

Agathe Balayn

Curriculum Vitae

Koornmarkt 82A, 2611EJ, Delft, the Netherlands
08/01/1995, Paris, France
☎ +33(0)6.99.55.72.23
✉ a.m.a.balayn@tudelft.nl

Education

- 04/2019 - **PhD candidate in Computer Science.**, *Delft University of Technology*, The Netherlands, Department: Web Information Systems.
Supervisors: Christoph Lofi, Alessandro Bozzon, Geert-Jan Houben
- 09/2016 - **MSc in Computer Science. Data Science and Technology track**, *Delft University of Technology*, The Netherlands, **GPA: 8.72/10, 9/10 for the thesis project.**
09/2018
 - Focus on Artificial Intelligence, Machine and Deep Learning, Human-Computer Interaction.
 - Completion of the Honours Programme of the university (additional 20 ECTS).
- 09/2014 - **MSc in Embedded Systems**, *ENSTA ParisTech (Top French Grande École, graduate school in Engineering)*, France, (Program leading to a "Diplôme d'ingénieur").
09/2018
 - Strong component of Control, Informatics and Signal.
 - Graduated first year 2nd of the class out of 144 students. **GPA: 4.0/4.0.**
- 2012-2014 **Preparatory class**, *Lycée du Parc*, France, 2 years of intensive undergraduate program (Mathematics, Physics, Engineering Sciences) preparing for the national competitive entrance examination to graduate schools in engineering.

Professional experiences

- 09/2018 - **Researcher at the IBM Center for Advanced Studies and at the TU Delft**, *the Netherlands*,
03/2019 Investigation of the fairness of Machine Learning pipelines for the inference of subjective labels.
 - Experimentations and writing of scientific publications and of a literature survey. Publication at the Human-Centered Machine Learning of CHI 2019.
 - Contribution to the IBM Research toolkit *AI Fairness 360* with newly designed metrics.
- 11/2017 - **Graduate Intern at IBM Benelux (Master's thesis)**, *Center for Advanced Studies, Amsterdam*,
09/2018 *the Netherlands*, Study of biases and fairness in crowdsourced data and Machine Learning models for the prediction of subjective properties, with the use-case of sentence toxicity prediction.
 - Design and implementation of a new fairness metric. Experiments on crowdsourcing techniques for training dataset collection, and on Machine and Deep Learning models for fairer predictions.
 - Publication at the CrowdBias workshop of HCOMP2018.
- 08/2017 - **Research Intern at the Honda Research Institute (HRI-JP)**, *Wako, Japan*.
10/2017
 - Creation of encoding schemes for sign language annotations. Design, implementation and evaluation of Deep Learning models (mainly RNN, LSTM neural networks) for Japanese sign language synthesis and recognition based on Kinect and motion capture data.
 - Publication at RO-MAN 2018.
- 05/2016 - **Research Assistant at the Research Institute for Cognition and Robotics (CoR-Lab)**, *Bielefeld University, Germany*.
07/2016
 - Design, implementation and evaluation of an active-compliance control mode using ELM neural networks and model-space learners for an industrial lightweight robotic arm (Universal Robots UR5), able to cope with unknown weights at the end-effector.
 - Publication at SIMPAR 2016.
 - Obtained one individual scholarship, based on merit (Erasmus Plus).
- 08/2015 **Summer trainee at the company Hakuba Lion Adventure**, *Hakuba, Japan*.
Accompanied groups of tourists to outdoor activities (canyoning, kayaking, basic ski lessons), set up air balloons, barista. Entirely Japanese-speaking team.

Publications

- VLDBJ 2021 Balayn, Agathe, et al. "Treatment of Bias in Data Management for Data-Driven Decision Support Systems [vision paper]" *at the Journal of Very Large Database Systems 2021* .
- WWW 2021 Balayn, Agathe, et al. "What do You Mean? Interpreting Image Classification with Crowdsourced Concept Extraction and Analysis" *at the Web Conference 2021* .
- Under review Balayn, Agathe, et al. "Automatic cyber extreme speech prediction: a Computer Science survey enlightened with Psychology" *to the ACM Transactions on Intelligent Systems and Technology , A literature survey on automatic methods for hate speech - and other related speeches - detection*.
- HCOMP workshop 2019 Balayn, Agathe, et al. "Designing Evaluations of Machine Learning Models for Subjective Inference. The Case of Sentence Toxicity." *at the Rigorous Evaluation of AI Systems workshop co-located with HCOMP 2019* .
- CHI workshop 2019 Balayn, Agathe, et al. "Unfairness towards subjective opinions in Machine Learning" *at the Human-Centered Machine Learning Perspectives workshop co-located with CHI 2019*.
- RO-MAN 2018 Balayn, Agathe, et al. "Data-driven development of Virtual Sign Language Communication Agents." *Robot and Human Interactive Communication, IEEE International Symposium on*. IEEE, 2018.
- Balayn, A. "On the fairness of crowdsourced training data and Machine Learning models for the prediction of subjective properties. The case of sentence toxicity: To be or not to be #\$\$%! toxic? To be or not to be fair?." Master thesis, TU Delft repository, 2018.
- HCOMP workshop 2018 Balayn, Agathe, et al. "Characterising and Mitigating Aggregation-Bias in Crowdsourced Toxicity Annotations." *CrowdBias workshop co-located with HCOMP 2018, AAAI Conference on*. AIII, 2018.
- SIMPAR 2016 Balayn, Agathe, et al. "Adaptive handling assistance for industrial lightweight robots in simulation." *Simulation, Modeling, and Programming for Autonomous Robots, IEEE International Conference on*. IEEE, 2016.

Other activities

External reviewer, *HCOMP 2019, IUI 2020*.

- 10/2019 **Presentation at the first symposium on Biases in Human Computation and Crowdsourcing (BHCC)**, *Sheffield, UK*, on Debugging dataset biases for data-driven decision support systems with human-in-the-loop and machine learning approach.
- 09/2019 - 11/2019 **Teaching Assistant for the Seminar on Web Information Systems**, *Delft University of Technology*, The Netherlands, Supervision of several students to write literature surveys on bias and privacy of structured and unstructured data.
- 07/2019 **Volunteer at the annual ACM Special Interest Group on Management of Data (SIGMOD 2019) conference**, *Amsterdam*, The Netherlands.
- 04/2019 - 06/2019 **Teaching Assistant for the Human Computation Course**, *Delft University of Technology*, The Netherlands, Grading of students. .

Technical Skills

Programming languages *Most experienced with:* **Python** (TensorFlow, Keras, Scikit-learn), **MATLAB**, **C++** (OROCOS, Gazebo simulation environment).

Some experience with: **C**, **Java**, **Maple**. **HTML**, **CSS**, **PHP**, **Javascript** (D3 library).

Basic knowledge of: **Prolog**, **OCaml**.

Others **Working knowledge of Linux**, **Git**, **common software suites (Office)**, **L^AT_EX**.