

Papa Séga WADE, PhD

Ingénieur de Recherche en IA | LLMs, Technologies Vocales, NLP | [Premier voicebot en Wolof](#).

📍 Paris, France | ✉ papa-sega.wade@m4x.org | ☎ +33 6 02 72 31 22

🌐 papasega.github.io | [in LinkedIn](#) | [GitHub](#) | 🤗 [Hugging Face](#)

PROFIL

Docteur en Informatique, spécialité Intelligence Artificielle (IMT Atlantique / Orange Innovation) et diplômé de l'École Polytechnique et Télécom Paris, je conçois et déploie des solutions d'IA vocale et de LLMs adaptées aux langues africaines. Créateur du premier voicebot en Wolof, j'ai porté ces innovations sur les scènes internationales les plus visibles du secteur (MWC Barcelone, GITEC Africa, VivaTech, AfricaTech). Je combine expertise technique de pointe et vision stratégique pour transformer l'IA en levier d'inclusion numérique et de création de valeur à l'échelle du continent.

Compétences Techniques

- **IA & LLMs** : Famille de modèles — Qwen-3, Mistral, GPT-OSS-20B (MoE), Qwen-2.5, LLaMA-3.2 | RAG | Ingénierie de prompts | Fine-tuning | LoRA/QLoRA | DPO | CFP | Adaptation de domaine.
- **Technologies vocales** : ASR (wolof-français), TTS, Diarisation, Speech Embeddings, Modèles SSL (WavLM, Wav2vec2.0, HuBERT)
- **Programmation & outils** : Python, PyTorch, TensorFlow, HuggingFace (HF), LangChain, Git, Gradio.
- **Autres** : Modélisation quantitative, Mathématiques fondamentales et appliquées, Data Science, Analyse statistique, HPC, Entraînement distribué, cluster GPU SLURM.
- **Langues** : Français (bilingue), Anglais (fluent), Wolof (natif).
- **Métriques d'évaluation** : F1-score, WER, BLEU, Perplexité, METEOR, Corrélation de Pearson, Exact Match.

Expériences

- **Chercheur en IA — Orange Innovation, Châtillon Sept. 2020 – Présent**
 - Conception et déploiement de solutions d'IA vocale en langues africaines (ASR, TTS, LLMs) pour les 17 pays d'Orange en Afrique.
 - **Rayonnement international** : Représentant Orange sur les salons internationaux majeurs :
 - MWC Barcelone 2026 : démo Maxit Easy Talk (assistant vocal multilingue pour l'inclusion numérique)
 - GITEC Africa (Marrakech, 2025), VivaTech (Paris, 2025), AfricaTech (Le Cap, 2025).
 - Conception de pipelines RAG avec LLMs open source pour l'automatisation du service client, réduisant le temps de traitement des conseillers.
 - Déploiement de [voicebots en Wolof](#), reliant langues sous-représentées et systèmes d'IA.
 - Construction de modèles ASR multilingues pour les langues subsahariennes (Wolof, lingala) avec les modèles/toolkit open source HuggingFace.

- **Thèse de doctorat en IA — Orange Innovation <> IMT Atlantique — Nov. 2021 – Juin 2025, Châtillon.**
 - Développement du premier LLM adaptatif en Wolof, s'appuyant sur des annotations sociolinguistiques pour ajuster les réponses selon le niveau de langue et le code-switching entre le wolof et le français.
 - Développement d'un pipeline de traduction automatique (anglais/français → wolof), s'appuyant sur la programmation parallèle et l'interface Google Translate. Ce système a permis de traduire efficacement en wolof des données d'instructions issues de Hugging Face, afin de pallier la pénurie de ressources linguistiques.
 - Proposition d'un système trimodal (audio, texte, marqueurs de fluence) pour l'évaluation automatique de la fluidité, améliorant la classification par rapport au baseline.
 - Introduction d'une segmentation basée sur les groupes de souffle pour une analyse plus fine, atteignant une corrélation >0,796 avec les jugements d'experts humains.
 - Publications à EUSIPCO, Speech Prosody et contributions à HAL-IMT et arXiv.
 - **Stages & Projets de recherche**
 - Orange Labs, Lannion (février 2020 – août. 2020) – NLP : modèle de reconnaissance automatique de la parole en Wolof.
 - Télécom Paris (2019) – Développement d'un [chatbot de soutien scolaire en mathématiques](#).
 - Société Générale (2018–2019) – Segmentation clients avec SQL/Teradata.
 - Total France (2018) – Préviation et analyse de comportements clients (LSTM, data mining).
 - Total France (2017) – Data mining sur les transactions de mobile money au Sénégal.
-

Projets Stratégiques

- **Maxit Easy Talk — Assistant vocal multilingue pour l'inclusion numérique**
 - Solution d'interaction vocale en langues locales permettant aux populations non-alphabétisées d'accéder aux services digitaux d'Orange (mobile money, telco).
 - [Présenté au MWC Barcelone 2026](#) sur le stand Orange, aligné sur la stratégie « Trust the Future ».
 - Levier d'acquisition client et de rétention sur des segments que les interfaces textuelles ne peuvent pas adresser.
 - **CCA – Contact Center Analysis**
 - Pipeline intégrant ASR + diarisation + segmentation + LLM pour analyser automatiquement les appels client-conseiller.
 - Le projet CCA a été présenté comme une innovation reconnue dans plusieurs salons internationaux : au Mobile World Congress (MWC) de Barcelone en mars 2025, au Gitex Africa de Marrakech en avril 2025, et au salon VivaTech de Paris en juin 2025.
 - Fournit aux équipes marketing et qualité des insights exploitables et des KPI sur les points de friction clients.
 - **LLM adaptatif en Wolof**
 - Fine-tuning de LLaMA-3.2, Qwen-2.5, GPT-OSS-20B (MoE) sur le Wolof avec des stratégies frugales (LoRA, quantification).
 - Premier LLM adaptatif au niveau de langue en Wolof, améliorant la pertinence des réponses en code-switching.
-

Éducation

- PhD en Intelligence Artificielle – IMT Atlantique & Orange Innovation (2021–2025)
 - Recherche axée sur les technologies vocales, le NLP et les LLMs pour les langues à faibles ressources, créant un environnement d'inclusion numérique adapté aux utilisateurs.
 - Diplôme d'Ingénieur en Data Science et Intelligence Artificielle – [École Polytechnique & Télécom Paris](#), France (2016–2020)
 - Data Science et Intelligence Artificielle. Mathématiques appliquées, Deep Learning, NLP, ASR.
 - Licence en Mathématiques – Université Assane Seck de Ziguinchor, Sénégal (2013–2015)
 - Mention Très Bien. Formation en mathématiques pures et appliquées.
-

Publications & Distinctions

- EUSIPCO 2025: [CBF-AFA: chunk-based multi-ssl fusion for automatic fluency assessment](#).
 - Workshop Speech Prosody 2024: [Holistic Trimodal Assessment of Fluency: Audio, Transcription, and Fluency Markers](#).
 - TLAN 2022: [Proof-of-Concept of a voicebot speaking wolof](#).
 - Bourse d'Excellence Fondation Mines-Télécom (2018–2019).
 - Prix d'Excellence dans l'Enseignement Supérieur, Sénégal (2015).
-

Leadership & Engagement

- Fondateur, [chaîne YouTube MathsPSW](#). 310 vidéos éducatives, 15,3k abonnés et plus de 850k vues. Démocratisation des mathématiques pour les lycéens sénégalais. Encadrement de plus de 50 élèves pour les Olympiades.
- Co-fondateur, [Institute of Algorithms of Senegal](#) (IAS). Initiative à but non lucratif pour démocratiser l'IA en Afrique. Organisation d'ateliers nationaux et de formations pour étudiants et ingénieurs.
- Référent [Animath](#) pour le Sénégal et membre de l'équipe pédagogique [TerangaMath](#). Coordination et mentorat pour les Olympiades de mathématiques, soutien à l'excellence STEM au lycée.
- Vice-Champion de France (2018) et Champion de France (2019), Boxe universitaire anglaise. Compétition au niveau national, démontrant résilience et leadership.