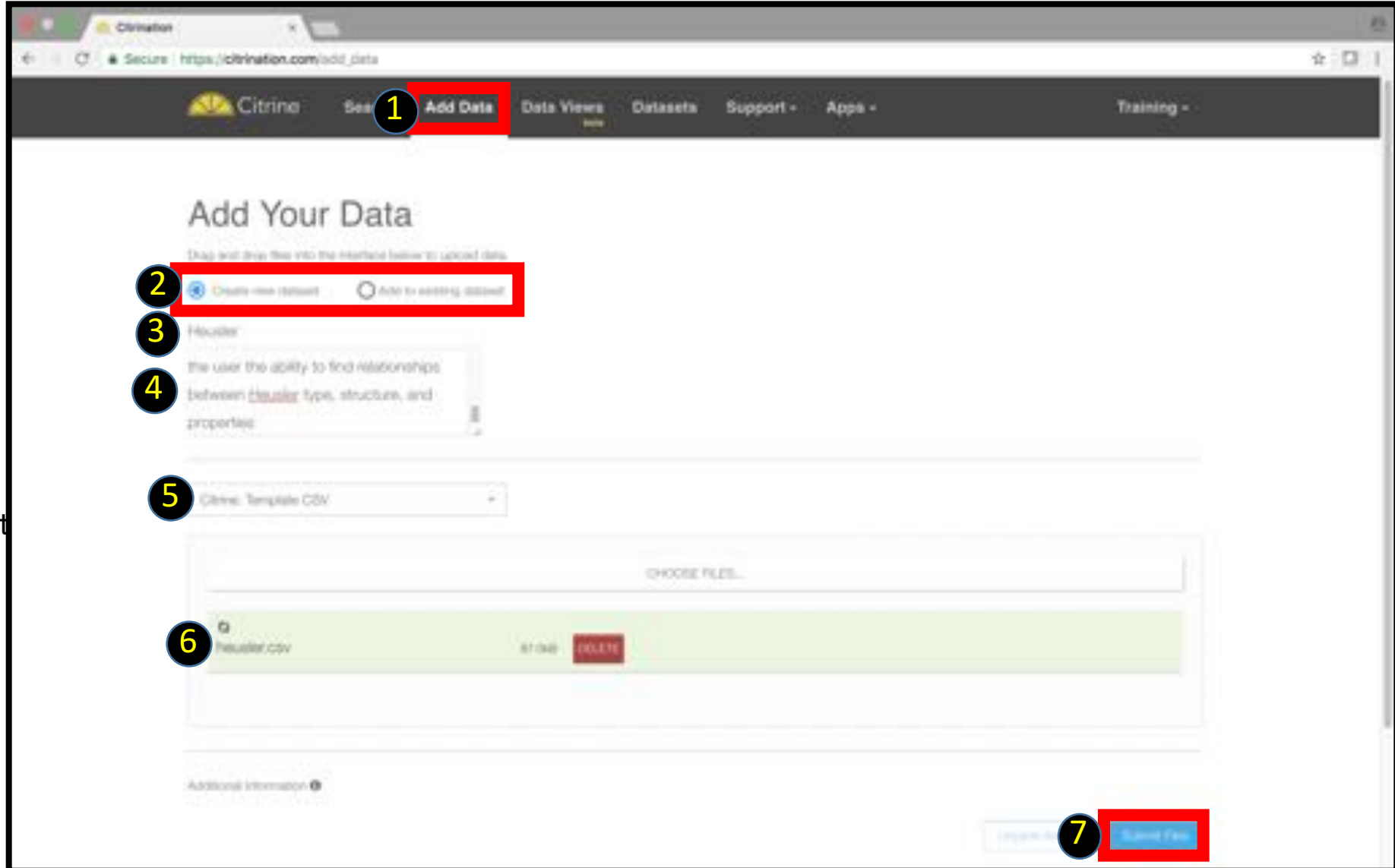
Upload your data to the secure Citrination site with the Add Data tab. You can create a new dataset or update an existing data set, (e.g. after experimentation).

Any file can be uploaded, but only some formats are parsed into materials records.

1. Click Add Data tab
2. Select new/existing dataset
3. Type Title

Each dataset for a given user must have a unique title

4. Type Description
5. Choose appropriate ingester (important)
6. Upload the file
7. Submit!



The screenshot shows the Citrination web application interface for adding data. The browser address bar shows the URL https://citrination.com/add_data. The navigation bar at the top includes the Citrine logo and tabs for 'Add Data', 'Data Views', 'Datasets', 'Support', and 'Apps'. The 'Add Data' tab is highlighted with a red box and a yellow circle containing the number 1.

The main content area is titled 'Add Your Data' and includes instructions: 'Drag and drop files into the workspace below to upload data.' Below this, there are two radio buttons: 'Create new dataset' (selected) and 'Add to existing dataset'. This section is highlighted with a red box and a yellow circle containing the number 2.

Below the radio buttons, there is a text input field for the dataset title, which contains the word 'Houster'. This is highlighted with a yellow circle containing the number 3.

Below the title field, there is a text area for the dataset description, which contains the text: 'the user the ability to find relationships between cluster type, structure, and properties'. This is highlighted with a yellow circle containing the number 4.

Below the description, there is a dropdown menu for selecting an ingester, which is currently set to 'Citrine Template CSV'. This is highlighted with a yellow circle containing the number 5.

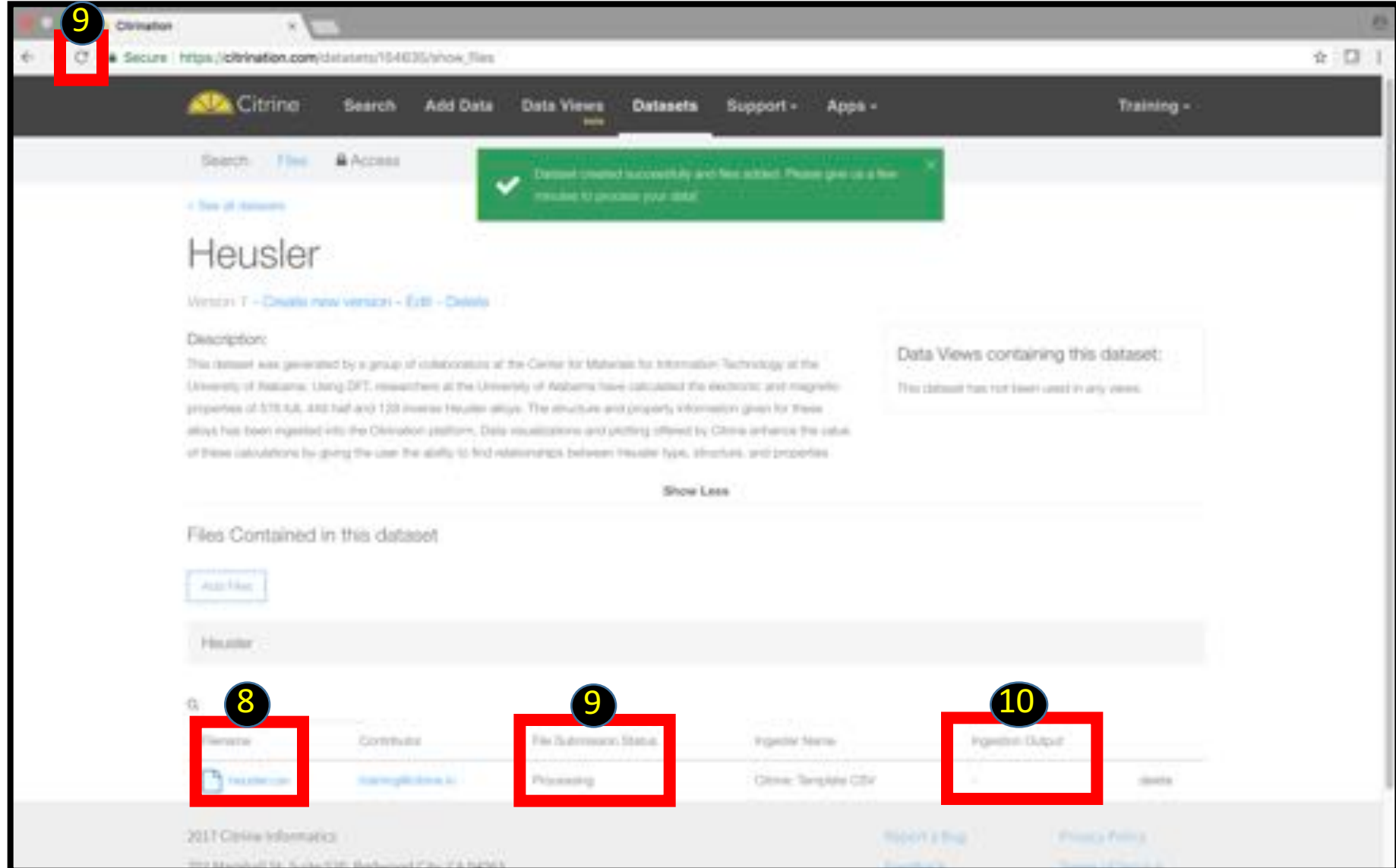
Below the ingester dropdown, there is a file upload area with a 'CHOOSE FILES...' button. A file named 'Houster.csv' has been uploaded and is shown in a green bar. This bar has 'STORE' and 'DELETE' buttons. This section is highlighted with a yellow circle containing the number 6.

At the bottom right of the form, there is a 'Submit File' button, which is highlighted with a red box and a yellow circle containing the number 7.

Add Data

(2/2)

8. Review uploaded file
9. Refresh to see the file's progress (may take up to <10m):
 - a. Initializing
 - b. Processing
 - c. Finished or Failed
10. Log file if data upload fails



The screenshot shows the Citrine Informatics web interface. At the top, a green banner indicates "Dataset created successfully and file added. Please give us a few minutes to process your data!". Below this, the dataset name "Heusler" is displayed, along with its version and a description. A table titled "Files Contained in this dataset" is shown, with columns for "File Submission Status" and "Ingestion Output". Red boxes and yellow numbers highlight the following elements:

- 9**: The browser's refresh button in the top left corner.
- 8**: The "Refresh" button in the file list table.
- 9**: The "File Submission Status" column header in the file list table.
- 10**: The "Ingestion Output" column header in the file list table.

Share Dataset

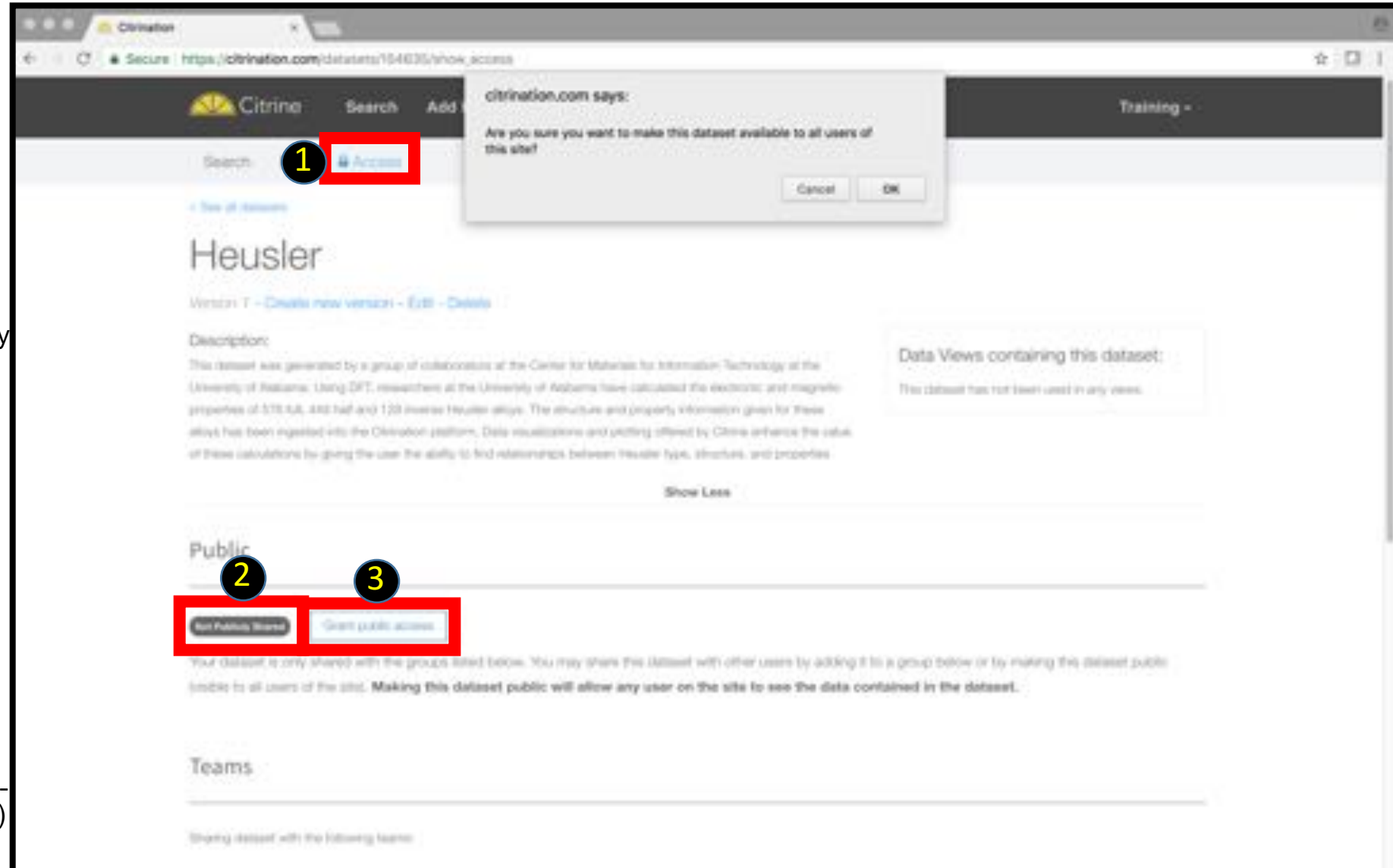
If you're a public Citrination user, this shares your data with everyone.

If you're a private Citrination user (you have your own site), only Citrine employees and users at your company can view it.

The process for sharing Data Views is similar- just click "Access"

1. Click Access Tab
2. Review current status
3. Click to Share

You can also share with Groups (short-term, can be added by your members) or Teams (long-term, managed by Citrine)



Create Models: Data Views

- **What is it?** This feature allows you to use Citrination's machine learning (ML) software to visualize and model data. You can predict the properties of untested materials or design new materials that meet your specifications. It also gives you a report on the ML model quality.
- **How does this help you?**
 - Use predict to assess material candidates of interest to you
 - Use design to suggest promising new candidates based on your target specifications
- **How it works:** The Citrination platform will build AI models to identify trends and make predictions.

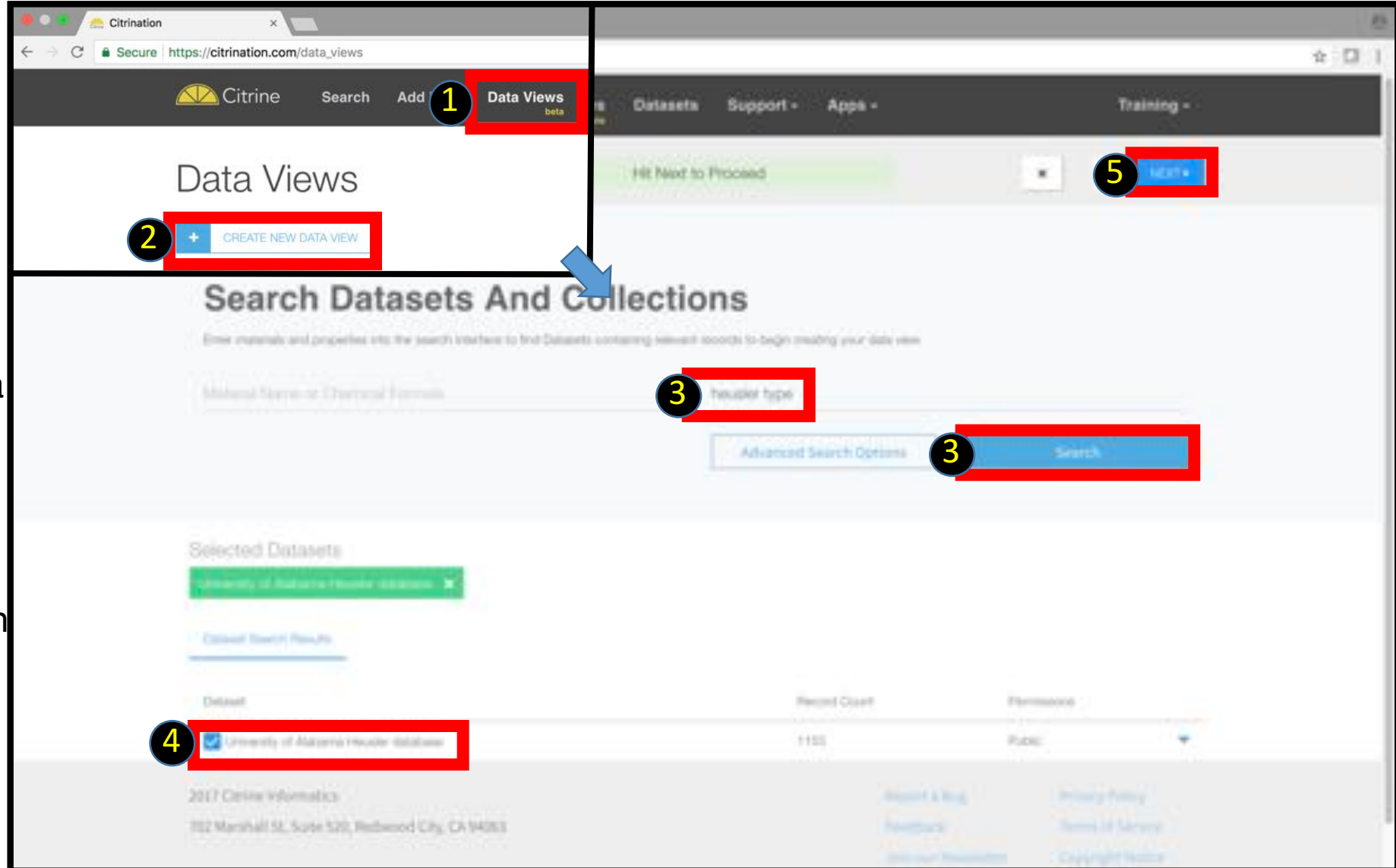
Data Views – Create (1/4)

Analyze your data in the Citrination platform by creating a Data View

1. Click Data Views tab
2. Click Create New Data View

Next, identify which datasets to use to train your model

3. Search based on properties contained in the dataset of interest
4. Select 1 or more dataset(s)
5. Click Next

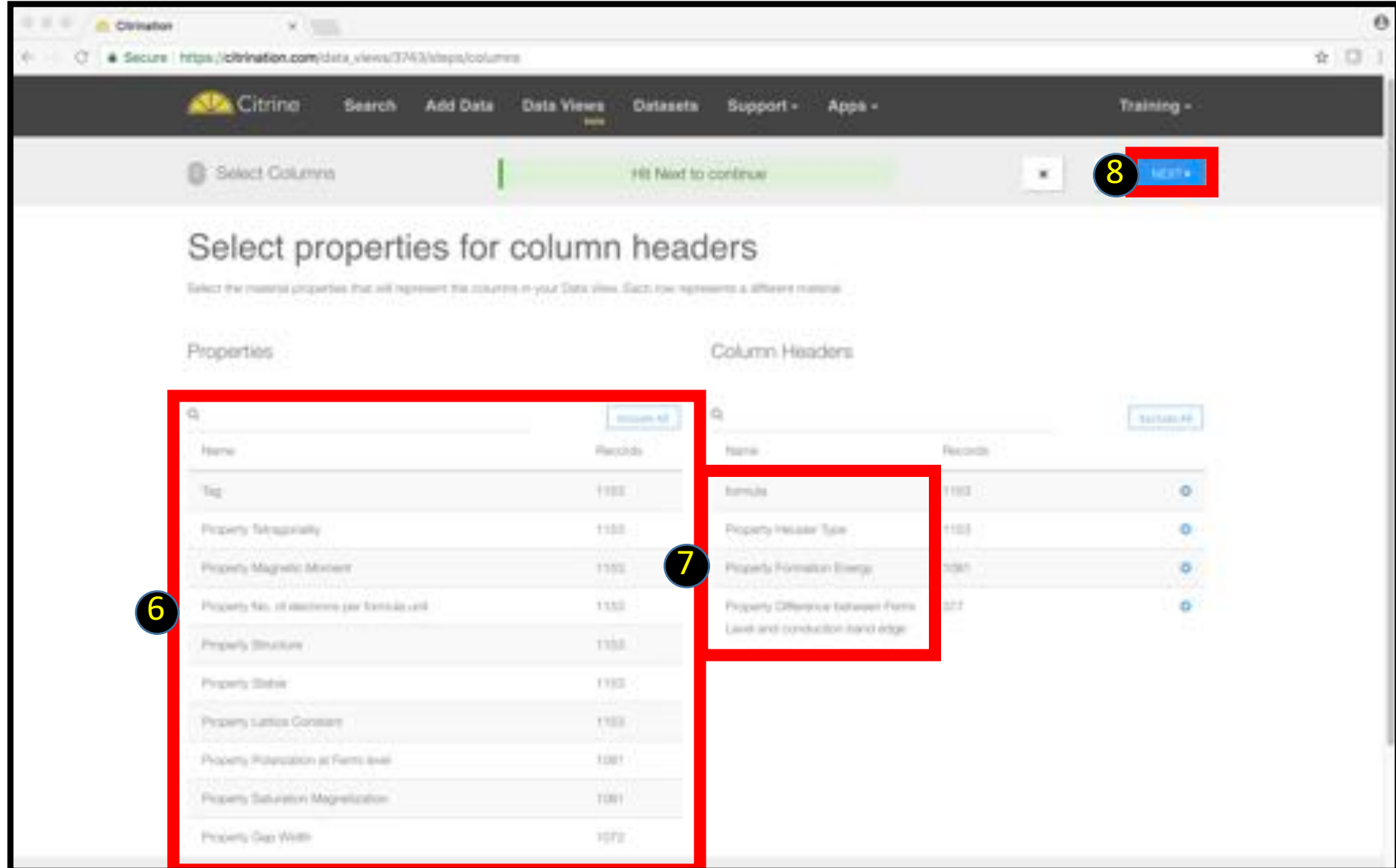


The screenshot shows the Citrination web application interface for creating a Data View. The steps are as follows:

- Step 1:** Click the "Data Views" tab in the top navigation bar.
- Step 2:** Click the "CREATE NEW DATA VIEW" button in the "Data Views" section.
- Step 3:** Enter a search term (e.g., "header type") in the search bar and click the "Search" button.
- Step 4:** Select the desired dataset(s) from the "Selected Datasets" list. In the example, "University of Arizona Header dataset" is selected.
- Step 5:** Click the "Next" button to proceed to the next step.

Data Views – Create (2/4)

6. Search/click properties you want to include as inputs or outputs to the Machine Learning (ML) model
7. Selected properties
8. Click Next



The screenshot shows the Citrine Informatics web interface for creating data views. The main heading is "Select properties for column headers". Below this, there are two tables: "Properties" and "Column Headers".

The "Properties" table is highlighted with a red box and a yellow circle with the number 6. It contains the following rows:

Name	Records
Tag	1100
Property Tetragonality	1100
Property Magnetic Moment	1100
Property No. of electrons per formula unit	1100
Property Structure	1100
Property State	1100
Property Lattice Constant	1100
Property Polarization at Fermi level	1001
Property Saturation Magnetization	1001
Property Gap Width	1070

The "Column Headers" table is also highlighted with a red box and a yellow circle with the number 7. It contains the following rows:

Name	Records
Formula	1100
Property Header Type	1100
Property Formation Energy	1001
Property Difference between Fermi Level and conduction band edge	1070

A "Next" button is highlighted with a red box and a yellow circle with the number 8.

Data Views – Create (3/4)

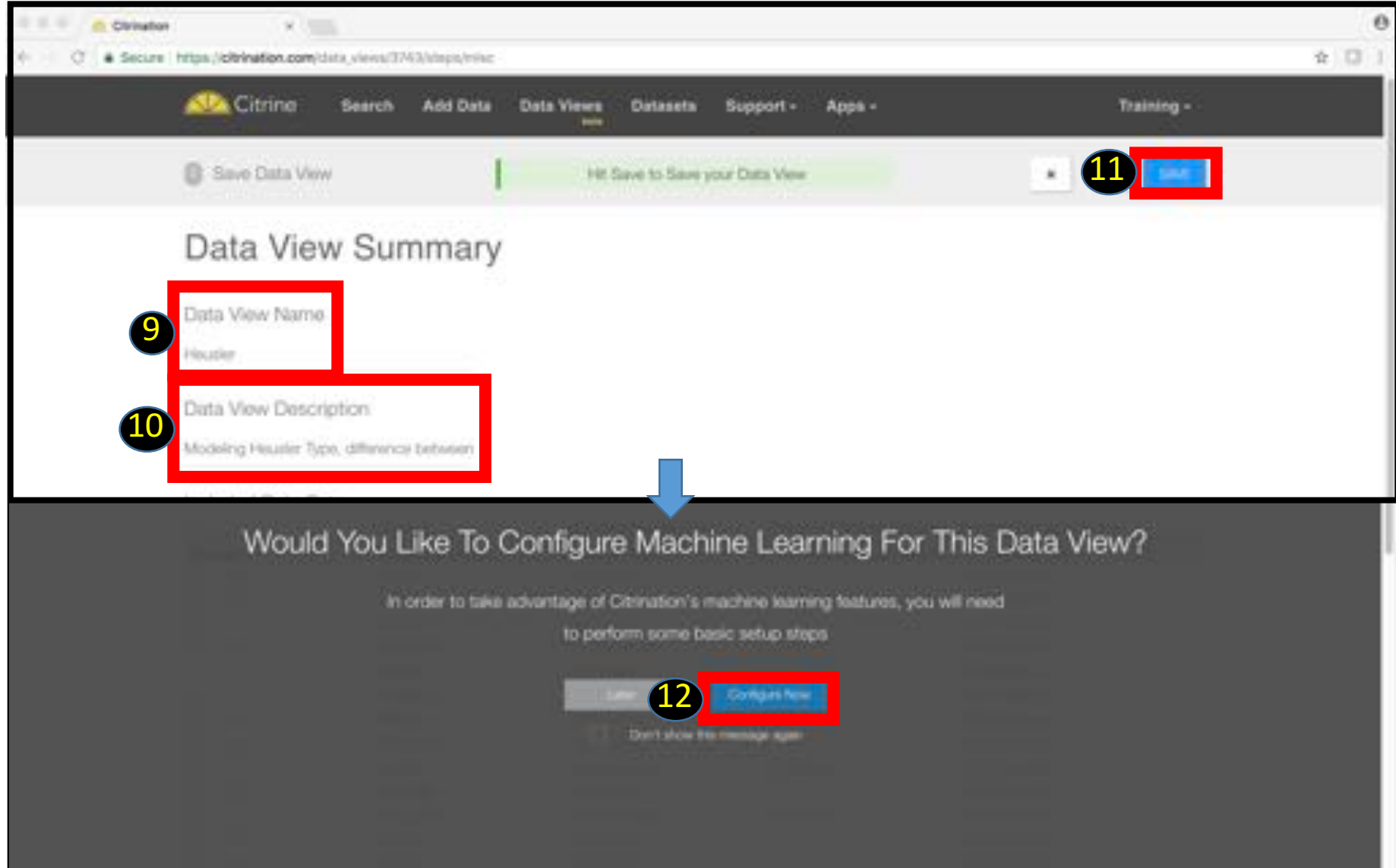
9. Type Data View Name

10. Type Description

11. Click Save

12. Click to Configure ML

(Once ML is configured, Citration can analyze the data and start to make predictions)



Save Data View | Hit Save to Save your Data View

11

Data View Summary

9 Data View Name
Husler

10 Data View Description
Modeling Husler Type, difference between

Would You Like To Configure Machine Learning For This Data View?

In order to take advantage of Citration's machine learning features, you will need to perform some basic setup steps

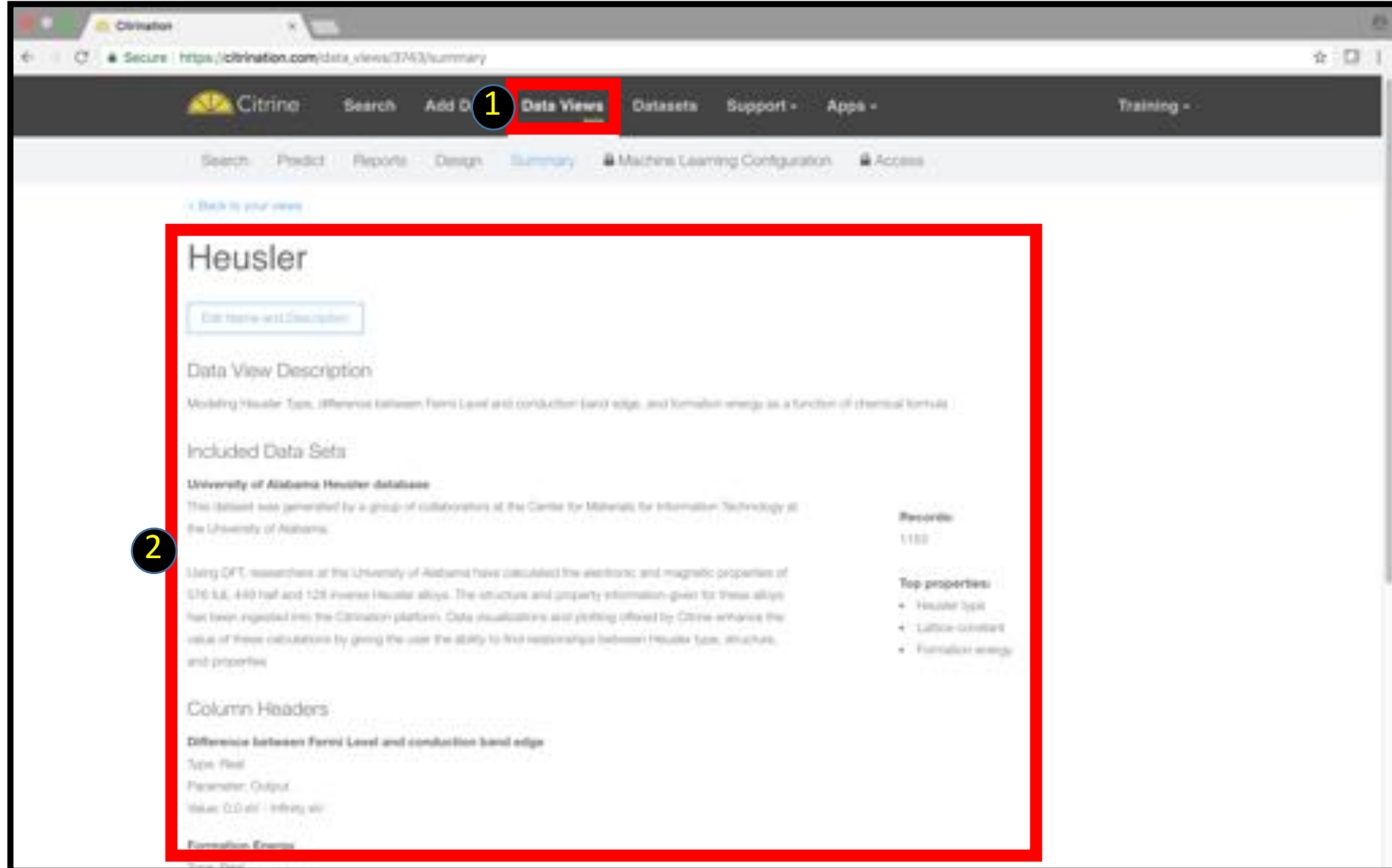
Later 12 Configure Now

Don't show this message again

Data Views – Summary

Review the important information on the Summary tab once you configure machine learning

1. Data View – Summary tab
2. Summary information



The screenshot shows the Citrine Informatics web interface. The top navigation bar includes the Citrine logo, a search bar, and links for 'Add Data View', 'Data Views' (highlighted with a red box and a yellow circle with the number 1), 'Datasets', 'Support', 'Apps', and 'Training'. Below this, a secondary navigation bar contains 'Search', 'Predict', 'Reports', 'Design', 'Summary' (highlighted with a red box and a yellow circle with the number 2), 'Machine Learning Configuration', and 'Access'. The main content area is titled 'Heusler' and contains a 'Data View Description' section, 'Included Data Sets' (listing the 'University of Alabama Heusler database'), 'Column Headers' (listing 'Difference between Fermi Level and conduction band edge'), and 'Top properties' (listing 'Heusler type', 'Lattice constant', and 'Formation energy').

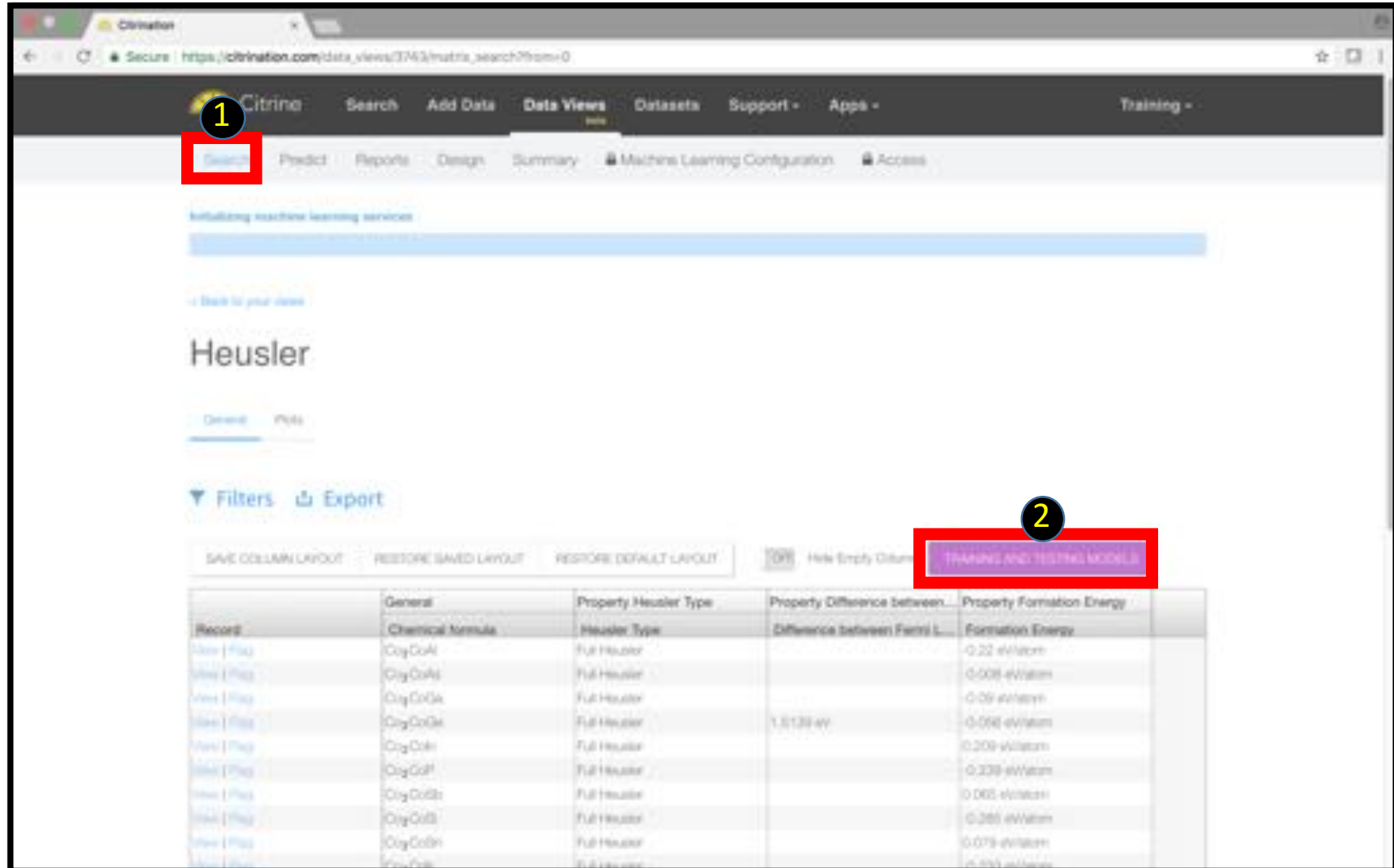
Data Views – Populate (1/2)

Fill in your data set's gaps with predicted values and uncertainty with the Populate button

1. Click Data Views-Search tab
2. Training and Testing Models

Citrine will begin training the models, and will display a purple "Training and Testing Data" button during this process. When the progress bars go away and the purple button is replaced by "Populate with Data" proceed to the next step.

Wait times are highly variable based on the amount and complexity of data. Most views train in 2-60 minutes. If it takes more than 12 hours, please contact Citrine.

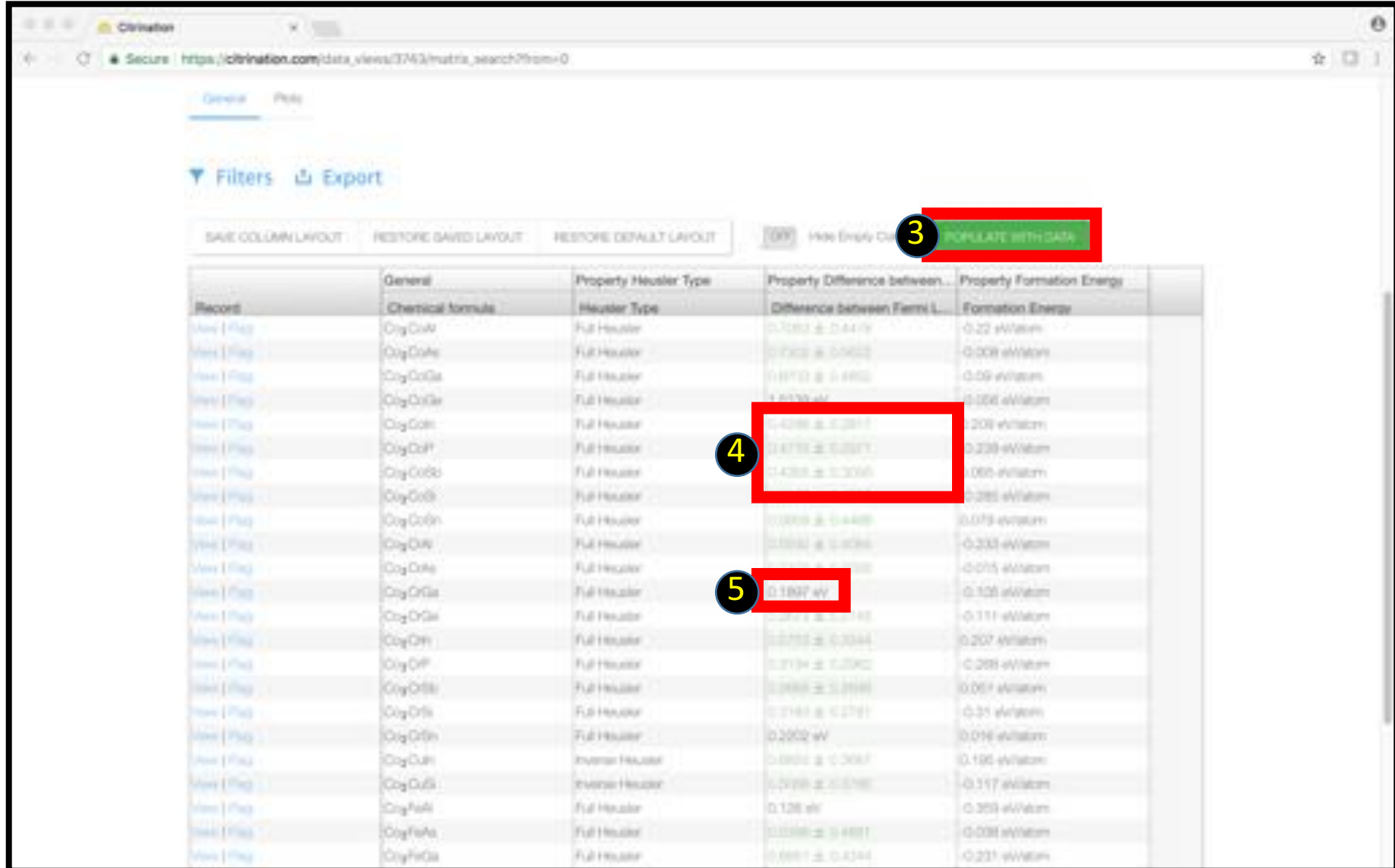


The screenshot shows the Citrine Informatics web application. The top navigation bar includes 'Citrine', 'Search', 'Add Data', 'Data Views', 'Datasets', 'Support', and 'Apps'. The 'Data Views' tab is active, and the 'Search' sub-tab is highlighted with a red box and a yellow circle with the number 1. Below the navigation bar, there is a progress bar for 'Initiating machine learning services'. The main content area is titled 'Heusler' and shows a table of data. The table has columns for 'Record', 'General', 'Property Heusler Type', 'Property Difference between...', and 'Property Formation Energy'. The 'Training and Testing Models' button is highlighted with a red box and a yellow circle with the number 2.

Record	General	Property Heusler Type	Property Difference between...	Property Formation Energy
	Chemical Formula	Heusler Type	Difference between Ferri L...	Formation Energy
View Plot	Co ₂ CoAl	Full Heusler		-0.22 eV/atom
View Plot	Co ₂ CoAl ₂	Full Heusler		-0.008 eV/atom
View Plot	Co ₂ CoGa	Full Heusler		-0.09 eV/atom
View Plot	Co ₂ CoGa ₂	Full Heusler	1.5129 eV	-0.066 eV/atom
View Plot	Co ₂ CoIn	Full Heusler		-0.209 eV/atom
View Plot	Co ₂ CoP	Full Heusler		-0.239 eV/atom
View Plot	Co ₂ CoSb	Full Heusler		-0.065 eV/atom
View Plot	Co ₂ CoSb ₂	Full Heusler		-0.265 eV/atom
View Plot	Co ₂ CoSn	Full Heusler		-0.079 eV/atom
View Plot	Co ₂ CoSn ₂	Full Heusler		-0.043 eV/atom

Data Views – Populate (2/2)

3. Click Populate with Data when it becomes visible
4. Predicted Values from the model are in **GREEN** with uncertainty
5. Recorded Values from your data source are in **BLACK**



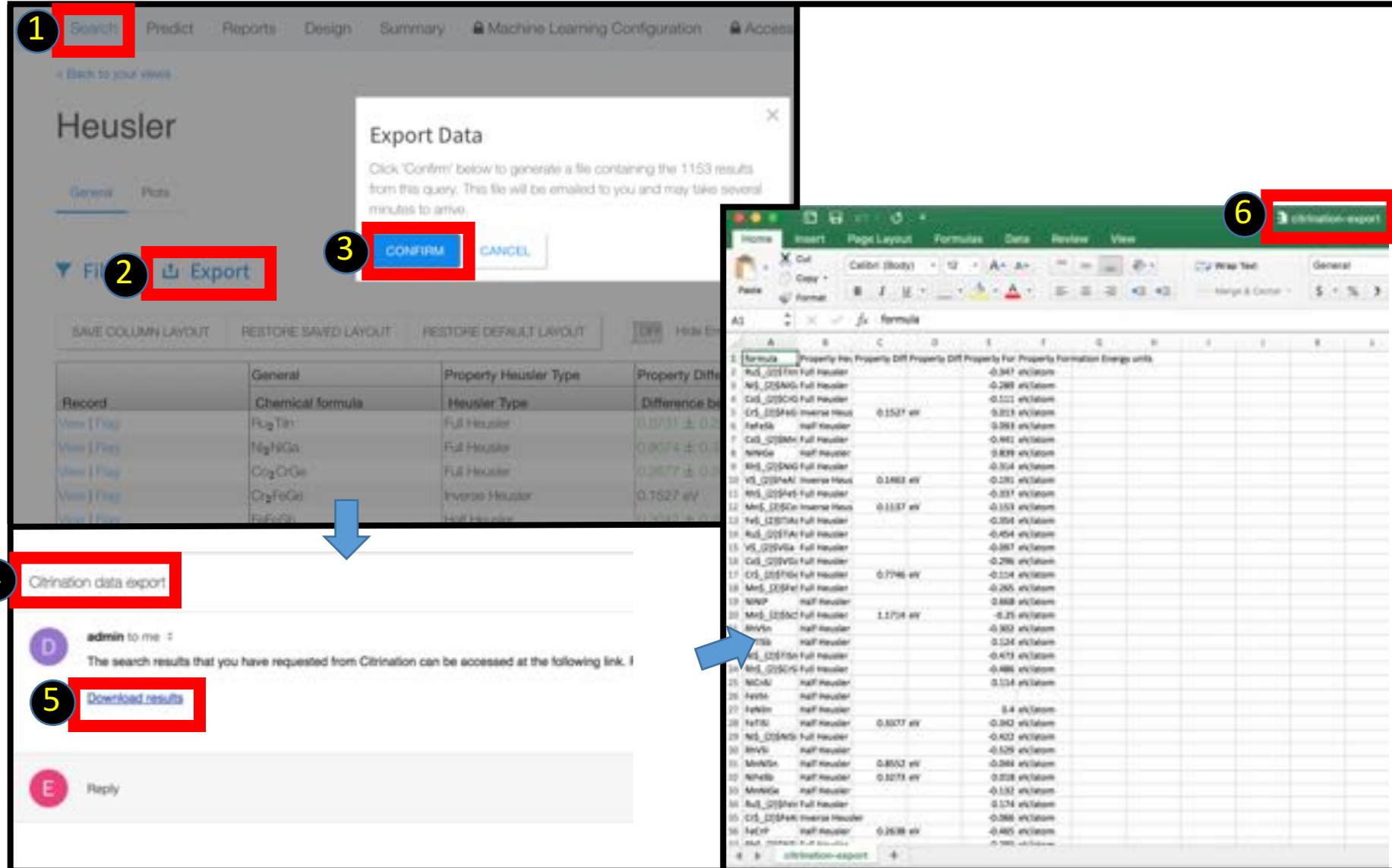
The screenshot shows the Citrine Informatics web interface. At the top, there are tabs for 'General' and 'Data'. Below the tabs are buttons for 'Filters' and 'Export'. A 'Populate with Data' button is highlighted with a red box and a yellow circle with the number 3. The main table has columns: 'Record', 'General', 'Property Header Type', 'Property Difference between...', and 'Property Formation Energy'. The table contains 20 rows of data. A red box with a yellow circle and the number 4 highlights a cell in the 'Property Difference between...' column. A red box with a yellow circle and the number 5 highlights a cell in the 'Property Formation Energy' column.

Record	General	Property Header Type	Property Difference between...	Property Formation Energy
	Chemical formula	Header Type	Difference between Fermi L...	Formation Energy
View Flag	C ₆ H ₆ CoW	Full Header	0.7082 ± 0.4478	0.22 eV/atom
View Flag	C ₆ H ₆ CoSe	Full Header	0.7302 ± 0.4602	0.008 eV/atom
View Flag	C ₆ H ₆ CoGa	Full Header	0.8173 ± 0.4892	0.09 eV/atom
View Flag	C ₆ H ₆ CoGe	Full Header	1.4135 eV	0.008 eV/atom
View Flag	C ₆ H ₆ CoIn	Full Header	1.4286 ± 0.2817	0.208 eV/atom
View Flag	C ₆ H ₆ CoPt	Full Header	1.4735 ± 0.2817	0.238 eV/atom
View Flag	C ₆ H ₆ CoSb	Full Header	1.4388 ± 0.3095	0.065 eV/atom
View Flag	C ₆ H ₆ CoSi	Full Header		0.285 eV/atom
View Flag	C ₆ H ₆ CoSn	Full Header	0.0009 ± 0.4488	0.078 eV/atom
View Flag	C ₆ H ₆ CoW	Full Header	0.0002 ± 0.4086	0.233 eV/atom
View Flag	C ₆ H ₆ CoSe	Full Header		0.015 eV/atom
View Flag	C ₆ H ₆ CoGa	Full Header	0.1807 eV	0.128 eV/atom
View Flag	C ₆ H ₆ CoGe	Full Header	0.2073 ± 0.2074	0.111 eV/atom
View Flag	C ₆ H ₆ CoIn	Full Header	0.0755 ± 0.2044	0.207 eV/atom
View Flag	C ₆ H ₆ CoPt	Full Header	0.2134 ± 0.2042	0.288 eV/atom
View Flag	C ₆ H ₆ CoSb	Full Header	0.2888 ± 0.2888	0.061 eV/atom
View Flag	C ₆ H ₆ CoSi	Full Header	0.2181 ± 0.2781	0.31 eV/atom
View Flag	C ₆ H ₆ CoSn	Full Header	0.2002 eV	0.016 eV/atom
View Flag	C ₆ H ₆ CoW	Inverse Header	0.0002 ± 0.3087	0.186 eV/atom
View Flag	C ₆ H ₆ CoSe	Inverse Header	0.2088 ± 0.2788	0.117 eV/atom
View Flag	C ₆ H ₆ CoGa	Full Header	0.128 eV	0.358 eV/atom
View Flag	C ₆ H ₆ CoGe	Full Header	0.0388 ± 0.4881	0.038 eV/atom
View Flag	C ₆ H ₆ CoIn	Full Header	0.0007 ± 0.4344	0.231 eV/atom

Data Views – Export

Manipulate and view the data in Excel by using the Export function

1. Click Data View-Search tab
2. Click Export
3. Click Confirm
4. Check email
5. Click Link
6. Download & Open File



1 Search

2 Export

3 CONFIRM

4 Citrine data export

5 Download results

6 citrine-export

Record	General	Property	Heusler Type	Property Difference
View Flag	Chemical formula		Heusler Type	Difference between
View Flag	Pu ₂ Ti ₃		Full Heusler	0.0731 ± 0.2
View Flag	Ni ₂ NiGa		Full Heusler	0.0074 ± 0.3
View Flag	Co ₂ CrGe		Full Heusler	0.0677 ± 0.3
View Flag	Cr ₂ FeGe		Inverse Heusler	0.1627 eV
View Flag	Fe ₂ Co ₂		Full Heusler	0.0007 ± 0.3

Data Views – Material Cards

View the properties of a material in the data view from the Data Views-Search tab

1. Click Data Views – Search tab
2. Click General
3. View material card

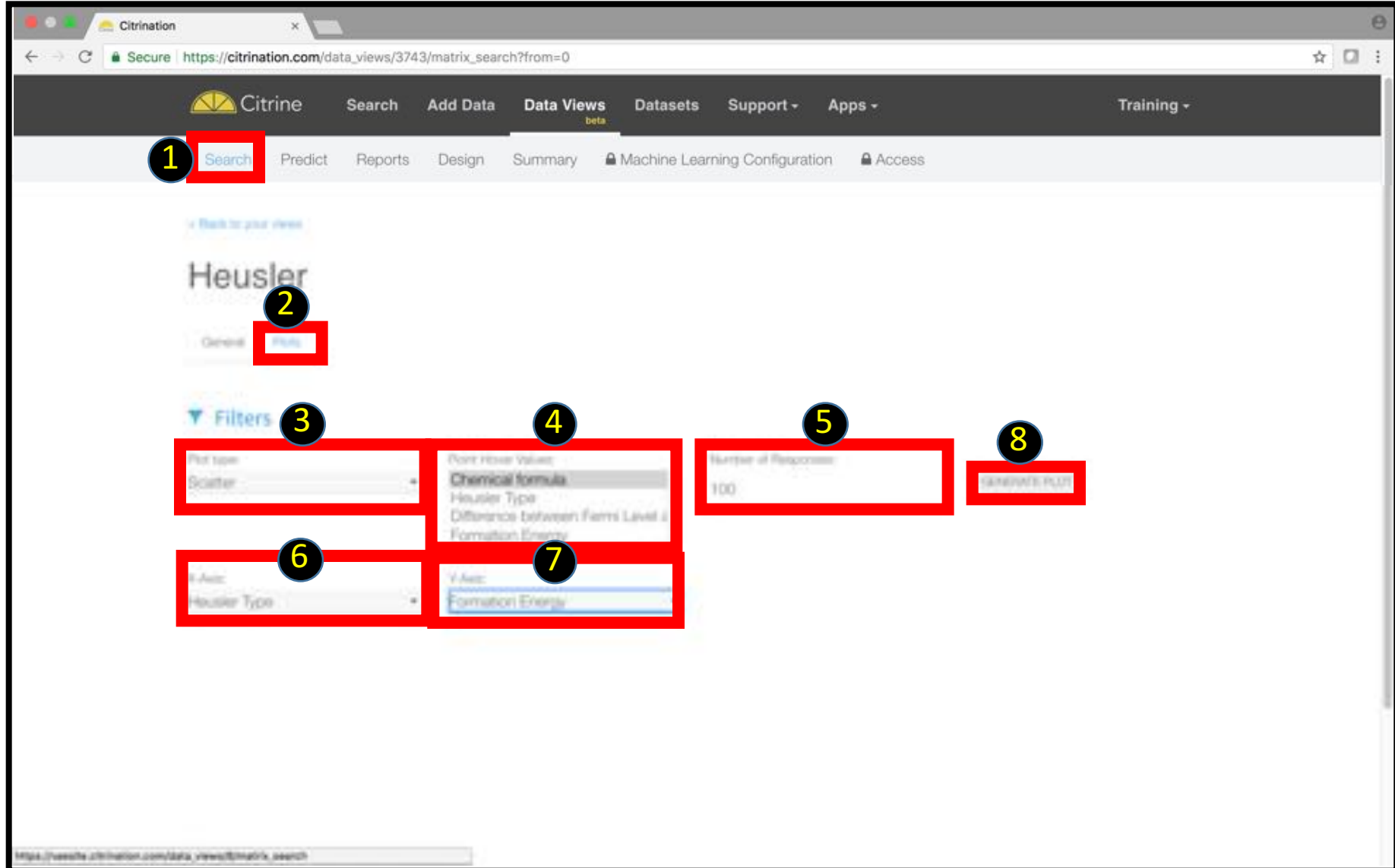
The screenshot shows the Citrine Data Views interface. The top navigation bar includes 'Citrine', 'Search', 'Add Data', 'Data Views', 'Datasets', 'Support', 'Apps', and 'Training'. A sidebar on the left has a 'Search' button highlighted with a red box and a yellow circle containing the number '1'. The main content area displays the search results for 'Heusler' materials. The interface includes a 'Back to your views' link, the search term 'Heusler', and tabs for 'General' and 'Plots'. Below the search term are 'Filters' and 'Export' buttons. A table of results is shown with columns: Record, General (Chemical formula), Property Heusler Type (Heusler Type), Property Difference between... (Difference between Fermi L...), and Property Formation Energy. The table lists 15 materials with their chemical formulas, Heusler types, and formation energies.

Record	General Chemical formula	Property Heusler Type Heusler Type	Property Difference between... Difference between Fermi L...	Property Formation Energy Formation Energy
View [Plot]	Pt ₂ Ti	Full Heusler	0.5731 ± 0.0291	0.347 eV/atom
View [Plot]	Pt ₂ NiGa	Full Heusler	0.0634 ± 0.0046	0.289 eV/atom
View [Plot]	Co ₂ CrGa	Full Heusler	0.2877 ± 0.0404	0.131 eV/atom
View [Plot]	Co ₂ FeGa	Inverse Heusler	0.1827 eV	0.013 eV/atom
View [Plot]	PdFeSi	Half Heusler	0.3042 ± 0.0008	0.030 eV/atom
View [Plot]	Co ₂ MnGa	Full Heusler	0.2853 ± 0.0018	0.441 eV/atom
View [Plot]	PdNiGa	Half Heusler	0.0530 ± 0.0030	0.608 eV/atom
View [Plot]	Pt ₂ NiGa	Full Heusler	0.0290 ± 0.0008	0.518 eV/atom
View [Plot]	Ni ₂ Ti	Inverse Heusler	0.1463 eV	0.181 eV/atom
View [Plot]	Pt ₂ NiSn	Full Heusler	0.0240 ± 0.0014	0.307 eV/atom
View [Plot]	Mn ₂ CoGa	Inverse Heusler	0.1137 eV	0.153 eV/atom
View [Plot]	Pt ₂ TiAs	Full Heusler	0.0381 ± 0.0048	0.354 eV/atom
View [Plot]	Pt ₂ TiAs	Full Heusler	0.3887 ± 0.1001	0.404 eV/atom
View [Plot]	Pt ₂ TiAs	Full Heusler	0.1000 ± 0.0000	0.767 eV/atom

Data Views – Plots

Use the Plots tab to visualize the data on a variety of plot types.

1. Click Data Views/ Search Tab
2. Click Plots
3. Select Plot Type
4. Select Point Hover values (if you put your cursor over the point, you see this data)
5. Type # responses
6. Select X Axis property
7. Select Y Axis property
8. Click Generate Plot

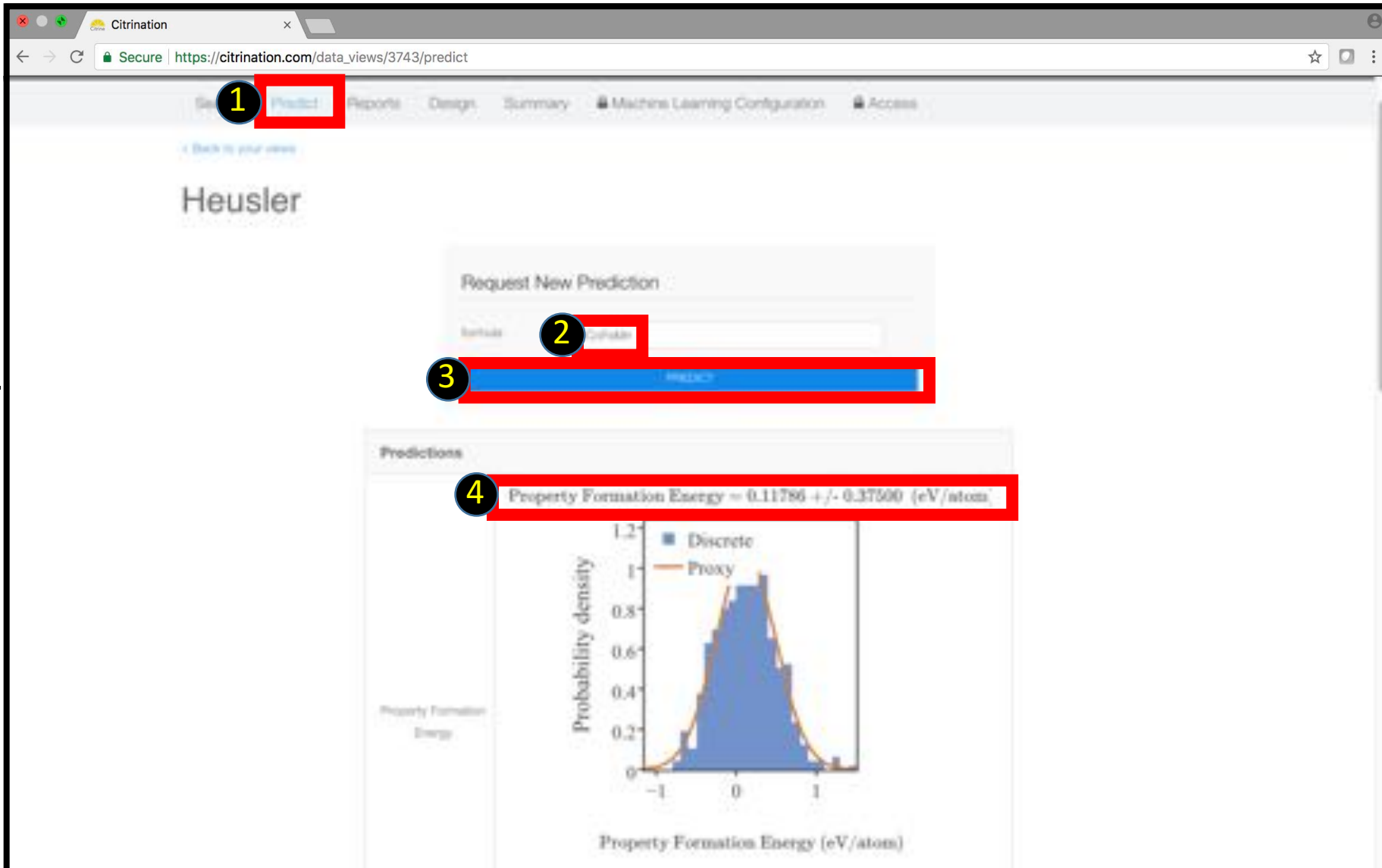


The screenshot shows the Citrine web interface for visualizing data. The top navigation bar includes 'Search', 'Add Data', 'Data Views', 'Datasets', 'Support', and 'Apps'. Below this is a secondary bar with 'Predict', 'Reports', 'Design', 'Summary', 'Machine Learning Configuration', and 'Access'. The main content area shows the 'Plots' tab selected, with a 'Filters' section and a 'Plot' section. The 'Plot' section has fields for 'Plot type' (Scatter), 'Point Hover Values' (Chemical formula, Heusler Type, Difference between Fermi Level & Formation Energy), 'Number of Responses' (100), and 'Generate Plot' button. The X and Y axes are set to 'Heusler Type' and 'Formation Energy' respectively.

Data Views – Predict

Use the Predict tab to predict the properties of a specific input (e.g. chemical formula)

1. Click Data Views – Predict tab
2. Type your input
3. Click Predict
4. Numerical prediction with uncertainty range of ± 1 standard deviation

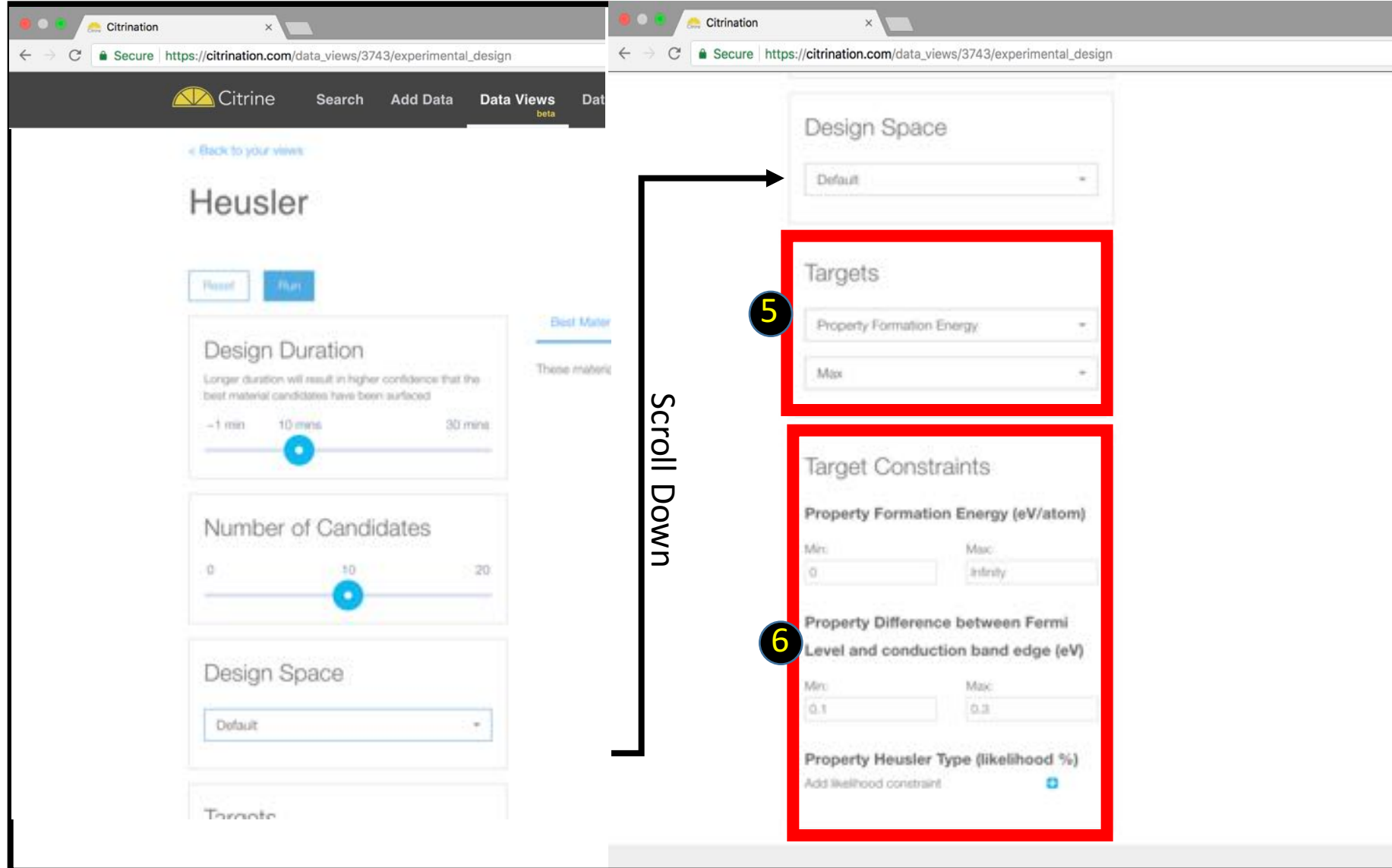


The screenshot shows the Citrine Informatics web interface for the 'Predict' tab. The interface includes a navigation bar with tabs: 'Select', 'Predict', 'Reports', 'Design', 'Summary', 'Machine Learning Configuration', and 'Access'. The 'Predict' tab is selected and highlighted with a red box and a yellow circle labeled '1'. Below the navigation bar, there is a 'Request New Prediction' section with a 'Formula' input field (containing 'Heusler') and a 'Predict' button (highlighted with a red box and a yellow circle labeled '3'). Below the 'Predict' button, the 'Predictions' section displays the result: 'Property Formation Energy = 0.11786 +/- 0.37500 (eV/atom)' (highlighted with a red box and a yellow circle labeled '4'). Below the prediction text is a plot showing the 'Probability density' on the y-axis (ranging from 0 to 1.2) and 'Property Formation Energy (eV/atom)' on the x-axis (ranging from -1 to 1). The plot includes a blue histogram labeled 'Discrete' and an orange line labeled 'Proxy'.

Data Views – Design (1/2)

Use the Design tab to generate candidate materials based on targets and constraints.

1. Click Data Views – Design tab
2. Select Maximum time for computer to explore options
3. Select Number of Candidates to return
4. Select space over which design will search for promising candidates
5. Select Optimized property and target
6. Select Constraints on the target properties
7. Click Run



The screenshot displays the Citrine web interface for the 'Design' tab. The left panel shows configuration options for a 'Heusler' material design, including sliders for 'Design Duration' (set to 10 mins), 'Number of Candidates' (set to 10), and a 'Design Space' dropdown (set to 'Default'). A 'Run' button is visible. The right panel shows a detailed view of the 'Targets' and 'Target Constraints' sections, which are highlighted with red boxes and numbered 5 and 6 respectively. A vertical arrow labeled 'Scroll Down' indicates the transition from the left panel to the right panel.

Design Space

Default

Targets

Property Formation Energy

Max

Target Constraints

Property Formation Energy (eV/atom)

Min: 0 Max: Infinity

Property Difference between Fermi Level and conduction band edge (eV)

Min: 0.1 Max: 0.3

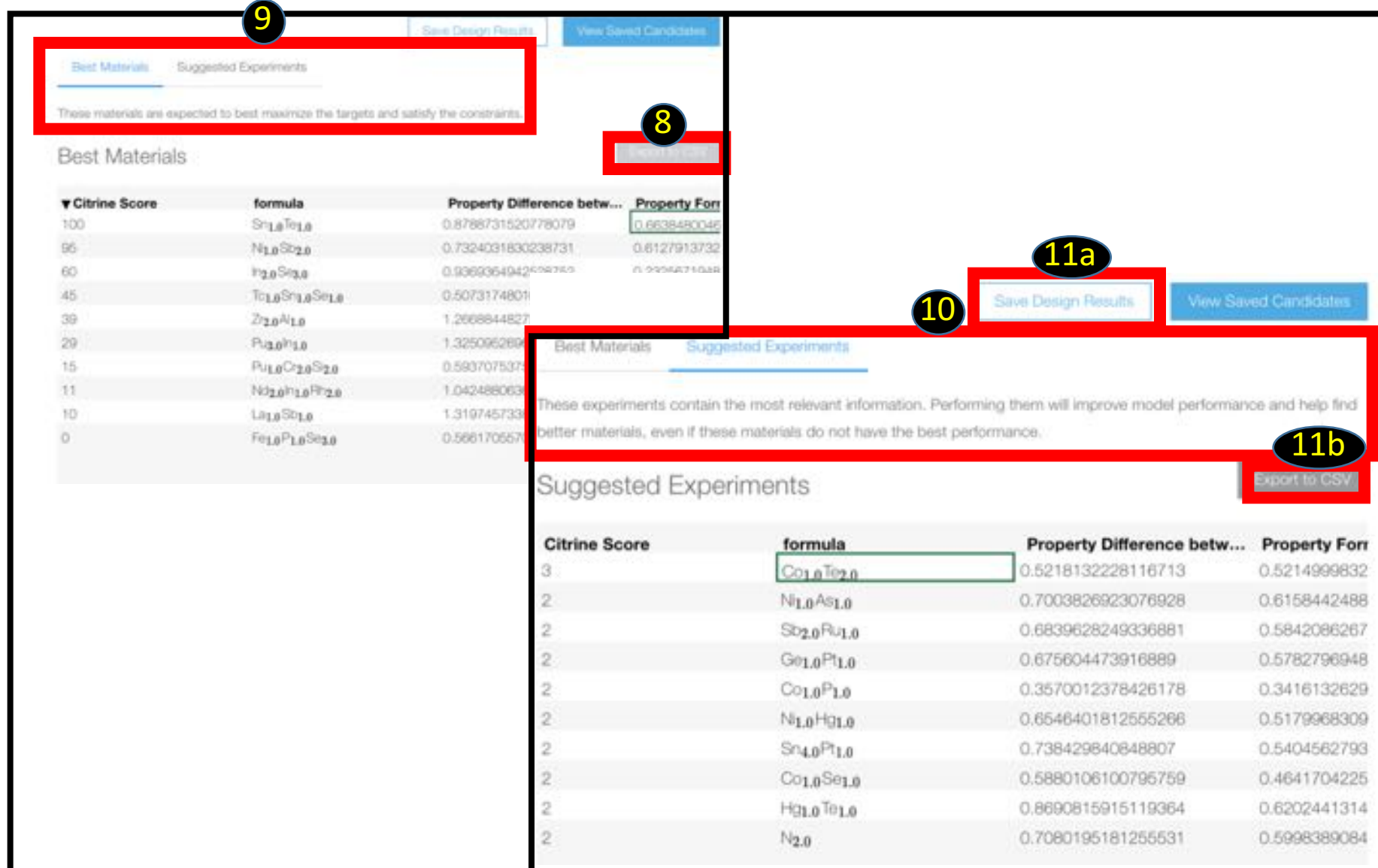
Property Heusler Type (likelihood %)

Add likelihood constraint

Data Views – Design (2/2)

8. Click Export to CSV to download results
9. Best Materials (short term success)
10. Suggested Experiments (long term success)
11. Click Save Design Results to save these candidates or Export to CSV to download results

(Citrination deletes unsaved design results when you leave the page)



Best Materials

These materials are expected to best maximize the targets and satisfy the constraints.

Citrine Score	formula	Property Difference betw...	Property For
100	$\text{Sn}_{1.0}\text{Te}_{1.0}$	0.8788731520778079	0.6638480045
95	$\text{Ni}_{1.0}\text{Se}_{2.0}$	0.7324031830238731	0.6127913732
60	$\text{In}_{2.0}\text{S}_{3.0}$	0.9369354942002753	0.5225271532
45	$\text{Te}_{1.0}\text{Sn}_{1.0}\text{Se}_{1.0}$	0.5073174801	
39	$\text{Zr}_{2.0}\text{Al}_{1.0}$	1.2668844827	
29	$\text{Ru}_{2.0}\text{Hf}_{1.0}$	1.3250952894	
15	$\text{Ru}_{1.0}\text{Cr}_{2.0}\text{Si}_{2.0}$	0.593707537	
11	$\text{Nd}_{2.0}\text{In}_{1.0}\text{Rh}_{2.0}$	1.042488063	
10	$\text{La}_{1.0}\text{Sb}_{1.0}$	1.319745733	
0	$\text{Fe}_{1.0}\text{P}_{1.0}\text{Se}_{1.0}$	0.566170557	

Suggested Experiments

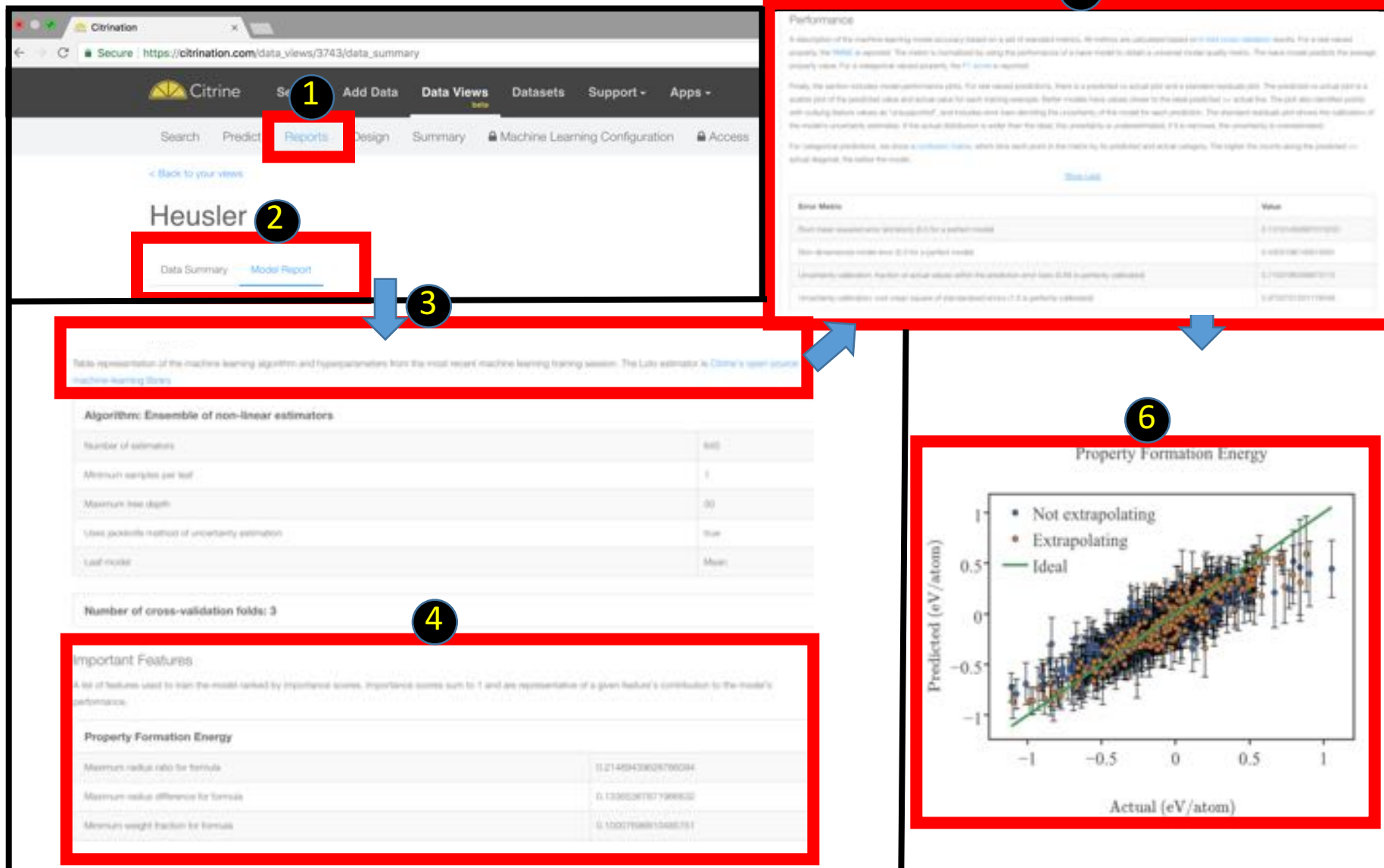
These experiments contain the most relevant information. Performing them will improve model performance and help find better materials, even if these materials do not have the best performance.

Citrine Score	formula	Property Difference betw...	Property For
3	$\text{Co}_{1.0}\text{Te}_{2.0}$	0.5218132228116713	0.5214999832
2	$\text{Ni}_{1.0}\text{As}_{1.0}$	0.7003826923076928	0.6158442488
2	$\text{Sb}_{2.0}\text{Ru}_{1.0}$	0.6839628249336881	0.5842086267
2	$\text{Ge}_{1.0}\text{Pt}_{1.0}$	0.675604473916889	0.5782796948
2	$\text{Co}_{1.0}\text{P}_{1.0}$	0.3570012378426178	0.3416132629
2	$\text{Ni}_{1.0}\text{Hg}_{1.0}$	0.6546401812555266	0.5179968309
2	$\text{Sn}_{4.0}\text{Pt}_{1.0}$	0.738429840848807	0.5404562793
2	$\text{Co}_{1.0}\text{Se}_{1.0}$	0.5880106100795759	0.4641704225
2	$\text{Hg}_{1.0}\text{Te}_{1.0}$	0.8690815915119364	0.6202441314
2	$\text{N}_{2.0}$	0.7080195181255531	0.5998389084

Data Views – Reports

You can use the Reports tab to understand your model quality.

1. Click Data Views-Reports tab
2. Click Model Report tab
3. Review ML settings
4. Review Features and their impact on the model
5. Review ML Model performance
6. Review predicted vs actual plot



The screenshot displays the Citrine web interface with the following components:

- Navigation Bar:** Includes tabs for Search, Predict, **Reports** (highlighted with a red box and number 1), Design, Summary, Machine Learning Configuration, and Access.
- Heusler Data Summary:** Shows a tab for "Data Summary" and a selected "Model Report" tab (highlighted with a red box and number 2).
- Performance Metrics:** A table showing various performance metrics (highlighted with a red box and number 5):

Score Metric	Value
Root-mean-square error (RMSE) for a perfect model	0.0000000000000000
Root-mean-square error (RMSE) for a perfect model	0.0000000000000000
Uncertainty estimation: Fraction of actual values within the prediction and less than 0.05 in probability (calibrated)	0.0000000000000000
Uncertainty estimation: Root-mean-square of standardized errors (RMSE) for a perfect model	0.0000000000000000
- Algorithm: Ensemble of non-linear estimators:** A table showing model settings (highlighted with a red box and number 3):

Parameter	Value
Number of estimators	100
Minimum samples per leaf	5
Maximum tree depth	30
Use probabilistic method of uncertainty estimation	True
Leaf model	Mean
- Important Features:** A table showing feature importance scores (highlighted with a red box and number 4):

Feature	Importance Score
Property Formation Energy	0.21459430626766094
Maximum radius ratio for formula	0.13065367907966632
Maximum weight fraction for formula	0.10007668815485751
- Property Formation Energy Plot:** A scatter plot showing Predicted (eV/atom) vs Actual (eV/atom) values (highlighted with a red box and number 6). The plot includes data points for "Not extrapolating" (blue dots) and "Extrapolating" (orange dots), along with an "Ideal" line (green line). The y-axis ranges from -1 to 1, and the x-axis ranges from -1 to 1.

Conclusion

We've learned how to:

- Search
- Upload Data
- Use Data Views to...
 - Create a view
 - Plot data
 - Predict properties of unknown materials
 - Design materials to meet parameters
 - Assess model quality



Citrine Informatics

The data analytics platform for the physical world

Thank you

elundberg@citrine.io