

Laurent Candillier
Français né en 1978
lcandillier@hotmail.com
(+34)615707752



1 Présentation rapide

- Ingénieur en Recherche & Développement Informatique - Intelligence Machine
- Docteur en Informatique, spécialité Machine Learning, Data Mining et Systèmes de Recommandations
- Connaissances avancées aussi en Recherche d'Information et Traitement Automatique des Langues
- Technologies informatiques : Python, SQL, Java, Apache Spark, Beam, GCP, Elastic-Search, Scala, PHP...
- Trilingue Français, Espagnol, Anglais

2 Cursus général

- **2017-2024** :
Senior Data Scientist, Ocado Technology, Barcelone, Espagne
- **2013-2016** :
Expert scientifique et technique en R&D informatique, TokTokTok, Digimind, Target2Sell, TheFamily, Charly.io, SonetIN, Refactor, France Consultants, Overblog, Ebuzzing - Teads
- **2010-2012** :
Directeur R&D informatique, Nomao - Overblog - Ebuzzing - Teads, Toulouse
- **2006-2009** :
Ingénieur R&D - post-doctorat (systèmes de recommandations), Orange Labs, Lannion
- **2003-2006** :
Ingénieur R&D - doctorat (clustering), Université de Lille - Société Pertinence, Paris
- **2001-2002** :
Ingénieur R&D - doctorat (profilng), Université de Lille - Société Rosebud, Paris
- **1996-2001** :
Master informatique, Université de Lille - Aston University, Birmingham, Angleterre

3 Travaux de recherche

La passion des mathématiques m'a valu une mention bien en Terminale Scientifique. Mais le côté ludique de l'informatique a pris le pas à l'université de Lille. La découverte du **Machine Learning** en stage de Master au sein du laboratoire GRAppA a fini d'établir ma vocation : je veux que l'humain et la machine collaborent pour gagner en intelligence mutuelle. En 2001

j'ai démarré une thèse sur le sujet des **Systèmes de Recommandations**. Le filtrage collaboratif avait récemment émergé et nous avons montré sa complémentarité avec les approches de profilage basées sur le contenu. Malheureusement, cette recherche a été avortée suite à des problèmes internes à la société Rosebud, et nous nous sommes redirigés vers une autre société, Pertinence, et un autre domaine, celui du **Clustering**. Notre axe de recherche consistait à trouver une solution qui soit compréhensible par l'humain, mais notre contribution majeure s'est située au niveau de la problématique ouverte importante de l'évaluation des méthodes de clustering (publication n°1). Après la soutenance de ma thèse, je suis retourné vers le domaine des systèmes de recommandations au sein d'Orange Labs. Nous avons participé au concours Netflix, pas gagné le million de dollars mais obtenu de très bons résultats avec notre algorithme de **Filtrage Collaboratif** optimisé par la pondération des mesures de similarité classiquement utilisées dans le domaine (publication n°2). Notre solution offrait également l'avantage d'être compréhensible par les utilisateurs. Après ce post-doctorat j'ai été nommé responsable de la Recherche & Développement dans le groupe Nomao - Overblog - Ebuzzing. Outre mes travaux en analyses des sentiments, j'y ai co-encadré deux thèses, l'une chez Nomao sur le thème du learning-to-rank appliqué aux moteurs de recherches géolocalisés et personnalisés, et l'autre chez Overblog sur l'injection de diversité dans les systèmes de recommandations (publication n°3). Nos travaux en **Apprentissage Actif** appliqué au problème de la déduplication de lieux chez Nomao nous ont également amenés à organiser un atelier et challenge au sein de la conférence internationale ECML-PKDD'2012 (publication n°4). En tant qu'expert scientifique et technique agréé par le Ministère de l'Enseignement Supérieur et de la Recherche, j'ai ensuite accompagné les Jeunes Entreprises Innovantes dans l'élaboration de leurs projets de R&D et l'étude et développement de leurs algorithmes intelligents. Dans le cadre de la mise en place du moteur de recherche de TokTokTok, nous avons publié l'article n°5 qui traite de l'interaction entre les systèmes de **Recherche d'Information** et le Traitement Automatique des Langues. Je travaille maintenant chez Ocado Technology sur tout ce qui touche à la prévision de la demande en environnement **Big Data** (présentation n°6). Davantage de détails sur mes recherches et publications sont accessibles en ligne sur <http://lcandillier.free.fr>

4 Publications (extrait 6/28)

1. **[2006] Cascade Evaluation of Clustering Algorithms**
 Laurent Candillier, Isabelle Tellier, Fabien Torre, Olivier Bousquet
 In Johannes Fürnkranz, Tobias Scheffer and Myra Spiliopoulou, editors
17th European Conference on Machine Learning
 ECML'2006, Berlin, Germany, 18-22 september 2006
 Lecture Notes in Computer Science, LNAI 4212, pages 574-581

2. **[2008] Designing Specific Weighted Similarity Measures to Improve Collaborative Filtering Systems**
 Laurent Candillier, Frank Meyer, Françoise Fessant
 In Petra Perner, editor
8th Industrial Conference on Data Mining
 ICDM'2008, Leipzig, Germany, 16-18 july 2008
 Lecture Notes in Computer Science, LNAI 5077, pages 242-255

3. **[2012] Multiple similarities for diversity in recommender systems**
 Laurent Candillier, Max Chevalier, Damien Dudognon, Josiane Mothe
 In *International Journal on Advances in Intelligent Systems*, Volume 5, Number 3 & 4
4. **[2012] Design and Analysis of the Nomao Challenge - Active Learning in the Real-World**
 Laurent Candillier, Vincent Lemaire
Workshop on Active Learning in Real-world Applications
 ECML-PKDD'2012, Bristol, UK, 28 september 2012
5. **[2016] RI-TAL : le TAL au service de la RI**
 Laurent Candillier, Julien Hénou
13ème Conférence en Recherche d'Information et Applications
 CORIA'2016, Toulouse, 9-11 march 2016
6. **[2018] Machine Learning with Scikit-Learn and Xgboost on Google Cloud Platform**
 Laurent Candillier, David Cournapeau, Steve Greenberg
Google Cloud Next'2018, San Francisco, 24-26 july 2018

5 Compétences en informatique

- Enseignement universitaire : programmation en Python, Ada, créations de sites web, usages d'internet, bases de données, tableurs
- Langues de programmation : Python, Java, Perl, C, C++
- Big Data : Google Cloud Platform, Big Query, Apache Spark, Beam, Scala
- Développement web : HTML, CSS, PHP, MySQL, PostgreSQL, ElasticSearch

6 Langues parlées

- Français, espagnol, anglais trilingue
- Des notions d'allemand

7 Loisirs

- Sports, danse, percussions, jonglerie, lecture, écriture, jeux d'échecs et de cartes, nature, culture, voyages, écologie, aide aux personnes