JOURNÉES D'ÉTUDES INED-SAGE

CE QUE L'ABONDANCE ET LA DIVERSITÉ DES DONNÉES NUMÉRIQUES FONT AUX PRATIQUES DE RECHERCHE EN SCIENCES SOCIALES

Journées d'études organisées par Arnaud BRINGÉ, Anne LAMBERT et Étienne OLLION, avec le concours de l'Ined et du SAGE-Université de Strasbourg.

VENDREDI 17 MARS

10H00-12H30 SESSION « DROIT ET ÉTHIQUE »

Président : **Alain TROGNON** (Insee)
Discutantes : **Béatrice CHERRIER** (Université de Caen/CREM)

PRÉSENTATIONS COURTES

- **Pierre-Antoine BILBAUT** (Ined/CERLIS), *Aspirer des données et après ? Questions d'éthique et de méthode*
- **Jérémie GARRIGUES** (STSLab/Université de Lausanne), *Accès aux données numériques et formalisation algorithmique : le cas du crawl d'une plateforme de livewebcam*
- **Bilel BENBOUZID** (LISIS/Université Paris-Est), **Nicolas BAYA-LAFFITE** (STSLab/Université de Lausanne), **Jean-Philippe COINTET** (Inra/LISIS), *Le sociologue et les dataproviders : quelles liaisons dangereuses ?*

TABLE RONDE DÉBAT *Quels droits pour l'enquête en sciences sociales ?*

- **Céline CASTETS-RENARD** (IRDEIC/Université Toulouse Capitole) [sous réserve]
- **Pierre-Yves GEOFFARD** (PSE/CNRS)
- **Florence WEBER** (CMH/École normale supérieure)

DÉJEUNER LIBRE

14H00-16H00 SESSION : « SAVOIRS »

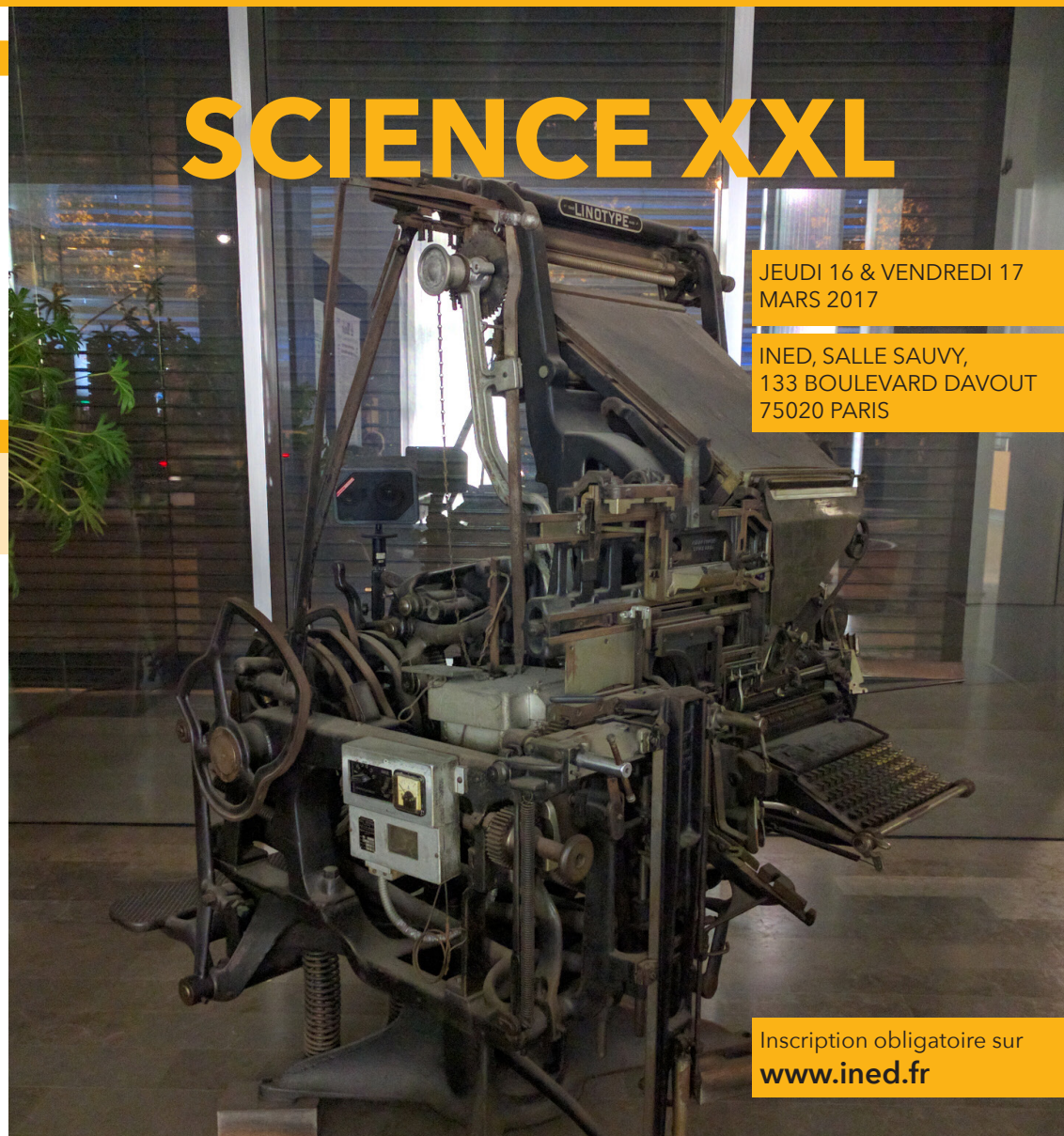
Présidente : **Anne LAMBERT** (Ined/CMH)
Discutant : **Rémy CAVENG** (CURAPP/Université de Picardie)

PRÉSENTATIONS COURTES

- **Paul HOBEIKA** (CSU/Université Paris 8), *Quelles catégories statistiques pour étudier les inégalités entre retraités ? Apports et limites des bases de données administratives*
- **Yoann DEMOLI** (Printemps/Université Versailles Saint-Quentin), *Enseigner les méthodes quantitatives grâce aux big data : retour d'expérience*
- **Olivier CABRIGNAC** (Global Life), **Arthur CHARPENTIER** (CREM/Université de Rennes 1), **Ewen GALLIC** (CREM/Université de Rennes 1), *Les données collaboratives généalogiques au service de la démographie historique*

PRÉSENTATIONS LONGUES

- **Marie BERGSTRÖM** (Ined), *Sites de rencontre : un nouveau site d'observation de la formation des couples*
- **Gilles BASTIN** (Pacte/Sciences Po Grenoble), *Le datascraper est-il un observateur participant comme un autre ? Un retour d'expérience sur le réseau LinkedIn*



JEUDI 16 & VENDREDI 17 MARS 2017

INED, SALLE SAUVY,
133 BOULEVARD DAVOUT
75020 PARIS

Inscription obligatoire sur
www.ined.fr

PROGRAMME

Données de l'Internet, données de capteurs, fichiers de gestion... Le volume d'informations numériques en apparence gratuites et immédiatement disponibles, ne cesse de croître. Comment ces données modifient-elles le travail scientifique, depuis sa conception jusqu'à sa restitution ?
Les communications qui rassemblent des chercheur-e-s de différentes disciplines (démographie, sociologie, économie, histoire, science politique, etc.), questionnent les reconfigurations de la recherche suite à cette massification et à cette diversification des sources, qu'il s'agisse de la transformation des méthodes ou des objets, des savoirs produits, des relations avec les enquêté-e-s et la place des chercheur-e-s, ou encore des enjeux juridiques et éthiques.
Les communications sont organisées en quatre sessions thématiques (données et sources ; outils d'analyse ; savoirs disciplinaires ; droit et éthique) et deux tables rondes transversales.

JEUDI 16 MARS

9H00-9H30 ACCUEIL CAFÉ

9H30 INTRODUCTION

Mot d'accueil : **Magda TOMASINI** (Directrice de l'Ined)
Introduction générale : **Anne LAMBERT** (Ined/CMH) et **Étienne OLLION** (SAGE/Université de Strasbourg)

10H00-12H00 SESSION 1 « SOURCES ET ENQUÊTES »

Président : **Étienne OLLION** (SAGE)
Discutant : **Baptiste COULMONT** (CSU/Université Paris 8)

PRÉSENTATIONS COURTES

- **Antoine BEVORT** (Lise/Cnam/CNRS), *La Webanalytics des sites socio-politiques : méthodologie et études empiriques*
- **Antoine COURMONT** (CEE/Sciences Po), *Ouvrir les données publiques : les données à l'épreuve de leur diffusibilité*
- **Laurent BEAUGUITTE** (IDEES/CNRS), **Hughes PECOUT** (CIST), **Marta SEVERO** (Dicen-IDF/Université Paris Nanterre), *Les flux RSS pour étudier l'actualité internationale*
- **Andrey GRUNIN** (CIHAM/Université d'Avignon), *Pour des données historiques massives : ambitions, enjeux, questionnements*
- **Léonard MOULIN** (Ined), *Étudier les trajectoires étudiantes à l'aide de données administratives*
- **Muriel EPSTEIN** et **Nicolas BOURGEOIS** (SAMM/Université Paris 1), *Des blogs enseignants comme « meilleurs entretiens non directifs » ? L'apport d'une analyse textuelle quantitative*
- **Alexandre HOBEIKA (CMH/SAGE),** *Nethnographie et jeux d'échelle : comment l'outillage méthodologique se modifie-t-il avec la taille du corpus ?*

DÉJEUNER LIBRE

14H00-16H00 SESSION 2 « OUTILS »

Présidente : **Eva LELIÈVRE** (Ined)
Discutant : **Arthur CHARPENTIER** (CREM/Université Rennes 1)

PRÉSENTATIONS LONGUES

- **Étienne OLLION** (SAGE/Université de Strasbourg) et **Julien BOELAERT** (SAGE/Université de Strasbourg), *Ce que peut le machine learning*
- **Églantine SCHMITT** (Costech/Université technologique de Compiègne), *Quelle ingénierie logicielle pour le traitement de données textuelles en sciences sociales ?*

PRÉSENTATIONS COURTES

- **Nicolas ROBETTE** (ENSAE), **Audrey MARIETTE** (CSU/Université Paris 8), **Laure DE VERDALLE** (Printemps//Université Versailles Saint Quentin), **Karim HAMMOU** (CSU/Universté Paris 8), *Heaven's gate? Construction d'une base de films par agrégation de diverses sources Internet*
- **Paul GIRARD**, **Mathieu JACOMY**, **Christophe LECLERCQ** (Sciences Po/médialab), *Explorer association et transformation à travers des datascapes*
- **Léo JOUBERT** (LEST/Université d'Aix-Marseille), *Global et local dans les communautés numériques. Le cas de Wikipedia*
- **Valérie BEAUDOUIN**, **Florence d'ALCHE-BUC**, **Adrien NOUVELLET**, **Christophe PRIEUR**, **François ROUEFF** (LTCI/Telecom ParisTech), *Ce que le web fait à une bibliothèque : étudier les usages de Gallica par les traces*
- **Pierre-Olivier CHASSET**, **Hadrien COMMENGES** (Géographie-Cités/Université Paris 1), **Clémentine COTTINEAU** (CASA/University College of London), **Christine KOSMOPOULOS** (Géographie-Cités/Université Paris 1), **Denise PUMAIN** (Géographie-Cités/Université Paris 1), **Juste RAIMBAULT** (LVMT/Université Paris-Est), *CybergeoNetworks : un outil interactif pour l'analyse de corpus bibliométriques massifs*

16H00-16H30 PAUSE

16H30-18H00 TABLE RONDE

Les sciences sociales et les données numériques : tour d'horizon disciplinaire

- **Claire LEMERCIER** (CSO/CNRS)
- **Pauline GIVORD** (Insee)
- **Dominique CARDON** (médialab/Sciences Po)

18H00-19H30 COCKTAIL