





ENZO DOYEN

PhD Candidate in Natural Language Processing

✉ contact@edoyen.com  edoyen.com  linkedin.com/in/enzo-doyen  github.com/spidersouris
 orcid.org/0009-0005-4742-5860

EDUCATION

PhD in Natural Language Processing

Strasbourg, France


 University of Strasbourg, LiLPa

Oct 2024 – Oct 2027

- **Thesis:** Mitigating Gender Biases in NLP Data for French
- **Supervisor:** Amalia Todirascu

Master's in Natural Language Processing

Strasbourg, France

 University of Strasbourg


Sep 2022 – Jun 2024

- **Grade:** 18.125/20 (GPA equiv.: 4.0). Highest Honors. Rank: 1/20.
- **Covered Topics:** DSA; OOP; data analysis; machine learning; stats; Python; Java; JavaScript; PHP; SQL; R; RegExp; Bash
- **Thesis:** Automatic Gender-Neutral Rewriting Using Human Collective Nouns in French

WORK EXPERIENCE

Teaching Assistant

On-site / Strasbourg, France

 University of Strasbourg

Sep 2024 – Present

Software Engineer Intern

Hybrid / Strasbourg, France


 University of Strasbourg, LiLPa

Feb 2024 – Apr 2024

- Engineered a web app (<https://thealtres.pages.unistra.fr/vis/>) for the Thealtres project to visualize theater data, allowing users to search, filter, sort, and analyze trends across ≈1,000 plays and 10,000+ characters
- Conceptualized and executed intuitive and appealing UI/UX design
- Created customized, interactive data visualizations using D3.js
- Created a custom Leaflet map leveraging the data

Bilingual Localizer (EN → FR)

Remote / Seattle, USA

 Valve Corporation

Nov 2017 – Present

- Translating Steam, the world's biggest video game digital distribution service (≈132 million monthly active users (MAU)); ≈3 million French MAUs) and major video games and hardware (Counter-Strike 2, Dota 2, Deadlock; Steam Deck, Valve Index).

FEATURED PROGRAMMING PROJECTS

touying-unistra-pristine | spidersouris/touying-unistra-pristine ★ 5

Aug 2024


- Created a custom presentation slides theme for the University of Strasbourg in Typst using Touying


termic | <https://termic.me/> | spidersouris/termic ★ 21


May 2023

- Designed and developed a full-stack web app replacement/alternative to Microsoft Terminology Search used by thousands of translators and terminologists around the world
- Developed open-source Python scripts to process and clean Microsoft-provided terminology files, and to extract terminologic data from VSCode

RESEARCH AND PUBLICATIONS

Doyen, E. & Todirascu, A. (2025). Man Made Language Models? Evaluating LLMs' Perpetuation of Masculine Generics Bias.  [10.48550/arXiv.2502.10577](https://arxiv.org/abs/2502.10577)

Doyen, E. & Ruiz Fabo, P. (2024). Thealtres visualization application (v1.0.0).  [10.5281/zenodo.11289443](https://zenodo.org/record/11289443)

Doyen, E. (2024). Neutralisation automatique du genre à travers l'utilisation de noms collectifs en français. University of Strasbourg. 

TALKS AND PRESENTATIONS

| | |
|---|--------------------|
| Analyzing use of masculine generics by LLMs in French | Aussois, France |
| Poster Session: ALPS 2025 (Doyen, E.) | Apr 2025 |
| Éviter les biais de genre dans les outils de TAL (Avoiding Gender Biases in NLP Tools) | Strasbourg, France |
| Talk: Séminaire LiRiC — Session 2 : inclusion et écriture (Doyen, E.) | Jan 2025 |
| Navigation structurée de métadonnées de théâtre (Structured Navigation of Theater Metadata) | Strasbourg, France |
| Poster Session: Humanités numériques en pédagogie et en recherche 2 (Doyen, E. & Ruiz Fabo, P.) | Oct 2024 |

HONORS AND AWARDS

| | |
|--|--------------------|
| Doctoral Fellowship | Strasbourg, France |
| Issued by the University of Strasbourg | Jun 2024 |
| • Selected as one of 8 recipients out of 18 applicants for a 3-year fellowship (≈€81,000) to support personal PhD research | |

SCHOOLS

| | |
|--|------------------|
| ALPS 2025 | Aussois, France |
| LIG (Univ. Grenoble Alpes) Naver Labs Europe Cohere | Apr 2025 |
| • Participated to a 28-hour program comprised of lectures, project sessions and poster sessions with world-class NLP researchers | |
| Lexicon Summer School | Almuñécar, Spain |
| University of Granada | Aug 2023 |
| • Participated to a 30-hour program of seminars and workshops in terminology and artificial intelligence with global experts | |

TEACHING RESPONSIBILITIES

| Course Name | Institution | Level | Dates |
|--|---|-------------------------------|-------------------------|
| ◊ Web, corpus et traduction (Web, corpus and translation) | Institut européen des métiers de la traduction (IEMT); University of Strasbourg | Master's (1st year) | 2024 – 2025 (S1) |
| ◊ Conception de pages Web (HTML/CSS) (Introduction to HTML/CSS) | University of Strasbourg | Bachelor's (2nd-3rd years) | 2024 – 2025 (S1, S2) |
| ⊙ Linguistique de corpus (Corpus linguistics) | University of Strasbourg | Master's (1st year) | 2024 – 2025 (S1) |
| ⊙ Analyse de corpus (Corpus analysis) | University of Strasbourg | Master's (2nd year) | 2024 – 2025 (S1) |

◊ Full responsibility | ⊙ Partial responsibility

SKILLS

- **Programming:** Python, TypeScript/JavaScript, SQL (Postgres), R, RegExp, LaTeX, Typst
- **AI/ML/DS/NLP:** PyTorch, scikit-learn, Keras, Pandas, NumPy, Matplotlib, seaborn, BeautifulSoup, Selenium, NLTK, spaCy
- **Web Development:** Next.js, Astro, Tailwind CSS, Node.js, Flask, Hugo
- **Tools:** Git/GitHub, Docker, Microsoft Azure, VSCode, Jupyter Notebook
- **Soft Skills:** Teamwork, team management, project management, problem solving, critical thinking, adaptability, time mangement, public speaking, technology trend awareness
- **Languages:** French (native), English (C2), Spanish (B2), Japanese (A1)