# Wafa **AISSA**

PhD in Artificial Intelligence

*UCLouvain, Louvain-la-Neuve, Belgium*

✉ wafa.aissa@uclouvain.be | ⊚ wafaAissaa | in wafa-aissa

## Con**tents**

## Edu**cation**

**Doctorate - Computer science -** Neural Module Networks for Compositional Visual Reasoning. ⬈          *Paris, France*

**Conservatoire National des Arts et Métiers (CNAM) - CEDRIC Lab - Vertigo**          *Jun 2019 - Dec 2023*

- Compositional Visual Reasoning.
- Visual Question Answering.
- Vision Language Models.
- Director: Michel Crucianu ✉ michel.crucianu@lecnam.net
- Co-supervisor: Marin Ferecatu ✉ marin.ferecatu@lecnam.net

**Masters - Artificial intelligence.** Grade: 16.23/20, Rank: 3/33          *Paris, France*

**Sorbonne Université UPMC - Faculty of Science and Engineering, 2nd-year Master** ⬈

**Université Paris Cité - Graduate School of Artificial Intelligence and Data Science, 1st-year Master** ⬈          *Sep 2016 - Sep 2018*

Relevant courses:
- Machine and Deep Learning.
- Natural Language Processing and Information retrieval.
- Computer Vision.
- Symbolic AI.

Relevant school projects:
- Deep Learning for Computer Vision.
- Paper Reproductibility challenge.
- Learning to play Hanabi using Reinforcement Learning.

**Licence - Computer Science and Business Administration.** Grade: 16.72/20, Rank: 1/60          *Tunis, Tunisia*

**Institut des hautes études commerciales de Carthage (IHEC)**          *Sep 2013 - Sep 2016*

# Experience

**Postdoc** *Louvain-la-Neuve, Belgium*
**Université Catholique de Louvain (UCLouvain)** *Nov 2024 - current*

- Machine/Deep learning for Text Complexity assessment and Simplification.
- AI for Accessibility and Inclusive Language Technologies.
- Member of the iRead4Skills project. Funded by the European Union. [↗]
- Promotor: Thomas François ✉ thomas.francois@uclouvain.be

**Teaching assistant** *Paris, France*
**Conservatoire National des Arts et Mériers (CNAM)** *Sep 2022 - Aug 2024*

- Machine/deep learning, AI for multimedia data, documentary and distributed databases, UML, SQL.

**Research internship:** Reinforcement Learning for Search Oriented Conversational Systems. *Paris, France*
**Laboratoire d'informatique de Paris 6 (LIP6) - MLIA team** *Mar - Aug 2018*

- State of the art bibliography of neural machine translation and conversational systems.
- Development of a reinforcement learning-driven translation model for search oriented conversational systems.
- Comparative evaluation of the proposed model on TREC datasets (Ad-hoc, Web, Terabyte...).
- Supervised by Laure Soulier and Ludovic Denoyer.

**Research internship:** Efficient word vector learning through corpus sampling. *Paris, France*
**Laboratoire d'informatique de Paris Descartes (LIPADE) - MLDS team** *Jun - Jul 2017*

- Bibliographic search on learning word embeddings.
- Investigating the impact of word frequency on similarity.
- Perspectives of improvement of "Word2Vec".
- Supervised by Severine Affeldt.

**Internship -** Social network platform for business. *Tunis, Tunisia*
**Informaticien Tunisien** *Feb–May 2016*

**Internship -** Operating System Administration. *Tunis, Tunisia*
**At-tounissiya Internet** *Jun–Jul 2014*

# Summer/Winter schools

**Participant** *Leuven, Belgium*
**Leuven.AI Conference** [↗] *Sep 2025*

- Keynote lectures and thematic tracks on reinforcement learning, reasoning, computer vision, and natural language processing.

**Participant** *Amsterdam, Netherlands*
**ELLIS Winter School on Foundation Models** [↗] *Mars 2025*

- Explore emerging European perspectives on foundation models, highlighting research directions.

**Participant** *Valletta, Malta*
**International Summer School on AI and Games 2024** [↗] *Jun 2024*

- Introduction to AI and games, LLMs for games, uses of AI for playing games, generating content for games, modeling players.
- Award: Winning team of the game AI jam.

**Participant** *Cambridge, England*
**International Summer School on AI and Games 2023** [↗] *Jun 2023*

- Award: Winning team of the game AI jam.

# Publications

## Assessing French Readability for Adults with Low Literacy: A Global and Local Perspective [↗]

Wafa Aissa*, Thibault Bañeras-Roux*, Elodie Vanzeveren, Lingyun Gao, Rodrigo Wilkens and Thomas François.

*Equal contribution

*EMNLP*

**Accepted at the 2025 main Conference on Empirical Methods in Natural Language Processing EMNLP (to appear)**

*2025*

Abstract: This study presents a novel approach to assessing French text readability for adults with low literacy skills, addressing both global (full-text) and local (segment-level) difficulty. We introduce a dataset of 461 texts annotated using a difficulty scale developed specifically for this population. Using this corpus, we conducted a systematic comparison of key readability modeling approaches, including machine learning techniques based on linguistic variables, fine-tuning of CamemBERT, a hybrid approach combining CamemBERT with linguistic variables, and the use of generative language models (LLMs) to carry out readability assessment at both global and local levels.

## The iRead4Skills Intelligent Complexity Analyzer [↗]

Wafa Aissa*, Raquel Amaro, David Antunes, Thibault Bañeras-Roux, Jorge Baptista, Alejandro Catala, Luís Correia, Thomas François, Marcos Garcia, Mario Izquierdo-Álvarez, Nuno Mamede, Vasco Martins, Miguel Neves, Eugénio Ribeiro, Sandra Rodriguez Rey, Elodie Vanzeveren.  *Alphabetical order.

*EMNLP - System demonstrations*

**Accepted at the 2025 Conference on Empirical Methods in Natural Language Processing EMNLP - System demonstrations (to appear)**

*2025*

Abstract: We present the iRead4Skills Intelligent Complexity Analyzer, an open-access platform specifically designed to assist educators and content developers in addressing the needs of low-literacy adults by analyzing and diagnosing text complexity. This multilingual system integrates a range of Natural Language Processing (NLP) components to assess input texts along multiple levels of granularity and linguistic dimensions in Portuguese, Spanish, and French. It assigns four tailored difficulty levels using state-of-the-art models, and introduces four diagnostic yardsticks—textual structure, lexicon, syntax, and semantics—offering users actionable feedback on specific dimensions of textual complexity. Each component of the system is supported by experiments comparing alternative models on manually annotated data.

## Modélisation de la lisibilité en français pour les personnes en situation d'illettrisme [↗]

Wafa Aissa*, Thibault Bañeras-Roux*, Elodie Vanzeveren, Lingyun Gao, Alice Pintard,  Rodrigo Wilkens and Thomas François

*TALN*

*Equal contribution

**La 32e Conférence sur le Traitement Automatique des Langues Naturelles (TALN)**

*2025*

Abstract: We present a new French readability formula specifically designed for people with low literacy. To this end, we built a corpus of 461 texts annotated according to a specialized difficulty scale tailored to this audience. We then conducted a systematic comparison of the main approaches to readability, including machine learning based on linguistic variables, CamemBERT fine-tuning, a hybrid approach combining CamemBERT and linguistic variables, and generative large language models (LLMs). An in-depth analysis of these models and their performance is carried out in order to assess their applicability in real-life contexts.

## Multimodal Representations for Teacher-Guided Compositional Visual Reasoning [↗]

Wafa Aissa, Marin Ferecatu and Michel Crucianu

*ACIVS*

**In Proceedings of the 21st international conference on Advanced Concepts for Intelligent Vision Systems**

*2023*

Abstract: Neural Module Networks (NMN) are a compelling method for visual question answering, enabling the translation of a question into a program consisting of a series of reasoning sub-tasks that are sequentially executed on the image to produce an answer. NMNs provide enhanced explainability compared to integrated models, allowing for a better understanding of the underlying reasoning process. To improve the effectiveness of NMNs we propose to exploit features obtained by a large-scale cross-modal encoder. Also, the current training approach of NMNs relies on the propagation of module outputs to subsequent modules, leading to the accumulation of prediction errors and the generation of false answers. To mitigate this, we introduce an NMN learning strategy involving scheduled teacher guidance. Initially, the model is fully guided by the ground-truth intermediate outputs, but gradually transitions to an autonomous behavior as training progresses. This reduces error accumulation, thus improving training efficiency and final performance. We demonstrate that by incorporating cross-modal features and employing more effective training techniques for NMN, we achieve a favorable balance between performance and transparency in the reasoning process.

### Curriculum Learning for Compositional Visual Reasoning ↗

Wafa Aissa, Marin Ferecatu and Michel Crucianu

**In Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications**

*VISAPP - VISIGRAPP*

*2023*

Abstract: Visual Question Answering (VQA) is a complex task requiring large datasets and expensive training. Neural Module Networks (NMN) first translate the question to a reasoning path, then follow that path to analyze the image and provide an answer. We propose an NMN method that relies on predefined cross-modal embeddings to "warm start" learning on the GQA dataset, then focus on Curriculum Learning (CL) as a way to improve training and make a better use of the data. Several difficulty criteria are employed for defining CL methods. We show that by an appropriate selection of the CL method the cost of training and the amount of training data can be greatly reduced, with a limited impact on the final VQA accuracy. Furthermore, we introduce intermediate losses during training and find that this allows to simplify the CL strategy.

### Modèle de compréhension du besoin en information pour la RI conversationnelle ↗

Wafa Aissa, Laure Soulier et Ludovic Denoyer

**La 16ème édition de CORIA (Conférence en Recherche d'Information et Applications)**

*CORIA*

*2019*

Abstract: IR is based on a standard framework that queries document collections through an in- formation need expressed as a set of keywords. Our contribution aims at overpassing this usual paradigm by starting the retrieval process from the natural language expression of the informa- tion need; giving rise to a new generation of IR systems based on conversational features (also called "search-oriented conversational systems"). Therefore, the first step focuses on the query formulation from the information need expressed in natural language. In this paper, we pro- pose a query formulation model able of 1) translating natural language expressions to keyword queries in a supervised manner, and 2) injecting relevance feedback in the learning process through reinforcement learning technics. To overcome the lack of training data, we consider the translation model as a word selection process. Our model is evaluated on two TREC collections to demonstrate its effectiveness.

### A Reinforcement Learning-driven Translation Model for Search-Oriented Conversational Systems ↗.

Wafa Aissa, Laure Soulier and Ludovic Denoyer

**Best paper award. In Proceedings of the 2018 EMNLP Workshop SCAI: The 2nd International Workshop on Search-Oriented Conversational AI**

*SCAI @ EMNLP*

*2018*

Abstract: Search-oriented conversational systems rely on information needs expressed in natural language (NL). We focus here on the understanding of NL expressions for building keyword-based queries. We propose a reinforcement-learning-driven translation model framework able to 1) learn the translation from NL expressions to queries in a supervised way, and, 2) to overcome the lack of large-scale dataset by framing the translation model as a word selection approach and injecting relevance feedback as a reward in the learning process. Experiments are carried out on two TREC datasets. We outline the effectiveness of our approach.

# Doctoral thesis summary

The context of my PhD thesis is compositional visual reasoning. When presented with an image and a question pair, our objective is to have neural networks models answer the question by following a reasoning chain defined by a program. We assess the model's reasoning ability through a Visual Question Answering (VQA) setup. Compositional VQA breaks down complex questions into modular easier sub-problems. These sub-problems include reasoning skills such as object and attribute detection, relation detection, logical operations, counting, and comparisons. Each sub-problem is assigned to a different module. This approach discourages shortcuts, demanding an explicit understanding of the problem. It also promotes transparency and explainability. Neural module networks (NMN) are used to enable compositional reasoning. The framework is based on a generator-executor framework, the generator learns the translation of the question to its function program. The executor instantiates a neural module network where each function is assigned to a specific module. We also design a neural modules catalog and define the function and the structure of each module. The training and evaluations are conducted using the pre-processed GQA dataset [3], which includes natural language questions, functional programs representing the reasoning chain, images, and corresponding answers. The research contributions revolve around the establishment of an NMN framework for the VQA task. One primary contribution involves the integration of vision and language pre-trained (VLP) representations into modular VQA. This integration serves as a "warm-start" mechanism for initializing the reasoning process. The experiments demonstrate that cross-modal vision and language representations outperform uni-modal ones. This utilization enables the capture of intricate relationships within each individual modality while also facilitating alignment between different modalities, consequently enhancing overall accuracy of our NMN. Moreover, we explore various training techniques to enhance the learning process and improve cost-efficiency. In addition to optimizing the modules within the reasoning chain to collaboratively produce accurate answers, we introduce a teacher-guidance approach to optimize the intermediate modules in the reasoning chain. This ensures that these modules perform their specific reasoning sub-tasks without taking shortcuts or compromising the reasoning process's integrity. We propose and implement several teacher-guidance techniques, one of which draws inspiration from the teacher-forcing method commonly used in sequential models. Comparative analyses demonstrate the advantages of our teacher-guidance approach for NMNs, as detailed in our paper [1]. We also introduce a novel Curriculum Learning (CL) strategy tailored for NMNs to reorganize the training examples and define a

start-small training strategy. We begin by learning simpler programs and progressively increase the complexity of the training programs. We use several difficulty criteria to define the CL approach. Our findings demonstrate that by selecting the appropriate CL method, we can significantly reduce the training cost and required training data, with only a limited impact on the final VQA accuracy. This significant contribution forms the core of our paper [2]. [1] W. Aissa, M. Ferecatu, and M. Crucianu. Curriculum learning for compositional visual reasoning. In Proceedings of VISIGRAPP 2023, Volume 5: VISAPP, 2023. [2] W. Aissa, M. Ferecatu, and M. Crucianu. Multimodal representations for teacher-guided compositional visual reasoning. In Advanced Concepts for Intelligent Vision Systems, 21st International Conference (ACIVS 2023). Springer International Publishing, 2023. [3] D. A. Hudson and C. D. Manning. GQA: A new dataset for real-world visual reasoning and compositional question answering. 2019.

# Paper presentations

**National conference: TALN 2025** *Marseille - France*
**Modélisation de la lisibilité en français pour les personnes en situation d'illettrisme** *July 2025*
Wafa Aissa*, Thibault Bañeras-Roux*, Elodie Vanzeveren, Lingyun Gao, Alice Pintard, Rodrigo Wilkens and Thomas François.
*Equal Contribution.

**International conference: ACIVS 2023** *Kumamoto, Japan*
**Multimodal Representations for Teacher-Guided Compositional Visual Reasoning** *Aug 2023*
Wafa Aissa, Marin Ferecatu and Michel Crucianu

**International conference: VISAPP 2023** *Lisbon - Portugal*
**Curriculum Learning for Compositional Visual Reasoning** *Feb 2023*
Wafa Aissa, Marin Ferecatu and Michel Crucianu

**National conference: CORIA 2019** *Lyon - France*
**Modèle de compréhension du besoin en information pour la RI conversationnelle** *Mar 2019*
Wafa Aissa, Laure Soulier and Ludovic Denoyer

**International workshop : SCAI workshop at EMNLP 2018** *Brussels - Belgium*
**A Reinforcement Learning-driven Translation Model for Search-Oriented Conversational Systems**
**Award: best paper** *Oct 2018*
Wafa Aissa, Laure Soulier and Ludovic Denoyer

# Skills

**Data Science and applications**
ML, DL, Visual Reasoning, Information Retrieval, NLP, CV, Learning representations, RL, LLM, VLM, Transformer, CNN, RNN, AE, Clustering.

**Machine learning and deep learning libraries**
Pytorch, Keras, scikit-learn, Pandas, Numpy, Ollama, Matplotlib, TensorFlow, Pyspark.

**Programming languages**
Python, R, Prolog, Matlab, sql, C++, java.

**Natural languages**
English, French, arabic.

# Volunteer activities

**AIESEC, International Association of Students in Economic and Commercial Sciences** *Tunis, Tunisia*
**Business Development and marketing manager.** *Oct 2015 - Aug 2016*
- Team work and leadership.
- Sales training.
- Graphic Design and Video Editing.
- Organizing committee member of seminars.