

Maria Boritchev

+33 (0)6 69 92 54 28 | maria.boritchev@telecom-paris.fr | <https://perso.telecom-paristech.fr/mboritchev/>

EXPÉRIENCE PROFESSIONNELLE

Position de maîtresse de conférences

Oct. 2023 – Aujourd’hui

Télécom Paris

Palaiseau, France

- Enseignement : Machine Learning, Traitement Automatique des Langues, Linguistique Computationnelle.
- Recherche : parsing sémantique du langage naturel, analyse d’erreur, applications en méta-recherche sur l’éthique de l’IA.
- Collaborations avec Tiphaine Viard, Mélanie Gornet, Simon Delarue.

Post-doctorat

Jan. 2023 – Août 2023

Orange Innovation

Lannion, France

- Parsing sémantique du langage naturel, combinant des méthodes de sémantique formelle et d’apprentissage profond.

Post-doctorat

Déc. 2021 – Déc. 2022

Mathematical Institute of the Polish Academy of Sciences

Varsovie, Pologne

- Développement et étude de modèles hybrides du raisonnement naturel, combinant des méthodes de logique formelle et d’apprentissage profond.

Position d’Attachée temporaire d’enseignement et de recherche (ATER)

Sept. 2020 – Sept. 2021

Institut des sciences du Digital, Management & Cognition, Université de Lorraine

Nancy, France

- 192 heures d’enseignement, sur l’année, auprès de publics de la première année de Licence à la dernière année de Master.
- Sujets enseignés : Algorithmique, Programmation, Traitement Automatique des Langues, Logique, Analyse de concepts formels, Méthodologie de la recherche, enseignements en français et en anglais, voir la dernière colonne de la table 2 pour les détails.
- Suite des projets de recherche autour du sujet de thèse dans l’équipe Sémagramme, LORIA, Inria Nancy Grand-Est.

Doctorat

Sept. 2017 – Nov. 2021

LORIA, Inria Nancy Grand-Est

Nancy, France

- Encadrants : MAXIME AMBLARD et PHILIPPE DE GROOTE
- Élaboration d’un modèle logique, compositionnel et dynamique du dialogue, en particulier pour l’articulation des questions et des réponses dans le dialogue oral.
- Construction d’un corpus du français oral (DinG – Dialogues in Games) composé de données réelles, transcriptions d’enregistrements de personnes jouant au jeu de société *Catane*.
- Élaboration et test de plusieurs schémas d’annotation multilingues pour les questions et les réponses dans le dialogue, développés d’abord sur l’anglais, puis testés sur des corpus en français, italien, espagnol, néerlandais, et chinois mandarin.
- Enseignement et création de supports d’enseignement et travaux dirigés, voir table 2.

Stage de recherche

Fév. 2017 – Juin 2017

LORIA, Inria Nancy Grand-Est

Nancy, France

- Encadrant : MAXIME AMBLARD
- Élaboration d’un modèle dynamique du dialogue dans un cadre théorique restreint.
- Mise en perspective du modèle sur des données provenant du corpus SLAM – Schizophrénie and Language, Analyse and Modélisation (en français).

Stage de recherche

Juin 2016 – Août 2016

Center for Argumentation Technology, University of Dundee

Dundee, Écosse

- Encadrant : CHRIS REED
- Extraction automatique de correspondances intertextuelles entre l’argumentation politique de candidats aux primaires présidentielles américaines des élections de 2016 et les réactions des usagers du réseau social Reddit.

Stage de recherche

Juin 2015 – Juil. 2015

Institut de Recherche en Informatique de Toulouse

Toulouse, France

- Encadrant : PATRICK ST. DIZIER
- Analyse des opinions et argumentations des catégories de consommateurs et consommatrices, suivant les données du site Internet Movie Database (IMDb).
- Développement d’une version anglophone de l’outil TextCoop (outil d’analyse du discours, en français).

Groupe	Cours	Langue	23-24
M2 Télécom & X	<i>DS-Telecom-20: NLP</i>	Anglais	3h30 CM + 6h TP
1 ^e année MS Télécom	<i>IA717: NLP</i>	Français	3h CM + 8h projet
1 ^e année MS Télécom	<i>BGD709 : Données du Web</i>	Français	3h CM + 6h TP
1 ^e année MS Télécom	<i>Projet fil rouge</i>	Français	
3 ^e année Ingénieur	<i>IA312: NLP</i>	Anglais	3h CM + 3h TP
1 ^e année Ingénieur	<i>Projet Artishow</i>	Français	
2 ^e année Ingénieur	<i>SD-TSIA214: ML for Text Mining</i>	Anglais	3h CM + 6h TP
M2 IPP DATA & AI	<i>Softskills seminar</i>	Anglais	0h30 TP
2 ^e année Ingénieur	<i>SI221-B : Bases de l'apprentissage approche algorithmique</i>	Français	3h CM + 6h TP
			Total : 18h30 CM + 27h30 TP

Table 1: NLP = Natural Language Processing, ML = Machine Learning, DS = Data Science, M2 = Master 2, M1 = Master 1, MS = Mastère spécialisé.

Groupe	Cours	Langue	17-18	18-19	19-20	20-21
M2 TAL et SC	<i>Méthodologie de la recherche</i>	Anglais			5h CM + 5h TD	5h CM + 5h TD
M1 et M2 TAL	<i>Remise à niveau – calcul matriciel</i>	Anglais		4h CM + TD		
M1 SC	<i>Innovation et technologies</i>	Anglais	15h TD			
M1 TAL	<i>Sémantique du web</i>	Anglais (+ russe)		15h TD		
M1 TAL	<i>Méthodes pour le TAL</i>	Anglais				4h TD
M1 TAL	<i>Calculabilité et complexité</i>	Anglais				10h TD
L3 SC	<i>Ingénierie linguistique</i>	Français				10h TD
L3 SC	<i>Formalismes et représentations de raisonnements</i>	Français	10h CM + 10h TD	10h CM + 10h TD	10h CM + 10h TD	10h CM + 10h TD
L2 MIASHS	<i>Algorithmique avancée</i>	Français				36h TD
L2 SC	<i>Traitement Automatique des Langues</i>	Français				10h TD
L1 MIASHS	<i>Algorithmique et programmation 1</i>	Français		20h TD		22h TD
L1 MIASHS	<i>Algorithmique et programmation 2</i>	Français				15h TD
L1 MIASHS	<i>Algorithmique et programmation 3</i>	Français		26h TD		26h TD
L1 SHS	<i>Outils et culture numérique</i>	Français				20h TD

Table 2: TAL = Traitement Automatique des Langues, SC = Sciences Cognitives, MIASHS = Mathématiques et Informatique appliquées aux Sciences Humaines et Sociales, SHS = Sciences Humaines et Sociales, M2 = Master 2, M1 = Master 1, L3 = Licence 3, L2 = Licence 2, L1 = Licence 1.

ÉDUCATION

Doctorat d'Informatique

Soutenu le 22 novembre 2021

Université de Lorraine, Nancy, France

Sep. 2017 – Nov. 2021

- * **Titre :** *Modélisation dynamique des dialogues*
- * **Encadrants :** Maxime Amblard (HdR) et Philippe de Groote (DR)
- * **Rapporteuse :** Farah Benamara Zitouna (IRIT, Université Paul Sabatier, Toulouse, France)
- * **Rapporteur :** Jonathan Ginzburg (Université Paris Diderot – Paris 7, France)
- * **Examinatrice :** Ellen Breitholtz (Université de Göteborg, Suède)
- * **Examinatrice :** Iris Eshkol-Taravella, (Université Paris Nanterre, France)
- * **Examineur :** Miguel Couceiro (Université de Lorraine, France)

Université de Lorraine

Master 2, Sciences Cognitives Appliquées, spécialité Traitement Automatique des Langues

Nancy, France

Sep. 2016 – Août 2017

École Normale Supérieure de Lyon

Master 1, Informatique Théorique

Lyon, France

Sep. 2015 – Août 2016

PARTICIPATION À LA VIE DE LA COMMUNAUTÉ :

Relectrice pour la revue Language Resources and Evaluation	2024
Relectrice pour la conférence SemDial (TrentoLogue 2024)	2024
Co-organisatrice de l'animation scientifique du département, IDS, Télécom Paris	2023-
Relectrice pour la revue TAFFC, Volume 14	2023
Relectrice pour la conférence SemDial (MariLogue 2023)	2023
Relectrice pour la revue Corpus (à paraître en janvier 2024)	2023
Relectrice pour la conférence IWCS 2023	2023
Relectrice pour la conférence Coling 2022	2022
Membre du comité de programme de la conférence SemDial (DubDial 2022)	2022
Membre du comité de programme de la conférence RECITAL	2020
Organisatrice du séminaire hebdomadaire d'équipe, équipe Sémagramme, Loria	2017-2019

PUBLICATIONS

Mapping AI ethics: a meso-scale analysis of its charters and manifestos, Mélanie Gornet, Simon Delarue, Maria Boritchev, Tiphaine Viard, ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT), Juin 2024, Rio de Janeiro, Brésil.

A formal and computational semantics approach to identifying underexplored topics: the case of artificial intelligence ethics charters and manifestos, Maria Boritchev, Tiphaine Viard, Proceedings of the first Undone Computer Science conference (UndoneCS), Février 2024, Nantes, France.

Towards dynamic inquisitive semantics of dialogue, Maria Boritchev, Theoretical Linguistics Vol. 49, Issue 1-2, Juin 2023.

Error Exploration for Automatic Abstract Meaning Representation Parsing, Maria Boritchev, Johannes Heinecke, Proceedings of the 15th International Conference on Computational Semantics (IWCS 2023), Juin 2023, Nancy, France.

Compositionality in a simple corpus, Manuel Vargas Guzmán, Maria Boritchev, Jakub Szymanik, Maciej Malicki, Journées Jointes des Groupements de Recherche « Linguistique Informatique, Formelle et de Terrain » et « Traitement Automatique des Langues », Novembre 2022, Marseille, France.

A Multi-Party Dialogue Resource in French, Maria Boritchev, Maxime Amblard, Proceedings of the Thirteenth International Conference on Language Resources and Evaluation (LREC 2022), Juin 2022, Marseille, France.

An inquisitive account of wh-questions through event semantics, Maxime Amblard, Maria Boritchev, Philippe de Groote, LACL 2021: Logical Aspects of Computational Linguistics, Décembre 2021, en ligne, France.

Picturing Questions and Answers – a formal approach to SLAM, Maria Boritchev, Maxime Amblard, (In)coherence of discourse – Formal and Conceptual issues of Language, Springer, Language, Cognition and Mind, pp 65-89, 2021.

On dialogue modeling: a dynamic epistemic inquisitive approach, Maria Boritchev, Philippe de Groote, New Frontiers in Artificial Intelligence, Springer International Publishing, pp 21-36, 2021.

There is as yet Insufficient Data for a Meaningful Answer, Maria Boritchev, Maxime Amblard, SemDial 2020 – WatchDial – 24th Workshop on the semantics and pragmatics of dialogue, Juillet 2020, en ligne, États Unis.

A Taxonomy of Real-Life Questions and Answers in Dialogue, Maxime Amblard, Maria Boritchev, Marta Carletti, Lea Dieudonat, Yiting Tsai, SemDial 2019 – LondonLogue – 23rd Workshop on the semantics and pragmatics of dialogue, Septembre 2019, Londres, Royaume Uni.

Toward Dialogue Modeling: A Semantic Annotation Scheme for Questions and Answers, Maria-Andrea Cruz-Blandón, Gosse Minnema, Aria Nourbakhsh, Maria Boritchev, Maxime Amblard, LAW XIII 2019 – The 13th Linguistic Annotation Workshop, Août 2019, Florence, Italie.

A compositional view of questions, Maria Boritchev, Maxime Amblard, WiNLP – Widening NLP – ACL Workshop, Juillet 2019, Florence, Italie.

Coffee or tea? Yes, Maria Boritchev, Maxime Amblard, The 22nd workshop on the Semantics and Pragmatics of Dialogue, Laurent Prévot, Magalie Ochs and Benoît Favre, Novembre 2018, Aix-en-Provence, France.

PRÉSENTATIONS

Semantics, reasoning & representation, journée de recherche du LTCI, Novembre 2023, Télécom Paris, France.

Logic, compositionality, and semantics for natural language modeling, séminaire de l'équipe LIRICA, Avril 2023, Marseille, France, en ligne.

Exploring compositionality and semantics towards hybrid approaches to natural language modeling, LISN, Mars 2023, Paris Saclay, France.

Compositionality and semantics for natural language modeling, séminaire de l'équipe S2A de Télécom Paris, Mars 2023, Paris, France, en ligne.

Dialogue modeling & reasoning, séminaire de l'équipe EXPRESSION, IRISA, Janvier 2023, Bretagne, France, en ligne.

Compositionality and logic in language, séminaire Jeunes Talents TALEP (JTT), Janvier 2023, Marseille, France.

DinG – dialogue, jeu, et questions, séminaire DELICORTAL, Décembre 2022, Grenoble, France, en ligne.

Compositionality, languages, and neural networks, séminaire Café TAL, Novembre 2022, Nancy, France, en ligne.

Hybrid models of natural reasoning, co-présentée avec Manuel Vargas Guzmán, Reasoning & natural language processing seminar, Juin 2022, Varsovie, Pologne.

Hybrid models of natural reasoning, co-présentée avec Manuel Vargas Guzmán, Workshop on reasoning & neural models, Mai 2022, Amsterdam, Pays-Bas.

Modeling Dialogue through the Question-Answer relationship, CLASP seminar, Mars 2022, Göteborg, Suède, en ligne.

An inquisitive account of wh-questions through event semantics, LACL 2021: Logical Aspects of Computational Linguistics, Décembre 2021, en ligne.

Dialogues in Games – DinG, Journées du GdR LIFT, Décembre 2021, Grenoble, France.

There is as yet Insufficient Data for a Meaningful Answer – Linking questions and answers, 24th Workshop on the semantics and pragmatics of dialogue, Juillet 2020, en ligne.

Semantic processing of (real-life) questions and answers, Workshop on Formal Approaches to (In)coherence and Dynamics in Dialogue, Octobre 2019, Göteborg, Suède.

Thé, café et licornes — Traitement de la structure conversationnelle dans la théorie dynamique des types, EJCIM 2019 – School for Young Researchers in Computer Science Mathematics, Mars 2019, Marseille, France.

On Politics and Argumentation, MALOTEC Seminar, Mars 2017, Nancy, France.

Grande L3 deviendra mini-chercheuse, SIESTE Seminar, Février 2017, Lyon, France.

RÉCOMPENSES ET FINANCEMENTS

Journée d'automne de l'école doctorale IAEM Déc 2019
Prix du meilleur poster

Programme de mobilité pour la recherche de l'institut français de Suède Oct 2019
Participante du programme GALAN

ENCADREMENTS

HNApp - Apprentissage machine avec les Humanités Numériques Jan. 2024 – Juin 2024

- *Projet Artishow, 1^e année Ingénieur*
- Étudiants : Etienne Pelissier, Jean-Claude Moissinac, Yacine Khalil, Daniel Berda, co-encadrement : JEAN-CLAUDE MOISSINAC
- Mise en place d'une plateforme web permettant de visualiser et d'explorer des données culturelles et de les mettre en relation avec des manuels scolaires.

Adaptation de Modèles de Langage au domaine Renault Jan. 2024 – Juin 2024

- *Projet fil rouge, 1^e année MS Télécom*
- Étudiants : Byong Hee Lee, Ahmed Belaaïj, Maxime Ledieu, Leo Barberis, co-encadrement : MATTHIEU LABEAU, FRANCOIS-PAUL SERVANT (Renault)
- Exploration d'approches et techniques de science des données possibles pour l'adaptation des modèles de langage généraux au domaine automobile et particulièrement aux entités Renault, à partir d'un modèle existant.

- What are you saying?** | *Projet tutoré, Master 1* Oct. 2020 – Août 2021
- Étudiants : Anar Yeginbergenova et Albert Millert, co-encadrement : MAXIME AMBLARD, CHUYUAN LI
 - Développement d'un parser dialogique pour des données cliniques.
- Dialogues in Game – DinG** | *Stage de recherche, License 3* Mai 2020 – Août 2020
- Étudiante : Lucille Dumont, co-encadrement: MAXIME AMBLARD
 - Étude des phénomènes d'incompréhension dans le dialogue à travers l'articulation des questions et des réponses.
- Does my question answer your answer?** | *Projet tutoré, Master 1* Oct. 2019 – Juin 2020
- Étudiants : Axel Didier et Alena Yakavets, co-encadrement : MAXIME AMBLARD
 - Automatisation de l'annotation des actes dialogiques (interrompu par la pandémie).
- DinG – élaboration du guide de transcription** | *Stage de recherche, Master 1* Juin 2019 – Août 2019
- Étudiant : Esteban Marquer, co-encadrement : MAXIME AMBLARD
- Tea or coffee?** | *Projet tutoré, Master 1* Oct. 2018 – Juin 2019
- Étudiantes : Léa Dieudonnat, Marta Carletti, Yiting Tsai, co-encadrement : MAXIME AMBLARD, MICHEL MUSIOL
 - Étude multilinguistique de la relation question-réponse dans le dialogue.
- What's The Answer: Dialogue Annotation** | *Projet tutoré, Master 1* Fév. 2018 – Juin 2018
- Étudiants : Maria Andrea Cruz Blandon, Gosse Minnema, Aria Nourbakhsh, co-encadrement : MAXIME AMBLARD
 - Création d'un schéma d'annotation multilingue pour les question et les réponses dans le dialogue, expérimentations sur l'automatisation de l'annotation.

MÉDIATION SCIENTIFIQUE

- “ASK A SCIENTIST”** en ligne, Août 2022 - Août 2023
- Interventions bi-mensuelles dans le cadre de l'association EARTHLINGS HUB auprès d'enfants déplacés au cours de la guerre en Ukraine. Réponse aux questions scientifiques des enfants, en russe et ukrainien.
- Informatique & Linguistique** Paris, France, Février 2022
- Présentation du domaine de la linguistique computationnelle à travers mes sujets de recherche auprès d'étudiantes et étudiants de la Licence 1 Sciences du Langage de l'Université de Paris, dans le cadre du cours d'ALICE CESBRON.
- CogTalk #2 : Faire sa thèse au Loria** Nancy, France, Février 2019
- Présentation de mon sujet de thèse et des procédures pour candidater en thèse, à destination des étudiantes et étudiants de l'Université de Lorraine (de la Licence 1 au Master 2).
- Samedi des Curieux, Enquête 2018** Drancy, France, Mars 2018
- Après-midi découverte de la logique et de la cryptographie pour des enfants de 6 à 10 ans et leurs parents.
- ApéroDOC #1** Nancy, France, Février 2018
- Présentation informelle de mon sujet de thèse aux doctorantes et doctorants de l'Université de Lorraine, tous sujets confondus.

COMPÉTENCES TECHNIQUES & LANGUES

- | | |
|---|---|
| Maitrise : Python | Bilingue : français, russe |
| Utilisation : OCaml, C, HTLM, CSS, Prolog | Courant : anglais |
| Notions : bash, Bootstrap, Coq, Java, Perl, SQL, Lex&Yacc | Débutant : allemand, polonais, ukrainien |
| Outils : Git, L ^A T _E X, Beamer, Anaconda, VS Code | |