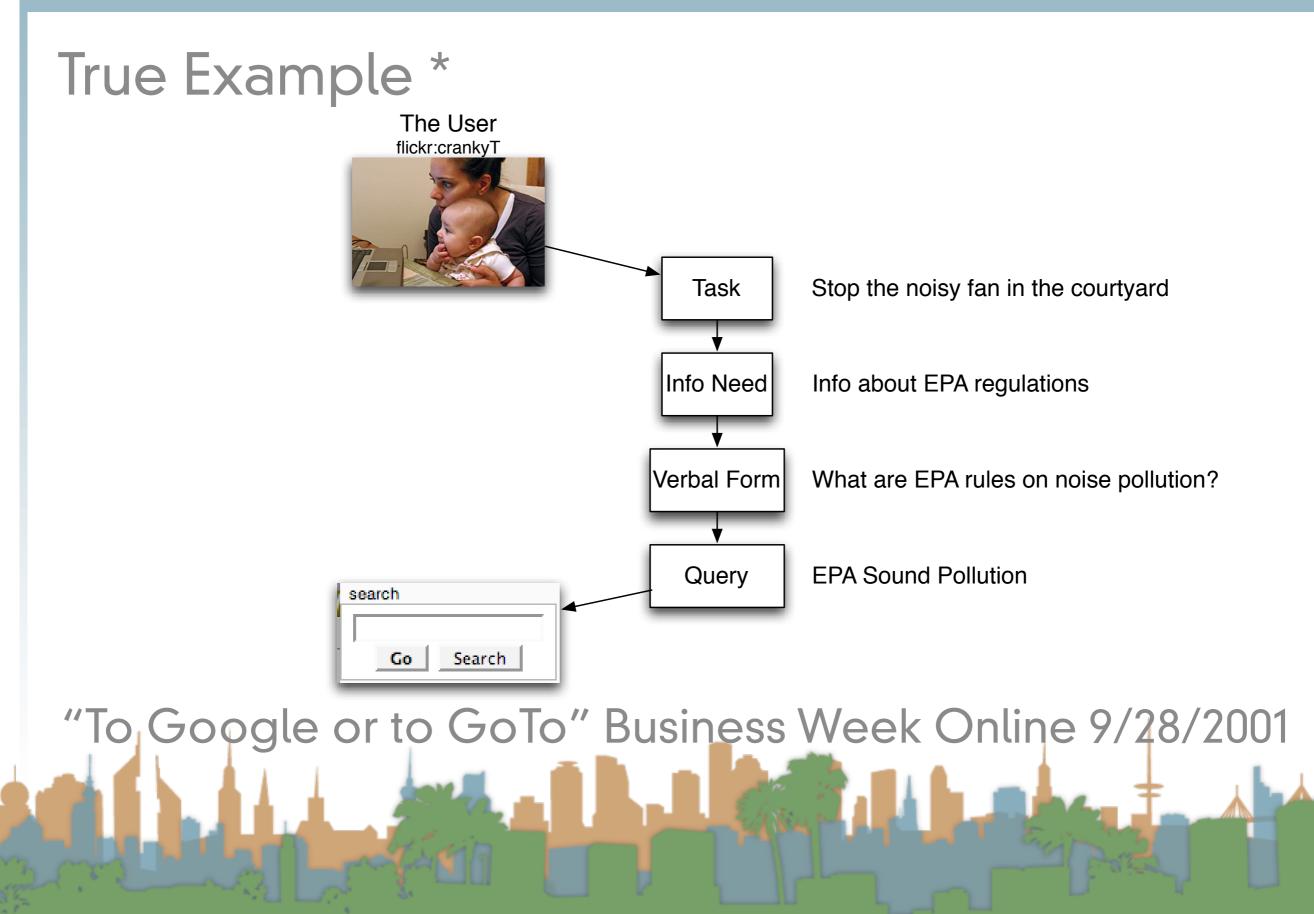
### Overview

## Overview

- Introduction
- Classic Information Retrieval
- Web IR
- Sponsored Search
- Web Search Basics
  - Size of the Web
- Web Users
- Spam
- Helping the User



# Answering "the need behind the query"

- The query is often an imprecise indicator of what the user really wants
- What can we do to get a better handle on the underlying information need?
- Query language
  - Adjust rank of English results for a Japanese query
- Use user context
  - In particular geographic context

# Answering "the need behind the query"

- Guess what type of information the user wants
  - a web page?
  - a map?
  - a stock price?
  - what else?
- Correct queries
  - Suggest correct spellings
  - Suggest related searches (google-fu)



# Examples - language

| 00  | kitakurihama - Google Search  |            |
|---|---|------------|
|   | 🛶 🗸 🥝 🏠 🖸 http://www.google.com/search?hl=en&cliv 🔻 🕨 💽 * keitei 🔍  | N.<br>Siri |
| Last o  | oen ▼ Class Development ▼ Research Links ▼ UCI ▼ Thesis pages ▼ DR ▼ Conferences ▼ Place Lab ▼  | ⁄≫         |
| Web   | Images Maps News Shopping Gmail more V  | -          |
|   | donald.j.patterson.iii@gmail.com   Web History   My Account   Sign out  | Ł          |
| G   | oogle kitakurihama Search Advanced Search Preferences   |            |
| Web   | Results 1 - 10 of about 276 for kitakurihama. (0.12 seconds)  |            |
| MySpa<br>informa<br>profile.<br><u>Cache</u><br>Kitaku<br>(ヘアー<br>p.bari3<br>ct1=&c<br><u>Cache</u> | ace.com - Nasty - Kitakurihama, Kanagawa - Hip Hop - www<br>ce music profile for Nasty with tour dates, songs, videos, pictures, blogs, band<br>ation, downloads and more.<br>myspace.com/index.cfm?fuseaction=user.viewprofile&friendid=1000891776 - 48k -<br>1 - Similar pages - Note this<br>urihama - [Translate this page]<br>rihama. 1 to 7 out of 7 Glee land (グリーランド) Karaoke box HAIR CLINIC TAIRA<br>·クリニック タイラ) Hairdressing and cosmetics TAKOCHU (たこ忠)<br>.net//lookup2.asp?agntcd=1802&<br>tt1_nm=&ct2=007&ct2_nm=Kitakurihama&lv=1 - 5k -<br>1 - Similar pages - Note this |            |
| 混合自   | <u>所ちびっこランド 北久里浜園</u> - [ <u>Translate this page</u> ]<br>由保育で三育法を実践.<br>dslink.jp/chibikko- <b>kitakurihama</b> / - 2k - <u>Cached</u> - <u>Similar pages</u> - <u>Note this</u>   |            |
| misaki  | ary<br>d for my friend on <b>kitakurihama</b> station at 1:00pm. there are a lot of fields in<br>guchi. i drove my car from <b>kitakurihama</b> station to his house<br>tajifu.blogspot.com/ - 61k - <u>Cached</u> - <u>Similar pages</u> - <u>Note this</u>  |            |

# Examples - query spelling

| 1 | 000                     | )                                     |                                     |                                  |                         |   | i                  | nfermatics – Go  | ogle Se   | arch                   |                     |                          |                  | $\bigcirc$  |
|---|-------------------------|---------------------------------------|-------------------------------------|----------------------------------|-------------------------|---|--------------------|--|-----------|------------------------|---------------------|--------------------------|------------------|-------------|
|   | <b>(</b> )              | ≥- (                                  | ଟି 😣                                |                                  | G http://               | www.google  | e.com/se           | arch?hl=en&clier   | t=firefo  | x-a&rls=org.mo         | zill 🔻 🕨 🧃          | G 🔹 keitei               | Q                | 340<br>1970 |
|   | Last open               | n▼ Cl                                 | ass Deve                            | lopment                          | t▼ Rese                 | arch Links 🔻  | UCI 🔻              | Thesis pages ▼   | DR 🔻      | Conferences <b>v</b>   | Place Lab ▼         | OpenPresence ▼           |                  | >>          |
|   | Web In                  | nages                                 | Maps                                | News                             | Shopping                | <u>Gmail</u> r  | <u>nore</u> ▼      | d  | onald.j.j | patterson.iii@gr       | mail.com   <u>W</u> | <u>eb History   My A</u> | ccount   Sign ou | t 💧         |
|   | Gc                      | 00                                    | gle                                 | infer                            | matics                  |   |                    |  |           | Search Advan<br>Prefer | ences               |                          |                  |             |
|   | Web                     |                                       |                                     |                                  |                         |   |                    |  |           | Results 1 -            | 10 of about 8       | 7 for infermatics        | (0.13 seconds)   |             |
|   | Did you                 | ı meai                                | n: <u>infori</u>                    | matics                           | 1                       |   |                    |  |           |                        |                     |                          |                  |             |
|   | dellal inte<br>Son Yazı | ernatıor<br>ılar. DÜ                  | al <b>infern</b><br>ŞÜNCEL          | n <b>atics</b> .<br>ER. 18/      | emrah del<br>6/2006 - [ | al inferm<br>lal. Bağlant<br>DÜŞÜNCEL<br>hed - <u>Simil</u> | ilar. Ana<br>ER    | Sayfa · Profilim   | · Arşiv · | RSS.                   |                     |                          |                  | U           |
|   | emrah de                | ellal,ark                             | adaş arar                           | na,soht                          | pet,google              |   | m tr tc, de        | ww dellal<br>ellal.com.tr.tc.                                  |           |                        |                     |                          |                  |             |
|   | dedicated               | ademic<br>d <b>Infern</b><br>g/lexwil | commur<br>natics for<br>ki.pl?actic | nity, pro<br>r both a<br>on=brow | cademic a               | id assessin<br>and general<br>1&id=Inferm                   | consum             |  | nd produ  | ucing                  |                     |                          |                  |             |
|   | Medical I<br>1-1-1 Min  | nferma<br>nami-Ko                     | a <b>tics</b> and<br>gushi, Ul      | Decisio<br>be, Yam               | n Science<br>nagnchi: 7 | es, Yamagu<br>'55–8505, J                                   | ichi Unive<br>apan | titis C virus<br>ersity, School of<br>ilar pages - <u>Note</u> |           | е,                     |                     |                          |                  |             |
|   | Ingenta                 | Conn                                  | ect Med                             | lical e                          | conomic                 | s-Hepatit   | tis C vir          | us infection a   | <u>nd</u> |                        |                     |                          |                  | *           |
| [ | Done                    |                                       |                                     |                                  |                         |   |                    |  |           |                        |                     |                          |                  | 1           |

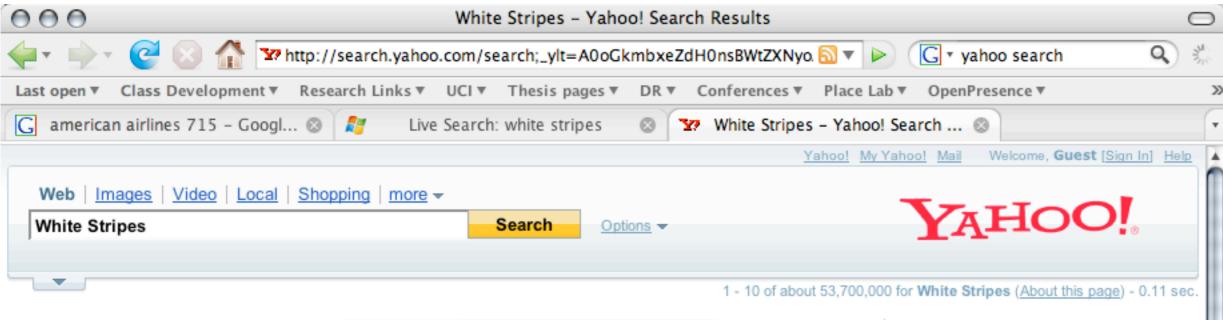
# Examples - query expansion

|  | Live Search   | : rock                 |  | 0       |
|--|---|------------------------|--|---------|
|  | com/results.aspx?q=rock&                            | &go=Search&mkt=en-u    | 🔊 🔻 🕨 🚺 🔭 keitei   | Q *     |
| Last open ▼ Class Development ▼ Research Links ▼   | UCI ▼ Thesis pages ▼                                | DR   Conferences       | Place Lab ▼ OpenPresence ▼   | 2       |
| Live Search  | Home Hotmail  | Spaces OneCare         |  | Sign in |
| rock   | Search  | Advanced · Options     |  |         |
| Web results 1-10 of 441,000,000<br>See also: Images, Video, News, Maps, MSN, More  | 2 V   | _                      | Related searches<br>Rock Music   |         |
| The Rock - Jellyfish.com/Movies<br>Shop and compare prices across stores. Rebates u  | up to \$1.10.                                       | Sponsored sites        | Rocks Minerals<br>Pictures Of Rocks  |         |
| Rock.com® - The Official Site of Rock Music<br>Free Email & Free Music Downloads by Rock.com.<br>more. Get your free @rock.com email today<br>www.rock.com · Cached page   |   | Music Downloads &      | Rock Star<br>Kid Rock<br>Rock Cycle<br>Hard Rock Cafe  |         |
| Rock music - Wikipedia, the free encyclopedia<br>Rock music is a form of popular music with a promi<br>drums, and bass. Many styles of rock music also<br>en.wikipedia.org/wiki/Rock_music · <u>Cached page</u>                                | inent vocal melody accom<br>use keyboard instrument |                        | Sponsored sites<br>Free Rock Music<br>Listen To Your Favorite Music<br>It's Free with the Music Toolba<br>Music.alot.com |         |
| Rock and roll - Wikipedia, the free encycle<br>Rock and roll is a genre of music that evolved<br>1950s, and quickly spread to the rest of the will<br>en.wikipedia.org/wiki/Rock_and_ roll · Cached<br>Show more results from en.wikipedia.org | d in the United States in the vorld.                | e late 1940s and early | The Rock<br>Interviews with Stars, Newsma<br>Icons at The Biography Chann<br>www.biography.com                           |         |
| Rock - Australia's Climbing Magazine<br>Subscription information. Overview of current issue<br>sale.   | e. Classifieds. Back issues                         | s and guide books for  | Rock Videos<br>Enjoy Savings & Selection on<br>Videos!<br>Shopzilla.com  | Rock    |

# Query shortcuts

- Map: "irvine, ca 92614"
- Calculation: "5+4"
- Flight Info: "american airlines 715"
- Stock price: "msft"
- Unit conversion: "1 dollar in euro"
- Music: "White Stripes"

## **Examples - query shortcuts**





#### The White Stripes

Official site for The White Stripes, the Detroit-based garage/blues duo featuring Jack and Meg White.

www.whitestripes.com - 2k - Cached

#### White Stripes v2

Jack and Meg White fansite includes news and tour information, forum, links, release details, and more.

SPONSOR RESULTS

White Stripes Ringtones Complimentary White Stripes Ringtones. Get Them Instantly. www.TheBombRingtones.com

White Stripes Ringtones White Stripes Ringtones. Get Them Complimentary Now.

ChartRingers.com/whitestripes

#### The White Stripes

Find, compare & buy. Compare & Buy from 1000's of Stores. www.Dealtime.com

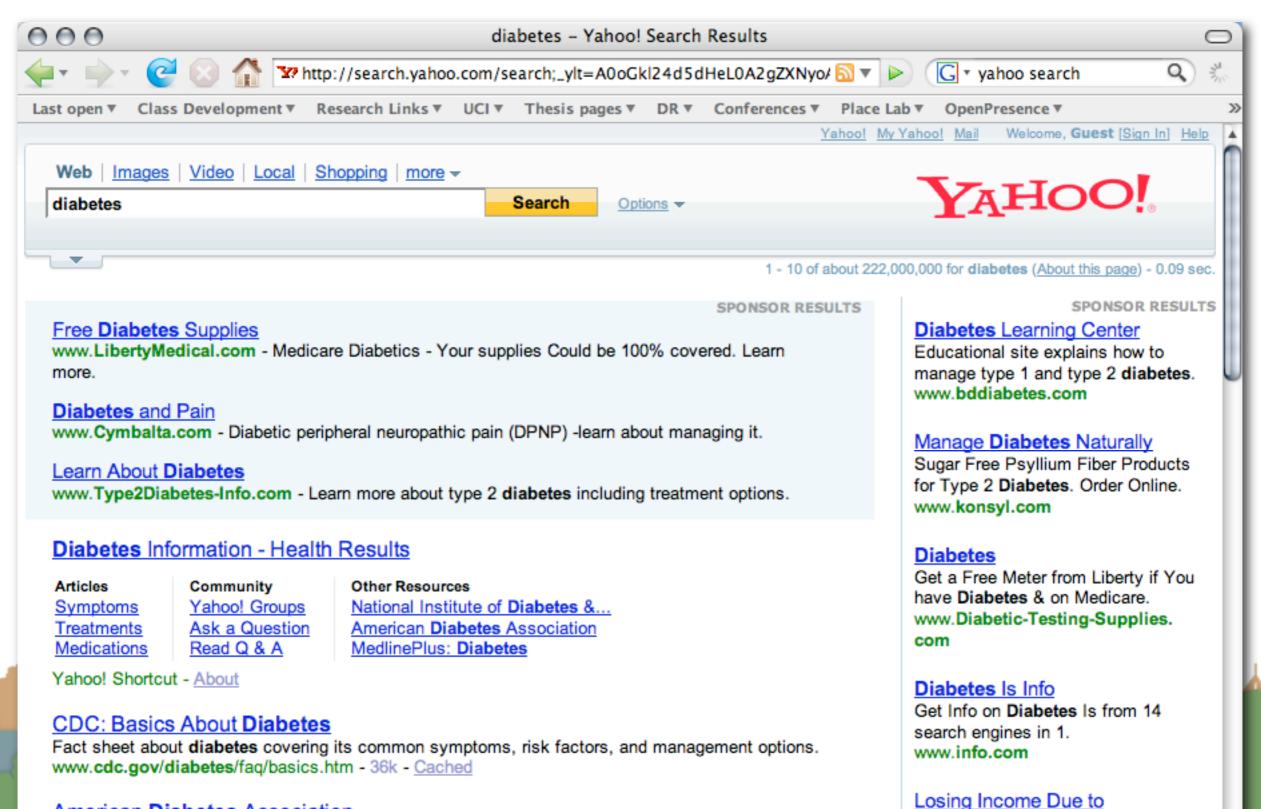
The White Stripes at Amazon.com

Low prices on new & used music. Qualified orders over \$25 ship free. Amazon.com/music

## Examples - query shortcuts

| Live Search: Don Patterson irvine ca  | 0                                       |
|---|---|
| 🖕 🔹 📄 🗸 🥝 🏠 ಶ http://search.live.com/results.aspx?q=Don+Patterson+irvine+ca&g( 🗟 🔻 🕨 💽 🔻 yahoo search   | Q 🐇                                     |
| Last open V Class Development Research Links UCI Thesis pages Conferences Place Lab OpenPresence Conferences  | >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>> |
| <u>G</u> american airlines 715 - Googl 🛛 ಶ Live Search: Don Patterson irvin 🛇 🐄 White Stripes - Yahoo! Search 🛇   |   |
| Live Search Home Hotmail Spaces OneCare   | Sign in                                 |
| Don Patterson irvine ca <u>Advanced</u> · <u>Options</u>  |   |
| Web results 1-9 of 253,000 Sponsored sites  |   |
| See also: Images, Video, News, Maps, MSN, More V  |   |
| Top local listings for Don Patterson near Irvine, CA (1 more) Is this useful? Yes No Visiting California & Need a Ho Reserve Rooms for Irvine CA  |   |
| www.calibex.com/hotels  | ·                                       |
| Don Patterson (626) 918-1974 15626 Dubesor St, La Puente  |   |
| Don Patterson (213) 389-7583 226 S Berendo St Apt 205, Los Angeles <u>Locate Don Patterson</u><br>Don Patterson (760) 729 6941 2768 Dundeo Ct. Carlshad Current address and phone available | ailable.                                |
| Don Patterson (760) 729-6941 2768 Dundee Ct, Carlsbad Instant results.  |   |
| DBLife: University of California-Irvine www.usa-people-search.com   |   |
| Jadwiga Indulska, University of Queensland, Australia Don Patterson, University of California Irvine Don Patterson at Amazon  | 1                                       |
| , USA Tom Rodden, Nottingham University, UK Program Committee: Low prices on new & used must  | sic.                                    |
| dblife.cs.wisc.edu/search.cgi?entity=entity-13655& amp;begin=60 · Cached page Qualified orders over \$25 ship amazon.com/music  | free                                    |
| DBLife: University of California-Irvine   |   |
| Jadwiga Indulska University of Queensland, Australia Don Patterson University of California   |   |
| Irvine, USA Tom Rodden Nottingham University, UK<br>UK  |   |
| dblife.cs.wisc.edu/search.cgi?entity=entity=13655& amp;begin=90 · Cached page www.nextag.com/hotels   |   |
| Show more results from dblife.cs.wisc.edu   |   |
| Jerry M. Patterson  |   |
| Professor: University of California at Irvine (2000-) Moulton Patterson Associates Partner See your message here  |   |
| (1967-75) State Bar of California 1967. Do you know something we don't?   |   |
| www.ppdb.com/people/406/000163014 - Cashed asso   |   |

# Examples - query aggregations



#### American Diabetes Association

# Context

- Context determination
  - Spatial
    - User Location
      - IP geocoding
      - cell phone localization
      - User input ("make this my home")
    - Web page location
    - Target Location
      - "SNA" maps to a latitude and longitude

# Context

- Query Stream
  - Previous queries have influence
- Personal Search
  - Your profile and history have influence
- Explicit
  - You choose to use a vertical search engine

# Context

- Implicit Context
  - google.fr is the default in france
    - User Location and Target Location

| 000                               | Google   |  | 0                   |
|-----------------------------------|--|--|---------------------|
| 🖕 - 🚽 🥑 🛞 🏠 🖸 http                | p://www.google.fr/   | 🔹 🕨 🤇 🕻 yahoo search   |                     |
| Last open ▼ Class Development ▼ R | esearch Links 🔻 UCI 🔻 Thesis pages 🔻 DR 🔻 Conferences 🔻 I                                | Place Lab ▼ OpenPresence ▼   | >>                  |
| G american airlines 715 - G ⊗     | 🝠 Live Search: Don Patterson 🛞 🍸 White Stripes – Yahoo! S                                | Sea 🛛 🖸 Goo  | ogle 🛛 🔍 🗸          |
| Web Images Maps Actualités V      | /idéo <u>Gmail</u> plus ▼  | ]  | iGoogle   Connexion |
|                                   | Google   |  |                     |
|                                   | Recherche Google J'ai de la chance<br>Rechercher dans :  Web  Pages francophones  Pages  | Recherche avancée<br>Préférences<br>Outils linguistiques<br>s : France |                     |
| Program                           | mes de publicité - <u>Solutions d'entreprise</u> - <u>À propos de Google</u> - <u>Go</u> | ogle.com in English  |                     |

# Context Use

- Restrict Results / Delete inappropriate results
- Rerank results based on context



| ( | 000   | Amazon.com: single cookbook  |                     |
|---|---|--|---------------------|
| ( | 🖕 - 🚽 🥑 🚳 🏠   | Inttp://www.amazon.com/exec/obidos/search-handle-url/index=blendl▼ ► G * yahoo search  | h 🔍 🕺               |
|   | Last open ▼ Class Developme   | nt▼ Research Links▼ UCI▼ Thesis pages▼ DR▼ Conferences▼ Place Lab▼ OpenPresence▼   | Blogs and Wikis ▼ ≫ |
|   | Amazon.com: single cook.  | 💿 💦 Live Search: Don Patterson 💿 🐄 White Stripes – Yahoo! Sear 💿 🔀 God   | ogle 🛛 🔍 🔻          |
|   | amazon.com  | Hello. Sign in to get personalized recommendations. New customer? Start here.         Your Amazon.com       If Today's Deals       Gifts & Wish Lists       Gift Cards       Yo  | ur Account   Help   |
|   | Shop All Departments 🛛 😒  | Search Amazon.com 🔄 single cookbook 💿 💓 Cart   | Your Lists 💌        |
|   | Category<br>Any Category  | "single cookbook"  |                     |
|   | Books (395)<br>Software (2)<br>Home & Garden (1)                          | Showing Top Results <ul> <li>Previous   Page: 1 2 3   Next &gt;</li> </ul> 1.              Serves One: Simple Meals to Savor When You're on Your Over Updecker (Paperback - Dec 25, 2005)           1.              Serves One: Simple Meals to Savor When You're on Your Over Updecker (Paperback - Dec 25, 2005) | <b>vn</b> by Toni   |
|   | Listmania!  | Buy new: \$16.95 \$11.53 26 Used & new from \$7.94<br>Get it by Thursday, Jan 24 if you order in the next 5 hours and choose of<br>Eligible for FREE Super Saver Shipping.<br>(8)<br>Books: See all 395 items  | ne-day shipping.    |
|   | Cooking - Just the two of<br>us/you: A list by Laura                      | <ol> <li>2. Eating in: The Official Single Man's Cookbook by Rich Lippma<br/>Maldonado, and Joe Azar (Paperback - Oct 1988)</li> <li>Buy new: \$8.95 <u>11 Used &amp; new</u> from \$5.92</li> <li>Not in stock; order now and we'll deliver when available<br/>Eligible for FREE Super Saver Shipping.</li> </ol> | in, Jose            |
| - | Best Books for Health and<br>Diet in 2007: A list by Moni<br>Jeddan "PWR" | Books: See all 395 items   | 25 1998)            |
| Ĩ | Create a Listmania! list Search Listmania!                                | 3. Cook INSIDE<br>Buy new: \$16.95 \$11.53 33 Used & new from \$7.45<br>Get it by Thursday, Jan 24 if you order in the next 5 hours and choose o<br>Eligible for FREE Super Saver Shipping.  |                     |

1