

Janaina SABINO

INFOLINGUISTE

ADRESSE: Courbevoie
TÉL: 06 35 18 57 28
LINKEDIN: www.linkedin.com/in/janasabino
GITHUB: <https://github.com/janasabino>
EMAIL: janasabino@gmail.com

EXPÉRIENCES PROFESSIONNELLES

- | | |
|---------------------|--|
| DEPUIS MAI-2018 | <p>Infolinguiste Consultante à AI&DATA, Paris
<i>Analyse de la Voix du client (mission au sein du Groupe Crédit Agricole)</i>
Développement d'un modèle de thématiques et d'émotions pour la catégorisation automatique des verbatims; rédaction du guide d'annotation; encadrement de la campagne d'annotation; annotation de données; normalisation de données; extraction d'information; analyse exploratoire de corpus (keywords, ngrams, LDA); prétraitement de données</p> <p><i>Extraction de thématiques d'insatisfaction via les emails de réclamation du Service Client (mission chez Digipost)</i>
Anonymisation de données (RGPD); prétraitement de données; clustering/identification des thématiques</p> <p><i>Mise en place du classement automatique des emails (mission chez Maisons du Monde)</i>
Prétraitement de données; développement d'un modèle de thématiques pour la catégorisation automatique des emails, annotation de données, affectation des nouveaux emails aux thématiques identifiées</p> |
| AVR 2017 - MAI 2018 | <p>Infolinguiste à VIAVOO, Boulogne-Billancourt
<i>Amélioration produit (Espagnol, Anglais, Portugais, Français)</i>
Création et maintenance de ressources linguistiques (dictionnaires, grammaires formelles); tests et recettes des évolutions du produit</p> <p><i>Création d'un module pour la fonctionnalité résiliation - Espagnol (stage de fin d'études)</i>
Création d'un dictionnaire sémantique de la résiliation en Espagnol; création de patrons syntaxiques pour l'extraction d'information; tests et recettes des évolutions du produit</p> |
| FÉVR-SEPT 2016 | <p>Infolinguiste (Stagiaire) à SYSTRAN, Paris
<i>Développement produit (Français, Anglais, Portugais)</i>
Élaboration de dictionnaires sémantiques spécialisés monolingues ou bilingues; tests et recettes des évolutions du produit</p> |
| MAI-SEPT 2015 | <p>Infolinguiste (Stagiaire) à SINEQUA, Paris
<i>Extraction d'entités nommées</i>
Élaboration de dictionnaires sémantiques monolingues; mise au point d'automates d'extraction de groupes nominaux et d'entités nommées; tests et recettes des évolutions du produit</p> |

FORMATION

2021 BOOTCAMP INTELLIGENCE ARTIFICIELLE, **Data Science Academy**
2021 DATA SCIENCE CERTIFICAT PROFESSIONNEL, **IBM**
2020 ADVANCED NLP WITH SPACY, **Spacy**
2020 ADVANCED NLP, **Data Science Academy**
2017 M1-M2 TRAITEMENT AUTOMATIQUE DES LANGUES, **INALCO**, Paris
2012 Double Licence LINGUISTIQUE ET LITTÉRATURE, **UFMG**, Belo Horizonte - Brésil

LANGUES

PORTUGAIS: langue maternelle
FRANÇAIS: courant - C1
ESPAGNOL: courant - C1
ANGLAIS: upper-intermediate (en cours)
ITALIEN: bonnes notions - A2

COMPETENCES TECHNIQUES

Langages:	PYTHON
Environnements:	JUPYTER, PYCHARM, GOOGLE COLAB, SPYDER, GIT
NLP:	NLTK, SPACY, GENSIM, TEXTBLOB, TOPIC MODELING
Manipulation de données:	PANDAS
Visualisation de données:	MATPLOTLIB, SEABORN
Machine Learning et Deep Learning:	Connaissance des concepts et applications des principaux algorithmes