

# Deliverable

## Rechtliche Anforderungen an die Erklärbarkeit, Nachvollziehbarkeit, Transparenz und Automatisierung von maschinellen Entscheidungen I



Eingereicht am: 31.12.2020

Eingereicht von: Monty-Maximilian Zühlke

Autor: Prof. Dr. Christian Heinze, LL.M

Zentrum für digitale Innovationen Niedersachsen (ZDIN)

Leibniz Universität Hannover

Institut für Rechtsinformatik IRI

Im vorliegenden Dokument wird der Zeitraum von 1.10.2019 bis 31.12.2020 dargestellt.

Das vorliegende Deliverable gehört zum Teilprojekt 3: Regulatorischer Rahmen und wirtschaftspolitische Instrumente

# 1 Generelle Informationen

## 1.1 Zusammenfassung

Im zurückliegenden Berichtszeitraum konnten für das Arbeitspaket 3.2. „Rechtliche Anforderungen an die Erklärbarkeit, Nachvollziehbarkeit, Transparenz und Automatisierung von maschinellen Entscheidungen“ zwei Doktoranden gewonnen werden. In der Projektarbeit gewonnene Erkenntnisse fließen in ihre Dissertationsvorhaben ein (hierzu siehe Punkt 2.2). Als konkrete Dissertationsthemen wurden identifiziert: „Zivilprozessuale Beweisführung und künstliche Intelligenz“ (J. Niemeyer) sowie „KI-gestützte Darlehensentscheidungen aus zivilrechtlicher, datenschutzrechtlicher und aufsichtsrechtlicher Sicht“ (S. Dalmer). Die Dissertationsthemen untersuchen die übergreifenden Fragen der Erklärbarkeit, Nachvollziehbarkeit, Transparenz und Automatisierung von maschinellen Entscheidungen bereichsspezifisch im materiellen Recht (Dalmer) und im Prozessrecht (Niemeyer). Zugrundegelegt werden die Fragen, inwieweit der vorhandene Regelungsrahmen den durch Technologieeinsatz geänderten Umständen gerecht wird und ob ggfs. ein Einschreiten des Gesetzgebers erforderlich ist. Im Berichtszeitraum konnte durch Literaturrecherche, die Konkretisierung der Themenausrichtung und durch die Erstellung einer Detailgliederung der Themen das Fundament für die weiteren Arbeiten gelegt werden.

Wissenschaftliche Erkenntnisse konnten zudem in Publikationen und Fachvorträgen verbreitet werden (hierzu siehe Punkt 2.3.) Zu nennen sind der von Prof. Dr. Heinze gehaltene Vortrag „Regulation of Artificial Intelligence“, die Vorstellung und Diskussion ihrer Dissertationsprojekte durch die Doktoranden, sowie die Veröffentlichung eines umfassenden KI- und Robotik-Rechtshandbuchs durch Prof. Dr. Heinze (Mitherausgeber) und die Erstellung eines in der Veröffentlichung befindlichen Fachaufsatzes betreffend das Verbot der automatisierten Einzelfallentscheidung bei der Darlehensvergabe (Heinze/Dalmer/Horstmann).

Für den kommenden Berichtszeitraum sind weitere Veröffentlichungen und Vorträge geplant und die Dissertationsprojekte sollen vorangetrieben werden.

## 1.2 Beitragende Arbeitspakete zum vorliegenden Deliverable

### AP3.2: Rechtliche Anforderungen an die Erklärbarkeit, Nachvollziehbarkeit, Transparenz und Automatisierung von maschinellen Entscheidungen

Beteiligte Partner*innen	Prof. Dr. Sternberg, LUH Prof. Dr. Kilian Bizer, Universität Göttingen
Bearbeiter*in	C. Heinze, S. Dalmer, J. Niemeyer
Praxispartner	-
Beginn der Arbeiten	M1 bis M60
Einreichungsdatum bzw. Frist	

### TA 3.2.1 Konkretisierung der Untersuchungsaufgabe und Entwurf einer Detailgliederung

Beteiligte Partner*innen	-
Bearbeiter*in	C. Heinze, S. Dalmer, J. Niemeyer
Praxispartner	-
Beginn der Arbeiten	M1 bis M6
Einreichungsdatum bzw. Frist	

### TA 3.2.2. Detailrecherche des Status quo zu Erklärbarkeit und Nachvollziehbarkeit maschineller Entscheidungen

Beteiligte Partner*innen	-
Bearbeiter*in	S. Dalmer, J. Niemeyer
Praxispartner	-
Beginn der Arbeiten	M1 bis M12
Einreichungsdatum bzw. Frist	

### TA 3.2.3. Identifikation typischer Quellen von Rechtsunsicherheit in Abstimmung mit der Praxis

Beteiligte Partner*innen	-
Bearbeiter*in	S. Dalmer, J. Niemeyer
Praxispartner	-
Beginn der Arbeiten	M13 bis M18
Einreichungsdatum bzw. Frist	

### TA 3.2.4 Ausarbeitung der Anforderungen an die Erklärbarkeit im geltenden Recht (Haftungsrecht, Verfahrensrecht, Datenschutzrecht)

Beteiligte Partner*innen	-
Bearbeiter*in	S. Dalmer, J. Niemeyer
Praxispartner	-
Beginn der Arbeiten	M13 bis M24
Einreichungsdatum bzw. Frist	

# 2 Ausführungen

## 2.1 Einleitung

Das Arbeitspaket „Rechtliche Anforderungen an die Erklärbarkeit, Nachvollziehbarkeit, Transparenz und Automatisierung von maschinellen Entscheidungen“ wird inhaltlich durch zwei Dissertationsprojekte bearbeitet, welche durch Vorträge und Einzelveröffentlichungen ergänzt werden. Als Dissertationsprojekte wurden die folgenden Themen identifiziert: „Zivilprozessuale Beweisführung und künstliche Intelligenz“ (J. Niemeyer) sowie „KI-gestützte Darlehensentscheidungen aus zivilrechtlicher, datenschutzrechtlicher und aufsichtsrechtlicher Sicht“ (Arbeitsthema) (S. Dalmer). Die Forschungsvorhaben wurden strukturiert und gegliedert. Diese Themen betreffen sowohl die materielle (S. Dalmer) als auch die prozessuale Rechtslage (J. Niemeyer) und ergänzen sich daher. Im Bereich der KI-gestützten Darlehensvergabe stellen sich Erklärbarkeits- und Transparenzprobleme hinsichtlich der verwendeten Scoring-Formel. Aus beweisrechtlichen Transparenz- und Erklärbarkeitsanforderungen im Zivilprozess lassen sich ggfs. Anforderungen für den beweissicheren Einsatz von KI in der Gesellschaft ableiten. Der Aufteilung in zwei verschiedene Themen wird die Arbeitshypothese zu Grunde gelegt, dass KI Regulierung sektorspezifisch erfolgen sollte, da sich das Risiko des KI Einsatzes und damit auch das Regulierungsbedürfnis in verschiedenen Gesellschafts- und Rechtsbereichen unterscheidet.<sup>1</sup> Übergreifend jedoch inhaltlich auf die Einzelthemen spezialisiert wird die Fragestellung verfolgt, inwieweit der vorhandene Regelungsrahmen den durch Technologieinsatz geänderten Umständen gerecht wird und dies ggfs. ein Einschreiten des Gesetzgebers erforderlich macht.

## 2.2 Erkenntnisgewinn

### Thema „Zivilprozessuale Beweisführung und künstliche Intelligenz“

Das Dissertationsthema wurde gewählt, da die zivilprozessualen Anforderungen zur Transparenz und Erklärbarkeit Anhaltspunkte für einen rechtssicheren Einsatz von KI liefern können.

Zunächst wurde der Literaturstand recherchiert. Die Literaturrecherche hat dabei ergeben, dass die zivilprozessualen Auswirkungen des Einsatzes intransparenter bzw. nicht erklärbarer KI nur in wenigen Beiträgen angesprochen wurden. Eine umfassende zivilprozessuale Würdigung dieses sog. Blackbox-Problems hat jedoch bislang nicht stattgefunden.

Auch eine Detailgliederung konnte fertiggestellt werden. Ein erster Schwerpunkt der Arbeit wird eine für die rechtswissenschaftliche Untersuchung aufbereitete Darstellung der technischen Grundlagen sein. Dabei wird insbesondere auch auf die Gründe für Blackbox KI, also für die fehlende Erklärbarkeit von KI bzw. für intransparente KI, eingegangen werden. Für diese konnten bislang sowohl rechtliche als auch technische (tatsächliche) Gründe identifiziert werden. So lassen sich als rechtliche Gründe Geheimnisschutz- oder Sicherheitsinteressen anführen. Als technisch bedingte Gründe sind hier wohl die Komplexität der Algorithmen, die Art der Algorithmen, die zu Grunde liegenden Lernprozesse und die Kombination verschiedener algorithmischer Elemente zu nennen. Zum Zwecke einer vergleichenden rechtlichen Bewertung sollen aber auch die bisher vorgestellten Ansätze erklärbarer KI (Explainable AI – XAI) betrachtet werden. Als Grundlage der zivilprozessualen Betrachtung werden

---

<sup>1</sup> vgl. Ebers in Ebers/Heinze/Krügel/Steinrötter, § 3, Rn. 194.



mögliche Haftungsszenarien des Einsatzes von KI dargestellt werden. Diese Darstellung wird sich an bereits erprobten KI Systemen orientieren (zum Beispiel im automobilen Bereich). In der rechtlichen Schwerpunktuntersuchung wird betrachtet werden, inwieweit KI-Daten bzw. -Entscheidungen in einen Haftungsprozess eingeführt und beweisrechtlich verwertet werden. Dabei werden sowohl KI-Daten bzw. -Entscheidungen betrachtet werden, deren zu Grunde liegenden Algorithmen eine Blackbox darstellen, als auch solche, deren Algorithmen erklärbar bzw. erklärbarer gestaltet werden. Die zivilprozessuale Untersuchung wird die Transparenzanforderungen der ZPO und die verfassungsrechtlichen Vorgaben herausarbeiten. Es wird hierzu auch verglichen werden, wie die Rechtsprechung bislang mit technischen oder naturwissenschaftlich begründeten Unwägbarkeiten umgegangen ist. Zudem werden Vergleiche mit den Möglichkeiten der Beweisführung in anderen komplexen technisch geprägten Sachverhalten gezogen. Aus dieser Untersuchung werden schließlich Anforderungen und Empfehlungen für eine aus zivilprozessualer Sicht erfreuliche Technikgestaltung bzw. Gestaltung der Rechtsverhältnisse abgeleitet werden.

## **Thema „KI-gestützte Darlehensentscheidungen aus zivilrechtlicher, datenschutzrechtlicher und aufsichtsrechtlicher Sicht“**

Das Dissertationsthema wurde gewählt, da sich hieran Fragen der Erklärbarkeit, Nachvollziehbarkeit, Transparenz und Automatisierung von maschinellen Entscheidungen juristisch sinnvoll aufbereiten lassen. Seit Jahren ist umstritten, inwieweit das Kredit scoring betreibende Akteure ihre zugrunde gelegten Modelle offenzulegen haben.<sup>2</sup> Zudem stellen sich unter Datenschutzgesichtspunkten rechtliche Probleme, wenn sich an das Scoring automatisierte Kreditentscheidungen anschließen. Zu dieser Frage wurde der unter Punkt 2.3 genannte Fachaufsatz verfasst.

Zur inhaltlichen Bearbeitung der Forschungsfragen wurden zunächst, soweit für die juristische Aufarbeitung der Thematik erforderlich, die technischen Grundlagen erarbeitet. Der Schwerpunkt liegt hierbei bei logistischen Regressionsmodellen, da diese regelmäßig beim bankinternen als auch beim durch Auskunftseien vorgenommenen Scoring zur Anwendung gelangen.<sup>3</sup> Sodann wurde der aktuelle Literaturstand ausgewertet und der Stoff gegliedert. Hierbei fiel auf, dass die aktuelle Diskussion schwerpunktmäßig das von Auskunftseien durchgeführte sog. „externe Scoring“ betrifft. Das eigenständig von den Kreditinstituten vorgenommene „interne Scoring“ wird nur am Rande behandelt. Aus diesem Grund wurde der Bearbeitungsschwerpunkt nunmehr auf das interne Scoring gelegt. Gegenüber dem externen Scoring ist hierbei besonders relevant, dass Kreditinstitute neben dem auch beim externen Scoring geltenden Datenschutzrecht besondere aufsichtsrechtliche Bestimmungen (bspw. §§ 10 Abs. 2, 18a KWG) und zivilrechtliche Bestimmungen (§§ 505a ff. BGB) zu beachten haben. Im Fortgang der Forschungsarbeit soll untersucht werden, ob in der Zusammenschau dieser Rechtsbereiche ein für das interne Scoring angemessener Regelungsrahmen existiert. Dies wird man indes nur bejahen können, wenn das Regelungsregime auch Fragen der Erklärbarkeit, Nachvollziehbarkeit, Transparenz und Automatisierung von KI-gestützten Darlehensentscheidungen angemessen berücksichtigt.

---

<sup>2</sup> Vgl. BGH NJW 2014, 1235 ff.

<sup>3</sup> Gesellschaft für Informatik, S. 32.

## 2.3 Erkenntnisdissemination

### **Rechtshandbuch „Künstliche Intelligenz und Robotik“**

Das Rechtshandbuch geht zurück auf die erste Fachtagung der Robotics and AI Law Society ([www.ai-laws.org](http://www.ai-laws.org)), die am 23. März 2018 im Leibniz Haus in Hannover stattfand. Deren Ziel nach ist wie vor aktuell, nämlich die rechtlichen Implikationen von KI und Robotik inter- und intradisziplinär aufzuzeigen und sowohl aktuell diskutierte als auch neue Fragen in den unterschiedlichen Rechtsgebieten zusammenzutragen. Die Fertigstellung des Handbuchs, dessen Mitherausgeber Prof. Dr. Christian Heinze ist, wurde durch das ZDIN Arbeitspaket unterstützt.

### **Vortrag Summer School 2020: Regulation of Artificial Intelligence**

Der am 10.09.2020 im Rahmen der Summer School des Zukunftslabors Gesellschaft und Arbeit durch Prof. Dr. Christian Heinze gehaltene Vortrag ging der Frage nach, inwieweit der deutsche und europäische Regulierungsrahmen angemessen ist. Es wurde die Leistungsfähigkeit des regulatorischen Instrumentariums erörtert und mögliche Schutzlücken wurden identifiziert. Zu diesem Zweck wurden ausgewählte Diskussionen, die in verschiedenen Rechtsbereichen geführt werden (zum Beispiel Haftung, Datenschutz, Diskriminierung), dargestellt.

### **Überblicksvorträge zu Dissertationsprojekten im Teilprojektreffen**

Der aktuelle Stand der Dissertationsprojekte sowie der jeweilige Forschungsansatz wurden durch die Doktoranden am 13.07.2020 im Rahmen eines Teilprojekttreffens den weiteren Projektbeteiligten vorgestellt. Es folgt eine Diskussion.

### **Fachaufsatz: Das Verbot automatisierter Einzelfallentscheidung bei der Darlehensvergabe**

Der von Prof. Dr. Christian Heinze und Stefan Dalmer gemeinsam mit Jan Horstmann erstellte Beitrag möchte zur Klärung der Funktion des Art. 22 DSGVO beitragen und den Anwendungsbereich der Norm konturieren, auch um so ein Fundament für die darauf aufbauende Transparenzdebatte zu bilden. Als konkretes Anwendungsbeispiel wurde hierfür die Darlehensvergabe gewählt, weil das Verbot automatisierter Einzelentscheidungen gerade dort eine zentrale Rolle einnimmt. Der Beitrag erscheint am Anfang des Jahres 2021 in einer Fachzeitschrift. Der Beitrag wurde in Kooperation mit dem von der Volkswagenstiftung geförderten Projekt „Bias and Discrimination in Big Data and Algorithmic Processing“ erstellt, in dessen Rahmen der Autor Jan Horstmann forscht.

### **Projektarbeit im Teilprojekt 3**

Der Berichtszeitraum wurde begleitet von regelmäßigen Treffen im Teilprojekt 3 „Regulatorischer Rahmen und wirtschaftspolitische Instrumente“. Dabei fand ein intensiver Austausch zu den jeweiligen Forschungsthemen, der gemeinsamen Außendarstellung und interdisziplinären Anknüpfungsmöglichkeiten statt.

## 3 Weiterführende Veröffentlichungen

Konferenz/ Journalname	Datum der Veröffentlichung	Autorenschaft	Ggf. Link zum Dokument
Künstliche Intelligenz und Robotik, Rechtshandbuch	30.09.2020	Ebers, Heinze, Krügel, Steinrötter et. al.	<a href="https://www.beck-shop.de/ebers-heinze-kruegel-steinroetter-kuenstliche-intelligenz-robotik/product/30097161">https://www.beck-shop.de/ebers-heinze-kruegel-steinroetter-kuenstliche-intelligenz-robotik/product/30097161</a>
Das Verbot automatisierter Einzelfallentscheidung bei der Darlehensvergabe	In Erscheinung	Heinze, Dalmer, Horstmann	

## 4 Literaturverzeichnis

Ebers/Heinze/Krügel/Steinrötter (Hrsg.)

Künstliche Intelligenz und Robotik, Rechtshandbuch, 1. Aufl., München, 2020.

Gesellschaft für Informatik

Technische und rechtliche Betrachtungen algorithmischer Entscheidungsverfahren, Berlin, 2018