The Chemist & the Web

Stephen R. Heller
MDL Information Systems Inc.
steveh@mdli.com
The slides from this presentation can be found at:

www.hellers.com/steve/pub-talks/chicago01/slide001.html
With apologies to Benjamin Disraeli...

There are four kinds of lies:

Lies
Damned Lies
Statistics
Internet statistics
Search Engines are like spouses

- if you have more than one things get complicated.
The Internet/World Wide Web contains a great deal of useful and valuable information for the practicing chemist.

While some of the information is free, a majority of the valuable information requires payment for access.
This presentation will provide an overview of what is on the web, both free and fee-based, and provides examples of what chemists do on an everyday basis when they access the web.

Examples will be taken from a variety of web resources such as the WebBook, ChemWeb, Chemindustry.com, chemical societies, and publishers.
Zeitgeist Explained...

zeit·geist | Pronunciation: 'tsIIt-"gIst, 'zIIt | Function: noun | Etymology: German, from Zeit (time) + Geist (spirit) | Date: 1884 | Meaning: the general intellectual, moral, and cultural climate of an era

12 Most Popular Search Terms at 10 Leading Portals and Search Engines from March 1999 to January 2001

<table>
<thead>
<tr>
<th>Rank</th>
<th>Term</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>sex</td>
<td>0.3289%</td>
</tr>
<tr>
<td>2</td>
<td>hotmail</td>
<td>0.2131%</td>
</tr>
<tr>
<td>3</td>
<td>yahoo</td>
<td>0.2044%</td>
</tr>
<tr>
<td>4</td>
<td>porn/pornography</td>
<td>0.1402%</td>
</tr>
<tr>
<td>5</td>
<td>chat/chatrooms</td>
<td>0.1233%</td>
</tr>
<tr>
<td>6</td>
<td>mp3</td>
<td>0.0935%</td>
</tr>
<tr>
<td>7</td>
<td>horoscopes/horoscope</td>
<td>0.0800%</td>
</tr>
<tr>
<td>8</td>
<td>ebay</td>
<td>0.0731%</td>
</tr>
<tr>
<td>9</td>
<td>aol</td>
<td>0.0714%</td>
</tr>
<tr>
<td>10</td>
<td>games</td>
<td>0.0659%</td>
</tr>
<tr>
<td>11</td>
<td>map/maps</td>
<td>0.0585%</td>
</tr>
<tr>
<td>12</td>
<td>Pokemon</td>
<td>0.0549%</td>
</tr>
</tbody>
</table>
Google Top 10 Search Terms - February, 2001
(Gaining Queries)

1. gnutella
2. greeting cards
3. valentine cards
4. mp3
5. dale earnhardt
6. all your base are belong to us
7. valentine
8. Napster
9. sms
10. nascar
Google Top 10 Search Terms - February, 2001
(Declining Queries)

1. xfl
2. annabel chong
3. survivor
4. eminem
5. groundhog day
6. earthquake
7. india
8. big brother
9. anna kournikova
10. ginger
Google Top 10 Search Terms – Week of August 13, 2001
(Gaining Queries)

1. baycol
2. default.ida
3. shannon elizabeth
4. fantasy football
5. ikea
6. nvidia
7. kid rock
8. jorge amado
9. james dean
10. umptn 2001

Where are “stem cells”?
Google Top 10 Search Terms – Week of August 13, 2001
(Declining Queries)

1. code red virus
2. tour de france
3. norton antivirus
4. planet of the apes
5. mariah carey
6. etna
7. Leilani Rios
8. orbitz
9. camping
10. rakhi cards
Lycos Top 12/50 (8/01)

1  Dragonball
2  Code Red Virus
3  Tattoos
4  Britney Spears
5  WWF
5  WWF
6  Pokemon
7  Las Vegas
8  Morpheus
9  IRS
10 Final Fantasy
11 Sircam Virus
12 Planet of the Apes
finite capacity scheduling
transport properties
adultery
feuerstein
conjugate gradient
lamb wave
antifreeze proteins
mars
antigravity
gene
cunliffe
broadband
linguistic categories
dna
enso
hazardous waste management and policies
brasseur
capillary optics
rechnernetze
pearson distribution
uranium recovery
variable frequency drives
celiac disease
mars
jadeite
cloning
cichlid?
starch
phase field
heat*shock response
free amino acids
chromatography
thamnophis
singlet oxygen
acidity
notes
double barrier tunneling
probiotics
protein stability
bitter taste
force constant
dissociative identity disorder
Goto.com - July 2001

Total: ~2 Billion searches

- prozac: 43,345
- chemistry: 35,087
- lipitor: 19,378
- organic chemistry: 6,779
- prozac side effects: 3,794
- benzene: 2,474
- analytical chemistry: 1,446
- journal of biological chemistry: 1,324
- mass spectrometry: 1,278
- cas number/s: 1,104
- acs: 848
- royal society of chemistry: 397
- journal of the american chemical society: 336
- chemistry molotov cocktail: 328
- career in chemistry: 168
- buy viagra online: 134
- chemistry web site: 128
Top terms from ACS journals web site
(April 2000 and June 2001)

in-1319 ; the-1173 ; alphabet (a: 5911 - q: 122)
acid-1109
synthesis-800
protein-683
polymer-558
water-504
dna-492
metal-442
molecular-432
chemical-404
analysis-399
chemistry-393
reaction-384
structure-361
steve-18 ...

Slide 19 of 394
ACS Web Site Usage

1. ACS meetings
2. Chicago meeting
3. ACS Education
4. Search of the web site
5. ACS careers
6. ACS library
7. ACS Employment
8. ACS membership
9. ACS Government Affairs
10. Join ACS
RSC Web Site Usage - May, 2001

- solid state reaction: 2599
- cation: 1130
- zeolite: 929
- xps: 747
- phosphors: 638
- vitamins fruit: 626
- rubber: 602
- water oxidation: 601
- hexavalent uranium: 588
- bimetallic: 553
- silver: 515
- battery: 511
- fuel cell: 483
ChemSoc Web Site Usage - May, 2001

watoc - 164
periodic table - 89
hplc - 70
elements - 50
green chemistry - 45
ursodeoxycolic acid - 33
royal society of chemistry - 32
medicinal chemistry - 31
periodic table of elements - 31
jobs in chemistry - 28
chemists - 27
acquisition solutions - 26
<table>
<thead>
<tr>
<th>Term</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iliev</td>
<td>1845</td>
</tr>
<tr>
<td>Chitosan</td>
<td>1723</td>
</tr>
<tr>
<td>Benzene</td>
<td>1171</td>
</tr>
<tr>
<td>Microwave</td>
<td>1120</td>
</tr>
<tr>
<td>Polyanaline</td>
<td>1116</td>
</tr>
<tr>
<td>Chemist</td>
<td>1077</td>
</tr>
<tr>
<td>Zeolite</td>
<td>1018</td>
</tr>
<tr>
<td>Polymer</td>
<td>1015</td>
</tr>
<tr>
<td>Lipase</td>
<td>976</td>
</tr>
<tr>
<td>Chromatography</td>
<td>891</td>
</tr>
<tr>
<td>Cyclodextrin</td>
<td>887</td>
</tr>
<tr>
<td>Catalysis</td>
<td>881</td>
</tr>
</tbody>
</table>
### May
- Chemical manufacturers
- Chemical distributors & dealers
- Job banks, recruiting & job seeking
- Pharmaceuticals, biochemicals & intermediates
- Polymers, plastics & additives
- Petrochemicals & hydrocarbons
- Chemical names, properties & structures
- Agrochemicals & fertilizers
- Chemistry
- Adhesives and sealants

### July
- Chemical manufacturers
- Pharmaceuticals, biochemicals & intermediates
- Job banks, recruiting & job seeking
- Chemical distributors & dealers
- Periodic Tables
- Polymers, plastics & additives
- Petrochemicals & hydrocarbons
- Agrochemical & fertilizers
- Chemical names, properties & structures
- Oil, gas, refineries
<table>
<thead>
<tr>
<th>May</th>
<th>June</th>
<th>July</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenol</td>
<td>Phenol</td>
<td>Soda Ash</td>
</tr>
<tr>
<td>Methanol</td>
<td>Soda Ash</td>
<td>Titanium Dioxide</td>
</tr>
<tr>
<td>Caustic Soda</td>
<td>Titanium Dioxide</td>
<td>Citric Acid</td>
</tr>
<tr>
<td>Citric Acid</td>
<td>Methanol</td>
<td>Carbon Black</td>
</tr>
<tr>
<td>Acetone</td>
<td>Caustic Soda</td>
<td>Chlorine</td>
</tr>
<tr>
<td>Sulfuric Acid</td>
<td>Acetic Acid</td>
<td>Red Phosphorus</td>
</tr>
<tr>
<td>Acetic Acid</td>
<td>Phosphoric Acid</td>
<td>Sodium Hydroxide</td>
</tr>
<tr>
<td>Chlorine</td>
<td>Ammonia</td>
<td>Ethanol</td>
</tr>
<tr>
<td>Titanium</td>
<td>Hydrogen Peroxide</td>
<td>Potassium Nitrate</td>
</tr>
</tbody>
</table>
Chemindustry.com - Most often Searched for Non-Chemical Terms

<table>
<thead>
<tr>
<th>May</th>
<th>June</th>
<th>July</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jobs</td>
<td>Jobs</td>
<td>Jobs</td>
</tr>
<tr>
<td>Chemical Engineer</td>
<td>MSDS</td>
<td>MSDS</td>
</tr>
<tr>
<td>Process Equipment</td>
<td>Values</td>
<td>Chemical Databases</td>
</tr>
<tr>
<td>Chemical Engineering</td>
<td>Chemical Reactions</td>
<td>Chemical Databases</td>
</tr>
<tr>
<td>Chemical Reactions</td>
<td>Chemical Engineering</td>
<td>Periodic Table</td>
</tr>
<tr>
<td>Lab Equipment</td>
<td>Periodic Table</td>
<td>Pumps</td>
</tr>
<tr>
<td>Periodic Table</td>
<td>Pumps</td>
<td>Chemical Reactions</td>
</tr>
<tr>
<td>Pumps</td>
<td>Process Equipment</td>
<td>Pumps</td>
</tr>
<tr>
<td>Heat Exchangers</td>
<td>Custom Manufacturing</td>
<td>Values</td>
</tr>
<tr>
<td>Used Equipment</td>
<td>Chemical Databases</td>
<td>Software</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Process Equipment</td>
</tr>
</tbody>
</table>
Dear Dr. Heller,

Thank you for contacting the ISI. Technical Help Desk. I'm sorry, but we are unable to provide you with the information you are seeking. The only information available close to your request is highly cited papers, but you stated that is not what you are seeking.

Thank you for your interest in ISI products.
We do not have any information available directly related to your request. We have recently released a new web site, Science Spotlight, which lists most frequently requested documents via ChemPort.

http://www.cas.org/spotlight
ChemGuide Usage

Name/Synonym - 53%
Name + Property (benzene+toxicity) - 10%
<table>
<thead>
<tr>
<th>Type of search</th>
<th>Searches</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical name</td>
<td>91762</td>
<td>47.9%</td>
</tr>
<tr>
<td>Chemical formula</td>
<td>68601</td>
<td>35.8%</td>
</tr>
<tr>
<td>CAS registry number</td>
<td>15990</td>
<td>8.3%</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>7346</td>
<td>3.8%</td>
</tr>
<tr>
<td>Energy levels: vibrational frequency</td>
<td>2814</td>
<td>1.5%</td>
</tr>
<tr>
<td>Chemical structure: applet interface</td>
<td>1439</td>
<td>0.8%</td>
</tr>
<tr>
<td>Ionization energy</td>
<td>823</td>
<td>0.4%</td>
</tr>
<tr>
<td>Chemical structure: form interface</td>
<td>610</td>
<td>0.3%</td>
</tr>
<tr>
<td>Energy levels: band head energy</td>
<td>589</td>
<td>0.3%</td>
</tr>
<tr>
<td>Electron affinity</td>
<td>393</td>
<td>0.2%</td>
</tr>
<tr>
<td>Author</td>
<td>352</td>
<td>0.2%</td>
</tr>
<tr>
<td>Proton affinity</td>
<td>349</td>
<td>0.2%</td>
</tr>
<tr>
<td>Acidity</td>
<td>248</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

...
<table>
<thead>
<tr>
<th>Web Site</th>
<th>Popularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>google.com</td>
<td>5</td>
</tr>
<tr>
<td>goto.com</td>
<td>149</td>
</tr>
<tr>
<td>lycos.com</td>
<td>426</td>
</tr>
<tr>
<td>chemindustry.com</td>
<td>3,282</td>
</tr>
<tr>
<td>chemweb.com</td>
<td>22,800</td>
</tr>
<tr>
<td>isinet</td>
<td>44,555</td>
</tr>
<tr>
<td>cas.org</td>
<td>61,929</td>
</tr>
<tr>
<td>scirus</td>
<td>62,851</td>
</tr>
<tr>
<td>rsc.org</td>
<td>63,294</td>
</tr>
<tr>
<td>acs.org</td>
<td>67,842</td>
</tr>
<tr>
<td>webbook.gov</td>
<td>111,333</td>
</tr>
<tr>
<td>chemguide.de</td>
<td>279,504</td>
</tr>
<tr>
<td>warr.com</td>
<td>1,289,376</td>
</tr>
</tbody>
</table>

Relative Popularity of Chemistry Web Sites
Conclusions

1. General Search engines are used primarily for current events, general searches. Chemistry/chemical searches are rare and lead to randomly useful results.

2. Specific chemistry web sites are used for specific and generally non-overlapping information.
Conclusions

3. Chemical Names, Chemical Formula, and CAS Registry Numbers are the main way to search for information.

4. Different sites are designed for and used by different users. E.g., The ACS Pubs and RSC sites tend to have academic/scientific queries; the chemindustry.com tends to have industry related/directed queries.
Acknowledgements

I would like to thank all the individuals and organizations which provided me with their data and information. Without their help this presentation would not have been possible.