Versuchsplan

# Beschreibung des Artefakts

Das Thema unserer Arbeit soll aus dem Themengebiet der Fair AI kommen, wir haben uns dabei für die Ermittlung von Verzerrungen (Bias) und Quellen potenzieller Diskriminierung in tabellarischen Datensätzen entschieden. Dieses Thema ist eher technischer Natur, daher haben wir uns für einen design-orientierten Forschungsansatz für den Verlauf dieser Arbeit entschieden. Geplant ist, ein Artefakt zu entwickeln, das tabellarische Daten und je nach Ausprägung des Artefakts einen Kontext des Datensatzes als Input nimmt und als Output gefundene Probleme im Datensatz im Bezug auf Fairness zurückgibt. Sollten keine Probleme gefunden worden sein, wird dementsprechend eine Meldung zurückgegeben, dass es in diesem Datensatz kein Problem mit Diskriminierung oder anderen Arten von Unfairness zu geben scheint. Der als Input übergebene Kontext wird durch das Artefakt begrenzt werden, also wird es nur eine gewisse Auswahl von Datensatzarten geben, die von dem Artefakt akzeptiert werden, für alle anderen Datensätze, die als Input übergeben werden, die keiner der vom Artefakt vorgegebenen Kategorie angehören, kann kein zuverlässiges Resultat garantiert werden. Die folgende Abbildung fasst die beschriebene Funktion des Artefaktes visuell nochmal zusammen.

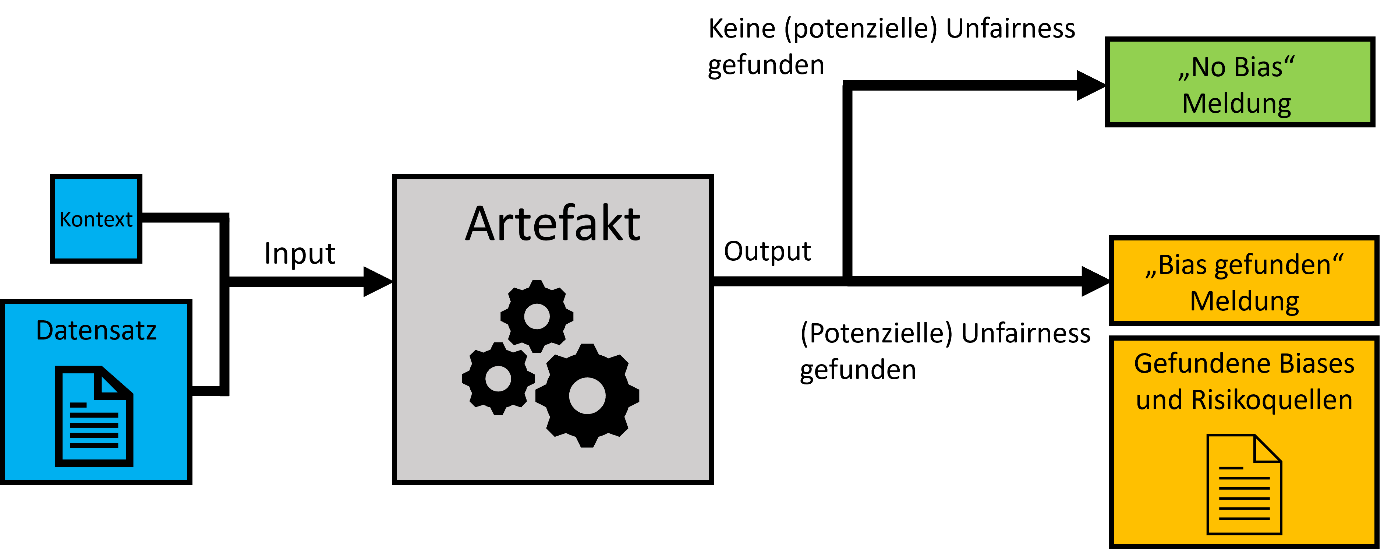


Abbildung Artefaktsfunktionsweise (abstrahiert)

# Geplante Experimente