

Why Powerset Misses the Point: Relevancy Matters!

On May 14th, Powerset™ launched its long awaited Website demonstrating its “NLP¹ Search” technology. While we are impressed with the user interface and experience they have developed, we believe that Powerset and the media are missing the point. **Relevancy matters.** Powerset’s technology, even though it includes advanced parsing, doesn’t deliver on the promise of providing the user with highly relevant search results. To meet this goal, Search requires *understanding*.

If a user is looking for an enhanced Wikipedia search for people or results that can be found by simple keywords, the Powerset application is good option. Their user tools are helpful, their interface is efficient and clean, and their information aggregation capability is useful. Used solely as a consumer-facing *enhanced Wikipedia search engine for simple searches*, with ties to other data sources, like Freebase™, Powerset provides an improved solution over what is currently available through Wikipedia or other engines that provide search for that dataset. However, if a user is looking to do a deep search with a more complex query or subject matter, Powerset is lacking. Only Cognition has the ability to understand the meaning behind the query and the results, and effectively return more precise and relevant results.

In this paper, we’ll compare Cognition’s Semantic NLP™ and Powerset’s NLP capabilities by examining a variety of complex (real world) queries that go well beyond proper names, proper nouns and keyword-type searches – all of which both Powerset and Wikipedia’s own native search do well. We’ll then provide an analysis on the Wikipedia results returned by both Cognition and Powerset.

It is important to note that Cognition was *not* developed as a Search engine, but rather an enabling technology which brings semantic NLP to other technologies and applications. Search is only one application for Cognition’s Semantic NLP. That being said, Search is an effective way to gauge the effectiveness of a semantic NLP technology’s ability to understand language and meaning.

PROOF POINTS

Objective:

Compare NLP capabilities between Cognition and Powerset in a variety of queries that go beyond proper nouns, names and keyword-type searches. Demonstrate complex real-world queries, then provide an analysis on the Wikipedia results that are returned.

Query #1: “Who won Wimbledon in 1956?”

Results:

Cognition: 20 documents – top 2 are lists of men's and women's champions for 1956

Powerset: 657 documents – no matches specific to the 1956 results

Top 5 Cognition results:

1956_Wimbledon_Championships_-_Men's_Singles:

http://en.wikipedia.org/wiki/1956_Wimbledon_Championships_-_Men's_Singles

1956_Wimbledon_Championships_-_Women's_Singles:

http://en.wikipedia.org/wiki/1956_Wimbledon_Championships_-_Women's_Singles

List_of_male_tennis_players:


http://en.wikipedia.org/wiki/List_of_male_tennis_players

¹ Natural Language Processing

Angela_Buxton: http://en.wikipedia.org/wiki/Angela_Buxton

Shirley_Fry: http://en.wikipedia.org/wiki/Shirley_Fry

The screenshot shows the Wikipedia Cognition search interface. At the top, it says "WIKIPEDIA.COGNITION" and "GIVING TECHNOLOGIES NEW MEANING™". There is a search bar with the text "Who won Wimbledon in 1956?" and a "SEARCH" button. Below the search bar, it says "1956 Wimbledon win: 20 files —". The search results are listed as follows:

- [1956 Wimbledon Championships - Men's Singles](#)
Ken Rosewall 6-2 4-6 7-5 6-4 in the final to win the Gentlemen's Singles title at the 1956 Wimbledon Championships . See also:
http://en.wikipedia.org/wiki/1956_Wimbledon_Championships_-_Men's_Singles
- [1956 Wimbledon Championships - Women's Singles](#)
Angela Buxton 6-3 6-1 in the final to win the Ladies' Singles title at the 1956 Wimbledon Championships . See also:
http://en.wikipedia.org/wiki/1956_Wimbledon_Championships_-_Women's_Singles
- [List of male tennis players](#)
1956 -) - ( Sweden) - winner of 11 Grand Slam singles titles • 1974/1975/1978/1979/1980/1981 French Open champion , 1976 quarter-finalist [...]
http://en.wikipedia.org/wiki/List_of_male_tennis_players
- [Angela Buxton](#)
French Championships and Wimbledon in 1956 with Althea Gibson . Buxton was the first
http://en.wikipedia.org/wiki/Angela_Buxton
- [Shirley Fry](#)
Althea Gibson in a quarterfinal of Wimbledon in 1956 and in the finals of the 1956 U.S. Championships and 1957 Australian Championships . Irvin was inducted [...]
http://en.wikipedia.org/wiki/Shirley_Fry
- [Ann Haydon-Jones](#)
, she won the Wimbledon girls' singles championship . Haydon-Jones played lawn tennis in a highly competitive era that incl
http://en.wikipedia.org/wiki/Ann_Haydon-Jones
- [List of Grand Slam Men's Singles champions](#)
also known as the "Big Four" of tennis

On the right side of the interface, there are buttons for "ADVANCED", "SAMPLES", "HELP", and "FEEDBACK". Below these buttons, there is a section for "The following word meanings were selected. Use dropdown menus to change the meanings." with dropdown menus for "win:" (selected: "1) n, v, tr: be victorious: he won th") and "Wimbledon:" (selected: "1) n: of southern Greater London, "). There is also a "SUBMIT" button at the bottom right.

Top 5 Powerset results:

World number one male tennis player rankings (refers to Connors, Becker in 1982 & 1989, not 1956)

Ken Rosewall (Rosewall won the Australian Davis cup in 1956, not Wimbledon)

Björn Borg (Borg won Wimbledon in 1976, not 1956)

Rod Laver (Context: ... giving Laver the "clay court triple" of Paris, Rome, and Hamburg that had been achieved previously only by Lew Hoad in 1956: no mention of Wimbledon.)

Angela Buxton (Only IDENTICAL HIT IN TOP 5)

Wikipedia Articles

Who won Wimbledon in 1956? SE

Freebase → Factz Articles

Wikipedia Articles: results 1 - 10 of 675

- [World number one male tennis player rankings](#) In 1982 and in 1989 respectively Connors and Becker both **winner**s of Wimbledon and the US Open were considered as World Champions even though the ATP ranked respectively McEnroe and Lendl as number 1. ... John McEnroe P. (USA) | (USA) | Björn Borg P. (Swe.) | (Swe.) | Tennis Magazine (France); (France); ITF; ; ATP Awards; ; McEnroe **won Wimbledon** and the U.S. Open...
- [Ken Rosewall](#) But considering that Wimbledon and the U.S. Open were the two big events of 1970 **Newcombe (Wimbledon winner)** and Rosewall (Forest Hills winner) remain to chose the number one player in the world. ... Rosewall was a member of the victorious Australian Davis Cup teams in 1953, 1955, 1956 and 1973, in all cases defeating USA in the final.
- [Björn Borg](#) Borg became the youngest male Wimbledon champion of the modern era at 20 years and 1 month (a record subsequently broken by Boris Becker, who **won Wimbledon** aged 17 in 1985). ... **Borg** is the first player to have **won two different Grand Slam singles tournaments (6 French Open and 5 Wimbledon)** at least five times.
- [Rod Laver](#) Among those titles were the Italian Championships and the German Championships, giving Laver the "clay court triple" of Paris, Rome, and Hamburg that had been achieved previously only by Lew Hoad in 1956. ... In 1967, Laver **won 18 titles**, including the **Wimbledon Pro**, the U.S. Pro Championships, the Wembley Pro Championship, and the French Pro Championship, which gave him a clean sweep of the most important professional titles.
- [Angela Buxton](#) Buxton played in Wightman Cup competition for Great Britain in 1954, 1955, and 1956. ... Buxton said that because of anti-Semitism, which she had to endure already during her childhood, even after **her Wimbledon victory** with Althea Gibson she was not invited to join the

Query #2: "Who won an Emmy for choreography in 1987?" [Answer: Michael Peters]

Results:

Cognition: 1 document – a perfect match: List of winners of the Emmy Award for Choreography
 Powerset: 259 documents – the perfect match is result #7

Cognition result:

List_of_winners_of_the_Emmy_Award_for_Choreography:
http://en.wikipedia.org/wiki/List_of_winners_of_the_Emmy_Award_for_Choreography

WIKIPEDIA.COGNITION GIVING TECHNOLOGIES NEW MEANING™ RETURN TO COGNITION

NEWSLETTER SIGN UP: WHO WE ARE | WHAT WE DO | CONTACT

Searching: English Wikipedia. Details

For Cognition to better understand you, enter a phrase with proper capitalization.

1987 Emmy choreography win: 1 file —

[List_of_winners_of_the_Emmy_Award_for_Choreography](#)
 Marguerite Derricks : 4 Nominations (1997-2000), 3 Wins (1997-1999)
http://en.wikipedia.org/wiki/List_of_winners_of_the_Emmy_Award_for_Choreography

1

ADVANCED SAMPLES

HELP FEEDBACK

The following word meanings were selected.
Use dropdown menus to change the meanings.

win:
1) n, v; fr: be victorious: he won th ▾

Emmy:
1) n: statuette awarded annually b ▾

choreography:
1) n: the art of arranging dance: a ▾

1987:
one definition.

Top 5 Powerset Results:

Bob Fosse -- He won a Tony (not Emmy) for choreography and didn't win in 1987, but died that year

58th Primetime Emmy Awards

Article on 58th Emmy awards in 2006, with a link to a stub only, for 1987: 39th Primetime Emmy Awards

Fred Astaire

He won Emmy's earlier, but never for choreography, and he died in 1987

List of persons who have won Academy, Emmy, Grammy, and Tony Awards list by people, not by year and no mention of Michael Peters

The Temptations:

Article about the group, The Temptations, and an Emmy-award winning documentary on them -- in 1987

Wikipedia Articles

Who won an Emmy for choreography in 1987?

Wikipedia Articles: results 1 - 10 of 259

- Bob Fosse** Fosse was in Washington D.C. On September 23, 1987, for a revival of his musical Sweet Charity which was opening at the National Theater. ... Bill Henry's 1990 documentary of Fosse's work (Dance In America: Bob Fosse Steam Heat), produced for an episode of the PBS programme Dance in America: Great Performances, **won an Emmy** that year.
- 58th Primetime Emmy Awards** The former tribute was presented by Simon Cowell, American Idol judge, with a performance by Barry Manilow, who **won an Emmy** later that evening. ... 1974 · 1975 · 1976 · 1977 · 1978 · 1979 · 1980 · 1981 · 1982 · 1983 · 1984 · 1985 · 1986 · **1987** · 1988 · 1989 · 1990 · 1991 · 1992 · 1993 · 1994 · 1995 · 1996 · 1997 · 1998 · 1999 · 2000 · 2001 · 2002 · 2003 · 2004 · 2005 · 2006 · 2007 · 2008
- Fred Astaire** That partnership, and the **choreography** of Astaire and Hermes Pan, helped make dancing an important element of the Hollywood film musical. ... Astaire **won an Emmy Award** for his performance.
- List of persons who have won Academy, Emmy, Grammy, and Tony Awards** International **Emmy** Award ... Bob Fosse - Fosse **won** all three awards in the same year, 1973.
- The Temptations** The miniseries was a ratings success and **won an Emmy** award for Best Direction; it was subsequently rerun on the VH-1 cable television network and released to VHS and DVD. ... 2007: The Temptations - Live In London (**1987**)
- Audrey Hepburn** | **1987** | Love Among Thieves | Baroness Caroline DuLac | Television movie. | ... Audrey Hepburn was one of the few people who have **won an Emmy**, a Grammy, an Oscar, and a Tony Award

Query #3: "politicians who went to UCLA"

Results:

Cognition: 175 documents – 3 of the top 5 matches are for specific politicians who went to UCLA, with the other two focusing on UCLA and political activities

Powerset: 1,200 documents – matches tend to focus on a variety of UCLA listings, with 1 of the top 5 matches being for a politician who attended UCLA

Top 5 Cognition results:

George Allen

[http://en.wikipedia.org/wiki/George_Allen_\(U.S._politician\)](http://en.wikipedia.org/wiki/George_Allen_(U.S._politician))

[Allen has a long history of interest in the Confederate flag, in spite of his never having lived in the South until his transfer from UCLA to the University of Virginia as a sophomore in college...]

Tom Bradley

[http://en.wikipedia.org/wiki/Tom_Bradley_\(politician\)](http://en.wikipedia.org/wiki/Tom_Bradley_(politician))

[One of the jobs that he had while at UCLA was as a photographer for comedian Jimmy Durante. He dropped out of UCLA during his junior year to attend the Los Angeles Police Academy, after placing near the top on a recruitment exam...]

Ted Brown

[http://en.wikipedia.org/wiki/Ted_Brown_\(politician\)](http://en.wikipedia.org/wiki/Ted_Brown_(politician))

[He graduated from UCLA in 1984, with a B.A. in political science with a concentration in Constitutional Law...]

Mike Davis

[http://en.wikipedia.org/wiki/Mike_Davis_\(politician\)](http://en.wikipedia.org/wiki/Mike_Davis_(politician))

[Davis founded The Images of Blacks in America Symposium at UCLA, a conference for high school students and professionals...]

Bill Morrow

[http://en.wikipedia.org/wiki/Bill_Morrow_\(California_politician\)](http://en.wikipedia.org/wiki/Bill_Morrow_(California_politician))

[He transferred to UCLA, where he graduated with honors in 1976...]

The screenshot shows the Wikipedia.Cognition search interface. At the top, it says "WIKIPEDIA.COGNITION GIVING TECHNOLOGIES NEW MEANING™". Below that is a search bar with the text "politicians who went to UCLA" and a "SEARCH" button. To the right of the search bar are links for "WHO WE ARE", "WHAT WE DO", and "CONTACT". Below the search bar, there are several search results, each with a title and a brief description. The results are: "George_Allen_(U.S._politician)", "Tom_Bradley_(politician)", "Ted_Brown_(politician)", "Mike_Davis_(politician)", "Bill_Morrow_(California_politician)", "History_of_the_University_of_California_Los_Angeles", and "Bruin_Alumni_Association". On the right side of the search results, there are buttons for "ADVANCED", "SAMPLES", "HELP", and "FEEDBACK". Below these buttons, there is a section for "The following word meanings were selected. Use dropdown menus to change the meanings." with a dropdown menu for "politician" and a dropdown menu for "UCLA". At the bottom right of the search results, there is a "SUBMIT" button.

Top 5 Powerset results:

UCLA-USC Rivalry

The UCLA-USC rivalry is the college rivalry between two universities located in Los Angeles, California: the University of California, Los Angeles and the University of Southern California.

1967 USC vs. UCLA Football Game

The 1967 USC vs. UCLA football game was an American College Football game played in the 1967 College Football Season on November 18, [1967](#).

Roberta Achtenberg

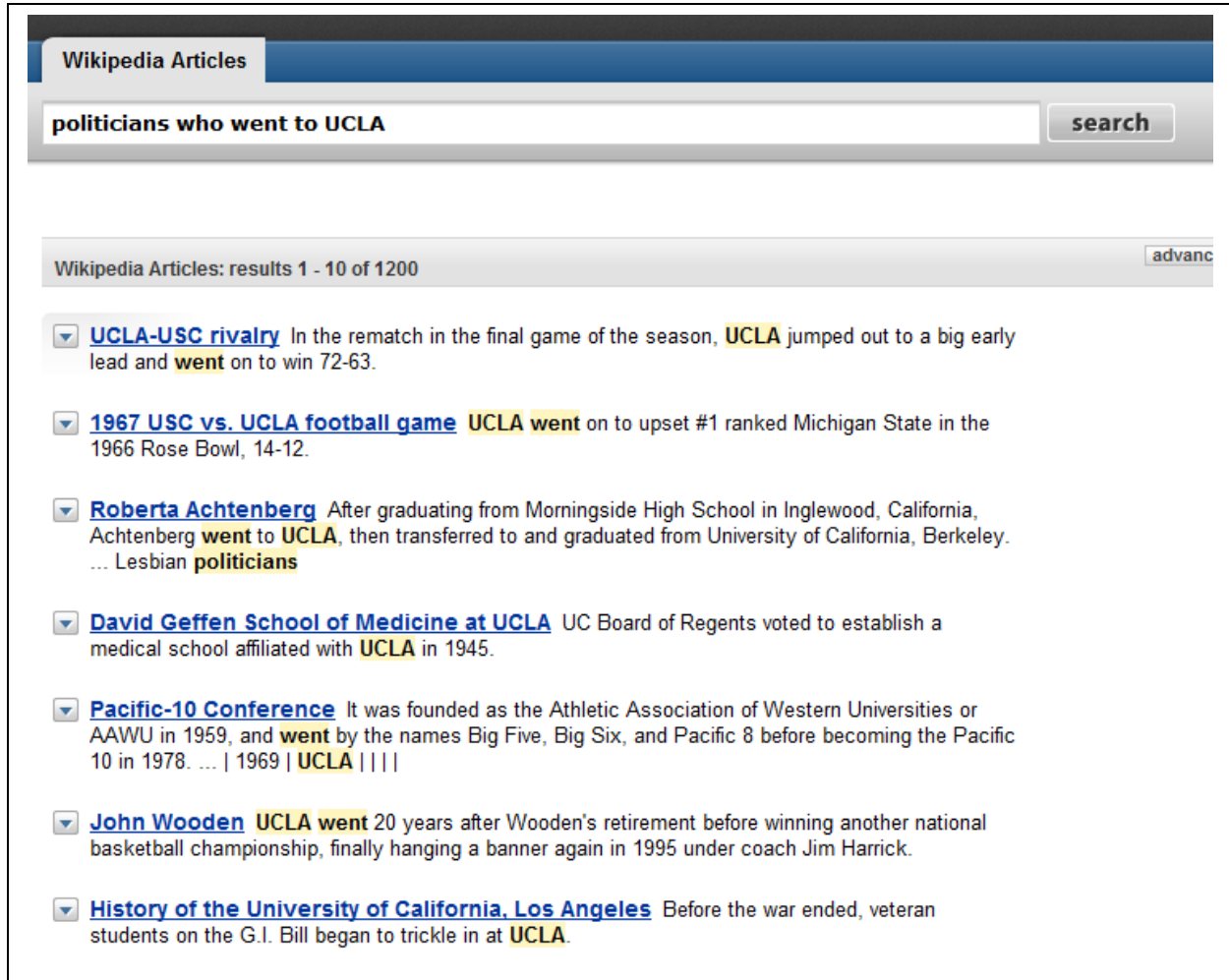
After graduating from Morningside High School in Inglewood, California, Achtenberg went to UCLA, then transferred to and graduated from University of California, Berkeley.

David Geffen School of Medicine at UCLA

UCLA School of Medicine or David Geffen School of Medicine at UCLA is an accredited allopathic medical school located in Los Angeles, California, United States.

Pacific 10 Conference

The Pacific-10 Conference (Pac-10) is a college athletic conference which operates in the western United States



The screenshot shows a search interface for Wikipedia articles. The search query is "politicians who went to UCLA". The results page shows "Wikipedia Articles: results 1 - 10 of 1200". The first seven results are listed below:

- [UCLA-USC rivalry](#) In the rematch in the final game of the season, **UCLA** jumped out to a big early lead and **went** on to win 72-63.
- [1967 USC vs. UCLA football game](#) **UCLA went** on to upset #1 ranked Michigan State in the 1966 Rose Bowl, 14-12.
- [Roberta Achtenberg](#) After graduating from Morningside High School in Inglewood, California, Achtenberg **went** to **UCLA**, then transferred to and graduated from University of California, Berkeley. ... Lesbian **politicians**
- [David Geffen School of Medicine at UCLA](#) UC Board of Regents voted to establish a medical school affiliated with **UCLA** in 1945.
- [Pacific-10 Conference](#) It was founded as the Athletic Association of Western Universities or AAWU in 1959, and **went** by the names Big Five, Big Six, and Pacific 8 before becoming the Pacific 10 in 1978. ... | 1969 | **UCLA** | | |
- [John Wooden](#) **UCLA went** 20 years after Wooden's retirement before winning another national basketball championship, finally hanging a banner again in 1995 under coach Jim Harrick.
- [History of the University of California, Los Angeles](#) Before the war ended, veteran students on the G.I. Bill began to trickle in at **UCLA**.

Query #4: "Texan wildflowers"

Results:

Cognition: 74 documents – Matches focus mainly on wildflowers in Texas

Powerset: 44 documents – Matches tend to focus only on the word "Texan," showing an inability to understand Texan and wildflowers in context.

Top 5 Cognition results:

Lady_Bird_Johnson_Wildflower_Center:

http://en.wikipedia.org/wiki/Lady_Bird_Johnson_Wildflower_Center

Bluebonnet: <http://en.wikipedia.org/wiki/Bluebonnet>

The_Daily_Texan: title="Lady Bird Johnson Wildflower Center">
http://en.wikipedia.org/wiki/The_Daily_Texan

Wild_Flower_Festival: (Richardson, Texas) http://en.wikipedia.org/wiki/Wild_Flower_Festival

Lady_Bird_Johnson: http://en.wikipedia.org/wiki/Lady_Bird_Johnson

The screenshot shows the Wikipedia.Cognition search interface. At the top, it says "WIKIPEDIA.COGNITION GIVING TECHNOLOGIES NEW MEANING™" and "RETURN TO COGNITION". Below that is a "NEWSLETTER SIGN UP:" field with a "GO" button. To the right are links for "WHO WE ARE | WHAT WE DO | CONTACT". The search bar contains "Texan wildflowers" and a "SEARCH" button. Below the search bar, it says "Searching: English Wikipedia. Details" and "For Cognition to better understand you, enter a phrase with proper capitalization." The search results are listed under "Texan wild-flower: 74 files —". The results include: "Lady_Bird_Johnson_Wildflower_Center" (botanical garden on La Crosse Avenue in Austin, Texas), "Bluebonnet" (title="Lupinus texensis">Lupinus texensis (Texas bluebonnet) emerged as the favorite of most Texans. The flowers' deep blue), "The_Daily_Texan" (title="Lady Bird Johnson Wildflower Center">Lady Bird Johnson Wildflower Center), "Wild_Flower_Festival" (Richardson, Texas which typically has a turn out of approximately 50,000 people. The Wildflower! Arts), "Lady_Bird_Johnson" (title=">[10] She developed her lifelong love of the environment as a child growing up in the tall pines and bayous of East Texas and watching the wildflowers [...]), "Climate_of_Dallas_Texas" (wildflowers (such as the bluebonnet , Indian paintbrush and other flora) which bloom in spring and are planted around the highways throughout Texas. [2] [...]), and "Dallas_Texas" (title="Wildflower">wildflowers (such as the bluebonnet , Indian paintbrush and other flora) bloom in spring and are planted around the highways throughout [...]). On the right side, there are buttons for "ADVANCED", "SAMPLES", "HELP", and "FEEDBACK". Below these buttons, it says "The following word meanings were selected. Use dropdown menus to change the meanings." and shows two dropdown menus: "Texan:" with "1) adj, n: a resident or native of Te" and "wild-flower:" with "1) n: a flowering plant that grows". A "SUBMIT" button is at the bottom right.

Top 5 Powerset results:

T-6 Texan

The T-6 Texan was a single-engine advanced trainer aircraft designed by North American Aviation, used to train fighter pilots of the United States Army Air Forces, United States Navy, Royal Air Force and other air forces of the British Commonwealth during World War II.

T-6 Texan II

The T-6A Texan II is a single-engine turboprop aircraft built by the Raytheon Aircraft.

Amarillo Texas Big Texan Steak

The Big Texan Steak Ranch is a steakhouse restaurant and motel located in Amarillo, Texas, United States which opened off of Route 66 in 1960 and moved to its present location on Interstate 40 in 1970.

The Daily Texan

general article on the U of Texas newspaper, not a specific article on flowers:


Highlighted section: In 1913, the student body voted to publish the paper each weekday, and the Daily Texan was born on September 14, 1913.

Lady_Bird_Johnson: http://en.wikipedia.org/wiki/Lady_Bird_Johnson

Wikipedia Articles

Texan wildflowers

Wikipedia Articles: results 1 - 10 of 44

- [T-6 Texan](#) The NA-88 design resulted in 2,970 AT-6C **Texans** and 2,400 as the SNJ-4. ... T-6 **Texan II**
- [T-6 Texan II](#) The T-6B variant of the **Texan II** was introduced in 2005. ... ↑ Global Security T-6 **Texan**
-  [The Big Texan Steak Ranch](#) The front view of the Big **Texan** Steak Ranch. ... ↑ OutdoorFlix.com - Big **Texan** Steak Ranch
- [The Daily Texan](#) In 1913, the student body voted to publish the paper each weekday, and the Daily **Texan** was born on September 14, 1913.
- [Lady Bird Johnson](#) This earned her the nickname of "Johnny Appleseed" of **Wildflowers**. ... "A former first lady leaves us her legacy", The Daily **Texan**, 2007-07-12.
- [2007 Texas Longhorn football team](#) Brown said that President Johnson used to enjoy discussing football with former UT coach Darrell K. Royal, while the first lady was known for her conservation efforts, particularly the spread of native **wildflowers**. ... "Law seeks out Texas football, again", The Daily **Texan**, 3 August 2007.
- [The Texan](#) The **Texan** (TV series), starring Rory Calhoun

Here is a comparison grid between the features and capabilities of Cognition's and Powerset's technologies:

Feature	Powerset	Cognition
Tokenization	x	x
Morphology	x	x
Phrases		x
Sense selection	x	x
Semantic map		x
Parsing	x	x
Integration with structured data		x
Enterprise version		x

Tokenization: Tokenization turns ASCII text, punctuation and spaces into words and sentences.

Morphology: Morphology recovers base word stems from inflected form (e.g., "baby" from "babies") and base stems from derived form (e.g., "digitate" from "interdigitate"). Powerset does not have the latter, only the inflectional morphology.

Phrases: Cognition recognizes 200,000 phrases such as “Bill of Rights” and “bok choy.”

Sense selection: Cognition handles the ambiguity of 17,000 ambiguous words of English. It works with meanings rather than strings. In "the union went on strike", the word “strike” has a different meaning than in the phrase "strike up a conversation."

Semantic map: Cognition maps each phrase and word meaning to its hypernyms (ancestors in an ontology), its synonyms, and its sense contexts (senses of other words that trigger this sense of this word).

Parsing: Parsing recovers the argument structure of a sentence. For example, in "The cat caught the mouse", "the cat" is the subject, "caught the mouse" is the predicate, "caught" is the verb and "the mouse" is the object.

Integration with structured data: Cognition can search in structured and unstructured data as exemplified on MEDLINE.cognition.com, where in the advanced search function it searches on structured information about dates, journals and authors, as well as the unstructured information within the MEDLINE abstracts.

Enterprise version: The linguistic meaning-based Search Engine capabilities of Cognition’s Semantic NLP enables an enterprise to dig deeper into their internal datasets with greater completeness and precision.

CONCLUSION:

Relevancy requires understanding. High precision and recall requires understanding. A more complete and meaningful user experience requires understanding.

One of the biggest barriers to building a natural language understanding system is to build the semantic map and the dictionary with details of the syntactic behavior of words (i.e. how words behave within context). Cognition’s Semantic NLP brings understanding to technologies and applications. We’ve spent more than 20 years building this capability into Cognition’s Semantic NLP for the English language.

Cognition Semantic NLP uniquely combines statistical algorithms with linguistically mapped coverage of the English language, employing more than 4 million semantic contexts (word meanings in context); 506,000 word stems (the base forms of words); 536,00 word senses (word and phrase meanings); 17,000 ambiguous word definitions; 200,000 phrases; and a robust ontology and taxonomy.

Powerset has done an excellent job re-thinking the user experience using Natural Language Processing technology with an advanced parser. However, in order to improve the actual results of a Search, there’s no getting away from the need for a system which understands word and phrase meanings within context. Cognition is the only complete semantic NLP technology which can deliver this understanding.

We encourage you to experience the impact of higher relevance by visiting Cognition's Semantic Wikipedia demonstration. It is the full Wikipedia dataset indexed using Cognition's Semantic NLP technology.

<http://WIKIPEDIA.cognition.com>

For more information regarding Cognition and its unique and powerful semantic technology, visit our Website at www.cognition.com. In addition, feel free to drop us an email at info@cognition.com.

COGNITION
GIVING TECHNOLOGIES NEW MEANING™