

## Curriculum Vitae for Julien Rabatel, PhD.

---

### CONTACT INFORMATION

GREYC - CNRS UMR 6072  
Université de Caen  
Département informatique  
Campus Côte de Nacre  
F-14032 Caen Cedex

*Voice:* +33(0) 231 56 73 93  
*Mobile:* +33(0) 677 51 95 34  
*Fax:* +33(0) 231 56 73 30  
*E-mail:* julien.rabatel@unicaen.fr  
*WWW:* <http://www.lirmm.fr/rabatel>

RESEARCH TOPICS PHD THESIS TITLE: Context-Aware Pattern Mining: Challenges and Applications in Sequential Data.

SUPERVISORS: Pascal Poncelet, Pr.; Sandra Bringay, PhD.

SUMMARY: In many application domains, such as medical diagnosis or equipment monitoring, data can be associated with contextual information describing the circumstances over which data have been collected. Taking into account this information can be of great interest for decision makers. For instance, the age or the gender of a patient can impact the diagnosis made by a medical expert. This thesis investigates the mining of sequential data in order to (1) provide interesting patterns by considering contextual information, and (2) exploit such patterns for other tasks such as classification, prediction or anomaly detection. The first part of this thesis aims at considering contextual information associated with data during the frequent pattern mining process, in order to provide the expert with patterns that are representative of a context. Existing work prior to this thesis could not reveal that some patterns strongly depend on context. We thus provide the notion of contextual frequent pattern, where a pattern is associated with a context. In addition, we generalize the notion of contextual pattern to various interestingness measures (other than frequency): information gain, growth rate, etc. In both cases, we unveil and exploit some essential theoretical properties of contextual patterns and provide efficient algorithms. The second part of this work concerns the use of contextual patterns to address various data mining tasks. We mainly focus here on sequential data in order to perform pattern-based classification, prediction and anomaly detection. Being able to consider contextual information is here of great interest. For instance, contextual patterns can highlight the fact that a behavior that is considered as anomalous in *summer* can be considered as normal in *winter*. All our approaches have been experimented on real-world datasets coming from various application domains such as sensor numerical data, textual data or health data and have been showed to be efficient in practical applications.

### RESEARCH INTERESTS:

- Frequent pattern mining
- Context-aware patterns
- Pattern-based applications (classification, prediction, anomaly detection)
- Sensor data mining
- Sequential data mining

**International Journal Articles**

*Mining Sequential Patterns: a Context-Aware Approach* J. Rabatel, S. Bringay and P. Poncelet. Advances in Knowledge Discovery and Management (AKDM), to appear in 2012. (extended version of an EGC 2011 paper)

*Anomaly detection in monitoring sensor data for preventive maintenance* J. Rabatel, S. Bringay and P. Poncelet. Expert Systems With Applications 38 (2011), pp. 7003-7015.

**National Journal Articles**

*Aide au diagnostic de pannes guidée par l'extraction de motifs séquentiels.* J. Rabatel, S. Bringay, P. Poncelet and M. Teisseire. Revue RNTI, Numéro Spécial Fouille de données complexes, Cépadués, 2010, pp. 87-112.

**International Refereed Conference and Workshop Papers**

*Contextual Sequential Pattern Mining.* J. Rabatel, S. Bringay and P. Poncelet. Domain Driven Data Mining Workshop (DDDM) in conjunction with IEEE ICDM 2010, Sydney, Australia. December, 2010, pp. 981-988.

*Fuzzy Anomaly Detection in Monitoring Sensor Data.* J. Rabatel, S. Bringay and P. Poncelet. 18th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2010), Barcelona, Spain. July, 2010, pp. 1-8.

*SOMAD: SensOr Mining for Anomaly Detection in Railway Data.* J. Rabatel, S. Bringay and P. Poncelet. In Proceedings of the 9th Industrial Conference on Data Mining (ICDM 2009), Leipzig, Germany. July, 2009, pp. 191-205. ***In nomination for the best paper award***

**National Refereed Conference and Workshop Papers**

*Extraction de motifs séquentiels contextuels.* J. Rabatel and S. Bringay. Actes des 11ièmes Journées Francophones Extraction et Gestion des Connaissances (EGC 2011), Brest, France. January, 2011, pp. 11-22. ***In nomination for the best academic paper award***

*Prédiction du grade d'un cancer du sein par la découverte de motifs séquentiels contextuels dans des puces à ADN* J. Rabatel, M. Fabrègue, S. Bringay, P. Poncelet and M. Teisseire. Atelier Extraction de Connaissances et Santé, in conjunction with the 11ièmes Journées Francophones Extraction et Gestion des Connaissances (EGC 2011), Brest, France. January, 2011, pp. 37-48.

*Aide à la décision pour la maintenance ferroviaire préventive.* J. Rabatel, S. Bringay and P. Poncelet. Actes des 10ièmes Journées Francophones Extraction et Gestion des Connaissances (EGC 2010), Hammamet, Tunisie. January, 2010, pp. 363-368.

*Visualisation des motifs séquentiels extraits à partir d'un corpus en Ancien Français.* J. Rabatel, Y. Lin, Y. Pitarch, H. Saneifar, C. Serp, M. Roche, A. Laurent. Actes des 8ièmes Journées Francophones Extraction et Gestion des Connaissances (EGC 2008), Nice, France. January, 2008, pp. 237-238.

**Other communications**

*Sensor Data Mining* in OSIDEAL: Workshop on statistical and computational tools for Ph.D students in Environmental studies and Agronomics, Montpellier, France. April, 2010.

*Main Challenges in Data Mining* in SeminDoc: Seminar of PhD Students at LIRMM, Montpellier, France. January, 2009.

## EDUCATION

**University of Montpellier 2**, Montpellier, France

Part-time Ph.D. Student, Computer Science, September 2008 - September 2011

- Thesis Topic: “*Context-Aware Pattern Mining: Challenges and Applications in Sequential Data*”
- Supervisors: Pascal Poncelet, Pr.; Sandra Bringay, PhD.

**University of Montpellier 2**, Montpellier, France

M.Sc., Computer Science, 2008

- Thesis Topic: “*Frequent pattern mining in railway monitoring data*”
- Supervisors: Pascal Poncelet, Pr.; Sandra Bringay, PhD; Maguelonne Teisseire, Senior Researcher.

PROFESSIONAL  
EXPERIENCE**University of Caen Basse-Normandie**, Caen, France - **University of Leuven**, Leuven, Belgium

*Post-doctoral fellow*

**November, 2011 - August, 2012**

“*Mining patterns in 3D molecular graphs in order to describe the affinity between a molecule and a protein*”

My mission within this project aims at extracting interesting subgraphs in a dataset of 3D molecular graphs, in order to describe the interaction phenomenon between a ligand (e.g., a potential medicine) and a protein. This mission includes an important multidisciplinary aspect by integrating the knowledge and the requirements of the domain experts (chemists). Currently in Caen for the first part of this project, I will work in Leuven for the second part from April to August 2012. The main partners of the project are the following:

**CoDaG Group, GREYC<sup>1</sup>, University of Caen Basse-Normandie**

- Pr. Bruno Crémilleux
- Dr. Bertrand Cuissart

**Machine Learning Group, Katholieke Universiteit Leuven, Belgium**

- Pr. Jan Ramon

**CERMN<sup>2</sup>, University of Caen Basse-Normandie**

- Pr. Ronan Bureau

**Tecnalia - France**, Montpellier, France

*Part-time Research Engineer*

**September, 2008 - October, 2011**

My responsibilities included conducting practically applied research, software implementation of the research and technology demonstrators, reports writing, and providing advice and guidance to different research related issues.

TEACHING  
EXPERIENCE**University of Caen Basse-Normandie**, Caen, France

*Teaching Assistant*

**February, 2012 - March, 2012**

Performed lecturing and tutoring duties for Master students (12h).

<sup>1</sup><https://www.greyc.fr>

<sup>2</sup>Centre d'Etudes et de Recherche sur le Médicament de Normandie

- Designed, lectured and assessed a *Text mining* course.

**University of Montpellier 3**, Montpellier, France

*Teaching Assistant*

**September, 2008 - October, 2008**

Performed tutoring duties for C2I undergraduate classes (24h).

- Demonstrated and assessed a Word, Excel, Web Programming course.

*Teaching Assistant*

**January, 2009 - May, 2009**

Performed tutoring duties for C2I undergraduate classes (22h).

- Demonstrated and assessed a Word, Excel, Web Programming course.

**CNAM**, Montpellier, France

*Teaching Assistant*

**November, 2009 - February, 2010**

Performed lecturing duties for CNAM (Life Long Learning) classes (30h).

- Designed, lectured and assessed a Software Engineering course.

PROFESSIONAL  
ACTIVITIES

**Organization committee**

*5th French Speaking Conference on Data Warehousing and Online analysis* (EDA 2009), Montpellier, France. June, 2009.

**Reviewer**

*21st International Conference on Database and Expert Systems Applications* (DEXA 2010)

*IEEE International Conference on Fuzzy Systems* (Fuzz-IEEE 2010)

*International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems* (IPMU 2010)

*IEEE International Conference on Data Mining* (ICDM 2009)

*Asia-Pacific Web Conference* (APWeb 2009)

*Scalable Fuzzy Algorithms for Data Management and Analysis* (Book chapter)

KEY SKILLS

**Computer Skills**

- *Languages*: Java, C/C++, Perl, PL/SQL.
- *Web Programming*: PHP, (X)HTML, Javascript, CSS.
- *Databases (DBMS)*: Oracle, PostgreSQL, MySQL.
- *Operating Systems*: Linux, Windows.

**Languages**

- French: mother tongue
- English: fluent
- Spanish: good notions