

# Chadi HELWE

✉ chadi.helwe@telecom-paris.fr | ☎ +33 7 54 17 39 45 | 🌐 chadihelwe | 📷 ChadiHelwe | 📄 Chadi Helwe

## EDUCATION

---

JULY 2024	<b><i>Doctor of Philosophy in Artificial Intelligence</i></b> - Institut Polytechnique de Paris	
2021	Ph.D's Thesis: Evaluating and Improving the Reasoning Abilities of Language Models Advisors: Prof. Fabian Suchanek and Prof. Chloé Clavel	
2017	<b><i>Master of Science in Computer Science</i></b> - American University of Beirut	GPA: 3.68/4
2015	Master's Thesis: Arabic Named Entity Recognition via Deep Co-learning Advisor: Prof. Shady Elbassuoni	
2014	<b><i>Bachelor of Science in Computer Science</i></b> - Notre Dame University - Louaize	GPA: 3.63/4
2010	Final Year Project: Educally - An Educational Social Network Advisor: Prof. Marie Khair	

## SELECTED WORK EXPERIENCE

---

PRESENT	<b><i>Graduate Research Assistant</i></b> - Institut Polytechnique de Paris, France	
JAN. 2021	<ul style="list-style-type: none"><li>• <b>MAFALDA: A Benchmark and Comprehensive Study of Fallacy Detection and Classification [1]:</b><ul style="list-style-type: none"><li>– Designed a novel taxonomy of fallacies.</li><li>– Proposed an annotation framework customized for fallacy classification and introduced an evaluation metric to address subjectivity within fallacy classification.</li><li>– Built a benchmark to assess the performance of language models in detecting and categorizing fallacies.</li></ul></li><li>• <b>TINA: Textual Inference with Negation Augmentation [2]:</b><ul style="list-style-type: none"><li>– Proposed a probabilistic definition of Textual Entailment (TE) and used it to augment TE datasets with new entailment relationships automatically.</li><li>– Explored various models fine-tuned with augmented datasets using an unlikelihood loss to enhance language model robustness in Textual Entailment tasks, particularly those involving negation examples.</li></ul></li><li>• <b>LogiTorch: A PyTorch-based Library for Logical Reasoning on Natural Language [3]:</b><ul style="list-style-type: none"><li>– Developed a Python library on top of Pytorch for logical reasoning with diverse benchmarks, models, and utility functions, simplifying dataset utilization and model training.</li></ul></li><li>• <b>Reasoning with Transformer-based Models: Deep Learning, but Shallow Reasoning [4]:</b><ul style="list-style-type: none"><li>– Wrote a survey paper on the performance of Transformers on various reasoning tasks, highlighting both their successes and limitations, including empirical and theoretical aspects.</li></ul></li></ul>	
DEC 2020	<b><i>Research Engineer</i></b> - Institut Polytechnique de Paris, France	
SEP 2020	Investigated the performance of Transformers on different reasoning tasks [4].	
AUG 2020	<b><i>Research Assistant</i></b> - American University of Beirut, Lebanon	
OCT 2017	<ul style="list-style-type: none"><li>• <b>Retrieving Textual Evidence for Knowledge Graph Facts using Deep Learning:</b><ul style="list-style-type: none"><li>– A project in collaboration with Aalborg University (Denmark) and Hacettepe University (Turkey).</li><li>– Investigated Transformer-based models trained with a dataset generated using distant supervision to rank passages based on their relevance to a given fact in the form of a Resource Description Framework (RDF) triple.</li></ul></li><li>• <b>Predicting Arabic Blog Credibility using Deep Co-learning [5]:</b><ul style="list-style-type: none"><li>– Implemented and evaluated a novel semi-supervised learning approach based on an algorithm called Co-training, which was adapted to the context of deep learning for the task of Arabic blog's credibility prediction, which can be trained using a small labeled dataset and a large unlabeled dataset.</li></ul></li><li>• <b>Automated Detection and Measurement of Corneal Haze and Demarcation Line in OCT Images [6, 7, 8, 9, 10]:</b><ul style="list-style-type: none"><li>– A project in collaboration with the Department of Ophthalmology (American University of Beirut Medical Center) and the ELZA Institute in Zurich (Switzerland).</li><li>– Redesigned an outdated software to detect and measure corneal haze and demarcation line in different types of Optical Coherence Tomography (OCT) images.</li><li>– Added new features and introduced deep learning methods to the software.</li><li>– Developed OCTAnalysis.com, a web interface of the software in Django/Python and Postgres SQL.</li><li>– Proposed and implemented a semi-weakly supervised learning approach to segment the area between the top boundary of a cornea and the demarcation line in OCT images.</li><li>– Developed a SegNet neural network to detect the boundaries of a cornea in OCT images and a VGG-16 neural network to detect artifacts in OCT images.</li><li>– Supervised two undergrad students who annotated a large dataset of OCT images.</li><li>– Developed an image segmentation tool for labeling.</li></ul></li></ul>	

SEPT 2017	<b>Graduate Research Assistant</b> - American University of Beirut, Lebanon
JUNE 2016	<ul style="list-style-type: none"> <li>• <b>Arabic Named Entity Recognition via Deep Co-learning</b> [11, 12]: <ul style="list-style-type: none"> <li>– Built a supervised deep learning model that infers the name entities' class in a Wikipedia article by classifying their Wikipedia pages into one of four classes: person, location, organization, or miscellaneous.</li> <li>– Generated a large dataset of partially annotated Wikipedia articles for the task of Arabic Named Entity Recognition (NER).</li> <li>– Proposed a novel semi-supervised learning approach based on an algorithm called Co-training, which was adapted to the context of deep learning for the task of Arabic NER and which can be trained using a small fully annotated dataset and a large partially annotated dataset.</li> <li>– Evaluated our proposed Deep Co-learning algorithm approach on three Arabic NER datasets.</li> </ul> </li> <li>• <b>ICD and CCS Coding using Deep Learning</b> [13]: <ul style="list-style-type: none"> <li>– A project in collaboration with the Department of Emergency Medicine (American University of Beirut Medical Center).</li> <li>– Designed and implemented a deep neural network architecture to predict the International Classification of Diseases (ICD) code and Clinical Classifications Software (CCS) single-level code of a discharge diagnosis.</li> </ul> </li> <li>• <b>Methodical Evaluation of Arabic Word Embeddings</b> [14]: <ul style="list-style-type: none"> <li>– A project in collaboration with Qatar University.</li> <li>– Built the first word analogy benchmark designed specifically for Arabic word embeddings.</li> <li>– Implemented different Long Short-term Memory recurrent neural network architectures to evaluate Arabic word embeddings on two NLP tasks: Document Classification and Named Entity Recognition.</li> </ul> </li> <li>• <b>Adaptive QoS for Spark Applications</b> [15]: <ul style="list-style-type: none"> <li>– Developed an adaptive quality management/selection method for Spark applications.</li> <li>– Implemented different QoS policies in Java.</li> </ul> </li> </ul>

## TEACHING EXPERIENCE

AUG 2023	<b>Teaching Assistant</b> - Institut Polytechnique de Paris, France
JULY 2021	<ul style="list-style-type: none"> <li>• Bases de données</li> <li>• Données du Web</li> <li>• Mise en pratique Données du Web</li> <li>• Mining of Large Datasets</li> </ul>
MAY 2020	<b>Teaching Assistant</b> - American University of Beirut, Lebanon
FEB 2015	<ul style="list-style-type: none"> <li>• Artificial Intelligence</li> <li>• Introduction to Programming</li> <li>• Compiler Construction (graduate course)</li> <li>• Machine Learning (graduate course)</li> </ul>
JUNE 2013	<b>Teaching Assistant</b> - Notre Dame University - Louaize, Lebanon
FEB 2013	<ul style="list-style-type: none"> <li>• Program Design and Data Abstraction I</li> <li>• Program Design and Data Abstraction II</li> </ul>

## OPEN SOURCE PROJECTS

PRESENT	<b>LogiTorch</b> - Creator and Maintainer
DEC 2021	LogiTorch [3] is a PyTorch-based library that includes different logical reasoning benchmarks, different models, as well as utility functions such as co-reference resolution. The library allows researchers and developers to easily use a logical reasoning dataset and train logical reasoning models with just a few lines of code.

## SKILLS

Python, Java, C++, PyTorch, Pandas, Numpy, Keras, L<sup>A</sup>T<sub>E</sub>X

## REVIEWER

AI Review, NAACL 2024, ACL 2023, EMNLP 2022, EACL 2022 (External Reviewer)

## AWARDS

2018	Recipient of the Best Computer Science Graduate Student Award from the American University of Beirut.
2015	Awarded a full graduate assistantship from the American University of Beirut.
2014	Graduated from Notre Dame University - Louaize with high distinction.
2013	Awarded a scholarship from Notre Dame University - Louaize.
2013	Dean's List for the Spring semester.
2012	Dean's List for the Spring and the Fall semesters.

## SUMMER SCHOOLS ATTENDED

---

- 2023 Oxford Machine Learning Summer School (OxML), Oxford, United Kingdom.
- 2021 Machine Learning Summer School (MLSS), Taipei, Taiwan.
- 2021 4th Advanced Course on Data Science and Machine Learning (ACDL), Tuscany, Italy.

## PUBLICATIONS

---

- [1] Chadi Helwe, Tom Calamai, Pierre-Henri Paris, Chloé Clavel, and Fabian Suchanek. “MAFALDA: A Benchmark and Comprehensive Study of Fallacy Detection and Classification”. In: *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. 2024.
- [2] Chadi Helwe, Simon Coumes, Chloé Clavel, and Fabian Suchanek. “TINA: Textual Inference with Negation Augmentation”. In: *Findings of the Association for Computational Linguistics: EMNLP 2022*. 2022.
- [3] Chadi Helwe, Chloé Clavel, and Fabian Suchanek. “LogiTorch: A PyTorch-based library for logical reasoning on natural language”. In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP): System Demonstrations*. 2022.
- [4] Chadi Helwe, Chloé Clavel, and Fabian M Suchanek. “Reasoning with Transformer-based Models: Deep Learning, but Shallow Reasoning”. In: *3rd Conference on Automated Knowledge Base Construction*. 2021.
- [5] Chadi Helwe, Shady Elbassuoni, Ayman Al Zaatari, and Wassim El-Hajj. “Assessing Arabic Weblog Credibility via Deep Co-learning”. In: *Proceedings of the Fourth Arabic Natural Language Processing Workshop*. Association for Computational Linguistics, 2019.
- [6] Shady Awwad, Lily Chacra, Chadi Helwe, Ahmad Dhaini, Farhad Hafezi, Emilio Torres, and Madeleine Yehia. “Accelerated Corneal Cross-linking Using 20 Minutes Riboflavin With Hydroxypropyl Methylcellulose Soaking Time Versus Conventional Cross-linking”. In: *International CXL Experts Meeting 2019*. 2019.
- [7] Shady Awwad, Lily Chacra, Chadi Helwe, Ahmad Dhaini, Farhad Hafezi, Emilio Torres, and Talar Telvizian. “Corneal Haze After Cross-linking for Keratoconus Eyes With and Without Mitomycin C Application”. In: *International CXL Experts Meeting 2019*. 2019.
- [8] Chadi Helwe, Shady Elbassuoni, Ahmad Dhaini, Lily Chacra, and Shady Awwad. “A Deep Learning Approach to Detect the Demarcation Line in OCT Images”. In: *Annual Conference on Medical Image Understanding and Analysis*. Springer. 2020.
- [9] Shady T Awwad, Lily M Chacra, Chadi Helwe, Ahmad R Dhaini, Talar Telvizian, Julien Torbey, Maamoun Abdul Fattah, Emilio A Torres-Netto, Farhad Hafezi, and Rohit Shetty. “Mitomycin C application after corneal cross-linking for keratoconus increases stromal haze”. In: *Journal of Refractive Surgery* (2021).
- [10] Lily M Chacra, Chadi Helwe, Jad F Assaf, Madeleine Yehia, Serge Baroud, Emilio A Torres-Netto, Farhad Hafezi, and Shady T Awwad. “Accelerated corneal crosslinking with 20-soaking hydroxypropyl methyl cellulose/riboflavin vs conventional crosslinking with 30-soaking dextran/riboflavin”. In: *Journal of Cataract & Refractive Surgery* 50.3 (2024), pp. 236–243.
- [11] Chadi Helwe and Shady Elbassuoni. “Arabic named entity recognition via deep co-learning”. In: *Artificial Intelligence Review* (2019).
- [12] Chadi Helwe, Ghassan Dib, Mohsen Shamas, and Shady Elbassuoni. “A Semi-Supervised BERT Approach for Arabic Named Entity Recognition”. In: *Proceedings of the Fifth Arabic Natural Language Processing Workshop*. Association for Computational Linguistics, 2020.
- [13] Chadi Helwe, Shady Elbassuoni, Mirabelle Geha, Eveline Hitti, and Carla Makhoul Obermeyer. “CCS coding of discharge diagnoses via deep neural networks”. In: *Proceedings of the 2017 International Conference on Digital Health*. 2017.
- [14] Mohammed Elrazzaz, Shady Elbassuoni, Khaled Shaban, and Chadi Helwe. “Methodical evaluation of arabic word embeddings”. In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 2017.
- [15] Bilal Abi Farraj, Wael Al Rahal Al Orabi, Chadi Helwe, Mohamad Jaber, Mohamad Omar Kayali, and Mohamed Nassar. “Reconfigurable and Adaptive Spark Applications”. In: *Proceedings of the 7th International Conference on Cloud Computing and Services Science*. 2017.