

Third Year Research Information Skills IT Practical

This practical introduces the concepts and methodology for searching the research literature using online databases. These databases allow sophisticated search queries to be performed on the millions of chemistry and scientific articles published over the last 100 years. The practical will compare and contrast several of the main databases and allow students to discover the most appropriate database to find the research papers and patents.

The databases considered in the practical are:

Google Scholar

<<http://scholar.google.com/>>

"Google Scholar provides a simple way to broadly search for scholarly literature. From one place, you can search across many disciplines and sources: articles, theses, books, abstracts and court opinions, from academic publishers, professional societies, online repositories, universities and other web sites. Google Scholar helps you find relevant work across the world of scholarly research"

Web of Knowledge

<<http://wok.mimas.ac.uk/>>

*"The ISI Web of Knowledge Service for UK Education provides a single route of access to Thomson Reuters's products subscribed to by an individual institution. It includes **Web of Science**; ISI Proceedings; Journal Citation Reports; Current Contents Connect; Derwent Innovations Index and many others"*

Please **NOTE** when you use Web of Knowledge select the tab **Web of Science** to focus just on the scientific literature.

Scopus <<http://www.scopus.com/>>

"Scopus is the largest abstract and citation database of peer-reviewed literature and quality web sources with smart tools to track, analyze and visualize research"

Reaxys <<https://www.reaxys.com/>>

"Offering a wealth of experimentally validated information, Reaxys brings a fresh look to synthetic chemistry with powerful functionality, combined content and relevant information"

Methodology for Searching

Consider your research topic, and break the topic into the most important concepts – these are your search terms. For example, if your research topic is "The use of chiral catalysts in asymmetric hydrogenation of alkenes", your significant search

terms are: chiral catalysts, asymmetric hydrogenation and alkenes.

You could enter this search as **chiral catalysts** and **asymmetric hydrogenation** and **alkenes**.

However, most databases allow us to improve the quality of search results by the use of the following simple techniques:

Similar terms or synonyms

Think of alternative or similar terms (synonyms) and put an **or** between them Eg. alkenes or olefins

Plurals

Use a * to find words starting with a common stem. The * is a truncation symbol replacing one or no letters Eg **alkene*** will find alkene and alkenes, **Cataly*** will find catalysts, catalyst, catalysis, catalyze etc.

Using terms and, or, & not

These are Boolean operators and combine search terms.

AND will find all search terms, **OR** will find one or another, **NOT** will exclude terms

Remember: if you combine the operators **and** and **or** in a search phrase you must enclose the **or** terms in brackets eg cats and (dogs or puppies). So, in our sample search, asymmetric hydrogenation and (olefin* or alkene*) will find references that include asymmetric hydrogenation, and either olefins or olefin or alkenes or alkene.

Using these tips our sample search will look like this: chiral cataly* and asymmetric hydrogenation and (olefin* or alkene*)

Using same – another Boolean operator

References that contain terms in the same sentence are likely to be more relevant than those that contain the terms in different parts of the reference. If your search returns too many hits, try replacing and with same.

Eg: catalyst same rhodium will find both terms in the same sentence.

Tasks 1

Lets start with searching the query mentioned above **chiral cataly* and asymmetric hydrogenation and (olefin* or alkene*)** in all four databases and compare their results.

	Google Scholar	Web of Science	Scopus	Reaxys
How many research papers are returned with this query?				
Identify the authors and University or Company for the most recent record returned?				
What journal has the most recent work been published in?				
What is the title of the research work?				
Can you identify many patents in this area?				

10 Marks

Task 2

Do you think there should be more or less research papers published in this area?

Are you surprised about how little or how many papers this search found?

Suggest a sensible number of research papers that this field would cover and give your reasons for this suggestion.

Based on what you might think are a sensible number of research papers, which cover this area, can you adjust the search query to find a more sensible number of research papers?

Just focus on Web of Science and Scopus for the results.

	Web of Science	Scopus
How many research papers are returned with this query?		
What was the search query you used?		

Do the two databases now agree with the number of research papers published in this area?

10 Marks

Task 3

Find recent articles in the current year searching for references on the **synthesis of carbon nanotubes**

	Google Scholar	Web of Science	Scopus
How many research papers are returned with this query?			
What is the title of most recent paper?			
Can you identify the most important paper published in this short time period?			
How many papers have been published in the last 4 Weeks			

10 Marks

Task 4

Choose one of our own members of staff from the Department of Chemistry at the University of Oxford, find all the articles written by that author and sort by times cited, we will just use Web of Science and Scopus for this search.

Who did you choose?

	Web of Science	Scopus
How many research papers this author has published?		
How many citations does their most cited research paper have?		
When was the most recent citation to the author work? For this you will need to generate a Citation Report and view citing articles.		
h-index score Web of Science or H-graph for Scopus		

Both the databases have tool called Create Citation Report (Overview) – this is shown after you searched the name of an author in the Web of Science. The Citation Report provides aggregate citation statistics and graphs for a set of search results. These statistics include: the total number of times all items have been cited; the average number of times an item has been cited; the number of times an item has been cited each year; the average number of times an item has been cited in a year.

The Citation report also shows the **h-index score**. The h-index is based on a list of publications ranked in descending order by the Times Cited. The value of h is equal to the number of papers (N) in the list that have N or more citations. This metric is useful because it discounts the disproportionate weight of highly cited papers or papers that have not yet been cited. **The h-index is often used to rank the importance of a scientist** and this kind of data is used to rank departments and whole universities.

10 Marks

Task 5

Find in Web of Science articles using scanning tunnelling microscopy written by researchers at the University of Oxford since 1995.

	Web of Science
How many research papers has the University published in this area?	
What was your search query to find these results?	
What would be a sensible number of publications in this area be?	
Can you adjust the search query to find a sensible number? What is the new search term you have used and how many papers have you found?	

10 Marks

Task 6

Find in Web of Science articles on the **design of hydrogen storage devices**, analyse the results by institution name and view those by the highest rated institution (use refine results function)

	Web of Science
How many research papers have been published in this area?	
What was the number of the top institution for this search?	
What was your search query to find these results?	
What would be a sensible number of publications in this area be and why?	
Can you adjust the search query to find a sensible number? What is the new search term you have used and how many papers have you found?	
Has this changed the top institution for this search?	

10 Marks

Task 7

The chemical names are problematic to search for since there are many variations and naming conventions. Find all the records in Web of Science that relate to the chemical **Undecane carboxylic acid**

	Web of Science
Draw its chemical structure.	
What is a suitable search query?	
How many research papers are returned with this selected query?	
What areas of chemistry is this chemical concerned with?	

10 Marks

Task 8

Reaxys is a powerful chemical synthesis database, for which you can use tools to draw chemical structures and find illustrated chemical reactions on research articles and synthetic procedures. Reaxys has an embedded java chemical drawing tool. Use this to search for synthetic approaches in the reaction tab to make the compound illustrated

The screenshot shows the Reaxys website interface. At the top, there is a navigation bar with tabs for Query, Results, Synthesis Plans, History, My Alerts, My Settings, Help, Register, and Login. Below this is a sub-navigation bar with tabs for Reactions, Substances and Properties, and Text, Authors and more. The main content area features a search interface with a text input field for 'Generate structure from name'. Below this is a chemical structure of a complex molecule, with a prompt 'Double click this frame and draw reaction query'. To the right of the structure are search filters under 'Search as / by', including options for Product, Starting material, Any role, Reagent/ Catalyst, As drawn, and Substructure (on heteroatoms, on all atoms). There are also checkboxes for 'Ignore stereo', 'No isotopes', 'No charges', 'No radicals', 'No additional rings', 'Keep Fragments separate', and 'Ignore Atom Mappings'. At the bottom of the search area are buttons for 'Copy to Substances Tab' and 'Clear'. Below the search area are tabs for 'Conditions (Form-based)' and 'Conditions (Advanced)', and a section for 'Reaction Data' and 'Bibliographic Data'. At the very bottom are buttons for 'Clear Query', 'Load Query/Batch', and 'Save Query'. A footer contains contact information and copyright details: 'Contact Us | Support | About Reaxys | Terms and Conditions | Privacy Policy. Copyright © 2009 Elsevier Properties SA. All rights reserved. Reaxys® is owned and protected by Elsevier Properties SA and used under license.'

How many different synthetic approaches have been published?

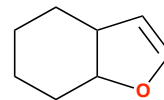
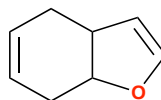
Draw the asymmetric catalysts involved.

Use the synthesize function and find out the lowest number of steps you can produce this end product from commercially available starting materials. What is the number of steps and what are starting materials?

10 Marks

Task 9

The chemical search tool allows you to investigate if a drawn compound is a substructure (included in) of other compounds.



Use the Substance and Properties tool and search for the compounds as drawn.

How many papers have been published about these compounds?

How is each compound synthesized?

Perform a substructure search. How many substructures are in the database for each of these compounds?

Select the no additional rings button. Now how many substructures are in the database for each of these compounds?

10 Marks

Task 10

The Reaxys database holds physical data and spectroscopic data on compounds. This is of course really useful when you synthesise compounds in the laboratory and you can compare your data with known information about the compound.

Ambruticin – is an important antifungal agent.

Draw the structure of Ambruticin

How many total synthesis approaches have been published?

When was the most recent total synthesis published and how many steps did it take.

Use the synthesize function and find out the lowest number of steps you can produce this end product from commercially available starting materials. What is the number of steps and what are these reactants ?

What physical data and NMR data are available about ambruticin?

10 Marks

All tasks done
