

Ahmed Hamdi

ahmed.hamdi@univ-lr.fr

Laboratoire L3i,

Faculté des Sciences et Technologies Avenue Michel Crépeau

17042, La Rochelle

<https://pageperso.univ-lr.fr/ahmed.hamdi/>

Parcours universitaire

- 2011-2015* **Diplôme :** Doctorat
Discipline : Sciences de langage
Lieu : Aix-Marseille université
Mention : Très honorable
- 2010-2011* **Diplôme :** Master
Discipline : Sciences de langage
Spécialité : Traitement automatique des langues
Lieu : Aix-Marseille université
Mention : Bien
- 2007-2010* **Diplôme :** Ingénieur
Discipline : Informatique
Lieu : École nationale de sciences de l'informatique (Tunisie)
Mention : Assez bien
- 2005-2007* **Diplôme :** Premier cycle
Discipline : Mathématique-Physique
Lieu : Institut préparatoire des études d'ingénieur d'El Monastir (Tunisie)
Mention : Assez bien

Expériences professionnelles

- 2022-Aujourd'hui* **Poste :** Enseignant chercheur
Lieu : La Rochelle Université
Unité de recherche : Laboratoire Informatique, Image et Interaction (L3i)
- 2017-2022* **Poste :** Post-doctorant et vacataire
Lieu : La Rochelle Université
Unité de recherche : Laboratoire Informatique, Image et Interaction (L3i)
- 2014-2016* **Poste :** Attaché temporaire d'enseignement et de recherche (ATER)
Lieu : Aix-Marseille Université
Unité de recherche : Laboratoire d'informatique fondamentale (LIF)

2011-2014 **Poste :** Doctorant et moniteur
Lieu : Aix-Marseille Université
Unité de recherche : Laboratoire d'informatique fondamentale (LIF)

Activités de recherche

2022-Aujourd'hui **Poste :** Post-doctorant
Unité de recherche : Laboratoire Informatique, Image et Interaction (L3i)
Lieu : La Rochelle Université

Responsabilités :

- Développement des relations entre le site socio-économique et culturel niortais et La Rochelle Université.

2017-2022 **Poste :** Post-doctorant
Unité de recherche : Laboratoire Informatique, Image et Interaction (L3i)
Lieu : La Rochelle Université
Projets : IDEAS, NewsEye, SecurDoc

Thèmes de recherche :

- Extraction d'informations à partir des documents administratifs.
- Reconnaissance et désambiguïsation d'entités nommées à partir de la presse ancienne.
- Détection d'opinions envers les entités nommées.
- Séparation de flux documentaires.

Autres responsabilités :

- Encadrement de l'étudiant Vinh-Nam Huynh en L3 informatique à l'université de Hanoï. L'encadrement a aboutit à une publication dans la conférence ICADL-2020.
- Encadrement de l'étudiant Joris Voerman en Master 2 informatique, qui s'est poursuivi en thèse. L'encadrement a aboutit à une publication dans le workshop ICDAR-WML-2017.

Mai 2014-Sep 2014

Poste : Stagiaire scientifique
Unité de recherche : Center for Computational Learning Systems (CCLS)
Lieu : Columbia University (New York - USA)
Superviseurs : Nizar Habash, Owen Rambow
Sujet de recherche : Analyse morphologique des dialectes arabes.

2011-2014 **Poste :** Doctorant
Unité de recherche : Laboratoire d'informatique fondamentale de Marseille
(LIF-CNRS)

Lieu : Aix-Marseille Université
Superviseurs : Alexis Nasr, Núria Gala
Sujet de recherche : Traitement automatique du dialecte tunisien à partir des ressources de l'arabe standard.

Février 2011-Juin 2011

Poste : Stage de fin d'étude de master
Unité de recherche : Laboratoire d'informatique fondamentale de Marseille
(LIF-CNRS)

Lieu : Aix-Marseille Université
Superviseurs : Alexis Nasr, Núria Gala
Sujet de recherche : Apport de la diacritisation dans l'analyse morphosyntaxique de l'arabe.

Février 2010-Juillet 2010

Poste : Stage de fin d'études d'ingénieur
Unité de recherche : laboratoire Langage, Langues et Cultures d'Afrique
(LLACAN-CNRS)

Lieu : École Nationale de Sciences de l'Informatique (ENSI - Tunisie)
Superviseurs : Fethi Debili
Sujet de recherche : Étiquetage grammatical de l'arabe standard.

Activités d'enseignement

2021-2022

Poste : Enseignant chercheur
Lieu : La Rochelle Université
Nombre d'heures : 104H
Modules enseignés : Réseaux et sécurité, Architecture orientée service, Modélisation, Bases du web, Technologies de l'information, Mise en oeuvre des systèmes Big data [[CM](#), [TP](#)], Informatique dans l'e-tourisme [[CM](#), [TP](#)].

2017-2021

Poste : Vacatire
Lieu : La Rochelle Université
Nombre d'heures : 50H
Modules enseignés : Technologies de l'information, Mise en oeuvre des systèmes Big data Informatique dans l'e-tourisme

2014-2016

Poste : ATER
Lieu : Aix-Marseille Université
Nombre d'heures : 384H
Modules enseignés : Algorithmique, Introduction informatique et programmation, Programmation C II, Outils informatiques et C2i, Programmation Python II, Développement web, Automates et circuits.

2011-2014

Poste : Moniteur
Lieu : Aix-Marseille Université
Nombre d'heures : 128H
Modules enseignés : Introduction informatique et programmation,

Récapitulation des activités d'enseignement :

Année	Intitulé	Public	Spécialité	Cours	TD	TP
2021-2022	Mise en œuvre des systèmes big data	Master 1	Informatique	4	–	6
	Architecture orientée service	Master 1	Informatique	3	9	–
	Réseaux et sécurité	Master 1	Informatique	18	12	–
	Informatique dans l'e-tourisme	toutes les licences		6	6	–
	Technologies de l'information	Licence 2	Gestion	–	28	–
	Modélisation	Licence 1	Informatique	–	7,5	15
	Bases du web	Licence 1	Informatique	–	30	–
2020-2021	Mise en œuvre des systèmes big data	Master 1	Informatique	4	–	6
	Informatique dans l'e-tourisme	toutes les licences		6	6	–
	Technologies de l'information	Licence 2	Gestion	–	28	–
2019-2020	Technologies de l'information	Licence 2	Gestion	–	28	–
2018-2019	Technologies de l'information	Licence 2	Gestion	–	28	–
2017-2018	Technologies de l'information	Licence 2	Gestion	–	14	–
2016-2017	Technologies de l'information	Licence 2	Gestion	–	14	–
2015-2016	Algorithmique	Licence 2	Mathématiques	–	24	20
	Introduction informatique et programmation	Licence 1	Informatique	20	24	21
	Outils informatiques et C2I	Licence 1	Informatique	–	24	–
	Programmation Python II	Licence 1	Informatique	–	24	–
	Automates et circuits	Licence 1	Informatique	–	24	18
2014-2015	Algorithmique	Licence 2	Informatique	–	24	20
	Outils informatiques et C2I	Licence 1	Informatique	–	48	–
	Introduction informatique et programmation	Licence 1	Informatique	–	–	63
	Programmation C II	Licence 1	Informatique	–	18	21
	Développement web I	Licence 1	Informatique	–	15	–
2013-2014	Introduction informatique et programmation	Licence 1	Informatique	–	–	42
	Outils informatiques et C2I	Licence 1	Informatique	–	24	–
2012-2013	Introduction informatique et programmation	Licence 1	Informatique	–	–	21
	Outils informatiques et C2I	Licence 1	Informatique	–	24	–
	Automates et circuits	Licence 1	Informatique	–	24	–

Publications**1. Articles de revue**

- NLE 2022 (Q1).
In-depth analysis of the impact of OCR errors on named entity recognition and linking. **Ahmed Hamdi**, Elvys Linhares Pontes, Nicolas Sidère, Mickaël Coustaty et Antoine Doucet. Natural Language Engineering Journal, pp 1 – 24.
- IJDL 2021 (Q2).
MELHISSA: A Multilingual Entity Linking Architecture for Historical Press Articles. Elvys Linhares Pontes, Luis Adrian Cabrera-Diego, Jose G. Moreno, Emanuela Boros, **Ahmed Hamdi**, Nicolas Sidère, Mickaël Coustaty et Antoine Doucet. International Journal on Digital Libraries, IJDL. pp 1 – 28.
- arXiv preprint 2021.
Named Entity Recognition and Classification on Historical Documents: A Survey. Maud Ehrmann, **Ahmed Hmadi**, Elvys Linhares Pontes, Matteo Romanello et

Antoine Doucet. Version courte acceptée à la revue ACM Computing Surveys (Q1 – impact factor = 7.99).

2. Conférences internationales

- SIGIR 2021 (Core A*).
A Multilingual Dataset for Named Entity Recognition, Entity Linking and Stance Detection in Historical Newspapers. **Ahmed Hmadi**, Elvys Linhares Pontes, Emanuela Boros, Tuyet Hai Nguyen Thi, Hackl Günter, Jose G. Moreno et Antoine Doucet. The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2021).
- ICDAR 2021 (Core A).
Information extraction from invoices. **Ahmed Hmadi**, Elodie Carel, Aurélie Joseph, Mickaël Coustaty et Antoine Doucet. The 16th International Conference on Document Analysis and Recognition.
- TPDFL 2020 (Core B).
Assessing and Minimising the Impact of OCR Quality on Named Entity Recognition. **Ahmed Hmadi**, Axel-Jean Caurant, Nicolas Sidère, Mickaël Coustaty et Antoine Doucet. Digital Libraries for Open Knowledge - 24th International Conference on Theory and Practice of Digital Libraries. Springer, 2020, pp. 87–101. DOI: 10.1007/978-3-030-54956-5_7. https://doi.org/10.1007/978-3-030-54956-5_7 [SLIDES]
- CONLL 2020 (Core A).
Alleviating Digitization Errors in Named Entity Recognition for Historical Documents. Emanuela Boros, **Ahmed Hamdi**, Elvys Linhares Pontes, Luis Adrian Cabrera-Diego, Jose G. Moreno, Nicolas Sidère et Antoine Doucet. Proceedings of the 24th Conference on Computational Natural Language Learning, pp. 431–441. <https://www.aclweb.org/anthology/2020.conll-1.35>
- ICADL 2020 (Core A).
When to Use OCR Post-correction for Named Entity Recognition? Vinh-Nam Huynh, **Ahmed Hamdi** et Antoine Doucet. Digital Libraries at Times of Massive Societal Transition - 22nd International Conference on Asia-Pacific Digital Libraries. Lecture Notes in Computer Science 12504 (2020), pp, 33-42. DOI: 10.1007/978-3-030-64452-9_3. https://doi.org/10.1007/978-3-030-64452-9_3. [SLIDES]
- ICADL 2020 (Core A).
Entity Linking for Historical Documents: Challenges and Solutions. Elvys Linhares Pontes, Luis Adrian Cabrera-Diego, Jose G. Moreno, Emanuela Boros, **Ahmed Hamdi**, Nicolas Sidère, Mickaël Coustaty et Antoine Doucet. Digital Libraries at Times of Massive Societal Transition - 22nd International Conference on Asia-Pacific Digital Libraries. Lecture Notes in Computer Science 12504 (2020), pp, 215-231. DOI: 10.1007/978-3-030-64452-9_19. https://doi.org/10.1007/978-3-030-64452-9_19.
- JCDEL 2019 (Core A*)
An Analysis of the Performance of Named Entity Recognition over OCRed Documents. Ahmed Hamdi, Axel Jean-Caurant, Nicolas Sidère, Mickaël Coustaty et Antoine Doucet. 19th ACM/IEEE Joint Conference on Digital Libraries, pp, 333-334. DOI: 10.1109/JCDEL.2019.00057. <https://doi.org/10.1109/JCDEL.2019.00057>.
- ICADL 2019 (Core A).
Impact of OCR Quality on Named Entity Linking. Elvys Linhares Pontes, **Ahmed**

Hamdi, Nicolas Sidere et Antoine Doucet. Digital Libraries at the Crossroads of Digital Information for the Future - 21st International Conference on Asia-Pacific Digital Libraries. Lecture Notes in Computer Science 11853 (2019) pp, 102-11. DOI: 10.1007/978-3-030-34058-2_11. https://doi.org/10.1007/978-3-030-34058-2_11,

- DAS 2018 (Core B).
Feature Selection for Document Flow Segmentation. **Ahmed Hamdi**, Mickaël Coustaty, Aurélie Joseph, Vincent Poulain D'Andecy, Antoine Doucet et Jean-Marc Ogier. 13th IAPR International Workshop on Document Analysis Systems. IEEE Computer Society, pp, 245–250. DOI:10.1109/DAS.2018.66. <https://doi.org/10.1109/DAS.2018.66>.
- MT SUMMIT 2013 (Core B).
The Effects of Factorizing Root and Pattern Mapping in Bidirectional Tunisian - Standard Arabic Machine Translation. **Ahmed Hamdi**, Rahma Boujelbane, Nizar Habash et Alexis Nasr. MT Summit 2013. <https://hal.archives-ouvertes.fr/hal-00908761>.

3. Workshops internationaux

- CEUR@CLEF 2020
Robust Named Entity Recognition and Linking on Historical Multilingual Documents. Emanuela Boros, Elvys Linhares Pontes, Luis Adrian Cabrera-Diego, **Ahmed Hamdi**, Jose G. Moreno, Nicolas Sidère et Antoine Doucet. Working Notes of CLEF 2020 - Conference and Labs of the Evaluation Forum, 2020. http://ceur-ws.org/Vol-2696/paper_171.pdf.
- WML@ICDAR 2017
Machine Learning vs Deterministic Rule-Based System for Document Stream Segmentation. **Ahmed Hamdi**, Joris Voerman, Mickaël Coustaty, Aurélie Joseph, Vincent Poulain D'Andecy et Jean-Marc Ogier. First Workshop of Machine Learning, 14th IAPR International Conference on Document Analysis and Recognition, WML@ICDAR 2017. IEEE, pp, 77-82. DOI: 10.1109/ICDAR.2017.332 <https://doi.org/10.1109/ICDAR.2017.332>.
- WANLP@ACL 2015
POS-tagging of Tunisian Dialect Using Standard Arabic Resources and Tools. **Ahmed Hamdi**, Alexis Nasr, Nizar Habash et Nuria Gala. Proceedings of the 2nd Workshop on Arabic Natural Language Processing Processing, ANLP@ACL 2015. Association for Computational Linguistics. DOI: 10.18653/v1/W15-3207. <https://doi.org/10.18653/v1/W15-3207>.
- VarDial@COLING 2014
Automatically building a Tunisian Lexicon for Deverbal Nouns. **Ahmed Hamdi**, Nuria Gala et Alexis Nasr. Proceedings of the 1st Workshop on Applying NLP Tools to Similar Languages, Varieties and Dialects, VarDial@COLING 2014. Association for Computational Linguistics and Dublin City University, pp, 95-102. DOI: 10.3115/v1/W14-5311. <https://doi.org/10.3115/v1/W14-5311>.

4. Livrables de projets

- *Named Entity Recognition and Linking*. **Ahmed Hamdi**, Elvys Linhares Pontes et Antoine Doucet. NewsEye 2020.

- *Stance detection*. **Ahmed Hamdi**, Thi Tuyet Hai Nguyen et Antoine Doucet. NewsEye 2020.

5. Rapport de thèse

- Ahmed Hamdi, 2015. *Traitement automatique du dialecte tunisien à l'aide d'outils et de ressources de l'arabe standard : application à l'étiquetage morphosyntaxique*. Aix-Marseille Université. <https://www.theses.fr/en/2015AIXM4089>.

6. Conférences nationales

- CORIA 2021
Atténuer les erreurs de numérisation dans la reconnaissance d'entités nommées pour les documents historiques. Emanuela Boros, **Ahmed Hamdi**, Elvys Linhares Pontes, Luis Adrian Cabrera-Diego, Jose G. Moreno, Nicolas Sidère et Antoine Doucet. Conférence francophone en Recherche d'Information et Application.
- TALN 2013
Un système de traduction de verbes entre arabe standard et arabe dialectal par analyse morphologique profonde. **Ahmed Hamdi**, Rahma Boujelbane, Nizar Habash et Alexis Nasr. Traitement Automatique des Langues Naturelles, TALN 2013. The Association for Computer Linguistics, pp, 395-406. <https://www.aclweb.org/anthology/F13-1029/>.
- RECITAL 2012
Apport de la diacritisation de l'analyse morphosyntaxique de l'arabe. **Ahmed Hamdi**. Proceedings of the Joint Conference JEP-TALN-RECITAL 2012. ATALA/AFCP pp, 247-254. <https://www.aclweb.org/anthology/F12-3019/>.

Responsabilités collectives et administratives

- Séminaire sur l'analyse du contenu digital avec la focalisation sur la reconnaissance et la désambiguïsation des entités nommées dans les textes historiques. Research Group in Computational Linguistics, Université de Wolverhampton, 19/05/2022.
- Participation à l'organisation de la conférence internationale CICLING (*International Conference on Intelligent Text Processing and Computational Linguistics*) en 2019 à La Rochelle. https://www.cicling.org/2019/#Conference_committees
- Relecture d'articles scientifiques (*Journal of Language Resources and Evaluation*).
- Participation à l'organisation de la conférence internationale CICLING (*International Conference on Intelligent Text Processing and Computational Linguistics*) 2019 à La Rochelle. https://www.cicling.org/2019/#Conference_committees
- Relecture d'articles scientifiques (*Journal of Language Resources and Evaluation*).
- Relecture de livrables de projets : NewsEye (<https://cordis.europa.eu/project/id/770299>) & EMBEDDIA (<https://cordis.europa.eu/project/id/825153>).
- Participation à l'organisation de la Journée des Ingénieurs, Doctorants, ATER et Post-docs (JIDAP) en 2018 et 2019. La JIDAP est un événement annuel au laboratoire L3i qui a pour objectif de favoriser les échanges entre les Ingénieurs, Doctorants, ATER et PostDoc qui permettrait à chacun d'avoir une idée sur les thèmes de recherche de ses

collègues et de découvrir plein de choses utiles sur le laboratoire particulièrement aux nouveaux venants.

- Participation aux fêtes de la science 2018, 2019 2020 et 2021: la fête de la science est un événement national proposé par le Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation où les chercheurs scientifiques de La Rochelle Université présentent leurs travaux au public (y compris les écoliers, collégiens et lycéens) afin de satisfaire toutes les curiosités. Dans ce cadre, j'ai pu présenter mes travaux de recherche dans la segmentation de flux documentaires en 2018 et l'extraction d'informations à partir de documents administratifs en 2019. Enfin, en 2020, j'ai présenté mes travaux sur la reconnaissance et la désambiguïsation des entités nommées à partir de la presse ancienne. <https://www.univ-larochelle.fr/wp-content/uploads/pdf/FDLS-2020-PROGRAMME-PUBLIC-2020.pdf>
- Participation aux journées portes ouvertes 2019 et 2020: les journées portes ouvertes sont des événements organisés par l'université afin de présenter aux futur(e)s étudiant(e)s les différents travaux dans les unités de recherche rattaché à La Rochelle université. Ces journées font aussi l'occasion pour visiter les locaux et échanger avec les enseignants sur les formations dispensées et les services proposés. [[SLIDES JPO 2020](#)]

Récompenses et prix

- Qualification au poste de maître de conférence CNU 27 (informatique) en 2021
- CLEF-HIPE 2022: Named Entity Recognition (NER) and Entity Linking (EL) on Historical Newspapers,
 - NER: **l'équipe L3i est première sur 20/26 classements**
 - EL: **l'équipe L3i est première sur 6/8 classements**
<https://hipe-eval.github.io/HIPE-2022/results>
- CLEF-HIPE 2020: Named Entity Recognition (NER) and Entity Linking (EL) on Historical Newspapers,
 - NER: **l'équipe L3i est première sur tous les classements**
 - EL: **l'équipe L3i est première sur 50/52 classements et deuxième sur 2 classements**
https://github.com/impresso/CLEF-HIPE-2020/blob/master/evaluation-results/ranking_summary_final.md
- ICADL 2019: "Impact of OCR quality on named entity linking", **best paper**
- Hackathon "La Rochelle ville connectée" 2017, **prix du Hackathon**
<https://opendata.larochelle.fr/1er-hackathon-une-grande-reussite/>