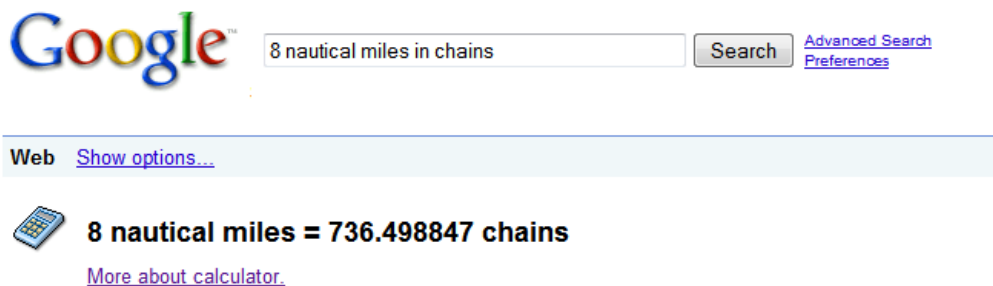


WolframAlpha - How it Works? [Part 1]

As there's been a lot of talk linking Wolfram|Alpha ('WA' from now on) with Google lately, let's start this 'how it works' hypothesis with a simple comparison.

In a sense, you can think of Wolfram Alpha as a kind of 'super' version of Google's Calculator.

If you don't know about Google's Calculator, go to Google's homepage, and enter something like '**8 nautical miles in chains**' - you'll see that the answer is 736.498847 (and if a 'chain' is an unknown to you - well, you've at least two places to look for it now!)



The image shows a screenshot of the Google search interface. The Google logo is on the left. To its right is a search input field containing the text "8 nautical miles in chains". To the right of the input field is a "Search" button. Further right are two links: "Advanced Search" and "Preferences". Below the search bar is a horizontal line. Underneath this line, on the left, is the word "Web" followed by a link "Show options...". Below that is a small calculator icon. To the right of the calculator icon is the text "8 nautical miles = 736.498847 chains". Below this text is a link "More about calculator."

So now, let's enter the same query in WA.

eight nautical miles in chains



Input interpretation:

convert 8 nmi (nautical miles) to Gunter's chains

Result:

736.5 ch (Gunter's chains)

Additional conversions:

48 609 feet

14.82 km (kilometers)

14 816 meters

1.482×10^6 cm (centimeters)

Interpretation:

length

Corresponding quantities:

Light travel time t in vacuum from $t = x/c$:

49 μ s (microseconds)

Sound travel time t at STP from $t = x/v$:

44 seconds

(assuming speed of sound at sea level and 15 °C \approx 340.3 m/s)

Computed by: [Wolfram Mathematica](#)

Download as: [PDF](#) | [Live Mathematica](#)

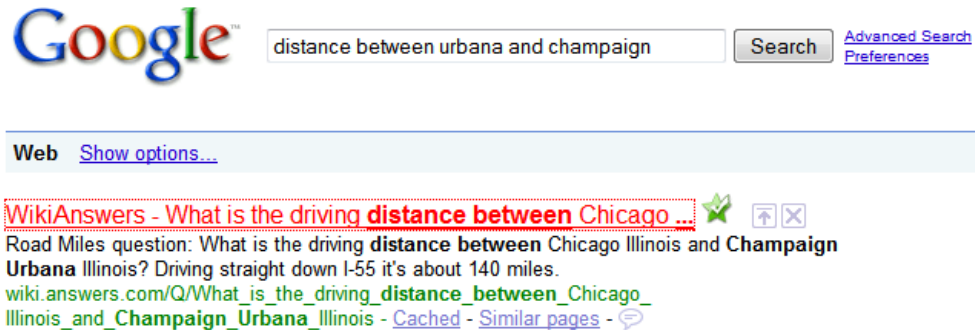
More or less the same answer - you'll notice that WA has 'rounded up', and has assumed that we're not too concerned with the missing inch!

Like Google, WA has carried out some sort of *look-up* to find out what a 'nautical mile' and a 'chain' have in common; and it's discovered that they're units of measure (the probable key word is 'mile' here).

WA then carried out some syntactic, and semantic analysis on the sentence as a whole, and came to the rapid conclusion that we'd like to know one (times eight), converted to the other. Apart from giving us some extra handy data, Google has done more-or-less the same thing of course.

Now, let's take another example.

If we enter; 'distance between urbana and champaign' (proper-case omitted on purpose here – no hints given, and none appreciated) into Google, we find that Google's Calculator can't help.



Google, and as you might expect, gives us a bunch of results drawn from web-text. The first result I got was taken from WikiAnswers where, someone's asked 'What is the driving distance between Chicago Illinois and Champaign Urbana Illinois?' – which is not what we're interested in. Result: there's no immediately obvious and satisfactory answer from a Google search for this query.

Ok, now let's try WA.

distance between urbana and champaign

Assuming Urbana (Illinois, USA) | Use [Urbana \(Ohio, USA\)](#) or [more](#) instead

Input interpretation:

distance from Urbana, Illinois, United States
to Champaign, Illinois, United States

Result:

3.027 miles

Unit conversions:

4.872 km (kilometers)

4872 meters

487176 cm (centimeters)

5328 yards

15983 feet

2.631 nmi (nautical miles)

Direct travel times:

[More](#)

aircraft (550 mph)	20 seconds
sound	15 seconds
light in fiber	25 μ s (microseconds)
light in vacuum	15 μ s (microseconds)

(assuming direct great-circle paths)

Map:



You'll see from the output that WA has assumed that, maybe because they're rather close together, we're asking about suburbs of Illinois (which I was). And, having given that a high probability, with a little semantic analysis, it has assumed that we'd like to know the distance between those two places (which I did); and you can see the result.

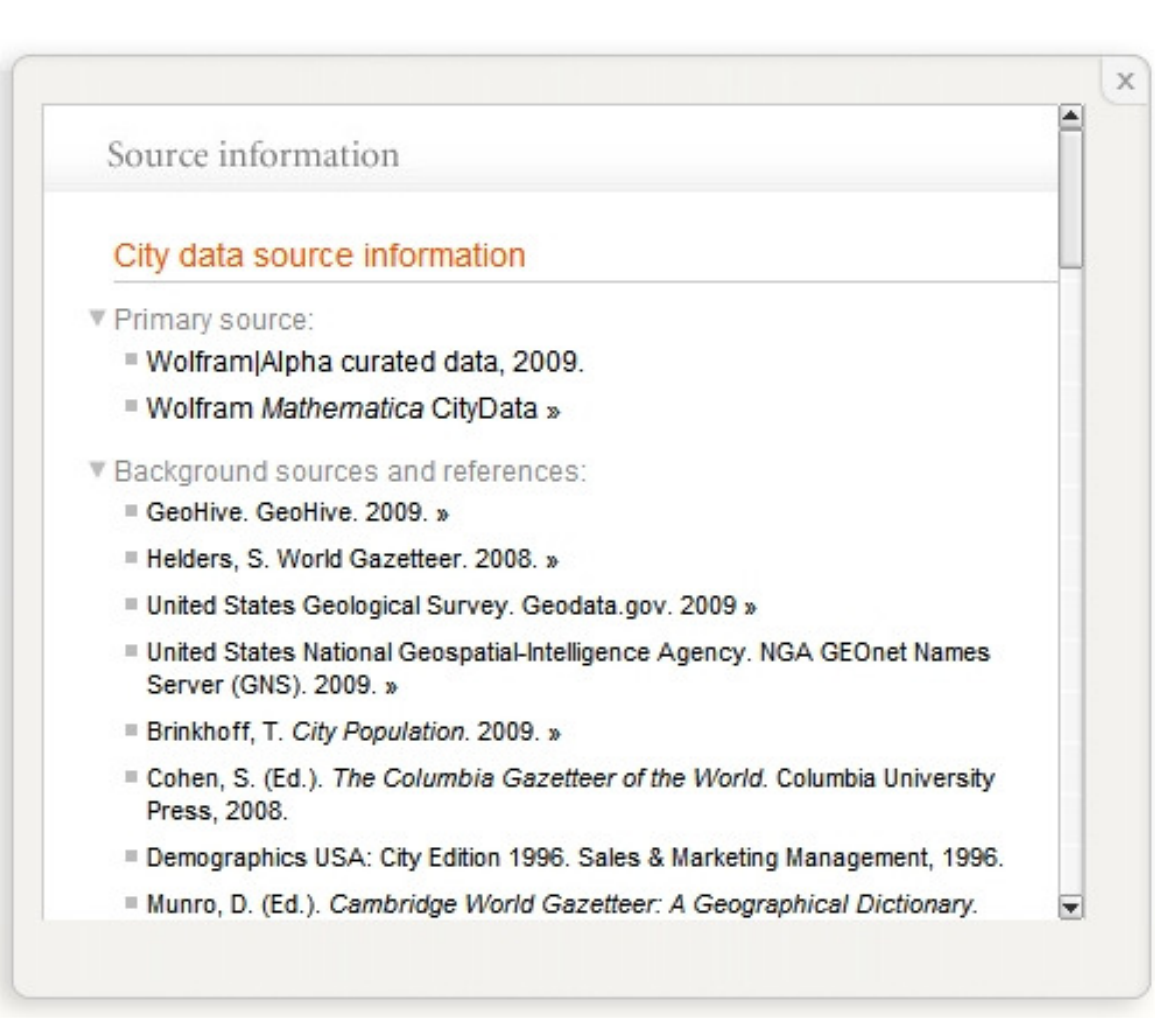
So, how did WA do it?

This is where we need to come back to the vague term I used previously - *look-up*.

Google, as far as we're concerned, gets its data from the public domain, i.e., from what's on the public web.

On the other hand, WA gets its data from, very importantly, the 'Deep Web'. That's data sources that either require a subscription, or at least some sort of *entry point* (an account maybe). In simple terms, some are **free**, and some are '**fee**' (guess which is more robust, structured, and reliable).

Considering our latest query, here is a partial list of data sources that *might* have been employed/analysed to answer our 'distance between urbana and champaign' query.



WA exploits data sources to determine links (relationships is probably a more accurate word) between search-terms, and, more importantly, how the likely important terms (might) relate to one another.

The words used in a query could supply WA (more so in WA than in Google say) with some useful hints – for example the preposition 'between' is certainly more interesting than verb 'driving'; and this is the part of the WA-engine where I suspect Mathematica (the language the whole platform is written in apparently) really earns its corn (the whys and wherefores are outside of the space provided here).

Now to the bottom line: As to how all this works (Google/WA) in detail – well, we don't really know (secrets and all). However, I for one suspect that at WA there's a rather clever ontological-database at work here. Indeed, one that will learn and evolve as more data is fed into it, and given that it has the right rules to link it all up.

So, there you go – not really an in-depth look at WA and what you can do with it, but hopefully something that will at least go some way to showing how WA isn't what we commonly call a 'search engine'.

By the way, I'll tell you why I used those two suburbs (Urbana and Champaign). One is where HAL (you know, 2001 and all that) was activated, whilst the other is where WA did the very same thing.



Dr. Peet Morris studied Software Engineering, Computational Linguistics, and Statistics at the University of Oxford. He is currently a researcher in the Department of Experimental Psychology, and a College Lecturer in Statistics [St. Hilda's College].

<http://www.st-hildas.ox.ac.uk/index.php/fellowsandlecturers/pmorris.html>