



Situation Professionnelle

Enseignement

- 2011-Présent Professeur des Universités (PREX1) à l'Université de Caen Normandie (UCN).
- 2024-Présent Professeur Associé à l'Université Technologique de Munster (MTU, Irlande).
- 2002-2011 Maître de Conférences à l'Université de la Beira Interior (UBI, Portugal).
- 2008-2011 Maître de Conférences invité à l'Université de Porto (Portugal).

Recherche

- 2022-Présent Chercheur associé du GIS BB@C.
- 2011-Présent Chercheur du laboratoire GREYC UMR 6072.
- 2004-2011 Chercheur au Centre de Technologie du Langage Humain et Bioinformatique de l'UBI.
- 2002-2011 Chercheur au Centre de Mathématiques de l'UBI.
- 2008-2011 Chercheur invité à l'INESC-ID de Lisbonne (Portugal).

Responsabilités

Enseignement

- 2021-Présent Co-Responsable du Master Informatique de l'UCN.
- 2021-Présent Membre élu du Conseil de Département d'Informatique de l'UCN.
- 2022-Présent Membre du Conseil de Perfectionnement du Master Informatique de l'UCN.
- 2012-2016 Co-Responsable du Master Traitement Automatique de l'Image et de la Langue de l'UCN.
- 2012-2016 Membre élu du Conseil de Département d'Informatique de l'UCN.
- 2007-2009 Responsable du Programme Doctoral en Informatique de l'UBI.
- 2005-2007 Vice-Président du Département d'Informatique de l'UBI.

Recherche

- 2019-Présent Directeur Adjoint du GREYC UMR 6072.
- 2025-Présent Responsable du Laboratoire Partenaire International INTELLIGENT
- 2020-Présent Membre du Comité de Conseil du *Technology Innovation Hub* de l'IIT-Patna (Inde).
- 2019-Présent Membre du Comité de Pilotage du Collège "Technologies du Langage Humain" de l'AFIA.
- 2022-2025 Membre du Groupe de Travail en TAL de la Stratégie Nationale pour l'Intelligence Artificielle.
- 2019-2023 Membre élu du CA de l'Association Française pour l'Intelligence Artificielle (AFIA).
- 2019-2021 Membre du Conseil de l'ED MIIS 590.
- 2012-2016 Responsable de l'équipe HULTECH du GREYC UMR 6072.
- 2004-2011 Directeur du Centre de Technologie du Langage Humain et Bioinformatique de l'UBI.

Formation

- 2010 Habilitation à Diriger des Recherches. *Information Digestion*. Université d'Orléans.
- 2002 Doctorat en Informatique. Extraction Automatique d'Associations Lexicales à Partir de Corpora. Université d'Orléans et Université Nouvelle de Lisbonne (Portugal).

1995 Diplômes d'Études Approfondies. *An Approach to Machine Translation: An Interlingua Machine Translation System*. Université d'Orléans.

Enseignement

- Licence Calcul Scientifique I et II (2021-), Introduction au Traitement du Langage Naturel (2013), UML (2013), Analyse Structurée des Systèmes d'Information (2002-2010), Analyse Objet des Systèmes d'Information (2003-2004), Intelligence Artificielle (2002-2010), Interfaces Homme-Machine (2007-2009), Publication et Recherche d'Information (2007-2010), Technologie du Langage Humain (2004-2007).
- Master Introduction au Traitement Automatique des Langues (2022-), Traitement Automatique des Langues (2022-), Traitement Automatique des Langues Avancé (2022-), Apprentissage Avancé (2022-), Ingénierie des Données (2022-), Ingénierie des Connaissances Textuelles (2012-2013), Ingénierie des Langues (2012-2013), Traitement Automatique du Langage Naturel (2012-2013, 2021-), Moteurs de Recherche (2012), Fouille de Données et Apprentissage (2012), Apprentissage Automatique (2012), Recherche d'Information Multimédia (2012), Interfaces Homme-Machine (2012), Sémantique et Documents (2011), Traduction Automatique (2011), Documents Composites (2011), Données Textuelles (2011), Web Sémantique (2011), Exploration et Entrepôt de Données (2006-2010).
- Doctorat Traitement Statistique du Langage Naturel (2010).

Recherche

- Thèmes Gaël Dias s'intéresse à (1) la Sémantique lexicale, (2) la Digestion de l'information, (3) l'Accessibilité du web pour les déficients visuels et (4) au traitement multimodal du langage pour les troubles mentaux. Dans (1), ses travaux portent sur la représentation sémantique des mots, l'extraction d'unités polylexicales, l'identification de relations sémantiques et la construction de taxonomies. En (2), il s'intéresse à la segmentation thématique, au résumé de texte, à la simplification des phrases et au partitionnement éphémère. En (3), il développe des interfaces haptiques et orales concurrentes pour l'écramage de textes. En (4), il propose des modèles de fusion, d'apprentissage hybride et par transfert, et développe des techniques de phénotypage digital.
- Publications (Co-)Auteur de plus de 170 publications en chapitres de livres, revues, conférences et ateliers internationaux dont 66 de classe Q1/A*/A (AAAI, ACL, AIM, BI, CIKM, COLING, CoNLL, CS, DL, EACL, ECAI, ECIR, ECML/PKDD, EMNLP, ESWC, ICDAR, ICDM, ICWSM, IEEE CIM, IEEE IS, IJCNN, INTERSPEECH, IPM, KDD, MM, NAACL, NLE, PLOS ONE, SIGIR, TOIS, WWW), 23 de classe Q2/B (ACIVS, AIME, CACLING, EAMT, EKAW, EPIA, IJCNLP, MICROMACHINES, MT, PACLIC, PACLING, SMC, SNAM, SPIRE, WI-IAT) et 19 de classe Q3/C (AWIC, COMPUTERS, EGC, FLAIRS, JADT, LREC, NLDB, RANLP, TALN, WIC) selon CORE et SJR.
- Encadrements (Co-)Directeur des thèses de Doctorat en Informatique de Sami Mechtouf (en cours), Amine Youcef Khodja (en cours), Søren Fomsgaard (en cours), Nilesh Tete (en cours), Soumaya Sabry (en cours), Navneet Agarwal (2024), Kirill Milintsevich (2024), Houssam Akhmouch (2022), João Paulo Cordeiro (2011), Dinko Lambov (2011), Rumen Moraliyski (2012), Ricardo Campos (2013), Sebastião Pais (2013) et José Moreno (2014). Co-directeur de la thèse de Doctorat en Lettres d'Isabel Marcelino (2012). Responsable de 23 stages de Master et de plus de 80 projets tutorés de Licence et Master.
- Projets Internationaux Collaborateur des projets HYBRIDS (HORIZON-TMA-MSCA-DN 101073351), AI4GREENBUS (ERASMUS-EDU-2025-PI-ALL-INNO).
- Projets Bilatéraux Responsable des projets LEILA (F-20/05) et LEILA II (F-48/07) avec l'Université d'Orléans et du projet Hubert Curien MULTILEXI avec l'Université de Ljubljana (Slovénie).

| | |
|-------------------|---|
| Projets Nationaux | Responsable des projets MENTAL.AI@CaeSAR (ANR-23-EXES-0001), SEMCOMP (CPER), VIPACCESS (PTDC/PLP/72142/2006), SUMO (POSC/PLP/57438/2004) et collaborateur des projets AI4GREENBUS (ERASMUS-EDU-2025-PI-ALL-INNO), MEDON (PTDC/EIA/80772/2006), PT-STAR (CMU-PT/HuMach/0039/2008), SITE-O-MATIC (POSC/EIA/58367/2004), MOVES (PTDC/EEI-AUT/28918/2017), DOCSCOPE (ANR-12-SECU-0006), ART-ADN (ANR-12-SOIN-0003-02), TAGTHUNDER (FSN), A2M2P (FHU), MUSE (GIP), HIGH-TECH (RIN), INDYX (BPI), 2E2MR (ANR) et PARTAGES (BPI). |
| Expertises | Membre du comité de rédaction de la revue TAL (2011-2021). Membre du comité de relecture de 11 revues internationales dont 4 de classe Q1/A (IPM, JCSL, NLE, TSLP) et 3 de classe Q2 (AICOMM, DKE, JLRE). Membre du comité de relecture de 23 conférences et ateliers internationaux dont 9 de classe A (AAAI, ACL, COLING, EAACL, ECIR, EMNLP, HLT-NAACL, IJCAI, SIGIR), 2 de classe B (EPIA, IJCNLP) et 6 de classe C (AWIC, EGC, FLAIRS, ICAART, RANLP, TALN). |
| Evaluations | Membre du comité d'évaluation auprès du Pôle de Compétitivité Numérique de Normandie, de l'ANRT, d'ECOS Nord, du Fonds de Recherche du Québec (Canada), du <i>Dutch CLARIAH Program for Digital Arts and Humanities</i> (Pays-Bas), de la <i>National Commission for Scientific and Technological Research</i> (Chili) et du Fonds de la Recherche Scientifique - FNRS (Belgique). |
| Jurys | Rapporteur et examinateur de 7 habilitations à diriger des recherches (section 27) et de 62 thèses de doctorat en Informatique en France et à l'étranger (Espagne, Finlande, Inde, Portugal). |
| Événements | Co-organisateur de la 22ème Conférence sur le Traitement Automatique des Langues Naturelles (TALN 2015), de la 12ème Conférence Portugaise en Intelligence Artificielle (EPIA 2005) et des ateliers internationaux MEMURA 2004, GAIW 2005, ELECTRA 2005, TEMA 2007/2009/2013/2015, TEXTGRAPH 2012, HISTOINFORMATICS 2013/2014/2017/2019/2021, WNLPe-Health 2022, EWS 2025. Co-organisateur de l'Ecole Saisonnière en Intelligence Artificielle (ESIA 2023) et de MENTAL.AI 2022/2024/2026. |
| Éditions | Co-éditeur de l'issue spéciale "Passage à l'échelle: complexité, algorithmique et architecture" de la revue TAL (2005), de l'issue spéciale "Histoinformatics : computational approaches to history" de la revue JDMDH (2020), de l'issue spéciale "Advances in machine learning methods for natural language processing and computational linguistics" de la revue MATHEMATICS (2021), de l'issue spéciale "Emotion processing: text and multi-modal data" de la revue <i>Frontiers in Artificial Intelligence</i> (2022) et de l'issue spéciale "Traitement automatique des langues pour la santé et l'accessibilité" de la revue TAL (2025). |
| Ressources | SENTA (Logiciel pour l'extraction d'associations textuelles N-aires) et TEMPOWORNET (Base de données lexicales pour l'analyse temporelle). |
| Distinctions | Prix du meilleur article du <i>9th International Symposium on Social Communication</i> , Santiago de Cuba, (2005). Prix du meilleur projet au concours Imagine Cup Microsoft Portugal (2006). Prix du meilleur projet au concours Parkurbis Portugal (2007). |
| Prime | Prime d'encadrement doctoral et de recherche (2013-2017), Régime indemnitaire des personnels enseignants et chercheurs (2023-). |

Bibliographie (uniquement A*/A et Q1/Q2 - hors ateliers)

Revue d'Audience Internationale avec Comité de Rédaction

- [1] Zahoor S., Liò P., Dias G., and Hasanuzzaman M. Integrating probabilistic trees and causal networks for clinical and epidemiological data. *Artificial Intelligence in Medicine*, 2026. [Q1, IF=6.20]. ISSN: 1873-2860.
- [2] Maggini M., Bassi D., Piot P., Dias G., and Gamallo-Otero P. A systematic review of automated

- hyperpartisan news detection. *PLOS One*, 2025. [Q1, IF=2.90]. eISSN: 1932-6203.
- [3] Muetunda F., Sabry S., Jamil M.L., Pais S., and Dias G. and Cordeiro J-P. Ai-assisted diagnosing, monitoring, and treatment of mental disorders: A survey. *ACM Transactions on Computing for Healthcare*, 2024. [Q1, IF=1.49]. ISSN: 2637-8051.
 - [4] Agarwal N., Dias G., and Dollfus s. Multi-view graph-based interview representation to improve depression level estimation. *Brain Informatics*, 2024. [Q1, IF=1.31]. ISSN: 2198-4018.
 - [5] Milintsevich K., Sirts K., and Dias G. Towards automatic text-based estimation of depression through symptom prediction. *Brain Informatics*, 2023. [Q1, IF=1.31]. ISSN: 2198-4018.
 - [6] Luqman Jamil M., Pais S., Cordeiro J.P., and Dias G. Detection of extreme sentiments on social networks with bert. *Social Network Analysis and Mining*, 2022. [Q2, IF=3.87]. ISSN: 1869-5450.
 - [7] Jayashree R.S., Dias G., Andrew J-J., Saha S., Maurel F., and Ferrari S. Multimodal web page segmentation using self-organized multi-objective clustering. *ACM Transactions on Information Systems*, 2022. [Q1, IF=2.89]. ISSN: 1046-8188.
 - [8] Qureshi S.A., Dias G., Hasanuzzaman M., and Saha S. Improving depression level estimation by concurrently learning emotion intensity. *IEEE Computational Intelligence Magazine*, 2020. [Q1, IF=5.86]. ISSN: 1556-603X.
 - [9] Maurel F., Dias G., Safi W., Routoure J-M., and Beust P. Layout transposition for non-visual navigation of web pages by tactile feedback on mobile device. *Micromachines*, 2020. [Q2, IF=2.43]. ISSN: 2072-666X.
 - [10] Qureshi S.A., Saha S., Hasanuzzaman M., and Dias G. Multi-task representation learning for multimodal estimation of depression level. *IEEE Intelligent Systems*, 2019. [Q1, IF=2.59]. ISSN: 1541-1672.
 - [11] Campos R., Dias G., Jorge A., and Nunes C. Identifying top relevant dates for implicit time sensitive queries. *Information Retrieval*, 2017. [Q2, CORE=B, IF=2.54]. ISSN: 1573-7659.
 - [12] Campos R., Dias G., Jorge A., and Nunes C. Gte-rank: a time-aware search engine to answer time-sensitive queries. *Information Processing and Management*, 2016. [Q1, IF=3.89]. ISSN: 0306-4573.
 - [13] Campos R., Dias G., Jorge A., and Jatowt A. Survey of temporal information retrieval and related applications. *ACM Computing Surveys*, 47(2):1–41, 2014. [Q1, IF=6.13]. ISSN: 0360-0300.
 - [14] Dias G., Moraliyski R., Cordeiro J.P., Doucet A., and Ahonen-Myka H. Automatic discovery of word semantic relations using paraphrase alignment and distributional lexical semantics analysis. *Natural Language Engineering*, 16(4):439–467, 2010. [Q2, CORE=A, IF=1.07].

Communications d'Audience Internationale avec Comité de Sélection

- [15] Maggini M., Merzougui D.E., Bandyopadhyay R., Dias G., Maurel F., and Gamallo-Otero P. Are llms enough for hyperpartisan, fake, polarized and harmful content detection? evaluating in-context learning vs. fine-tuning? In *Proceedings of the 20th International AAAI Conference on Web and Social Media (ICWSM)*, 2026. [CORE=A].
- [16] Ledoyen F., Dias G., Pantin J., Maurel F., Lechervy A., and Chahir Y. Facilitating cognitive accessibility with llms: A multi-task approach to easy-to-read text generation. In *Proceedings of the 30th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2025. [CORE=A*].
- [17] Lemaire R., Kaibaldiyev A., Mariette E., Viglieri D., Lechervy A., Maurel F., Dias G., Pantin J., Blaiot G., Agin V., Poirel N., Bui B., Platel H., Vivien D., and Chahir Y. Wysiwyg: What you see is where your gaze. In *Proceedings of the 33rd ACM International Conference on Multimedia (MM)*, 2025. [CORE=A*].
- [18] Ledoyen F., Dias G., Maurel F., Pantin J., Chahir Y., Lechervy A., Gouzonnat E., Berthelot M., Moravac S., Altinier A., and Khairalla A. Inclusive easy-to-read generation for individuals with cognitive impairment. In *Proceedings of the 28th European Conference on Artificial Intelligence (ECAI)*, 2025. [CORE=A].
- [19] Bournonville A., Rioult F., Maurel F., Dias G., Pitel G., Largillier T., and Marchand E. Multilingual evaluation of main content extractors for web pages. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2025. [CORE=A*].
- [20] Merzougui D.E., Tete N., Maurel F., Dias G., Hasanuzzaman M., Bournonville A., Madelaine E., Berthelin Le Tellier T., Ledoyen F., Poutrain-Lejeune L., Rioult F., and Pantin J. Concurrent speech and

- auditory tag clouds for non-visual web interaction. In *Proceedings of the 26th Interspeech Conference (INTERSPEECH)*, 2025. [CORE=A].
- [21] Agarwal N., Milintsevich K., Métivier L., Rotharmel M., Dias G., and Dollfus S. Analyzing symptom-based depression level estimation through the prism of psychiatric expertise. In *Proceedings of the Joint International Conference on Computational Linguistics, Language Resources and Evaluation (COLING-LREC)*, 2024. [CORE=A].
- [22] Jha P., Dias G., Lechervy A., Jangra A., Moreno J., Pais S., and Sriparna S. Combining vision and language representations for patch-based identification of lexico-semantic relations. In *Proceedings of the 30th ACM International Conference on Multimedia (ACM MM)*, 2022. [CORE=A*].
- [23] Akhmouch H., Dias G., and Moreno J. Understanding feature focus in multitask settings for lexico-semantic relation identification. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021)*, 2021. [CORE=A*].
- [24] Rane C., Dias G., Lechervy A., and Ekbal A. Improving neural text style transfer by introducing loss function sequentiality. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2021)*, 2021. [CORE=A*].
- [25] Qureshi S.A., Dias G., Saha S., and Hasanuzzaman M. Gender-aware estimation of depression severity level in a multimodal setting. In *Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN 2021)*, 2021. [CORE=A].
- [26] Kumar R., Chauhan D., Dias G., and Ekbal A. Modelling personalized dialogue generation in multi-party settings. In *Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN 2021)*, 2021. [CORE=A].
- [27] Campos R., Duque J, Cândido T., Mendes J., Dias G., Jorge A., and Nunes C. Time-matters: Temporal unfolding of texts. In *Proceedings of the 43rd European Conference on Information Retrieval (ECIR 2021)*, 2021. [CORE=A].
- [28] Bannour N., Dias G., Chahir Y., and Akhmouch H. Patch-based identification of lexical semantic relations. In *Proceedings of the 42nd European Conference on Information Retrieval (ECIR 2020)*, 2020. [CORE=A].
- [29] Balikas G., Dias G., Moraliyski R., Akhmouch H., and Amini M-R. Learning lexical-semantic relations using intuitive cognitive links. In *Proceedings of the 41st European Conference on Information Retrieval (ECIR 2019)*, 2019. [CORE=A].
- [30] Hasanuzzaman M., Dias G., Ferrari S., Mathet Y., and Way A. Identifying temporal orientation of word senses. In *Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning (CoNLL 2016)*, 2016. [CORE=A].
- [31] Hasanuzzaman M., Sze W.L., Salim M.P., and Dias G. Collective future orientation and stock markets. In *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI 2016)*, 2016. [CORE=A].
- [32] Hasanuzzaman M. and Ferrari S. and Dias G. Transductive learning for the identification of word sense temporal orientation. In *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI 2016)*, 2016. [CORE=A].
- [33] Hasanuzzaman M., Dias G., Ferrari S., and Mathet Y. Iterative tempowordnet. In *Proceedings of the 13th Extended Semantic Web Conference (ESWC 2016)*, 2016. [CORE=A].
- [34] Cleuziou G. and Dias G. Learning pretopological spaces for lexical taxonomy acquisition. In *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2015)*, 2015. [CORE=A].
- [35] Moreno J. and Dias G. Adapted b-cubed metrics to unbalanced datasets. In *Proceedings of the 38th Annual ACM SIGIR Conference (SIGIR 2015)*, 2015. [CORE=A*].
- [36] Hasanuzzaman M., Saha S., Dias G., and Ferrari S. Understanding temporal query intent. In *Proceedings of the 38th Annual ACM SIGIR Conference (SIGIR 2015)*, 2015. [CORE=A*].
- [37] Campos R., Dias G., Jorge A., and Nunes C. Gte-rank: Searching for implicit temporal query results. In *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management (CIKM 2014)*, pages 1–41, 2014. [CORE=A].

- [38] Moreno J. and Dias G. Pagerank-based word sense induction within web search results clustering. In *Proceedings of Digital Libraries (DL 2014)*, 2014. [CORE=A*].
- [39] Acharya S., Saha S., Moreno J., and Dias G. Multi-objective search results clustering. In *Proceedings of the 25th International Conference on Computational Linguistics (COLING 2014)*, 2014. [CORE=A].
- [40] Moreno J., Dias G., and Cleuziou G. Query log driven web search results clustering. In *Proceedings of the 37th Annual ACM SIGIR Conference (SIGIR 2014)*, pages 777–786, 2014. [CORE=A*].
- [41] Hasanuzzaman M., Dias G., Ferrari S., and Mathet Y. Propagation strategies for building temporal ontologies. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2014)*, pages 6–11, 2014. [CORE=A].
- [42] Moreno J. and Dias G. Easy web search results clustering: When baselines can reach state-of-the-art algorithms. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2014)*, pages 1–5, 2014. [CORE=A].
- [43] Campos R., Dias G., Jorge A., and Nunes C. Gte-cluster: A temporal search interface for implicit temporal queries. In *Proceedings of the 35th European Conference on Information Retrieval (ECIR 2014)*, pages 775–779, 2014. [CORE=A].
- [44] Moreno J. and Dias G. Using text-based web image search results clustering to minimize mobile. In *Proceedings of the 34th European Conference on Information Retrieval (ECIR 2013)*, pages 775–779, 2012. [CORE=A].
- [45] Moreno J., Dias G., and Cleuziou G. Post-retrieval clustering using third-order similarity measures. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013)*, pages 153–158, 2013. [CORE=A*].
- [46] Campos R., Dias G., Jorge A., and Nunes C. Gte: A distributional second-order co-occurrence approach to improve the identification of top relevant dates in web snippets. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM 2012)*, pages 2035–2039, 2012. [CORE=A].
- [47] Moreno J. and Dias G. Image results exploration using ephemeral clustering. In *Proceedings of the 2012 World Wide Web Conference (WWW 2012)*, 2012. [CORE=A*].
- [48] Cleuziou G., Buscaldi D., Levorato V., and Dias G. A pretopological framework for the automatic construction of semantic-lexical structures from texts. In *Proceedings of the 20th ACM Conference on Information and Knowledge Management (CIKM 2011)*, pages 2453–2456, 2011. [CORE=A].
- [49] Grigonyté G., Cordeiro J.P., Moraliyski R., Dias G., and Brazdil P. A paraphrase alignment for synonym evidence discovery. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, pages 403–411, 2010. [CORE=A].
- [50] Dias G., Mukelov R., and Cleuziou G. Unsupervised graph-based discovery of general-specific noun relationships from web corpora frequency counts. In *12th International Conference on Natural Language Learning (CoNLL 2008)*, pages 97–104, 2008. [CORE=A].
- [51] Dias G., Alves E., and Lopes J.G.P. Topic segmentation algorithms for text summarization and passage retrieval: An exhaustive evaluation. In *22nd Conference on Artificial Intelligence (AAAI 2007)*, pages 1334–1340, 2007. [CORE=A*].