Measuring and Improving User Engagement

Ph. D. Proposal, Janette Lehmann
Advisors: Dr. Ricardo Baeza-Yates, Dr. Mounia Lalmas

UNIVERSITAT POMPEU FABRA

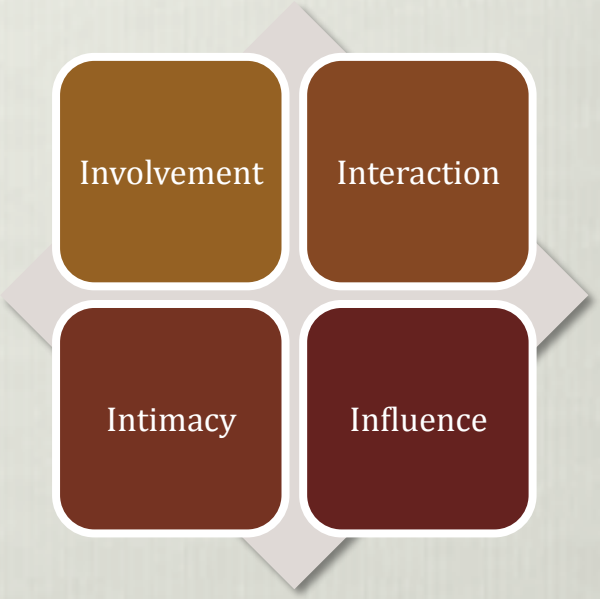
Overview

What is user engagement?

- **User engagement** is an estimate of depth of visitor interaction against a clearly defined set of goals [Peterson, 2008]

- **In-the-moment engagement**:
  users stay on a web site over a long time

- **Long-term-engagement**:
  users come back frequently and over a long-term



State of the Art

How to evaluate it?

	Measures	Characteristics
Self-reported engagement	Structured *(e.g. questionnaire, product reaction cards)* Semi-structured *(e.g. interview, report)*	Subjective, user study (lab/online)

State of the Art

How to evaluate it?

	Measures	Characteristics
Self-reported engagement	Structured *(e.g. questionnaire, product reaction cards)* Semi-structured *(e.g. interview, report)*	Subjective, user study (lab/online)
Cognitive engagement	Task-based *(e.g. dual-task)* Neurological *(e.g. EEG)* Physiological *(e.g. eye tracking)*	Objective, user study (lab/(online))

State of the Art

How to evaluate it?

	Measures	Characteristics
Self-reported engagement	Structured *(e.g. questionnaire, product reaction cards)* Semi-structured *(e.g. interview, report)*	Subjective, user study (lab/online)
Cognitive engagement	Task-based *(e.g. dual-task)* Neurological *(e.g. EEG)* Physiological *(e.g. eye tracking)*	Objective, user study (lab/(online))
Interaction engagement	Web analytics *(e.g. dwell time, return rate)* IR metrics *(e.g. precision, recall)*	Objective, data study

State of the Art

How to evaluate it?

	Measures	Characteristics
Self-reported engagement	Structured *(e.g.* ***questionnaire****, product reaction cards)* Semi-structured *(e.g. interview, report)*	Subjective, user study (lab/online)
Cognitive engagement	Task-based *(e.g. dual-task)* Neurological *(e.g. EEG)* Physiological *(e.g.* ***eye tracking****)*	Objective, user study (lab/(online))
Interaction engagement	**Web analytics *(e.g. #users, dwell time)*** IR metrics *(e.g. precision, recall)*	Objective, data study

State of the Art

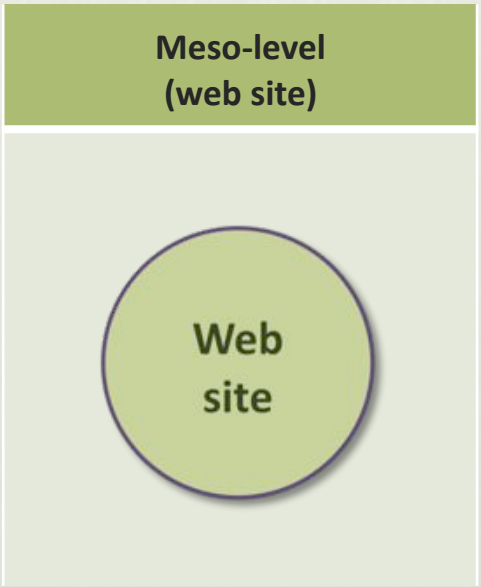
What influences user engagement?

- Influence of **service features** (e.g. #links, #pictures, text length) on **quality attributes** (e.g. usability, satisfaction, aesthetic, engagement)
- Relations between **service features** and **dwell time**
- Evaluation of web services with the **same content** but a **different presentation/functionality**

- *Importance of service features depends on the type of service [4]*
- *Influence on quality attributes can be contradictory [4]*
- *Quality attributes have dependencies (Halo Effect) [3,4]*

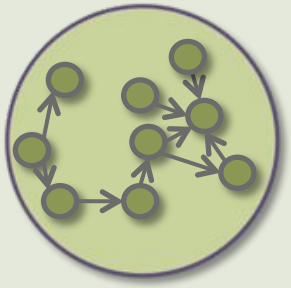
Research Problem

Three perspectives of engagement



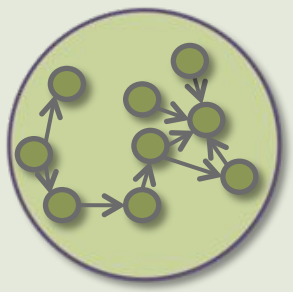
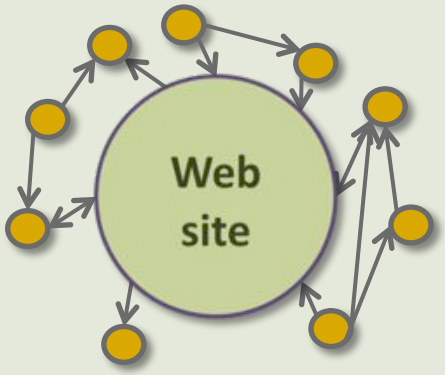
Research Problem

Three perspectives of engagement

Micro-level (web site parts)	Meso-level (web site)
	

Research Problem

Three perspectives of engagement

Micro-level (web site parts)	Meso-level (web site)	Macro-level (further web sites)
		

Research Problem

Goals of the thesis:

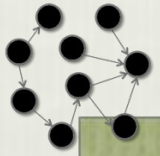
1. Investigation of the engagement of a web site in a more fine-grained way
2. Enhancement and reformulation of existing engagement metrics
3. Investigation of the effect of user demographics
4. Study of the aspects of page features of web sites that influence engagement

Methodology

Techniques of the thesis:

Web analytics

- Web metrics
  - #Users
  - Dwell time
  - Click-depth
  - Return rate
  - …



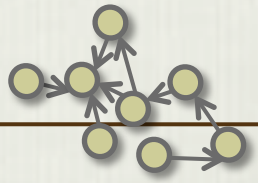
Complex networks

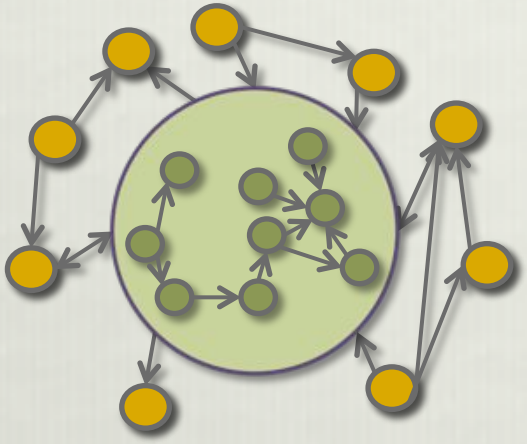
- Representation of traffic data as networks
- Network/Node metrics
- Bow-tie structure
- Clustering

Interaction patterns

- Typical user navigation paths
- (Non)-successful paths
- Entry-, exit-, return-points

Methodology





Engagement networks

- **Nodes** represent services, topics or functionalities
- **Edges** represent the traffic between nodes
- **Special nodes** represent external web sites
- **Labels** save special node/edge characteristics (e.g. location, user engagement metrics)

Methodology

Network metrics

- Node count
- Edge count
- Density
- Diameter
- Clustering coefficient
- …

Node metrics

- Degree
- Betweenness
- Page rank
- Strength
- Local clustering coefficient
- …

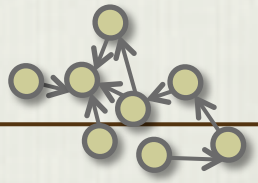
Engagement metrics

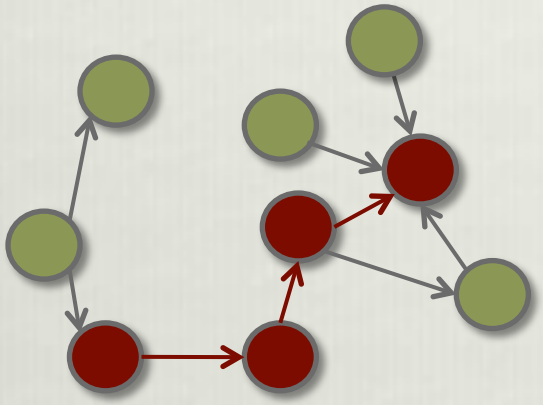
- #Users
- Dwell time
- Click-depth
- Return rate
- …

Networks analysis

- Calculation of network, node and engagement metrics
- Find interpretations for user engagement
- Study of the relationships between metrics
- Clustering of the network

Methodology

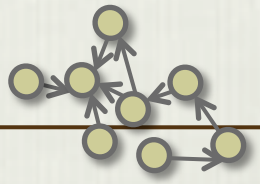




Interaction pattern

- Identification and evaluation of typical navigation paths of the web site
- Consider entry-, exit- and return-points
- Identification of different types of users

Status of Research



Engagement networks

- Creating **engagement networks** based on traffic data from Yahoo!
- Nodes: Yahoo! services, Edges: Number of clicks between services
- Movements from/to the environment (EXT -> INT, INT -> EXT) and returns (INT -> EXT -> INT)

Proportion of nodes	
INT	0.09%
EXT	99.91%
Proportion of edges	
INT -> INT	0.01%
EXT -> INT	49.78%
INT -> EXT	48.48%
INT -> EXT -> INT	1.73%

Status of Research

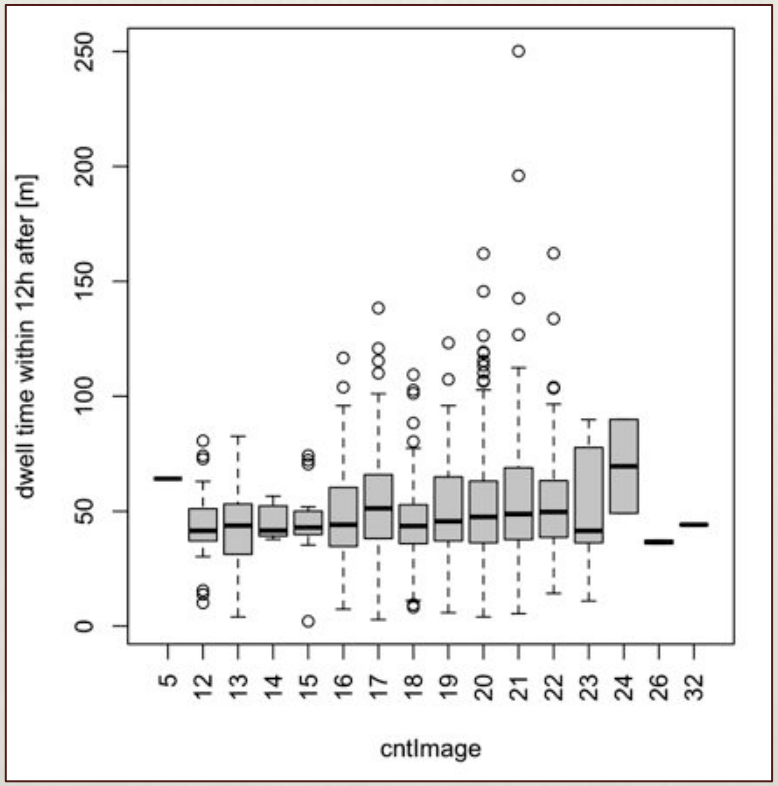
Engaging content

- Study the influence of **story features** on **downstream engagement**
  - Story features: editorial, entity, image, link, sentiments, style, syntax, textual
  - Downstream engagement: dwell time within Yahoo! afterwards (1h, 12h, 24h)

- **Pearson's correlation coefficient** is used to find relationships between story features and downstream engagement

Status of Research

Engaging content



Yahoo! News, Canada

Conclusion

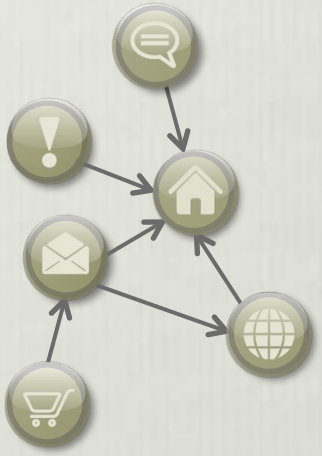
- **Objective:**
  Improving the evaluation and manipulation of user engagement

- **Field of activity I:**
  Creation and study of engagement networks based on the micro-, meso- and macro-level

- **Field of activity II:**
  Relationships between service features (e.g. #links, #pictures, text length) and engagement

Thank you for your attention!

Any Questions?

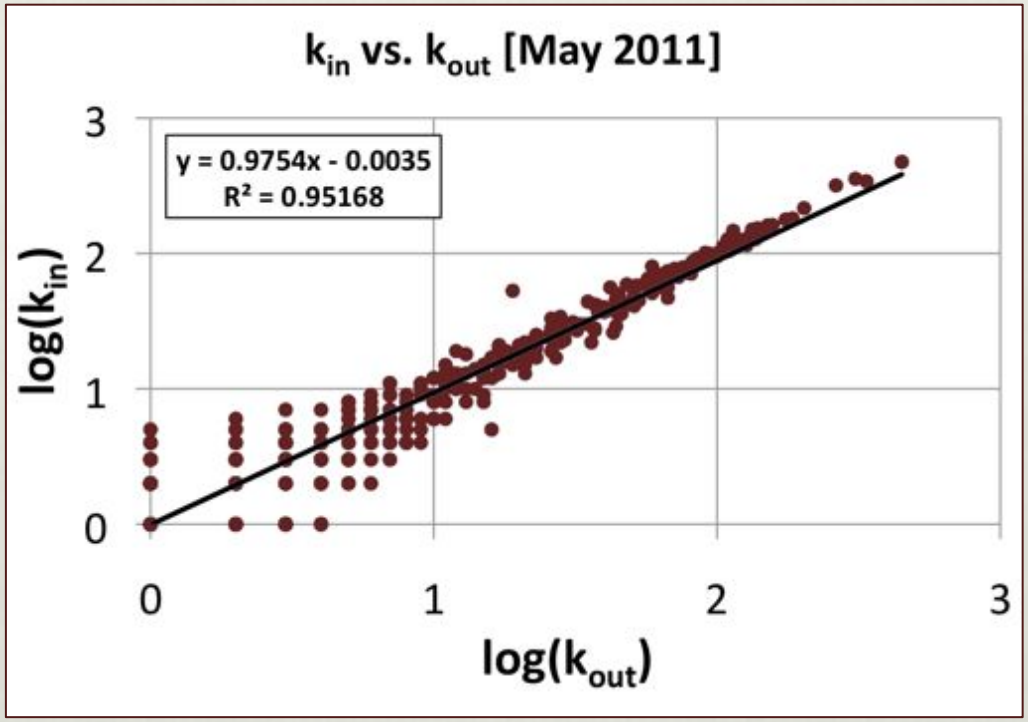


References

1. Brian Haven and Suresh Vittal. *Measuring engagement.* Forrester Research, June 2008.
2. Brian Haven. *Marketing's new key metric: Engagement*. Forrester Research, August 2007.
3. N Tractinsky. *What is beautiful is usable*. Interacting with Computers, 13(2):127–145, 2000.
4. Jan Hartmann, Alistair Sutcliffe, and Antonella De Angeli. *Towards a theory of user judgment of aesthetics and user interface quality.* ACM Transactions on Computer-Human Interaction (TOCHI), 15:15:1–15:30, December 2008.
5. Elizabeth Churchill. *Enticing engagement.* interactions, 17(3):82–87, 2010.
6. Eric T. Peterson and Joseph Carrabis. *Measuring the immeasurable: Visitor engagement.* Technical report, Web Analytics Demystified, 2008.
7. AL Barabasi and R Albert. *Emergence of scaling in random networks.* Science (New York, N.Y.), 286(5439):509–512, 1999.
8. Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. *Graph structure in the web: experiments and models.* 2000.
9. Angeles Serrano, Ana Maguitman, Mari´an, Santo Fortunato, and Alessandro Vespignani. *Decoding the structure of the www: A comparative analysis of web crawls*. ACM Trans. Web, 1(2), 2007.
10. Mark R. Meiss, Filippo Menczer, Santo Fortunato, Alessandro Flammini, and Alessandro Vespignani. *Ranking web sites with real user traffic.* In Proceedings of the international conference on Web search and web data mining, WSDM '08, pages 65–76, New York, NY, USA, 2008. ACM.
11. Chao Liu, Ryen W. White, and Susan Dumais. *Understanding web browsing behaviors through weibull analysis of dwell time*. In Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval, SIGIR '10, pages 379– 386, New York, NY, USA, 2010. ACM.
12. G. Paliouras, C. Papatheodorou, V. Karkaletsis, C. Spyropoulous, and P. Tzitziras. *From web usage statistics to web usage analysis*. In IEEE International Conference on Systems Man and Cybernetics, pages 159–164, 1999.

Status of Research

Engagement networks



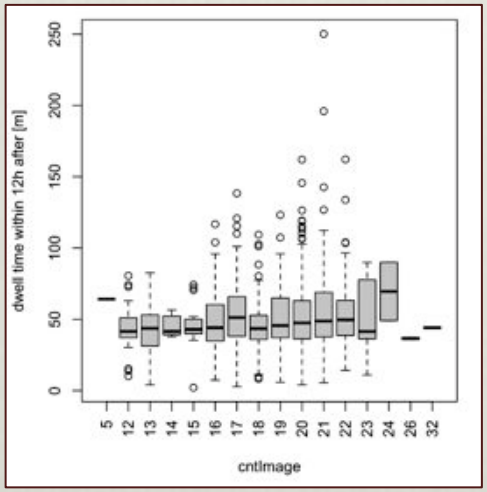
Degree distribution, Yahoo! network

Status of Research

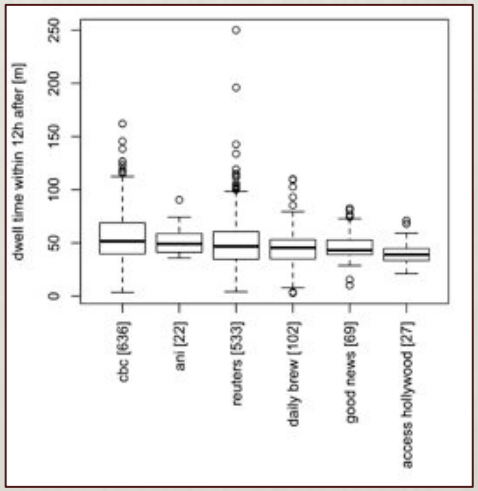
Engaging content

- Study the influence of **story features** on **downstream engagement**
  - Downstream engagement: dwell time within Yahoo! afterwards (1h, 12h, 24h)
  - Story features: editorial, entity, image, link, sentiments, style, syntax, textual
- **Pearson's correlation coefficient** is used to find relationships between story features and downstream engagement

Yahoo! News, Canada



Yahoo! News, UK



State of the Art

Common metrics for user engagement:

Involvement	Interaction	Intimacy	Influence
• #Users • #Visits • Dwell time • Click-depth • Return rate • Site logins	• Contributed content to blogs • Quantify/ frequency of written reviews, blog comments, etc.	• Sentiment tracking on third-party sites (blogs, reviews, etc.) and internal customer contributions • Like-Buttons	• Product/service ratings • Content forwarded to friends • Service invitations sent to friends

[Brian Haven, 2007]

Research Problem

Motivation

- Actual studies are **not precise** enough
  - Web sites can contain several services, functionalities and topics
  - Movements from and to other web sites have to be considered
  - Investigation of navigation paths of users
  - Differences concerning demographics
- Influence of page features on the whole web site is unknown
  - Study of the **downstream engagement**