Ansaf Salleb-Aouissi, Ph.D.

Columbia University, New York
Phone: 212-939-7117
Email: `as2933@columbia.edu` or `ansaf@cs.columbia.edu`
Homepage: `http://www1.ccls.columbia.edu/~ansaf/`
Google Scholar: `https://scholar.google.com/citations?user=77kOUvYAAAAJ&hl=en`

————————RESEARCH INTERESTS————————

Machine learning, data science, real-world applications. Classification, frequent patterns, unsupervised learning, crowd-sourcing, document analysis, clinical informatics, big data, educational data science.

————————EDUCATION————————

INRIA Rennes, France	Machine Learning	Postdoc fellow 2005
University of Orléans, France	Machine Learning	Ph.D. 2003 (joint scholarship from the French government and industry)
University of Orléans, France	Computer Science	M.S. 1999
USTHB Algiers, Algeria	Computer Science	Engineer 1996

————————EMPLOYMENT————————

2015-present   **DEPARTMENT OF COMPUTER SCIENCE, COLUMBIA UNIVERSITY.**
LECTURER IN DISCIPLINE.

2015-2016   **DATA SCIENCE INSTITUTE, COLUMBIA UNIVERSITY.**
LECTURER IN DISCIPLINE.

2006-2015   **CENTER FOR COMPUTATIONAL LEARNING SYSTEMS(CCLS) COLUMBIA UNIVERSITY.**
ASSOCIATE RESEARCH SCIENTIST.

2014-2015   **DEPARTMENT OF COMPUTER SCIENCE AT COLUMBIA UNIVERSITY.**
ADJUNCT ASSISTANT PROFESSOR.

ě   **2012-Present**   **HEALTH ANALYTICS CENTER. COLUMBIA UNIVERSITY DATA SCIENCE INSTITUTE.**
AFFILIATED FACULTY MEMBER.

2004-2005   **FRENCH NATIONAL INSTITUTE OF COMPUTER SCIENCE AND CONTROL (INRIA) RENNES, FRANCE.**
INRIA POSTDOCTORAL RESEARCH FELLOW.

2002-2004   **UNIVERSITY OF ORLÉANS, COMPUTER SCIENCE DEPARTMENT AND LABORATOIRE D'INFORMATIQUE FONDAMENTALE D'ORLÉANS (LIFO).**
ADJUNCT ASSISTANT PROFESSOR AND RESEARCHER.

1999-2002   **UNIVERSITY OF ORLÉANS, COMPUTER SCIENCE DEPARTMENT AND LABORATOIRE D'INFORMATIQUE FONDAMENTALE D'ORLÉANS (LIFO).**
TEACHING ASSISTANT.

EDUCATION ACTIVITIES

- **Teaching**:

  1. Columbia University: Discrete Mathematics COMS 3203 – Fall 2016.

  2. Columbia University: Introduction to Computing for engineers and applied sciences ENGIE1006 – Fall 2016.

  3. Columbia University: Artificial Intelligence COMS 4701 – Summer 2016.

  4. Columbia University: EdX Data Science and Analytics XSeries DS102X – Spring 2016.

  5. Columbia University: Data Science Capstone & Ethics ENGIE4800 – Spring 2016.

  6. Columbia University: Discrete Mathematics COMS 3203 – Fall 2015 and Spring 2016. `http://www1.ccls.columbia.edu/~ansaf/3203/`

  7. Columbia University: Introduction to Data Science COMS 4242 – Spring 2015. `http://www1.ccls.columbia.edu/~ansaf/4242/`

  8. Columbia University: Artificial Intelligence COMS 4701 – Fall 2014. `http://www1.ccls.columbia.edu/~ansaf/4701/`

  9. Columbia University: Machine Learning for Data Science COMS 4721 – Spring 2014. `http://www1.ccls.columbia.edu/~ansaf/4721/`

  10. Columbia University: Advanced Machine Learning COMS 6772 (Two lectures - 4h - Fall 2008).

  11. University of Orléans (France): courses in topics including: data mining, operating systems, discrete mathematics, databases, computer programming.

- **Graduate advisees**:

  1. Antonio Moretti, Ph.D. candidate, starting January 2015, Columbia University.

  2. Faiza Khan Khattak, Ph.D. candidate, recipient of a Fulbright fellowship. Columbia University. Defended June 29th, 2016.

  3. Bert Huang, Ph.D. Defended July 2011. Columbia University. Now Assistant professor at Virginia Tech Department of Computer Science.

  4. Ronan Trépos, Ph.D., co-supervision between 2004-2005 with Prof. M.-O. Cordier. Defended in 2008. University of Rennes I, France. Now researcher at INRA, France.

MAJOR PUBLICATIONS

- **Press**

  [P2] S. Conova, Why Mothers Deliver Early - And How To Stop It
  Columbia Medicine Magazine Volume 35 No. 2, 2016.
  http://www.columbiamedicinemagazine.org/features/fall-2015/why-mothers-deliver-early-Ð-and-how-stop-it

  [P1] Columbia Office of the Executive Vice President for Research Newsletter Spring 2015. Page 3 `http://columbia.us9.list-manage.com/track/click?u=3676dddb84cdd81111f6ba8ec&id=514c035f6e&e=d5755b524b`

- **Journals and Book Chapters**

  [J8] Faiza Khan Khattak and **Ansaf Salleb-Aouissi** A Tour Through Crowd-labeling. Submitted to XRDS The ACM Magazine for Students.

[J7] Faiza Khan Khattak and **Ansaf Salleb-Aouissi**
Toward Robust Crowd Labeling using Expert Evaluation and Pairwise Comparison
Under submission.

[J6] **Ansaf Salleb-Aouissi**, Christel Vrain, Cyril Nortet, Xiangrong Kong, Daniel Cassard
QuantMiner for Mining Quantitative Association Rules
Journal of Machine Learning Research Open source software. 14(Oct):3153-3157, 2013.
`https://github.com/QuantMiner/QuantMiner`

[J5] R. Trepos, **Ansaf Salleb-Aouissi**, M-O. Cordier, V. Masson and c. Gascuel
Building actions from classification rules
Knowledge and Information Systems: Volume 34, Issue 2 (2013), Page 267-298.

[J4] Rebecca J. Passonneau, Vikas Bhardwaj, **Ansaf Salleb-Aouissi** and Nancy Ide
Multiplicity and Word Sense: Evaluating and Learning from Multiply Labeled Word Sense Annotations. Language Resources and Evaluation 46(2): 219-252 (2012).

[J3] Cynthia Rudin, David Waltz, Roger Anderson, Albert Boulanger, **Ansaf Salleb-Aouissi**, Maggie Chow, Haimonti Dutta, Philip Gross, Bert Huang, Steve Ierome
Machine Learning for the New York City Power Grid," in the IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 2, pp. 328-345, Feb. 2012

[J2] **Ansaf Salleb-Aouissi**, C. Vrain
A Contribution to the Use of Decision Diagrams for Loading and Mining Transaction Databases
Fundamenta Informaticae Journal 2007. Volume 78, Number 4, pp. 575-594. IOS Press.

[J1] B. Duval, **Ansaf Salleb-Aouissi**, C. Vrain
On the Discovery of Exception Rules: A Survey. Reviewed **Book chapter** in "Quality Measures in Data Mining" book. In F. Guillet and Howard J. Hamilton editors Springer's Lecture Notes in Artificial Intelligence, Volume 43/2007 pp. 77-98. Springer January 2007.

- **Conferences**

[C30] Ilia Vovsha, **Ansaf Salleb-Aouissi**, Anita Raja, Axinia Radeva, Ashwath Rajan, Alex Rybchuk, Thomas Koch, Yiwen Huang, Hatim Diab, Ashish Tomar, and Ronald Wapner
Using Kernel Methods and Model Selection for Prediction of Preterm Birth
To appear in the Machine Learning for Healthcare 2016 JMLR conference track proceedings. August 2016.

[C29] Faiza Khan Khattak, **Ansaf Salleb-Aouissi** and Anita Raja
Accurate Crowd-labeling using Item Response Theory.
Accepted. To appear in Collective Intelligence Conference 2016

[C28] Antonio Moretti, Kathy McKnight and **Ansaf Salleb-Aouissi**
Application of Sentiment and Topic Analysis to Teacher Evaluation Policy in the U.S.. Educational Data Mining (EDM) conference 2015.

[C27] **Ansaf Salleb-Aouissi**, Christel Vrain and Daniel Cassard
Learning Characteristic Rules in Geographic Information Systems. The 9th International Web Rule Symposium (RuleML) 2015.

[C26] Ilia Vovsha , Ashwath Rajan , **Ansaf Salleb-Aouissi**, Anita Raja , Axinia Radeva, Hatim Diab, Ashish Tomar and Ronald Wapner Predicting preterm birth is not elusive: machine learning paves the way to individual wellness. *AAAI Spring Symposium – Big Data Becomes Personal: Knowledge into Meaning - For Better Health, Wellness and Well-Being.*, 2014. Paper: `http://www1.ccls.columbia.edu/~ansaf/CING/AAAI2014.pdf`

[C25] Faiza Khan Khattak and **Ansaf Salleb-Aouissi**
Robust Crowd Labeling using Little Expertise
Proceedings of the Sixteenth International Conference on Discovery Science DS 2013, LNAI 8140, pp. 94-109, 2013.

[C24] Anita Raja, Mohammad Hasan, Shalini Rajanna, **Ansaf Salleb-Aouissi**
A Scalable Approach to Modeling Risk in the MDAP Network Proceedings of Naval Postgraduate Schools 10th Annual Acquisition Research Symposium, pp 293-318, Monterey, CA.

[C23] Faiza Khan Khattak and **Ansaf Salleb-Aouissi**. Improving Crowd Labeling through Expert Evaluation. In the 2012 AAAI Spring Symposium Series SS-12-06 Wisdom of the Crowd.

[C22] Cynthia Rudin, Benjamin Letham, **Ansaf Salleb-Aouissi**, Eugene Kogan, David Madigan
Sequential Event Prediction with Association Rules. Journal of Machine Learning Research - Proceedings Track 19: 615-634 (2011)

[C21] Haimonti Dutta, David L. Waltz, Karthik M. Ramasamy, Philip Gross, **Ansaf Salleb-Aouissi**, Hatim Diab, Manoj Pooleery, Catherine A. Schevon, Ronald Emerson
Patient-Specific Seizure Detection from Intra-cranial EEG Using High Dimensional Clustering.
In ICMLA 2010: 782-787

[C20] Rebecca J. Passonneau, **Ansaf Salleb-Aoussi**, Vikas Bhardwaj and Nancy Ide
Word Sense-Annotation of Polysemous Words by Multiple Annotators
In the Seventh International Conference on Language Resources and Evaluation **LREC 2010**.

[C19] B. Huang and **Ansaf Salleb-Aouissi**
Maximum Entropy Density Estimation with Incomplete Presence-Only Data
In the 12th International Conference on AI and Statistics **AISTATS 2009**. JMLR Proceedings Volume 5, pages 240-247.

[C18] **Ansaf Salleb-Aouissi**, Bert Huang, David Waltz
Discovering Characterization Rules from Rankings
In the IEEE proceedings of the International Conference on Machine Learning and Applications **ICMLA 2009**, pages 154-161, December 2009.

[C17] Bert Huang, **Ansaf Salleb-Aouissi**, Phil Gross
Alive on Back-feed Culprit Identification via Machine Learning
In the IEEE proceedings of the International Conference on Machine Learning and Applications **ICMLA 2009**, pages 359-365, December 2009.

[C16] Phil Gross, **Ansaf Salleb-Aouissi**, Haimonti Dutta and Albert Boulanger
Ranking Electrical Feeders of the New York Power Grid
In the IEEE proceedings of the International Conference on Machine Learning and Applications **ICMLA 2009**, pages 725-730, December 2009.

[C15] Haimonti Dutta, David Waltz, Alessandro Moschitti, Daniele Pighin, Philip Gross, Claire Monteleoni, **Ansaf Salleb-Aouissi**, Albert Boulanger, Manoj Pooleery and Roger Anderson
Estimating the Time Between Failures of Electrical Feeders in the New York Power Grid
In Next Generation Data Mining Summit, : Dealing with the Energy Crisis, Greenhouse Emissions, and Transportation Challenges **NGDM 2009**.

[C14] **Ansaf Salleb-Aouissi**, Bert C. Huang and David L. Waltz
Vers des Machines a Vecteurs de Support "Actionables": Une Approche Fondee sur le Classement
Toward Actionable Support Vector Machines: A Ranking-based Approach
Best Paper Award in *Knowledge Extraction and Management (Extraction et Gestion des Connaissances)* **EGC 2008**, *pp. 285-295* Sophia Antipolis, France.

[C13] **Ansaf Salleb-Aouissi**, C. Vrain and C. Nortet
QuantMiner: A Genetic Algorithm for Mining Quantitative Association Rules
In the Proceedings of the *20th International Conference on Artificial Intelligence* **IJCAI 2007**, *pp. 1035-1040. Hyberadad, India.*

[C12] R. Trepos, **Ansaf Salleb**, M.-O. Cordier, V. Masson, C. Gascuel
Une approche fondee sur une distance pour la recommandation d'actions.
In *15e congres francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle (15th French Speaking Conference on Pattern Recognition and Artificial Intelligence)* **RFIA 2006** *Tours, France.*

[C11] **Ansaf Salleb** and C.Vrain

Estimation of the Density of Datasets with Decision Diagrams

In Proceedings of 15th International Symposium on Methodologies for Intelligent Systems **ISMIS 2005**. In M.-S. Hacid, N. V. Murray, Z. W. Ras, editors, Springer's Lecture Notes in Artificial Intelligence LNAI 3488, pp. 688-697, 2005.

[C10] R. Trepos, **Ansaf Salleb**, M.-O. Cordier, V. Masson, C. Gascuel

A Distance-Based Approach for Action Recommendation.

In Proceedings of the *9th European Conference on Machine Learning* **ECML 2005** *Porto portugal* Springer's Lecture Notes in Artificial Intelligence LNAI 3720, pp. 425-436, 2005.

[C9] C. Nortet, **Ansaf Salleb**, T. Turmeaux and C. Vrain

Extraction de Règles d'Association Quantitatives Application à des Données Médicales.

Proceedings of the *5th French speaking conference on Knowledge Discovery and Knowledge Management* **EGC 2005**. Revue des Nouvelles Technologies de l'Information RNTI, Ed. Cépaduès, Volume II, Pages 495-506. January 2005, Paris France.

[C8] Z. Maazouzi **Ansaf Salleb** and C. Vrain

BoolLoader : un chargeur efficace dédié aux bases denses.

In the *4th French speaking conference on Knowledge Discovery and Knowledge Management* **EGC 2004**. Revue des Nouvelles Technologies de l'Information RNTI, pp. 71–81, Ed. Cépaduès, January 2004, Clermont-Ferrand, France.

[C7] T. Turmeaux, **Ansaf Salleb**, C. Vrain and D. Cassard

Learning Characteristic Rules Relying on Quantified Paths

In the *7th European Conference on Principles and Practice of Knowledge Discovery in Databases,* **PKDD 2003**, Lavrač et al. (Eds.) LNAI 2838, pp. 471-482, 2003. Springer-Verlag

[C6] T. Turmeaux, D. Cassard, **Ansaf Salleb** and C. Vrain

Apprentissage de Règles Caractéristiques

In the *3rd French speaking conference on Knowledge Discovery and Knowledge Management* **EGC 2003** Hermès science publications, RSTI - RIA revue, volume 17, pages 437-448, January 2003, Lyon, France.

[C5] **Ansaf Salleb**, Z. Maazouzi and C. Vrain.

Mining Maximal Frequent Itemsets by a Boolean Approach.

In the *15th European Conference on Artificial Intelligence*, **ECAI 2002**, F. van Harmelen (Ed.) IOS Press Amsterdam, 2002, pages 385-389, Lyon, France, July 2002.

[C4] **Ansaf Salleb**, Z. Maazouzi

Approche Booléenne pour l'Extraction des Itemsets Fréquents Maximaux

In the *French speaking conference on machine learning* **CAp 2002**, PUG Press, pages 111-122, June 2002, Orléans, France.

[C3] A. Lips, D. Cassard, V. Bouchot, M. Billa, **Ansaf Salleb**, M. Gonzalez, L. Bougrain, C. Vrain, B. Tourlière, G. Stein and F. Alexandre.

Quantitative Assessments of a Continent-Scale Metallogenic GIS by Data-Driven and Knowledge-Driven Approaches to Construct Decision-Aid Documents.

Extended abstract **GIS in Geology 2002**, Moscou, November, 2002.

[C2] **Ansaf Salleb**, Z. Maazouzi and C. Vrain

Search Space Splitting for Mining Maximal Frequent Itemsets

In the International Conference on Information and Knowledge Engineering **IKE'02**, Las Vegas, USA, CSREA Press ed., June, 2002, pages 482-488

[C1] **Ansaf Salleb** and C. Vrain

An Application of Association Rule Discovery to Geographic Information Systems

In the *4th European Conference on Principles and Practice of Knowledge Discovery in Databases,* **PKDD 2000**, D.A. Zighed et al. (Eds.) LNAI 1910, pp. 613-618. Springer-Verlag ed.

- **Workshops**

[W11] Antonio Moretti, Jose Gonzalez-Brenes, Katherine McKnight and **Ansaf Salleb-Aouissi**. Mining Student Ratings and Course Contents for Computer Science Curriculum Decisions. *Data Mining for Educational Assessment and Feedback (ASSESS 2014), KDD 2014 workshop.*, New York 2014.

[W10] Faiza Khan Khattak and **Ansaf Salleb-Aouissi**. Quality Control of Crowd Labeling through Expert Evaluation. *Nips 2nd Workshop on Computational Social Science and the Wisdom of Crowds, Granada, Spain.*, 2011.

[W9] **Ansaf Salleb-Aouissi**, Axinia Radeva, Rebecca Passonneau, Boyi Xie, Faiza Khan Khattak, Ashish Tomar, David Waltz, Mary McCord, Harriet McGurk, and Noémie Elhadad.
Diving into a Large Corpus of Pediatric Notes.
ICML 2011 Learning from Unstructured Clinical Text Workshop Bellevue, WA, USA, July 2, 2011.

[W8] Leon Wu, Gail Kaiser, Cynthia Rudin, David Waltz, Roger Anderson, Albert Boulanger, **Ansaf Salleb-Aouissi**, Haimonti Dutta, and Manoj Poolery
Evaluating Machine Learning for Improving Power Grid Reliability
In the International Conference on Machine Learning ICML 2011 workshop on Machine Learning for Global Challenges.

[W7] **Ansaf Salleb-Aouissi**, Axinia Radeva, Rebecca Passonneau, Ashish Tomar, David Waltz, Mary McCord, Harriet McGurk, Noémie Elhadad
A Perspective on Understanding Infantile Colic
In NIPS 2010 Workshop on Learning and Planning from Batch Time Series Data, Whistler, British Columbia, Canada.

[W6] Vikas Bhardwaj, Rebecca Passonneau, **Ansaf Salleb-Aouissi**, Nancy Ide
Anveshan: A Framework for Analysis of Multiple Annotators' Labeling Behavior
In the fourth linguistic annotation workshop LAW IV, held in conjunction with ACL 2010 Upsala, Sweden.

[W5] Haimonti Dutta, David Waltz, **Ansaf Salleb-Aouissi**, Catherine Schevon and Ronald Emerson
Designing Patient-Speciþc Seizure Detectors From Multiple Frequency Bands of Intra-cranial EEG Using Support Vector Machines", DMHM, Workshop on Data Mining for HealthCare Management held in conjunction with PAKDD, 2010, Hyderabad, India.

[W4] B. Passoneau, **Ansaf Salleb-Aouissi**, and N. Ide
Making Sense of Word Sense Variation
In **SEW-2009** - Semantic Evaluations: Recent Achievements and Future Direction. A NAACL HLT 2009 Workshop, pages 2-9, June 4, 2009, Boulder, Colorado, USA.

[W3] C. Nortet, **Ansaf Salleb**, T. Turmeaux and C. Vrain
Le role de l'utilisateur dans un processus d'extraction de règles d'association.
Proceedings of the *French speaking workshop on Data and Knowledge Quality* **DKQ 2005** *co-located with EGC 2005.* pages 35-42. January 2005, Paris France.

[W2] **Ansaf Salleb**, T. Turmeaux, C. Vrain and C. Nortet
Mining Quantitative Association Rules in a Atherosclerosis Dataset.
PKDD Discovery Challenge 2004 (co-located with the *6th European Conference on Principles and Practice of Knowledge Discovery in Database*) Pisa, Italy, pp. 98-103, Sept. 2004.

[W1] D. Cassard, M. Billa, V. Bouchot, **Ansaf Salleb**, G. Stein and B. Tourliere
Predictive Mapping with GIS Andes Datasets.
Abstract and poster in the *European Union of Geosciences, Symposium FMF5* **EUG XI**, Cambridge publications ISSN 1362-0886, Strasbourg France, April 2001.

- **December 2015 - December 2018**
  "Deep Content Classification" Funded by SAP SE – Germany (Role: Principal Investigator).

- **April 2015 - June 2016**
  "Optimal Path to Knowledge" Funded by Pearson Education, Center for Educator Learning & Effectiveness (Role: Principal Investigator).

- **September 2014 - August 2016**
  "EAGER: Collaborative Research: Advanced Machine Learning for Prediction of Preterm Birth" Funded by the National Science Foundation (Role: Principal Investigator). `http://www.nsf.gov/awardsearch/showAward?AWD_ID=1454855`

- **September 2014 - June 2015**
  "Media Portrayal of Teachers" Funded by Pearson Education, Center for Educator Learning & Effectiveness (Role: Principal Investigator).

- **June 10, 2014**
  US patent 8,751,421 "Machine Learning for the power grid". Role: co-inventor.

- **June - September 2014**
  "Reinforcement learning to develop an adaptive stochastic controler for optimizing energy consumption in buildings." PI: Roger Anderson. Role: Research Scientist.

- **July 2013 - September 2013**
  National Science Foundation. RI: Medium: Collaborative Research: From Text to Pictures (Role: Research Scientist; PIs: Owen Rambow & Julia Hirschberg).

- **January 2013 - December 2013**
  Naval Postgraduate School grant "Leveraging Structural Characteristics of Interdependent Networks to Model Non-linear Cascading Risks" in response to NPS-BAA-12-002. (Role: Consultant PI: Anita Raja UNC Charlotte).

- **2011- 2013**
  Principal Investigator of "Understanding Baby Colic via Machine Learning" project involving CCLS and Columbia University Medical School. This project funded starting April 2011 by a Research Initiatives in Science and Engineering (RISE) funding from the Columbia University Executive Vice President for Research, (Role: Principal Investigator).

- **2009-2010**
  Columbia University, Alliance Grant for Faculty Joint Project travel grant. "Supervised Ranking: A Distance-based Approach" PI Ansaf Salleb-Aouissi (Columbia University) and Frank Nielsen (Ecole Polytechnique, Paris). `http://www.columbia.edu/cu/alliance/research.html` (Role: Principal Investigator).

- **2010- 2012**
  Member of the Smart Grid project, funded by the US Department of Energy (Role: Research Scientist; PI: Roger Anderson).

- **2006-2010**
  Member of the *CALM project* involving Consolidated Edison of New York and CCLS (Columbia University). This collaboration aims at developing and applying advanced Artificial Intelligence and Machine Learning methods toward a modern grid technology to improve public safety and system reliability. (Role: Research Scientist; PI: Roger Anderson)

- **2008-2010**
  Member of the *Epilepsy project* involving CCLS and the Medical school (Columbia University). This collaboration aims at developing an "Early Warning" Device to Allow Epilepsy Patients to Live a More Normal Life. This project was funded by Research Initiatives in Science and Engineering (RISE) at Columbia University. (Role: Research Scientist; PI: David Waltz)

- **2010**
  Member of the *Epilepsy project* involving CCLS and Columbia University Medical School. This project was funded by the Epilepsy Foundation (Role: Research Scientist; PI: Catherine Schevon).

- **2004-2005**
  Member of the *SACADEAU project* involving IRISA-INRIA and INRA (French Institute for Agronomy Research) in France. This project aims at "Building a Knowledge Acquisition System for Decision-Aid to Improve Stream-water Quality". (Role: Postdoc)

- **2003-2004**
  Member of the project "French-Russian Research Hub for the Search and Discovery of Super-Large Metallic Deposits". It is a collaboration between the French Geological Survey BRGM and the Russian Academy of Sciences. This collaboration was supported by a NATO Science Program. (Role: PhD student)

- **2000-2004**
  During my Masters and my Ph.D. thesis on "Data Mining in Geographic Information Systems (GIS): Application to the Andes Cordillera Mineral Deposits" I was a member of a collaboration between the University of Orleans in France and the *French Geological Survey (*BRGM*) Service of Mineral Resources.* (Role: PhD student)

PROFESSIONAL ACTIVITIES

Administrative activities, fairs, conferences, and meetings:

- Starting May 2016: Associate Editor of the Springer Journal Signal, Image and Video Processing.

- Local arrangement committee member for the International Joint Conference for Artificial Intelligence IJCAI 2016 New York City. Responsibility Website and program committee member. http://ijcai-16.org/

- Data Science Education Committee. Role: participant. 2016.

- Administrative responsibilities at CCLS include creating and running the CCLS technical report committee since 2007 and CCLS internal evaluation of PhD students every semester since Fall 2014.

- Member of the research group GAfoDonnees at CNRS, France (2001-2005).

- Member of the working group on "Data Mining and Quality Measures: GaFoQualité", sponsored by STIC-CNRS, Nantes, France. 2002.

- Member of the organization team for the French speaking conference on machine learning CAp 2002, Orléans France.

- Member of the working group on the "PKDD Discovery Challenge 2004" sponsored by STIC-CNRS. Lyon and Paris, France.

- Member of the Columbia University Research Officers committee 2006-2007.

- Organizer of the CCLS journal Club 2008.

- Co-organizer of the CCLS retreat 2008. October 22, 2008 at the Wave Hill Center, New York.

- Initiation of children to counting in the binary system at "Initiation to science", University of Orleans, France (2003).

- 2011- 2012 Introducing kids (3-5 years old) to Science (four hours): Christ the King Daycare Middle Village NY.

PC Membership and Journal Reviewing:

- PC member of the 2015 International Joint Conference on Neural Networks **IJCNN 2017**.

- PC Member the International Joint Conference for Artificial Intelligence **IJCAI 2016** conference

- PC Member the **ECML-PKDD 2016 and 2015** conferences.

- PC member of the 2015 International Joint Conference on Neural Networks **IJCNN 2015**.

- PC Member the **ECML-PKDD 2014** conference

- Reviewer for the Transactions on Knowledge and Data Engineering journal **TKDE 2014**.

- PC Member the fourteen SIAM International Conference on Data Mining, **SDM'14**.

- NSF Division of Information and Intelligent Systems (**IIS**), 2013.

- PC member of the 2013 International Joint Conference on Neural Networks **IJCNN 2013**.

- PC member of the Knowledge Extraction and Management (Extraction et Gestion des Connaissances), **EGC 2010-2015**.

- PC member of the 2012 International Joint Conference on Neural Networks **IJCNN 2012** special session "Unsupervised model-based learning from high dimensional and functional data".

- Reviewer for Data Mining and Knowledge Discovery Journal **DMKD** 2009, 2012.

- PC member of the European Conference on "Machine Learning" and "Principles and Practice of Knowledge Discovery in Databases" **ECML PKDD 2011**.

- PC member of the International Conference on Knowledge Discovery and Information Retrieval **KDIR 2011**.

- PC member of the International Conference on Data Engineering, **ICDE 2011**.

- PC member of the 1st Maghrebian Conference on the Extraction and management of Knowledge **EGC-Maghreb 2010**.

- "Advances in Knowledge Discovery and Management", a refereed book of chapters, published by Springer, **AKDM 2009-2012**.

- NSF Division of Information and Intelligent Systems (IIS), 2009.

- The 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining **KDD 2009**,

- European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases **ECML/PKDD 2008**,

- Quality Measures for Data Mining Book, Springer-Verlag **QMDM2006**,

- Data and Knowledge Quality Workshops **DKQ2005, 2006**,

- Qualité des Données et des Connaissances workshops co-located with EGC **QDC2007-2010**.