

Souhail Bakkali

Rennes, France | souhail.bakkali@irisa.fr | +33 753 420 036 | LinkedIn | GitHub | Google Scholar

Professional Summary

Associate Professor and Research Scientist working at the intersection of Multimodal Artificial Intelligence, Natural Language Processing, and Computer Vision. I design and lead applied research that translates state-of-the-art methods into robust, scalable systems for document understanding and multimodal learning. My contributions include vision-language architectures and efficient training/inference strategies, evidenced by 20 peer-reviewed publications (223 citations) in venues such as Pattern Recognition, WACV, and ICDAR. I have secured over 432,000 EUR in competitive funding, coordinated international collaborations, and managed large-scale computing to deliver production-grade outcomes while maintaining scientific rigor and effective cross-functional leadership.

Technical Skills

Multimodal AI & Deep Learning: Expert (5+ years) in multimodal learning, integrating language, vision, and layout, structured and heterogeneous data. Expert (5+ years) in vision-language models (CLIP, BLIP, LayoutLM, LLaVA). Experienced in multimodal fusion techniques for document understanding.

Computer Vision: Expert (5+ years) in image processing and document image analysis. Expert (5+ years) in scene-text detection and recognition. Experienced in object detection, classification, image retrieval and few-shot learning using CNNs, ViTs, and GANs.

Natural Language Processing (NLP): Expert (5+ years) in large language models (LLMs) (GPT, BERT, T5, RoBERTa). Expert (5+ years) in text classification, named entity recognition (NER), information extraction. Experienced in semantic search, information retrieval, and topic modeling.

Machine Learning & AI Frameworks: Expert (5+ years) in deep learning frameworks: TensorFlow, PyTorch, Keras. Experienced in scikit-learn, OpenCV, Hugging Face Transformers, NLTK, spaCy.

MLOps & Research Model Management: Experienced in managing and tracking deep learning experiments using MLflow and TensorBoard. Familiar with reproducible AI workflows, model versioning, and hyperparameter tuning. Knowledgeable in containerized environments (Docker) for research scalability. Experienced in setting up and maintaining SLURM-based compute clusters for large-scale AI experiments.

Information Retrieval & Search Systems: Experienced in vector search and document indexing with Elasticsearch, FAISS. Experienced in semantic search and question-answering systems.

High-Performance Computing: Administrator and manager of a cluster of servers equipped with GPUs (16x2080Ti, 10xA40, 7xA6000, 10xH100, 2xGP100) for 5 years, overseeing resource allocation, system optimization, and deep learning model deployment for large-scale multimodal AI experiments. Setup of L3i-calcul cluster based on SLURM.

Software Development & Deployment: Experienced in designing AI-driven applications using Flask, FastAPI. Experienced in model deployment with Docker.

Programming Languages: Python (expert, 5+ years), SQL (experienced), JavaScript, HTML, CSS (experienced), C++ (known).

Tools & Platforms: Git, GitHub, GitLab, Jupyter Notebook, VS Code, PyCharm, PostgreSQL, MySQL, MongoDB, Apache Spark, AWS, Google Cloud, Hugging Face Inference API

Operating Systems: Linux (expert, Ubuntu and CentOS), Windows

Research Leadership & Project Management: Proven track record leading cross-functional AI research teams across international institutions (9+ collaborations). Secured 432,000+ EUR in competitive funding. Managed end-to-end project lifecycle from ideation to creation. Experience coordinating multi-stakeholder projects with academic and industry partners (Siren Enterprise). Strong communication skills with technical and non-technical audiences.

Production Systems & Scalability: Experience deploying AI models in production environments. Managed GPU cluster infrastructure for 5+ years supporting large-scale deep learning experiments. Implemented reproducible ML workflows, model versioning, and experiment tracking. Knowledge of cloud platforms (AWS, Google Cloud) for scalable AI deployment.

Agile Development & Collaboration: Experience working in collaborative research environments with version control (Git, GitHub, GitLab). Strong problem-solving skills and adaptability to rapidly evolving technologies. Proven ability to deliver research insights and prototypes aligned with industry needs.

Foundation Models & LLMOps: Experienced with foundation models (GPT family, T5, BERT, vision-language models) across evaluation, deployment, and lifecycle management (guardrails, monitoring, observability). Focus on prompt strategy design, model governance, and reproducible pipelines for production-grade usage.

Retrieval-Augmented Generation (RAG) & Knowledge Indexing: Building semantic indices and hybrid search with FAISS/Elasticsearch, context chunking and reranking, citation grounding, and evaluation of faithfulness/factuality. Experience integrating RAG with LLMs for reliable, source-aware outputs.

AI Agents & Tool Use: Orchestration of agentic workflows (planning, memory, tool/function calling) for multi-step tasks and automation. Coordination of data retrieval, transformation, and decision stages aligned with real-world constraints.

Prompt Engineering & Context Engineering: System/instruction prompt design, few-shot strategies, structured prompting templates, and context construction (metadata, retrieval windows, formatting) to improve controllability, robustness, and task generalization.

Efficient Learning & Model Optimization: Knowledge distillation, pruning, mixed precision/quantization (FP16/INT8), and parameter-efficient fine-tuning (LoRA/PEFT). Experience with early exiting and compute-aware training/inference for performance-efficiency trade-offs.

Professional Experience

Associate Professor (Researcher), IRISA-Lab, ESIR, Dept. of Computer Science, Rennes, France Sep 2025 – Present

- Co-supervising 2 PhD students on handwritten text recognition in documents with self-supervised learning.
 - Building a foundation model for handwritten recognition using self-supervised learning.
 - Explainable multimodal document understanding by fusing knowledge graphs with large language models, augmented with federated learning for secure, interpretable processing.
 - Multimodal and Cross-Modal Document Image Classification.
 - Content-Based Document Image Retrieval.
 - Few-Shot Learning for Document Classification.

Associate Professor (Researcher), L3i-Lab, Dept. of Computer Science, La Rochelle University, France Sep 2023 – Aug 2025

- Led research initiatives in Multimodal Document Analysis and Understanding with measurable impact:
 - Document Abstractive Summarization.
 - Early Multimodal Exits for Fast Adaptive Inference.
 - Multimodal and Cross-Modal Document Image Classification.
 - Content-Based Document Image Retrieval.
 - Few-Shot Learning for Document Classification.
- Developed Blockchain-Based Reputation Systems for LLM Evaluation.
- Built biometric fraud detection systems for digital identity verification.
- Created benchmark datasets for Khmer scene-text recognition (WildKhmerST, KhmerST).
- Coordinated AI for Marine Ecology project analyzing video/image data from bio-recording technology.
- Co-supervising 3 PhD students and mentored 5+ Master's/Bachelor's thesis projects (3 resulting in publications).
- Secured 432,000+ EUR in research funding through 7+ competitive grants.
- Published 16+ papers in top-tier venues: Pattern Recognition, WACV, ICDAR, ACCV, IJDAR.
- Built collaborations with 9+ international institutions and 2 industry partners.

Postdoctoral Researcher & Lecturer, L3i-Lab, Dept. of Computer Science, La Rochelle Université, France Jan 2023 - Aug 2023

- Deep Learning for Multimodal Document Analysis and Understanding

- Document Image Classification
- Content-Based Document Retrieval, Indexing and Search
- Few-Shot Document Classification
- Multimodal Fusion Approaches
- Supervisor: Mickaël Coustaty - Associate Professor, La Rochelle Université

PhD Student & Lecturer, L3i-Lab, Dept. of Computer Science, La Rochelle Université, France Dec 2019 – Dec 2022

- PhD in Computer Science and Applications with European Label (CVC, Universitat Autònoma de Barcelona)
- Thesis: Inter-Modal and Unified Learning of Vision and Language for Multimodal Document Understanding
- Developed efficient multimodal AI frameworks for document analysis and understanding
- Collaborated with international institutions: L2TI (France), CVC (Spain), LIIR (Belgium), CompuCore-lab (Ireland)
- Collaborated with industry partner: Allread Machine Learning Technologies
- Published 5 papers during PhD including journals (Pattern Recognition, IJDAR), conferences (WACV, ICIP), and a workshop (CVPRW)

Information Technology Engineer (Research Intern), L3i Laboratory, La Rochelle Université, France Feb 2019 – Jul 2019

- Research Thesis: Face Authentication using Deep Learning to Merge Biometric Anti-Fraud Methods with Face Verification Algorithm
- Developed facial expression recognition methods for biometric authentication

Data Science Engineer (Intern), Inlog Solutions, Rabat, Morocco Jun 2018 – Aug 2018

- Fingerprint verification and recognition with deep learning
- Built Data Mining models and performed descriptive data analysis for knowledge extraction
- Client Segmentation and Recommendation of new possible offers

Software Developer (Intern), Lotus IT, Tangier, Morocco Jun 2017 – Aug 2017

- Web Development Project

Education

Ph.D. in Computer Science, La Rochelle Université, France Dec 2019 – Dec 2022

- **Thesis:** Multimodal Document Understanding with Unified Vision and Language Cross-Modal Learning
- European Doctorate Label from Computer Vision Center (CVC), Universitat Autònoma de Barcelona
- Specialization: Multimodal AI, Computer Vision, Natural Language Processing, Document Analysis
- Key Contributions: Vision-language models, cross-modal learning, few-shot learning, document classification

Engineering Degree, Institut National des Postes et Télécommunications (INPT), Morocco Sep 2016 – Sep 2019

- Major: Telecommunications Engineering and Information Technology
- Master Thesis: Face Authentication using Deep Learning for Biometric Anti-Fraud and Face Verification

Key Publications & Research Impact

Publication Metrics: 20 publications, 223 citations, H-index 6 (Google Scholar)

Selected Publications

GlobalDoc: Cross-Modal Vision-Language Framework for Document Retrieval 2025
S. Bakkali et al.

IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2025)

Multimodal Adaptive Inference with Anytime Early Exiting 2024

O. Hamed, S. Bakkali et al.

International Conference on Document Analysis and Recognition (ICDAR 2024)

VLCDoC: Vision-Language Contrastive Pre-training for Cross-Modal Classification 2023

S. Bakkali et al.

Pattern Recognition (Impact Factor: 8.0)

EAML: Ensemble Self-Attention-Based Mutual Learning for Document Classification 2021

S. Bakkali et al.

International Journal on Document Analysis and Recognition (IJ DAR)

Research Grants & Funding

Accepted Projects:

- **FELM (EU-CONEXUS 2025-2027):** 153,000 EUR with SETU Ireland for lightweight multimodal learning models with fairness and transparency - Co-Scientific Director 50%
- **FlexChain (2025-2027):** 124,000 EUR joint PhD thesis with Siren Enterprise on federated learning and blockchain for secure document processing - Co-Scientific Director 50%
- **ModamPadoc (RNA-LRU 2025):** 120,000 EUR joint PhD thesis with USTH Vietnam on multimodal document understanding for heritage documents using LLMs - Co-Scientific Director 50%
- **DocInnovate (ACI 2024):** 7,000 EUR for digital platform development for historical archives using LLMs and LMMs - Scientific Director 100%
- **AI for ECO-WELLNESS (PHC PERIDOT 2024-2026):** 28,000 EUR with MNS University Pakistan for handwritten text recognition in medical prescriptions - Co-Scientific Director 25%

Rejected:

- **MuLDMoBE (RNA 2025-2027):** 130,000 EUR with Josef Stefan Institute Slovenia for multilingual document structures using LLMs - Co-Scientific Director 50%

Research Collaborations & Projects

International Collaborations (9+ Institutions):

- **Josef Stefan Institute, Slovenia:** Developing lightweight multimodal learning models for multilingual documents (1 grant under review: 130,000 EUR)
- **Siren Company, France:** FlexChain project on federated learning and blockchain for secure document processing (124,000 EUR funded)
- **CompuCore-lab, SETU, Ireland:** Compression and knowledge distillation for efficient multimodal models (1 grant under review: 153,000 EUR)
- **LIIR Laboratory, KU Leuven, Belgium:** Performance-efficiency tradeoffs for document classification (1 publication at ICDAR 2024)
- **Computer Vision Center (CVC), UAB, Spain:** Multimodal document understanding (5 publications, 1 under review in Pattern Recognition)
- **L2TI Laboratory, Sorbonne Paris Nord, France:** Published 5 papers on multimodal document understanding
- **AllRead ML Technologies, Barcelona, Spain:** Industry collaboration on text detection and recognition (4 publications)
- **Cambodia Academy of Digital Technology:** Low-resource language text recognition (2 joint PhD theses, 120,000 EUR)
- **USTH Vietnam:** Heritage document understanding (joint PhD, 120,000 EUR)

Collaboration Impact Metrics:

- 9+ co-authored publications with international partners.
- 432,000+ EUR combined funding from collaborative projects.

- 5+ joint PhD supervisions with international institutions.
- 1 industry partnership resulting in practical AI solutions.

Technical Focus Areas:

- Vision-language models for efficient visually-rich document analysis and understanding
- Cross-modal learning and multimodal fusion techniques
- Scene-text detection and recognition for low-resource languages
- Fairness, transparency, and explainability of Large Multimodal Models
- Data privacy, federated learning, and blockchain-based document security
- Few-shot learning for document classification with limited labeled data

Awards & Recognition

- **European Doctorate Label (Doctor Europaeus):** Awarded for international research collaboration at Computer Vision Center, Universitat Autònoma de Barcelona (4 months stay)
- **Research Impact:** 20 publications, 223 citations, H-index 6
- **Funding Success:** Secured 432,000+ EUR in competitive research grants (7+ projects)
- **Dataset Contributions:** Created benchmark datasets (WildKhmerST with 10,000+ labeled images, KhmerST with 1,500+) advancing multimodal AI research and adopted by international research community
- **Conference Presentations:** Presented research at 6+ international venues including WACV, ICDAR, ACCV, ICIP
- **Industry Recognition:** Collaboration with Siren Enterprise on production AI systems
- **Technical Leadership:** Managed GPU cluster infrastructure (45+ GPUs) supporting 20+ researchers for 5 years

Professional Activities

Peer Review & Technical Leadership

- **Journals:** Pattern Recognition (PR), International Journal of Computer Vision (IJCV), Engineering Applications of Artificial Intelligence (EAAI), NeuroComputing.
- **Conferences:** ICDAR, ICPR, ICIP, LREC-COLING - Served on program committees and reviewed 20+ submissions
- **Professional Membership:** GRCE (Written Communication Research Group)

Research Leadership & Community Building

- Established and direct Reading Group at L3i Laboratory (2023-present) - 20+ participants, bi-weekly sessions
- Organized 5+ seminars on deep multimodal learning at L3i Laboratory with international speakers
- Member of organizing committee: ICPR 2024 Call for Competitions, DAS 2022 Workshop (200+ attendees)

Mentorship & Team Leadership

- Co-supervising 3 PhD students at La Rochelle Université (ongoing, expected completion 2026-2027)
- Co-supervising 1 PhD student at CompuCore-lab (ongoing, expected completion 2026-2027)
- Co-supervising 1 PhD student at IRISA-lab (ongoing, expected completion 2028-2029)
- Mentored 5+ Master's and Bachelor's thesis projects
- Technical advisor for cross-functional research teams spanning 4 countries
- Course coordinator: Advanced Database Management Systems (50+ students annually)

References

Available upon request