

Curriculum Vitae

Description

I am a data enthusiast, hacking and researching solutions to problems at the intersection of Information Retrieval (search), Natural Language Processing, and Machine Learning in general. I hold a PhD in computer science, where the focus of my research was to improve the effectiveness of search systems by modelling and identifying the different concepts of simple - and often ambiguous - keyword queries. I worked for a variety of academic (Universities of Glasgow, Toulouse, and Avignon) and industry (Yellow Pages Canada, OpenEdition) organisations, where I developed over the years the many practical and theoretical skills needed for the modern Research Scientist. I authored more than 20 papers published in top tier journals and conferences related to Data Mining, Information Retrieval and NLP (ACM TOIS, CIKM, SIGIR, ACL), and participated several times in data challenges and workshops mainly related to Information Retrieval topics (TREC, INEX, NTCIR). I enjoy learning new things every day, and can easily adapt to new tools and technologies if needed. I live in the French countryside with my family, passionate about riding sports, and advocating for working remotely in a calm and healthy environment.

Information

Romain Deveaud

French nationality

Citations : 450 — h-index : 11 (source: Google Scholar)

6, Menardeix

23140, Pionnat, France

✉ romain.deveaud@gmail.com

📄 <https://romaindeveaud.github.io>

☎ +33 (0)6 19 22 20 28 — +33 (0)9 72 52 62 32

Professional Experience

Postdoc researcher

10/2018 - 10/2019

OpenEdition, CNRS, Aix-Marseille Université

Appropriation du Savoir Ouvert project, publicly funded. Researching and analysing web logs data in order to describe how people visit and consume online open access academic documents. Fully remote. Analysis of web server logs to understand how people interact and consume the content of an open access platform for academic content. Implemented predictive models of dwell time, session-based document embeddings, taxonomy of returning users...

Supervised by: *Marin Dacos* and *Patrice Bellot*.

Postdoc researcher

05/2018 - 08/2018

Laboratoire Informatique d'Avignon (Computer science lab of Avignon)

AMIS project, publicly funded by the French government. Indexing and summarisation of news audio files (github.com/amis-anr/audio-summariser). Fully remote.

Unsupervised identification of salient and relevant audio parts from multi-lingual news reports. Development of an interface for human annotation and further supervised training based on the informativeness of the previously extracted audio segments.

Supervised by: *Eric SanJuan*.

Postdoc researcher

01/2016 - 12/2016

IRIT, Université Paul Sabatier, Toulouse

Projet ANR CAAS, *Contextual Analysis and Adaptive Search*; Learning to Rank applied to predicting the effectiveness of document retrieval systems configurations. Fully remote, worked part-time (50%).

Supervised by: *Josiane Mothe*, collaborating with *Jian-Yun Nie*.

Senior developer

01/2015 - 05/2015

Yellow Pages Canada, Montréal, Canada.

Team Search & Relevance.

Implemented an unsupervised categorisation of unknown words and queries using word embeddings and an in-house ontology.

Romain Deveaud

Postdoc researcher

10/2013 - 11/2014

Terrier team, University of Glasgow, UK.

European project FP7 SMART, information retrieval and recommendation in the context of smart cities. Open source development (see github.com/SmartSearch).

Development of unsupervised (probabilistic) and supervised (Learning to Rank) models for the recommendation of venues and points-of-interest in a zero-query setting. Supervised by: *Iadh Ounis* and *Craig Macdonald*.

Graduate student

10/2010 - 09/2013

Laboratoire Informatique d'Avignon (Computer science lab of Avignon).

ANR project CAAS, *Contextual Analysis and Adaptive Search*.

Modelling of the latent concepts of keyword queries using query-biased topic models (LDA). Participation in several workshops. Supervised by: *Eric SanJuan* and *Patrice Bellot*.

Web developer

09/2009 - 08/2010

Direction of Information Systems, Université d'Avignon et des Pays de Vaucluse.

Developing free and open source web applications designed to help collaborative work (see github.com/UAPV).

Supervised by: *Thierry Spriet*.

Software developer

06/2009 - 08/2009

Laboratoire Informatique d'Avignon (Computer science lab of Avignon).

European project FP6 LUNA, *Spoken language understanding in multilingual communication systems*.

Developing research prototypes and interfaces.

Supervised by: *Frédéric Béchet*.

Education**Ph.D. in Computer Science**

2013

Vers une représentation du contexte thématique en Recherche d'Information (Representing Topical Query Contexts in Information Retrieval), Very honorable mention.

Supervisors: *Eric SanJuan* and *Patrice Bellot*.

Viva passed at l'Université d'Avignon et des Pays de Vaucluse, November 29th, 2013.

Jury president : *Josiane Mothe*.

Reviewers : *Jian-Yun Nie* and *Philippe Mulhem*.

Examinators : *Jaap Kamps*, *Benjamin Piwowarski*, and *Jacques Savoy*.

M.Sc. in Computer Science

2010

Université d'Avignon et des Pays de Vaucluse.

Working part-time during last year, thesis: *Open-source software at the University of Avignon*.

B.Sc. in Computer Science

2008

Université d'Avignon et des Pays de Vaucluse.

DUT in Computer Science (2-year French diploma)

2007

Université de la Méditerranée, Aix-en-Provence.

Specialised in software engineering.

Skills & technologies

These are just a few highlighted skills, but being a fast learner I can quickly be proficient in new technologies if needed.

Programming : Python, R, PHP, Ruby, bash, ...

Frameworks & tools : pandas, pytorch, scikit-learn, matplotlib, ...

Other : L^AT_EX, scientific communication, vulgarisation

Romain Deveaud

Awards

Best paper award: CORIA 2013 (an English version of this paper currently has more than 150 citations)

Best paper runner-up award: MajecSTIC 2009

Publications

Journals

- [1] **Romain Deveaud**, Josiane Mothe, Md Zia Ullah and Jian-Yun Nie. Learning to Adaptively Rank Document Retrieval System Configurations. *ACM Transactions on Information Systems (TOIS)*, Volume 37 Issue 1, January 2019.
- [2] **Romain Deveaud** and Florian Boudin. De quoi parle ce Tweet? Resumer Wikipédia pour contextualiser des microblogs. *Interaction Intelligence Information*, 2014.
- [3] **Romain Deveaud**, Eric SanJuan, and Patrice Bellot. Accurate and Effective Latent Concept Modeling for Ad Hoc Information Retrieval. *Document Numérique*, 2014.

Conferences

- [4] Carlos-Emiliano González-Gallardo, **Romain Deveaud**, Eric SanJuan, and Juan-Manuel Torres-Moreno. Audio Summarization with Audio Features and Probability Distribution Divergence. *Proceedings of CICLing*, 2019.
- [5] **Romain Deveaud**, Josiane Mothe and Jian-Yun Nie. Learning to Rank System Configurations. *Proceedings of CIKM*, pages 2001–2004, 2016. Acceptance rate: 23%.
- [6] **Romain Deveaud**, Véronique Moriceau, Josiane Mothe and Eric Sanjuan. Informativeness for Adhoc IR evaluation: a measure that prevents assessing individual documents In *Proceedings of ECIR*, pages 818–823, 2016. Acceptance rate: 25%.
- [7] **Romain Deveaud**, M-Dyaa Albakour, Craig Macdonald and Iadh Ounis. Experiments with a Venue-Centric Model for Personalised and Time-Aware Venue Suggestion In *Proceedings of CIKM*, pages 53–62, 2015. Acceptance rate: 16%.
- [8] Graham McDonald, **Romain Deveaud**, Richard McCreddie, Craig Macdonald and Iadh Ounis. Tweet Enrichment for Effective Dimensions Classification in Online Reputation Management. In *Proceedings of ICWSM*, pages 654–657, 2015. Acceptance rate: 27%.
- [9] Thibaut Thonet, **Romain Deveaud**, Craig Macdonald and Iadh Ounis. Suggestion contextuelle composite. In *Actes de CORIA*, pages 89–104, 2015. Acceptance rate: 24%.
- [10] **Romain Deveaud**, M-Dyaa Albakour, Craig Macdonald and Iadh Ounis. On the Importance of Venue-Dependent Features for Learning to Rank Contextual Suggestions In *Proceedings of CIKM*, pages 1827–1830, 2014. Acceptance rate: 25%.
- [11] **Romain Deveaud**, M-Dyaa Albakour, Jarana Manotumrukka, Craig Macdonald and Iadh Ounis. Diversifying Contextual Suggestions from Location-based Social Networks In *Proceedings of CIKM (demo)*, pages 125–134, 2014. Acceptance rate: 39%.
- [12] M-Dyaa Albakour, **Romain Deveaud**, Craig Macdonald and Iadh Ounis. Diversifying Contextual Suggestions from Location-based Social Networks In *Proceedings of IiX*, pages 125–134, 2014. Acceptance rate: 46%.
- [13] **Romain Deveaud**, Eric SanJuan, and Patrice Bellot. Are Semantically Coherent Topic Models Useful for Information Retrieval? In *Proceedings of ACL*, pages 148–152, 2013. Acceptance rate: 24%.
- [14] **Romain Deveaud**, Eric SanJuan, and Patrice Bellot. Estimating Topical Context by Diverging from External Resources. In *Proceedings of SIGIR*, pages 1001–1004, 2013. Acceptance rate: 35%.
- [15] **Romain Deveaud**, Eric SanJuan, and Patrice Bellot. Unsupervised Latent Concept Modeling to Identify Query Facets. In *Proceedings of OAIR*, pages 93–100, 2013. Acceptance rate: 22.5%.

- [16] Ludovic Bonnefoy, Vincent Bouvier, **Romain Deveaud**, and Patrice Bellot. Vers une détection en temps réel de documents web centrés sur une entité donnée. In *Actes de CORIA*, pages 21–36, 2013. Acceptance rate: 20%.
- [17] **Romain Deveaud**, Ludovic Bonnefoy, and Patrice Bellot. Quantification et identification des concepts implicites d’une requête. In *Actes de CORIA*, pages 159–174, 2013. Acceptance rate: 20%, **Best paper award**.
- [18] **Romain Deveaud** and Florian Boudin. Contextualisation automatique de tweets à partir de wikipédia. In *Actes de CORIA*, pages 125–140, 2013. Acceptance rate: 20%.
- [19] **Romain Deveaud** and Patrice Bellot. Combinaison de ressources générales pour une contextualisation implicite de requêtes. In *Actes de TALN*, pages 479–486, 2012.
- [20] **Romain Deveaud**, Eric SanJuan, and Patrice Bellot. Ajout d’informations contextuelles issues de Wikipédia pour la recherche de passages. In *Actes de TALN*, pages 63–68, 2011.
- [21] **Romain Deveaud**, Florian Boudin, Eric SanJuan, and Patrice Bellot. Correction de césures et enrichissement de requêtes pour la recherche de livres. In *Actes de CORIA*, pages 89–96, 2011.
- [22] Ludovic Bonnefoy, **Romain Deveaud**, and Eric Charton. Interrogations de moteurs de recherche par des requêtes formulées en langage naturel. In *Actes de MajecSTIC*, 2009. **“Best hope” special award**.

Workshops

- [23] Richard McCreadie, **Romain Deveaud**, M-Dyaa Albakour, Stuart Mackie, Craig Macdonald, Iadh Ounis, Thibaut Thonet and Bekir Taner Dinçer. University of Glasgow at TREC 2014: Experiments with Terrier in Contextual Suggestion, Temporal Summarisation and Web Tracks. In *Proceedings of TREC*, 2014.
- [24] Graham McDonald, **Romain Deveaud**, Richard McCreadie, Timothy Gollins, Craig Macdonald and Iadh Ounis. University of Glasgow Terrier Team / Project Abacá at RepLab 2014: Reputation Dimensions Task. In *Proceedings of RepLab @ CLEF 2014*, 2014.
- [25] **Romain Deveaud**, M-Dyaa Albakour, Craig Macdonald, and Iadh Ounis. Challenges in Recommending Venues within Smart Cities. In *Proceedings of the ECIR 2014 Workshop on Information Access in Smart Cities*, i-ASC, 2014.
- [26] **Romain Deveaud** and Florian Boudin. Effective Tweet Contextualization with Hashtags Performance Prediction and Multi-Document Summarization. In *Proceedings of INEX*, 2013.
- [27] **Romain Deveaud** and Eric SanJuan. LIA at the NTCIR-10 INTENT Task. In *Proceedings of NTCIR*, 2013.
- [28] **Romain Deveaud**, Eric SanJuan, and Patrice Bellot. LIA at TREC 2012 Web Track : Unsupervised Search Concepts Identification from General Sources of Information. In *Proceedings of TREC*, 2012.
- [29] Ludovic Bonnefoy, **Romain Deveaud**, and Patrice Bellot. Do Social Information Help Book Search? In *Proceedings of INEX*, 2012.
- [30] **Romain Deveaud** and Florian Boudin. LIA/LINA at the INEX 2012 Tweet Contextualization track. In *Proceedings of INEX*, 2012.
- [31] Pierre Joulin, **Romain Deveaud**, Eric SanJuan-Ibekwe, Jean-Marc Francony, and Françoise Para. Design, implementation and experiment of a YeSQL Web Crawler. In *Proceedings of the SIGIR 2012 Workshop on Open Source Information Retrieval*, OSIR, 2012.
- [32] **Romain Deveaud**, Eric SanJuan, and Patrice Bellot. Social Recommendation and External Resources for Book Search. In *Proceedings of INEX*, pages 68–79, 2011.
- [33] **Romain Deveaud**, Eric SanJuan, and Patrice Bellot. LIA at TREC 2011 Web Track: Experiments on the Combination of Online Resources. In *Proceedings of TREC*, 2011.
- [34] **Romain Deveaud**, Florian Boudin, and Patrice Bellot. LIA at INEX 2010 Book Track. In *Proceedings of INEX*, pages 118–127, 2010.

PhD thesis

[35] **Romain Deveaud**. Vers une représentation du contexte thématique en Recherche d'Information, 2013.