Literature search

Sebastian Jentschke

UNIVERSITY OF BERGEN

Welcome to a lecture on literature search. Often you begin a project such as your master's thesis with getting an overview regarding the exisiting literature in the field. You then either summarize what you found for your introduction. A substantial proportion of students writes their theses purely theoretical (i.e., not empirically / without collecting data) as literature reviews or metaanalyses. Especially for the latter students, it is important that the your approach to the literature search is well documented and reproducible.

In this lecture, I will try to provide some hints how to do this and what possible pitfalls might be.

Agenda

- introduction: search engines vs. databases
- search terms: selection, combination (boolean), operators
- systematic review and meta-analysis: aims and procedure
- use and practical hints: Google Scholar, Oria, Web of Science, PubMed
- reference management: Zotero, Mendeley, EndNote

What we will cover in todays lecture is first an introduction what «machinery» you can use for your literature search. Those fall along a continuum with two «extremes»: search engines (e.g., Google Scholar, Oria) vs. databases (e.g., PsycINFO, PubMed).

The next part deals with how to search: I will try to give some advice on the choice of search terms and how these can be combined using booleans (AND, OR, NOT) or other possible operators (wildcards, etc.).

I will then present some basics on systematic reviews and meta-analyses. Even though it is typically beyond what you aim for when preparing an introduction, the methodology for conducting them provides useful general principles for you. If your thesis is a literature review or a meta-analysis, you should follow this methodology quite strictly.

Agenda

- introduction: search engines vs. databases
- search terms: selection, combination (boolean), operators
- systematic review and meta-analysis: aims and procedure
- use and practical hints: Google Scholar, Oria, Web of Science, PubMed
- reference management: Zotero, Mendeley, EndNote

I will then provide some practical hints for for large search engines / databases we might use for our search: Google Scholar, Oria, Web of Science and PubMed. One area of focus will be how to export data from them so that you can use them further.

Finally, I will say a few words on different software packages for reference management. You possibly will already have recognized that I have quite strong opinions about open-source software and that it should be used whenever possible. Therefore, my recommendation is Zotero and I will say a little bit why and how it compares with other packages such as Mendeley (also free of charge) and EndNote (you might be taught about earlier in your B.Sc., the online version is free-of-charge). The point with that part is less to inflict my opinion on you but to enable you to make a better informed decision when deciding what to use for your own work.
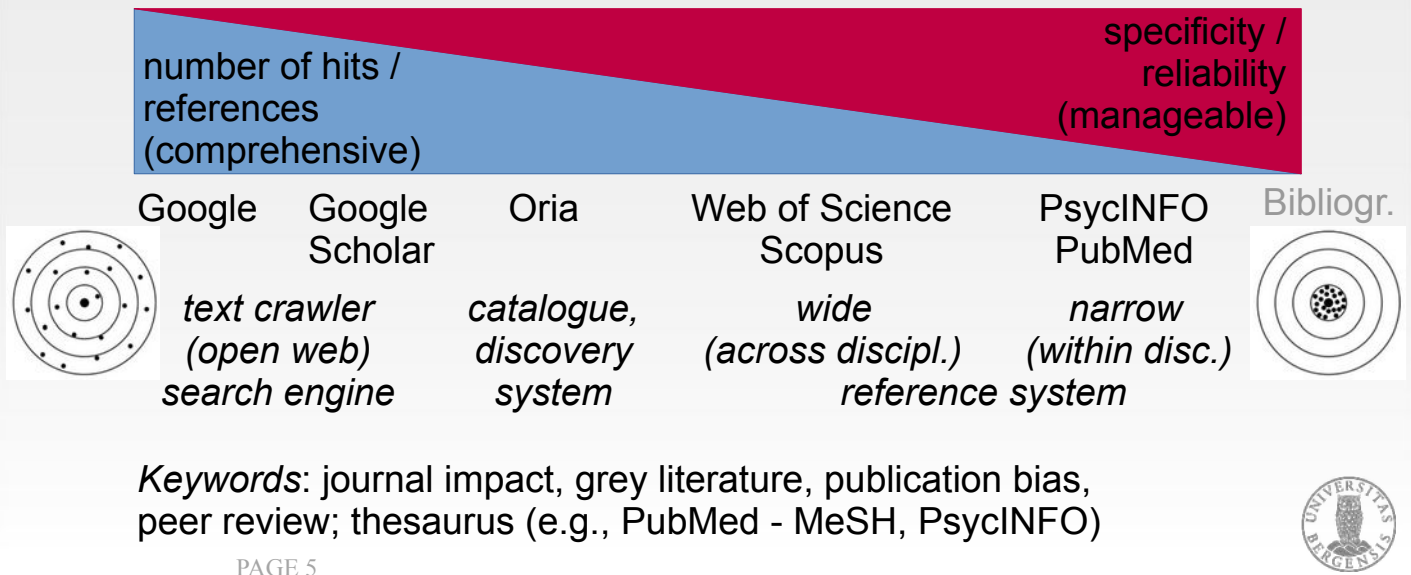
Search engines vs. databases

This part gives an introduction which search engines or databases you could use for your search. It also introduces how they differ and what their relative advantages and disadvantages are. The advantages are relative since it is very much depending on the state of the field you are interested in or the question you would like to answer which search engine or database is most appropriate (primarily how many publications exist in that field or on that question).

Search engines vs. databases

number of hits / references (comprehensive)				specificity / reliability (manageable)
Google	Google Scholar	Oria	Web of Science Scopus	PsycINFO PubMed
	text crawler (open web) search engine	catalogue, discovery system	wide (across discipl.) reference system	narrow (within disc.)

Bibliogr.

Keywords: journal impact, grey literature, publication bias, peer review; thesaurus (e.g., PubMed - MeSH, PsycINFO)

A general word on literature search: When working scientifically, one of the cornerstones of literature search is that the search can be **documented** and (ideally) be **replicated**. Replication has limits to it as there is new literature published every day, so as a consequence, seach results will change on a (more or less) day-to-day basis.

When doing a proper literature search, we should therefore document: (1) which search engine we used (and possibly why), (2) your search terms, (3) the date of the search, (4) how many results you obtained, and (5) how we continued to select from those results. Details on these points are given step-by-step over the course of the lecture.

Generally, when we decide what tool or engine to use for our search we can choose them along a conti-nuum with two extremes. At the one end stands comprehensiveness, on the other end specifity.

Search engines vs. databases

number of hits / references (comprehensive)				specificity / reliability (manageable)
Google	Google Scholar	Oria	Web of Science Scopus	PsycINFO PubMed	Bibliogr.
text crawler (open web) search engine		*catalogue, discovery system*	*wide (across discipl.) reference system*	*narrow (within disc.)*

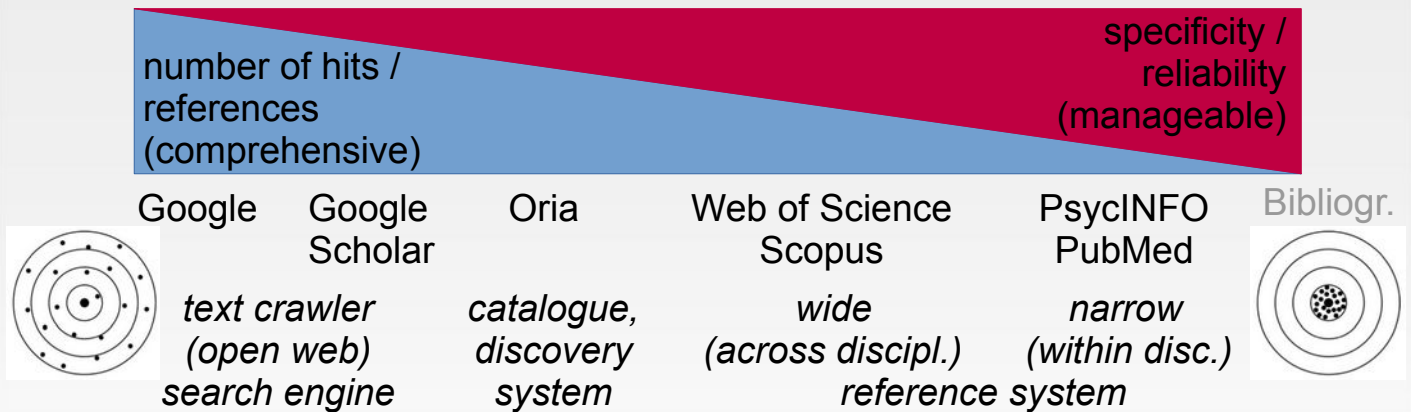*Keywords*: journal impact, grey literature, publication bias, peer review; thesaurus (e.g., PubMed - MeSH, PsycINFO)

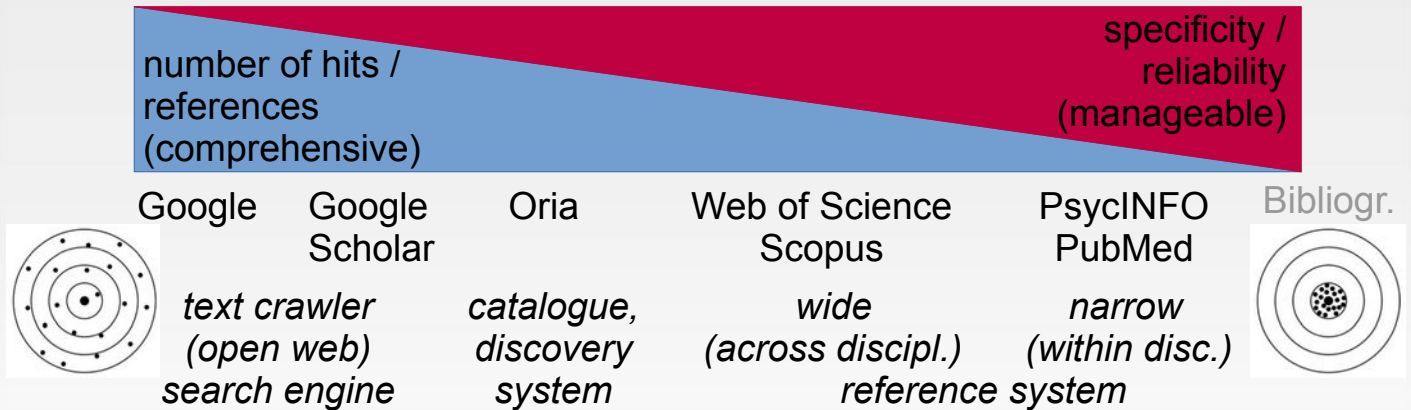Google (proper) might ensure that you don't miss anything, but at the same time would provide you with such wealth of information that is between difficult and impossible to handle that amount.

Using Google for searching already makes replicability impossible as the search results are adjusted to your individual preferences and another person wouldn't get the same set of results as you got.

Another reason why Google is not appropriate for doing proper searches in science, is the selection strategy most of us would adopt. Typically, we limit our selection of articles to what we find on page 1 or 2 of the results, since that is about what we can handle. It is impossible to replicate.

To not give a wrong impression: It is not my aim to «bash» Google. It is often perfectly fine for searching in private contexts. It just is not a suitable tool for literature search in scientific contexts.

Search engines vs. databases
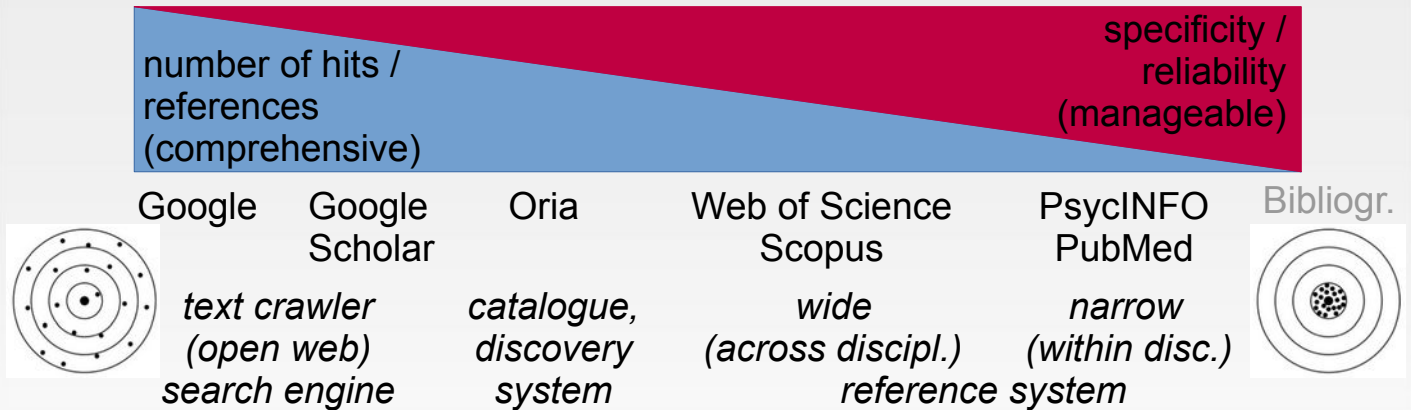
number of hits / references (comprehensive)				specificity / reliability (manageable)	
Google	Google Scholar	Oria	Web of Science Scopus	PsycINFO PubMed	Bibliogr.
	text crawler (open web) search engine	*catalogue, discovery system*	*wide (across discipl.)*	*narrow (within disc.)*	
			reference system		

Keywords: journal impact, grey literature, publication bias, peer review; thesaurus (e.g., PubMed - MeSH, PsycINFO)
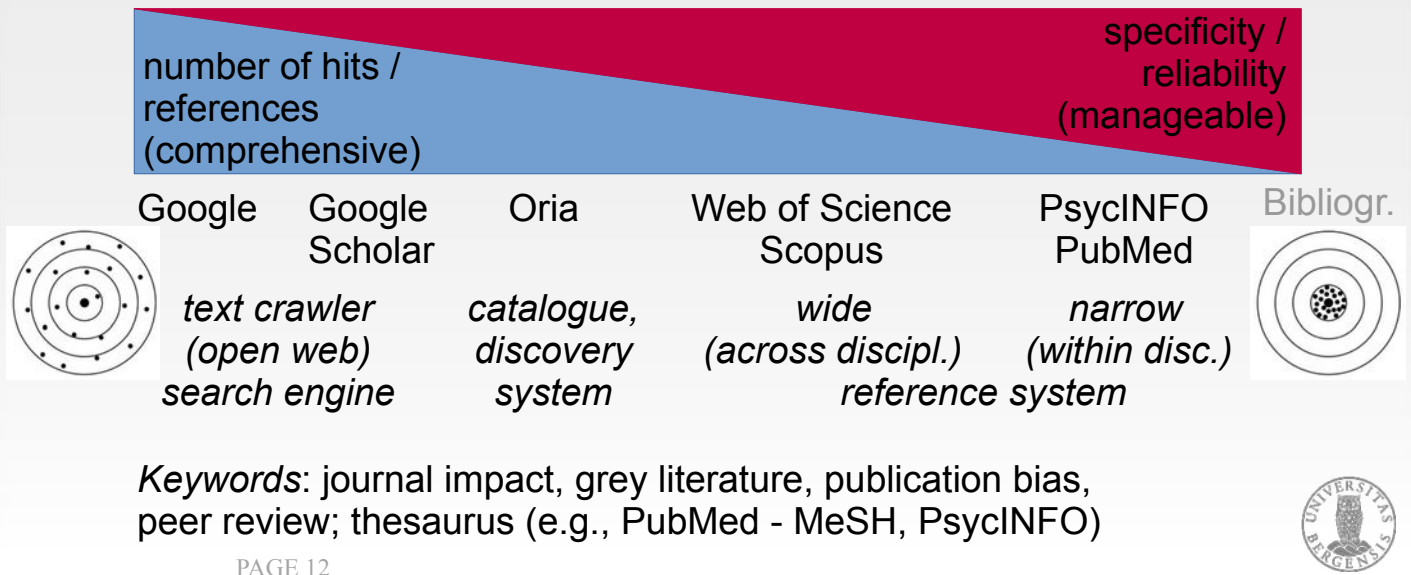
If we move on on the continuum the next stop is Google Scholar. It is already a bit «curated» or «selected» by limiting what is covered to scientific works (or what Google assumes to belong into that category).

Both Google and Google Scholar is based upon what is called text crawlers. They hoover through the web convert everything that they can get hold of into simple text which is stored in a database which again is used for the search. By learning from which articles are chosen most often when people search for certain terms, the search algorithm learns over time and puts the «most relevant» hits first.
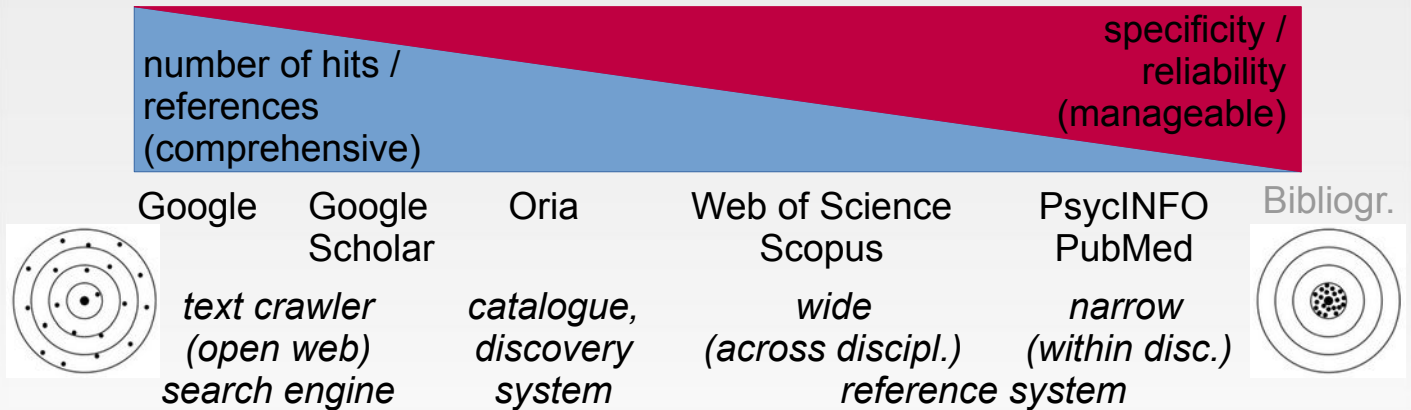
Oria is a catalogue or discovery system. It takes the same approach (text crawling: text → database → search) but is limited to text resources the Norwegian university system has access to. It is often very helpful and efficient to find a specific article, but to my taste still to comprehensive and «uncurated» for literature search. Another (supposed) disadvantage is that the search results can't be exported.

Web of Science and Scopus are large commercial databases who compile curated databases from several disciplines. Web of Science is owned by Clarivate, Scopus by Elsevier.

UiB has a contract with Clarivate, i.e., we can use Web of Science without charge while affiliated with UiB (you might need to be on campus or connected via VPN to have access).

Web of Science typically gives you a comprehensive but well-curated or characterized set of search results and lets you arrange your list of list of results (e.g., after which result has been cited most).

Search engines vs. databases

number of hits / references (comprehensive)				specificity / reliability (manageable)
Google	Google Scholar	Oria	Web of Science Scopus	PsycINFO PubMed	Bibliogr.
text crawler (open web) search engine		*catalogue, discovery system*	*wide (across discipl.)*	*narrow (within disc.)*
			reference system		

Keywords: journal impact, grey literature, publication bias, peer review; thesaurus (e.g., PubMed - MeSH, PsycINFO)

Maybe I should introduce two terms here: Curated databases and meta-data.

Curated databases are typically hosted at certain institutions, e.g., APA for PsycInfo. Typically, those institutions select a list of what they deem the most relevant journals in the field. One criterium for that can be the impact factor, i.e. how often articles from that particular journal are on average cited.

From these journals, meta-data are extracted. Meta-data accompany an article (i.e., the full-text): Title, authors, journal, abstract, etc. In curated databases typically keywords are added for each article. To make finding it easier and more consistent. Based upon these meta-data your your reference management software creates your reference list. This is why you want them to be of high quality otherwise you have to do to much manual editing. Web of Science and PubMed typically provide meta-data of high quality, Oria often less so.
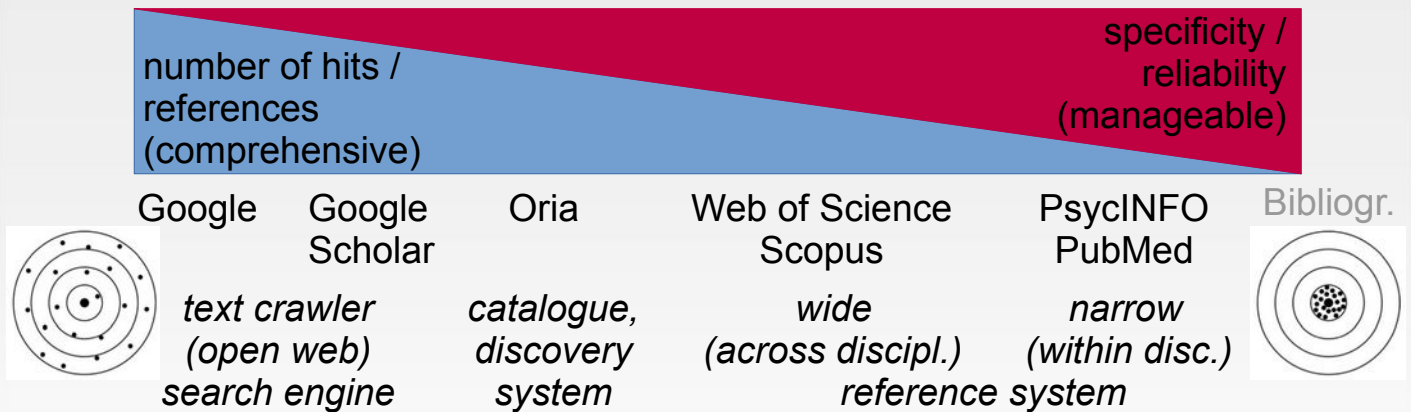
We reached the right end of the continuum and field-specific databases. Field-specific is a maybe a little imprecise, e.g., contains PubMed (which has medicine as a main focus) also covers major psychology journals. Which field-specific databas you use, is typically dependent on what your area of work is. In my case – cognitive neuroscience – PubMed is best fitting. For other areas in psychology, PsycINFO (which you can get through either Web of Science or as individual database) might be the better choice. That said, if I did a parellel search (i.e., same search terms, same date), the results were well overlapping (~ 80 – 90 %) which indicates that it doesn't matter much which one you use.

Search engines vs. databases

number of hits / references (comprehensive)				specificity / reliability (manageable)
Google	Google Scholar	Oria	Web of Science Scopus	PsycINFO PubMed	Bibliogr.
text crawler (open web) search engine		catalogue, discovery system	wide (across discipl.) reference system	narrow (within disc.)

Keywords: journal impact, grey literature, publication bias, peer review; thesaurus (e.g., PubMed - MeSH, PsycINFO)

What I also like about PubMed is that it is free-of-charge (other than, e.g., PsycInfo that you have to pay for). It also has a well-maintained system of subject terms / keywords (MeSH – medical subject headings) in combination with a thesaurus which finds synonyms (other fitting keywords). I will later demonstrate how these keywords are used when showing a search in PubMed. The original search term is quite extended by using them which ensures that you don't miss references that match the concepts of your search term.

Finally, in my opinion PubMed has the meta-data of the highest quality, i.e. when you import references from PubMed into your reference management software, you typically need the least amount of checking and correcting them.

Regardless of my preferences (PubMed), I would still urge you to try out all options for search engines and choose what is most appropriate for you.

Search engines vs. databases
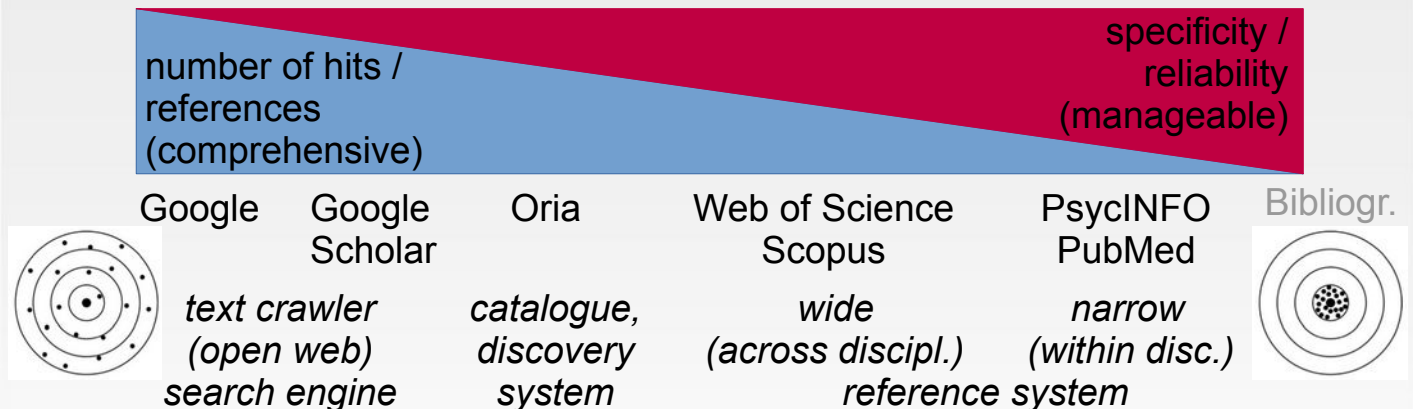
number of hits / references (comprehensive)				specificity / reliability (manageable)
Google	Google Scholar	Oria	Web of Science Scopus	PsycINFO PubMed
	text crawler (open web) search engine	catalogue, discovery system	wide (across discipl.) reference system	narrow (within disc.)

Bibliogr.

Keywords: journal impact, grey literature, publication bias, peer review; thesaurus (e.g., PubMed - MeSH, PsycINFO)

Continuing that thought with some practical advice: What search engine or database you need is not only subject to preferences but typically also very much dependent on your field or topic of work.

If you are working on a topic that is rather new or where many studies exist (yet), you rather want a search engine from the left part of the continuum (i.e., you would like to get as many publications as you could get). Google Scholar also includes grey literature (unpublished manuscripts, reports, etc.). If you want all existing literature and the field doesn't have many publications you possibly want to include all of them.

A problem with grey literature is that it often hasn't been through a peer-review process which (should) provide some quality control. So, you have a stronger responsibility to judge whether if it worth including, given it's quality. That can be difficult.

Search engines vs. databases

| number of hits / references (comprehensive) | | | | | specificity / reliability (manageable) |

Google	Google Scholar	Oria	Web of Science Scopus	PsycINFO PubMed	Bibliogr.
text crawler (open web) search engine		*catalogue, discovery system*	*wide (across discipl.) reference system*	*narrow (within disc.)*	

Keywords: journal impact, grey literature, publication bias, peer review; thesaurus (e.g., PubMed - MeSH, PsycINFO)

For the databases at the right end of the continuum, curation and selection are advantages: These databases would typically select only journals with a high impact factor. A high impact means that articles from that journal are cited often. This provides another layer of quality control beyond the peer-review process: If a substantial amount of researchers finds it worth to cite publications from a particular journal, it indicates that those publications are of good quality (indicating a sensible approach which articles are selected for publication and good quality control from the journal). It doesn't mean that you wouldn't see studies of bad quality in those journals but generally, the quality is quite high.

So, if there are few studies in the field: rather use search tools from the left end. If there are many published studies, rather choose your tools from the right end.

Search terms

After deciding which search tool or engine might be most appropriate, we have to choose HOW to search, i.e., which keywords we have to use when searching.

Selecting search terms – PICO

- **P** (problem / patient / population): patient, [demographic] group, diagnosis
- **I** (intervention): research method, intervention, exposition, application
- **C** (comparison): evt. other intervention / exposition
- **O** (outcome): class / category of results / outcome
- *Example: elderly (P) – "technical aids" OR robots (I) – opinions, attitudes (O)*

The so-called PICO-system provides a good guideline how to choose them.

One would typically start with P, a certain problem problem, patient group or population (or strictly-speaking a sub-population or demographic group). An example from my own work would be «elderly» or «adolescents», «SLI» (Specific Language Impairment) or «hypoxia» (lack of ogygene supply).

We continue with some form of intervention (I). This can also be a certain research method or a certain application. Using my example, I would for example use «technical aids» or robots. The reason why «technical aids» is in quotation marks is that aids might be mistaken for the immune disease (especially when searching in a medical database like PubMed).

Selecting search terms – PICO

- **P** (problem / patient / population): patient, [demographic] group, diagnosis
- **I** (intervention): research method, intervention, exposition, application
- **C** (comparison): evt. other intervention / exposition
- **O** (outcome): class / category of results / outcome
- *Example: elderly (P) – "technical aids" OR robots (I) – opinions, attitudes (O)*

Often, but not in my example, we have a third term (C) to which we compare the second part (I). For example, as psychologists we might be interested in comparing two interventions / therapies (e.g., CBT vs. pharmacological interventions).

Finally, we might be interested in a certain outcome, even tough it can be go much beyond and include e.g., behaviour or attitudes. In my example, I am looking for opinions or attitudes towards technical aids or robots.

Generally, also given that PICO has a quite strong clinical focus, it might not always be perfectly appropriate to include all terms or you might be required to bend the meaning of the different categories (P, I, C, O) a bit (as I did in my example).

PICO is rather thought as a framework that helps to generate and structure such search terms.

Extending search terms – Thesaurus

reference book in which **words with similar meanings** are grouped together (www.collinsdictionary.com)

- https://www.collinsdictionary.com/dictionary/english-thesaurus
- https://www.macmillanthesaurus.com
- https://www.merriam-webster.com/thesaurus

Synonyms (concept 1)

Synonyms (concept 2)

(X **OR** X **OR** X) **AND** (Y **OR** Y **OR** Y **OR** Y)
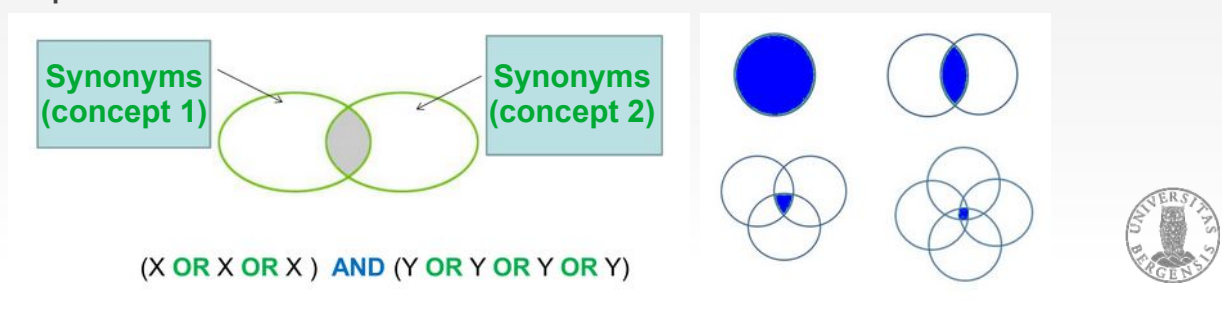
Sometimes, if it is not done automatically as in, e.g., PubMed, you have to extend your search terms with synonyms.

In the case of English the slide lists web pages where you could search for synonyms. Typically, you use English terms for your search.

However, when searching in Oria, you might be in need for sources for Norwegian synonyms such as https://synonymer.no/ or https://ordbok.uib.no/.

Typically, it is wise to OR-combine these synonyms. More on Booleans (like OR) the next slide.
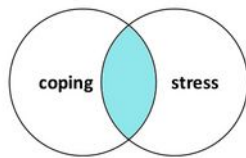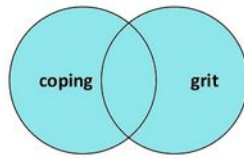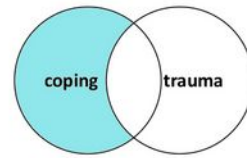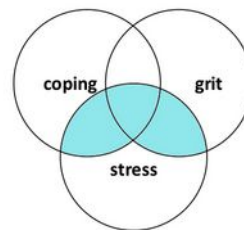
We have three so-called Boolean operators: AND which requires that both search terms have to appear, e.g., in title or abstract. If only one is contained it is not included in the results.

OR includes the search terms in the results if either the first or the second appears, e.g., in title or abstract. That means that your list of results will be more comprehensive if you used OR as compared to if used AND.

NOT means that the first but not the second term has to be contained. Sometimes you have to write AND NOT instead of NOT.

These operators can be combined. If you use parentheses, the operation in parentheses is carried out first before continuing with the other parts. In the left example (coping AND stress) are handled first (i.e., both coping and stress have to be contained in, e.g., in title or abstract, and afterwards those occurences that contain trauma are removed.

Search terms operators

- **wildcards:**
  truncation (child* → children, child, childhood, childish, etc.; *oxide → peroxide, sulfoxide, etc.)
  one char. (colo?r → different spellings [US / UK])
- **phrase searches:** "AN EXPRESSION"
- **proximity operators:** NEAR/5 (WoS), N5 / W8 (near / within, followed by; SocIndex), ADJ4 (Ovid)

Finally, we have so-called wildcards. * truncates the word and searches for everything containing the stem child (child, children, etc.). * can be also be used at the begin of the word, e.g., to substitute for different prefixes (per-, sulf- + oxide).

The search enclosed with quotation marks includes only the terms or phrases enclosed. You saw it earlier with «technical aids» which I wanted to distinguish from aids (as disease).

Some search engines, including Web of Science permit proximity operators. They allow to define that search terms have to appear within a certain distance of each other. This is helpful when combining relative generic words, e.g., cognition or cognitive and memory. «cognition N5 memory» would for example find «spatial memory and cognition» but not if those terms were more than five words apart.

Limiting your search results

- ***publication year***: last 10 years (or 5 years in quickly changing fields, e.g., genetics, neuroscience)

- ***geographical / cultural*** *– if culture / society supposedly has an impact*: Norway → Scandinavia → Europa → industrialized countries (OECD)? → world / no limits

- ***article type*** *– only offered by some databases (e.g., PubMed)*: *validation* studies, reviews / meta-analyses (if there are to many hits)

- ***species or age range*** *– only offered by some databases (e.g., PubMed)*: humans, children vs. adults

Often, you end up with far too many search results. For example results the search terms «elderly AND ("technical aids" OR robots) AND (opinions OR attitudes)» in 510 results in PubMed, 72 in Web of Science and 33200 in Google Scholar.

Typically, to obtain between 50 and 200 search results is a good target. It is large enough to be a solid basis for being comprehensive whereas it is not too overwhelming such that the results can't be handled anymore.

Often, the number of results is larger. To get it down to a amount that can be handled, there are different strategies for limiting your search terms.

Limiting your search results

- **publication year**: last 10 years (or 5 years in quickly changing fields, e.g., genetics, neuroscience)

- **geographical / cultural** – *if culture / society supposedly has an impact*: Norway → Scandinavia → Europa → industrialized countries (OECD)? → world / no limits

- **article type –** *only offered by some databases (e.g., PubMed)*: *validation* studies, reviews / meta-analyses (if there are to many hits)

- **species or age range** – *only offered by some databases (e.g., PubMed)*: humans, children vs. adults

The easiest and most obvious is to limit based upon the publication date. Typically, one wishes to describe the current state of the field, so covering the last 10 years or even 5 years in fields that change or develop quickly (neuroscience, genetics, health technology) is sufficient.

If you suppose that culture or society has an impact, another option could be chosen: It involves limiting the search to certain countries. For example is it reasonable to assume that the certain characteristics are typical for health care systems in Scandinavian countries (such as having care for all people and most / all diseases included in the health insurance and having comprehensive health registers). What is covered by health insurance might still be similar in most European countries (whereas those typically don't have as many registers). The U.S. health system, in contrast, doesn't share either of those characteristics.

Limiting your search results

- ***publication year***: last 10 years (or 5 years in quickly changing fields, e.g., genetics, neuroscience)

- ***geographical / cultural** – if culture / society supposedly has an impact*: Norway → Scandinavia → Europa → industrialized countries (OECD)? → world / no limits

- ***article type** – only offered by some databases (e.g., PubMed)*: *validation* studies, reviews / meta-analyses (if there are to many hits)

- ***species or age range** – only offered by some databases (e.g., PubMed)*: humans, children vs. adults

PAGE 23

The third option – article type – very much depends on how far developed a field is. If it is a field where very much research was conducted on a certain topic it is recommended to limit the search to meta-analyses and reviews. If the field is progressing rather fast, it is better to rely on recent studies. Another distinction (offered, e.g., by PubMed) is to limit the results to experimental studies (leaving out, e.g., qualitative research) or randomized clini-cal trials.

Finally (and this is again an option only PubMed offers) it is possible to limit the results to certain age groups or species. One might be only interes-ted in human adolescents or in rodents (e.g., if one is interested in the research by Edvard and May-Britt Moser on grid cells in the hippocampus).

Literature search

- saving / exporting to text files
- saving searches (e.g., Oria, PubMed) or receiving alerts
- expanding and limiting your search: WoS or PubMed: #1 AND *keyword;* Oria (material type): books, online access
- Oria → Databaser

PAGE 24

| Nytt søk | Databaser | Finner ikke? | Hjelp | E-tidsskrifter | Bla i lister | ... |

Universitetsbiblioteket i Bergen ▼   elderly AND ("technical aids" OR robots) AND (opinions OR at ✕   🔍   AVANSERT SØK
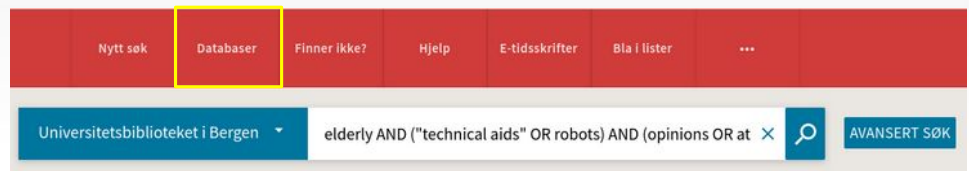
Your choice of search engine might be according to how easy it is to save / export your searches in a format the allows to work with them in other pro-grams (e.g., your reference manager or Excel).

In some search engines it is possible to create yourself an account (or in the case of Oria, to log in with your UiB-account) where previous searches are kept. You may also sign up for alerts (when a new publication appears fitting your search terms).

Some engines allow to expand or limit your search, e.g., with combining searches (#1 AND keyword would AND-combine your first search with one or more other keywords). Alternatively, one could limit your results to certain material types (e.g., in Oria: books, online material, publications in Norwegian).

Finally, if you choose «Databaser» from the menu in Oria, you have direct access to databases like PsychInfo or ERIC (education-related material). Currently, there are 30 databases for psychology.

Aims and procedures when conducting systematic reviews and meta-analyses

The key for conducting systematic reviews and meta-analyses is to make informed decisions (e.g., choosing appropriate search terms) and to document every step of the process well to enable readers of your thesis or manuscript to understand what you did and (ideally) to replicate the search (if needed).

Systematic review – meta-analysis

Aims:

- gain a general idea of your field of research
- place your own work, justify its relevance
- identify main developments and gaps in research
  - → redefine scope / research question
  - → explain your focus and justify delimitations
- discuss different methods and theoretical approaches used regarding your topic

Let's begin with the aims of systematic reviews and meta-analyses. Meta-analyses are likely beyond what you need for a M.Sc. thesis (they more typically represent a first step / the first publication in a Ph.D.). The basic procedure for systematic reviews and meta-analyses is quite similar, execpt that you different means for summarizing your search results: statistical means (typically effect sizes) in a meta-analysis vs. a verbal summary in systematic review (typically the verbal summaries would contain aim / hypotheses and main results).

The first aim with those methods is to give you an idea about the current status of your field of work. This helps you to place your own question within that context and identify main developments and also possible gaps. This is how you typically structure your introduction: what are existing studies on the topic, which aspects did they answer and where are gaps that your study can fill in).

Systematic review – meta-analysis

Aims:

- gain a general idea of your field of research
- place your own work, justify its relevance
- identify main developments and gaps in research
  → redefine scope / research question
  → explain your focus and justify delimitations
- discuss different methods and theoretical approaches used regarding your topic

Knowing about the main developments and the gaps also helps you define the scope of your own work, e.g., what aspect you would like to focus on and possibly why you chose to limit the scope (e.g., it might be far too comprehensive to assess memory in a certain age group, you might decide to just look at a sub-type / -category like implicit memory).

Finally, the overview also gives you some ideas about methods you could use to explore your research questions. This should help avoid that you re-invent the wheel when choosing your study paradigm or that you answer research questions that were already covered by numerous studies. Nowadays (i.e., with the replication crisis) it became more acceptable to do replication studies; earlier, it would have been difficult to get results published if they weren't «new» enough.

Systematic review – meta-analysis

[1] Literature search:

- select appropriate keywords (+ find synomyns?)
- select appropriate databases (+ tracing references and find grey literature?)
- define inclusion and exclusion criteria (e.g., reviews, English, published since..., etc.)
- **→ make these decisions explicit; keep records (register terms and date; export bibliography)**

Most of the practical aspects of what is listed on the slide were covered in the section on literature search. The main focus here is to list up which questions you should document in your methods section. In any case, you should keep records of what you did.

The first aspect is the choice of your search term and how you decided which keywords to use in this search term.

The second aspect is to document decision about search engines / data bases you used. In some cases, if there is not enough literature available via those databases, you would have to go beyond by searching for grey literature or tracing references (i.e., looking up references from the reference list in articles you read).

Finally, you should carefully document which inclusion and exclusion criteria you chose (e.g. only literature in English from the last 5 years).

Systematic review – meta-analysis

[2] Evaluate, organize and analyze results:

- title and abstract: further refine inclusion and exclusion criteria (e.g., only experimental studies)
- full-text: read in depth; register key findings or extract statistics *(N, mean, std.-dev.)*
- **→ make these decisions explicit; keep records (register terms and date; export bibliography)**

Once you conducted and documented your search, it is time to evaluate, organize and analyze your results.

Typically, a good strategy is to import your search results into a speadsheet where you add columns for each inclusion / exclusion criterium you choose (use 1 for studies you would keep given a certain criterium and 0 for studies you would exclude).

If you have «hard» criterias, e.g., the publication language being English, it is wisest to start with them. Afterwards, you should have two rounds of review, one screening the titles and de-select those not fitting your research question, and then a second round reading the abstract (you save time by not having to look at the abstract if the title already indicates that the article is not suitable).

If you do a systematic review, it is recommended to note down some key words or a brief summary for each article to help you to summarize it later.

Systematic review – meta-analysis

[2] Evaluate, organize and analyze results:

- title and abstract: further refine inclusion and exclusion criteria (e.g., only experimental studies)
- full-text: read in depth; register key findings or extract statistics *(N, mean, std.-dev.)*
- **→ make these decisions explicit; keep records (register terms and date; export bibliography)**

If you do a meta-analysis you have to extract the required statistics (such as number of participants in the study, means and standard deviations for the different conditions or effect sizes where available).
Again, it is paramount that you document your decisions and keep records for every step you did. Often, a table like the spreadsheet that you used for selecting your studies can go into the supplementary materials of your submission.

Within the main document (your thesis or an article manuscript) you might provide a flow chart as the one that you can see on the left hand side. On the right hand side, you find an example how such a spreadsheet for selection could look like.

There are four columns, the first one selecting if the language wasn't English, and three other columns assessing whether the three main topics used in the search term were covered according to title and abstract.

When using such a spreadsheet, you would start with having «1» in all rows of the first column and set those articles not in English to «0». Afterwards, you choose your complete table and sort it by the selection column for language so that the «0» end up either at the beginning or the end (I personally would recommend the end, i.e., sorting in «Descen-ding» order).

Systematic review – meta-analysis
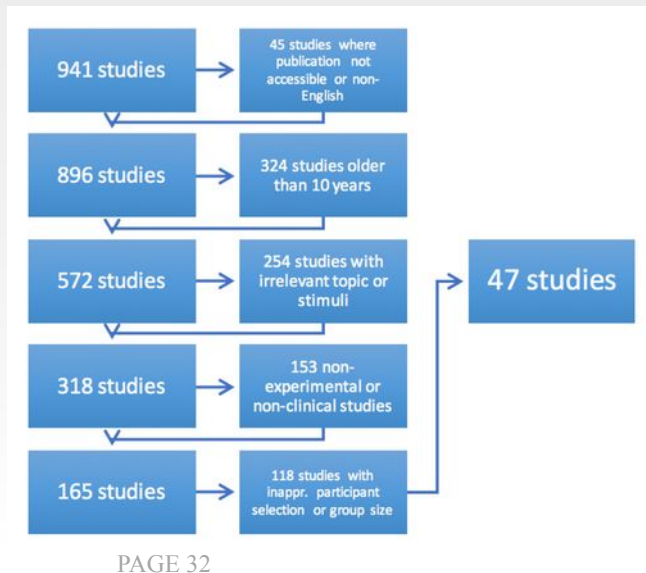
a typical way to illustrate inclusion and exclusion criteria and their effect on the number of articles

L: Language
T: Technical aid / robot
P: Population (elderly)
A: Attitudes / opinions

Then you create a new column and copy the existing column into it. For those rows marked with a one you assess the second criterium. Afterwards, you sort again and continue with this process until you checked all your criteria. I prepared a video showing the process that you can find on MittUiB.

Once you are finished, you can start creating a figure like the one on the right hand. This is most easily done with starting with your number of lines at outset are (i.e., in the original file you imported). Then, you sum up the values for each of your selection columns: Only those articles included (i.e., who got a «1») go into that sum, so that you can always find out how many are still included. By calculating the difference to the number of included references at the previous stage you can also say how many have been removed.

Systematic review – meta-analysis

[3] Summarize:

- ***narrative and systematic reviews:***
  organize key findings / concepts along broader
  categories (e.g., functional vs. social robots;
  humanoid x tasks; amount of exposition; influence
  of cognitive decline

- ***meta-analysis:*** overview, sub-groups/-categories

A side effect of that you have organized and selected your articles is that you already should have got a good idea of what the main topics are, especially after you went through the abstracts. As I said, it is wise to take notes during that process.

Given that you started with a type of structure when selecting your search terms (PICO), this could be a first step in organizing or structuring your findings. You can add topics that you recognized when reading the abstracts. Often it is wise, to let it sink for a day or two (this is at least what I try to do). During such period, and with a bit of thinking and reflection, a lot organizes itself.

When actually writing up your summary, it is detri-mental that you always begin with such a structure otherwise mess is guaranteed. And this is neither pleasant for you to work with nor for any reader.

Systematic review – meta-analysis

[3] Summarize:

- ***narrative and systematic reviews:***
  organize key findings / concepts along broader
  categories (e.g., functional vs. social robots;
  humanoid x tasks; amount of exposition; influence
  of cognitive decline

- ***meta-analysis:*** overview, sub-groups/-categories

After the structure stands, you can fill in bits from the articles under the headings you created. Don't forget to properly cite them, it is often easier to handle references already at that stage to prevent that you «forget» where they came from.

In principle, a similar type of structuring also happens when summarizing statistical measures in a meta-analysis. Here, we are maybe less thinking in terms of topics but rather in terms of sub-groups or different measures that could be summarized (e.g., results from a specific questionnaire; or coming back to our search term: attitudes of nurses or relatives who take care in comparison to the elderly actually confronted with the robot).

Practical hints for several search engines

The next part contains some practical hints and advice when dealing with the four large search engines / databases we may use most regularly: Google Scholar, Oria, Web of Science, and PubMed.
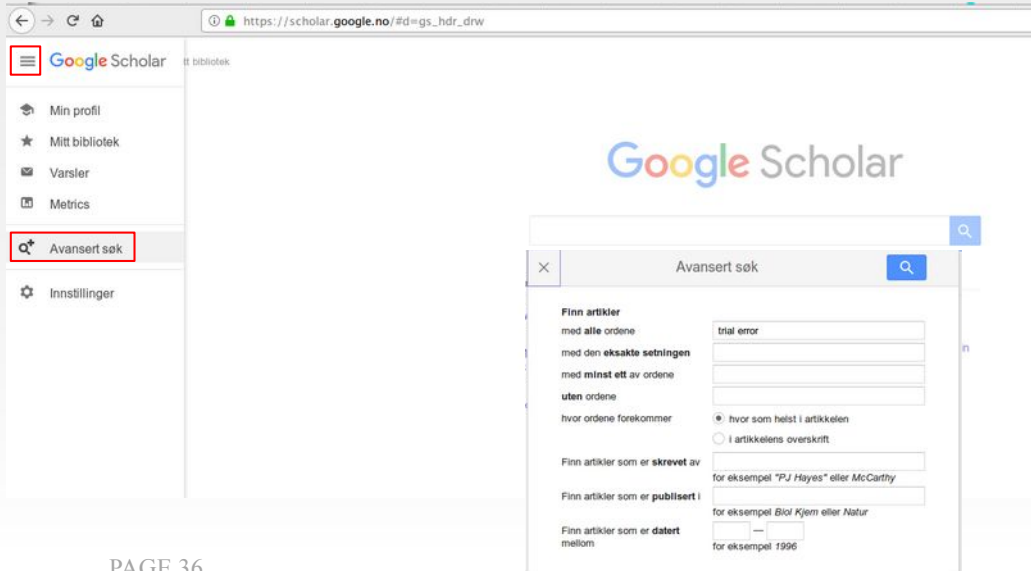
In Google Scholar, you can open the Hamburger menu (three vertical bars; top-left on the screen) and choose «Advanced Search» / «Avansert søk». You will get a new window where you have several options. Each of these input fields represents and works quite similar to one of the Boolean operators we introduced earlier. «med alle ordene» / «with all of the words» works like an AND, «med minst ett av ordene» / «with at least one of the words» like OR, and «uten ordene» / «without the words» like NOT.
In addition, you have «med den eksakte setningen» / «with the exact phrase» which works as if you enclosed the term in the input field in quotation marks. You can also determine where Google Scholar looks for those terms: in the title or in the whole article. Finally, you can look for author names, journal names and limit the publication years.

Google Scholar

outside UiB:

When you get your results in Google Scholar you have further options how to proceed.

The red boxes on the right-hand side show options how to further limit your search results (e.g., after publication year or with excluding citations or patents). In addition you can set a warning where you get an e-mail whenever a new publication relevant to that search appears.

At the bottom (the boxes marked in orange), possible alternative search terms are listed. Other users, looking for similar concepts, used those searches. You could use them to modify your own search.

Finally, in blue are marked changes between what you see depending on whether you are inside or outside the UiB network range (if you use a VPN you will be inside). Inside UiB, your results are ex-tended, e.g. with the number of citations for that article from Web of Science. Furthermore, you get a direct link to a fulltext, if available.
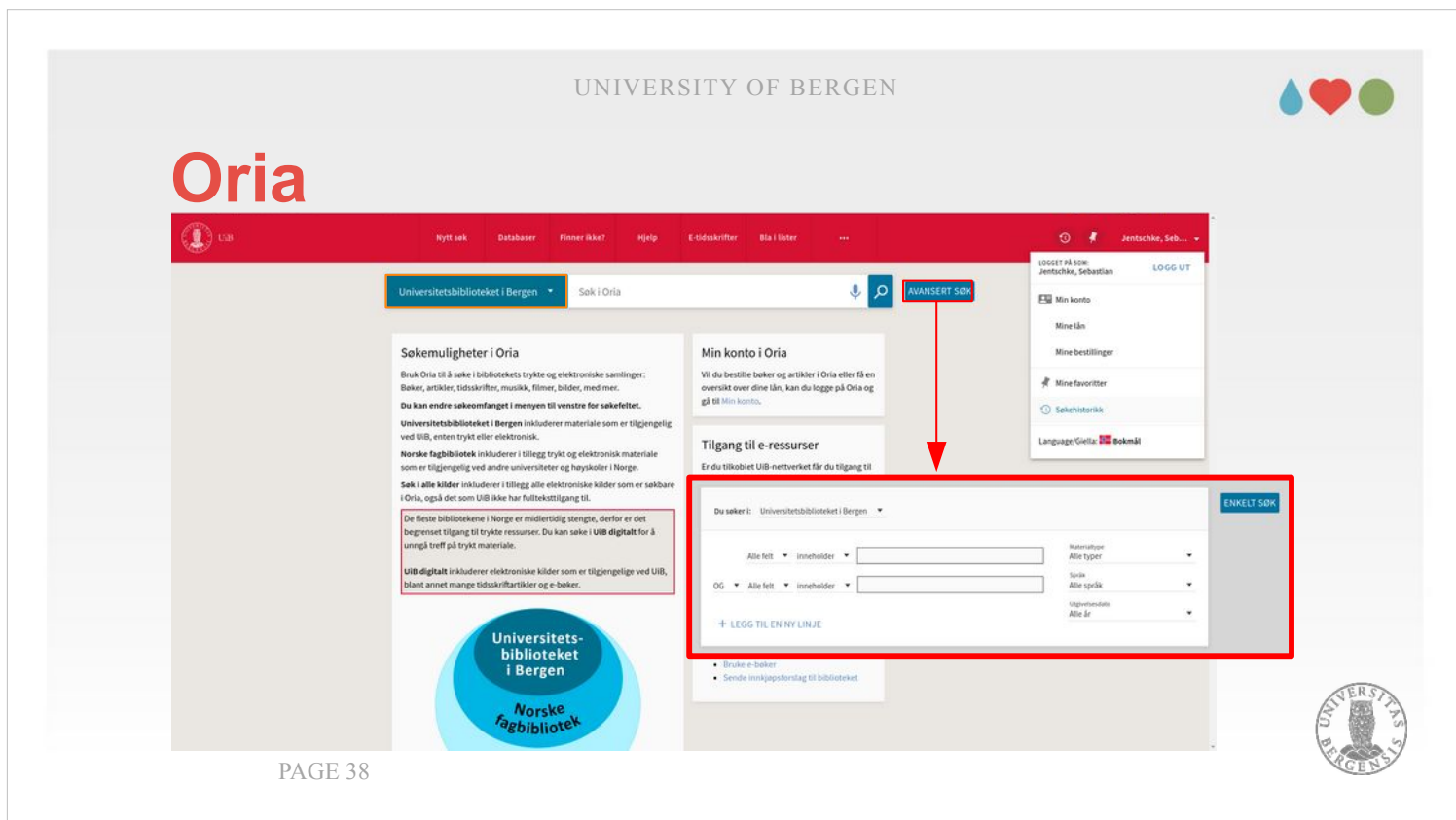
When you open the oria-main-page (oria.no), you will see a picture like the one above. If you log in (top-right corner), your search history can be stored and you can use the same search terms again later. In addition, you can choose to get personally adjusted search results (explained on the next slide).

You have the opportunity to switch between a simple search («ENKELT SØK») and an advanced search («AVANSERT SØK»). In the advanced search you can combine several search terms using Booleans (OR or AND). In addition you can select to which fields (all fields or only some like title, author, etc.) the search in that input field is applied to. On the right hand side, different material types, languages and years can be chosen. The orange encircled field next to the input for the search terms permits switching collections (UIB, Norway-wide, etc.).

UNIVERSITY OF BERGEN

Oria

Once you got your search results, you got further options to choose from.

The switch marked in green permit you to adjust the searches to your personal interests. With the entries on the right-hand side you can see what areas of interest you already set (Psykologi) and underneath you have the opportunity to modify them with «Rediger fagområder». When you click on that entry, the white window shown on the right opens.

The orange marked box lets you set whether you want your results sorted be relevance (or by date, author or title).

The red area lets you limit your search results to certain collections or material types, authors, years, etc.

If you press the «quotation marks»-icon, a new window opens with details on this reference. Here, you have several options to export the reference and get a link to access the fulltext («Tilgjengelig online»).

Unfortunately, the reference in APA-style which you can copy-and-paste is often wrong. Oria is at least aware of this error-proneness: «Husk å sjekke referansens kvalitet før du bruker den i ditt arbeid». Take this advice serious.

This nicely illustrates what difference the quality of meta-data makes and how it can differ among different search engines / databases.
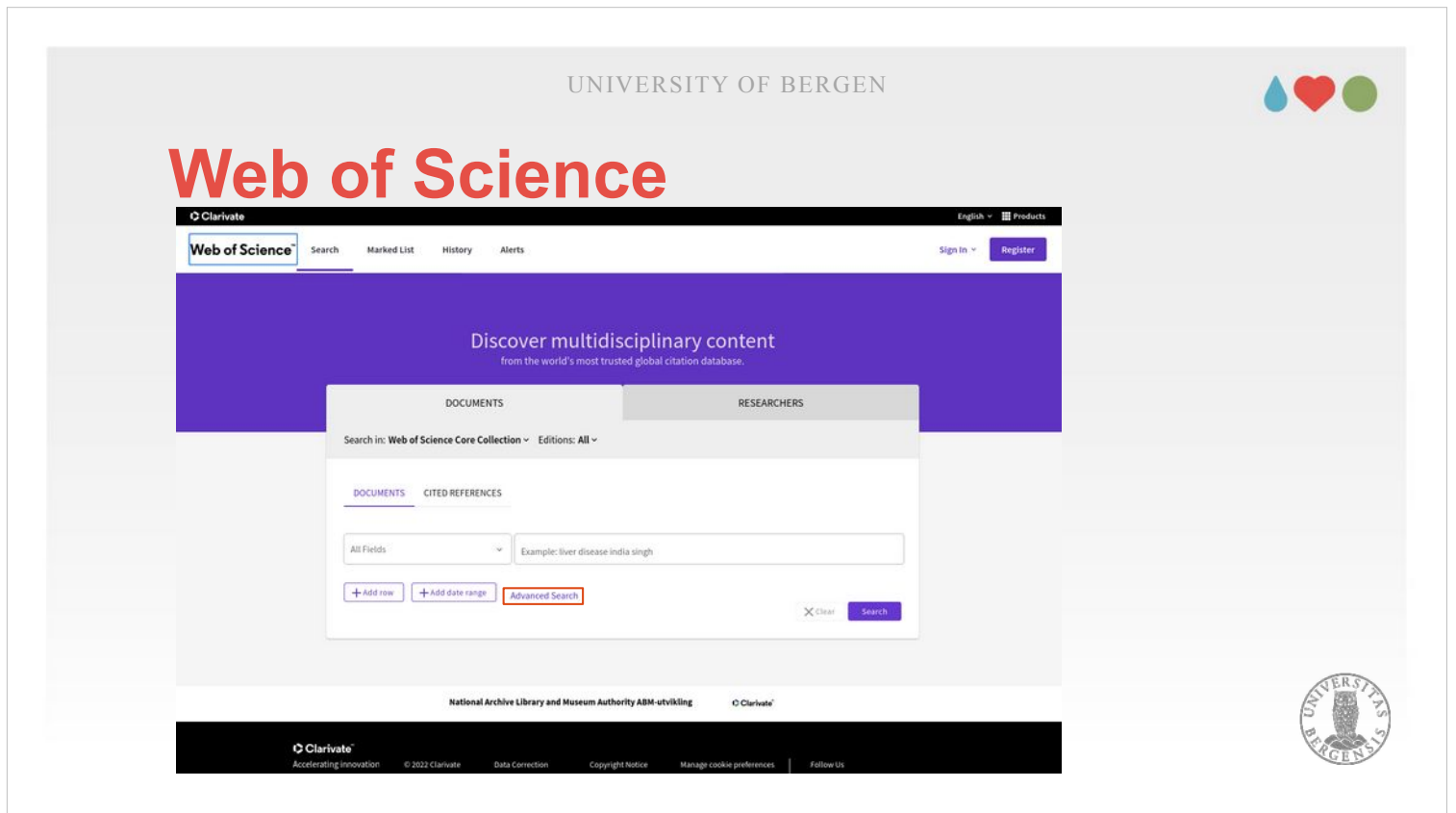
Web of Science

Web of Science also has two options for searching (strictly speaking four, of which I will explain two, since the other two are not relevant).

On this slide, the simple search is shown. Here you would just enter a search term like the one we assembled on an earlier slide (16 / 17: PICO).

The simple search has two further options: (1) Setting a data range (since 1945 or shorter periods, e.g., last 5 years). (2) Which databases to include in the search (under «Search in:»). Science Citation Index should be included for our searches. The other ones can be included in addition.

Using «Add row», you can provide several search criteria (those are combined using Boolean AND, OR and NOT).

Web of Science

In the Advanced Search you get the full monty. In the box marked with red on the right hand side, you find field names that you can use to create quite specific searches. «AU=Jentschke-S AND PY=2021» would give you my publications from last year, or «AU=Jentschke-S AND PY=2012-2022» those from the last ten years. You could, of course, also look for more interesting people, e.g, for «AU=Moser EI OR AU=Moser MB»

You can combine several of those searches with their number, e.g., #3 AND PY=2012-2022 (for getting the Moser-references from the last 10 years).

When you click on the number of results in the Search History (at the bottom of the search page, scroll down to see it), a new window opens. There, you can inspect and export your references. The field names (e.g., AU) in the green rectangle are also used to mark the entries in the files that you export.

In the window with the search results, you can inspect the individual publications or export them.

To export them, you click on the button «Export». Then you have the choice of several reference managers: For EndNote you can choose one of the list entries (desktop or online). For other reference management software (e.g., Zotero and Mendeley) RIS or BibTex are maybe most suitable.

If you want to further process the data in a spread-sheet (e.g., Excel, LibreOffice Calc or Google Sheets), you have to chose «Tab delimited file». This opens a new window, shown on the right. Here, you need to set «Full Record» under «Record Content». Alternatively, you can set «Custom Selection» and press «Edit» (next to it) to select which fields to export. If you want to export several literature searches it may be a good idea to adjust those fields (they are kept).

PubMed also provides you with a simple and an advanced search option. In the input field for simple search you just write your search term.

When choosing advanced search, you get a list with different field options, enabling you to search for author names, journal, publication year, etc.

Like in Web of Science you can generate separate searches and then combine them (#1 AND ...). The Advanced search page also contains a detailed overview of the search history, e.g., which synonyms have been used for your search terms (more on that later).

On the left hand side of the page with your search results you will find a button for «Additional filters» which brings up the window on the right where you can filter your results for (almost) any possible limitation you can imagine (species, age groups, sex, article type, language of the publication, etc.).

I said that you can open the advanced search with clicking on «Advanced» under the simple search field in the results (the oragne rectangle).

This opens the advanced search window shown on the left side. At the bottom of that page you find your search history. If you do a «quick-and-dirty» search, it is possible not required to check all the details about what PubMed did, but when you do a more systematic search you should record what PubMed made of your original search: You see from the figure on the right-hand side that a lot of synonyms was added. Typically, those synonyms work well. This time, however, there was an issue: aids (in technical aids) was resolved to «acquired immunodeficiency syndrome». This is the reason why I mentioned on slide 16 that «technical aids» has to be enclosed in quotation marks to avoid that issue (with quotation marks, we search for the whole expression «technical aids»).

PubMed

You should record the complete search term (yes, it is long), shown in the right figure on the previous page and put it in the supplementary materials.

Finally, once we got our search results, possibly applied further filters and checked whether the search term was applied correctly (or corrected it if not), we are ready to properly look at our results. «Display options» in the top-left of the window permits you to sort the results in different ways (Relevance or date).

To export the results, you press «Save» at the top of your search results. There, you choose «All results» and «CSV» as format.

Finally, you have a quotation mark-icon for each reference that allows you to export that specific reference or the copy the formatted output to the clipboard. Other than in Oria, I did not find mistakes this time. This is what I ment when mentioning the superior quality of the meta-data in PubMed.

Outside the trodden paths...

- look up references from «central» articles
- following researchers (ResearchGate, Google Scholar)
- checking articles that cite your own article (follow the development)

Finally, you could search outside the trodden paths. Whereas that sounds a little romantic, it should be regarded as last resort, not as a typical strategy.

You should use the strategies listed on this page only if there exist very few publications in the particular area that you are interested in.

A typical starting point would be to take one or several central publications from that field and check what is listed in their reference lists.

You might also take one or more of the researchers of such a publications (or a specific person where you know that he/she works in that field) and check their records on ResearchGate or Google Scholar. Authors often provide preprints or conference presentations on ResearchGate.

Finally, if you already published in that field, you could do some «reverse engineering» and check who cited your publication to stay up-to-date about the development in that research area.

Overview over reference manage- ment systems

I would like to conclude the lecture with an overview over reference management systems. Typically, once you have chosen the references that you would like to include in your literature review or meta-analysis, you have to organize them.

Of the three alternatives you likely have been introduced to EndNote in the classes that the library offers. Since more recently, the library also offers support for Zotero, though not as extensive as for EndNote (https://www.uib.no/ub/94096/zotero). Given that support you could get, it is maybe wisest to choose from those two when selecting the reference management you are going to use.

Reference management

- Zotero (open-source; zotero.org)
  Mendeley (free of charge; Elsevier; mendeley.com)
  EndNote (commercial; Clarivate; endnote.com; Mac/Win)
- all provide meta-data-extraction, browser integration and PDF annotation
- all provide Word plugins; Zotero and Mendeley also for LibreOffice; Zotero also for GoogleDocs; Mendeley also for Office 365 (done not well; reviews: 1.4 / 5; I could not install it)

Zotero is open-source and hosted by the George Mason University (Vienna, VI, U.S.). Given the affiliation, it is likely that the project doesn't run out of funding (which can be a problem with open-source-projects, leading to that they are not further developed). A second option is Mendeley (free-of-charge; Elsevier). The third option is EndNote (0 NOK for Online; 3500 NOK for the desktop version; Clarivate / Thomson Reuters).

Generally, they are pretty similar in terms of what platforms they are available on (all Mac / Win; Zotero and Mendeley for Linux in addition) and what functionality they provide: All offer meta-data extraction (you import a PDF and it looks up the reference information automatically), browser integration (you have a a plug-in that imports materials directly into your reference manager when clicking an icon) and PDF annotation (mark test passages, write comments into PDFs, etc.)

Reference management

- Zotero (open-source; zotero.org)
  Mendeley (free of charge; Elsevier; mendeley.com)
  EndNote (commercial; Clarivate; endnote.com; Mac/Win)
- all provide meta-data-extraction, browser integration and PDF annotation
- all provide Word plugins; Zotero and Mendeley also for LibreOffice; Zotero also for GoogleDocs; Mendeley also for Office 365 (done not well; reviews: 1.4 / 5; I could not install it)

They differ a bit in what text processors they support. All support local installations of Word. Zotero and Mendeley also support for LibreOffice. Zotero in addition supports GoogleDocs, Mendeley supports Office365 (currently as beta version). Here, your choice should depend on what software or web interface you typically use for writing: If you run a local installation of Office, it doesn't matter. For LibreOffice, it is Zotero or Mendeley. If you prefer web interfaces, then it is Zotero if you use Google Docs or Mendeley if you use Office 365 (and get the plugin to install...).

Reference management

Strength and weaknesses:

- Zotero: integration with Word, Writer, GoogleDocs; browser integration; LaTeX / BibTeX; apps for mobile platforms (iOS, Android in preparation)
- Mendeley: integration with Word, Writer, Office365 [?]; better lookup; LaTeX / BibTeX
- EndNote: integration with ISI Web of Knowledge; best supported by the UiB library [but: they offer Zotero too]

In addition to the platform / availability and the general functionality which is rather similar, there are certain aspects where one software package has and edge over the others. Zotero has perhaps the best browser integration. In Mendeley, it is easiest to look up meta-data from PubMed. EndNote has the best integration with Web of Science and Oria (and you might easiest find somebody to help and answer questions about EndNote in the UiB library).

In terms of sustainability, EndNote introduced an online version that you not have to pay for. For the other two packages (Zotero and Mendeley), the desktop programs are free-of-charge as well. Test-drive all of them and choose which agrees most with your workflow. Fortunately, it is rather easy to switch: All of them have options to import references from the other programs.

Further information:

Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review* (2nd ed.). Sage Publications.

https://www.phdontrack.net/search-and-review (especially «Search examples»)

https://training.cochrane.org/interactivelearning
https://training.cochrane.org/handbook/current

http://www.prisma-statement.org/

If you want to have a more thorough introduction into systematic literature reviews and meta-analyses, please have a look at the first reference.

The second reference extends the pratical hints for choosing search terms that I gave.

There is also a fantastic introduction into literature reviews / meta-analyses on cochrane.org. Cochrane is an organization that summarizes scientific evidence focussing on healthcare knowledge. Their summaries are often very much worth reading.

On the PRISMA-web page will you find a checklist, a template for a flow diagram, and their guidelines (top-left on the page are the links). These materials could also help and guide you when creating systematic reviews or a meta-analyses.

Thank you for your interest and your attention!

Thanks for reading and I am happy to answer questions in the discussion for this lecture on MittUiB.

UNIVERSITY OF BERGEN

Eco-Lighthouse