

# Mathilde HUGUIN

## Adresse professionnelle

Inist-CNRS, 2 rue Jean Zay  
54519 Vandœuvre-lès-Nancy



Page web : [mathildehuguin.fr](http://mathildehuguin.fr)

Email : [mathilde.huguin@inist.fr](mailto:mathilde.huguin@inist.fr)

Née le 6 mars 1990

## SITUATION ACTUELLE

---

- Depuis déc. 2023 **Ingénieure de recherche**, [Inist-CNRS](#) (UAR 76), Vandœuvre-lès-Nancy  
Fonction : responsable du traitement et de l'analyse des données [Istex](#).  
Mission : promouvoir l'offre de services Istex auprès de la communauté scientifique française.
- Depuis fév. 2022 **Qualifiée aux fonctions de Maître de Conférences** par le CNU pour la section 7 (n°22207321420)
- Depuis déc. 2021 **Jeune docteure**, membre associée au laboratoire [ATILF](#) (UMR 7118) et à l'Université de Lorraine

## EXPÉRIENCES PROFESSIONNELLES

---

- Sept. 2022–Nov. 2023 **Ingénieure de recherche (CDD)**, [Inist-CNRS](#) (UAR 76), Vandœuvre-lès-Nancy  
Fonction : responsable du traitement et de l'analyse des données [Istex](#).  
Mission : promouvoir l'offre de services Istex auprès de la communauté scientifique française.
- Sept. 2022–Déc. 2022 **Vacataire** pour l'Université de Lorraine (cumul d'activités), Dépt. Sciences du langage, Nancy  
Mission : enseignement de méthodologie de la recherche (masters, 12h/an).
- Oct. 2021–Août 2022 **Ingénieure (CDD)**, projet [Demonext](#) ANR-17-CE23-0005, [ATILF](#) (UMR 7118), Nancy  
Mission : création d'une ontologie, analyse et annotation sémantique d'une base de données linguistiques.
- Sept. 2018–Août 2020 **ATER** à l'Université de Lille (CDD), laboratoire [STL](#) (UMR 8163), Villeneuve-d'Ascq  
Mission : enseignement de morphologie, sémantique et préparation aux concours (licences, 192h/an).
- Sept. 2015–Août 2018 **Doctorante Contractuelle Chargée d'enseignement** à l'Université de Lorraine (CDD), Nancy  
Mission : enseignement de grammaire, morphologie, syntaxe et lexique (licences, 64h/an).
- 2014-2015 **Vacataire** pour la revue linguistique [Verbum](#), [ATILF](#) (UMR 7118), Nancy  
Mission : édition, relecture et mise en page des numéros XXXIV n°2 et XXXVII n°2.
- 2014 **Vacataire** dans le projet [ORFEO](#) ANR 12-CORP-0005, [ATILF](#) (UMR 7118), Nancy  
Mission : élaboration (transcription et alignement texte/son) d'un corpus de français parlé.
- 2013 **Stagiaire** dans le projet [SEQUOIA](#) ANR 08-EMER-013, [LORIA](#) (UMR 7503), Villers-lès-Nancy  
Mission : annotations syntaxiques manuelles d'un corpus, correction d'une analyse syntaxique automatique.

## FORMATION UNIVERSITAIRE

---

- 2015-2021 **Doctorat** en Sciences du Langage, [ATILF](#) (UMR 7118), Université de Lorraine, Nancy  
TITRE : "[Analyse morphologique des mots construits sur base de noms de personnalités politiques](#)"  
Description : analyse linguistique de mots issus de noms de personnalités politiques.
- 2015 **Master** en Sciences du Langage, option TAL, Université de Lorraine, Nancy
- 2013 **Licence** en Sciences du Langage, Université de Lorraine, Nancy

## FORMATION CONTINUE

---

- Déc. 2023 **Lodash**, formation interne Inist-CNRS  
Objectif : initiation à la bibliothèque Lodash au travers d'exemples concrets.
- Nov. 2023 **Atelier terminologie, thésaurus et ontologies**, formation interne Inist-CNRS  
Objectif : introduction aux notions de bases (terminologie, concept, ontologie computationnelle) et présentation d'outils.
- Sept. 2023 **Linux et shell**, formation interne Inist-CNRS  
Objectif : présentation de Linux, principes et commandes de base, initiation à la programmation de scripts shell.
- Mai 2023 **Json**, formation interne Inist-CNRS  
Objectif : présentation du format, historique et comparaison avec d'autres formats (XML).
- Avr. 2023 **OpenRefine**, formation interne Inist-CNRS  
Objectif : curation et formatage de données, alignement avec une Wikibase.

- Sept. 2022 **Mise en place de la gestion Agile au sein d'une équipe**, formation OXiane  
Objectif : gérer efficacement ses projets en appliquant un framework Agile.
- Sept. 2022 **Nouvelles méthodes pour la fouille et l'analyse des corpus textuels**, école thématique CNRS  
Objectif : maîtriser des outils d'analyse de données (R/R Studio, Python, GarganText, TXM, Hyperbase).
- Juin 2022 **Python par la pratique**, CNRS - M2i  
Objectif : connaître les usages courants du langage.
- Avr. 2021 **Culture de l'intégrité scientifique**, ED SLTC  
Objectif : sensibilisation aux enjeux épistémiques, éthiques et méthodologiques de la recherche.
- Déc. 2016 **Innovations pédagogiques et outils numériques**, ED SLTC  
Objectif : préparer, animer et évaluer un cours en utilisant les outils numériques.
- Nov. 2016 **Quantifier en SHS : méthodologie et traitement de données**, ED SLTC  
Objectif : utiliser des options de calculs avancées dans les tableurs (Excel), introduction à R.

## PUBLICATIONS

---

- À paraître **Typage sémantique des noms en français pour la ressource lexicale morphologique Démonette-2.0.** *Lexique*.
- 2023 **Le corpus « Machine Translation » : une exploration diachronique des (méta)données Istex.** F. Boudin et al. (éds), Actes ARTS, CORIA TALN RJCRI RECITAL 2023, 5-9 juin 2023, 54-59.
- 2023 **MaTOS : traduction automatique pour la science ouverte.** F. Boudin et al. (éds), Actes ARTS, CORIA TALN RJCRI RECITAL 2023, 5-9 juin 2023, 8-15.
- 2022 **The MoNoPoli Database: Extragrammatical and Subverted Processes in French Words Based on Proper Names of Politicians.** *The Prague Bulletin of Mathematical Linguistics* (PBML) 119, 93-119.
- 2022 **Analyse morphologique des mots construits sur base de noms de personnalités politiques.** *L'Information Grammaticale* 174, 56-58.
- 2021 **The MoNoPoli database. Or how to catch Macron-itis.** F. Namer et al. (eds), *Proceedings of DERIMO*, 9-10 septembre 2021, 76-85.
- 2018 **Anthroponyms and paradigmatic derivation in French.** N. Hathout & F. Namer (eds), *Defining paradigms in word formation, Lingue e Linguaggio XVII* (2), 217-232.
- 2018 **Anthroponymes et création lexicale. Ou pourquoi Moranono s'éloigne du canon dérivationnel.** V. Balnat & C. Gérard (éds), *Néologie et noms propres, Cahiers de Lexicologie* 113 (2), 61-77.
- 2018 **Morphologie et organisation du lexique : la catégorie nominale, comparaison des déonomastiques et des dénominatifs.** F. Neveu et al. (éds), *Actes du 6ème CMLF*, 9-13 juillet 2018, Mons.
- 2018 **Les Noms d'Humains comme base de construction morphologique. Lexèmes construits sur noms propres de personnalités politiques françaises.** C. Schnedecker & W. Mihatsch (éds), *Les noms d'humains – théorie, méthodologie, classification*, 163-184, De Gruyter.

## PROJETS ET PARTENARIATS

---

- Depuis 2023 **Membre du projet de recherche Machine Translation for Open Science ANR-22-CE23-0033**  
Objectif : développer de nouvelles méthodes pour la traduction automatique de documents scientifiques.
- 2021-2022 **Membre du projet de recherche Demonext, Dérivation en extension ANR-17-CE23-0005**  
Objectif : construire une base de données qui décrit les propriétés des mots de manière systématique.
- 2016-2021 **Création du partenariat CORMORAN, CORpus web de mots MORphologiquement construits sur ANthroponymes, entre l'ATILF et Data Observer**  
Objectif : élaborer une méthodologie de collecte et de veille pour construire un corpus web.
- 2014-2018 **Membre du projet de recherche international NHUMA, Les noms d'humains : de la description linguistique aux applications lexicographiques**  
Objectif : analyser les noms qui dénotent des humains.

## CORPUS ET GUIDE

---

- Oct. 2023 **Corpus Machine Translation V2.** [Istex - data.istex.fr](#), corpus spécialisé  
Objectif : retracer l'aspect diachronique de la traduction automatique en augmentant la couverture chronologique des ressources du corpus.
- Mai 2023 **Corpus Machine Translation.** [Istex - data.istex.fr](#), corpus spécialisé  
Objectif : retracer l'aspect diachronique de la traduction automatique.
- Déc. 2022 **Corpus Nom propre & Morphologie.** [Lodex - lodex.inist.fr](#), corpus spécialisé  
Objectif : modéliser les prémisses d'une recherche doctorale grâce à la fouille de documents Istex.
- Nov. 2022 **Corpus Végane & Végétarien.** [Istex - data.istex.fr](#), corpus d'actualité  
Objectif : examiner l'évolution des aspirations des personnes véganes et végétariennes dans le temps.
- Juin 2022 **Guide d'annotation sémantique de données morphologiques (livrable 1.1 - ANR Demonext)**  
Objectif : permettre l'annotation sémantique de lexèmes nominaux hors contexte.
- Déc. 2021 **Corpus MoNoPOLI (Mots construits sur Noms propres de personnalités Politiques)**  
Description : MoNoPOLI, disponible sur [Ortolang](#), contient 6 500 dérivés de noms propres soit 55 000 contextes.

## RESPONSABILITÉS ADMINISTRATIVES ET PÉDAGOGIQUES

---

- 2023 **Membre du comité d'organisation du colloque ISMo**, ATILF & Université de Lorraine  
*International Symposium of Morphology.*
- Depuis sept. 2022 **Membre du réseau d'ingénierie documentaire de l'Inist-CNRS**  
Objectif : partager les savoir-faire et les compétences entre les équipes de l'unité.
- 2021 **Membre du comité d'organisation du colloque DeriMo**, ATILF & Université de Lorraine  
*Third International Workshop on Resources and Tools for Derivational Morphology.*
- Depuis 2018 **Membre du conseil de perfectionnement du Master SDL**, Université de Lorraine  
Mission : représentante des anciens étudiants.
- 2015 **Encadrement d'un stage de L3 SDL**  
Objectif : collecter des contextes de néologismes, procéder à une annotation morphologique.

## VALORISATION DE LA RECHERCHE, VULGARISATION

---

- Juin 2023 **Conférence TALN 2023 : Istex pour le TAL et le TDM** Billet de blog [Istex.fr](#).
- Mai 2023 **Exploration diachronique de la traduction automatique** Billet de blog [Istex.fr](#).
- Nov. 2022 **Un nouvel exemple d'usage des données Istex dans data.istex** Billet de blog [Istex.fr](#).
- Nov. 2022 **Les aspirations des véganes et végétariens au travers d'un corpus Istex** Billet de blog [Istex.fr](#).
- Juill. 2022 **Pourquoi dit-on plus souvent macroniste que macronien ?** Intervention sur Radio Declic.
- Juin 2022 **Lumière sur les néologismes politiques** Podcast audio, Sciences en Lumière (CNRS-UL).

## PRIX ET BOURSE

---

- 2023 **Prix scientifique de thèse L'Harmattan**
- 2023 **Prix de thèse 2022 de l'Académie Lorraine des Sciences**
- 2016 **Premier prix des communications doctorantes, AFLS Conferences 2016**, Belfast, UK
- 2015 **Bourse doctorale du MESR délivrée sur concours**

## COMPÉTENCES ET DOMAINES D'EXPERTISE

---

CULTURE SCIENTIFIQUE : LINGUISTIQUE

- **Compétences** : fouille de textes & TAL, morphologie, sémantique, phonologie, terminologie

CULTURE SCIENTIFIQUE : TRAITEMENT DE DONNÉES ET DE CORPUS

- **Compétences** : constitution/exploration de corpus, extraction de données, programmation, statistiques, principes FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable)

- **Outils** : Lodex, OpenRefine, R/R Studio, AntConc, Transcriber, Praat, Elan, Clan, moissonnage d'API, XML/XSL, TXM, bases en Python, Perl, JavaScript

#### CONNAISSANCE DES MÉTHODOLOGIES DE CONDUITE DE PROJET

- **Compétences** : méthodes Agiles (Scrum), méthodologie UX (expérience utilisateur), gestion de PGD

- **Outils** : Trello, Kanboard

#### ANIMATION DE FORMATIONS ET PÉDAGOGIE

- **Expériences** : 6 ans d'enseignement à l'Université (licences et masters), formations à la constitution de corpus Istex (doctorants, étudiants conservateurs), formée aux pédagogies innovantes (classe inversée)

- **Outils** : Moodle, Frama, Zoom, Teams, Prezi

#### CONNAISSANCE DE L'ENVIRONNEMENT INFORMATIQUE : BUREAUTIQUE & WEB

- **Outils** : Microsoft Office, Open Office, Inkscape,  $\LaTeX$ , Wordpress, HTML, Mattermost

LANGUES Français (langue maternelle), Anglais (B2), Espagnol (B1), Langue des Signes Française (A2)