Gautier Dagan

London, UK
gautier.dagan@gmail.com
gautier.tech
+44-742-383-6797

EDUCATION

University of Amsterdam — Amsterdam, The Netherlands
Master of Science in Artificial Intelligence; 8.55/10; Cum Laude — *Sept. 2017 – Aug. 2019*

Northwestern University — Evanston, IL
BA in Computer Science, Economics and International Studies (Triple Major); GPA: 3.59 — *Aug. 2013 – June 2017*

EXPERIENCE

Research Grid — London, UK
Machine Learning Engineer (Part-Time) — *Jan. 2021 - Present*
I am leading the Machine Learning at Research Grid. I architect text summarization and classification pipelines, redesign and refactor existing projects, manage interns and assist in the hiring process.

Entrepreneur First — London, UK
London Cohort (LD15) — *Sept. 2020 - Dec. 2020*
EF funds individuals and help build co-founding teams, develop ideas, and accelerate through fundraising from the world's best investors. Throughout the programme, I conducted 50+ customer interviews and validated 3 business ideas over the course of 3 months.

Mansa — London, UK
Machine Learning Engineer — *Sept. 2019 - Sept. 2020*
As the first employee at Mansa, I was heavily involved in all aspects of development. During my time there, I created a simple credit score, with a current 0% default rate, to estimate risk based on combined personal and business financial health indicators. I developed and architected the data pipeline from scratch using AWS, Docker, FastAPI, and Serverless. I also implemented an in-house OCR solution for bank statements of 10 major French banks to reduce on-boarding attrition and augment Open Banking coverage.

Institute of Logic, Language and Computation — Amsterdam, The Netherlands
Graduate Student Researcher under the i-machine-think research group — *Jan. 2019 – Sept. 2019*
As a graduate student researcher at the ILLC, I collaborated with Dr Elia Bruni and Dr Dieuwke Hupkes on several research projects. I conducted an individual project on attentive RNN architectures for language models. I assisted in the development and testing of a positional attention mechanism with Yann Dubois, which was later published in ACL 2020. I conducted my Master's thesis on language emergence in multi-agent referential games, published in EACL 2021. Finally, I also supervised peers and guided research into the design of analysis methods for emergent languages.

University of Amsterdam — Amsterdam, The Netherlands
Teaching Assistant in Information Retrieval — *Feb. 2019 – Apr. 2019*
As a TA, I covered topics such as distributed crawling, Term/Semantic Search and Learning to Rank. I also supervised students to create a search engine from scratch and assisted in the grading of projects.

Zcruit — Evanston, IL
Co-founder — *Feb. 2016 – Jan. 2018*
During my Bachelor's, I co-founded Zcruit – a predictive analytics and data-aggregation start-up focusing on college football recruiting. Since then, the team has grown to 5 employees and 70+ enterprise clients.

PUBLICATIONS

- **Gautier Dagan**, Dieuwke Hupkes, and Elia Bruni *Co-evolution of language and agents in referential games* EACL 2021 |preprint |code
  We introduce the Language Transmission Simulator to model both cultural and architectural evolution in a population of agents. As our core contribution, we empirically show that the optimal situation is to also take into account the learning biases of language learners and thus let language and agents *co-evolve.*When

we allow the agent population to evolve through architectural evolution, we achieve across the board improvements on all considered metrics and surpass the gains made with *cultural transmission*. These results stress the importance of studying the underlying agent architecture and pave the way to investigate the *co-evolution* of language and agent in language emergence studies.

- Yann Dubois, **Gautier Dagan**, Dieuwke Hupkes, and Elia Bruni *Location Attention for Extrapolation to Longer Sequences* ACL 2020 |paper
  Neural networks are often unable to extrapolate patterns beyond the seen data, even when the abstractions required for such patterns are simple. In this paper, we first review the notion of extrapolation, why it is important and how one could hope to tackle it. We then focus on a specific type of extrapolation which is especially useful for natural language processing: generalisation to sequences that are longer than the training ones. We hypothesise that models with a separate content- and location-based attention are more likely to extrapolate than those with common attention mechanisms. We empirically support our claim for recurrent seq2seq models with our proposed attention on variants of the Lookup Table task.

- Gabriele Bani, Davide Belli, **Gautier Dagan**, Alexander Geenen, Andrii Skliar, Aashish Venkatesh, Tim Baumgartner, Elia Bruni, Raquel Fernández *Adding object detection skills to visual dialogue agents* ECCV Workshop 2018 |paper
  We equip a dialogue agent that asks questions about a visual scene with object detection skills. We use Mask R-CNN object features as a replacement for ground-truth annotations in a Guesser module. We show that our system exploits the Mask R-CNN object features, in contrast to the original Guesser augmented with global, VGG features. Furthermore, by automating the object detection in GuessWhat?!, we open up a spectrum of opportunities, such as playing the game with new, non-annotated images and using the more granular visual features to condition the other modules of the game architecture.

Projects

- **RIIID Knowledge Tracing**
  Kaggle Competition |code   Dec. 2020 - Jan. 2020
  Finished 21st out of 3395 teams (top 0.6%) in the Kaggle RIIID competition. The Knowledge Tracing task consists of understanding how a student learns, through their past interactions (answers), to predict the outcome of future questions. In this competition, I implemented both Light Gradient Boosting Machines and custom Transformer architectures to obtain state-of-the-art results.

- **Machine**
  Open Source Library |code   Jan. 2019 - Aug. 2019
  Contributor to the *i-machine-think* machine repository, an open source python library for Sequence-to-Sequence based on the IBM library pytorch-seq2seq.

- **Global and Local Critical Policy Learning for Abstractive Summarization**
  Information Retrieval II Research Project |paper |code   Dec. 2018
  Traditional Abstractive Summarization models suffer from training directly using a maximum likelihood approach, which is known to decrease the models' abstractive power and therefore generate summaries which are less human-like in abstraction. We attempt to solve this by incorporating a local (word level) and global (sentence level) loss weighting using the ROUGE metric directly as a reward through Policy Gradient optimisation. By evaluating the sub-sequences of the generated summary, we obtain the gain they provide in ROUGE score for the entire summary and weigh the loss locally and globally to reflect this.

Skills

- **Programming Languages**: Python, NodeJS, Typescript, SQL, Latex
- **Technologies**: Git, Pytorch, AWS, Serverless, Docker
- **Certifications**: AWS Machine Learning Speciality (Issued August 2020), AWS Solutions Architect Associates (Issued June 2020)
- **Languages**: English (native), French (native), Spanish (A2)
- **Competitions**: Kaggle Competition Expert (**top 1%**), RoboCup IranOpen 2018, Autonomous Robot Design Competition (2016 and 2017)