

PROGRAMME
JOURNÉE DU LUNDI 26 JUIN 2017
ESPACE RECHERCHE SCD UNIVERSITÉ LILLE 3

10H45 : ACCUEIL

11H15 : OUVERTURE DES JOURNÉES

11H30 : FOUILLE D'INFORMATIONS SUR LE MÉDICAMENT
CONFÉRENCE INVITÉE NATALIA GRABAR (CNRS, LABORATOIRE STL,
UNIVERSITÉ LILLE)

12H30 : REPAS

14H00 : FOUILLE DE TEXTES POUR LES SHS

- Exploration Contextuelle (EC3) : une stratégie de détection, d'extraction et de visualisation des données cibles
Christophe Jouis, Hammou Fadili, Jean-Gabriel Ganascia, Bin Yang, Marc Bertin et Iana Atanassova – LIP6 / CNAM / GFZ German Research / URFIST / CRIT
- La fouille de texte pour la valorisation du patrimoine industriel textile dans les Hauts-de-France,
Rémi Cardon et Eric kergosien, GERiiCO, Lille

15H00 : HISTOIRE DE L'ART

- Retracer le processus de création d'un spectacle à partir de milliers de documents
Clarisse Bardiot – laboratoire DeVisu, Valenciennes
- Bio-bibliographies de critiques d'art francophone, une analyse quantitative à visée prosopographique
Gérald Kembellec et Marie Gispert, CNAM / Université Sorbonne Paris 1, Paris

16H00 : PAUSE CAFÉ

16H15 : VISUALISATION DES DONNÉES VOLUMINEUSES

- Visualisation des données de prêts en bibliothèque : La plateforme Prévu,
Camille Monnier, Gaétan Darquié et Mehdi Bourgeois – Université Paris 8
- Interactive Visualization of User Groups
Sihem Amer-Yahia, João Comba, Viviane Moreira, Behrooz Omidvar-Tehrani, Fabian et Colque Zegarra – CNRS, Université Grenoble Alpes / The Federal University of Rio Grande do Sul (Brazil)
- Datascape Privacy : méthode d'exploration visuelle d'un large corpus de données web
Maxime Crepel et Mathieu Jacomy, Sciences Po - médialab

17H45 : FIN DE JOURNÉE

19H30 : REPAS (LIEU ENCORE À DÉFINIR)

PROGRAMME
JOURNÉE DU MARDI 27 JUIN 2017
ESPACE RECHERCHE SCD UNIVERSITÉ LILLE 3

9H00 : ANALYSE DE COMMUNAUTÉS PAR LES GRAPHS

- Application de la maximisation des traits aux données volumineuses : graphes de contraste et identification des rôles communautaires dans les graphes
Jean-Charles Lamirel – Laboratoire LORIA, Nancy
- La détection de réseaux sociaux à partir de leurs traces : l'exemple d'un site de pétitions en ligne
Thomas Soubiran – CERAPS, CNRS Lille 2
- Expression politique sur Twitter : les communautés comme indicateurs de point de vue
Ophélie Fraisier, Guillaume Cabanac, Yoann Pitarch, Romaric Besançon et Mohand Boughanem – IRIT / CEA, Toulouse

10H30 : PAUSE CAFÉ

10H45 : ANALYSE DE DONNÉES MULTIMEDIA

- Système multi-agents pour l'analyse statistique et économétrique de données massives
Quentin Baert, Anne-Cecile Caron, Virginie Delsart, Maxime Morge and Jean-Christophe Routier – CRISTAL / LIFL /CNRS, Lille, France
- Traitement de l'Information Narrative dans le Contexte de l'Héritage Culturel
Pierrick Bruneau, Luxembourg Institute of Science and Technology, Luxembourg
- Conditions et effets du scraping massif en sciences sociales : le cas d'une plateforme de live webcams
Jérémy Garrigues – STSLab / Géopolis, Suisse

12H15 : REPAS

13H45 : GENRE, DONNÉES SEXUÉES : ETUDES PLURIDISCIPLINAIRES

- Du bon usage de l'interdisciplinarité pour l'analyse de données sexuées : le cas des données bibliographiques de la conférence EGC
Cécile Favre, Guillaume Cabanac, Gilles Hubert et Cyril Labbé – ERIC / IRIT / LIG, France
- Les femmes sont-elles des geeks comme les autres ? Genre et représentations médiatiques des Technologies de l'Information et de la Communication
Marion Dalibert et Simona De Lulio, GERiiCO, Université de Lille

15H : CLÔTURE DES JOURNÉES

INSCRIPTION GRATUITE MAIS OBLIGATOIRE VIA LE LIEN SUIVANT :

<https://goo.gl/forms/aFmDTOWCXoXWLbbz1>

CONFÉRENCE INVITÉE :

NATALIA GRABAR, CNRS, LABORATOIRE STL, UNIVERSITÉ LILLE 3

Les informations sur les médicaments sont cruciales dans le processus de santé. De plus, elles sont présentes dans différentes sources : dossier médical du patient, littérature scientifique, réseaux sociaux. Les informations qui y sont disponibles varient et peuvent être complémentaires dans certains contextes. Dans notre présentation, nous allons nous concentrer sur la fouille et l'extraction d'information sur les prescriptions, sur les effets indésirables et les interactions de médicaments, et sur l'usage de médicaments par les patients. Le travail effectué dans le cadre de différents projets (ANR MIAM, IRESP PACHA, FP7 PROTECT, compétitions I2B2...) sera présenté.

THÉMATIQUES DES JOURNÉES

Le laboratoire GERiCO et les groupes de travail de l'association EGC – « Fouille de Données Complexes » (GT-FDC), « Visualisation d'informations, interaction et fouille de données » (GT-VIF, commun avec l'AFIHM) et « Gestion et Analyse de données Spatiales et Temporelles » (GT-GAST, commun avec l'action prospective EXCES, GDR MAGIS) - organisent la 6ème édition des journées thématiques visant d'une part à poursuivre les activités des groupes et d'autre part à développer des axes communs autour de la prise en compte, la gestion, l'analyse et la visualisation des données massives (big data). Ces journées sont également soutenues par le GDR MADICS (<http://www.madics.fr>).

Les problématiques abordées lors de ces deux journées peuvent concerner les processus (acquisition, structuration, extraction d'information et de connaissances et la visualisation) ou les données elles-mêmes. L'objectif de ces journées est, d'une part, de rassembler l'ensemble des acteurs de la communauté scientifique intéressés par ces nouvelles approches de gestion de données massives ainsi que par les concepts manipulés autour de ces données. D'autre part, l'objectif est d'animer et de dynamiser une communauté scientifique pluridisciplinaire. Pour ce faire, un appel à communications est donc lancé, et des personnes spécialistes de ces questions seront également invitées à présenter leurs travaux.

Pour cette nouvelle édition, un focus est fait sur des approches pluridisciplinaires (Informatique, Sciences de l'Information, Mathématiques, Sciences du langage, Géographie, Biologie, Droit, éthique, Histoire...) du phénomène des data dans leurs différentes acceptions (open data, big data, smart data...) et sur différents terrains (analyse prédictive, veille, politiques publiques, humanités numériques, e-science, données de la recherche et données scientifiques...). Les Arts, Lettres, et les Sciences Humaines et Sociales dans leur globalité y trouvent un double intérêt. Le premier est d'ouvrir de nouveaux espaces de recherche et de nouvelles thématiques impactant considérablement la manière d'effectuer ces recherches. Le second est de fournir à ces disciplines les moyens de participer aux innovations technologiques et scientifiques, de

produire des biens et des services à forte valeur socio-économique et, par-delà à accroître leur apport dans la chaîne de création de valeur socio-économique et culturelle.

Ainsi, pour faire suite à la 4^{ième} édition organisée en juin 2015 sur le thème des Humanités Numériques à Lyon, ces journées permettent de proposer un espace de rencontres et d'échanges, tout d'abord sur les concepts manipulés autour des données, et ensuite sur les problématiques d'analyse et de visualisation de données massives, en particulier dans le contexte des Arts, des Lettres, des Sciences Humaines et les Sciences Sociales.

Comité d'organisation

Eric Kergosien (GERiiCO, Université Lille)
Bernard Jacquemin (GERiiCO, Université Lille)
Amel Fraisse (GERiiCO, Université Lille)
Stéphane Chaudiron (GERiiCO, Université Lille)

Comité de coordination

Correspondant EGC avec les groupes de travail :

Mustapha Lebbah (LIPN, Université Paris 13)

GT-GAST :

Thomas Guyet (IRISA, Agrocampus Ouest)
Eric Kergosien (GERiiCO, Université Lille 3)
Cyril de Runz (CReSTIC, Université de Reims Champagne-Ardenne)

GT-FDC :

Cécile Favre (ERIC, Université Lyon 2)
Germain Forestier (MIPS, Université de Haute Alsace)
Camille Kurtz (LIPADE, Université Paris Descartes)
Alexandre Blansché (LITA – Université de Lorraine)

GT-VIF :

Hanene Azzag (LIPN, Université Paris 13)
David Bihanic (CALHISTE, Université de Valenciennes et du Hainaut-Cambrésis)
Monique Noirhomme (FUNDP, Namur, Belgique)
Fabien Picarougne (LINA, Université de Nantes)
Pierrick Bruneau (LIST, Luxembourg)