

FRANÇOIS BUET

23 septembre 1994

Campus universitaire bât. 507, rue du Belvédère, 91405 Orsay, France

IdHAL : francois-buet ◊ ORCID : 0000-0002-7405-9182 ◊ IdRef : 266180167

francois.buet@lisn.upsaclay.fr

FICHE RÉCAPITULATIVE

PARCOURS

Postdoctorat

Université Paris-Saclay, CNRS, LISN (Orsay)

Sous la supervision de Cyril Grouin et Sahar Ghannay, dans l'équipe SEME, depuis février 2023.

Doctorat

Université Paris-Saclay, CNRS, LISN (Orsay)

Sous la supervision de François Yvon, dans l'équipe TLP, d'octobre 2018 à octobre 2022.

Diplôme de Master / Diplôme d'Ingénieur Civil des Mines

Université de Lorraine, Mines de Nancy (Nancy)

Stage de recherche sous la supervision de Mathieu Constant, à l'ATILF dans l'équipe Ressources, de février à août 2018.

TRAVAUX ET THÉMATIQUES DE RECHERCHE

Mots-clefs :

- profilage d'auteur, détection de genre
- simplification de phrase, compression de phrase
- sous-titrage automatique

Publications :

- Revues nationales : 1
- Conférences internationales : 2
- Conférences nationales : 2
- Ateliers internationaux : 1

ENSEIGNEMENTS

Intitulé	Établissement	Niveau	Heures	Année
Remise à niveau Informatique	UPSaclay, Polytech	M2	14	2019-2022
Algorithmique	UPSaclay, Polytech	L2	124	2018-2022
Introduction à l'Informatique	UPSaclay, Polytech	L1	26	2019-2021
Introduction à l'IA	UPSaclay, UFR Sciences	L3	105	2018-2022
Mathématiques pour l'Informatique	UPSaclay, UFR Sciences	L2	14	2021-2022
Outils logiques et algorithmiques	UPSaclay, UFR Sciences	L2	6	2021-2022
Total heures statutaires			289	

(UPSaclay pour Université Paris-Saclay)

ÉTUDES

2022 Thèse en Informatique à l'Université Paris-Saclay, préparée au sein du Laboratoire Interdisciplinaire des Sciences du Numérique (LISN / CNRS, Orsay) sous la direction de François Yvon (LISN / CNRS, Université Paris-Saclay), débutée le 1^{er} octobre 2018 et soutenue le 21 octobre 2022.

Titre : *Modèles neuronaux pour la simplification de parole, application au sous-titrage.*

Jury :

- Annelies Braffort (Présidente, LISN / CNRS, Université Paris-Saclay)
- Christophe Cerisara (Rapporteur, LORIA / CNRS, Université de Lorraine)
- Benoit Favre (Rapporteur, LIS / CNRS, Aix-Marseille Université)
- Yannick Estève (Examineur, LIA, Avignon Université)
- Thierry Etchegoyhen (Examineur, Vicomtech)

Notice Sudoc : <https://www.sudoc.fr/266678904>

2018 Diplôme de Master en Traitement Automatique des Langues à l'Université de Lorraine.

Diplôme d'Ingénieur Civil des Mines à l'École Nationale Supérieure des Mines de Nancy (Nancy), Département Informatique.

Stage sous la supervision de Mathieu Constant (ATILF / CNRS, Université de Lorraine).

Titre : *Préparation de l'annotation syntaxique d'une base de textes littéraires.*

2016 Bachelor en Sciences de l'Ingénieur à l'École Nationale Supérieure des Mines de Nancy (Nancy).

2015 CPGE Mathématiques - Physique au Lycée Thiers (Marseille).

EXPÉRIENCES PROFESSIONNELLES

février 2023 - juin 2024

Chercheur postdoctoral

CNRS, LISN (Orsay)

Sous la supervision de Cyril Grouin et Sahar Ghannay, dans l'équipe SEME.

Financement : projet ANR GEM

octobre 2021 - septembre 2022

ATER mi-temps

Université Paris-Saclay

Enseignement à Polytech Paris-Saclay et à la Faculté des Sciences d'Orsay.

octobre 2018 - septembre 2021

Doctorant contractuel

CNRS, LISN (Orsay)

Sous la direction de François Yvon, dans l'équipe TLP.

Financement : projet ROSETTA (Bpifrance)

Doctorant enseignant

Université Paris-Saclay

Enseignement à Polytech Paris-Saclay et à la Faculté des Sciences d'Orsay.

février - août 2018

Stagiaire de recherche

CNRS, ATILF (Nancy)

Sous la supervision de Mathieu Constant, dans l'équipe Ressources.

juin - septembre 2017

Stagiaire assistant ingénieur

StarWear (Esch-sur-Alzette, Luxembourg)

Stage de 12 semaines dans une *start'up*, avec pour objectif de créer des applications pouvant démontrer le potentiel d'un produit faisant usage d'un capteur d'orientation.

PUBLICATIONS

Revue nationale avec comité de lecture

- [BY22] François BUET et François YVON. "Sous-titrage automatique : étude de stratégies d'adaptation aux genres télévisuels". In : *Revue TAL : traitement automatique des langues*. Varia 63.1 (déc. 2022), p. 11-35. URL : <https://hal.science/hal-03890594>.

Conférences internationales avec comité de lecture

- [BY21a] François BUET et François YVON. "Toward Genre Adapted Closed Captioning". In : *Interspeech 2021*. Brno (virtual), Czech Republic : ISCA, août 2021, p. 4403-4407. DOI : 10.21437/interspeech.2021-1762. URL : <https://hal.archives-ouvertes.fr/hal-03329488>.

Présentation orale, Taux d'acceptation: 42,3 % (963/2277)

- [Kar+22] Alina KARAKANTA, François BUET, Mauro CETTOLO et François YVON. "Evaluating Subtitle Segmentation for End-to-end Generation Systems". In : *13th Language Resources and Evaluation Conference (LREC)*. ELDA. Marseille, France, juin 2022, p. 3069-3078. URL : <https://hal.science/hal-03783891>.

Poster, Taux d'acceptation: 61,75 % (804/1302)

Conférences nationales avec comité de lecture

- [Bue20] François BUET. "Analyse de la régulation de la longueur dans un système neuronal de compression de phrase : une étude du modèle LenInit". In : *6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 3 : Rencontre des Étudiants Chercheurs en Informatique pour le TAL*. Sous la dir. de Christophe BENZITOUN, Chloé BRAUD, Laurine HUBER, David LANGLOIS, Slim OUNI, Sylvain POGODALLA et Stéphane SCHNEIDER. Nancy, France : ATALA, juin 2020, p. 57-70. URL : <https://hal.archives-ouvertes.fr/hal-02786187>.

Présentation orale, Taux d'acceptation: 72,7 % (16/22)

- [BY21b] François BUET et François YVON. "Vers la production automatique de sous-titres adaptés à l'affichage". In : *Traitement Automatique des Langues Naturelles*. Sous la dir. de Pascal DENIS, Natalia GRABAR, Amel FRAISSE, Rémi CARDON, Bernard JACQUEMIN, Eric KERGOSIEN et Antonio BALVET. Lille, France : ATALA, 2021, p. 91-104. URL : <https://hal.archives-ouvertes.fr/hal-03265891>.

Article long, Présentation orale, Taux d'acceptation: 77,6 % (45/58)

Ateliers internationaux avec comité de lecture

- [Xu+22] Jitao XU, François BUET, Josep CREGO, Elise BERTIN-LEMÉE et François YVON. “Joint Generation of Captions and Subtitles with Dual Decoding”. In : *19th International Conference on Spoken Language Translation (IWSLT 2022)*. Dublin, Ireland, mai 2022, p. 74-82. URL : <https://hal.archives-ouvertes.fr/hal-03666567>.

Scientific paper, Présentation orale, Atelier de ACL 2022, Taux d’acceptation: 50 % (9/18)

Résumés

- [Add+23] Gilles ADDA, François BUET, Sahar GHANNAY, Cyril GROUIN, Camille GUINAUDEAU, Lufei LIU, Aurélie NÉVÉOL, Albert RILLIARD et Uro RÉMI. “Projet Gender Equality Monitor (GEM)”. In : *18e Conférence en Recherche d’Information et Applications, 16e Rencontres Jeunes Chercheurs en RI, 30e Conférence sur le Traitement Automatique des Langues Naturelles, 25e Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues*. Sous la dir. de Christophe SERVAN et Anne VILNAT. Paris, France : ATALA, 2023, p. 21-21. URL : <https://hal.science/hal-04208588>.

ACTIVITÉS D’ENSEIGNEMENT

Intitulé	Établissement	Niveau	Heures	Nature
2021-2022 : ATER mi-temps				
Remise à niveau Informatique	UPSaclay, Polytech	M2	12	TD/TP
Algorithmique	UPSaclay, Polytech	L2	40	TD/TP
Introduction à l’IA	UPSaclay, UFR Sciences	L3	24	TD/TP
Mathématiques pour l’Informatique	UPSaclay, UFR Sciences	L2	14	TD/TP
Outils logiques et algorithmiques	UPSaclay, UFR Sciences	L2	6	TP
2020-2021 : Doctorant enseignant				
Algorithmique	UPSaclay, Polytech	L2	28	TD/TP
Introduction à l’Informatique	UPSaclay, Polytech	L1	13	TD
Introduction à l’IA	UPSaclay, UFR Sciences	L3	24	TD/TP
2019-2020 : Doctorant enseignant				
Remise à niveau Informatique	UPSaclay, Polytech	M2	2	TD
Algorithmique	UPSaclay, Polytech	L2	28	TD/TP
Introduction à l’Informatique	UPSaclay, Polytech	L1	13	TD
Introduction à l’IA	UPSaclay, UFR Sciences	L3	27	TD/TP
2018-2019 : Doctorant enseignant				
Algorithmique	UPSaclay, Polytech	L2	28	TD/TP
Introduction à l’IA	UPSaclay, UFR Sciences	L3	30	TD/TP
Total heures statutaires			289	

ÉCOLE D’ÉTÉ SUIVIE PENDANT LA THÈSE

Ninth Lisbon Machine Learning School (LxMLS 2019)

Instituto Superior Técnico (IST, Lisbonne, Portugal)

École centrée sur l’apprentissage automatique, les statistiques, et particulièrement sur leur application dans le domaine du TAL.

ACTIVITÉS AUTOUR DE LA RECHERCHE

Comités de programme

2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2024)

Relecteur pour quatre articles.

Domaine : Summarization, NLP Applications

Appel à soumissions : https://2024.naacl.org/calls/first_call_for_papers/

INTERSPEECH 2023

Sous-relecteur pour Sahar Ghannay, pour un article.

Appel à soumission : <https://interspeech2023.org/call-for-papers/>

Comité éditorial

Numéro 237 de la revue *Langages*

Relecteur pour une proposition.

Appel à contribution : <https://doi.org/10.58079/1bi4/>