

VIIème colloque international franco-espagnol du réseau E.GRAPHELES

Corpus spécialisés et Intelligence artificielle / humaine

Organisé par le laboratoire Grammatica en collaboration avec le laboratoire Textes et Cultures et son équipe CoTraLiS

Du 23 au 25 octobre 2024

A l'Université d'Artois – Maison de la Recherche à Arras

Le réseau interuniversitaire E-GRAPHELES (Grammaire, phraséologie et enseignement des langues étrangères et de spécialité) est né en 2012 d'un accord de collaboration signé entre l'Université d'Artois et l'Université Polytechnique de Valence. Il comprend désormais, outre les universités d'Artois et de Valence, co-fondatrices, l'université de Perpignan, l'université de Grenade, l'université de Murcie, l'université de Majorque et l'université de Neuchâtel.

Le réseau se fixe comme objectif scientifique de réunir les chercheurs spécialistes en analyse du discours et didactique des langues de spécialité, afin qu'ils/elles partagent leurs recherches, construisent des corpus spécialisés et des programmes de formation en contextes universitaire et professionnel, rédigent et diffusent des publications conjointes et organisent régulièrement des colloques internationaux, en français et en espagnol. Ces manifestations scientifiques se tiennent tous les deux ans alternativement en France et en Espagne.

Organisé par le laboratoire Grammatica en collaboration avec le laboratoire Textes et Cultures, en particulier son équipe CoTraLiS, le colloque international franco-espagnol « Corpus spécialisés et Intelligence artificielle/humaine » s'inscrit dans le prolongement du colloque organisé à l'université Polytechnique de Valence en octobre 2022 qui était consacré à la Communication et à l'Enseignement - apprentissage des langues dans les environnements numériques. Il traitera de l'utilisation de l'intelligence artificielle et humaine pour l'analyse des corpus linguistiques et littéraires existants ou en cours de construction.

L'Intelligence artificielle étend sa toile et son influence de façon exponentielle dans la recherche scientifique et l'enseignement supérieur et remet en question toutes les pratiques des enseignants-chercheurs ainsi que la relation pédagogique avec les étudiants.

Pour autant, elle demeure le fruit de l'intelligence humaine par laquelle passent sa régulation, son contrôle et son usage. Le rapport aux textes et aux discours qui constitue la trame des

recherches en Linguistique, Didactique et Littérature, s'en trouve profondément affecté et nécessite une réflexion de fond pour faire de l'IA un formidable outil d'analyse et de formation respectueux des principes déontologiques propres aux chercheurs.

S'agissant de la construction des corpus spécialisés, l'IA peut permettre une analyse rapide et poussée couplée à des dispositifs de formation en langue de spécialité et littérature plus adaptés aux différents contextes professionnels et académiques et aux besoins d'apprenants toujours plus spécifiques (Deep learning, Machine learning...). En Analyse de discours, l'IA peut aussi mettre en relation les différents corpus à des fins d'intercompréhension et de contextualisation. Une utilisation régulée de l'IA dans l'apprentissage des langues constitue donc une valeur sûre pour préparer la mobilité et l'accès au marché du travail.

Ainsi depuis plusieurs années le laboratoire Grammatica élabore des référentiels de compétences en langue professionnelle dans les secteurs du BTP, l'hôtellerie - restauration, l'hygiène - propreté, l'Aide à la personne (EHPAD), qui comportent un important corpus de textes oraux retranscrits faisant l'objet d'analyses lexicale, syntaxique et discursive à partir notamment de programmes informatiques d'analyse de discours (Sketch Engine, Tropes...). Ces référentiels ont ainsi permis d'élaborer des programmes de formation en français professionnel adaptés aux besoins des secteurs en tension. C'est pourquoi l'université d'Artois a pu obtenir un appel d'offre du ministère de l'intérieur en 2023 pour la formation des réfugiés pour le secteur de la restauration. C'est dans ce cadre que Grammatica mène un programme d'analyse de discours et de formation recourant à l'IA avec la conception d'une « Machine Learning » appliquée aux corpus professionnels issus des référentiels.

Quant au domaine de la phraséologie, le laboratoire Grammatica élabore des méthodes hybrides linguistique/informatique pour le traitement des expressions polylexicales ou multiwords, la constitution de corpus à l'aide d'outils informatiques ainsi que l'extraction et la constitution de ressources lexicales (collocations et pragmatèmes) dans l'objectif de mieux comprendre le langage humain (théories) et le langage artificiel (ChatGPT).

Des programmes de recherche recourant à l'IA et au « Deep Learning » sont également menés en Espagne à Valence et Grenade en particulier notamment pour l'analyse des lexiques de spécialité.

L'utilisation de l'IA et des « Machines Learning » doit être interrogée à l'aune de la recherche en analyse de discours et enseignement des langues de spécialité et, dans cette perspective, ce colloque permettra de comparer les expériences, de mutualiser les compétences et d'envisager des collaborations scientifiques.

Le colloque se déclinera en 3 axes d'études :

- Axe n°1 :** Intelligence artificielle / humaine en Analyse de Discours, outils, démarches et méthodologies, figement /défigement.
- Axe n°2 :** Exploitation didactique et Intelligence artificielle / humaine, conception de ressources pédagogiques, ressources lexicales de type phraséologique et ingénierie de formation.
- Axe n°3 :** Intelligence artificielle / humaine et corpus littéraires et journalistiques, web, spécialisés, etc., linguistique, linguistique de corpus, linguistique textuelle, phraséologie et stylistique, analyse littéraire des textes.

Bibliographie

- Goes, J., Mangiante, J.-M., Meneses, L., Olmo, F. y Pineira-Tresmontant, C. (eds), 2019, *Apports et limites des corpus numériques en analyse de discours et didactique des langues de spécialité*, Craiova : Editura Universitaria.
- Mangiante, J.-M. et Parpette, C., 2023, *Le Français sur objectif spécifique : ingénierie de formation et conception de programme*, Coll. F, Vanves : Hachette FLE.
- Mayaffre, Damon, Vanni, Laurent, *L'intelligence artificielle des textes. Des algorithmes à l'interprétation*, Paris, Honoré Champion, « Lettres Numériques », 15, 2021.
- Mejri S., Meneses-Lerín L. & Buffard-Moret B. (dir.), *La phraséologie française. La phraséologie française en questions ?*, Paris, Hermann, coll. Vertige de la langue, 2020, 464 p.
- Meneses-Lerín L., *Corpus et ressources numériques : nouveaux paradigmes de recherche en linguistique, en didactique et en traduction*, Studii de lingvistică, Vol. 7, Editura Universității din Oradea, 2017, 257 p.
- Meneses-Lerín, L., Molina Romero, C. y Montes Villar, L. (Coord.), 2020, *Ecrire, traduire et enseigner les langues à l'ère du numérique*, Comares.
- Olmo-Cazevieille, F. (ed.), 2023, *Investigación lingüística en entornos digitales*, Tirant Lo Blanch.
- Olmo-Cazevieille, F. et Aubin, S. (Coord.), 2023, *Usages collaboratifs d'outils numériques en didactique de la langue-culture française*, *Synergies Espagne*, n°16.
- Olmo-Cazevieille, F. et Mangiante, J.-M. (Coord.), 2015, *Digital Technologies: Research and Educational Practices in Languages for Specific Purposes*. @tic, revista d'innovació educativa, n°15, <https://ojs.uv.es/index.php/attic/issue/view/N.15>
- Olmo-Cazevieille, F. et Mangiante, J.-M. (eds), 2014, *Lenguas, comunicación y tecnologías digitales. II Coloquio Franco-Español de análisis del discurso y de enseñanza de lenguas para fines específicos*. València : Universitat Politècnica de València.
<https://riunet.upv.es/handle/10251/45936>
- Perinán-Pascual, C. et Mangiante, J.-M. (eds), 2023, *Enseñanza-aprendizaje en entornos digitales*, Tirant Lo Blanch.
- Sabah, Gérard, « Intelligence artificielle, linguistique et cognition », in Catherine Fuchs (dir.), *La Linguistique cognitive*, Paris, Ophrys, 2004, p. 195-221.
- Tyne, H., Bilger, M., Buscali, L., Leray, M., Pérez, C. et Curr, N. (eds), 2022, *A la découverte de la langue : apprentissages et affordances*, Bruxelles : Peter Lang.

Les propositions de communication s'effectueront sous la forme d'un résumé de 2500 signes maximum, bibliographie comprise (1500 signes minimum). Il comportera le prénom, nom de l'auteur ou de l'autrice, le nom de l'organisme et du laboratoire d'appartenance, l'adresse mél, un titre et l'axe dans lequel la communication s'inscrit.

Le cadre théorique, le terrain de recherche, la problématique devront être indiqués, ainsi qu'une courte bibliographie. Le résumé devra être rédigé en police Times New Roman 12, intervalle 1,5, avec paragraphes justifiés à gauche et à droite.

Les propositions seront envoyées **avant le 15 mars 2024**, dernier délai à l'adresse mél :

jmarc.mangiante@univ-artois.fr.

Comité d'Organisation

Jan GOES, Jean-Marc MANGIANTE, Luis MENESES, Martine LAVAUD, Carmen PINEIRA, Mélanie TREDEZ-LOPEZ.

Comité scientifique :

Jan GOES, Jean-Marc MANGIANTE, Luis MENESES, Martine LAVAUD, Carmen PINEIRA, Mélanie TREDEZ-LOPEZ, Françoise OLMO-CAZEVIEILLE, Henry TYNE, Salah MEJRI, Lichao ZHU, Zied BOURAOUI, Céline CORTEEL, Cécile AVEZARD-ROGER, M^a Carmen MOLINA ROMERO, Simona RUGGIA, Anne BESSON, Stéphane PATIN, Denis VIGNERON, Antonio PAMES, Hoai TRAN, Catherine COUTURIER, Marie LEFELLE, Cindy CHARNEAU, Stéphanie LABBE-RIOU, Maria Luisa CARRIÓ PASTOR, Milagros DEL SAZ RUBIO, Mouny MODELAR, Sabrina ROYER, Elodie LANG, Marie BEILLET, Carlos PERINAN-PASCUAL