

Nicolas Gutehrle

Ph.D. in Natural Language Processing

36 boulevard Eugène Spuller
Dijon, 21000

☎ 07 86 13 07 18

✉ nicolas.gutehrle@gmail.com

🌐 <https://nicolasgutehrle.github.io>



Formations

- 2020 – 2024 **Ph.D. in Natural Language Processing**, *Université de Franche-Comté*, Besançon, France, Centre de Recherches Interdisciplinaires et Transculturelles (C.R.I.T.).
PhD thesis title :
"Information extraction from unstructured documents for the valorisation of historical periodicals. Application to the heritage of the Bourgogne Franche-Comté Region in France" under the supervision of Dr. Iana Atanassova
- 2016–2018 **MSc "Languages and Foreign Cultures" (LLCER), specialty Natural Language Processing**, *Université de Franche-Comté*, Besançon, France, With Distinction.
— MSc thesis title : *"Langue contrôlée pour un système de messages et alertes dans un environnement de mobilité : gestion de l'ambiguïté phonologique"* under the supervision of Pr. Sylviane Cardey and Dr. Iana Atanassova
— End-of-studies internship : *"Perception of smart objects and cities"* under the supervision of Pr. Isam Shahrour et Dr. Natalia Grabar at *Université de Lille*, France
- 2013–2016 **BSc "Languages and Foreign Cultures", specialty English and Natural Language Processing**, *Université de Franche-Comté*, Besançon, France, With honours.
— Third year of study spent abroad at *National University of Ireland Galway* with the Erasmus program

Experiences

Work experiences

- July 2024 – Current **Post-doctoral researcher in the "ANR InScim" project**, *Centre de Recherches Interdisciplinaires et Transculturelles (C.R.I.T.)*, Université de France-Comté, France.
- February 2024 – June 2024 **Study engineer in the "ANR InScim" project**, *Centre de Recherches Interdisciplinaires et Transculturelles (C.R.I.T.)*, Université de France-Comté, France.
- September – December 2023 **Secretary**, *Centre de Recherches Interdisciplinaires et Transculturelles (C.R.I.T.)*, Université de France-Comté, France.
— Administrative management and communication
— Logistics and help with organising scientific events
— Budget management and procurement

- April – July 2022 **Research Internship**, *Digital Science Center*, University of Innsbruck, Austria.
 Under the supervision of Professor **Adam Jatowt**
- Research work in an interdisciplinary and international environment
 - Creation of new international collaborations
 - Discovery of the culture of a foreign country
- August 2019 – July 2020 **Data Analyst**, *OnlineFormaPro*, Vesoul, France.
- Research and Development in Deep Learning
 - Integration of the xAPI norm
 - Research in developing chatbots and speech synthesis
- March 2019 – May 2019 **Temporary worker in "AME" project**, *Université de Franche-Comté*, Besançon, France, MSHE C.N. Ledoux.
- Digitizing and indexing archive collections from French author Jean-Paul Goux (photographs, interview transcriptions, handwritten notes, audio tapes)
- September 2018 – March 2019 **Study engineer in the "ORTEP Revitalisation" project**, *Université de Franche-Comté*, Besançon, France, MSHE C.N. Ledoux.
- Digitization and automatic analyses of local newspapers archives published between 1840 and 1939
 - Development of a web interface for the recognition, extraction and geo-resolution of Named Entities in newspaper ads
 - Presentation of a research poster at the "*HumaSpacia*" workshop in MSH Dijon

Other experiences

- May 2021 **Helsinki Digital Humanities Hackathon 2021**, Online.
 "*Space Wars*" project
- March 2020 **Collaborathon COVID-19**, Online.
 "*Inventaire des masques*" project
- October 2019 **Hacking Health 2019**, Besançon, France.
 "*Pic'Peaucket*" project (management and analyses of photographs in dermatology)

Teaching Activities

- 2021, 2022, 2024 **Initiation to document redaction with LaTeX**, *URFIST*, Lyon, France.
- September 2020 – Current **Teaching assistant in MSc "Languages and Foreign Cultures" (LLCER), specialty Natural Language Processing**, *Université de Franche-Comté*, Besançon, France, Centre de Recherches Interdisciplinaires et Transculturelles (CRIT).
- Phonetic modelling (tutorial and lecture classes)
 - Cognitive methods and Machine Learning (tutorial classes)
 - Software engineering for NLP (tutorial classes)
 - Scientific writing workshop (tutorial classes)
 - Advanced Programming (tutorial classes)
- September – December 2018 **Teaching assistant in MSc "Languages and Foreign Cultures" (LLCER), specialty Natural Language Processing**, *Université de Franche-Comté*, Besançon, France, Centre de Recherches Interdisciplinaires et Transculturelles (CRIT).
- Phonetic modelling (lectures and tutorial classes)

Supervision

Master thesis

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, supervised by Dr. Iana Atanassova, co-supervised with Aurélie Nombrot

Maya Mathié : *Exploration des paroles de chansons en français pour la détection automatique de contenus explicites*, 2022-2024, supervised by Dr. Iana Atanassova

Jacob Tadei : *Evaluer le niveau d'un joueur de Counter-Strike par sa maîtrise langagière*, 2022-2024, supervised by Dr. Iana Atanassova

Fany Montay : *Analyse morphologique des langues agglutinantes par machine-learning : application au finnois*, 2022-, supervised by Dr. Iana Atanassova, co-supervised by Yağmur Öztürk

Internships

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

Engagements

Article reviewing

— Reviewer for *Journal on Computing and Cultural Heritage* (2021-2023), *NLP4DH* (2022, 2023) and *ICTIR* (2023)

Seminar organisation

- Co-organiser of the NLP seminars of C.R.I.T.
- Co-organiser of the *TEDonnées* study day
- Co-organiser of the doctoral seminars of C.R.I.T.

Popularisation of science events

- Participation to the OVNI science popularisation festival in Arbois (18/11/2023) : <https://actu.univ-fcomte.fr/agenda/ovni-debarque-arbois>
- Organiser of the C.R.I.T. laboratory stand at the *Fête de la Science* event (14/10/2023) : <https://actu.univ-fcomte.fr/agenda/besancon-fete-de-la-science-2023>
- Participation to the *Nuit des Chercheurs* event (29/09/2023) : <https://nuitdeschercheurs-france.eu/?2023>
- Participation to the online *Speed Searching* event (12/09/2023) : <https://nuitdeschercheurs-france.eu/?2023>
- Participation to the OVNI science popularisation festival in Morteau (01/04/2023) : <https://actu.univ-fcomte.fr/agenda/ovni-debarque-morteau#.ZCMNOBXP23I>
- Participation to the online *Speed Searching* event (20/09/2022) : <https://nuitdeschercheurs-france.eu/?2022>
- Participation to the *Expérimentarium* formations in Besançon (December 2021) : <https://www.experimentarium.fr/regions/franche-comte>

Popularisation of science articles

- *Where did it happen ? : Spatial imaginaries of World War I* Axel Jean-Caurant, Jani Marjanen, Lidia Pivovarova, Nicolas Gutehr , 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 s mantique des documents*, 4 me Printemps de la donn e en BFC : <https://tinyurl.com/8as576vf>

Other involvements

- Doctoral PhD student representatives of C.R.I.T. (2023-Current)
- Scientific Mediation Coordinator for C.R.I.T. (2023-Current)
- In charge of communication and website management for the Natural Language Processing department at C.R.I.T. : communication, management and updating of the timetable, room bookings, seminar announcements, updating of the department's scientific publications : <http://tesniere.univ-fcomte.fr/>
- Responsible for maintaining the NLP department's equipment (computers, printers, projectors)

Skills

Franais	Native	Programming	Python, Java
English	C2	Libraries	spaCy, pandas, NLTK, numpy, scikit-learn, Keras, HuggingFace
Spanish	B1		Streamlit, Panel, Flask, itertools, networkx
German	A1	Web	HTML, CSS, JavaScript
Irish	Notions	Databases	SQL, MongoDB, BaseX, Solr, Neo4J
		Markup languages	L ^A T _E X, XML (TEI, ALTO)
		IDE	Visual Code, PyCharm, Jupyter
		Other	Git, TXM

Hobbies

- Cooking
- Music
- Reading, cinema, video games, role playing games
- Learning languages
- Sport, writing