

Atelier du projet ANR “AlgoPol”

Centre Marc Bloch, Georg-Simmel-Saal, Berlin, 24-25.09.13



Digital
Humanities

Le projet Algotopol (<http://algotopol.fr>) vise à comprendre la manière dont l'information est filtrée et référencée sur Internet et ses diverses communautés — blogs, forums et médias en ligne de l'espace public numérique français et allemand, ou encore plateformes grand public comme facebook, twitter ou wikipedia — ainsi que de déconstruire (pour reconstruire) certains des algorithmes sur lesquels repose ce filtrage.

Mardi 24.09, 10h-18h30

horaire	terrain	contenu
10h	facebook	Irène Bastard (Orange Labs/EHESS) et Stéphane Raux (Linkfluence/U. Paris-VII) feront un point complet sur le projet de collecte de données de réseaux égo-centrés sur facebook à travers l'application développée à cet effet, puis lanceront la discussion sur leur utilisation: dialogue avec la CNIL, liens avec le “CSA” et son panel, réutilisation des données par les autres participants du projet, notamment afin d'étudier le relais d'information au sein de réseaux sociaux locaux (le “clair-obscur” du Net).
11h30, pause		
11h45	ICWSM'13	Compte-rendu d'ICWSM'13 (<i>Intl. Conf. on Weblogs & Social Media</i>), Boston, MA (Jul 8-11, 2013) (Joyce)
12h	wikipedia	Chih-Chun Chen (U. Cambridge) présentera les résultats et travaux en cours sur le comportement d'édition et de citation sur Wikipedia, en vue de comprendre, plus largement, comment les contributeurs ajoutent contenu et références de manière collaborative. La discussion se poursuivra plus globalement sur la possibilité d'esquisser un comportement général de citation sur internet.
13h, déjeuner		
14h30	twitter	Joyce Bessis (CMB/IEP Paris) rappellera l'état de l'art sur l'étude des motifs temporels dans les grands corpus, puis présentera des résultats empiriques sur la dynamique de propagation des <i>hashtags</i> sur twitter, notamment via les retweets. Une discussion suivra sur les algorithmes de recommandation temporels et les questions liées à la temporalité de la propagation du contenu.
15h30	Blogs/médias	Dominique Cardon (Orange Labs/EHESS) et Camille Roth (CNRS/CMB) présenteront les premiers résultats sur la structure des communautés de l'espace public politique numérique français (voire allemand), notamment sur la position des acteurs en termes de réputation, d'autorité, et d'activité au sein de leur territoire thématique, et en dehors.
16h20, pause		
16h45	*	Jean-Philippe Cointet (INRA/IFRIS) présentera un outil d'extraction d'entités nommées (e.g. noms de personnes et de lieux) et illustrera ses possibilités sur divers corpus textuels du projet. Ce pourra être l'occasion, dans un second temps, de faire le point sur le traitement des données au sein du projet: différents scripts développés, obstacles rencontrés (e.g. normalisation des sources), nouvelles données qui vont prochainement être collectées sur de nouveaux terrains (dont les forums de discussion) ou bien en élargissant les terrains existants (dont la blogosphère).
17h30	Blogs/médias et al.	Sébastien Lerique (CMB/EHESS) nous présentera son projet de thèse qui vise à explorer empiriquement l'épidémiologie/interprétation/reformulation des représentations, notamment dans le cas particulier des citations, à la fois sur des corpus du web (<i>in vivo</i>) et expérimentaux (<i>in vitro</i>).
18h	github	Antoine Mazières (LIAFA) présentera un projet sur les logiques de pouvoirs au sein de la plate-forme de développement collaborative github.

Mercredi 25.09, 10h-12h30

10h	*	Telmo Menezes (CNRS/CMB) re-présentera et illustrera empiriquement le projet “GraphBrain” de collecte d'informations décentralisée à partir de grands corpus textuels (e.g. internet, bases bibliographiques) de manière à appréhender son potentiel en termes de mesures d'autorité et de certification de savoirs à partir d'une multitude de sources (internet étant à nouveau un prototype idéal).
10h45	*	Christophe Prieur (LIAFA/U. Paris 7) fera le point sur les questions liées aux algorithmes de recommandation locale, en vue, notamment, d'améliorer leur conception en s'inspirant des résultats expérimentaux sur le partage d'information et sur la structure des réseaux égo-centrés.
11h45	Blogs/médias et al.	Camille Roth (CNRS/CMB) introduira des questions liées aux “horizons épistémiques”: comment quantifier la diversité des sources, la balkanisation des communautés, la concentration ou le polycentrisme des sources/références. La question se pose notamment en vue de qualifier les communautés politiques “radicalisées” vs. “centristes”, ou bien les acteurs du débat sur le réchauffement climatique, en France et en Allemagne.