

Evaluation zum BMBF-geförderten Projekt „RIKI – Risikokommunikation zur Künstlichen Intelligenz“ im Rahmen des Wissenschaftsjahres 2019

Autoren:

Carolin Otte

Gwendolyn Rohsaint

Yannic Scheuermann

Moritz Schmid

Dr. Philipp Niemann

Gesamtleitung der Evaluation:

Dr. Philipp Niemann

Nationales Institut für Wissenschaftskommunikation gGmbH

Schloss-Wolfsbrunnenweg 33

69118 Heidelberg

+49 721 608 41650

niemann@nawik.de

Inhalt

I.	Evaluationskonzeption & Aufbau des Berichts	1
II.	Kurzzusammenfassung der Evaluationsergebnisse	3
	Fishbowl-Debatten	3
	Youtube-Kanal	3
	Kompendium „Risikokommunikation zur Künstlichen Intelligenz“	4
III.	Evaluation der RIKI-Fishbowl-Debatte zum Thema „Autonomes Fahren und Künstliche Intelligenz“	5
III.1.	Hintergrund und Datenerhebung	5
III.2.	Soziodemografie der Teilnehmer*innen der Veranstaltung	6
III.2.1.	Geschlecht	6
III.2.2.	Alter	6
III.2.3.	Höchster formaler Bildungsabschluss	7
III.2.4.	Berufliche Tätigkeit	8
III.3.	Formatbezogene Erfahrungen und Erwartungen der Teilnehmer*innen	8
III.3.1.	Teilnahme an Veranstaltungen zu einem wissenschaftlichen Thema	8
III.3.2.	Teilnahme an öffentlichen Diskussionen mit Bürgerbeteiligung	9
III.3.3.	Wichtigster Grund für den Besuch der Veranstaltung	10
III.4.	Bezug, Vorkenntnisse und Einschätzungen der Teilnehmer*innen zum Thema autonomes Fahren	10
III.4.1.	Beruflicher oder privater Bezug zum Thema autonomes Fahren	10
III.4.2.	Vorkenntnisse zum Thema autonomes Fahren	11
III.4.3.	Inhaltliche Einschätzungen zum Thema autonomes Fahren	12
III.5.	Grundlegende Ex-Post-Einschätzungen der Teilnehmer*innen	13
III.5.1.	Gesamteindruck von der Diskussion	13
III.5.2.	Inhalt der Diskussion	14

III.5.3.	Erwerb neuen Wissens zum Thema autonomes Fahren	15
III.5.4.	Inhaltliche Einschätzungen zum Thema autonomes Fahren.....	15
III.5.5.	Erneute Teilnahme an einer öffentlichen Diskussion zum Thema autonomes Fahren.....	18
III.6.	Formatbezogene Einschätzungen der Teilnehmer*innen	19
III.6.1.	Zentrale Aspekte der Veranstaltung.....	19
III.6.2.	Diskussion mit den Expert*innen auf Augenhöhe	20
III.6.3.	Anmerkungen zum Stil der Veranstaltung.....	21
III.6.4.	Erneute Teilnahme an einer öffentlichen Diskussion im Stil der Veranstaltung	22
III.7.	Ergebnisse und Diskussion	23
III.8.	Anhänge	25
III.8.1.	Anhang 1) Fragebogen	25
III.8.2.	Anhang 2) Fragen der Leitfadeninterviews.....	30
IV.	Evaluation der RIKI-Fishbowl-Debatte zum Thema „Künstliche Intelligenz in der Medizin“	31
IV.1.	Einleitung.....	31
IV.2.	Soziodemografie der Teilnehmer*innen der Veranstaltung	31
IV.2.1.	Geschlecht	31
IV.2.2.	Alter	32
IV.2.3.	Höchster formaler Bildungsabschluss.....	33
IV.2.4.	Berufliche Tätigkeit.....	33
IV.3.	Formatbezogene Erfahrungen und Erwartungen der Teilnehmer*innen.....	34
IV.3.1.	Teilnahme an Veranstaltungen zu einem wissenschaftlichen Thema.....	34
IV.3.2.	Teilnahme an öffentlichen Diskussionen mit Bürgerbeteiligung.....	35
IV.3.3.	Wichtigster Grund für den Besuch der Veranstaltung	35
IV.4.	Bezug, Vorkenntnisse und Einschätzungen der Teilnehmer*innen zum Thema Künstliche Intelligenz in der Medizin.....	36

IV.4.1.	Beruflicher oder privater Bezug zum Thema Künstliche Intelligenz in der Medizin.....	36
IV.4.2.	Vorkenntnisse zum Thema Künstliche Intelligenz in der Medizin.....	37
IV.4.3.	Inhaltliche Einschätzungen zum Thema Künstliche Intelligenz in der Medizin.....	38
IV.5.	Grundlegende Ex-Post-Einschätzungen der Teilnehmer*innen	39
IV.5.1.	Gesamteindruck der Diskussion.....	39
IV.5.2.	Inhalt der Diskussion	40
IV.5.3.	Erwerb neuen Wissens zum Thema Künstliche Intelligenz in der Medizin	41
IV.5.4.	Inhaltliche Einschätzungen zum Thema Künstliche Intelligenz in der Medizin..	41
IV.5.5.	Erneute Teilnahme an einer öffentlichen Diskussion zum Thema Künstliche Intelligenz in der Medizin.....	44
IV.6.	Formatbezogene Einschätzungen der Teilnehmer*innen	45
IV.6.1.	Zentrale Aspekte der Veranstaltung.....	45
IV.6.2.	Diskussion mit den Expert*innen auf Augenhöhe	46
IV.6.3.	Anmerkungen zum Stil der Veranstaltung.....	47
IV.6.4.	Erneute Teilnahme an einer öffentlichen Diskussion im Stil der Veranstaltung.....	48
IV.7.	Ergebnisse und Diskussion	48
IV.8.	Anhänge	51
IV.8.1.	Anhang 1) Fragebogen.....	51
IV.8.2.	Anhang 2) Fragen der Leitfadeninterviews.....	56
V.	Vergleich der beiden Fishbowl-Debatten.....	58
V.1.	Veranstaltungsqualität.....	58
V.1.1.	Aspekte der Veranstaltung im Vergleich.....	59
V.1.2.	Wie wurde die Qualität der Diskussion empfunden?	60
V.1.3.	Konnten die Diskussionen auf Augenhöhe geführt werden?.....	63
V.1.4.	Wurde das Veranstaltungsformat vom Publikum angenommen?.....	63

V.1.5.	Wie werden öffentliche Diskussionen zu Themen aus dem Bereich Künstliche Intelligenz angenommen?.....	63
V.2.	Problembewusstsein im jeweiligen Themenfeld: Mediale Berichterstattung	64
VI.	Evaluation des Youtube-Kanals des RIKI-Projekts.....	67
VI.1.	Der YouTube-Kanals „Risikokommunikation zur KI“	67
VI.2.	Evaluation des Kanals	68
VI.2.1.	Allgemeine Metriken des Kanals	68
VI.2.2.	Nutzungskontext & Nutzer*innen	75
VI.2.3.	Entwicklung.....	79
VI.2.4.	Detailbetrachtung der Metriken.....	80
VI.3.	Zusammenfassung und Fazit	84
VI.4.	Anhänge	86
VI.4.1.	Anhang 1: Veröffentlichungsdaten der Videos und Aufrufe.....	86
VI.4.2.	Anhang 2: Kommentare.....	87
VII.	Evaluation des Kompodiums „Risikokommunikation zur Künstlichen Intelligenz“	88
VII.1.	Hintergrund.....	88
VII.2.	Methodisches Vorgehen	88
VII.3.	Ergebnisse	90
VII.3.1.	Inhalt.....	90
VII.3.2.	Relevanz.....	91
VII.3.3.	Umsetzung.....	93
VII.4.	Zusammenfassung.....	94
VII.5.	Anhang.....	95
VII.5.1.	Anhang 1: Leitfaden.....	95

I. Evaluationskonzeption & Aufbau des Berichts

Als zentrale Elemente des Projekts „RIKI – Risikokommunikation zur Künstlichen Intelligenz“ mit gesellschaftlichem Wirkungspotenzial werden in dieser Evaluation die beiden öffentlichen Diskussionen im Fishbowl-Format in Karlsruhe und Berlin, der Youtube-Kanal und das Kompendium „Risikokommunikation zur Künstlichen Intelligenz“ betrachtet.

Hinsichtlich der **Fishbowl-Debatten** ist neben der grundsätzlichen Einschätzung der Debatten- und Veranstaltungsqualität von besonderem Interesse, inwieweit sich das Problembewusstsein der Teilnehmer*innen und ihre grundsätzliche Haltung zum Thema Künstliche Intelligenz (KI) durch die Debatten verändert haben. Somit wird erfasst, inwieweit die Veranstaltung den Teilnehmer*innen aus persönlicher Perspektive wie auch inhaltlich etwas gebracht hat. Zu diesem Zweck kam ein Pretest-Posttest-Design zum Einsatz, bei dem die Teilnehmer*innen sowohl vor als auch nach der jeweiligen Veranstaltung an einer standardisierten schriftlichen Befragung teilnahmen (Debatte Karlsruhe: 26 von etwa 30 Teilnehmer*innen; Debatte Berlin: 28 von etwa 30 Teilnehmer*innen). Zudem wurden leitfadengestützte Interviews mit einer Auswahl von Teilnehmer*innen geführt (jeweils sieben bei beiden untersuchten Debatten). Die Untersuchung basiert auf zwei Qualifikationsarbeiten im Studiengang „Wissenschaft – Medien – Kommunikation“ am Karlsruher Institut für Technologie (KIT).

Der **Youtube-Kanal** des Projekts, auf dem die Video-Statements zu KI-Themen von Workshopteilnehmer*innen sowie die filmischen Dokumentationen der Fishbowl-Diskussionen veröffentlicht wurden, wurde einer quantitativen Resonanzanalyse unterzogen. Dabei wurden neben den Abrufzahlen auch andere Metadaten wie die Zahl der Kommentare oder der Likes/Dislikes analysiert.

Das **Kompendium** „Risikokommunikation zur Künstlichen Intelligenz“, das auf den Ergebnissen der Workshops und Fishbowl-Debatten des Projekts aufbaut, soll Wissenschaftler*innen und Öffentlichkeitsarbeiter*innen den Zugang zur KI-Kommunikation erleichtern. Inwieweit dies in den beiden anvisierten Zielgruppen tatsächlich gelingen konnte, wurde durch qualitative Leitfadeninterviews mit KI-Forschenden (elf Befragte) und Öffentlichkeitsarbeiter*innen aus dem KI-Bereich (fünf Befragte) im Rahmen dieser Evaluation ermittelt. Geplant waren ursprünglich 16 Befragungen, die sich gleichmäßig auf die beiden Gruppen verteilen. Durch die Veränderungen aufgrund der Covid19-Pandemie gestaltete sich die Rekrutierung von Öffentlichkeitsarbeiter*innen für die Teilnahme an der Evaluation als überaus schwierig, weshalb aus dieser Gruppe lediglich fünf Personen befragt werden konnten. Durch die

Hinzunahme von drei zusätzlichen KI-Forschenden wurde insgesamt die geplante Zahl von 16 Befragungen erreicht.

Der vorliegende Evaluationsbericht gliedert sich wie folgt: Im Anschluss an eine Kurzzusammenfassung der Ergebnisse (Kapitel II) werden zunächst die Ergebnisse der Evaluation der beiden Fishbowl-Diskussionen „Autonomes Fahren und KI“ in Karlsruhe sowie „KI in der Medizin“ in Berlin dargelegt. Die Analyse basiert auf den BA-Arbeiten von Carolin Otte (Autonomes Fahren und KI) und Gwendolyn Rohsaint (KI in der Medizin) im Studiengang „Wissenschaft – Medien – Kommunikation“ am Karlsruher Institut für Technologie (KIT). Zu beiden Fishbowl-Diskussionen erfolgt zunächst eine detaillierte Betrachtung der erhobenen Einzeldaten (Kapitel III und IV). In Kapitel V werden ausgewählte Untersuchungsaspekte dann noch einmal explizit vergleichend betrachtet. Kapitel VI des Berichts widmet sich der Evaluation des Youtube-Kanals zum RIKI-Projekt, Kapitel VII schließlich der des Kompendiums „Risikokommunikation zur Künstlichen Intelligenz“.

II. Kurzzusammenfassung der Evaluationsergebnisse

Fishbowl-Debatten

Insgesamt zeigt die Evaluation der beiden Diskussionsveranstaltungen „Reden Sie mit: Autonomes Fahren und Künstliche Intelligenz“ (Karlsruhe) und „Künstliche Intelligenz in der Medizin – Reden Sie mit!“ (Berlin), dass Diskussionen im Fishbowl-Format eine sinnvolle Möglichkeit sind, um das Publikum in eine Debatte mit einzubeziehen. Der direkte Austausch zwischen Expert*innen und Publikum funktioniert, wird als „auf Augenhöhe“ wahrgenommen, und fast alle Befragten sind mit der Veranstaltung an sich und auch mit dem Stil der Debatte zufrieden.

Bei der Veranstaltung in Karlsruhe hatten die wenigsten Teilnehmer*innen einen persönlichen oder beruflichen Bezug zum Thema, es war jedoch ein starkes Interesse an der Thematik und ein gewisses Vorwissen vorhanden. Passend dazu gaben über die Hälfte der Teilnehmer*innen an, neues Wissen aus der Veranstaltung mitgenommen zu haben. Bei der Einschätzung des Themas an sich durch die Besucher*innen konnten tatsächliche Veränderungen festgestellt werden.

Die Veranstaltung in der Charité konnte vor allem Besucher*innen aus dem Berufsfeld und mit einem Bezug zum Thema anlocken. Entsprechend hoch schätzten die Besucher*innen selbst ihr Vorwissen ein und entsprechend gering war der Wissenszuwachs durch die Debatte. Auch in diesem Fall hat sich die Einschätzung des Themas durch die Besucher*innen nach der Veranstaltung verändert.

Youtube-Kanal

Die Auswertung des YouTube-Kanals „Risikokommunikation zur KI“ zeigt, dass die Videos vor allem von am Thema interessierten Nutzer*innen angesehen werden. Die verwendeten Suchbegriffe machen deutlich, dass die Nutzer*innen nicht primär nach RIKI suchen, sondern nach generellen Themen im Kontext KI. Dies zeigt sich auch an den Themen der am häufigsten und am längsten betrachteten Videos des Kanals.

Die Videos konnten keine große Anzahl an Abonnenten generieren. Dies lässt darauf schließen, dass zwar die Einzelthemen zum Themenfeld KI interessant sind, das Thema „Risikokommunikation“ im Allgemeinen jedoch weniger.

Insgesamt zeigen die Aufrufe vor allem von der Webseite des RIKI-Projekts, aber auch über die YouTube-Suche, dass in der Öffentlichkeit ein Interesse am Thema KI generell, und in etwas geringerem Maße auch an der Risikokommunikation dazu besteht. Besonders die Möglichkeit,

die öffentliche Fishbowl-Diskussion in Karlsruhe zum Thema Autonomes Fahren und Künstliche Intelligenz nachzuverfolgen, wird von vielen Zuschauer*innen genutzt.

Kompodium „Risikokommunikation zur Künstlichen Intelligenz“

Die 16 aus den beiden Zielgruppen des Kompodiums befragten Forschenden und Kommunikator*innen im KI-Bereich waren mit dem Inhalt des Produkts sehr zufrieden. Insbesondere die ausgewählten Themen, deren Einordnung und ihre Aktualität waren ausschlaggebend dafür, dass das Kompodium mehrfach als gute Übersicht über das Themengebiet KI gelobt wurde. Auch mit der Korrektheit der Darstellung der einzelnen Themenschwerpunkte waren die Befragten, bis auf eine einzelne Kritik an einer Quelle, zufrieden.

Diejenigen Interviewten, die selbst bereits im Bereich Risikokommunikation tätig waren, waren überzeugt davon, dass das Kompodium auch in zukünftigen Kommunikationssituationen zur Rekapitulation von Kommunikationsstrategien bzw. insgesamt als Vorbereitungsmaterial relevant sein wird. Ein Großteil der Befragten wollte das Kompodium weiterempfehlen oder hatte dies bereits getan. Als besonders hilfreich wurde es für Personen mit noch wenig Kommunikationserfahrung eingestuft.

Optimistisch war die Mehrheit der Befragten auch dahingehend, dass das Kompodium zu mehr öffentlicher Kommunikation über KI beitragen wird. Die meisten Interviewten fühlten sich durch das Produkt besser auch auf schwierige Kommunikationssituationen vorbereitet und an die Relevanz öffentlicher Kommunikation erinnert.

Auch mit der gestalterischen Umsetzung waren alle Befragten zufrieden. Lediglich die zusätzlichen Online-Inhalte wurden von den Befragten überhaupt nicht genutzt.

Insgesamt konnte das Kompodium die befragten Forschenden und Kommunikator*innen im KI-Bereich somit in den untersuchten Kategorien Inhalt, Relevanz und Umsetzung überzeugen.

III. Evaluation der RIKI-Fishbowl-Debatte zum Thema „Autonomes Fahren und Künstliche Intelligenz“

III.1. Hintergrund und Datenerhebung

Im Rahmen des Wissenschaftsfestival *EFFEKTE* unter dem Motto „KI – Karlsruher Innovationen“ veranstaltete das Nationale Institut für Wissenschaftskommunikation (NaWik) am 06. Juli 2019 im Existenzgründerzentrum *Perfekt Futur* des Alten Schlachthofs in Karlsruhe eine Fishbowl-Diskussion. Diese trug den Titel: „Reden Sie mit: Autonomes Fahren und Künstliche Intelligenz“. Veranstaltet wurde die Diskussion als Teil des Projekts „Risikokommunikation zur Künstlichen Intelligenz“ (RIKI).¹

Eine Fishbowl-Diskussion ist eine besondere Art der öffentlichen Diskussion, welche die Interaktion von Lai*innen und Expert*innen fördern soll.² Dabei besteht der Diskussionsraum aus zwei Stuhlkreisen. Im inneren, kleineren Kreis sitzen die Expert*innen, es wird ein Stuhl für einen Laien oder eine Laiin (in diesem Falle Bürger*innen der Stadt Karlsruhe) freigehalten. Die Besucher*innen sitzen im äußeren Stuhlkreis und können für einige Zeit in den inneren wechseln, um sich an der Expert*innendiskussion zu beteiligen. Sobald ihre Fragen beantwortet sind, oder sie ihren Diskussionsbeitrag geleistet haben, wechseln die Lai*innen wieder in den äußeren Stuhlkreis, so dass ein ständiger Austausch stattfinden kann.³

An der Diskussion waren folgende Personen als Expert*innen beteiligt: Prof. Dr. Anne Koziolok (Karlsruher Institut für Technologie), Prof. Dr. Marius Zöllner (Forschungszentrum Informatik, Karlsruhe) und Christian Hubschneider (Forschungszentrum Informatik, Karlsruhe). Moderiert wurde die Diskussion von der Journalistin Eva Wolfangel. Während der Veranstaltung waren circa 30 Teilnehmer*innen anwesend.

Zur Evaluation der Diskussion wurde sowohl vor als auch nach der Veranstaltung eine schriftliche Befragung unter den Teilnehmer*innen durchgeführt. Der Fragebogen ist in Anhang 1 (Kapitel III.8.1) beigefügt. Insgesamt wurden 26 Fragebögen ausgefüllt.

Nach der Veranstaltung wurden leitfadengestützte Interviews ergänzend zu den Fragebögen durchgeführt. Es konnten sieben Teilnehmer*innen befragt werden. Die Fragen der Interviews

¹ Programmflyer zum Wissenschaftsfestival *EFFEKTE* 2019, https://www.effekte-karlsruhe.de/fileadmin/templates/web/pdf/Programmheft-EFFEKTE2019_Versand_neu.pdf, Aufgerufen 27.11.2019

² Ebenda

³ Scholz, Lothar; Kelch, Linda: Fish-Bowl-Methode. In: Bundeszentrale für Politische Bildung: Methoden-Kiste. 8. Auflage, 2018, Bonifatius GmbH, Paderborn. S. 28, <http://www.bpb.de/shop/lernen/thema-im-unterricht/36913/methoden-kiste>, Aufgerufen 27.11.2019

lehnten sich thematisch an die Fragen der schriftlichen Befragung an und dienten der inhaltlichen Vertiefung.

III.2. Soziodemografie der Teilnehmer*innen der Veranstaltung

III.2.1. Geschlecht

Von den Besucher*innen beantworteten 19 (73%) die Frage nach ihrem Geschlecht mit „männlich“, fünf (19%) mit „weiblich“ (Abb. III.1).⁴

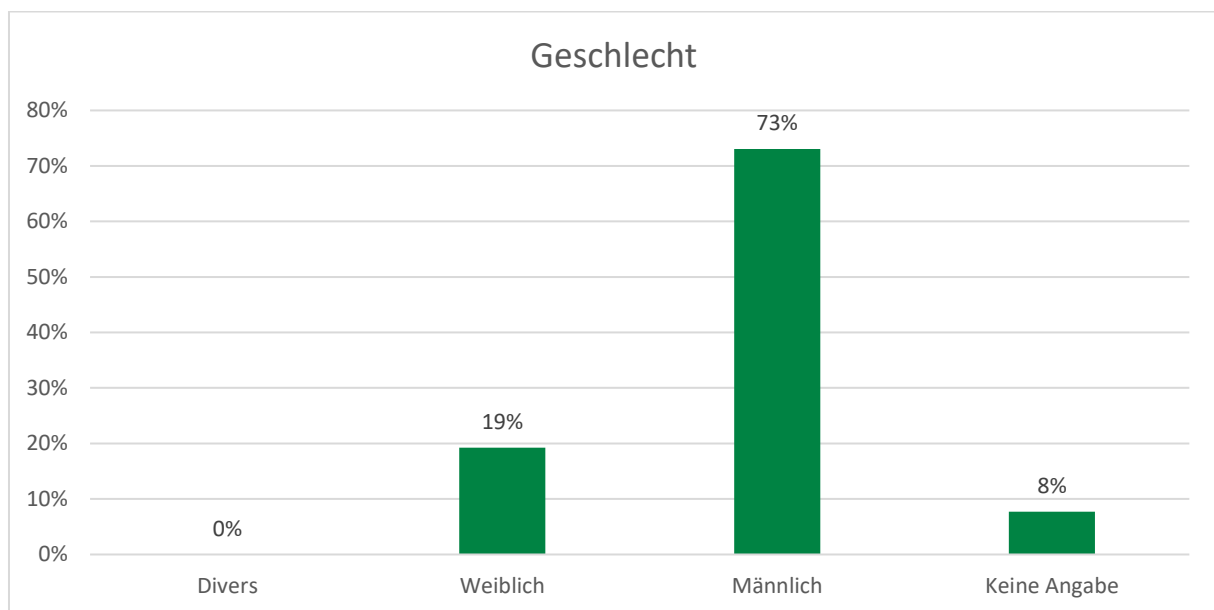


Abbildung III.1: Prozentuale Antworten auf die Frage: „Ihr Geschlecht“. N=26

III.2.2. Alter

Die meisten der Befragten waren zwischen 25 und 34 Jahre alt (46%, n=12). Die zweitgrößte Altersgruppe bildeten sechs (23%) Besucher*innen, die 65 Jahre und älter waren, genauso viele Personen entfielen auf die Gruppen 35-44 und 45-54 Jahre zusammen. Es waren keine Besucher*innen jünger als 25 Jahre oder in der Altersgruppe zwischen 55 und 64 anwesend (Abb. III.2).

⁴ Die Antwort „Keine Angabe“ bedeutet hier und im Folgenden eine nicht ausgefüllte Frage. Dabei konnte nicht bestimmt werden, ob dies bewusst geschah oder ob die Teilnehmer*innen die Frage übersehen haben.

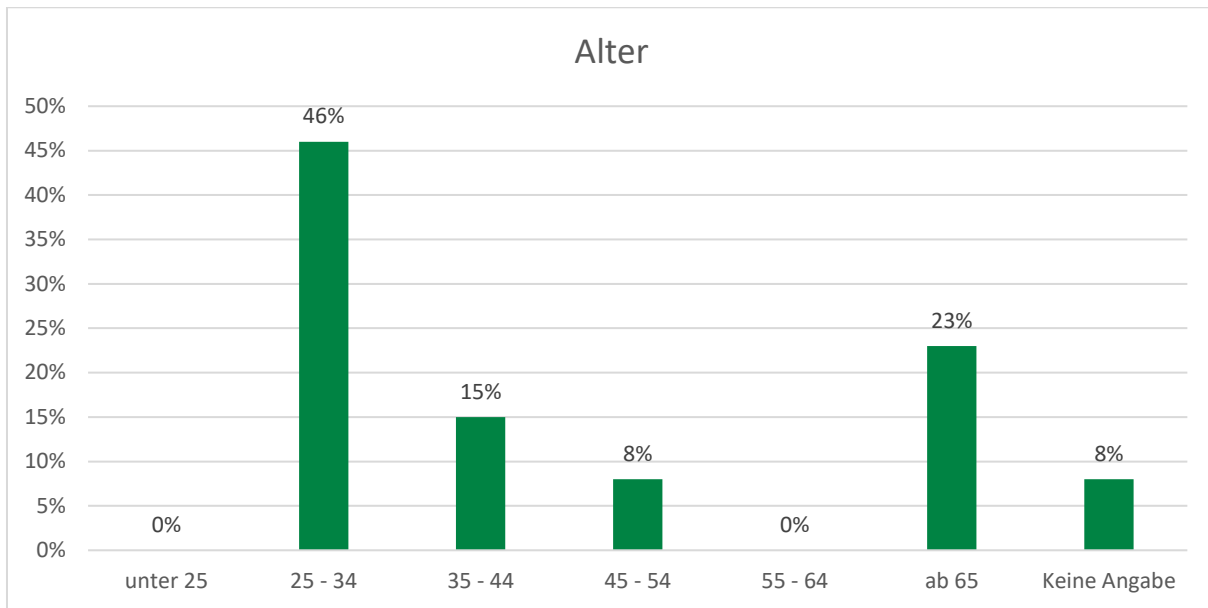


Abbildung III.2: Prozentuale Antworten nach dem Alter der Besucher*innen. N=26

III.2.3. Höchster formaler Bildungsabschluss

Das formale Bildungsniveau der Befragten war überdurchschnittlich hoch. 85% (n=22) hatten einen akademischen Abschluss: 62% (n=16) besaßen einen Hochschulabschluss, 23% (n=6) der Befragten eine Promotion (Abb. III.3).

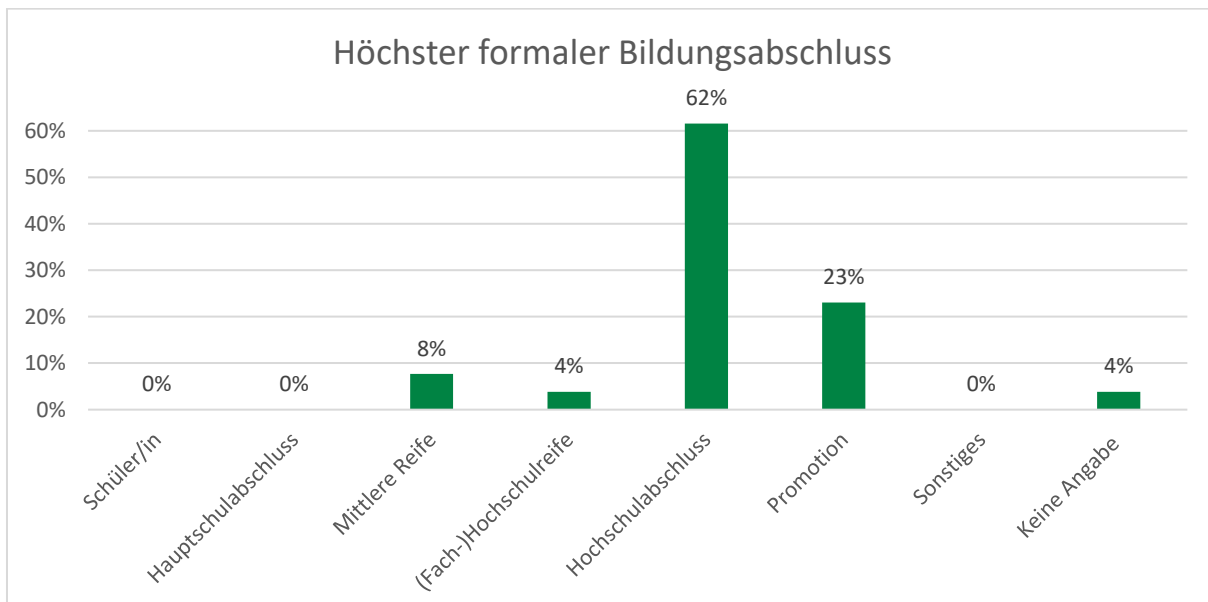


Abbildung III.3: Prozentuale Antworten auf die Frage: „Was ist Ihr höchster Bildungsabschluss?“. N=26

III.2.4. Berufliche Tätigkeit

Die Hälfte der Teilnehmer*innen arbeitete im wissenschaftlichen Bereich (50%, n=13), ein Viertel im kaufmännischen Bereich (23%, n=6) und 15% (n=4) im Ingenieurbereich (Abb. III.4).

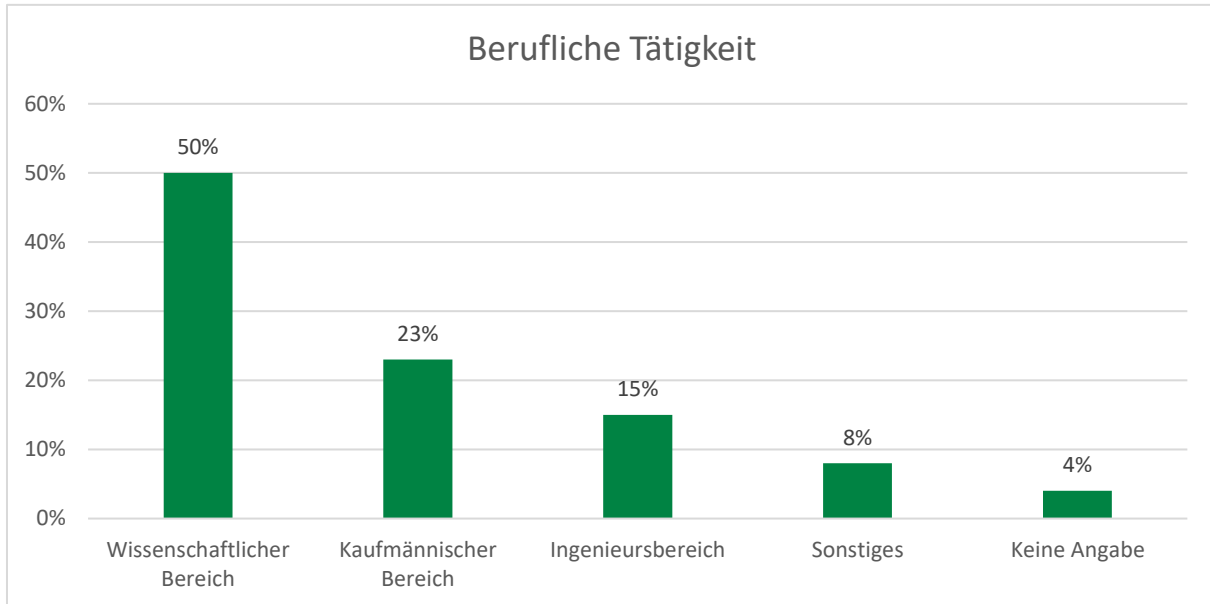


Abbildung III.4: Prozentuale Antworten auf die Frage: „Was ist Ihre berufliche Tätigkeit?“. N=26

III.3. Formatbezogene Erfahrungen und Erwartungen der Teilnehmer*innen

III.3.1. Teilnahme an Veranstaltungen zu einem wissenschaftlichen Thema

Die Antworten aus den Fragebögen zeigen, dass 69% (n=18) der Befragten schon mindestens einmal an einer öffentlichen Veranstaltung zu wissenschaftlichen Themen teilgenommen haben, 54% (n=14) aller Besucher*innen sogar mehrmals. Für knapp ein Viertel der Befragten (23%, n=6) war es die erste öffentliche Veranstaltung zu einem wissenschaftlichen Thema (Abb. III. 5).

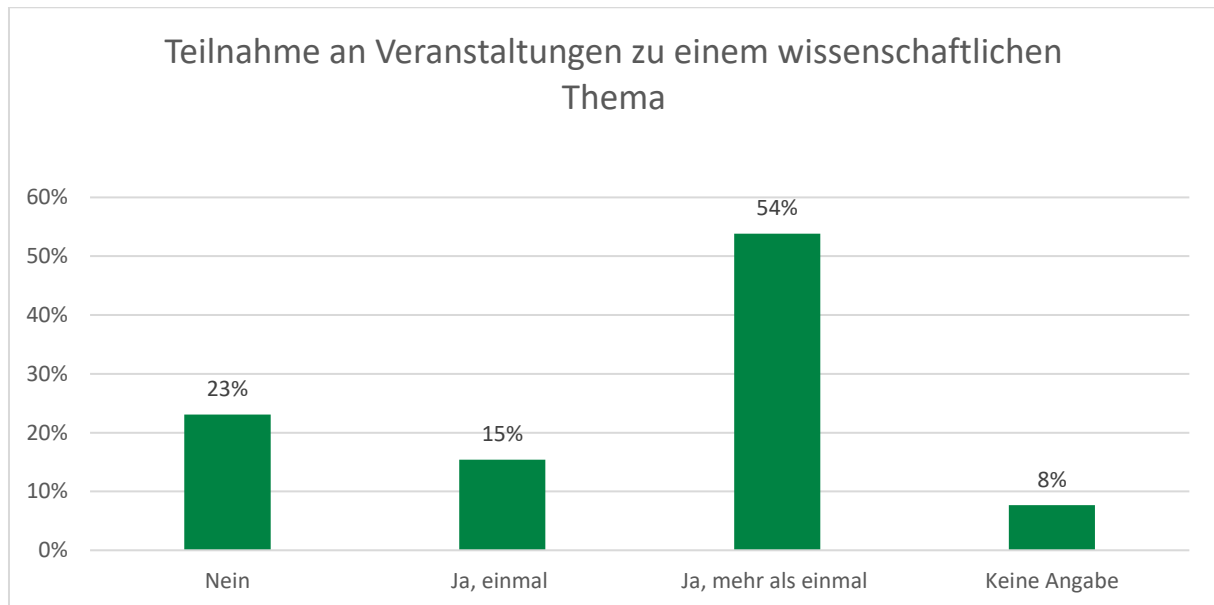


Abbildung III.5: Prozentuale Antworten auf die Frage: „Haben Sie bereits an einer öffentlichen Veranstaltung zu einem wissenschaftlichen Thema teilgenommen?“. N=26

III.3.2. Teilnahme an öffentlichen Diskussionen mit Bürgerbeteiligung

Für 65% (n=17) der Befragten war die Fishbowl-Debatte die erste öffentliche Diskussion mit Bürgerbeteiligung. Knapp unter einem Viertel der Teilnehmer*innen (23%, n=6) haben schon mehrmals an solchen Diskussionen teilgenommen und 8% (n=2) zumindest einmal (Abb. III.6).

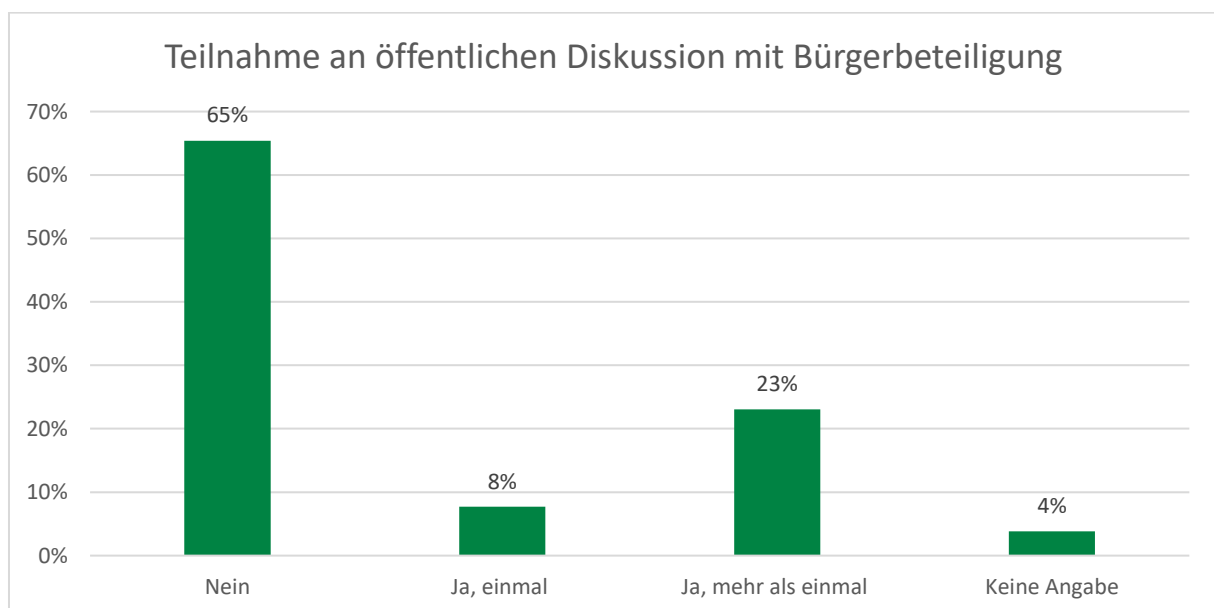


Abbildung III.6: Prozentuale Antworten auf die Frage: „Haben Sie bereits an einer öffentlichen Diskussion mit Bürgerbeteiligung teilgenommen?“. N=26

Alle in den zusätzlichen Interviews befragten Besucher*innen gaben an, dass sie zum ersten Mal an einer Veranstaltung dieses Formats teilgenommen haben.

III.3.3. Wichtigster Grund für den Besuch der Veranstaltung

Die Hälfte der Befragten (50%, n= 10) kamen zu der Veranstaltung, weil sie mehr über autonomes Fahren wissen wollten. Sieben (35%) der Besucher*innen kamen aus einem generellen Interesse an wissenschaftlichen Themen. Bemerkenswert ist, dass ein*e Besucher*in angab, aus Interesse an der Methode – einer Fishbowl-Debatte im öffentlichen Raum – zu der Veranstaltung gekommen zu sein (Abb. III.7). Einige Besucher*innen gaben, trotz Aufforderung nur eine Antwort abzugeben, mehrere Gründe an. Diese konnten nicht mit in die Berechnung aufgenommen werden.

Von den für die Interviews befragten Teilnehmer*innen gaben alle als Grund für den Besuch an, ihr Wissen über das Thema vertiefen zu wollen.

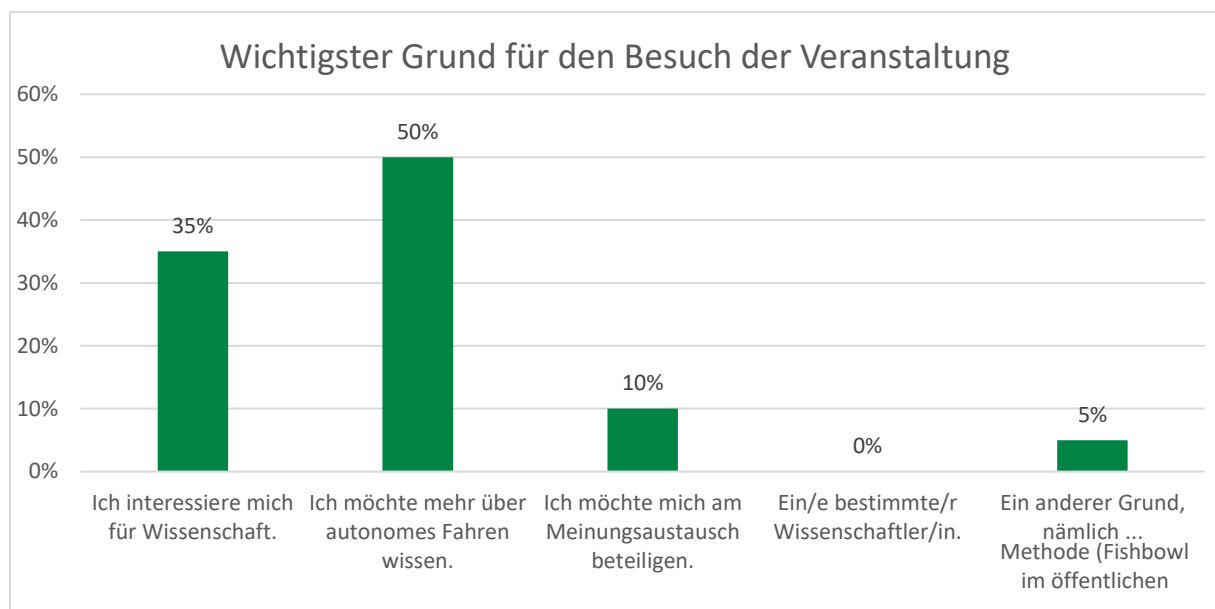


Abbildung III.7: Prozentuale Antworten auf die Frage: „Was war für Sie der wichtigste Grund zur heutigen Veranstaltung zu kommen?“. N=20

III.4. Bezug, Vorkenntnisse und Einschätzungen der Teilnehmer*innen zum Thema autonomes Fahren

III.4.1. Beruflicher oder privater Bezug zum Thema autonomes Fahren

Etwa ein Drittel der Besucher*innen (35%, n=9) hatten einen beruflichen oder privaten Bezug zum Thema autonomes Fahren. 62% (n=16) der Befragten gaben an, keinen solchen Bezug zu haben (Abb. III.8). Dabei teilt sich der berufliche und private Bezug zum Thema in etwa gleiche Anteile auf. Vier Personen (15%) gaben einen Bezug über Studium oder ihre Arbeit in der Forschung an, fünf (19%) einen Bezug über sonstige Berufe zum Thema an (Abb. III.8).

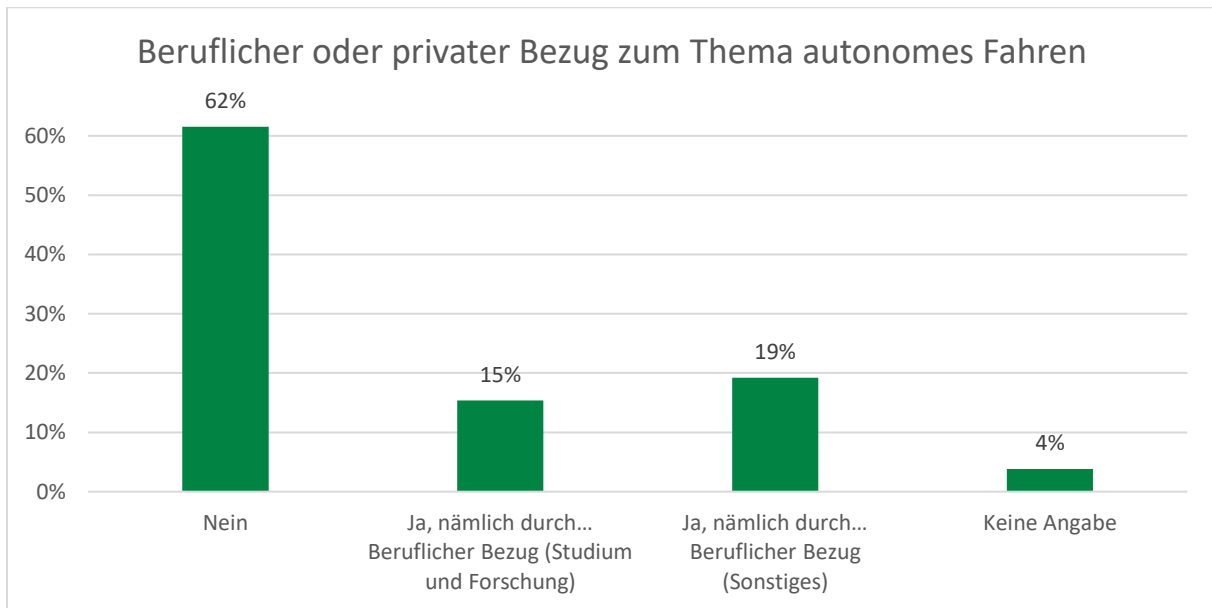


Abbildung III.8: Prozentuale Antworten auf die Frage: „Haben Sie einen beruflichen oder privaten Bezug zum Thema autonomes Fahren?“. N=26

Bis auf eine Person haben alle Teilnehmer*innen an den Interviews einen beruflichen oder studienbedingten Bezug zum Thema Künstliche Intelligenz, Autonomes Fahren oder Ingenieurswesen angegeben.

III.4.2. Vorkenntnisse zum Thema autonomes Fahren

Über die Hälfte der Teilnehmer*innen gaben an, zumindest grundlegend informiert zu sein. Neun Teilnehmer*innen (35%) gaben an, „weder Experte/in noch Laie/in“ zu sein, acht (31%), dass sie „eher Experte/in“ seien, jedoch gab keiner der Besucher*innen an, ein*e „Experte/in“ zu sein (Abb. III.9).

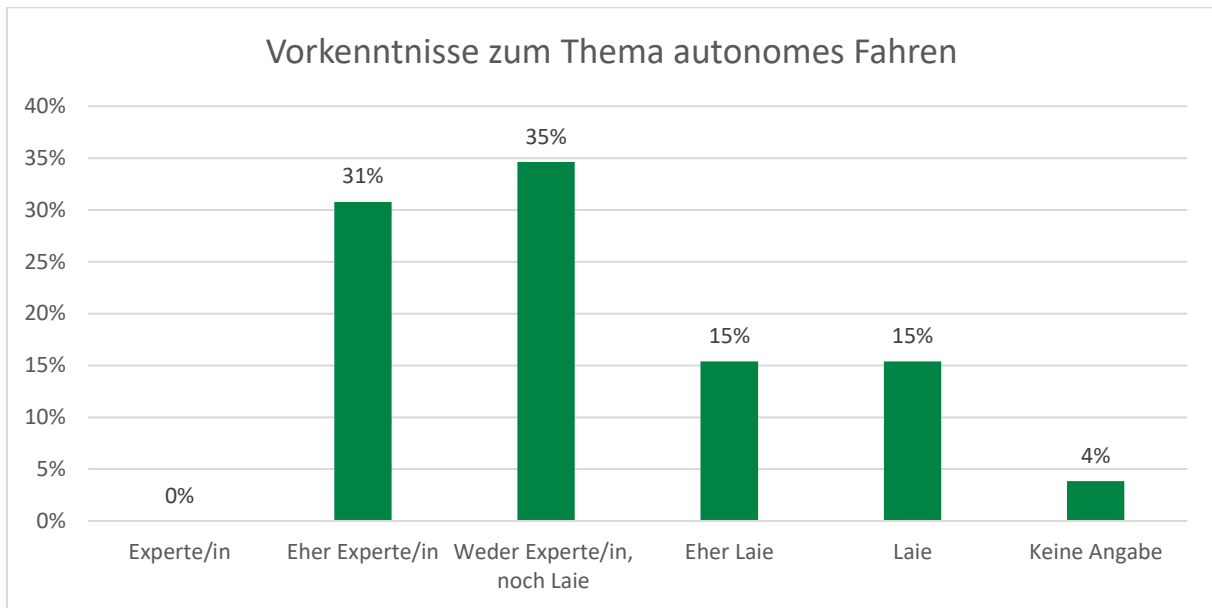


Abbildung III.9: Prozentuale Antworten auf die Frage: „Wie schätzen Sie selbst Ihre Kenntnisse zum Thema autonomes Fahren ein?“. N=26

III.4.3. Inhaltliche Einschätzungen zum Thema autonomes Fahren

Im Vorfeld der Fishbowl-Debatte sollten die Teilnehmer*innen die folgenden Aussagen von „trifft überhaupt nicht zu“ in fünf Stufen bis „trifft voll und ganz zu“ bewerten:

- In wenigen Jahren fahren autonome Fahrzeuge ohne Sicherheitsfahrer/in auf deutschen Straßen.
- Ich fühle mich als Verkehrsteilnehmer/in auf dem Testfeld Autonomes Fahren in Karlsruhe sicher.
- Das Thema autonomes Fahren wird durch die Medien stark problematisiert.
- Die Datensicherheit ist beim autonomen Fahren gewährleistet.
- Die Steuerungskomplexität beim autonomen Fahren stellt kein Sicherheitsproblem dar.
- Ich fühle mich durch die mediale Berichterstattung ausreichend über den Stand der Forschung zum Thema autonomes Fahren informiert.
- Das Thema autonomes Fahren wird durch die Medien stark positiv dargestellt.

Die Befragung zeigt, dass die Teilnehmer*innen die Aussagen eher neutral sahen oder meinten, dass sie nichtzutreffend seien. Auf dem Testfeld Autonomes Fahren (TAF) Karlsruhe fühlte sich jedoch eine deutliche Mehrheit sicher (Abb. III.10).

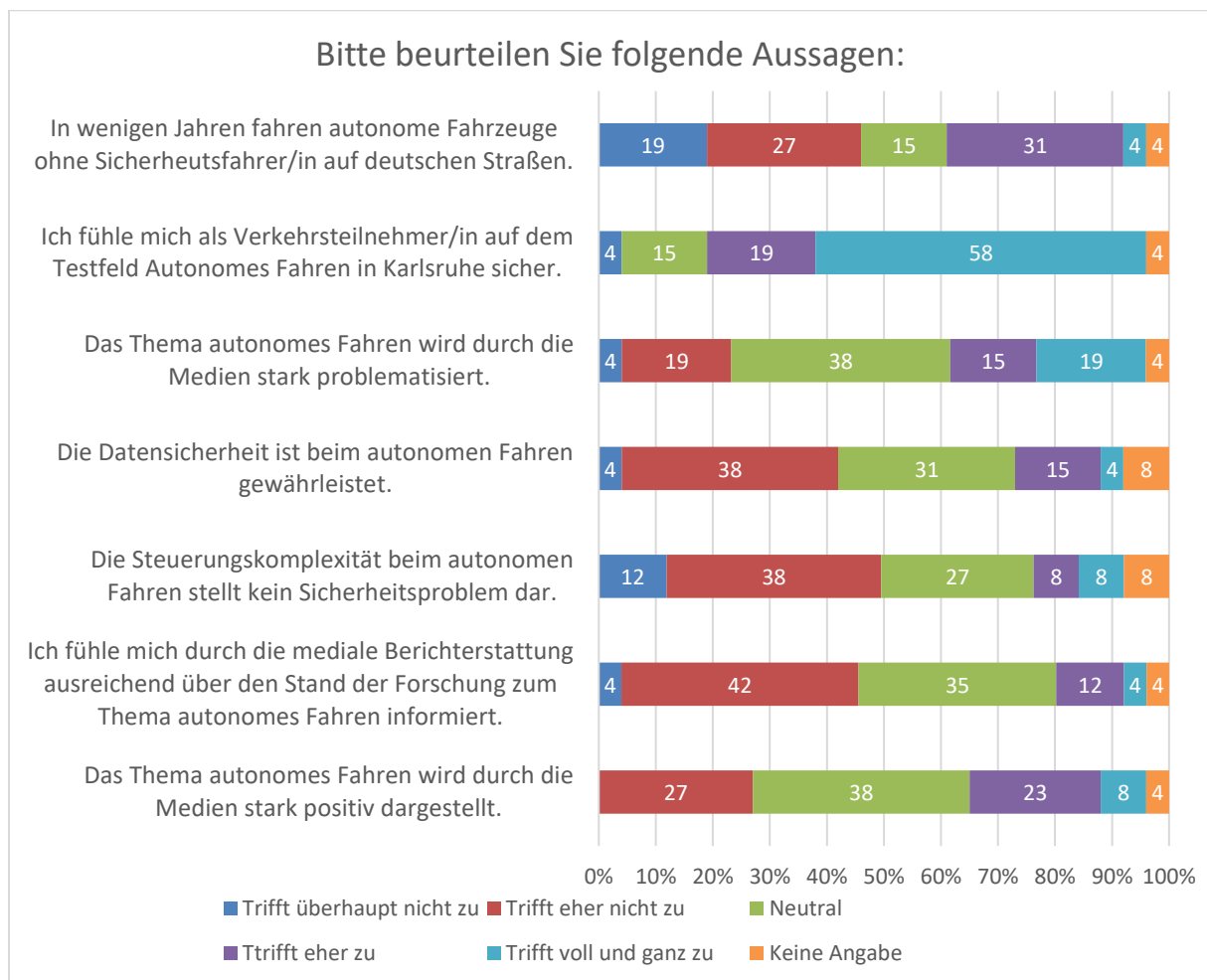


Abbildung III.10: Prozentuale Antworten zur Fragebatterie „Bitte beurteilen Sie die folgenden Aussagen“. Die Befragten sollten sieben inhaltliche Aussagen bewerten. N=26

III.5. Grundlegende Ex-Post-Einschätzungen der Teilnehmer*innen

III.5.1. Gesamteindruck von der Diskussion

Die Besucher*innen sollten im Anschluss an die Veranstaltung angeben, wie ihr Gesamteindruck der Diskussion war (Abb. III.11). Dazu wurden „Schulnoten“ von „sehr gut“ bis „mangelhaft“ vergeben.

Die Fishbowl-Diskussion wurde von der Hälfte der Befragten als „gut“ eingestuft (50%, n=13), von 38% (n=10) als „sehr gut“. Zwei Personen (8%) gaben der Debatte die Note „befriedigend“, niemand empfand sie als „ausreichend“ oder „mangelhaft“.

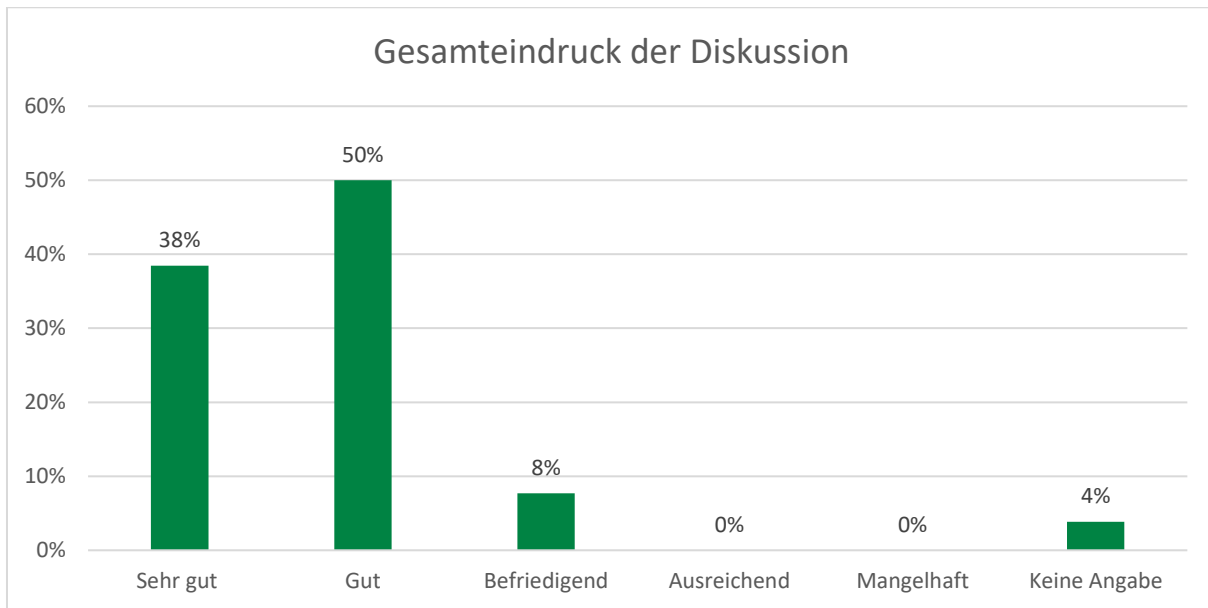


Abbildung III.11: Prozentuale Antworten auf die Frage: „Was ist Ihr Gesamteindruck von der heutigen Diskussion?“. N=26

III.5.2. Inhalt der Diskussion

Die Befragten wurden darum gebeten, kurz wiederzugeben, worum es in der Diskussion ging (Abb. III.12). Die Antworten wurden für die Auswertung thematisch geordnet und zusammengefasst.

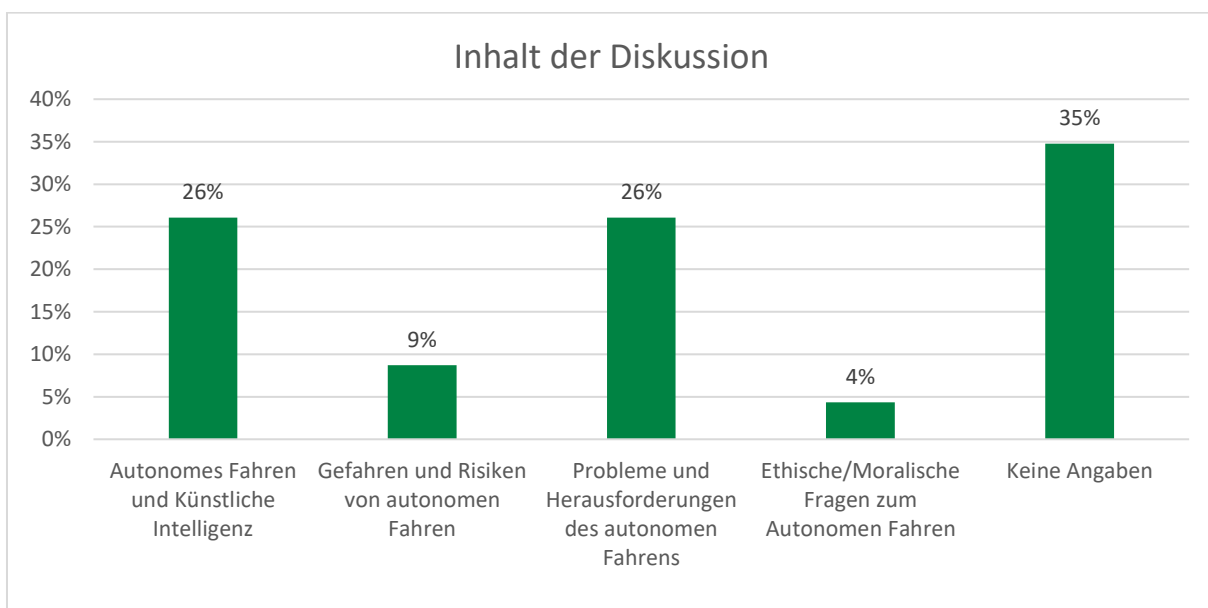


Abbildung III.12: Prozentuale Antworten auf die Frage „In einem Satz zusammengefasst: Worum ging es in der heutigen Diskussion?“. Zur Auswertung wurden die Antworten in Themengruppen zusammengefasst. Mehrfachnennungen waren möglich. N=23

26% (n=6) der Befragten gaben eine Antwort aus dem Bereich „Autonomes Fahren und Künstliche Intelligenz“ als Thema der Diskussion an, was dem Titel der Veranstaltung entspricht. Die gleiche Anzahl an Besucher*innen gab eine Antwort aus dem Bereich „Probleme und Herausforderungen des autonomen Fahrens“. Über ein Drittel (35% n=8) machten keine Angaben zum Thema.

III.5.3. Erwerb neuen Wissens zum Thema autonomes Fahren

Die Teilnehmer*innen sollten die Aussage „Ich habe aus dieser Veranstaltung viel neues Wissen über autonomes Fahren mitgenommen.“ in fünf Stufen von „trifft überhaupt nicht zu“ bis „trifft voll und ganz zu“ beantworten. Für 54% (n=14) der Befragten traf die Aussage „eher zu“, 35% (n=9) bewerteten sie mit „neutral“. Jeweils eine Person (je 4%) beantwortete die Frage mit „trifft überhaupt nicht zu“ und „trifft eher nicht zu“. Niemand gab an, dass die Antwort „voll und ganz“ zutreffen würde (Abb. III.13).

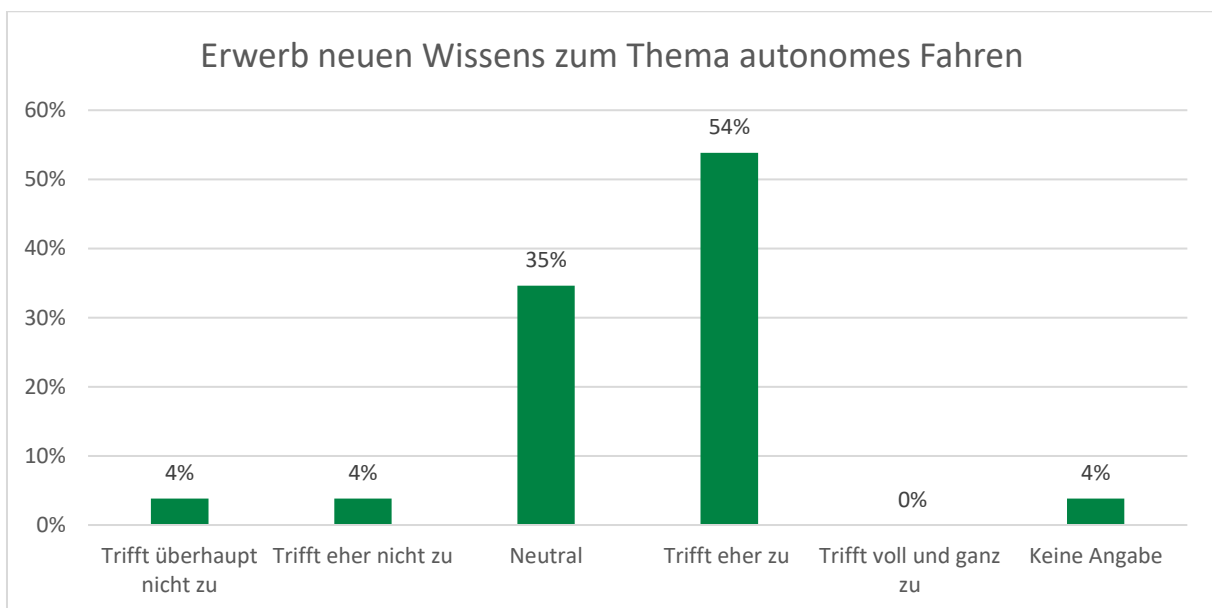


Abbildung III.13: Prozentuale Antworten auf die Frage: „Bitte beurteilen Sie folgende Aussage: Ich habe aus dieser Veranstaltung viel neues Wissen über autonomes Fahren mitgenommen.“. N=26

III.5.4. Inhaltliche Einschätzungen zum Thema autonomes Fahren

Die Besucher*innen sollten nach der Veranstaltung erneut eine Einschätzung zu den bereits im Vorfeld abgefragten inhaltlichen Aussagen zum Thema autonomes Fahren geben (vergleiche Abschnitt 4.3). Die Bewertung erfolgt erneut in fünf Stufen von „trifft überhaupt nicht zu“ bis „trifft voll und ganz zu“ (Abb. III.14).

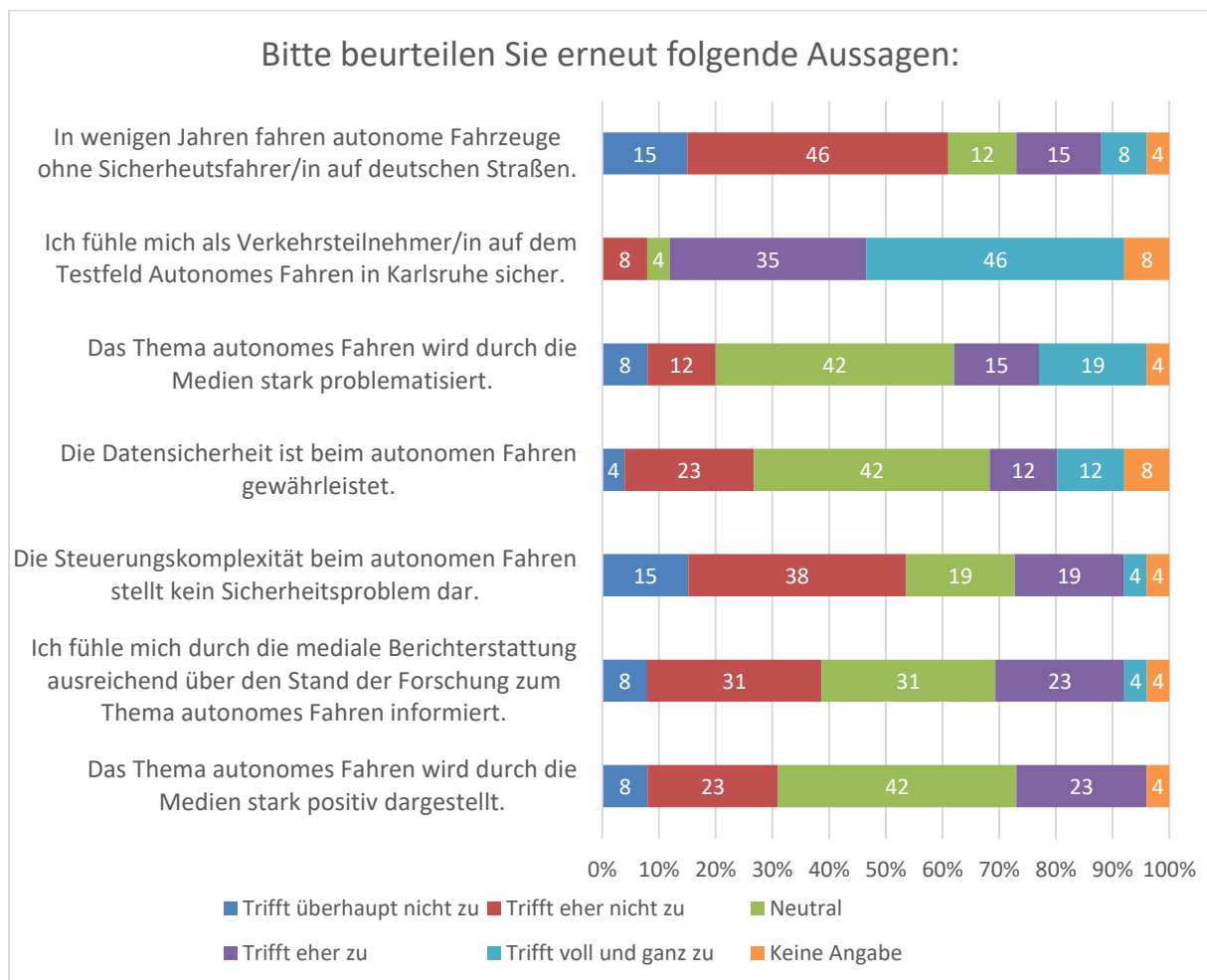


Abbildung III.14: Prozentuale Antworten Fragebatterie „Bitte beurteilen Sie erneut folgende Aussagen“. Erneut wurden sieben inhaltliche Aussagen bewertet. N=26

Bei allen Aussagen lässt sich im Vergleich zu den Aussagen vor der Debatte eine Verschiebung feststellen (Abb. III.15).

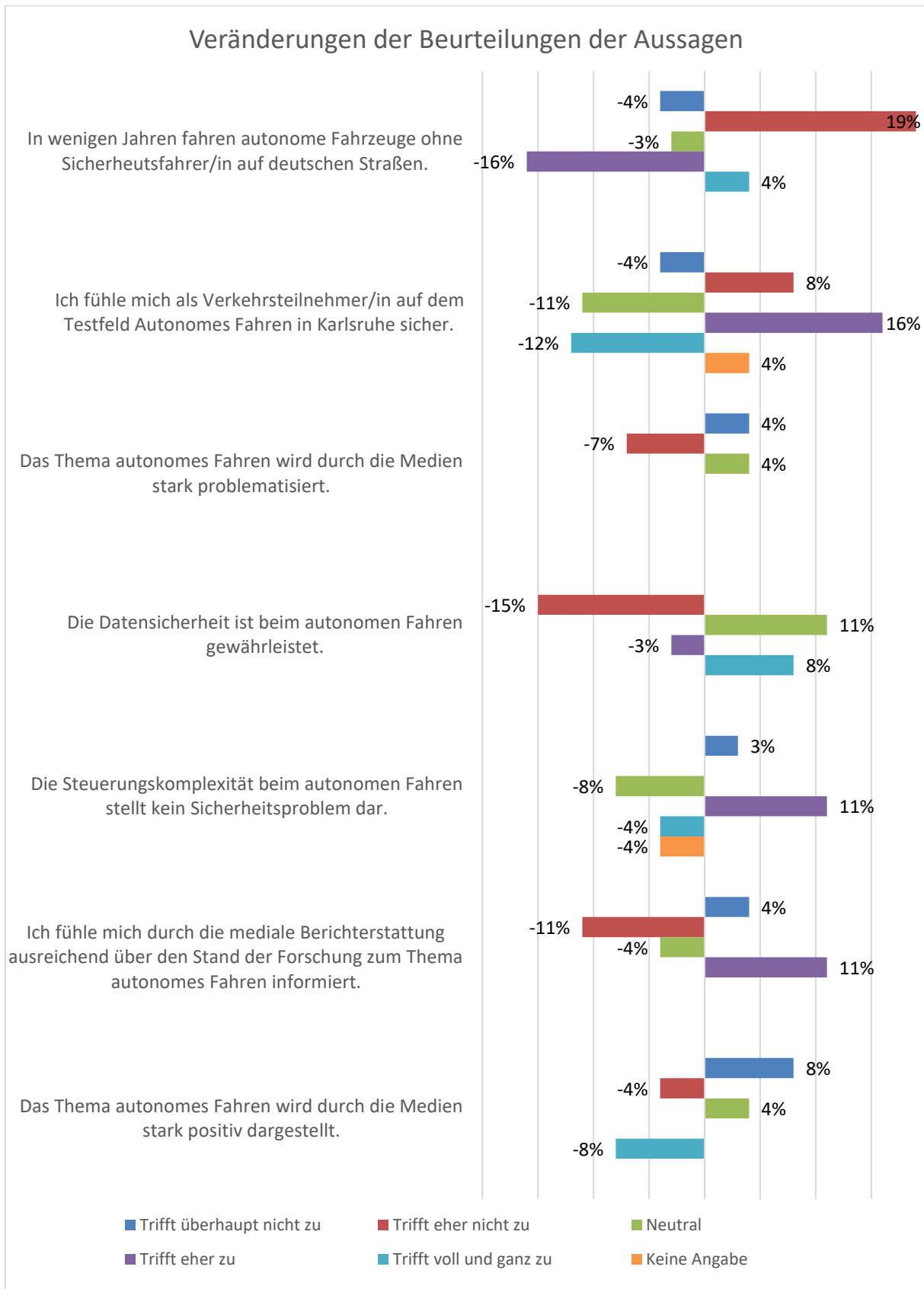


Abbildung III.15: Veränderungen der Beurteilungen der Aussagen zum Thema Autonomes Fahren (Pretest vs. Posttest). N=26

Besonders auffällig ist, dass die Aussage „In wenigen Jahren fahren autonome Fahrzeuge ohne Sicherheitsfahrer/in auf deutschen Straßen“ nun deutlich kritischer eingeschätzt wird: Der Prozentsatz an „trifft eher nicht zu“-Einschätzungen ist von 27% (n=7, pretest) um 19 Prozentpunkte auf 46% (n=12, posttest) gestiegen. Auch die Aussage „Ich fühle mich als Verkehrsteilnehmer/in auf dem Testfeld Autonomes Fahren in Karlsruhe sicher“ wird mit mehr Skepsis gesehen: Die Einschätzung „trifft voll und ganz zu“ fiel um zwölf Prozentpunkte von 58% auf 46%, die Einschätzung „neutral“ um elf Prozentpunkte von 15% auf 4%. Gleichzeitig stieg die Einschätzung „trifft eher nicht zu“ um acht Prozentpunkte von 0% auf 8%. Die Einschätzung „trifft eher zu“ stieg allerdings auch, um 16 Prozentpunkte von 19% auf 35%.

Die Einschätzung der Datensicherheit wurde nach der Diskussion tendenziell positiver bewertet, was sich durch einen Abfall der Aussage „trifft eher nicht zu“ von 38% auf 23% zeigt. Zur Einschätzung der medialen Vermittlungsleistung im Themenbereich autonomes Fahren ist die Aussage „Ich fühle mich durch die mediale Berichterstattung ausreichend über den Stand der Forschung zum Thema autonomes Fahren informiert“ von besonderem Interesse. Hier zeigt sich ein eher positives Bild: Die Einschätzung „trifft eher nicht zu“ fiel um elf Prozentpunkte, während die Einschätzung „trifft eher zu“ um elf Prozentpunkte anstieg.

In den Interviews wurden die Teilnehmer*innen gefragt, welche neuen Denkanstöße sie durch die Veranstaltung erhalten hatten. Folgende Themenbereiche wurden als Antworten genannt:

- Moral/ Ethik/ Philosophie
- Die Zukunftsaussichten von KI
- Der Vergleich zwischen KI und dem Menschen
- Rechtliche und gesetzliche Herausforderungen

Ebenso wurde die Vertiefung von Wissen und das Aufzeigen alternativer Blickwinkel als generelle Auswirkung der Diskussion genannt.

III.5.5. Erneute Teilnahme an einer öffentlichen Diskussion zum Thema autonomes Fahren

24 der Befragten (92%) gaben an, dass sie erneut an einer Veranstaltung zum Thema autonomes Fahren teilnehmen würden. Eine Person (4%) gab an, dass sie keine weitere Veranstaltung zu dem Thema besuchen würde, da es „gegenwärtig keine neue Information“ gebe (Abb. III.16).

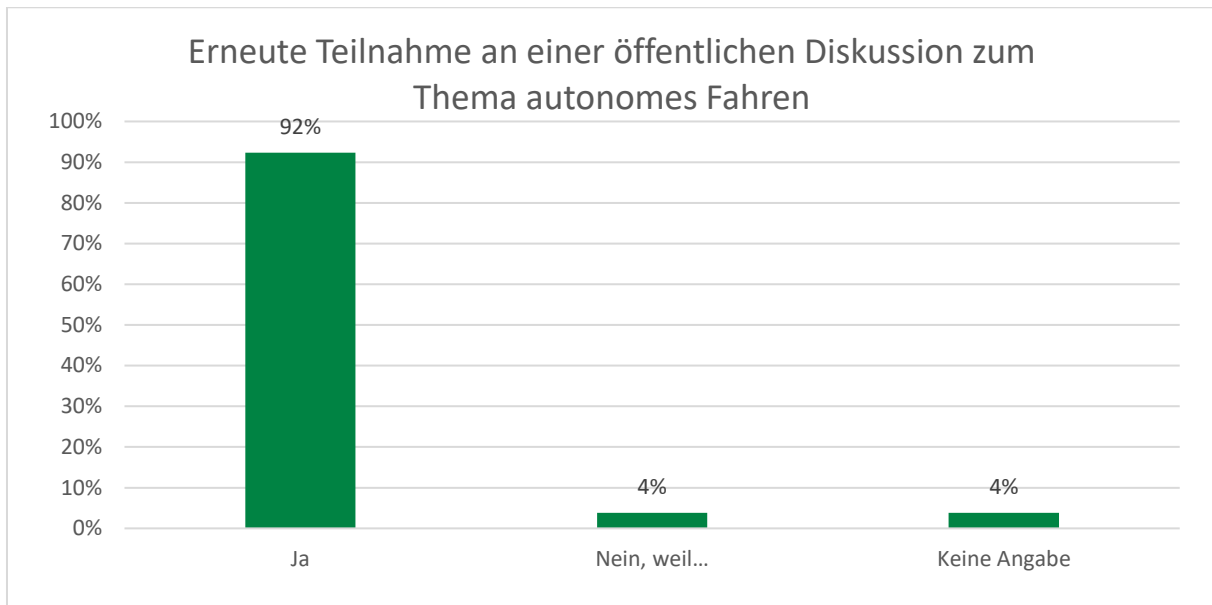


Abbildung III.16: Prozentuale Antworten auf die Frage: „Würden Sie wieder an einer öffentlichen Diskussion zum Thema autonomes Fahren teilnehmen?“. N=26

III.6. Formatbezogene Einschätzungen der Teilnehmer*innen

Die Teilnehmer*innen wurden auch explizit um Einschätzungen zu Form und Stil der Diskussion gebeten.

III.6.1. Zentrale Aspekte der Veranstaltung

Die Besucher*innen sollten die Debatte nach unterschiedlichen Aspekten bewerten (Abb. III.17). Dazu sollten sie „Schulnoten“ von „sehr gut“ bis „mangelhaft“ vergeben. Die Aspekte waren:

- Einleitung und Vorstellung der Experten/innen
- Auswahl der Experten/innen
- Moderation und Diskussionsführung
- Beteiligung der Bürger/innen an der Diskussion
- Interaktion zwischen Experten/innen und Bürger/innen
- Vertrauenswürdigkeit der Experten/innen
- Unterhaltungswert der Veranstaltung
- Veranstaltungsgröße

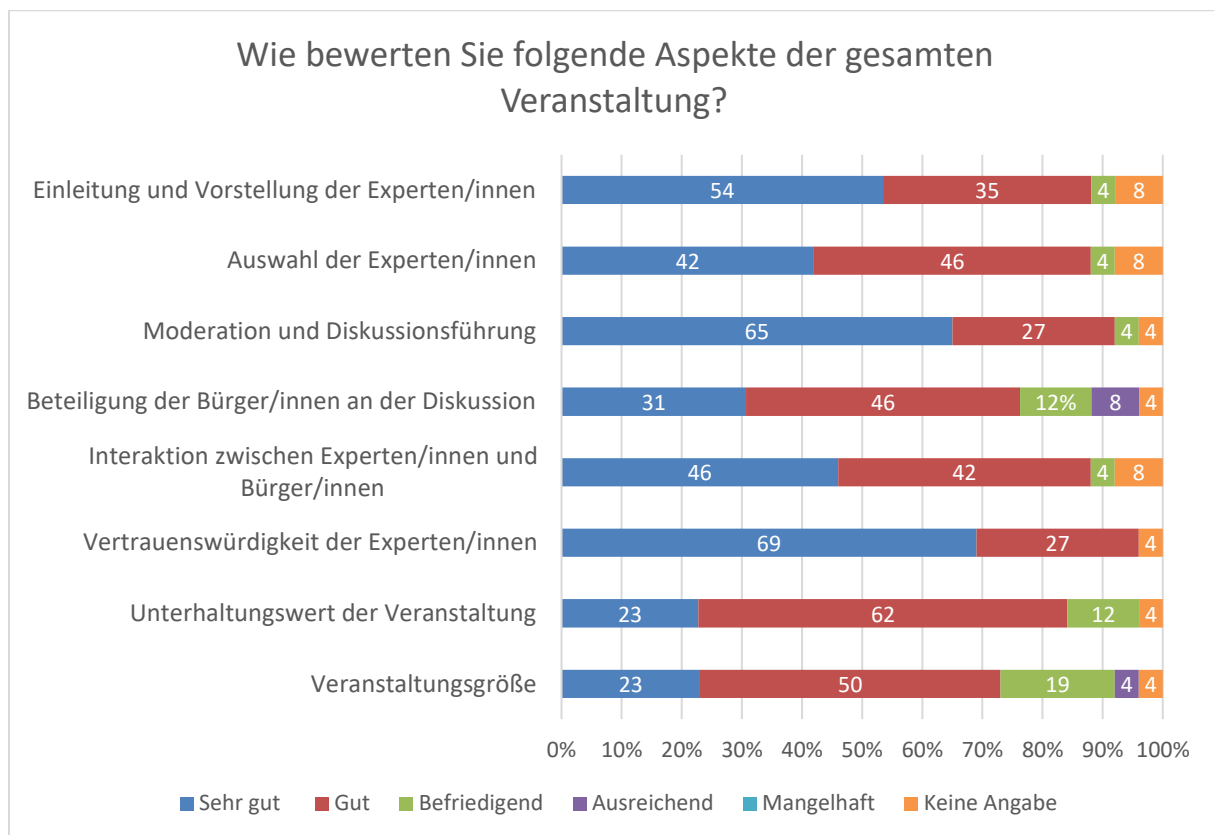


Abbildung III.17: Prozentuale Antworten zur Fragebatterie „Wie bewerten Sie folgende Aspekte der gesamten Veranstaltung“. N=26

Insgesamt bewerteten die Besucher*innen die Veranstaltung als „sehr gut“ oder „gut“. Die Note „mangelhaft“ wurde nicht vergeben. Besonders häufig als „sehr gut“ bewertet wurden die Aspekte „Moderation und Diskussionsführung“ (65%, n=17) und „Vertrauenswürdigkeit der Experten/innen“ (69%, n=18).

In den Interviews gab eine Person sogar an, das Format anderen Diskussionsformaten vorzuziehen, da man als Lai*in die Möglichkeit zur Meinungsäußerung bekäme. Des Weiteren wurden in den Interviews der Veranstaltungsort, die Atmosphäre, die Auswahl der Expert*innen, der thematische Umfang der Diskussion und die Interaktionsmöglichkeit besonders gelobt oder zumindest als angemessen eingestuft.

III.6.2. Diskussion mit den Expert*innen auf Augenhöhe

Die Frage „Hatten Sie das Gefühl, dass die Diskussion mit den Experten/innen auf Augenhöhe stattfand?“ wurde fast vollständig bejaht (96%, n=25), mit nur einer Enthaltung (4%) (Abb. III.18).

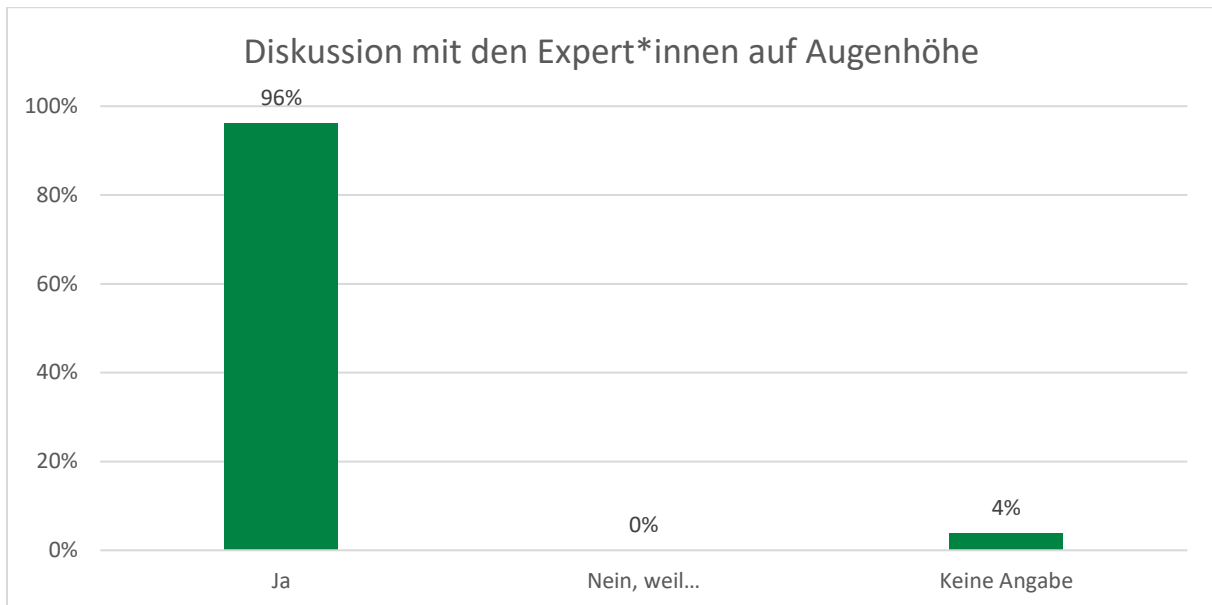


Abbildung III.18: Prozentuale Antworten auf die Frage: „Hatten Sie das Gefühl, dass die Diskussion mit den Experten/innen auf Augenhöhe stattfand?“. N=26

Auch die leitfadengestützten Interviews mit sieben Teilnehmer*innen der Fishbowl-Debatte bestätigten die Aussage, dass die Diskussion auf Augenhöhe mit den Expert*innen stattfand. Die Befragten gaben an, dass es sie Überwindung gekostet hätte, sich an der Diskussion zu beteiligen. Gründe dafür waren Unsicherheit und mangelndes Wissen zum Thema KI und autonomes Fahren, sowie der zeitliche Aufwand, den das Formulieren einer Frage benötigte. Die Diskussion sei häufig bereits bei einem anderen Thema gewesen, bevor die Frage in den Köpfen der Teilnehmer*innen formuliert werden können. Ebenso stellten sich den Teilnehmer*innen oft kleinere Fragen, die sie als nicht diskussionswürdig erachteten und deshalb nicht stellten. Eine weitere Hemmschwelle war die Scheu, nicht den ersten Diskussionsbeitrag stellen zu wollen; diese Hemmung nahm erst im Laufe des Gesprächs ab, als einige Fragen dann schon beantwortet waren. Später stellte sich der Mangel an freien Plätzen im Innenkreis als Hindernis heraus, da nicht für alle Interessent*innen ein Platz im Kreis vorhanden war, nachdem die anfängliche Zurückhaltung überwunden worden war.

III.6.3. Anmerkungen zum Stil der Veranstaltung

Die Teilnehmer*innen wurden gefragt, ob sie Anmerkungen zum Stil der Veranstaltung hätten. Für die Auswertung wurden die Antworten in positive und negative Bemerkungen zusammengefasst. Die Mehrheit der Teilnehmer*innen (82%, n=18) hatte keine Anmerkungen zum Stil der Veranstaltung. Unter den wenigen erfolgten Anmerkungen wurden die Expert*innenauswahl sowie die "schöne Location" positiv hervorgehoben. Die Diskussion selbst wurde als offen und transparent gelobt. Zwei Teilnehmer*innen wünschten sich allerdings

längere Expert*inneninputs. Außerdem wurde in einem Fragebogen die Veranstaltungsankündigung als unzureichend kritisiert und der Debatte vorgeworfen, dass sie sich "leicht in wenig relevante Fragen" verlief. Ein*e Teilnehmer*in kritisierte das Zeitmanagement.

In den leitfadengestützten Interviews wurden zusätzlich verschiedene Vorschläge zur Verbesserung gemacht. Darunter wurde genannt:

- Die Anzahl der Plätze im äußeren Kreis verkleinern.
- Der Diskussion mehr Zeit einräumen.
- Die Teilnehmer*innen stärker aufzufordern und unterstützen, ihre anfängliche Zurückhaltung zu überwinden.

Die Interviewpartner*innen empfanden es als negativ, dass einige Fragen im Diskussionsverlauf untergingen und merkten an, dass es interessant gewesen wäre, mehr gegensätzliche Statements zu diskutieren.

III.6.4. Erneute Teilnahme an einer öffentlichen Diskussion im Stil der Veranstaltung

Die große Mehrheit der Befragten (96%, n=25) würde wieder an einer Fishbowl-Debatte teilnehmen, niemand antwortete mit „nein“ (Abb. III.19). Nur eine Person (4%) machte keine Angabe zu der Frage.

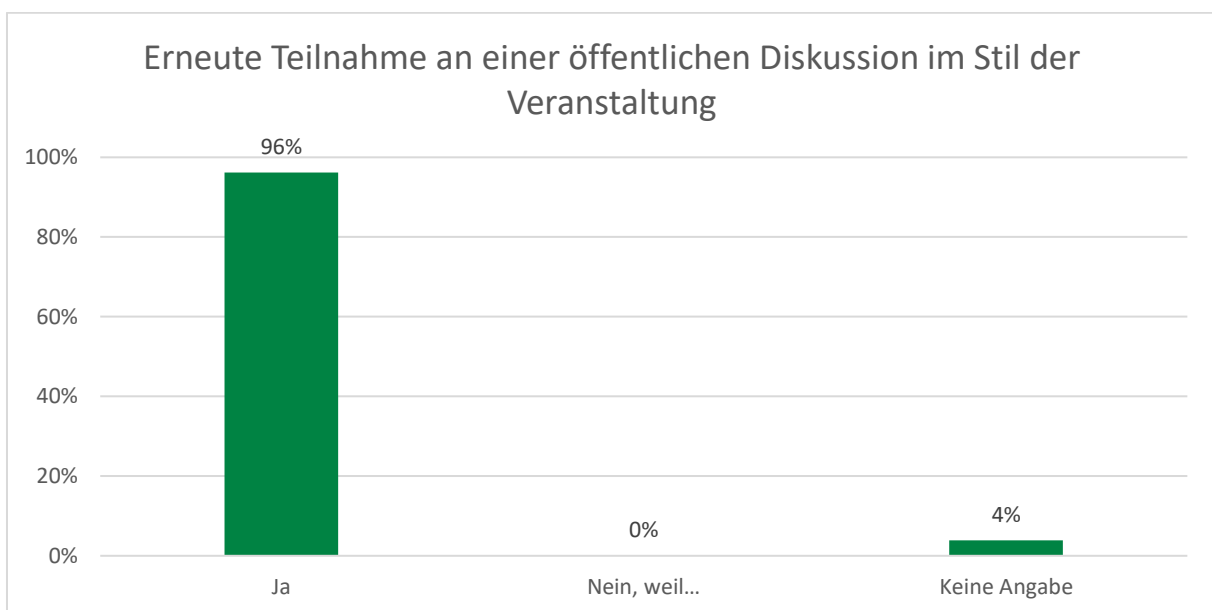


Abbildung III.19: Prozentuale Antworten auf die Frage „Würden Sie wieder an einer öffentlichen Diskussion im Stil der heutigen Veranstaltung teilnehmen. N=26“

In den Interviews bekräftigten die Befragten, dass sie die Fishbowl-Debatte als interessante und lehrreiche neue Erfahrung ansahen und auch ein weiteres Mal teilnehmen würden.

III.7. Ergebnisse und Diskussion

Die Analyse hat deutlich gemacht, dass die meisten der Teilnehmer*innen der evaluierten Fishbowl-Debatte männlich waren und aus dem Hochschulumfeld (durch Arbeit oder Studium) kamen. Die größte einzelne Altersgruppe mit zwölf Personen war zwischen 25-35 Jahren, der Großteil der Gäste war allerdings älter als 35. Die meisten Besucher*innen der Debatte hatten bereits an einer öffentlichen Veranstaltung zu wissenschaftlichen Themen teilgenommen, jedoch war dies für die Mehrheit die erste öffentliche Veranstaltung mit Bürgerbeteiligung.

Insgesamt hatten die wenigsten Teilnehmer*innen einen persönlichen oder beruflichen Bezug zum Thema autonomes Fahren. Dies steht in Einklang damit, dass die meisten Besucher*innen die Veranstaltung mit dem Wunsch besucht haben, mehr über autonomes Fahren zu lernen. Der zweithäufigste Grund war ein generelles Interesse an wissenschaftlichen Themen. Dennoch gab der Großteil der Befragten an, ein Wissen zumindest auf der Stufe „weder Experte/in noch Laie/in“ zu besitzen, ein Drittel schätzte sein Wissen sogar als „eher Experte/in“ ein. Das bedeutet, dass trotz fehlendem Bezug ein Interesse am Thema und ein gewisses Vorwissen vorhanden waren.

Die Bewertung der verschiedenen Aussagen zum Thema autonomes Fahren zeigt besonders, dass sich die Teilnehmer*innen auf dem TAF Karlsruhe sicher fühlten. Bei allen anderen Aussagen waren die Besucher*innen eher gespalten und die Aussagen wurden weiter gestreut eingeschätzt. Bei den meisten Fragen gab grob ein Drittel immer „neutral“ an, was darauf hindeutet, dass sie sich im Vorfeld keine Gedanken darüber gemacht haben.

Die Befragung nach der Diskussion zeigt, dass das Modell der Fishbowl-Debatte im Grundsatz gut funktioniert. Der Gesamteindruck der Veranstaltung wurde von beinahe allen Teilnehmer*innen als „gut“ oder „sehr gut“ bezeichnet. Passend dazu ist die Selbsteinschätzung der Teilnehmer*innen, „viel neues Wissen“ aus der Veranstaltung mitgenommen zu haben.

Wie gut die Diskussion von den Befragten aufgenommen wurde, zeigt besonders die Einzelbenotung der verschiedenen Aspekte der Veranstaltung. Sie wurde fast durchgängig mit „gut“ oder „sehr gut“ bewertet. Besonders gut wurde die Vertrauenswürdigkeit der Expert*innen bewertet.

Alle Besucher*innen, die dazu eine Antwort abgegeben haben, fanden, dass die Gespräche auf Augenhöhe stattfanden. Die meisten Besucher*innen würden sowohl erneut an Debatten zum Thema autonomes Fahren als auch an einer Diskussion im Stil einer Fishbowl-Debatte

teilnehmen. Die meisten Teilnehmer*innen hatten darüber hinaus keine weiteren Anmerkungen zum Stil der Debatte. Negative Anmerkungen betrafen das Zeitmanagement und die Relevanz der Fragen. Die positiven Anmerkungen lobten unter anderem die Örtlichkeit sowie die Atmosphäre der Diskussion.

Die zweite, der Diskussion nachgelagerte Befragung zu den inhaltlichen Aussagen zum Thema autonomes Fahren macht deutlich, dass sich die Einschätzungen der Teilnehmer*innen nach der Diskussion geändert hat, wobei diese nicht unbedingt positiver geworden ist. So fühlten sich mehr der Befragten auf dem TAF Karlsruhe weniger sicher, gleichzeitig wurde aber die Datensicherheit beim autonomen Fahren positiver eingeschätzt. Deutlich gestiegen ist die Skepsis gegenüber der Vorstellung, auf deutschen Straßen in absehbarer Zeit autonome Fahrzeuge ohne Sicherheitsfahrer anzutreffen.

Insgesamt zeigt die Evaluation der Veranstaltung *„Reden Sie mit: Autonomes Fahren und Künstliche Intelligenz“*, dass Diskussionen im Fishbowl-Format eine sinnvolle Möglichkeit sind, um das Publikum in eine Debatte mit einzubeziehen. Der direkte Austausch zwischen Expert*innen und Publikum funktioniert, wird als „auf Augenhöhe“ wahrgenommen und fast alle Befragten sind mit der Veranstaltung an sich und auch mit dem Stil der Debatte zufrieden.

III.8. Anhänge

III.8.1. Anhang 1) Fragebogen

Öffentliche Diskussion: Autonomes Fahren und Künstliche Intelligenz

Vielen Dank im Voraus für Ihre Mithilfe bei unserer Forschung!

Wir möchten Sie bitten, **vor** Beginn der Veranstaltung den **Fragebogen 1** und **nach** der Veranstaltung den **Fragebogen 2** auszufüllen.

Die Fragebogenerhebung findet im Rahmen einer Evaluation des Projekts **„Risikokommunikation zur Künstlichen Intelligenz“** durch das Nationale Institut für Wissenschaftskommunikation (NaWik) statt.

Alle im Fragebogen erhobenen Daten fließen ausschließlich anonymisiert in die Studie ein. Eine Weitergabe der Daten an Dritte erfolgt nicht.

Kontakt

Carolin Otte	carolin@otte.info
Gwendolyn Rohsaint	grohsaint@gmail.com
Dr. Philipp Niemann	niemann@nawik.de

Öffentliche Diskussion: Autonomes Fahren und Künstliche Intelligenz

Fragebogen 1 – Bitte beantworten Sie diese Fragen vor der Veranstaltung.

1) Haben Sie bereits an einer öffentlichen Veranstaltung zu einem wissenschaftlichen Thema teilgenommen?

Ja, einmal Ja, mehr als einmal Nein

2) Haben Sie bereits an einer öffentlichen Diskussion mit Bürgerbeteiligung teilgenommen?

Ja, einmal Ja, mehr als einmal Nein

3) Haben Sie einen beruflichen oder privaten Bezug zum Thema autonomes Fahren?

Nein Ja, nämlich.....

4) Was war für Sie der wichtigste Grund zur heutigen Veranstaltung zu kommen?
(bitte nur eine Antwortmöglichkeit ankreuzen)

Ich interessiere mich für Wissenschaft.

Ich möchte mehr über autonomes Fahren wissen.

Ich möchte mich am Meinungsaustausch beteiligen.

Ein/e bestimmte/r Wissenschaftler/in

Ein anderer Grund, nämlich.....

5) Was ist Ihr höchster Bildungsabschluss?

Schüler/in Hauptschulabschluss Mittlere Reife Sonstiges:

(Fach-) Hochschulreife Hochschulabschluss Promotion

6) Was ist Ihre berufliche Tätigkeit?

.....

7) Ihr Geburtsjahr: **8) Ihr Geschlecht:**

..... weiblich männlich divers

Öffentliche Diskussion: Autonomes Fahren und Künstliche Intelligenz**Fragebogen 1** – Bitte beantworten Sie diese Fragen vor der Veranstaltung.

9) Wie schätzen Sie selbst Ihre Kenntnisse zum Thema autonomes Fahren ein?

Experte/in				Laie	
1	2	3	4	5	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

10) Bitte beurteilen Sie folgende Aussagen:

	trifft voll und ganz zu			trifft überhaupt nicht zu	
a) In wenigen Jahren fahren autonome Fahrzeuge ohne Sicherheitsfahrer/in auf deutschen Straßen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Ich fühle mich als Verkehrsteilnehmer/in auf dem Testfeld Autonomes Fahren in Karlsruhe sicher.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Das Thema autonomes Fahren wird durch die Medien stark problematisiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Die Datensicherheit ist beim autonomen Fahren gewährleistet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Die Steuerungskomplexität beim autonomen Fahren stellt kein Sicherheitsproblem dar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Ich fühle mich durch die mediale Berichterstattung ausreichend über den Stand der Forschung zum Thema autonomes Fahren informiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Das Thema autonomes Fahren wird durch die Medien stark positiv dargestellt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vielen Dank für Ihre Teilnahme!**Bitte füllen Sie den Fragebogen 2 erst nach der Veranstaltung aus.**

Öffentliche Diskussion: Autonomes Fahren und Künstliche Intelligenz

Fragebogen 2 – Bitte beantworten Sie diese Fragen nach der Veranstaltung.

1) Wie ist Ihr Gesamteindruck von der heutigen Diskussion?

sehr gut gut befriedigend ausreichend mangelhaft

2) In einem Satz zusammengefasst: Worum ging es in der heutigen Diskussion?

.....

3) Wie bewerten Sie folgende Aspekte der gesamten Veranstaltung?

	sehr gut	gut	befriedigend	ausreichend	mangelhaft
a) Einleitung und Vorstellung der Experten/innen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Auswahl der Experten/innen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Moderation und Diskussionsführung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Beteiligung der Bürger/innen an der Diskussion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Interaktion zwischen Experten/innen und Bürger/innen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Vertrauenswürdigkeit der Experten/innen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Unterhaltungswert der Veranstaltung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
h) Veranstaltungsgröße	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4) Hatten Sie das Gefühl, dass die Diskussion mit den Experten/innen auf Augenhöhe stattfand?

Ja Nein, weil

5) Bitte beurteilen Sie die folgende Aussage:

Ich habe aus dieser Veranstaltung viel neues Wissen über autonomes Fahren mitgenommen.

trifft voll und ganz zu trifft überhaupt nicht zu



Öffentliche Diskussion: Autonomes Fahren und Künstliche Intelligenz

Fragebogen 2 – Bitte beantworten Sie diese Fragen nach der Veranstaltung.

6) Würden Sie wieder an einer öffentlichen Diskussion zum Thema autonomes Fahren teilnehmen?

Ja Nein, weil

7) Würden Sie wieder an einer öffentlichen Diskussion im Stil der heutigen Veranstaltung teilnehmen?

Ja Nein, weil

8) Haben Sie Anmerkungen zum Stil der heutigen Veranstaltung?

+ -

9) Bitte beurteilen Sie erneut folgende Aussagen:

	trifft voll und ganz zu		trifft überhaupt nicht zu		
a) In wenigen Jahren fahren autonome Fahrzeuge ohne Sicherheitsfahrer/in auf deutschen Straßen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Ich fühle mich als Verkehrsteilnehmer/in auf dem Testfeld Autonomes Fahren in Karlsruhe sicher.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Das Thema autonomes Fahren wird durch die Medien stark problematisiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Die Datensicherheit ist beim autonomen Fahren gewährleistet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Die Steuerungskomplexität beim autonomen Fahren stellt kein Sicherheitsproblem dar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Ich fühle mich durch die mediale Berichterstattung ausreichend über den Stand der Forschung zum Thema autonomes Fahren informiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Das Thema autonomes Fahren wird durch die Medien stark positiv dargestellt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

III.8.2. Anhang 2) Fragen der Leitfadeninterviews

Fragenblock 1

Wie schätzen Sie die Veranstaltung ein? (Posttest, Frage 2)

Was hat Ihnen besonders gut gefallen?

Was hat Ihnen gar nicht gefallen?

Was haben Sie von dieser Veranstaltung erwartet?

Was haben Sie vermisst?

Fragenblock 2

Sie haben im Fragebogen angegeben, dass Sie **bereits** an einer öffentlichen Diskussion mit Bürgerbeteiligung teilgenommen haben. (Pretest, Frage 5) Welches Format hatte diese Diskussion?

Sie haben im Fragebogen angegeben, dass Sie **noch nie** an einer öffentlichen Diskussion mit Bürgerbeteiligung teilgenommen haben. (Pretest, Frage 5) In welcher Form haben Sie sich die Bürgerbeteiligung vorgestellt?

Fragenblock 3

Haben Sie sich an der Diskussion beteiligt?

Wenn **nein**, warum haben Sie sich nicht beteiligt?

Wenn **ja**, wie haben Sie sich dabei gefühlt?

Hat es Sie Überwindung gekostet, aufzustehen und sich in den Experten/innenkreis zu setzen?

Fragenblock 4

Sie haben sich eingestuft als Experte/in beziehungsweise Laie/in Stufe X. (Pretest, Fragen 1, 2 und 9)

Warum?

Wie viel Vorwissen haben Sie zum Thema autonomes Fahren?

Fragenblock 5

Ihre Haltung zum Thema autonomes Fahren hat sich geändert. (Vergleich zwischen Pretest, Frage 10 und Posttest, Frage 9)

Warum beziehungsweise durch was wurde das ausgelöst?

IV. Evaluation der RIKI-Fishbowl-Debatte zum Thema „Künstliche Intelligenz in der Medizin“

IV.1. Einleitung

Das Nationale Institut für Wissenschaftskommunikation (NaWik) richtete am 26. Oktober 2019 im Spreepalais in Berlin eine Fishbowl-Diskussion⁵ mit dem Titel: „Künstliche Intelligenz in der Medizin – Reden Sie mit!“ aus. Veranstaltet wurde die Diskussion als Teil des Projekts „Risikokommunikation zur Künstlichen Intelligenz“ (RIKI) zusammen mit Experten und Expertinnen der Fraunhofer Gesellschaft und der Charité.⁶

An der Diskussion waren folgende Personen als Expert*innen beteiligt: Dr. med. Alexander Meyer (BIH-Charité Berlin, Deutsches Herzzentrum Berlin), Dr. Andrea Schenk (Fraunhofer MEVIS) und Prof. Dr. Thomas Berlage (Fraunhofer FIT). Moderiert wurde die Diskussion von Frau Nicola Kuhrt. Während der Veranstaltung waren von den 60 bereitgestellten Sitzplätzen ungefähr 30 von Zuschauer*innen besetzt.

Zur Evaluation der Diskussion wurde sowohl vor als auch nach der Veranstaltung eine schriftliche Befragung unter den Teilnehmer*innen durchgeführt. Der Fragebogen ist in Anhang 1 (Kapitel IV.8.1) beigefügt. Insgesamt wurden 28 Fragebögen ausgefüllt.

Nach der Veranstaltung wurden leitfadengestützte Interviews ergänzend zu den Fragebögen durchgeführt. Es konnten sieben Teilnehmer*innen befragt werden. Insgesamt wurden 14 Fragen in fünf Frageblöcken (siehe Anhang 2 in Kapitel IV.8.2) gestellt. Daraus konnten aus Zeitgründen (je ca. fünf Minuten pro Befragter/Befragtem) zwischen sechs und elf Fragen beantwortet werden. Die Fragen der Interviews lehnten sich thematisch an die Fragen der Fragebögen an und dienten der inhaltlichen Vertiefung.

IV.2. Soziodemografie der Teilnehmer*innen der Veranstaltung

IV.2.1. Geschlecht

Von den Besucher*innen beantworteten 16 (57%) die Frage nach ihrem Geschlecht mit „männlich“, elf (39%) mit „weiblich“ (Abb. IV.1).⁷

⁵ Zur Fishbowl-Debatte als Veranstaltung vergleiche Kapitel III.1.

⁶ Übersicht zu der Veranstaltung, <https://www.nawik.de/riki/fishbowl-session-ki-und-medizin/>, Aufgerufen 04.02.2020.

⁷ Die Antwort „Keine Angabe“ bedeutet immer eine nicht ausgefüllte Frage. Dabei konnte nicht bestimmt werden, ob dies bewusst geschah oder ob die Teilnehmer*innen die Frage übersehen haben.

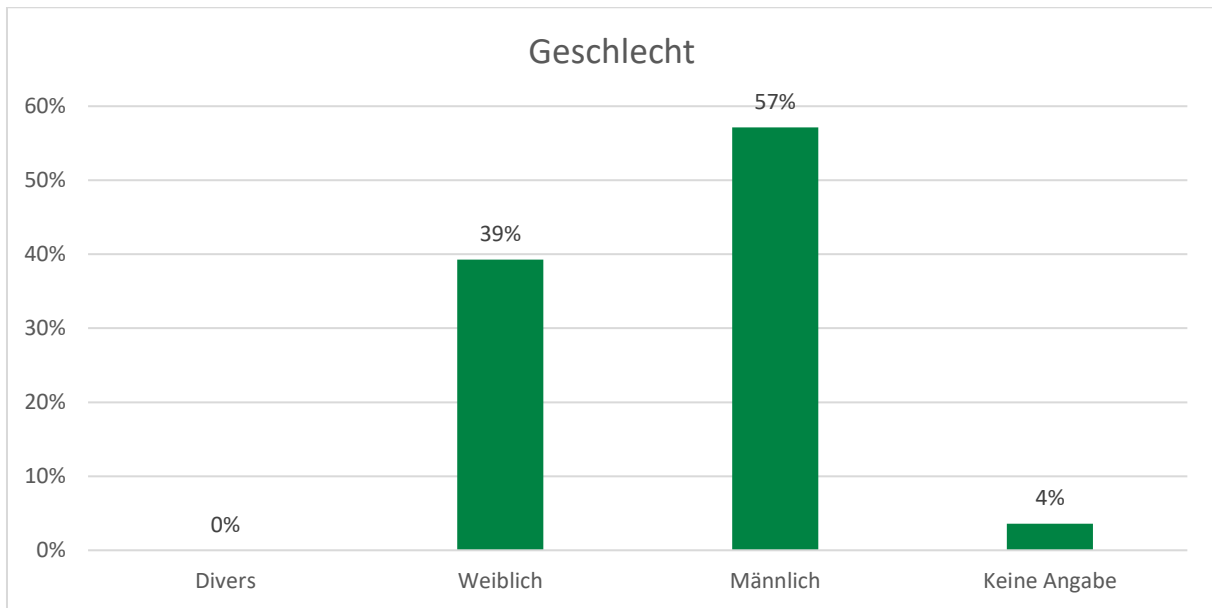


Abbildung IV.1: Prozentuale Antworten auf die Frage: „Ihr Geschlecht“. N=28

IV.2.2. Alter

Die Hälfte der Befragten war zwischen 25 und 34 Jahre alt (50%, n=14). Die zweithäufigste Altersgruppe bildeten die unter 25-jährigen mit 18% (n=5), die restlichen Teilnehmer*innen verteilten sich auf die anderen Altersgruppen. Es waren keine Besucher*innen in der Altersgruppe zwischen 35 und 44 anwesend (Abb. IV.2).

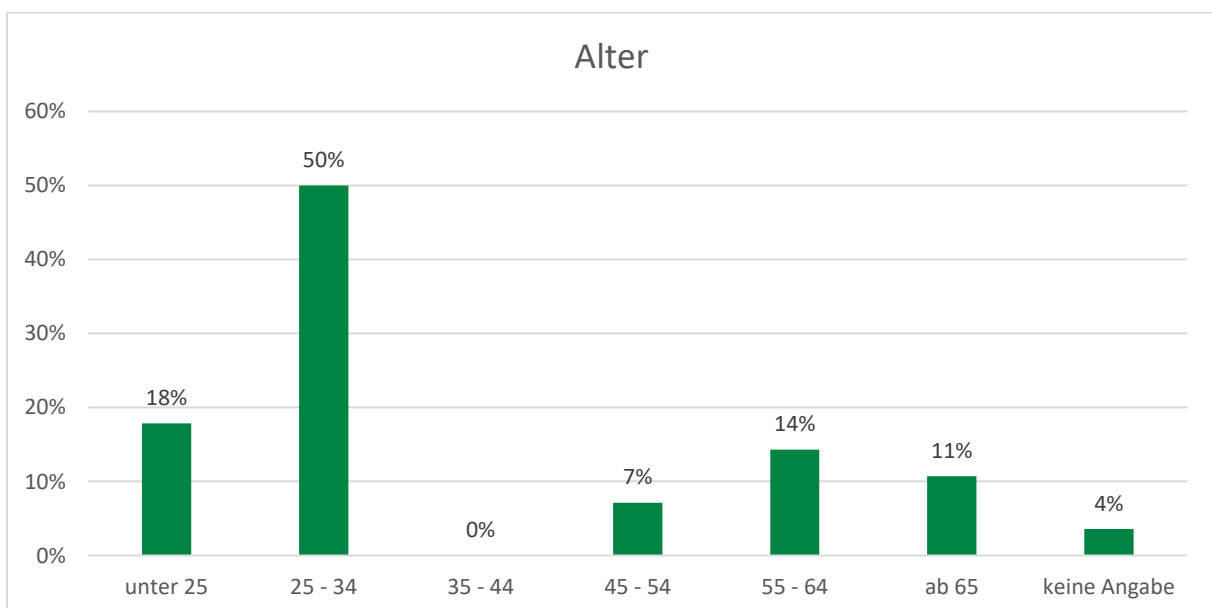


Abbildung IV.2: Prozentuale Antworten auf die Frage nach dem Alter der Besucher*innen. N=28

IV.2.3. Höchster formaler Bildungsabschluss

Eine knappe Mehrheit von 57% (n=16) aller Befragten hatte einen akademischen Bildungsabschluss: 46% (n=13) besaßen einen Hochschulabschluss, 11% (n=3) der Befragten eine Promotion (Abb. IV.3). Acht Befragte (28%) gaben die (Fach-)Hochschulreife als höchsten Bildungsabschluss an.

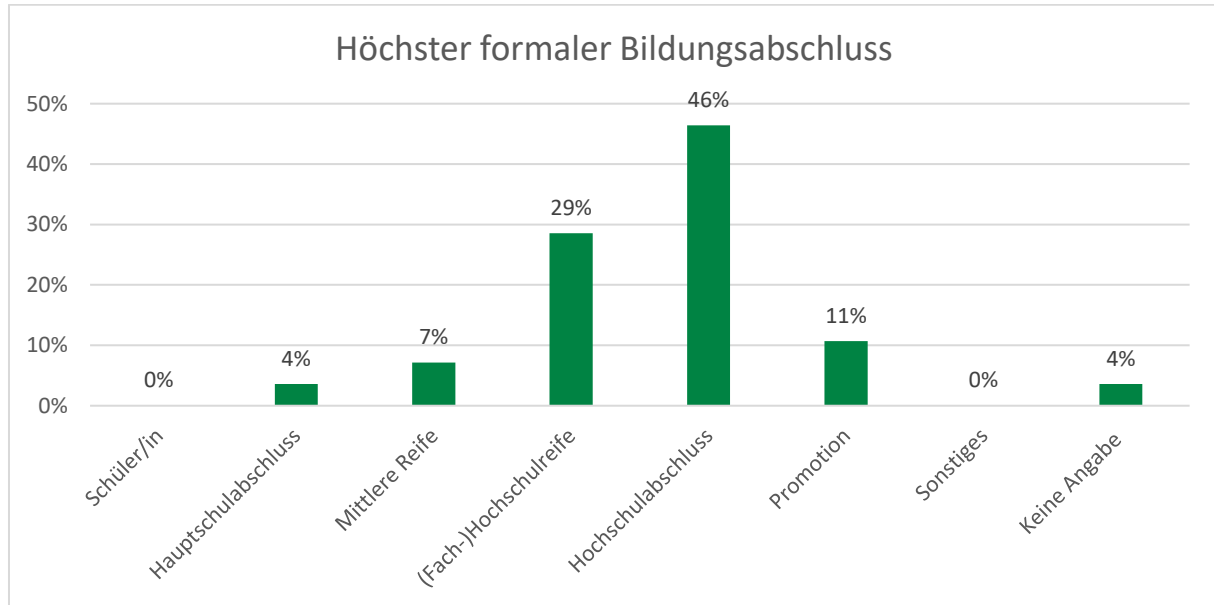


Abbildung IV.3: Prozentuale Antworten auf die Frage: „Was ist Ihr höchster Bildungsabschluss?“. N=28

IV.2.4. Berufliche Tätigkeit

Knapp die Hälfte der Teilnehmer*innen kam aus dem medizinischen Bereich (43%, n=12), gefolgt vom wissenschaftlichen Bereich (18%, n=5) (Abb. IV.4). Fast ein Drittel der Besucher*innen (29%, n=8) übte einen „sonstigen“ Beruf aus, beispielsweise als Journalist*in oder Künstler*in.

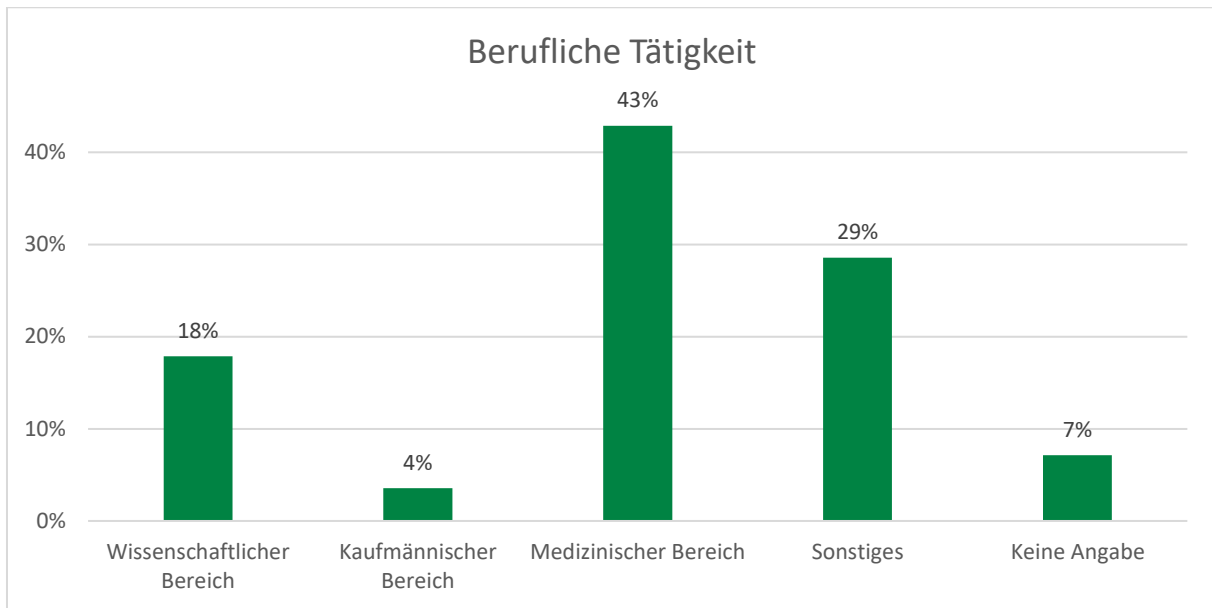


Abbildung IV.4: Prozentuale Antworten auf die Frage: „Was ist Ihre berufliche Tätigkeit?“. N=28

IV.3. Formatbezogene Erfahrungen und Erwartungen der Teilnehmer*innen

IV.3.1. Teilnahme an Veranstaltungen zu einem wissenschaftlichen Thema

Die Antworten aus den Fragebögen zeigen, dass 75% der Befragten (n=21) schon mindestens einmal an einer öffentlichen Veranstaltung zu wissenschaftlichen Themen teilgenommen haben, 61% aller Besucher*innen (n=17) sogar mehrmals. Für knapp ein Fünftel der Befragten (21%, n=6) war es die erste öffentliche Veranstaltung zu einem wissenschaftlichen Thema (Abb. IV.5).

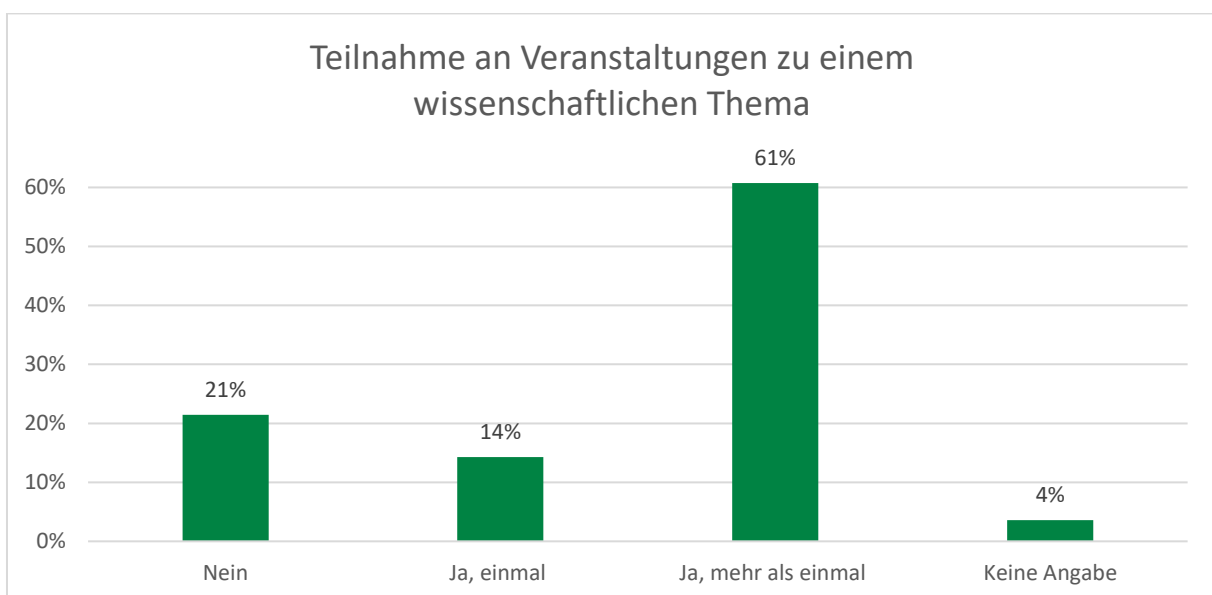


Abbildung IV.5: Prozentuale Antworten auf die Frage: „Haben Sie bereits an einer öffentlichen Veranstaltung zu einem wissenschaftlichen Thema teilgenommen?“. N=28

IV.3.2. Teilnahme an öffentlichen Diskussionen mit Bürgerbeteiligung

Für die Hälfte der Befragten (50%, n=14) war die Fishbowl-Debatte die erste öffentliche Diskussion mit Bürgerbeteiligung. 29% (n=8) haben schon mehrmals an solchen Diskussionen teilgenommen, 18% (n=5) erst einmal (Abb. IV.6).

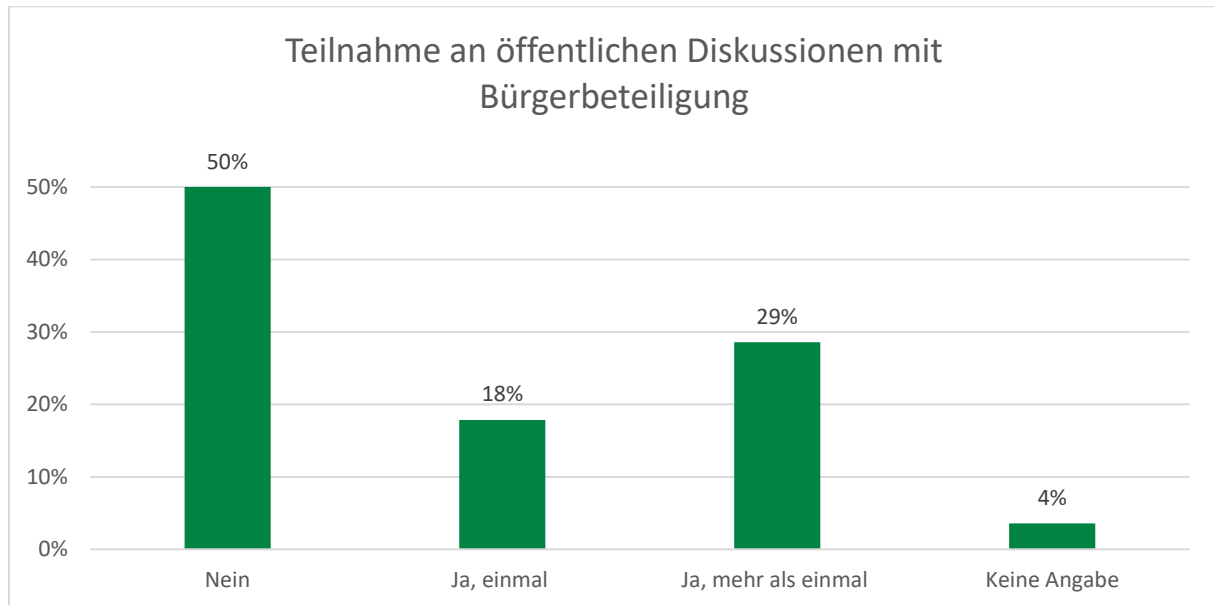


Abbildung IV.6: Prozentuale Antworten auf die Frage: „Haben Sie bereits an einer öffentlichen Diskussion mit Bürgerbeteiligung teilgenommen?“. N=28

Alle in den zusätzlichen Interviews befragten Besucher*innen gaben an, dass sie zum ersten Mal an einer Veranstaltung dieses Formats teilgenommen haben.

IV.3.3. Wichtigster Grund für den Besuch der Veranstaltung

Die meisten Befragten (70%, n=14) kamen zu der Veranstaltung, weil sie mehr über Künstliche Intelligenz in der Medizin wissen wollten. Zehn Prozent der Besucher*innen (n=2) kamen aus einem generellen Interesse an wissenschaftlichen Themen. Ebenso viele Teilnehmer*innen kamen, um sich an dem Meinungs austausch zum Thema zu beteiligen (Abb. IV.7). Einige Besucher*innen gaben, trotz Aufforderung nur eine Antwort abzugeben, mehrere Gründe an. Diese konnten nicht mit in die Berechnung aufgenommen werden.

Von den für die Leit fadeninterviews befragten Teilnehmer*innen gaben alle als Grund für den Besuch an, ihr Wissen über das Thema vertiefen zu wollen. Keiner gab als Grund das Format der Veranstaltung an.

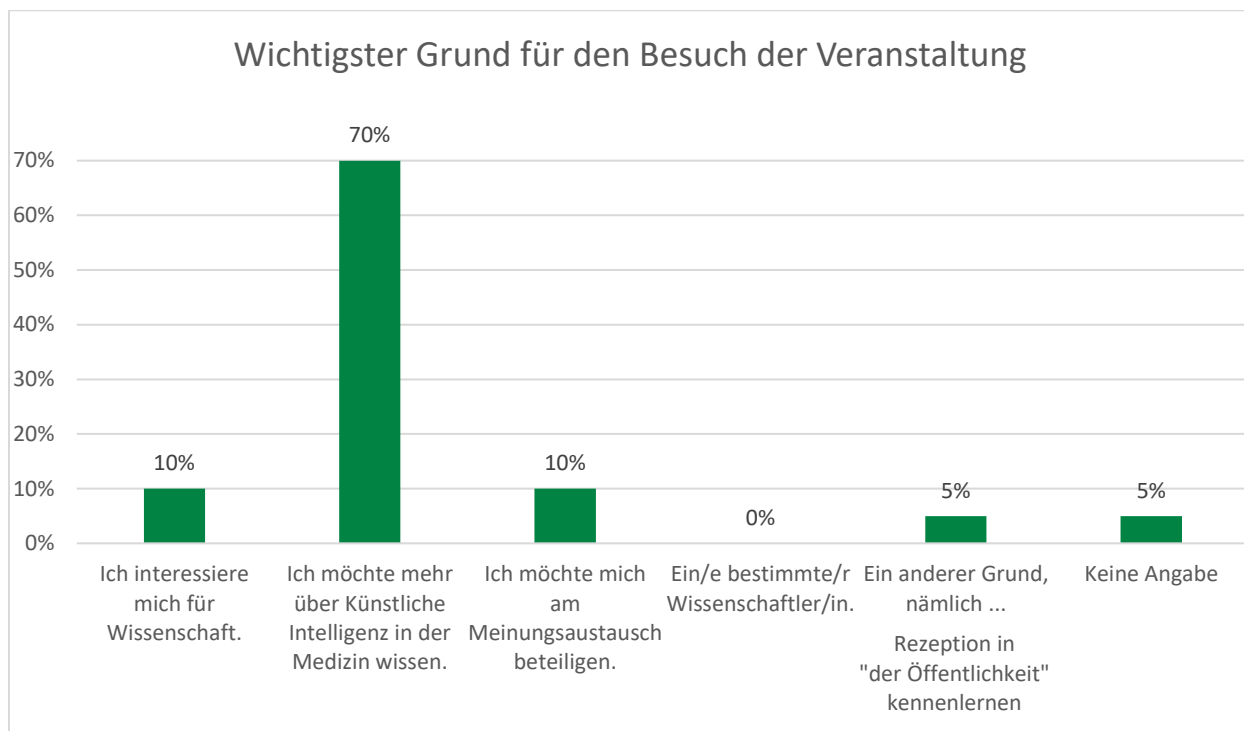


Abbildung IV.7: Prozentuale Antworten auf die Frage: „Was war für Sie der wichtigste Grund zur heutigen Veranstaltung zu kommen?“. N=20

IV.4. Bezug, Vorkenntnisse und Einschätzungen der Teilnehmer*innen zum Thema Künstliche Intelligenz in der Medizin

IV.4.1. Beruflicher oder privater Bezug zum Thema Künstliche Intelligenz in der Medizin

Die meisten der Besucher*innen (82%, n=23) hatten einen beruflichen oder privaten Bezug zum Thema Künstliche Intelligenz in der Medizin. Nur 14% (n=4) der Befragten gaben an, keinen solchen Bezug zu haben (Abb. IV.8).

Dabei teilt sich der berufliche und private Bezug zum Thema in etwa gleiche Anteile auf. Zwölf Personen (43%) gaben einen Bezug über Studium oder ihre Arbeit in der Forschung an, elf (39%) einen Bezug über sonstige Berufe (z. B. „Krankenpfleger“) zum Thema an (Abb. IV.8).

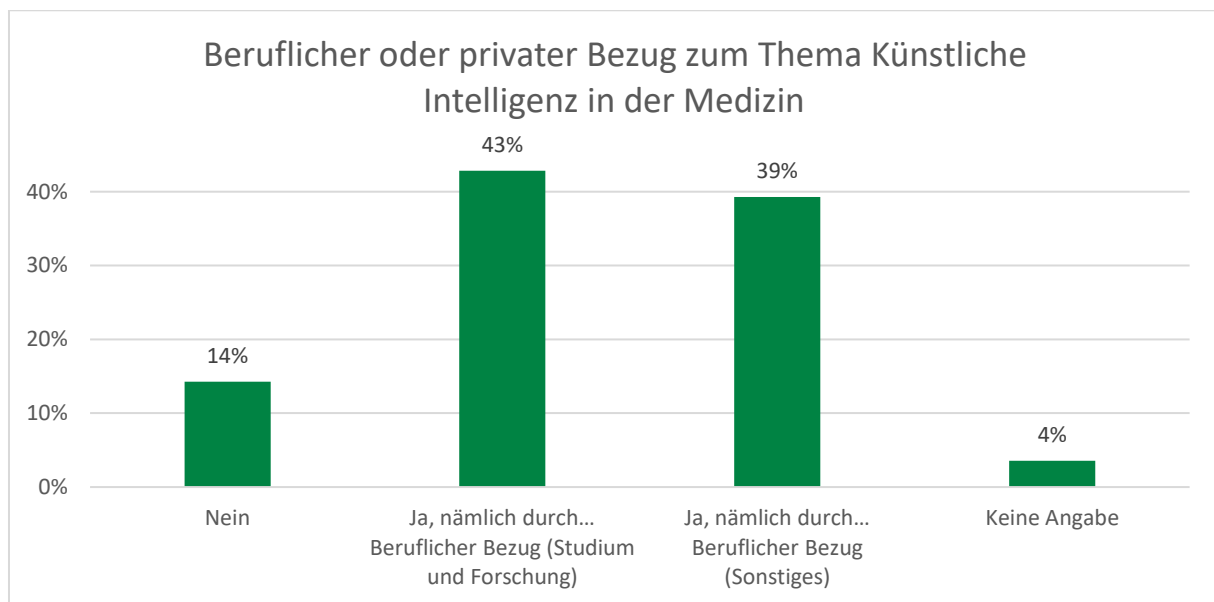


Abbildung IV.8: Prozentuale Antworten auf die Frage: „Haben Sie einen beruflichen oder privaten Bezug zum Thema Künstliche Intelligenz in der Medizin?“. N=28

IV.4.2. Vorkenntnisse zum Thema Künstliche Intelligenz in der Medizin

Über die Hälfte der Teilnehmer*innen gab an, zumindest grundlegend informiert zu sein. Etwa ein Drittel der Teilnehmer*innen (32%, n=9) gab an, „weder Experte/in noch Laie/in“ zu sein, acht, dass sie „eher Laie*in“ seien (29%). 25% (n=7) der Besucher*innen gaben an, „eher Experte/in“ zu sein. Zwei Personen (7%) schätzten sich als „Experte/in“ ein. Eine Person sah sich als „Laie/in“ (4%) (Abb. IV.9).

Keine der interviewten Personen machte Angaben zur Menge ihres Vorwissens.

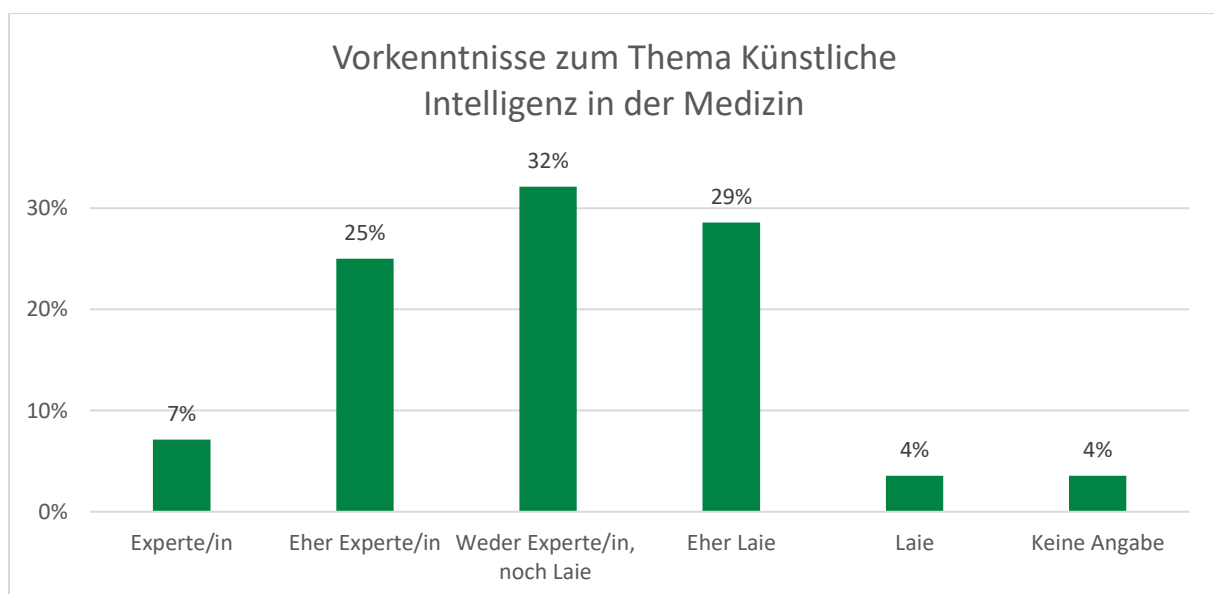


Abbildung IV.9: Prozentuale Antworten auf die Frage: „Wie schätzen Sie selbst Ihre Kenntnisse zum Thema Künstliche Intelligenz in der Medizin ein? N=28

IV.4.3. Inhaltliche Einschätzungen zum Thema Künstliche Intelligenz in der Medizin

Im Vorfeld der Fishbowl-Debatte sollten die Teilnehmer*innen die folgenden Aussagen von „trifft überhaupt nicht zu“ in fünf Stufen bis „trifft voll und ganz zu“ bewerten:

- In wenigen Jahren wird Künstliche Intelligenz Ärzte z. B. im OP-Saal oder bei der Diagnose ersetzen.
- Ich halte die Anwendung von Künstlicher Intelligenz im OP-Saal für ethisch bedenklich.
- Ich halte die Anwendung von Künstlicher Intelligenz in der Diagnostik für ethisch bedenklich.
- Das Thema Künstliche Intelligenz in der Medizin wird durch die Medien stark problematisiert.
- Die Datensicherheit ist beim Umgang mit Patientendaten gewährleistet.
- Ich fühle mich durch die mediale Berichterstattung ausreichend über den Stand der Forschung zu Künstlicher Intelligenz in der Medizin informiert.
- Das Thema Künstliche Intelligenz in der Medizin wird durch die Medien stark positiv dargestellt.

Die Befragung zeigt, dass die Teilnehmer*innen die Aussagen eher neutral sahen oder meinten, dass sie nicht zutreffend seien (Abb. IV.10).

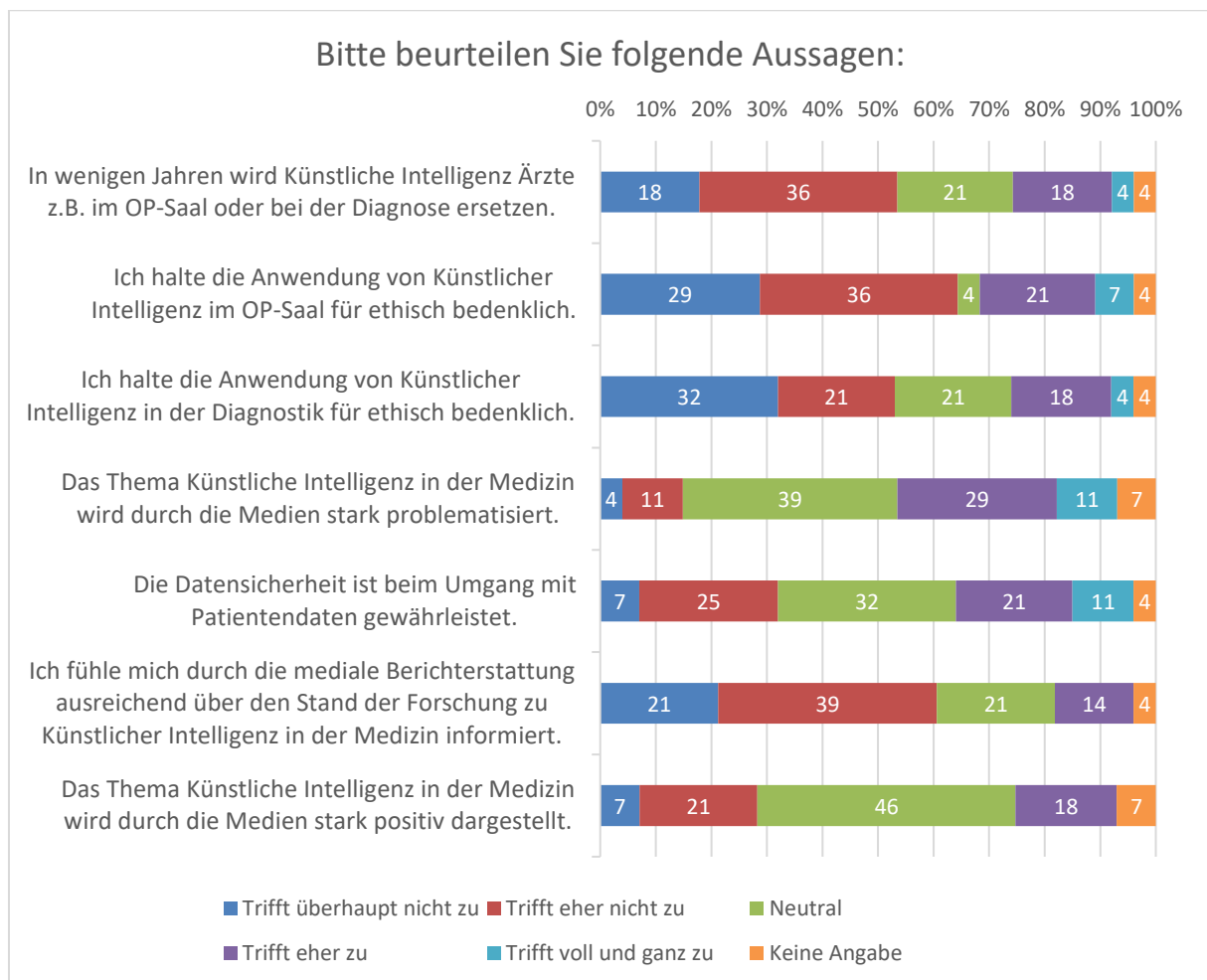


Abbildung IV.10: Prozentuale Antworten zur Fragebatterie „Bitte beurteilen Sie die folgenden Aussagen“. Die Befragten sollten sieben inhaltliche Aussagen bewerten. N=28

IV.5. Grundlegende Ex-Post-Einschätzungen der Teilnehmer*innen

IV.5.1. Gesamteindruck der Diskussion

Die Besucher*innen sollten im Anschluss an die Veranstaltung angeben, wie ihr Gesamteindruck der Diskussion war (Abb. IV.11). Dazu wurden „Schulnoten“ von „sehr gut“ bis „mangelhaft“ vergeben.

Die Fishbowl-Diskussion wurde von 43% (n=12) der Besucher*innen als „gut“ bewertet, fünf (18%) bewerteten sie als „sehr gut“. Die Note „befriedigend“ wurde von vier Personen (14%) vergeben, niemand empfand sie als „ausreichend“ oder „mangelhaft“. Sieben Teilnehmer*innen (25%) gaben keine Wertung ab.

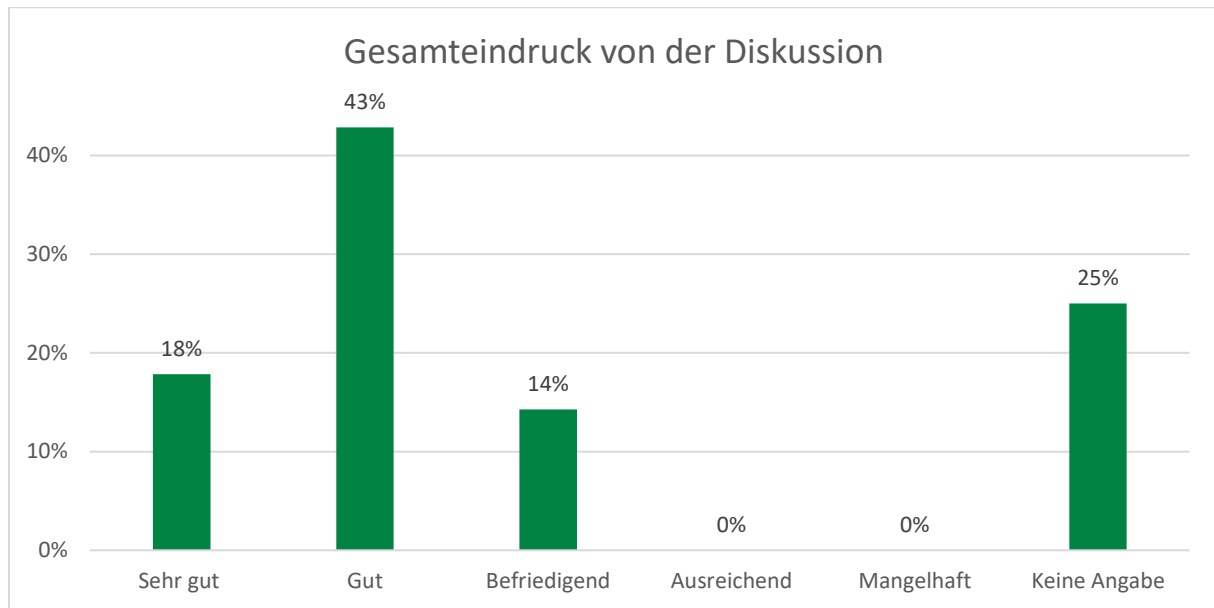


Abbildung IV.11: Prozentuale Antworten auf die Frage: „Was ist Ihr Gesamteindruck von der heutigen Diskussion?“. N=28

IV.5.2. Inhalt der Diskussion

Die Befragten wurden darum gebeten, kurz wiederzugeben, worum es in der Diskussion ging (Abb. IV.12). Die Antworten wurden für die Auswertung thematisch geordnet und zusammengefasst.

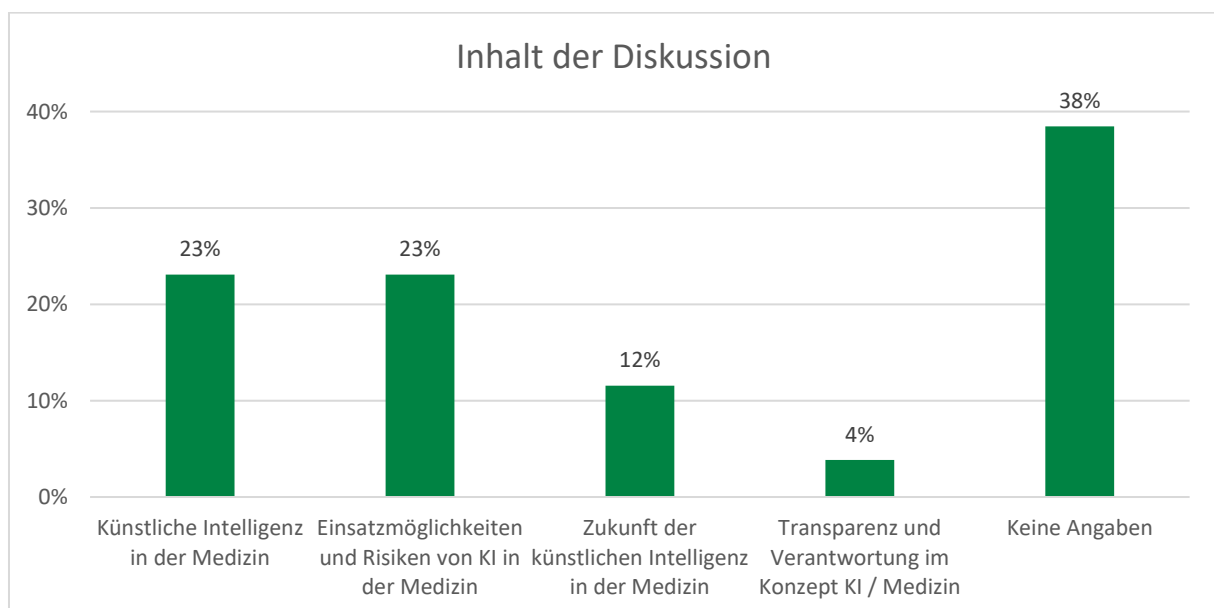


Abbildung IV.12: Prozentuale Antworten auf die Frage „In einem Satz zusammengefasst: Worum ging es in der heutigen Diskussion?“. Zur Auswertung wurden die Antworten in Themengruppen zusammengefasst. Mehrfachnennungen waren möglich. N=26

23% (n=6) der Befragten gaben als Antwort eine Aussage ähnlich des Titels der Veranstaltung „Künstliche Intelligenz in der Medizin“. Ebenso viele Besucher*innen gaben eine Antwort aus dem Bereich „Einsatzmöglichkeiten und Risiken von KI in der Medizin“ an. Über ein Drittel (38% n=10) machte keine Angaben zum Thema.

IV.5.3. Erwerb neuen Wissens zum Thema Künstliche Intelligenz in der Medizin

Die Teilnehmer*innen sollten die Aussage „Ich habe aus dieser Veranstaltung viel neues Wissen über Künstliche Intelligenz in der Medizin mitgenommen.“ in fünf Stufen von „trifft überhaupt nicht zu“ bis „trifft voll und ganz zu“ beantworten. Für 39% (n=11) der Befragten traf die Aussage „eher zu“. Jeweils vier Teilnehmer*innen (14%) bewerteten sie mit „neutral“ und „trifft eher nicht zu“. Zwei Personen stimmten „voll und ganz“ zu, niemand gab „trifft überhaupt nicht zu“ an (Abb. IV.13).

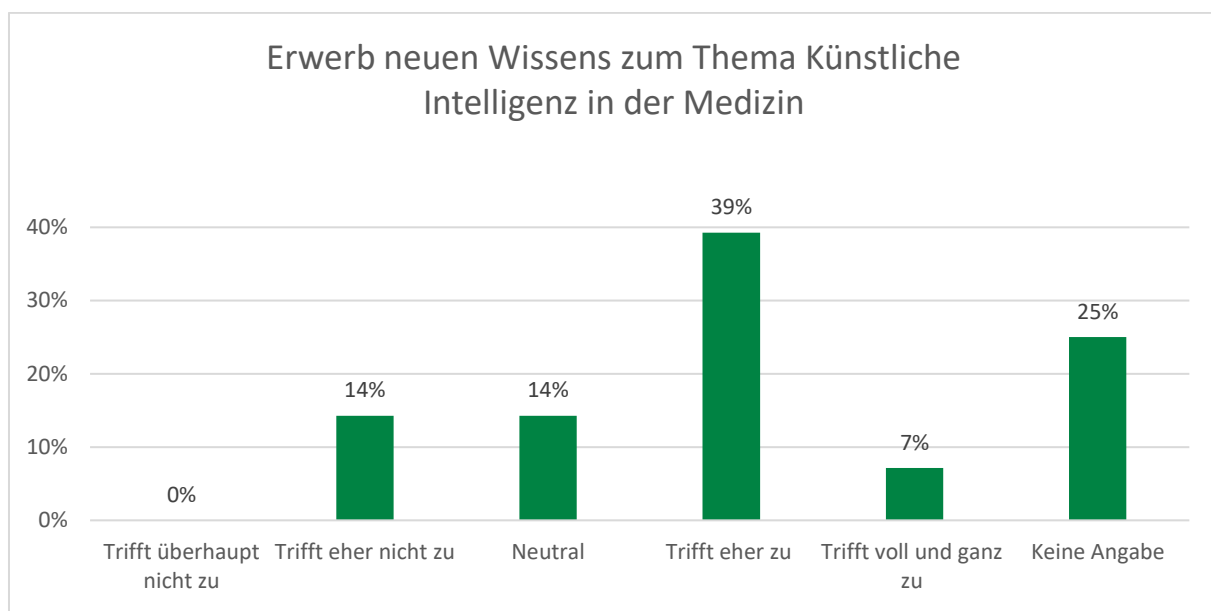


Abbildung IV.13: Prozentuale Antworten auf die Frage: „Bitte beurteilen Sie folgende Aussage: Ich habe aus dieser Veranstaltung viel neues Wissen über Künstliche Intelligenz in der Medizin mitgenommen.“. N=28

IV.5.4. Inhaltliche Einschätzungen zum Thema Künstliche Intelligenz in der Medizin

Die Besucher*innen sollten nach der Veranstaltung erneut eine Einschätzung zu den bereits im Vorfeld abgefragten inhaltlichen Aussagen zum Thema Künstliche Intelligenz in der Medizin geben (vergleiche Abschnitt 4.3). Die Bewertung erfolgt erneut in fünf Stufen von „trifft überhaupt nicht zu“ bis „trifft voll und ganz zu“ (Abb. IV.14).

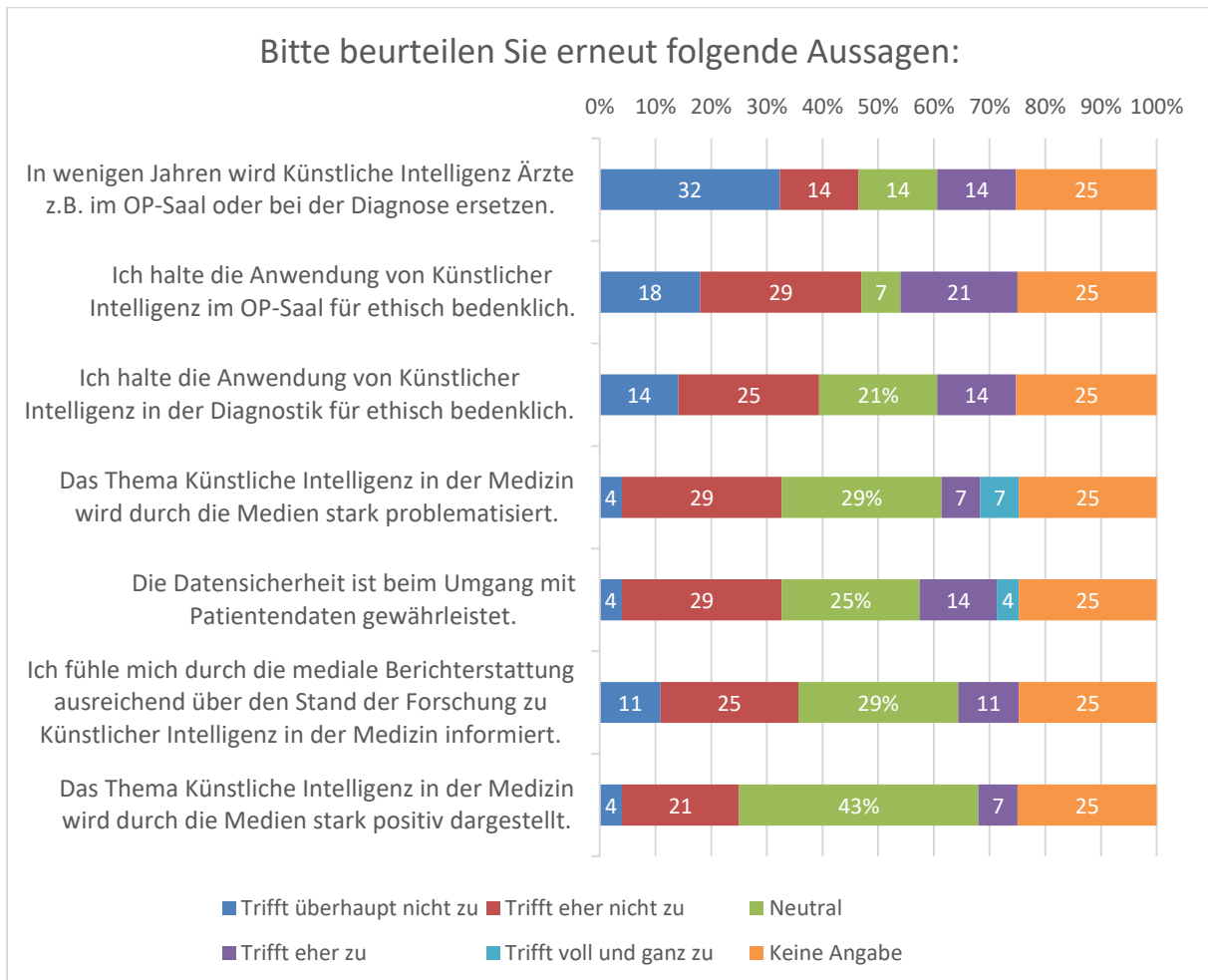


Abbildung IV.14: Prozentuale Antworten Fragebatterie „Bitte beurteilen Sie erneut folgende Aussagen“. Erneut wurden sieben inhaltliche Aussagen bewertet. N=28

Bei allen Aussagen lässt sich im Vergleich zu den Aussagen vor der Debatte eine Verschiebung feststellen (Abb. IV.15).

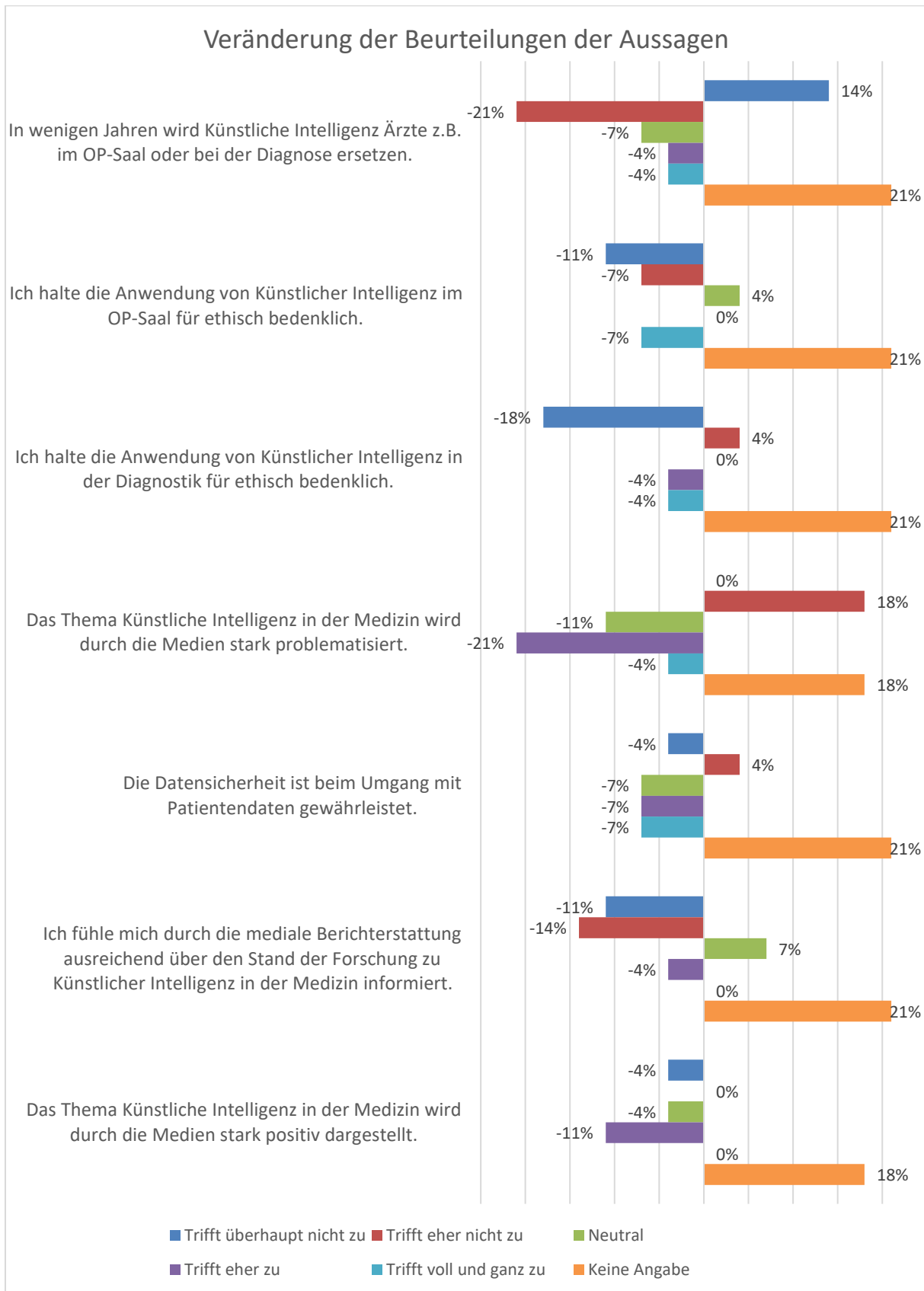


Abbildung IV.15: Veränderungen der Beurteilungen der Aussagen zum Thema Künstliche Intelligenz in der Medizin (Pretest vs. Posttest). N=28

Am auffälligsten ist der starke Anstieg der nicht gemachten Angaben („keine Angabe“) bei allen Fragen. Da dies uniform geschah ist anzunehmen, dass es immer die gleichen Personen waren, die keine Angaben gemacht haben. Dies kann verschiedene Gründe haben. Entweder haben es sich diese Personen nicht zugetraut oder sich schwer damit getan, eine Aussage zu treffen, oder sie haben die Frage mit Absicht übersprungen, etwa aus Zeitnot.

Abgesehen davon zeigen die Antworten, dass die Debatte die Meinung der Besucher*innen zu den Aussagen änderte. Besonders die Frage *„In wenigen Jahren wird Künstliche Intelligenz Ärzte z. B. im OP-Saal oder bei der Diagnose ersetzen“* zeigt eine Verschiebung von *„trifft eher nicht zu“* (-21 Prozentpunkte) zu *„trifft überhaupt nicht zu“* (+14 Prozentpunkte). Bei der Aussage *„ich halte die Anwendung von Künstlicher Intelligenz in der Diagnostik für ethisch bedenklich“* fiel die Einschätzung *„trifft überhaupt nicht zu“* um 18 Prozentpunkte. Bei der Aussage *„Das Thema Künstliche Intelligenz in der Medizin wird durch die Medien stark problematisiert“* fielen die Einschätzungen *„neutral“* und *„trifft eher zu“* um elf bzw. 21 Prozentpunkte, wohingegen die Aussage *„trifft eher nicht zu“* um 18 Prozentpunkte stieg.

Zur Einschätzung der medialen Vermittlungsleistung im Themenbereich KI in der Medizin ist die Aussage *„Ich fühle mich durch die mediale Berichterstattung ausreichend über den Stand der Forschung zu Künstlicher Intelligenz in der Medizin informiert“* von besonderem Interesse. Hier fiel die Einschätzung *„trifft eher nicht zu“* um 14 Prozentpunkte und die Aussage *„trifft überhaupt nicht zu“* um elf Prozentpunkte, während die Einschätzung *„neutral“* um sieben Prozentpunkte stieg.

In den leitfadengestützten Interviews wurden die Teilnehmer*innen gefragt, welche neuen Denkanstöße sie durch die Veranstaltung erhalten hatten. Folgende Themenbereiche wurden als Antworten genannt:

- Moral/ Ethik im Bereich der Diagnostik
- Datensicherheit/ Datenfreigabe bei Patientendaten

IV.5.5. Erneute Teilnahme an einer öffentlichen Diskussion zum Thema Künstliche Intelligenz in der Medizin

71% der Befragten (n=20) gaben an, dass sie erneut eine Veranstaltung zum Thema Künstliche Intelligenz in der Medizin besuchen würden. Eine Person (4%) gab an, dass sie keine weitere Veranstaltung zu dem Thema besuchen würde, allerdings ohne eine Nennung eines Grundes (Abb. IV.16).

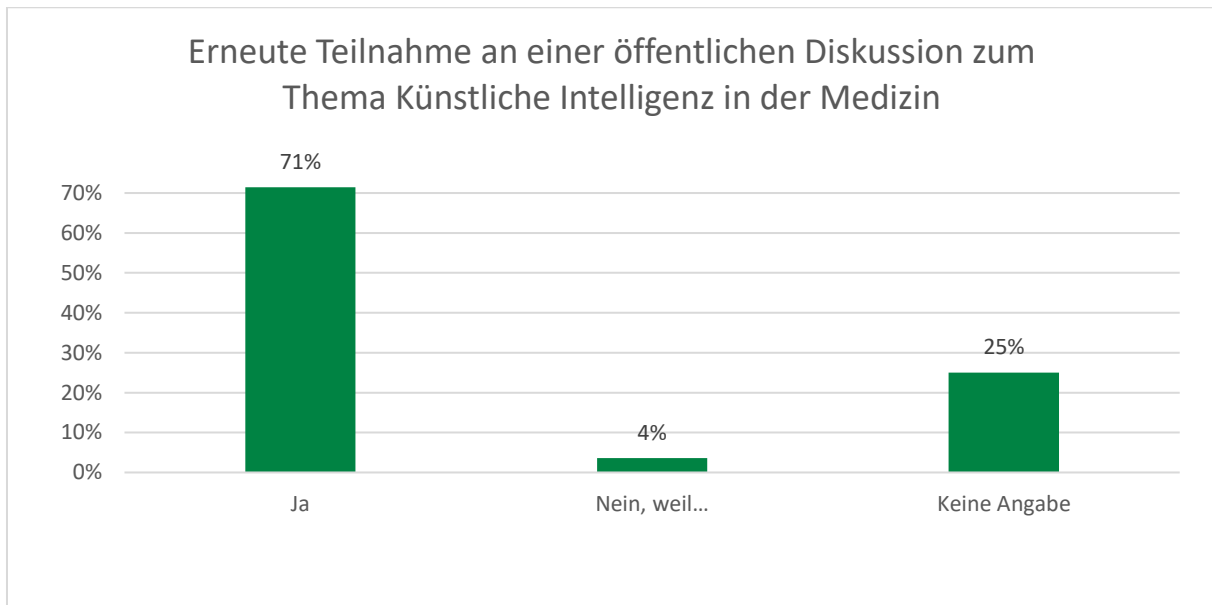


Abbildung IV.16: Prozentuale Antworten auf die Frage: „Würden Sie wieder an einer öffentlichen Diskussion zum Thema Künstliche Intelligenz in der Medizin teilnehmen?“. N=28

In den Interviews gaben die Befragten an, dass sie die Fishbowl-Debatte als interessante und lehrreiche neue Erfahrung ansahen und auch ein weiteres Mal teilnehmen würden.

IV.6. Formatbezogene Einschätzungen der Teilnehmer*innen

Neben der inhaltlichen Auswertung wurden die Teilnehmer*innen auch explizit um Einschätzungen zu Form und Stil der Diskussion gebeten.

IV.6.1. Zentrale Aspekte der Veranstaltung

Die Besucher*innen sollten die Debatte nach unterschiedlichen Aspekten bewerten (Abb. IV.17). Dazu sollten sie „Schulnoten“ von „sehr gut“ bis „mangelhaft“ vergeben. Die Aspekte waren:

- Einleitung und Vorstellung der Experten/innen
- Auswahl der Experten/innen
- Moderation und Diskussionsführung
- Beteiligung der Bürger/innen an der Diskussion
- Interaktion zwischen Experten/innen und Bürger/innen
- Vertrauenswürdigkeit der Experten/innen
- Unterhaltungswert der Veranstaltung
- Veranstaltungsgröße

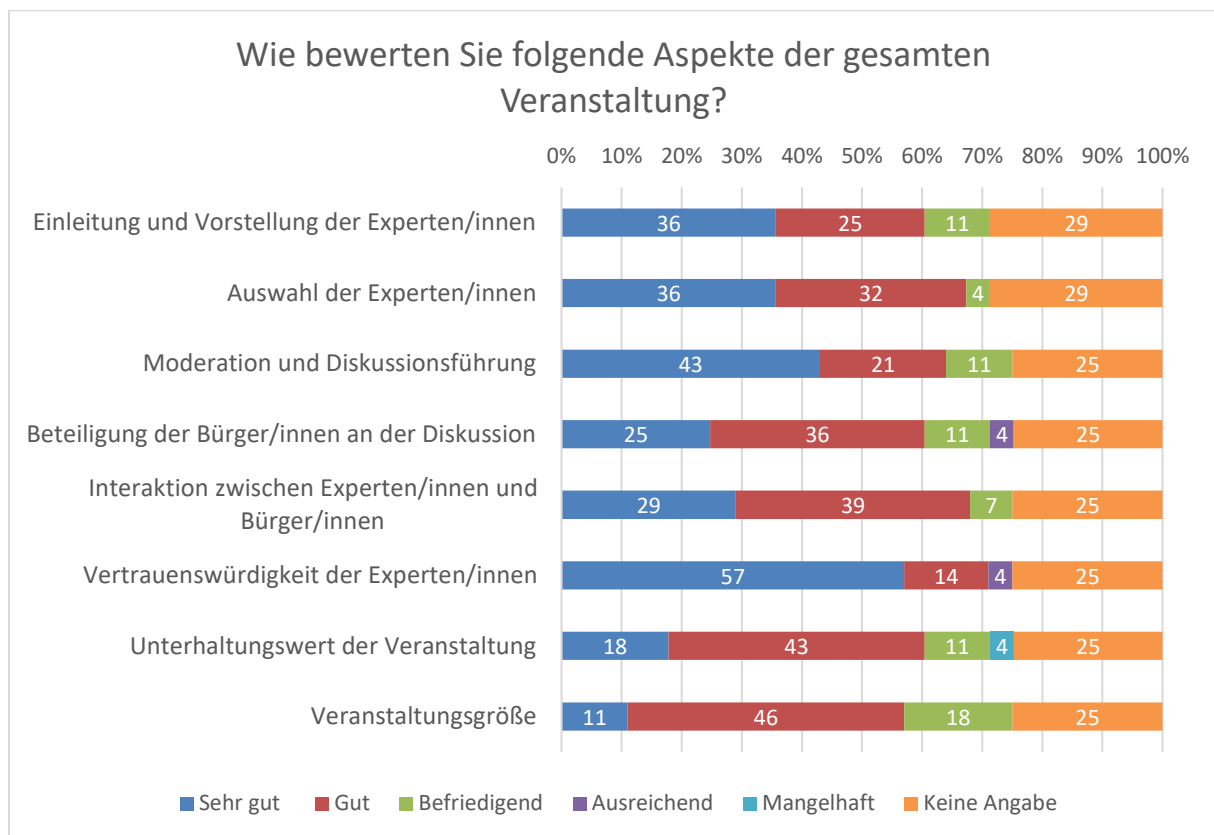


Abbildung IV.17: Prozentuale Antworten zur Fragebatterie „Wie bewerten Sie folgende Aspekte der gesamten Veranstaltung“. N=28

Insgesamt bewerteten die Besucher*innen die Veranstaltung als „sehr gut“ oder „gut“. Besonders häufig als „sehr gut“ bewertet wurde die „Vertrauenswürdigkeit der Experten/innen“ (57%, n=16). Die Note „mangelhaft“ wurde nur einmal (4%) für den „Unterhaltungswert der Veranstaltung“ vergeben.

IV.6.2. Diskussion mit den Expert*innen auf Augenhöhe

Die Frage „Hatten Sie das Gefühl, dass die Diskussion mit den Experten/innen auf Augenhöhe stattfand?“ wurde von 19 Personen bejaht (68%), eine Person (4%) verneinte die Frage ohne Angabe von Gründen (Abb. IV.18).

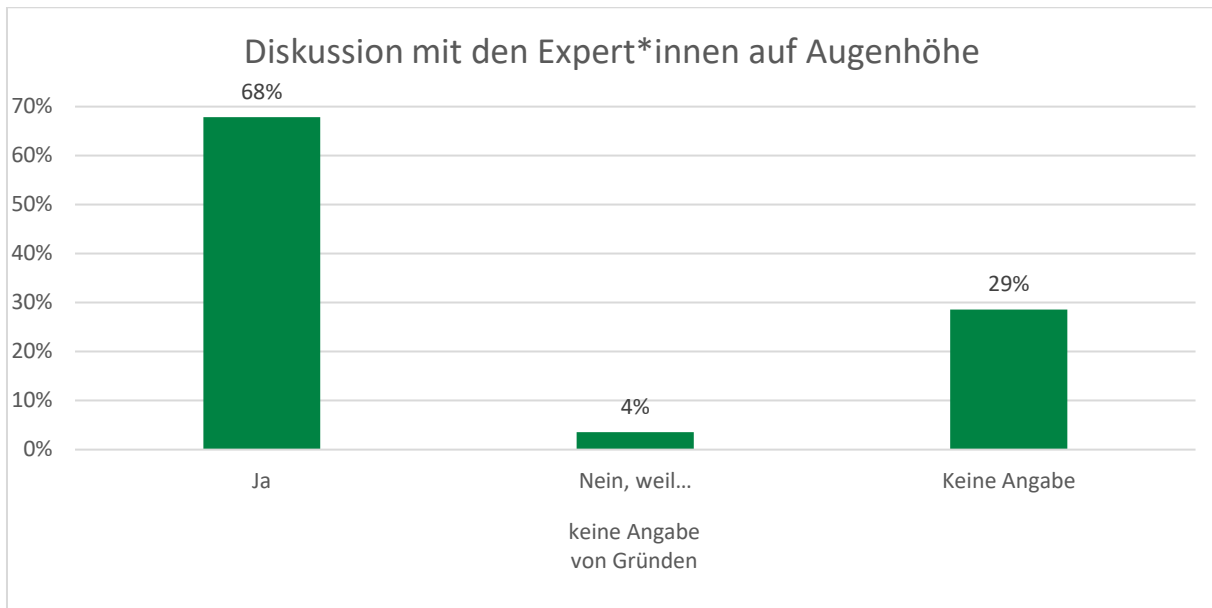


Abbildung IV.18: Prozentuale Antworten auf die Frage: „Hatten Sie das Gefühl, dass die Diskussion mit den Experten/innen auf Augenhöhe stattfand?“. N=28

Auch die leitfadengestützten Interviews bestätigten die Aussage, dass die Diskussion auf Augenhöhe mit den Expert*innen stattfand. Von einer befragten Person wurde die Videoaufzeichnung als mögliche Hemmschwelle für eine aktive Beteiligung an der Diskussion genannt. Eine weitere Person wies auf Nachfrage darauf hin, dass das Format an sich eventuell für Personen ohne Konferenzerfahrung mit einer gewissen Beteiligungs-Hemmschwelle verbunden sei. Konkret sprach jedoch nur eine Person davon, dass sie tatsächlich erst nach anfänglicher Überwindung an der Diskussion teilgenommen habe.

IV.6.3. Anmerkungen zum Stil der Veranstaltung

Die Teilnehmer*innen wurden gefragt, ob sie Anmerkungen zum Stil der Veranstaltung hätten. Für die Auswertung wurden die Antworten in positive und negative Bemerkungen zusammengefasst.

Die Mehrheit der Teilnehmer*innen (88%, n=22) hatte keine Anmerkungen zum Stil der Veranstaltung. Von den restlichen Befragten wurden mit je einer Anmerkung die Möglichkeit zur Beteiligung, die "stille Räumlichkeit", die kostenfreien Getränke, die Verständlichkeit sowie die Kompetenz der Expert*innen hervorgehoben.

Die einzige negative Anmerkung bat darum, bessere Mikrofone für die Expert*innen zu verwenden.

Weitere Vorschläge aus den Anmerkungen lauteten, Patientenvertreter*innen und Kostenträger*innen zu dem Format einzuladen, die Fragen im Voraus zu sammeln und

thematisch vor zu selektieren. Außerdem wünschte sich ein*e Teilnehmer*in "Tische zum Ablegen".

In den Leitfadeninterviews wurden zusätzlich einige Vorschläge zur Verbesserung gemacht. Darunter wurde genannt:

- Visuelle Präsentation zu Grundbegriffen der diskutierten Thematik
- Offensivere Kommunikation der geplanten thematischen Struktur der Diskussion an das Publikum
- Abfrage und Kommunikation der themenbezogenen Hintergründe der Personen im Publikum
- Reduktion der thematischen Sprünge während der Diskussion

IV.6.4. Erneute Teilnahme an einer öffentlichen Diskussion im Stil der Veranstaltung

Die Mehrheit der Befragten (75%, n=21) würde wieder an einer Fishbowl-Debatte teilnehmen, niemand antwortete mit „nein“ (Abb. IV.19).

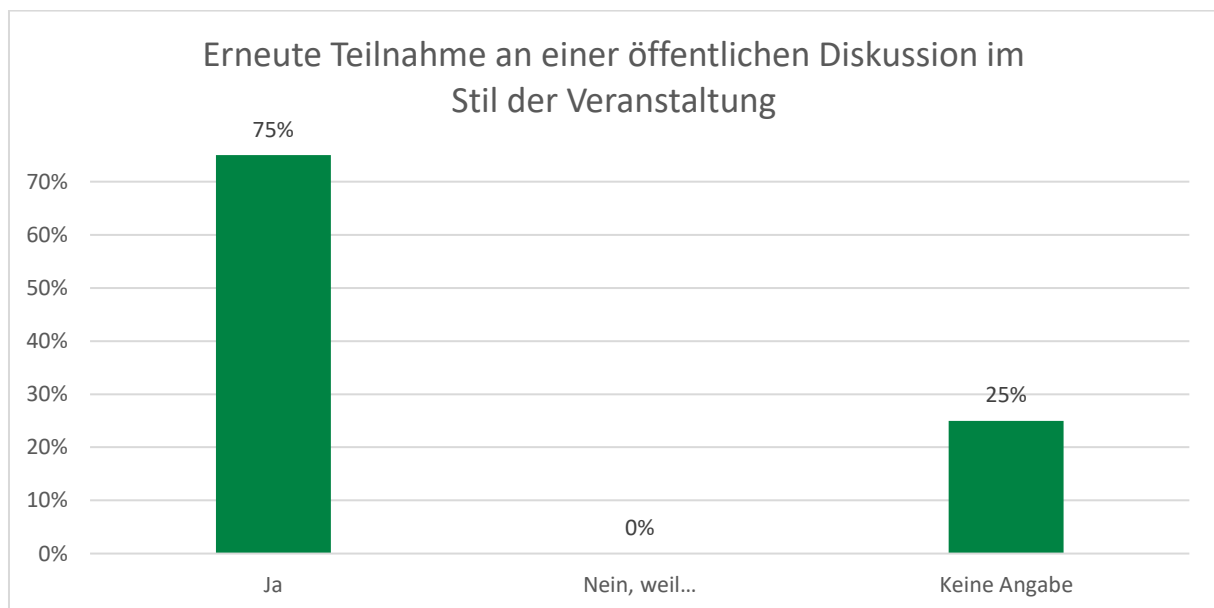


Abbildung IV.19: Prozentuale Antworten auf die Frage „Würden Sie wieder an einer öffentlichen Diskussion im Stil der heutigen Veranstaltung teilnehmen. N=28“

IV.7. Ergebnisse und Diskussion

Die Befragung zu Geschlecht, persönlichem Bildungshintergrund, Alter und zu den Vorerfahrungen der Besucher*innen zeigt, dass die Mehrheit der Teilnehmer*innen männlich war und aus dem Hochschulumfeld (durch Arbeit oder Studium) kam. Die größte einzelne

Altersgruppe mit vierzehn Personen lag zwischen 25 und 35 Jahren. Die meisten Besucher*innen der Debatte hatten bereits an einer öffentlichen Veranstaltung zu wissenschaftlichen Themen teilgenommen, jedoch war es für die Hälfte die erste öffentliche Veranstaltung mit Bürgerbeteiligung.

Insgesamt hatten die meisten der Teilnehmer*innen einen persönlichen oder beruflichen Bezug zu dem Thema Künstliche Intelligenz in der Medizin. Dennoch gaben die meisten Besucher*innen an, die Veranstaltung mit dem Wunsch besucht zu haben, mehr über Künstliche Intelligenz in der Medizin zu lernen. Der Großteil der Befragten gab an, ein Wissen zumindest auf der Stufe „weder Experte/in noch Laie/in“ zu besitzen, gefolgt von der Einschätzung „eher Laie/in“ und „eher Experte/in“. Das bedeutet, dass ein Interesse am Thema und Vorwissen vorhanden war.

Die Bewertung der verschiedenen Aussagen nach der Diskussion zu Künstlicher Intelligenz in der Medizin zeigt besonders, dass die Teilnehmer*innen weniger ethische Bedenken bezüglich des Themas haben. Sie denken allerdings nicht, dass Künstliche Intelligenz bald in der Medizin Ärzte ersetzen wird. Die Sicherheit von Patientendaten halten sie für eher problematisch.

Die Befragung nach der Diskussion zeigt, dass das Modell der Fishbowl-Debatte im Grundsatz funktioniert. Der Gesamteindruck der Veranstaltung wurde von beinahe allen Teilnehmer*innen als „gut“ oder „sehr gut“ bezeichnet. Bei der nachträglichen Frage nach dem Thema der Veranstaltung machten die meisten Befragten keine Angabe. Zu gleichen Teilen wurde der Titel der Veranstaltung (Künstliche Intelligenz in der Medizin) und „Einsatzmöglichkeiten und Risiken von KI in der Medizin“ genannt. Dennoch darf auf Grund der restlichen Antworten angenommen werden, dass die Themen das Publikum erreicht haben.

Wie gut die Diskussion von den Befragten aufgenommen wurde, zeigt besonders die Einzelbenotung der verschiedenen Aspekte der Veranstaltung. Sie wurde fast durchgängig mit „gut“ oder „sehr gut“ bewertet. Besonders gut wurde die Vertrauenswürdigkeit der Expert*innen bewertet.

Die meisten Besucher*innen, die dazu eine Antwort abgegeben haben, fanden, dass die Gespräche auf Augenhöhe stattfanden. Etwas weniger positiv sind die Aussagen zu der Frage, ob viel neues Wissen aus der Veranstaltung mitgenommen wurde. Nicht ganz die Hälfte der Befragten (46 %) bejahte dies, wobei bei der Einordnung dieses Ergebnisses das relativ hohe Vorwissen der Befragten zu beachten ist. Dennoch würden die meisten Besucher*innen sowohl erneut an Debatten zum Thema Künstliche Intelligenz in der Medizin als auch an einer Diskussion im Stil einer Fishbowl-Debatte teilnehmen. Die meisten Teilnehmer*innen hatten darüber hinaus keine weiteren Anmerkungen zum Stil der Debatte. Die negativen

Anmerkungen überwogen geringfügig, wobei es insgesamt nur zwei rein negative Anmerkungen gab.

Die Auswertung der Veranstaltung „Künstliche Intelligenz in der Medizin – Reden Sie mit!“ zeigt, dass Diskussionen im Fishbowl-Format eine gute Möglichkeit sind, um das Publikum einer Debatte mit einzubeziehen. Der direkte Austausch zwischen Expert*innen und Publikum wird gefördert und auf Augenhöhe geführt.

Die Veranstaltung in der Charité konnte vor allem Besucher*innen aus dem Berufsfeld und mit einem Bezug zum Thema anlocken. Entsprechend hoch schätzten die Besucher*innen selbst ihr Vorwissen ein und entsprechend gering war der Wissenszuwachs durch die Debatte. Dennoch zeigten sich fast alle Teilnehmer*innen zufrieden mit der Fishbowl-Diskussion. Bei der Einschätzung des Themas an sich durch die Besucher*innen konnten zudem tatsächliche Veränderungen festgestellt werden.

IV.8. Anhänge

IV.8.1. Anhang 1) Fragebogen

Öffentliche Diskussion: Künstliche Intelligenz in der Medizin

Vielen Dank im Voraus für Ihre Mithilfe bei unserer Forschung!

Wir möchten Sie bitten, **vor** Beginn der Veranstaltung den **Fragebogen 1** und **nach** der Veranstaltung den **Fragebogen 2** auszufüllen.

Die Fragebogenerhebung findet im Rahmen einer Evaluation des Projekts „**Risikokommunikation zur Künstlichen Intelligenz**“ durch das Nationale Institut für Wissenschaftskommunikation (NaWik) statt.

Alle im Fragebogen erhobenen Daten fließen ausschließlich anonymisiert in die Studie ein. Eine Weitergabe der Daten an Dritte erfolgt nicht.

Kontakt

Gwendolyn Rohsaint grohsaint@gmail.com

Carolin Otte carolin@otte.info

Dr. Philipp Niemann niemann@nawik.de

Öffentliche Diskussion: Künstliche Intelligenz in der Medizin

Fragebogen 1 – Bitte beantworten Sie diese Fragen vor der Veranstaltung.

1) Haben Sie bereits an einer öffentlichen Veranstaltung zu einem wissenschaftlichen Thema teilgenommen?

- Ja, einmal Ja, mehr als einmal Nein

2) Haben Sie bereits an einer öffentlichen Diskussion mit Bürgerbeteiligung teilgenommen?

- Ja, einmal Ja, mehr als einmal Nein

3) Haben Sie einen beruflichen oder privaten Bezug zum Thema Künstliche Intelligenz in der Medizin?

- Nein Ja, nämlich.....

4) Was war für Sie der wichtigste Grund zur heutigen Veranstaltung zu kommen?
(bitte nur eine Antwortmöglichkeit ankreuzen)

- Ich interessiere mich für Wissenschaft.
- Ich möchte mehr über Künstliche Intelligenz in der Medizin wissen.
- Ich möchte mich am Meinungsaustausch beteiligen.
- Ein/e bestimmte/r Wissenschaftler/in
- Ein anderer Grund, nämlich.....

5) Was ist Ihr höchster Bildungsabschluss?

- Schüler/in Hauptschulabschluss Mittlere Reife Sonstiges:
- (Fach-) Hochschulreife Hochschulabschluss Promotion

6) Was ist Ihre berufliche Tätigkeit?

.....

7) Ihr Geburtsjahr:

.....

8) Ihr Geschlecht:

- weiblich männlich divers



Öffentliche Diskussion: Künstliche Intelligenz in der Medizin

Fragebogen 1 – Bitte beantworten Sie diese Fragen vor der Veranstaltung.

9) Wie schätzen Sie selbst Ihre Kenntnisse zum Thema Künstliche Intelligenz in der Medizin ein?

Experte/in				Laie	
1	2	3	4	5	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

10) Bitte beurteilen Sie folgende Aussagen:

	trifft voll und ganz zu			trifft überhaupt nicht zu	
a) In wenigen Jahren wird Künstliche Intelligenz Ärzte z.B. im OP-Saal oder bei der Diagnose ersetzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Ich halte die Anwendung von Künstlicher Intelligenz im OP-Saal für ethisch bedenklich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Ich halte die Anwendung von Künstlicher Intelligenz in der Diagnostik für ethisch bedenklich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Das Thema Künstliche Intelligenz in der Medizin wird durch die Medien stark problematisiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Die Datensicherheit ist beim Umgang mit Patientendaten gewährleistet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Ich fühle mich durch die mediale Berichterstattung ausreichend über den Stand der Forschung zu Künstlicher Intelligenz in der Medizin informiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Das Thema Künstliche Intelligenz in der Medizin wird durch die Medien stark positiv dargestellt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vielen Dank für Ihre Teilnahme!

Bitte füllen Sie den Fragebogen 2 erst nach der Veranstaltung aus.



Öffentliche Diskussion: Künstliche Intelligenz in der Medizin

Fragebogen 2 – Bitte beantworten Sie diese Fragen nach der Veranstaltung.

1) Wie ist Ihr Gesamteindruck von der heutigen Diskussion?

sehr gut gut befriedigend ausreichend mangelhaft

2) In einem Satz zusammengefasst: Worum ging es in der heutigen Diskussion?

.....

3) Wie bewerten Sie folgende Aspekte der gesamten Veranstaltung?

	sehr gut	gut	befriedigend	ausreichend	mangelhaft
a) Einleitung und Vorstellung der Experten/innen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Auswahl der Experten/innen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Moderation und Diskussionsführung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Beteiligung der Bürger/innen an der Diskussion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Interaktion zwischen Experten/innen und Bürger/innen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Vertrauenswürdigkeit der Experten/innen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Unterhaltungswert der Veranstaltung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
h) Veranstaltungsgröße	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4) Hatten Sie das Gefühl, dass die Diskussion mit den Experten/innen auf Augenhöhe stattfand?

Ja Nein, weil

5) Bitte beurteilen Sie die folgende Aussage:

Ich habe aus dieser Veranstaltung viel neues Wissen über Künstliche Intelligenz in der Medizin mitgenommen.

trifft voll und ganz zu trifft überhaupt nicht zu



Bitte beachten Sie **die zweite Seite.**

Öffentliche Diskussion: Künstliche Intelligenz in der Medizin

Fragebogen 2 – Bitte beantworten Sie diese Fragen nach der Veranstaltung.

6) Würden Sie wieder an einer öffentlichen Diskussion zum Thema Künstliche Intelligenz in der Medizin teilnehmen?

- Ja Nein, weil

7) Würden Sie wieder an einer öffentlichen Diskussion im Stil der heutigen Veranstaltung teilnehmen?

- Ja Nein, weil

8) Haben Sie Anmerkungen zum Stil der heutigen Veranstaltung?

+ **-**

9) Bitte beurteilen Sie erneut folgende Aussagen:

	trifft voll und ganz zu			trifft überhaupt nicht zu	
a) In wenigen Jahren wird Künstliche Intelligenz Ärzte z.B. im OP-Saal oder bei der Diagnose ersetzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Ich halte die Anwendung von Künstlicher Intelligenz im OP-Saal für ethisch bedenklich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Ich halte die Anwendung von Künstlicher Intelligenz in der Diagnostik für ethisch bedenklich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Das Thema Künstliche Intelligenz in der Medizin wird durch die Medien stark problematisiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Die Datensicherheit ist beim Umgang mit Patientendaten gewährleistet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Ich fühle mich durch die mediale Berichterstattung ausreichend über den Stand der Forschung zu Künstlicher Intelligenz in der Medizin informiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Das Thema Künstliche Intelligenz in der Medizin wird durch die Medien stark positiv dargestellt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Vielen Dank für Ihre Teilnahme!

IV.8.2. Anhang 2) Fragen der Leitfadeninterviews

Fragenblock 1

Wie schätzen Sie die Veranstaltung ein? (Posttest, Frage 2)

Was hat Ihnen besonders gut gefallen?

Was hat Ihnen gar nicht gefallen?

Was haben Sie von dieser Veranstaltung erwartet?

Was haben Sie vermisst?

Fragenblock 2

Sie haben im Fragebogen angegeben, dass Sie **bereits** an einer öffentlichen Diskussion mit Bürgerbeteiligung teilgenommen haben. (Pretest, Frage 5) Welches Format hatte diese Diskussion?

Sie haben im Fragebogen angegeben, dass Sie **noch nie** an einer öffentlichen Diskussion mit Bürgerbeteiligung teilgenommen haben. (Pretest, Frage 5) In welcher Form haben Sie sich die Bürgerbeteiligung vorgestellt?

Fragenblock 3

Haben Sie sich an der Diskussion beteiligt?

Wenn **nein**, warum haben Sie sich nicht beteiligt?

Wenn **ja**, wie haben Sie sich dabei gefühlt?

Hat es Sie Überwindung gekostet, aufzustehen und sich in den Experten/innenkreis zu setzen?

Fragenblock 4

Sie haben sich eingestuft als Experte/in beziehungsweise Laie/in Stufe X. (Pretest, Fragen 1, 2 und 9)

Warum?

Wie viel Vorwissen haben Sie zum Thema Künstliche Intelligenz in der Medizin?

Fragenblock 5

Ihre Haltung zum Thema Künstliche Intelligenz in der Medizin hat sich geändert. (Vergleich zwischen Pretest, Frage 10 und Posttest, Frage 9)

Fragenblock 6

Warum beziehungsweise durch was wurde das ausgelöst? Waren Sie der Meinung, Künstliche Intelligenz würde Ärzte z. B. im OP-Saal oder bei der Diagnose in wenigen Jahren ersetzen? (Vergleich Pretest 10 und Posttest 9 a))

Wie hat die heutige Diskussion etwas an Ihrer Meinung dazu geändert?

Fragenblock 7

Sie haben die Veranstaltung in folgendem Satz zusammengefasst: _____. (Posttest 1)

Warum was das für Sie der zentrale Aspekt der Diskussion?

V. Vergleich der beiden Fishbowl-Debatten

Im abschließenden Kapitel der Evaluation der Fishbowl-Debatten des RIKI-Projekts werden die Einschätzungen der Teilnehmer*innen der beiden Debatten in den beiden zentralen Bereichen „Veranstaltungsqualität“ und „Problembewusstsein im jeweiligen Themenfeld: Mediale Berichterstattung“ miteinander verglichen:

V.1. Veranstaltungsqualität

Die Teilnehmer*innen wurden nach der Veranstaltung gebeten, diese auf einer Likert-Skala von sehr gut (1) über gut (2), befriedigend (3) und ausreichend (4) bis hin zu mangelhaft (5) zu bewerten. Der Mittelwertvergleich gibt Aufschluss darüber, wie die beiden Veranstaltungen jeweils im Durchschnitt eingeschätzt wurden.

In den folgenden Betrachtungen wurden die Befragten, die keine Angaben machten, nicht mehr miteinbezogen. Die Fishbowl-Debatte *Künstliche Intelligenz in der Medizin* in Berlin wurde von den Besucher*innen im Schnitt mit 1,95 (gut) bewertet unter einer Standardabweichung von 0,669, wobei 21 von 28 Teilnehmer*innen eine Wertung abgaben. Die Diskussionsveranstaltung *Autonomes Fahren und Künstliche Intelligenz* wurde durchschnittlich mit 1,68 (gut) bewertet, mit ähnlicher Standardabweichung von 0,627, wobei 25 der 26 Teilnehmer*innen eine Wertung abgaben (Tabelle V.1).

Tabelle V.1: Mittelwertvergleich des Gesamteindrucks

	Mittelwert	N	Std.-Abweichung
Berlin	1,95 (gut)	21	0,669
Karlsruhe	1,68 (gut)	25	0,627

Die häufigste Bewertung war bei beiden Veranstaltungen gut (2,0). Die Fishbowl *KI in der Medizin* wurde von 81 % mit gut und besser bewertet, die Fishbowl *Autonomes Fahren* sogar von 92 % der Teilnehmer*innen. Bei beiden Veranstaltungen wurde keine Bewertung abgegeben, die schlechter als befriedigend ausfiel (Abb. V.1).

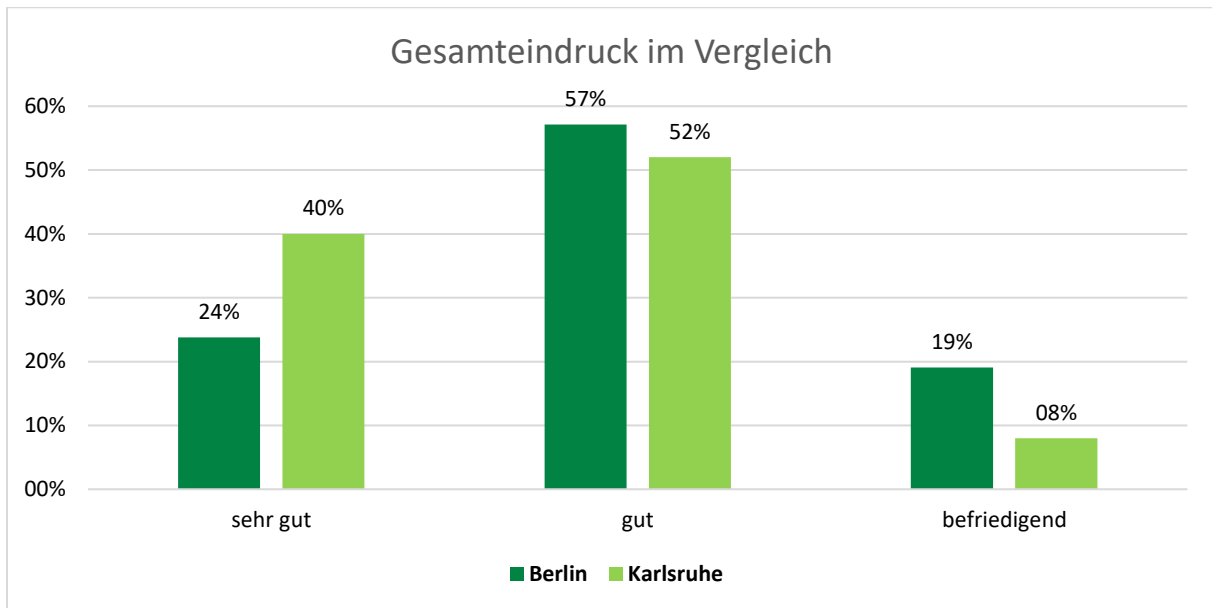


Abbildung V.1: Gesamteindruck der Fishbowl-Debatten. Berlin: N=21; Karlsruhe: N=25

V.1.1. Aspekte der Veranstaltung im Vergleich

Die folgenden Aspekte wurden herangezogen, um die Veranstaltungsqualität zu messen: Einleitung und Vorstellung der Expert*innen, Auswahl der Expert*innen, Moderation und Diskussionsführung, Beteiligung der Bürger*innen an der Diskussion, Interaktion zwischen Expert*innen und Bürger*innen, Vertrauenswürdigkeit der Expert*innen sowie der Unterhaltungswert der Veranstaltung und die Bewertung der Veranstaltungsgröße. Wie der Tabelle V.2 zu den Aspekten der Veranstaltungsqualität im Vergleich zwischen Berlin und Karlsruhe zu entnehmen ist, wurden alle Aspekte sowohl in Karlsruhe als auch in Berlin durchschnittlich gut oder besser bewertet. Die durchschnittlichen Bewertungen in Berlin fielen ein wenig besser aus als in Karlsruhe, nur die Auswahl der Expert*innen sowie die Beteiligung der Bürgerinnen und Bürger an der Diskussion wurden in Karlsruhe besser bewertet.

Tabelle V.2: Aspekte der Veranstaltungsqualität im Vergleich zwischen Berlin und Karlsruhe

	Berlin			Karlsruhe		
	Mittelwert	N	Standard-abweichung	Mittelwert	N	Standard-abweichung
Einleitung und Vorstellung der Experten/innen	1,65 (gut)	20	0,745	1,46 (sehr gut)	24	0,588

Auswahl der Experten/innen	1,55 (gut)	20	0,605	1,58 (gut)	24	0,584
Moderation und Diskussionsführung	1,57 (gut)	21	0,746	1,36 (sehr gut)	25	0,569
Beteiligung der Bürger/innen an der Diskussion	1,90 (gut)	21	0,831	1,96 (gut)	25	0,889
Interaktion zwischen Experten/innen und Bürger/innen	1,71 (gut)	21	0,644	1,54 (gut)	24	0,588
Vertrauenswürdigkeit der Experten/innen	1,33 (sehr gut)	21	0,730	1,28 (sehr gut)	25	0,458
Unterhaltungswert der Veranstaltung	2,05 (gut)	21	0,921	1,88 (gut)	25	0,600
Veranstaltungsgröße	2,10 (gut)	21	0,625	2,04 (gut)	25	0,790

V.1.2. Wie wurde die Qualität der Diskussion empfunden?

Im Posttest bewerteten die Befragten die Moderation und Diskussionsführung, die Beteiligung der Bürger*innen an der Diskussion und die Interaktion zwischen Expert*innen und Bürger*innen. Diese drei Indikatoren geben Aufschluss darüber, wie die Qualität der Diskussionsveranstaltungen von den Teilnehmer*innen empfunden wurde.

Bei beiden Veranstaltungen wurde die Moderation und Diskussionsführung mehrheitlich als sehr gut bewertet. 68 % der Befragten in Karlsruhe und 57,1 % der Befragten in Berlin empfanden diese als sehr gut. Bei keiner der beiden Veranstaltungen wurde eine schlechtere Bewertung als befriedigend abgegeben. Die Moderation von Eva Wolfangel bei der Fishbowl in Karlsruhe bewerteten 96 % der Befragten als gut und besser, zur Moderation von Nicola Kuhrt bei der Fishbowl in Berlin kreuzten 85,7 % der Befragten gut oder sehr gut an (Abb.V.2).

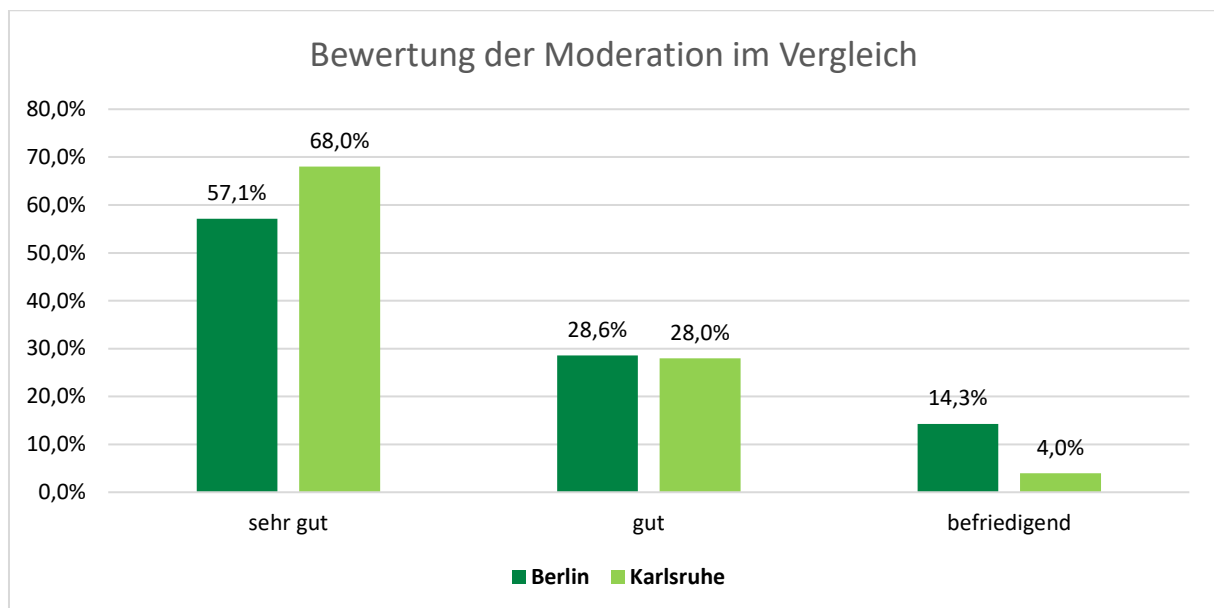


Abbildung V.2: Bewertung der Moderation. Berlin: N=21; Karlsruhe: N=25

Im Durchschnitt wurde die Moderation der Berliner Fishbowl *Künstliche Intelligenz in der Medizin* mit 1,57 (gut) bei einer Standardabweichung von 0,746 und die der Karlsruher Fishbowl *Autonomes Fahren und Künstliche Intelligenz* mit 1,36 (sehr gut) bei einer Standardabweichung von 0,569 bewertet (Tabelle V.3).

Tabelle V.3 Mittelwertvergleich der Bewertung der Moderation

	Mittelwert	N	Std.-Abweichung
Berlin	1,57 (gut)	21	0,746
Karlsruhe	1,36 (sehr gut)	25	0,569

Bei beiden Veranstaltungen wurde die Beteiligung der Bürgerinnen und Bürger an der Diskussion mehrheitlich als *gut* oder *sehr gut* bewertet. 80,9 % der Befragten in Berlin gaben *sehr gut* oder *gut* an, unter den Befragten in Karlsruhe taten dies ähnlich viele (80 %). Nur 20 % der Berliner und 19 % der Karlsruher Befragten empfanden die Bürgerbeteiligung an der jeweiligen Diskussion als *befriedigend* oder *ausreichend* (Abb. V.3)

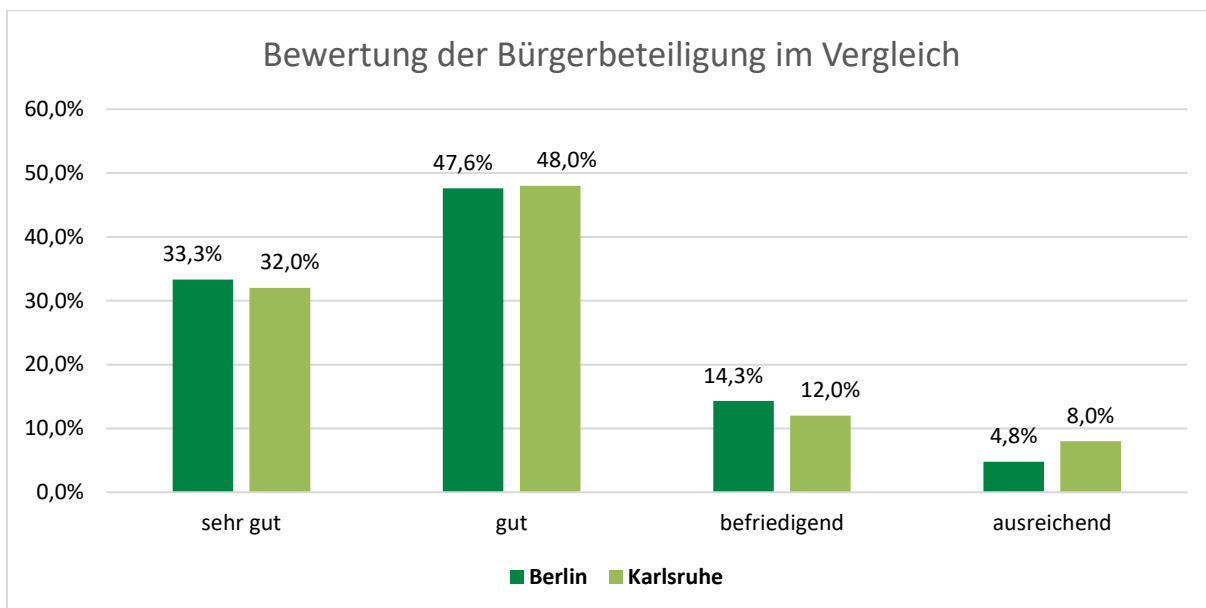


Abbildung V.3: Bewertung der Bürgerbeteiligung. Berlin: N=21; Karlsruhe: N=25

Der dritte Indikator zur Veranstaltungsqualität, die Interaktion zwischen Expert*innen und Bürger*innen, wurde ähnlich gut wie die vorherigen beiden bewertet. 95 % der Berliner und 95,8 % der Karlsruher Teilnehmer*innen bewerteten sie als gut oder besser (Abb V.4).

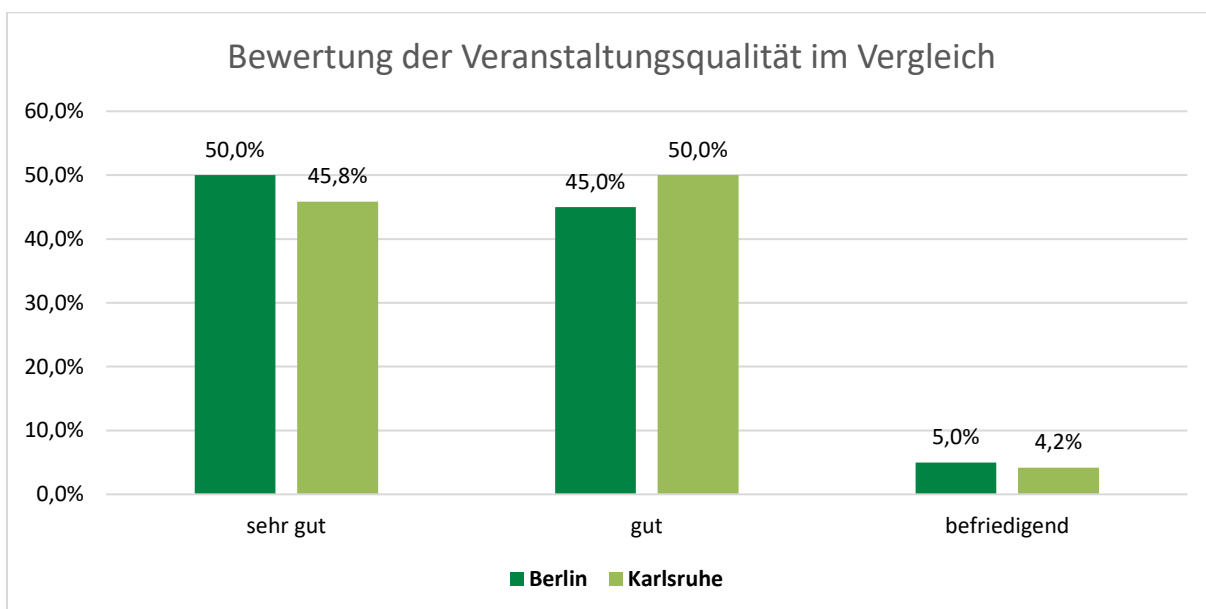


Abbildung V.4: Bewertung der Veranstaltungsqualität. Berlin: N=21; Karlsruhe: N=24

Die Befragten in Berlin bewerteten die Interaktion mit der Expertin Dr. Andrea Schenk und den Experten Dr. med. Alexander Meyer und Prof. Dr. Thomas Berlage durchschnittlich gut mit 1,71 bei einer Standardabweichung von 0,644. Die Befragten in Karlsruhe bewerteten die Interaktion mit der Expertin Prof. Anne Koziolk und den Experten Prof. J. Marius Zöllner und

Christian Hubschneider durchschnittlich mit 1,54 (gut) bei einer Standardabweichung von 0,588 (Tabelle V.4).

Tabelle V.4 Mittelwertvergleich der Bewertung der Interaktion zwischen Experten/innen und Bürger/innen

	Mittelwert	N	Std.-Abweichung
Berlin	1,71 (gut)	21	0,644
Karlsruhe	1,54 (gut)	24	0,588

V.1.3. Konnten die Diskussionen auf Augenhöhe geführt werden?

Alle Befragten in Karlsruhe gaben an, dass die Diskussion in der Fishbowl auf Augenhöhe stattfand. Diesen Eindruck hatten auch fast alle Teilnehmer*innen in Berlin (95%).

V.1.4. Wurde das Veranstaltungsformat vom Publikum angenommen?

25 der Berliner Teilnehmer*innen und 21 der Karlsruher Teilnehmer*innen gaben an, dass sie wieder an einer Fishbowl-Diskussion teilnehmen würden. Keiner gab an, nicht noch einmal an einer Fishbowl-Diskussion teilnehmen zu wollen, es gab jedoch Enthaltungen. In Karlsruhe wurde ein Fragebogen dazu nicht ausgefüllt, während sich in Berlin sieben von 28 Teilnehmer*innen nicht dazu äußerten.

V.1.5. Wie werden öffentliche Diskussionen zu Themen aus dem Bereich Künstliche Intelligenz angenommen?

Fast alle Befragten würden wieder an einer öffentlichen Diskussion zum Thema Künstliche Intelligenz in der Medizin beziehungsweise Autonomes Fahren teilnehmen wollen. 95,2 % der Teilnehmer*innen der Fishbowl *KI in der Medizin* und 96 % der Teilnehmer*innen der Fishbowl *Autonomes Fahren* machten entsprechende Angaben (Abb. V.5).

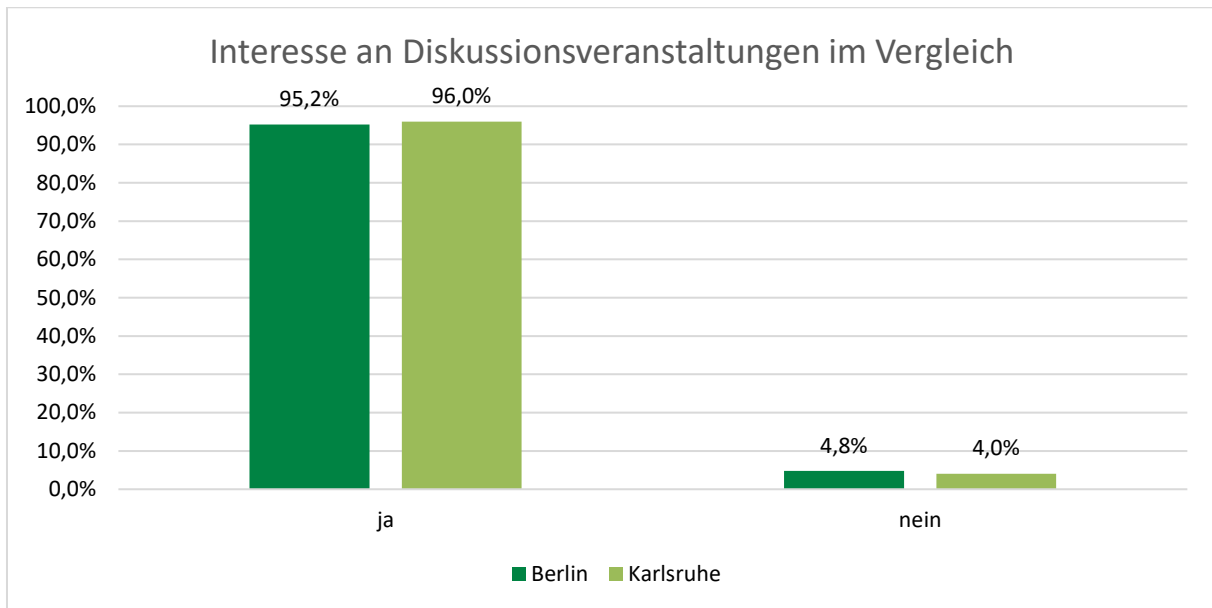


Abbildung V.5: Interesse an Diskussionsveranstaltungen zum Thema. Berlin: N=21; Karlsruhe: N=25

V.2. Problembewusstsein im jeweiligen Themenfeld: Mediale Berichterstattung

Beim Aspekt Problembewusstsein im jeweiligen Themenfeld konzentriert sich die folgende Darstellung auf den Bereich, der unmittelbar mit dem Mediensystem verknüpft ist. Verglichen wird das Antwortverhalten zur Frage, wie die mediale Berichterstattung zu Themen im Bereich der Künstlichen Intelligenz von den Teilnehmer*innen der beiden Debatten eingeschätzt wird. Drei Aspekte, die die mediale Darstellung der Themen KI in der Medizin bzw. Autonomes Fahren beleuchten, werden sowohl in der Pretest-Befragung als auch im Posttest mittels entsprechender Aussage-Items erfasst. Um die Verschiebungen zwischen Pre- und Posttest bei den jeweiligen Veranstaltungen zu untersuchen, wird ein Mittelwertvergleich angestellt. Eine dieser Aussagen lautet: „Das Thema Künstliche Intelligenz in der Medizin bzw. Autonomes Fahren wird durch die Medien stark problematisiert.“ (vergleiche jeweils Anhang 1 in den Kapiteln III.8.1 & IV.8.1). Der im Folgenden als *Problematisierung durch die Medien* aufgeführte Aspekt untersucht, wie sehr die Befragten den Eindruck haben, dass das Thema KI in den Medien negativ dargestellt wird. Die gegenteilige Einschätzung einer starken positiven Darstellung wird mit der Aussage „Das Thema Künstliche Intelligenz in der Medizin bzw. Autonomes Fahren wird durch die Medien stark positiv dargestellt.“ geprüft (vergleiche jeweils Anhang 1 in den Kapiteln III.8.1 & IV.8.1). Im Folgenden wird sie als *Überhöht positive Darstellung durch die Medien* bezeichnet. Außerdem wurde um eine Einschätzung gebeten, wie gut sich die Befragten durch die Medien zur KI-Forschung informiert fühlen: „Ich fühle mich durch die mediale Berichterstattung ausreichend über den Stand der Forschung zu Künstlicher

Intelligenz in der Medizin bzw. Autonomes Fahren informiert.“ (vergleiche jeweils Anhang 1 in den Kapiteln III.8.1 & IV.8.1). Betitelt wird dieser Aspekt im Folgenden mit *Medien informieren zum Thema*. Die Zustimmung oder Ablehnung der Aussagen ist Likert-skaliert mit 0= trifft überhaupt nicht zu, 1=trifft eher nicht zu, 2=neutral, 3=trifft eher zu und 4=trifft voll und ganz zu. Ein Mittelwert von 2 bedeutet also, dass die Teilnehmer*innen im Durchschnitt neutral zu der Aussage stehen. Eine Erhöhung des Mittelwerts zwischen Pretest und Posttest bedeutet, dass die Teilnehmer*innen nach der Veranstaltung im Schnitt eher zustimmen.

Tabelle V.5: Mittelwertvergleich der Veränderungen in der Wahrnehmung der medialen Berichterstattung durch die Fishbowl-Debatten Quelle: Eigene Darstellung

		Berlin			Karlsruhe		
		Mittelwert	N	Standard-abweichung	Mittelwert	N	Standard-abweichung
Problematisierung durch die Medien	Pretest	2,3	20	1,031	2,28	25	1,137
	Posttest	1,8	20	1,056	2,28	25	1,173
Überhöht positive Darstellung durch die Medien	Pretest	1,7	20	0,865	2,12	25	0,927
	Posttest	1,7	20	0,733	1,84	25	0,898
Medien informieren zum Thema	Pretest	1,19	21	1,030	1,68	25	0,900
	Posttest	1,52	21	0,928	1,84	25	1,028

Beim Mittelwertvergleich zwischen Pre- und Posttest zum Aspekt *Problematisierung durch die Medien* zeigt sich in Karlsruhe kein Unterschied. In Berlin hingegen zeigt Tabelle V.5 eine deutliche Verschiebung um 0,5 nach unten, d. h. die Teilnehmer*innen stimmen nach der Debatte der Aussage im Schnitt weniger zu, dass die Medien das Thema KI in der Medizin stark problematisieren.

Umgekehrt stellt sich die Situation bei der Frage, ob es eine überhöht positive mediale Berichterstattung gibt. Während in Berlin im Mittel keine Verschiebung zu sehen ist, haben die Karlsruher*innen nach der Debatte im Durchschnitt eher weniger den Eindruck, dass das Thema autonomes Fahren und KI in den Medien zu positiv dargestellt wird.

Die Einschätzung der allgemeinen Frage danach, ob man durch die Medien überhaupt ausreichend über das jeweilige Thema und den aktuellen Forschungsstand informiert ist, ändert sich in beiden Veranstaltungen leicht. In beiden Fällen nimmt die mittlere Überzeugung zu, dass man selbst durch die Medien ausreichend informiert ist.

VI. Evaluation des Youtube-Kanals des RIKI-Projekts

VI.1. Der YouTube-Kanals „Risikokommunikation zur KI“

Der YouTube-Kanal des Projekts „Risikokommunikation zur Künstlichen Intelligenz“ (RIKI) wurde am 16. Mai 2019 eröffnet. Unter dem Kanalnamen „Risikokommunikation zur KI“ werden dort die für das Projekt produzierten Videos veröffentlicht.⁸ Bei diesen Videos handelt es sich zum größten Teil um Interviews mit KI-Wissenschaftlern, die Fragen zu Befürchtungen und Risiken rund um das Thema KI beantworten. Ebenso finden sich dort Videos der in dem Projekt stattfindenden Veranstaltungen, wie der öffentlichen Diskussion „Reden Sie mit: Autonomes Fahren und Künstliche Intelligenz“ am 06. Juni 2019 in Karlsruhe oder der öffentlichen Diskussion „Künstliche Intelligenz in der Medizin – Reden Sie mit!“ am 26. Oktober 2019 in Berlin.⁹

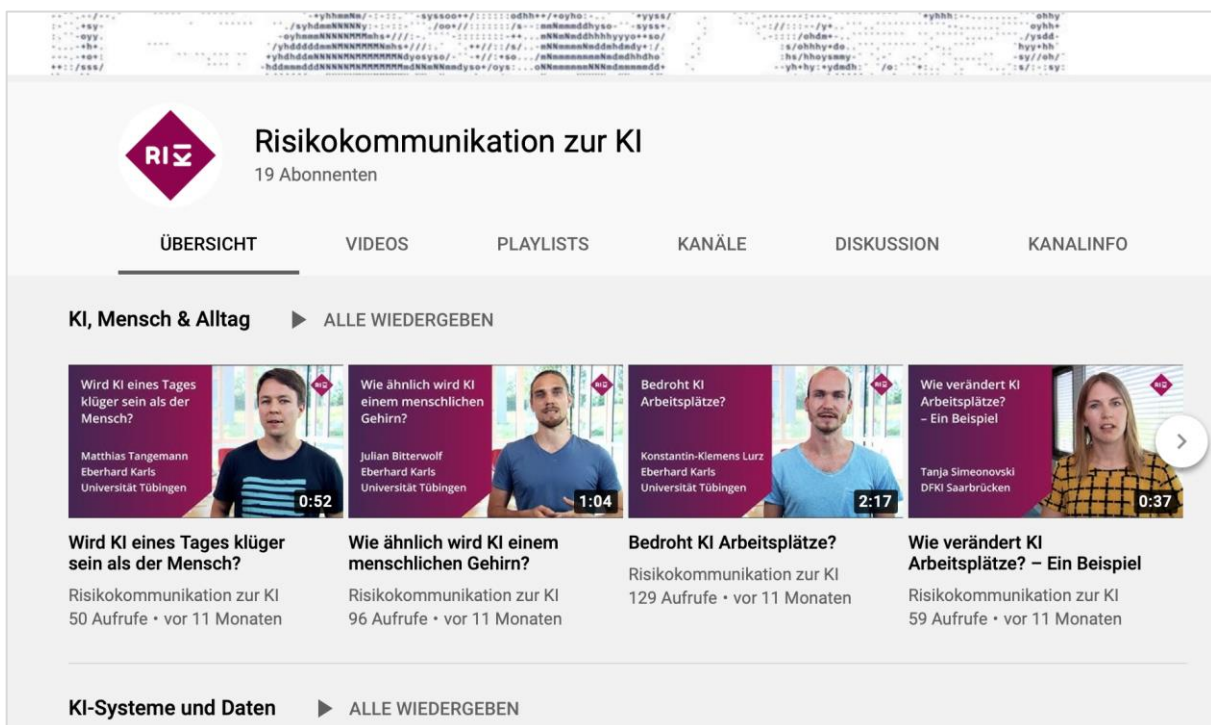


Abbildung VI.1: Der YouTube-Kanal „Risikokommunikation zur KI“, Screenshot vom 14.09.2020

Dieser Bericht evaluiert die ersten neun Monate der Laufzeit des YouTube-Kanals (16. Mai 2019 bis 16. Februar 2020). Als Datenbasis dienen Zahlen aus dem YouTube-Analysetool „YouTube Analytics“¹⁰.

⁸ <https://www.youtube.com/channel/UCmy0-mw9LDakaKM0Xln0FMQ/about> Abgerufen: 18.02.2020.

⁹ <https://www.youtube.com/channel/UCmy0-mw9LDakaKM0Xln0FMQ/videos> Abgerufen: 18.02.2020.

¹⁰ https://support.google.com/youtube/answer/9002587?hl=de&ref_topic=9313691 Abgerufen: 18.02.2020.

Im Folgenden werden zunächst die grundlegenden Metriken zum Kanal ausgewertet. Darunter fallen die Zahlen und Entwicklungen zu Abonnenten, Aufrufzahlen, Wiedergabezeiten (durchschnittlich und gesamt) sowie die „Mag ich“- und „Mag ich nicht“-Angaben. Im nächsten Schritt werden der genauere Nutzungskontext sowie die Nutzer*innen der Videos betrachtet, soweit dies aus den von YouTube zur Verfügung gestellten Daten hervorgeht. Dies beinhaltet, woher die Videos abgerufen, wie die Videos gefunden und welche Suchbegriffe dazu verwendet werden, sowie von welcher Zugriffsquelle aus die längste Abspielzeit generiert wird.

In einem dritten Schritt werden zentrale Metriken im Detail betrachtet. Dazu werden die jeweiligen „Top 5“-Videos des Kanals zu den Analyseaspekten „Gesamtaufrufe“, „durchschnittliche Wiedergabedauer“ und „Wiedergabezeit gesamt“ zusammengestellt und verglichen. Schließlich folgt eine detaillierte Einzelbetrachtung des Videos „Autonomes Fahren – die Schuldfrage“, das in vielfacher Hinsicht als exemplarisch für die Videos des Kanals gelten kann.

In Kapitel VI.3 werden die Ergebnisse der Evaluation noch einmal zusammengefasst.

VI.2. Evaluation des Kanals

VI.2.1. Allgemeine Metriken des Kanals

Zu Beginn der Auswertung wird das Leistungsverhalten des Kanals insgesamt betrachtet. Dazu werden die wichtigsten Kernmetriken eines YouTube-Kanals herangezogen: Abonnenten (Abb. VI.2), Aufrufzahlen (Abb. VI.3), Wiedergabezeiten (Abb. VI.4 und Abb. VI.5) sowie „Mag ich“- und „Mag ich nicht“-Angaben (Abb. VI.6). Alle Daten wurden für einen Zeitraum von neun Monaten erhoben, von der Kanalgründung am 16. Mai 2019 bis zum 16. Februar 2020.

VI.2.1.1. Abonnenten

Als Abonnenten zählen die Nutzer*innen, die den Kanal „Risikokommunikation zur KI“ auf YouTube abonniert haben. Dies bedeutet, dass neue Videos des Kanals direkt auf der Startseite der Nutzer*innen angezeigt werden, und dass – sofern diese Möglichkeit aktiviert wurde – die Nutzer*innen auch durch eine Benachrichtigung auf neue Videos aufmerksam gemacht werden.

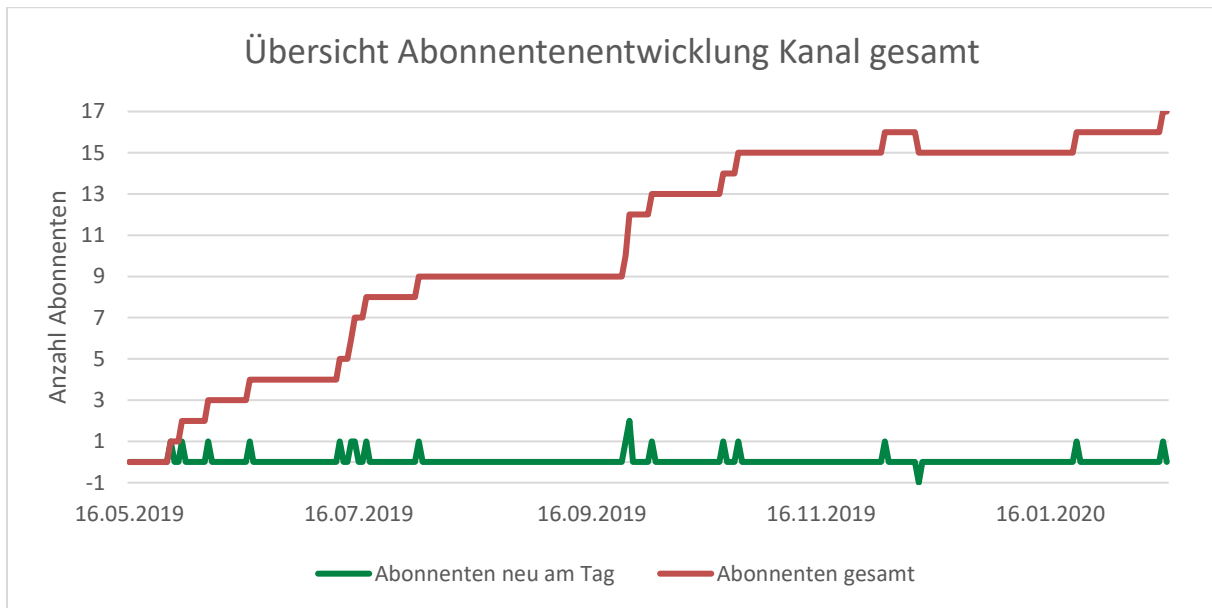


Abbildung VI.2: Übersicht über die Abonnentenentwicklung für den Kanal insgesamt (Stand 16.02.2020). Zu sehen ist der Zuwachs an neuen Abonnenten pro Tag (grün) sowie die Gesamtzahl der Abonnenten (rot).

Der Kanal konnte im Untersuchungszeitraum nur vereinzelt neue Abonnenten gewinnen (Abb. VI.2), am Ende des Zeitraums hatten 17 Nutzer*innen den Kanal abonniert, wobei im Laufe der Kanal-lebenszeit ein*e Nutzer*in das Abonnement kündigte.

VI.2.1.2. Aufrufzahlen

Jeder Klick auf ein Video gilt als ein Aufruf. Während früher Aufrufzahlen als eine der wichtigsten Metriken für den YouTube-Algorithmus und für Werbeeinnahmen galten, werden heute andere Metriken herangezogen, um Videos weit oben in den Suchergebnissen oder den Empfehlungen zu platzieren. Der YouTube-Algorithmus wird stetig weiterentwickelt und seine genaue Funktionsweise ist nicht öffentlich bekannt.

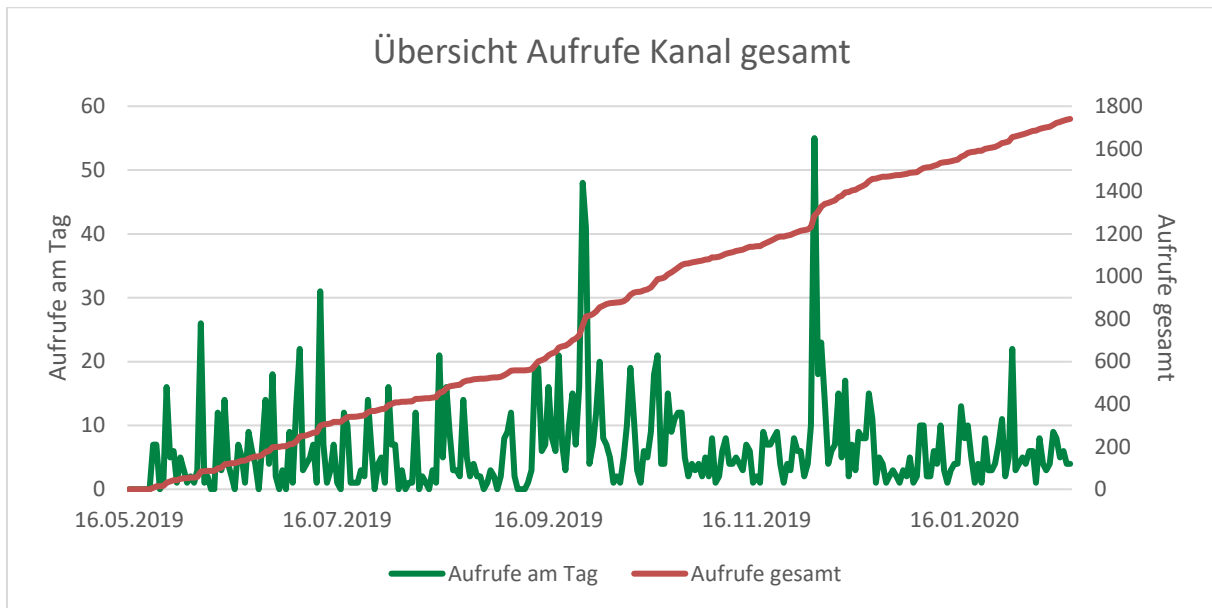


Abbildung VI.3: Übersicht über die Aufrufe des gesamten Kanals. Die Anzahl der Aufrufe am Tag (grün, linke y-Achse) ist gegenüber der Gesamtzahl der Aufrufe (rot, rechte x-Achse) abgebildet.

Die Aufrufe folgen keinem deutlichen Muster, es können aber zwei größere Ausschläge festgestellt werden, mit je über 45 Aufrufen: am 26.09.2019 (Veröffentlichung Video: „Wie ähnlich wird KI einem menschlichen Gehirn?“, 48 Aufrufe) und am 03.12.2019 (Veröffentlichung Video: „Gesundheitsdaten – was können und wie wollen wir sie sammeln?“, 55 Aufrufe). Die starken Ausschläge der Aufrufzahlen und das Fehlen einer gleichbleibend hohen „Grundmenge“ an Aufrufen zeigen, dass der Kanal keine Nutzer*innen generieren konnte, die täglich auf den Kanal zugreifen (Abb. IV.3), sondern dass der Kanal von Nutzer*innen dann aufgerufen wird, wenn sie ein spezielles Interesse an den Themen haben. Vergleicht man die Ausschläge der Aufrufzahlen mit den Daten der Videoveröffentlichungen (Abb. IV.17 in Kapitel VI.4.1), so zeigt sich, dass der Kanal vor allem aufgerufen wird, wenn neue Videos freigeschaltet werden.

VI.2.1.3. Wiedergabezeit gesamt und im Durchschnitt

Die Wiedergabezeit der Videos ist die Zeit, die ein Video aktiv von den Nutzer*innen angesehen wird. Soweit bekannt, nutzt YouTube diese Metrik inzwischen als eines der Hauptmerkmale, um einzuschätzen, wie „gut“ ein Video ist, und damit, wie hoch es in den Suchergebnissen und Empfehlungen einzuordnen ist.

Neben der Gesamt-Wiedergabezeit (Abb. VI.4) stellt YouTube auch die durchschnittliche Wiedergabedauer pro Tag (Abb. VI.5) zur Verfügung. Dies ist die durch YouTube Analytics geschätzte durchschnittliche Wiedergabezeit der Videos pro Aufruf pro Tag. Sie gibt also an,

wie lange die Nutzer*innen im Durchschnitt an diesem Tag Videos auf dem Kanal angesehen haben.

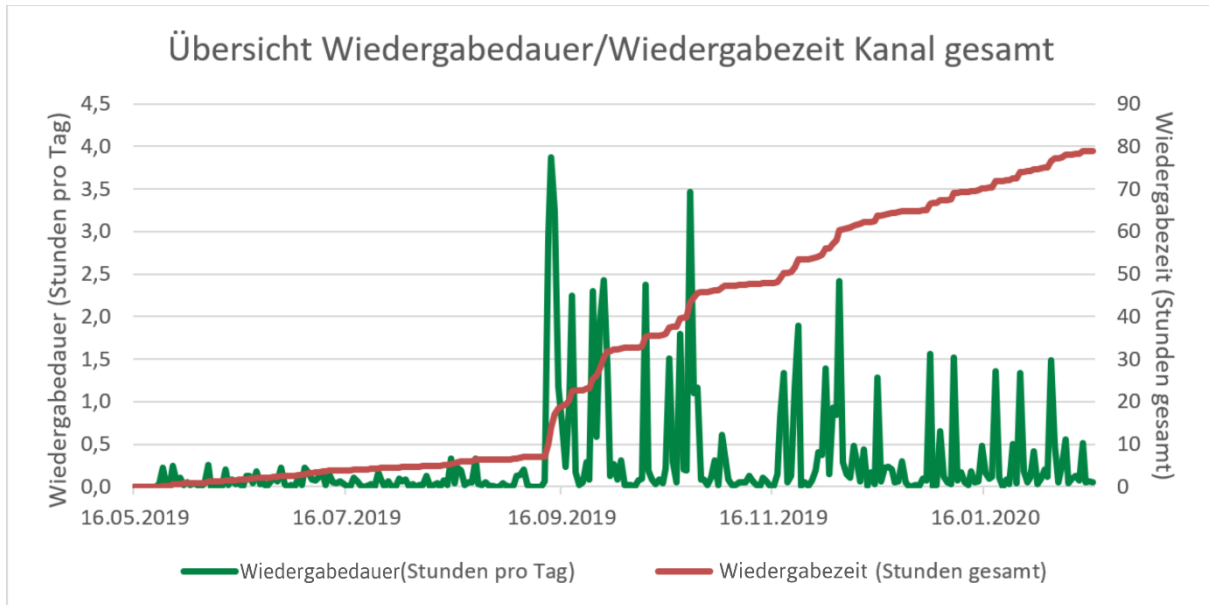


Abbildung VI.4: Übersicht über die Wiedergabedauer für den gesamten Kanal. $N = 79$ Stunden. Auf der linken y-Achse ist die Wiedergabedauer aller Videos zusammen pro Tag abgebildet (in Minuten, grün), auf der rechten die gesamte Wiedergabezeit bis zum 16.02.2020 (in Minuten, rot). Deutlich zu sehen ist der Anstieg an Wiedergabezeit ab dem 12.09.2019 mit dem Maximum am 13.09.2019.

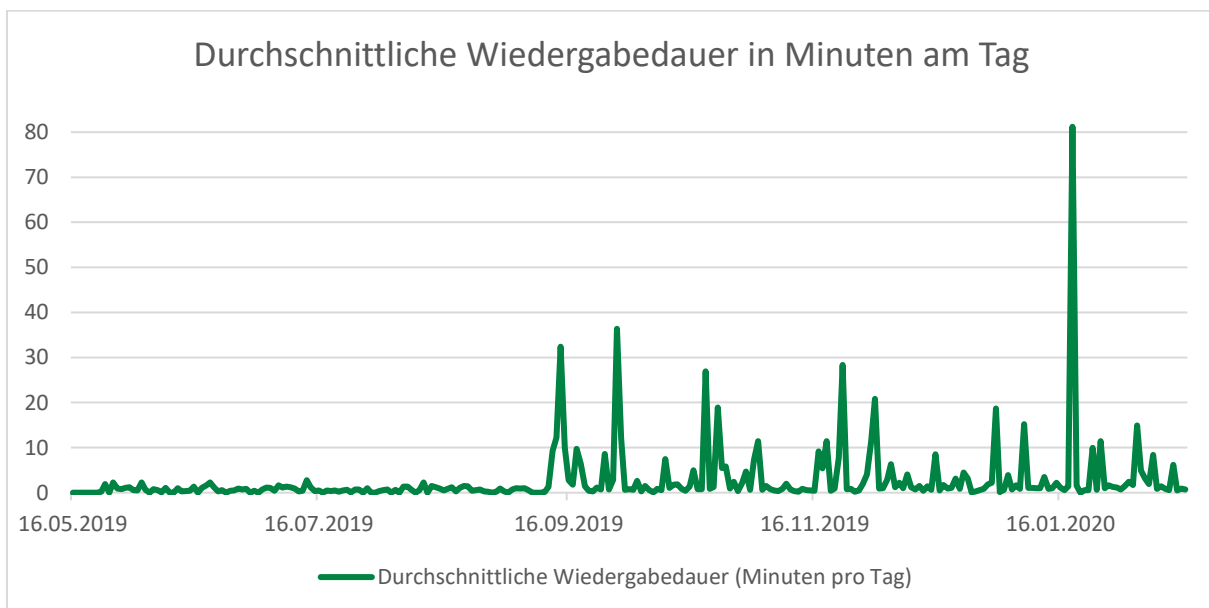


Abbildung VI.5: Durchschnittliche Wiedergabedauer der Videos auf dem Kanal in Minuten am Tag. Dieser Wert gibt an, wie lange Nutzer*innen im Durchschnitt Videos auf dem Kanal angesehen haben. Besonders auffällig ist das Maximum am 19.01.2020.

Besonders deutlich ist der Anstieg sowohl in der gesamten Wiedergabedauer als auch in der durchschnittlichen Wiedergabedauer ab dem 12.09.2019. An diesem Tag wurde die Aufnahme der Fishbowl-Diskussion „Reden Sie mit – Autonomes Fahren und Künstliche Intelligenz“ aus Karlsruhe veröffentlicht. Mit einer Länge von 1:21:21 (81,35 min) ist es das bis dato längste (bis zum 16.02.2020 zweitlängste) auf dem Kanal veröffentlichte Video. Dieses Video wurde am häufigsten aufgerufen (Abb. VI. 7 und Tabelle VI.1) und generierte die meiste Wiedergabezeit und durchschnittliche Wiedergabedauer. Dies deutet darauf hin, dass die Länge des Videos die Nutzer*innen des Kanals nicht abschreckt hat und diese ein Interesse haben, die Fishbowl-Diskussionen für mindestens einen gewissen Zeitraum nachzuverfolgen.

VI.2.1.4. „Mag ich“ und „Mag ich nicht“ Angaben

Neben Kommentaren ist die von YouTube zur Verfügung gestellte „Mag ich“- und „Mag ich nicht“-Funktion der direkteste Weg, um eine Rückmeldung zu erhalten, wie gut oder schlecht ein Video und dessen Inhalt von den Nutzer*innen aufgenommen wird. Durch einen Button können die Zuschauer*innen des Videos einen virtuellen „Daumen hoch“ (Mag ich) oder „Daumen runter“ (Mag ich nicht) verteilen und so ihre Meinung äußern (Abb. VI.6).

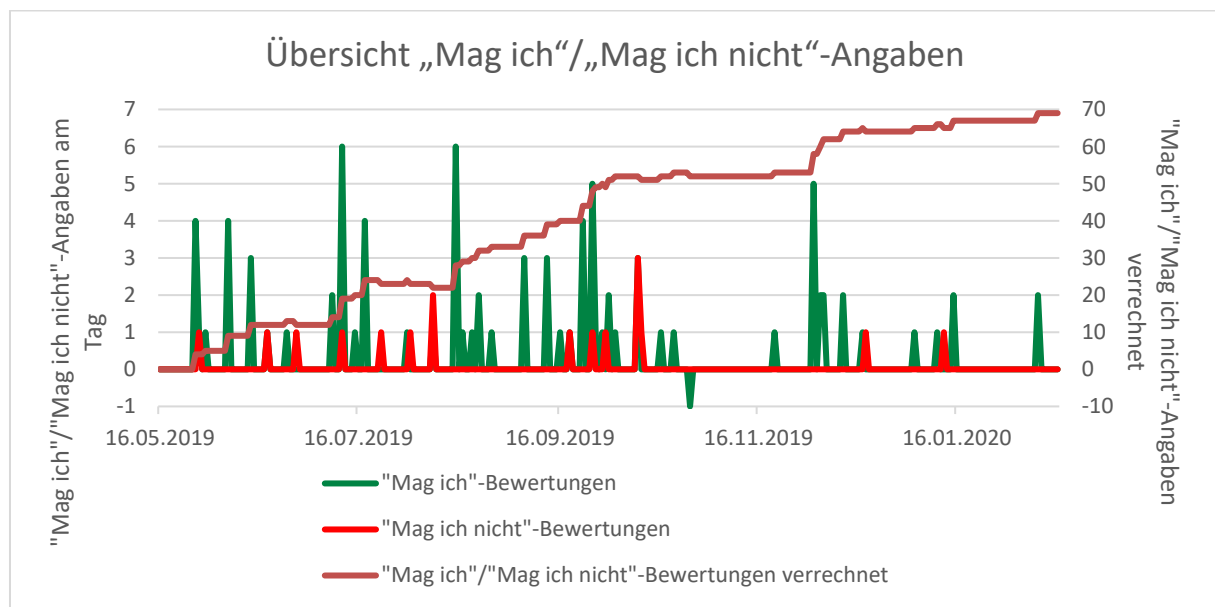


Abbildung VI.6: Übersicht über die „Mag ich“- und „Mag ich nicht“-Angaben des gesamten Kanals. Die Anzahl der „Mag ich“ und „Mag ich nicht“-Angaben („Mag ich“: grün, „Mag ich nicht“: rot) pro Tag ist auf der linken y-Achse abgetragen. Die rechte y-Achse zeigt die Verrechnung beider Angaben insgesamt. Dazu wurde den „Mag ich“-Angaben ein Wert von +1 zugewiesen, den „Mag ich nicht“-Angaben ein Wert von -1. Danach wurde die Summe aus beiden gebildet. Der negative „Mag ich“-Wert bedeutet, dass ein Nutzer*innen seine „Mag ich“-Angabe zurückgezogen hat. Insgesamt hat der Kanal so bis zum Ende des Untersuchungszeitraums 69 positive Angaben erhalten.

Ähnlich dem über die Zeit verteilten Gewinn von Abonnements, sind die „Mag ich“- und „Mag ich nicht“-Angaben relativ gleichmäßig über die Videos und den Untersuchungszeitraum verteilt. Vergleicht man, wann „Mag ich“-Angaben gemacht wurden und wann User*innen den Kanal abonniert haben (Abb VI.2.), zeigt sich keine Übereinstimmung. Es lässt sich vermuten, dass ein „mögen“ der Videos nicht automatisch mit einem Abonnement des Kanals einhergeht. Da im Untersuchungszeitraum nur zwei Kommentare abgegeben wurden, lässt sich feststellen, dass die Interaktion der Nutzer*innen mit dem Kanal gering ist (vollständige Kommentare in Kapitel VI.4.2).

VI.2.1.5. Videos nach Aufrufzahlen

Es folgt eine Auflistung aller veröffentlichten Videos nach der Anzahl ihrer Aufrufe (Abb. VI.7). Vergleicht man die Veröffentlichungszeitpunkte mit der Aufrufzahl, kann keine Übereinstimmung festgestellt werden. Wie oft ein Video aufgerufen wird, hängt nicht direkt damit zusammen, wie lange ein Video schon veröffentlicht ist.

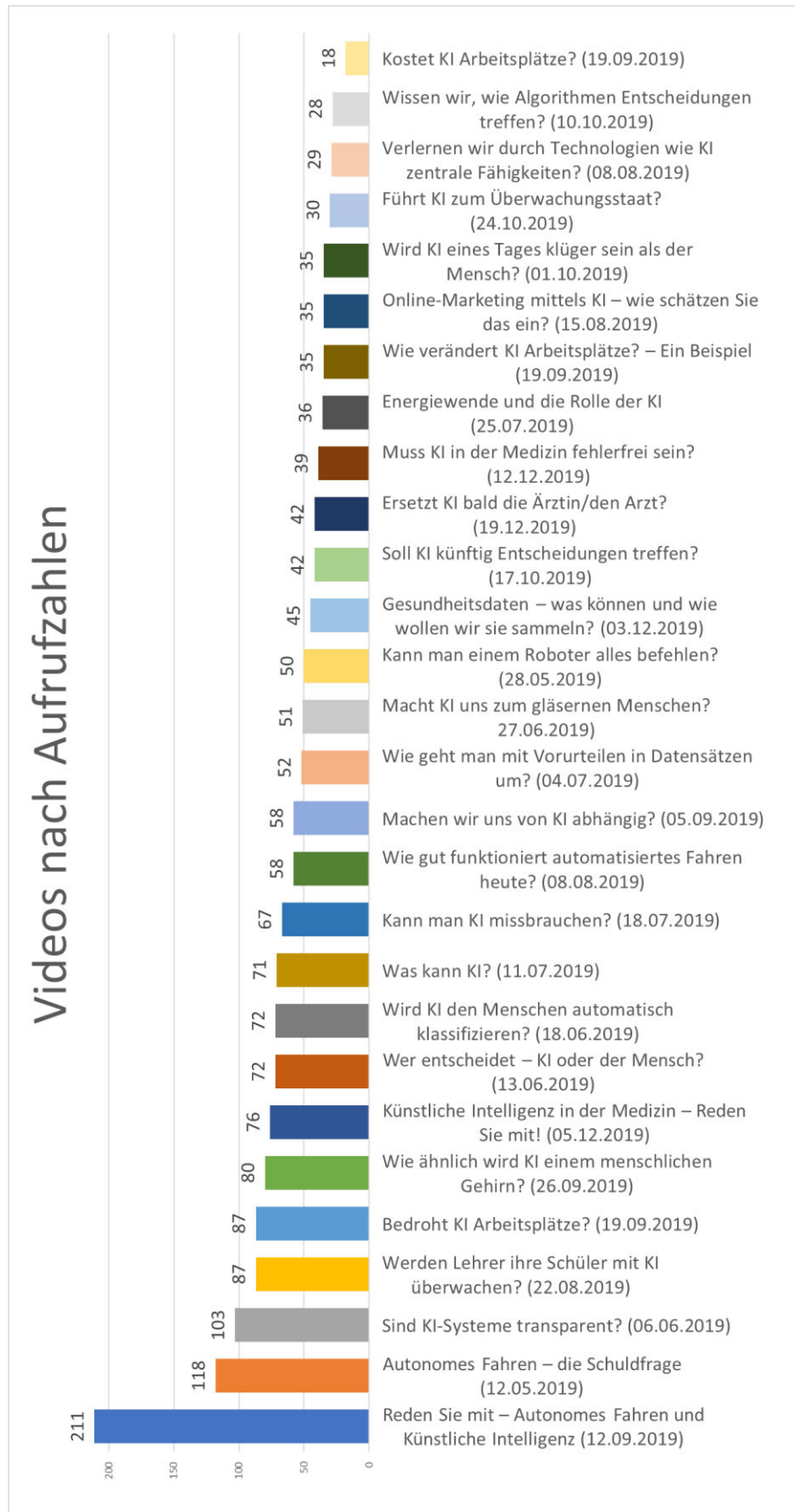


Abbildung VI.7: Videos nach Aufrufzahlen. In Klammern das Datum der Veröffentlichung der Videos.

VI.2.2. Nutzungskontext & Nutzer*innen

Neben den Daten zu der Nutzung des Kanals stellt YouTube auch Daten zu den Nutzer*innen selbst zur Verfügung. Diese Daten werden, soweit ersichtlich, durch die von den Browsern der Nutzer*innen zur Verfügung gestellten Informationen und den von ihnen selbst in ihren YouTube-Profilen hinterlegten Angaben erstellt. YouTube ist dadurch theoretisch in der Lage, ein relativ genaues Bild des durchschnittlichen Kanal-Besuchenden zu geben. YouTube sammelt dabei Daten unter anderem zu Alter und Geschlecht des Nutzers, dem Endgerät, auf dem die Videos abgespielt werden, sowie dem Land, aus dem die Videos abgerufen werden. Dazu kommen Informationen, ob die Videos auf der YouTube-Seite direkt oder über auf externen Seiten eingebundene Videos angesehen werden. Diese Daten lassen sich wieder nach Aufrufen, Wiedergabezeit etc. aufschlüsseln.

VI.2.2.1. Zugriffsquelle

Die Zugriffsquelle gibt an, von wo die Videos des Kanals abgerufen werden bzw. auf sie zugegriffen wird. YouTube verwendet hierfür verschiedene Begriffe:

- Extern: Das Video ist auf einer externen Seite (z. B. der Webseite von RIKI) eingebunden oder von extern abgerufen worden (z. B. über die Google-Suche) (Abb. VI.11).
- YouTube-Suche: Den Zuschauer*innen wurde das Video über die YouTube-Suche angezeigt, nachdem sie einen passenden Suchbegriff eingegeben hatten (Abb. VI.10).
- Kanalseiten: Die Besucher*innen greifen auf das Video über die Startseite des Kanals oder eine der Unterseiten zu.
- Direkt oder unbekannt: Ein*e Nutzer*in hat zum Beispiel einen Link auf einer Webseite angeklickt oder direkt als Adresse eingegeben, oder es stehen keine Daten zur Quelle zur Verfügung.
- Videovorschläge: Vorschläge für Videos, die nach dem Abschluss von anderen Videos angezeigt werden.
- Funktionen zur Auswahl von Inhalten: Hierzu zählen beispielsweise Zugriffe über die Startseite des Nutzers oder den Abo-Feed.
- Sonstige: Hierunter fallen unter anderem Zugriffe über Benachrichtigungen durch YouTube oder Aufrufe über Playlists, die ein Video des Kanals beinhalten.

Wie bei den sonstigen Metriken lassen sich die Zugriffsquellen nach Aufrufen (Abb. VI.8) und nach Wiedergabezeit (Abb. VI.9) aufschlüsseln.

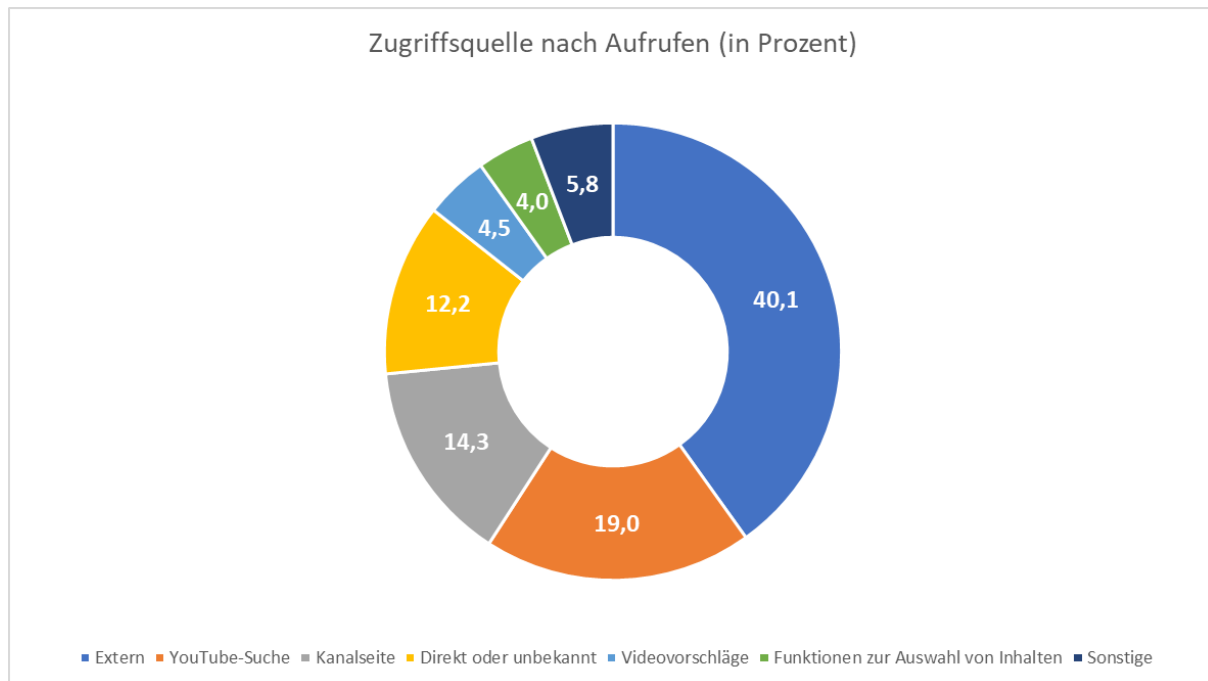


Abbildung VI.8: Zugriffsquellen des Kanals nach Aufrufen geordnet. N = 1740 Aufrufe. Extern (blau): 698 Aufrufe, 40,1%; YouTube-Suche (orange): 331 Aufrufe, 19,0%; Kanalseite (grau): 249 Aufrufe, 14,3%; Direkt oder unbekannt (gelb): 212 Aufrufe, 12,2%; Videovorschläge (hellblau): 79 Aufrufe, 4,5%; Funktionen zur Auswahl von Inhalten (grün): 70 Aufrufe, 4,0%; Sonstige (dunkelblau): 101 Aufrufe, 5,8%.

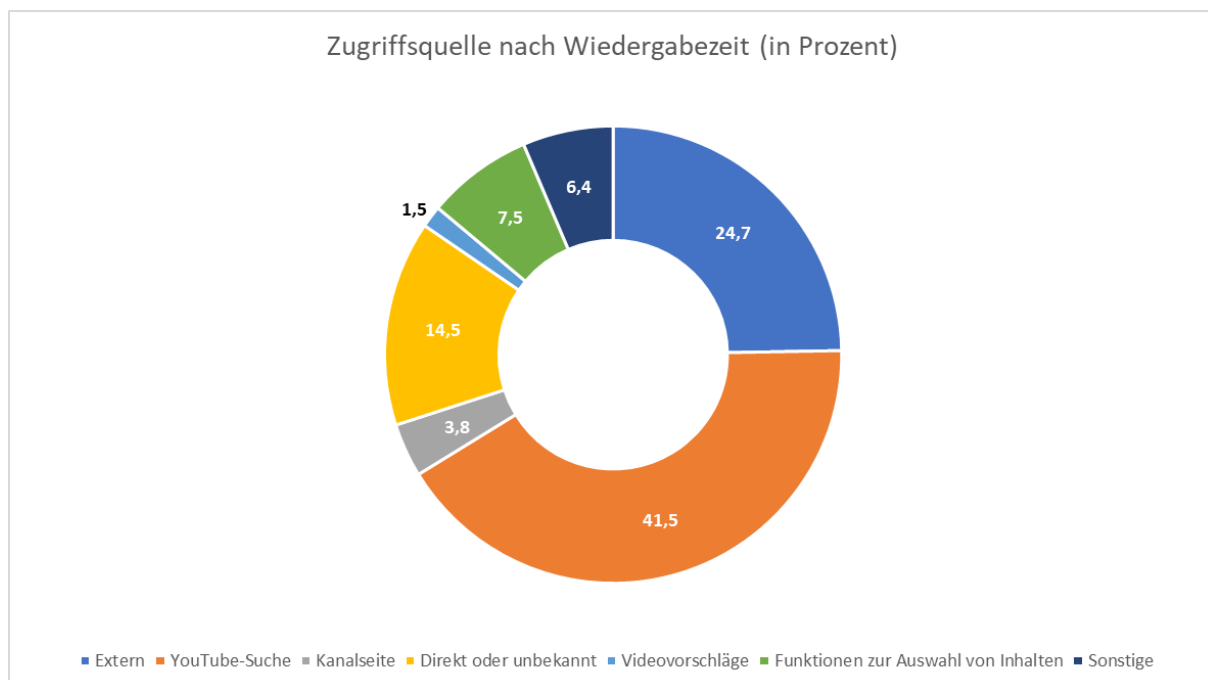


Abbildung VI.9: Zugriffsquellen des Kanals nach Wiedergabezeit in Prozent. N = 79 Stunden. Extern (blau): 19,5h, 24,7%; YouTube-Suche (orange): 32,8h, 41,5%; Kanalseite (grau): 3h, 3,8%; Direkt oder unbekannt (gelb): 11,5h, 14,5%; Videovorschläge (hellblau): 1,2h, 1,5%; Funktionen zur Auswahl von Inhalten (grün): 5,9h, 7,5%; Sonstige (dunkelblau): 5,1h, 6,4%.

Während die meisten Videos von einer externen Quelle aufgerufen werden (40,1%), wurde jedoch nur 24,7% der gesamten Wiedergabezeit aus dieser Quelle generiert. 41,5% der

Wiedergabezeit ergibt sich aus Aufrufen aus der YouTube-Suche (bei lediglich 19% aller Aufrufe).

VI.2.2.2. Suchbegriffe

Eine der Kernfunktionen von YouTube ist die Möglichkeit, Videos anhand von Schlagwörtern zu suchen. Diese müssen nicht zwangsweise von den Ersteller*innen der Videos eingegeben werden, da YouTube auch Videotitel und -beschreibungen in die Suche mit einbezieht. Die Suchbegriffe können dabei auf das Thema hindeuten („künstliche intelligenz“), aber auch auf andere Merkmale des Videos („tübingen“) sein. Die Suchbegriffe (Abb. VI.10) beziehen sich dabei auf die Aufrufzahlen (Abb. VI.8) und die Wiedergabezeit (Abb. VI.9) aus der Quelle „YouTube-Suche“, die 19% aller Kanalaufrufe und 41,5% der Wiedergabezeit generiert.

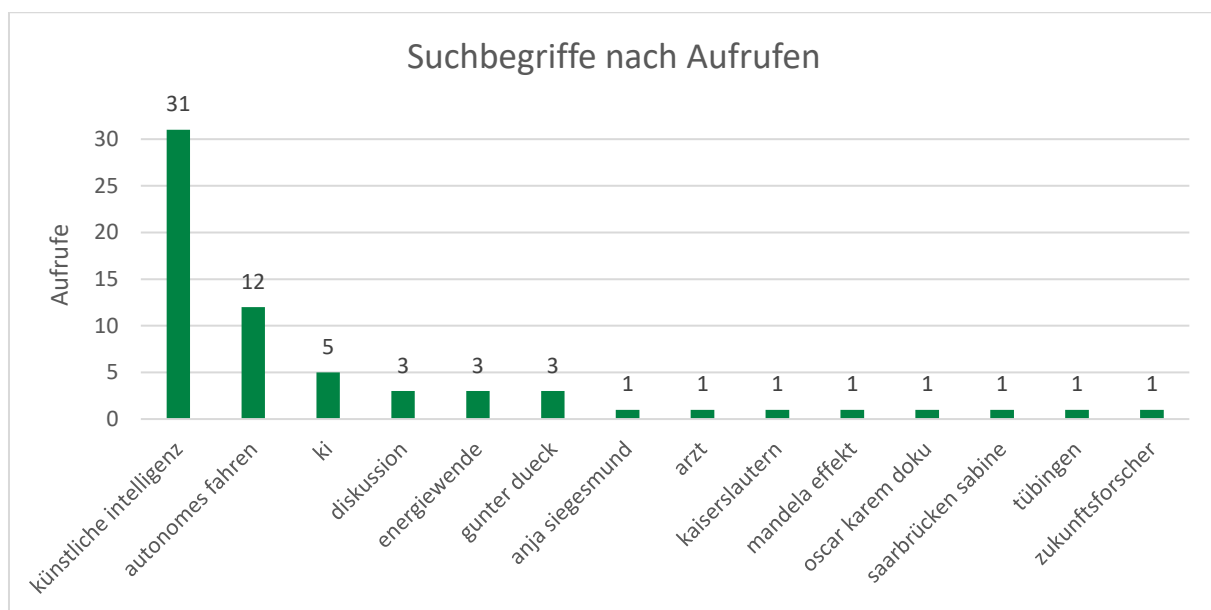


Abbildung VI.10: Suchbegriffe nach Aufrufen. N = 331 Aufrufe. Auflistung der Suchbegriffe, die in der YouTube-Suche eingegeben wurden und eines der Videos des Kanals als Ergebnis hatten, dass in der Folge angeklickt wurde.

Erwartbare Begriffe, die direkt auf das Projekt Hinweisen („NaWik“, „RIKI“, „Risikokommunikation“ etc.), finden sich nicht unter den von YouTube angegebenen Suchbegriffen. Dies zeigt, dass die Nutzer*innen, die über die YouTube-Suche auf den Kanal zugreifen, eher nach spezifischen Themen (besonders „künstliche intelligenz“) suchen als nach dem Projekt an sich.

VI.2.2.3. Externe Zugriffsquellen

YouTube bietet die Möglichkeit, die externen Zugriffsquellen aufzulisten, soweit diese Informationen vorliegen (Abb. VI.11). Dies kann zum Beispiel bei direkte Links und Einbettungen der Videos der Fall sein.

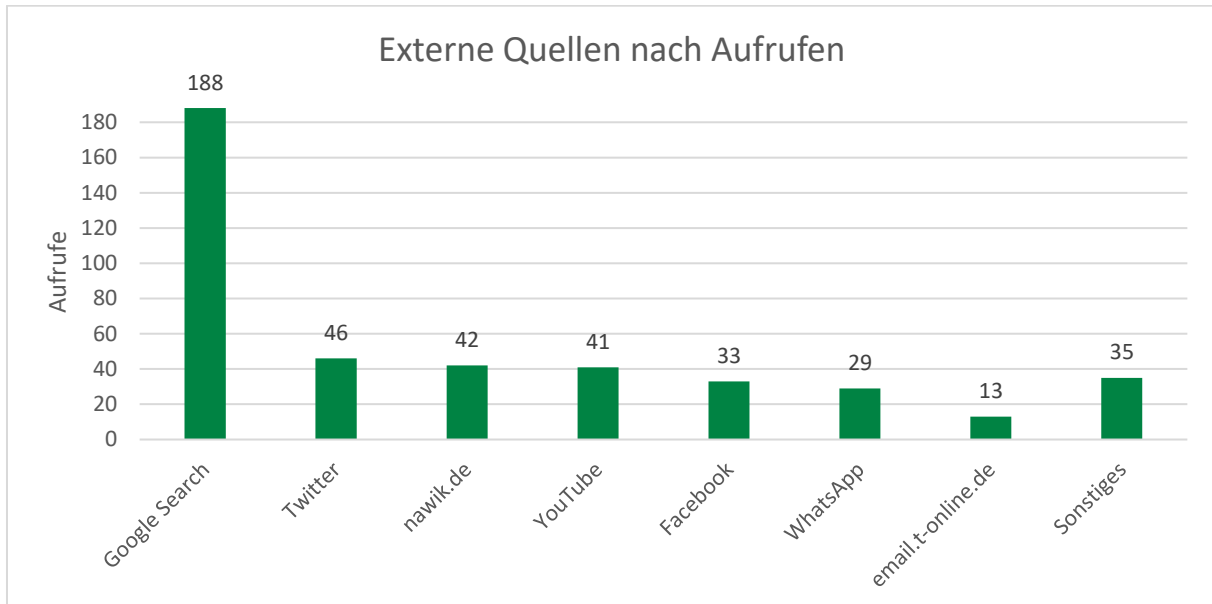


Abbildung VI.11: Externe Zugriffsquellen nach Aufrufen. N = 698. Auflistung der externen Websites, von denen die Videos auf dem RIKI-Kanal aufgerufen wurden.

Dies beinhaltet sowohl Suchmaschinen (Google Search), soziale Medien (Twitter, Facebook) und Messenger-Apps (WhatsApp) sowie andere Websites (nawik.de). Die Quelle „YouTube“ wird angegeben, da in einigen Fällen eine genaue externe Quelle nicht ermittelt werden kann, das Video aber über den (eingebetteten) YouTube-Player abgespielt wurde.

VI.2.2.4. Weitere demographische Daten

Für den Kanal „Risikokommunikation zur KI“ stehen auf Grund der geringen Aufrufzahlen und Abonnenten sowie der vornehmlich externen Zugriffsquellen keine Daten zu Alter und Geschlecht der Zuschauer*innen zur Verfügung.

92,4% der Wiedergabezeit des Kanals (73 Stunden) werden von Nutzer*innen ohne Abonnement generiert, die restlichen 7,6% (6 Stunden) durch Nutzer*innen mit Abonnement.

583 (33,5%) der Aufrufe stammen aus Deutschland, soweit diese zurückverfolgt werden konnten. Für 1157 Aufrufe (66,5%) kann YouTube keinen Aufrufort zur Verfügung stellen.

98,9% (1720 Aufrufe) der Videos werden ohne Untertitel angesehen, 1,1% (20 Aufrufe) mit deutschem Untertitel. Dabei handelt es sich um von YouTube automatisch erzeugte Untertitel.

VI.2.3. Entwicklung

Neben dem Ist-Zustand des Kanals wurde auch die Entwicklung des Kanals über den betrachteten Zeitraum untersucht. Dies soll beispielhaft an den Aufrufzahlen im Vergleich zu den Video-Veröffentlichungen im Rahmen von RIKI deutlich gemacht werden. Dazu wurden besonders hervorstechende Anstiege von Aufrufzahlen („Spikes“, über 25 Aufrufe am Tag) mit den an oder nahe an diesem Datum veröffentlichten Videos markiert (Abb. VI.12). Eine vollständige Übersicht, wann welches Video veröffentlicht wurde und wie sich dies auf die Aufrufzahlen ausgewirkt hat, befindet sich in Anhang 1 (Kapitel VI.4.1).

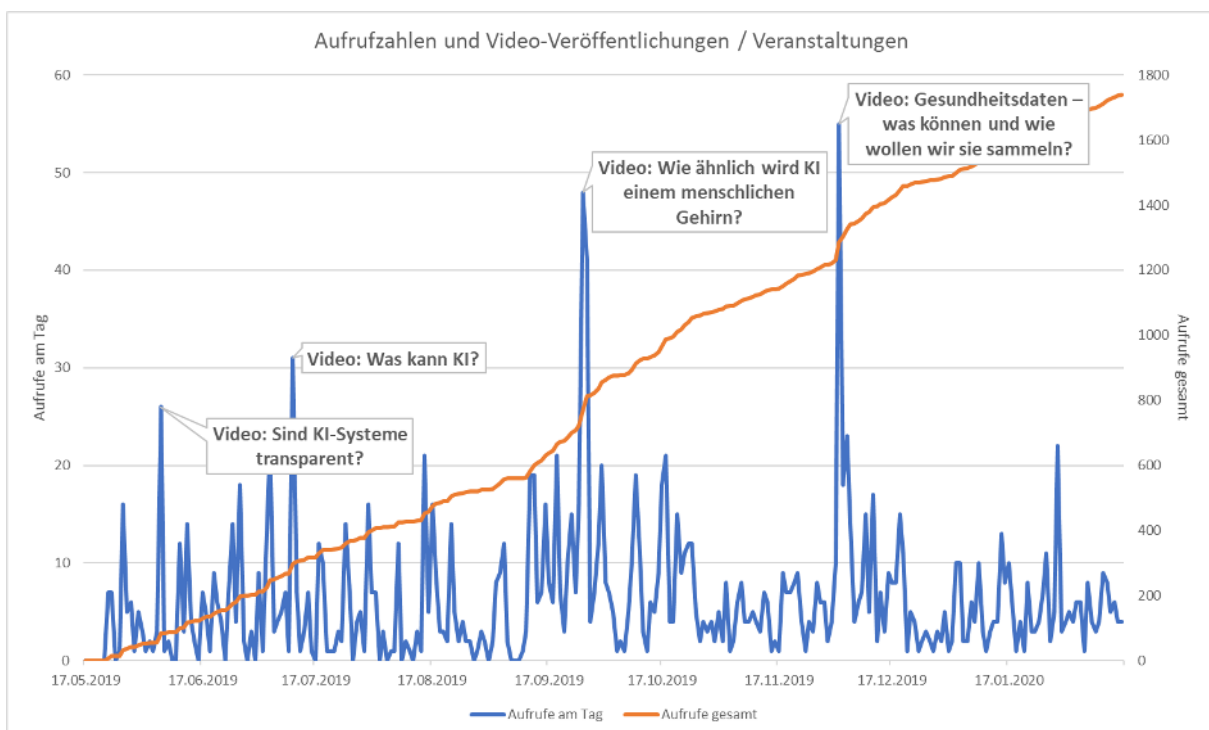


Abbildung VI.12: Aufrufzahlen der Videos des Kanals „Risikokommunikation zur KI“ im Vergleich zu den Veröffentlichungsterminen der Videos und Veranstaltungen des Projekts. Markiert sind Anstiege über 25 Aufrufe am Tag. Für den Spike am 05.07.2019 wurde das darauffolgende Ereignis markiert.

Die vier größten Spikes mit über 25 Aufrufen am Tag ergaben sich zu sehr speziellen und anthropozentrischen Themen („Wie ähnlich wird KI einem menschlichen Gehirn“ und „Gesundheitsdaten – was können und wie wollen wir sie sammeln?“) und grundlegenden technischen Fragen („Sind KI-Systeme transparent?“ und „Was kann KI?“).

VI.2.4. Detailbetrachtung der Metriken

VI.2.4.1. Top-5-Videos

Um einen tieferen Einblick in das Leistungsverhalten der Videos zu bekommen, werden die Top-5-Videos verschiedener Kategorien betrachtet (Tabelle VI.1). Dabei ist zu beachten, dass die Videos verschiedene Laufzeiten und Längen haben. Auf Grund der wenigen generierten Abonnenten und „Mag ich“/ „Mag ich nicht“-Angaben ist eine Top-5-Aufstellung in diesen Kategorien nicht sinnvoll. Im Untersuchungszeitraum vom 16.05.2019 bis zum 16.02.2020 entfallen die meisten „Mag ich“-Angaben auf das Video „Wie ähnlich wird KI einem menschlichen Gehirn?“ (9 „Mag ich“-Angaben, Veröffentlicht: 26.09.2019, Videolänge: 1,07 min). Die meisten „Mag ich nicht“-Angaben erhielt „Wissen wir, wie Algorithmen Entscheidungen treffen?“ (3 „Mag ich nicht“-Angaben, Veröffentlicht: 10.10.2019, Videolänge: 1,83 min). Insgesamt konnten elf Videos jeweils einen Abonnenten für den Kanal gewinnen (Tabelle VI.2). Die restlichen sieben Abonnenten stammen aus unbekannter Quelle.

Tabelle VI.1: Top-5-Videos in den Kategorien Aufrufe, durchschnittliche Wiedergabezeit und Wiedergabezeit gesamt im betrachteten Zeitraum vom 16.05.2019 bis zum 16.02.2020). Bei den Ergebnissen ist die jeweilige Laufzeit (ab Veröffentlichungsdatum) sowie die Länge des Videos zu beachten.

Top 5 nach Aufrufen	Videotitel	Veröffentlichungszeitpunkt des Videos	Videolänge (min)	Aufrufe
1	Reden Sie mit – Autonomes Fahren und Künstliche Intelligenz	12.09.2019	01:21:21	211
2	Autonomes Fahren – die Schuldffrage	23.05.2019	00:02:18	118
3	Sind KI-Systeme transparent?	06.06.2019	00:00:59	103
4	Werden Lehrer ihre Schüler mit KI überwachen?	22.08.2019	00:03:10	87
5	Bedroht KI Arbeitsplätze?	19.09.2019	00:02:17	87
Top 5 nach Durchschnittlicher Wiedergabedauer				Durchschnittliche Wiedergabedauer
1	Reden Sie mit – Autonomes Fahren und Künstliche Intelligenz	12.09.2019	01:21:21	00:14:51
2	Künstliche Intelligenz in der Medizin – Reden Sie mit!	05.12.2019	01:30:29	00:05:14
3	Werden Lehrer ihre Schüler mit KI überwachen?	22.08.2019	00:03:10	00:02:07
4	Wissen wir, wie Algorithmen Entscheidungen treffen?	10.10.2019	00:01:50	00:01:16
5	Bedroht KI Arbeitsplätze?	19.09.2019	00:02:17	00:01:11
Top 5 nach Wiedergabezeit gesamt				Wiedergabezeit (Stunden)
1	Reden Sie mit – Autonomes Fahren und Künstliche Intelligenz	12.09.2019	01:21:21	52,2
2	Künstliche Intelligenz in der Medizin – Reden Sie mit!	05.12.2019	01:30:29	6,6
3	Werden Lehrer ihre Schüler mit KI überwachen?	22.08.2019	00:03:10	3,1
4	Autonomes Fahren – die Schuldffrage	23.05.2019	00:02:18	2,1
5	Bedroht KI Arbeitsplätze?	19.09.2019	00:02:17	1,7

Tabelle VI.2: Gewonnene Abonnenten. Die in der Tabelle aufgelisteten Videos haben im Untersuchungszeitraum jeweils einen Abonnenten für den Kanal gewonnen. Die restlichen der insgesamt 18 Abonnenten stammen aus unbekannter Quelle.

Videotitel	Veröffentlichungszeitpunkt des Videos	Videolänge (min)	Gewonnene Abonnenten
Kann man einem Roboter alles befehlen?	28.05.2019	00:01:15	1
Sind KI-Systeme transparent?	06.06.2019	00:00:59	1
Wird KI den Menschen automatisch klassifizieren?	18.06.2019	00:02:45	1
Was kann KI?	11.07.2019	00:00:31	1
Kann man KI missbrauchen?	18.07.2019	00:00:31	1
Wie gut funktioniert automatisiertes Fahren heute?	08.08.2019	00:00:49	1
Bedroht KI Arbeitsplätze?	19.09.2019	00:02:17	1
Wie ähnlich wird KI einem menschlichen Gehirn?	26.09.2019	00:01:04	1
Wird KI eines Tages klüger sein als der Mensch?	01.10.2019	00:00:52	1
Führt KI zum Überwachungsstaat?	24.10.2019	00:01:18	1
Gesundheitsdaten – was können und wie wollen wir sie sammeln	03.12.2019	00:00:59	1

Es zeigt sich, dass die Metriken Aufrufe, durchschnittliche Wiedergabezeit und Gesamtwiedergabezeit unabhängig voneinander sind. Es scheint einen mindestens schwachen Zusammenhang zwischen den Aufrufzahlen und der Gesamtwiedergabezeit zu geben, da vier der fünf Videos in beiden Top-5-Listen übereinstimmen.

VI.2.4.2. Einzelbetrachtung eines Videos

Wie genau das Leistungsverhalten der Videos auf dem Kanal ist, soll anhand eines Beispiel-Videos untersucht werden. Dazu dient das Video „Autonomes Fahren – die Schuldfrage“ (YouTube-Link <https://youtu.be/fDf58xaVBCC>). Das Video war das erste, das auf dem Kanal veröffentlicht wurde (Veröffentlichungsdatum: 23.05.2019) und hatte bis zum Ende des Untersuchungszeitraums 118 Aufrufe. Es hat eine Abspieldauer von 2,3 Minuten (entspricht 2:18 Minuten). Die Metriken des Videos geben sehr gut die Leistung eines durchschnittlichen Videos auf dem Kanal wieder. Dies beinhaltet beispielsweise starke Aufrufzahlen bei der Veröffentlichung, die dann langsam abfallen. Auch der Verlauf der Zuschauerbindung und der Wiedergabezeit kann für die Videos des Kanals als exemplarisch gelten. Das Format und der Inhalt sind ebenfalls ein repräsentatives Beispiel für die anderen Veröffentlichungen auf dem Kanal – mit Ausnahme der beiden Videos der Fishbowl-Debatten.

Neben der Entwicklung der Aufrufe (Abb. VI.13), der Wiedergabezeit gesamt (Abb. VI.14) und dem durchschnittlichen Prozentsatz der Wiedergabedauer (Abb. VI.15) wird auch die Zuschauerbindung betrachtet (Abb. VI.16). Die Wiedergabezeit gesamt gibt an, wie lange das Video insgesamt an einem Tag gesehen wurde, der durchschnittliche Prozentsatz der Wiedergabedauer hingegen, wie viel des Videos im Schnitt pro Aufruf angesehen wurde. Als Zuschauerbindung bezeichnet YouTube, wie und wie lange sich Zuschauer*innen das Video ansehen.

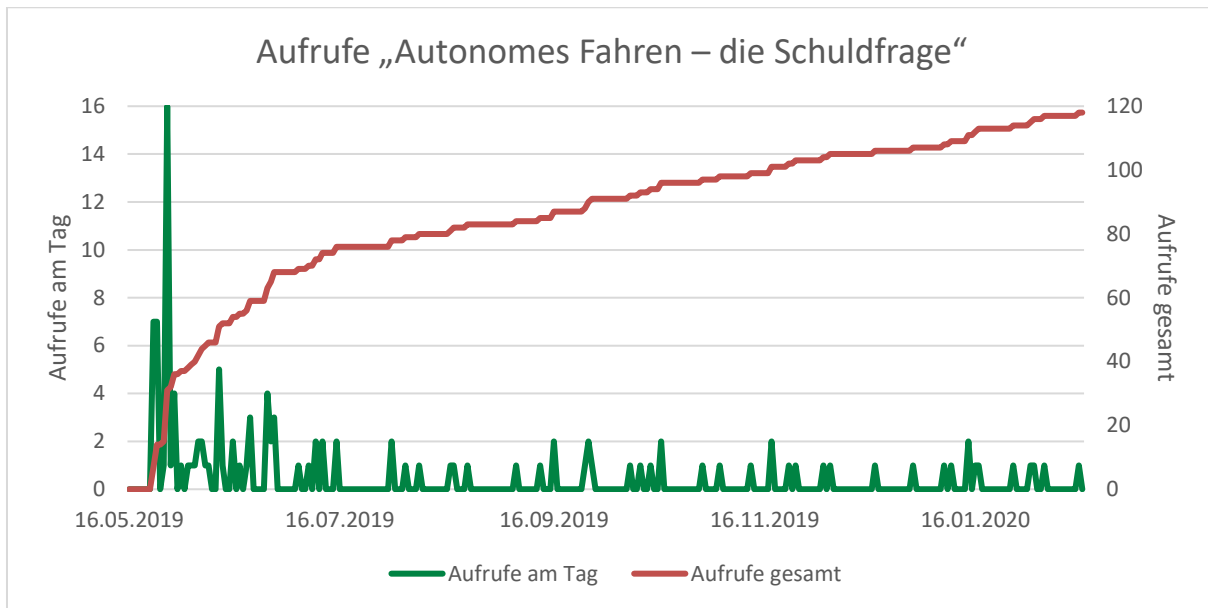


Abbildung VI.13: Entwicklung der Aufrufe des Videos „Autonomes Fahren – die Schuldfrage“. N = 118. Abgetragen sind die Aufrufe am Tag (grün, linke y-Achse) und die gesamten Aufrufe (rot, rechte y-Achse). Die meisten Aufrufe verzeichnete das Video kurz nach Veröffentlichung, danach fallen die Aufrufe stark ab.

Das Video wurde am häufigsten in der ersten Woche nach der Veröffentlichung aufgerufen (Abb. VI.13, Aufrufe am Tag), danach wurde das Video seltener und sporadischer aufgerufen.

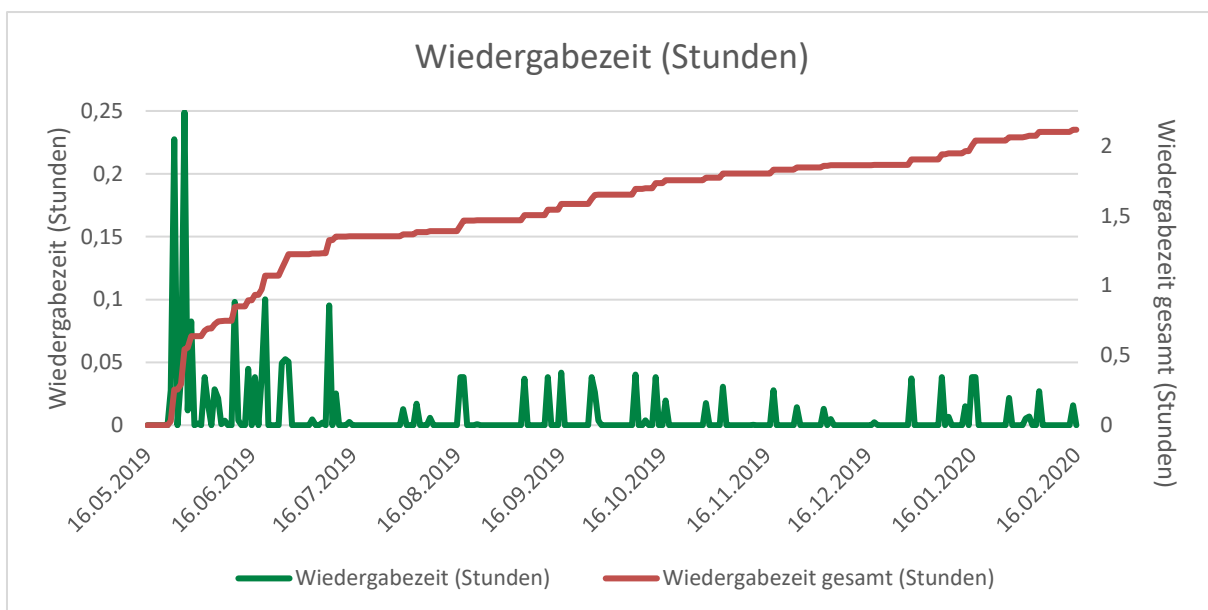


Abbildung VI.14: Entwicklung der Wiedergabezeit des Videos „Autonomes Fahren – die Schuldfrage“. Abgetragen ist die tägliche Wiedergabezeit (in Minuten, grün, linke y-Achse) und die gesamte Wiedergabezeit (in Minuten, rot, rechte y-Achse). Kurz nach der Veröffentlichung erreichte das Video eine hohe Wiedergabezeit pro Tag, die im Laufe der Zeit weniger wird. Dies ähnelt dem Verlauf der Aufrufe.

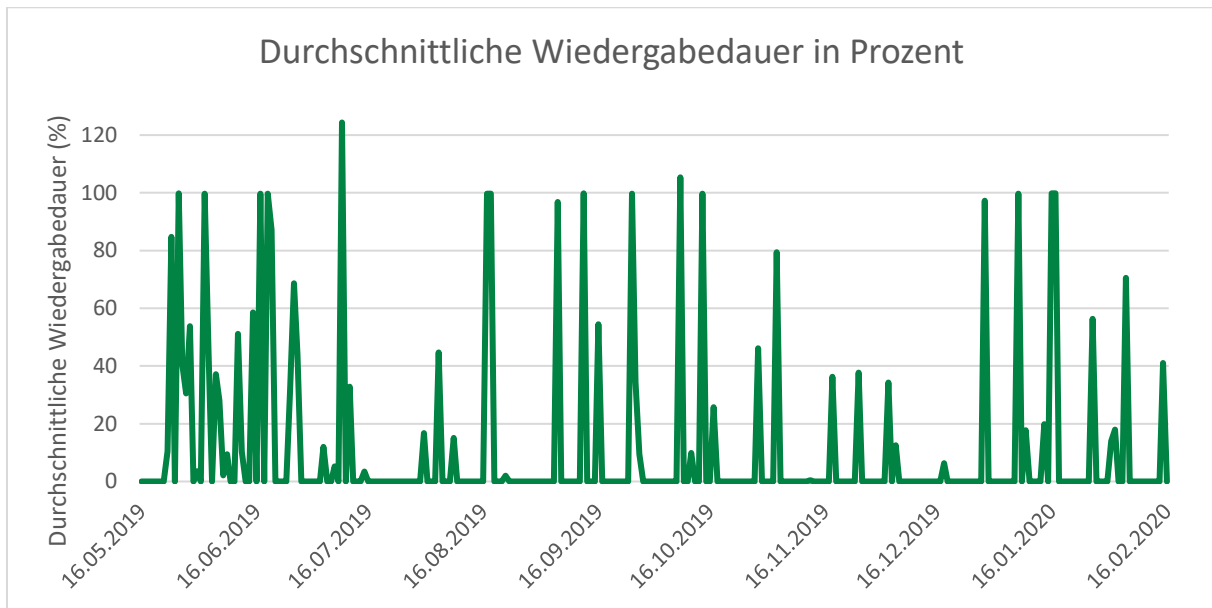


Abbildung VI.15: Durchschnittlicher Prozentsatz der Wiedergabedauer des Videos „Autonomes Fahren – die Schuldfrage“. Der Wert gibt an, wie viel des Videos in Prozent an dem Tag pro Aufruf betrachtet wurde. Werte über 100% ergeben sich, wenn Nutzer*innen Teile des Videos mehrfach ansehen, zum Beispiel durch zurückspulen.

Der Verlauf Gesamtwiedergabedauer des Videos pro Tag (Abb. VI.14) ähnelt dem Verlauf der Aufrufzahlen (Abb. VI.13). Wenn ein Video oft aufgerufen wird, werden auch mehr Minuten an Wiedergabezeit generiert. Wie bei den Aufrufen sieht man, dass dieses vor allem nach der Veröffentlichung des Videos vermehrt wiedergegeben wurde. Der durchschnittliche Prozentsatz der Wiedergabe (Abb. VI.15) zeigt, dass das Video in vielen Fällen zu 100% angesehen wurde.

Die drei Metriken im Vergleich zeigen, dass das Video vor allem kurz nach der Veröffentlichung besonders häufig aufgerufen wurde, entsprechend hoch ist die Wiedergabezeit in der ersten Woche.

Die Zuschauerbindung ist besonders in den ersten Sekunden des Videos sehr groß, fällt dann aber in den ersten 30 Sekunden stark ab, bevor sie bis zum Ende des Videos weiter, jedoch flacher, abnimmt. 26,09% aller Nutzer*innen sehen das Video bis zum Ende an (Abb. VI.16).

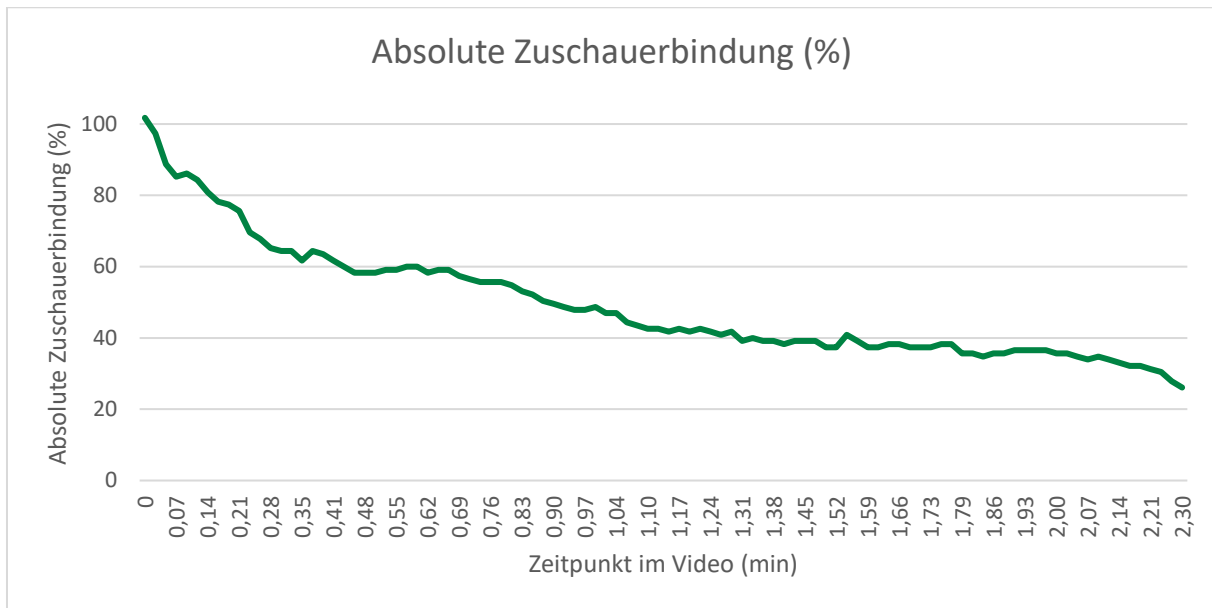


Abbildung VI.16: Absolute Zuschauerbindung des Videos „Autonomes Fahren – die Schuldfrage“. Das Schaubild zeigt in Prozent wie häufig der jeweilige Zeitpunkt des Videos angesehen wurde, bezogen auf die gesamten Wiedergaben des Videos. Ausschläge nach oben zeigen, dass Nutzer*innen zu diesem Zeitpunkt direkt gesprungen sind (etwa per Link), Einbrüche nach unten, dass Teile des Videos übersprungen wurden.

VI.3. Zusammenfassung und Fazit

Die Auswertung des YouTube-Kanals „Risikokommunikation zur KI“ zeigt, dass die Videos vor allem von am Thema interessierten Nutzer*innen angesehen werden. Obwohl fast die Hälfte der Nutzer*innen von externen Seiten auf die Videos zugreift, schauen die durch die YouTube-Suche auf den Kanal gekommenen Nutzer*innen die Videos doppelt so lange an.

Die verwendeten Suchbegriffe zeigen, dass die Nutzer*innen nicht primär nach RIKI suchen, sondern an generellen Themen zu KI interessiert sind. Dies zeigt sich auch an den Themen der am häufigsten und am längsten betrachteten Videos des Kanals. Ein Ausreißer hierbei ist das Video der Fishbowl-Debatte in Karlsruhe, das auch bedingt durch seine überdurchschnittliche Länge hohe Wiedergabezeiten generiert hat.

Die Analyse der durchschnittlichen Wiedergabedauern und die exemplarische Betrachtung der Nutzerbindung zeigen, dass die Zuschauer*innen relativ schnell das Interesse an den Videos verlieren können, was damit zusammenhängen mag, dass die gesuchten Begriffe, die zum Aufruf eines RIKI-Videos führen, häufig unspezifisch sind, und somit stets die Gefahr besteht, dass ein angeklicktes Video nicht den tatsächlichen Erwartungen der Nutzer*innen entspricht. Die Videos konnten keine große Anzahl an Abonnenten generieren. Dies lässt darauf schließen, dass zwar die Einzelthemen zum Themenfeld KI interessant sind, das Thema „Risikokommunikation“ im Allgemeinen jedoch weniger. Erst seit September, also etwa nach

fünf Monaten, konnte der Kanal eine höhere Zahl von Aufrufen und wiedergegebenen Minuten verzeichnen. Die meisten Aufrufe der einzelnen Videos werden kurz nach deren Veröffentlichung erzeugt.

Auch wenn der Kanal bis jetzt nur wenige Abonnenten verzeichnen konnte, so zeigen die Aufrufe vor allem von der Webseite des RIKI-Projekts, aber auch über die YouTube-Suche, dass in der Öffentlichkeit ein Interesse am Thema KI generell und in etwas geringerem Maße auch an der Risikokommunikation dazu besteht. Besonders die Möglichkeit, die öffentliche Fishbowl-Diskussion in Karlsruhe zum Thema Autonomes Fahren und Künstliche Intelligenz nachzuverfolgen, wird von vielen Zuschauer*innen genutzt.

VI.4. Anhänge

VI.4.1. Anhang 1: Veröffentlichungsdaten der Videos und Aufrufe

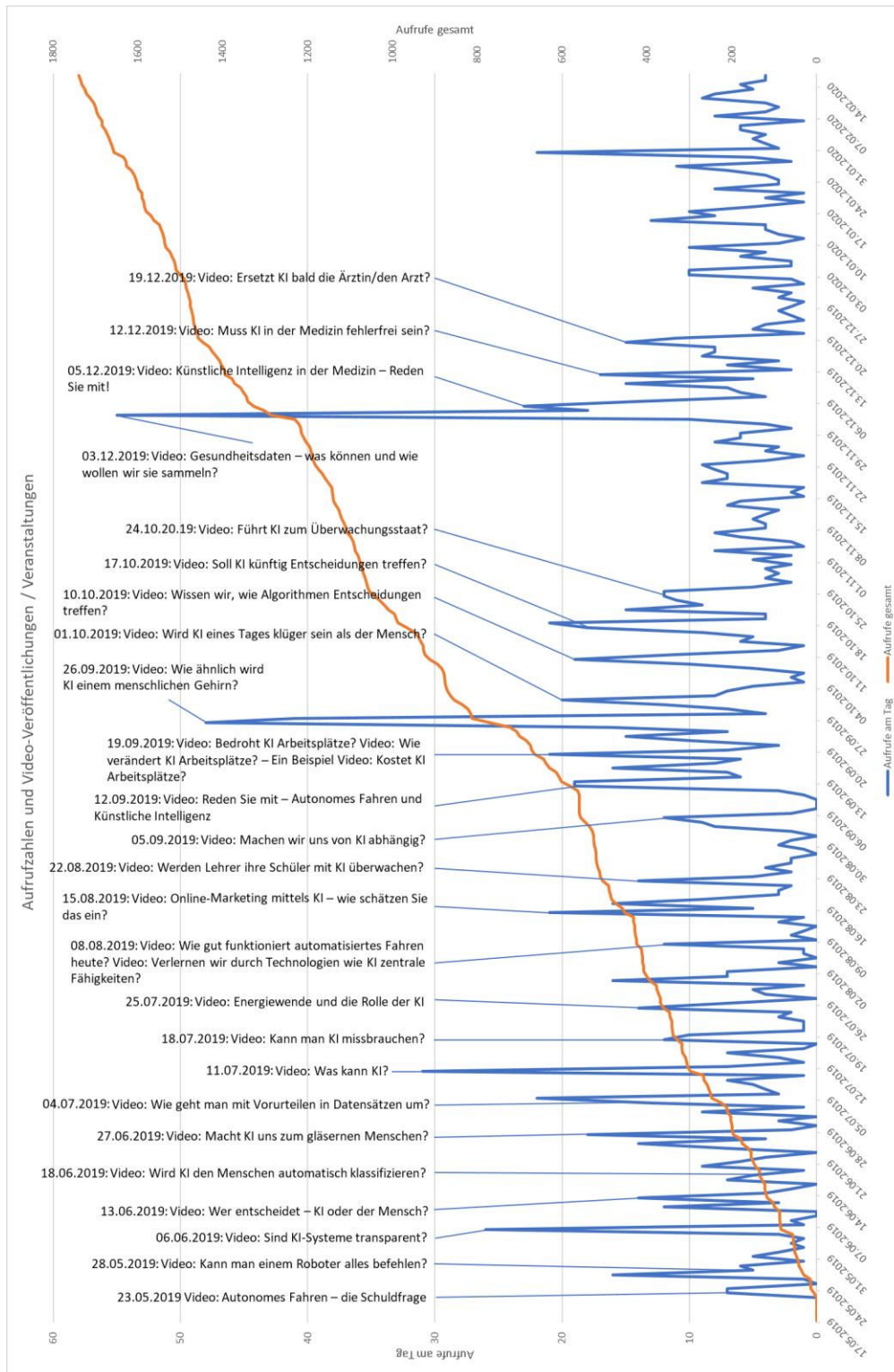


Abbildung VI.17: Anhang 1, Veröffentlichungsdaten der Videos und Veranstaltungen im Vergleich zu den Aufrufzahlen (blau) und der Summe aller Aufrufe (orange). Es zeigt sich, dass eine Veröffentlichung eines Videos zu einem Anstieg an Aufrufen an diesem Tag führt.

VI.4.2. Anhang 2: Kommentare

Kommentar 1:

Video: „Was kann KI?“¹¹

Gepostet am: Januar 2020

Nutzer*in: o0o0o0sam0o0o0o

Wortlaut:

Ja, aber was kann KI?

Kommentar 2:

Video: „Energiewende und die Rolle der KI“¹²

Gepostet am: August 2020

Nutzer*in: Werner Müller

Wortlaut:

Das kWh von meiner über 20 Jahre alten PV-Anlage verursacht kostet von unter 0 Cent/kWh und es kommen immer noch leicht über 1000 kWh von jedem kWp im Jahr. Wir haben so ca. von Sonnenaufgang bis Sonnenuntergang, also an 3500 bis 4000 Stunden im Jahr habe ich PV-Strom.

Das kann kein Kernfusionsreaktor kein Kernkraftwerk oder Kohlekraftwerk oder auch Gaskraftwerk zu dem Preis an meine Steckdose liefern.

Solarstrom ist für den Bürger und auch Betrieb die günstigste Energiequelle, aber nicht die einzige.

¹¹ <https://www.youtube.com/watch?v=-Ai3RtVQuwc>, Abgerufen: 18.02.2020.

¹² <https://www.youtube.com/watch?v=nEtaFzgUTKU>, Abgerufen: 18.02.2020.

VII. Evaluation des Kompendiums

„Risikokommunikation zur Künstlichen Intelligenz“

VII.1. Hintergrund

Ein wesentliches Ziel des Projekts *RIKI* war es, im Bereich der Künstlichen Intelligenz (KI) Forschende für kritische Themen zu sensibilisieren und ihnen Kommunikationskompetenzen im Umgang mit der Öffentlichkeit zu vermitteln. Dazu wurden deutschlandweit sechs Kommunikationsworkshops für Nachwuchswissenschaftler*innen aus diesem Forschungsfeld veranstaltet, bei denen sie ermutigt und unterstützt wurden, sich in die öffentliche Debatte rund um ihr Forschungsthema einzubringen. Neben den Workshops wurde die Veranstaltungsreihe des Projekts *RIKI* durch zwei öffentliche Diskussionsrunden ergänzt, die im Format von Fishbowl-Debatten stattfanden.

Die Erkenntnisse aus diesem einjährigen Kommunikationsprojekt sind im Kompendium „Risikokommunikation zur Künstlichen Intelligenz“ zusammengefasst. Das Kompendium richtet sich an Forschende im Bereich KI, denen es als Handreichung zur gesamtgesellschaftlichen Auseinandersetzung mit kritischen Fragestellungen in der KI dienen soll, und an Kommunikatorinnen und Kommunikatoren von KI-Themen, die diese Kommunikation begleiten und mitgestalten.

Um dieses Kompendium zu evaluieren, wurden mit Forschenden und Kommunikator*innen aus dem Feld der Künstlichen Intelligenz Expert*inneninterviews geführt, deren Ergebnisse im Folgenden dargestellt werden.

VII.2. Methodisches Vorgehen

Zur Evaluierung des Kompendiums „Risikokommunikation zur Künstlichen Intelligenz“ wurden mit 16 Expert*innen telefonisch Leitfadeninterviews durchgeführt. Die Zielgruppe des Kompendiums setzt sich aus Forschenden und Kommunikator*innen im KI-Bereich zusammen. Daher wurden aus beiden Gruppen Expert*innen befragt. Einerseits elf Wissenschaftler*innen mit verschiedenen Forschungsschwerpunkten und auf unterschiedlichen Karrierestufen und andererseits fünf Wissenschaftskommunikator*innen mit unterschiedlicher Berufserfahrung aus der Unternehmens- und Institutskommunikation. Das Ungleichgewicht bei der Zahl der Befragten in den beiden Zielgruppen erklärt sich durch Rekrutierungsschwierigkeiten aufgrund der Covid19-Pandemie. Der Leitfaden bestand aus 14 Leitfragen und weiterführenden Fragen zum Kompendium, mit denen die Expert*innen zur Qualität der Inhalte, der Relevanz des Kompendiums und zu ihrem Eindruck von der gestalterischen

Umsetzung befragt wurden (Vergleich Anhang 1 in Kapitel VII.5.1 für den vollständigen Leitfaden). Die Fragen lassen sich 13 Kategorien zuordnen, die im Folgenden aufgeführt sind.

Tabelle VII.1: Struktur der Befragung anhand verschiedener Kategorien

Inhalt	Leitfragen
Thematische Grundlagen	Haben Sie das Kapitel Risikokommunikation genutzt? Haben Sie das Kapitel zu KI-Definitionen genutzt?
Themenauswahl	Wo hatten Sie Schwerpunkte bei der Nutzung?
Qualität der Themenschwerpunkte	Wie bewerten Sie die Qualität in dem Themenbereich, der Ihrem Arbeitsalltag am nächsten steht?
Einordnung	Entspricht der Überblick, der gegeben wird, Ihrem Eindruck vom Thema?
Richtigkeit	Wird die Darstellung der Komplexität gerecht?
Aktualität	Entspricht der Überblick, der gegeben wird, Ihrem Eindruck vom aktuellen Stand der Debatte?
Relevanz	Leitfragen
Risikokommunikation im Arbeitskontext	Wie relevant ist das Thema „Risikokommunikation zur Künstlichen Intelligenz“ für Ihren Arbeitskontext?
Nutzung des Kompendiums für die Risikokommunikation	Ist das Kompendium etwas, das Sie verwenden werden? Würden Sie es weiterempfehlen bzw. weitergeben?
Nutzen des Kompendiums über die Arbeit hinaus	Werden Aspekte behandelt, die schon außerhalb des Arbeitskontextes diskutiert wurden?
Seriosität	Ist für Sie relevant, dass das Kompendium aus Bundesmitteln des BMBF finanziert wurde? Was verbinden Sie mit dem NaWik?
Einfluss darauf, Forschende für KI-Kommunikation zu begeistern	Inwieweit kann das Kompendium dazu beitragen, dass Forschende mehr Kommunikation über KI in der Öffentlichkeit, auch in den Medien, wagen?
Umsetzung	Leitfragen
Multimedialität	Haben Sie die Online-Inhalte in irgendeiner Form genutzt? Hatte diese Verknüpfung einen Mehrwert für Sie?
Gestaltung	Was sagen Sie zum Erscheinungsbild des Kompendiums?
Weitere Aspekte	Haben Sie noch Anmerkungen zum Kompendium?

VII.3. Ergebnisse

VII.3.1. Inhalt

Das Kapitel *KI-Definitionen* und das Kapitel *Risikokommunikation* des Kompendiums vermitteln thematische Grundlagen, auf die die fünf darauffolgenden Themenschwerpunkte aufbauen. Sowohl die Wissenschaftler*innen als auch die Kommunikator*innen wurden zu ihrer Nutzung dieser Kapitel befragt.

Alle 16 Befragten haben das Kapitel *Risikokommunikation* (S.12-19) genutzt. Es wurde mehrheitlich als nützlicher Überblick empfunden. Allerdings wiesen sowohl zwei der Kommunikator*innen als auch zwei der Forschenden darauf hin, dass bekanntes Wissen wiederholt wurde und sie wenig bis nichts neues erfuhren, sondern eher ihr Basiswissen auffrischen konnten. Vier der Interviewten gaben an, das Kapitel nur überflogen zu haben, da sie sich entweder nicht angesprochen gefühlt (2 Befragte) oder für andere Kapitel stärker interessiert hätten (2 Befragte). Der folgenden Aussage eines Forschenden stimmen elf der 16 Probanden inhaltlich zu:

*„Das waren Sachen, die man allgemein weiß, aber es war interessant, das so konkret zu lesen.“
(Interview Nr. 11)*

Eine Kommunikatorin äußerte, sich zwar mit den Inhalten der Risikokommunikation zu identifizieren, ihre Arbeit aber nie als Risikokommunikation bezeichnet zu haben. Sie war bisher wenig mit dem Konzept in Kontakt gekommen und fand die Bezeichnung einleuchtend.

Das Kapitel *KI-Definitionen* erfuhr noch mehr Zuspruch als das zur Risikokommunikation. Es wurde ebenso von allen 16 Befragten gelesen, wobei zwei erneut angaben, es nur überflogen zu haben. Sowohl Forschende als auch Kommunikator*innen gaben an, es als besonders nützlich empfunden zu haben:

„Wir haben intern für unsere Kommunikation versucht, festzuzurren, was die einzelnen Begriffe sind, sodass man ähnlich kommuniziert. Daher fand ich das sehr aufschlussreich als Überblick über die verschiedenen Definitionen.“ (Interview Nr. 14)

Fünf Themen aus dem Bereich der Künstlichen Intelligenz werden im Kompendium besonders beleuchtet und unter den Aspekten der medialen Darstellung, dem Stand der Forschung und der Risikokommunikation spezifisch im jeweiligen Themenbereich dargestellt. Diese fünf Themenbereiche sind: KI und Umgang mit Daten, KI und Arbeit, Autonomes Fahren, KI in der Medizin sowie KI und Robotik.

Alle Befragten stimmten zu, dass der im Kompendium gelieferte Überblick ihrem Eindruck des Standes der Debatte entspricht und die relevanten Themenbereiche der KI in ausreichendem Umfang dargestellt wurden. Zweimal wurde angemerkt, dass das Thema Datensicherheit noch ausführlicher hätte behandelt werden können.

„Das Kompendium ist sehr gut gegliedert, vor allem wie die Anwendungsbereiche voneinander abgegrenzt werden. Wenn man z. B. im Bereich Automotive Security forscht, kann man sich die acht Seiten zum Autonomen Fahren anschauen.“ (Interview Nr. 13)

Die Qualität der Aufbereitung in jenen Themenschwerpunkten, die dem Arbeitsumfeld der Befragten am nächsten standen, wurde von allen als gut bewertet, wobei in einem Fall konkrete Kritik zur Ausgewogenheit und Belastbarkeit einer spezifischen Quelle geäußert wurde. Überwiegend empfanden die Befragten die Schwerpunktthemen im Kompendium jedoch als guten Überblick zum jeweiligen Bereich.

„Hatte Hand und Fuß, aber war trotzdem verständlich.“ (Interview Nr. 14)

Ebenso bescheinigten alle Befragten dem Kompendium, dass die Darstellung der jeweiligen Komplexität gerecht wird, und lobten in diesem Zusammenhang wiederholt die Themenauswahl sowie die gelungene Veranschaulichung.

Besonders die Kommunikationstipps wurden stark von den Forschenden und den Kommunikator*innen genutzt. Die meisten Befragten empfanden diesen Teil des Kompendiums als besonders wichtig.

„Das war für mich die Essenz.“ (Interview Nr. 6)

Für Forschende stellen sie konkrete Handlungsempfehlungen dar (acht Befragte), die sie auch als Nachschlagewerk oder zur Reflexion ihres eigenen Kommunikationsverhaltens nutzen möchten. Eine Kommunikatorin merkte an, dass die Tipps noch weiterer Bereiche abdecken könnten:

„Die Tipps bieten konkrete Beispiele, an die Forschende andocken können. Aber mir fehlt: Manche Kommunikationssituationen muss man verlassen, weil der Konsens nicht gegeben ist. Wie beende ich Kommunikation, die in die Aggression driftet?“ (Interview Nr. 3)

VII.3.2. Relevanz

Da das Kompendium unter anderem als Handreichung für potenziell kritische Kommunikationssituationen im Themenfeld KI konzipiert wurde, wurden die Interviewten explizit gefragt, ob sie das Kompendium in ihrem jeweiligen Arbeitskontext verwenden werden. Diejenigen Experten, die bereits selbst im Bereich Risikokommunikation tätig waren

(elf von 16 Befragten), waren sich einig, dass sie das Kompendium in Zukunft verwenden werden. Das Kompendium wurde dabei von sechs der Interviewten als guter Einstieg gerade für unerfahrene Kommunikator*innen beschrieben:

„Wenn ich zum Beispiel junge Mitarbeiter bei einem Tag der offenen Tür hätte, die etwas unsicher sind, aber in ihrem Thema gut. Dann würde ich sagen: Guck mal, nimm das mal und lies das Kapitel zu dem Thema durch.“ (Interview Nr. 3)

„Habe es auch schon einem Mitarbeiter gegeben, der hier angefangen hat. Guck da mal rein.“ (Interview Nr. 1)

Als weiterer wichtiger Aspekt wurde in diesem Kontext hervorgehoben, dass durch die Kommunikationstipps das Bewusstsein für Experten-Laien-Kommunikation verbessert werden könne:

„Man wird in der Wissenschaft immer darauf getrimmt, rational zu argumentieren. Aber gerade in öffentlichen Diskussionen, z. B. Podiumsdiskussionen, spielen Emotionen rein, und das sollte man sich vor Augen führen, bevor man sich in eine solche Situation begibt.“ (Interview Nr. 13)

Vier der Expert*innen meinten darüber hinaus, dass sie das Kompendium als Vorbereitungsmaterial heranziehen würden, wenn sie vor einer konkreten Kommunikationssituation stünden, um sich vor allem die Kommunikationsstrategien und Empfehlungen noch einmal ins Gedächtnis zu rufen. Zwei Forschende empfanden zusätzlich den Überblick über jene Themen hilfreich, in denen sie selbst keine Expert*innen sind, um sich auch auf weniger fachspezifische Fragen vorzubereiten.

Von den fünf Interviewten, die bisher zwar im Themenfeld arbeiten, aber selbst noch nicht öffentlich als Kommunikator*innen aufgetreten sind, empfanden drei die Kommunikationstipps im Kompendium trotzdem als sehr hilfreich mit Verweis auf private Kommunikationssituationen.

Die Hälfte der Interviewten wurde zudem gefragt, ob es für sie relevant war, dass das Projekt aus Mitteln des Bundesministeriums für Bildung und Forschung finanziert wurde. Sechs der Befragten formulierten, dass sie dies als positiv und als Qualitätsmarker empfanden. Das Gefühl, dass das Projekt durch eine solche Finanzierung unabhängig arbeiten könne, war dabei die häufigste Begründung. Die beiden anderen Befragten empfanden die Finanzierung durch das BMBF als neutral bzw. nicht relevant.

Das Nationale Institut für Wissenschaftskommunikation (NaWik) war vier der sechs dazu Befragten vorher überhaupt nicht bekannt. Zwei der Befragten kannten das NaWik vom Namen

her, verbanden aber, wie die Übrigen, weder etwas Positives noch etwas Negatives mit der Institution.

Zwölf der 16 Befragten wurde die Frage gestellt, ob sie das Kompendium weitergeben oder weiterempfehlen würden. Alle Befragten bejahten dies. Fünf Befragte hatten das Kompendium sogar bereits weiterempfohlen. Eine Person schränkte ihre Antwort dahingehend ein, dass sie es nur an extern Kommunizierende weiterempfehlen würde, aber nicht an ihre direkten Forschungskolleg*innen. Vier Interviewteilnehmer*innen konnten aus Zeitgründen nicht zu diesem Aspekt befragt werden.

Die Mehrheit der Interviewten (zwölf Befragte) ist überzeugt, dass das Kompendium dazu beitragen kann, mehr Kommunikation über KI in der Öffentlichkeit zu ermöglichen. Aspekte, die aus Sicht der Expert*innen dafür sprechen, beziehen sich darauf, dass das Kompendium zum einen ein Bewusstsein für die Relevanz von Kommunikation erzeugt und ihnen zum anderen durch die aufgeführten Kommunikationsstrategien Fähigkeiten und Selbstvertrauen vermitteln kann. Auch an dieser Stelle des Interviews wurde mehrfach darauf verwiesen, dass das Kompendium gerade jüngeren Forschenden oder sogar Studierenden den Einstieg in die (öffentliche) Kommunikation zum Thema KI vereinfachen könnte. Von denjenigen Befragten, die nicht überzeugt waren, dass das Kompendium zu mehr Kommunikation beiträgt, waren drei der Meinung, dass gerade öffentliche Kommunikation auch eine „Typfrage“ sei.

VII.3.3. Umsetzung

Alle Befragten hatten einen guten bis sehr guten generellen Eindruck von dem Kompendium, empfanden das Design als gelungen (sechs Befragte) und lobten vor allem dessen Struktur (fünf Befragte). Des Weiteren wurde die Gestaltung als „ansprechend“ (zwei Befragte), „hochwertig“ (zwei Befragte) und „professionell“ (zwei Befragte) bezeichnet.

Die Online-Inhalte, auf die im Kompendium verwiesen wird, wurden von keinem Befragten intensiv genutzt. Drei der Interviewten gaben sogar an, dass ihnen diese gar nicht aufgefallen wären.

Am Ende des Leitfadeninterview hatten die Befragten die Gelegenheit, frei noch weitere Anmerkungen zum Kompendium zu geben. Der Großteil dieser Anmerkungen bezog sich auf die weitere Verbreitung des Kompendiums. Vier Befragte merkten an, dass das Kompendium auch für die breite Öffentlichkeit interessant bzw. nützlich wäre. So äußerte eine Kommunikatorin:

„Dem Kompendium ist die Gratwanderung sehr gut gelungen, Fachinformationen korrekt zu übermitteln, sie aber in Sprache zu übertragen, die auch dem Ottonormalverbraucher verständlich ist.“ (Interview Nr. 15)

Zwei Forschende waren hingegen der Meinung, dass die Inhalte des Kompendiums zwar einem breiteren Publikum zur Verfügung gestellt werden sollten, empfahlen aber eine Kürzung bzw. einen zielgruppengerechteren Zuschnitt. Ein Forschender fragte zudem nach einer englischen Version, um das Kompendium an nicht-deutschsprachige Kollegen weitergeben zu können. Zwei Befragte merkten an, dass sie sich das Kompendium in der vorliegenden Form als Reihe vorstellen könnten. Entweder zu weiteren Themen wie beispielsweise zur „Risikokommunikation in der IT“ oder auch als aktualisierte Fassungen zu Risikokommunikation und Künstliche Intelligenz.

VII.4. Zusammenfassung

Die 16 aus den beiden Zielgruppen des Kompendiums befragten Forschenden und Kommunikator*innen im KI-Bereich waren mit dem Inhalt des Produkts sehr zufrieden. Insbesondere die ausgewählten Themen, deren Einordnung und ihre Aktualität waren ausschlaggebend dafür, dass das Kompendium mehrfach als gute Übersicht über das Themengebiet KI gelobt wurde. Auch mit der Korrektheit der Darstellung der einzelnen Themenschwerpunkte waren die Befragten, bis auf eine einzelne Kritik an einer Quelle, zufrieden.

Diejenigen Interviewten, die selbst bereits im Bereich Risikokommunikation tätig waren, waren überzeugt davon, dass das Kompendium auch in zukünftigen Kommunikationssituationen zur Rekapitulation von Kommunikationsstrategien bzw. insgesamt als Vorbereitungsmaterial relevant sein wird. Ein Großteil der Befragten wollte das Kompendium weiterempfehlen oder hatte dies bereits getan. Als besonders hilfreich wurde es für Personen mit noch wenig Kommunikationserfahrung eingestuft.

Optimistisch war die Mehrheit der Befragten auch dahingehend, dass das Kompendium zu mehr öffentlicher Kommunikation über KI beitragen wird. Die meisten Interviewten fühlten sich durch das Produkt besser auch auf schwierige Kommunikationssituationen vorbereitet und an die Relevanz öffentlicher Kommunikation erinnert.

Auch mit der gestalterischen Umsetzung waren alle Befragten zufrieden. Lediglich die zusätzlichen Online-Inhalte wurden von den Befragten überhaupt nicht genutzt.

Insgesamt konnte das Kompendium die befragten Forschenden und Kommunikator*innen im KI-Bereich somit in den untersuchten Kategorien Inhalt, Relevanz und Umsetzung überzeugen.

VII.5. Anhang

VII.5.1. Anhang 1: Leitfaden

Leitfrage	Nachfragen & relevante inhalte
Sie haben das Kompendium „Risikokommunikation zur Künstlichen Intelligenz“ seit einigen Wochen. Was ist Ihr genereller Eindruck ?	<ul style="list-style-type: none"> ➤ Was ist Ihnen positiv aufgefallen? ➤ Was ist Ihnen negativ aufgefallen?
Haben Sie das Kompendium im Print oder digital genutzt?	<ul style="list-style-type: none"> ➤ Was sagen Sie zum Umfang? ➤ Was zur Druckqualität? ➤ Es gibt das Kompendium sowohl als Print-, als auch als digitale Version. <ul style="list-style-type: none"> • Welche bevorzugen Sie? • Beeinflusst das Print- bzw. PDF-Format, wie Sie das Kompendium in der Praxis verwenden?
Wie relevant ist das Thema „ Risikokommunikation zur Künstlichen Intelligenz“ für Ihren Arbeitskontext?	<ul style="list-style-type: none"> ➤ Ist das Kompendium etwas, das Sie dabei verwenden werden? ➤ In welchem Kontext? ➤ Würden Sie es weiterempfehlen /-geben?
Das Kompendium behandelt verschiedene Themenbereiche. Wo hatten Sie Schwerpunkte bei der Nutzung?	<ul style="list-style-type: none"> ➤ Haben Sie das Kapitel Risikokommunikation genutzt? Warum (nicht)? ➤ Haben Sie das Kapitel zu KI-Definitionen genutzt? <ul style="list-style-type: none"> • Ja: Wie schätzen Sie es ein? Ist es hilfreich für Debatten zu KI-Themen? • Werden dabei Aspekte behandelt, die schon außerhalb des Arbeitskontextes diskutiert wurden? • Nein: Warum haben Sie es nicht genutzt? Würden Sie es in anderem Zusammenhang noch nutzen?
Es gibt die fünf Themenbereiche; KI und Umgang mit Daten, KI und Arbeit, Autonomes Fahren, KI in der Medizin sowie KI und Robotik. Welchem der Themenbereiche würden Sie sich zuordnen? Welcher Themenbereich steht Ihrem Forschungsthema/ Arbeitsalltag am nächsten?	<ul style="list-style-type: none"> ➤ Wie bewerten Sie die Qualität in diesem Themenbereich? <ul style="list-style-type: none"> • Wird die Darstellung der Komplexität gerecht? • Entspricht der Überblick, der gegeben wird, Ihrem Eindruck vom Thema und dem aktuellen Stand der Debatte? ➤ Wie schätzen Sie die Kommunikationstipps ein? ➤ Die Kommunikationstipps entstammen den Workshops, sind also von Forschenden für Forschende formuliert. <ul style="list-style-type: none"> • Ist Ihnen das aufgefallen? • Spielt die Quelle der Tipps eine Rolle für Sie?
Haben Sie den Eindruck, dass auch die Darstellung der anderen	<ul style="list-style-type: none"> ➤ Welche Erwartung hatten Sie beim Lesen? ➤ Wurde diese erfüllt? ➤ In KI-Diskussionen wird KI oft auch breiter behandelt. Haben Sie das Gefühl, jetzt besser für solche Diskussionen gewappnet zu sein?

Themenbereiche für Sie wichtig ist? („fachfremde“ Themenbereiche)	
Ein Ziel ist es, mehr Kommunikation über KI in der Öffentlichkeit, auch in den Medien, zu wagen. Inwieweit kann das Kompendium dazu beitragen?	<ul style="list-style-type: none"> ➤ Welche Inhalte waren dabei hilfreich? ➤ Was hätten Sie sich sonst noch gewünscht? ➤ Würden Sie das RIKI-Kompendium heranziehen, wenn Sie sich auf öffentliche Kommunikation z.B. ein Interview vorbereiten?
<i>Die folgenden Fragen stellen, falls dazu noch genügend Zeit bleibt:</i>	
Das Projekt „Risikokommunikation zur Künstlichen Intelligenz“, kurz RIKI, und damit auch das RIKI-Kompendium ist gefördert aus Mitteln des Bundesministeriums für Bildung und Forschung.	<ul style="list-style-type: none"> ➤ Ist für Sie relevant, dass es aus Bundesmitteln finanziert wurde? ➤ Ist das ein Qualitätsmarker? Oder neutral/negativ?
Das Projekt „Risikokommunikation zur Künstlichen Intelligenz“ wurde vom Nationalen Institut für Wissenschaftskommunikation, kurz NaWik , durchgeführt.	<ul style="list-style-type: none"> ➤ Assoziieren Sie etwas mit diesem Absender? ➤ Was verbinden Sie mit dem NaWik?
Im RIKI-Kompendium wird auf Online-Inhalte hingewiesen. Haben Sie diese abgerufen oder in irgendeiner Form genutzt?	<ul style="list-style-type: none"> ➤ Falls nein: Warum nicht? ➤ Falls ja: Warum und wie? • Hatte diese Verknüpfung einen Mehrwert für Sie? • Haben sich das Kompendium und die Online-Inhalte sinnvoll ergänzt? • Könnten Sie sich vorstellen, diese Online-Inhalte an interessierte Kollegen weiterzuempfehlen?
Das Kompendium wurde Ihnen im Print zugeschickt. Was sagen Sie zum Erscheinungsbild ?	<ul style="list-style-type: none"> ➤ Was sagen Sie zum Design? • Bilder • Aufbau und Struktur • Spalten • Fettgedruckte Schlagworte • Unterüberschriften • Kapitelstruktur konsequent
<i>Zur folgender, abschließender Frage springen, falls die Zeit knapp wurde:</i>	
Vielen Dank, haben Sie noch irgendwelche Anmerkungen zum Kompendium?	
<i>Folgendes abfragen / recherchieren. In beiden Fällen auf anonymisierte Verwendung hinweisen:</i>	
Alter	
Geschlecht	
F: Schwerpunkt/ K: Arbeitsfeld	
Akademischer Grad/ Position	