

Nicolas Gutehrlé

Docteur en TAL

36 boulevard Eugène Spuller
Dijon, 21000

☎ 07 86 13 07 18

✉ nicolas.gutehrle@gmail.com

📄 <https://nicolasgutehrle.github.io>



Formations

2020–2024 **Doctorat en Traitement Automatique des Langues**, Université de Franche-Comté, Besançon, France, Centre de Recherches Interdisciplinaires et Transculturelles (C.R.I.T.), financé par le projet EMONTAL, région Bourgogne Franche-Comté (2020-2023).

Sujet de thèse :

"Information extraction from unstructured documents for the valorisation of historical periodicals. Application to the heritage of the Bourgogne Franche-Comté Region in France" sous la direction de Dr. Iana Atanassova

2016–2018 **Master Langues, Littératures, et Civilisations Etrangères et Régionales (LL-CER) - parcours Traitement Automatique des Langues (TAL)**, Université de Franche-Comté, Besançon, France, Mention Très Bien.

— **Sujet du mémoire** : "Langue contrôlée pour un système de messages et alertes dans un environnement de mobilité : gestion de l'ambiguïté phonologique" sous la direction de Pr. Sylviane Cardey et Dr. Iana Atanassova

— **Stage de fin d'études (6 mois)** : "Perception des villes et objets intelligents" : sous la direction de Pr. Isam Shahrour et Dr. Natalia Grabar, Université de Lille, France

2013–2016 **Licence Langues, Littératures et Civilisations Etrangère (LLCE) spécialité Anglais**, Université de Franche-Comté, Besançon, France, Mention Bien.

— Troisième année d'études effectuée en Erasmus à l'université de Galway (Irlande)

Expériences

Professionnelles

Juillet 2024 – Actuel **Post-doctorat, projet "ANR InScim"**, Centre de Recherches Interdisciplinaires et Transculturelles (C.R.I.T.), Université de France-Comté, France.

Février 2024 – Juin 2024 **Ingénieur d'études, projet "ANR InScim"**, Centre de Recherches Interdisciplinaires et Transculturelles (C.R.I.T.), Université de France-Comté, France.

Septembre – Décembre 2023 **Secrétariat**, Centre de Recherches Interdisciplinaires et Transculturelles (C.R.I.T.), Université de France-Comté, France.

— Gestion administrative et communication

— Logistique et aide à l'organisation d'événements scientifiques

— Gestion budgétaire et approvisionnement

- Avril – Juillet 2022 **Stage de recherche (3 mois)**, *Digital Science Center*, Université d'Innsbruck, Autriche.
 sous la direction du professeur **Adam Jatowt**
- Travail de recherche dans un milieu interdisciplinaire et international
 - Initiation de multiples collaborations internationales (voir Publications)
- Août 2019 – **Data Analyst**, *OnlineFormaPro*, Vesoul, France.
 Juillet 2020
- Recherche et Développement en Deep Learning
 - Intégration de la norme xAPI
 - Recherches en chatbots et synthèse vocale
- Mars 2019 – **Vacataire, projet "AME"**, *Université de Franche-Comté*, Besançon, France,
 Mai 2019 MSHE C.N. Ledoux.
- Numérisation et indexation des archives de l'auteur Jean-Paul Goux (photographies, transcriptions d'entretiens, notes manuscrites, cassettes audio)
- Septembre 2018 – **Ingénieur d'études, projet "ORTEP Revitalisation"**, *Université de Franche-Comté*, Besançon, France, MSHE C.N. Ledoux.
 2019
- Numérisation et analyse automatisée d'archives de journaux locaux publiés entre 1840 et 1939
 - Création d'une interface web de reconnaissance, d'extraction et de géo-résolution d'entités nommées à partir de petites annonces
 - Présentation des résultats de recherches à la journée d'études "*HumaSpatia*" à la MSH Dijon

Autres expériences

- Mai 2021 **Helsinki Digital Humanities Hackathon 2021**, En ligne.
 Projet "*Space Wars*"
- Mars 2020 **Collaborathon COVID-19**, En ligne.
 Projet "*Inventaire des masques*"
- Octobre 2019 **Hacking Health 2019**, Besançon, France.
 Projet "*Pic'Peaucket*" (gestion et analyse de photographies en dermatologie)

Enseignements

- 2021, 2022, 2024 **Rédaction de documents scientifiques avec LaTeX**, 12h au total, URFIST.
 Lyon, France
- Septembre 2020 – Actuel **Chargé de cours en master LLCER parcours Traitement Automatique des Langues**, *Université de Franche-Comté*, Besançon, France, Centre de Recherches Interdisciplinaires et Transculturelles (C.R.I.T.).
- Modélisation phonétique (TD et CM)
 - Approches cognitives (TD)
 - Techniques de programmation 2 (TD)
 - Atelier d'écriture scientifique (TD)
 - Génie Logiciel (TD)
- Septembre – **Chargé de cours en master LLCER parcours Traitement Automatique des Langues**, *Université de Franche-Comté*, Besançon, France, Centre de Recherches Interdisciplinaires et Transculturelles (C.R.I.T.).
 Décembre 2018
- Modélisation phonétique (TD et CM)

Encadrements

Mémoires de master

- Bruno de Brito** : *Extraction de thématiques (topic modelling) dans la presse française du 20ème siècle*, 2021-2023, encadré par Dr. Iana Atanassova
- Lilian Rodrigues de Almeida** : *Détection automatique de dialectes avec un système hybride segmental basé sur des traits de phonétique acoustique : une étude de cas différenciant le français breton du français standard*, 2022-2023, encadré par Dr. Iana Atanassova
- Maya Mathié** : *Exploration des paroles de chansons en français pour la détection automatique de contenus explicites*, 2022-2024, encadré par Dr. Iana Atanassova
- Jacob Tadei** : *Evaluer le niveau d'un joueur de Counter-Strike par sa maîtrise langagière*, 2022-2024, encadré par Dr. Iana Atanassova
- Fany Montay** : *Analyse morphologique des langues agglutinantes par machine-learning : application au finnois*, 2022-, encadré par Dr. Iana Atanassova, co-encadré avec Yağmur Öztürk

Stages

- Bruno de Brito** : *Traitement automatique d'une base bibliographique*, 01/02/23 - 31/03/23, encadré par Dr. Iana Atanassova

Engagements

Evaluation d'articles

- Evalueur d'articles pour *Journal on Computing and Cultural Heritage* (2021-2024), *NLP4DH* (2022-2024) et *ICTIR* (2023)

Organisation de séminaires

- Co-organisateur des séminaires en TAL du C.R.I.T.
- Co-organisateur des journées d'études TEDonnées
- Co-organisateur des séminaires doctoraux du C.R.I.T.

Événements de vulgarisation scientifique

- Participation au festival de vulgarisation scientifique OVNI à Arbois (18/11/2023) : <https://actu.univ-fcomte.fr/agenda/ovni-debarque-arbois>
- Organisateur du stand du laboratoire C.R.I.T. à la *Fête de la Science* (14/10/2023) : <https://actu.univ-fcomte.fr/agenda/besancon-fete-de-la-science-2023>
- Participation à la *Nuit des Chercheurs* (29/09/2023) : <https://nuitdeschercheurs-france.eu/?2023>
- Participation au *Speed Searching* en ligne dans le cadre de la *Nuit des Chercheurs* (12/09/2023) : <https://nuitdeschercheurs-france.eu/?2023>
- Participation au festival de vulgarisation scientifique OVNI à Morteau (01/04/2023) : <https://actu.univ-fcomte.fr/agenda/ovni-debarque-morteau#.ZCMNOBXP23I>
- Participation au *Speed Searching* en ligne dans le cadre de la *Nuit des Chercheurs* (20/09/2022) : <https://nuitdeschercheurs-france.eu/?2022>
- Participation à la formation *Expérimantium* à Besançon (Décembre 2021) : <https://www.experimentarium.fr/regions/franche-comte>

Articles de vulgarisation scientifique

- *Where did it happen ? : Spatial imaginaries of World War I* Axel Jean-Caurant, Jani Marjanen, Lidia Pivovarova, Nicolas Gutehrlé, Oleg Harlamov, Farimah Karimi, Suphan Kirmizialtin, David Mayer, Maelle Le Roux, and Haoyu Wei. 2021. : <https://tinyurl.com/5bhnykna>
- *Donner un nouveau sens aux documents historiques*, journal En Direct, n° 312, 2024 : <https://tinyurl.com/3p9ytdrw>
- *Comprendre les archives : vers de nouvelles interfaces de recherche reposant sur l'annotation sémantique des documents*, 4ème Printemps de la donnée en BFC : <https://tinyurl.com/8as576vf>

Autres engagements

- Représentants des doctorants du C.R.I.T. (2023-Actuel)
- Référent Médiation Scientifique du C.R.I.T. (2023-Actuel)
- Responsable de la communication et de la gestion du site web du département Traitement Automatique des Langues (TAL) du C.R.I.T. : communication, gestion et mise à jour de l'emploi du temps, réservation des salles, annonces des séminaires, mise à jour des publications scientifiques du département : <http://tesniere.univ-fcomte.fr/>
- Responsable de la maintenance des appareils du département TAL (ordinateurs, imprimantes, projecteur)

Compétences

Français	Maternelle	Programmation	Python, Java
Anglais	C2	Librairies	spaCy, pandas, NLTK, numpy, scikit-learn, Keras, HuggingFace
Espagnol	B1		Streamlit, Panel, Flask, itertools, networkx
Allemand	A1	Web	HTML, CSS, JavaScript
Irlandais	Notions	Bases de données	SQL, MongoDB, BaseX, Solr, Neo4J
		Langages à balisage	LaTeX, XML (TEI, ALTO)
		IDE	Visual Code, PyCharm, Jupyter
		Autres logiciels	Git, TXM

Centre d'intérêts

- Cuisine
- Apprentissage de langues étrangères
- Composition musical
- Sport, écriture
- Lecture, cinéma, jeux vidéos, jeux de rôles