

JANAINA SABINO

Infolinguiste

Analyste de données textuelles informatisées



CONTACT

✉ janasabino@gmail.com
☎ 06 35 18 57 28
📍 Courbevoie, France
🏠 janasabino.github.io
🌐 @janasabino
🌐 Janaina Sabino

COMPÉTENCES TECHNIQUES

Langages:

Python ●●●●●
SQL ●●●●●
HTML/CSS ●●●●●

Environnements:

Git ●●●●●

IDEs

Jupyter ●●●●●
Google Colab ●●●●●
Pycharm ●●●●●

NLP

NLTK ●●●●●
spaCy ●●●●●

Software & Tools

Matplotlib/Seaborn ●●●●●
Pandas ●●●●●

Modélisation

Machine Learning ●●●●●
Deep Learning ●●●●●

LANGUES:

Portugais ●●●●●
Français ●●●●●
Espagnol ●●●●●
Anglais ●●●●●
Italien ●●●●●

LOISIRS:

Littérature du 19ème siècle, natation, musées et cinéma classique

EXPÉRIENCES PROFESSIONNELLES

📅 MAI 2018 - MAR 2021
📍 AI&DATA, Paris

Infolinguiste Consultante

Analyse de la Voix du client (mission au sein du Groupe Crédit Agricole)

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

Extraction de thématiques d'insatisfaction via les emails de réclamation du Service Client (mission chez Digipost)

Anonymisation de données (RGPD); prétraitement de données; clustering/identification des thématiques

Mise en place du classement automatique des emails (mission chez Maisons du Monde)

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

📅 Avr 2017 - Mai 2018
📍 VIAVOO, Boulogne

Infolinguiste

Développement produit (Espagnol, Anglais, Portugais, Français)

Création et maintenance de ressources linguistiques (dictionnaires, grammaires formelles); tests et recettes des évolutions du produit

Création d'un module pour la fonctionnalité résiliation - Espagnol (stage de fin d'études)

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

📅 Fév 2016 - Sep 2016
📍 Systran, Paris

Infolinguiste (Stagiaire)

Développement produit (Français, Anglais, Portugais)

Création et maintenance de ressources linguistiques (dictionnaires, grammaires formelles); élaboration de dictionnaires sémantiques spécialisés monolingues ou bilingues; tests et recettes des évolutions du produit

📅 Mai 2015 - Sep 2015
📍 Sinequa, Paris

Infolinguiste (Stagiaire)

Extraction d'entités nommées

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

FORMATION

📅 2022

📍 Data Science Academy (à distance)

Bootcamp Intelligence Artificielle

📅 2022

📍 Data Science Academy (à distance)

Advanced NLP

📅 2021

📍 Spacy (à distance)

Advanced NLP with spaCy

📅 2017

📍 INALCO, Paris

M1-M2 Traitement Automatique des Langues

📅 2012

📍 UFMG, Belo Horizonte (Brésil)

Double Licence Lettres et Édition