

Patrice Bellot

Professeur des Universités en Informatique (PR1)

Directeur adjoint du Laboratoire d'Informatique et des Systèmes LIS (UMR 7020)

Délégué scientifique « Fouille de données textuelles – Information scientifique » CNRS



Aix-Marseille Université — CNRS — Polytech Marseille

Domaine Univ. St Jérôme – Av. Escadrille Normandie Niemen 13397 Marseille Cedex 20

(+33/0)6 43 18 44 07 / patrice.bellot@univ-amu.fr / patrice.bellot@cnrs-dir.fr

ORCID : orcid.org/0000-0001-8698-5055

Domaines : Recherche d'information, Traitement automatique des Langues

Synthèse de la carrière

2018 – ... : Professeur des Universités 1^{ère} classe Aix-Marseille Université / Polytech

2017 – ... : Chargé de mission puis Délégué Scientifique « Information Scientifique et Technique et Fouille de données textuelles » — Délégation institutionnelle 50%, CNRS Sciences Informatiques

2011-2018 : Professeur des Universités 2^e classe, Aix-Marseille Université

2000-2011 : Maître de conférences à l'Université d'Avignon et des Pays de Vaucluse

1999-2000 : ATER à l'Université d'Avignon

1996-1999 : Moniteur de l'ESR – Université d'Avignon

Titulaire de la PEDR (2004 – 2012, 2014 – 2022) et RIPEC (2023 – ...)

Diplômes

HDR en Informatique (2008) – Université d'Avignon

Doctorat en Informatique (01.2000) – Université d'Avignon (allocation MESR)

DEA en Informatique et Ingénieur (1996) – Université Aix-Marseille II

Licence de Mathématiques (1993) — Université de Provence

Responsabilités pédagogiques

2002 – 2004 : Responsable DESS Traitement Automatique de l'Information, Université d'Avignon

2004 – 2010 : Responsable Master Informatique (spécialité puis mention), Université d'Avignon

2010 et 2011 : Directeur des Etudes IUP Génie Mathématiques et Informatique, Université d'Avignon (du L2 au M2)

2021– ... : Membre élu du Conseil d'Ecole de Polytech Marseille

2014 – ... : Membre élu du Conseil du département Génie Industriel et Informatique de Polytech Marseille

2006 – 2011 : Membre élu du CEVU de l'Université d'Avignon

2005 – 2009 : Membre élu du Conseil d'Administration de l'IUP GMI de l'Université d'Avignon

09.2020— 09.2022 : Responsable des stages Initiation à la Recherche de 4^e année du dép. GII

2018 – 2020 : co-porteur du projet de création de la filière en apprentissage « Systèmes Numériques » de Polytech Marseille (validation CTI en 2021, démarrage en 2022)

2019 – ... : Membre des conseils de perfectionnement du Master Humanités Numériques (Universités de Lyon) et du Master Informatique (Aix-Marseille Université)

Enseignements (depuis 2017)

Au sein d'Aix-Marseille Université :

- 90% de mon service annuel : Ingénieur Polytech : Analyse de données (3^e année), Ingénierie documentaire (5^e année), Recherche d'information (spécialités « Génie Industriel et Informatique » et « Systèmes numériques »)
- Master Informatique : Recherche d'information (2021)

- Master Psychologie : Ergonomie des Systèmes d'Information (2018-2021)
- Licence Lettres : Ecriture automatique (2018 et 2019)
- Formation continue : Traitement des données en Python (2021, 2022, 2023)

Extérieur AMU :

- Ecole Centrale Marseille : Fouille de données et traitement en Python – 18 h (2023, 2024, 2025)
- Université de Limoges : Séminaires Humanités Numériques (Master) – 12h (2024, 2025)
- Sorbonne Université : Fouille de données et de textes (DU Data Stratégie) (2019)
- CNRS : Panorama de la fouille de textes + atelier « Fouille de textes par apprentissage automatique » (Action Nationale de Formation ANF) (2020, 2021, 2022)
- Espagne, Université de Vigo : PhD Programme on Information and Communications Technology (2017)
- Tunisie, Université de la Manouba : « Analyse des réseaux sociaux et Recommandation » (2017)

Thèmes de recherche

Mes travaux se situent au croisement de deux domaines que sont la Recherche d'Information (RI) et le Traitement Automatique des Langues (TAL). Ils concernent la recherche et l'extraction d'information au sein de grandes collections documentaires (dont les bibliothèques numériques), les systèmes de questions-réponses, la simplification textuelle, la classification thématique et l'analyse des sentiments et des émotions.

Animation de la Recherche

Directeur-adjoint du LIS depuis janvier 2024

Responsabilité d'équipes :

2013 – 2018 : équipe *Data, Information & content Management Group* (DIMAG) (18 E/C, 17 doct.)

2018 – 2019 : équipe *Recherche d'Information et Interactions* (R2I) (12 E/C, 10 doct.)

Direction, comités et conseils :

2015 – 2018 : Directeur Adjoint de la FRIAM (*Fédération de Recherche en Informatique et Interactions d'Aix-Marseille*), FR CNRS

2011 – 2020 : Co-Directeur scientifique de l'OpenEdition Lab au sein de l'USR CLEO-OpenEdition.org

2014 – ... : Comité de pilotage de la TGIR *Huma-Num*, représentant le Président d'AMU

2019 – ... : Conseil scientifique du GDR TAL (*Traitement Automatique des Langues*) – CNRS

2011 – 2013 : co-porteur du thème "Information, Langue et Documents Numériques" au sein du GDR I3

2015 – 2018 : représentant de la France dans l'infrastructure européenne (ERIC) CLARIN (*Common Language Resources and Technology Infrastructure*)

2018 – ... : Conseil scientifique de la Fédération SFERE-Provence (Structure Fédérative d'Études et de Recherches en Éducation de Provence, FED 4238)

2018 – ... : Membre du comité Intelligence Artificielle du Pôle de compétitivité SCS

2018 – ... : CNRS : Représentant de l'INS2I au sein de la DIST puis de la DDOR (Direction des données ouvertes de la recherche) et de l'Atelier Opérationnel et Stratégique EOSC (European Open Science Cloud)

2016 – 2020 : Conseil de direction de l'Institut Convergence ILCB (*Language, Communication & the Brain*)

2017 – 2022 : Conseil scientifique de l'Agence Bibliographique de l'Enseignement Supérieur (ABES)

2015 : référent du champ scientifique "Information, Données, Communication" pour le Comité d'Orientation Stratégique (COS) d'AMU

Société savante :

2015 – 2023 : Vice-Président de l'ARIA (*Association Recherche d'Information et Applications*)

Direction de thèses

18 thèses de Doctorat en Informatique ont été soutenues sous ma direction dont 1 cotutelle (Tunisie) et 3 co-encadrements internationaux (Espagne, Algérie, Tunisie) et 1 allocation ministérielle (Doctorat Handicap — Univ. Toulon) .

4 thèses en cours en co-direction, dont 1 RAPID DGA, 1 contrat privé (Naval Group) , 1 contrat AMU, 1 thèse associée à une CPJ CNRS avec OpenEdition).

Projets collaboratifs (principaux depuis 2011)

2011 – 2014 : responsable LSIS projet ANR CAAS *Contextual Analysis and Adaptive Search*

2012 – 2015 : projet Région PACA en analyse de sentiments (porteur)
 2012 – 2020 : PIA Equipex DILOH (*Digital Library for Open Humanities*) : porteur pour OpenEdition et le LIS pour la partie recherche en informatique (3 thèses financées),
 2012 – 2016 : PIA « Technologies des contenus numériques » Intertextes : responsable pour le LIS
 2016 – 2018 : PIA ISTE.fr (inscrit dans la feuille de route nationale des infrastructures de recherche) : Président du comité technique (porté par le CNRS et Univ. Lorraine)
 2020 – ... : PIA AMPIRIC « Education et numérique » : représentant du LIS, membre du Conseil scientifique interne
 2020 – ... : PIA IDeAL « *Integration and Development at Aix-Marseille through Learning* », membre du Comité exécutif du CEntre de formation et de soutien aux Données de la REcherche CEDRE
 2022 – 2024 : PIA GraphoGameAdapt (entraînement à la lecture): responsable pour le LIS (porté par le LPC AMU-CNRS)
 2022 – 2025 : Rapid DGA RAFFAL (*Résumé abstraitif francophone fluide et adaptatif pour l'amélioration de la lisibilité*) : responsable LIS
 2022 – 2025 : projet européen FAIR-IMPACT (science ouverte et impact de la *fairisation* des données) : responsable LIS

Sélection de publications

- Liana Ermakova, Patrice Bellot, Jaap Kamps, Diana Nurbakova, Irina Ovchinnikova, Eric SanJuan, Élise Mathurin, Sílvia Araújo, Radia Hannachi, Stéphane Huet, Nicolas Poinssu: *Automatic Simplification of Scientific Texts: SimpleText Lab at CLEF-2022*. ECIR 2022: 364-373
- M. Chaa, O. Nouali, P. Bellot, “ *Leveraging App Features to Improve Mobile App Retrieval* ” IJIDS International Journal of Intelligent Information and Database Systems, 2021
- A. Htaït, S. Fournier, P. Bellot, L. Azzopardi, G. Pasi, « *Using Sentiment Analysis for Pseudo-Relevance Feedback in Social Book Search* », ACM SIGIR ICTIR International Conference on the Theory of Information Retrieval, September 14–17, 2020, Virtual Event, Norway. ACM, New York, NY, USA, 2020
- Gaël Guibon, Magalie Ochs, Patrice Bellot “*From Emoji Usage to Categorical Emoji Prediction*”, 19th International Conference on Computational Linguistics and Intelligent Text Processing CiCLING 2018, Hanoi (Vietnam), mars 2018. **Best software award**
- Amal Htaït, Sébastien Fournier, Patrice Bellot “*Unsupervised Creation of Normalization Dictionaries for Micro-Blogs in Arabic, French and English*”, 19th International Conference on Computational Linguistics and Intelligent Text Processing CiCLING 2018, Hanoi (Vietnam), mars 2018. **Best short paper award**
- Messaoud Chaa - Patrice Bellot - Omar Nouali "New Technique to Deal With Verbose Queries in Social Book Search", 2017 IEEE/WIC/ACM International Conference on Web Intelligence (WI) , pp. 799-806, Leipzig, Germany
- W. Bakari, P. Bellot, M. Neji “*A logical representation of Arabic questions toward automatic passage extraction from the Web*” , in : *International Journal of Speech Technology* (Kluwer Academic Publishers), Vol. June 2017, pp. 23
- P. Bellot – V. Moriceau – J. Mothe – E. San Juan - Xavier Tannier, "INEX Tweet Contextualization Task : Evaluation, Results and Lesson Learned" , in : *Information Processing and Management* (Elsevier), Vol. 2016, pp. 23 p., 2016.
- R. Deveaud, E. San Juan, P. Bellot, "Are Semantically Coherent Topic Models Useful for Ad Hoc Information Retrieval ?", 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013), Sofia, Bulgaria, August 2013.
- L. Bonnefoy, V. Bouvier, P. Bellot, "A weakly-supervised detection of entity central documents in a stream", The 36th Annual ACM SIGIR Conference SIGIR'13, Dublin (Ireland), July 2013.
- Romain Deveaud, Eric San Juan, Patrice Bellot, "Estimating Topical Context by Diverging from External Resources", The 36th Annual ACM SIGIR Conference SIGIR'13, Dublin (Ireland), July 2013.

DBLP : 176 publications référencées, dont 25 revues
 Google Scholar : 2649 citations, dont 1459 depuis 2018

Jurys de Doctorats et d'HDR

- HDR : 15 jurys depuis 2015, dont 8 en tant que rapporteur (et 1 en Sciences du Langage et 1 en Science de l'éducation)
- Doctorats : 65 jurys depuis 2009, dont 48 en tant que rapporteur (dont 2 en Suisse, 2 en Espagne, 1 aux Pays-Bas, 1 en Algérie, 1 en Norvège)

Comités de programme et comités éditoriaux

Président du comité de programme de la conférence CORIA 2019 (Recherche d'Information et Applications)
 Conference Chair de CLEF 2018 (*Conference and Labs of the Evaluation Forum*)

Président du comité d'organisation de la conférence CORIA 2011 (Recherche d'Information et Applications) (Université d'Avignon)

Membre des comités d'organisation des conférences :

- ECIR 2018 (*European Conference on Information Retrieval*) (Grenoble)
- ACM SIGIR 2019 (Paris)

Membre régulier des comités de programme des principales conférences en Recherche d'information et domaines connexes telles que : *ACM SIGIR, ACM CIKM, ACM WDSM, ECIR, Interspeech, IEEE ICASSP, IJCAI, IEEE-ACM Web Intelligence, ICDL*

Relecteur pour les **revues** : *ACM Trans. On Information Systems, ACM Transactions on Social Computing, ACM Transactions of Affective Computing, ACM Computing Surveys, Pattern Recognition Letters (Elsevier), Information Science (Elsevier), Int. Journal on Digital Libraries (IJDL), IEEE Access, Information Processing & Management (IPM), IEEE Transactions on Knowledge and Data Engineering...*

Membre du comité de rédaction permanent de la revue *Traitement Automatique des Langues* depuis 2009

Co-organisateur des campagnes internationales d'évaluation :

- "Tweet Contextualization" de la conférence Cross-Language Evaluation Forum (CLEF) & Initiative for XML Retrieval (INEX) 2012, 2013 et 2014
- "Question Answering" des conférences INEX 2009, 2010, 2011 (QA@INEX)
- "SimpleText" (*Automatic Simplification of Scientific Texts*) de la conférence CLEF (2020, 2021, 2022)

Expertises

- ANR : membre des comités d'évaluation des programmes *Contenus et Interactions* CONTINT (2013), *appel générique* (2014, 2015, 2016, 2017), *science ouverte* (2019), *recherche action COVID* (2020)
- ANRT, Régions : Grand-Est, Normandie, Ile-de-France, Normandie, Occitanie
- Suisse : Expertise pour la *Swiss National Science Foundation* (SNSF) : 1 expertise par an depuis 2017
- Canada : programme *Mitacs Acceleration* (annuel, depuis 2021)
- Royaume-Uni : *Expert Advisory Group of Wellcome* (2020)
- Europe : membre du *College of Expert Reviewers* de l'*European Science Foundation* depuis 2019

Prix

Lauréat en 2011 puis en 2012 de *Google Research Awards* en Humanités Numériques