



Retour d'expérience sur des collaborations pluridisciplinaires entre Informatique et SHS

Julien Velcin
Université de Lyon, Lyon 2, ERIC UR 3083
julien.velcin@univ-lyon2.fr

Journée Humanités Numériques & IA
BnF, 3 mai 2024



Contexte



- L'Université Lumière Lyon 2 est l'un des principaux centres en Sciences Humaines et Sociales (SHS)
- De nombreux structures :
 - *Maison des Sciences de l'Homme Lyon St-Etienne*
 - *Institut Rhône-Alpin des Systèmes Complexes*
 - *Pôle de spécialité sur les Humanités Numériques*
- Le laboratoire ERIC a une longue tradition de collaborations pluridisciplinaires, notamment avec les SHS

Plan

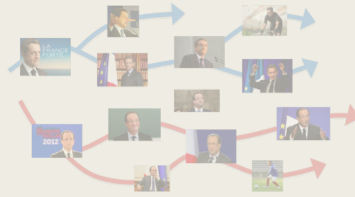
- Contexte de travail
- Un voyage en SHS
- Des points communs
- La recette du succès ?

Voyage en SHS



- **Infocom**
 - *analyses des controverses*
 - *Journalisme A l'ère Du Numérique (JADN)*
- **Linguistique**
 - *analyse du vocabulaire médical*
 - *dictionnaires de Trévoux*
- **Science politique**
 - *ImagiWeb*
 - *collaboration avec EDF*
- **Littérature**
 - *LIFRANUM*

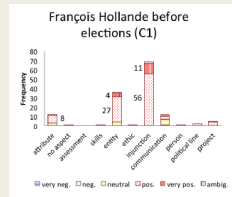
Projet ImagiWeb



- Financement ANR (2012-2015)
- Extraire l'image d'entités, par exemple :



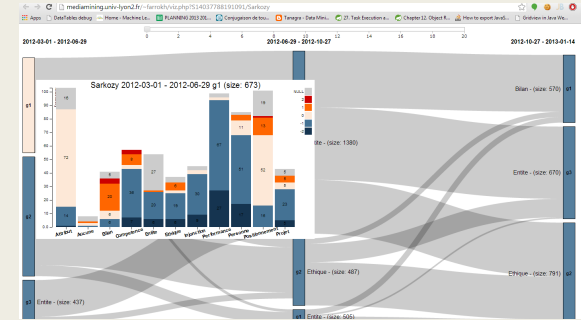
- Cibles :**
- Bilan
 - Projet
 - Ethique
 - Injonction
 - Communication
 - (...)



- Le consortium : ERIC, LIA, CEPEL, AMI Software, XRCE, EDF

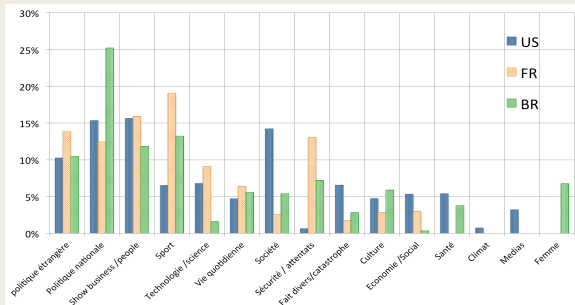
Projet ImagiWeb (suite)

- Evolution des images dans le temps (Kim et al., 2015)

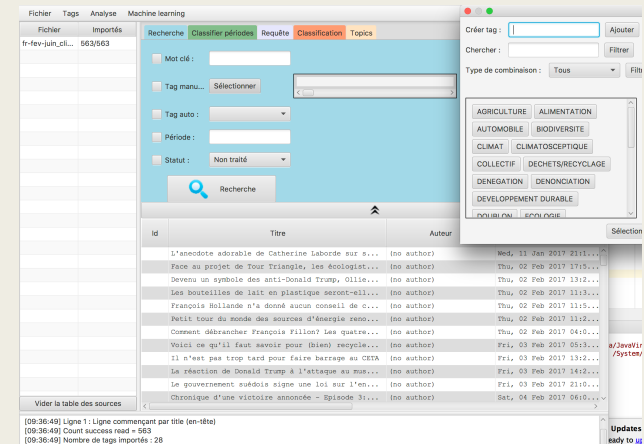


JADN (2016-2017)

- Etudier le *mediascape* (Appadurai, 1990), comparer l'information diffusée par plusieurs médias suivant le contexte (Velcin et al., 2017) :

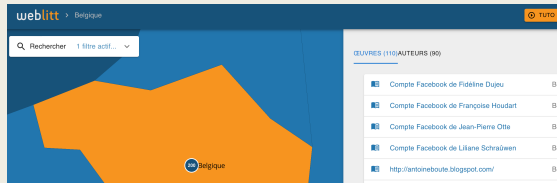


JADN (suite) : Newsbrowser



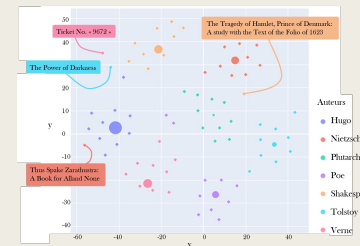
(pas de publication)

- Financement ANR (2020-2024)
- Identifier, rassembler et analyser la littérature nativement numérique

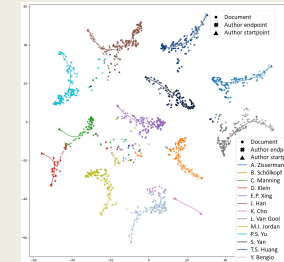


- Le consortium : MARGE, ERIC, BnF

- Apprentissage de représentations d'auteurs



représentation statique avec VADES (non publié)



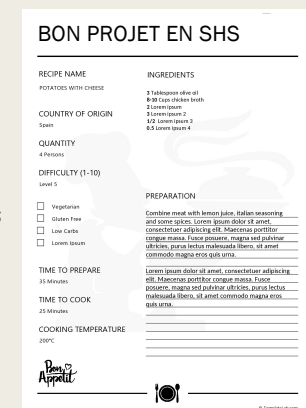
représentation temporelle avec BARL (Terreau et al., 2024)

Des points communs

- Une **ambition** importante pour des moyens qui ne sont pas à la hauteur (en termes de financement ou de temps)
- Beaucoup de **bonne volonté** pour trouver des solutions, parfois efficaces, souvent insuffisantes
- Un **alignement** parfois problématique entre les objectifs et les attentes des uns et des autres

La recette du succès ?

- Les bons ingrédients
 - une *vision lucide des objectifs et attentes de chacun des partenaires*
 - un *bon dimensionnement des besoins (RH, matériel...)*
 - un *respect des codes en usage dans les différentes disciplines*
- Ce qu'il faut éviter
 - des *projets montés trop vite*
 - absence de *travaux pilotes*



Publications

ImagWeb

- Y.M. Kim + 3 auteurs. Temporal multinomial mixture for instance-oriented evolutionary clustering, ECIR 2015.
- J. Velcin + 11 auteurs (**pluri**). Investigating the Image of Entities in Social Media: Dataset Design and First Results, LREC 2014.
- A. Hasnat + 3 auteurs. Simultaneous Clustering and Model Selection for Multinomial Distribution: A Comparative Study. IDA 2015: 120-131
- A. Hasnat + 3 auteurs. Evolutionary clustering for categorical data using parametric links among multinomial mixture models. Econometrics and Statistics, vol.3: pp.141-159, Elsevier, 2017.
- J. Boyadjian et J. Velcin (**pluri**). De l' « opinion mining » à la sociologie des opinions en ligne. Pour une approche interdisciplinaire de l'étude du web politique. Question de communication, 2017.
- J. Boyadjian + 2 auteurs (**pluri**). Le Web politique au prisme de la science des données, numéro spécial de la revue Réseaux 204, 2017, numéro coordonné par J. Boyadjian, A. Olivesi et J. Velcin. Ed. La Découverte.

JADN

- J. Velcin + 5 auteurs (**pluri**). Fouille de textes pour une analyse comparée de l'information diffusée par les médias en ligne : une étude sur trois éditions du Huffington Post. Atelier Journalisme computationnel @EGC, 2017.

LIFRANUM

- E. Terreau, Antoine Gourru, Julien Velcin: Writing Style Author Embedding Evaluation. Eval4NLP 2021: 84-93
- E. Terreau, Julien Velcin: Building Brownian Bridges to Learn Dynamic Author Representations from Texts. IDA (1) 2024: 230-241

13

Conclusions personnelles

- Collaborations toujours **très stimulantes**
- Il faut du **temps** pour travailler ensemble
- Il manque souvent un **chaînon** pour faire le **lien** entre les disciplines (peut-être un problème lié à la formation ?)
- Il est toujours difficile de **valoriser** le travail réalisé, en particulier le travail collectif

14