

## Marian-Andrei RIZOIU

**address:** CECS, The Australian National University, Canberra, ACT 2601, Australia

**email:** [marian-andrei@rizoiu.eu](mailto:marian-andrei@rizoiu.eu)

**telephone:** +61 (0)2 6125 8605

**website:** <http://www.rizoiu.eu/>



# CURRICULUM VITAE

## Employment history

- 03/2016 – **Research Fellow, College of Eng. and Comp. Science, Australian National University.**  
current  
*Research:* online popularity prediction models, stochastic point processes, machine learning and social media analysis.  
*Teaching:* Document Analysis, Advanced Databases and Data Mining (Bachelors/Masters).  
*Leadership:* course convenor for the Document Analysis course.  
*Supervision:* 2 PhD students, 5 Masters/Honours students.
- 2014 – 2016 **Research Scientist, Optimisation Research Group, NICTA Canberra, Australia.**  
*Research:* online privacy, big social data analytics, data mining.  
*Teaching:* Social Media Analysis (Bachelors/Masters) at the ANU.  
*Leadership:* assist with grant proposals writing.  
*Supervision:* 1 PhD student, 1 Masters/Honours student.
- 2013 – 2014 **Postdoctoral Fellow, ERIC laboratory, Lumière University Lyon, France.**  
*Research:* ontology construction from text, knowledge graphs evolution for brand management.  
*Teaching:* Machine Learning (Masters), Software Development Methodologies (Masters), Data Mining (Bachelors)  
*Leadership:* developed (half of) the Machine Learning course.
- 2009 – 2013 **Adj. Assistant Professor, Dept. of Computer Science, Lumière University Lyon, France.**  
*Teaching (Bachelors):* Object Oriented Programming, Unix Operating Systems, Calculus software.  
*Supervision:* 4 Honours thesis, 9 industrial internships.  
*Leadership:* PhD student representative, course convenor for Calculus software.
- summer 2013 **Research intern, Machine Learning group, NICTA Sydney, Australia.**

## Education background

- 2009 – 2013 **PhD in Computer Science**, ERIC laboratory, Lumière University Lyon, France.  
(June 24<sup>th</sup> 2013) “Semi-supervised structuring of complex data”, supervision S. Lallich and J. Velcin.
- 2008 – 2009 **Masters Degree in Data Mining and Knowledge Engineering**,  
(double diploma) Polytechnic School, Nantes University, France.
- 2004 – 2009 **Engineer Degree in Systems and Computer Engineering**  
School of Computer Science, Polytechnic University of Bucharest, Romania.

## Competences, activities, services and awards

**Languages** **English** (professional, IELTS 8.5/9), **French** (professional), **Romanian** (native).

**Technical skills** **Keywords:** big data mining and analysis, machine learning, natural language processing, online social media analysis, statistical analysis, academic software development, research student supervision, industrial project collaboration, international double-diploma teaching  
**Programming Languages:** strong skills in R, Matlab, Python, Java; good skills in C/C++, Visual Basic, PL / SQL, Xquery, Xpath, SpartQL, PHP, Bash, Perl *etc.*  
**Developing:** Lead developer of academic open-source web-based softwares: a) *HIPie*<sup>1</sup> to describe and predict the online popularity of Youtube videos (2017); b) *CommentWatcher*<sup>2</sup> to analyse online discussion forums and visualise their online social networks of users (2014).

**Awards and grants** **Awards:** ERASMUS International student exchange award (2009), Best student paper award ICTAI, Athens, Greece (2012).

**Travel Grants:** Rhône-Alpes local government (2013), ECR Travel Awards (2015).

**Material Grant:** NVIDIA GPU Grant Program (2015).

1 Presentation and live system: [http://cm.cecs.anu.edu.au/post/expecting\\_to\\_be\\_HIP/](http://cm.cecs.anu.edu.au/post/expecting_to_be_HIP/)

2 Presentation site and source code: <http://rizoiu.eu/commentwatcher>

- Reviewing** Journal of Machine Learning Research, Journal IEEE Transactions on Multimedia, Computational Intelligence, Transactions on Information Systems, Journal Transactions on Knowledge Discovery from Data; conferences ICWSM, NIPS, FLAIRS etc
- Social and organizing skills** **2010 – 2012:** PhD student representative of the ERIC laboratory.  
**October 2012:** Head of the volunteer team and member in the organisation committee of the international conferences: *Discovery Science 2012* and *Algorithm Learning Technology 2012*, Lyon.  
**August 2011:** Organisation committee member of the international conferences: *Web Intelligence 2011* and *Intelligent Agent Technology 2011*, Lyon.

## Grants and projects

- 2015 – present **US Air Force Research Office, "The Anatomy of Social Media Popularity", USD\$555K**  
 Role: **co-PI**, Develop theoretical models for predicting the popularity of online content.
- 2014 – present **Data61 CSIRO, "Human Dynamics", role: key personnel**  
 Analyse and predict social and financial networks, urban analytics from crowd-sourced images.
- 2012 – 2014 **French National Research Agency, "Images on the Web", role: key personnel**  
 Analyse the image life cycles of politicians and companies through online media; apply natural language processing for ontology construction.
- 2012 – 2014 **University Lumière Lyon (internal funding), "Evolution of discourse on nuclear medicine", role: lead-PI**  
 Develop natural language models for specialised discourse in nuclear medicine.
- 2010 – 2011 **University Lumière Lyon (internal funding), "Detecting business owners from mortuary texts", role: lead-PI**  
 Apply image and text processing for 18<sup>th</sup> and 19<sup>th</sup> century business owners identification.
- 2009 **Rhône-Alpes local government, "CONVERSESSION", role: key personnel**  
 Design of a novel platform to analyse online debates; incubation of a start-up company.

## Research interests

- Interests** My current research is focused on: a) user privacy in the online media, b) modelling online human behaviour and interactions, c) explaining and predicting societal phenomena (such as the influence of online social bots on the democratic process or the adoption of disruptive technologies).
- My previous research dealt with: a) natural language processing and semantic image representation, b) knowledge injection into non-supervised learning algorithms, c) spatial-temporal clustering, and d) textual clustering.
- Keywords**
- Big data mining, applied machine learning, natural language processing;
  - online privacy, social network analysis, stochastic modelling, semi-supervised learning.

## Recent invited talks

- 06/2017 **Hawkes Intensity Processes for Social Media Popularity.** Invited seminar, University of Sydney, Australia.
- 03/2016 **Evolution of Privacy Loss in Wikipedia.** Invited talk, March session of the Monthly Wikimedia Research Showcase, San Francisco, USA.
- 06/2015 **Challenges of the social media: Information diffusion in social networks.** Invited speaker, Romanian Scientific and Business Diaspora Symposium, Canberra, Australia.
- 05/2015 **Semi-supervised analysis of complex data.** Invited talk. University of Nice, Sophia-Antipolis, France.
- 10/2013 **CommentWatcher: an open-source platform for analyzing online forum discussions.** Invited speaker, BLEND 2013 conference, Lyon, France.

## Divers

I love to play the piano and to explore the wild Australian outback.

# Publications

## International conferences:

- [1] Rizoïu, M.-A., Mishra, S., Kong, Q., Carman, M., & Xie, L. (2018). *SIR-Hawkes: on the Relationship Between Epidemic Models and Hawkes Point Processes*. In: Proceedings of International Conference on World Wide Web (**WWW '18**), Lyon, France, pp. 1–9, 2018. (**CoRE Rank: A\***, **a.r.:** 14%, **h5:** 77)
- [2] Rizoïu, M.-A., & Xie, L. *Online Popularity under Promotion: Viral Potential, Forecasting, and the Economics of Time*. In: Proceedings of International AAAI Conference on Web and Social Media (**ICWSM '17**), Montreal, Canada, pp. 1–10, 2017. (**a.r.:** 14%, **h5:** 59)
- [3] Rizoïu, M.-A., Xie, L., Sanner, S., Cebrian, M., Yu, H., & Van Hentenryck, P., *Expecting to be HIP: Hawkes Intensity Processes for Social Media Popularity*. In: Proceedings of International Conference on World Wide Web (**WWW '17**), Perth, Australia, pp. 735-744, 2017. (**CoRE Rank: A\***, **a.r.:** 17%, **h5:** 74)
- [4] Mishra, S., Rizoïu, M.-A., & Xie, L., *Feature Driven and Point Process Approaches for Popularity Prediction*. In: Proceedings of International Conference on Information and Knowledge Management (**CIKM '16**), Indianapolis, USA, p. 1069–1078, 2016. (**CoRE Rank: A**, **a.r.:** 17%, **h5:** 42)
- [5] Rizoïu, M.-A., Xie, L., Caetano, T., & Cebrian, M., *Evolution of Privacy Loss in Wikipedia*. In: Proc. International Conference on Web Search and Data Mining (**WSDM '16**), San Francisco, USA, pp. 215-224, February 2016. (**CoRE Rank: A\***, **a.r.:** 18%, **h5:** 58)
- [6] Kim, Y.-M., Velcin, J., Bonnevey, S., & Rizoïu, M.-A., *Temporal Multinomial Mixture for Instance-Oriented Evolutionary Clustering*. In: 37<sup>th</sup> European Conference on Information Retrieval (**ECIR '15**), Vienna, Austria, pp. 593–604, Springer, March 2015. (**CoRE Rank: B**, **a.r.:** 23%, **h5:** 25)
- [7] Rizoïu, M.-A., *Semi-Supervised Structuring of Complex Data*. In: Doctoral Consortium of the 23<sup>rd</sup> International Joint Conference On Artificial Intelligence (**IJCAI '13**). Beijing, China. AAAI Press. 2013. (**CoRE Rank: A\***, **h5:** 55)
- [8] Rizoïu, M.-A., Velcin, J. & Lallich, S., *Structuring typical evolutions using Temporal-Driven Constrained Clustering*. In: 24<sup>th</sup> International Conference on Tools with Artificial Intelligence (**ICTAI '12**), Athens, Greece, pp. 610-617, IEEE, November, 2012. **Best Student Paper Award**. (**CoRE Rank: B**, **h5:** 17)
- [9] Musat, C., Velcin J., Trausan-Matu, S., & Rizoïu M.-A. *Improving Topic Evaluation Using Conceptual Knowledge*. In: 22<sup>nd</sup> International Joint Conference On Artificial Intelligence (**IJCAI '11**). Barcelona, Spain. pp. 1866-1871, AAAI Press. July, 2011. (**CoRE Rank: A\***, **h5:** 55)
- [10] Musat, C., Velcin J., Rizoïu M.-A., & Trausan-Matu, S. *Improving Topic Models using Conceptual Data*. In: 19<sup>th</sup> International Symposium on Methodologies for Intelligent Systems (**ISMIS '11**). Warsaw, Poland. pp. 133–142. June 2011. (**CoRE Rank: C**, **h5:** 11)

## International journals:

- [11] Rizoïu, M.-A., Velcin, J., Bonnevey, S., & Lallich, S. *ClusPath: A Temporal-driven Clustering to Infer Typical Evolution Paths*. Data Mining and Knowledge Discovery (**DAMI**), 30(5), pp. 1324-1349, (2016). (**ERA Rank: A**, **h5:** 33)<sup>3</sup>
- [12] Rizoïu, M.-A., Velcin, J. & Lallich, S., *How to use Temporal-Driven Constrained Clustering to detect typical evolutions*. International Journal of Artificial Intelligence Tools (**IJAIT**), vol. 23(4), (2014). (**h5:** 13)
- [13] Rizoïu, M.-A., Velcin, J. & Lallich, S., *Semantic-enriched Visual Vocabulary Construction in a Weakly Supervised Context*. Intelligent Data Analysis (**IDA**), vol. 19(1), pp. 161-185, (2015). (**ERA Rank: B**, **h5:** 16)
- [14] Rizoïu, M.-A., Velcin, J. & Lallich, S., *Unsupervised Feature Construction for Improving Data Representation and Semantics*. Journal of Intelligent Information Systems (**JIIIS**), vol. 40(3), pp. 501-527, (2013). (**ERA Rank: C**, **h5:** 21)

## National scope journals and conferences:

- [15] Rizoïu, M.-A., Velcin, J., & Chauchat, J.-H., *Regrouper les données textuelles et nommer les groupes à l'aide des classes recouvrantes*. In: 10<sup>ème</sup> conférence Extraction et Gestion des Connaissances (**EGC '10**), Hammamet, Tunisia; Vol. E-19, p. 561-572. January, 2010. (**h5:** 6) (in French)
- [16] Musat, C., Rizoïu M.-A., & Trausan-Matu, S., *An Intra and Inter-Topic Evaluation and Cleansing Method*. Romanian Journal of Human-Computer Interaction (**RRIOC**), vol. 3 (2010) pp. 81-96. (in Romanian)

## Book chapters:

- [17] Rizoïu, M.-A., Lee, Y., Mishra, S., & Xie, L. *A Tutorial on Hawkes Processes for Events in Social Media*. In “Research Frontiers of Multimedia”, S.-F. Chang (Ed.), (2017), pp. 1–26, ACM Books.
- [18] Rizoïu, M.-A. & Velcin, J. *Topic Extraction for Ontology Learning*. In book: “Ontology Learning and Knowledge Discovery Using the Web: Challenges and Recent Advances” (2011), chapter 3, pp. 38-61.

## Theses

- [19] Rizoïu, M.-A. *Semi-Supervised Clustering of Complex Data*. PhD thesis, University Lumière Lyon, France. June 2013.
- [20] Rizoïu, M.-A. *Textual Data Clustering and Cluster Naming*. Masters Thesis, University of Nantes, France, June 2009.

## Under review

- [21] Wu, S., Rizoïu, M.-A., & Xie, L. (2017). *Beyond Views: Measuring and Predicting Engagement on YouTube Videos*. *arXiv Preprint*, 1–9. <https://arxiv.org/abs/1709.02541>
- [22] Kong, Q., Rizoïu, M.-A., Wu, S., & Xie, L. (2018). *Will This Video Go Viral? Explaining and Predicting the Popularity of Youtube Videos*.

3 Metrics key. **ERA/CoRE**: Australian Publication Ranking; **a.r.**: acceptance rate for conferences; **h5**: the h5 metric of the venue (source Google Scholar)