Gaël Lejeune | Ph.D

✉ gael.lejeune@sorbonne-universite.fr • 🖆 /www.lejeunegael.fr
in fr.linkedin.com/pub/57/836/a77 • ⊕ https://github.com/rundimeco
Deputy director of the CERES unit

Experience

Sorbonne University	**France**
Associate professor	*since Sep. 2017*
Research : Microblogs Classification, Document Datation and String Algorithmics	
University Paris 13	**France**
Postdoctorate Fellow	*Jan. 2017 – Aug. 2017*
Research : Multilingual Neologism detection and Tracking	
University of Caen	**France**
Postdoctoral Fellow	*Sep. 2015 – Dec. 2016*
Research : Event Prediction for Dialogue Modelling	
University of Nantes	**France**
Assistant Lecturer	*Sep. 2014 – Aug. 2015*
Research : Multilingual Text Mining	
University of Caen	**France**
Assistant Lecturer	*Sep. 2013 – Aug. 2014*
Research : Genre-based Natural Language Processing	
University of Caen	**France**
PhD Student	*Oct. 2010 – Aug. 2013*
Multilingual Epidemic Surveillance	
GREYC, University of Caen	**France**
Research Engineer	*Sep. 2009–Aug. 2010*
Discourse rules for text mining	
University of Helsinki	**Finland**
Research Assistant	*Feb. 2009 – Jun. 2009*
Multilingual extension of IE systems, aggregation of facts across languages	

Education

Associate Professor Qualification	**2014-2017**
National Council of Universities	*France*
Qualification in two fields: Computer Science and Linguistics	
PhD in Computer Science	**2010-2013**
University of Caen Basse-Normandie	*France*
Multilingual epidemic surveillance : a parsimonious character-based approach	
Master in Translation studies and Cognitive Science	**2006-2009**
University of Caen Basse-Normandie/Ionian University	*France/Greece*

- Ontological aspects in machine translation
- Towards multilingual information extraction

Bachelor in Linguistics	**2005-2006**
University of Caen Basse-Normandie	*France*
Bachelor of Public Administration	**2001-2004**
University of Caen Basse-Normandie	*France*

Skills and Interests

Programming languages and tools...

- Python, Php, Javascript, Sql
- Bash, Linux
- Weka, Pytorch, Nltk
- Svn, LaTeX

Research interests
○ Natural Language Processing, COmputational Linguistics
○ Multilingual Information Access
○ Text Mining and Text Analytics
○ Information Extraction and Information Retrieval
○ Health Informatics
○ Machine Learning, Deep Learning
○ Text Algorithmics

Natural Languages
○ French (mother tongue)
○ English (written, spoken, fluent)
○ Spanish and Russian (written, spoken)
○ Finnish, Greek and Polish (basic knowledge)

Selected Publications

Multilingual Epidemiological Text Classification: A Comparative Study
Proceedings of COLING 2020 (CORE A conference)
Stephen Mutuvi and Emanuela Boros and Antoine Doucet and Adam Jatowt and Gaël Lejeune and Moses Odeo

Impact Analysis of Document Digitization on Event Extraction
Natural Language for Artificial Intelligence NL4AI (AI*AI 2020 workshop, CORE B)
Nhu Khoa Nguyen and Emanuela Boros and Gaël Lejeune and Antoine Doucet

Out-of-the-Box and into the Ditch? Multilingual Evaluation of Generic Text Extraction Tools
12th Web as Corpus WAC (LREC 2020 Workshop CORE C)
Adrien Barbaresi and Gaël Lejeune

Daniel@FinTOC-2019 Shared Task : TOC Extraction and Title Detection
Second Financial Narrative Processing Workshop (FNP 2019)
Emmanuel Giguet and Gaël Lejeune

Analyse automatique de documents anciens : tirer parti d'un corpus incomplet, hétérogène et bruité
RIDOWS Journal (Recherche d'Information, Document, Web sémantique), vol 19-1, 2019
Karine Abiven and Gaël Lejeune

Analyse Diachronique de Corpus: le cas du poker
Journées d'Analyse des Données Textuelles (JADT 2018, CORE C conference)
Lichao Zhu and Gaël Lejeune

Modèles en Caractères pour la Détection de Polarité dans les Tweets
Défi Fouille de Textes 2018 (TALN 2018)
Davide Buscaldi, Joseph Le Roux and Gaël Lejeune

Character Based Pattern Mining for Neology Detection
Sub Word and Character Level models for NLP (EMNLP 2017 workshop, CORE A)
Gaël Lejeune and Emmanuel Cartier

Highlighting Psychological Features for Predicting Child Interjections During Story Telling
Interspeech 2016 (CORE A conference)
Gaël Lejeune, François Rioult and Bruno Crémilleux

Ambiguity Diagnosis for Terms in Digital Humanities
Language and Resources Conference 2016 (CORE C conference)
Béatrice Daille, Evelyne Jacquey, Gaël Lejeune, Luis F. Melo and Yannick Toussaint

Multilingual Event Extraction for Early Epidemic Detection
Journal of Artificial Intelligence in Medicine 2015 (CORE A jorunal), Yearbook of Medical Informatics Best Papers
Gaël Lejeune, Romain Brixtel, Antoine Doucet and Nadine Lucas

Any Language Early Detection of Epidemic Diseases from Web News Streams
International Conference on Healthcare Informatics 2013 (acceptation arate <20%)
Romain Brixtel, Gaël Lejeune, Antoine Doucet and Nadine Lucas

Added-value of automatic multilingual text analysis for epidemic surveillance
Artificial Intelligence in Medicine 2013 (CORE A conference)
Gaël Lejeune, Romain Brixtel, Charlotte Lecluze, Antoine Doucet and Nadine Lucas

Complete list: `http://www.lejeunegael.fr/#publications`