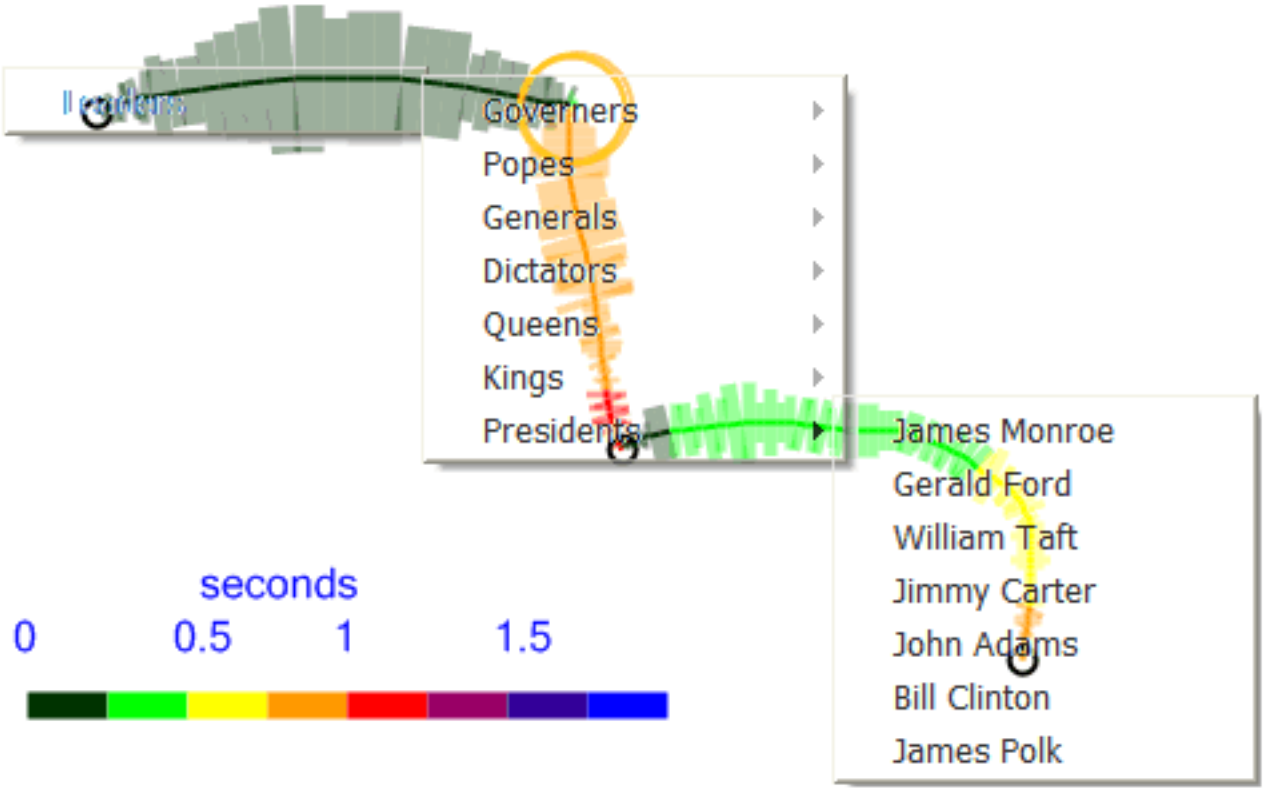
Designing and Using Analytics to Create (and Iterate) Better Products

Andy Edmonds, Sr Director Product Management Search Science, eBay

For Silicon Valley Product Camp 2011

Where I'm Coming From....



Stroke width maps velocity.

Large circles are pauses.

Small black circles are clicks.

Read more at my blog http://surfmind.com -- post frequency is low, but the archives go back to 1999

My Day Job…

How do you search and sort among everything that mankind has ever valued?

Evaluating the the quality of a search experience is *hard*.

And… bringing new methods of customer listening at eBay including the Garden, our “labs” environment for new experiences.

Garden by eBay

Plant the seeds of new ideas. Try out these features and tell us what you think.

~~Ten~~ Nine Take-Aways

1. Match User & Business Goals
2. Model the Success Path
3. Question Non-Optimal Patterns
4. Triangulate
5. Combine Your Methods
6. Track the UI Origin of User Actions
7. Client Side Logging is Essential
8. Know Your Algorithms
9. Data is Code

Your Job: Optimize the Overlap between Business and User Goals

UX Dimensions

- Ease of use
- Error rate
- Usefulness
- Learnability
- Memorability

(Nielsen, Jakob (1990). *Hypertext and Hypermedia*. Academic Press.)

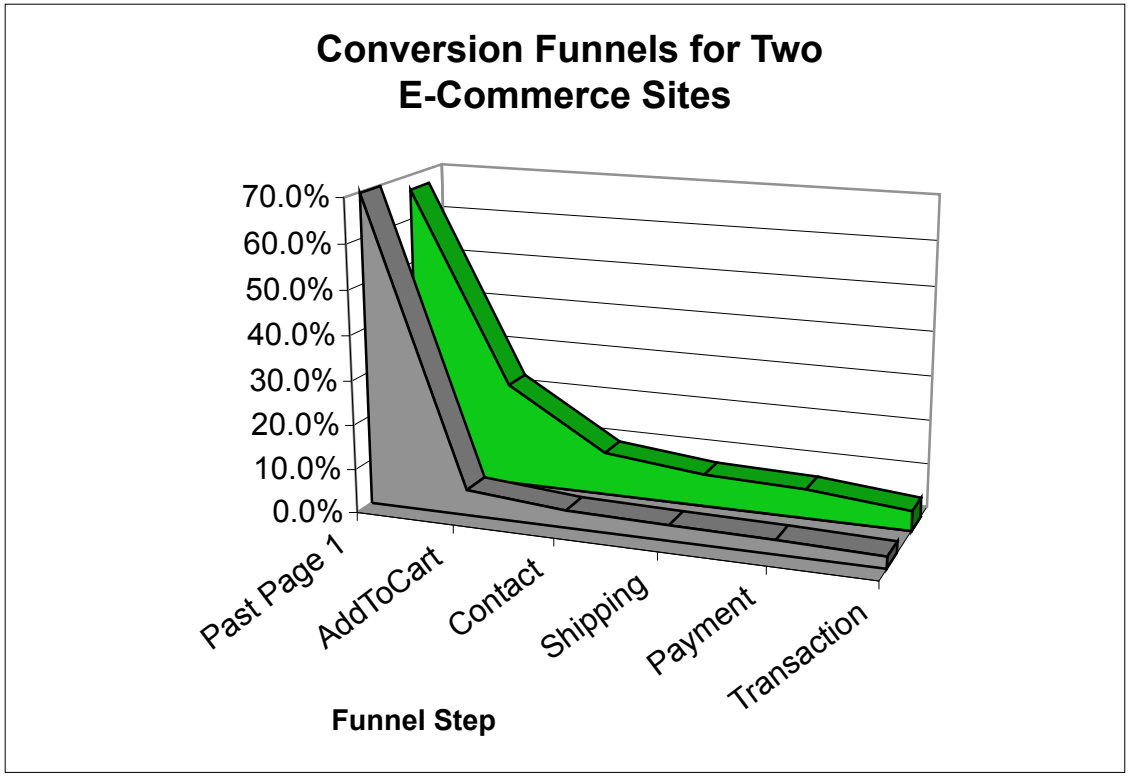
Business Dimensions

- User Engagement
- Conversion Rate
- Retention
- Customer Lifetime Value

MODEL THE SUCCESS PATH

The E-Commerce Funnel

- Model your tasks
- Focus where the ROI is greatest
- Track sub-metrics



Analytic Tools Offer Support for Basic Task Modeling

- Google Analytics :: Goals

- Or try my conversion funnel modeler:
  http://alwaysbetesting.com/abtest/tools/funnel/

Step Name (click to edit)	Conversion %	Raw Numbers	Manage Steps
Starting Point - Homepage	100%	1000	++
Homepage Clickthrough	66 %	660	++ --
Add to Cart	33 %	218	++ --
Checkout Start	58 %	126	++ --
Checkout End	60 %	76	
Total Conversion	7.6 %	Save Model	

Beyond the Success Path: Consider the Optimal Path

- P.A. Smith 1996, "Lost in Hypertext"
  - (citation space)

- From MS work at Univ. of Dalhousie

Lower values of lostness indicate more confusion and lower values of efficiency are better.

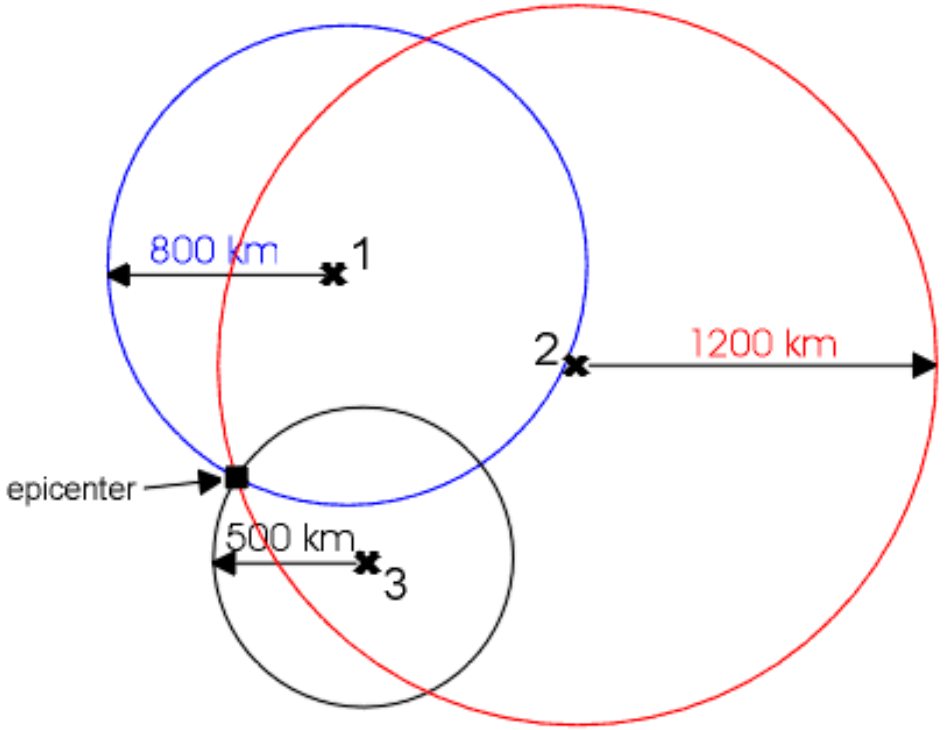
PageViews	Distinct Page Views	Lostness ↓	Efficiency
49	18	0.37	0.77
38	16	0.42	0.69
58	25	0.43	0.83
44	20	0.45	0.74
48	23	0.48	0.77
58	28	0.48	0.83
30	15	0.50	0.60

TRIANGULATE BETWEEN METHODS

Triangulation

Sources

- Event Logs
- A/B Testing
- Usability Lab
- Ethnography
- Feedback Forms
- Support Logs
- Sales Calls
- Bug Tracking

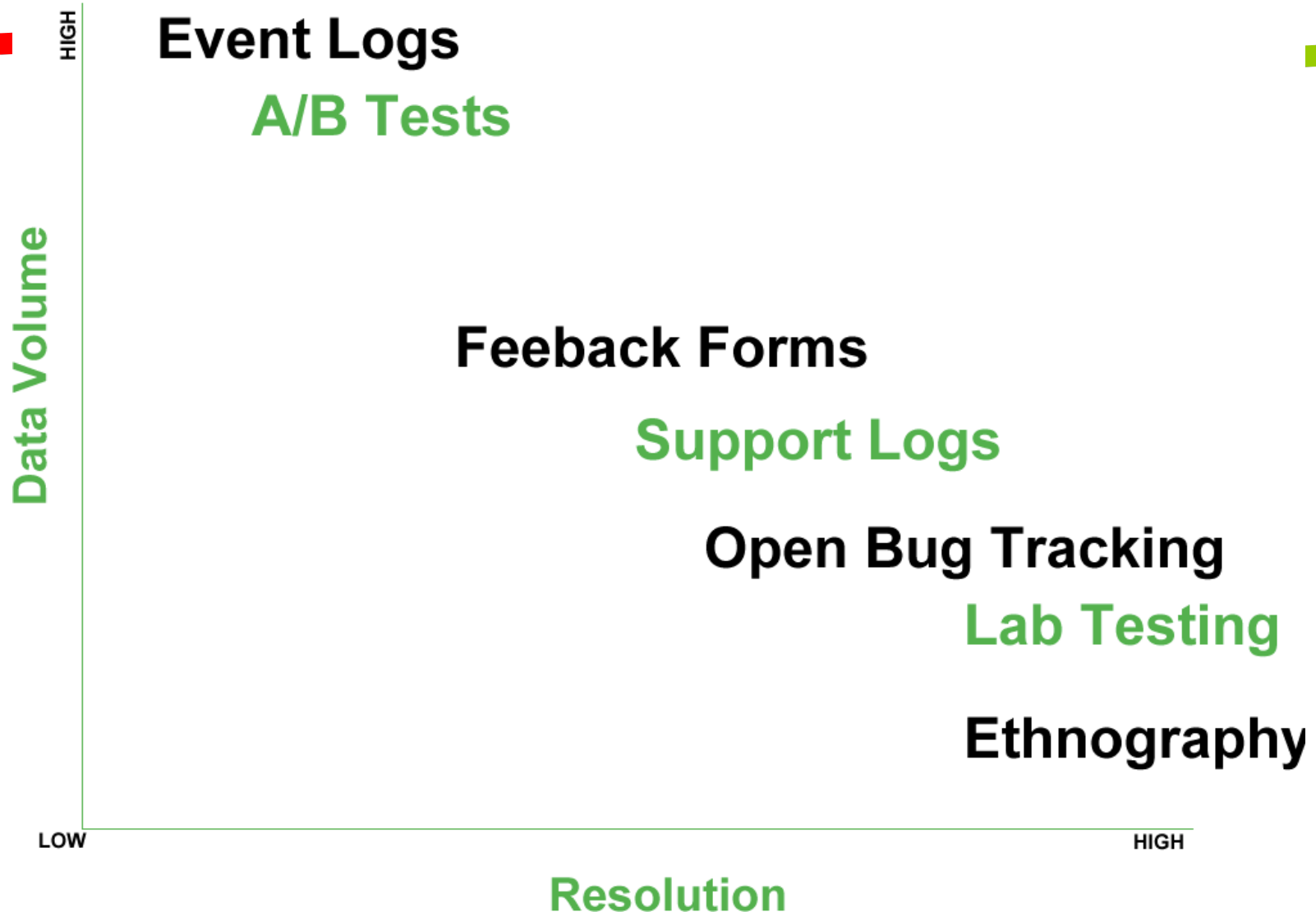


Distance to epicenter

Seismograph 1 - 800 km
Seismograph 2 - 1200 km
Seismograph 3 - 500 km

Image with permission from David R. Jessey, Geological Sciences, Cal Poly Pomona

Volume vs Resolution

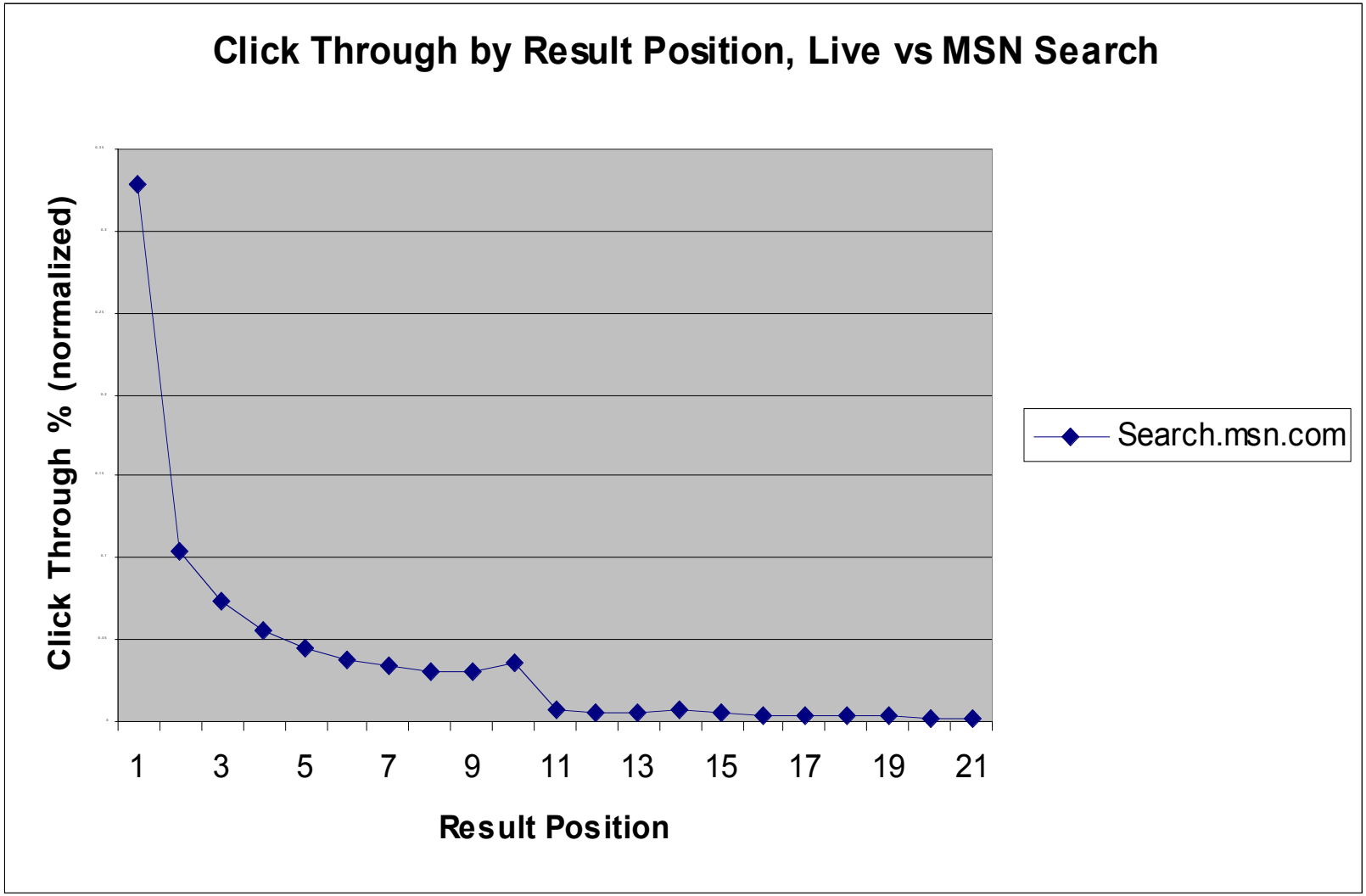


Triangulation at eBay: Pogosticking

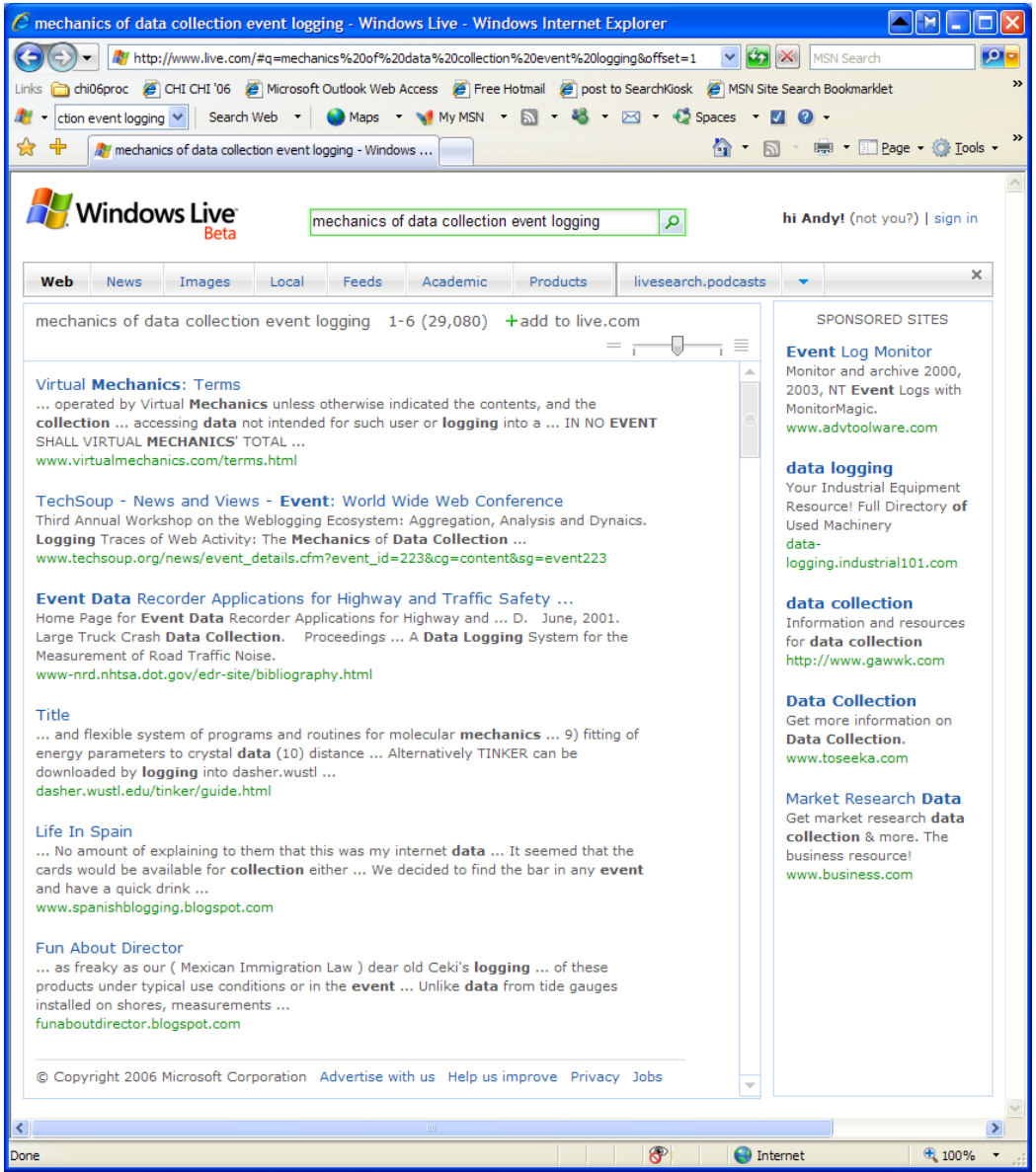
- Observation: Pogo-sticking between the Search Result Page and the View Item is a problem.
- Methods
  - Data-mining research: what categories have highest pogo-sticking?
  - Prototypes & Usability Testing: do users appreciate designs which streamline this flow?
  - The Garden: do opt-in users respond well?
  - A/B Testing: does the average user get it?

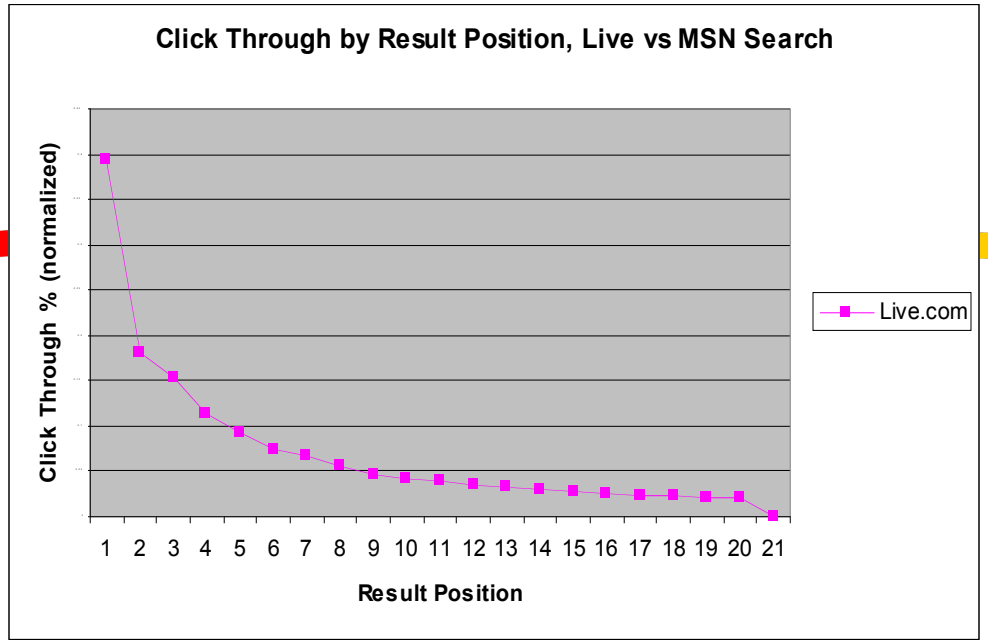
QUESTION NON-OPTIMALITIES

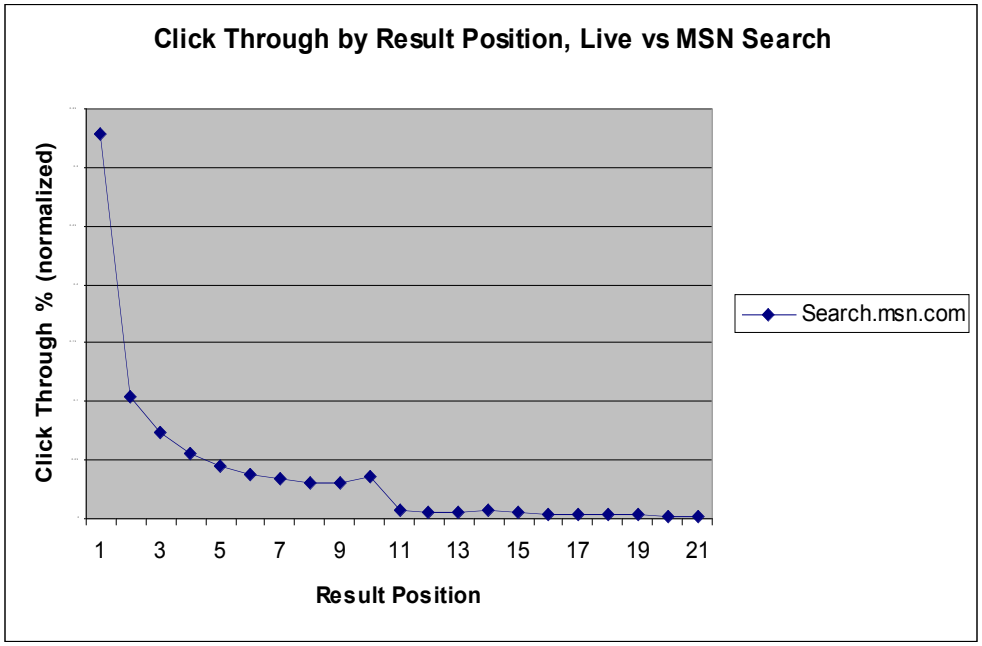
Is Rank 11 Really That Much Worse than Rank 10?



“Infinite Scroll” at Live.com







The effect on click-through by position is dramatic.

The removal of the next page link results in a much smoother function crossing the 10-11 boundary.

COMBINE YOUR METHODS

Got Usability Testing?

- Don’t let your UX lab and your Quants get out of synch!
- Make sure patterns observed in lab testing are easy to identify in web analytics.

- Using Google Analtyics, tag Usability Test Browsers with a Custom Variable
  - http://code.google.com/apis/analytics/docs/tracking/gaTrackingCustomVariables.html

Extreme Methodological Mashups

- Got a complex task where user success is hard to infer?
  - Use machine learning to predict user responses on satisfaction questions from analytic data
- Site Feedback Forms
  - With system state (page, query term, browser, etc)
  - At eBay, we even augment the user feedback with the surround page trail

CLIENT SIDE LOGGING IS ESSENTIAL

In the beginning, there was the access log

- 66.249.65.107 - - [08/Oct/2007:04:54:20 -0400] "GET /support.html HTTP/1.1" 200 11179 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"

- But then came the Browser Cache

Browser Cache Creates Blind Spots

- User visits
  - Homepage -> Article 1 -> Homepage -> Article 2
- Log says
  - Homepage -> Article 1 -> Article 2

Client Side Logging Solutions

- Mechanism
  - Javascript collects data like screen resolution
  - Makes a call to logging server, often loading an IMG
  - May make a call “onUnload” to get page timing, etc.
- Options
  - Google Analytics
  - Omniture
  - Piwik (open source)

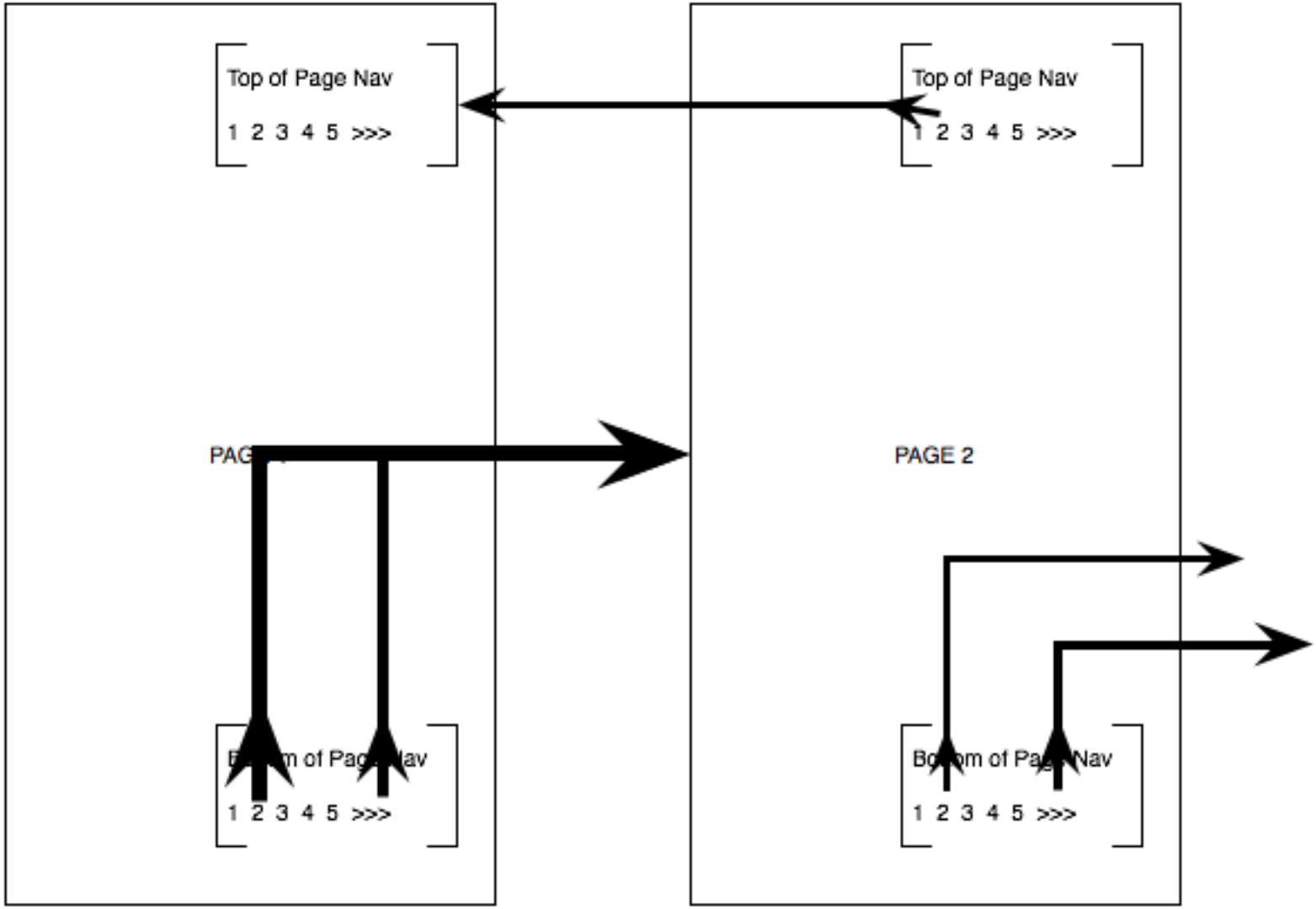
MULTIPLE UI METHODS

A Design Conundrum

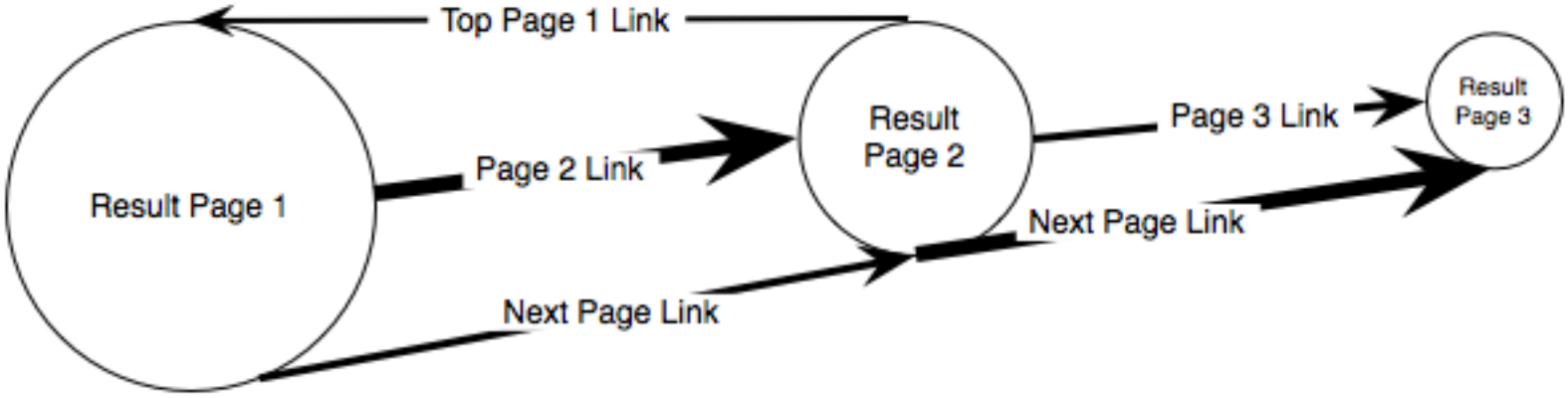
- Hick's Law, 1868
  - Time to choose an option is inversely related to # of options
  - http://en.wikipedia.org/wiki/Hick%27s_law
- Milliseconds Matter:
  - An introduction to micro-strategies and to their use in describing and predicting interactive behavior
  - With learning, users will optimize their UI interaction
  - http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.2442&rep=rep1&type=pdf

Sequence Modeling of Pagination

Comparing use of top and bottom of page navigation at MSN Search



Modeling Sequences



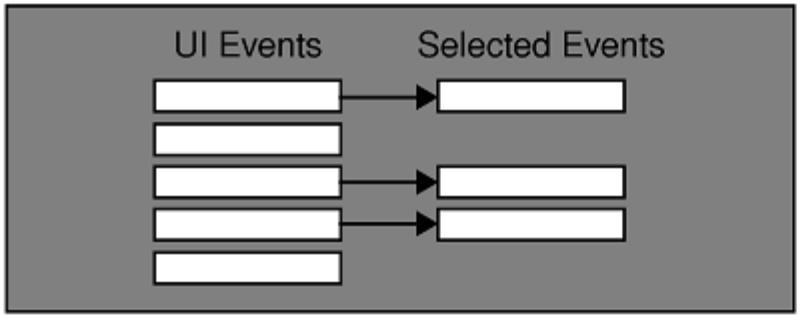
- Identify the states and transitions of interest
- Walk session ordered data
- Compute node to node transition frequencies for each path
- Present cumulative transition probability & partial probabilities

So how do you insure you can identify paths accurately?

- Pass-through parameters
  - http://www.bing.com/images?**FORM=Z9LH**
  - http://www.bing.com/images?**FORM=Z9FD**
  - Google Analytics Implementation:
    - http://alwaysbetesting.com/abtest/index.cfm/2008/4/22/Tracking-UI-Level-Links-An-Open-Source-Script
- Unload client side logging
  - For Google Analytics, use Custom Events
  - Ex. https://github.com/christianhellsten/jquery-google-analytics

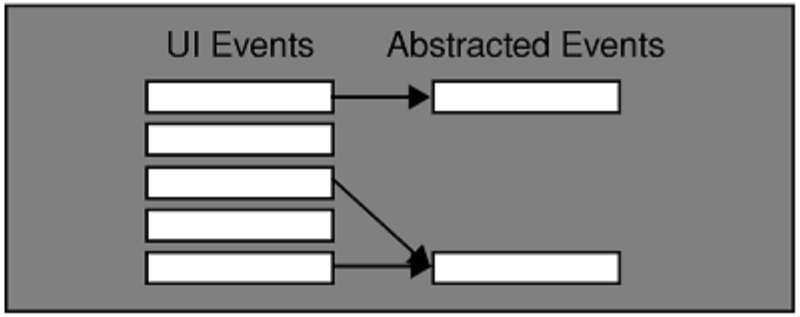
KNOW YOUR ALGORITHMS

Abstractions on Event Data



Description: Selection is the process of separating events of interest from the rest of the event stream. Recoding is the process of generating a new event stream based on selected events.

Examples: Incident Monitoring; CHIME; Hawk; MacSHAPA; User-Identified CIs; EDEM.



Description: Abstraction is the process of generating new events based on existing, or patterns of existing, events. Recoding is the process of generating a new event stream based on abstracted events.

Examples: CHIME; Hawk; MacSHAPA; User-Identified CIs; EDEM.

Fig. 5. Transformation.

– Hilbert, D.M. and D.F. Redmiles. 1999. Extracting Usability Information from User Interface Events. David M. Hilbert and David F. Redmiles. Technical Report UCI-ICS-99-40, Department of Information and Computer Science, University of California, Irvine.

Sequence Detection and Comparison

- Sequence detection:
  - Identifying bots
- Sequence Comparison:
  - Navigational vs informational queries

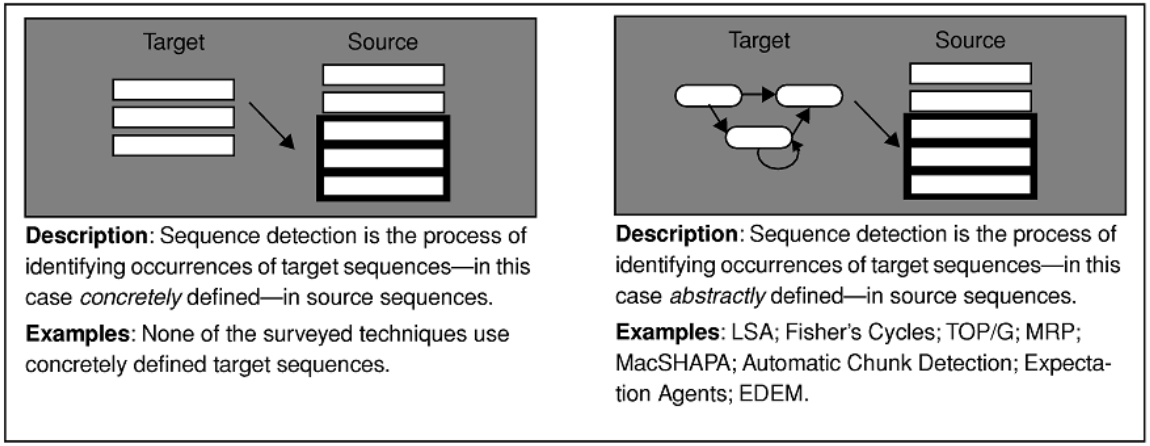


Fig. 7. Sequence detection.

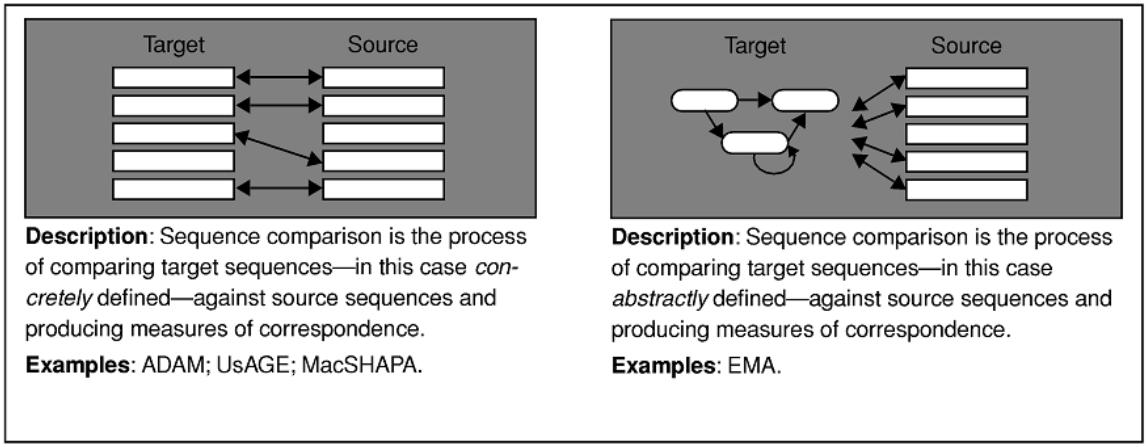


Fig. 8. Sequence comparison.

Advanced Topic: Good Structural Markup Can Make Abstraction Easier

Sample Mouse Move Event			
DOM Path	**Element ID**	**Element Class**	**Element Offset**
		dataNumber	TD(2);
	dataRow2	dataRow	TR(2);
	conversionTableBody		TBODY(0);
	conversionTable	conversionTable	TABLE(0);
			FORM(0);
	conversionForm		DIV(0);
timestamp	1529 milliseconds from page render		
Type	mousemove		
X	748 pixels		
Y	381 pixels		

From a research publication, http://research.microsoft.com/pubs/53546/edmondsjwe2007.pdf

DATA IS CODE

2008: Data is Code.
2010: Data begets Features

- Data is Code
  - http://videolectures.net/cikm08_norvig_slatuad/
- Data begets Features
  - Example: LinkedIn InMaps
  - http://blog.linkedin.com/2011/01/24/linkedin-inmaps/

Andy Edmonds, aedmonds@ebay.com
http://friendfeed.com/andyed
http://linkedin.com/in/andyed

We're hiring!

THANKS!