shelf

Why Your Company Search Isn't Google

Contents

You're not the only one feeling this way. In a world where we're able to find whatever we need with just a few keystrokes, company searchaes fail spectacularly again and again.

But why exactly is that? And most importantly, what can we do about it?

01. Your Company's Search Isn't Google

Everyone wants their company search to work like the search engine Google. Why can't document search work that well? It can be a real head scratcher.

After this section, you'll understand exactly why the two systems don't work the same way and why company search is much more difficult to nail.

We all know that Google provides a surefire way to find what or who you need on the internet. But how, exactly?
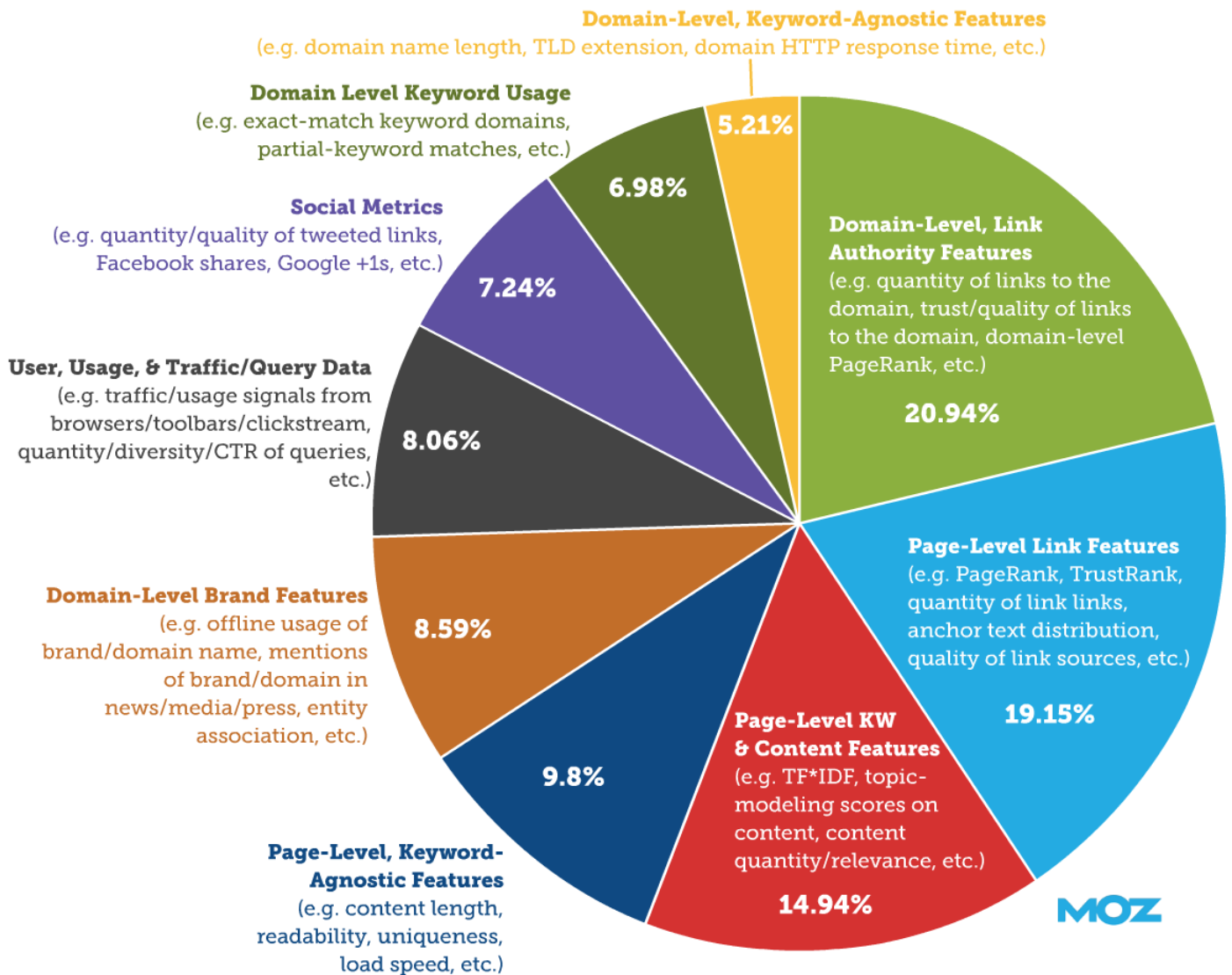
Google's power comes from the World Wide Web; it harnesses the ability to draw from a sea of global online resources, searches, and users.

The online giant has the enormous benefit of learning from over 3.5 billion searches every day. This kind of reach is invaluable for giving people what they want. Imagine you work for a marketing firm that's doing research on the buying habits of millennials. If your research covered 500 people, you'll certainly uncover telling habits and data. But what if you could expand your research to 50,000,000 people? 500,000,000? Your results will be more specific, accurate, and actionable. The same goes for Google's use of the world wide web.

Search engine optimization experts all agree that there are 100s of factors that go into how Google ranks a website and decides whether it will end up on the first page of search results or on page 5. A report by the company MOZ summarizes all these key factors into a more manageable 10. See the chart below.

Weighting of Thematic Clusters of Ranking Factors in Google

(based on survey responses by 128 SEO professionals in June 2013)

Domain-Level, Keyword-Agnostic Features
(e.g. domain name length, TLD extension, domain HTTP response time, etc.)

Domain Level Keyword Usage
(e.g. exact-match keyword domains, partial-keyword matches, etc.)

Social Metrics
(e.g. quantity/quality of tweeted links, Facebook shares, Google +1s, etc.)

User, Usage, & Traffic/Query Data
(e.g. traffic/usage signals from browsers/toolbars/clickstream, quantity/diversity/CTR of queries, etc.)

Domain-Level Brand Features
(e.g. offline usage of brand/domain name, mentions of brand/domain in news/media/press, entity association, etc.)

Page-Level, Keyword-Agnostic Features
(e.g. content length, readability, uniqueness, load speed, etc.)

5.21%

6.98%

7.24%

8.06%

8.59%

9.8%

Domain-Level, Link Authority Features
(e.g. quantity of links to the domain, trust/quality of links to the domain, domain-level PageRank, etc.)

20.94%

Page-Level Link Features
(e.g. PageRank, TrustRank, quantity of link links, anchor text distribution, quality of link sources, etc.)

19.15%

Page-Level KW & Content Features
(e.g. TF*IDF, topic-modeling scores on content, content quantity/relevance, etc.)

14.94%

MOZ

As you can see, most of these factors are external of the search engine itself — meaning they rely on things that the site isn't directly in control of, like inbound links, traffic data, and social metrics.

This is what makes replicating a "Google-level" search within a company so challenging. Let's take a look at the chart below of how Google ranks results as compared to the capabilities of a typical internal search such as Dropbox, Google Drive, Box, SharePoint, or OneDrive.

Search Factors	Google	Company Search
Page-Level Keywords & Content-Content quantity, content relevance, topic modeling scores, etc.	✔	✔
Page-Level Link Features-PageRank, anchor text distribution, quality of link source, etc.	✔	✘
Domain Level Link Authority-Quantity and Quality of links to the domain, Domain-level PageRank, etc.	✔	✘
Page-Level Keyword Agnostic Features-Content length, readability, uniqueness, load speed, etc.	✔	✘
Domain Level Brand Features-Mentions of brand/domains in news, press and media, etc.	✔	✘
User, Usage & Traffic Data-Quantity and quality of incoming website traffic, website usage, and website clicks from browsers, toolsbars, etc.	✔	✘
Social Metrics-Quantity and quality of tweeted links, Facebook shares, Google+1's, etc.	✔	✘
Domain Level Keyword Usage-Exact-match keyword domains, partial-keyword matches, etc	✔	✘
Domain-Level Keyword-Agnostic Features-Domain name length, top level domain extensions, domain response time, etc.	✔	✘

It's not that the company search engine is inferior, it's that there is so much more available criteria to put into the search algorithm with publicly facing content. Inputs like the domain authority, backlinks, social media mentions, press mentions, anchor text distribution, link quality, domain traffic quality, and domain response time are all factors available to a World Wide Web search engine but striped away from a company search.

A company search can't rely on the volumes of data generated on the World Wide Web everyday. It can't crawl other sites and weigh content based on relevancy. It's an unfortunate truth but most company searches rely purely on the content.

This can explain the findability frustrations that people have with platforms like SharePoint, OneDrive, Box, Dropbox, and Google Drive.

But there is hope; there are proven ways to improve company findability by implementing best practices from knowledge management to make the company search experience more like going to a local bookstore or library than wading through piles of clutter in your attic or basement.

While you can't turn a World Wide Web search engine into your company search, you can pick a better one with robust features and capabilities to help your team find what it needs faster.

Qualities of an Incredible Company Search

To counterbalance your company search's inability to crawl the web and learn from the 100s of criteria generated by 3.5 billion people, you need a platform that supplements and complements Findability with other search & browse best practices. It's by leveraging these best practices of Findability that you're able to get closer to the power of a World Wide Web search like Google.

Not all enterprise content management systems are the same. Some are designed for optimizing storage, others are designed for optimizing collaboration. And others are designed to optimize for Findability.

According to the McKinsey Global Institute, the average company sees employees spending 19% of their time looking for information or tracking down colleagues who can help with a specific task.

High performing organizations can reduce the average time employees spend on information retrieval tasks by 5.5% - 6.5%. This represents a 30-35% improvement over average performing companies and equates to saving each employee 2.5 hours a week and 125 hours a year.

The improvement can be realized by any organization simply by implementing best practices in enterprise search technology.

Qualities of Best Practice Enterprise Search

Keyword, Best Match First   ✓

Phrase Match Search   ✓

In-Document Search   ✓

Image Search   ✓

Harnesses Metadata   ✓

Federated Search   ✓

Faceted Navigation ✓

Unified Discovery (Keyword + Faceted Search) ✓

Recommendations ✓

Connections ✓

Tag Library ✓

Badges ✓

Flexible Taxonomy ✓

Let's take a closer look at how these features work and why they're important to have as part of your company's content management system.

In the next section, we'll break down each feature to give you a more complete understanding of these unfamiliar terms and discuss their value to businesses of all sizes.

The Winning Features of Enterprise Search

For most, our experience with company search usually starts with a keyword. Seems like an easy to understand concept, but it's more complicated than that. Powerful keyword searches are more sophisticated than just the ability of an algorithm to look for words. There are many different types of keyword searches which cover a wide-range of uses and the more powerful the search, the more types of keyword combinations and variations it simultaneously covers.

While it seems reasonable to expect that every company search would have exact phrase match—the ability to search for a sequence of words instead of the individual words, most don't. You'd also think that all content management platforms would include the ability to search inside documents like PDFs and Excel files, but they don't.

If you're scratching your head and wondering why, so were we. It's how we came to learn so much about what can help make a company search great.

In this section we're going to cover the best practice features of a company keyword search and provide examples on why each one is important and should be included as an overall search platform.

A. Best Results, Not All

The primary quality of a great search is providing only the best results, not every possible result. Too many content management platforms either don't return any results at all, or provide every possible result, even though they're not relevant. This is a classic Goldilocks example where a lack of results is too cold, and every result under the sun is too hot. Both options hurt the user's ability to find what they need.

Box search can be too cold. This is an example of the notorious, dead-end sorry we couldn't find what you're looking for message.
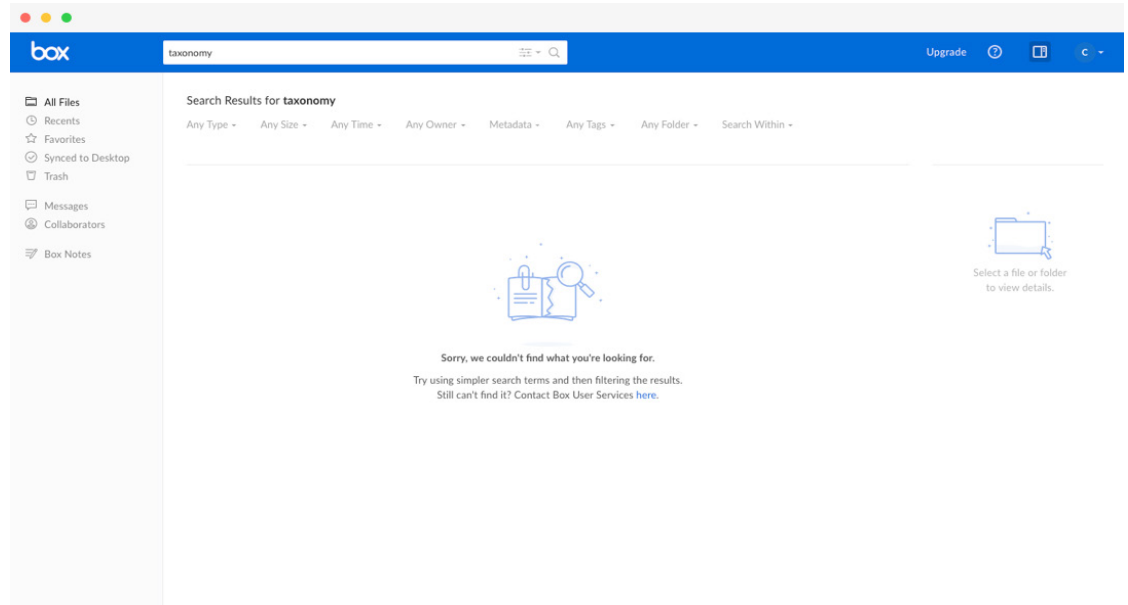

This type of message causes the searcher to reword their search again and again, or start clicking through their folders to find what they need.

Google Drive's search is too hot. It does the opposite and returns way too many results. Here we have to comb through every result to find what we're looking for. In reality, it's no more efficient than clicking through your filters and folders manually.
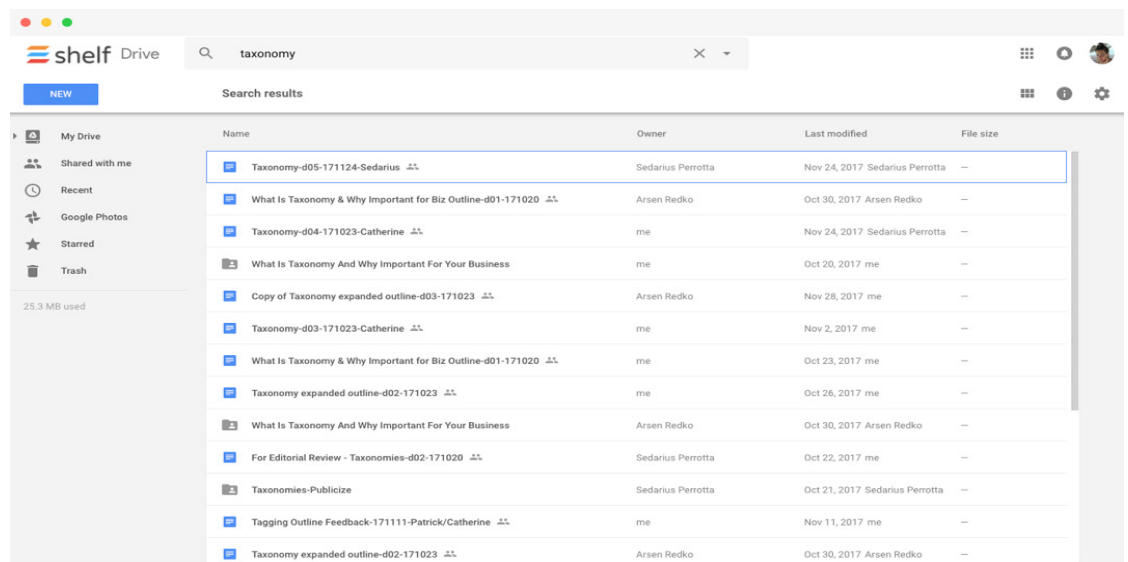
In both cases above we're still not getting what we need. A too cold search gives us too few results and creates a dead end for the searcher, while a too hot search gives tons of irrelevant results which causes stress and frustration. A great company search is *just right* and only provides a targeted list of relevant results.
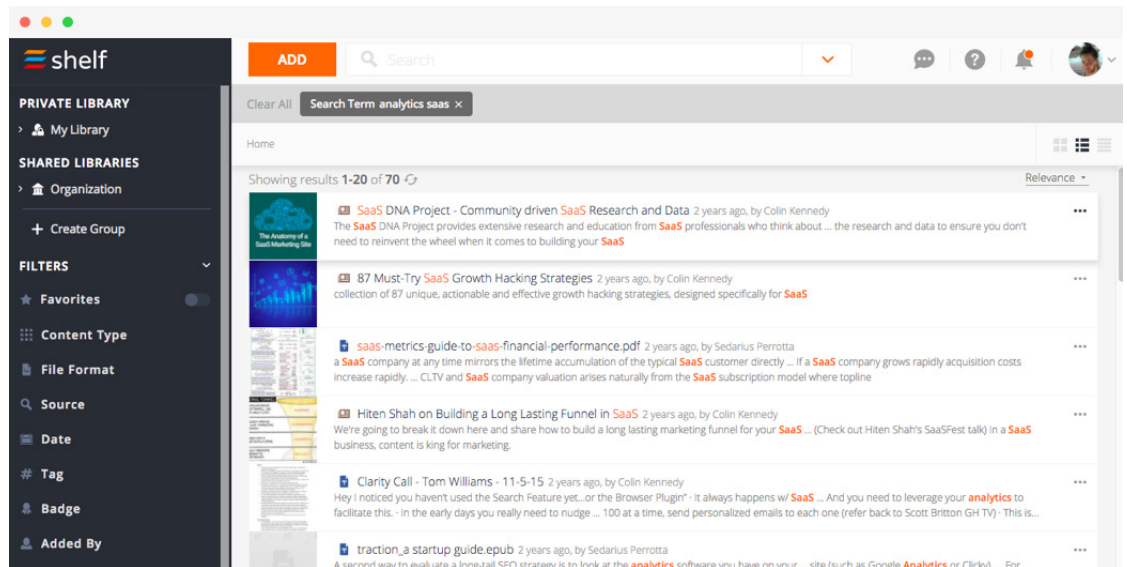
B. Exact Phrase Match Search

The best searches need to be able to search by individual words but also by an exact sequence of words, or phrases. Actually, according to Jumpshot over 86% of all searches are 2 words or more. See graph below;
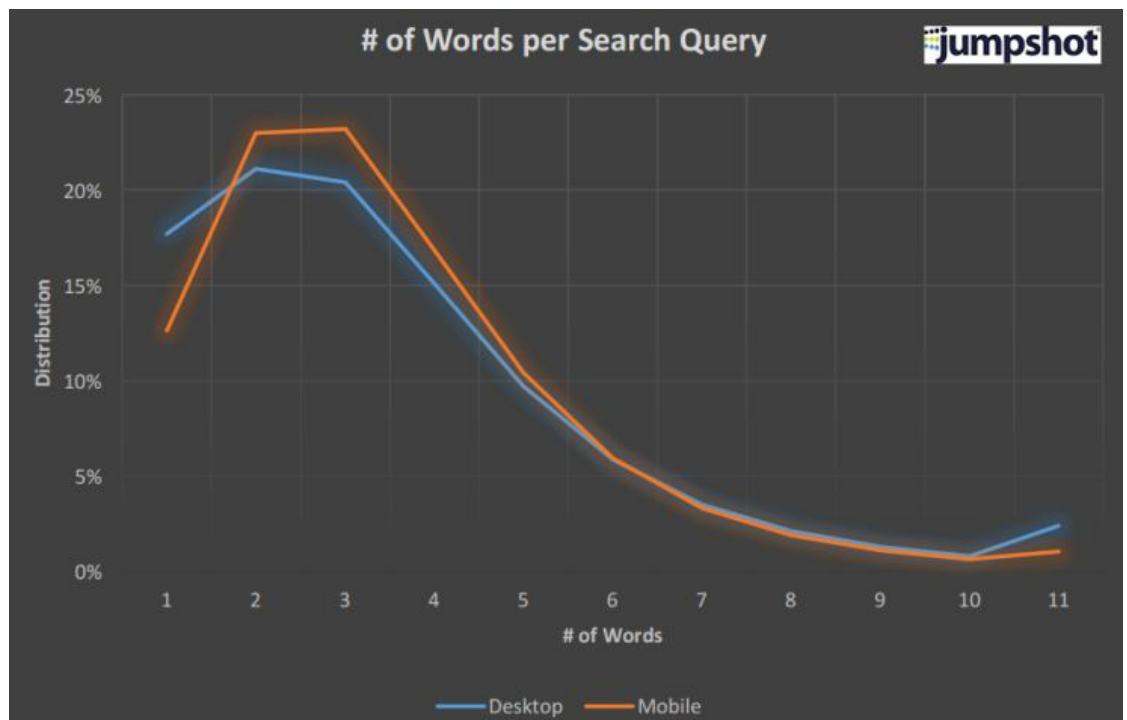
For example, if you were doing a typical search for the keywords "email marketing," a non-phrase match search algorithm would return all results that had both "email" and "marketing" in the documents. While some of these materials may be relevant to your search, chances are the search engine is not returning the most relevant, phrase-matched results first.

A phrase match algorithm would return results that had "email marketing" together in the documents first. This provides a more precise and practical way to locate exactly what you're looking for. It also covers the vast majority of use cases of the average person searching. Without an exact phrase match algorithm only 14% of all user searches will be hundred percent relevant.

Surprisingly, many platforms, including Sharepoint, have trouble with this feature. In fact, sometimes the search is so poor, it is hard to understand what the the results have to do with the keyword search. In the example below, Sharepoint returned a bunch results for email marketing but does not indicate where the result is coming from. All that can be seen is a document title that may or may not have the keyword.
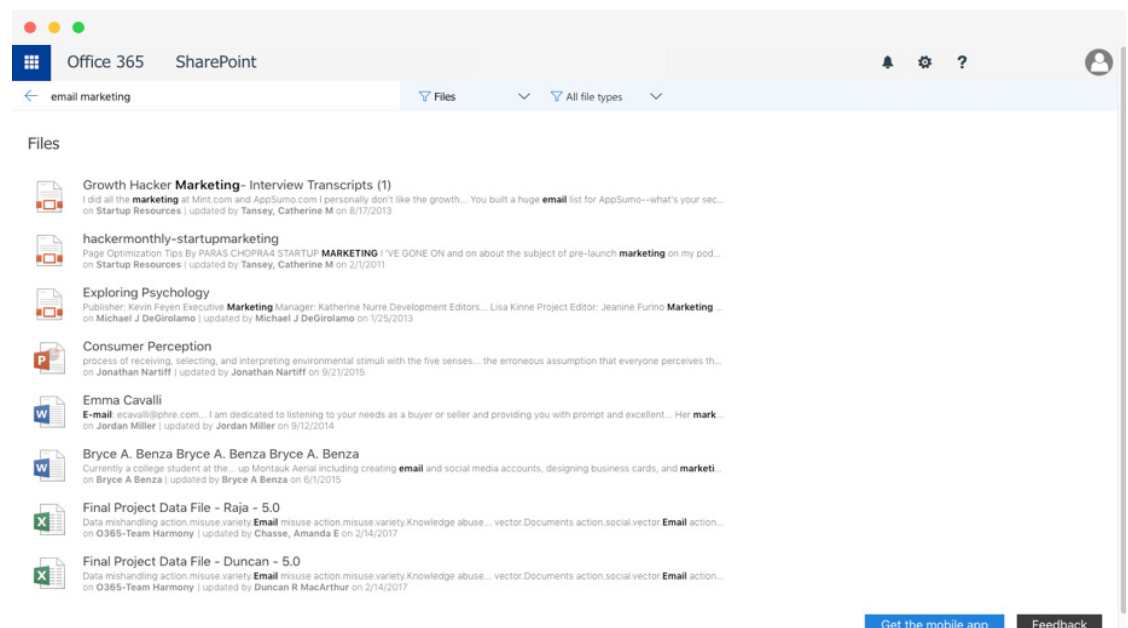
Alternatively, the same search in a platform with exact phrase match, like Shelf, provides much more targeted, focused results matching "email marketing" first.
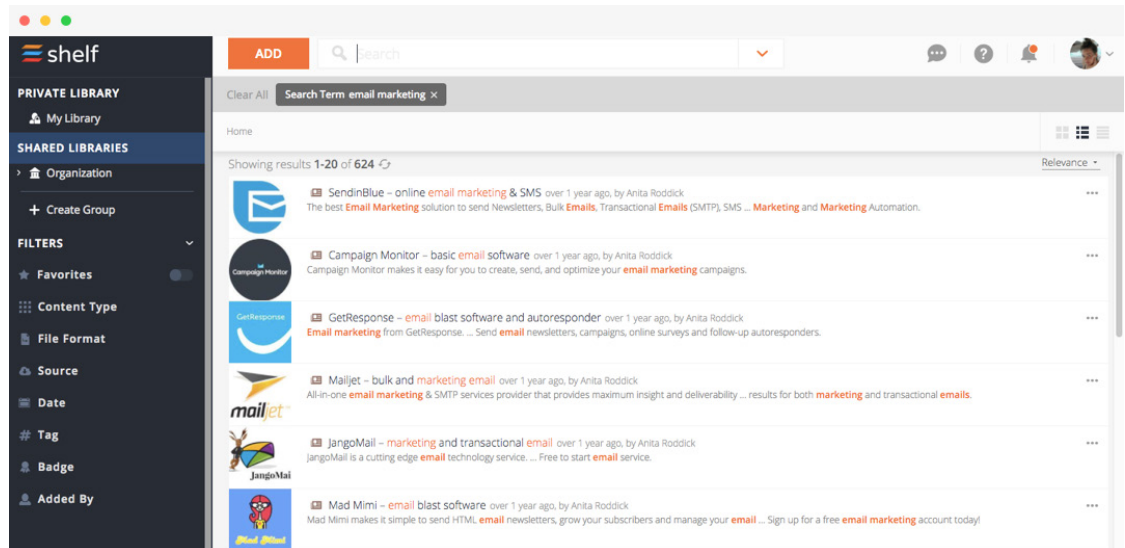
The bottomline is that you want to have phrase match as a key feature of your search engine because it covers 86% of all user searches. Without it, any search algorithm is incomplete.

C. In-Document

In-document search is the ability of a search engine to search within ALL types of documents, including; PDFs, PowerPoint, Word Docs, Excel spreadsheets, CSVs, and more.

This is a powerful and necessary component of the best search engines because it allows you to go beyond the basic title search that most platforms rely on.

For example, say you heard about an interesting HR study that involved the company Avanti. You want to show some colleagues but can't remember any other identifying details so you type "Avanti" into Dropbox.
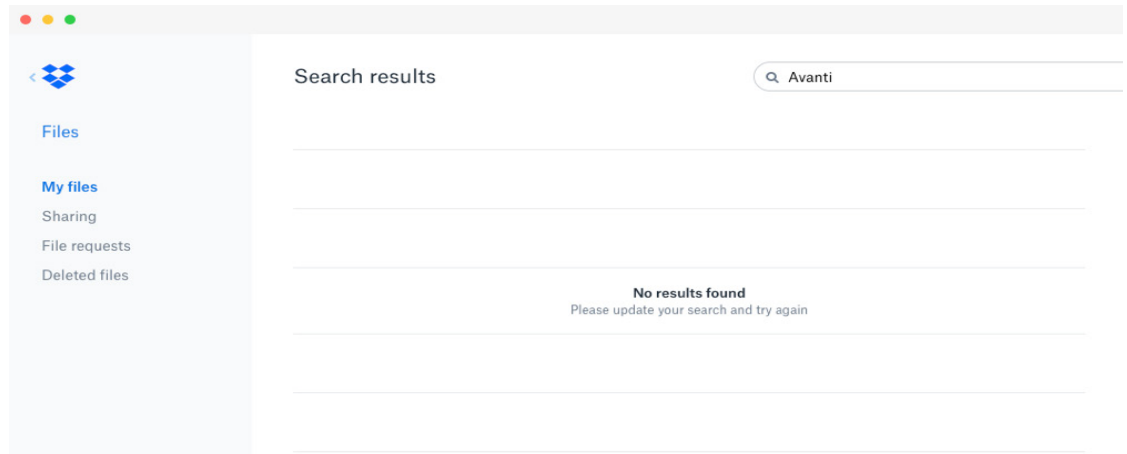
Most document platforms essentially rely on a title search, and because "Avanti" doesn't appear in the title, the platform doesn't correctly match the search with the results even though the same file has been uploaded.

Now try the same search in a platform that has the technology to search inside various document formats. It turns out the title of the document was, "What is Talent Intelligence", but the user only remembered that it was by Avanti.

With in-document search the software is able to locate the file correctly.
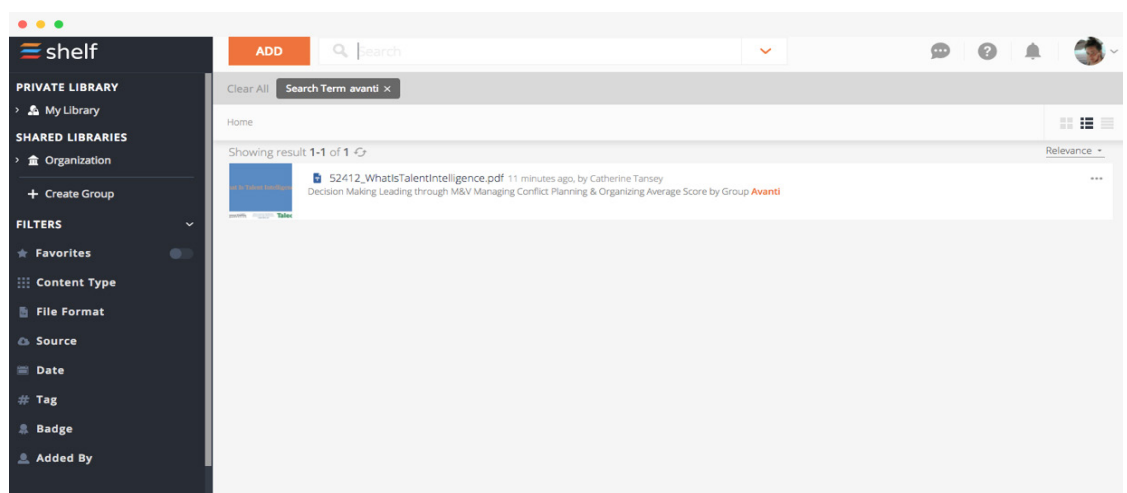
In-document search lets you go beyond searching by simply by title. You can search inside the body of any document, whether it's 50 pages or 500 pages. With this feature you call pull specific material out of PDFs, word docs, excel spreadsheets and more with a few simple keystrokes, and continue on with your work.

E. Image Search

Image search is much more than the ability to search for an image or graphic. Rather, image search, known as optical character recognition (OCR) is technology that allows you to search the text within images, infographics, and photos.

This is helpful because you're able to look for information contained within images like never before.

Say you've become well acquainted with Max Marmer's Startup Genome project and you'd like to share a chart of his with a colleague. You can't remember what it's called but you know "Marmer" is contained within the graphic somewhere.

With OCR, a quick search for Marmer returns four results:
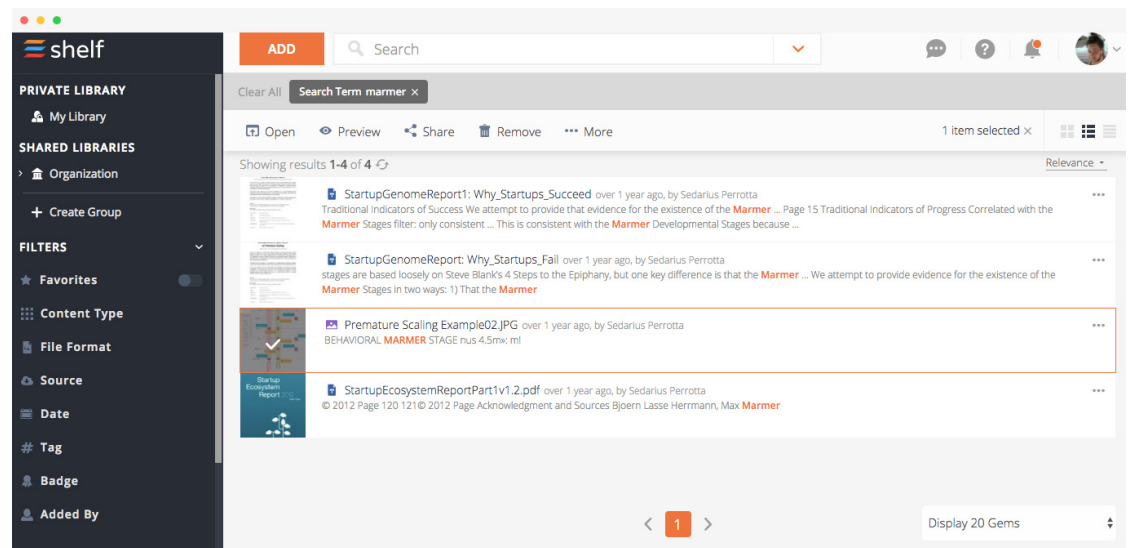
Knowing it's the graphic you're after, you open up the third down to find just what you're looking for:
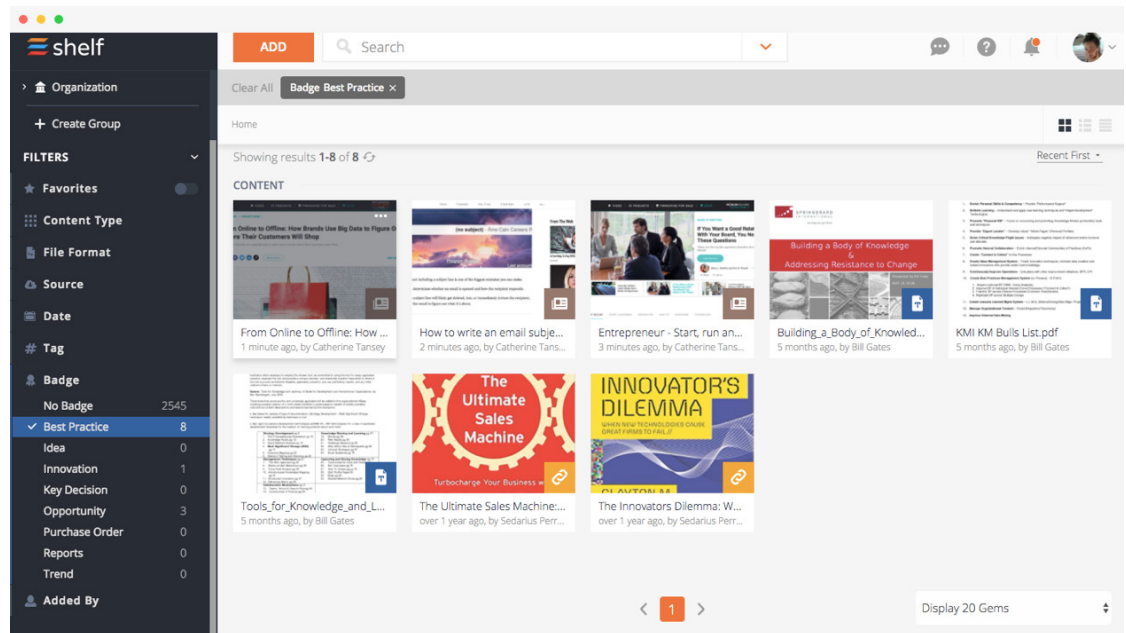
OCR is a powerful tool to have in your arsenal and brings a new level of granularity to your search by allowing you to look within the text of images, photos, infographics, maps, menus, or graphs.

Search isn't Enough:
The Power of Browse & Discover

- **Power of Metadata**
- **Federated Search**
- **Faceted Navigation**
- **Unified discovery**
- **Advanced Search**
- **Recommendations**
- **Connections**
- **Tag Library**
- **Badges**
- **Taxonomy**
- **Section Summary**

When a user doesn't know the right search term, keywords alone don't help them find what they're looking for. The best company searches go beyond keyword search and incorporate other ways to pinpoint exactly what the user needs. The qualities discussed next are all ways to help your users browse and discover.

Power of Metadata

A great search is built off of processing rich metadata, so it must be capturing the metadata of the source material. Many platforms just collect basic information like title and file type.

What happens when user's don't remember the title? Knowing the file type alone isn't going to help them find anything.

Most likely, users will remember other things about the file, such as; who added it, when it was added, and where were they when it was added. By remembering any combination of these other factors a user can quickly pinpoint information. This is why harnessing rich metadata is a foundational element of great findability.

For example with metadata capture, you could filter by "Added by" and see all the content contributed by a selected user.
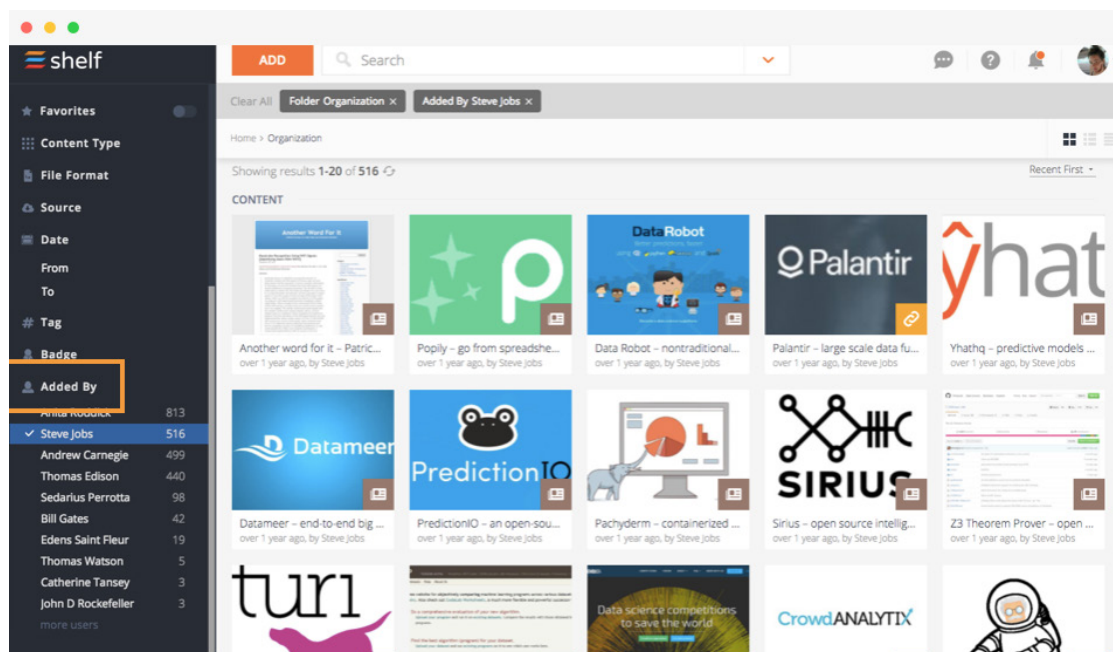
or filtering by Content type, videos, see videos in a certain folder.
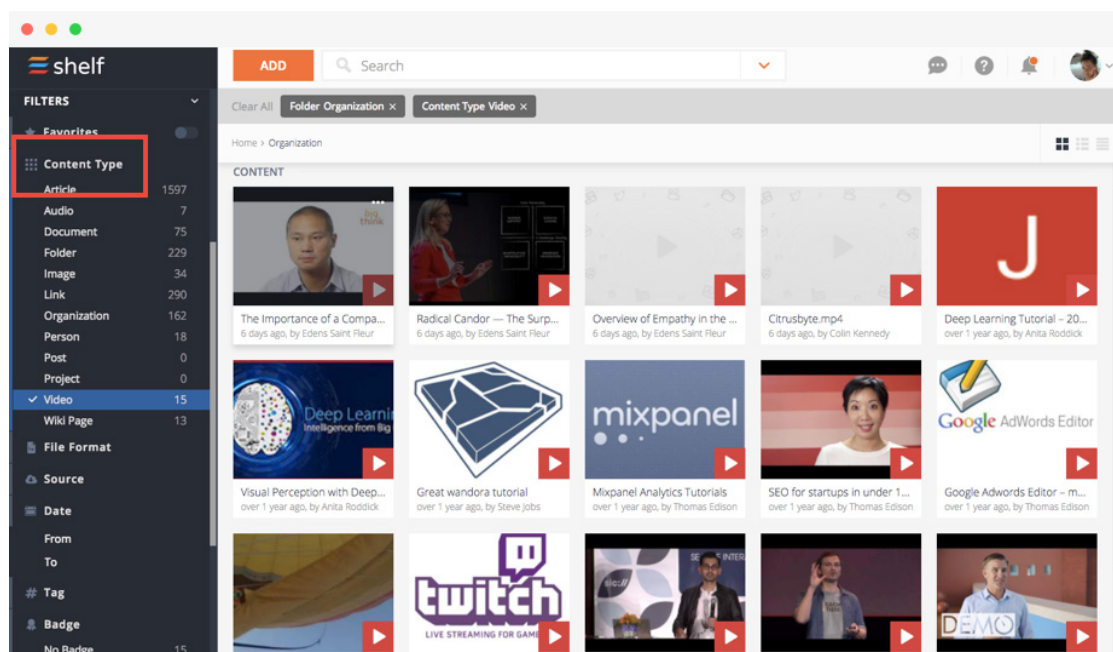
When a content platform leverages Metadata it gives users additional ways to find the things they need, when they need it most.

Federated Search

Modern workplaces have lots of different programs and platforms to manage their digital stuff. It's amazing that we live in such a diverse tech friendly workplace, but what about when you go to search for something? Where do you look?

This is why you need federated search. Federated search is the ability to search across different platforms through a single search bar.

For example, with Shelf you can search Google Drive, Dropbox, and OneDrive in a single place after syncing your accounts.
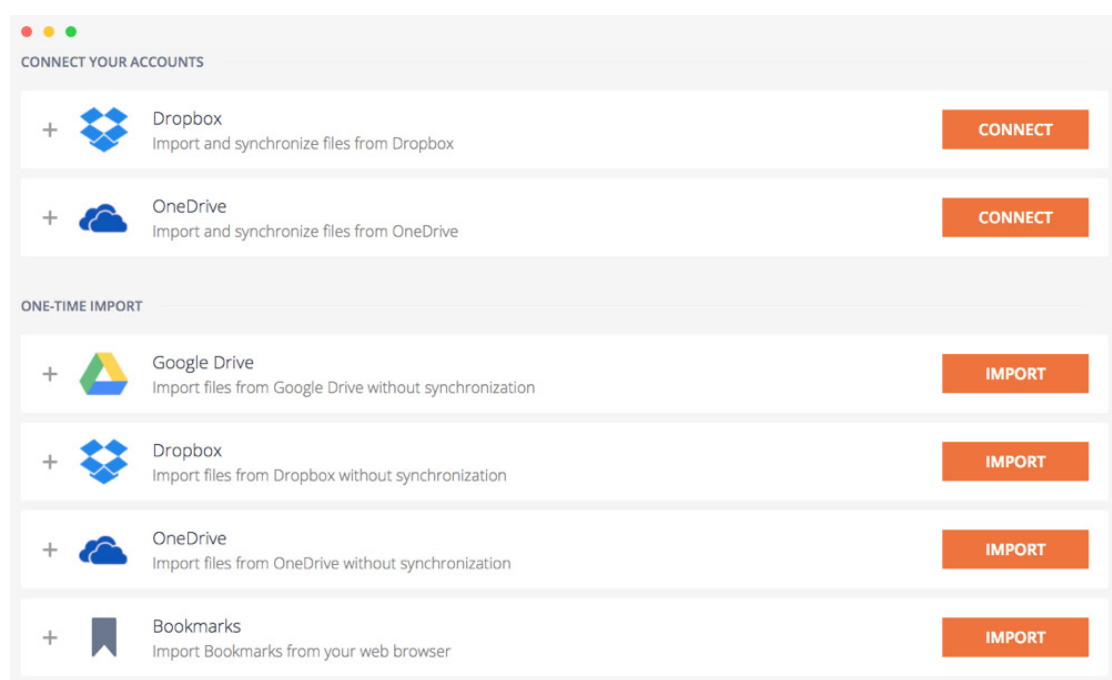
Federated search is powerful because it allows you to draw information from a variety of sources into a single place.

Siloed content can be an achilles heel of findability because users don't know where to search. Federated search alleviates this issue by allowing users to centralize and search through one bar.

Faceted Navigation

Peter Moreville, the author of Search Patterns heralded faceted search as "arguably the most significant search innovation of the past decade". Its done by leveraging metadata to provide users with visible options for continually refining their search queries.

Amazon does this, perhaps, better than anyone.

When traditional search lets you say, "I'd like to purchase a laptop" faceted search lets you say, "I'd like to purchase a black laptop with a 15 inch screen that's under $900, was made in 2015, is four-star rated, and available with express shipping."

Faceted navigation is more powerful than traditional search because the filters help the user continually narrow their search. It also offers options which may have been previously unknown to the user.

Companies can also take advantage of faceted search.

For example, by using the filters on the left panel a user can narrow their search by see by content type, file format, source, date, tag, Badge, or by the person who added it.
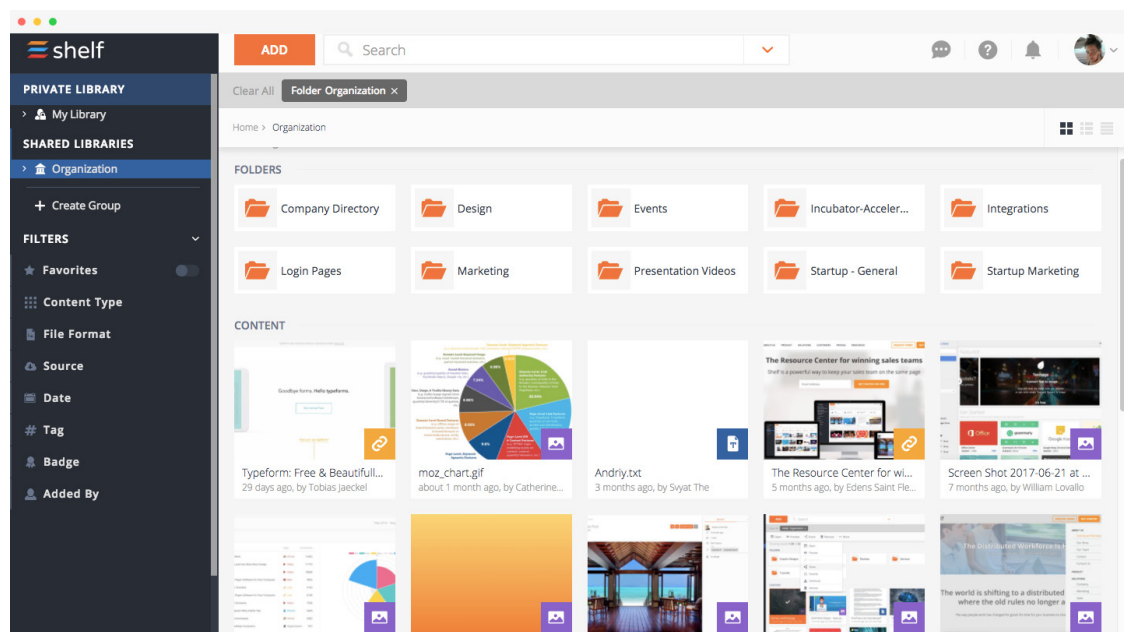
The power of faceted search is you can go from 1000's of results down to less than 10 in three clicks or less without even knowing the name of the title or keywords.

For example, say you're looking for a piece of content added by a coworker but you can't remember exactly what it's called. You know who added it, that it's an article, and that it's about fundraising.

You can first filter by user, Colin, and get 309 results:
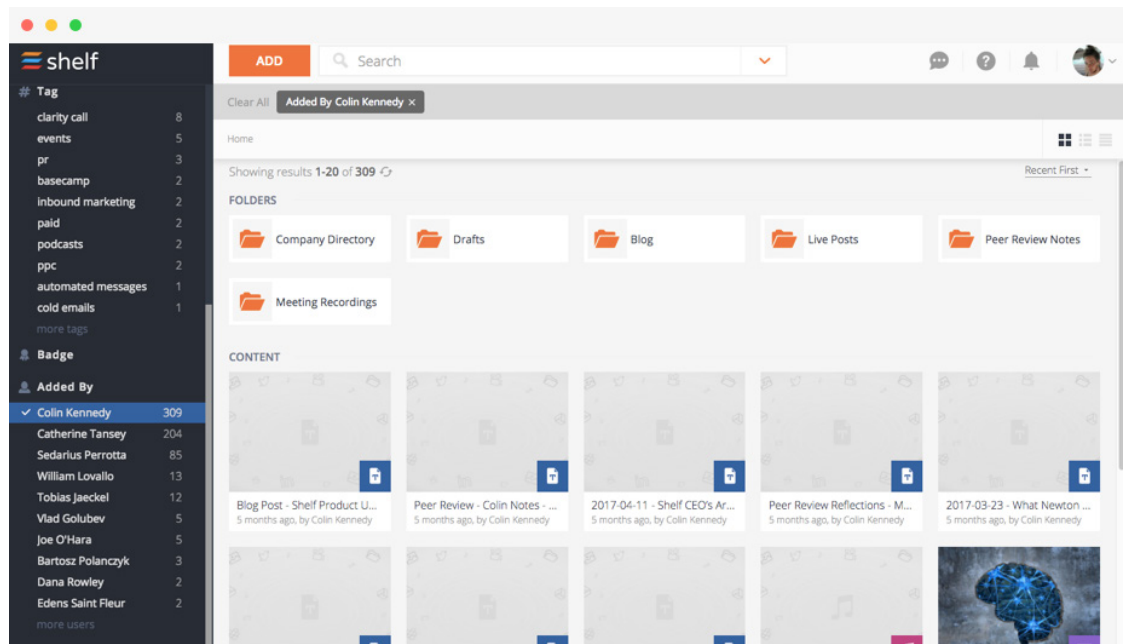
Then by content type, Article, and narrow down to 26:
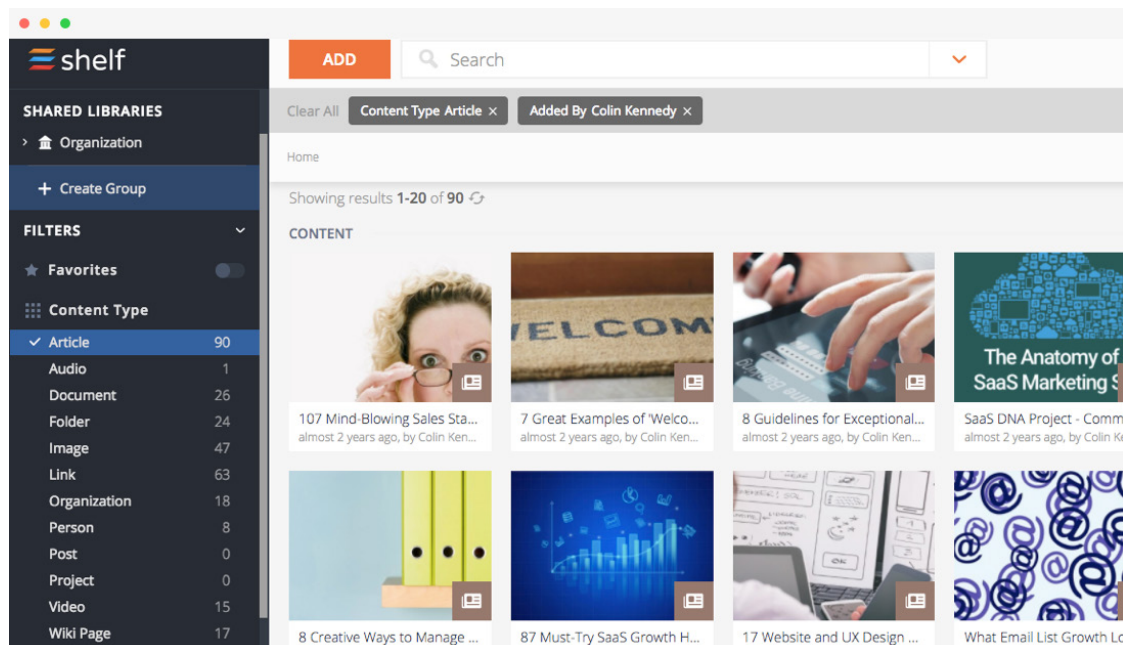
And lastly by tag, Fundraising, and refine to 1:
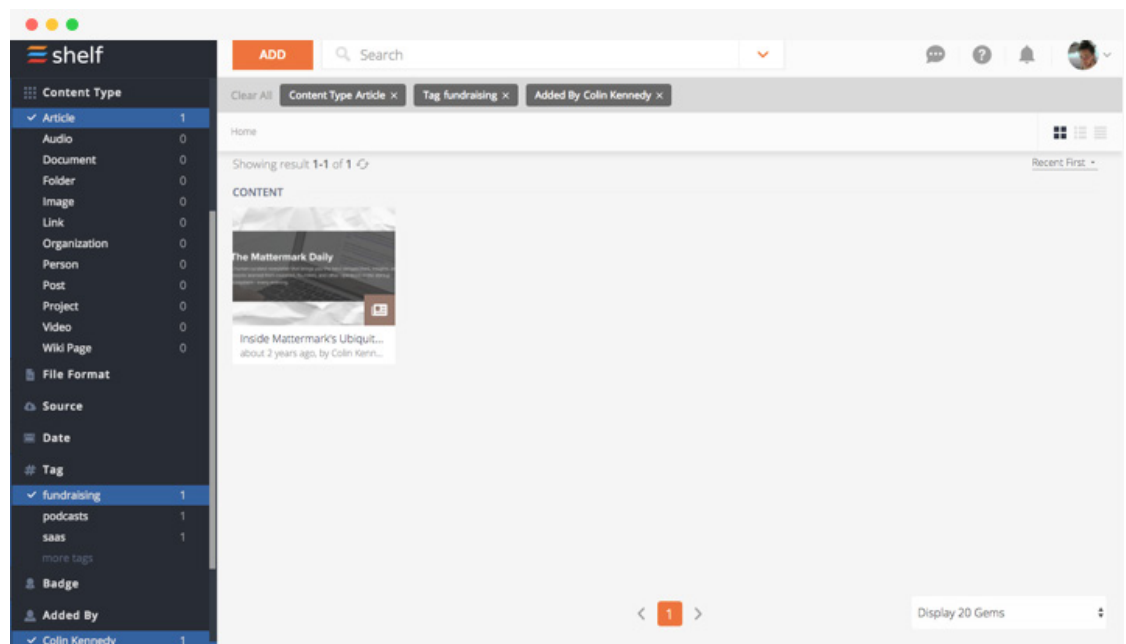
This is the true power of faceted navigation: the ability to filter 1,000s or 10,000s of results with just three clicks. With so many different options to narrow down what they need, your employees will be finding what they need in no time.

Advanced Search

Advanced search is usually available as a dropdown menu from most search engines or databases for those who are confident using it.

Advanced search is helpful because it gives the user options for specificity that basic search doesn't offer and gives you the option to quickly filter and find what you're looking for.

For example, say you're trying to find a specific web link. You know it was about an innovative way to deal with content sprawl. You also remember that it originated from a web clipping and know that it was added by George from Sales.
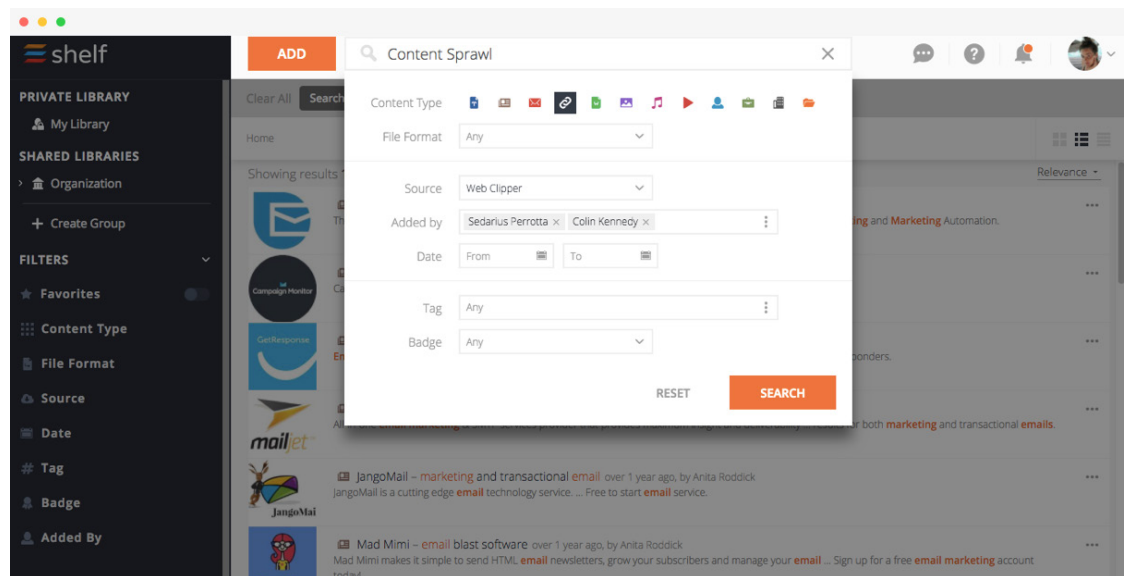
By using advanced search you can leverage all 3 pieces of information at one time through a single search. This enables you narrow down possible results quickly and more effectively then searching via any one method along.

Advanced search lets you plug in this information and what what you need quickly.

Unified discovery

Unified discovery is an interface design where you're able to search, browse, and view results simultaneously. This is helpful because it allows you to browse within folders, use the search bar and filters, and see results at the same time, preventing the annoying back and forth between the results and the content you're actually looking for.

Many knowledge management softwares don't include this feature in their interface. For example, in Google Drive you can't search inside the "recent" filter. Instead, Google Drive's searches across your entire collection of content stored. This is Inefficient and just plain annoying, as you can end up scrolling and scrolling to find what you need.
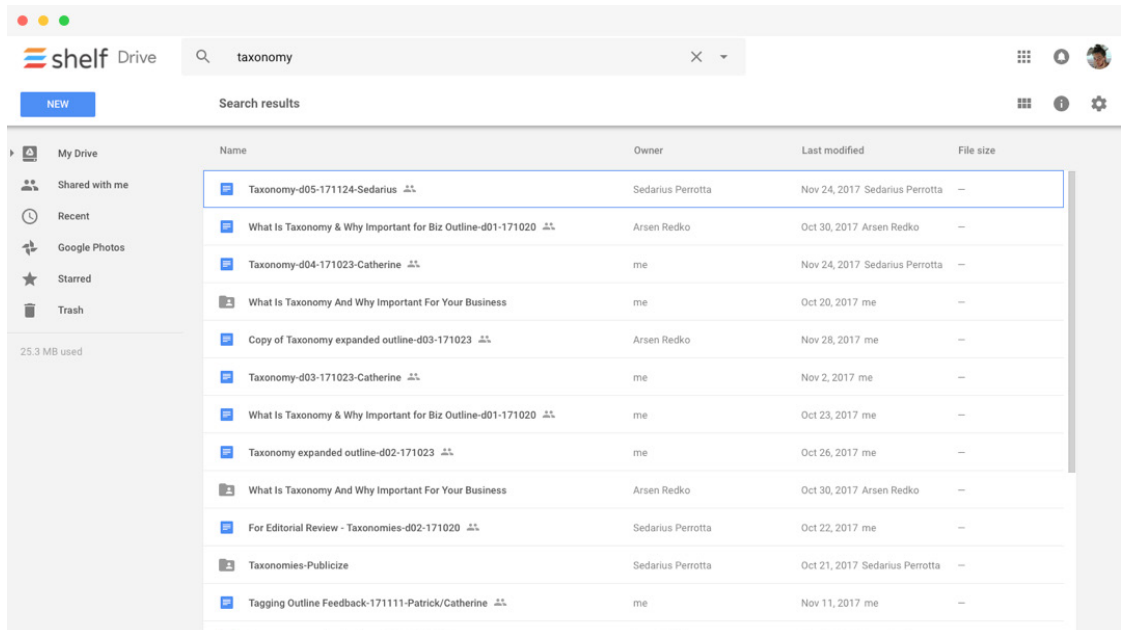
With Unified Discovery this problem doesn't exit as you can "search" and "browse" at the same time.

For example, below the search "knowledge management" was done and then the folder "Organization" was clicked.
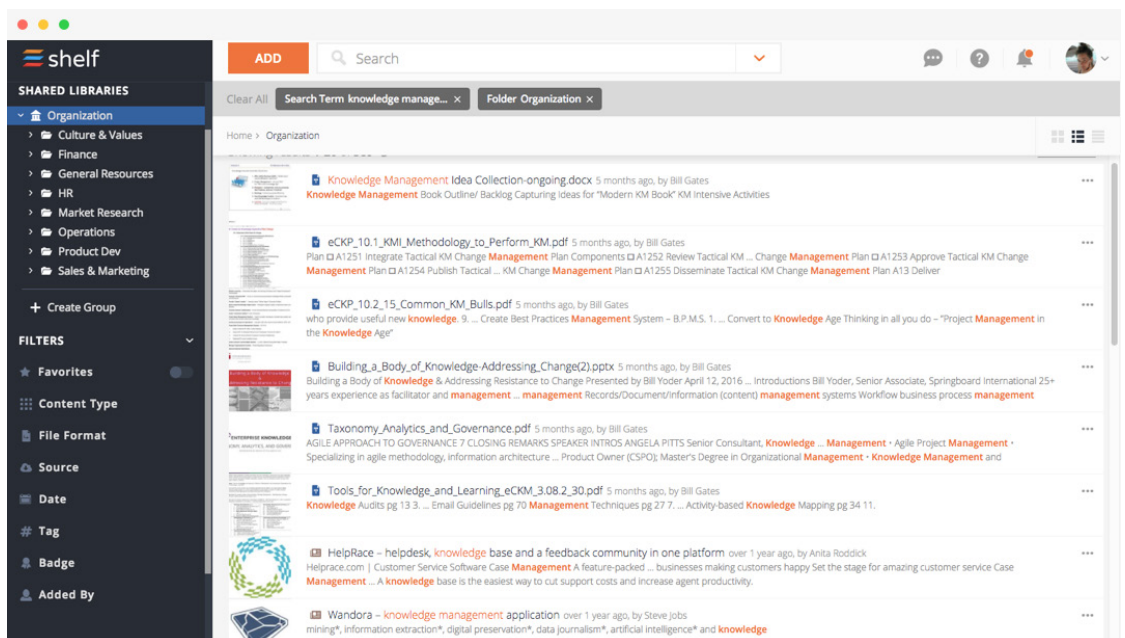
We can then filter by "Recent First" to find the document(s) we've been working on with ease.
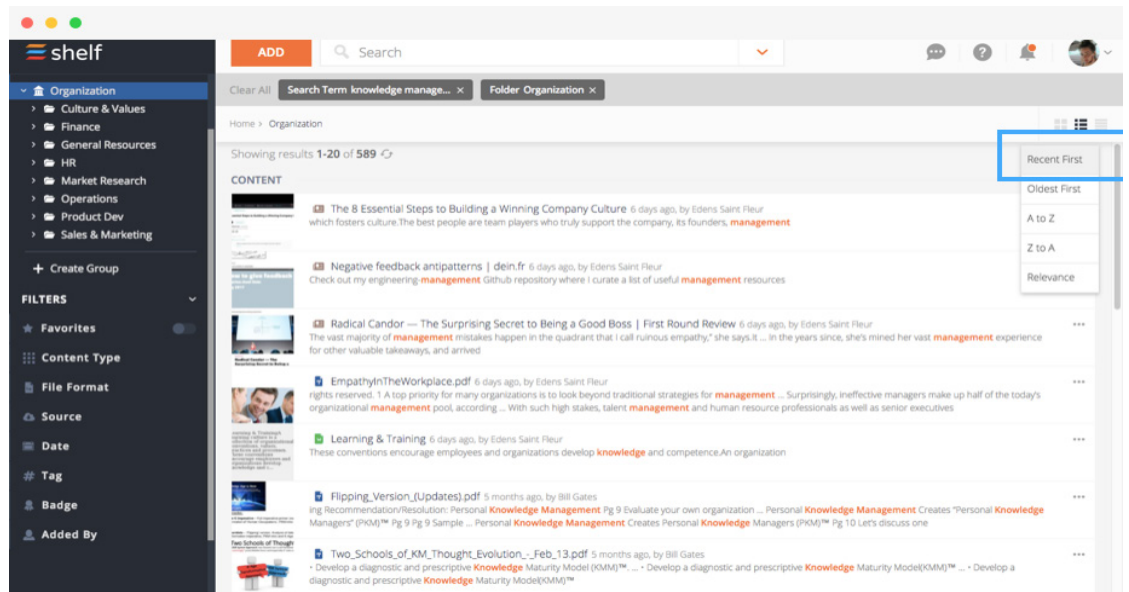
Unified Discovery is one of the most important features of a company content management platform because it reduces the number of abandoned searches and allows users to leverage the power of both search and browse together.

Recommendations

The best search platforms recommend other items, content, or services you're likely to be interested in based on your current search or page. Recommendations help people find things they don't even know exist, and the feature is widely seen in ecommerce and online streaming services.

For example, by searching for the book Good to Great on Amazon, you'll see recommendations that look like this:

An internal company search engine that incorporates recommendations will be equally as helpful for your employees to discover valuable content that they may not know exists.
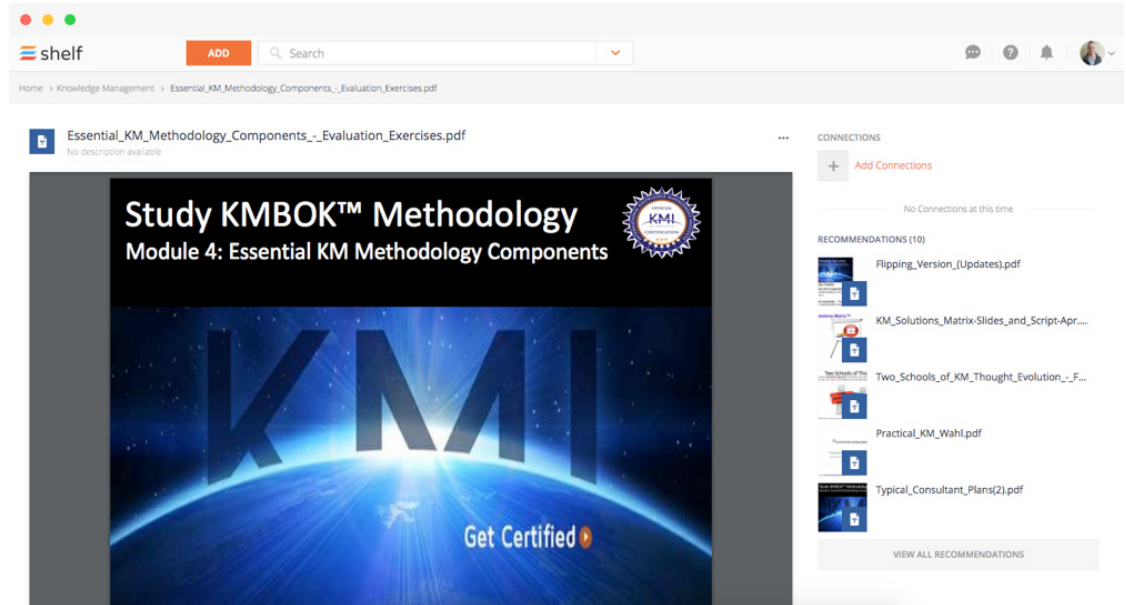
Think of recommendations as a similar experience to browsing a bookshelf at Barnes & Nobles. Although users didn't look for a particular title or document it's a great way to help them find new content that may be more valuable than what they were looking for.

Connections

Connections is a feature of search platforms which let administrators manually recommend content to users by "connecting" them. With a few clicks, admins are able curate the content that users see when on a particular page.

Connections is a powerful tool as it helps ensure users are seeing all the information necessary to a project, client, or training program.
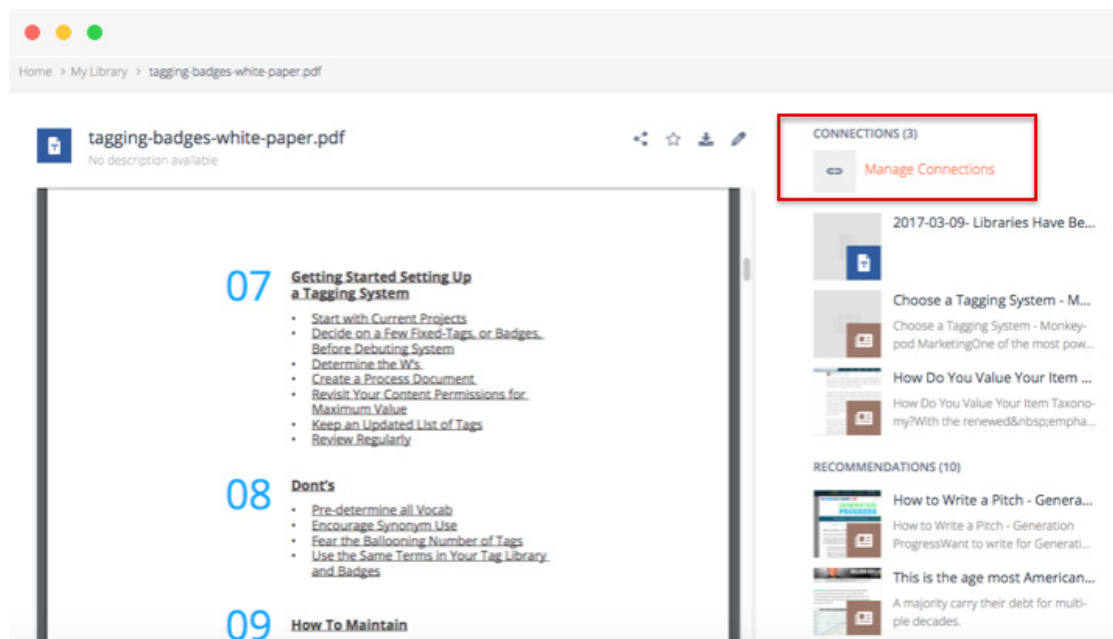
Take the onboarding process, for example. You have a number of different materials which need to be read or viewed.
With a connections feature, you're able to manually group these together to make sure your new employee sees all the necessary information.

Connections is similar to recommendations, but much more personal as the admin get to link the content instead of an algorithm.

Tag Library

Great searches harness the power of rich metadata through search, but also allow you to add metadata to content in the form of tags and view them in a library.

Tags are pieces of metadata you add to documents, images, videos, etc, in the form of your own keywords. Having the option of adding tags to your content is important for your search because it gives you and your users another way to categorize your content so you can more easily find it later.
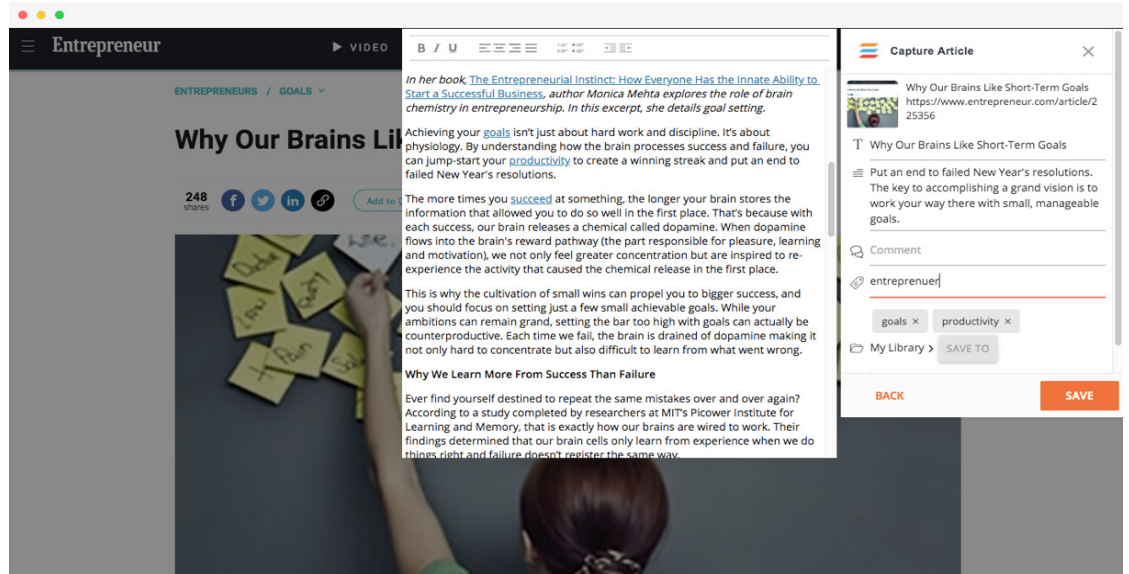
In the image above, we've tagged an article about setting short-term goals. We've added metadata to it that we found useful and appropriate: "productivity," "goals," and "entrepreneur."

Later we're able to search by the keywords we've applied in the form of tags, or filter by the tags themselves in the tag library.
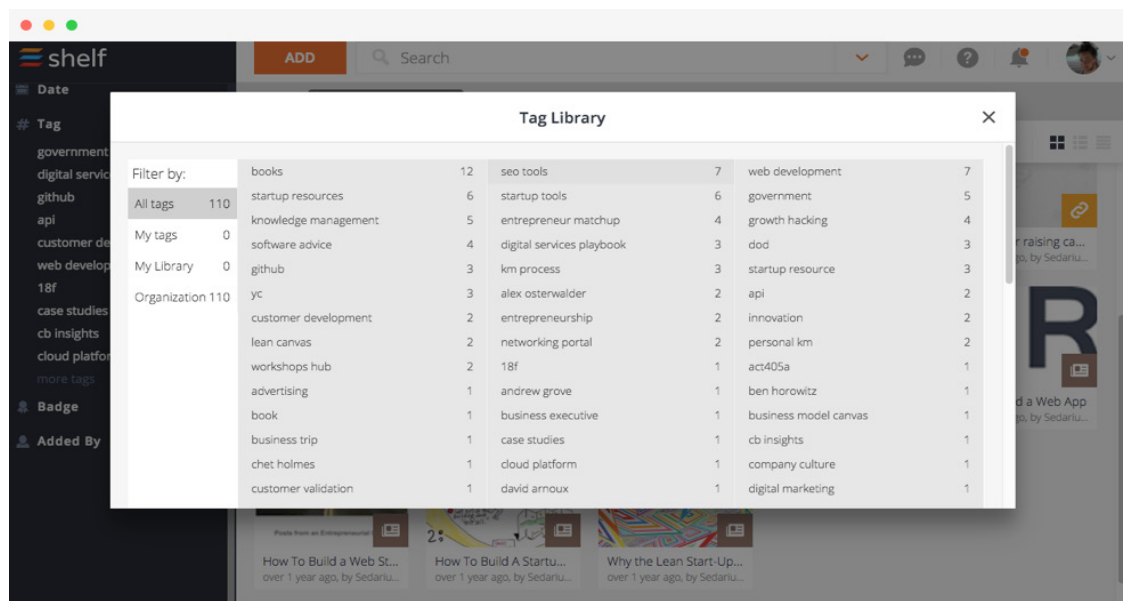
Users can view the entire tag library to see the distribution and abundance of tags, or gain inspiration for their own.

Badges

Badges are another type of metadata users can add to content, but they differ from tags in that they are pre-determined by the system administrator instead of phrased in a users own words.

Badges also differ from tags in that they're best utilized when they center around work-flow indicators (Completed, To Do, Postponed), conceptual topics (Best Practice, Strategy, Trend), or industry-specific commonly used terms.

Like Tags, Badges help make a company search great because they're another piece of identifying information for users to add to content and search by.
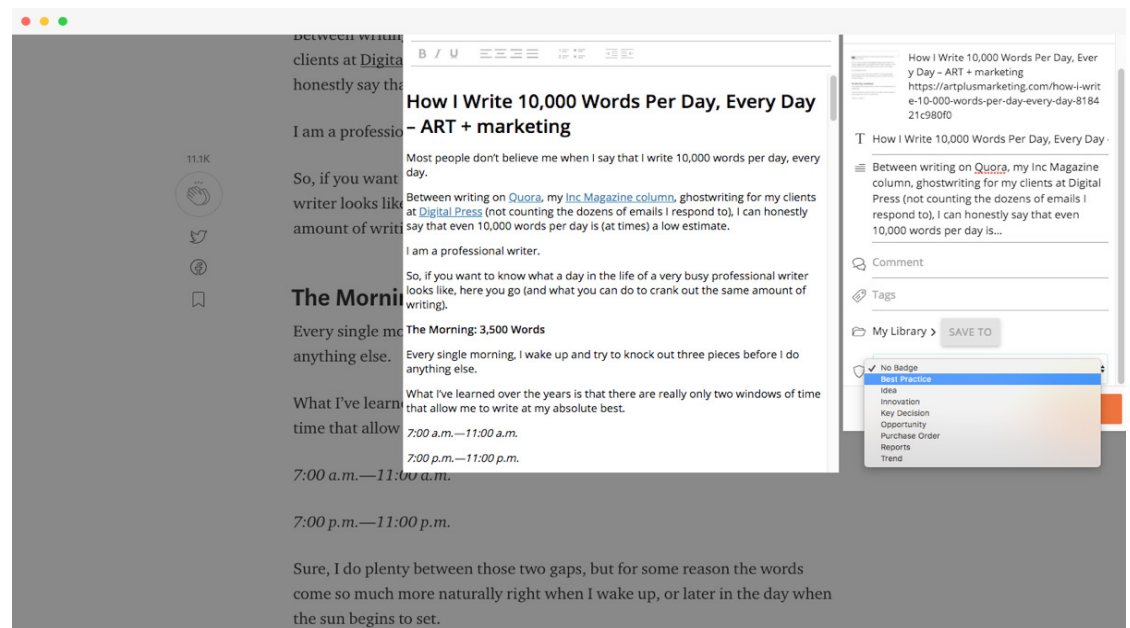
For example, say you run the marketing department of a SaaS company. You've created a blog to help position your company as thought leaders in your industry. In collecting similar web articles which mirror the tone and format you're hoping to strike in your own pieces, you clip them from the web and add the Badge "Best Practice."

Your employees can then filter by the Badge "Best Practice" as they prepare to write their own articles, giving them a focused set of results to draw from.
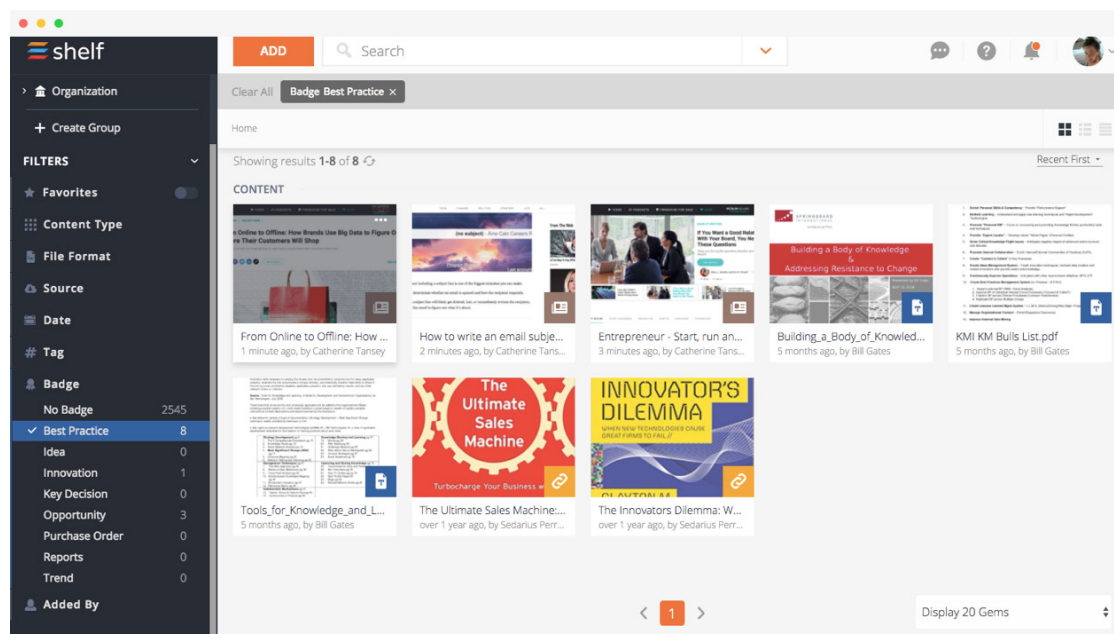
Badges add another level of organization and specificity for you and your team to work from. A quick way to categorize content in predetermined terms, Badges will help make clipping, organizing, and finding easy for all.

Taxonomy

Online search engines have trained us to think of the term search as a keyword search, but that's not really all it entails. Search is everything you or your employees do to find what you need.

A taxonomy, or a structured organization to your content and data, is vital to increasing the findability of your company's content by search. Having your files scattered all over the place in different folders and platforms results in chaos where documents are lost, work needs to be redone, and unnecessary mistakes are made. A good taxonomy helps avoid all of this, but you're still left with a ton of files and folders to search through.

In order to best harness the power of the taxonomy you've designed, you need to be able to refine your search with filters like content type, file format, added by, and so on.
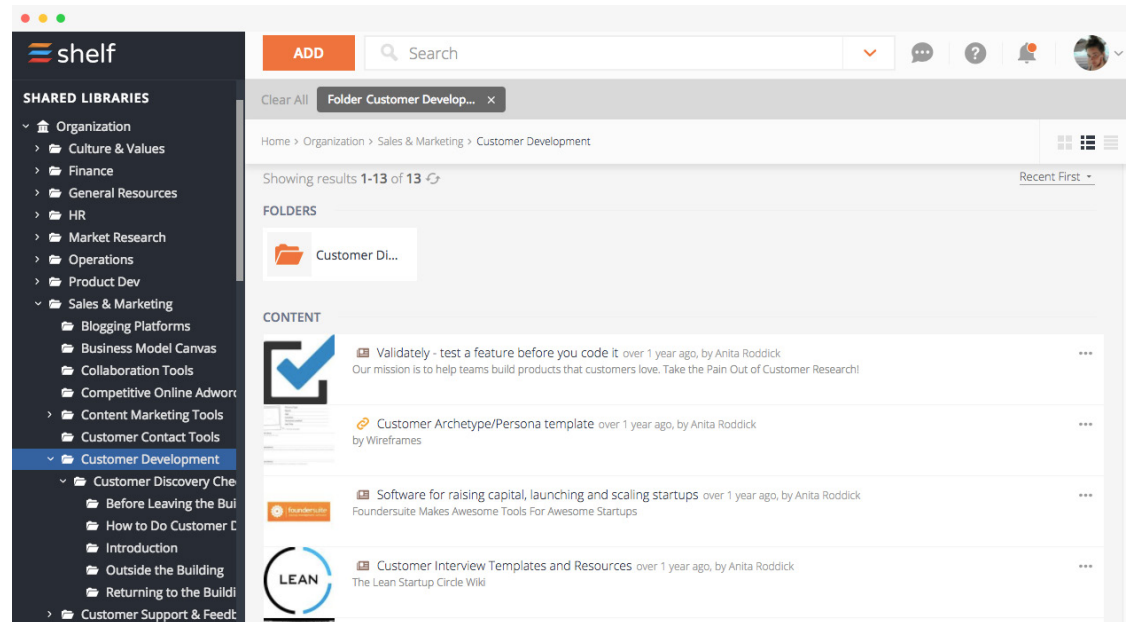
Taxonomies make this easier for you and your team by providing a clear hierarchical structure of your data. They help present a clear and easily navigable map of a company and its resources.

A structured taxonomy ensures that everyone on your team knows where final drafts are expected to be saved for easier retrieval later, where to find last year's marketing budget, and how you should name your documents — but the real power comes from your ability to filter results within the taxonomy.

Section Conclusion

Complementing search with browse and discovery form the foundation of a resilient and scalable content system which will help your company save 1000's of hours a year, prevent work from being unnecessarily redone, and improve your overall efficiency and performance.

Not Google, But What Your Company Needs

While we're accustomed to easily finding what we need with an online search engine, the same can't be said of most internal company searches. We end up manually searching through various files and folders, or calling, texting, and emailing colleagues for help in locating the missing content. This is a waste of time and causes friction amongst employees.

If enterprise search could imitate a world wide web search like Google, it would. But that's not possible, and for all the reasons we discussed here — Google's ability to learn from 3.5 billion searches per day, and it's reliance on 100s of external factors to determine relevant results.

Instead we have to create new and innovative ways to get the same great results and help people find and organize their company's most valuable content.

This white paper was written to demonstrate how search can work in a company. Our platform, Shelf, is much more than a cloud based content management software with a sophisticated set of search capabilities. It is also designed to make organizing easy and intuitive so your content platform can grow and change with your business.

Upgrade your company search and capitalize on the valuable knowledge and content your company has created and collected.

How Do Other Content Platforms Stack Up?

Search is synergistic — all of the components come together to form something much more powerful than they are on their own. Let's review the elements of search, browse and discover and see how different commonly used content platforms stack up:

Search Comparison

Qualities of an Incredible Company Search	Shelf	SharePoint	Box	Drive	Dropbox
Keyword, Best First	✔	✔	✔	✔	✔
Flexible Taxonomy	✔	✔	✔	✔	✔
Advanced Search	✔	✔	✔	✔	✘
In-Document Search	✔	✘	✔	✔	✘
Phrase Search	✔	✘	✔	✔	✘
Image Search	✔	✘	✘	✘	✘
Federated Search	✔	✘	✘	✘	✘
Faceted Navigation	✔	✘	✘	✘	✘
Unified Discovery (Keyword + Faceted Search)	✔	✘	✘	✘	✘
Recommendations	✔	✘	✘	✘	✘
Connections	✔	✘	✘	✘	✘
Tag Library	✔	✘	✘	✘	✘
Badges	✔	✘	✘	✘	✘

According to the McKinsey Global Institute, the average company sees employees spending 19% of their time looking for information or tracking down colleagues who can help with a specific task.

High performing organizations can reduce the average time employees spend on information retrieval tasks by 5.5% - 6.5%. This represents a 30-35% improvement over average performing companies and equates to saving each employee 2.5 hours a week and 125 hours a year.

The improvement can be realized by any organization simply by implementing best practices in enterprise search technology.

Consider this, by increasing efficiency by just 5%, a company of 100 employees could save 15,625 hours a year. In real dollars that is a savings of approximately $625,000 each year. This is make or break money for many organizations.

While it's easy to get lost in the day to day logistics and stress of running a business, it's important to consider just how much time and money you're losing because your employees aren't able to find what they need easily.

Consider just how much money you could save by helping your people do their jobs more efficiently. Book a demo with Shelf and we'll show you how.

Please visit: shelf.io/request-demo
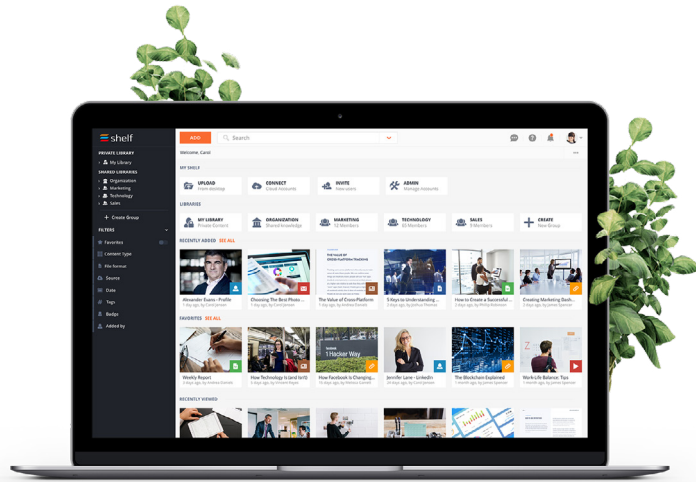
Shelf is the Enterprise Content Platform with the best search in the industry. Quickly find what you need in seconds with an intuitive, easy to use interface.

Trusted by Top Brands

We have made Shelf our entire practice's document repository... amazing customer service gives us confidence that we will be supported using Shelf if/when anything comes up. If you're looking for an easy to use, effective content repository, this is by far the way to go!

Dave Visser Consultant at Slalom Consulting

Use code: **BESTCOMPANYSEARCH** by March 21st and get 50% off your first 3 months.

Contact us for a demo today.

Get Demo