

# Georgios Balikas

---

## Research Interests

Natural language processing, machine learning, information retrieval and data mining theory and applications.

I am especially interested in multi-view representation learning. In a part of my research this translates to multi-lingual and cross-lingual representation learning. In particular, distributional representations and topic modeling as well as their extensions to multi/cross-lingual settings.

From a theoretic point-of-view I am interested in performance bounds, with a focus on how one can leverage unlabeled data.

---

## Education

Nov 2014 - **PhD in Artificial Intelligence**

*University of Grenoble-Alpes*

*France*

Thesis: Multiview Learning for Sequence Extraction Tasks. Industry-funded PhD in collaboration with Coffreo, a company situated in Clermont-Ferrand.

Sep 2013 - Sep 2014 **MSc in Artificial Intelligence**

*Université Pierre & Marie Curie - Paris VI*

*France*

Specialization: Machine Learning and Decision Making

Sep 2005 - Jul 2011 **MEng in Electrical & Computer Engineering**

*Aristotle University of Thessaloniki*

*Greece*

Specialization on Electronics and Computer Engineering, Grade: 7.89/10, "Very Good"

2005 **High School Diploma,**

*1<sup>st</sup> High School of Kavala, Grade: 19.5/20, "Excellent"*

*Greece*

---

## Publications

- Journal G. Tsatsaronis, G. Balikas, I. Partalas, M. Zschunke, M. R. Alvers, D. Weissenborn, A. Krithara, S. Petridis, D. Polychronopoulos, Y. Almirantis, P. Malakasiotis, I. Androutsopoulos, J. Pavlopoulos, N. Baskiotis, P. Gallinari, T. Artières, A. Ngonga, N. Heino, E. Gaussier, L. Barrio-Alvers, M. Schroeder and G. Paliouras: **An overview of the BioASQ large-scale biomedical semantic indexing and question answering competition**, BMC Bioinformatics, 2015
- Conference G. Balikas, M-R. Amini: **On a Topic Model for Sentences**, short paper to appear in Proceedings of the 39th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), Pisa, Italy, 17-21 July, 2016
- Conference G. Balikas, M-R. Amini: **Multi-label, Multi-class Classification Using Polylingual Embeddings**, short paper in Proceedings of the 38th European Conference on Information Retrieval (ECIR), Padua, Italy, 20-23 March, 2016
- Conference G. Balikas, M-R. Amini: **Language-Independent Embeddings for Text Summarization**, PhD session of the 14th International Symposium on Intelligent Data Analysis (IDA), Saint-Etienne, France, 22-24 October, 2015

- Conference G. Balikas, I. Partalas, E. Gaussier, R. Babbar, M-R. Amini: **A New, Efficient Model Selection Method for Regularized Classification**, full paper in Proceedings of the 14th International Symposium on Intelligent Data Analysis (IDA), Saint-Etienne, France, 22-24 October, 2015 IDA, 2015
- Conference G. Balikas, A. Kosmopoulos, A. Krithara, G. Paliouras and I. Kakadiaris: **Results of the BioASQ Track of the Question Answering Lab at CLEF 2015**, working notes of the Conference and Labs of the Evaluation Forum (CLEF), Toulouse, France, 8-11 September, 2015
- Conference G. Balikas, N. Baskiotis: **Detecting Influence in User-Activity Logs**, 1st International Conference on Computational Social Science (ICCSS), Helsinki, Finland, 8-11 June, 2015
- Conference G. Balikas, I. Partalas, A.N. Ngomo, A. Krithara, E. Gaussier and G. Paliouras: **Results of the BioASQ Track of the Question Answering Lab at CLEF 2014**, working notes of the Conference and Labs of the Evaluation Forum (CLEF), Sheffield, UK, 15-18 September, 2014
- Workshop G. Balikas and M-R Amini: **TwISE at SemEval-2016 Task: Twitter Sentiment Classification**, to appear in Proceedings of 10th International Workshop on Semantic Evaluations (SemEval), San Diego, California, U.S.A., 12-17 June, 2016
- Workshop G. Balikas, A. Krithara, I. Partalas, G. Paliouras: **BioASQ: A challenge on large-scale biomedical semantic indexing and question answering**, in the Proceedings of the 1st International Multimodal Retrieval in the Medical Domain Workshop in connection with the 37th European Conference on Information Retrieval (ECIR), Vienna, Austria, 29 March-2 April, 2015
- National Conference G. Balikas and M-R Amini: **Learning language-independent sentence representations for multi-lingual, multi-document summarization**, in Proceedings of Conférence sur l'Apprentissage Automatique (CAp), Lille, France, 5 July, 2015
- National Conference A. Goyal, G. Balikas, P. Goswami, M-R Amini and E. Gaussier: **Transfer Learning for IR Model Parameter Tuning** in Proceedings of Conférence sur l'Apprentissage Automatique (CAp), Lille, France, 5 July, 2015

---

## Professional Experience

- Nov 2014- **Software Engineer & Research Associate**  
*Coffreo, Clermont-Ferrand* *France*  
 • Transfer knowledge from academia to the company. • Benchmark document classification systems. • Extract knowledge from plain-text. • Develop a system to collect documents from a repository which are described in a plain-text call.
- Dec 2012-Sep 2014 **Software Engineer in the BioASQ project**  
*Laboratoire d'Informatique de Paris 6 (L.I.P. 6)* *France*  
 • Development of a MySQL-based Web platform using Python and Django ([participants-area.bioasq.org](http://participants-area.bioasq.org)), • development of JSON-based Web APIs, • pre-processing of large volumes of text data, • implementation of evaluation measures, • user support.
- Apr 2014-Sep 2014 **Detecting influence from user-activity logs**  
*Laboratoire d'Informatique de Paris 6 (L.I.P. 6)* *France*  
 • Given action logs of a music platform, investigate if influence exists. • implementation of twitter crawlers for data collection. • Estimate the probability of influence between users by proposing a generative model. • Design an influence graph. • Investigate the role of the temporal dimension in the data.

Jan 2011-Jul 2011 **MEng Project: Design, development and implementation of controllers for robotic constructions using Lego NXT**  
*Robotics and Automation Laboratory, Aristotle University* *Greece*  
R&D of PID controllers in Matlab applied on Lego Mindstorms NXT for: • position control, • trajectory tracking, • grasping and handling of objects with parallel, rigid sides.

---

## Awards

SemEval 2016 Twitter Sentiment Evaluation (Task 4): ranked **1/11** in subtask C "Tweet classification according to a five-point scale" and **2/14** in subtask D "Tweet quantification according to a two-point scale".

cDiscount challenge Large scale classification of short product titles ([www.datascience.net](http://www.datascience.net)). Ranked **10/175**.

Travel Awards Travel grants to present my work and attend the International Conference on Computational Social Science (ICCSS 2015) and the European Conference on Information retrieval (ECIR 2015).

---

## Summer Schools

ESSIR 2015 European Summer School on Information Retrieval

---

## Technical Skills

Programming:	Python, Bash, Matlab, C/C++	Databases:	MySQL, Oracle, SQLite, MongoDB, Neo4j
Web Development:	Django, (X)HTML, CSS, jQuery, XML, RDF/SPARQL	Web services:	REST APIs, JSON,
Progr. Paradigms:	MVC, Map/Reduce	Misc:	L <sup>A</sup> T <sub>E</sub> X, Git, subversion

---

## Languages

Greek:	Native	German:	Basic (B <sub>2</sub> : Goethe Institute)
English:	Fluent	French:	Intermediate

---

## Personal Information

Hobbies: Cycling, table tennis, basketball, greek traditional dances, cinema, classic literature

Volunteer: President of the Committee of Residents in Fondation Hellénique, 2013-2014

Experience: Greek dances teacher, 2007-2014