



Linked Learning 2011

Using Linked Data to Reduce Learning Latency for e-Book Readers

Julien Robinson, Johann Stan, Myriam Ribière
29 May 2011



introduction

- reduce learning latency = time to integrate knowledge
- assumptions:
 - social learning is efficient
 - learn with other learners: need of common environment
 - the book is central to the learning process
- ⇒ conclusion: use web2.0 e-books for learning
- objective:
 - create a social book, the "sBook"
 - social features for learning:
 - annotate, share annotations
 - awareness of other learners
 - social navigation

state-of-art domains, **our focus**

analyse annotations
⇒ navigation

LibraryThing
Shelfari
BookGlutton,
GoodReads
Copia

shared annotations

share annotation

short text
+ unstructured data

Ali Harb & al. (2010)
H.-L. Kim & al. (2008)

semantic analysis
document classification

text analysis

semantic analysis

university of Pittsburg
Rosta Farzan, Peter Brusilovsky
Jae K. Kim & al. (2008)
Jill Freyne & al. (2007)

experiments in education
interactive design
social navigation
symbolic annotations

social navigation

the sBook



the sBook: overview

The screenshot displays the sBook application interface. At the top, there is a navigation bar with links for Dashboard, Tools, Settings, and Administration, along with a search bar and a Log out button. The main content area is titled "Frankenstein, by Mary Wollstonecraft Shelley" and shows "Chapter 1". The text of the chapter is displayed in a large font. A blue box highlights the phrase "community of the paragraph" within the text. To the right of the text, there is a vertical sidebar with a "View notes for" dropdown menu set to "All". Below this, there are three sections labeled "to annotation board", each containing small thumbnail images of people. A blue circle highlights the middle section of the annotation board. On the far right, there are vertical tabs for "Readers", "Annotations", and "Videos".

the heatmap: social navigation

...based on analysis of annotations...

The screenshot shows the sBook interface with a heatmap overlay. The heatmap highlights three chapters in the table of contents: Chapter 6, Chapter 11, and Chapter 24. A blue box at the top left contains the text "user goal: study the theme of 'children'", with blue arrows pointing to the highlighted chapters. A blue box at the bottom left contains the text "where readers evoke 'children'", with blue arrows pointing to the highlighted chapters. The main content area shows Chapter 1, with a heatmap overlay on the text. The right sidebar shows a "View notes for:" section with three annotations, each with a "to annotation board" link and a small image. The sidebar also has sections for "Readers", "Annotations", and "Videos".

where readers evoke "children"

analysis of short unstructured annotations

- short unstructured messages: mail, tweets, comments, annotations,...

Linked Data: structured information (e.g. DBpedia) about...
 people, companies, books, geographic locations, entertainment media
 & broadcasting, medical & biological data, online communities,...

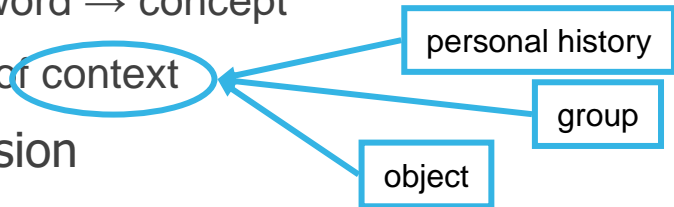
- our aim:

- annotation → "tag cloud" = vector of ponderated concepts from Linked Data sources
- ⇒ enable to leverage annotations as structured data

- our analysis:

- vocabulary extraction (NLP)
- disambiguation
 - keyword → concept
 - use of context
- expansion

NLP: Natural Language Processing



example

"I like Facebook, the social network"

Facebook: 62.962962962963
 Networks: 60.740740740741
 Sociology: 57.058823529412
 American Jews: 53.492063492063
 American billionaires: 44.694862979068
 Social systems: 44.444444444444
 Student culture: 44.426406926407
 Social networks: 43.846120984279
 Web 2.0: 42.5
 Social sciences: 42.261904761905
 Facebook employees: 41.107103422893
 Social media: 41.025641025641
 Network theory: 36.947712418301
 American atheists: 36.899415204678
 Cultural economics: 36.842105263158
 Systems theory: 35.614379084967
 Social network services: 33.477650063857
 Value: 33.333333333333
 Social psychology: 31.642512077295
 Graph theory: 31.632837750485
 Living people: 29.65873015873
 Community building: 28.800915331808
 Harvard University alumni: 27.501780626781
 American computer programmers: 24.750196625197

heatmap construction

- semantic heatmap:
 1. get user goal (keyword, phrase)
 2. user goal → concept vector
 3. concept vector → expansion → keywords
 4. define $\text{heat}(\text{annotation}) = \text{similarity}(\text{annotation}, \text{expanded keywords})$
 5. $\text{heat}(\text{parag}) = \text{sum}(\text{heat}(\text{annotations}))$

- other heatmaps
 - statistical
 - keyword search
 - **dictionary-based**
 - user group (class, friends)
 - time-based (most recent)
 - contextual (location,...)
 - all of the above!

The screenshot shows a web interface for heatmap construction. It features several sections: 'Activate heatmap' with a checked checkbox; 'Restrict to group' with an empty text input; 'Search for' with a text input containing 'children'; 'Use semantic expansion' with a checked checkbox and a 'Clear semantic cache!' button; 'Use dictionary' with three sub-options: 'Synonyms', 'Derivatives etc.', and 'Antonyms', each with an unchecked checkbox; 'String comparison method' with a dropdown menu set to 'cosine similarity' and a note below it: 'cosine = case-aware, num occ = caseless'; and 'Canonize annotations' with an unchecked checkbox. A 'hide' button is located at the bottom left. A blue circle highlights the 'Search for' input field, and a blue arrow points from the first step of the list above to this field.

the sBook: heatmap settings

The screenshot displays the sBook interface with several key elements highlighted by blue callouts:

- user goal**: A box pointing to the search input field containing the text "children".
- expansion & comparison options**: A box pointing to the settings panel on the right, which includes:
 - Use semantic expansion
 - Use dictionary
 - String comparison method: cosine similarity
 - Options for Synonyms, Derivatives etc., and Antonyms.
- where readers evoke "children"**: A box pointing to the left sidebar's table of contents, specifically highlighting Chapter 24.

The main content area shows search results for "children" within a text passage from a book. The text includes phrases like "I am by t...", "many yea...", "was resp...", "younger c...", "early, nor...", "As the cir...", "friends w...", "name was...", "same cou...", "in the mo...", "wretched...", "unfortunate circumstances. He bitterly deplored the false pride which led his friend to a conduct so little worthy affection that united them. He lost no time in endeavouring to seek him out, with the hope of persuading him to world again through his credit and assistance. Beaufort had taken effectual measures to conceal himself, and it months before my father discovered his abode. Overjoyed at this discovery, he hastened to the house, which w situated in a mean street near the Reuss. But when he entered, misery and despair alone welcomed him. Beaufort had saved but a very small sum of money from the wreck of his fortunes, but it was sufficient to provide him with sustenance for some months, and in the meantime he hoped to procure some respectable employment in a merchant's house. The interval was, consequently, spent in inaction; his grief only became more deep and rankling when he had leisure for reflection, and at length it took so fast hold of his mind that at the end of three months he lay on a bed of sickness, incapable of any exertion.

His daughter attended him with the greatest tenderness, but she saw with despair that their little fund was rapidly decreasing and that there was no other prospect of support. But Caroline Beaufort possessed a mind of an uncommon mould, and her courage rose to support her in her adversity. She procured plain work; she plaited straw and by various means contrived to earn a pittance scarcely sufficient to support life.

the sBook: find relevant annotations

The screenshot displays the sBook interface with several key elements:

- Text:** A passage from a book is shown, with several lines highlighted in pink. The highlighted text includes: "at the last evening with us. He had endeavoured to persuade his fellow student, but in vain. His father was a narrow-minded ambition of his son. Henry deeply felt the misfortune of being when he spoke I read in his kindling eye and in his animated glance miserable details of commerce.", "ch other nor persuade ourselves to say the word 'Farewell!' It repose, each fancying that the other was deceived; but when at to convey me away, they were all there—my father again to abeth to renew her entreaties that I would write often and to friend.", and "way and indulged in the most melancholy reflections. I, who had ally engaged in endeavouring to bestow mutual pleasure—I was form my own friends and be my own protector. My life had his had given me invincible repugnance to new countenances. I old familiar faces," but I believed myself totally unfitted for the
- Note Editor:** A window titled "Note editor" is open, showing a text input field with the text: "then promises to write often. All of these events display Frankenstein's great love for his family and friends." The word "family" is circled in blue. Below the text field are buttons for "Upload a related video", "Browse...", "Upload Your Video", and "Capture with your webcam".
- Annotations:** On the right side, there are three "to annotation board" buttons, each with a small image icon.

someone's annotation

where readers evoke "children"

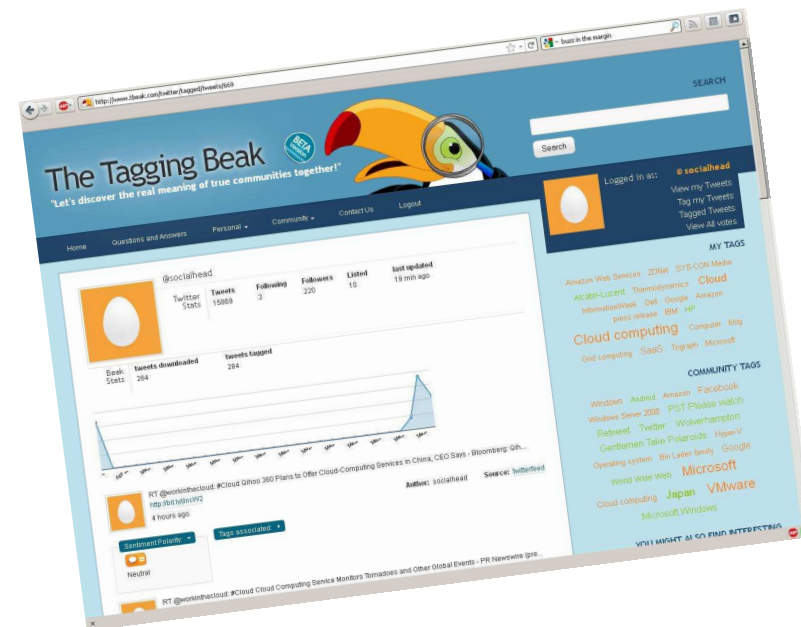
"family" = concept linked to "children"

need a keyword?

**build learning communities: user
interaction profiles**

interaction profile construction tool: the “tagging beak”

- prototype: the Tagging Beak
 - web service based on Twitter
 - analysis of short unstructured messages
 - “group” context: based on followers (social)
- currently online (beta version)
 - produces tag clouds
 - examine user engagement in topics



the Tagging Beak: overview

http://www.tbeak.com

The screenshot displays the Tagging Beak interface with three tweet analysis cards and a large tag cloud overlay. Each tweet card includes a profile picture, text, a link, a timestamp, a sentiment polarity indicator, and associated tags.

Tweet 1: RT @workinthecloud: #Cloud Qihoo 360 Plans to Offer Cloud-Computing Services in China, CEO Says - Bloomberg: Qih...
<http://bit.ly/jIncW2>
 4 hours ago
 Sentiment Polarity: Neutral
 Tags associated: Semantic Tags: StarHub TV, American businesspeople, China, Blo... Offer Cloud-Computing Services; Hashtags: Cloud

Tweet 2: RT @workinthecloud: #Cloud Cloud Computing Service Monitors Tornadoes and Other Glob...
<http://bit.ly/kSm1sM>
 4 hours ago
 Sentiment Polarity: Negative
 Tags associated: Semantic Tags: Cloud computing, Preston, cloud computing servic... Newswire; Hashtags: Cloud

Tweet 3: RT @workinthecloud: #Cloud No Wonder Small Businesses Are Confused About the Cloud...
<http://bit.ly/kDkL6p>
 8 hours ago
 Sentiment Polarity: Negative
 Tags associated: Semantic Tags: Small business, Business, Europe, Wonder Smal... UKNo, PCWorld, eWeek; Hashtags: Cloud

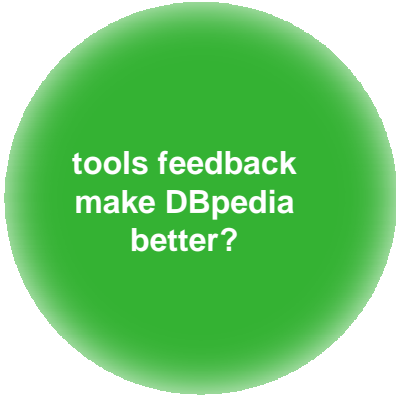
Tag Cloud (MY TAGS):
 IBM, Amazon Web Services, Microsoft, InformationWeek, HP, Alcatel-Lucent, ZDNet, SaaS, Amazon, Trigraph, Cloud computing, Cloud, Grid computing, Google, Computer, press release, Thermodynamics, Dell, SYS-CON Media, blog

COMMUNITY TAGS:
 Twitter, Microsoft Windows, Retweet, Android, Gentlemen Take Polaroids, Bin Laden family, World Wide Web, Hyper-V, Operating system, Google, Microsoft, Cloud computing, Amazon, Windows, Wolverhampton, Facebook, PST Please watch, Japan, VMware, Windows Server 2008

issues & perspectives

(controversial stuff inside)

issues & perspectives: DBpedia



- issues with use of DBpedia (based on wikipedia):
 - based on hyperlinks, the less verified content of Wikipedia?
 - based on constant popular edition; unpredictable: size and precision of articles, ambiguity management.
 - example:
 - in September 2010, a search for "Android" did not yield the concept "Android" as result, because the label was not defined in English; in March 2011 two of the results had disappeared ("Android Lust" and "Paranoid Android"). The concept "Android" did not have a `dcterms:subject`.
 - March 2011, search for "Appearance" yields only one result: disambiguation leading to a film called "appearances" (plural!). No link, disambiguation or redirection, to Visual Appearance, which would seem to be the logical choice, and is the first hit of the wikipedia disambiguation page; BTW, we lost a musical album compared to September 2010.

- volume of DBpedia (many link types and data!)

- choose appropriate graph exploration
- determine distances between concepts
- **determine popularity of concepts** (à la TF/IDF)

e.g. a town is linked to all the famous people born, dead or living in that town; many keywords lead to music groups or albums

- perspectives: other Linked Data sources?

issues & perspectives: expansion algorithm

- choice of expansion algorithm for students
 - we would like "Paris" to be linked to "France" ⇒ use of ontologies & Linked Data
 - The dictionary yields no results
 - we would also like "character" to be linked to "personage", "reputation" ⇒ use of dictionaries
 - DBpedia yields no such results
 - how to combine results?
 - how to restrict to meaningful results, or ponderate results? How to evaluate?
- to-do: **live experiment** to compare several algorithms

called "named entities"
in the short paper

called "literary theory
terms" in the short paper

conclusion

- we are researching social navigation, in particular the use of free text analysis, for efficient education with e-books
- heatmap for social navigation
 - utility & performance testing with ACU students scheduled in october 2011
- interaction profiles
 - reader profiles
 - book profiles




references & acknowledgements

- references:

- Jae K. Kim, Rosta Farzan, and Peter Brusilovsky. "Social navigation and annotation for electronic books". In "BooksOnline '08: Proceeding of the 2008 ACM workshop on Research advances in large digital book repositories", pages 25-28, New York, NY, USA, 2008. ACM.
- John Lees, "The world in your own notebook". In "The Best of Creative Computing Volume 3", pages 5-7, 1980
- Jens Lehmann, Chris Bizer, Georgi Kobilarov, Sren Auer, Christian Becker, Richard Cyganiak, and Sebastian Hellmann. "DBpedia - a crystallization point for the web of data". Journal of Web Semantics, 7(3):154-165, 2009.
- Ali Harb, Michel Beigbeder, Jean-Jacques Girardot, "Classification Supervisée de Questions: Rôle de l'Expansion Sémantique", EMSE-G2I, 2010
- H.-L. Kim, J. Breslin, S.-K. Yang, and H.-G. Kim. Social semantic cloud of tag: Semantic model for social tagging. Agent and Multi-Agent Systems: Technologies and Applications, pages 83-92, 2008
- Myriam Ribière, Jérôme Picault, and Sylvain Squedin. "The sbook: towards social and personalized learning experiences". In "Proceedings of the third workshop on Research advances in large digital book repositories and complementary media", BooksOnline '10, pages 3-8, New York, NY, USA, 2010. ACM.

- acknowledgements:

- thanks to all in Bell Labs who participate or provide items for this work
- thanks to BookGlutton for their dataset, in particular Aaron Miller (CTO) 
- The set of ideas about the future of e-books we have presented in this paper are currently being addressed in the scope of a collaborative project between Alcatel-Lucent, Abilene Christian University and Cambridge University Press



Thank you!

Questions?

AT
THE
SPEED
OF
IDEAS

AT
THE
SPEED
OF
IDEAS

www.alcatel-lucent.com