

PowerSearchingWithGoogle.com

Lesson 3.1: Web organization

[Access Lesson 3.1 slides here](#)

Contents:

Understanding how the web is organized


Deep web vs. Visible web

Using site: operator to filter results

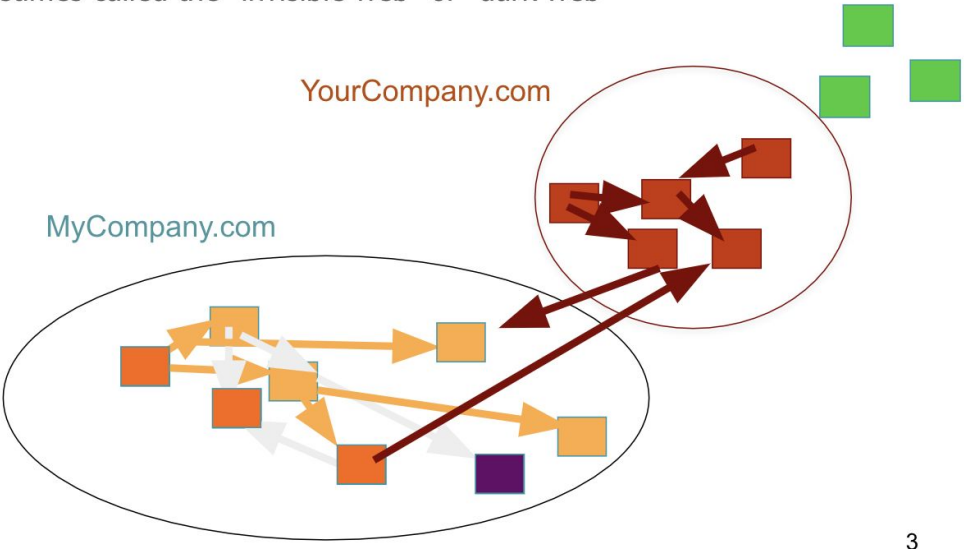
Country codes and top level domains

Welcome to class 3. In this next set of lessons we're going to be talking about advanced techniques for doing Google searches; basically we're going to learn ways to be more precise about what we say, what we ask for and how to take the big set of results and filter it down to just what you really want.

Lesson 3.1 Web organization



- **Visible web:** everything indexed by search engines
- **Deep web:** everything on the web, but not indexed!
 - Sometimes called the “invisible web” or “dark web”

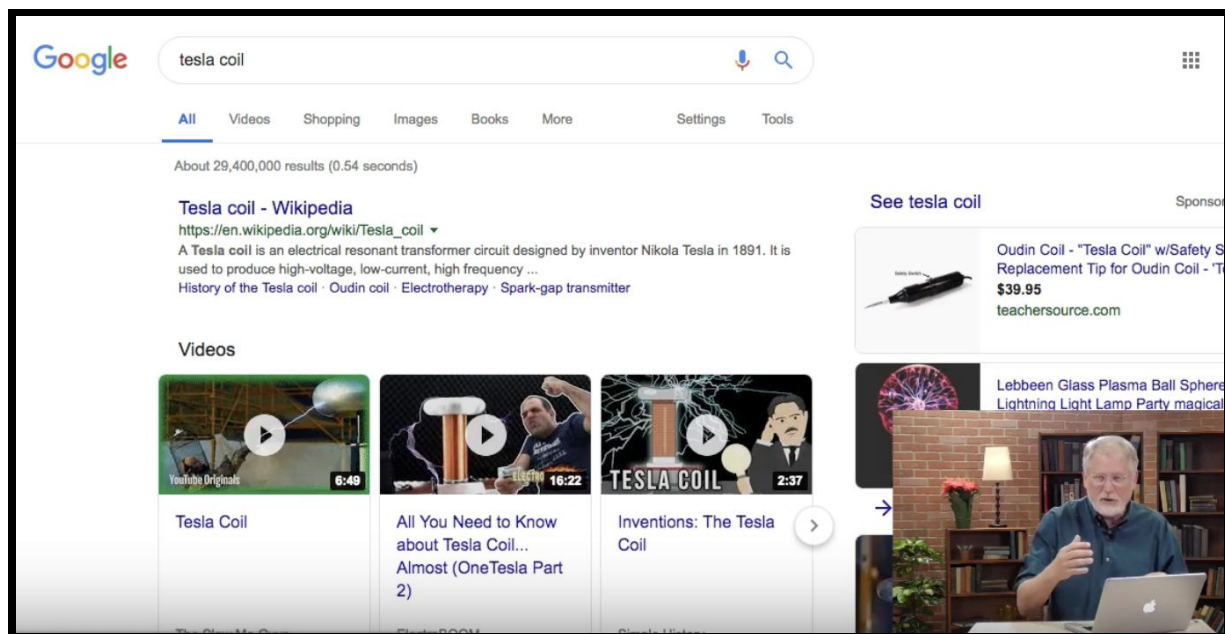


To do this effectively, first we need to learn a little bit about how the web is organized. Basically you can think about each webpage as being a page of content, videos, text, images whatever, and the visible web are all those pages that get indexed by search engines. What the spiders do is they go around to look at each page, then they look at the links, and then they say how does that link connect to another page.

So in this illustration let's look at my company, so what you see here are all those yellow squares each representing one webpage and then there are links that go from one web page to another web page. The spiders crawl those links and then

index the content of that. Links that go outside of the company, mycompany.com to say yourcompany.com are those red arrows that go in and out of those two different circles, those two different ellipses you see there, that's a website, that's an important concept and we'll use that in just a second, but the visible web is the stuff that Google searches over and indexes on each site, my company, your company, whatever.

The Deep Web is stuff that is not here, so for example you see the green boxes in the upper right; those are webpages that are lonely, they're forgotten, there are no links going to them, they're no links coming out of them, we don't see it, Google can't see that stuff, so that's dark web or deep web. When you're searching you often will want to get results just from one website, one of these ellipses will show you how to do that, so the way we get more refinement on our searches is we add an operator. We'll talk a lot about operators in this whole lesson, in this whole class, but we'll start very simply. Suppose you do a search say for [tesla coil].



It would look like this [tesla coil] and guess what? You get lots of tesla coils, those are the big tesla coils that generate lots of lightning bolts, and so on, so that's great, but remember what we just talked about in the site? What you can do is to say [site:] as you see here [site:] and then domain name, like say Stanford University which is stanford.edu, so now when I search for [tesla coil site:stanford.edu] what that does is it searches only inside of that site, only inside the ellipse that's called stanford.edu. Let me modify this query [site:stanford.edu] like that, now when I hit enter what will happen is it will still do tesla coil as a search, but only return results from stanford.edu and all these results, all these images, all these regular web results or articles whatever are all as you can see from Stanford.edu. In this case large.stanford.edu or slac.stanford.edu or news.stanford.edu, you see how this works.

tesla coil site:stanford.edu


Feedback

Wireless Power Transfer - Stanford University
large.stanford.edu/courses/2014/ph240/ho1/ ▼
by b Ho - Related articles
Nov 14, 2014 - 1: Tesla's system of resonant coils for wireless power transfer. The coils are "tuned" to resonate using a capacitor, labeled "condenser".

Contactless Power Transfer - Stanford University
large.stanford.edu/courses/2016/ph240/arshavsky1/ ▼
by A Arshavsky - Related articles
Dec 1, 2017 - [1] The tesla coil in this demonstration produced a very high voltage that oscillated at a high frequency, but with a low current. Electricity was ...

Wireless Charging - Stanford University
large.stanford.edu/courses/2017/ph240/liang1/ ▼
Nov 11, 2017 - In 1891, Tesla built the first device that could transmit electricity without the use of cords with the invention of the Tesla coil. [2] Currently, the two ...

[PDF] **The TESLA Detector - SLAC**
www.slac.stanford.edu/econf/C010630/papers/E3029.PDF ▼
by K Mönig - Related articles
For the superconducting linear collider TESLA a multi purpose detector has been designed. This ... are situated inside a 4T superconducting coil. Sections 2 and ...



There are many, many different kinds of operators and as I said we'll talk about more in just a little bit, but what they all do is filter the results, that is they make the results set smaller in a way that you want, so if the big ellipse here is the entire universe of all results about Tesla coil the smaller enclosed ellipse is stanford.edu results about the topic Tesla coils.

3.1 Operators

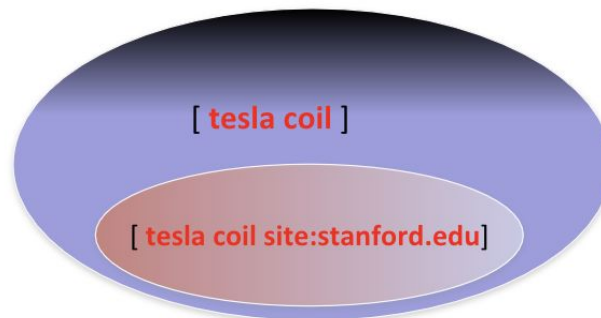


- An **operator** is something extra you add into your query

[**tesla coil**]

[**tesla coil site:stanford.edu**]

- Many different kinds of operators, but they all **filter** results

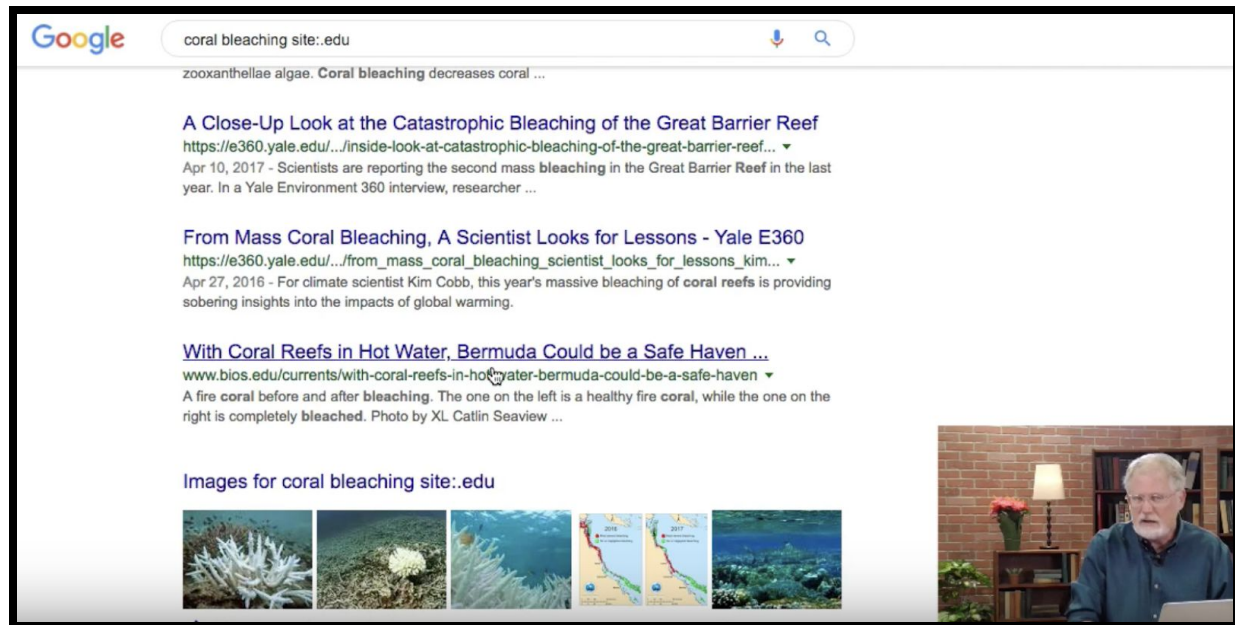


4

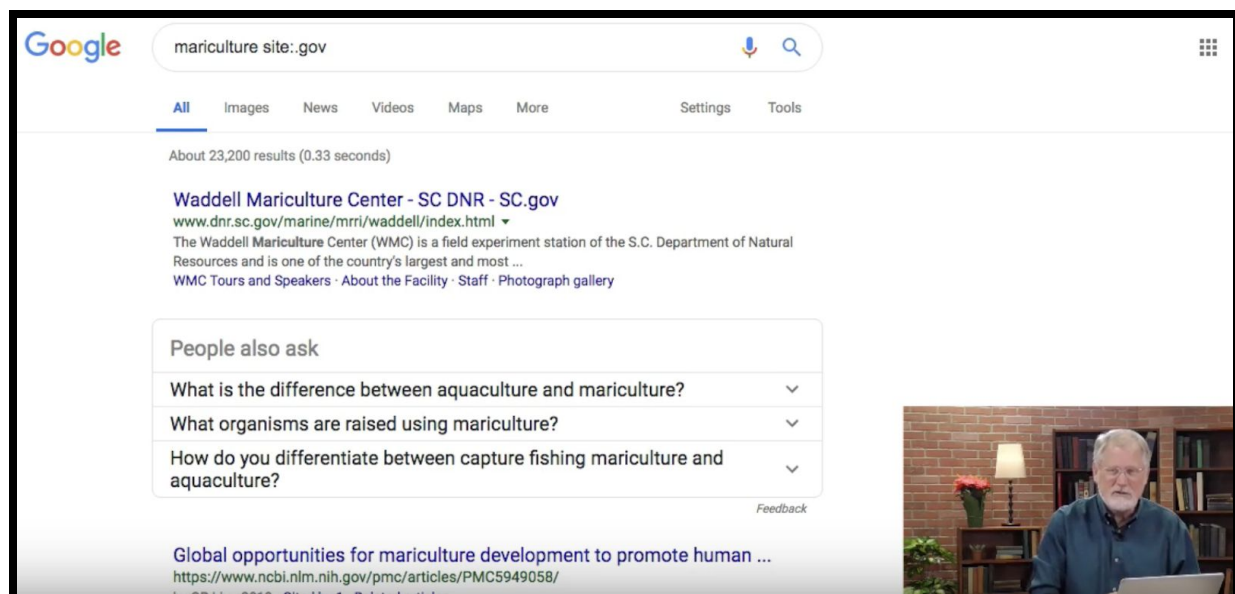
Now there are lots of ways to express what the site is. Country codes like .IN or .ES allow you to search just India so you could say for example site:IN Tesla coil and you get just results from India, same thing with ES which is Spain, BR and so on.

- **Country codes make effective site: restrictions:**
 - Examples: .IN or .ES or .BR (etc.)
- **Be aware of top-level domain (TLD) names:**
 - .EDU -- educational institutions... but ONLY in the US
 - .MIL -- military servers, but again, only in the US
 - .CO.UK -- equivalent to .COM , but with the country domain
 - .GO.KE -- similar to .GOV, but for different countries

Now there are also top-level domains like these TLDs, .com, .edu, .mil and so on. These are also other sites that you can restrict your results to, so for example I could say [site:.edu] and search only education resource sites. Let me give you an example of that. So if we're researching on say the topic of coral bleaching right I can do obviously site:stanford.edu like that, but let's talk about just edu sites, now edu sites are the education institutions so stanford.edu, Berkeley, Florida State, all these different institutions are all edu, so if I now say edu all these results will come just from educational resources.



So you can see here all the edu sites in the US, here you get a great set of results. Now let's take this and switch the topic just a little bit. I'm going to search for mariculture that is growing fish and shellfish and so on in the ocean and I'm going to look at site.gov which is US government resources and we can see there are in fact a bunch of different sites, a bunch of different resources that are owned and run by the government, the US federal government and this is the set of pages all about mariculture from those resources.



If you're trying to get into a business of mariculture, this is a great way to search for that. So now you know how the [site:] restriction operator works. We'll learn more in just a little bit, but right now go ahead and do the activity and explore what site: can do for you. Keep in mind the different kinds of domains you can search for and what that would mean for your searches.

Power Searching with Google © 2019 Google, Inc. CC-BY-SA
Updated 3/1/19 A. Awakuni Fernald