



# **Seven Strategies for Search**

## **Transportation Edition**

A.J. Million, MoDOT Librarian

This presentation does not provide rules. Only strategies...

Sorry!

# The “Deep Web”

## Just what is it?

The Internet is not as open as you think. Only some files are freely available. Others are limited to specialized databases.

Google, Bing and other search engines can only index open web directories. That’s about **20%**. The rest is considered the “Deep Web” and include:

- Material from special databases.
- Files on local area networks.
- Protected files on web servers.
- Data from web applications.



# The “Deep Web”

## What this means...

When you use a search engine only public files come up. Search engines can't communicate with specialized databases or index private directories.

A good strategy is to ***start using a search engine*** since it's quick and easy. Next, if the results don't come up, then check everywhere else.

Images are hard since they don't have characters to match up against.

***Knowing what is public and not*** should be a major factor when making all of your search decisions.



# The “Deep Web”

... for Transportation

**TRIS** (now TRID), **RiP** and **TRB’s** databases are mostly indexed. But, Google’s algorithm limits results to about 50%.

Want for-purchase resources? It’s tough getting access through a search engine. You can still use one to hone in on them though...

Economic and other data generated on-the-fly (like from BLS or the Census Bureau) usually has to be **manually pulled**.



# RSS Feeds

## What are they?

RSS stands for “Really Simple Syndication” feeds. They are links to web addresses and let you set up specialized data streams based on what you want.

They are identifiable by an orange logo with lines coming from a center dot. They also update in real-time. You can use them to avoid checking for multiple updates.

Feeds are composed of a certain markup language (XML) that most computers can read.



# RSS Feeds

## How can you use them?

There are several ways to use RSS feeds, but popular ones include feed readers, web browsers, and operating systems.



# RSS Feeds

## ... for Transportation

**TRIS** (my favorite example), along with most other academic databases offer RSS functionality. Just look for the icon.

Optimally, you can use feeds in instances where news and/or information is regularly updated. Some good examples:

- Legislative efforts at: [thomas.loc.gov](http://thomas.loc.gov)
- Expert blogs like: [transportation.nationaljournal.com](http://transportation.nationaljournal.com)
- News and other current events from the media.



# Special Search Operators

## What are they?

Most search engines operate using a complex set of algorithms. They match keyword characters to identify what users want to find.

In special situations, specific commands make more sense to use than “plain language.” These shortcuts let you manipulate search results.

Operators vary, but there are a few common ones that are especially useful. Their value is mostly selective and not worth using all the time.

Want a laundry list? Visit: [tinyurl.com/custom-commands](https://tinyurl.com/custom-commands)



# Special Search Operators

## ... for Transportation

All of this holds true in almost any environment. Not just specific tools.

**Phrase Searching:** By putting quotes around a phrase [i.e. **“bituminous pavement in Missouri”**] you tell the computer to find that exact set of words. This is good for titles.

**Site Specific:** If you search with a keyword and site query [i.e. **asphalt site:modot.mo.gov**] you eliminate others. Taking this approach is great if you only want results from one website.



# Search Operators... Again

## ... for Transportation

**Exclude Word:** To get rid of an improper keyword add a minus before it [i.e. **traffic merging -late**] to eliminate returns. A good time to use this is when a phrase has dual meanings.

**Wildcards:** To return variants on a keyword just add an asterisks before or after the uniform string [i.e. **road\***]. This tells the computer to look for all forms of the specified word and **not** the entered word alone. This is great when terminology is imprecise.



# Social Media Tools

## What are they?

Twitter, Facebook, and MySpace are online tools that enhance regular communication. They aren't considered to be "search" tools but still offer an alternative to query based systems.

How? Think of social media as a way to digitally ask someone in your network a question. They're also great for following trends.

Some tools that can be useful include:

- *Facebook* • *LinkedIn* • *Twitter* • *Flickr* • *Digg* • *Etc...*



# Social Media Tools

What can some of them do?

**Twitter:** [www.twitter.com](http://www.twitter.com)

Lets you set up feeds, send them to others, etc. As a research tool it emphasizes “finding” over “searching.”

**Facebook:** [www.facebook.com](http://www.facebook.com)

A very general site, Facebook lets you do a little of everything. It’s good for staying in touch and communicating with others.

**LinkedIn:** [www.linkedin.com](http://www.linkedin.com)

Basically, LinkedIn is Facebook but focused on professionals



# Social Media Tools

## ... for Transportation

Just about every State DOT has some sort of social media presence. A friend of mine actually **got a job** using Twitter.

MoDOT is the lead state for an FHWA Pooled Fund. One of the contributors is inaccessible via e-mail but easy to get in touch with through Facebook.

Having contacts never hurts...



# Cloud Computing

## What is it?

How many of you remember the old computer terminals that you used to use? Could computing is similar.

Because storage is cheap and the Internet is so fast there's a movement to run applications and store data online. The Internet is moving that way.

Nobody is sure where Cloud Computing will lead, but online publishing tools help hint at what's to come.

IBM estimates that data online *doubled* from 2007 to 2011.



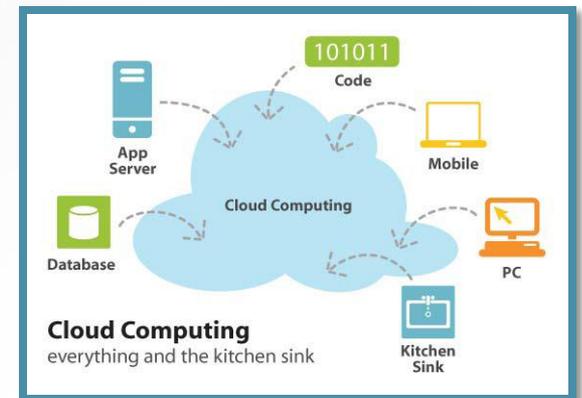
# Cloud Computing

How does the relate to search again?

As more and more “stuff” moves to the Internet, much of it is going to become part of the “Deep Web.”

You can exploit this! If you identify places you routinely visit you’ll have more success than if you search ***everywhere***. Don’t cast a wide net unless you don’t know what you’re looking for...

Oh, and this is a work in progress too.



# Controlled Vocabularies

## What are they?

Controlled vocabularies are pre-defined subject headings used in the place of natural language. They help eliminate redundancies and double-meanings.

Terms vary, but most databases offer special subject headings to users. They cut out problems inherent in language.

A few examples include :

- “**Cats**” and “**Felines**.” Substitute “**Cats**” instead in subject.
- “**Tabby**.” Along with “**Calico**” is a defined a “**cat**” type.



# Controlled Vocabularies

... for Transportation

Two useful examples to keep in mind include:

**TRT: [trt.trb.org](http://trt.trb.org)**

Hundreds of controlled vocabularies exist but most are subject specific. The **Transportation Research Thesaurus** used by TRB is a relevant example.

**LCSH: [authorities.loc.gov](http://authorities.loc.gov)**

Another example is the **Library of Congress Subject Headings**. It isn't great for engineering but can be considered to be a well rounded option.



# Pathfinding and Usability

## What is it?

Pathfinding and site usability refer to design. Well designed websites and databases all *share features that make them easy to navigate*.

These features operate in the same way that road signs do. If you know what a certain feature means – or does – then you'll find your way much easier.

Since this is such a broad topic there's no way to know everything. I'll limit my discussion to TRIS.

Many design elements are universal.



# Pathfinding and Usability

What do I look for?

Breadcrumbs, e-mail features, controlled vocabularies, and check boxes... All of these are used for a reason.



TRIS > Search Results 0 Marked Records: Print | Email | Save | View | Clear

You searched on "missouri" in keywords

Limit Results   [Modify](#) | [New Search](#) | [Search History](#)

Showing 1 - 25 of 1254 records 1 of 51 >>

 **RSS Feed**    

<input type="checkbox"/> Mark: All   None   Show	Record Summary	Published
<input type="checkbox"/>	<b>State of Bridges. Shockingly High Number of Bridges Remain Sub-Standard</b> <i>Better Roads</i> , 2009, pp 8-19 Actions: <a href="#">View Record</a>   <a href="#">View Full Document</a> 	2009/11/1

Showing 1 - 25 of 1254 records 1 of 51 >>

[Home](#) | [Filter search](#) | [Modify search](#) | [Help](#)

# Quick Takeaways

## What should I remember?

1. The “Deep Web” means that Google can’t find everything.
2. RSS feeds can be a useful way to bring information to you.
3. Special commands offer extra flexibility.
4. Social media is a quick way to way to “ask a friend.”
5. Controlled vocabularies standardize language.
6. Cloud computing means check portals.
7. Design is king! Learn the way things are formatted.



**Thank you!**

A.J. Million, MoDOT Librarian  
*anthony.million@modot.mo.gov*  
Tel: 573-751-3075