

Curriculum Vitae

Alexia Schneider

Contact

alexia.schneider@umontreal.ca

[github](#), [gitlab](#)

[ORCID](#)

Formation

2025 - en cours

Doctorat en Humanités Numériques sous la direction de Marcello Vitali-Rosati,
Université de Montréal

Titre: Enrichissement sémantique de la recherche documentaire en SHS par
l'hybridation de systèmes d'IA symbolique et statistique.

2022 - 2024

Maîtrise/Master Technologie des Langues : mention "Traitement Automatique des
Langues" *Université de Strasbourg*

2016 - 2017

Baccalauréat/Licence Professionnelle Métiers du Document en Environnement
Numérique *IUT Robert Schuman, Strasbourg*

2015 - 2016

Maîtrise/Master 1 de Littératures *Université Paris-Sorbonne* en Erasmus à *St. An-*
drews University, Écosse

2014 - 2015

Baccalauréat/Licence de Lettres Modernes *Université Paris-Sorbonne*

2012 - 2014

Classes préparatoires aux grandes écoles : section A/L *Lycée Fustel de Coulanges, Strasbourg*

Stages

Avril - Juillet 2024

Reconnaissance des entités nommées pour l'encodage automatique en XML-TEI de notices bibliographiques d'ouvrages des XVIIe et XVIIIe siècles. *Bibliothèque Nationale et Universitaire, Strasbourg*

- annotation manuelle en XML-TEI d'un sous-corpus pour l'entraînement de modèles
- création et entraînement de modèles de reconnaissance d'EN
- alignement automatique (désambiguïsation) avec le référentiel IdRef

Janvier - Mars 2024

Développement d'un "assistant IA" avec des LLMs génératifs *noota, Paris*

- développement d'un système RAG
- *prompt engineering* et évaluation qualitative des réponses générées
- évaluation des LLM et SLM pour la mise en production

Août - Septembre 2023

Classification automatique du type de didascalie ([Projet Thealtres](#)) *Laboratoire "Linguistique, Langue, Parole" LiLPa UR 1339, Unistra, Strasbourg*

- extraction depuis le French Drama Corpus (TEI) des didascalies via le langage de requête XQuery.
- composition d'une typologie à 13 classes généralisable
- entraînement d'algorithmes d'apprentissage automatique (RandomForest, LogisticRegression, Naive Bayes) avec la bibliothèque Python scikit-learn
- affinage de grands modèles de langue (LLMs) de type BERT
- évaluation de l'impact de la réduction du nombre d'exemples sur la qualité de l'apprentissage des modèles

Mai - Juillet 2023

Annotation manuelle d'un corpus de vaudeville et identification automatique des "airs" ([Projet Thealtres](#)). *Laboratoire "Linguistique, Langue, Parole" LiLPa UR 1339, Unistra, Strasbourg*

- annotation manuelle et mise en accord d'une prosopographie avec une autre annotatrice
- développement d'un programme pour l'identification de titres d'"airs" à base de règles ("regex" ou motifs dans la chaîne de caractères).
- alignement des titres avec un référentiel à partir d'une recherche de similarité soit sur une mesure de distance d'édition (Levenshtein) soit une mesure de distance (cosinus) à partir de représentation du contenu avec des grands modèles de langue de type Transformers
- intégration du projet sur la plateforme de gestion de code et de projet informatique GitLab

Publications et communications scientifiques

Schneider, A. (2025, mai 30). Hybridation de l'IA et remise en question de l'utilisation générale et généraliste de l'IA. Zenodo. <https://zenodo.org/doi/10.5281/zenodo.15641286>

Schneider, A., & Vitali-Rosati, M. (2025, mai 8). Atelier IA et revues scientifiques—Explorer l'automatisation de la recherche d'évaluateur · ice · s [Atelier]. États généraux du réseau Circé, Montréal. <https://reseauircce.org/evenements/etats-generaux-des-revues-scientifiques-du-quebec/>

Ruiz Fabo, P., Boissard, F., & Schneider, A. (2024b, octobre). Paysage social représenté dans le théâtre populaire en alsacien, allemand et français (1800-1929). Faculté des langues, Université de Strasbourg. <https://hal.science/hal-04869157>

Schneider, A., & Wingert, R. (2024, octobre 17). Reconnaissance d'Entités Nommées pour des données bibliographiques. Humanités numériques en pédagogie et en recherche à la faculté des langues 2e ed. <https://doi.org/10.34847/nkl.eb726q1s>

Ruiz Fabo, P., Boissard, F., & Schneider, A. (2024a, août). Comparative metadata for popular theater in Alsatian, German and French. ADHO - Alliance of Digital Humanities Organizations. <https://zenodo.org/records/11306738>

Schneider, A., & Wingert, R. (2024, juin 3). Le projet Ritter augmenté: Reconnaissance d'Entités Nommées pour les données bibliographiques. Rencontre régionale IA4LAM « Intelligence Artificielle, patrimoine et Humanités Numériques », BNU, Strasbourg.

Schneider, A. (2024). Analyse computationnelle de textes de théâtre français: Classification automatique du type de didascalies [Mémoire de Master]. Université de Strasbourg.

Schneider, A., & Ruiz, P. (2024). Stage Direction Classification in French Theater: Transfer Learning Experiments. Proceedings of LaTeCH-CLfL 2024, 278-286. St. Julians, Malta. Association for Computational Linguistics. <https://aclanthology.org/2024.latechclfl-1.28/>

Schneider, A. (2023, novembre 29). Classification automatique du type de didascalie. Journée d'étude: « Données de théâtre, petit et grand : Le bon, le brut et le futé », MISHA, Strasbourg. <https://doi.org/10.5281/zenodo.10644192>

Schneider, A., & Nugues, L. (2023, novembre 29). Détecter automatiquement des airs de vaudeville. Journée d'étude: 'Données de théâtre, petit et grand : Le bon, le brut et le futé', MISHA, Strasbourg. <https://doi.org/10.5281/zenodo.10702988>

Programmes et modèles entraînés

Schneider, A. (2025). Programme : Atelier IA et revues scientifiques—Explorer l'automatisation de la recherche d'évaluateur · ice · s des États généraux du réseau Circé [Jupyter Notebook]. <https://colab.research.google.com/drive/1OyrkTWUCmOBv58ALNOvf6oE5jp03h3BW#scrollTo=qFFu-kVKqq2w>

Schneider, A. (2024). Lab-bnu/ritter_ner [Jupyter Notebook]. lab-bnu. https://github.com/lab-bnu/ritter_ner (Édition originale 2024)

Schneider, A. (2024). Fr_camembert_ritter (Revision 4819b24). Hugging Face. <https://doi.org/10.57967/hf/2721>

Schneider, A., & Ruiz Fabo, P. (2024). Code and data for paper «Stage Direction Classification in French Theater : Transfer Learning Experiments» (Version v1.0) [Python]. Zenodo. <https://doi.org/10.5281/zenodo.10594105>

Langues

Français: langue maternelle

Anglais: deuxième langue (940/1000 TOEIC)

Espagnol: vécu et travaillé au Mexique et en Colombie pendant un an.