# FairAlgos

Fairness, Bias und Transparenz von (Computer Vision) Algorithmen

| | | | |
|---|---|---|---|
| **Programm / Ausschreibung** | Bridge, Bridge_NATS, Bridge_NATS 2019 | **Status** | abgeschlossen |
| **Projektstart** | 01.05.2020 | **Projektende** | 31.01.2023 |
| **Zeitraum** | 2020 - 2023 | **Projektlaufzeit** | 33 Monate |
| **Keywords** | Computer Vision Algorihmen, Machine Learning, Bias, Transparency | | |

## Projektbeschreibung

Computer Vision ist ein Fachgebiet der Künstlichen Intelligenz, dessen (Forschungs)-Ergebnisse in den letzten Jahren weltweit verstärkt zur Anwendung kamen: Lernende Bildanalyseverfahren in der Medizin tragen zur Diagnose bei, Kameras in mobilen Endgeräten werden für den Aufbau von Trainingsdaten für Deep Learning genutzt, Mustererkennung in Verbindung mit lernenden Algorithmen wird von der Strafverfolgung über die Kontrolle der Grenzen mit Gesichtserkennung bis hin zu Smart Homes im Gesundheitswesen und Pflegebereich in allen Bereichen der Gesellschaft eingesetzt.
Es ist daher an der Zeit, ethischen Überlegungen bei Anwendungen von Computer Vision eine besondere Aufmerksamkeit zukommen zu lassen. Fairness, Bias und Transparenz sind grundlegende Themen des Maschinellen Lernens, deren Behandlung besonders bei der Nutzung von vortrainierten Modellen von Bedeutung ist.
Es ist das Ziel dieses Projektes, Techniken und Modelle für Fairness bewusste Algorithmen zu entwickeln. Darüberhinaus werden Fairness Guidelines erstellt, die das Verständnis für entscheidungs-treffende Systeme durch mehr Transparenz und Interpretierbarkeit fördern.

## Abstract

Computer Vision is a field of Artificial Intelligence, whose (research) results have been used extensively in the last few years worldwide: Learning techniques in medicine contribute to the diagnosis, cameras in mobile devices are used for the development of training data for deep learning, Pattern recognition in conjunction with learning algorithms is used by law enforcement, from the control of borders with face recognition to smart homes in healthcare and nursing in all areas of society.
Therefore, it is time to pay particular attention to ethical considerations in Computer Vision applications. Fairness, bias and transparency are fundamental topics of machine learning, especially when using pre-trained models.
It is the goal of this project to develop techniques and models for fairness-conscious algorithms. Fairness guidelines are created to improve the understanding of decision-making systems through more transparency and interpretability.

## Projektkoordinator

- Technische Universität Wien

## Projektpartner

- VICESSE Research GmbH
- CogVis Software und Consulting GmbH
- Vienna Centre for Societal Security (VICESSE) Wiener Zentrum für sozialwissenschaftliche Sicherheitsforschung