

3 Wissen, Kontrolle und Risiken im Datenökosystem

In diesem Kapitel wird analysiert, welche Probleme Datenökosysteme bergen. Denn die soeben beschriebene Struktur von Datenökosystemen sorgt durch die Rollenverteilung unter den Akteuren für asymmetrisches Wissen sowie für ungleiche Kontrollmöglichkeiten. Diese Tatsache sorgt dafür, dass die Akteure mit mehr Wissen und Kontrolle den Akteuren mit weniger Wissen und Kontrolle Risiken auferlegen. Da diese Risikoauferlegung ethisch relevant ist und somit eine Rolle spielt für sowohl die ethische Bewertung als auch die ethische Gestaltung von Datenökosystemen, soll in diesem Kapitel erklärt werden, warum es dazu kommt.

3.1 Epistemische Asymmetrie im Datenökosystem

In diesem Teilkapitel wird die soeben analysierte Struktur von Datenökosystemen aus einer epistemischen Perspektive betrachtet – das heißt, es wird gefragt welche Akteure was wissen über die Abläufe im Datenökosystem. Es wird sich zeigen, dass durch die Rollenverteilung im Datenökosystem den Datenquellen und Datennutzenden essentielles Wissen fehlt und dass ihnen dadurch Risiken auferlegt werden.

Dem Epistemologen Paul Humphreys zufolge ist ein Prozess für eine Person epistemisch undurchsichtig (epistemically opaque), wenn die Person nicht alle epistemisch relevanten Elemente des Prozesses kennt.²³² Das kann zum Beispiel dann der Fall sein, wenn das Wissen über den Prozess ungleich verteilt ist. Solche Asymmetrien des Wissens oder auch epistemische Asymmetrien gibt es immer dann, wenn einer der Akteure nicht über das Wissen der

232 Humphreys 2009:618

anderen Akteure verfügt. Bei dem nur einseitig vorhandenen Wissen kann es sich um spezialisiertes technisches Wissen oder Wissen über relevante organisatorische Abläufe von Fachleuten handeln, aber es kann auch in Alltagssituationen auftreten, zum Beispiel zwischen Ärzt:innen und Patient:innen.²³³ Aus der Sicht des Sprachwissenschaftlers Paul Drew hängt das Wissen in jedem Fall von Rollenidentitäten ab. Außerdem werde die Asymmetrie des Wissens mit der Benachteiligung eines der Beteiligten und mit kommunikativen Schwierigkeiten in Verbindung gebracht.²³⁴

Eine Asymmetrie des Wissens bedeute jedoch nicht, dass der Akteur mit weniger Wissen nichts weiß. Bestimmtes Wissen wird nur Fachleuten zugeschrieben. So gelten zum Beispiel medizinische Diagnosen als normativ dem Arzt oder der Ärztin zugehörig. Die Patient:innen mögen trotzdem ihre eigenen Diagnosen kennen, aber sie haben das Wissen, weil es ihnen mitgeteilt wurde, und deshalb behandeln sie dieses Wissen nicht so, als wäre es ihr eigenes. In solchen Situationen handelt es sich auf Seiten der Patient:innen um vermitteltes Wissen und nicht um Wissen aus erster Hand. Wenn man aber Wissen aus erster Hand hat, ist man eine maßgebliche Quelle des Wissens (authoritative sources of knowledge) und hat epistemische Autorität.²³⁵

Laut Drew ist die Asymmetrie des Wissens eine natürliche oder notwendige Folge der ungleichen Verteilung des Wissens. Aber diese natürliche Ungleichverteilung könnte von dem Akteur mit epistemischer Autorität als strategisches Mittel eingesetzt werden, um mit dem Akteur mit weniger Wissen beliebig umzugehen – das nennt Drew eine interaktionelle (interactional) Asymmetrie. Die Möglichkeit, das asymmetrische Wissen auszunutzen, heißt jedoch nicht, dass die Asymmetrie mit Kontrolle oder Dominanz gleichzusetzen ist. Der Akteur mit vermitteltem Wissen wird also nicht notwendigerweise von dem Akteur mit epistemischer Autorität kontrolliert oder dominiert.²³⁶

233 Drew 1991:22, 25; Russo et al. 2024:1592

234 Drew 1991:25

235 Ariss 2009:910; Drew 1991:39f.; Snoek Henkemans, Wagemans 2012:347, Fußnote 2

236 Drew 1991:30, 32, 41, 43

Epistemische Dominanz erfordert dem Philosophen Keith Harris zufolge vielmehr die nicht wechselseitige Kompetenz eines Akteurs, die für einen anderen Akteur verfügbaren Beweise zu kontrollieren. Diese Kompetenz kann dadurch ausgeübt werden, dass Beweise vorgelegt, zurückgehalten oder gestaltet und verfügbar gemacht werden. Die Kompetenz der epistemischen Dominanz kann aber auch in Beziehungen vorhanden sein, in denen sie nicht ausgeübt wird. Der epistemisch dominante Akteur hat damit ein gewisses Maß an Kontrolle über die Fähigkeiten, das Wissen, die gerechtfertigten Überzeugungen und das Verständnis des dominierten Akteurs. Das heißt aber nicht, dass diese Kontrolle notwendigerweise bösartig genutzt wird. Die epistemische Dominanz muss zudem nicht beabsichtigt sein und kann aus der Erfüllung gesellschaftlicher Erwartungen oder anderen Zwängen entstehen.²³⁷

Es gibt verschiedene Abstufungen der epistemischen Dominanz: Wenn eine Person für die Wissensbildung vollkommen abhängig ist von nur einem Akteur, wie zum Beispiel ein Kind von seinen Eltern, dann ist der Einfluss dieses Akteurs sehr groß. Wenn es aber mehrere Akteure gibt, die Einfluss nehmen auf die einer Person zur Verfügung stehenden Beweise, ist jeder einzelne Akteur weniger einflussreich, und die epistemische Dominanz (epistemic domination) jeder der Akteure nimmt ab.²³⁸

Zudem gibt es verschiedene Möglichkeiten, epistemische Dominanz auszuüben. So können die verfügbaren Beweise eines Akteurs bei einem Ungleichgewicht durch eine Stellungnahme (testimony) direkt kontrolliert werden.²³⁹ Ein solches Ungleichgewicht kann dadurch entstehen, dass Stellungnahmen nur einseitig möglich sind, wenn zum Beispiel Politiker:innen der Öffentlichkeit Stellungnahmen präsentieren, während die einzelnen Mitglieder der Öffentlichkeit aber nicht in der Lage sind, Politiker:innen ebenfalls mit einer Stellungnahme zu erreichen. Epistemische Dominanz kann aber auch ausgeübt werden, indem man die Stellungnahmen kontrolliert, welche der dominierte Akteur von anderen erhält. Außerdem können die Medien kontrolliert werden, zu denen der dominierte Akteur Zugang hat. Oder es kann das weitere epistemische Umfeld der

237 Harris 2022:134f.

238 Ebd.:135

239 Ebd.:135f.

dominierten Person gestaltet werden, um die Informationen zu kontrollieren, die sie erreichen. Ein Beispiel sind Eltern, die bestimmen, mit wem ihre Kinder interagieren. Als epistemische Dominanz gilt laut Harris auch die Beschränkung des Zugangs zu Informationen oder Beweisen. Zudem können gefälschte Beweise vorgelegt werden, oder es können Informationen so bearbeitet werden, dass sie ohne Kontext irreführend sind.²⁴⁰

Laut Harris ist es unvermeidlich, dass ein gewisses Maß an epistemischer Dominanz vorhanden ist. Man kann nicht all sein Wissen aus erster Hand gewinnen und muss sich deshalb auch auf vermitteltes Wissen verlassen. Wenn die dominierenden Akteure wohlmeinend und gut informiert sind, dann kann das den dominierten Akteuren epistemisch zugutekommen – sie können etwas lernen und aus der Beziehung der epistemischen Dominanz einen Vorteil ziehen. Epistemische Dominanz kann aber auch epistemisch schädlich sein für den dominierten Akteur. Dieser Schaden kann unabsichtlich verursacht werden durch einen Akteur, der selbst epistemisch dominiert wird und diese Dominanz unbeabsichtigt auf andere ausdehnt, über die er Kontrolle hat. Aber es ist auch möglich, dass dominante Akteure die Unwissenheit anderer absichtlich ausnutzen, um ihre eigene epistemische Dominanz auszuweiten. Eine Situation, in welcher der dominierte Akteur absichtlich daran gehindert wird, epistemisch wertvolle Zustände zu erreichen, nennt Harris bössartige epistemische Dominanz (*malignant epistemic domination*).²⁴¹

Diese bössartige epistemische Dominanz kann laut Harris negative Konsequenzen haben: „If the dominating party is ill-intentioned, the dominated party is likely to form false beliefs.”²⁴² Insbesondere, wenn der dominierende Akteur schlecht informiert ist oder schlechte Absichten hat, kann der dominierte Akteur verzerrten Ergebnissen ausgesetzt werden und somit Schwierigkeiten haben, epistemisch wertvolle Zustände zu erreichen. Außerdem können Überzeugungen, welche der dominierte Akteur bisher hatte, untergraben werden. Eine weitere Gefahr liegt darin, dass bössartige epistemische Dominanz bei dem dominierten Akteur Orientierungslosigkeit auslöst. Das heißt, der dominierte Akteur weiß nicht mehr, was er

240 Harris 2022:136

241 Ebd.:136f. 139

242 Ebd.:137

glauben soll, wem er vertrauen kann und wie diese Unsicherheit aufgelöst werden kann. Orientierungslosigkeit kann das Ergebnis eines Wettbewerbs um die epistemische Vorherrschaft sein, aber sie kann auch das ausdrückliche Ziel bestimmter Akteure sein. Im letzteren Fall kann sie durch Manipulation der epistemischen Landschaft verstärkt werden und die Evidenz selbst kann ihre Bedeutung verlieren. Orientierungslosigkeit kann aber auch schon dadurch ausgelöst werden, dass der dominierte Akteur erkennt, dass er epistemisch dominiert wird. Diese Erkenntnis erweckt Misstrauen in die eigenen Überzeugungen und auch gegenüber vermeintlich verlässlichen Instanzen wie Wissenschaftler:innen.²⁴³

3.1.1 Übertragung auf Datenökosysteme

In Datenökosystemen besteht eine epistemische Asymmetrie, weil Datengenerierende aus erster Hand wissen, was mit den gesammelten und gespeicherten Daten passiert und weil Datenverarbeitende aus erster Hand wissen, wie die Ergebnisse zustande gekommen sind. Datenquellen und Datennutzende hingegen sind darauf angewiesen, auf das durch Datengenerierende und Datenverarbeitende vermittelte Wissen zurückgreifen. Dadurch, dass Datengenerierende und Datenverarbeitende Wissen aus erster Hand haben, sind sie eine maßgebliche Quelle des Wissens und haben epistemische Autorität.²⁴⁴ Die Akteure mit epistemischer Autorität können die Ungleichverteilung von Wissen als strategisches Mittel einsetzen, um mit dem Akteur mit weniger Wissen beliebig zu verfahren.²⁴⁵ Es ist also möglich, dass Datengenerierende und Datenverarbeitende ihr Wissen gegenüber Datenquellen und Datennutzenden als strategisches Mittel einsetzen, dennoch werden Letztere nicht automatisch von Ersteren kontrolliert oder dominiert.²⁴⁶

Um epistemische Dominanz handelt es sich aber dann, wenn ein Akteur dazu fähig ist, die für einen anderen Akteur verfügbaren Beweise zu kontrollieren und damit zu beeinflussen, zu welchen

243 Harris 2022:134, 137–139

244 Drew 1991:39f.

245 Ebd.:32

246 Ebd.:41

Überzeugungen der andere Akteur kommt.²⁴⁷ Datengenerierende sind zu einem gewissen Grad dazu in der Lage, die für Datenquellen verfügbaren Beweise in Bezug darauf, was mit den Daten passiert, zu kontrollieren. Und Datenverarbeitende sind zu einem gewissen Grad dazu in der Lage, die für Datennutzende verfügbaren Beweise in Bezug darauf, wie das Ergebnis zustande gekommen ist, zu kontrollieren. Datenquellen und Datennutzende sind dennoch in Bezug auf ihre Daten und die algorithmischen Ergebnisse nicht vollkommen epistemisch dominiert, da ihre Wissensbildung nicht ausschließlich abhängig ist von jeweils nur dem einen Akteur. Denn Datenquellen können auch an anderer Stelle zumindest allgemeine Informationen darüber bekommen, was passieren kann, wenn personenbezogene Daten zentral gespeichert werden und Datennutzende können ebenfalls aus anderen Quellen allgemeine Informationen darüber beziehen, was passieren kann, wenn Daten von Algorithmen analysiert werden.

Aber in Bezug auf das konkrete Datenökosystem wissen nur die Datengenerierenden, wie genau die Daten gespeichert werden und welche Maßnahmen ergriffen wurden, um der Zweckentfremdung durch Unbefugte vorzubeugen – nur dieser eine Akteur weiß, wie sicher die personenbezogenen Daten der Datenquellen sind und zu welchen Zwecken sie verarbeitet werden sollen. Die Informationen darüber, welche Daten und welcher Algorithmus genutzt wurden, um das Ergebnis zu generieren, haben nur die Datenverarbeitenden. Durch diese epistemische Dominanz können Datengenerierende und Datenverarbeitende die verfügbaren Beweise der Datenquellen und Datennutzenden durch ihre Stellungnahmen direkt kontrollieren und könnten zum Beispiel auch den Zugang zu Informationen beschränken.²⁴⁸ Das heißt aber noch nicht, dass diese epistemische Dominanz auch bösartig eingesetzt wird. Solange die dominierenden Akteure wohlmeinend und gut informiert sind, kann das den dominierten Akteuren epistemisch zugutekommen – in dem Fall erhalten Datenquellen und Datennutzende alle Informationen, die sie brauchen um zu epistemisch wertvollen Überzeugungen zu gelangen. Bei schlechten Absichten der Datengenerierenden oder Datenverarbeitenden kann diese epistemische Dominanz schädlich sein

247 Harris 2022:134, 137

248 Ebd.:135f.

für Datenquellen oder Datennutzende – zum Beispiel, wenn die Sicherheit der Daten beschönigt wird, um möglichst viele Daten zu sammeln. Ein solcher Schaden kann aber auch unabsichtlich dadurch verursacht werden, dass Datengenerierenden oder Datenverarbeitenden es selbst nicht besser wissen.²⁴⁹

Die Prozesse in Datenökosystemen sind in einem solchen Fall für Datenquellen und Datennutzende epistemisch undurchsichtig (epistemically opaque), weil sie nicht alle epistemisch relevanten Elemente des Prozesses kennen.²⁵⁰ Denn die epistemisch relevanten Elemente des Prozesses bei der Datenverarbeitung im Datenökosystem beziehen sich auf Wissen über den Datenfluss. Auf Grund der epistemischen Asymmetrie besteht aber die Gefahr, dass Datenquellen nicht darüber informiert werden, zu welchem Zweck ihre Daten verarbeitet werden, und Datennutzenden nicht wissen, wie die Ergebnisse, die ihnen zur Verfügung gestellt werden, zustande gekommen sind.

Datengenerierende und Datenverarbeitende haben zwar Wissen aus erster Hand über die epistemisch relevanten Elemente des Prozesses bezüglich der Nutzung und Verarbeitung der Daten im Datenökosystem, aber sie haben kein vollkommenes Wissen darüber. Ob personenbezogene Daten zweckentfremdet werden, kann nur mit einer gewissen Wahrscheinlichkeit ausgeschlossen werden, und auch ein Datendiebstahl kann den Verantwortlichen vorerst verborgen bleiben. Die Entstehung der Ergebnisse von Datenverarbeitung ist ebenfalls nicht vollkommen durchsichtig, wenn komplexe KI-Systeme eingesetzt werden.²⁵¹ Dann können selbst die Akteure mit mehr Wissen nicht mehr nachvollziehen, wie genau die Ergebnisse zustande gekommen sind, und es besteht ein Risiko, dass sie verzerrt sind.

Dieses Risiko können dann selbst Expert:innen nicht perfekt einschätzen, da es immer eine epistemische Unsicherheit gibt.²⁵² Wenn zudem die Einschätzung des Risikos von denjenigen Akteuren kommt, die das Risiko den Datenquellen und Datennutzenden auferlegen, kann es zu einem Interessenskonflikt kommen. Luhmann formuliert das so: „Diejenigen, für die das Risikoverhalten anderer

249 Harris 2022:137–139

250 Humphreys 2009:618

251 Russel, Norvig 2021:801

252 Möller 2009:219f.

zur Gefahr wird, werden anders darüber urteilen als die, die die Entscheidung selbst treffen oder an ihr beteiligt sind.“²⁵³ Die Einstellung der Akteure, die anderen ein Risiko auferlegen, widerspricht aber nicht automatisch der der Risikoexponierten. Denn Erstere könnten entweder ein Interesse daran haben, dass das Risiko als gering eingeschätzt und ohne Protest getragen wird oder aber daran, das durch ihre Arbeit entstehende Risiko so genau wie möglich einzuschätzen, um nicht für mögliche Folgeschäden aufkommen zu müssen.

Selbst wenn Datengenerierende und Datenverarbeitende das durch ihre Arbeit entstehende Risiko so gut wie möglich einschätzen wollen, bleibt immer eine Restunsicherheit.²⁵⁴ Denn die risikoauferlegenden Akteure haben ja nicht die volle Kontrolle über die Risiken, die sie verursachen. So können Datengenerierende zum Beispiel personenbezogene Daten sammeln und so sicher wie möglich speichern, doch es ist schwer vorauszusehen, wie groß das Interesse an den Daten ist, und mit welchen Methoden womöglich Angreifer versuchen könnten, die Sicherheitsvorkehrungen zu umgehen. Gleichmaßen können Datenverarbeitende sich größte Mühe geben, korrekte algorithmisch generierte Ergebnisse zur Verfügung zu stellen, aber insbesondere, wenn KI-Systeme eingesetzt werden, ist das Risiko verzerrter Ergebnisse nicht leicht einzuschätzen.

Es besteht bei allen Akteuren epistemische Unsicherheit über die Folgen des Datenökosystems. Aber nur Datenquellen und Datennutzende tragen das Risiko, dass ihre Daten ohne ihr Wissen zweckentfremdet werden, oder dass ihnen verzerrte Ergebnisse zur Verfügung gestellt werden, auf deren Grundlage sie zu falschen Überzeugungen gelangen. Deshalb argumentiere ich, dass Datenquellen auf Grund der Struktur eines Datenökosystems epistemisch benachteiligt sind, da sie nicht über die Verwendung ihrer personenbezogenen Daten Bescheid wissen und damit für sie das Risiko der Verletzung der Privatsphäre besteht. Für Datennutzende besteht auf Grund der epistemischen Asymmetrie die Möglichkeit, einem verzerrten Ergebnis zu glauben, weshalb fehlende Transparenz für sie ein Risiko birgt.

253 Luhmann 1990:163

254 Ebd.

3.1.2 Risiko: Definition

Der Begriff „Risiko“ kann unterschiedliche Bedeutungen haben.²⁵⁵ Nach Ansicht des theoretischen Philosophen Sven Ove Hansson bezieht sich der Begriff Risiko auf Situationen, in denen es möglich, aber nicht sicher ist, dass ein unerwünschtes Ereignis eintreten wird. Mit Risiko kann aber auch die Ursache, die Wahrscheinlichkeit oder der statistische Erwartungswert (Schadenhöhe mal Eintrittswahrscheinlichkeit) eines unerwünschten Ereignisses gemeint sein, welches eintreten oder nicht eintreten kann. Zudem wird Risiko im Rahmen der Entscheidungstheorie interpretiert als die Tatsache, dass eine Entscheidung unter den Bedingungen bekannter Wahrscheinlichkeiten getroffen wird.²⁵⁶ Der Soziologe Niklas Luhmann hingegen meint: „Üblicherweise spricht man von Risiko immer dann, wenn ein möglicher Schaden um eines Vorteils willen in Kauf genommen wird.“²⁵⁷ Seiner Meinung nach können Risiken dementprechend durch Verzicht auf die Vorteile vermieden werden.

Auf eins können sich jedoch laut Luhmann alle Definitionen einigen: „(...) der Risikobegriff verweist auf Zukunft.“²⁵⁸ Fest steht, dass das in den Definitionen betrachtete unerwünschte Ereignis, beziehungsweise der mögliche Schaden, immer in der Zukunft liegt. Dem Wirtschaftswissenschaftler Frank H. Knight zufolge unterscheidet die Fähigkeit, sich auf eine Situation einstellen zu können, bevor diese Situation eintritt, uns Menschen von Tieren. Je weiter man vorausschauen könne, desto besser könne man sich anpassen und desto kompetenter könne man leben.²⁵⁹ Der Medizinethiker Joschka Haltaufderheide meint, dass dies das menschliche Zeitverständnis in der Neuzeit prägt: „Wir leben und handeln in der Erwartung und auf das hin, was noch nicht ist. Die Kenntnis der Zukunft, zumindest ihrer Struktur nach, ist daher essentiell, um eine Antwort auf die Frage formulieren zu können, was zu tun ist.“²⁶⁰ Somit sei auch jegliche ethische Einschätzung abhängig von dem Blick auf potenziell eintretende unerwünschte Ereignisse.

255 Hansson 2005:7f.

256 Hansson 2004:10

257 Luhmann 1990:135

258 Ebd.:140

259 Knight 1921:200

260 Haltaufderheide 2015:14

Der Technologiephilosoph Günter Ropohl charakterisiert den Begriff des Risikos anhand des Schadens, den ein unerwünschtes Ereignis verursacht, und anhand der Eintrittswahrscheinlichkeit dieses unerwünschten Ereignisses. Dabei unterscheidet er verschiedene Formen des Risikos anhand von folgenden Merkmalen: Der Schaden, der riskiert wird, kann unterschiedlich hoch sein, und diese Schadenshöhe kann unterschiedlich leicht bestimmbar sein. Zudem variiert, inwiefern die Schadenshöhe beeinflussbar ist durch zum Beispiel Vorkehrungen in Form von Regeln. Die Ursache des Schadens könne natürlichen Ursprungs sein oder durch Menschen verursacht werden – durch die Risikoexponierten selbst oder durch Außenstehende. Die Bekanntheit des Schadenspotenzials spielt laut Ropohl eine Rolle bei seiner Akzeptanz. Auch die Eintrittswahrscheinlichkeit des Risikos ist mit unterschiedlicher Gewissheit bestimmbar und manchmal durch die Betroffenen, durch Außenstehende oder gar nicht beeinflussbar.²⁶¹

Luhmann würde Ropohl in dem Punkt widersprechen, dass der Schaden, der riskiert wird, auch natürlichen Ursprungs oder durch Außenstehende verursacht sein kann. Denn er ist der Meinung, dass man sich nur selbst einem Risiko aussetzen kann. Wenn Außenstehende einen möglichen Schaden „(...) um eines Vorteils willen in Kauf [nehmen] (...)“²⁶², dann sei man hingegen einer Gefahr ausgesetzt. Das Wesentliche bei seiner begrifflichen Unterscheidung ist die Möglichkeit, selbst eine Entscheidung zu treffen: Trifft man eine Entscheidung, die potenziell zu zukünftigen Schäden führt, „(...) geht man mit der Entscheidung ein Risiko ein. Schäden, die außerhalb dieses Einflußbereiches liegen, werden, solange sie noch unsicher sind, als Gefahr angesehen.“²⁶³ So seien auch natürlich verursachte Katastrophen keine Risiken, sondern Gefahren. Hat man jedoch Entscheidungsmöglichkeiten, „die einen etwaigen Schadenseintritt beeinflussen bzw. ihn vermeiden helfen könnten.“²⁶⁴, so kann man laut Luhmann wieder von einem Risiko sprechen, wenn die Möglichkeiten bewusst nicht ergriffen werden.

261 Ropohl 2016:12ff.

262 Luhmann 1990:135

263 Ebd.:149

264 Ebd.:150

Luhmann unterscheidet insbesondere deshalb zwischen Risiko und Gefahr, weil es Situationen gibt, „(...) in denen das Risikoverhalten des einen zur Gefahr für den anderen wird.“²⁶⁵ Dabei ist es möglich, dass das Risiko zwar rational kalkuliert wurde, aber für andere, die an der Entscheidung nicht beteiligt waren, die jedoch von den potenziellen Schäden zukünftig betroffen sein könnten, daraus eine Gefahr entsteht: „Das riskante Verhalten der einen wird zur Gefahr für die anderen (...)“.²⁶⁶ Luhmann behauptet, dass man in Gefahr ist, wenn potenzieller Schaden durch „(...) Ursachen außerhalb der eigenen Kontrolle (...)“²⁶⁷ entsteht, also von Außenstehenden oder natürlichen Umständen verursacht wird. Während man sich gegenüber natürlichen Gefahren entweder risikoavers oder risikoaffin verhalten könne, sei man den Gefahren, die durch die Entscheidungen von Außenstehenden entstehen, hilflos ausgesetzt.²⁶⁸ Auch risikoaffine Menschen seien nicht unbedingt bereit dazu, diese, durch riskante Entscheidungen von Außenstehenden verursachten Gefahren, hinzunehmen.²⁶⁹ Dass wir Menschen aber im Alltag nicht umhin können, das von Außenstehenden verursachte Risiko hinzunehmen (zum Beispiel im Straßenverkehr), wird in Kapitel 5 dargestellt.

Betrachten wir nun die Perspektive der Außenstehenden, welche durch ihre Entscheidungen Risiken entstehen lassen. Die Entscheidungstheorie geht davon aus, dass den Entscheidenden bekannt ist, welche Ergebnisse möglich sind und mit welcher Wahrscheinlichkeit sie eintreten werden.²⁷⁰ Sicher bekannt sind die Wahrscheinlichkeiten der möglichen Ergebnisse allerdings nur in idealisierten Fallbeispielen. So schreibt Hansson: „Only very rarely are probabilities known with certainty.“²⁷¹ Denn in der Realität ist es laut Haltaufderheide oft nicht möglich, „(...) sich der vollständigen Bekanntheit aller kausalen Einflüsse auf ein zukünftiges Ereignis sicher zu sein (...)“.²⁷² Deshalb handelt es sich laut Hansson meist tatsächlich um Entscheidungen ‚unter Unsicherheit‘: „In decision-making under un-

265 Luhmann 1990:152

266 Ebd.:161

267 Ebd.:149

268 Ebd.:153

269 Ebd.:154

270 Hansson 1999:539

271 Ebd.

272 Haltaufderheide 2015:153

certainty, probabilities are either not known at all or are only known with insufficient precision.”²⁷³ Dann gibt es keine Möglichkeit, die Auswirkungen einer Entscheidung abzuschätzen.

Um in dem von Unsicherheiten geprägten Alltag Entscheidungen treffen zu können, müssen wir Menschen Unsicherheiten reduzieren und Überzeugungen fixieren. Das heißt, Aussagen über die Welt werden nicht nur Wahrscheinlichkeiten zugewiesen, sondern sie werden entweder für wahr oder falsch gehalten. Auf diese Weise ist die unsichere Welt kognitiv handhabbar und Entscheidungen können leichter getroffen werden. Eine solche Reduzierung der Komplexität darf allerdings nur vorübergehend sein, denn wenn es Gründe dafür gibt, müssen wir von der Überzeugung zur Wahrscheinlichkeit oder sogar zur Unsicherheit zurückkehren können.²⁷⁴

Wenn also Entscheidungen ‚unter Risiko‘ getroffen werden, ist nicht tatsächlich vollständig bekannt, mit welcher Wahrscheinlichkeit unerwünschte Ereignisse eintreten. Vielmehr wird die Beschreibung der Entscheidungsprobleme vereinfacht, indem sie als Fälle mit bekannten Wahrscheinlichkeiten behandelt werden.²⁷⁵ Laut Luhmann transformiert die Wahrscheinlichkeitsrechnung Unsicherheit in sichere Entscheidungen.²⁷⁶

Das Gegenteil von Risiko ist jedoch laut dem praktischen Philosophen Niklas Möller, dem theoretischen Philosophen Sven Ove Hansson und dem Ethiker Martin Peterson nicht absolute Sicherheit. Sie sind der Meinung, dass die Annahme je geringer das Risiko, desto höher die Sicherheit auf einer unzureichenden Definition von Sicherheit basiert. Um diese Meinung zu begründen, unterscheiden Möller et al. zwischen einem absoluten und einem relativen Sicherheitskonzept. Absolute Sicherheit bedeutet demnach, dass das Risiko eines Schadens vollständig beseitigt ist, während relative Sicherheit bedeutet, dass das Risiko auf ein bestimmtes (tolerierbares) Niveau reduziert wurde. Die Autoren sind sich einig, dass das Konzept von absoluter Sicherheit nur ein Ideal ist, das niemals erreicht werden kann. Auch Luhmann stimmt zu: „Es gibt aber keine risikofreie

273 Hansson 1999:539

274 Hansson 2004:12–15; Hansson 1999:540

275 Hansson 1999:539; Hansson 2003:293; Hansson 2004:11f.

276 Luhmann 1990:136

Sicherheit.²⁷⁷ Vielmehr lassen Sicherheitskonzepte in der Realität meist Risikowahrscheinlichkeiten über Null zu. Gemäß einem solchen relativen Sicherheitskonzept bedeutet also die Aussage „diese Daten sind sicher“ das gleiche wie „die Sicherheit der Daten ist so hoch, wie es unter Berücksichtigung angemessener Kosten und vorbeugender Maßnahmen erwartet werden kann, und das Risiko einer Zweckentfremdung der Daten ist sehr gering“.²⁷⁸

Sicherheit ist aber auch aus einem weiteren Grund nicht unbedingt immer der Gegenspieler von Risiko. Denn der Schaden, den ein Risiko auslöst, kann laut Luhmann „(...) auch im Ausbleiben eines Vorteils [bestehen], in dessen Erwartung man investiert hatte.“²⁷⁹ Zum Beispiel besteht bei der Teilnahme an einer Lotterie das Risiko, dass man eine Niete zieht – dies wäre ein unerwünschtes Ereignis. Die Vermeidung dieses Ereignisses ist aber keine Frage der Sicherheit.²⁸⁰ Sicherheit kann nur dann im Angesicht eines Risikos erstrebenswert sein, wenn das unerwünschte Ereignis einen Schaden verursacht und damit das Gegenteil von Sicherheit erzeugt.

Ich gehe im Folgenden davon aus, dass ein Risiko eine Situation beschreibt, in welcher ein unerwünschtes Ereignis mit einer gewissen Wahrscheinlichkeit eintritt. In solche Situationen kann sich ein Individuum selbst bringen und sich damit selbst einem Risiko aussetzen. Aber Risiken können einem Individuum auch von Außenstehenden auferlegt werden – dadurch entsteht für das Individuum die Gefahr, dass das unerwünschte Ereignis tatsächlich eintritt. In erster Linie wird das Individuum aber, durch die Entscheidungen von Außenstehenden, in eine Situation gebracht, in welcher unsicher ist, ob das unerwünschte Ereignis eintritt oder nicht.

Datenquellen und Datennutzende wissen auf Grund der epistemischen Asymmetrie im Datenökosystem nicht Bescheid, wie mit den sie betreffenden Daten umgegangen wird. So fehlt Datenquellen das Wissen darüber, zu welchem Zweck ihre Daten verarbeitet werden und Datennutzenden fehlt das Wissen über die Entstehung des Ergebnisses, dem sie ausgesetzt sind. Das heißt, im Datenökosystem können zwei unerwünschte Ereignisse mit einer gewissen Wahrscheinlichkeit eintreten – es entstehen Risiken: Für Datenquellen

277 Luhmann 1990:136

278 Möller et al. 2006:420

279 Luhmann 1990:138

280 Möller et al. 2006:421

besteht das Risiko, dass ihre Daten zu Zwecken verwendet werden, denen sie nicht zugestimmt haben oder dass sogar Unbefugte Zugriff auf ihre Daten bekommen. Für Datennutzende besteht das Risiko einem verzerrten Ergebnis zu glauben und somit zu falschen Überzeugungen zu gelangen, weil sie die Entstehung nicht nachvollziehen können. Im nächsten Teilkapitel soll gezeigt werden, dass Datenquellen und Datennutzende neben dem epistemischen Nachteil auch weniger Kontrolle haben über die Vorgänge im Datenökosystem und dass ihnen dadurch die Risiken auferlegt werden.

3.2 Verteilung der Kontrollmöglichkeiten im Datenökosystem

Wie in Teilkapitel 3.1 festgestellt, haben die epistemisch dominanten Akteure ein gewisses Maß an Kontrolle über die Fähigkeiten, das Wissen, die gerechtfertigten Überzeugungen und das Verständnis der epistemisch dominierten Akteure.²⁸¹ In diesem Teilkapitel sollen die Kontrollmöglichkeiten der Akteure über die Vorgänge im Datenökosystem genauer betrachtet werden. Zur Veranschaulichung möchte ich nun mein Modell der ethisch relevanten Akteure eines Datenökosystems auf die öffentliche Verwaltung anwenden und den hypothetischen Fall eines digitalisierten Wohngeldantrages betrachten.

Im Falle der öffentlichen Verwaltung ist das Ziel der Datenverarbeitung im Datenökosystem eine effektivere Bearbeitung von Anträgen, z.B. auf Wohngeld.²⁸² Bei der digitalen Gestaltung des Wohngeldantrages in Form eines Datenökosystems, fließen die Daten der Antragstellenden (Datenquellen) zur Wohngeldbehörde. Diese sammelt die Daten ein und nimmt damit die Rolle als Datengenerierende ein. Sie leitet die Daten weiter an das Statistische Landesamt, an das Statistische Bundesamt sowie an das Bundesministerium des Innern, für Bau und Heimat und an das Bundesamt für Bauwesen und Raumordnung, damit diese die Daten im Rahmen einer Wohngeldstatistik verarbeiten können.²⁸³ Gleichzeitig ist die Wohngeldbehörde auch für die Verarbeitung der Daten zuständig, da sie den Antragstellenden eine Rück-

281 Harris 2022:134f.

282 §1a, Abs. 1, Satz 1, OZG

283 §§ 34–36 WoGG

meldung über deren Anspruch auf Wohngeld und dessen Höhe geben muss – sie repräsentiert somit auch die Akteursgruppe der Datenverarbeitenden. Datennutzende sind dann sowohl die Antragstellenden, die erfahren, ob sie Wohngeld bekommen, als auch der Staat, der anhand der Wohngeldstatistik erkennen kann, an welchen Orten in Deutschland bezahlbare Wohnungen fehlen.

Von einem solchen Datenökosystem profitieren sowohl Antragstellende (Datenquellen und Datennutzende), die Wohngeldbehörde (Datengenerierende und -verarbeitende) als auch der Staat (ebenefalls Datennutzende). Die Antragstellenden müssen nicht mit all ihren Dokumenten zum Amt, sondern sie können schnell und von Zuhause aus erfahren, ob sie einen Anspruch auf Wohngeld haben. Die Wohngeldbehörde hat einen geringeren Arbeitsaufwand, wenn sie alle notwendigen Dokumente digital abrufen kann und bei der Einschätzung des Anspruchs durch einen Algorithmus unterstützt wird. Der Staat profitiert von der Weitergabe und Auswertung der Daten sowie einer kürzeren Bearbeitungszeit.

Denn wir nehmen an, dass ohne schnellere Antragsbearbeitung mehrere Risiken bestehen: Auf Grund des Personalmangels in den Wohngeldbehörden können nicht alle Wohngeldanträge zeitnah bearbeitet werden – so müssen Personen mit Anspruch auf Wohngeld länger ohne dieses auskommen. Den Antragstellenden droht also der Wohnungsverlust, falls die Miete ohne Unterstützung nicht mehr bezahlt werden kann. Den Mitarbeitenden in den Wohngeldbehörden droht berufliche Überlastung und dem Staat fehlen aussagekräftige Wohngeldstatistiken und somit wertvolle Planungsgrundlagen für den öffentlichen Wohnungsbau.

Aber es bestehen auch Risiken durch die Datenverarbeitung im Datenökosystem. Um das Ziel der schnelleren Antragsbearbeitung erreichen zu können, ist es notwendig, die Dokumente der Antragstellenden digital zu speichern und zwischen den verschiedenen Behörden auszutauschen. Viele dieser Dokumente enthalten sensible Informationen über die Antragstellenden.²⁸⁴

Durch das digitale Speichern werden die Antragstellenden dem Risiko des Zugriffs durch Unbefugte, z.B. Hacker, und damit der

284 Zum Beispiel: Personalien, Adresse, Verdienstbescheinigung (auszufüllen vom Arbeitgeber), Einkommensnachweise (Ausbildungs- oder Arbeitsvertrag, Lohnabrechnungen), Schwerbehindertenausweis, Bescheide über Kindergeld/Kinderzuschlag/Elterngeld

Zweckentfremdung ihrer sensiblen Daten ausgesetzt. Das heißt, es kann potenziell ein Schaden auftreten. Der Schaden wird nicht zwangsläufig verursacht, aber es besteht ein Risiko. Und zum schnelleren Bearbeiten der Wohngeldanträge benötigt man noch ein weiteres Mittel: Den Algorithmus, der die Daten analysiert und berechnet, ob ein Anspruch auf Wohngeld besteht. Insbesondere wenn der Regelsatz des Algorithmus nicht durchschaubar und dadurch eventuell unbemerkt falsch ist, entsteht ein weiteres Risiko.

Es fällt auf, dass von den Risiken insbesondere die Antragstellenden in ihrer Rolle als Datenquellen und als Datennutzende betroffen sind. Datengenerierende haben die Kontrolle darüber, wie die Daten der Datenquellen gespeichert werden und an wen sie weitergegeben werden. Die Datenquellen hingegen müssen im Fall des Wohngeldantrages der Weitergabe ihrer Daten an unter anderem das Statistische Bundesamt zustimmen, um den Antrag überhaupt stellen zu können. Ebenso haben Datenverarbeitende die Kontrolle darüber, wie die Daten verarbeitet werden und mit welcher Erklärung die Ergebnisse zur Verfügung gestellt werden. So wissen die Antragstellenden in ihrer Rolle als Datennutzende nicht, mit welchem Algorithmus ihr Antrag bearbeitet wurde und können nicht überprüfen, ob dieser Algorithmus ein fehlerhaftes Regelwerk erlernt oder vorgegeben bekommen hat.

3.2.1 Drei-Parteien-Modell

Um besser zu verstehen, welche Rollen die Akteure in Bezug auf das Risiko spielen und wieviel Kontrolle sie jeweils darüber haben, wird nun das Drei-Parteien-Modell von der Technologiephilosophin und Ethikerin Hélène Hermansson und dem bereits erwähnten theoretischen Philosophen Sven Ove Hansson vorgestellt. Dieses Modell soll eine systematische Beschreibung der ethischen Aspekte des Risikos ermöglichen.²⁸⁵ Die Autor:innen identifizieren in Situationen, in denen Risiken entstehen, drei Parteien: die Risikoexponierten, die Entscheidungstragenden und die durch das Risiko Begünstigten. Durch das Betrachten der Beziehungen zwischen diesen drei Parteien werden die ethischen Aspekte von Risiken identifiziert. Um diese

285 Hermansson, Hansson 2007:130

Beziehungen analysieren zu können, entwickeln die Autor:innen sieben Fragen. Diese Fragen betrachten das Verhältnis zwischen den durch das Risiko Begünstigten und den Risikoexponierten und die Verteilung von Risiken und Nutzen zwischen ihnen – insbesondere, ob diese Verteilung durch Umverteilung oder Ausgleich weniger ungerecht gestaltet werden kann.

Die Fragen betrachten außerdem die Beziehung zwischen den Risikoexponierten und den Entscheidungstragenden und den Einfluss, den die Risikoexponierten auf die Entscheidungsfindung haben. Der Einfluss der Risikoexponierten ist abhängig von den Informationen, die ihnen zur Verfügung gestellt werden: „(...) the most common ethical problem in relation to information is the degree to which experts and decision-makers provide the risk-exposed with adequate and comprehensible information about the risk (...).“²⁸⁶ Zuletzt betrachten die Fragen auch die Beziehung zwischen den Entscheidungstragenden und den durch das Risiko Begünstigten. Denn wenn derselbe Akteur zu beiden Gruppen gehören sollte, stellt das seine Legitimität in Frage: „A person who will herself reap the benefits if someone else is exposed to a risk does not have much credibility as a judge of whether the benefits outweigh the risks.“²⁸⁷ Sollten die Entscheidungstragenden selbst von dem Risiko profitieren können, so ist die Gefahr groß, dass sie aus Eigennutz den Risikoexponierten das Risiko auferlegen.

Eine Möglichkeit ein solches eigennütziges Verhalten zu vermeiden ist, die Entscheidung, ob das Risiko eingegangen wird, den Risikoexponierten zu überlassen – so können sie das Risiko frei wählen: „(...) for a risk to be freely chosen, the risk-exposed has to have full information about the risk and has to be capable of making the decision whether or not to accept the risk.“²⁸⁸ So lässt sich unterscheiden, ob die Risikoexponierten das Risiko akzeptiert haben, um einen Vorteil zu erlangen, den sie sonst nicht hätten erlangen können, oder ob sie gegen ihren Willen dem Risiko ausgesetzt wurden.²⁸⁹

286 Hermansson, Hansson 2007:132

287 Ebd.

288 Ebd.:142

289 Ebd.:130

3.2.2 Anwendung des Drei-Parteien-Modells auf das Wohngeldantrag-Beispiel

Die Anwendung des Drei-Parteien-Modells ermöglicht die Analyse der ethischen Aspekte von (un-)zulässiger Auferlegung von Risiken am Beispiel des Wohngeldantrages. Zuerst müssen die drei Parteien identifiziert werden: Die Risikoexponierten sind die Antragstellenden, da ihre Daten potenziell zweckentfremdet werden können bzw. der zur Antragbearbeitung genutzte Algorithmus ihren Antrag fälschlicherweise ablehnen könnte. Die Entscheidungstragenden sind die Wohngeldbehörden oder im weitesten Sinne der Staat, welcher sich für die Digitalisierung der Wohngeldanträge entscheidet. Die durch das Risiko Begünstigten sind sowohl Antragstellende, die Wohngeldbehörde als auch der Staat. Dabei sind die Entscheidungstragenden im Kontext vom Datenökosystem die Akteure mit mehr Kontrolle und Risikoexponierte sind meist die Akteure mit weniger Kontrolle (außer Entscheidungstragende legen sich selbst ein Risiko auf).

Zuerst wird das Verhältnis zwischen den Risikoexponierten und den durch das Risiko Begünstigten betrachtet – das ist in diesem Fall das Verhältnis zwischen den Antragstellenden einerseits und der Gruppe der Antragstellenden, der Wohngeldbehörde und dem Staat andererseits. Die Verteilung von Risiken und Nutzen zwischen ihnen ist klar, denn die Risiken tragen allein die Antragstellenden, da sie nicht die Kontrolle über den Einsatz des Algorithmus haben. Der Nutzen der Digitalisierung der Wohnanträge scheint jedoch gleichmäßig verteilt zu sein – jede beteiligte Partei profitiert auf ihre Weise. Man könnte es als Ausgleich bezeichnen, dass die Antragstellenden als Risikoexponierte auch gleichzeitig Teil der Partei der durch das Risiko Begünstigten sind.

Betrachten wir die Beziehung zwischen den Risikoexponierten und den Entscheidungstragenden so fällt auf, dass sich die Akteure in diesen Parteien nicht mehr überschneiden. Die Antragstellenden sind als Risikoexponierte nicht Teil der Entscheidungstragenden, welche durch die Wohngeldbehörde und den Staat vertreten sind. Die Antragstellenden entscheiden zwar, ob sie der digitalen Sammlung ihrer Daten zustimmen, sind aber dazu gezwungen, wenn sie einen Wohngeldantrag stellen wollen. Wie sicher die Daten ausgetauscht und gespeichert werden und ob ein KI-System genutzt wird,

entscheiden die Wohngeldbehörden und der Staat. Selbst wenn die Antragstellenden über die Risiken der digitalen Datenspeicherung und der Nutzung eines Algorithmus zur Bearbeitung der Anträge informiert würden, wären sie immer noch darauf angewiesen, diese Risiken einzugehen, um sich die finanzielle Unterstützung holen zu können, die ihnen zusteht.

Bei der Betrachtung der Beziehung zwischen den Entscheidungstragenden und den durch das Risiko Begünstigten fällt auf, dass die Wohngeldbehörden und der Staat zu beiden Parteien gehören. Hier zeigt sich, dass die Wohngeldbehörden und der Staat davon profitieren, dass die Antragstellenden einem Risiko ausgesetzt werden. Dies ließe befürchten, dass die Wohngeldbehörden und der Staat den Antragstellenden das Risiko aus Eigennutz auferlegen. Da allerdings auch die risikoexponierten Antragstellenden von der Situation profitieren können, kann die Befürchtung widerlegt werden. Das Risiko wird den Antragstellenden auferlegt, damit sie selbst, sowie alle weiteren beteiligten Parteien, davon profitieren können.

Die ausgeglichene Verteilung des Nutzens zwischen Risikoexponierten und Begünstigten zeigt auf, dass den risikoexponierten Antragstellenden das Risiko nicht aus Eigennutz, sondern zur Erreichung eines Allgemeinwohls auferlegt wird. Jedoch ist die Verteilung von Risiken zwischen Risikoexponierten und Begünstigten noch immer sehr ungleich. Nicht nur die Risiken sind ungleich verteilt, sondern auch die Möglichkeit, die Risiken zu kontrollieren. Denn die Informationen, die den risikoexponierten Antragstellenden zur Verfügung gestellt werden, verleihen ihnen nicht die Kontrolle über ihre Daten oder den Verlauf der Antragsbearbeitung. Diejenigen, die das Risiko eingehen, können dieses Risiko also nicht beeinflussen.

Die Entscheidung, ob das Risiko eingegangen wird, wird nicht den Risikoexponierten überlassen – sie können das Risiko nicht frei wählen: „(...) for a risk to be freely chosen, the risk-exposed has to have full information about the risk and has to be capable of making the decision whether or not to accept the risk.”²⁹⁰ Die risikoexponierten Antragstellenden können zwar über die Risiken informiert werden, sie haben deshalb aber trotzdem keine gute Alternative als das Risiko zu akzeptieren. Wer einen Wohngeldantrag stellen will, um ihm zustehende staatliche Leistungen ausgezahlt zu bekommen,

290 Hermansson, Hansson 2007:142

der muss die Risiken in Kauf nehmen. Antragstellende werden dabei nicht gegen ihren Willen dem Risiko ausgesetzt – sie akzeptieren das Risiko, um einen Vorteil zu erlangen, den sie sonst nicht hätten erlangen können.²⁹¹ Das kann in dem Fall dieses Beispiels allerdings nicht als freie Wahl bezeichnet werden.²⁹²

Die Risiken im Datenökosystem wurden zwar durch Datengenerierende und Datenverarbeitende rational kalkuliert, aber für Datenquellen und Datennutzende, die an der Entscheidung nicht beteiligt waren, die jedoch von den potenziellen Schäden zukünftig betroffen sein könnten, entsteht daraus eine Gefahr.²⁹³ Für Datenquellen besteht die Gefahr, dass ihre Daten ohne ihr Wissen zweckentfremdet werden, und für und Datennutzende besteht die Gefahr, dass ihnen verzerrte Ergebnisse zur Verfügung gestellt werden, auf deren Grundlage sie zu falschen Überzeugungen gelangen. In Luhmanns Worten sind Datenquellen und Datennutzende deshalb in Gefahr, weil der potenzielle Schaden durch „(...) Ursachen außerhalb der eigenen Kontrolle (...)“²⁹⁴ entsteht, also von Außenstehenden verursacht wird. Dieser Gefahr, die durch die Entscheidungen von Außenstehenden entstanden ist, seien Datenquellen und Datennutzen- de hilflos ausgesetzt.²⁹⁵

Deshalb können Datenökosysteme verstanden werden als ein Zusammenspiel von Akteuren, in welchem die einen den anderen ein Risiko auferlegen. Datengenerierende haben nicht nur mehr Wissen, sondern auch mehr Kontrolle als die Datenquellen darüber, was mit den Daten passiert, und Datenverarbeitende haben mehr Wissen und Kontrolle darüber, welche Ergebnisse generiert werden als die Datennutzenden. Die Risiken, die dabei entstehen, wie zum Beispiel Zweckentfremdung von personenbezogenen Daten oder missverständliche algorithmische Ergebnisse, werden also durch die Akteure mit mehr Kontrolle verursacht – die Risiken werden allerdings durch die Akteure mit weniger Kontrolle getragen. Deshalb soll in Kapitel 5 erörtert werden, wann es zulässig ist, jemand anderem ein Risiko aufzuerlegen.

291 Hermansson, Hansson 2007:130

292 Im Folgenden wird erörtert, warum man aber die freie Wahl haben sollte und wie sich dies im Kontext von Datenökosystemen umsetzen ließe.

293 Luhmann 1990:161

294 Ebd.:149

295 Ebd.:153

3.3 Die Risiken der Akteure im Datenökosystem

Die Risiken entstehen auf Grund der epistemischen Ungewissheit über die Folgen des Umgangs mit Daten im Datenökosystem. Für alle vier Akteure ist es auf Grund möglicher Zweckentfremdung sowie undurchsichtiger Algorithmen schwierig abzusehen, welche Folgen das Sammeln und Verarbeiten der Daten haben kann. Die epistemische Asymmetrie verschärft die Ungewissheit ausgerechnet für die Akteure, die die Risiken tragen. Es ist aber erst die Kontrollungleichheit, die es den Akteuren mit weniger Wissen und Kontrolle (Datenquellen und Datennutzenden) verwehrt die Entscheidung zu treffen, ob sie die Risiken tragen wollen. Die Akteure mit mehr Wissen und Kontrolle (Datengenerierende und Datenverarbeitende) aber können die Entscheidung treffen, ob die Risiken eingegangen werden. Datengenerierende entscheiden, dass die Daten der Datenquellen gesammelt werden und wie viel Information sie über die Zwecke erhalten. Datenverarbeitende entscheiden, wie die Daten verarbeitet werden und wie die Ergebnisse für die Datennutzenden aufbereitet werden. Auf diese Weise legen Datengenerierende und Datenverarbeitende den Datenquellen und Datennutzenden Risiken auf.

Im Datenökosystem können durch die Kontrollungleichheit ebenfalls zwei unerwünschte Ereignisse mit einer gewissen Wahrscheinlichkeit eintreten. Sie sind den Risiken, die unter 3.1.1 aus einer rein epistemischen Perspektive betrachtet wurden, sehr ähnlich: Für Datenquellen besteht das Risiko, dass sie keine Kontrolle darüber haben, wer zu welchem Zweck auf ihre Daten zugreift. Für Datennutzende besteht das Risiko einem verzerrten Ergebnis zu glauben und somit zu falschen Überzeugungen zu gelangen, weil sie die Entstehung nicht überprüfen können. In beiden Fällen können die Akteure mit mehr Wissen und Kontrolle die Wahrscheinlichkeit, mit welcher das unerwünschte Ereignis eintritt, erhöhen oder verringern.

3.3.1 Risiko für Datenquellen

Ohne Datenquellen könnte es kein Datenökosystem geben – es gäbe keine Daten zum Sammeln und Verarbeiten. Trotzdem haben

Datenquellen kaum Kontrolle über den Datenfluss. Sie haben aber Interesse daran, zu kontrollieren zu welchen Zwecken ihre Daten genutzt werden und wer auf die Daten Zugriff hat. Dies ist ein Interesse an der eigenen informationellen Privatsphäre. Der Begriff der Privatsphäre wird in Kapitel 6 ausführlich definiert und seine Relevanz begründet. An dieser Stelle sei zunächst gezeigt, warum Datenquellen überhaupt um ihre Privatsphäre bangen müssen.

Während die Datenquellen ein Interesse daran haben, die Kontrolle über ihre Daten und damit ihre Privatsphäre zu behalten, haben Datengenerierende und Datenverarbeitende ein Interesse daran, möglichst viele Daten zu sammeln und zu verarbeiten. Das ist darin begründet, dass Datenverarbeitende sehr viele Daten benötigen, um ihre Algorithmen damit trainieren zu können – je mehr Daten, desto besser trainiert ist der Algorithmus. Diese Interessen geraten dadurch in Konflikt miteinander, dass es Zeit kostet, jede Person nach ihrer Zustimmung zu fragen und weil Personen ihre Daten vorenthalten und dadurch die Quantität der verfügbaren Daten verringern könnten.

3.3.2 Risiko für Datennutzende

Ohne die Datennutzenden müsste es kein Datenökosystem geben – wofür sollten Daten gesammelt und verarbeitet werden, wenn die Ergebnisse niemand nutzt? Trotzdem fehlt den Datennutzenden meist essentielles Wissen über die Entstehung der Ergebnisse, denen sie Glauben schenken. Es ist aber wichtig für sie zu erfahren, wie die Ergebnisse eines Datenökosystems zustande gekommen sind, um Fehler erkennen zu können. Dies ist ein Interesse an der transparenten Darstellung der Ergebnisse im Datenökosystem. Was genau Transparenz im Kontext eines Datenökosystems bedeutet und warum sie wichtig ist, wird in Kapitel 7 erörtert. An dieser Stelle sei zunächst gezeigt, warum die Transparenz im Datenökosystem nicht garantiert ist.

Hier besteht ein Konflikt zwischen dem Interesse an Transparenz der Datennutzenden und dem Interesse an wirtschaftlicher Rentabilität der Datenverarbeitenden. Laut dem Datenethiker Brent Mittelstadt und seinen Kolleg:innen ist dieser Interessenskonflikt geprägt von einer Informationsasymmetrie, da Datenverarbeitende

mehr Informationen darüber haben, wie die Ergebnisse zustande kommen.²⁹⁶ Laut Mittelstadt et al. wird oft absichtlich geheim gehalten, wie genau der eingesetzte Algorithmus funktioniert, damit er nicht kopiert oder das Ergebnis des Algorithmus manipuliert wird und somit der Wettbewerbsvorteil verloren gehen kann (Gaming the System): „The commercial viability of data processors in many industries (e.g. credit reporting, high frequency trading) may be threatened by transparency.“²⁹⁷ Um marktfähig zu bleiben, seien manche Datenverarbeitende auf Geheimhaltung angewiesen – Transparenz würde ihr Geschäftsmodell gefährden.

Ein weiteres Problem ist, dass das Offenlegen der Funktionalität von verwendeten Algorithmen für die Datennutzenden, die selbst keine Programmiererfahrung haben, nicht zum Verständnis des Vorgangs beiträgt.²⁹⁸ Trotzdem haben die Datennutzenden ein Interesse daran zu verstehen, wie die Ergebnisse, welche ihre Möglichkeiten oder ihre Entscheidungen beeinflussen, zustande kommen. Insbesondere weil durch dieses Verständnis Fehler durch verzerrte (biased) Daten erkannt werden können und Diskriminierung verhindert werden kann.

3.3.3 Zwischenfazit

Zusammengefasst kann man sagen, dass Datengenerierende durch das Sammeln und Speichern von personenbezogenen Daten das *Risiko* eingehen, dass digital gespeicherte Daten von Datenquellen zweckentfremdet werden und dass die Datenquellen die Kontrolle über ihre Daten verlieren – dadurch entsteht für Datenquellen eine *Gefahr*. Datenverarbeitende gehen durch das Analysieren der Daten mit Hilfe von KI-Systemen und das anschließende Bereitstellen von Ergebnissen das *Risiko* ein, dass die Ergebnisse für Datennutzende nicht mehr nachvollziehbar sind – bei essenziellen Themen, wie Wohngeld oder Berufswahl, kann dadurch eine *Gefahr* für die Datennutzenden entstehen.

296 Mittelstadt et al. 2016:6

297 Ebd.

298 Ebd.

Es ist wichtig zu betonen, dass die *Gefahr* im Datenökosystem dadurch entsteht, dass Datenquellen und Datennutzende bewusst durch Datengenerierende und Datenverarbeitende einem *Risiko* ausgesetzt werden. Denn Datengenerierende und Datenverarbeitende verursachen eine Situation, in welcher es möglich, aber nicht sicher ist, dass ein unerwünschtes Ereignis für Datenquellen und Datennutzenden eintritt – laut Hansson verursachen sie dementsprechend ein Risiko für andere.²⁹⁹ Nun soll erörtert werden, wie diese Risikoauferlegung im Datenökosystem ethisch bewertet werden kann.

299 Hansson 2004:10