# EvilSeed: A Guided Approach to Finding Malicious Web Pages

L. Invernizzi[1] S. Benvenuti[2] M. Cova[3,5]
P. Milani Comparetti[4,5] C. Kruegel[1] G. Vigna[1]

[1]UC Santa Barbara

[2]University of Genova

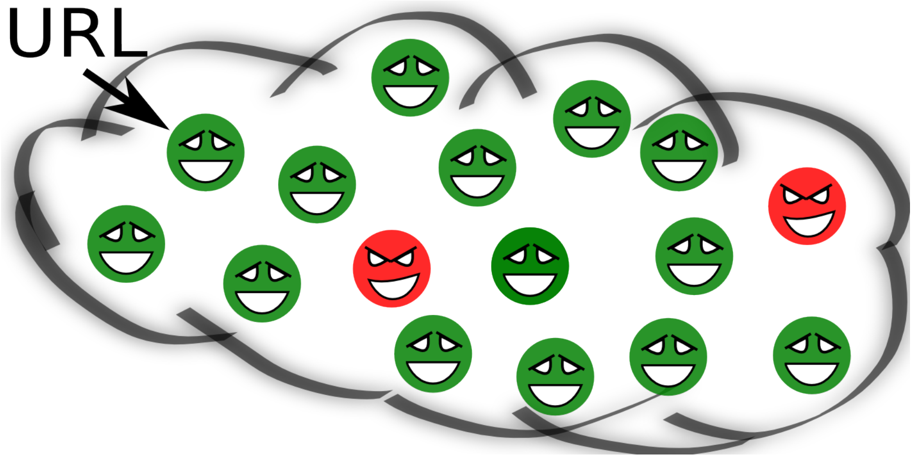
[3]University of Birmingham

[4]Vienna University of Technology

[5]Lastline, Inc.

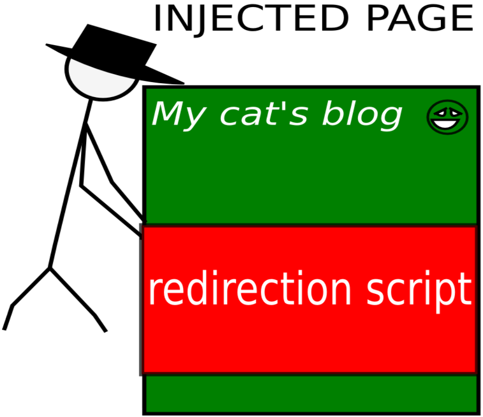
IEEE Security & Privacy 2012

# Finding malicious URLs

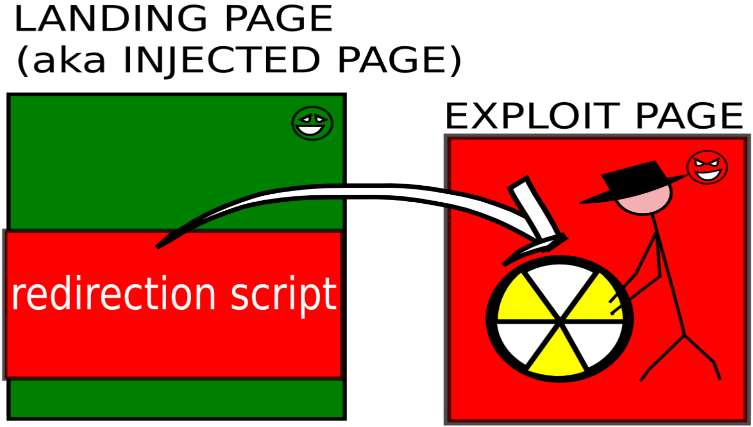


# Landing and exploit pages

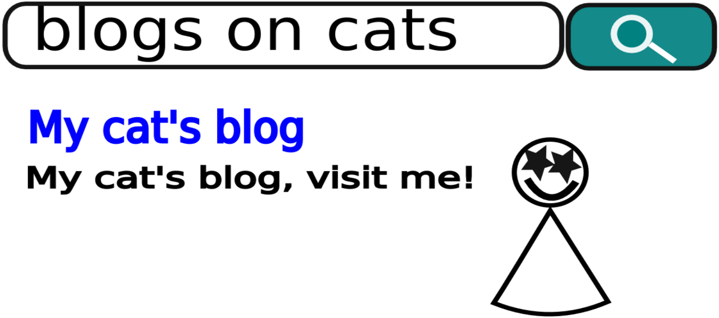
# Landing and exploit pages



# Landing and exploit pages



# Landing and exploit pages

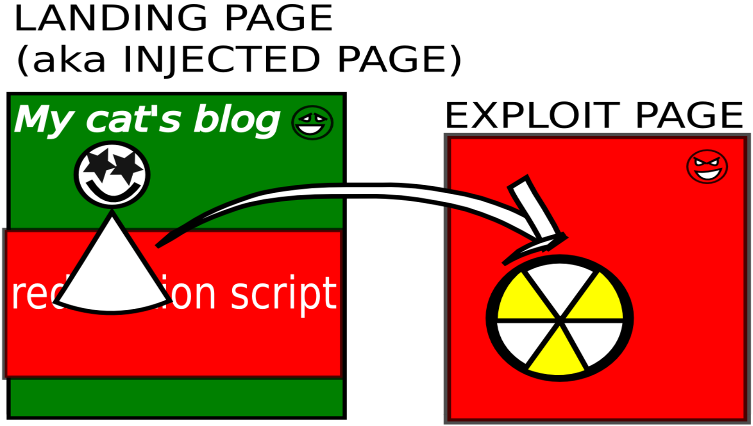


blogs on cats

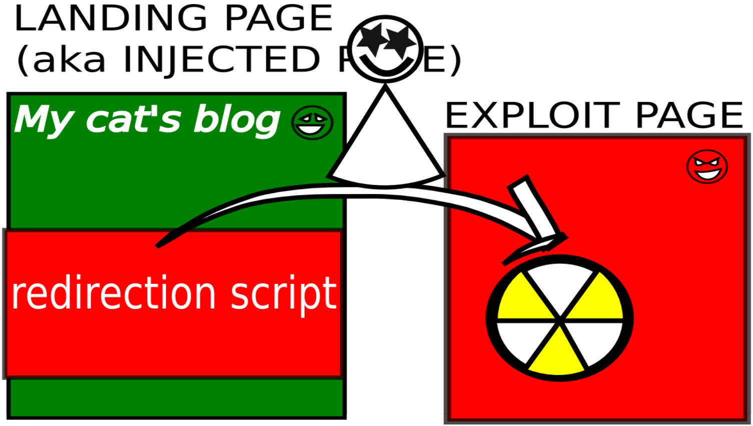
**My cat's blog**

**My cat's blog, visit me!**

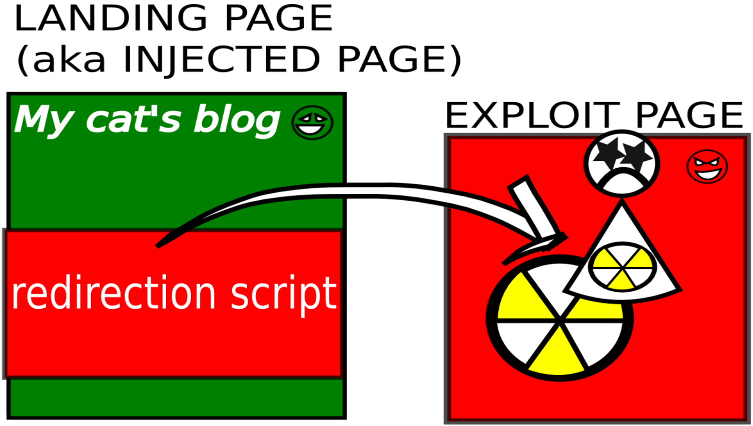
# Landing and exploit pages



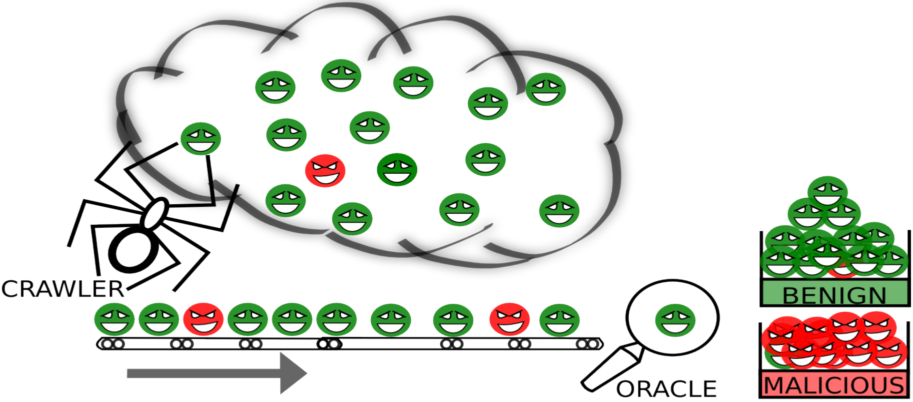
# Landing and exploit pages



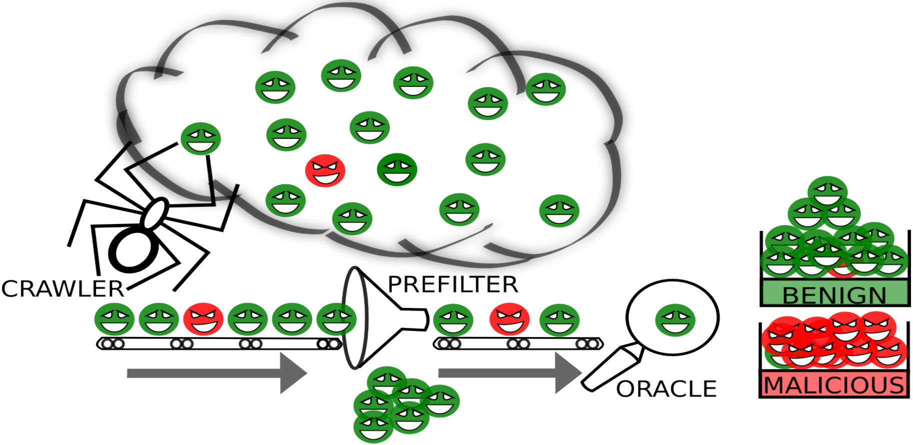
# Landing and exploit pages



# Finding malicious URLs



# Finding malicious URLs



# Finding malicious URLs is hard!

## Wepawet

- Over 120 thousand URLs analyzed per day by the oracle.
- Available online: http://wepawet.cs.ucsb.edu

## The problem

$0,138\%$ of the URLs reached with a random crawl are malicious

# Our goal

## Finding malicious URLs efficiently

WE HAVE

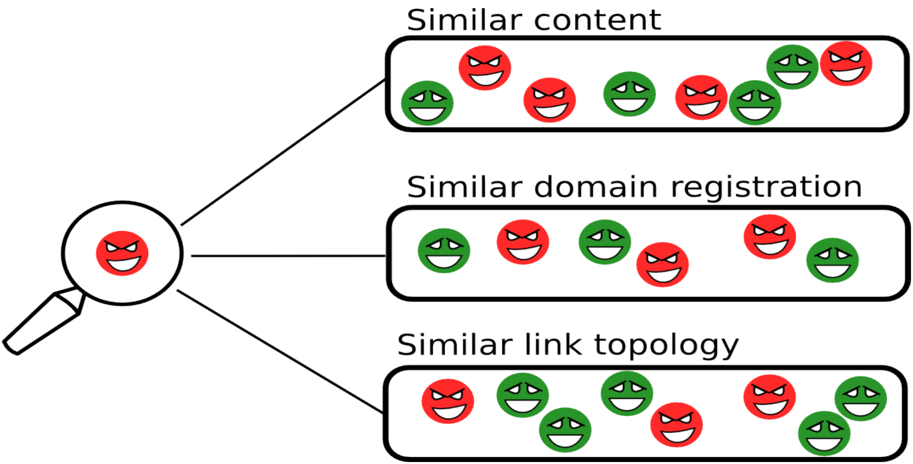
A set of pages that our oracle labeled malicious: our **EVIL SEED**

WE WANT

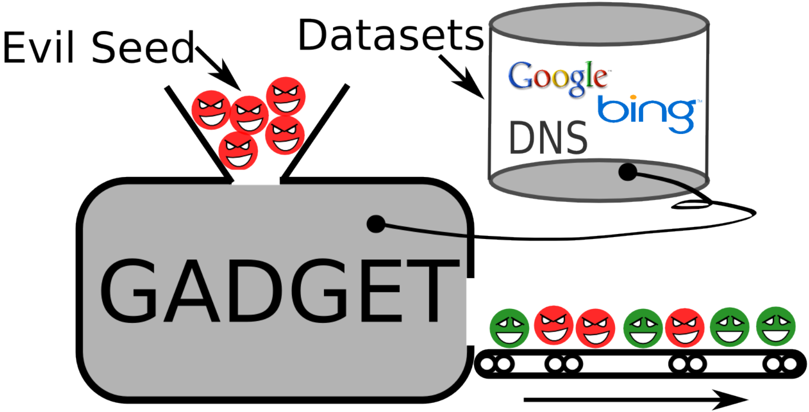
**More** landing pages

**More** exploit pages

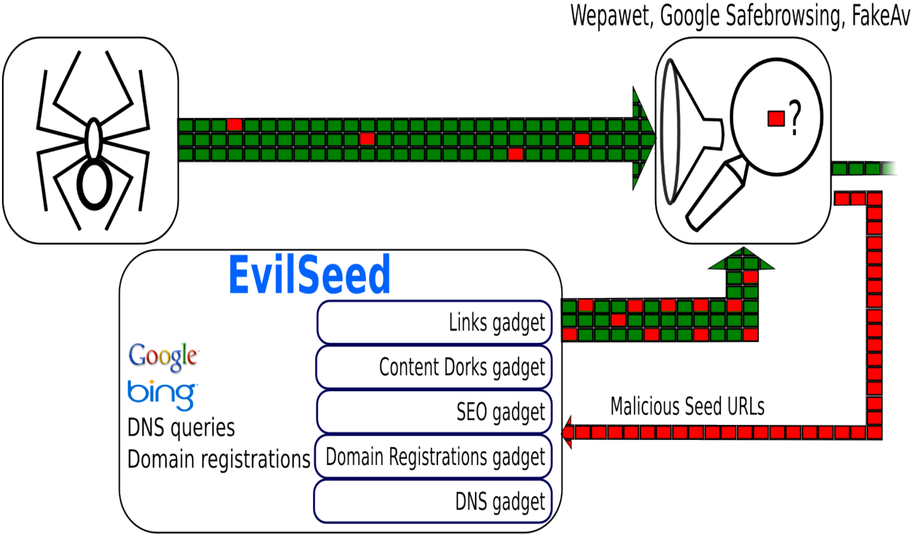
# What can a malicious URL tell us?



# Gadgets

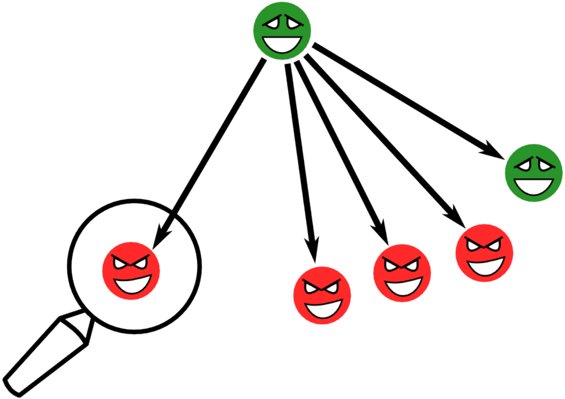


# EvilSeed


Wepawet, Google Safebrowsing, FakeAv
EvilSeed
Google
bing
DNS queries
Domain registrations
Links gadget
Content Dorks gadget
SEO gadget
Domain Registrations gadget
DNS gadget
Malicious Seed URLs

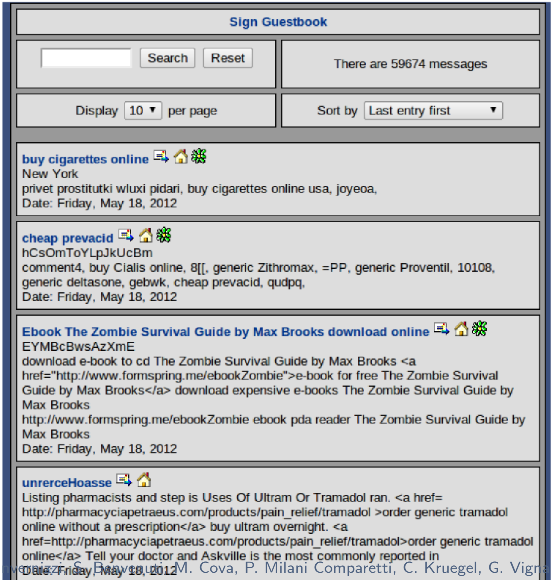
# Links gadget

Designed to locate *malware hubs*



Example query: `link:http://malicious-url.com`

# Links gadget



There are 59674 messages

Display 10 per page

Sort by Last entry first

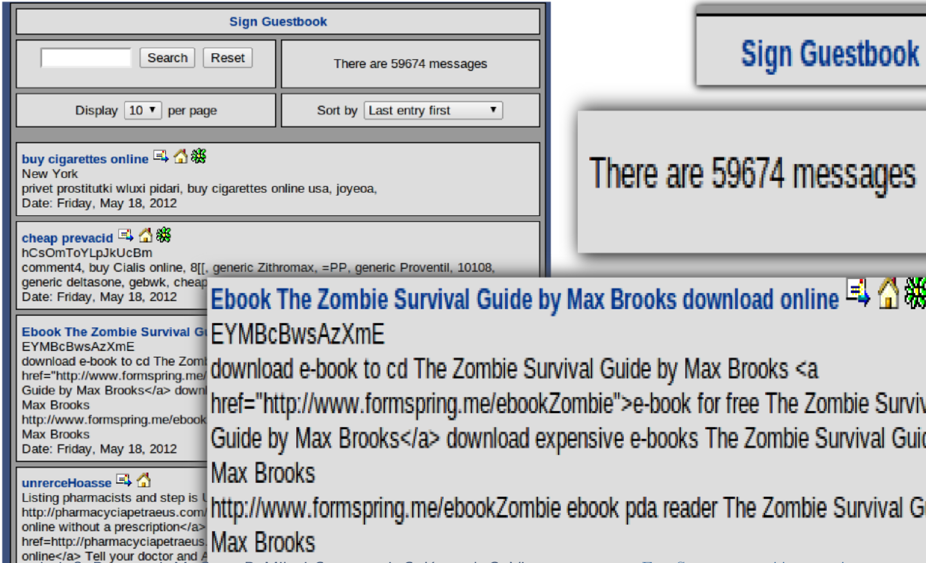
**buy cigarettes online**
New York
privet prostitutki wluxi pidari, buy cigarettes online usa, joyeoa,
Date: Friday, May 18, 2012

**cheap prevacid**
hCsOmToYLpJkUcBm
comment4, buy Cialis online, 8[[, generic Zithromax, =PP, generic Proventil, 10108, generic deltasone, gebwk, cheap prevacid, qudpq,
Date: Friday, May 18, 2012

**Ebook The Zombie Survival Guide by Max Brooks download online**
EYMBcBwsAzXmE
download e-book to cd The Zombie Survival Guide by Max Brooks <a href="http://www.formspring.me/ebookZombie">e-book for free The Zombie Survival Guide by Max Brooks</a> download expensive e-books The Zombie Survival Guide by Max Brooks http://www.formspring.me/ebookZombie ebook pda reader The Zombie Survival Guide by Max Brooks
Date: Friday, May 18, 2012

**unrerceHoasse**
Listing pharmacists and step is Uses Of Ultram Or Tramadol ran. <a href= http://pharmacyciapetraeus.com/products/pain_relief/tramadol >order generic tramadol online without a prescription</a> buy ultram overnight. <a href=http://pharmacyciapetraeus.com/products/pain_relief/tramadol>order generic tramadol online</a> Tell your doctor and Askville is the most commonly reported in
Date: Friday, May 18, 2012

# Links gadget



Sign Guestbook

There are 59674 messages

Display 10 per page

Sort by Last entry first

**buy cigarettes online**
New York
privet prostitutki wluxi pidari, buy cigarettes online usa, joyeoa,
Date: Friday, May 18, 2012

**cheap prevacid**
hCsOmToYLpJkUcBm
comment4, buy Cialis online, 8[[, generic Zithromax, =PP, generic Proventil, 10108,
generic deltasone, gebwk, cheap
Date: Friday, May 18, 2012

**Ebook The Zombie Survival Guide by Max Brooks download online**
EYMBcBwsAzXmE
download e-book to cd The Zombie Survival Guide by Max Brooks <a
href="http://www.formspring.me/ebookZombie">e-book for free The Zombie Surviv
Guide by Max Brooks</a> download expensive e-books The Zombie Survival Gui
Max Brooks
http://www.formspring.me/ebookZombie ebook pda reader The Zombie Survival G
Max Brooks
Date: Friday, May 18, 2012

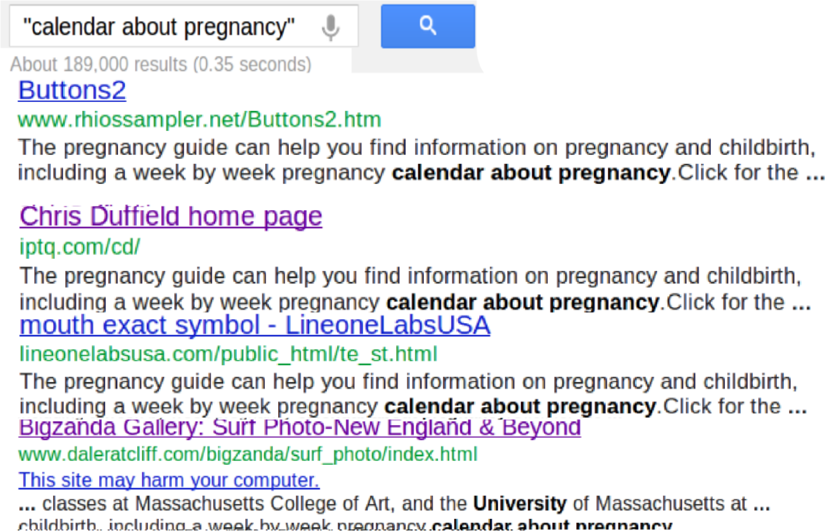
**unrerceHoasse**
Listing pharmacists and step is
http://pharmacyciapetraeus.com/
online without a prescription</a>
href=http://pharmacyciapetraeus.
online</a> Tell your doctor and
Date: Friday, May 18, 2012

# Content Dorks gadget

Creates signatures from the content of landing pages.
Two methods:

- n-gram extraction
- term-extraction (e.g., `cnn.com` yields: Eurozone recession, gay wedding, Facebook attack, graphic content)

# Content Dorks gadget



"calendar about pregnancy"

About 189,000 results (0.35 seconds)

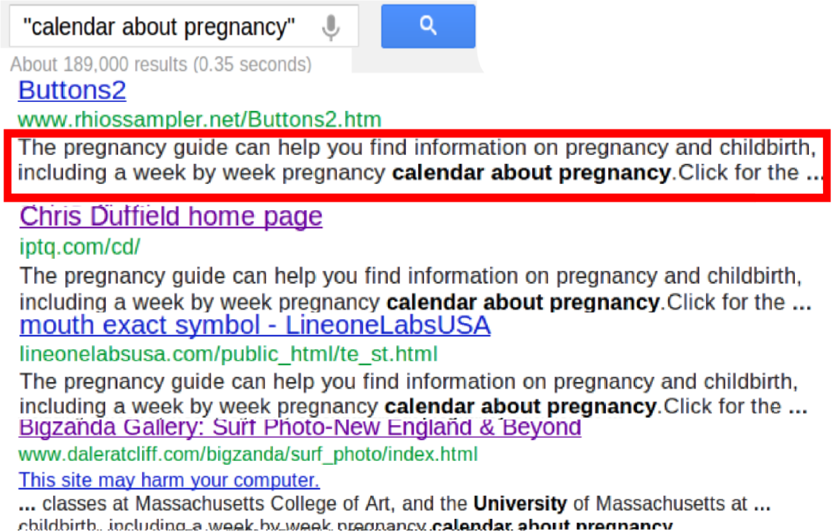
Buttons2
www.rhiossampler.net/Buttons2.htm
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Chris Duffield home page
iptq.com/cd/
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

mouth exact symbol - LineoneLabsUSA
lineonelabsusa.com/public_html/te_st.html
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Bigzanda Gallery: Surf Photo-New England & Beyond
www.daleratcliff.com/bigzanda/surf_photo/index.html
This site may harm your computer.
... classes at Massachusetts College of Art, and the **University** of Massachusetts at ...
childbirth, including a week by week pregnancy **calendar about pregnancy**

# Content Dorks gadget



"calendar about pregnancy"

About 189,000 results (0.35 seconds)

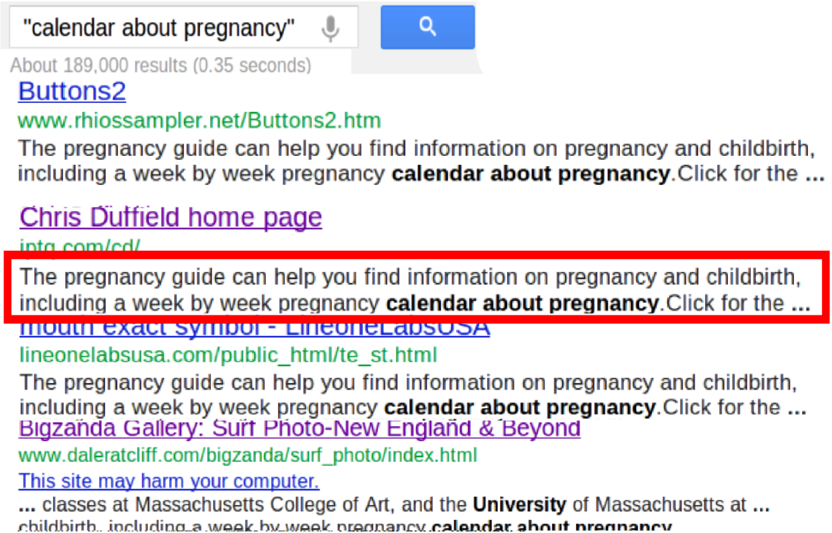
Buttons2
www.rhiossampler.net/Buttons2.htm
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Chris Duffield home page
iptq.com/cd/
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

mouth exact symbol - LineoneLabsUSA
lineonelabsusa.com/public_html/te_st.html
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Bigzanda Gallery: Surf Photo-New England & Beyond
www.daleratcliff.com/bigzanda/surf_photo/index.html
This site may harm your computer.
... classes at Massachusetts College of Art, and the **University** of Massachusetts at ...
childbirth, including a week by week pregnancy **calendar about pregnancy**

# Content Dorks gadget



"calendar about pregnancy"

About 189,000 results (0.35 seconds)

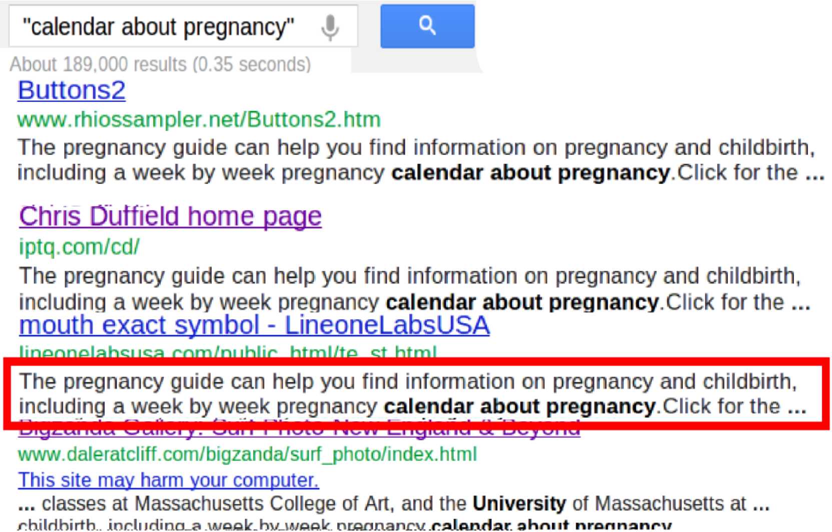
Buttons2
www.rhiossampler.net/Buttons2.htm
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Chris Duffield home page
iptg.com/cd/
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

mouth exact symbol - LineoneLabsUSA
lineonelabsusa.com/public_html/te_st.html
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Bigzanda Gallery: Surf Photo-New England & Beyond
www.daleratcliff.com/bigzanda/surf_photo/index.html
This site may harm your computer.
... classes at Massachusetts College of Art, and the **University** of Massachusetts at ...
childbirth, including a week by week pregnancy **calendar about pregnancy**

# Content Dorks gadget



"calendar about pregnancy"

About 189,000 results (0.35 seconds)

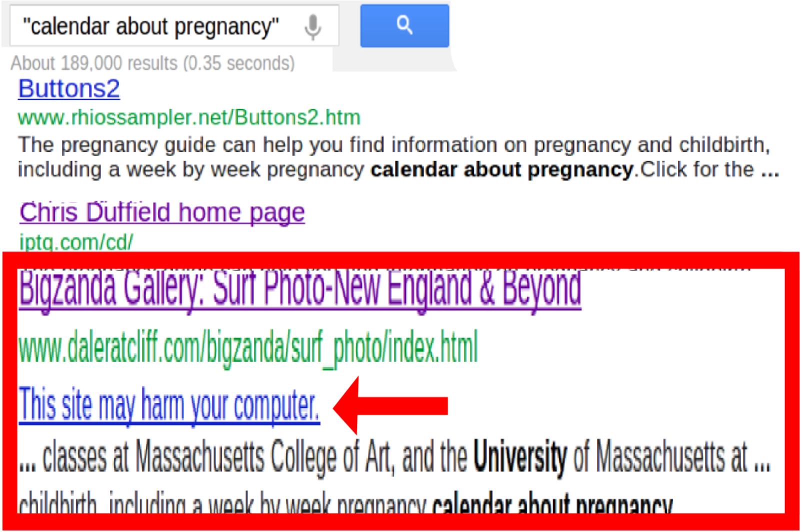
Buttons2
www.rhiossampler.net/Buttons2.htm
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Chris Duffield home page
iptq.com/cd/
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

mouth exact symbol - LineoneLabsUSA
lineonelabsusa.com/public_html/te_st.html
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Bigzanda Gallery: Surf Photo New England & Beyond
www.daleratcliff.com/bigzanda/surf_photo/index.html
This site may harm your computer.
... classes at Massachusetts College of Art, and the **University** of Massachusetts at ...
childbirth, including a week by week pregnancy **calendar about pregnancy**

# Content Dorks gadget



"calendar about pregnancy"

About 189,000 results (0.35 seconds)

Buttons2

www.rhiossampler.net/Buttons2.htm

The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Chris Duffield home page

iptq.com/cd/

Bigzanda Gallery: Surf Photo-New England & Beyond

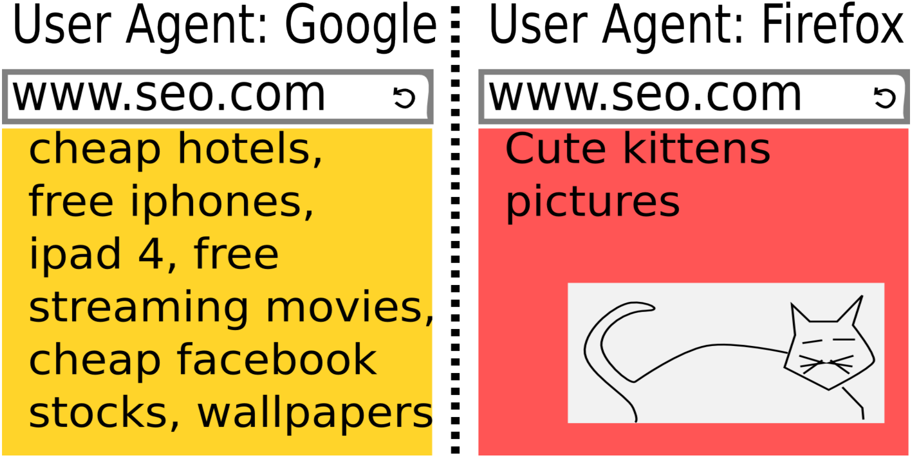
www.daleratcliff.com/bigzanda/surf_photo/index.html

This site may harm your computer.

... classes at Massachusetts College of Art, and the **University** of Massachusetts at ...

childbirth, including a week by week pregnancy **calendar about pregnancy**

# SEO gadget



User Agent: Google

www.seo.com

cheap hotels, free iphones, ipad 4, free streaming movies, cheap facebook stocks, wallpapers

User Agent: Firefox

www.seo.com

Cute kittens pictures

# SEO gadget

Expansion strategies:

- Find pages with similar content as Google sees it (e.g., query for `title:"free iphones"`)
- Find pages hosted on the same domain (e.g., query for `site:seo.com`)
- Follow links

# Domain Registrations gadget

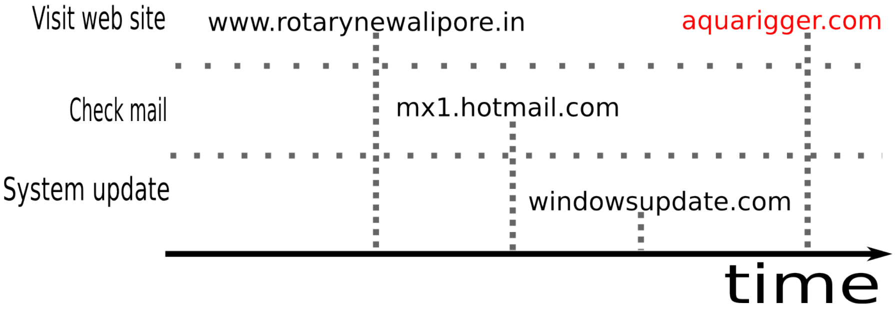
We know that:

- `http://a.com/exploit` is malicious.
- `a.com` has been registered moments before `b.com`

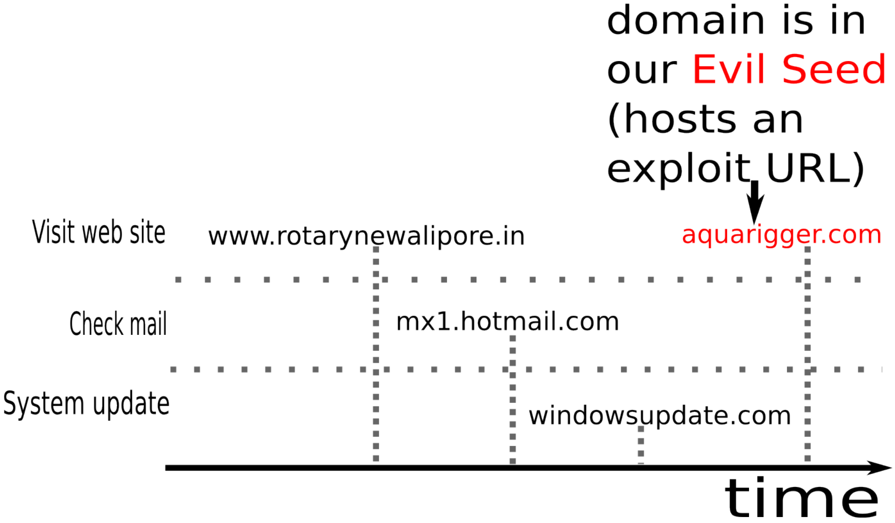
We suspect that:

- `http://b.com/exploit` is also malicious

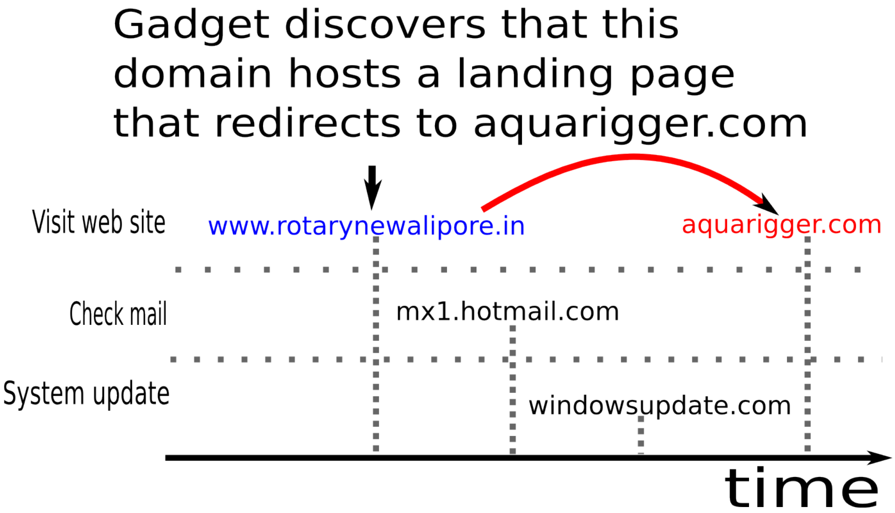
# DNS Queries gadget



# DNS Queries gadget



# DNS Queries gadget



# Evaluation metrics

$$\text{Toxicity} = \frac{\text{URLs classified as malicious}}{\text{URLs submitted to the Oracle}}$$

$$\text{Seed Expansion} = \frac{\text{malicious URLs found by }\textsc{EvilSeed}}{\text{seed size}}$$

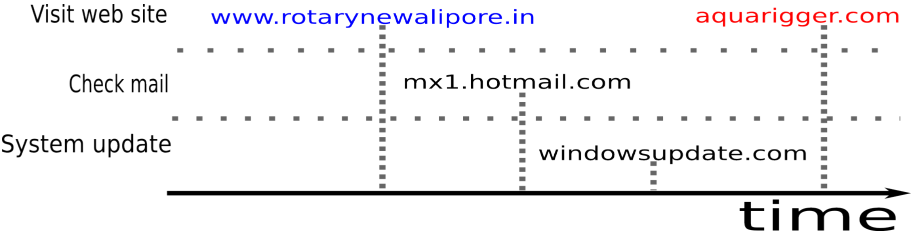
# Online evaluation: URLs

| Source | Seed | Analyzed | Malicious | Toxicity | Expansion |
|---|---|---|---|---|---|
| **Crawler w/ Prefilter** | | 437,251 | 604 | **0.14%** | |
| EvilSeed | | | | | |
| Links | 604 | 71,272 | 1,097 | 1.53% | 1.81 |
| SEO | 604 | 312 | 16 | 5.12% | 0.02 |
| Keywords | 604 | 13,896 | 477 | 3.43% | 0.78 |
| Ngrams | 604 | 140,660 | 1,446 | 1.02% | 2.39 |
| Total | | 226,140 | 3,036 | **1.34%** | **5.02** |
| **Web Search** | | | | | |
| Random Strings | | 24,137 | 68 | **0.28%** | |
| Random Dictionary | | 27,242 | 107 | **0.39%** | |
| Trending Topics | | 8,051 | 27 | **0.33%** | |
| Manual Dorks | | 4,506 | 17 | **0.37%** | |

# Online evaluation: domains

| **Source** | **Seed** | **Analyzed** | **Malicious** | **Toxicity** | **Expansion** |
|---|---|---|---|---|---|
| **Crawler w/ Prefilter** | | 53,445 | 98 | **0.18%** | |
| EvilSeed | | | | | |
| Links | 98 | 7,664 | 107 | 1.39% | 1.09 |
| SEO | 98 | 7 | 5 | 71.42% | 0.07 |
| Keywords | 98 | 3,245 | 119 | 3.66% | 1.22 |
| Ngrams | 98 | 33,510 | 263 | 0.78% | 2.68 |
| Total | | 44,426 | 494 | **1.12%** | **5.04** |
| **Web Search** | | | | | |
| Random Strings | | 4,227 | 16 | **0.37%** | |
| Random Dictionary | | 9,285 | 35 | **0.37%** | |
| Trending Topics | | 1,768 | 8 | **0.45%** | |
| Manual Dorks | | 3,032 | 13 | **0.42%** | |

# DNS evaluation

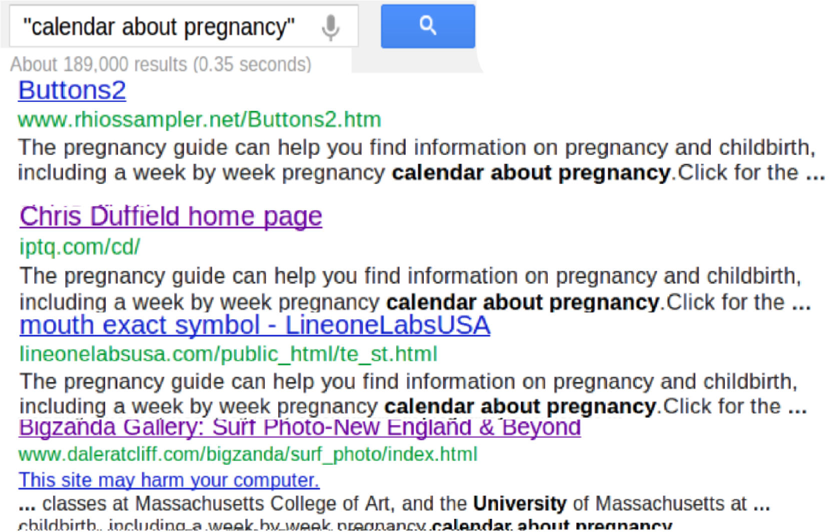


Data:

- 377,472,280 DNS resolutions
- 115 malicious seeds

Resulting in 3.5% toxicity, 1.48 seed expansion

# EvilSeed for Search Engines



"calendar about pregnancy"

About 189,000 results (0.35 seconds)

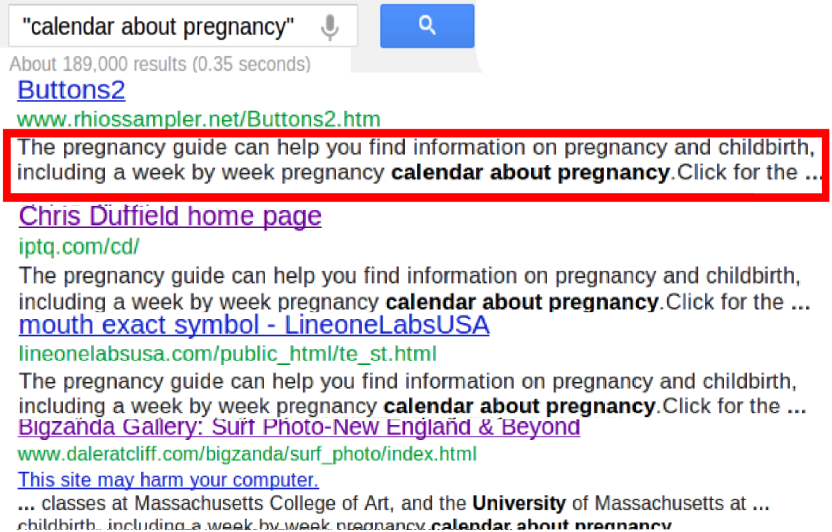
Buttons2
www.rhiossampler.net/Buttons2.htm
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Chris Duffield home page
iptq.com/cd/
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

mouth exact symbol - LineoneLabsUSA
lineonelabsusa.com/public_html/te_st.html
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Bigzanda Gallery: Surf Photo-New England & Beyond
www.daleratcliff.com/bigzanda/surf_photo/index.html
This site may harm your computer.
... classes at Massachusetts College of Art, and the **University** of Massachusetts at ...
childbirth, including a week by week pregnancy **calendar about pregnancy**

# EvilSeed for Search Engines



"calendar about pregnancy"

About 189,000 results (0.35 seconds)

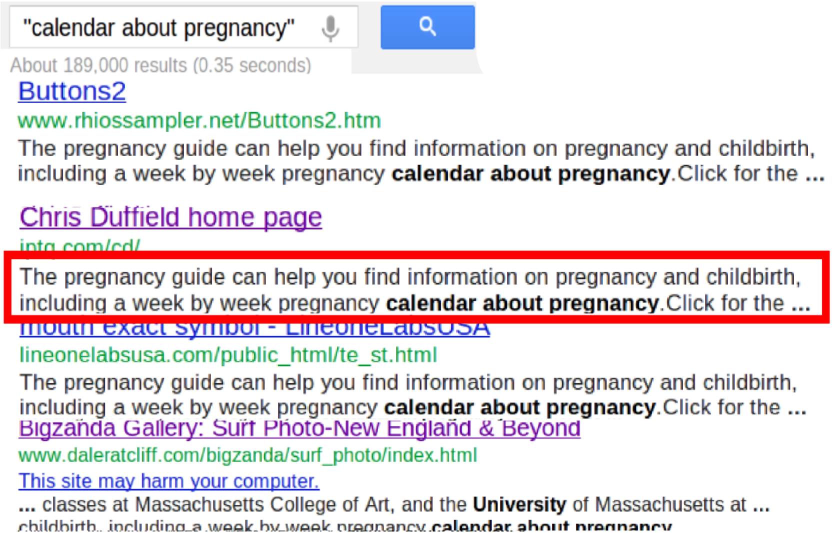
Buttons2
www.rhiossampler.net/Buttons2.htm
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Chris Duffield home page
iptq.com/cd/
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

mouth exact symbol - LineoneLabsUSA
lineonelabsusa.com/public_html/te_st.html
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Bigzanda Gallery: Surf Photo-New England & Beyond
www.daleratcliff.com/bigzanda/surf_photo/index.html
This site may harm your computer.
... classes at Massachusetts College of Art, and the **University** of Massachusetts at ...
childbirth, including a week by week pregnancy **calendar about pregnancy**

# EvilSeed for Search Engines



"calendar about pregnancy"

About 189,000 results (0.35 seconds)

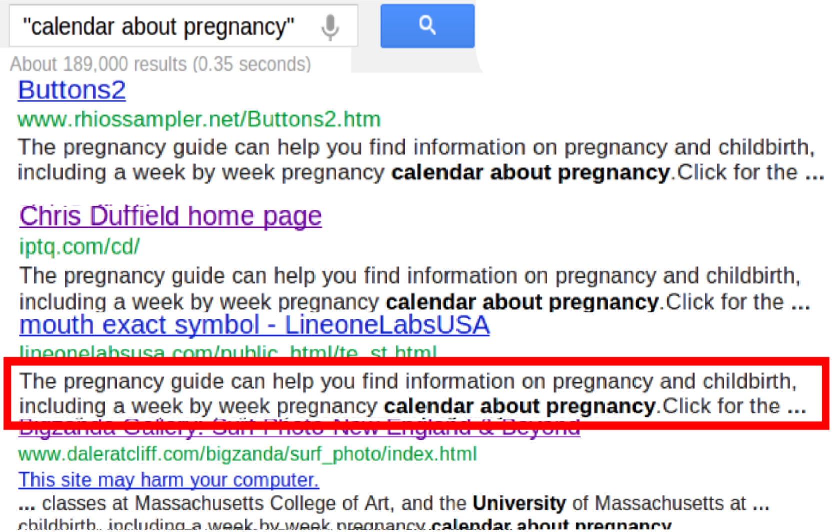
Buttons2
www.rhiossampler.net/Buttons2.htm
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Chris Duffield home page
iptq.com/cd/
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

mouth exact symbol - LineoneLabsUSA
lineonelabsusa.com/public_html/te_st.html
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Bigzanda Gallery: Surf Photo-New England & Beyond
www.daleratcliff.com/bigzanda/surf_photo/index.html
This site may harm your computer.
... classes at Massachusetts College of Art, and the **University** of Massachusetts at ...
childbirth, including a week by week pregnancy **calendar about pregnancy**

# EvilSeed for Search Engines



"calendar about pregnancy"

About 189,000 results (0.35 seconds)

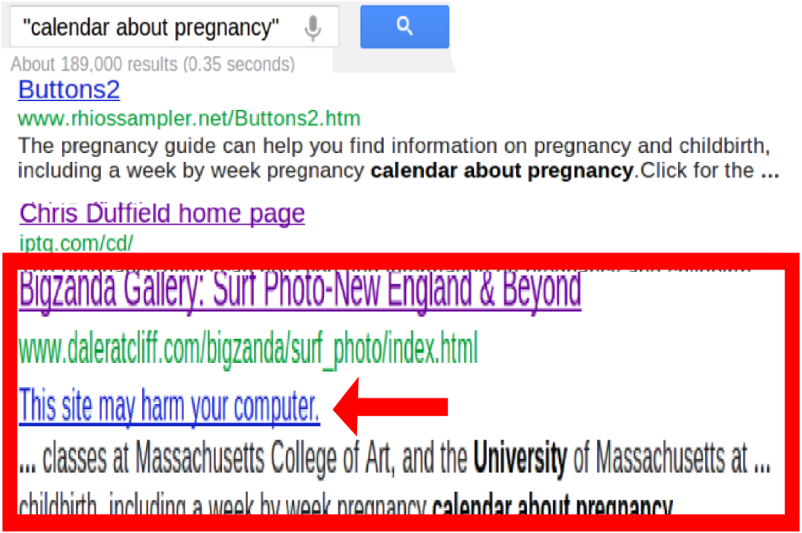
Buttons2
www.rhiossampler.net/Buttons2.htm
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Chris Duffield home page
iptq.com/cd/
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

mouth exact symbol - LineoneLabsUSA
lineonelabsusa.com/public_html/te_st.html
The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Bigzanda Gallery: Surf Photo New England & Beyond
www.daleratcliff.com/bigzanda/surf_photo/index.html
This site may harm your computer.
... classes at Massachusetts College of Art, and the **University** of Massachusetts at ...
childbirth, including a week by week pregnancy **calendar about pregnancy**

# EvilSeed for Search Engines



"calendar about pregnancy"

About 189,000 results (0.35 seconds)

Buttons2

www.rhiossampler.net/Buttons2.htm

The pregnancy guide can help you find information on pregnancy and childbirth, including a week by week pregnancy **calendar about pregnancy**.Click for the ...

Chris Duffield home page

iptq.com/cd/

Bigzanda Gallery: Surf Photo-New England & Beyond

www.daleratcliff.com/bigzanda/surf_photo/index.html

This site may harm your computer.

... classes at Massachusetts College of Art, and the **University** of Massachusetts at ...

childbirth, including a week by week pregnancy **calendar about pregnancy**

# Conclusions

- Finding malicious urls is important to protect the users

# Conclusions

- Finding malicious urls is important to protect the users, but it's hard

# Conclusions

- Finding malicious urls is important to protect the users, but it's hard
- It's critical to generate feeds with high toxicity ($\Rightarrow$ high efficiency)

# Conclusions

- Finding malicious urls is important to protect the users, but it's hard
- It's critical to generate feeds with high toxicity (⇒ high efficiency)
- We designed EvilSeed, a guided search approach that is a ten-fold efficency improvement over crawling

# Conclusions

- Finding malicious urls is important to protect the users, but it's hard
- It's critical to generate feeds with high toxicity (⇒ high efficiency)
- We designed EvilSeed, a guided search approach that is  a ten-fold efficency improvement over crawling
- But crawling is needed nontheless, to generate the evil seed

# Thanks!

`http://bit.ly/evilseed`

invernizzi@cs.ucsb.edu

# Thanks!

`http://bit.ly/evilseed`

invernizzi@cs.ucsb.edu

# SEO evaluation

| | URLs | |
|---|---|---|
| **Cloaking Seeds** | 248 | |
| **Visited** | 1,219,090 | |
| **Analyzed** | 12,063 | |
| **Malicious** | 11,384 | |
| **Toxicity** | **94.37%** | (Crawler's: 0.14%) |
| $\frac{\text{Malicious}}{\text{Visited}}$ | **0.93%** | |
| **Expansion** | **45.90** | |