Evaluation of web sites

Laura Hollink

Web Technology
23 Januari 2007

The goal of evaluation studies (1)

I want to...

- Attract more visitors
- Sell more products
- Decide which web application to use
- Get better ‘rates’ from visitors
- Etc.

The goal of evaluation studies (2)

.. and therefore I want to know...

- Is it easy for a beginner to learn to use my website?
- Is my search engine better than the competitors?
- How much do people enjoy my web site?
- How well does my website support people in their task?
- Is it easy for visitors to find what they are looking for on my website?
- Does my website stimulate buying my products?
- Does a 'product of the month' display stimulate buying my products?
- Etc.

Outline

Outline

Research Question

Always start with a clear research question!

A research question is:

- Also called Problem Statement
- Always a question (ends with a "?")
- Sometimes more than one.
- Has practical and/or theoretical relevance.
- Is feasible (time, money, people).


Three types of studies

- Explorative (what is related?)
  - What factors determine if people come back to visit a website a second time?
  - Is there a correlations between characteristics of my visitors and the types of errors that they make?
- Descriptive (what happens?)
  - What percentage of people find my website through Google?
  - What percentage by typing in the URL directly?
- Explanatory (why does it happen?)
  - Does the addition of a login function mean that more people come back to my website a second time?
  - Did the revision of the link structure of my website make visitors find what they were looking for quicker?

Three types of studies

- Explorative (what is related?)
  - What factors determine if people come back to visit a website a second time?
  - Is there a correlations between characteristics of my visitors and the types of errors that they make?
- Descriptive (what happens?)
  - What percentage of people find my website through Google?
  - What percentage by typing in the URL directly?
- Explanatory (why does it happen?)
  - Does the addition of a login function mean that more people come back to my website a second time?
  - Did the revision of the link structure of my website make visitors find what they were looking for quicker?

Three types of studies

- Explorative (what is related?)
  - What factors determine if people come back to visit a website a second time?
  - Is there a correlations between characteristics of my visitors and the types of errors that they make?
- Descriptive (what happens?)
  - What percentage of people find my website through Google?
  - What percentage by typing in the URL directly?
- Explanatory (why does it happen?)
  - Does the addition of a login function mean that more people come back to my website a second time?
  - Did the revision of the link structure of my website make visitors find what they were looking for quicker?

Three types of studies

- Explorative (what is related?)
  - What factors determine if people come back to visit a website a second time?
  - Is there a correlations between characteristics of my visitors and the types of errors that they make?
- Descriptive (what happens?)
  - What percentage of people find my website through Google?
  - What percentage by typing in the URL directly?
- Explanatory (why does it happen?)
  - Does the addition of a login function mean that more people come back to my website a second time?
  - Did the revision of the link structure of my website make visitors find what they were looking for quicker?

Hypothesis

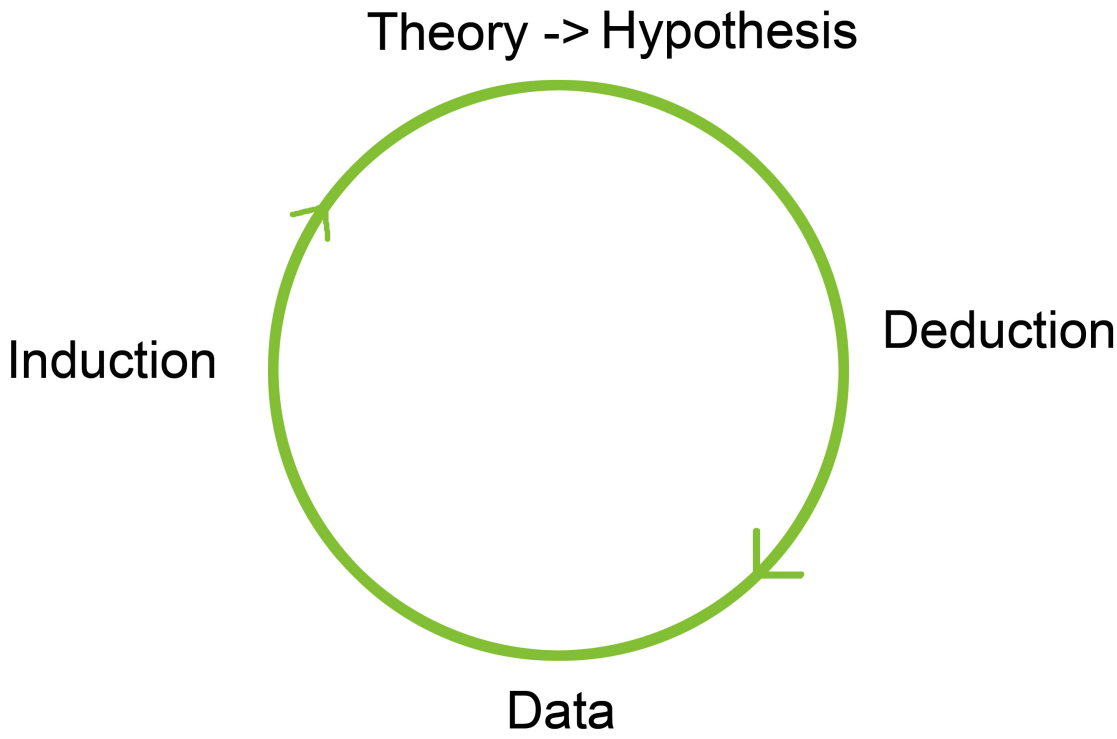
A hypothesis is:

- A prediction of the outcome of your test
- Deduced from theory of observations
- This is whay you actually test
- Necessary in explanatory studies, often used in descriptive, rarely in explorative studies.

Examples:

- Website A with 10 adds per page is rated lower by visitors than website B with no adds.
- Visitors who have returned to the website more than 10 times use more shortcuts than first-time visitors.
- Using shortcuts reduces the time to reach the target page.

Empirical Cycle



Collecting data

- Data collection methods
  - laboratory experiment
  - analysing texts
  - survey
  - interview
  - etc...
- Qualitative vs. Quantitative data
  - Qualitative: non numerical, e.g. analysis of words (interview), pictures or objects.
  - Quantitative: analysis of numerical data

Collecting data

- Data collection methods
  - laboratory experiment
  - analysing texts
  - survey
  - interview
  - etc...
- Qualitative vs. Quantitative data
  - Qualitative: non numerical, e.g. analysis of words (interview), pictures or objects.
  - Quantitative: analysis of numerical data

Variables

Independent variables

- the variable that you vary
  - level of expertise (expert/novice)
  - website A or website B
  - with or without login function

Dependent variables

- the variable that you measure
  - number of mistakes
  - time to reach the required page
  - satisfaction rate

Mini Quiz

Hypothesis: Young people make less mistakes than old people.

Independent variable :

Dependent variable :

Mini Quiz

Hypothesis: Young people make less mistakes than old people.

Independent variable : Age

Dependent variable : number of mistakes

Mini Quiz

Hypothesis: Young people make less mistakes than old people.

Independent variable : Age

Dependent variable : number of mistakes

Mini Quiz

Hypothesis: Young people make less mistakes than old people.

Independent variable : Age

Dependent variable : number of mistakes

Mini Quiz

Hypothesis: More then half of the visitors find my website via Google.

Independent variable :

Dependent variable :

Mini Quiz

Hypothesis: More then half of the visitors find my website via Google.

Independent variable : Entry point

Dependent variable : Number of visitors

Mini Quiz

Hypothesis: More then half of the visitors find my website via Google.

Independent variable : Entry point

Dependent variable : Number of visitors

Mini Quiz

Hypothesis: More then half of the visitors find my website via Google.

Independent variable : Entry point

Dependent variable : Number of visitors

Population

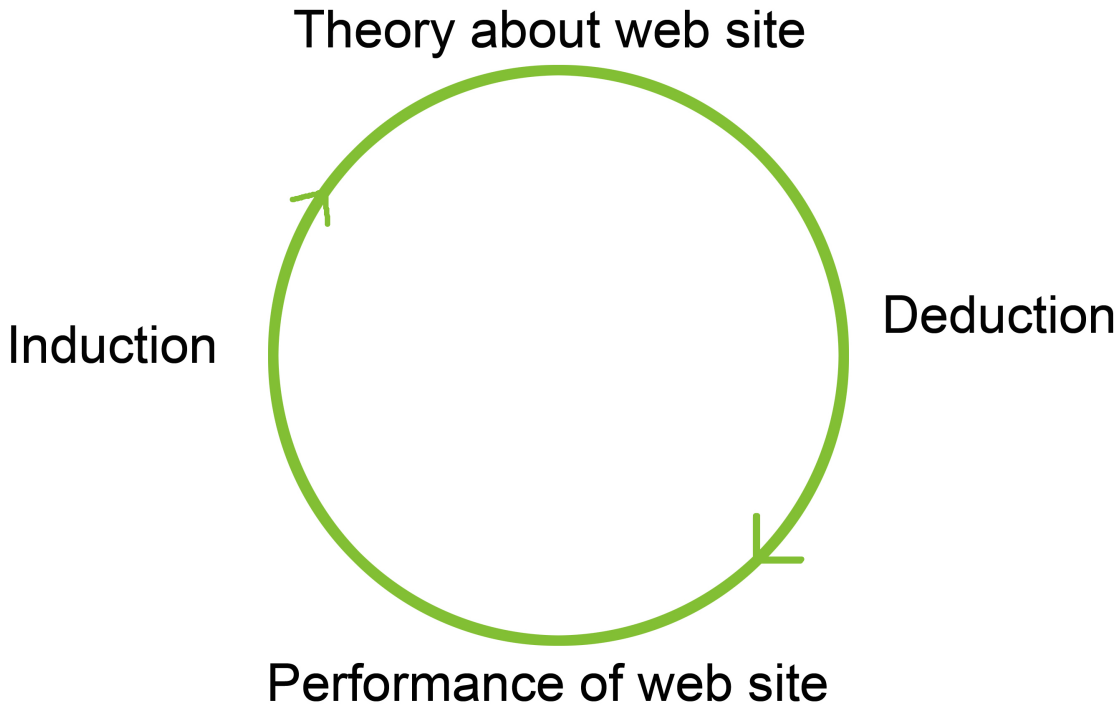
- What is the population?
  - Customers
  - All the web users
  - Registered users
- What is my sample?
  - random sample
  - convenience sample
  - voluntary response sample

Population

- What is the population?
  - Customers
  - All the web users
  - Registered users
- What is my sample?
  - random sample
  - convenience sample
  - voluntary response sample

Outline

Empirical cycle of web evaluation



Web sites vs. other applications

- Remote and largely unknown user group
- Navigation through hyperlinks

Web sites vs. other applications

- Remote and largely unknown user group
- Navigation through hyperlinks

Intermezzo: How People Navigate

People tend to minimize:

- Time to get to target
  - minimize number of links to scan, but also
  - minimize time spend on clicking and waiting for the page to load
- Mental effort
  - remember as little as possible (breadcrumbs)
  - reason as little as possible

Intermezzo: Navigation structures of web sites

- Hierarchy (tree)
- Linear
- Matrix (grid)
- Full mesh
- Arbitrary network
- Hybrid

- Does the navigation-structure of your web site match the mental model of your users?
- Does it follow the internal structure of your data?

Intermezzo: Navigation structures of web sites

- Hierarchy (tree)
- Linear
- Matrix (grid)
- Full mesh
- Arbitrary network
- Hybrid

- Does the navigation-structure of your web site match the mental model of your users?
- Does it follow the internal structure of your data?

Methods for web evaluation I

Common methods for evaluation of web sites:

- Mockups
- Prototypes
- Focus groups and card sorting
- Usability inspection
- Group walkthrough
- (Remote) User testing
- (Online) Survey

Methods for web evaluation II

Specific **web**-evaluation methods:

- Web analytics
- Online experiments

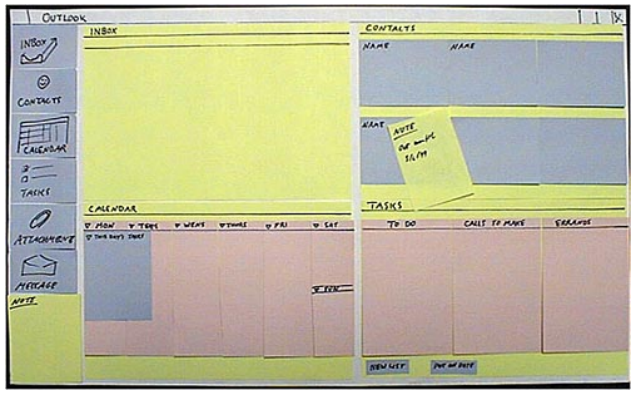
Mockups

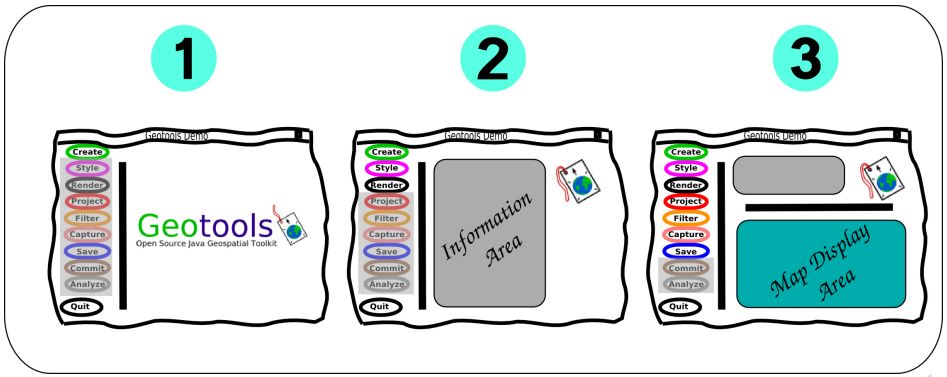
What Quick, static model of a web page.

Goal Facilitate communication across team of designers, developers, user, manager, clients.

Low fidelity mockup

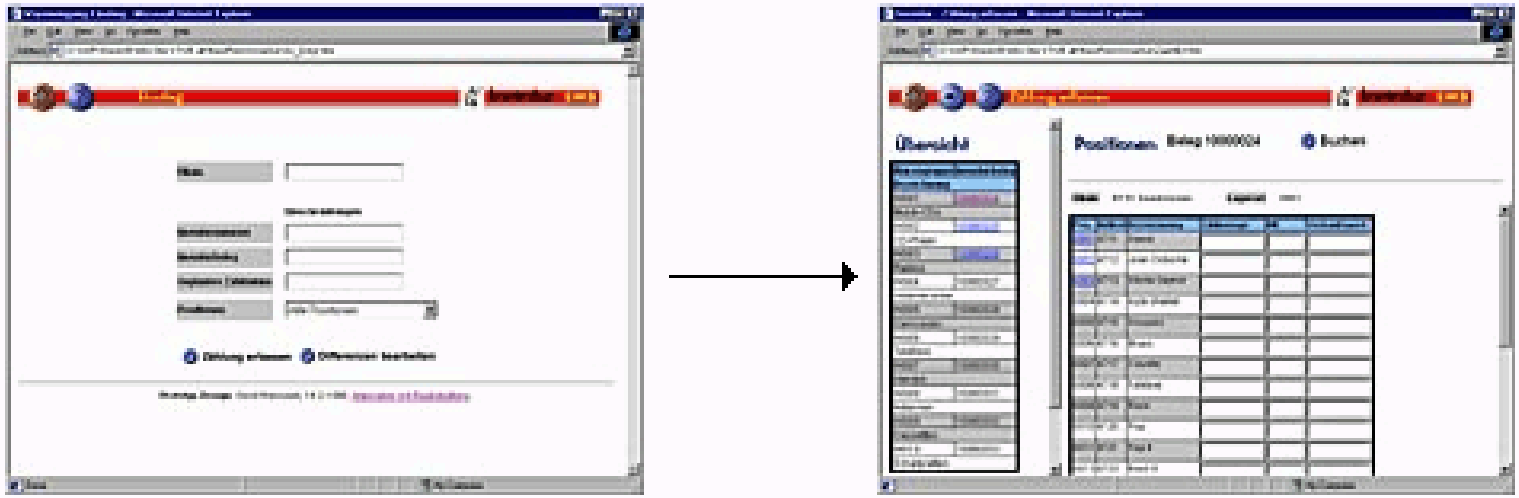
- Early in the design phase
- Only basic functionality or visual layout
- Cheap
- E.g thumbnail sketch, paper mockup
- Focus on conceptual design





High fidelity Mockups

- Later in the design phase
- Refined details of the design
- Expensive
- E.g digital mockup (html or image)

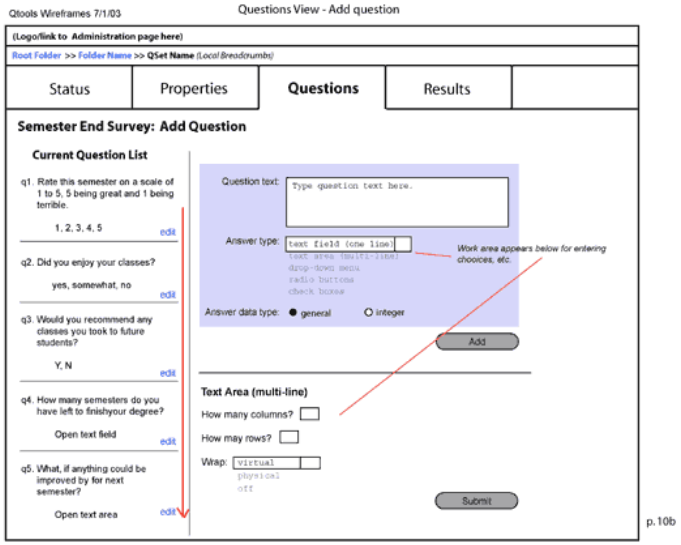


Prototypes

What A working example of the web site.

Goal Allows to do a user test / usability inspection before building the actual web site

Early stage: paper storyboard
Midway: digital storyboard, wireframe
Later in design: high end prototype



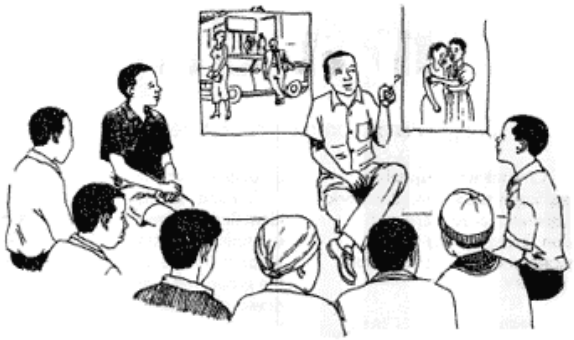
Focus groups

What Moderated group discussion

Goal To elicit user's views and opinions

How People comment on a presented idea, a mockup, etc.

When Early in the design stage



Carefully chosen pictures can be used in FGDs to provoke detailed discussion.

Card sorting

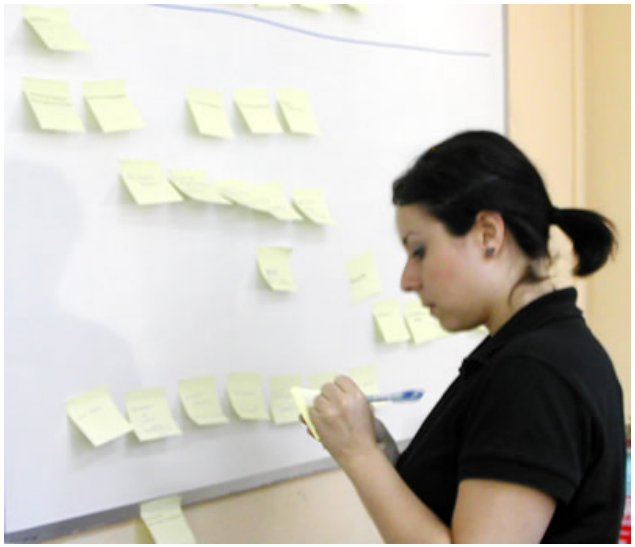
Special type of focus group

What Group of people sort items

Goal To obtain an intuitive structure for the web site

How People place cards with item names in clusters, and name the clusters

When Once the items that will appear on the web site are identified.



Card sorting on Del.icio.us tags
(look below for details of method)

Usability inspection I

Checklist of usability guidelines,
e.g. **The Ten Web Guidelines**:

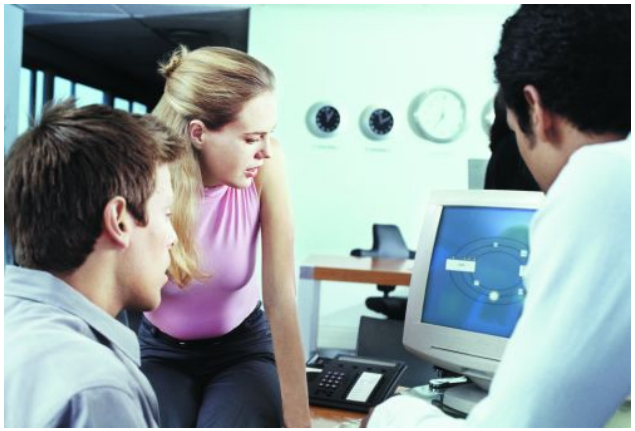
1. Content and scope
2. Speed
3. Navigation
4. Appropriateness to task
5. Visual design
6. Compatibility
7. Simplicity
8. Consistency and contrast
9. Error handling
10. Respect for the user

Usability inspection II

- Go systematically through a web site and check if everything complies with the guidelines.
- Performed by development team (designer, developer, tester, manager, usability expert, domain expert).
- Note: Don’t neglect your first impression.
- Can be partly automated

Group walkthrough

- A group of people walk though the web site as if they are performing the primary tasks.
- Give comments along the way.
- Ideally a mixed group: designers, different types of users, developers.



(Remote) User testing

Observe user while she performs primary tasks on the website.
Observe and record:

- Watch and take notes
- Record video and/or audio
- Log actions
- Record keystrokes and mouseclicks
- Eye tracking
- Discussion afterwards
- Questionnaire

(Online) Survey I

Surveys or questionnaires are a widely used technique.

Online survey: adapt questions to previous answers.

Collect information about, for example:

- Demographics of visitors
- Needs and preferences of visitors
- Other web sites visited by your visitors.

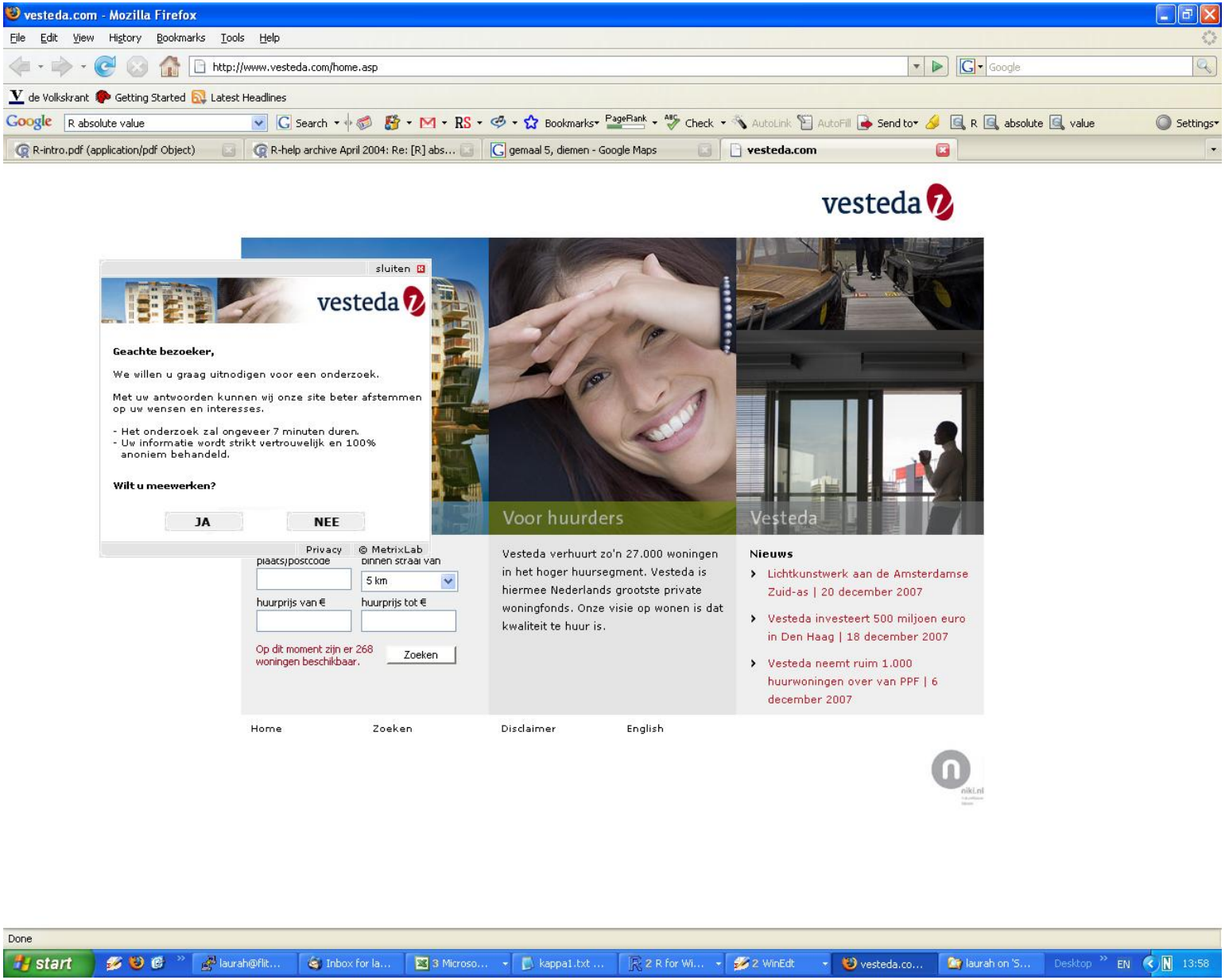
(Online) Survey II

Open questions: respondents can enter any response they like
Closed questions: respondent choose from a predefined set of answers.
Likert scale: rate agreement on a numbered scale

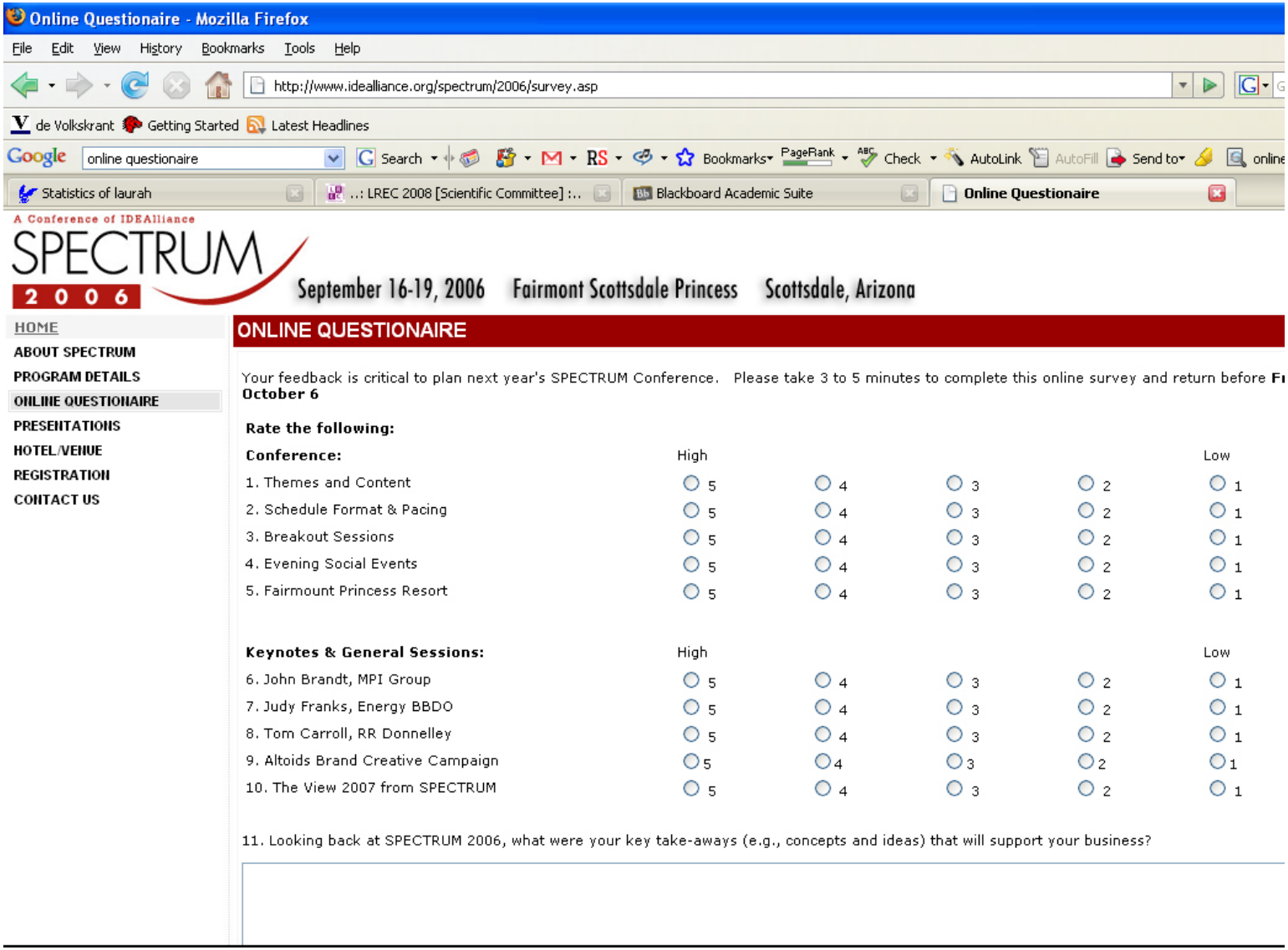
1. It is the duty of doctors to keep people alive for as long as possible.

- ☐ Strongly Agree
- ☐ Agree
- ☐ Agree somewhat
- ☐ Undecided
- ☐ Disagree somewhat
- ☐ Disagree
- ☐ Strongly disagree

Online survey: example I



Online survey: example II



Outline

Logfiles

What Records of what happens on a websites.

Goal To capture visitor data such as number of hits, navigation, conversion rate, number of errors, where did they browse from, which browser, etc.

How Every time a request is made to the server of the website, that request is added to the logfile.

Use Amongst many other things: examine traffic patterns by time of day, day of week, referrer, or user agent

Example: cs.vu.nl logfiles

```
89.0.93.210 - - [06/Jan/2008:03:56:35 +0100]
"GET /~laurah/VO/vt_data3.rdf HTTP/1.1"
200 220855 "-" "Mozilla/5.0 (X11; U; Linux
x86_64; en-US; rv:1.8.1.11) Gecko/20071128
Iceweasel/2.0.0.11 (Debian-2.0.0.11-1)"
```

Example: cs.vu.nl logfiles

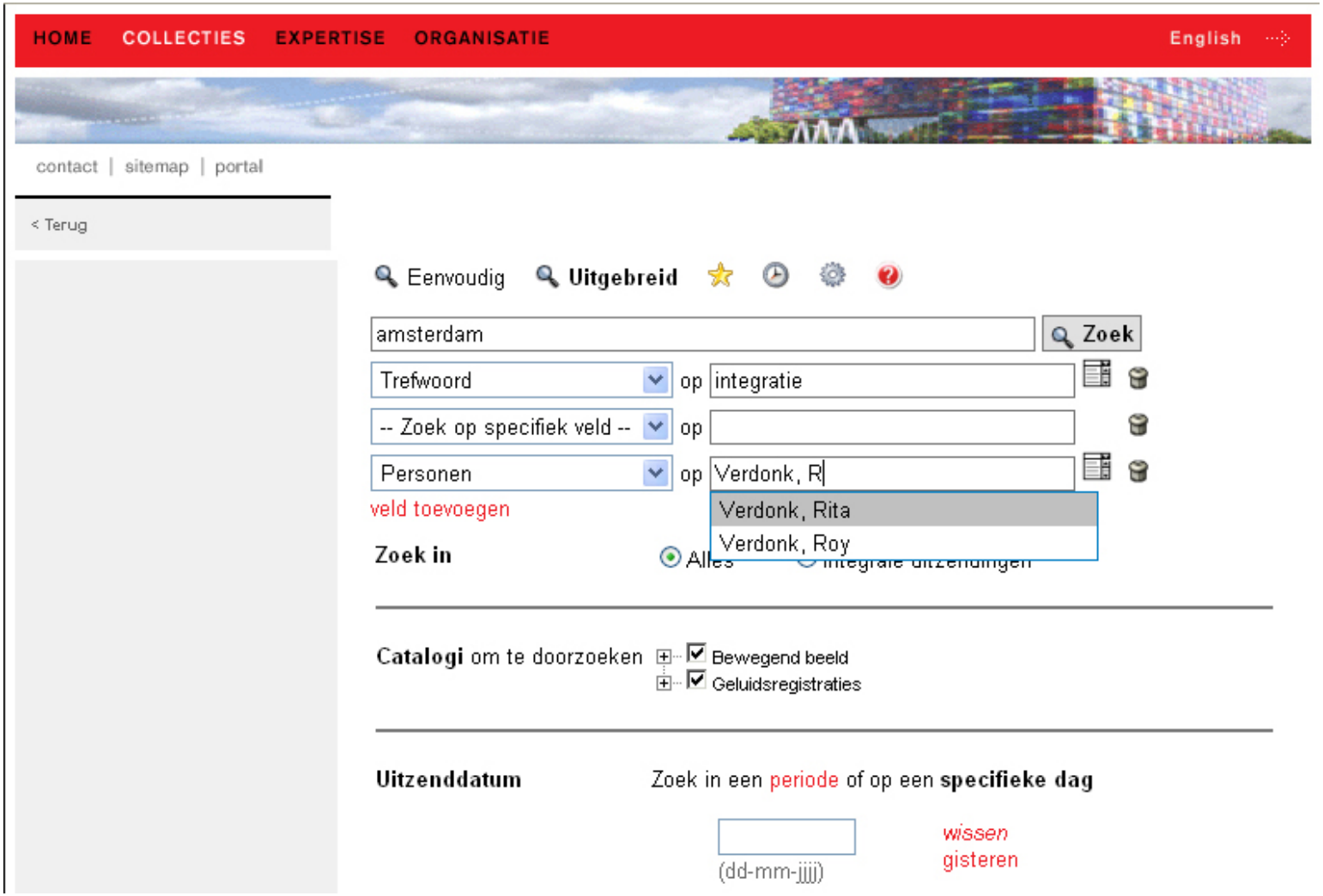
```
171.64.75.130 - - [06/Jan/2008:03:58:06
+0100] "GET /~laurah/foaf.rdf HTTP/1.0" 200
643 "-" "WebVac
(webmaster@pita.stanford.edu WebVac.org )"
```

Meaning: IP adres of visitor - user id - time - the request - the status code - size of the page - the ‘referrer’ - browser of the visitor.

Example: cs.vu.nl logfiles

See AWSTATS page.

Example: Beeld en Geluid logfiles



Example: Beeld en Geluid logfiles

dat_DateTime	vch_Action	Query	StartR	Res#
11/18/08 13:39	open advanced search form		0	0
11/18/08 13:40	view search results		1	0
11/18/08 13:40	view item details		0	1
11/18/08 14:08	open advanced search form		0	0
11/18/08 14:10	login		0	0
11/18/08 14:13	open advanced search form		0	0
11/18/08 14:13	view search results	rondom tien	1	0
11/18/08 14:13	open DRM popup		0	0
11/18/08 14:16	new order list		0	0
11/18/08 14:16	add item to order list		0	0
11/18/08 14:17	open advanced search form		0	0
11/18/08 14:17	view search results	achterwerk in de kast	1	0
11/18/08 14:17	view item details	achterwerk in de kast	0	1
11/18/08 14:18	view search results	achterwerk in de kast	1	0
11/18/08 14:18	open DRM popup		0	0
11/18/08 14:19	add item to order list		0	0
11/18/08 14:19	view shopping cart		0	0
11/18/08 14:22	submit order list		0	0
11/18/08 14:26	open advanced search form		0	0
11/18/08 14:26	view search results	RTL nieuws	1	0
11/18/08 14:26	open DRM popup		0	0
11/18/08 14:31	open advanced search form		0	0
11/18/08 14:31	new search	sex met angela	0	0

Example: Beeld en Geluid logfiles

Sessions:

- Track coherent sequences of actions
- Frequent patterns/combinations of actions
- Search - Watch program - Login - Buy
- Login - Search - Refine - Watch program - Buy

Example: Beeld en Geluid logfiles

- Does our indexing cover user needs?
- For which programs do people search
- How many people actually buy what they find?
- Do people buy more if they find it easily?
- Which user actions lead to buying programs?
- Can we link search terms to bought programs?

Page tagging

What A small piece of JavaScript on a web page

Goal To capture visitor data similar to that stored in logfiles

How Every time the page is requested it automatically runs the JavaScript in the web browser and sends information to a remote server.

Advantage I Possibility to get information about visitor: screen resolution, screen colour depth and the java version they are running.

Advantage II Information even if a page is cached.

Example: Page tagging

Script on html page from Google Analytics:

```
<script src="
  http://www.google-analytics.com/urchin.js"
type="text/javascript">
</script>
<script type="text/javascript">
  _uacct="UA-12345-X"; urchinTracker();
</script>
```

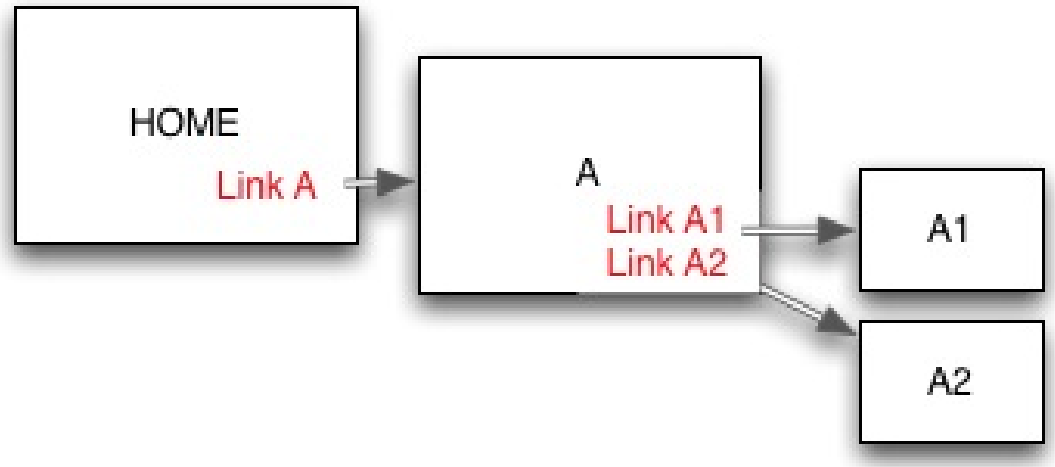
Further use of logfiles

Examples

- Predict where a user will go based on where she has been.
- Optimize structure of web site based on how users navigate.
- Debug labels of links

Further use of logfiles: debug labels of links:

- If you observe this pattern frequently:
  Click Link A - Click Link A1 - Go Back - Click Link A2
- The name of Links A1 or A2 might be misleading.
- E.g People - Contact



Pros of Web Analytics

- Many statistics, such as number of clicks, time spend on page, drop off.
- Origin of visiters.
- Objective measurements
- Data from the actual visitors.
- Continuous, realtime feedback.
- Trends visible.

Cons of Web Analytics

- No insight into motivation and opinion of visitor.
- No information on pages that have not been visited.
- Dependent on cookies for visitor identificaion.
- Privacy issues.

Outline

Online experiments

Compare different online versions of one web site

- Distribute visitors over versions
- Measure which version perform better

When?

- Evaluation **after** release
- Also used for e-mail or banners

Experimental designs:

- Online experiment with 2 versions: A/B testing
- Multiple versions is harder to interpret

Online experiments

Compare different online versions of one web site

- Distribute visitors over versions
- Measure which version perform better

When?

- Evaluation **after** release
- Also used for e-mail or banners

Experimental designs:

- Online experiment with 2 versions: A/B testing
- Multiple versions is harder to interpret

Online experiments

Compare different online versions of one web site

- Distribute visitors over versions
- Measure which version perform better

When?

- Evaluation **after** release
- Also used for e-mail or banners

Experimental designs:

- Online experiment with 2 versions: A/B testing
- Multiple versions is harder to interpret

Cycle of online experiments

1. Build web site
2. Test
3. Analyse
1. Adapt web site (or not)
2. Test
3. Analyse
1. Adapt web site (or not)

   Etc...

Goals of online experiments

Examples:

- Increase ‘conversion rate’
- Increase the amount of registered users (information)
- Decrease time spend on answering customer questions

Pros and cons of online experiments

+ Measure real users in real world context
+ Easy to get a lot of participants
- External factors
- Not every goal is measurable, e.g. branding, PR.
- Issues that are not tested will not appear
- No information on why one version is better
- Cookies needed for indentification

Cookies

- Needed for user identification in online experiments
- 39% Of internet users delete their cookies at least once a month.
- Some people disable cookies altogether.

Intermezzo: The privacy blunder of AOL

August 4, 2006

- AOL published 20 million web queries from about 500.000 AOL users in the course of three months (march to may 2006).
- The AOL username was replaced by a unique ID, everything else was kept unchanged in the logs.
- BUT: People frequently search their own name, address, social security number, names of friends, etc.

`Zoekresultaten AOL op straat.`

Intermezzo: The privacy blunder of AOL

August 4, 2006

- AOL published 20 million web queries from about 500.000 AOL users in the course of three months (march to may 2006).
- The AOL username was replaced by a unique ID, everything else was kept unchanged in the logs.
- BUT: People frequently search their own name, address, social security number, names of friends, etc.

`Zoekresultaten AOL op straat.`

Intermezzo: The privacy blunder of AOL

Many AOL users could be identified:

"how to change brake pads on scion xb
2005 us open cup florida state champions
how to get revenge on a ex
how to get revenge on a ex girlfriend
*how to get revenge on a friend who f***ed you over*
replacement bumper for scion xb
florida department of law enforcement
crime stoppers florida

"how to kill your wife
pictures of dead people
photo of dead people
car crash photo"

Intermezzo: The privacy blunder of AOL

August 6 AOL takes down the web site but there is still google cache copy available.

August 7 NEWS.com

AOL apologizes for release of user search data: "This was a screw-up, and we're angry and upset about it. It was a mistake, and we apologize."

September 25 Slashdot
News for Nerds. Stuff that matters.

AOL Subscribers Sue Over Release Of Search Data