

## dAlalog.at

dAlalog.at - Förderung von Fairness und Vertrauen in KI durch Formate der partizipativen Technikgestaltung

|                                 |  |                        |               |
|---------------------------------|--|------------------------|---------------|
| <b>Programm / Ausschreibung</b> | Ideen Lab 4.0, Ideen Lab 4.0, Ideen Lab4.0 - Ausschreibung 2019              | <b>Status</b>          | abgeschlossen |
| <b>Projektstart</b>             | 01.04.2020   | <b>Projektende</b>     | 30.06.2021    |
| <b>Zeitraum</b>                 | 2020 - 2021  | <b>Projektlaufzeit</b> | 15 Monate     |
| <b>Keywords</b>                 | Künstliche Intelligenz, Fairness, Vertrauen, Partizipative Technikgestaltung |                        |               |

### Projektbeschreibung

In den letzten Monaten und Jahren wurden zahlreiche Fälle von KI-Systemen medial diskutiert, die wegen ihren Datensammelwut, ihren diskriminierenden Algorithmen und ihrer Beeinflussung von WählerInnenverhalten das Vertrauen in künstliche Intelligenz weltweit schwächen. Auch auf österreichischem Boden wurden entsprechende Fälle ausgemacht und medial rezipiert.

Vor diesem Hintergrund hat es sich das Projekt dAlalog.at zur Aufgabe gemacht, einen nutzerbasierten Ansatz zu sondieren, um die Fairness von KI-Systemen zu fördern und das Vertrauen von österreichischen Unternehmen sowie der österreichischen Gesellschaft in Künstliche Intelligenz zu steigern. Hierzu werden partizipative Technikentwicklungsformate entwickelt, die wertebasierte Diskursräume zwischen Unternehmen, Organisationen, NutzerInnen/Betroffenen sowie Forschung öffnen und auf die Förderung von Fairness von und Vertrauen in KI-Systeme abzielen. Die im Projekt entwickelten Partizipationsformate machen KI-Technologie für den EndanwenderInnen/Betroffene greifbar und zugänglich. Einem partizipativen Technikgestaltungsansatz folgend, ermöglichen sie Selbstgestaltung von KI-Systemen auf Seite der NutzerInnen bzw. Betroffenen, tragen zur Fairness der Systeme bei und, erhöhen so deren gesellschaftliche und unternehmerische Akzeptanz.

In einem initialen Projektschritt soll mittels ExpertInnen-Einbindung identifiziert werden, welche Aspekte der Fairness künstlicher Intelligenz im Kontext Österreich als zentral erachtet werden. Die darauf aufbauend entwickelten partizipativen Technikgestaltungsformate werden anschließend Rahmen mehrerer Workshops erstmals erprobt und evaluiert. Mit dieser Vorgangsweise adressiert das Projekt dAlalog.at ein zentrales Erfordernis für KI-Systeme in Österreich und sondiert einen möglichen Lösungsansatz zur Herstellung von Fairness und Vertrauen in künstliche Intelligenz.

### Abstract

In recent months and years, numerous cases of AI systems have been discussed in the media that have weakened confidence in artificial intelligence worldwide. Amongst those were cases of unlimited data collection, discriminatory algorithms and AIs which influenced voters' behaviour. Similar cases have also happened in Austria, resulting in significant media coverage.

Against this background, the project dAlalog.at aims at exploring a user-centred approach for increasing the fairness of AI

systems as well as the trust of Austrian companies and society in artificial intelligence. To this end, participative technology development formats will be developed that open value-based discourse spaces between companies, organisations, users, society and research and aim at promoting fairness of and trust in AI systems. The participation formats developed in the project make AI technology tangible and accessible for the end users and society. Following a participative technology design approach, they enable self-design of AI systems on the part of the users or other persons affected by the systems. Thereby, they contribute to the fairness of the AI systems and increase their social and entrepreneurial acceptance. In an initial project step, experts will be involved to identify which aspects of the fairness of artificial intelligence are considered central in the context of Austria. The participatory technology design formats developed on this basis will then be tested and evaluated for the first time within the framework of several workshops. With this approach, the project dAlalog.at addresses a central requirement for AI systems in Austria and explores a possible solution for establishing fairness of and trust in artificial intelligence.

## **Endberichtkurzfassung**

Künstliche Intelligenz (KI) ist auch in Österreich auf dem Vormarsch und beschäftigt die breite Öffentlichkeit genauso wie politische Entscheidungstragende in zunehmenden Ausmaß. Ein zentrales Thema ist hierbei die Frage danach, wie KI-Systeme möglichst fair gestaltet werden können und keine von KI-Algorithmen betroffenen Personen Nachteile davontragen müssen. Das Projekt dAlalog.at hat sich dieser Frage angenommen und konzipierte Workshopformate für das Verstehen, die Entwicklung und den Einsatz von Künstlicher Intelligenz. Die Workshopformate sollen einerseits den Diskurs der interessierten Öffentlichkeit über faire Künstliche Intelligenz anregen und andererseits Betriebe bei der Entwicklung und Einbindung fairer Künstlicher Intelligenz unterstützen. Die Workshopformate können kostenfrei auf [www.dAlalog.at](http://www.dAlalog.at) heruntergeladen werden.

## **Projektkoordinator**

- Interdisziplinäres Forschungszentrum für Technik, Arbeit und Kultur (IFZ)

## **Projektpartner**

- Universität Salzburg
- Research & Data Competence OG
- JOANNEUM RESEARCH Forschungsgesellschaft mbH
- Universität Wien