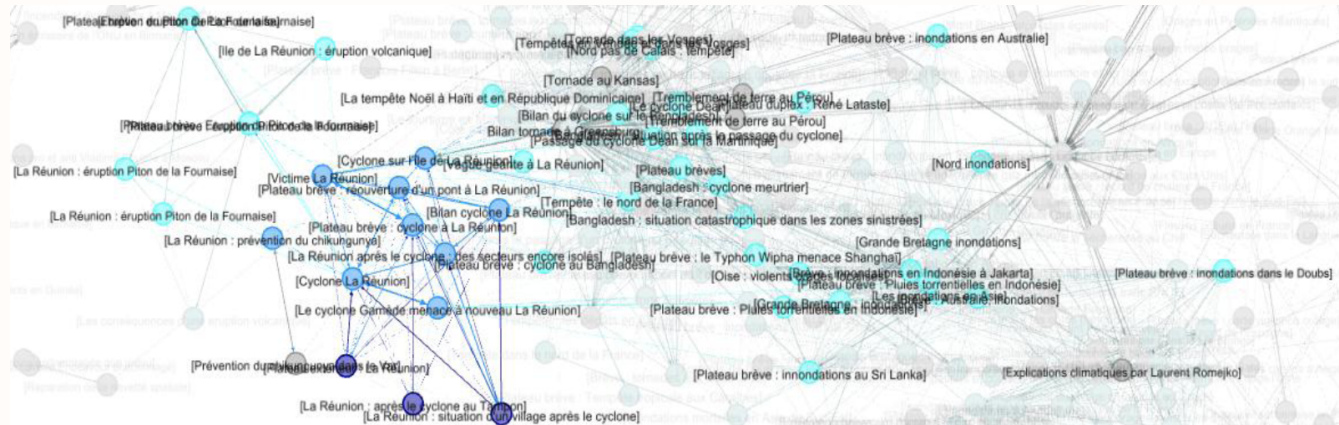


Linking Media in Acceptable Hypergraphs

(LIMAH = )



PREFics EA 4246



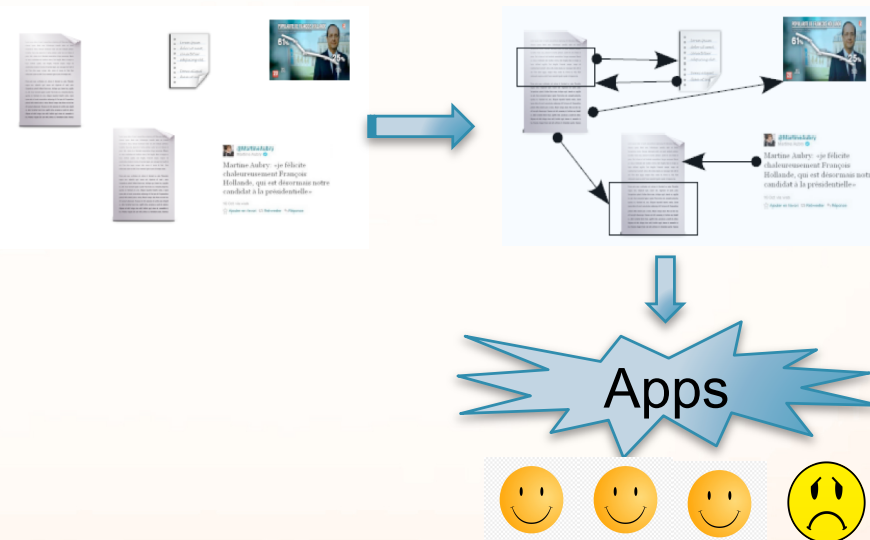
Period: 2014—2018



Project overview

General idea: content-based hypergraph construction aware of acceptability and usage

- From a collection, how to automatically build an hypergraph that provides exploitable links?
 - ➔ *segmentation, linking*
- How collections with explicit links modify usage?
 - ➔ *acceptability, law, practice*



Two application domains:

- **Education** = facilitate learning from audiovisual material
 - ➔ *investigation on segmentation and impact*
- **News** = explore collections to gain insight on a topic/event
 - ➔ *investigation on linking and impact*



Main scientific contributions



Segmentation

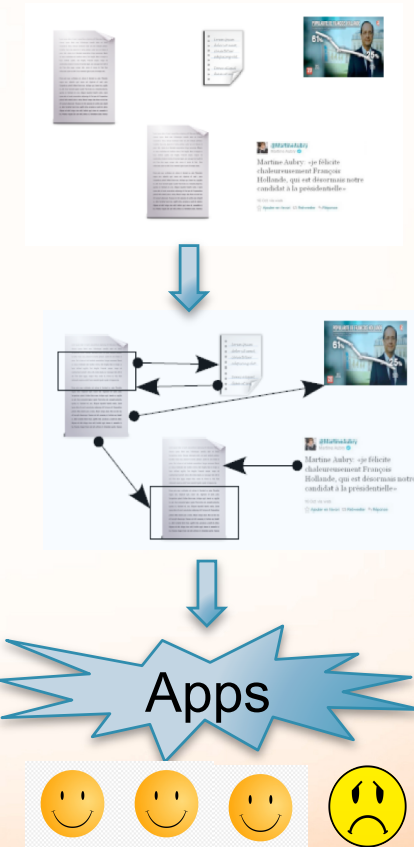
- Micro and macro scaffolding in learning

Description and linking

- Opinion mining
- Multimodal hyperlinking
- Navigable multimedia graphs

Usage and practices

- LIMAH news navigation interface
- Study of practices with linked media
- Characterization of the legal status of hyperlinks



Scaffolding for educational videos



- Specific difficulties encountered by students
 - Important to perceive content organization inside each video
 - Not crucial to link different videos as for news (no different techno)
- The technology used in LIMA H also allows segmenting and visualizing the semantic structure of a video
- **Are these scaffolding tools useful to improve information seeking or learning?**

Table des matières

- 1: Introduction
- 2: Paroles de parents
- 3: Paroles de chercheurs
- 4: Diagnostiquer : de l'orthophoniste au centre de référence
- 5: La théorie phonologique
- 6: La théorie visuelle
- 7: La théorie de la migration des neurones
- 8: La théorie génétique
- 9: Nécessité de repérer, dépister au diagnostiquer
- 10: Les méthodes de rééducation et d'entraînement
- 11: Les adaptations en milieu scolaire
- 12: Informer les professionnels et le public
- 13: Conclusion
- 14: Génétique

Passer au questionnaire

07:47/18:03

1 2 3 4 5 6 7 8 9 10 11 12 13 14

Inserm, Scientifilms (2009). Expertise collective : Dyslexie, dysorthographe, dyscalculie.



Scaffolding for educational videos



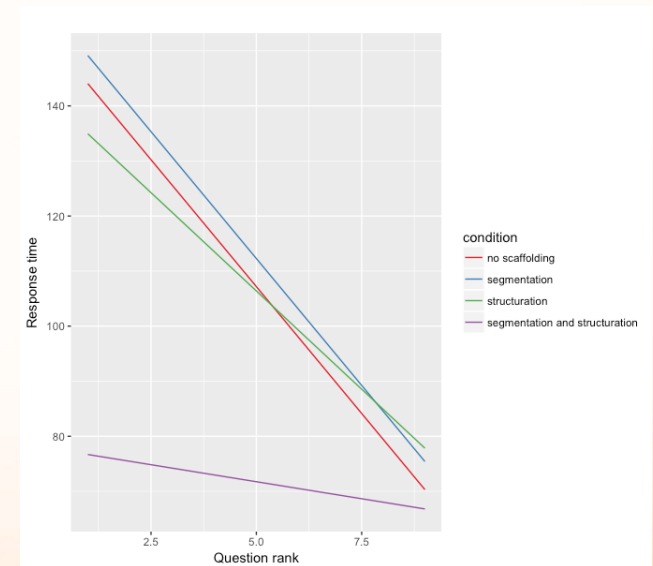
- 5 experimental studies with a total of 410 students involved
 - Controlled environment with pre-segmented videos
- Evaluation of the effects of segmentation and structure visualization for

1. Information seeking (studies 1 to 3)

- Positive effect on response time, task success, etc.
- Only when segmentation and structure both displayed

2. Learning task (studies 4 & 5)

- Positive effect on memorization
- Only when segmentation and structure both displayed
- Only when learner is engaged in an active process
 - i.e., seeking task before learning task



Linking/enrichment implemented in a companion project with (almost) standard text retrieval techniques



Opinion mining



- Exploiting comments on news articles
- Polarity
 - ✓ positive/negative/neutral/mixed
- Target and aspect
 - ✓ economical, political, migration, etc.

Yet another [#Brexit](#) problem only gaining traction now: the consequences EU [#copyright](#) law no longer applying in UK



Brexit: important changes if EU copyright law lapse...

The UK is at risk of falling out of the EU's copyright framework, with such a change likely to affect many areas, practitioners have told WIPR.

worldipreview.com

Theoretical contributions

- Inconsistencies in predefined opinion lexicons → domain dependent
- Inability to capture opinion properties with word embedding

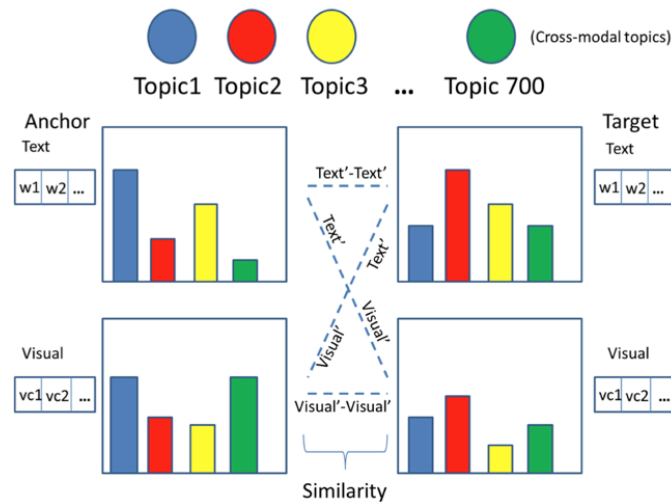
Experimental results

- Reference system for French, robust to degraded language
- Machine learning based approaches yielding state-of-the-art results for French
- Deep learning approach proposed, soon on SaaS platform allgo.inria.fr
- 3 participations to text-mining challenge
 - *top 5 for every tasks (>15 participants)*



Multimodal hyperlinking

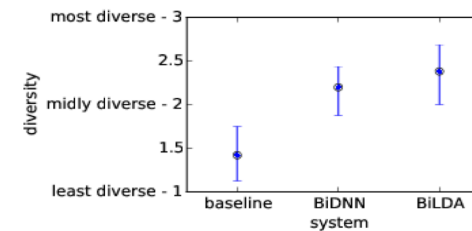
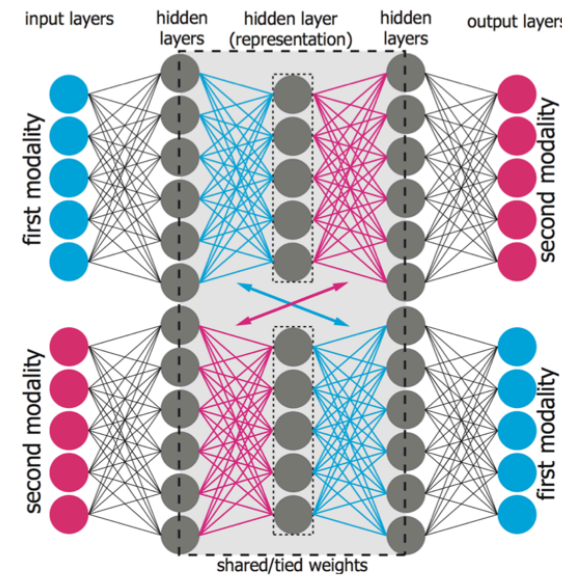
Bimodal extension of LDA mapping words and visual concepts via topic-specific distributions



	transcripts			concepts		
	n_u	\bar{d}_a	\bar{d}_i	n_u	\bar{d}_a	\bar{d}_i
baseline	29.8	0.51	0.61	35.6	0.61	0.71
BiDNN	40.8	0.20	0.12	46.7	0.42	0.31
BiLDA	40.0	0.25	0.16	38.0	0.48	0.41

Intrinsic measures of diversity

Crossmodal matching with symmetrical bi-directional auto-encoders



Perceived diversity (25 subjects)

[Bois et al. Exploiting multimodality in video hyperlinking to improve target diversity. MMM 17]



Adaptive nearest neighbors



We want **nearest neighbors graphs that are *navigable***

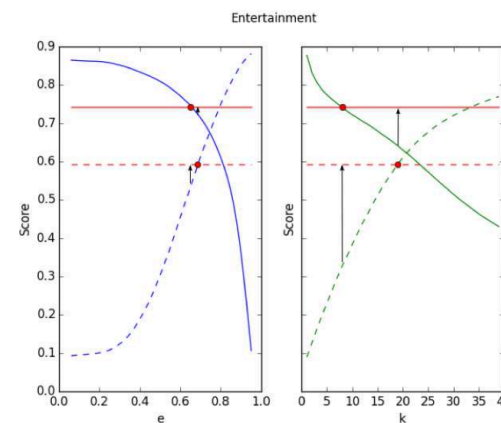
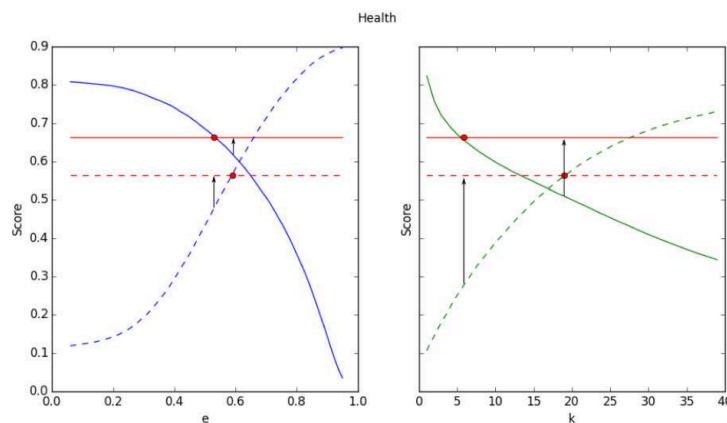
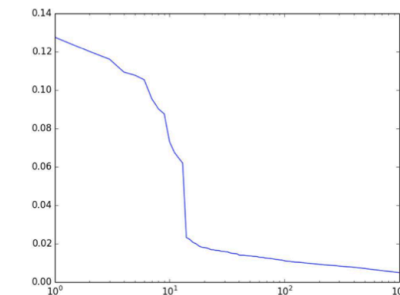
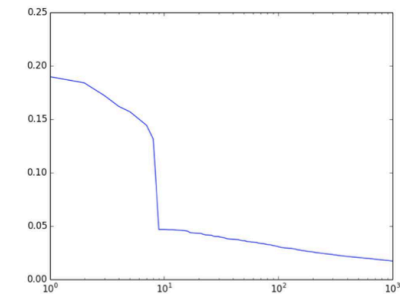
→ links are meaningful, diverse but not too many, with one component where the path between related items is short;

k and ϵ NN graphs are not suited

Idea: exploit local properties of the neighborhood

→ link each item to a varying number of neighbors determined from the local density

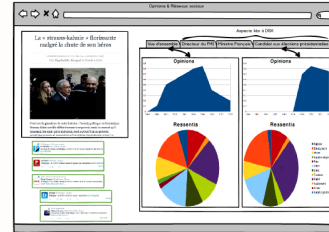
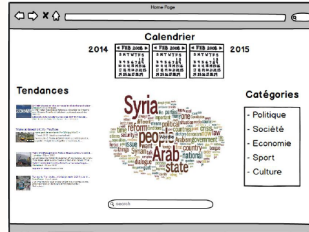
Better navigability properties and improved relevance of links on News Aggregator Dataset



The News Navigation interface



- **Upstream study** using mockups with journalists



- **User performance and acceptability study**

- 3 weeks of paper, radio and TV news
 - ~5000 documents, transcription + topic segmentation for radio/TV
- 3 versions = 3 groups, 25 infocom students & 7 journalists (v3) in total

version 2

version 3

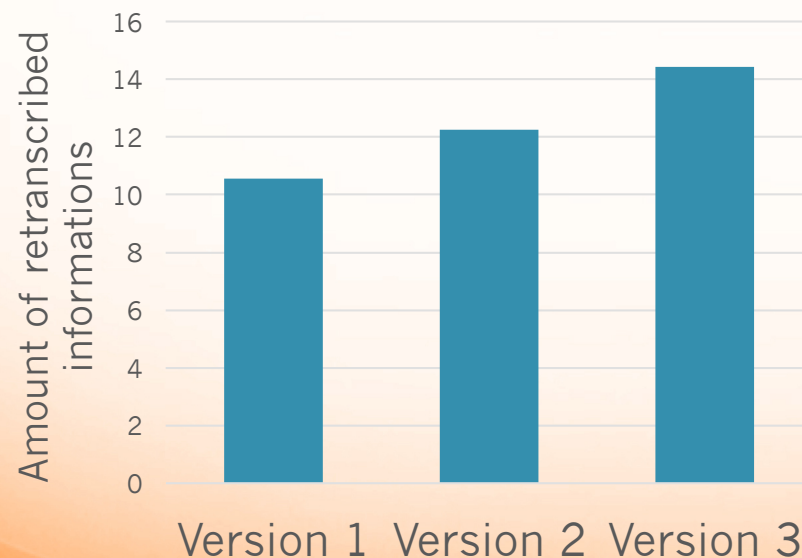


User tests of the interface

Task. Write a synthetic note on a given topic (Solar Impulse 2) in a limited amount of time (20 min.), including as many pieces of information as possible.

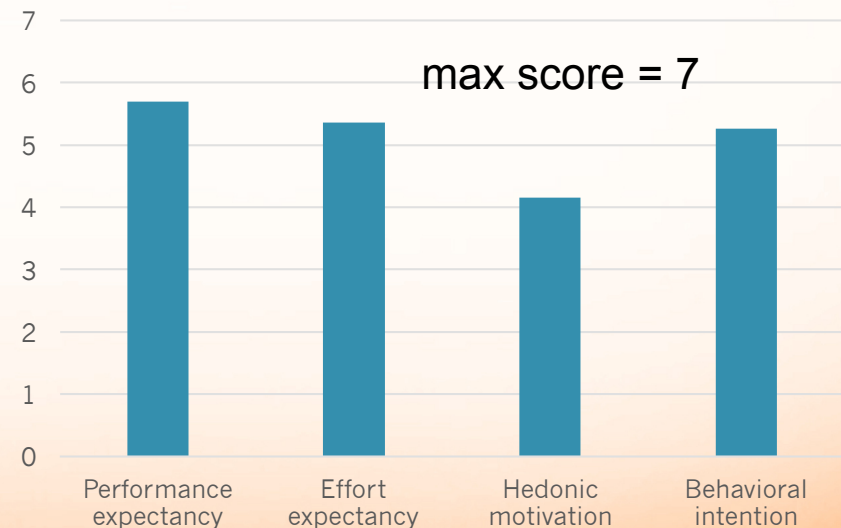
Study 1. Evaluating performance.

- 25 journalists & infocom students
- 3 distinct groups



Study 2. Evaluating acceptance.

- 6 journalists + 1 researcher
- version 3 only



Many thanks to



and



Study of practices

How people do actually use the interface for research purposes, what for and how does it change their way of problematizing their research?



One screen allow to watch how the user is doing his research-production in real time.
A camera to record how the dynamic between his draft activities and the research-production he is working on.

- Longitudinal study
- No duration limit
- Production consistent with professional activity

A 3 steps study for each user:

- Step 1 & 2: test on either LIMAH or their usual tools
- Step 3: mixing tools



Study of practices

Study performed over 9 professionals: journalists, community managers, librarians, PhD students...

- Hypergraph tool more consistent with an exploration activity than with a search one
- Less time (~50%) to accomplish research-production with LIMAH than with their usual tools (small size of the test database makes it difficult to separate the influence of the size from that of the hypergraph representation)
- Interest in the hypergraph organization of a collection for daily practice, provided the possibility to ingest their own database
- Work still in progress on the dynamic between draft & research-productions (sensemaking as an iterative process)



Legal status of hyperlinks

Freedom to link (to decide if one can link or not; power of self-determination) vs.

1. The appropriation of content protected by intellectual property
2. The liability of the link provider for the publication of information

The EU approach is characterized by:

- The freedom to link regarding EU IP law
- A differentiated framework regarding legal liability for the diffusion of information through the link

The thesis attempts to identify the **overall coherence of the current and future EU laws** as applied on hyperlinking



Summary and new opportunities



- Objectives globally achieved
- Fully interdisciplinary work
- Operational proof of concept, with extensive user tests
- International visibility
 - 13 international publications (~40% multipartners)
 - TRECVID international video hyperlinking benchmark
 - Opinion mining software (to be) released
- close collaboration with the media world
- Opportunities gained with LIMAH
 - follow-up Inria Project Lab *Knowledge-mediated integrated content and data analytics* (Linkmedia, 2017-2021)
 - adaptation of the news navigation interface for the DGA (Linkmedia)
 - project proposal on editorial strategies based on data-rich augmented forms (IMTA)

Thank you!

