

Teil 2:
Daten schützen – Kinder schützen.
Ethik der Überwachung in Familie und Alltag

Zur Bedeutung anthropomorpher Zuschreibungen für das Vertrauen in vernetztes Kinderspielzeug

Ricarda Moll

Vernetztes Kinderspielzeug als soziale Roboter

Seit einigen Jahren sind Spielzeuge für Kinder auf dem deutschen wie internationalen Markt verfügbar, die durch die Möglichkeit der Vernetzung mit anderen Geräten neue Formen spielerischer Interaktivität zwischen Kind und Spielzeug ermöglichen sollen. Konzeptuell lassen sich mindestens drei Eigenschaften nennen, die in Verbindung mit dem sog. *Internet of Toys* relevant sind:¹

Erstens beinhalten zahlreiche der betreffenden Spielzeuge Sensoren, die Informationen aus der Umgebung aufnehmen können. Typischerweise sind dies Mikrofone oder Kameras, wobei auch andere Formen von Sensoren wie zum Beispiel Berührungssensorik zum Einsatz kommen. Zweitens sind die betreffenden Spielzeuge vernetzt; sie verfügen über eine Funkschnittstelle, über die das Produkt zur Übertragung der aufgezeichneten Daten in der Lage ist. Drittens zeichnen sich die betreffenden Produkte dadurch aus, dass sie eine physisch anfassbare, körperliche Form annehmen (sog. *Embodiment*),² wobei sich die Art der Verkörperung von bereits bekannten Spielzeugformen wie Teddybären oder Puppen unterscheiden kann, aber nicht muss. Das Merkmal des Embodiment unterscheidet vernetztes Spielzeug in bedeutender Weise von Sprachassistenten wie Siri oder Alexa, indem es eine physische Bindung an das Objekt ermöglicht. So kann ein vernetztes Spielzeug im Alltag durch die Wohnung getragen werden, Zuwendung erhalten, oder nachts mit ins Bett genommen werden, wie es auch bei herkömmlichen Spielzeugen durchaus üblich ist.

Die Kombination von eingebauter Sensorik, Funkschnittstelle und Verkörperung ermöglicht verhältnismäßig komplexe Interaktionen. Eine besondere Rolle spielt hierfür die Integration von Spracherkennung, mit Hilfe derer das Spielzeug menschenähnliche Kommunikation simuliert. Im Zuge

1 Vgl. Peter et al. 2019.

2 Vgl. ebd. 2019.

dessen werden also Informationen auf eine sozial wirkende Art und Weise ausgetauscht. Insofern sind derzeit zumeist noch als Spielzeuge vermarktete Produkte perspektivisch als soziale Roboter anzusehen, deren inhärente Funktion die Fähigkeit zu sozialer Interaktion in menschenähnlicher Art und Weise ist.³

Gegenstand dieses Beitrags ist die Frage, inwieweit Kinder auf Basis der sozial erscheinenden Eigenschaften sozialer Roboter *Vertrauen* in das vernetzte Spielzeug entwickeln. Ausgangspunkt hierfür ist das Phänomen sogenannter anthropomorpher Zuschreibungen, die als Grundlage für das Entstehen einer sozialen Beziehung gesehen werden können (s. Unterabschnitt „Anthropomorphe Zuschreibungen“). Das Vertrauen des Kindes in einen sozialen Roboter spielt in mindestens zweierlei Hinsicht eine bedeutende Rolle für den weiteren Informationsaustausch: Erstens ist Vertrauen (s. Unterabschnitt „Vertrauen in soziale Roboter“) ausschlaggebend in einer Situation, in der das Kind eine Information vom Roboter empfängt und deren Glaubwürdigkeit beurteilen muss (s. Unterabschnitt „Epistemisches Vertrauen“). Zweitens ist Vertrauen entscheidend für die Frage, inwieweit ein Kind dem sozialen Roboter gegenüber Informationen preisgibt (s. Unterabschnitt „Selbstoffenbarung“). Zu beiden Aspekten wird – ohne Anspruch auf Vollständigkeit – ein Einblick in damit verbundene Forschung gegeben sowie praktische und ethische Implikationen diskutiert.

Anthropomorphe Zuschreibungen

Menschen neigen dazu, Verhaltensweisen von nicht menschlichen Objekten mit Attributen und Zuständen zu erklären, die als spezifisch für die Gattung des Menschen gelten. Hierzu gehören beispielsweise rationale Gedanken, Selbstreflektion, Kreativität, Emotionen und Empathie.⁴ Anthropomorphismus beschreibt in diesem Sinne die Zuschreibung von mentalen Zuständen und Prozessen zu nicht menschlichen Akteuren (im Folgenden: anthropomorphe Zuschreibungen).⁵ Die grundsätzliche Tendenz, das Gegenüber in einem „sozialen Modell“ mental zu repräsentieren⁶ basiert auf kognitiver Ebene auf einer induktiven Inferenz, bei der das früh er-

3 Vgl. Breazeal 2003, S. 167.

4 Vgl. Waytz/Cacioppo/Epley 2010, S. 220 für einen Überblick.

5 Vgl. ebd., S. 221.

6 Vgl. Breazeal 2003, S. 168.

lernte und somit auch mental hoch verfügbare Wissen über menschliche Eigenschaften auf andere Akteure übertragen wird.⁷ Evolutionär betrachtet scheint dies ein adaptiver Mechanismus zu sein, der soziale Interaktion und Kooperation erleichtert.⁸

Anthropomorphe Zuschreibungen nehmen tendenziell zu, wenn das in Frage stehende Objekt beobachtbare Eigenschaften aufweist, die an einen Menschen erinnern. Hierzu zählen beispielsweise physische Eigenschaften (z. B. Gesicht, Bewegungen), aber auch geäußerte Emotion und soziale Motivation.⁹ Soziale Roboter sind in diesem Sinne genau dazu konzipiert, den Eindruck zu erwecken, sie seien menschenähnliche Wesen; dies soll in der Logik der Roboter-Entwicklung den sozialen Austausch zwischen Mensch und Maschine vereinfachen.¹⁰ Empirische Untersuchungen zeigen, dass diese Konzeption Menschen durchaus beeinflusst, selbst wenn die physisch-soziale Form des Roboters ein Tier ist. So zeigte beispielsweise die Analyse von Nutzerkommentaren in einem Diskussionsforum von AIBO, einem sozialen Roboter-Hund der Firma Sony, dass (erwachsene) Nutzer von AIBO diesen auf der einen Seite zwar als Technologie ansahen, andererseits jedoch eine emotionale Bindung zu ihm aufbauten und ihm teilweise Gefühle und Absichten zuschrieben.¹¹

Da die Schwelle zu anthropomorphen Zuschreibungen bereits für erwachsene Anwender_innen niedrig zu sein scheint, liegt es nahe, dass auch Kinder einer Maschine, vor allem in Form eines sozialen Roboters, menschenähnliche Eigenschaften zuschreiben. Die grundsätzliche Fähigkeit zur Attribution mentaler Zustände ist Teil des (mental)en Heranwachsens des Kindes, welches mit zunehmendem Alter eine sogenannte *Theory of Mind* entwickelt, also eine Vorstellung von und ein Verständnis für mentale Zustände von sich selbst und anderen Personen.¹² Eine Theory of Mind kann somit als Voraussetzung für anthropomorphe Zuschreibungen in der Interaktion mit sozialen Robotern gesehen werden.

7 Vgl. Shin/Kim 2020, S. 446 für einen Überblick.

8 Vgl. Damiano/Dumouchel 2018, S. 1.

9 Vgl. z. B. Fink 2012.

10 Vgl. Damiano/Dumouchel 2018. Der Zusammenhang zwischen menschenähnlichen Eigenschaften und Akzeptanz seitens des_r Nutzers_in ist hierbei nicht linear. So scheinen Menschen eine zu große Ähnlichkeit eines Roboters mit einem Menschen unheimlich zu finden; dieses Phänomen ist unter dem Begriff *Uncanny Valley* bekannt, s. Mori/MacDorman/Kageki 2012.

11 Vgl. Friedman/Kahn/Hagman 2003, S. 275.

12 Vgl. Wellman 2018 für einen Überblick.

Erste Untersuchungen stützen die These, dass Kinder zu anthropomorphen Zuschreibungen neigen, wenn sie mit einem sozialen Roboter interagieren. In einer einschlägigen Studie aus dem Jahr 2012¹³ wurden Interaktionssequenzen zwischen Kindern und einem (in diesem Fall von einem Menschen ferngesteuerten) sozialen Roboter namens *Robovie* initiiert (z. B.: Begrüßung, prosoziale Bitte, Rollenwechsel im Spiel, Umarmung). Jede Interaktion endete damit, dass der Versuchsleiter das Spiel zwischen Kind und Robovie unterbrach und Robovie in einen Schrank sperrte. Robovie protestierte dagegen, unter anderem indem er sich auf moralische Grundsätze wie Gerechtigkeit und Fairness bezog. So sagte er beispielsweise, dass er Angst habe, allein im dunklen Schrank zu sein, und es ungerecht sei, dass er das Spiel nicht fortsetzen dürfe. Im Folgenden wurde das Kind im Rahmen eines strukturierten Interviews zu Robovie befragt. Interessanterweise betrachteten Kinder Robovie teilweise als eine Art moralisches Wesen mit eigenen Rechten; so fanden 54 Prozent der Kinder, dass es nicht in Ordnung sei, Robovie einfach in den Schrank zu sperren – bei einer tatsächlichen Person fanden dies 100 Prozent, wohingegen jedoch alle Kinder sich einig waren, dass es in Ordnung sei, einen Besen in den Schrank zu sperren.¹⁴ Insofern scheinen soziale Roboter für Kinder in eine neue ontologische Kategorie zu fallen, als Systeme, die zwar nicht lebendig, aber eben auch nicht *nicht* lebendig seien.¹⁵ Hypothesenkonform glaubte die Mehrheit der befragten Kinder darüber hinaus, Robovie könne kognitive Zustände wie Intelligenz, Interesse, Erwartungen und Gedanken haben. Sie glaubten auch, dass Robovie Gefühle hat, beispielsweise stimmten 81 Prozent der befragten Kinder der Aussage zu, dass sie Robovie trösten müssten, wenn dieser sagen würde, er sei traurig und dass sie ebenfalls Trost bei Robovie suchen würden, wenn sie selbst traurig seien.¹⁶

Eine Implikation anthropomorpher Zuschreibungen ist, dass hierdurch die Entstehung einer scheinbar sozialen Beziehung möglich wird.¹⁷ Diese Beziehung ist einseitig und insofern auch eine Illusion vonseiten des

13 Vgl. Kahn et al. 2012.

14 Vgl. Kahn et al. 2012, S. 310. Hierzu sei angemerkt, dass der Vergleich zwischen Robovie und einem Besen konfundiert sein könnte, denn im Gegensatz zu Robovie gehört ein Besen tatsächlich in den Schrank. Eine geeignetere Vergleichskategorie könnte hier ein unanimiertes Spielzeug sein, oder – im Zuge eines experimentellen Settings – ein unanimierter, nicht responsiver Roboter.

15 Vgl. Kahn Jr/Gary/Shen 2013, S. 35.

16 Vgl. Kahn et al. 2012, S. 309.

17 Van Straten/Peter/Kühne 2020 für einen Überblick.

Kindes, denn der soziale Roboter ist nicht tatsächlich zu einer sozialen Bindung fähig. Diese Einseitigkeit schließt jedoch nicht aus, dass vonseiten des Kindes emotionale und kognitive Prozesse in Gang gesetzt werden, die typisch für eine soziale Beziehung zu einem Menschen sind. Ein Beispiel für einen solchen im zwischenmenschlichen Bereich relevanten Prozess ist die Entstehung von Vertrauen.

Vertrauen in soziale Roboter

Für den gegenwärtigen Zweck kann Vertrauen definiert werden als eine Einstellung oder Haltung gegenüber einem anderen Akteur, die auf Verhaltensebene mit der Delegation von Kontrolle¹⁸ und auf kognitiver Ebene mit der Zuschreibung von Eigenschaften wie Kompetenz, Wohlwollen und Integrität bzw. Verlässlichkeit einhergeht.¹⁹ Um jemandem zu vertrauen, muss eine Person also davon überzeugt sein, dass der in Frage stehende Akteur a) kompetent in Bezug auf das Anliegen ist („können“) und b) das Anliegen auch erfüllen will („wollen“), beispielsweise weil er wohlwollend und integer ist.²⁰

Das Vertrauen des Kindes in einen sozialen Roboter kann hier in mindestens zweierlei Hinsicht relevant werden: Zum einen spielt Vertrauen eine Rolle, wenn das Kind die Glaubwürdigkeit einer Information einschätzen muss, die es vom sozialen Roboter erhält (s. Unterabschnitt „Epistemisches Vertrauen“); zum anderen ist Vertrauen relevant, wenn das Kind einschätzen muss, ob der soziale Roboter eine Information für sich behalten wird, die es ihm gegenüber offenbart (s. Unterabschnitt „Selbstoffenbarung“).

Epistemisches Vertrauen

Ein Großteil von dem, was wir über unsere Umwelt wissen und über sie lernen, basiert auf Informationen, die wir von anderen bekommen.²¹ Eine Voraussetzung dafür, dass Menschen tatsächlich von anderen lernen, ist

18 Vgl. Castelfranchi/Falcone 2000.

19 Vgl. Mayer/Davis/Schoorman 1995.

20 Vgl. Castelfranchi/Falcone 2000.

21 Vgl. Origi 2014; Nurmsoo/Robinson 2009a, S. 23.

ihr Vertrauen darin, dass ihr Gegenüber ihnen die Wahrheit sagt.²² Hierzu muss dem Gegenüber einerseits zugetraut werden, dass er korrektes Wissen auf dem in Frage stehenden Gebiet hat, insofern also kompetent ist, und andererseits die relevanten Informationen auch korrekt weitergeben will, weil er der vertrauenden Person wohlwollend gegenüber eingestellt ist.²³

Empirische Untersuchungen zeigen, dass Kinder soziale Roboter grundsätzlich als Informationsquelle betrachten, insbesondere, wenn diese auch nonverbal menschlich wirken, beispielsweise über ein sozial-responsives Blickverhalten.²⁴ Darüber hinaus lernen dreijährige Kinder die Namen neuartiger Objekte eher von sozialen Robotern als von einer unanimierten Maschine, sofern sie ersteren mentale Eigenschaften zuschrieben.²⁵ Insofern scheinen anthropomorphe Zuschreibungen prinzipiell Annahmen über die Gültigkeit der empfangenen Information zu beeinflussen.

Unklar ist hierbei jedoch, inwieweit Kinder differenziertere Zuschreibungen von Wohlwollen und Kompetenz machen, wenn sie mit einem sozialen Roboter interagieren. Aus Studien zu kindlichem Vertrauen in erwachsene Personen („informants“) gibt es Hinweise darauf, dass Kinder diese beiden Marker für Vertrauenswürdigkeit unterschiedlich stark gewichten. Beispielsweise glaubten Kinder einem Erwachsenen, der zuvor ein bekanntes Objekt falsch bezeichnet hatte, weil er die Augen verbunden hatte, auch dann nicht, wenn dieser ein dem Kind unbekanntes Objekt mit *unverbundenen* Augen benannte.²⁶ Dies könnte darauf hindeuten, dass Kinder Fehler des Informanten in erster Linie auf dessen Wohlwollen zurückführen und in diesem Fall situative Grenzen der Kompetenz weniger schwer in ihren Urteilen gewichten. Hierfür spricht auch ein Befund aus dem Jahr 2013. Hier zeigte sich, dass Kinder erst ab einem Alter von ca. fünf Jahren Expertise bzw. Kompetenz in bedeutender Weise in ihre Vertrauensurteile einbeziehen; kleinere Kinder vertrauten den Aussagen von Informanten auch dann, wenn diese offensichtlich keine Expertise aufweisen, solange sie jedoch wohlwollend („nett“) erschienen.²⁷ Diese Tendenz scheint ab einem bestimmten Alter abzunehmen, wobei selbst Erwachsene zumindest bei

22 Vgl. Sperber et al. 2010, S. 360.

23 Vgl. ebd., S. 369.

24 Vgl. Breazeal et al. 2016.

25 Vgl. Brink/Wellman 2020.

26 Vgl. Nurmsoo/Robinson 2009b. Vgl. Nurmsoo/Robinson 2009a für ein gegenteiliges Ergebnis.

27 Vgl. Landrum/Mills/Johnston 2013; Lane/Wellman/Gelman 2013; für einen Überblick: Tong/Wang/Danovitch 2020.

emotional schwierigen Entscheidungen das Wohlwollen ihres Gegenübers stärker in ihrem Vertrauensurteil gewichten als seine Expertise.²⁸

Für den gegenwärtigen Kontext sind die geschilderten Befunde insofern relevant, als dass die körperliche Form sozialer Roboter – ggf. in Form von vernetztem Spielzeug – anthropomorphe Zuschreibungen unterstützt, die wiederum eine sozial-emotionale Bindung zur Maschine fördern.²⁹ Wenn Kinder – und anteilig auch Erwachsene – dem Wahrheitsgehalt von Informationen vor allem aus sozioemotionalen Gründen („Ist er nett? Mag er mich?“) vertrauen und dabei in geringerem Maße die Kompetenz des Akteurs berücksichtigen („Ist er kompetent auf diesem Gebiet?“), könnten Kinder ihrem technischen Gefährten auch dann glauben, wenn er ihnen objektiv gesehen eine verzerrte oder gar unzutreffende Information gibt. Insofern erhöht die anthropomorphe Form eines vernetzten Spielzeugs das Potenzial für Fehlinformation und Täuschung des Kindes, woraus sich wiederum ethische sowie praktische Implikationen ergeben (s. Unterabschnitt „Implikationen“).

Selbstoffenbarung

Selbstoffenbarung im weitesten Sinne findet in jeder Art von kommunikativem Akt statt, unabhängig davon, ob dies beabsichtigt ist oder nicht.³⁰ Schon diese Feststellung ist in Bezug auf die Kind-Roboter-Interaktion relevant, weil das Kind durch jede Verhaltensweise Informationen über sich selbst offenbart. Insoweit ist das Kind potenziell Gegenstand von Live-Tracking, dessen Ausmaß mit davon abhängt, welche der erhobenen Daten standardmäßig erhoben und weiterverarbeitet werden. Die Besonderheit in Bezug auf einen sozialen Roboter ist jedoch vielmehr, dass auch explizite, bewusste Selbstoffenbarung ihm gegenüber wahrscheinlicher werden könnte, wenn das Kind ihn als soziales Wesen wahrnimmt (vgl. Unterabschnitt „Anthropomorphe Zuschreibungen“). Das Kind könnte dem sozialen Roboter hierdurch beispielsweise Informationen anvertrauen, die den Status eines (kindlichen) Geheimnisses haben. In diesem Sinne könnte der technische „Gefährte“ des Kindes in noch bedeutenderer Weise in dessen

28 Vgl. White 2005.

29 Studien zeigen darüber hinaus, dass anthropomorphe Zuschreibungen auch beeinflusst, inwieweit Kinder einen sozialen Roboter als zuständig für bestimmte Wissensgebiete wahrnehmen, s. Oranç/Küntay 2020.

30 Vgl. Greene/Derlega/Mathews 2006, S. 411.

Privatsphäre eindringen, als es über das Nutzer-Tracking ohnehin schon möglich ist.

Empirische Beispiele dafür, dass Menschen sich einer Maschine gegenüber offenbaren, existieren spätestens seit den 60er-Jahren, in denen der erste soziale Chatbot mit dem Namen *Eliza* entwickelt wurde.³¹ Eliza ermöglichte eine natürliche Art von Unterhaltung zwischen einem Menschen und einem Computer und war so angelegt, dass es die Gedanken und Gefühle im Sinne des aktiven Zuhörens nach Carl Rogers spiegelt und damit den Eindruck erweckt, aufmerksam und empathisch zu sein.³² Da sich Eliza in der Unterhaltung scheinbar wie ein therapeutisch agierender Mensch verhielt, provozierten die programmierten Verhaltensweisen anthropomorphe Zuschreibungen; in die (scheinbar) empathischen Äußerungen wurden vonseiten der sprechenden Person Absichten und Hintergrundwissen interpretiert.³³ Eliza schien aus eben diesem Grund einen „starken emotionalen Effekt auf viele der Nutzer“³⁴ zu haben, die große Bereitschaft zeigten, sich dem Computerprogramm gegenüber zu offenbaren.

In Bezug auf soziale Roboter, die anthropomorphe Zuschreibungen schon aufgrund ihrer äußeren Erscheinung provozieren (sollen), liegt es nahe, dass die Integration interaktiver Dialogsysteme in ein vernetztes Spielzeug dazu führen kann, dass Kinder dem Spielzeug private Informationen offenbaren. Diese Vermutung stützen auch verschiedene empirische Untersuchungen. Beispielsweise antworteten 57 Prozent der 90 befragten Kinder in der Studie von Kahn et al., dass sie dem sozialen Roboter Robovie ein Geheimnis anvertrauen würden.³⁵ Ähnliches legt eine Fallstudie aus dem Jahr 2017 nahe, in der neun Kinder zum Umgang mit sozialen Roboter-Spielzeugen befragt wurden.³⁶

Den Kommunikationspartner_innen im Kontext der bewussten Selbstoffenbarung zu vertrauen, kann insbesondere die Erwartung beinhalten, dass dieser die offenbarte Information für sich behalten. Insofern kann Vertrauen auch hier zunächst als Zuschreibung von Kompetenz verstanden werden: Ein Kind könnte beispielsweise auf der einen Seite davon ausge-

31 Vgl. Weizenbaum 1966.

32 Vgl. Turkle 2007, S. 502.

33 Vgl. Weizenbaum 1966, S. 42 Die Zuschreibung menschlicher Eigenschaften zu künstlicher Intelligenz wird daher in informationstechnischen Disziplinen auch Eliza-Effekt genannt; vgl. Kim/Schmitt/Thalmann 2019, S. 2.

34 Turkle 2007, S. 502 (eigene Übersetzung).

35 Vgl. Kahn et al. 2012, S. 309.

36 Vgl. McReynolds et al. 2017.

hen, dass der soziale Roboter dazu fähig ist, ein Geheimnis für sich zu behalten, oder auf der anderen Seite unfähig dazu ist, es weiterzugeben (Zuschreibung von Kompetenz bzw. Inkompetenz). Insbesondere letztere Zuschreibung zeigt sich in der bereits angeführten Fallstudie, in der die befragten Kinder ihrem vernetzten Spielzeug teilweise die Fähigkeit absprechen, sich an eine Unterhaltung erinnern zu können.³⁷ Hierbei ist allerdings unklar, ob diese Antworten auf einem mangelnden technischen Verständnis oder vielmehr auf einer ausbleibenden Anthropomorphisierung beruhen. Denn während Erinnerung eine menschliche Fähigkeit ist, hat sie im Gegensatz zu anderen zugeschriebenen mentalen Eigenschaften eine technische Entsprechung – so ist es sehr wahrscheinlich, dass ein sozialer Roboter mit Spracherkennung die Äußerungen des Kindes aufzeichnet und sich insoweit auch „erinnern“ kann, selbst wenn er diese Erinnerung dem Kind gegenüber nicht artikulieren kann.

Neben der Zuschreibung von Kompetenz kann für die Selbstoffenbarung des Kindes auch eine Rolle spielen, inwieweit es den sozialen Roboter als Freund wahrnimmt, der ihm wohlwollend gegenüber eingestellt ist. So geht man bei einem wohlwollenden, integren Freund nicht davon aus, dass dieser ein anvertrautes Geheimnis weitererzählt. Dass ein Kind sein vernetztes Spielzeug als wohlwollend wahrnehmen könnte, liegt nahe, denn neben der physisch-sinnlichen Vertrautheit ist die mentale Repräsentation des sozialen Roboters von anthropomorphen Zuschreibungen geprägt. Diese werden bereits durch physische Eigenschaften des Roboters hervorgerufen; sie werden darüber hinaus jedoch durch fortgeschrittene Dialogsysteme weiter gefördert werden, sofern diese die Simulation menschlicher Unterhaltungen immer weiter perfektionieren und individualisieren. Moderne soziale Chatbots wie ChatGPT der Firma OpenAI³⁸ übertreffen bereits heute Elizas Fähigkeiten zur Simulation menschlicher Unterhaltungen.³⁹ Insofern ist es wahrscheinlich, dass ein Kind einem vernetzten Spielzeug, in das ein solches Dialogsystem integriert ist, private Informationen anvertraut. Dies könnte es in der Annahme tun, dass der vernetzte Freund das anvertraute Geheimnis entweder nicht weitergeben kann oder dies aufgrund der freundschaftlichen Beziehung zum Kind nicht tun würde.

37 Vgl. McReynolds et al. 2017, S. 5202.

38 Vgl. Bubeck et al. 2023.

39 Vgl. Shah et al. 2016.

Implikationen

Die fortschreitende Optimierung von dem, was schon heute als künstliche Intelligenz bezeichnet wird, wird es irgendwann selbstverständlich erscheinen lassen, dass auch Kinder mit sozialen Robotern interagieren. Insofern sind die aktuell vermarkteten vernetzten Kinderspielzeuge als der Anfang einer längerfristigen Entwicklung zu sehen.

Die vorausgegangenen Darstellungen deuten insbesondere auf das Problem hin, dass Kinder einem sozialen Roboter vertrauen, weil sie ihn als soziales Wesen wahrnehmen. Ihr Vertrauen bildet also nicht Erwartungen an die Zuverlässigkeit eines technischen Systems oder die wohlwollenden Absichten des für die Programmierung (mit-)verantwortlichen Herstellers ab, sondern das Vertrauen in den verkörpert und als anteilig lebendig wahrgenommenen sozialen Roboter selbst.⁴⁰ Eine Dynamik, für die nach den vorangehenden Ausführungen sowohl die technischen als auch die psychischen Voraussetzungen unter Umständen erfüllt sind, könnte demnach wie folgt aussehen:

Ein Kind vertraut seinem „Spielzeug“ persönliche Informationen an, die es gegebenenfalls sogar niemandem sonst erzählt. Der soziale Roboter, der keine mentalen Eigenschaften wie Absicht oder Wohlwollen abseits seiner programmierten Informationsverarbeitung besitzt, überträgt diese Informationen an Server des Anbieters, wo sie unter Umständen nicht für den maximalen Nutzen für das Kind, sondern zur Gewinnmaximierung für das Unternehmen weiterverarbeitet werden. Die gesammelten Informationen ermöglichen es, dem Kind über den sozialen Roboter eine personalisierte Empfehlung für bestimmte Produkte zu übermitteln. Das Kind, das dem sozialen Roboter innerhalb dieser als freundschaftlich wahrgenommenen Beziehung epistemisch vertraut, übernimmt in der Konsequenz die Empfehlung und lässt sich somit in seinen Wünschen und seiner Entwicklung beeinflussen.

Das skizzierte Manipulationspotenzial sozialer Roboter sollte insbesondere im Kontext kindlicher Vertrauensbeziehungen berücksichtigt werden. Eine grundlegende Frage ist hier beispielsweise, inwieweit es entwicklungspsychologisch bedeutsam und insofern auch ethisch vertretbar ist, in einem Kind die Illusion von sozialer Beziehung und Freundschaft zu erwecken.⁴¹

40 Vgl. Culley/Madhavan 2013.

41 Vgl. Turkle 2007. Für Ausführungen zu ethischen Gesichtspunkten anthropomorpher Zuschreibungen im Kontext sozialer Roboter s. Damiano/Dumouchel 2018.

Diese Frage stellt sich insbesondere, weil Kinder im Gegensatz zu Erwachsenen die Art der sozialen Beziehung, die sie zu ihrem vernetzten Spielzeug haben, unter Umständen weniger gut auf metakognitiver Ebene einordnen können. Eine solche Einordnung schützt zwar nicht notwendigerweise vor dem Eingehen einer sozialen Beziehung mit dem sozialen Roboter, wie aus Berichten von erwachsenen Nutzer_innen bekannt ist,⁴² macht es theoretisch jedoch möglich, vigilant in Bezug auf die tatsächliche Funktionsweise der Maschine und damit verbundenen Risiken zu sein.

Relevant scheint hier auch der Befund, dass Aufklärungsversuche mit Kindern in Bezug auf soziale Roboter nicht notwendigerweise zielführend sind. Beispielsweise berichtet die Ethnologin Sherry Turkle von Kindern, die mit dem sozialen Roboter *Cog* interagierten.⁴³ Nachdem die Kinder sich mit ihm vertraut gemacht und ihm gegenüber soziale Verhaltensweisen gezeigt hatten, wurde *Cog* gezielt entmystifiziert. *Cogs* Entwickler erklärte ihnen hierzu detailliert, wie der soziale Roboter auf technischer Ebene funktioniert. Laut Turkle verhielten sich die Kinder jedoch einige Minuten später *Cog* gegenüber wieder so, als sei er ein soziales Wesen.

Diese Fallschilderung deutet darauf hin, dass die erläuterten psychologischen Mechanismen – insbesondere anthropomorphe Zuschreibungen und darauf aufbauende Vertrauenshandlungen – verhältnismäßig stabil und damit auch gewissermaßen immun gegenüber Aufklärung sind. Es scheint insofern eine noch offene empirische Frage zu sein, wie Kinder im Kontext sozialer Robotik effektiv aufgeklärt und geschützt werden können. Beim Design sozialer Roboter sollte aus diesem Grund nicht ausschließlich die größtmögliche Akzeptanz des Roboters als wichtigstes Kriterium gelten. Vielmehr sollte es auch Ziel sein, das Vertrauen des_r Nutzers_in auf ein angemessenes Maß zu begrenzen, beispielsweise indem man anthropomorphe Hinweise so weit wie möglich reduziert.

Da Kinder aufgrund ihres Alters und Entwicklungsstands besonders vulnerabel sein können (s. auch Erwägungsgrund 38 DS-GVO), sollten über die vorangehenden Überlegungen hinaus Instrumente geschaffen werden, die insbesondere den Markt für vernetzte Kinderspielzeuge regulieren. Gegenstand der Regulierung kann perspektivisch auch das Ausmaß der Beziehungstäuschung sein, die im Rahmen einer solchen Kind-Roboter-Interaktion erzeugt werden darf. Um das Missbrauchspotenzial einer möglicherweise entstehenden Kind-Roboter-Beziehung frühzeitig zu begrenzen,

42 Friedman/Kahn/Hagman 2003.

43 Vgl. Turkle 2007, S. 504.

scheinen darüber hinaus konservative Vorschriften zur Erhebung, Weiterverarbeitung und Nutzung von Daten über das Kind zu kommerziellen Zwecken unabdingbar.

Literatur

- Breazeal, Cynthia (2003): Toward sociable robots. In: *Robotics and Autonomous Systems* 42, 3, S. 167-175.
- Breazeal, Cynthia et al. (2016): Young Children Treat Robots as Informants. In: *Top Cogn Sci* 8, 2, S. 481-491.
- Brink, Kimberly A./Wellman, Henry M. (2020): Robot teachers for children? Young children trust robots depending on their perceived accuracy and agency. In: *Developmental psychology* 56, 7, S. 1268-1277.
- Bubeck, Sébastien/Chandrasekaran, Varun/Eldan, Ronen et al. (2023): Sparks of Artificial General Intelligence: Early experiments with GPT-4. Online: <https://arxiv.org/pdf/2303.12712> (letzter Zugriff: 02.06.2023).
- Castelfranchi, Cristiano/Falcone, Rino (2000): Trust and control: A dialectic link. In: *Applied Artificial Intelligence* 14, 8, S. 799-823.
- Culley, Kimberly E./Madhavan, Poornima (2013): A note of caution regarding anthropomorphism in HCI agents. In: *Computers in Human Behavior* 29, 3, S. 577-579.
- Damiano, Luisa/Dumouchel, Paul (2018): Anthropomorphism in Human-Robot Co-evolution. In: *Frontiers in Psychology* 9, S. 468.
- Fink, Jul (2012): Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction. In: Ge, S. S., et al. (Hrsg.): *Social Robotics*, Berlin, Heidelberg, S. 199-208.
- Friedman, Batya/Kahn, Peter H./Hagman, Jennifer (2003): Hardware Companions? What Online AIBO Discussion Forums Reveal about the Human-Robotic Relationship: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, S. 273-280.
- Greene, Kathryn/Derlega, Valerian J./Mathews, Alicia (2006): Self-Disclosure in Personal Relationships. In: Vangelisti, A. L./Perlman, D. (Hrsg.): *The Cambridge Handbook of Personal Relationships*, Cambridge, S. 409-428.
- Kahn, Peter H., et al. (2012): "Robovie, you'll have to go into the closet now": children's social and moral relationships with a humanoid robot. In: *Developmental psychology* 48, 2, S. 303-314.
- Kahn Jr, Peter H./Gary, Heather E./Shen, Solace (2013): Children's Social Relationships With Current and Near-Future Robots. In: *Child Dev Perspect* 7, 1, S. 32-37.
- Kim, Seo Young/Schmitt, Bernd H./Thalmann, Nadia M. (2019): Eliza in the uncanny valley: anthropomorphizing consumer robots increases their perceived warmth but decreases liking. In: *Marketing Letters* 30, 1, S. 1-12.
- Landrum, Asheley R./Mills, Candice. M./Johnston, Angie. M. (2013): When do children trust the expert? Benevolence information influences children's trust more than expertise. In: *Dev Sci* 16, 4, S. 622-638.

- Lane, Johnathan D./Wellman, Henry M./Gelman, Susan A. (2013): Informants' Traits Weigh Heavily in Young Children's Trust in Testimony and in Their Epistemic Inferences. In: *Child Development* 84, 4, S. 1253-1268.
- Mayer, Roger C./Davis, James H./Schoorman, F. David (1995): An Integrative Model of Organizational Trust. In: *The Academy of Management Review* 20, 3, S. 709-734.
- McReynolds, Emily et al. (2017): Toys That Listen: A Study of Parents, Children, and Internet-Connected Toys: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. New York, S. 5197-5207.
- Mori, Masahiro/MacDorman, Karl F./Kageki, Norri (2012): The Uncanny Valley [From the Field]. In: *IEEE Robotics & Automation Magazine* 19, 2, S. 98-100.
- Nurmsoo, Erika/Robinson, Elisabeth J. (2009a): Children's Trust in Previously Inaccurate Informants Who Were Well or Poorly Informed: When Past Errors Can Be Excused. In: *Child Development* 80, 1, S. 23-27.
- Nurmsoo, Erika/Robinson, Elisabeth J. (2009b): Identifying unreliable informants: do children excuse past inaccuracy? In: *Dev Sci* 12, 1, S. 41-47.
- Oranç, Cansu/Küntay, Aylin C. (2020): Children's perception of social robots as a source of information across different domains of knowledge. In: *Cognitive Development* 54.
- Origgi, Gloria (2014): Epistemic trust. In: Capet P./Delavallade, T. (Hrsg.): *Information Evaluation*, London, S. 35-54.
- Peter, Jochen, et al. (2019): Asking Today the Crucial Questions of Tomorrow: Social Robots and the Internet of Toys. In: Mascheroni, Giovanna/Holloway, Donell (Hrsg.): *The Internet of Toys. Practices, Affordances and the Political Economy of Children's Smart Play*. Cham, S. 25-46.
- Shah, Huma et al. (2016): Can machines talk? Comparison of Eliza with modern dialogue systems. In: *Computers in Human Behavior* 58, S. 278-295.
- Shin, Hong I./Kim, Juyoung (2020): My computer is more thoughtful than you: Loneliness, anthropomorphism and dehumanization. In: *Current Psychology* 39, 2, S. 445-453.
- Sperber, Dan et al. (2010): Epistemic Vigilance. In: *Mind & Language* 25, 4, S. 359-393.
- Tong, Yu/Wang, Fuxing/Danovitch, Judith (2020): The role of epistemic and social characteristics in children's selective trust: Three meta-analyses. In: *Developmental Science* 23, 2, e12895.
- Turkle, S. (2007): Authenticity in the age of digital companions. In: *Interaction Studies* 8, 3, S. 501-517.
- Van Straten, Caroline L./Peter, Jochen/Kühne, Rinaldo (2020): Child-Robot Relationship Formation: A Narrative Review of Empirical Research. In: *International Journal of Social Robotics* 12, 2, S. 325-344.
- Waytz, Adam/Cacioppo, John/Epley, Nicholas (2010): Who Sees Human?: The Stability and Importance of Individual Differences in Anthropomorphism. In: *Perspectives on Psychological Science* 5, 3, S. 219-232.
- Weizenbaum, Joseph (1966): ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine. In: *Commun. ACM* 9, 1, S. 36-45.

- Wellman, Henry M. (2018): Theory of mind: The state of the art, in: *European Journal of Developmental Psychology* 15, 6, S. 728-755.
- White, Tiffany B. (2005): Consumer trust and advice acceptance: the moderating roles of benevolence, expertise and negative emotions. In: *Journal of Consumer Psychology* 15, 2, S. 141-148.