

Dr. Jean-Valère Cossu

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

Professional Experience

- **Since January 2017– Data Scientist / Research Engineer**
My Local Influence, Marseille (France)
- **October 2015– Data Scientist / Research Engineer**
Vodkaster/RIPLAY, Paris (France)
- **2012 September to 2015 August – PhD candidate and teaching assistant**
Laboratoire Informatique d'Avignon, Avignon (France)
- **Novembre 2011 to August 2012 - Research Engineer / Trainee**
Laboratoire Informatique d'Avignon, Avignon (France)
- **January 2011 – Co Webmaster of l'Alliance Francophone Sudden-Strike**
Website, forum, community management
Historical Mod development for real-time strategy game

Studies

- **2012-2015, PhD in Computer Science (First Class Honours), University of Avignon**
Analyse de l'image de marque sur le Web 2.0, under the supervision of Pr Marc El-Bèze, Dr Juan-Manuel Torres Moreno and Dr Eric SanJuan.
Defend on December 16th 2015. Comity members: Josianne Mothe (President), Philippe Mulhem (Referee), Rocio Abascal Mena, Julien Velcin, Paolo Rosso, Frédéric Gimello-Mesplomb, Jian-Yun Nie (Examiners) and Marc El-Bèze (PhD Advisor)
- **2009-2012, Bachelor's and Master's degree in computer science**
Computer science and networking, (CERI) University of d'Avignon (France)
- **2007-2009, Higher National Diploma in Computer Science**
Network Management (Lycée Bonaparte) Ajaccio (France)

Teaching Activities

- **Teaching assistant at CERI, Université d'Avignon**
 - Database and social networks (Social Network Analysis / Complex Networks), M2 (21H) 2015
 - Social Networks and Infrastructure (Social Network Analysis with NLP), M2 (21H) 2014
 - C++ Base programming (Compiling, creating class and functions), L1 (65H) 2013-2015
 - C++ Avanced programming (Linked lists, makefile and backtracking), L2 (76H) 2013-2015
 - C++ Resit programming, L1 (20H) 2013
 - Networking programming (Sockets, client-server), M1 (21H) 2013

Advising Activities

- **Project and Internship advisor at CERI, Université d'Avignon**
 - 2015 Co-advisor MSc Project: Virtualization, hypervisor and container for scientific computing
 - 2014 Co-advisor MSc Intership Mohamed El-Hamdaoui (publication ECML-PKDD'2014 student session) An opinion mining Partial Least Square Path Modeling for football betting
 - 2013-2014 Advisor LIA participation @CLEF RepLab International challenges
 - 2013 Co-advisor MSc Project: Deep Learning Neural Networks for Reputation Monitoring

Research Topics

- **MYLI**
 - Data mining for business intelligence
 - Automatic document mining and categorization
 - Infrastructure and data management
- **Vodkaster**
 - Recommender system
 - Video encoding and streaming mechanism
 - Data mining for business intelligence
 - Fraud detection
- **PhD Thesis : Online Reputation Monitoring**
 - Automatic statistical social medias categorization (blogs, tweets, Twitter profiles)
 - Extractive statistical summarization
 - Reputation Modelling, Author Profiling
 - Manual, semi-automatic and automatic data annotations of social medias
 - Metrics impact on tasks, systems and learning strategies
- **Scientific Evaluation Campaigns**
 - Online Reputation Monitoring and Author Profiling (RepLab) 2013-2014
 - Automatic contextualization and summarization (INEX 2014)
 - Categorization of receipt and ingredients retrieval (DEFT 2013)
- **Internship**
 - Recommender system for movies, based on NLP comments (Vodkaster)
 - Semi-automatic summary and modeling of movies topic using opinions in comments
- **Reviewer for:**
 - Journal: ACM TOIT, ELSEVIER IP&M
 - Conferences: TALN 2014, CORIA 2015, CICILING 2015, ICEIS 2018
- **Lab MC² « Multilingual Cultural Mining and Retrieval »** <https://mc2.talne.eu/>

Skills

- **Research skills:** Automatic Categorization, Automatic summarization, Information Retrieval, Business Intelligence, Information Modeling, Machine Learning
- **Reporting and administrative procedure:** french taxes related documents JEI/CIR/CII, BPM, data management
- **Misc:** MySQL/MariaDB, Linux, Perl, C/C++, Java, PHP, Analytics Web, Dashboards, Signal Processing, CDN, P2P, HAProxy, L^AT_EX
- **Languages:** Fluent scientific / technical English, French native, Spanish and Italian (basics)

Interests

• Photography	• Modeling	• Hiking / Diving
• South Koreanculture	• Online Real Time Strategy videos games	

Publications

- **Journal**

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>

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>

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/>

- **International conferences and workshop with proceedings**

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

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://iris.cnrs.fr/ismis15/index.php/program>

Jean-Valère Cossu, Nicolas Dugue, et Vincent Labatut, " *Detecting Real-World Influence Through Twitter*" – The Second European Network Intelligence Conference, Karlskrona (Suède) – 2015

<http://enic.cse.bth.se/conference-program/>

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

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/

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>

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, Passau (Allemagne) – 2015

<http://nldb2015.org/programme/>

Jean-Valère Cossu, Rocío Abascal-mena, Alejandro Molino, Juan-Manuel Torres-Moreno, et Eric San-Juan, " *Bilingual and Cross Domain Politics Analysis*", International Symposium on Language & Knowledge Engineering, Puebla (Mexique) – 2014

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

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, Montpellier (France) – 2014

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

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 (PhD Session), Nancy (France) – 2014

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

- **Peer reviewed French conferences and workshop with proceedings**

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, Caen (France) – 2015

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

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, Toulouse (France) – 2013

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

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, Sables d'Olonne (France) – 2013

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

- **International challenges participation**

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, (UK) – 2014 (**Best participation award**)

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

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, (Espagne) – 2013

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

Xavier Bost, Ilaria Brunetti, Luis-Adrian Cabrera-Dieg, 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, Sables d'Olonne (France) – 2013

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

- **Technical reports**

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>

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>

- **Others**

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>

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>

- **Talks**

Table ronde Agor@ntic-LBNC-LIA (Politique(s), transparence et éthique) : Gouvernance des corpus de données. Université d’Avignon, Novembre 2017

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, June 2017

<https://icca.univ-paris13.fr/wp-content/uploads/ProgrammeVF.pdf>

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, February 2017

NLP-based classifiers for Reputation assessments, Lessons and Outcome for RepLab, CLEF Best of the labs Univ. Toulouse, September 2015

Tweets Contextualization for Reputation, Séminaire de Statistical NLP and Contextualization, LIG Grenoble, June 2015

Tweets Contextualization for Reputation Monitoring, CLEF INEX, Univ. Sheffield, Septembre, 2014

LIA@RepLab2014, CLEF INEX, Univ. Sheffield, Septembre 2014

LIA@RepLab2013, CLEF INEX, Univ. Politècnica de València, Septembre 2013