

# CURRICULUM VITÆ

Dr. ANASTASIA KRITHARA

---

## PERSONAL DETAILS

Date of Birth: 21/12/1981

Telephone: +30 210 6503204

Email: akrithara@iit.demokritos.gr, nkrithara@gmail.com

---

## RESEARCH INTERESTS

Machine Learning, Information Retrieval, Artificial Intelligence, Knowledge Discovery

---

## WORKING EXPERIENCE

- 10/2008– **Associate Researcher.** Institute of Informatics & Telecommunications(IIT), NCSR "Demokritos".
- 04/2005– 04/2008 **Research Engineer** (under a Cifre grant). Xerox Research Center Europe (XRCE), Grenoble, France. Member of the Text and Visual Pattern Analysis (TVPA) group.  
*Project: Learning Aspect Models with Partially Labeled data.*
- 05/2004–10/2004 **Research intern.** Hewlett-Packard Research Labs, Bristol, UK. Department of Semantic and Adaptive Systems, BICAS group.  
*Project: Semantic Classification.*
- 08/2001–09/2001 **Intern** in IT Department of "Water Supply Company of Athens" (EYDAP).  
*Post:* Assistant of Database Administrator. *Object:* Installation and administration of relational databased in Microsoft SQL Server environment. Monitoring of Local Area Networks (LAN) and Wide Area Networks (WAN).

---

## EDUCATION

- 04/2005–06/2008 **Ph.D. in Informatics**, University Pierre et Marie Curie, Paris VI, France.  
Thesis: *Learning Aspect Models with Partially Labeled data*.  
The motivation is to explore new methods which take advantage of both semi-supervised and active learning and try to solve the annotation cost problem. In the framework of the current work, conducted in Xerox Research Centre Europe, the focus is text classification, and in particular, extensions of the existing classification system.  
Supervisors: Prof. P.Gallinari, Ass. Prof. M. Amini and Dr. J.M. Renders
- 10/2003–10/2004 **MSc. in Advanced Computing**, Machine Learning and Data Mining, University of Bristol, UK  
Thesis: *Semantic Classification*. A new approach was investigated, constructing and using semantic features as rich input to machine learners. The approach was tested in HP problem and solution descriptions dataset.  
Supervisors: Dr. S. Cayzer and Prof. P. Flach.
- 10/1999–09/2003 **Degree in Informatics**, Athens University of Economics and Business, Athens Greece  
Thesis: *Network Management: monitoring Access Server AS5300*  
Supervisor: Prof. T. Apostolopoulos.

---

## PROJECTS

- **SYNC3 (Synergetic Content Creation and Communication)**.  
Objective: SYNC3 will provide an intelligent framework for making more accessible the vast quantity of user comments on news issues. The project will structure the part of blogosphere that refers to running news stories, rendering it accessible, manageable and re-usable. The immediate target of SYNC3 is the news industry and social networks, but domains like commerce, tourism, e-science and business intelligence are likely to benefit from the resulting technology.  
Participants: Athens Technology Center SA (EL), NCSR "Demokritos" (EL), Gottfried Wilhelm Leibniz Universität Hannover -L3S (DE), The University of Edinburgh (UK), Xerox Research Centre Europe (FR), European Journalism Centre- EJC (NL), i-Sieve (EL), RIA Novosti (RS), Google (UK).  
Programme: Information and Communication Technologies (Seventh Framework), European Union.  
Start: April 2009. End: March 2012.
- **CLARIN-EL (Common Language Resources and Technologies)**.

Objective: CLARIN-EL is the Greek national counterpart of the CLARIN Research Infrastructure (RI); as such, it seeks to implement CLARIN's goals and visions focusing on the Greek language and the needs of the Greek research community. CLARIN is a RI project in the Social Sciences and Humanities (SSH) thematic domain of the ESFRI (European Strategy Forum on Research Infrastructures) Roadmap. It is a large-scale pan-European collaborative effort aiming to create an infrastructure that will enhance and innovate research in SSH deploying Language Resources & Technology (LRT). More specifically, CLARIN intends to integrate digital content of any type (text, audio, image, video) and the relevant Language Resources (annotated recordings, texts, lexica, ontologies etc.) and Language Technologies (speech recognizers, lemmatisers, parsers, summarizers, information extractors etc.) into a coherent repository which will be available to scholars of all disciplines via web services.

Network Coordinator: Institute for Language and Speech Processing (ILSP).  
Start: 2008. End: 2011

- **BOEMIE (Bootstrapping Ontology Evolution with Multimedia Information Extraction).**

Objective: BOEMIE advocates a synergistic approach that combines multimedia extraction and ontology evolution in a bootstrapping process involving, on the one hand, the continuous extraction of semantic information from multimedia content in order to populate and enrich the ontologies and, on the other hand, the deployment of these ontologies to enhance the robustness of the extraction system.

Participants: NCSR "Demokritos" (EL), Fraunhofer-Gesellschaft / IMK (DE), University of Milano (IT), Hamburg University of Technology (DE), CERTH-ITI (EL), Tele Atlas (NL)

Programme: Information Society Technologies (Sixth Framework), European Union.

Start: March 2006. End: February 2009.

---

## PUBLICATIONS

- **A. Krithara**, M. Amini, C. Goutte and J.M. Renders. Learning Aspect Models with Partially Labeled Data. Pattern Recognition Letters -Elsevier, September 2010.
- **A. Krithara**, M. Amini, C. Goutte and J.M. Renders. An extension of the Aspect PLSA Model to Active and Semi-supervised Learning for Text Classification. 6th Hellenic Conference on Artificial Intelligence (SETN), Athens, Greece, May 2010.
- G. Petasis, V. Karkaletsis, **A. Krithara**, G. Paliouras and C.D. Spyropoulos. *Semi-automated ontology learning: the BOEMIE approach*. International Workshop on Inductive Reasoning and Machine Learning on the Semantic Web (IRMLoS), European Semantic Web Conference (ESWC), June 2009.
- **A. Krithara** *Learning Aspect Models with Partially Labeled Data*. PhD

thesis. University Pierre et Marie Curie (Paris VI). Defended the 30th June 2008.

- **A. Krithara**, M. Amini, J.M. Renders and C. Goutte. *Semi-Supervised Document Classification with a Mislabeling Error Model*. 30th European Conference on Information Retrieval (ECIR 2008), Glasgow, Scotland.
- **A. Krithara**, C. Goutte, J.M. Renders, and M.R. Amini. *Reducing the annotation burden in text classification*. In Proceedings of the 1st International Conference on Multidisciplinary Information Sciences and Technologies (InSciT 2006), Merida, Spain, Oct 2006.
- **A. Krithara**, C. Goutte, M. Amini and J.-M. Renders. *Active, Semi-Supervised Learning for Textual Information Access*. International Workshop on Intelligent Information Access (IIIA 2006), Helsinki, Finland, 2006.
- **A. Krithara**. *Semantic Classification*. Technical report. HPL-2004-182. HP Labs, Bristol, September 2004.

---

## PRESENTATIONS

- |          |   |
|----------|---|
| 07/05/10 | <i>An extension of the Aspect PLSA Model to Active and Semi-supervised Learning for Text Classification</i> . SETN 2010, Athens, Greece.  |
| 03/07/09 | <i>Semi-automated ontology learning: the BOEMIE approach</i> . IRMLeS, ESWC 2009, Heraklion, Greece.  |
| 30/06/08 | <i>Apprentissage de Modèles Latents avec des Données partiellement étiquetées</i> . Computer Science Laboratory of Paris 6 (LIP6), Paris, France. Defence of the PhD (in French). |
| 11/06/08 | <i>Learning Aspect Models with Partially Labeled Data</i> . NCSR “Demokritos”, Institute of Informatics & Telecommunications (IIT). Athens, Greece.                               |
| 02/04/08 | <i>Semi-Supervised Document Classification with a Mislabeling Error Model</i> . ECIR 2008, Glasgow, Scotland  |
| 24/08/07 | <i>Semi-Supervised Learning for Text Classification</i> . Machine Learning Summer School 2007. Student session. Tuebingen, Germany.   |
| 10/07/07 | <i>Apprentissage semi-supervisé et actif pour la classification du texte</i> . Computer Science Laboratory of Paris 6 (LIP6), Paris, France. Pre-defence of the PhD (in French).  |
| 05/07    | <i>Semi-Supervised and active learning for textual information access</i> . Xerox Seminars, Grenoble, France.   |
| 28/10/06 | <i>Reducing the Annotation Burden in Text Classification</i> . In-SciT 2006, Merida, Spain.   |

07/07/06      *Active, Semi-Supervised Learning for Textual Information Access*. IIIA 2006, Helsinki, Finland.

---

#### ATTENDANCES

30/03/08      Advanced language modeling approaches (Case study: Expert search). Tutorial.  
Search and Discovery in User-Generated Text Content. Tutorial.  
Glasgow, Scotland.

20-31/08/07      Machine Learning Summer School 2007 (MLSS07). Tuebingen, Germany.

06-08/07/06      International Workshop in Intelligent Information Access, Helsinki, Finland.

11/08/05      ROC Analysis in Machine Learning Workshop, Bonn, Germany.

07/08/05      Workshop in Partially Labeled data, Bonn, Germany.

---

#### TECHNOLOGIES/ PLATFORMS/ TOOLS

**Scientific languages:** Matlab, Java, C/C++, Perl, Tcl.

**Operating Systems:** Linux, Unix, Windows platforms.

---

#### LANGUAGE KNOWLEDGE

**Greek**      Native

**English**      Fluent

**French**      Fluent

---

#### ACTIVITIES

- Member of the European Networks of Excellence **PASCAL** and **PASCAL-2**  
(Pattern Analysis, Statistical Modelling and Computational Learning)
- External reviewer for several journals and conferences: JASIST (Journal of the American Society for Information Science and Technology), 5th IFIP Conference on Artificial Intelligence Applications & Innovations (AIAI2009), 6th Annual European Semantic Web Conference (ESWC2009), Knowledge and Information Systems (KAIS), Health Informatics Journal (HIJ), 12th Conference on Artificial Intelligence in Medicine (AIME'09).

---

## REFERENCES

Available upon request