

V I S U

A L I Z A

T I O N F O R

I T - S E C U R I T Y

L. Aaron Kaplan (kaplan@cert.at)

<http://CERT.AT>

OVERVIEW



- Motivation
- Target Group
- 5 Minutes of design background for techies
- Tools
- DNSviz and Flows

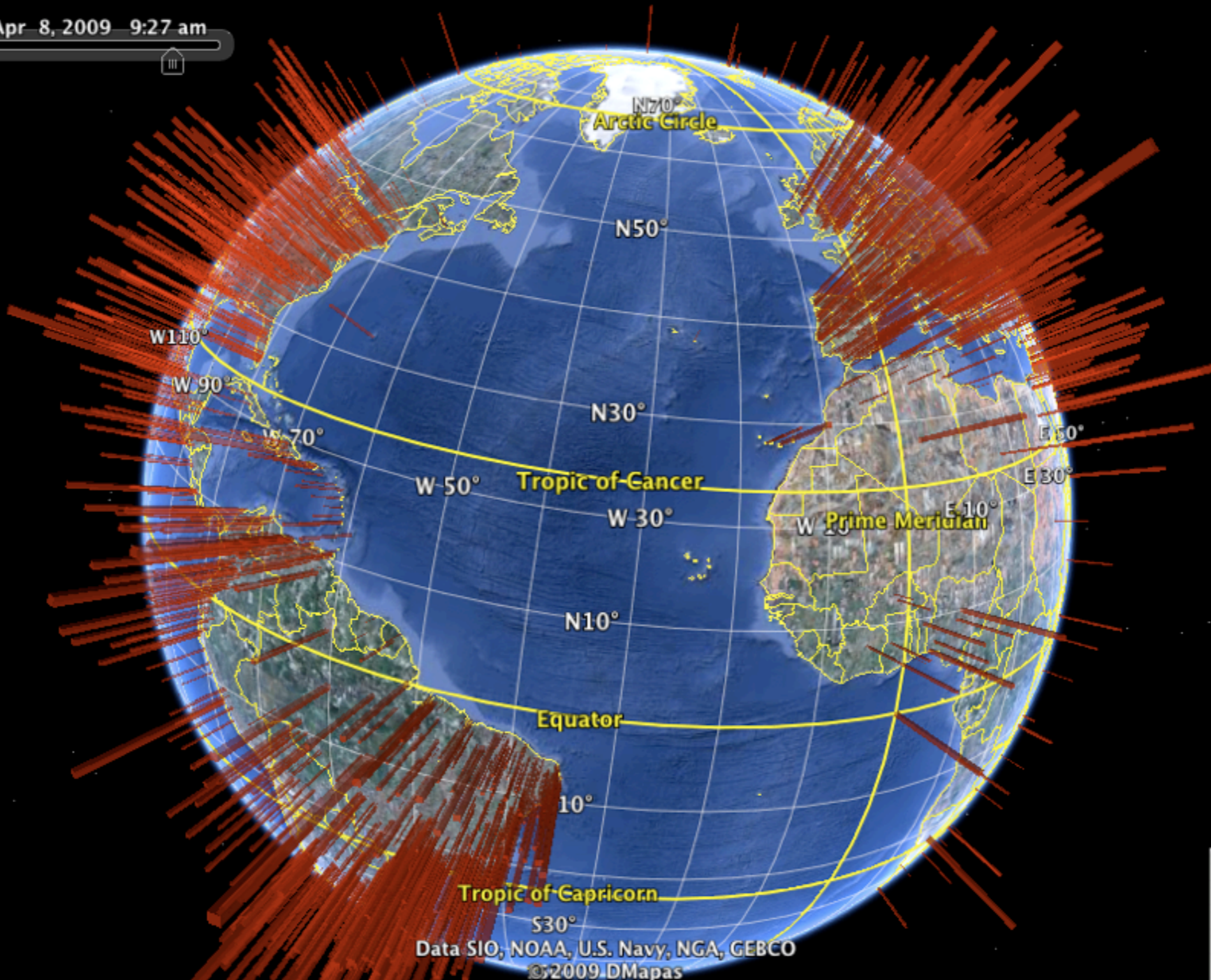
Why am I holding this talk?



Why am I holding this talk?



Apr 8, 2009 9:27 am



Data SIO, NOAA, U.S. Navy, NGA, GEBCO
©2009 DMapas



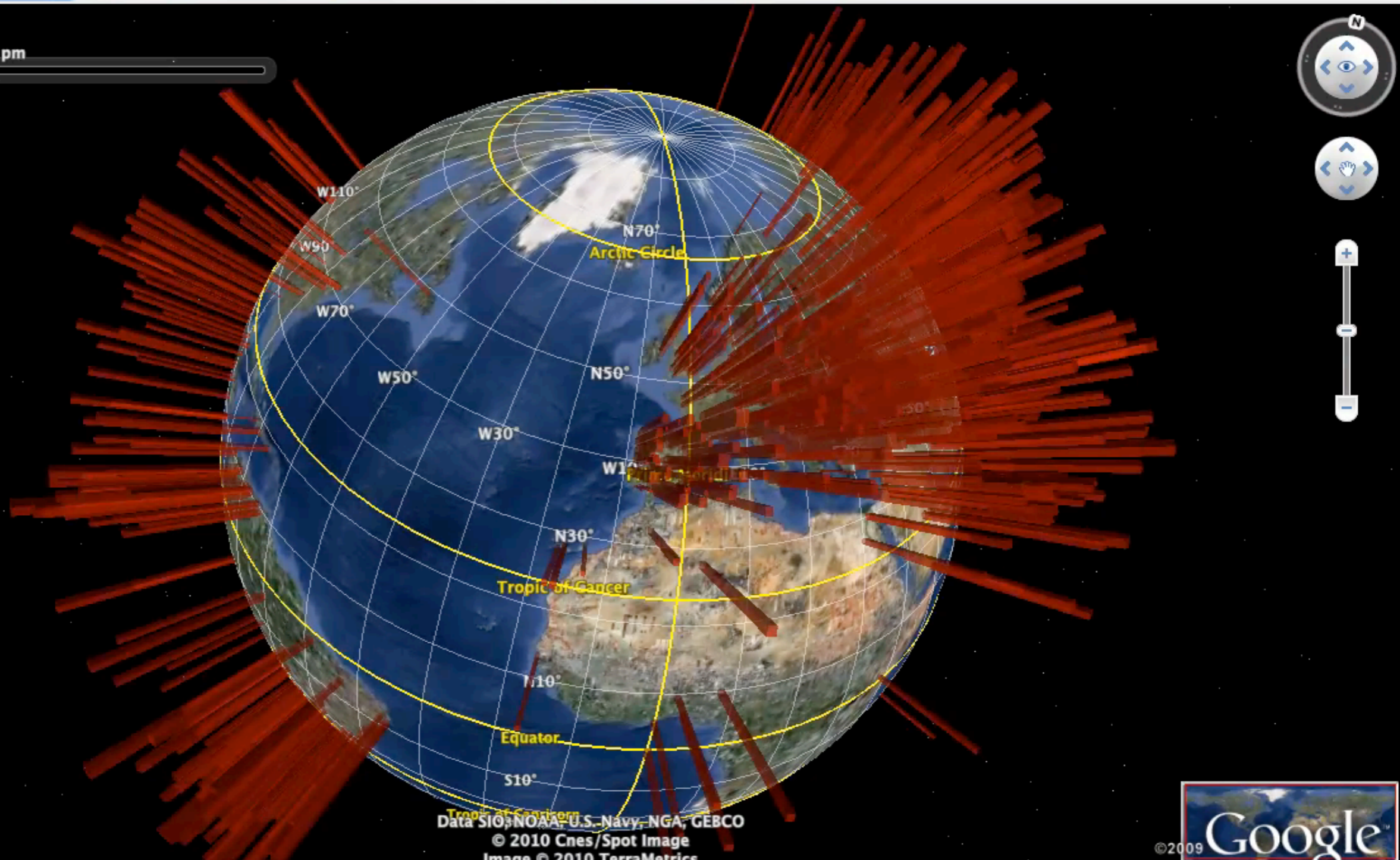
Why am I holding this talk?



Google Earth File Edit View Tools Add Window Help Google Earth

Google Earth toolbar icons

Mar 31, 2009 11:58 pm



Data SIO, NOAA, U.S. Navy, NGA, GEBCO
© 2010 Cnes/Spot Image
Image © 2010 TerraMetrics



Why am I holding this talk?

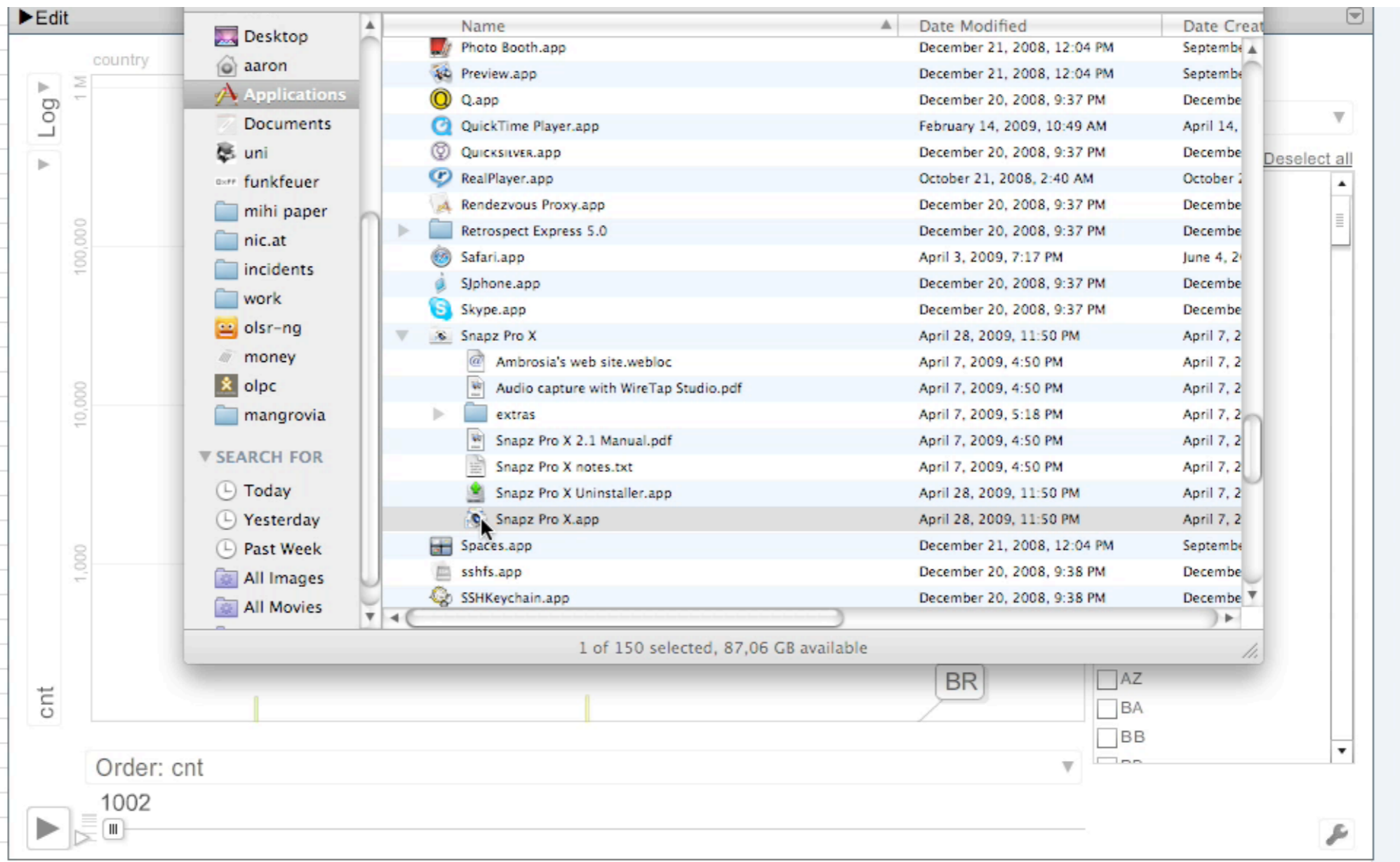


Why am I holding this talk?



Why am I holding this talk?

country	date	cnt	cnt2
MX	1000	4	1
KR	1001	2	1
BR	1002	2	1
VE	1003	1	1
CA	1004	1	1
KH	1005	1	1
RO	1006	1	1
TT	1007	1	1
RU	1008	1	1
PK	1009	1	1
AR	1010	1	1
ES	1011	1	1
CN	1012	1	1
US	1013	1	1
PH	1014	1	1
EC	1015	1	1
CN	1016	762166	1
BR	1017	644646	1
RU	1018	582881	1
VN	1019	255845	1
UA	1020	245734	1
IN	1021	233908	1
KR	1022	230627	1
ID	1023	210525	1
IT	1024	177121	1
ES	1025	171711	1
PH	1026	153319	1
US	1027	150700	1
TH	1028	132307	1
RO	1029	111490	1
TW	1030	103898	1
MY	1031	93251	1
AR	1032	84714	1
VE	1033	84454	1
MX	1034	79976	1



The screenshot shows a Mac OS X desktop environment. In the foreground, a file browser window is open, displaying a list of applications and files. The list includes items like Photo Booth.app, Preview.app, Q.app, QuickTime Player.app, Quicksilver.app, RealPlayer.app, Rendezvous Proxy.app, Retrospect Express 5.0, Safari.app, Siphone.app, Skype.app, Snapz Pro X, and Spaces.app. The Snapz Pro X application is selected. The file browser also shows a sidebar with a navigation pane and a search pane. In the background, a spreadsheet application is visible, showing a table with columns for country, date, cnt, and cnt2. The spreadsheet data is the same as the table provided in the previous block. The spreadsheet interface includes a toolbar with a play button and a filter icon, and a status bar at the bottom showing 'Order: cnt' and '1002'. The bottom of the screen shows a navigation bar with 'Add Sheet', 'Sheet 1', and 'Gadget1' buttons, and a dropdown menu for 'country'.

- What will you get out of it?
 - Quick IT–security visualization skills with 5 tools
 - Understanding the basic visualization cycle
 - Initial good results in < 1 day
 - Really good results in 10+ years ;–)

CERT.at, Austria



CERT.at, Austria



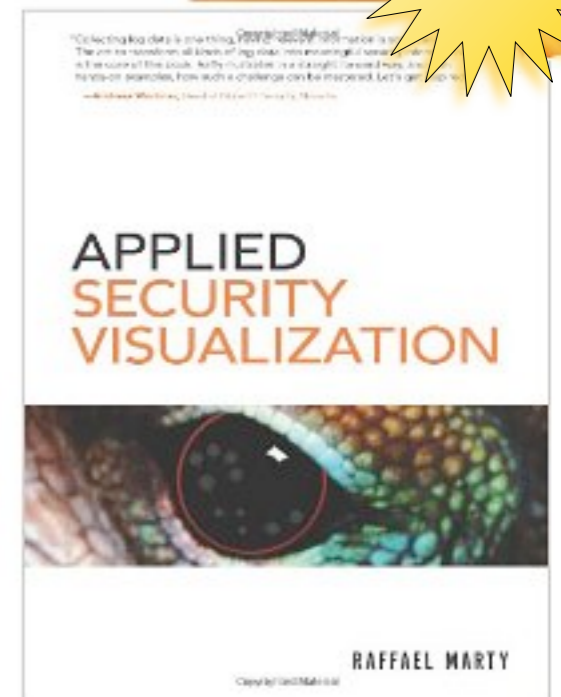
- CERT.at is part of **NIC.at**, the Austrian domain registry. CERT.at is the national CERT
- Austria is in **Europe**, but we definitely like the friends from **AUSCert** and down under
- Vienna, Austria is where we will have our next **FIRST conference 2011**
- **German** is spoken in Austria
- Our neighbouring countries are: Hungary, Slovenia, Germany, Switzerland, Slovakia, Czech Republic, Italy, Liechtenstein

- Motivation
- Target Group
- 5 Minutes of design background for techies
- Tools
- DNSviz and Flows

Motivation



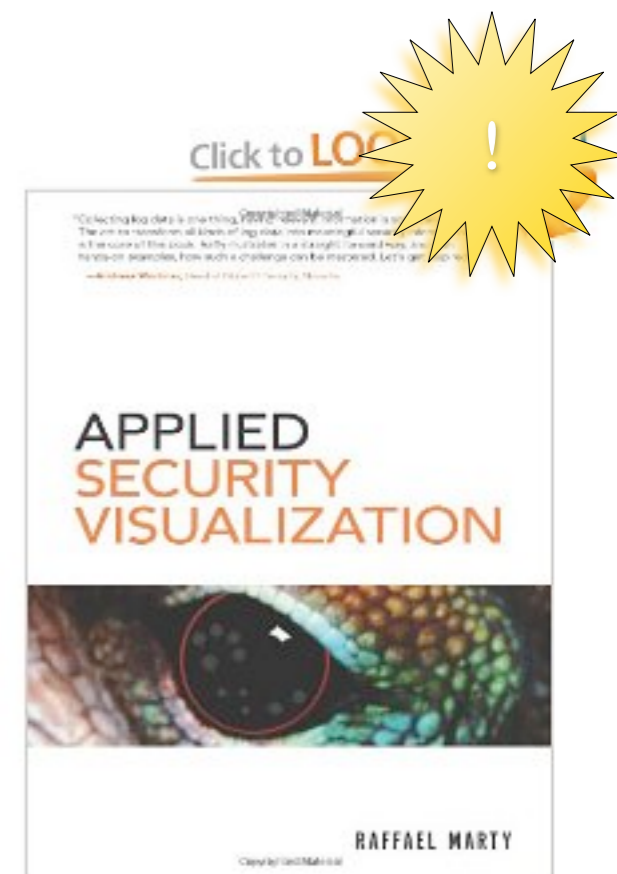
Click to LOG



Motivation

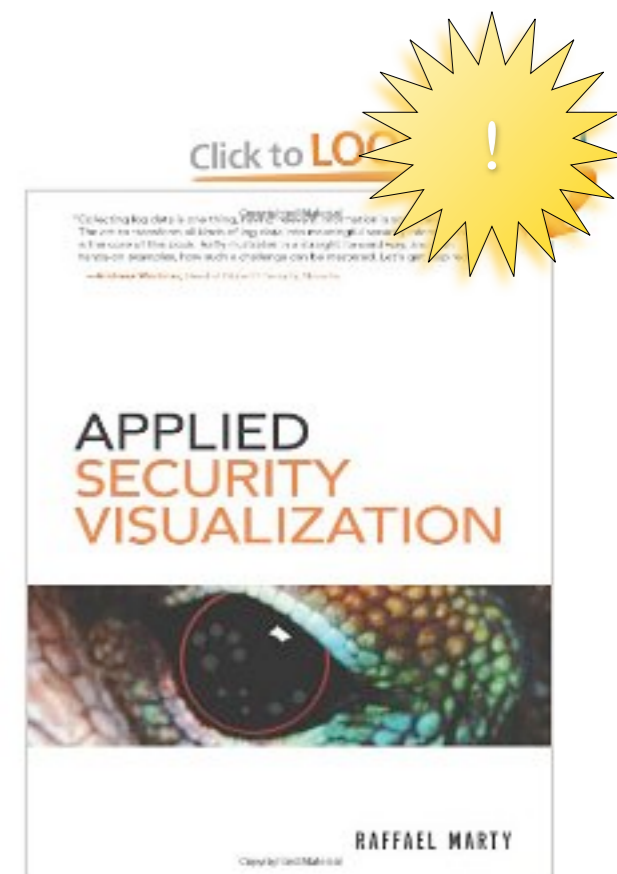


- “A picture is worth 1000 log records” (R. Marty)



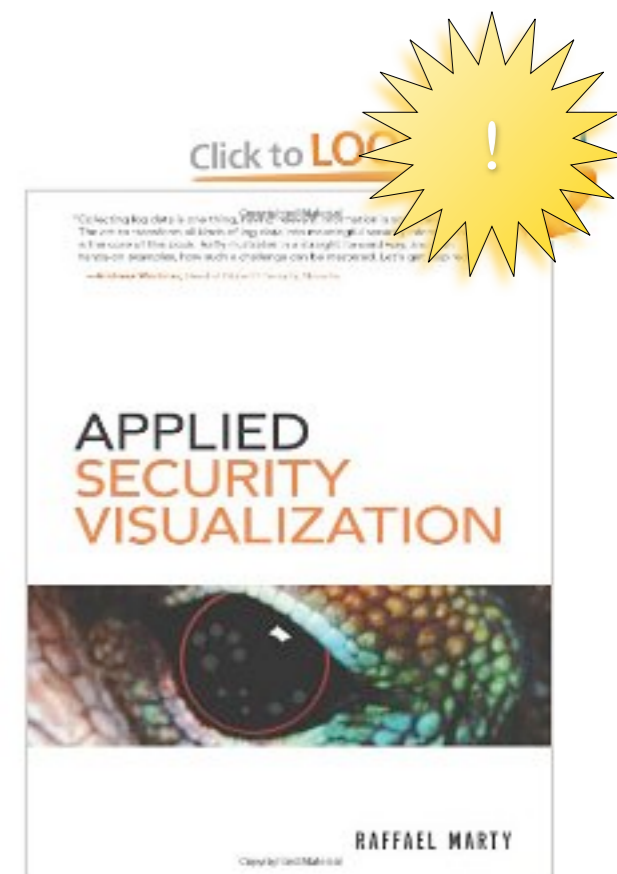
Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion



Motivation

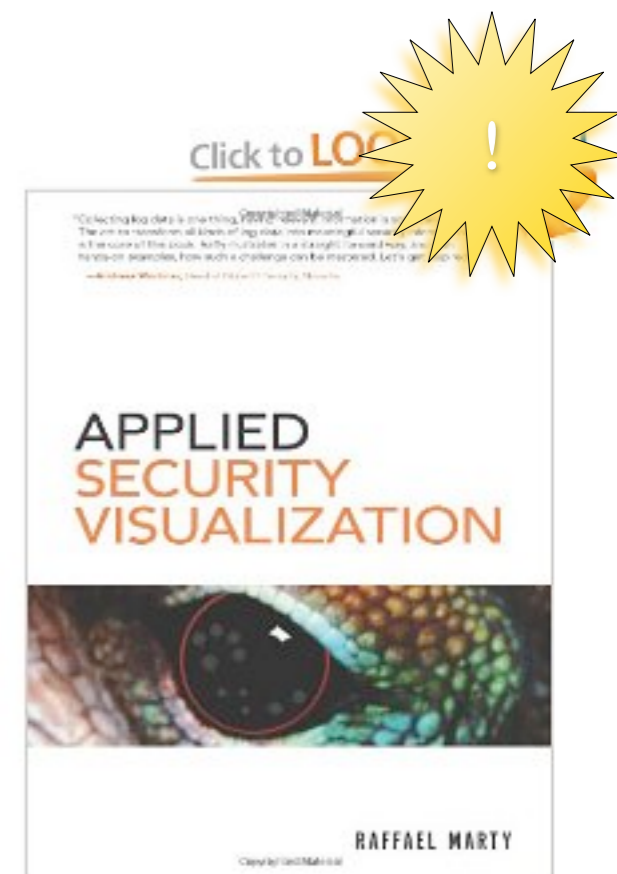
- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**



Motivation

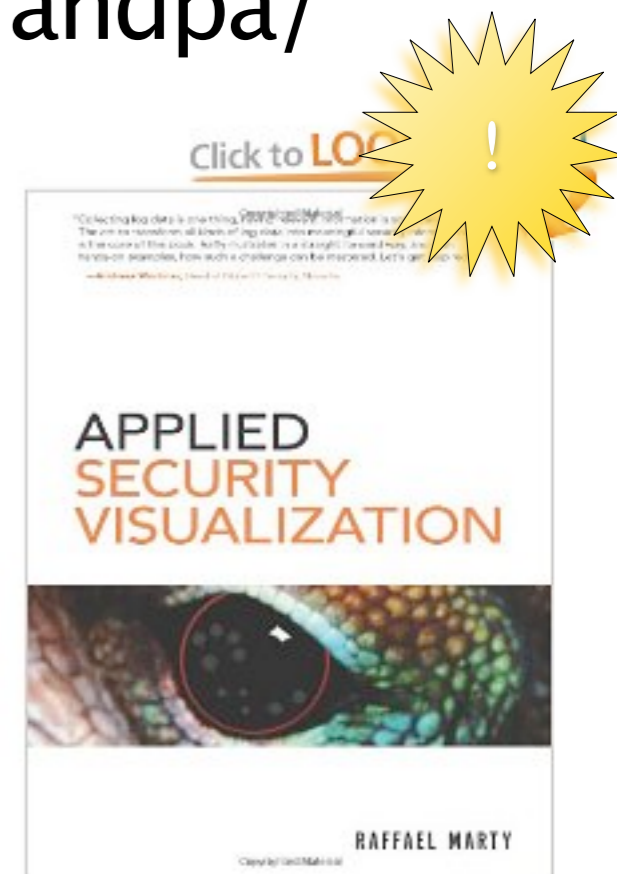


- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**
- People “get it”



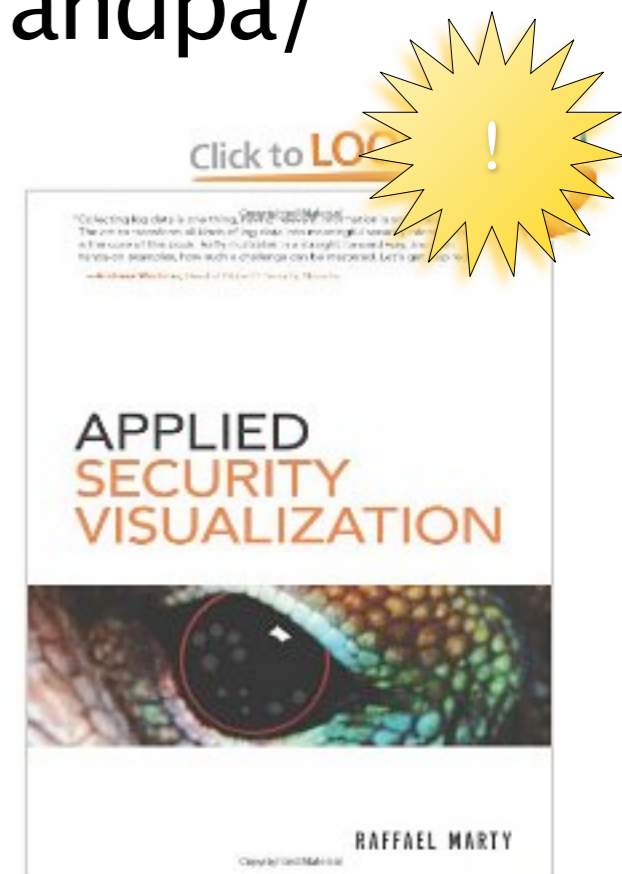
Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**
- People “get it”
- Visualization can explain it all to your grandpa/
father/mother/partner...



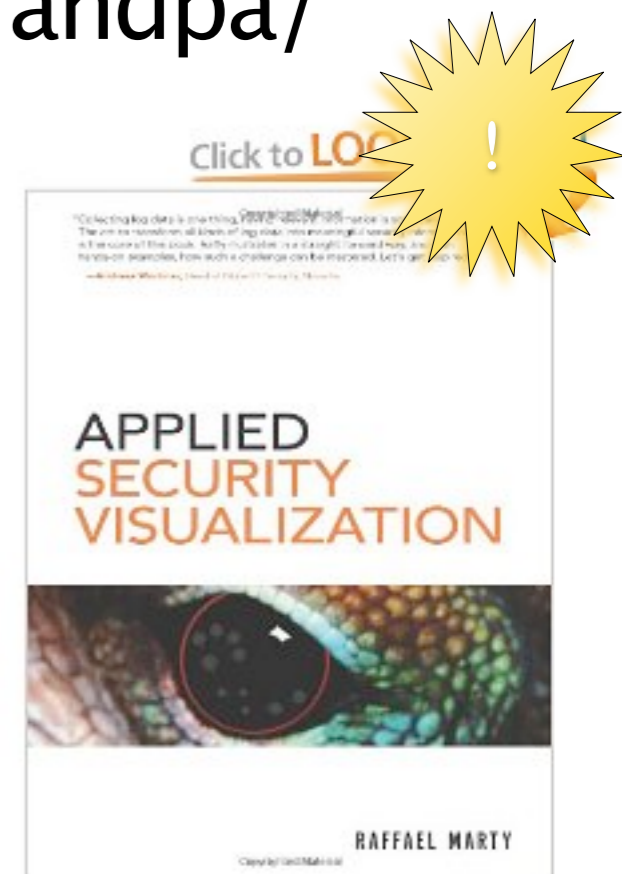
Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**
- People “get it”
- Visualization can explain it all to your grandpa/
father/mother/partner...
- ... and helps them understand that
you need to save the internet first



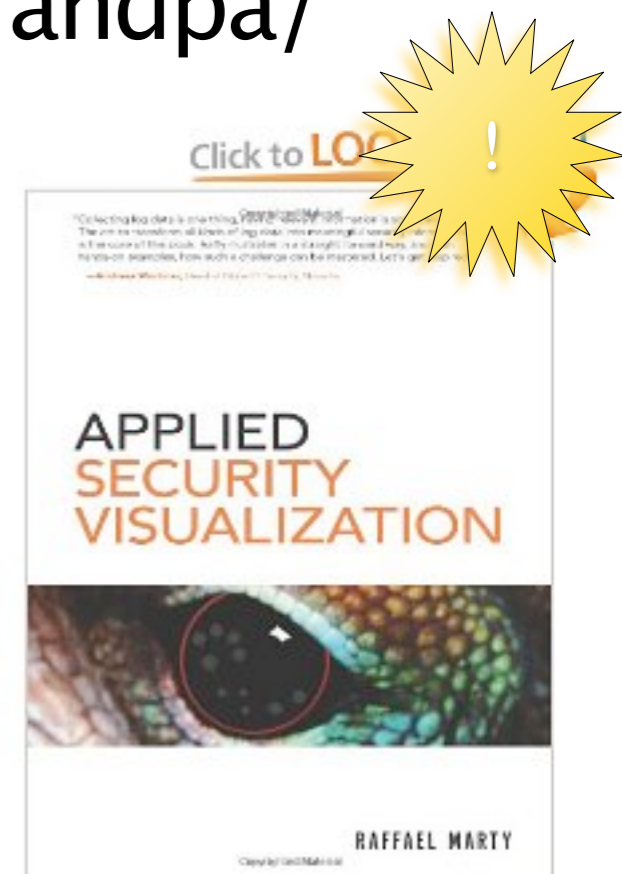
Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**
- People “get it”
- Visualization can explain it all to your grandpa/
father/mother/partner...
- ... and helps them understand that
you need to save the internet first
- gives new insights -> explore data



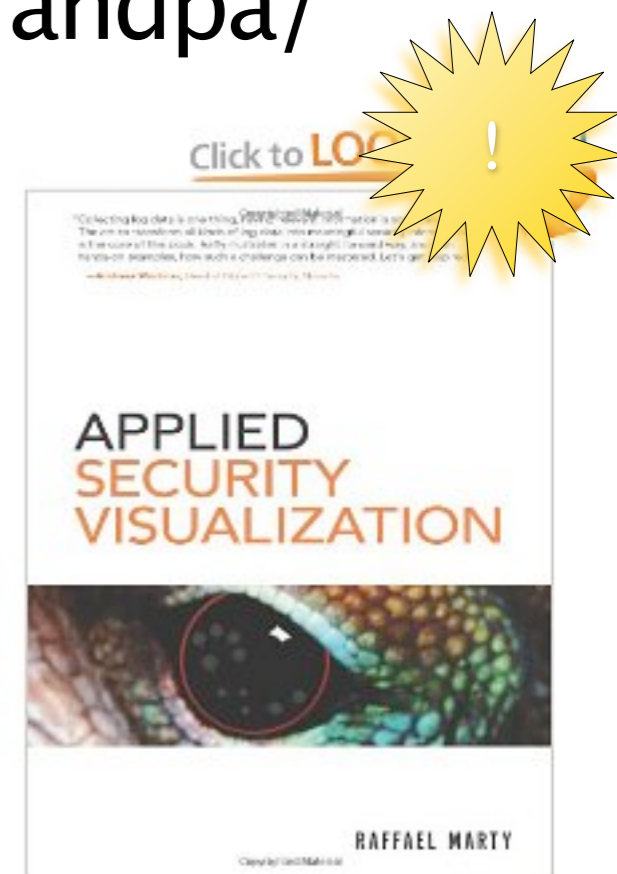
Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**
- People “get it”
- Visualization can explain it all to your grandpa/
father/mother/partner...
- ... and helps them understand that
you need to save the internet first
- gives new insights -> explore data
- gives us an overview



Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**
- People “get it”
- Visualization can explain it all to your grandpa/
father/mother/partner...
- ... and helps them understand that
you need to save the internet first
- gives new insights -> explore data
- gives us an overview
- **sells your services**



OVERVIEW



- Motivation
- Target Group
- 5 Minutes of design background for techies
- Tools
- DNSviz and Flows

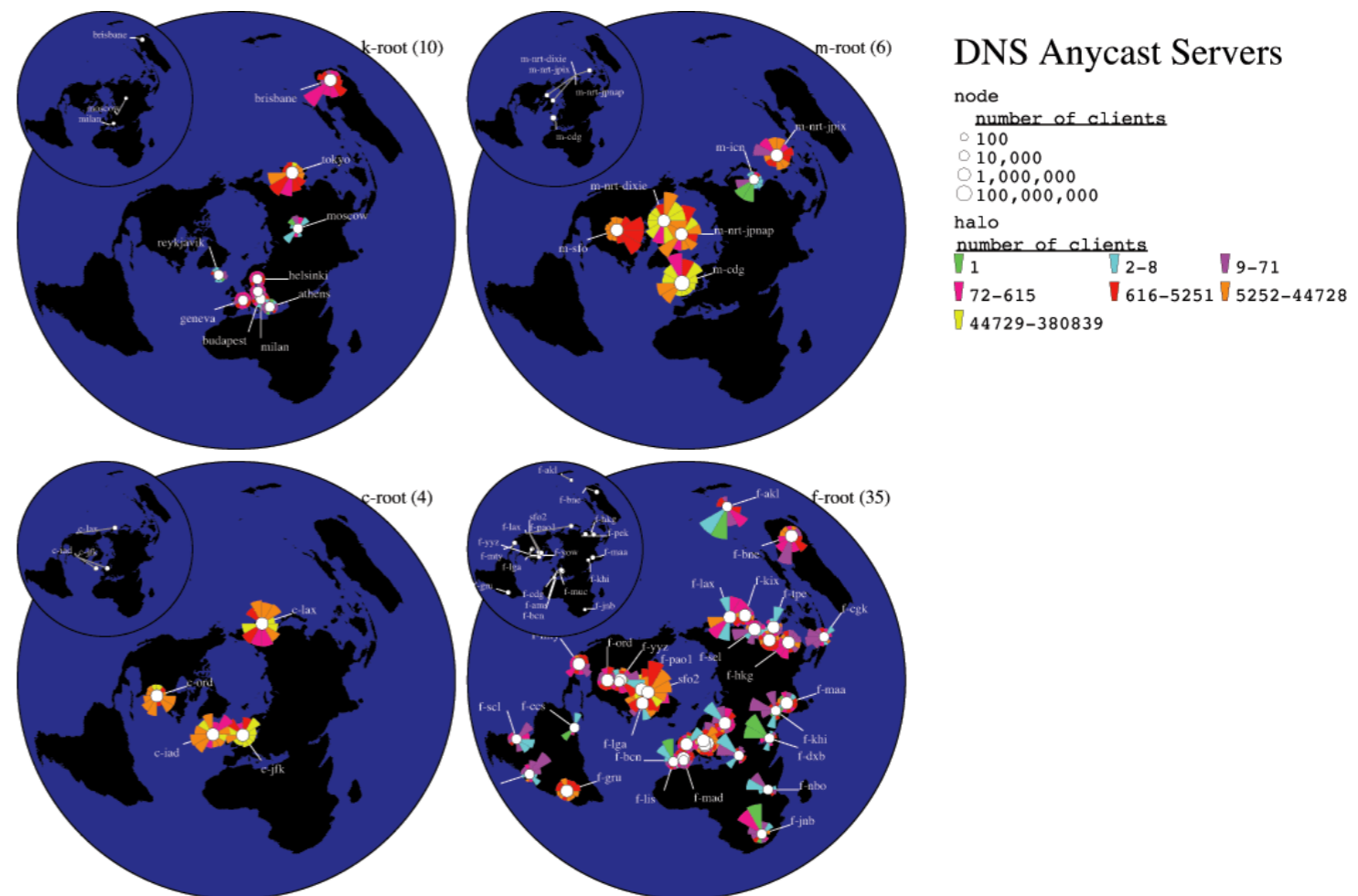
Target groups



- Users
- Management, Sales, Politicians
- Operational staff
- Researchers

Target groups

- Users
- Management, Sales, Politicians
- Operational staff
- Researchers



source: CAIDA.org

Target groups

- Users
- Management, Sales, Politicians
- Operational staff
- Researchers

Conficker Eye Chart



- Motivation
- Target Group
- 5 Minutes of design background for techies
- Tools
- DNSviz and Flows

Some design background

- One of the leading persons in the field right now:
Edward Tufte
- Learned a lot from **Otto Neurath**: “Isotypes” in Vienna in the early 1900s
- First invention of “**icons**”.
Idea: educate the illiterate working class population in Europe w.r.t basic economics relationships



Otto Neurath's Isotype

Kriegsverluste

1525 Bauernkrieg

Bauern

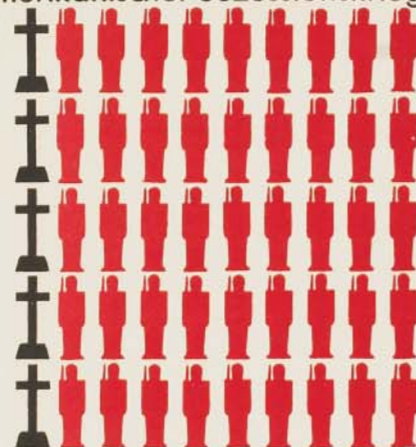


1812 Zug Napoleons nach Russland



1861-65 Amerikanischer Sezessionskrieg

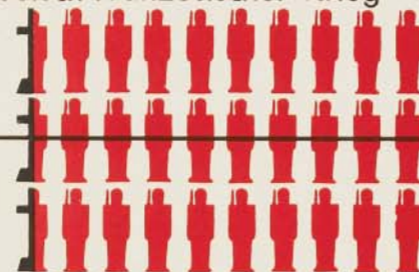
Nordstaaten



1870-71 Deutsch-Französischer Krieg

Deutsche

Franzosen



Jede einzelne Figur 50 000 Mann Überlebende

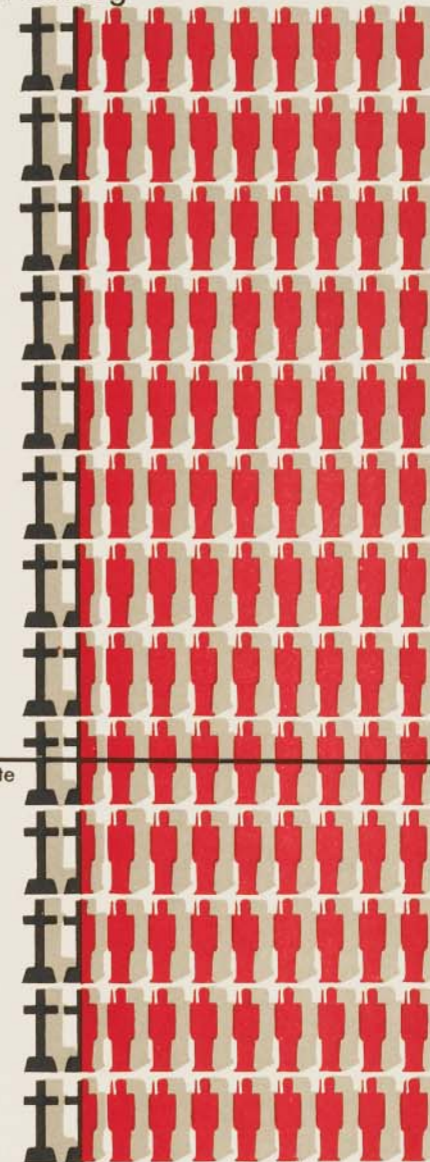
Jedes einzelne Kreuz 50 000 Tote

Die Heere sind auf 500 000 Mann abgerundet, die Verluste auf 5 Prozent

1914-18 Weltkrieg

Entente

Zentralmächte



Jede Figur mit Grau 500 000 Mann Überlebende

Jedes Kreuz mit Grau 500 000 Tote

Die Heere sind auf 5 Millionen Mann abgerundet, die Verluste auf 5 Prozent

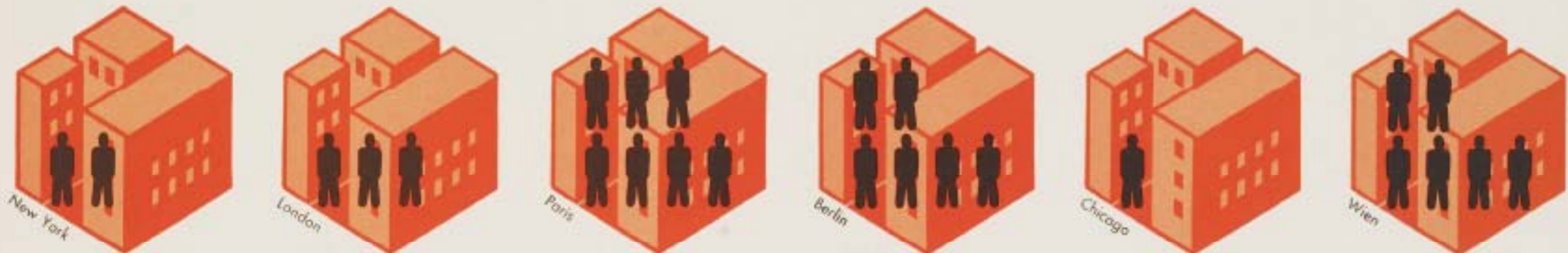
Angefertigt für das Bibliographische Institut AG., Leipzig
Gesellschafts- und Wirtschaftsmuseum in Wien ©

Otto Neurath's Isotype

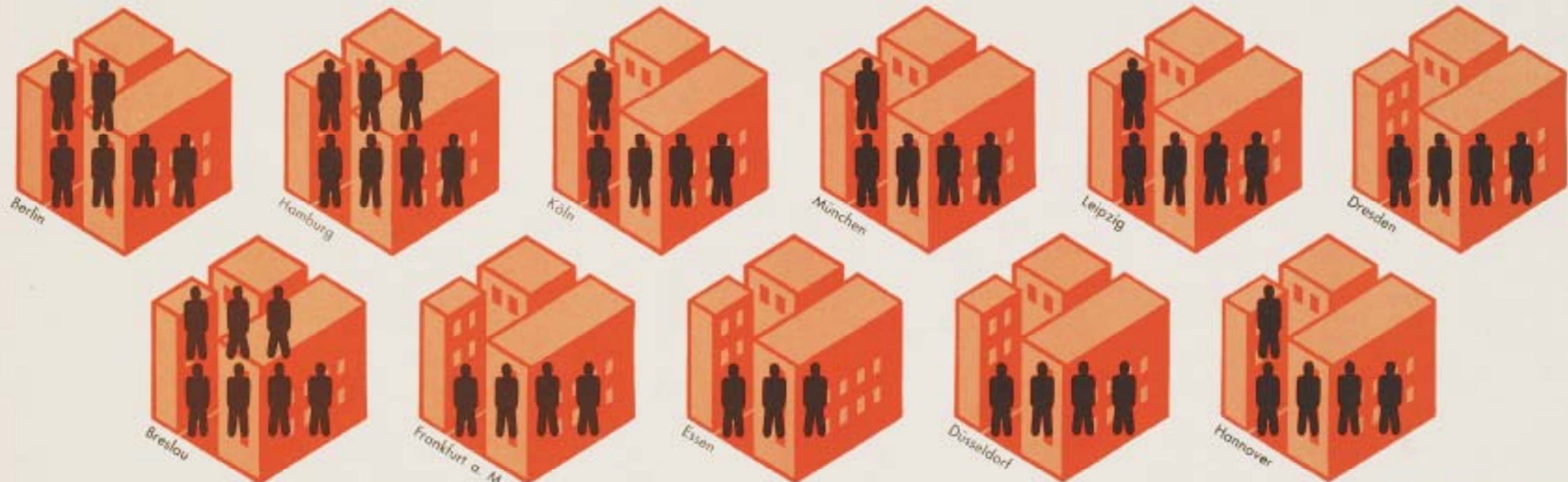
Wohndichte in Großstädten

Bewohner auf 200 m² verbauter Fläche (Gebäudegrundstücke einschl. Strassen, ausschl. grosser Parkanlagen)

Einige Weltstädte



Die deutschen Großstädte über 400.000 Einwohner



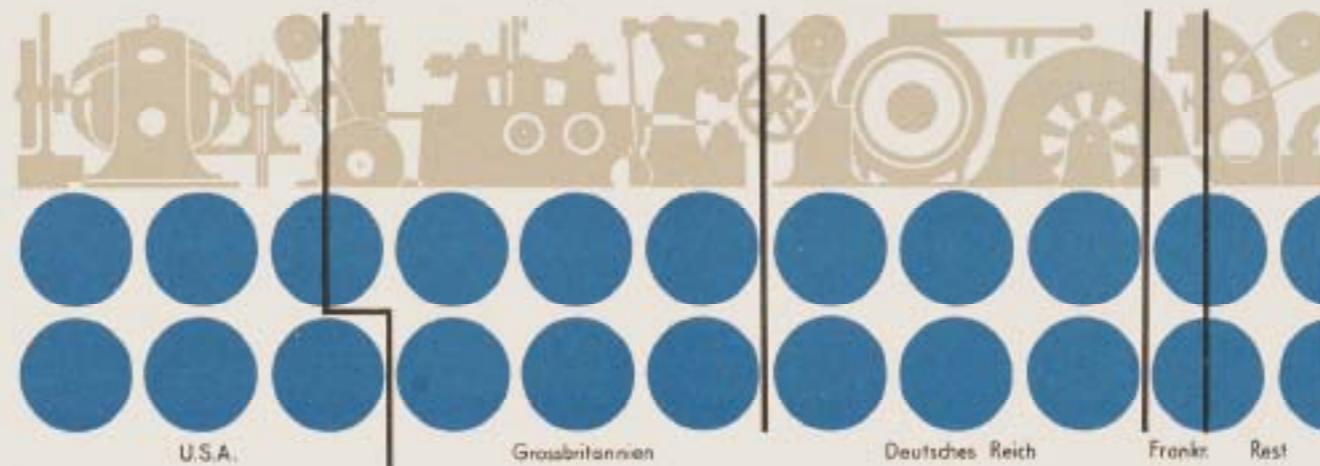
Anordnung der Städte nach ihrer Grösse. Anfang 1929

Angehört für das Bildgraphische Institut AG, Leipzig
Gesellschafts- und Wirtschaftswissenschaften in Wien ©

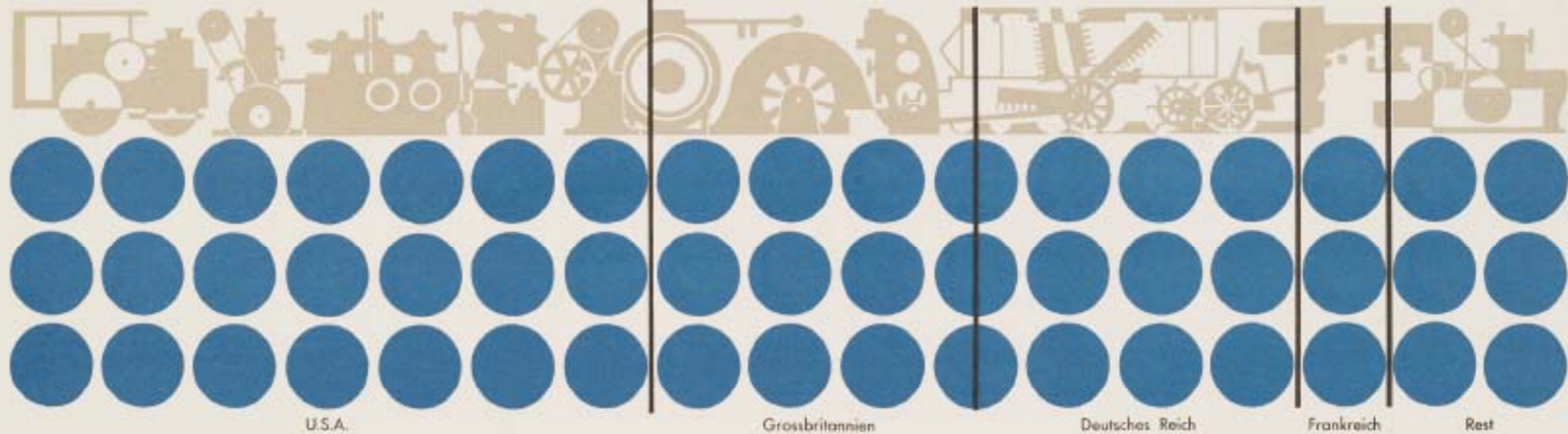
Otto Neurath's Isotype

Maschinenausfuhr vor dem Krieg und jetzt

Durchschnitt
1909 - 1913



1928



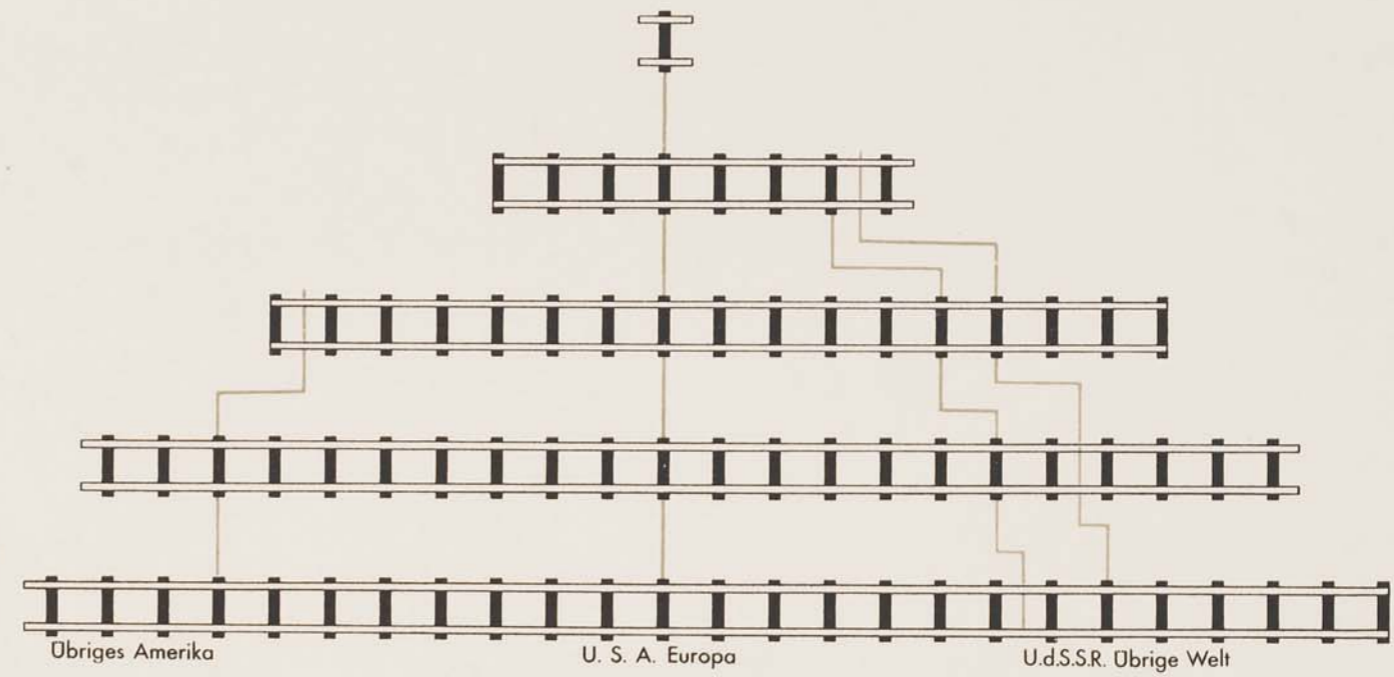
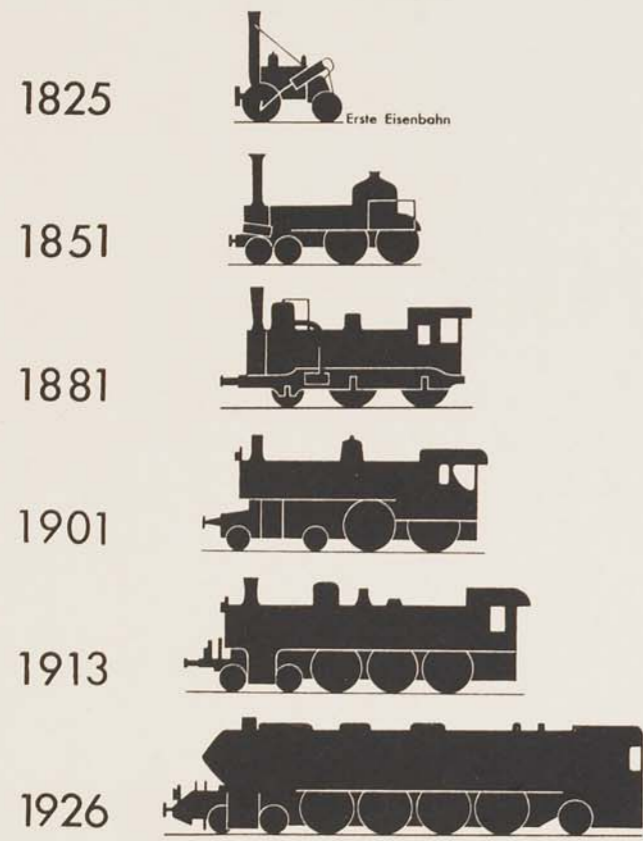
Jeder Kreis 100 Millionen Mark

Die Längen der Führungsbänder geben einen ungefähren Vergleich der Mengen der ausgeführten Maschinen. Die Kaufkraft des Goldes ist auf zwei Drittel gesunken.

Angefertigt für das Bibliographische Institut AG., Leipzig
Gesellschafts- und Wirtschafts-Museum in Wien

Otto Neurath's Isotype

Entwicklung der Eisenbahnen



50 000 Streckenkilometer

Angefertigt für das Bibliographische Institut AG., Leipzig
Gesellschafts- und Wirtschaftsmuseum in Wien

Otto Neurath's Isotype

Handelsmarinen der Erde

1850



1900



1913



1929

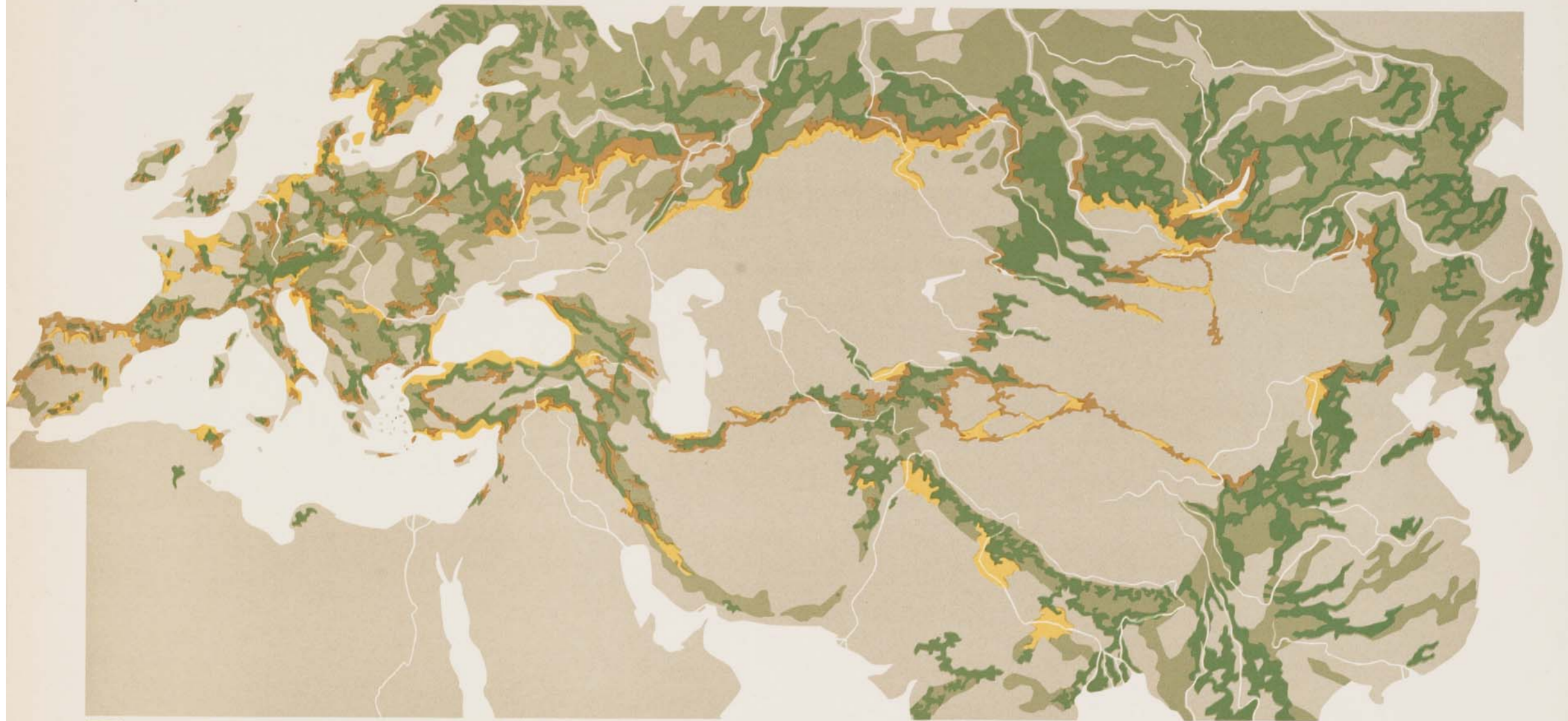


Jedes Schiff 5 Millionen Bruttoregistertonnen

Angefertigt für das Bibliographische Institut AG., Leipzig
Gesellschafts- und Wirtschaftsmuseum in Wien ©

Otto Neurath's Isotype

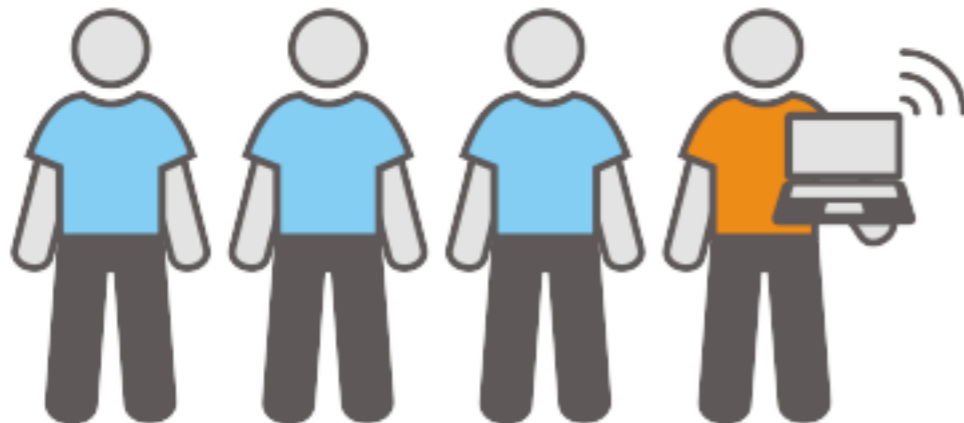
Waldbestand in Eurasien



- Bekannte Verbreitung des gegenwärtig geschlossenen Hochwaldbestandes der Nutzhölzer
- Nachweisbar abgeholzter geschlossener Hochwaldbestand
- Vermutlich ehemals geschlossener Hochwaldbestand
- Zone lockeren Waldbestandes, einschließlich Buschwald, Waldsteppe u. s. w.

Angefertigt für das Bibliographische Institut AG., Leipzig
Gesellschafts- und Wirtschaftsmuseum in Wien ©

Modern day examples



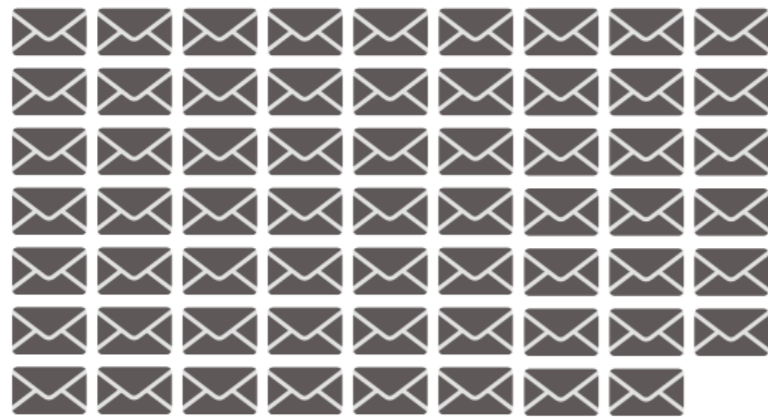
How many people are connected to the internet?

In 2009, we had approximately 6,767,805,208 people on the earth from those, 1,802,330,457 have internet access which makes it 26.6% or one quarter of the world population.

(source: <http://www.internetworldstats.com/stats.htm>)

Modern day examples

Handling, sending, receiving and filtering out spam



Each mail represents 1 trillion spam mails

takes up the power of 2.4 million US houses



Each house represents 100 000 US houses

or 1 nuclear power plant



An average nuclear power plant produces 2500 Mwatts



- ☐ Good Emails 2%
- ☒ Spam 98%
- ☑ Orders from Spam 0,02%



US household per month

\$ Each sign represents \$1000

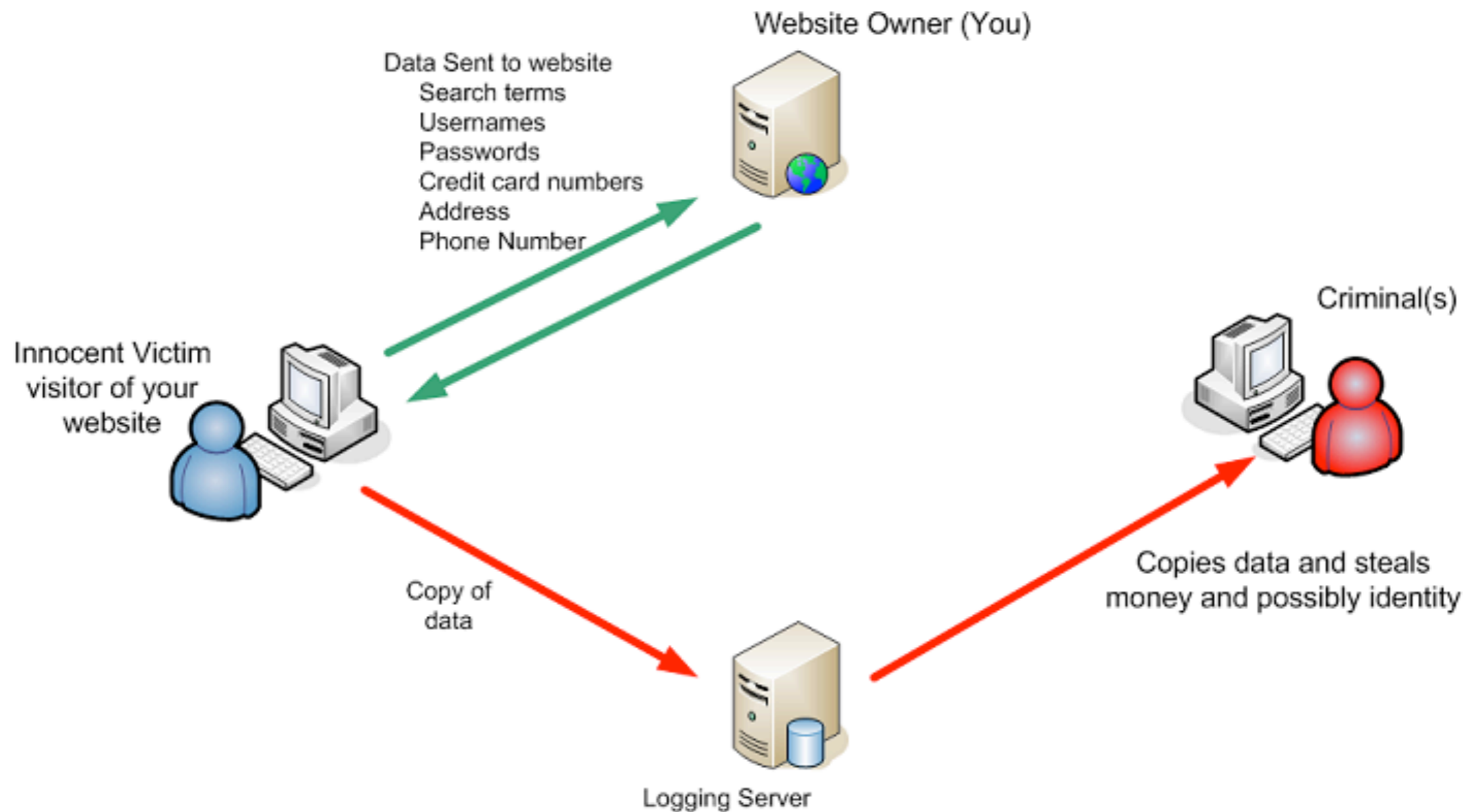


spammer per day

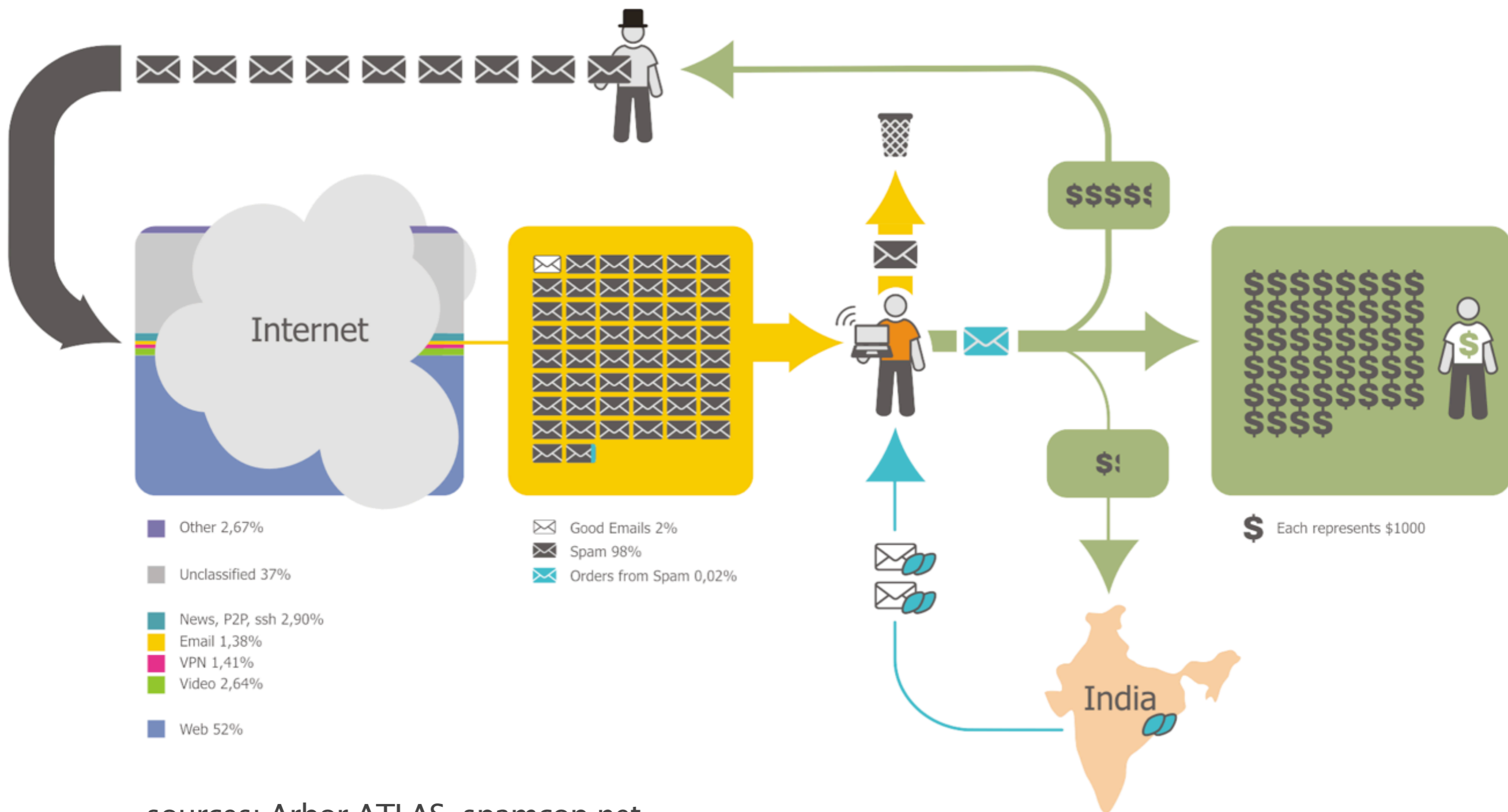
waste of resources by spam and a spammer's income

(source: McAfee CO2 Impact of Spam + NY Times)

Making users understand IT security

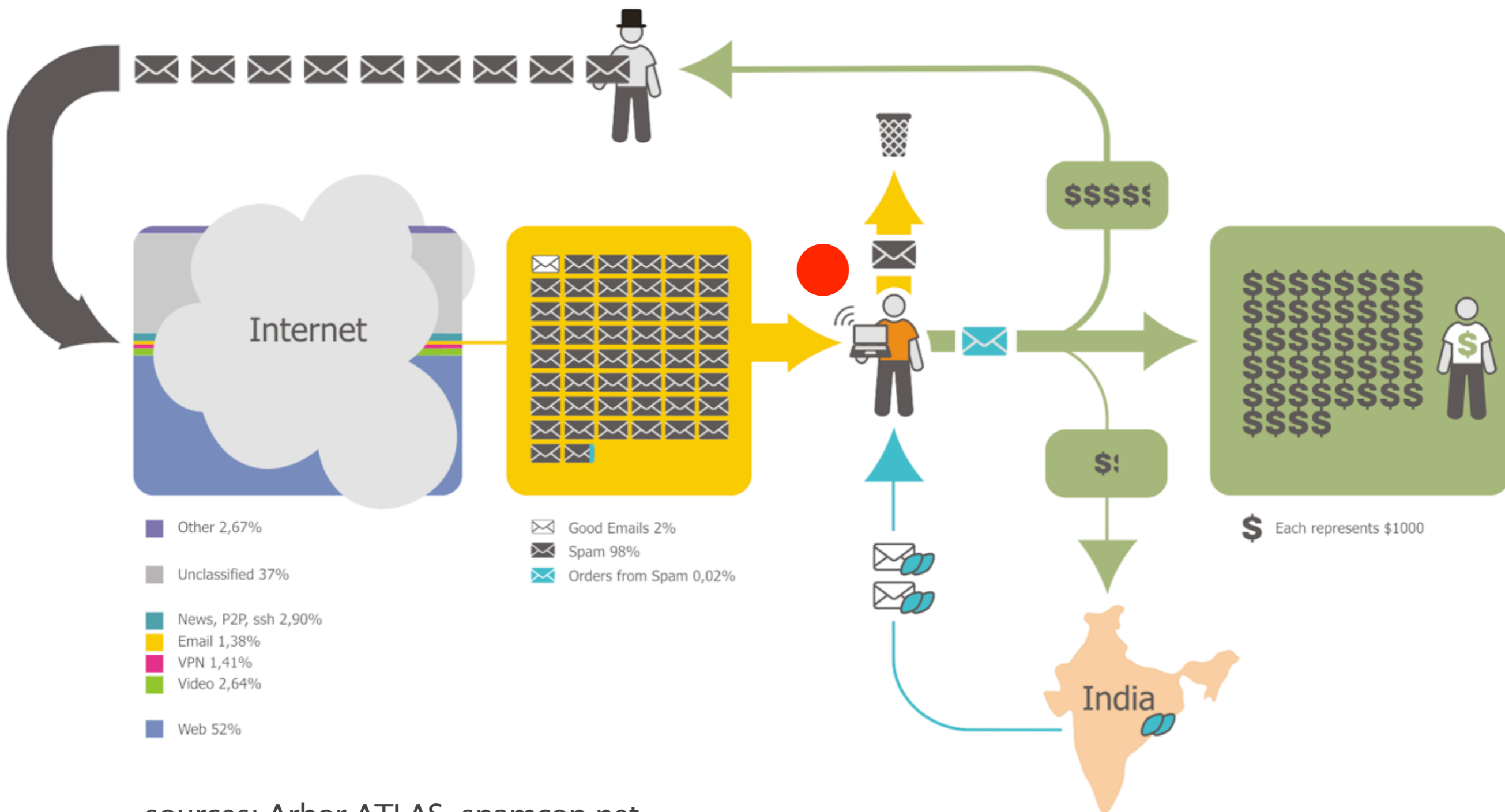


Making users understand IT security



sources: Arbor ATLAS, spamcop.net

Making users understand IT security



sources: Arbor ATLAS, spamcop.net

OVERVIEW



- Motivation
- Target Group
- 5 Minutes of design background for techies
- Tools
- DNSviz and Flows

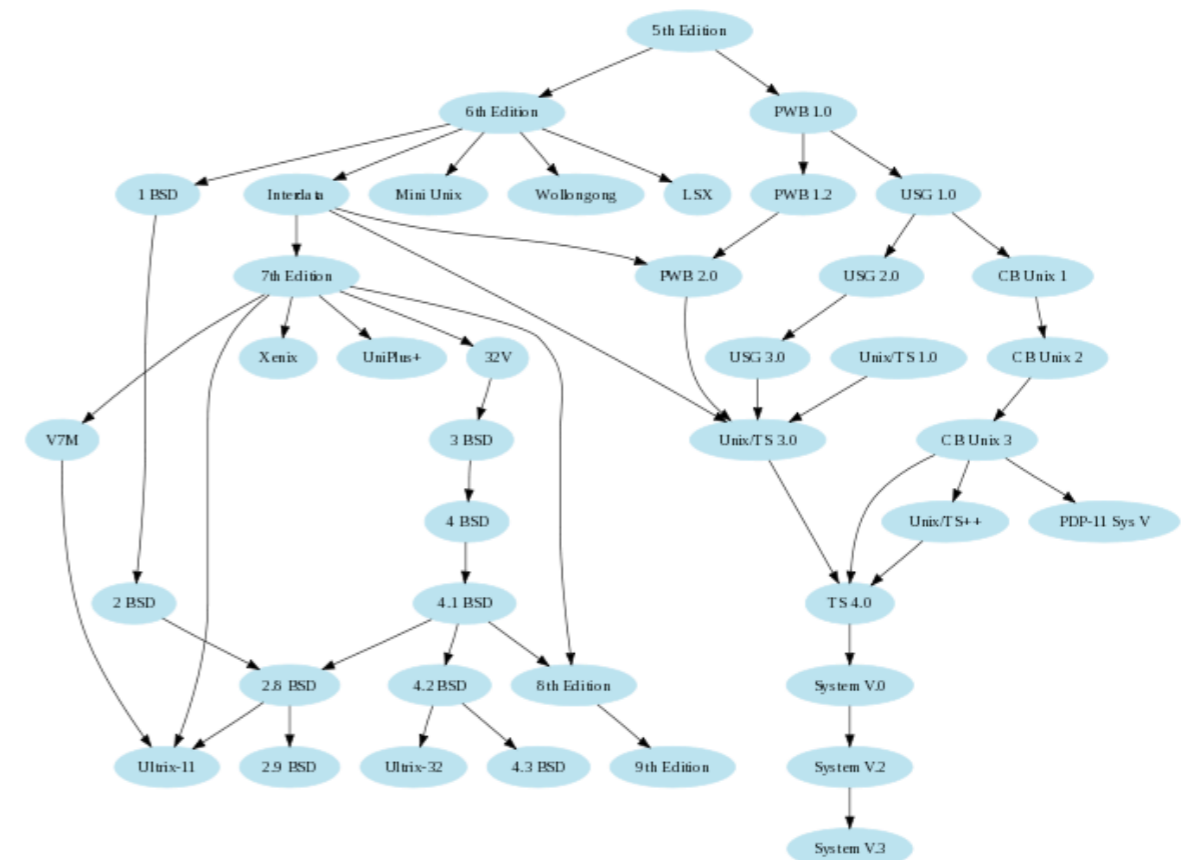
- Graphviz
- Maxmind GeolP
- Logster
- Unix wizardry
- Google Earth
- Gapminder
- [Processing.org](https://processing.org)
- Outlook: Davix

Graphviz

- based on research @ AT&T Labs
- Syntax:

```
digraph {  
  A -> B;  
  A -> C [label="foo"];  
}
```

```
dot -T png -o out.png \  
  inputfile.dot
```



Maxmind GeolIP



- <http://maxmind.com>
- cityLite DB is usually enough

```
my $gi = Geo::IP->open("/home/aaron/GeoLiteCity.dat",
GEOIP_STANDARD);
# ----- functions -----
# input : ip
# output: array [countrycode, city, lat, lon]
sub ip2geolocate {
    my $ip = $_[0];
    my @ret;
    my $record = $gi->record_by_name("$ip");
    @ret = ( $record->latitude , $record->longitude) ;
    return(@ret);
}
```

Tools: Logster



- Logster by Clarified Networks
- Input format: Apache log file format
- output: movie. Can screen capture

Tools: Logster



QuickTime Player File Edit View Share Window Help [System Tray Icons] (Charged) Wed 16:18 Aaron Kaplan

20090402-s100.log

Screen Recording—Inspector

Screen Recording

Format: H.264, 1280 x 800



2009-04-01 23:59:59 GMT

Tools: Gapminder



Google docs conficker hits by country

Autosaved on 3:31 PM GMT+02:00

File Edit View Insert Format Form Tools Help

10pt B

	A	B	C	D	E
1	country	date	cnt	cnt2	
2	MX	1000	4	1	
3	KR	1001	2	1	
4	BR	1002	2	1	
5	VE	1003	1	1	
6	CA	1004	1	1	
7	KH	1005	1	1	
8	RO	1006	1	1	
9	TT	1007	1	1	
10	RU	1008	1	1	
11	PK	1009	1	1	
12	AR	1010	1	1	
13	ES	1011	1	1	
14	CN	1012	1	1	
15	US	1013	1	1	
16	PH	1014	1	1	
17	EC	1015	1	1	
18	CN	1016	762166	1	
19	BR	1017	644646	1	
20	RU	1018	582881	1	
21	VN	1019	255845	1	
22	UA	1020	245734	1	
23	IN	1021	233908	1	
24	KR	1022	230627	1	
25	ID	1023	210525	1	
26	IT	1024	177121	1	

conficker.C by country

Color: cnt

Select: FM, FO, FR, GA, GB, GD, GE, GF, GH, GI, GL

Deselect all

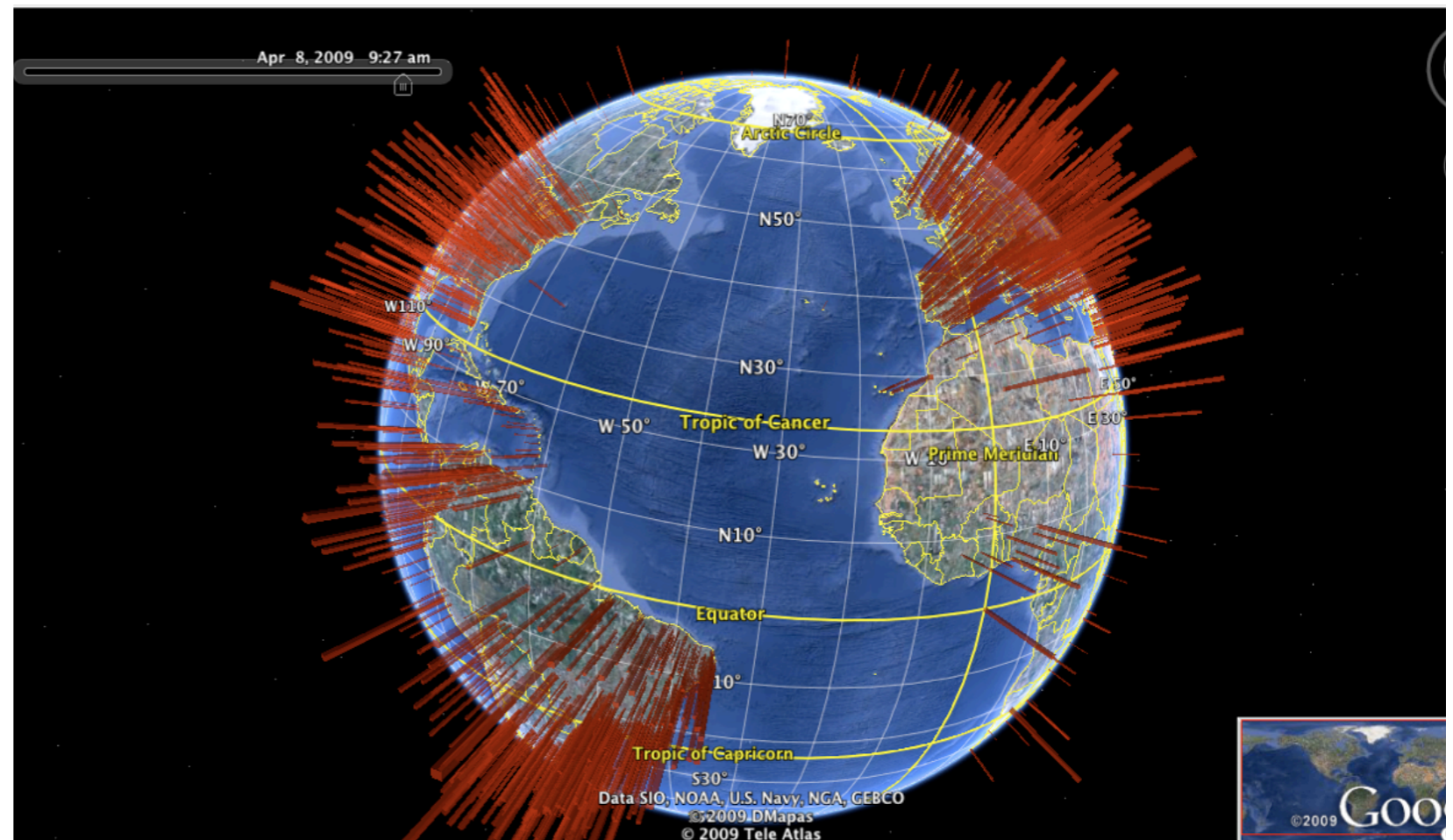
Order: cnt

1427

1

Tools: Google Earth

- format: KML. Well documented.
- Head section
- Placemarks



Tools: Unix filters

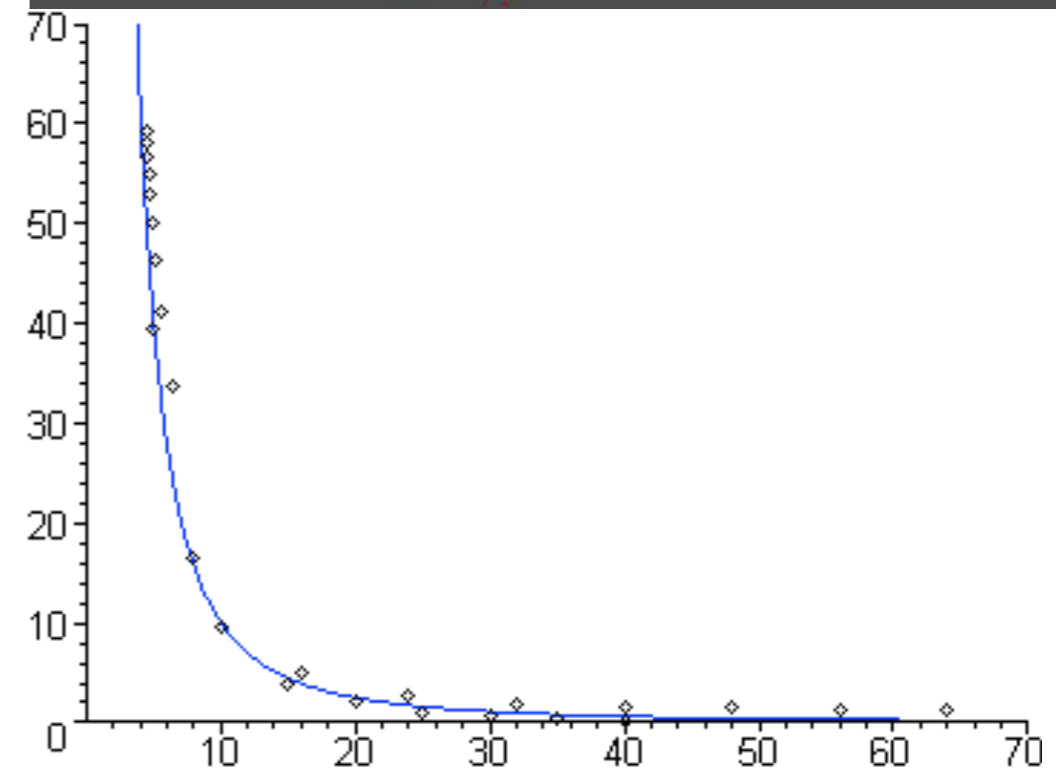
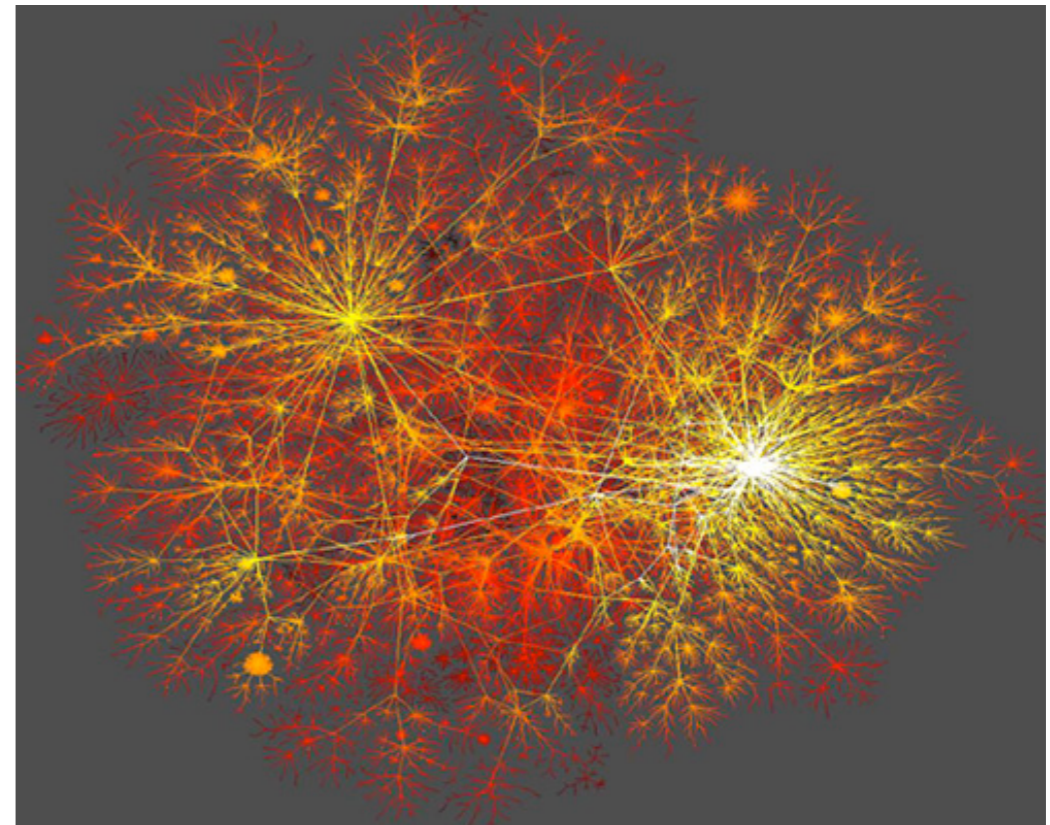


- Use Unix tools to quickly get a grasp of the trends
- `cut -d “;” -f 5 | sort | uniq -c | sort -rn`
- `gnuplot`

`plot “myfile.csv” using 1 with boxes`

Scale-freeness

- Albert–László Barabási made them famous.
- Some property is distributed by an inverse power law formula:
$$P(k) \sim 1/k^\gamma \quad (2 < \gamma < 3)$$
- “fractal”
- “internet-ish”
- “biological”
- “not again-ish”



TOOLS: Processing.org



- Invented by Ben Fry, Casey Reas @MIT
- Basic idea: easy IDE for Java 3D/OpenGL programming. Lots of examples, openprocessing.org
- Includes a rich API:
 - sockets
 - DB connections
 - serial I/O
 - sound, etc.

```
circlePolarCoordLayout | Processing 1.1
// XXX FIXME STUFF
// 0) ignore spaces. Be flexible with 't', spaces, ';' etc
// 1) maxCircleArea is a bad overall parameter -> think of smthg better
// 2) transition from keypressed 'a' -> 'b'
// 3) in 'a' mode: text width == bar height
// 4) read from URL / java applet param

import processing.opengl.*;

boolean bLogarithmic = false;

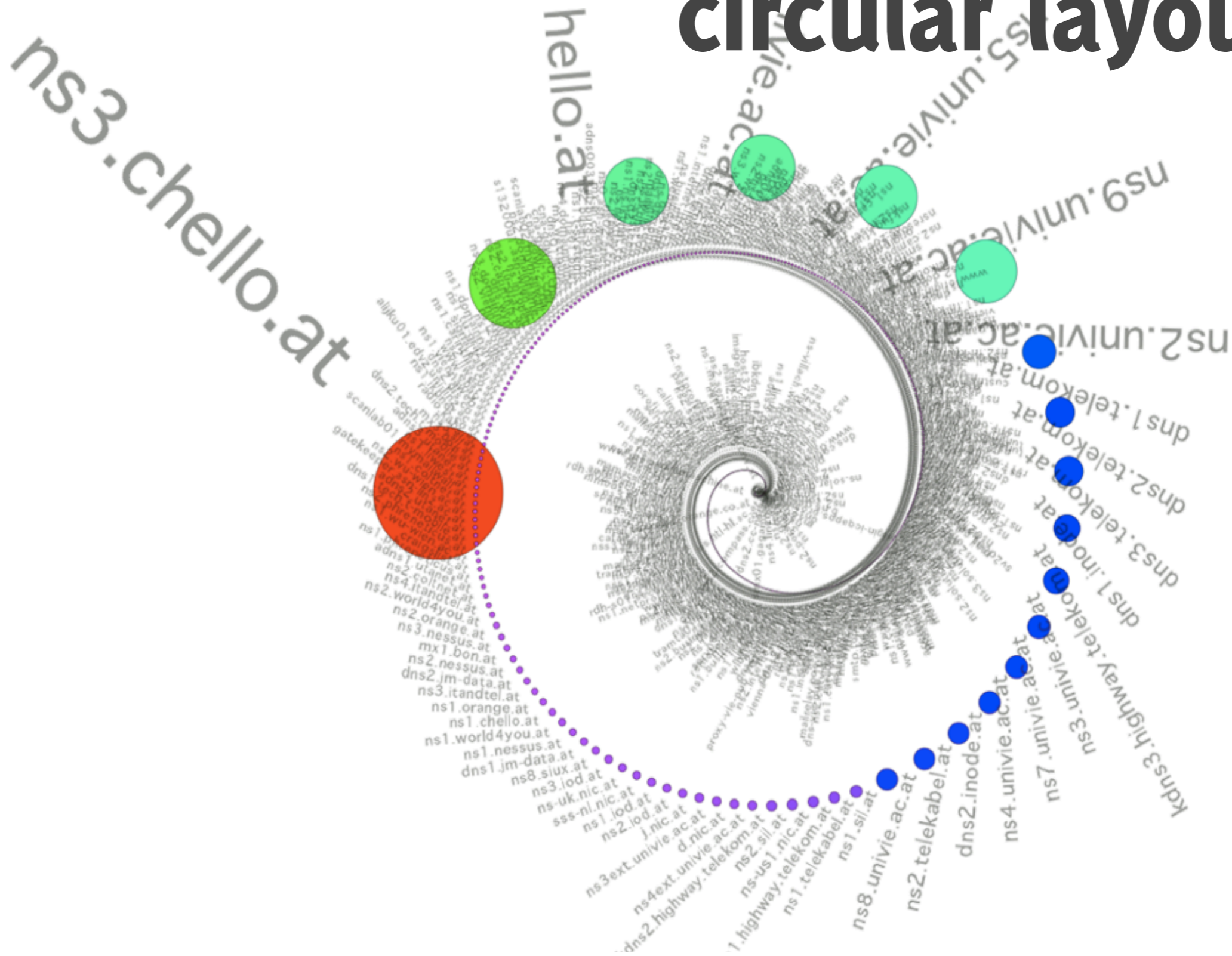
int border=10;
int num_circles=-1;

float maxCircleArea = 40000.0;
float maxCircleRadius;
float maxDistance;
float angularStep = 2*PI/360.0*3; // 3 degrees
float defaultTextAngle = PI; // how to rotate the text by default

// arrays for the circles
```



Processing example: circular layout



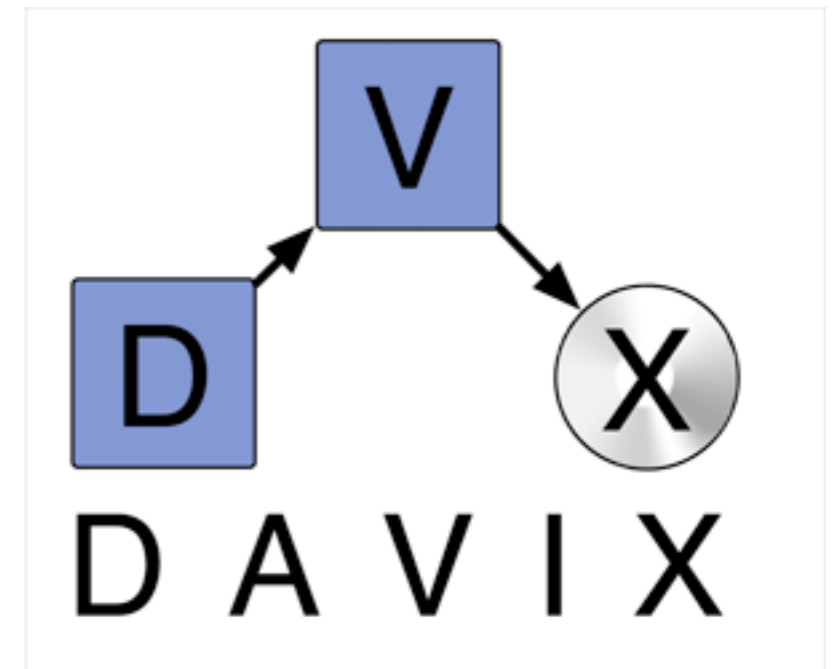
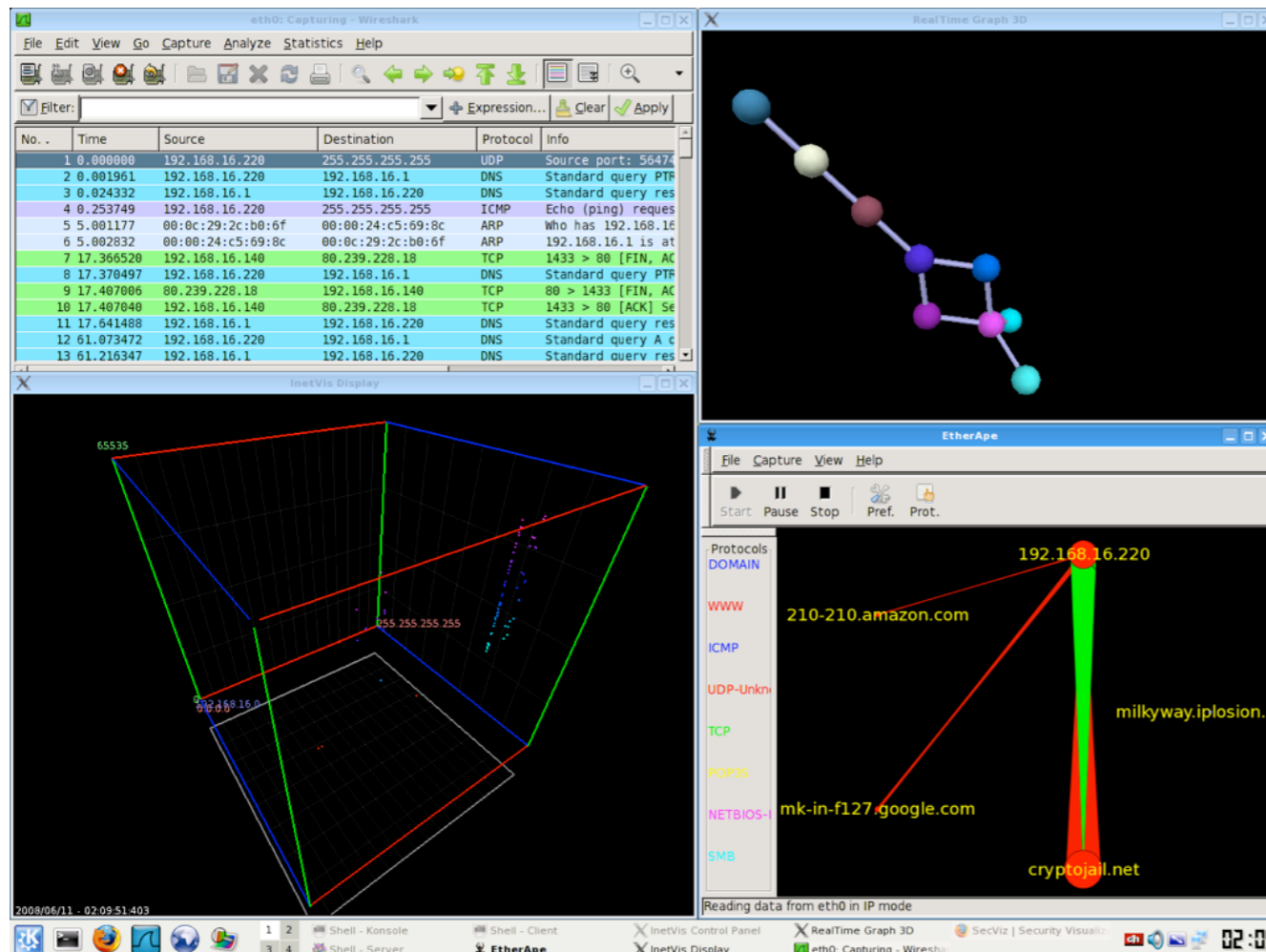
Other processing Examples



- Esfera
- Registrymon

Outlook: DAVIX

- ISO image on <http://www.secviz.org>



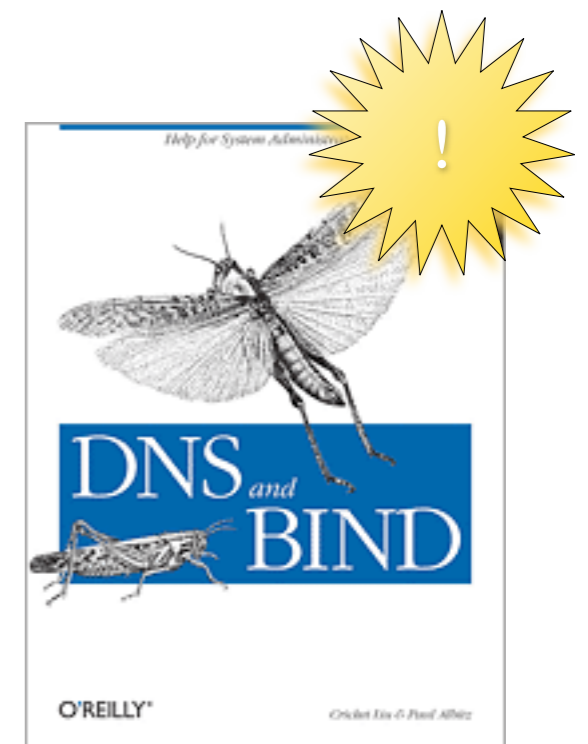
OVERVIEW



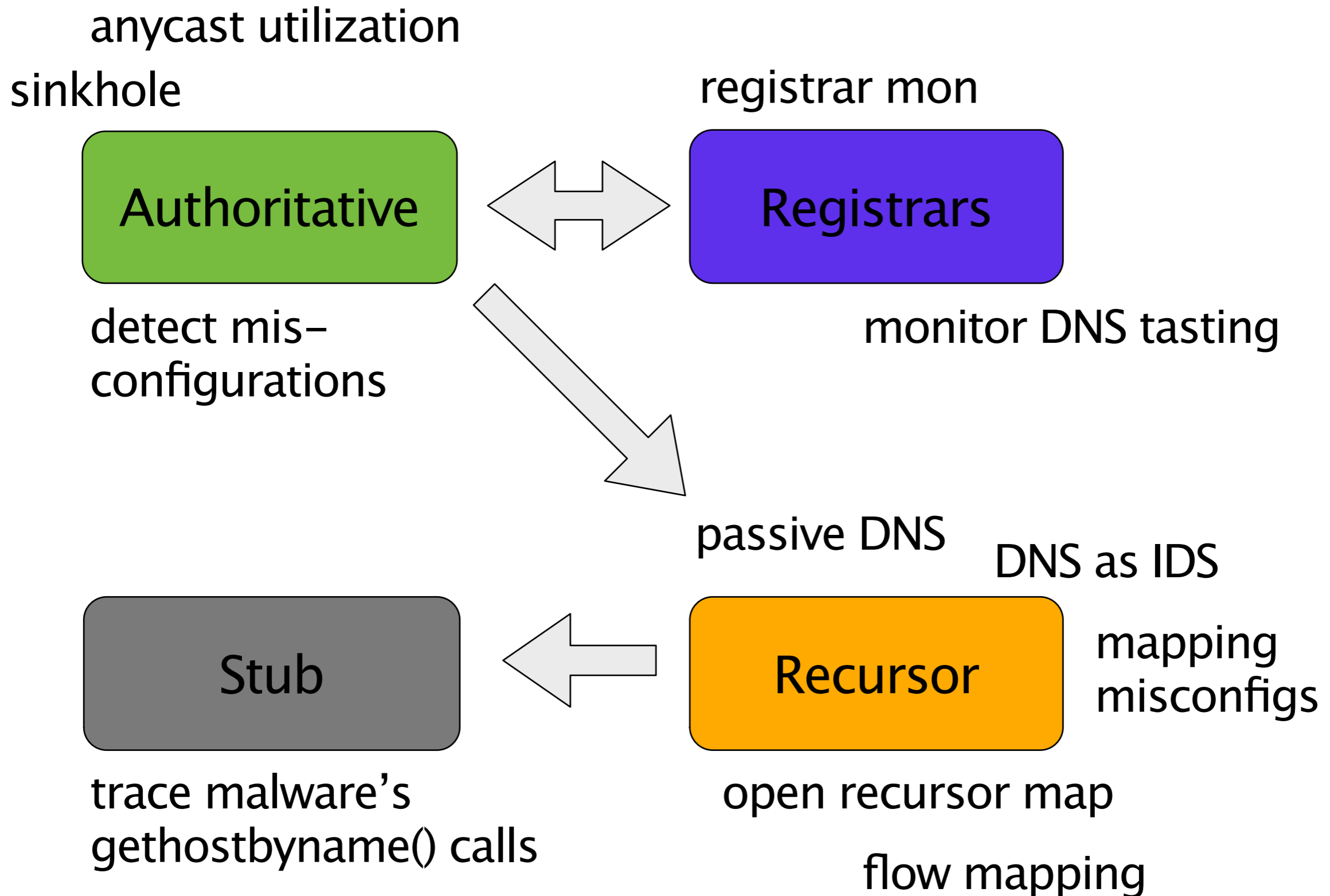
- Motivation
- Target Group
- 5 Minutes of design background for techies
- Tools
- DNSviz and Flows

DNS

DNS



DNS for IT security viz



Idea list DNS and IT security viz



- **Authoritative Nameservers:**
 - you don't see much at the authoritative NS
 - TTLs are wrong
 - other misconfigurations
 - But – idea: Spam for a newly registered domain should be a spike. But can we filter it out from the noise?
 - Anycast effectiveness (c.f. CAIDA paper)
 - **Sinkholing works!**

Idea list DNS and IT security viz



- Registry / Registrars:
 - from registry's perspective: track your resellers. How "clean" is a registrar?
 - monitor DNS tasting. Find domain catchers.
- Recursors:
 - passive DNS
 - DNS "netflow" ("passive DNS++")
 - DNS as IDS (<- Google talk today!)
 - log/visualize localhost/bogus/bogon answers!
 - fastflux
 - monitor TXT record answers
 - map (maliciously) open recursors

Idea list DNS and IT security viz



- Stub resolvers:
 - trace malware's `gethostbyname()` syscalls (Minibis)
 - idea: outgoing FW + logster for the stub / PC

DNS netflow example



- Done in Processing
- data: tcpdump -ni eth0 port 53 and src = ...
- filter out local queries
- find all nameservers which are queried
- aggregate(!) + transform via perl script to...
- format:
lat srcip; lon srcip; lat dstip; lon dstip; amount
- aggregation factor:
 - aaron@lair:~\$ wc -l outgoing-without-ports.txt
 - 100000 outgoing-without-ports.txt
 - aaron@lair:~\$ wc -l flows-lat-lon.txt
 - 28948 flows-lat-lon.txt
- source code demo?

DNS netflow

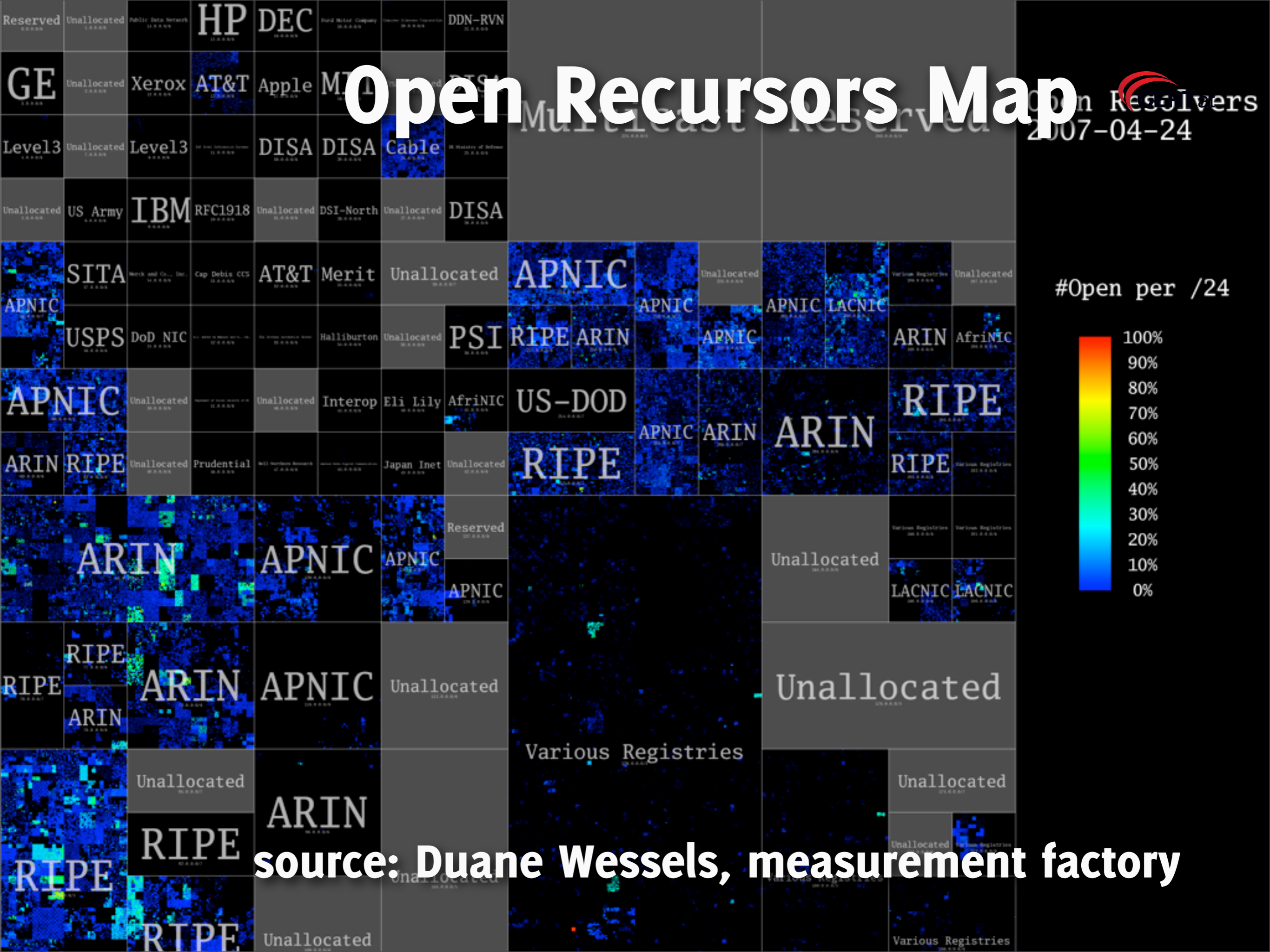


25.15 fps
rotation = ((0.6660568, 0.0, -0.11140848))

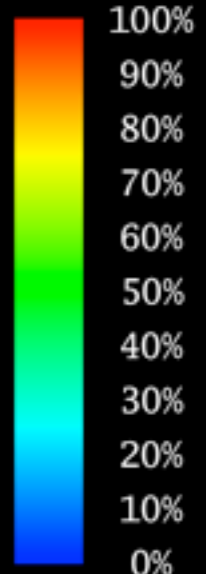


Open Recursors Map

Open Recursors
2007-04-24



#Open per /24



source: Duane Wessels, measurement factory

SIG? Data exchange?

annapetukhova.com, processing.org, Otto Neurath



Thanks!

annapetukhova.com, processing.org, Otto Neurath