

Dr. Jean-Valère Cossu

Data Scientist - Chercheur, MY LOCAL INFLUENCE, 45 Rue Frédéric Joliot Curie MYLI – 13013 Marseille (France)
28 ans, Nationalité Française +33 6 65 63 07 28; jvcossu@gmail.com www.jeanvalerecossu.fr

Expériences professionnelles

- **Depuis Janvier 2017 : Ingénieur de Recherche** My Local Influence, Marseille (France)
- **Octobre 2015 à Octobre 2016 : Ingénieur de Recherche** Vodkaster, Paris (France)
- **Septembre 2012 à Aout 2015 : Doctorant chargé d'enseignement** UAPV LIA, Avignon (France)
- **Novembre 2011 à Aout 2012 : Ingénieur de Recherche** UAPV LIA, Avignon (France)
- **Janvier 2011 : Co Administrateur de l'Alliance Francophone Sudden-Strike**

Thématiques de Recherches

- **My Local Influence** : Data Mining, analyse d'opinions et outils d'aide à la décision
 - Analyse automatique sessions
 - Résumé, classification et Catégorisation automatique de documents
 - Infrastructure data et gestion des données personnelles
- **Vodkaster** : Data Mining, recommandation et outils d'aide à la décision
 - Analyse automatique de logs (systèmes, sessions de recherches et transactionnels)
 - Recommandation de films et séries, profilage d'utilisateurs et analyse de communautés
 - Encodage et diffusion de flux multimédias, Détection de fraude
- **Thèse** : Analyse de la réputation en ligne d'entité
 - Résumé, classification et Catégorisation automatique (opinions et thématiques) de profils d'utilisateurs et de textes issus des médias sociaux en utilisant des méthodes statistiques
 - Influence des métriques d'évaluations sur les tâches, algorithmes et stratégie d'apprentissage
- **Campagnes d'Évaluation Scientifique**
 - Analyse de la réputation en ligne d'entités (Replab) en 2013 et 2014
 - Détection de l'influence et de la profession d'un utilisateur de Twitter (RepLab/PAN 2014)
 - Résumé automatique de texte (INEX 2014)
 - Catégorisation de recettes de cuisines et recherche d'ingrédients (DEFT 2013)
- **Alternance M2**
 - Système de catégorisation et recommandation de films basé sur l'analyse du langage naturel dans des critiques courtes d'un réseau social spécialisé : Vodkaster
- **Relecteur pour** TALN 2014, CORIA 2015, CICLING 2015, ACM TOIT, IP&M, ICEIS 2018
- **Lab MC²** « *Multilingual Cultural Mining and Retrieval* » <https://mc2.talne.eu/>

Formations

- **2012-2015, Doctorat en informatique (mention très honorable), Université d'Avignon**
Analyse de la dynamique de l'image de marques sur le Web 2.0, encadrée par Pr. Marc El-Bèze, Dr. Juan-Manuel Torres Moreno et Dr. Eric SanJuan
Soutenu publiquement le 16 décembre 2015 devant un jury composé de : Josianne Mothe (Présidente/Rapporteur), Philippe Mulhem (Rapporteur), Rocio Abascal Mena, Julien Velcin, Paolo Rosso, Frédéric Gimello-Mesplomb, Jian-Yun Nie (Examineurs) et Marc El-Bèze (Directeur)
- **2009-2012, Licence et Master en informatique** Université d'Avignon (France)
- **2007-2009, BTS Informatique de Gestion (Opt. réseaux), Ajaccio (France)**

Compétences Techniques

- **Compétences scientifiques** : Fouille des données ; Résumé, classification, catégorisation et apprentissage automatique ; Recherche d'Information, Optimisation et Aide à la décision, Modélisation et Visualisation d'information
- **Rédactionnel et administratif**: Dossiers JEI/CIR/CII, Documentation fonctionnelle et scientifique, BPM, gestion des données personnelles, benchmark de solutions SaaS, Reporting
- **Technologies, systèmes, réseaux & télécoms** : MySQL/MariaDB, Linux, Perl, C/C++, Java, PHP, Analytics Web, Dashboards, Traitement du signal, CDN, P2P, HAProxy, L^AT_EX

Enseignements

- **2015 (21h) – Base de données et réseaux sociaux, Master 2 Informatique, CERI**
Analyse de réseaux sociaux avec approche réseaux complexes et théorie des graphes analyse d'applications P2P, traitement automatique des langues et intégration en base données.
- **2014 (21h) – Infr. Réseaux et médias sociaux, Master 2 Informatique, CERI**
Analyse de réseaux sociaux avec une approche traitement de la langue, analyse d'applications P2P, utilisation d'une API et crawling.
- **2013-2015 (161h) – Programmation C++, Licence 1 & 2 Informatique, CERI**
Compilation, classes, fonctions, listes, pointeurs, récursivité et backtracking.
- **2013 (21h) – Programmation réseau, Master 1 Informatique, CERI**
Programmation réseau (sockets) en Java, C/C++, architectures clients-serveurs et distribuées.

Encadrements

- **2015- Co-encadrement groupe de projet Master 2**
Virtualisation, hyperviseur et container, quel choix pour du calcul scientifique ?
- **2014 – Co-encadrement stage de Master 2 – Mohamed El-Hamdaoui**
Modélisation de paris sportifs par PLS-PM et analyse d'opinions (publication ECML-PKDD)
- **2013-2014 – Encadrement et coordination de la participation LIA à RepLab**
- **2013 – Co-encadrement groupe de projet de Master 1**
L'apprentissage par réseaux de neurones profonds appliqué à l'analyse de réputation

Langues

Français (natif)

Anglais (BULATS 74/100)

Espagnol (expression élémentaire, compréhension intermédiaire)

Italien (intermédiaire)

Loisirs et autres

Photographie

Modélisme

Randonné

Plongée sous-marine

Culture coréenne

Jeux vidéo de simulation

Cuisine

Histoire & Géopolitique (Antiquité, 1^{er} empire, Guerres Mondiales)

Liste des Publications

• Journaux internationaux – revues avec comité de lecture

1. Jean-Valère Cossu, Alejandro Molina Villegas et Mariana Tello-Signoret – "*Active learning in annotating micro-blogs dealing with e-reputation*" – Journal of Interdisciplinary Methodologies and Issues in Science - Vol. 3 Contextualisation Numérique (2017)
<https://jimis.episciences.org/3970>
2. Jean-Valère Cossu, Nicolas Dugue et Vincent Labatut – "*Review of Features for the Discrimination of Twitter Users: Application to the Prediction of Offline Influence*" – SNAM Special Issue on Diffusion of Information and Influence in Social Networks (2016).
<https://www.springerprofessional.de/en/social-network-analysis-and-mining-1-2016/7032556>
3. Jean-Valère Cossu, Juan-Manuel Torres-Moreno, Eric San-Juan et Marc El-Bèze – "*Intweeitive Text Summarization*" – International Journal of Computational Linguistics and Applications - Vol. 7 No. 1 (January-June 2016)
<https://www.gelbukh.com/ijcla/2016-1/>

• Conférences et ateliers internationaux avec comité de lecture

1. Jean-Valère Cossu et Liana Ermakova – "*Lexical Context for Profiling Reputation of Corporate Entities*" - 19th International Conference on Enterprise Information Systems (ICEIS), Porto (Portugal) (2017)
http://www.iceis.org/Abstracts/2017/ICEIS_2017_Abstracts.htm
2. Jean-Valère Cossu, Eric San-Juan, Juan-Manuel Torres-Moreno et Marc El-Bèze – "*Multi-Dimensional Reputation Modeling using Micro Blog contents*" – 22nd International Symposium on Methodologies for Intelligent Systems (ISMIS), Lyon (France) (2015)
<http://liris.cnrs.fr/ismis15/index.php/program>
3. Jean-Valère Cossu, Nicolas Dugue, et Vincent Labatut – "*Detecting Real-World Influence Through Twitter*" – The Second European Network Intelligence Conference (ENIC), Karlskrona (Suède) – 2015
<http://enic.cse.bth.se/conference-program/>
4. Jean-Valère Cossu, Killian Janod, Emmanuel Ferreira, Julien Gaillard et Marc El-Bèze, "*NLP-based classifiers to generalize experts assessments in E-Reputation monitoring*" – Sixth International Conference of the CLEF initiative, Toulouse (France) (2015)
http://clef2015.clef-initiative.eu/CLEF2015/accepted_papers.php#ps1
5. Rocío Abascal-Mena, Jean-Valère Cossu, Alejandro Molina et Juan-Manuel Torres-Moreno – "*Anotación automática de datos acerca de la reputación de los políticos en redes sociales*" – Research in Computing Science (ISSN 1870-4069) Issue 97 (2015), page 81-99.
http://www.rcs.cic.ipn.mx/rcs/2015_97/
6. Jean-Valère Cossu, Rocío Abascal-mena, Alejandro Molino, Juan-Manuel Torres-Moreno et Eric San-Juan – "*Machine Learned Annotation of tweets about politicians' reputation during Presidential Elections: the cases of Mexico and France*" – International Workshop on Replicability and Reproducibility in NLP : Adaptive methods, resources and software at IJCAI 2015, Buenos Aires (Argentina) (2015)
<https://sites.google.com/site/adaptivenlp2015/home/list-of-accepted-papers>
7. Jean-Valère Cossu, Eric San-Juan, Juan-Manuel Torres-Moreno et Marc El-Bèze – "*Automatic Classification and PLS-PM Modeling for Profiling Reputation of Corporate Entities on Twitter*" – Application of Natural Language to Information Systems (NLDB), Passau (Allemagne) (2015)
<http://nldb2015.org/programme/>

8. Jean-Valère Cossu, Rocio Abascal-mena, Alejandro Molino, Juan-Manuel Torres-Moreno, et Eric San-Juan – “*Bilingual and Cross Domain Politics Analysis*” – Research in Computing Science (ISSN 1870-4069) Issue 85 (2014), page 9-19.

http://www.rcs.cic.ipn.mx/2014_85/

9. Mohamed El-Hamdaoui et Jean-Valère Cossu – “*An opinion mining Partial Least Square Path Modeling for football betting*” – European Conference on Machine Learning and Practice of Knowledge Discovery in Databases (ECML-PKDD PhD Session), Nancy (France) (2014)

<https://phdsession-ecmlpkdd2014.greyc.fr/node/7.html>

10. Jean-Valère Cossu, Benjamin Bigot, Ludovic Bonnefoy et Gregory Senay – “*Towards the improvement of topic priority assignment using various topic detection methods for e-reputation monitoring on Twitter*” – Application of Natural Language to Information Systems (NLDB), Montpellier (France) (2014)

http://www.lirmm.fr/~mroche/NLDB2014/Web/accepted_paper.html

• **Conférences francophones avec comité de lecture**

1. Leila Khouas, Caroline Brun, Anne Peradotto, Jean-Valère Cossu, Julien Boyadjian et Julien Velcin, “*Etude de l’image de marque d’entités dans le cadre d’une plateforme de veille sur le Web social*” – Conférence sur le Traitement Automatique des Langues Naturelles (TALN), Caen (France) (2015)

<https://taln2015.greyc.fr/session-demos/>

2. Jean-Valère Cossu, Juan-Manuel Torres-Moreno et Marc El-Bèze – “*Recherche et utilisation d’entités nommées conceptuelles dans une tâche de catégorisation*”, Conférence sur le Traitement Automatique des Langues Naturelles (TALN), Sables d’Olonne (France) (2013)

http://www.atala.org/taln_archives/TALN/TALN-2013

3. Jean-Valère Cossu, Julien Gaillard, Juan-Manuel Torres-Moreno et Marc El-Bèze – “*Contextualisation de messages courts : l’importance des métadonnées*” – Conférence Francophone sur l’Extraction et la Gestion des Connaissances (EGC), Toulouse (France) (2013)

<https://www.irit.fr/~Josiane.Mothe/EGCCM.htm>

• **Rapports techniques**

1. Jean-Valère Cossu, Ludovic Bonnefoy, Xavier Bost et Marc El-Bèze – “*How to merge three different methods for information filtering*” – arXiv preprint arXiv :1510.07385

<https://arxiv.org/abs/1510.07385>

2. Jean-Valère Cossu, Marc El-Bèze, Juan-Manuel Torres-Moreno et Eric San-Juan – “*E-reputation monitoring on Twitter with active learning automatic annotation*” – <hal-01002818> (2014)

<https://hal-univ-avignon.archives-ouvertes.fr/hal-01002818>

• **Participations aux challenges nationaux et internationaux**

1. Jean-Valère Cossu, Killian Janod, Emmanuel Ferreira, Julien Gaillard et Marc El-Bèze – “*lia@RepLab 2014 : 10 methods for 3 tasks*” – Replab : An evaluation campaign for Online Reputation Management Systems Fifth International Conference of the CLEF initiative, Sheffield (UK) (2014) **(Best participation award)**

<http://nlp.uned.es/replab2014/#program>

2. Jean-Valère Cossu, Benjamin Bigot, Ludovic Bonnefoy, Mohamed Morchid, Xavier Bost, Gregory Senay, Richard Dufour, Vincent Bouvier, Juan-Manuel Torres-Moreno et Marc El-Bèze – “*lia@RepLab 2013*” – Replab : An evaluation campaign for Online Reputation Management Systems Fourth International Conference of the CLEF initiative, Valencia (Espagne) (2013)

<http://ceur-ws.org/Vol-1179/>

3. Xavier Bost, Ilaria Brunetti, Luis-Adrian Cabrera-Diego, Jean-Valère Cossu, Andrea Linhares, Mohamed Morchid, Juan-manuel Torres-Moreno, Marc El-Bèze et Richard Dufour – “Systèmes du LIA à DEFT’13” – Défi de Fouilles de Textes Conférence sur le Traitement Automatique des Langues Naturelles (TALN), Sables d’Olonne (France) – 2013

<https://deft.limsi.fr/actes/2013/auteurs.html>

- **Autres**

1. Mohamed Morchid, Yonathan Portilla, Didier Josselin, Richard Dufour, Eitan Altman, Marc El-Beze, Jean-Valère Cossu, Georges Linarès, et Alexandre Reiffers-Masson – “An Author-Topic based Approach to Cluster Tweets and Mine their Location” – Procedia Environmental Sciences 27, 2015

<http://www.sciencedirect.com/science/article/pii/S1878029615003217>

2. Julien Velcin, Anne Peradotto, Leila Khouas, Jean-Valère Cossu, Jean-Yves Dormagen et Caroline Brun – “Analyser l’image de marque d’entités sur le web. Revue du projet ImagiWeb” – Revue des sciences et technologies de l’information, 1633, 1311.

<https://isi.revuesonline.com/article.jsp?articleId=19659>

- **Interventions orales**

1. Table ronde Agor@ntic-LBNC-LIA (Politique(s), transparence et éthique) : Gouvernance des corpus de données. Université d’Avignon, Novembre 2017
2. Cinéphilies/Sériephilies 2.0. Perspectives internationales “Cinéphilie 2.0 ou 1.1 ? La recommandation sociale à l’ère des micro-critiques et des taux de satisfaction d’utilisateurs. Mise en perspective du modèle Vodkaster” – Labex ICCA, Université Sorbonne Nouvelle, Juin 2017
<https://icca.univ-paris13.fr/wp-content/uploads/ProgrammeVF.pdf>
3. Ressorts et enjeux des algorithmes dans la recommandation de contenus culturels, Cours Economie Numérique et Création du M2 Médias. publics et cultures numériques – Institut français de presse Panthéon Assas, Février 2017
4. NLP-based classifiers for Reputation assessments, Lessons and Outcome for RepLab, CLEF Best of the labs Univ. Toulouse, Septembre 2015
5. Tweets Contextualization for Reputation, Séminaire de Statistical NLP and Contextualization, LIG Grenoble, 2015
6. Tweets Contextualization for Reputation Monitoring, CLEF INEX, Univ. Sheffield, Septembre 2014
7. LIA@RepLab2014, CLEF INEX, Univ. Sheffield, Septembre 2014
8. LIA@RepLab2013, CLEF INEX, Univ. Politècnica de València, Septembre 2013