



D E F C O N X I

Online

CORPORATE

INTELLIGENCE

What we'll cover

- ◆ How online intelligence gathering differs from traditional intelligence gathering
- ◆ The difference between intelligence and espionage
- ◆ Corporate “Dash boards”
- ◆ Tips from the field
- ◆ Opportunities for the community



Who am I?

- ◆ MikeSchrenk@yahoo.com
- ◆ Minneapolis-based consultancy:
Michael Schrenk Ltd. (www.schrenk.com)
- ◆ Write Webbots and Spiders for corporate clients
- ◆ DEFCON X, “Introduction to Writing Spiders and Web Agents”.
- ◆ Also write, speak and teach



Intelligence = Information

- ◆ In a business sence, you want to know:
- ◆ What can you learn about your competition?
- ◆ What do people know about you?
- ◆ Are people stealing from you?
- ◆ *The most important thing is...*
- ◆ *What can you predict?*
 - ◆ Collect library of information
 - ◆ Compare changes



Definitions

- ◆ Intelligence is not necessarily:
- ◆ **Espionage**
 - ◆ spying,
 - ◆ launching trojans
 - ◆ Wiretapping
- ◆ **A covert action**
 - ◆ Tampering with a situation to change
 - ◆ A Strategy
 - ◆ An Election



Traditional sources for collecting corporate intel

- ◆ Go to conferences,
- ◆ *Hire your competition's employees,*
- ◆ *Lookup patent records,*
- ◆ *Use secret shoppers,*
- ◆ *Study help wanted ads,*
- ◆ *Read trade publications,*
- ◆ *Talk to vendors*



Disadvantages of Traditional corporate intel

- ◆ *Mostly after-the-fact*
- ◆ *Requires contact*
- ◆ *Mostly one-time activities—must be repeated*
- ◆ *Cannot be done anonymously*
- ◆ *Can be expensive*



Gathering online intel means learning new habits

- ◆ Web agents determine how you use the Internet
 - ◆ Browsers
 - ◆ Mail clients
 - ◆ *News readers*
 - ◆ *Telnet*
 - ◆ *Competitive advantages come when you perform better and differently.*
 - ◆ *Webbots/Spiders*



Advantages of online corporate intelligence

- ◆ Can be done from a distance (*with stealth*)
- ◆ Can be automated
- ◆ *Can be done anonymously (for the most part)*
- ◆ *Reduces or eliminates latency between when an event happens and a decision can be made*
- ◆ *Can be interactive*
- ◆ *Can create relevance that traditional methods cannot*



Online Corporate Intel most effective when...

- ◆ It is automated
- ◆ Data can be parsed and stored in a database
- ◆ *Stores data over a period of time*
- ◆ *Creates context by combining various data from a variety of sources*
- ◆ *Uses statistical analysis to make recommendations*
- ◆ Allows user to configure



Creating relevance

- ◆ Cross reference **multiple sources**
- ◆ Gather information **periodically** over time
 - ◆ *Trend analysis*
- ◆ *Show relationships between data*
- ◆ *Online intelligence applications can automate the evaluation process*



(Some) sources for online intelligence

- ◆ Corporate web sites
 - ◆ Job postings
 - ◆ Product pricing
 - ◆ *News*
- ◆ *Government web sites*
 - ◆ *Court records*
 - ◆ *SEC filings*
 - ◆ *Patent records*
 - ◆ *Census data*



(More) sources for online intelligence

- ◆ Online auctions

 - What's a good buy?

 - What's a good selling price?

- ◆ Whois servers

- ◆ News servers

- ◆ HTTP headers

- ◆ Mail

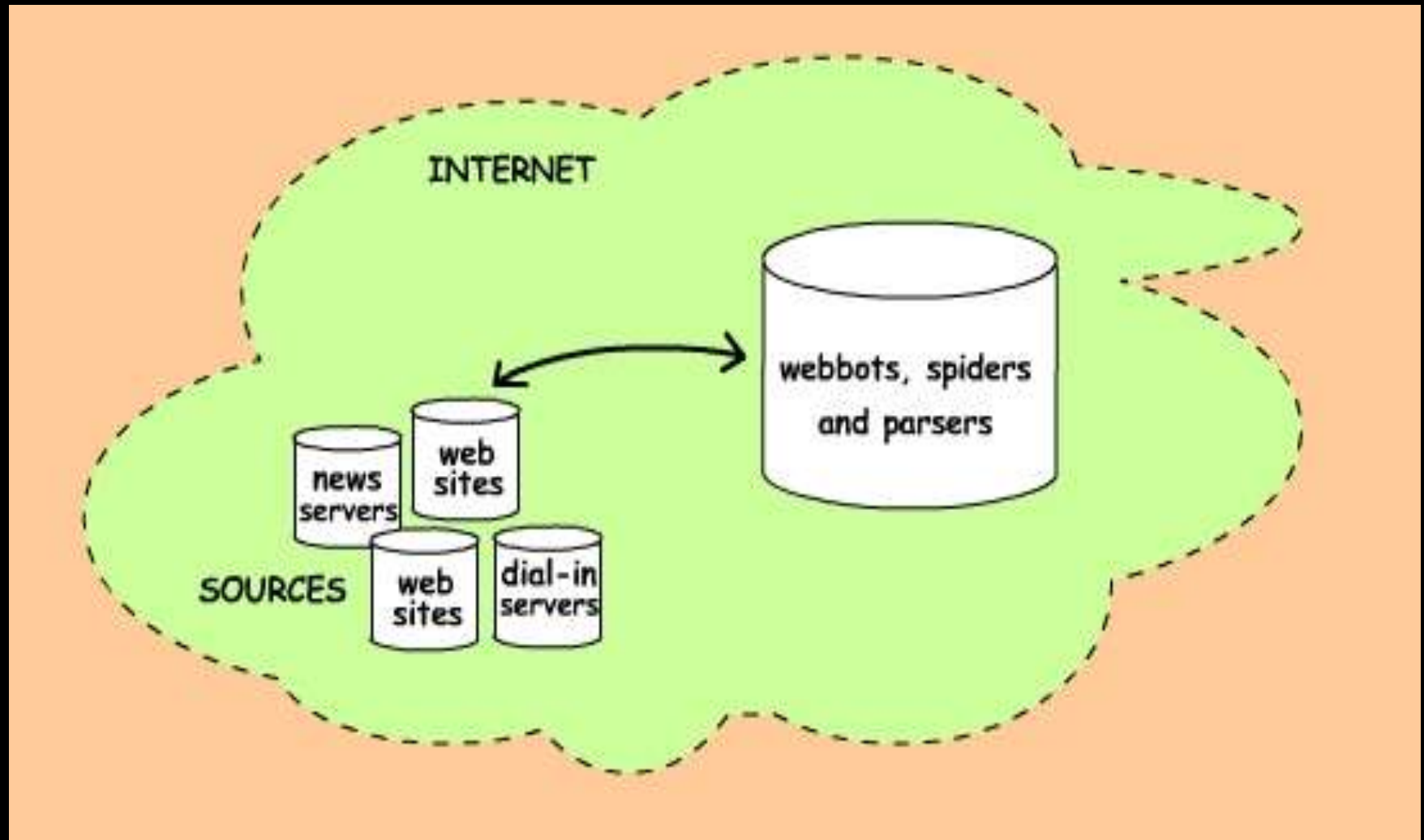


Technology

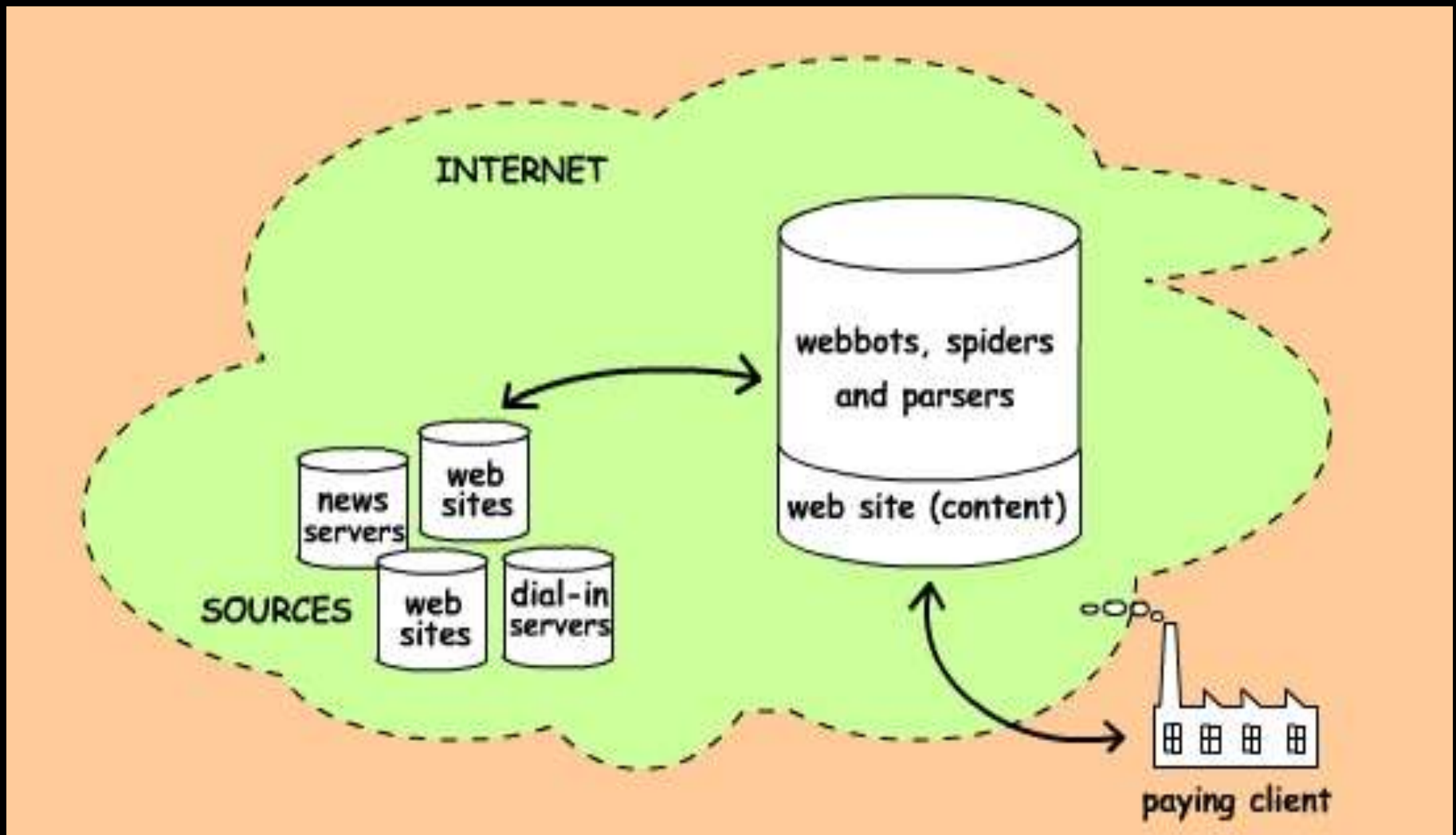
(the basics of collecting and using intelligence in three steps)



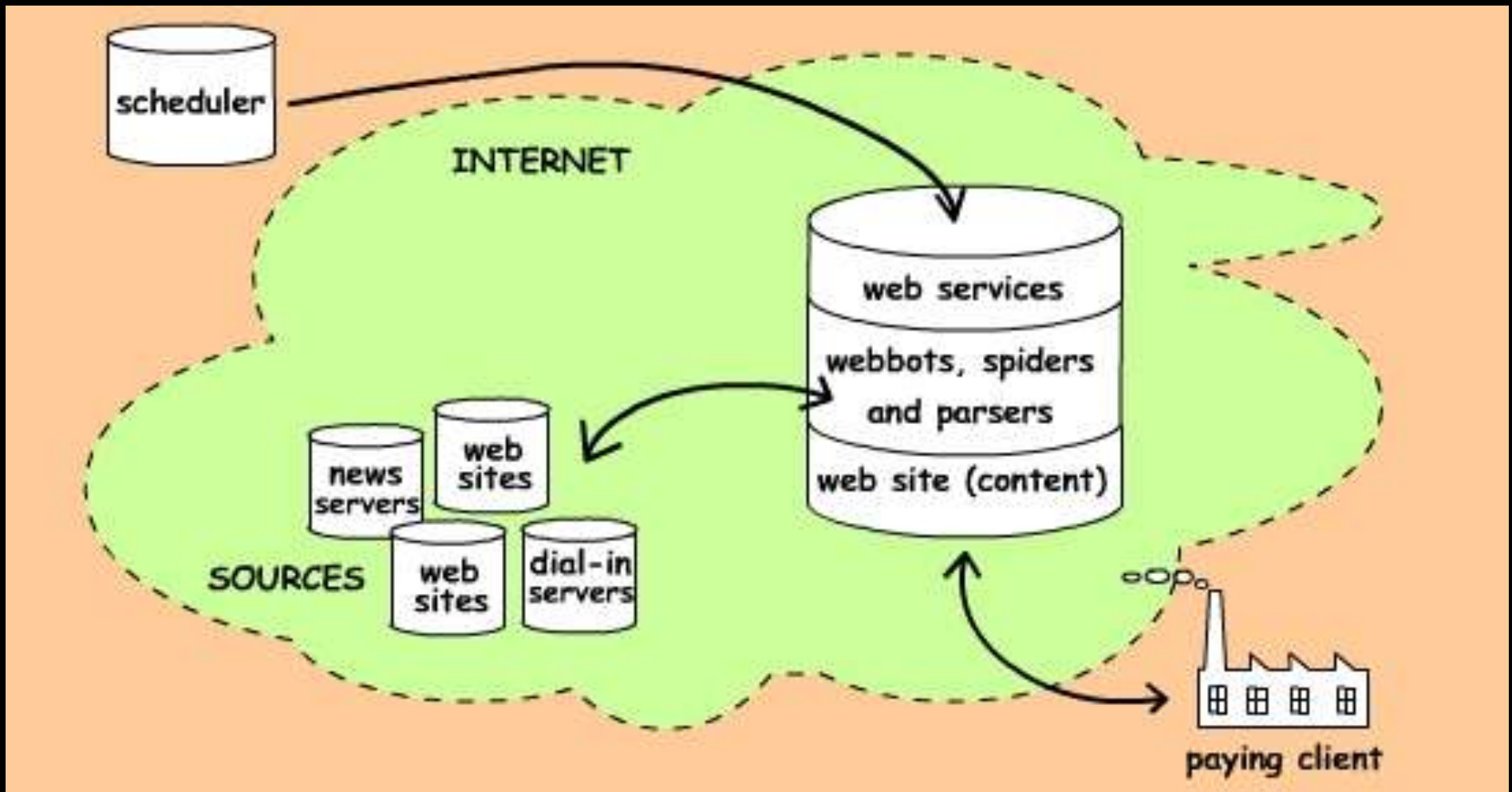
Identify sources, write bots and store data



Write a data-driven web site for the customer of the data



Create a scheduler for the webbots and spiders



Simple examples



Corporate Intelligence “Dashboards”

- ◆ *Provide “big picture” of data within some defined context*
- ◆ *Adds context to data*
 - ◆ *Filters*
 - ◆ *Statistics*
 - ◆ *Show trends*
- ◆ *Creates branding opportunities*



Comparitive Room Pricing of Competitors in Your Region

Month: July 2003

This chart shows your **competitors' pricing for standard rooms** for the month of July 2003. These prices are identical to the advertised prices on the hotels' official web pages and are **automatically updated** every 48 hours. In addition to the price, the chart also indicates **if the price has changed** in the past 48 hours and 10 days. (as shown in the key). Clicking on either **+** or **-** will display price history for that day.

KEY

\$128

PRICE

+

48 HOUR
TREND

-

10 DAY
TREND

The lower portion of the table shows how your prices are positioned among your competitors.

July 2003	TUE	WED	THU	FRI	SAT	SUN	MON	TUE	WED	THU	FRI	SAT	SUN	MON	TUE	WED	THU	FRI	SAT	SUN	MON	TUE	WED	THU	FRI	SAT	SUN	MON	TUE	WED	THU
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Competitor #1	102	102	120	161	159	163	89	100	103	124	157	158	163	87	104	101	119	158	158	162	91	103	102	120	157	164	164	90	86	98	124
Competitor #2	99	99	119	159	159	159	89	99	99	119	159	159	159	89	99	99	119	159	159	159	89	99	99	119	159	159	159	89	89	99	119
Competitor #3	103	100	117	163	164	154	88	97	95	119	155	159	156	84	98	104	114	164	156	155	89	103	95	122	158	163	156	87	92	96	124
Competitor #4	99	98	114	159	157	149	85	96	96	112	156	150	159	87	95	91	109	149	151	154	80	93	96	111	158	150	156	79	81	96	116
Competitor #5	72	72	92	132	132	132	62	72	72	92	132	132	132	62	72	72	92	132	132	132	62	72	72	92	132	132	132	62	62	72	92
Competitor #6	99	94	119	148	154	150	79	90	93	108	151	147	149	83	88	99	110	146	155	152	87	90	88	118	157	155	149	76	74	86	112
Competitor #7	93	92	111	150	158	154	83	93	95	114	153	158	152	82	94	91	116	156	150	151	88	92	91	115	158	152	152	80	81	96	117
Competitor #8	90	94	113	150	154	153	80	99	98	116	159	159	151	87	98	98	113	155	150	159	81	93	96	111	154	159	157	83	79	97	118
Competitor #9	90	94	110	152	151	151	80	98	97	119	159	149	159	82	95	95	109	149	149	151	87	96	92	117	152	157	155	89	86	97	119
Competitor #10	90	94	117	158	152	154	87	94	90	115	159	154	159	82	92	89	116	155	150	154	83	90	99	117	156	154	151	89	87	89	109
Lowest Prices	72	72	92	132	132	132	62	72	72	92	132	132	132	62	72	72	92	132	132	132	62	72	72	92	132	132	132	62	62	72	92
Highest Prices	103	102	120	163	164	163	89	100	103	124	159	159	163	89	104	104	119	164	159	162	91	103	102	122	159	164	164	90	92	99	124
Average Prices	93	93	113	153	154	151	82	93	93	113	154	152	153	82	93	93	111	152	151	152	83	93	93	114	154	154	153	82	81	92	115
YOUR PRICES	108	107	111	156	166	156	95	106	113	108	172	148	161	90	110	98	105	164	162	165	75	102	97	108	172	164	156	99	100	93	128
YOUR POSITION	13	13	1	2	12	5	13	14	20	5	21	3	8	9	17	5	5	12	11	13	7	8	5	4	20	11	3	18	18	1	13
	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG

Mail example

(Are people stealing from you?)



Sign In

[? Need Help?](#)

New to eBay?

or

Already an eBay user?

If you want to sign in, you'll need to register first.

Registration is fast and **free**.

[Register >](#)

Dear eBay member.
Your account has been chosen accidentally. Since there are a lot of cases of cheating, we'd like you to visit your account. If you are not going to do that, your account will be removed away from our site system.

eBay User ID

[Forgot](#) your User ID?

Password

[Forgot](#) your password?

[Sign In >](#)

[Keep me signed in](#) on this computer unless I sign out.

[Account protection tips](#) | [Secure sign in \(SSL\)](#)

Policing the Internet

- ◆ Problem:
 - ◆ People steal things and want to liquidate them quickly.
 - ◆ Online auctions are an attractive alternative to pawn shops.



Policing the Internet

(continued)

- ◆ Solution:
 - ◆ Create an online interface that allows law enforcement to enter groups of items stolen at the same time and place.
 - ◆ Write webbots that looks for individuals selling similar groupings from similar place online auctions.



What are people reading?

The screenshot shows a Netscape browser window displaying the Amazon.com website. The address bar shows the URL: <http://www.amazon.com/exec/obidos/tg/browse/-/224203/104-2948484-6664>. The page title is "Amazon.com Purchase Circles -- U.S. House of Representatives - Netscape".

The Amazon.com header includes the logo, navigation links (VIEW CART, WISH LIST, YOUR ACCOUNT, HELP), and a search bar. The search bar contains "All Products" and "GO!". The "Purchase Circles" section is active, showing "U.S. House of Representatives" and "See What's Selling".

The main content area features a list of books. The first item is "BUSH AT WAR" by Bob Woodward. The list price is \$28.00, and the current price is \$19.60. The average customer rating is 4.5 stars. A link to "more reviews" is provided.

On the right side, there are two boxes: "Explore Other Purchase Circles" with links to Geography, Government, Companies, Organizations, and Education; and "Purchase Circles Favorites" with an "ADD TO FAVORITES" button and a note: "(If you're not Michael G Schrenk, click)".

The browser's status bar at the bottom shows "Document: Done (4.39 secs)".

What are employess of Apple Computer reading?

1. Mac OS X in a nutshell

2. Mac OS X Hacks

3. Mac OS X: The Missing Manual

4. Pattern Recognition

5. Harry Potter and the Order of the Phoenix

6. What should I do with My Life?



Interactive Intelligence

- ◆ Sniping agents:
 - ◆ Software that places last-second bids on online auctions.
 - ◆ Prevents the bidding process from raising the auction prices.
 - ◆ Somewhat limited by proxy bidding



Interactive Intelligence

- ◆ Intelligent shopping software:
 - ◆ Webbots collect market information on select items.
 - ◆ When an items meet criteria (determined by collected market intelligence) those items are either automatically purchased or reccomended for purchase.
 - ◆ Stocks, online auctions, etc.



Online Sources

- Respect their bandwidth
- Be as stealthy as you can

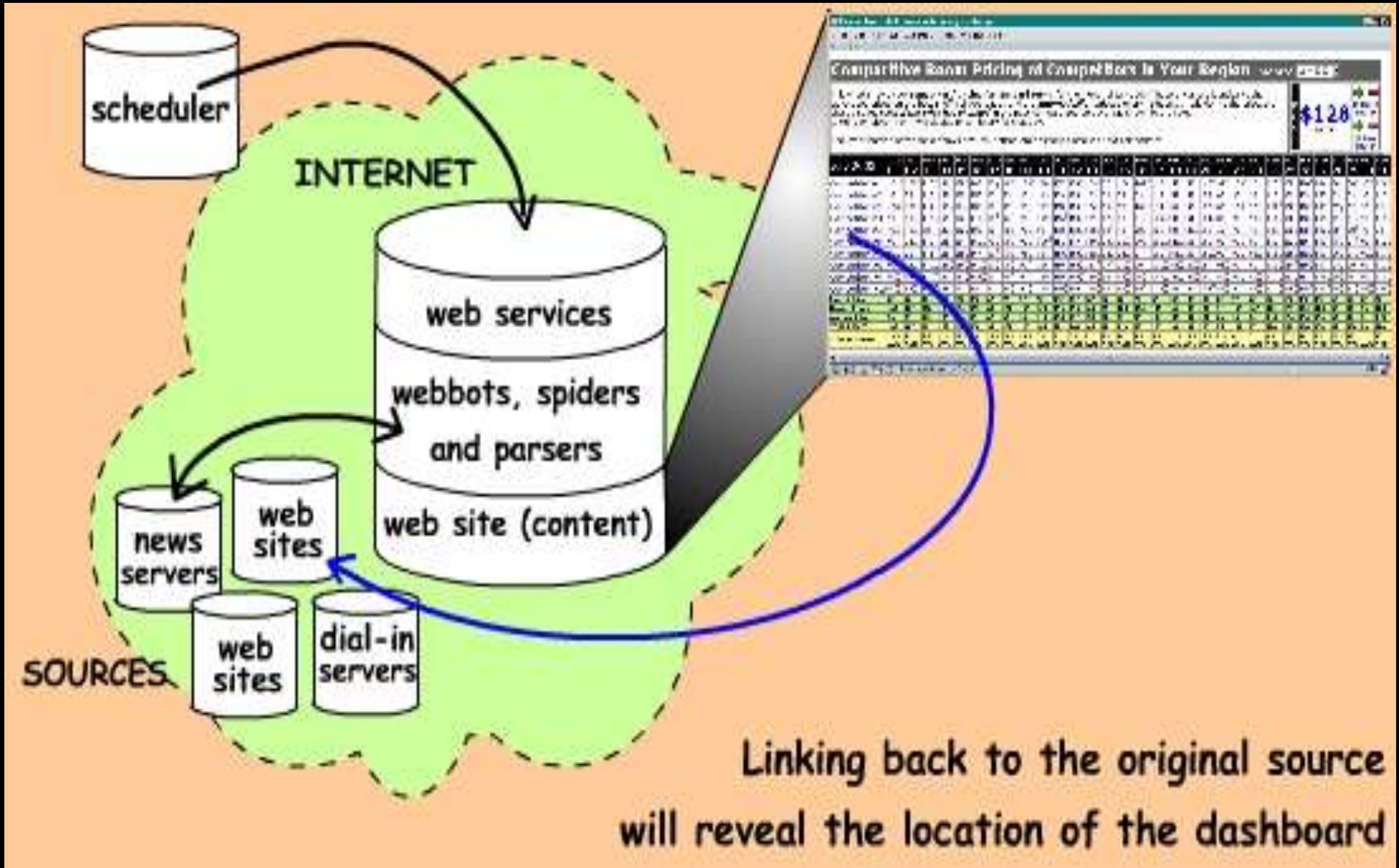


Tips on writing stealthy bots and spiders

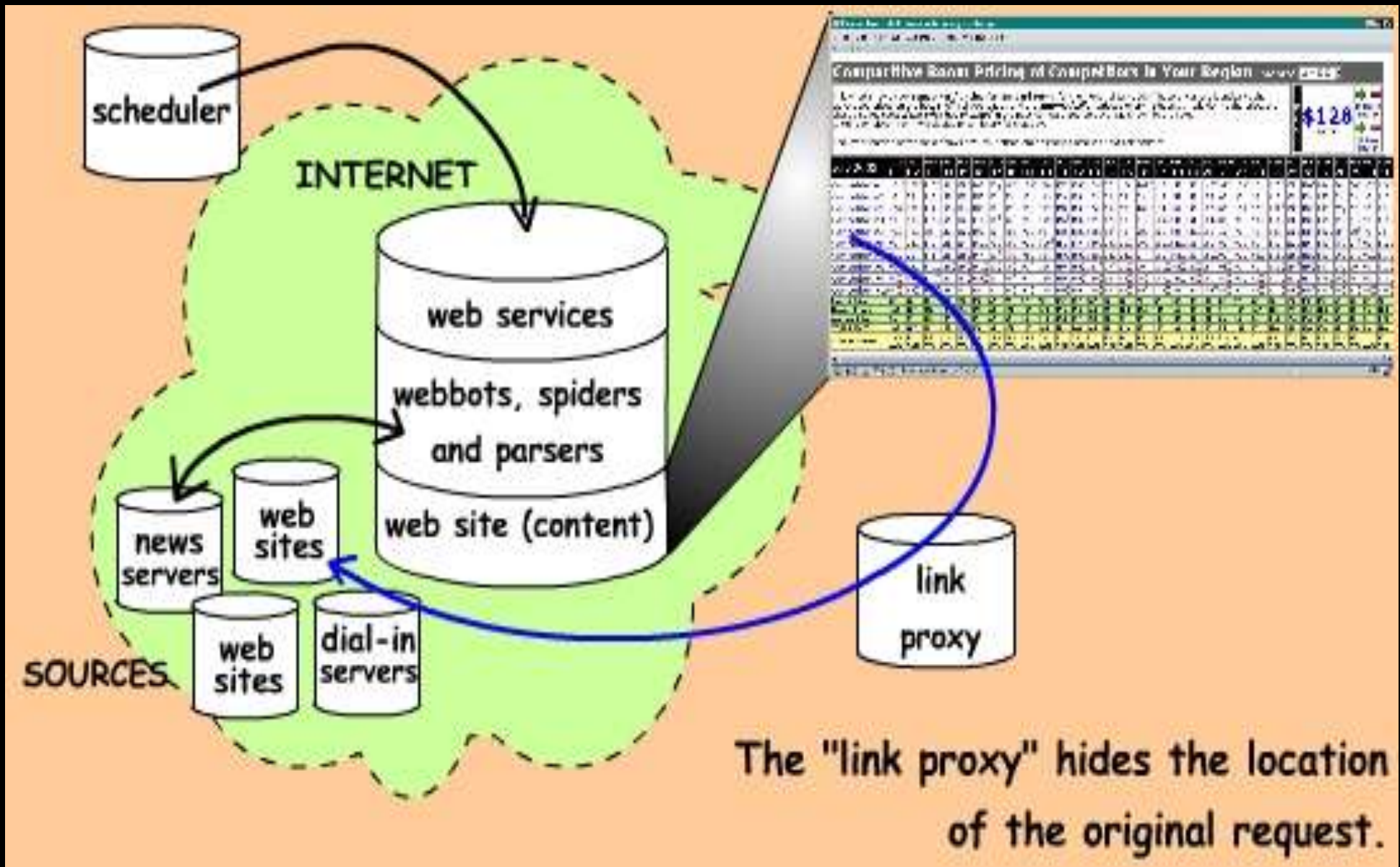
- Treat bandwidth with respect
 - (But don't forget to download both HTML & images)
- Introduce randomness
 - Change your start times
 - Randomize time periods between downloads
 - Randomize sequence of page downloads
 - Rotate IP addresses
 - Use a “link proxy:



Destroying the REFERER



Destroying the REFERER



Interfacing to the link proxy

- Change links from:
- ``
- `link`
- ``

- to:
- `<a href="http:www.some_safe_pace.com/link_proxy.php?`
- `url=https://www.some_online_resource.com/members/link.html">`
- `link`
- ``



Link proxy code

- `<BASE`
`HREF="https://www.some_online_resource.com/members">`
- `<?php`
- `$ch = curl_init();`
- `curl_setopt ($ch, CURLOPT_URL, $url);`
- `curl_setopt ($ch, CURLOPT_HEADER, 0);`
- `curl_setopt ($ch, CURLOPT_USERPWD, "username:password");`
- `curl_setopt ($ch, CURLOPT_REFERER, "");`
- `curl_setopt ($ch, CURLOPT_USERAGENT, "Mozilla/7.01");`
- `curl_exec ($ch);`
- `curl_close ($ch);`
- `?>`



Thank you (Q&A)

• Mike Schrenk

• MikeSchrenk@yahoo.com

• www.schrenk.com

