INITIATIVE FORTBILDUNG FÜR WISSENSCHAFTLICHE SPEZIALBIBLIOTHEKEN UND VERWANDTE EINRICHTUNGEN e.V.

2020 Vision: Info Pro Skills for a New Decade

Search Skills for Today's Info Pros

and

Thriving in the New Information Landscape

Presented by:
Mary Ellen Bates
Bates Information Services
BatesInfo.com

Presented for:
Initiative Fortbildung e.V.
9 and 10 May 2019

2020 VISION

DAY 1: Search Skills for Today's Info Pros

DAY 2: Thriving in the New Information Landscape

Search Skills for Today's Info Pros

Inside a Searcher's Mind: Bringing the Detective to the Search

Start every project with a clear focus. Even if you have done many similar projects, each one is different. Make sure you understand what the purpose is behind each project – what's the "why"? Find out who will be using the results and what they will be doing with the information. Think about what the answer is likely to look like – will it be a table, a chart, a collection of articles, an analysis of trends? And think about what you as a researcher need to know before you can begin your search. What background research do you need to conduct?

Stay focused during your search. As you wade through information, watch for what is essential, what is interesting but tangential, and what is an important outlier. Watch your time and be sure to allow enough time for doing something with the information *after* you have found it. I recommend spending 25% of your project time on post-processing of the results.

Watch for confirmation bias! Google is designed to looks for answers that match your question, which means that you will get different results if you search for `health effects coffee`, `is coffee harmful` or `is coffee healthy`.

Approach your search differently. Keep in mind that your clients only ask you to do what they *think you can do*. Unless they know you can provide data visualization of results or a dataset for a data mining project, they won't ask you. Use the reference interview to *expand* your clients' vision of what an information professional can do.

Look for *clues*, not *answers*. The answer isn't likely to be easily findable – if it were, our client would have found it already! – but we can always find a lead to the eventual answer. Look for context in each client project. What is driving this request? What is behind their inquiry? What else puts this in perspective? Watch for 80:20 rule, in which you will find 80% of the information in the first 20% of your research time, and it could take the other 80% of your time to find the last 20% of information. Expect to try multiple queries and multiple approaches before you find your answer.

Techniques of a Detective

Look for hidden signals! I had a project in which I had to identify the market for an office furniture "upcycling" service. The signals that might indicate a need included

Sustainability in job titles

Recent rebranding campaign

Fast growth

Change of CXO, and so on

Sometimes, you don't know where to start. For a recent project, I needed criteria for evaluation of searching online service. I Couldn't describe what I was looking for – I did't know what the deliverable was called. I just searched with a few criteria that my client mentioned ("ease of use", navigation, interface), and explored all the results I found, looking at the search suggestions that search engines provided for ideas.

Think about what the result will look like. Are you looking for industry-specific information or coverage from general news sources. For example, you will get different results when searching for `"hydraulic fracturing" wastewater` as opposed to `fracking wastewater.` Look at search terms that address process vs. outcome. For example, a search for `customer surveys` (process) will get different results than a search for `customer retention` (outcome).

Think about what you've learned when you find nothing. WHY did you get no results? Did you mistype the query? Did you use the wrong logic or syntax for this system? Are you not using the terms that most writers would use? Do you have some unexamined assumptions about what the answer will look like or where the information will come from? Are you searching an inappropriate source?

Think about what finding nothing tells you, and tell the client about this. Do the gaps correlate to an event? What related information did you find? What do you recommend next if your client wants you to pursue this further?

Consider looking for case studies if you need examples of who's doing something well. Include the phrase "case study" in your search.

When searching in a fee-based online service, use the *search filters* to survey the terrain and get a sense of where you need to dig deeper. A search filter for Publication Date shows spikes of news during a particular period of time. A filter for Subject shows alternative terms you can use. A filter for Database shows unexpected sources of coverage in industries or areas you might not have thought of.

Different Search Approaches

Today, we have more options and more choices:

Open-access content

Content from government digitization efforts

Online vendors offering text and data mining

Geolocated (mapped) content

Expand your search to new areas. Check job aggregator sites such as Indeed.de or Glassdoor.de for where and who an organization is hiring. Glassdoor also tells you about the work environment and hiring process, which can offer useful insight into an organization's culture.

Talk to people! If you can't find the answer, find someone who knows. Don't use the "Contact us" mystery box on a web site; do some research to find the individual within an organization who is likely to know the answer. Use LinkedIn, SlideShare.net (limiting by format and date) or conduct a literature search to find people writing about your topic. Finding experts takes work – search for your topic AND `filetype:pptx.`

Don't get too complex when searching on the web. Using too many terms often gets you lower results. Don't use long OR statements; instead, think of the most likely words to appear in the results. Graze, don't just dig deep; look for the next clue, not the answer.

Use `site:` to narrow to a type of resource. For example: beekeeping site:gov –site:nlm.nih.gov looks for pages from a government web site other than Pnlm.nih.gov that mention beekeeping. (Of course, this only works in where countries that use top-level domains such as .gov, .edu or .ac). You can also use site: to expand your search to dig deeper in a resource. For example: `site:rwth-aachen.de "artificial intelligence"` returned 4990 results, whereas searching within RWTH's site only retrieved 52 results.

Use `site:` to explore deeper into a site and find *subdomains* to explore within a web site. A query for `site:nasa.gov –site:www.nasa.gov` will find other areas of web site, such as sealevel.nasa.gov, www.jpl.nasa.gov, exoplanets.nasa.gov and so on.

If you want to find information from resources outside Germany, find ways to "leave" the country. If you simply go to the country-specific Google site, you will still see German-centric results. Instead, use a VPN to that country before going to that country's Google site. Change your location in your search engine's Region Settings. And with Google, you can add `?gl=`*TLD* to the results URL (e.g., ?gl=uk)

Be mindful of the different results you get from spoken as opposed to typed queries. Mobile (spoken) queries tend to have a conversational structure and use more words. A query on a laptop (typed) tends to use more structured search logic and fewer words.

If you aren't finding what you want, consider escaping search engines' filter bubble by using a private search engine. Duckduckgo.com and Startpage.com are two popular privacy-oriented search engines, but I tested 4 *other* private search engines to see how they handled research queries. All had good, relevant results, most of which did not appear in the first page of Google search results. These four include:

Qwant.com, based in France with its own search index. It filters search results for news, social, images, video, and "freshness" (most recent).

Mojeek.com, based in UK and also with its own search index. One of its advantages is that you can set your location to any country, or specify "none" to get results not filtered by region.

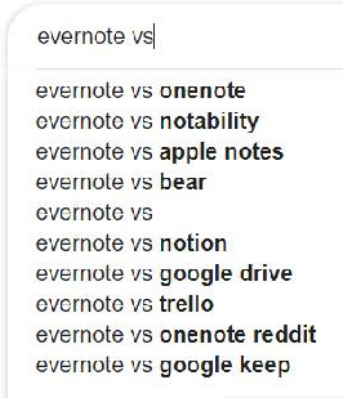
Yippy.com is the surviving entity of Clusty and Vivisimo, two long-time search resources. It supports clustering of search results; you can even "remix" the clusters if you want a different view of the results. Yippy is particularly useful for ambiguous questions where you can use help in figuring out which aspects of a topic to approach.

Gibiru.com offers "Uncensored Anonymous Search". It includes ads in search results, unlike other private search engines, although they appear not to be customized. Its *Uncensored* version surfaces the weird that Google normally suppresses; I found 0% overlap of uncensored results with Google results.

Getting Creative

Use your search tools creatively. For example, use online maps to view the street address of a questionable supplier – are they located in a shady part of town?

Search engines' auto-complete feature taps into what "most" people are thinking; if you are looking for alternatives to a product or service, try the following to see what products people compare it to.



Use image searches to find text resources! If you are researching an individual, search for images of the person and then follow those links to find unexpected connections.

Use reverse image search (searching for similar images to a particular image) to see if anyone is using your institution's photos or graphics. Likewise, if you identify a useful graphic, search for *other* copies of that image to find discussions of that topic. You can also use image search to identify market reports or get ideas on how to search.

Use Tineye.com or Google Images; ImgOps.com offers a list of reverse image search tools.

Extract text from audio. If you find a useful video interview or news program, download the audio and use a transcription service to convert the audio to text. Temi.com charges just €5/audio hour and provides a high-quality transcript.

Tap into the expertise of your colleagues! Teach your staff to reach out to each other; reduce the fear of NKIA (Not Knowing It All). Encourage brainstorming, particularly for difficult or ambiguous projects. If you need to ask for help from outside colleagues, be clear and concise, and be sure to maintain confidentiality. For example: *I am looking for stats on XYZ. I've tried these sources. What else?* Or *I'm researching topic XYZ but not finding much. Is there a better way of expressing that concept?*

What's New (or at least useful) with Google: Tips and Tools for Today's Google

Embrace Google's brain! Yes, Google monitors your searches – that's both a feature AND a bug. It lets you ex*pand* your search to find related information you wouldn't find otherwise, and lets you *fillter* search to limit parameters.

Beware of Google's "Featured Snippets". Google can't read and it gets lazy. Show this to your clients to build F.U.D. (Fear, Uncertainty & Doubt). And make sure your library info is search-engine-friendly!

Shake up your routine. You will get different results when you search while logged in, in Incognito mode, when you are using a VPN. Your results when searching on a mobile device will be very different than a search on your laptop. When searching Google on your mobile device, the search tools are hidden, you can't view more than 10 results at a time, you sometimes see fewer ads, and you get distinctly different results than a laptop search.

When you use the "Search by voice" feature (click the microphone on the Google search box) and speak your query, Google reads the first result out loud. Note that the query is treated just like a typed search, so notice how you express yourself.

Watch for Google's tools and advanced search features, which vary from page to page. These tools are usually on search results page; look for "Tools" or the ☐ icon, or pull down the Settings menu and select "Advanced search".

Note that advanced search is not available in Google News, and sort-by-relevance/date is gone. To regain those advanced search features, go to Google.de and type in your query. On the results page, click News. Now you have the tools to sort by date or relevance, and to limit to the past hour, day week, etc.

<u>Google Tricks</u>

Tricks that sometimes work include:

Repeat a search term – I get different results when I search `botanic library` and `library botanic library`.

Change the order of the search terms – I get different results when I search for `"hydraulic fracturing" wastewater` and `wastewater "hydraulic fracturing"`.

Other tips and tricks:

Use * as a place-holder -- `site:bibliothek.*.de OR site:library.*.de` finds web sites for *any* library.

SOME punctuation makes a difference -- `marketers'` is different than `marketer's`; `on-site` also retrieves `onsite` and `on site`; `café` is different than `cafe`.

Test searching with plural vs. singular nouns -- `autonomous vehicles statistics` gets different results than `autonomous vehicle statistic`.

Remember the range search (..) -- `2018..2020` retrieves pages that mention 2018, 2019 or 2020

Use `intitle:` to get focused content

For terms you want closely related, try all these:

word1 word2
word1–word2
“word1 word2”
“word1 * word2”

Watch for missing search terms – Google will often return results with some of your search terms missing. Watch for missing words and for queries exceeding 32 words.

Try a simple date search – `before:yyyy` and `after:yyyy`. This strategy gets more results than using the date tool but Google still can’t handle dates well.

Consider using Google to hack a site. If you can’t find something in a web site, consider googling it with `site:`. You may find unindexed pages or archived copies of pages.

Other (possibly useful) Google Tools

Use Google Trends search (trends.google.de) to look for hidden signals. By searching for queries that people make, you can see when people are thinking about a topic, and what words they use.

To use Google’s Reverse Image search, go to images.google.de and click the camera icon. Ignore the top of the results page and scroll down until you see the similar images.

Google Lens (app) lets you conduct searches through your camera. You can search for similar products to one you are looking at, get information on local landmarks, identify a plant or animal, read barcodes and QR codes, and link to where you can buy a book that you see.

Google’s Dataset Search (toolbox.google.com/datasetsearch) lets you search for datasets. It includes content from open-access aggregators such as dryad and figshare. Note that it only searches metadata, not the content of the dataset directly. The depth and quality of the metadata varies wildly; a minimally indexed dataset will be difficult to retrieve.

Searching the Deep Web / Grey Literature

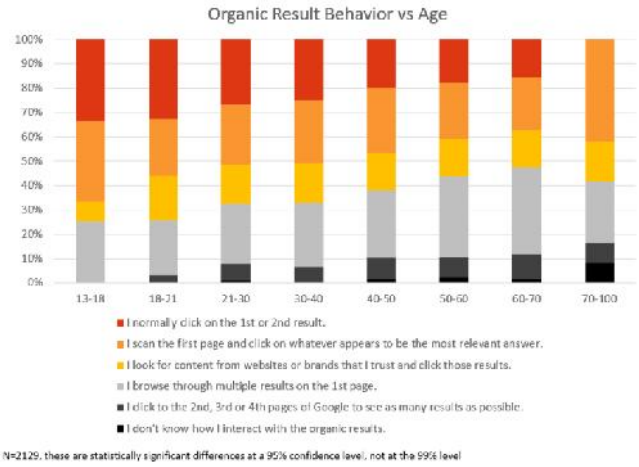
The deep web ≠ the dark web. We are looking at "deep web" content - what search engines can't access, can't "read", choose not to index or otherwise do not easily surface. Deep web content can sometimes still be found through a search engine, when searched with a different mindset.

What's in the deep web? – Databases; images, video and other multimedia; books, articles, documents, and other printed reports; and much of social media (more on that later).

What's grey literature? - Publications not produced by commercial publishers. These might include dissertations and article preprints, clinical trial results, analyses from institutes, think tank policy papers and reports, governmental and NGO working papers and studies, professional and trade association reports and newsletters, conference proceedings, and corporate white papers. Less traditional grey literature includes slidedecks, social media, patents, standards, software, survey responses, and videos.

Content from more than 10 years ago is often not retrievable in Google, although Duckduckgo and millionshort might be able to find it. If you know the URL, you can also try looking it up in archive.org.

The second page of results may also count as "grey literature"; a recent survey found that younger Internet users often just look at the snippets from a search result and do not even click through to the full record. (https://searchengineland.com/younger-users-rely-on-snippets-and-knowledge-panel-often-dont-click-survey-says-315963)



Why search dark/grey content? Often, the content is more current than published content. It offers outliers and perspectives not covered in traditional media. You can find a niche focus not well covered elsewhere. Finding grey literature helps fight the perception that "it's all on the web for free".

Search Strategies for Grey Literature

Keep thinking like a detective - look for clues, not the answer. Search for the page BEFORE the dataset that you need; look for the landing page that describes that dataset. Use your peripheral vision as you search and watch for unexpected results. Be prepared to wade through LOTS of results; grey literature can be buried in the fifth page of search results.

When Googling for grey literature, try searching both Google and scholar.google.de. Use words that are likely to be used to describe the content – `"technical report" OR "conference proceedings" OR "white paper"`, for example, or `"institutional repository" OR "open access repository"`.

Look directly at likely sources: government agencies for statistics and market research; the World Bank, UN and other NGOs; and associations in your organization's field. Find sites that monitor deep web resources, such as a infodocket.com and researchbuzz.me.

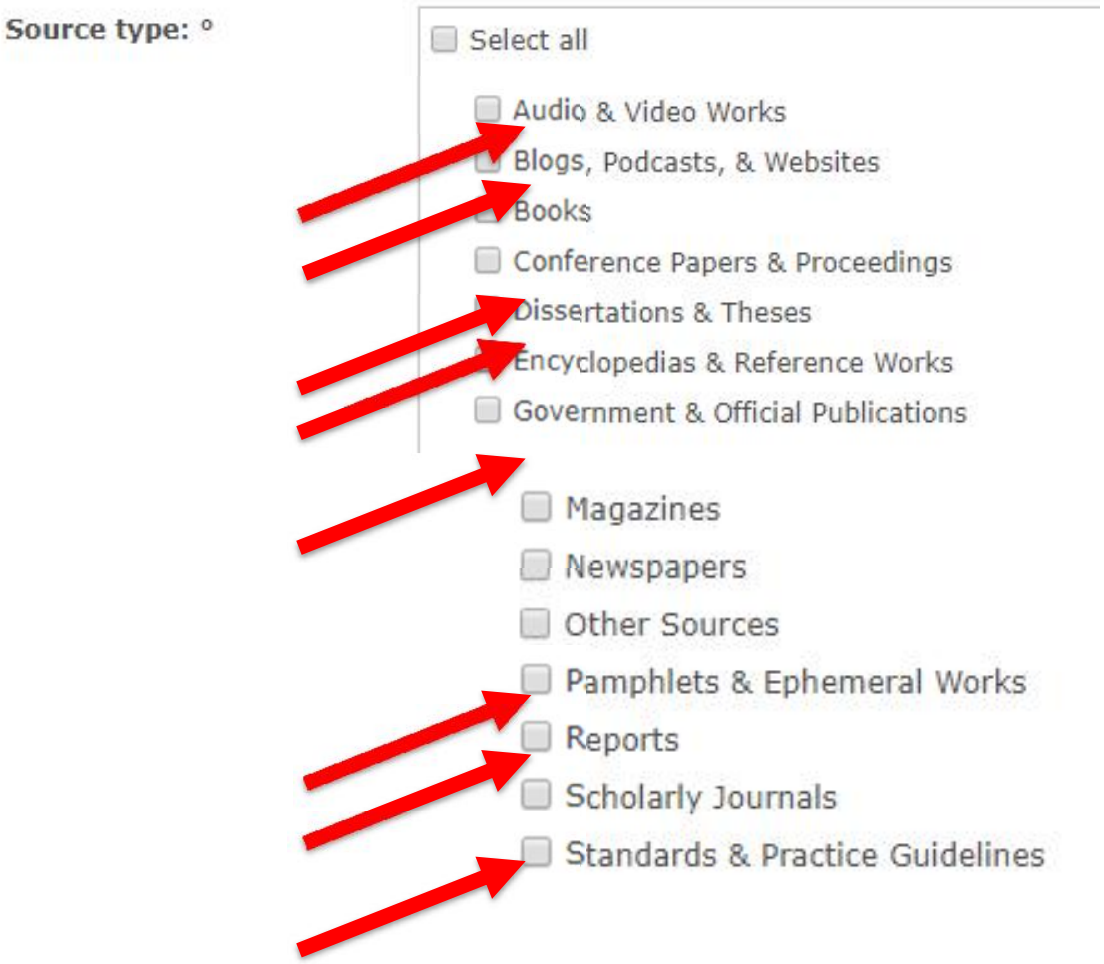
Build your own collection of grey literature – monitor ssrn.com for preprints and other early scholarly research; search arXiv.org for hard and social sciences. Consider monitoring Researchgate.com, at least to see if any of your organization's materials appear in it.

Use Millionshort.com to dig deeper. It eliminates the most popular sites from search results, enabling you to find obscure or less-commercial sites.

Know your search criteria and think about whether you need to limit your search by format, source, timeliness, reliability and breadth/depth. You may want to start your search by finding a libguide on the topic. Include `inurl:libguides` in your query to find a libguide on your topic. Once you find one good source and look for mentions of it elsewhere on the web.

Identify relevant conferences through a general search, and then search the conference page for papers, slides, etc. Many conferences make their speakers' presentations or papers available through their web site. Try searching for prior years' conference papers in archive.org .

Use fee-based services' "Source type" to identify grey literature.



Some Grey Lit/Deep Web Tools

Look for theses & dissertations in EBSCO Open Dissertations (biblioboard.com/opendissertations), Open Access Theses & Dissertations (oatd.org), and the Networked Digital Library of Theses & Dissertations (search.ndltd.org). Note the authors of useful dissertations and see where they have written and spoken at for additional information

Use Zanran.com to find statistics and factual information. This tool uses computer vision to find graphs and other statistical information in published matererial, making it a useful tool to identify sources to deep deeper into. You can mouse over each result to see the graphic being referenced. This is a particularly useful source if you don't know where to start or what resource to consult first.

Finding datasets and statistics can be tricky. Use search engines for leads – try searching:

Keywords `(database OR dataset)`

`Keywords (filetype:xls OR filetype:xlsx)`

Try searching Datacite (search.datacite.org), which specializes in open-access datasets, articles, images, etc. Unfortunately, it only searches the source's metadata, not the underlying data.

Also consider the Registry of Research Data Repositories (re3data.org), which is Datacite's directory of data collections. Again, you are only searching metadata, so search broadly to locate information on datasets and collections.

To locate tart with one known source – an association, NGO, non-profit, library or museum, for example. Try Googling *keywords* `(portal OR resources OR "online tool")`

Build your own curated collection of open access resources.

Open-access journals: doaj.org

Open-access books: doabooks.org

Open-access repositories: opendoar.org

Public Library of Science: plos.org

OpenGrey (Grey Lit in Europe) opengrey.eu

BASE – Bielefeld Academic Search Engine

base-search.net

Gleaning Insight from Social Media

Start with a clear focus – ask yourself what's the purpose of this project? Who's going to see the results? What will the answer look like? What do I need to find/do first?

DO remember what you're searching -- people's thoughts and pointers, not articles or other lengthy text.
DO plan to browse, not just search. Finding useful material on social media is a multi-step process.
DO watch for clues. What hashtags are people using when they discuss a topic? What are they saying? What words/phrases do they use that you wouldn't have thought of otherwise?
Try searching social media on Google with the site: option; it's not comprehensive but you can construct more complex searches in Google than you can in most social media platforms.
DON'T just lurk – participate! You get a much better feel for a service if you use it yourself.
DON'T look for THE answer… look for leads to the answer.

Searching LinkedIn

Use LinkedIn to find experts by searching with advanced filters. To get to the advanced search page, click the search box, then click "Search for People", then click "All Filters". (Note that, to find librarians, search for "libraries" in the industry field.)
Search LinkedIn posts for useful content; it's sort like Facebook for professionals. You can find LinkedIn posts by searching for Content (instead of People, Jobs, Companies, Groups, etc.).

Searching Twitter

What's Twitter good for? Pointers to useful resources, monitor interesting people and staying up on current trends. You can find posts that were particularly influential or popular by including in your query either min_retweets:n (the post was retweeted at least n times) or min_faves:n (the post was liked at least n times). You can also limit your search to find contemporaneous news coverage by using the syntax since:2019-04-01 until:2019-04-30
Use Twitter to monitor trade and professional conferences to identify speakers' slides and other resources and to spot new ideas and trends. Find the show's hashtag (organizers often list it on their web site) and search by the hashtag.
You can search just within the tweets of people you follow by including `filter:follows` in your query. If you build lists of experts, thought leaders and influential publications in your area, this can be a particularly useful resource. You can locate other people's public lists in Twitter. Use the syntax `site:twitter.com/*/lists` along with subject words.

You can search within your lists by the syntax list:*user-handle/list-name.* (For example, if I search list:mebs/people-i-read AI, I will find posts from people in my list called People I Read that mention AI.

Searching Facebook

Good luck! Try enclosing phrases in quotation marks, and include the city or other identifiers to focus your search. Be sure to select ALL results (rather than People, People, and so on). Just a few weeks ago, Facebook announced that it is is changing (again). Its planned platform redesign will emphasize private groups and visual stories in an effort to fight misinformation. The long-term impact of this isn't clear -- moderators will be responsible for more policing of content, and this may lead to groups becoming more insular and toxic.

Searching YouTube

What's YouTube good for? Use it to find tutorials and how-to videos. Consider browsing YouTube to get ideas for internal video tutorials describing how to search library resources. YouTube is useful for professional development – use it to find lectures, speeches or interviews on a topic you need to get up to speed on.

To find content in YouTube, start with a broad search, then use the Filters to narrow, focus and sort the results. Note YouTube terminology – a channel is all the videos uploaded by a user, and a playlist is a curated collection of videos.

Final thoughts for searching social media – be creative, manage your (and your client's) expectations about what you can find, and participate in social media yourself.

# # #

Thriving in the New Information Landscape

Reexamining Free vs Fee in an Open-Access World

As Stewart Brand noted back in 1985, information wants to be free because it has become so cheap to distribute, copy, and recombine – too cheap to meter. But it also wants to be expensive because it can be immeasurably valuable to the recipient.

Can't I just Google that?

It's impossible to get unfiltered Google results, and we need to remind our clients about this. Factors that influence the results of your search include your IP address, your search and click-through history, the speed of reviewing and clicking links on the search results page, the type of search query you use, your browser and operating system, the number of search queries you've made, your typing speed, demographics, the time of day, and so on.

Can't I just use Google Alerts?

Try a side-by-side feature comparison of Google News Alert and a fee-based service to demonstrate to your clients where the trade-offs lie. Show that a search engine or free news alert service often offers fewer search and filtering tools, fewer power search tools, and fewer format and delivery options.

In fact, dare to compare typical queries on a search engine or other free source and on a fee-based online service. Chart the results and show your clients the content and features they are missing.

<u>Choosing Free or Fee</u>

When evaluating an enterprise subscription, the free-or-fee checklist will include considerations such as:

- Price!
- Subject coverage for your clients' practice areas
- Self-service user support (to minimize library staff time spent supporting end users)
- What your clients trust and will use
- Functionality for the types of queries your users are likely to try

A checklist for an individual librarian deciding on whether to use a fee-based service or free resources will include considerations such as:

What's the budget for this?

Do I need to bill out the expense?

What will the end result look like?

How much is at stake?

What do I need in order to get smart?

Who is likely to care about this?

What kind of resources will I need?

What is "authoritative" to this client?

What do I need now in order to add value later?

Do I have expertise in this area?

How much time do I have?

How specialized is the topic?

Getting The Most From Fee-Based Sources

Make sure you're using the service's power tools. Do you remember the service's syntax and features? How do you specify adjacency? How do you limit your query by field? How do you truncate terms? What other special search tools are available, such as atleast*n*, word count or allcaps.

Use the "pearl-culturing" technique when you aren't sure how to approach a problem. In your initial search, limit your search terms to the title field. Review the subject keywords of the retrieved material, and then use those keywords for a broader, better search.

Use the search result filters for quick data visualization. A list of the most frequently mentioned executives, for example, may indicate the leading players in an field.

Getting The Most From Free Sources

The web tool I use the most today is my brain! Strategies to use before you even begin your search include:

Do your own background research first (yes, even Wikipedia. Remember that the first search you do is often for yourself, not your client.

Run a quick search in Google News just to see if there is a time period you should pay attention to or any unexpected developments.

A hammer isn't just for nails – think creatively about web finding tools!

Archive.org can be used to find information on organizations that no longer exist, to get a snapshot of how an organization was presenting itself in the past, and to support prior-art patent searches.

Use free sources to find snippets of information from expensive sources by including words that are likely to retrieve mentions of reports, such as `“market research” OR “research report” OR “market study”` or `forecast OR trends OR outlook OR statistics.`

Remember local search for in-depth coverage not found elsewhere. Searching in local newspapers and other sources often gets you coverage of small / privately-held companies, news about employee layoffs, in-depth coverage of natural disasters, and a sense of local insights.

“Bonsai searching” is what I call those narrow or niche requests that require information on a topic that is not well-covered in the literature. These requests require a different approach and a different standard for what you consider to be “good enough”. You have to look way beyond the usual sources - look for slide decks on SlideShare.net, look for experts to contact, and be prepared to just flail around on Google for a while before you find anything. For these types of research, expect to spend more time on your results than usual. Give a one-paragraph answer summarizing what you could find and what wasn’t available; give an “information topography” report about what kinds of information are even available.

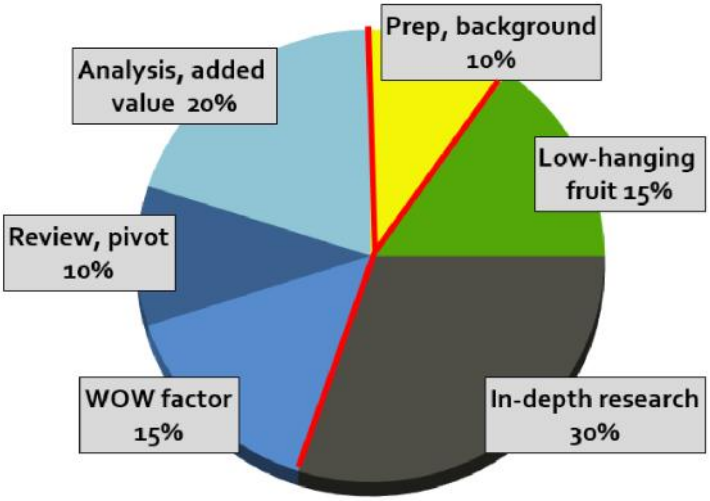
Budgeting Your Time

Even free resources aren’t truly free, as they still require you to spend time using them. When budgeting your time, be clear on how much your client wants and when good enough will be good enough. Beware of data dumps and providing more information than your client will read.
Build in "pause points" during your search process to review, reevaluate and revise your approach. Pause every 15-30 minutes and ask yourself “Is this avenue fruitful? Should I pivot? How much more time do I have? Are there any outliers I should follow up on? What am I missing?”
In addition to pausing every 15-30 minutes during a project, pause when your search is not turning up what you expected. Ask yourself: Am I using the right approach & works? What does NOT finding anything tell me? What are my unquestioned assumptions?

Pause when you have found what may be "good enough" and consider whether you should stop at that point. Will more information actually improve the outcome or just overwhelm your client? What can you do to make the information more useful to your client?

Be sure to budget time for all media options – the Web, social media, premium services, primary research, internal resources, etc. Spend time on the WOW factor – adding something to the results that make it clearly valuable than and superior to a simple web search.

For a 1-hour project, I recommend breaking out your time as follows:
5 minutes preparation, getting organized and taking a deep breath
15 minutes for low-hanging fruit
5 minutes to review, pivot
20 minutes additional focused research
5 minutes to identify a WOW
10 minutes reviewing, polishing, packaging, writing an executive summary



Getting the Most from Your Info Budget: Leveraging Digital Content for Greater Value

It's not just info pros… Many professions struggle to demonstrate their ROI. For example, the ROI for a trade show booth may be measured by the number of prospects spoken with, the number of sales leads, the number or sales appointments or actual new contracts. The ROI for a social media marketing campaign may look at the cost per lead, the cost vs. Google Adwords, or the conversion rate.

Are we talking about the right things? According to a study by Taylor & Francis (goo.gl/tTujFS), 93% of librarians think usage is a key value metric, whereas only 24% of administrators think usage is an important metric. It's crucial that we count what matters to upper management!

What are you to the bottom line? Overhead or contributing to revenue-generation? (Why not both?) Think about your administration's concerns. Show that your budget is spent on resources that are being used. Identify and highlight library services/resources that valued by valuable employees. Show tangible ties between library activities and strategic goals of your organization.

Link expenses to operations that tangibly advance your organization's goals, whether that is supporting an endowment campaign, providing decision support for a project team, helping the organization pivot to new strategic focus and audience or to support staff professional development and retention.

Link your service to outcomes using the formula:
Key client group's goal:_____
Info needs:_____
Measurable impact/outcome:_____
Examples:
Goal for CFO: Reduce corporate risk
Info needs: Due diligence research, workshops on copyright use
Impact/outcome: Risks identified, increased use of licensed photo collection

Goal for hospital: “Improved clinical care”
Info needs: Patient care questions, clinical guidelines
Impact/outcome: Revision of clinical procedures, provide new knowledge

Goal for university: Increase graduation rates & job success
Info needs: Analytical search skills, skills in researching employers
Impact/outcome: Improved student success

Look for metrics that connect your operations to the success of your organization:

- % of awarded patents supported by library research
- # of regulatory approvals supported
- # of strategic initiatives supported
- Key sales goals met

Show the impact of the library on your organization:

- Supporting professional development leads to improved employee retention
- Effective outreach to patient groups leads to better health outcomes

Show impact of the library on your organization’s staff. Look at information flows and pain points, and show how the library helps reduce the time spent searching rather than finding information, duplication of information-gathering efforts within a team, and underutilization of resources.

<u>Looking at Your Time Strategically</u>

Find new ways to add value. Identify new initiatives and projects within your organization; they may need extra support and you can set up good information management and retention practices. Identify information-intensive programs in your organization and reach out with help on information evaluation, acquisition, enhancement and management.

Leverage your fee-based services. Look for ways to raise value and awareness with simple data-mining and data visualization tools or a deliverable that is designed professionally. Invest time to customize your users’ settings or profile to improve strategic searching and enhance their results. Regularly remind users of library resources they can’t access themselves; raise users’ DISsatisfaction with their search results from free resources.

Use the online services' tools. Install library/search widgets or APIs at informational pain points where you know that users are seeking information. Embed content with project groups; assign a librarian to a high-profile project or team.
Build strategic library usage by reviewing usage statistics and identifying underused resources. Promote those sources in internal communications, in training classes, on your library web page, etc. Remember that face-to-face interactions have impact, especially in a digital world. Plan onsite awareness-raising events where your clients are; offer weekly drop-by sessions; cultivate the people in groups who are particularly interested in information resources.

Calculating the True Cost of Your Time

Annual salary * 1.3 = fully-loaded salary (the full cost to your employer)
52 weeks – 6 weeks = 1840 work hours/year
Full salary / total work hours = full hourly rate
So, for example, a €65.000/year salary = €46/hour; a €100.000/year salary = €71/hour

A client's minutes add up! If you save your client 15 minutes twice a day by teaching him how to use a specialized resource rather than Google, you have saved your organization €8165/year.
If you customize a client's profile on an online service so he can search more efficiently and save 15 minutes a day, you have saved your organization €4083/year.
If you create a shared news dashboard for a 6-member team and save each team member one hour a week, you have created a €20.000 annual savings for your organization.

Also highlight the easy ways of showing the library's cost-effectiveness. You can compare the cost of an individual acquiring a resource vs. the library acquiring it and circulating it multiple times (one formula is to take 20% of a material's cost {reflecting the cost of a used copy of a book), multiply that by the number of times it is circulated). Even interlibrary loan operations can be shown to be cost effective when compared to paying €50 to purchase an article or to the cost of NOT having the information.

According to a 1993 study by the Special Libraries Association (https://www.sla.org/wp-content/uploads/2013/11/Special-Libraries-Increasing-the-Information-Edge.pdf), 75% of users said they would not read an article if it were not easily accessible. This study also found that professionals said they got €275 of value per article read, giving an ROI of 8:1.

Outsell Inc. found that a library interaction saved a user 9 hours (" Information Management Under Fire: Measuring ROI for Enterprise Libraries", 2007). If a typical library interaction requires 1 hour of info pro time, every library interaction saves €593! What are your library's savings?

Evaluating Digital Content

When evaluating information providers, cost and content drive the conversation. This misses important considerations such as features, functionality, ease of use, user support, license restrictions, etc.

When evaluating total cost, consider:

Does the pricing work for your organization's employees and work flow?

How much IT support will you need to deploy, modify and maintain this service?

How much will the vendor do?

How much user/searcher support and training does vendor provide? In multiple formats and media? With 24/7 coverage? And do you want users calling the vendor?

Can you get usage information by department, by resource, by region?

How much can the user interface be customized for your users?

How easy is the platform for new users? For experienced users?

Does the platform include DOIs to refer users to already-licensed content?

How agile is the vendor? Do they cover non-traditional media? How do they respond to the perception that everything is available on the web for free?

What can you do with the content? Are there data visualization or data analytics tools? Are there license restrictions? What about redistribution rights?

Assign value to ROI factors. The "sharing economy" leads to a low awareness of copyright; having strong licensing means risk reduction. Privacy laws are leading to more gaps in web news, and search engines contribute to confirmation bias, resulting in less reliable search results.

Challenging Your Value Assumptions

To help identify your value, ask yourself:

What outcome is this research supporting?

How can I make this research more useful?

Where do our services have greatest impact?

And ask your clients:

Where would you go if you couldn't call us?

How much do you pay for ad hoc info purchases?
How much time did this project save you?
How do you describe our services to a colleague?

What new services can you strategically take on? Where are your clients' work flows and where can you have the biggest impact? You can use your information skills to become the User Whisperer for internal web sites, helping your web designers understand what questions users are trying to answer. Show your users better ways to monitor the news or social media, or track the impact of your organization's intellectual property.

Likewise, look at what services you can strategically drop. What service is no longer adding value? What resource is no longer being used? Who else is doing this? What can be done instead?

Big Data and Info Pros: Let's own the conversation

Where's our role in big data? We can locate, compare and evaluate data sources for clients, even if we aren't involved in the design of big data projects.
Information has evolved from print indexes and card catalogs to bibliographic databases, then to full-text databases and now to text and data mining of "semantic triples" (a.k.a. info bits)

Structured information forms a *semantic triple*, expressed as Subject — predicate — object. So, for example, the sentence "the sky is blue" can be interpreted as "sky — has_the_color — blue".
Semantic triples from a Wikipedia article on shingles might include:
shingles — is_also_called — herpes_zoster
shingles — is_caused_by — varicella_zoster_virus
varicella_zoster_virus — is_treated_with — acyclovir
immunosupression — is_risk_factor_for — shingles

Semantic triples from a bibliographic citation might include:
Article_X — has_author — Doe,_John
Article_X — published_in — *Heredity*
Doe,_John — has_affiliation — Drexel_University
Article_X — funded_by — grant_123
Article_X — has_subject — Alzheimer's_Disease

Google Books Ngram Viewer (books.google.com/ngrams) is an example of a large TDM project. Millions of books were digitized and each word and sentence parsed. While it uses a non-intuitive syntax, you can search for ANY word modifying another word.

Linked open data enables meaningful connections across content. Normalized data (made consistent with metadata) leads to enhanced discovery beyond full-text. For example, Springer Nature Journal Suggester (journalsuggester.springer.com) lets you provide the manuscript title and abstract, and it recommends where to submit your manuscript.

Libraries can use big data to help searchers gauge article impact, by adding an API that looks up article DOIs in a search result, and calculates the number of citations to that article. Libraries can build an internal open-access image repository and monitor open access journals for specific types of images needed for their clients' research. Librarians can help discern high-quality conferences from

predatory ones by building a tool to chart the number of institutions represented by the conference speakers, and to chart citation and reference metrics of the speakers. Life science librarians can ensure more comprehensive searches by building an API that looks up search terms for the MeSH equivalent, and automatically appending all terms within that concept to the user's query. Librarians can create a dashboard for business intelligence - monitoring key publications to identify what institutions are publishing research, who are the most cited researchers at an institution, and what institutions are receiving grants.

Linked open data examples
Springer Nature SciGraph (springernature.com/gp/researchers/ scigraph) lets you view visual patterns in large datasets and see relationships across disciplines and formats.
PubChem open chemistry database (https://pubchem.ncbi.nlm.nih.gov/) uses an API to expand search for all names for a substance.
The Library of Congress Linked Data Service (id.loc.gov) offers datasets of LC's subject headings, name authority files, etc.
DBpedia.org is a dataset of structured data extracted from Wikipedia

Info pros can play a wide range of roles with respect to big data:
We can think creatively about info
We understand uses of structured data
We know what metadata is needed
We care about quality of taxonomies
We understand our clients' info needs, search behavior
We can evaluate and curate datasets, develop a quality and cost checklist, identify what's missing from open access and monitor government agencies, open-data initiatives

Tools for monitoring newly available datasets include dataDryad.org (repository of research data) and Google Dataset Search (toolbox.google.com/datasetsearch). Monitor re3data.org (Registry of Research Data Repositories) to identify new repositories to monitor.
Our challenge is in identifying where text and data mining could help behind the scenes within our organization. It is critical that we conduct *strategic* reference interviews. Clients only ask for what they think we can get, so they often do not think to bring us in on TDM or big data projects, where we can make a big impact. Info pros understand the client's use case; we ask "What's essential? What's nice to have?", and we think creatively about finding answers.

Examples of where info pros can add value:
A university wants to calculate the value of a college degree in order to show the ROI for attending their university. What data sets do they need? Starting & mid-range salaries, by profession; cost of living data; student loan payment data, etc. (Source: *The Accidental Data Scientist* (2015))

A B2C company notices product sales down. The info pro builds an API to monitor news, identifies a dataset of weather, imports it into Tableau, and creates a dashboard enabling users to understand trends.

Info pros collaborate with other groups. Big data projects involve expertise from collection development, cataloging/metadata, IT and outreach.
Info pros see the bigger picture. We think beyond the tool or resource to delivering insight.

You can lead the big data discussion! Create an internal libguide on big data, including licensed sources and licensing restrictions, internal subject experts to consult, best practices, and a collection of tools and tutorials. Try googling `inurl:libguides ("data mining" OR "big data")` for ideas.

Build a TDM checklist, highlighting the library's role, including:

- ✓ Indentify the questions being addressed
- ✓ Identify data elements, functionality needed
- ✓ Identify best resources *for the budget & project*
- ✓ Retrieve, extract & enhance the data
- ✓ Build tools to query the data
- ✓ Analyze the results
- ✓ Archive/curate the data

You can lead the big data discussion! You can create a big-data sandbox, build relationships with your publishers and your legal department, ask your clients what they are working on, and actively seek out ways to engage your clients with big data projects.

Artificial Intelligence and The Future of Librarianship

Artificial Intelligence is here already! According to Deloitte, 25% of businesses have deployed AI in at least a test capacity, and 72% have or will deploy AI within 2 years. According to IDC, 61% of executives see AI as a top risk factor (up from 6% in 2017).

AI is actually a whole toolbox:

- Computer vision
- Facial recognition
- Speech recognition
- Pattern recognition
- Machine translation
- Text analysis
- Robotics
- And more!

What can we teach AI to do?

- What you can do *with 5 seconds of thought* -- recognize an image, identify email spam, translate a phrase.
- What you could do with enough time – reviewing content and detecting patterns

Deep learning is where AI really gets weird…

For example, researchers can feed a deep learning machine an enormous dataset of images from the web. The system creates its own classification scheme for what it "sees". Researchers show the system a cat image and the machine extrapolates from there, identifying other images of cats in its dataset. However, we don't know or control *how* it has recognized what cats look like.

AI isn't thinking (yet…) AI can find unexpected patterns, but humans can:

- Create *new* product
- Design something a *new* way
- Question *why* we do something
- Respond to *unrecognized* needs

What AI can't do:

AI doesn't empathize

AI doesn't have creativity

AI can't tell you the so-what?

AI can't suss out the "question behind the question"

Asking "How can AI help do my work better?" is the wrong question. Transformative technology doesn't do what we do better - it does something we can't do. Disruptive technology disrupts! See, for example, the impact of AirBnB on:

Hospitality and travel industry

Neighborhoods

Local government revenue

Impact of Lyft and Uber

Taxi industry & taxi medallions

Independence for non-drivers

Electronic publishing didn't kill the book. Rather its impact include:

Disintermediated publishers, authors and readers

Created a long-tail self-publishing market

Drove the growth of audiobooks

Changed library patrons' expectations

Self-driving cars affect:

Vehicle design: no need for driver or just-in-case space; different safety concerns

Manufacturing: fewer cars, more variety

Neighborhoods: increased mobility

Cities: drop in public transit use and parking spaces, increased night life

Insurance industry: reduced risk, fewer policies

Privacy: sensors everywhere, easy to track you

AI has had an impact on a wide range of professionals. Journalists use AI story-writing tools for semi-structured data (company earnings, local election results). Lawyers use AI to scan case law and discovery material. Radiologists and pathologists use AI to review images.

AI in the Library World

Behind-the-scenes AI examples include:

Springer Bibanalyzer, which creates structured data from cited references

Computer Science Ontology, a taxonomy built with data mining

Smart Topic Miner, which iImproves classification and metadata of scholarly pubs

Librarians can serve as AI interpreters and guides. See, for example, the University of Rhode Island Library's AI Lab (web.uri.edu/ai/). It includes AI workstations (with open source tools) and curated open data sets, they host meet-ups, workshops and discussions, and they curate online courses, journals and other AI resources. When designing an AI lab, address the different levels of needs of users:

Novices need a basic introduction, a place to play, lots of resources

Practitioners need just-in-time help, want expert chat

Experts want peer group meetings and to be kept updated on new technology & resources

AI can be used in the library to enhance discoverability with tools to allow entity recognition, image and audio recognition, taxonomy development, sentiment analysis, and a recommendation engine. Info pros can use machine learning to extract data within articles, books and other content, and create a tool to find local information or a specialized collection.

Key AI concerns in the library include:

Reliance on deep learning - can we trust algorithms we can't understand?

Need to teach information literacy – users often trust Google more than Wikipedia

Need to advocate for "explainable AI" systems and demand transparent processes

Need to negotiate content licenses and partner with information providers

Need to evaluate and acquire OA content

Need to leverage and share our special collections - do we have the resources to digitize? Should we?

Do libraries have to hire data specialists? Do we provide analysis or "just" the tools?

How do libraries keep staff AI-literate?

Need to work with IT, advocate for data quality and tools, and manage turf wars

AI may drive our clients' expectations, leading them to expect more resources form the library.

Underlying bias in data sets is an ongoing problem and one that researchers may not focus on. This is why we're so needed!

Will AI change the library's form and function? Do we design our services and products in formats and media for people to view the info or in formats for AI to access on behalf of humans? Do users bring their own tools and access the digital version of the library?

What functions could AI take over in your library in 3-5 years?
What are the biggest opportunities and threats of AI and libraries?

# # #

Mary Ellen Bates

Mary Ellen Bates is the owner of Bates Information Services, providing strategic business research to business professionals, and consulting services to the information industry.

She is a frequent keynote speaker, and is the author of seven books and innumerable articles on the information industry. In 2005, she was honoured with the John Jacob Astor Award in LIS, given by the Checkpoint Charlie Foundation in conjunction with Initiative Fortbildung e.V.

She blogs at reluctant–entrepreneur.com, and tweets as mebs.

She, her dogs and her bees live in Colorado, US.

Bates Information Services, Inc.
mbates@BatesInfo.com
Twitter: mebs
LinkedIn: maryellenbates
Blog: Reluctant–Entrepreneur.com