Searching Inorganic Chemistry
Agenda: Focus Inorganic Chemistry

• 1. Content in SciFinder
• 2. Indexing of Inorganic Substances
• 3. Appropriate Search Strategies
  – Chemical Name Searching
  – Molecular Formula
  – Structure Searching
  – Topic Searching
• 4. Combining Answers
• 5. Sources, Thanks and Questions
1. Content: Inorganic Chemistry is a cornerstone of the CAS databases

- The following CA Sections are for Inorganic Chemistry
  - 21: Organometallic and organometalloidal compounds
  - 49: Industrial inorganic chemicals
  - 54: Extractive metallurgy
  - 55: Ferrous metals and alloys
  - 56: Nonferrous metals and alloys
  - 57: Ceramics
  - 58: Cement, concrete, and related building materials
  - 67: Catalysts, reaction kinetics and inorganic reaction mechanism
  - 78: Inorganic chemicals and reactions
  - 79: Inorganic analytical chemistry
- They represent about 13% of CAS’ coverage
1. Content: Registry lists different types of Inorganic Compounds

- 1,079,682 Alloys
- 2213862 Coordination Compounds (2,579,146 when you count also salts including a coordination compound)
- 645,444 Tabular Inorganics
- 9,046 Minerals

- Some of the following Class Identifiers also describe Inorganic Compounds:
  - 286,384 Incompletely Defined Substances
  - 184,919 Mixtures
2. Indexing Inorganic Compounds

- Elements and Isotopes
- Minerals
- Salts
- Alloys
- Tabular Inorganic Compounds
- Incompletely Defined Substances
- Mixtures
- Coordination Compounds
2.1 Indexing Elements and Isotopes

Elements and their isotopes receive their own CAS RN. The same is true for all kind of ions!
## 2.1 Indexing Elements and Isotopes

### Ions of different Boron isotopes receive their own CAS RN.

<table>
<thead>
<tr>
<th>Isotopes</th>
<th>Mass</th>
<th>CAS RN</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 B</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>10 B</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>Boron, isotope of mass 10</td>
</tr>
<tr>
<td>B +</td>
<td></td>
<td>Boron, ion (B1+)</td>
</tr>
<tr>
<td>B 2+</td>
<td></td>
<td>Boron, ion (B2+)</td>
</tr>
<tr>
<td>9 B</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>8 B</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>Boron, isotope of mass 8</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>Boron</td>
</tr>
<tr>
<td>7440-42-8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.2 Indexing Minerals

Minerals are naturally occurring salts or mixtures of salts. Due to the fact that their composition may vary, there are many „Incompletely Defined“ or even „Manual Registrations“. There are always (at least) two CAS RNs for a mineral: one under the mineral name and one with the exact same molecular formula with the chemically synthetic version.

Please keep in mind that all CAS databases are in English! Lapilazuli or Lasurit (both German) will give no answers: you need to search Lazurite.[1]
2.2 Indexing Minerals

Minerals might be registered in different ways!

Even an unspecified manual registration for the group of Apatites
2.3 Indexing Salts: Salts That Do Not Contain Carbon

Simple salts from acids not containing Periodic Table Group VI atoms (e.g. oxygen and sulfur) are registered in the way normally drawn by the chemist. For example, sodium chloride and calcium bromide are represented as NaCl and CaBr2 respectively.

**Salts from acids containing Periodic Table Group VI atoms**

However, salts from acids containing Periodic Table Group VI atoms and bases containing Periodic Table Group I and II atoms (e.g. sodium, potassium, calcium, barium) are registered as the free acid combined with the base (which is given the metal symbol).

For example, although chemists write the formula for the calcium phosphate as Ca3(PO4)2 the substance is registered as if the hydrogens were still attached to the phosphate group, i.e. as phosphoric acid, H3PO4. As the actual salt has three Ca atoms Registration of Substances 193 to two phosphate groups, the substance is considered as Ca3(H3PO4)2, and the molecular formula, which has the first component in the alphabet as a single atom, then becomes Ca. 2/3 H3 O4 P.

Actually, there are two substances called calcium phosphate in the database and only one is shown here (the Molecular Formula Field entry for the second substance is Ca. x H3 O4 P). The issue is the chemical description of the substances in the original literature.

If the particular form is not specified in the article, the substance is registered as that in which the ratio is not specified. There are currently more than 10,000 entries for this second substance, so a search for information on calcium phosphate should probably include its CAS Registry Number (10103-46-5) as well. [1]
2.3 Indexing Salts: Salts That Do Not Contain Carbon

The most simple case: Sodium Chloride

CAS Registry Number: 7647-14-5

Sodium chloride (NaCl)
Salt (6Cl,7Cl); Sodium chloride (BCD); Adsorben; Arm-A-Vial; Ayr; BCD; Brinekote Superfine; Canners 999; Cap Shure 70 Salt; Cap-Shure 85 Salt; Common salt; D 02056; Dendritic salt; Dendrite; Extra Fine 200 Salt; Extra Fine 325 Salt; Flexivale; Gingivyl; H.G. Blending; Hyperseal; Iodized salt; Kochsalz; LS 1700; Mahr; NES 77264; Nakuru N; Naruto Shokuan; Natraphic chloride; Natrum mur; Nisoi-no Yakishio E; SS Salt; Seas salt; Slow Sodium; Sodium chloride; Sodium monochloride; Soft Salt S 30; Solose; Special Salt 100/25; Suprasel; Suprasel Nitrite; Table salt; Titrisol; Top flake; Uzushio Biryou M; Walesal A; White
2.3 Indexing Salts: Salts That Do Not Contain Carbon

Caution: Confusion possible! In the molecular formula, the first component in the alphabet must be a single atom (or substance)! But not in the structure representation ...
2.3 Indexing Salts: Organic Acids

Salts from Organic Acids and Periodic Table Group I and II Bases

The registration is similar to the registration described above. Note that the molecular formula for sodium acetate is listed as C2H4O. Na, whereas a chemist would represent the formula for the substance as C2H3ONa. [1]

“Dot Disconnected Molecular Formula“ = Component based Registration!
2.3 Indexing Salts: Nitrogen-Containing Bases

Salts with Nitrogen-Containing Bases

When the base is an amine the salts are represented as the free base and the free acid. Salts from nitrogen-containing bases involving other acids (e.g. sulfuric acid and phosphoric acid) are registered similarly (e.g. ammonium sulfate is the multicomponent substance with the molecular formula $\text{H}_3\text{N} \cdot \frac{1}{2} \text{H}_2 \text{O} \cdot \text{S}$) although the molecular formula for ammonium chloride is $\text{ClH}_4 \text{N}$. These substances are most easily found through searches based on names. [1]
2.4 Indexing Alloys

Alloys
When specified in the original article, the composition of the alloy is listed in the name and composition fields and the constituent elements are listed in the formula field. A more generic description (see “Stainless Steel”) is applied when the elemental composition is not precisely specified in the original article. [1]

All alloys with a component table also have a molecular formula containing all elements separated by dots!
2.5 Indexing Tabular Inorganics

Tabular Inorganics
Very similar to Alloys, Tabular Inorganic registrations are Multi Component registrations, which have also a Component Table. Here however the component ration is not given in percent weight, but in element ratios.
2.5 Indexing Tabular Inorganics

All Tabular Inorganics with a component table also have a molecular formula containing all elements separated by dots and an additional molecular formula with the ratios!
2.6 Indexing Mixtures

Mixtures
Mixtures are registered where two or more chemically discrete components have been mixed together for a specific use (e.g. formulations involving pharmaceutical and agricultural chemicals). In general, host–guest complexes are also considered as mixtures. [1]

Biological active mixtures containing CuSO4 (eg. for fungicides)
2.7 Indexing Incomplete Defined Substances

Incompletely Defined Substances

Incompletely defined substances are those that have a known molecular formula but for which the complete valence bond structure was not fully described in the original article. For example, while $o$, $m$, and $p$-xylene are the specific dimethylbenzenes, if only ‘xylene’ is mentioned in the original article, then the incompletely defined substance is indexed.

Similar issues are encountered with salts in which ions have different possible ratios. [1]
2.8 Indexing Coordination Compounds

The representation of the structures of some coordination compounds requires modifications to normal valence bond definitions. Generally, the electrons involved in bonding the organic groups to the metal are provided by the organic groups as either $\sigma$–donors and $\pi$-donors. In broad terms, these are distinguished in that the electrons in the former case come from atoms, whereas in the latter case they come from double bonds.

Pay attention to the fact, that several substances (which a chemist would consider to be a coordination compound) are not classified as Coordination Compounds by CAS. CAS only assigns them to be Coordination Compounds if there are more connections to the metal atom than valences. See next slide!

$\sigma$-Complexes
Generally, charges in structures relate to the species involved in the preparation of the complex. For example, cisplatin (Figure A4.25) is made from Pt$^{2+}$, Cl$^-$, and NH$_3$.

$\pi$-Complexes
In $\pi$-complexes an extra ‘bond’ is drawn between the atoms involved in the complex.

[1]
2.8 Indexing Coordination Compounds

Coordination Compound according to CAS rules

No Coordination Compound according to CAS rules
2.8 Indexing Coordination Compounds

The salt is not classified as a „Coordination Compound“! But it will be retrieved in SF even if you limit to „Coordination Compound“.
3. Appropriate Search Strategies

- 3.1 Chemical Name Searching
- 3.2 Molecular Formula
  - Exact MF
  - Open MF
- 3.3 Structure Searching
  - For Salts
  - For Coordiantion compounds
- 3.4 Topic Searching
3.1 Chemical Name Searching

- Works very nice with simple names like minerals, easy salts, trade names ...
  - Try the following examples: zirconia, diamond, potassium bromide, talcum, Nirosta

- In Searching by Substance Identifier it’s for sure also possible to search for a CAS RN
  - Please be aware that many CAS RN in the Internet or in catalogues are simply wrong!
  - [http://www.cas.org/expertise/cascontent/registry/checkdigit.html](http://www.cas.org/expertise/cascontent/registry/checkdigit.html)
3.1 Chemical Name Searching

In SciFinder, you can search for substances by chemical name. The search bar allows you to enter chemical names, such as "zirconia; diamond; potassium bromide; talcum," and then select "Search." The system will display substances matching the entered names, as shown in the image. This feature is useful for quickly finding specific chemicals in a large database.
3.1 Chemical Name Searching

Element modifications have own CAS Registry numbers!
3.2 Molecular Formula Searching

• Always remember the first rule of MF searching: „Never trust the MF you’ve calculated!“
  – 0 results are often the consequence of a wrong entry 😞

• Remember the different conventions of MF indexing - as outlined before in this lesson
  – DotDisconnected Formulas for Salts, Mixtures etc.
  – Registration rules for Salts

• It is clever to use MF searching, if you look for Alloys or Tabular Inorganics if you only know the elements which should be present
3.2 Molecular Formula Searching

Searching for the MF might be tricky! Searching Potassium permanganate and forget the dot – no result, forget the H – no result!
3.2 Molecular Formula Searching - open

In order to search for the brand new superconductors (Iron pnictides ...no typo 😊), just leave the MF open by using only the dots!

It’s not possible in SciFinder to search for ratios or ratio ranges.
3.2 Molecular Formula Searching - open

Sort Substances by „Number of References“ to bring most cited substances to the front!
3.2 Molecular Formula Searching - open

Get the references for your Iron Pniclides and save them as a "Saved Answer Set" for Refining and Combining!
3.3 Structure Searching: General remarks

- Always switch on „Precision analysis“ when you do substructure searches
- Don‘t be too sophisticated in the first step, always start as simple as possible
- Don‘t be shy to draw fragments and don‘t connect them in the first step
- Substructure searches in Inorganic Chemistry almost always run to completion as long as you have any element other than C,N,O,S... involved
3.3 Structure Searching for Salts

Use „Undefined Bonds“ if you are not sure about the correct bond order!
3.3 Structure Searching for Salts

Check „Single Component“ to retrieve exactly as many components as you have drawn (here 3).
3.3 Structure Searching for Salts

SciFinder is a trademark of the American Chemical Society

SciFinder®
Welcome Thomas Haubrich | Sign Out

Chemical Structure exact with limits > substances (9)

Substances
Get References  Get Reactions  Get Commercial Sources  Combine Answer Sets

Save  Print  Export

View:  |  |  |  |  |

1. Substance Detail
20670-20-6
(Component: 7664-38-2)

Cu(II)

Na

Cu·9/9 H3O4·P·2/3 Na
Phosphoric acid, copper(2+) sodium salt (9:9:2) (PCI)

2. Substance Detail
63257-73-6
(Component: 7664-38-2)

Cu(II)

Na

Cu·3/2 H3O4·P·3/2 Na
Phosphoric acid, copper(2+) sodium salt (3:2:3) (PCI)

3. Substance Detail
477779-78-5
(Component: 7664-38-2)

Cu(II)

Na

Cu·7/9 H3O4·P·9/9 Na
Phosphoric acid, copper(2+) sodium salt (7:9:9) (PCI)

4. Substance Detail
141262-16-0
(Component: 7664-38-2)

Cu(II)

Na

Cu·3/2 H3O4·P·3/2 Na
Phosphoric acid, copper(2+) sodium salt (3:2:3) (PCI)
3.3 Structure Searching for Coordination Compounds

1. Bonds to metals in queries may be ignored in the initial structure search. In these cases, “Show precision analysis” is necessary to obtain more precise answer sets. Always check this box before you start a search!

2. It is not necessary to insert charges in structure queries.

3. This representation of $\pi$-complexes may produce structures where the normal valencies of atoms are exceeded (e.g. ferrocenes derived from pentamethylcyclopentadiene), but structure searches proceed as expected.
3.3 Structure Searching for Coordination Compounds

Draw a bond from every \( \pi \)-donor to the central metal – that’s how CAS stores the structure of coordination compounds! (even if it looks funny)
3.3 Structure Searching for Coordination Compounds

Forget about overlaps!
3.3 Structure Searching for Coordination Compounds

Leave „Single Component“ open to retrieve all substances of which titanocene is a component.
3.3 Structure Searching for Coordination Compounds

Precision analysis is important – check out the results and compare them! (use browser back button – as an exception!)
3.3 Structure Searching for Coordination Compounds
3.4 Topic Searching for Inorganic Chemistry

• Always perform a Topic Search in addition to your substance search based strategy (MF, Substance Identifier and/or Structure Search)

• Always save your results as a “Saved Answer Set” in SFoW, so that you can combine it with other answer sets later
3.4 Topic Searching for Pnictides

Search for „Iron Pnictides“ – if you are not sure what SciFinder does in the background, just enter additional synonyms in brackets!
3.4 Topic Searching for Pnictides

In this case the synonym was already found by SciFinder: iron=Fe for SciFinder!
3.4 Topic Searching for Pnictides

This area of research is so brand new, that you would have certainly missed relevant articles, if you would have only searched for substances!
4. Combining Answer Sets

- **Especially in hard to index areas of science, like in some of the fields of Inorganic Chemistry, several strategies might reveal unique hits.**

- **Use the power of the „Combine Answer Sets“ feature of SFoW to combine the answers before further drill down with „Refine“, „Analyze“ or „Categorize“**
4. Combine Answer Sets from Topic Searching and MF Searching
4. Intersect the sets to see all answers included in both sets
4. Intersect result in 41 answers

Only 41 references where found with both strategies. 100 additional articles from substance searching and several hundreds from the topic search!
4. Use Categorize to explore the Combined Answers

Categorize is a perfect tool to drill down to specific index terms assigned by CAS scientists.
5. Sources and Thanks!

• Various SciFinder Tutorials and E-Seminars can be found on

• A very good, problem oriented textbook for a deeper understanding of SciFinder and the CAS databases:
  – I cited several definitions from the appendices of that book for this presentation (marked with [1]).