

# MERS-MRI – Methodische, ethische, rechtliche und soziale Implikationen für die Mensch-Roboter-Interaktion in Alltagswelten

Andreas Bischof<sup>1</sup>, Robert Herms<sup>1</sup>, Maximilian Eibl<sup>1</sup>

Professur Medieninformatik, Technische Universität Chemnitz<sup>1</sup>

[andreas.bischof@informatik.tu-chemnitz.de](mailto:andreas.bischof@informatik.tu-chemnitz.de), [robert.herms@...](mailto:robert.herms@...), [maximilian.eibl@...](mailto:maximilian.eibl@...)

## Zusammenfassung

Der Workshop thematisiert die methodischen, ethischen, rechtlichen und sozialen Herausforderungen bei der Entwicklung von Mensch-Roboter-Interaktion für Alltagswelten. Als Alltagswelten werden dabei unstrukturierte Umgebungen verstanden, die durch soziale Interaktionen und alltägliche Praktiken geprägt sind, wie etwa Supermärkte, Privatwohnungen oder Krankenhäuser. Der Workshop fokussiert im Besonderen auf Gestaltungs- und Evaluationsmethoden, die in der avisierten Nutzungs situation verankert sind und die Nutzerinnen und Nutzer auf Augenhöhe integrieren. Das Ziel des Workshops ist es, Expertinnen und Experten aus Disziplinen wie Soziologie, Philosophie und Rechtswissenschaften und Konstrukteurinnen und Konstrukteure von robotischen Anwendungen für Alltagswelten zu vernetzen und eine gemeinsame Diskussion zu ermöglichen.

## 1 Einleitung

Roboter für Alltagswelten wie Supermärkte, Wohnungen oder Pflegeheime zu bauen, kann als „wicked problem“ verstanden werden (Rittel & Webber, 1973). Alltagswelten sind für Roboter nicht nur physisch schlecht kontrollierbare Umwelten, auch soziale Situationen und Interaktionen sind in einem technischen Sinn oftmals unvollständig und widersprüchlich. Gleichzeitig wirkt auf der Seite der Entwicklung der Zwang, die soziale Komplexität solcher Situationen auf diskrete Mathematik zu reduzieren (Meister, 2014; Lindemann, 2016). Dieser Kontrast aus einem schwer zu standardisierenden Phänomen wie Mensch-(Roboter-)Interaktion einerseits und standardisierten Problemlösungsstrategien andererseits ist in anderen Domänen

der Mensch-Computer-Interaktion und Software-Entwicklung bereits thematisiert worden. Das Konzept des „*wicked problem*“ wurde deshalb schon früher zur Konzeption von Designentscheidungen herangezogen (DeGrace & Stahl, 1990; Courtney, 2008; Berger et al., 2017). Der heuristischer Wert der Beschreibung der Herausforderung von Technikentwicklung für Alltagswelten als Planungsproblems durch Rittel und Webber besteht unseres Erachtens vor allem in der **Verschränkung der Problemdimensionen Mensch, Prozess und Technik**.

Eben dort, wo die nun auf Alltagswelten statt auf Fabrikanwendungen zielenden Roboter funktionieren sollen, ist es nicht nur die Komplexität der Maschinen, die zur Herausforderung wird. Für solche „soziale‘ Maschine gilt es zunächst zu bestimmen, was „Funktionieren“ in der konkreten Anwendung überhaupt bedeuten soll – und auf welchem Weg dies erreicht werden soll. Der empirische Blick auf den Forschungsstand des interdisziplinären Feldes Human-Robot Interaction (HRI) hat gezeigt, dass die Reflexion über diese Fragen weder methodisch noch institutionell verankert ist (Bischof, 2017). Daraus folgt, dass die konkreten Bedingungen und Methoden der Entwicklung und Evaluation von HRI für Alltagswelten wissenschaftlich besser begleitet werden müssen. Insbesondere im Hinblick auf die Methodologie der Evaluation (2.) sowie die die konkreten ethischen, rechtlichen und sozialen Themen (3.) ist eine **Forschungslücke** offensichtlich.

Im Umgang mit dem „wicked problem“ Roboter für Alltagswelten zu bauen, sind Sozialrobotik und HRI mit Phänomenen konfrontiert, die etablierte Wissensformen, Fähigkeiten und Perspektiven beispielsweise der Informatik und KI-Forschung irritieren und neue erforderlich machen. Die Entwicklung und Evaluation von Robotern für Alltagswelten muss dabei aus unserer Sicht als eine hybride Praxis zwischen wissenschaftlich-scientistischen und praxeologisch-entwerfenden Polen verstanden werden. Das Vorgehen enthält Elemente von naturwissenschaftlich-experimentellen Methoden, designerische Praktiken und typisch ingenieurale Problemlösungsstrategien. Diese unterschiedlichen Einflüsse implizieren unterschiedliche Theorie- und Methodenbestände, die sich teilweise auf konkurrierende wissenschaftstheoretische Traditionen und Wissenskulturen beziehen. Der Workshop versucht diese Heterogenität abzubilden und interdisziplinär, empirisch verankerte Impulse insbesondere zu den in (3.) aufgeworfenen Fragen zu diskutieren und zu dokumentieren, um diese hybride Praxis mit einem höheren Reflexionspotential zu unterstützen.

## 2 Methodologische Herausforderungen der Evaluation von Mensch-Roboter-Interaktion in Alltagswelten

Die Interaktion zwischen Robotern und Menschen ist nicht nur technisch-konzeptionell, sondern auch methodisch eine große Herausforderung. Die bisherigen Ergebnisse im interdisziplinären Forschungsfeld HRI (Bischof, 2017; Sabanovic, 2010; Dautenhahn, 2007) zeigen, dass die Evaluationskriterien eines aus Nutzersicht adäquaten und erwünschten Einsatzes von Robotern in Alltagswelten nicht ausschließlich mit standardisierten Methoden in Laborsettings zu messen sind (2.1). Stattdessen erfordert die Gestaltung einer erfolgreichen

HRI eine Verankerung der Evaluation in der konkreten Situation des Einsatzes. Die Bestimmung und Kontrolle der für die Nutzerinnen und Nutzer dabei relevanten Faktoren wiederum ist oftmals erst durch den probeweisen Einsatz und die Konfrontation mit Prototypen möglich (Weiss et al., 2008; Sabanovic et al., 2006), was eine besondere methodischen Herausforderung darstellt (2.2).

## 2.1 Grenzen laboratisierter Verfahren

Die häufigste Variante der empirischen Evaluation von HRI sind standardisierte Verfahren wie Laborexperimente und quantitative Fragebogen-Studien. Diese fokussieren zumeist auf Performance-Metriken bei der Durchführung von kooperativen Aufgaben oder Dialogen (Goodrich & Schultz, 2007) und subjektive Einstellungen zur erlebten Qualität der HRI aus Sicht der Teilnehmenden (Takayama & Pontafaru, 2009). Diese Fragestellungen dienen einerseits als direkte Evaluation der entwickelten HRI, oder werden andererseits als Implikationen zur weiteren Gestaltung, etwa ob weitere Schutzmechanismen im Verhalten des Roboters in Hinblick auf die Personensicherheit angebracht sind, verwendet.

Auch wenn die Entwicklung der Fragebögen und Test-Protokolle, sowie die Evaluation der Frage-Items und die Durchführung dieser standardisierten Verfahren mittlerweile qualitativ oftmals auf dem Niveau psychologischer Forschung stattfindet (Rosenthal-von der Pütten et al., 2013), führen standardisierte Evaluationsmethoden in der HRI eine methodologische Einschränkung mit: Die Kontingenz – also prinzipielle Offenheit und teilweise auch Zufälligkeit – und Komplexität sozialer Situationen wird hierbei reduziert, indem spezifische Ausschnitte sozialer Welten dekontextualisiert und in laborierte Ordnungen gefasst werden. Statt beispielsweise dem Einkauf als Teil der alltäglichen Routine und den dabei möglicherweise auftretenden Problemen bei der Interaktion mit einem Roboter zu erheben und auszuwerten, wird eine dem nachempfundene Labor-Interaktion mit einem Roboter hergestellt und gemessen. Es handelt sich also um eine Rekonfiguration der ‚natürlichen‘ Ordnung einer bestimmten Alltagswelt. Das daraus resultierende Labor-Experiment muss erkenntnistheoretisch als eine Realität eigener Ordnung verstanden werden (Knorr-Cetina, 2002, S. 45). Die größte methodische Herausforderung standardisierter Evaluationsverfahren der Mensch-Roboter-Interaktion ist deshalb ein methodologisches: Ein im Labor erfolgreich evaluerter Roboter kann in einer unstrukturierten Alltagsumgebung komplett versagen, der Grad der Adäquatanz einer Labor-Evaluation für konkrete Alltagswelten ist – unabhängig von der Güte der Durchführung – zunächst vollkommen unklar.

## 2.2 Formative Evaluations-Ansätze als Chance

Die beschriebene Herausforderung verweist auf zwei Dimensionen der Dekontextualisierung von HRI in Evaluationsmethoden: die räumliche Transformation in abgeschlossene Labore, und die zeitliche in eine der eigentlichen Entwicklung nachgeordnete Evaluation der fertigen Maschine. Beiden Problemen lässt sich mit formativen, also die Entwicklung begleitenden und informierenden Evaluations-Ansätzen, begegnen.

Die Mensch-Computer-Interaktion und ihre Forschungsfelder, insbesondere partizipative Designmethoden und andere Formen des intensiven „user involvement“ haben bereits eine ganze Reihe formativer Evaluationsmethoden und Entwicklungskonzepte hervorgebracht und diskutiert. Gemeinsam ist ihnen, dass erarbeitete Interaktionsstrategien und Interfaces nicht erst **summativ, nach Beendigung der Entwicklungsarbeit, evaluiert, sondern bereits während des Konstruktionsprozesses unter Einbindung zukünftiger Nutzerinnen und Nutzer bewertet und ggf. angepasst werden**. Dieser formative Modus der Evaluation kann sich beispielsweise an den Prinzipien der partizipativen Gestaltung orientieren (Sanders, 2003), bei der die Menschen, die von einem Design betroffen sind, **im Prozess der Entwicklung eine Stimme erhalten**.

Allerdings sind diese Ansätze durchaus mit erhöhtem Personal- und Zeitaufwand verbunden und können nicht „nebenbei“ erledigt werden. Gleichzeitig besteht die Möglichkeit, auf einem etablierten Kanon partizipativer Methoden aufzubauen und sich pragmatischer Methoden, wie etwa der „Living Lab“-Methode, zu bedienen, die bereits seit mehr als zehn Jahren erfolgreich eingesetzt wird (Eriksson et al., 2005). Das Prinzip eines „Living Labs“ besteht zum Beispiel darin, die adressierten Nutzerinnen und Nutzer mit den technischen Systemen in möglichst lebensechten Umgebungen, sogar den avisierten Anwendungsorten, zusammenzubringen. Dadurch können die Nutzerinnen und Nutzer die Technik praxisnah explorieren und selbst Vorschläge zur Verbesserung anführen. Es handelt sich also um eine Methode mit hoher Nutzerzentrierung, die etablierte Methoden der Systementwicklung, wie Pilotstudien, Laborexperimente oder Feldtests in zweierlei Hinsicht ergänzt (Følstad, 2008, S. 117): Zum einen lassen sich „Living Labs“ bereits zu frühen Zeitpunkten des Entwicklungsprozesses, in denen das System noch nicht vollständig umgesetzt ist, durchführen. Zum anderen begegnen sich **Entwicklerinnen und Entwickler und adressierte Nutzerinnen und Nutzer in diesen Formaten in dialogischer Form gewissermaßen auf Augenhöhe** und nicht als anonymer Proband und Testleiter, oder beobachtete Person und Beobachter.

Formative Evaluationskonzepte, egal in welcher Methode umgesetzt, erlauben es, jeweils dem aktuellen Stand der technischen Plattform und der entwickelten Interaktionskonzepte angepasste Evaluationsmethoden durchzuführen. Sie zielen darauf, die Sicht der adressierten Nutzerinnen und Nutzer teils auch niedrigschwellig zu erheben, um bereits zu frühen Zeitpunkten im Entwicklungsprozess Annahmen und Hypothesen des Designs und der Designer zu prüfen.

### 3 Ethische, rechtliche und soziale Aspekte in der Evaluation von Mensch-Roboter-Interaktion in Alltagswelten

**Akzeptanz und Nutzerfreundlichkeit** sind für eine gute Mensch-Roboter-Interaktion in Alltagswelten **wichtige, aber nicht ausreichende Maße**. Entwicklungs- und Forschungsprojekte der HRI werfen eine Reihe umfassenderer ethischer, rechtlicher und sozialer Aspekte auf. Die verantwortungsvolle Auseinandersetzung mit solchen Aspekten war das Ziel der Etablierung verpflichtender, sogenannter ELSI-Arbeitspakte (ethical, legal and social issues) in Forschungs- und Entwicklungsprojekten (Zwart et al., 2014). Mit der begrüßenswerten

Verpflichtung zur Bearbeitung solcher Fragestellungen gehen aber erneut methodische und methodologische Fragen einher. So implizieren übergreifende ELSI-Fragestellungen u.a. die Hinzuziehung fachlicher Expertise beispielsweise aus Domänen wie Rechtswissenschaft, angewandte Philosophie, Psychologie, Soziologie und Technikfolgeabschätzung. Im Folgenden sollen beispielhafte Aspekte für die drei genannten Bereiche vorgestellt werden.

### 3.1 Ethische Aspekte der HRI in Alltagswelten

Die Evaluation ethischer Aspekte der Mensch-Roboter-Interaktion in Alltagswelten besteht ausdrücklich nicht darin, durch das Projekt eventuell betroffene ethische und gesellschaftliche Fragestellungen aufzuzählen. Stattdessen besteht Ethik darin, konkrete Kriterien für gutes und schlechtes Handeln und die Bewertung seiner Motive und Folgen aufzustellen. In Bezug auf die Entwicklung von Mensch-Roboter-Interaktion besteht eine ethische Bewertung also vor allem darin, Kriterien guten Handelns für die Entwicklungsziele, den Entwicklungsprozess, den Umgang mit avisierter Nutzerinnen und Nutzern und die resultierende HRI zu identifizieren und diese wirkungsvoll und nachhaltig in das eigene Vorgehen zu integrieren.

Für diese Bewertung gibt es unterschiedliche Leitlinien und Orientierungsrahmen. Einerseits sind durch das Rechtssystem und die Selbstverpflichtungen guter wissenschaftlicher Praxis (DFG, 2013) bereits rechtliche Normen gesetzt (3.2.). Andererseits gibt es bereits methodische Angebote, um die inhärenten Werte von Entwicklungszielen und -praktiken zu bewerten. Dazu gehören beispielsweise das „Value Sensitive Design“ (Friedman et al., 2002), das MEESTAR-Modell (Weber, 2015; Manzeschke, 2015) oder das AMICAI-Modell (Wille et al., 2016).

Die ethische Evaluation muss dabei auch die Absichten und Ziele der Forschenden und des Forschungsprojekts beinhalten und diese auf Konflikte zwischen den Werten der avisierter Nutzer diskutieren. Ohne eine solche Überprüfung auf Konflikte können fehlgeleitete Entwicklung und falsche Annahmen im Forschungsprozess nicht ausgeschlossen werden. So ist es beispielsweise denkbar, dass das Ziel der Automatisierung von einzelnen Handlungsschritten des Einkaufs einen wesentlichen Konflikt mit dem Selbstverständnis und den Praktiken der Verkäuferinnen und Verkäufer in einem Supermarkt bedeutet. In diesem Fall kann eine ethische Bewertung auch mögliche Lösungswege für die Konfliktlösung aufzeigen. Im gewählten Supermarkt-Beispiel könnte eine Lösung darin bestehen, das Personal in die Bedienung des Roboters zu integrieren und eine hybride Assistenz zu designen.

### 3.2 Rechtliche Aspekte der HRI in Alltagswelten

Für Roboter im öffentlichen Raum und Assistenzroboter wurden in den vergangenen Jahren erste Rechtsvorschriften und Normen, wie bspw. die ISO für persönliche Assistenzroboter (DIN, 2014), erlassen. Dennoch zieht der Einsatz (teil-)autonom agierender Maschinen in Alltagswelten zahlreiche, teilweise bislang ungelöste rechtliche Folgefragen nach sich. Zu den grundlegenden Fragen im Umgang mit Robotern gehören im Hinblick auf das Strafrecht die Fragen der Haftung (Beck, 2009). Eine wichtige Frage bei der Evaluation rechtlicher Aspekte lautet deshalb: Welche Sicherheitsrisiken bestehen bei dem Einsatz in öffentlich zugänglichen Räumen und wie können sie minimiert werden?

Ein weiterer wichtiger rechtlicher Aspekt ist der **Datenschutz von Nutzerinnen und Nutzern, die mit bildverarbeitenden Systemen**, wie Robotern mit visueller Erkennung, interagieren. Für die Aufzeichnung, Speicherung und Verarbeitung von Videodaten bestehen sowohl in öffentlichen wie auch privaten Räumen Gesetzesregelungen, die zu berücksichtigen sind?

### 3.3 Soziale Aspekte der HRI in Alltagswelten

Oftmals herrscht in der Evaluation von HRI in Alltagswelten eine dyadische Konzeption der Nutzung vor: Mensch-Roboter-Interaktion wird verstanden als Phänomen zwischen *einem Nutzer und einer Schnittstelle*. In der sozialen Wirklichkeit geraten Roboter in Alltagswelten aber in ein bereits bestehendes Netzwerk aus unterschiedlichen Akteuren und Akteurinnen. So ist eine robotische Einkaufsassistent nicht nur mit Kundinnen und Kunden konfrontiert, sondern noch viel häufiger mit Mitarbeiterinnen und Mitarbeitern des Supermarkts. Zwar zählen Zweitäre nicht zur Primärzielgruppe der Entwicklung, sie sind für das Gelingen der robotischen Anwendung aber oftmals Gatekeeper, wie Studien beispielsweise zu Museumsführern und Museums-Robotern gezeigt haben (Nourbakhsh et al., 2003).

Auf der meso- und makrosoziologischen Ebene stellen sich neben der Akteurskonstellation in der Nutzungssituation Folgefragen. Im Beispiel des Supermarktroboters ließe sich fragen, welche flankierenden arbeitsmarktpolitischen Maßnahmen (Weiterbildungen, Anpassungen in der Ausbildung) hilfreich oder notwendig werden, wenn robotische Assistenzsysteme eingeführt werden. Übergreifend kann und muss nach Veränderungen oder Verschwinden ganzer Berufsbilder, wie bspw. das des Einzelhandelskaufmanns bzw. der Einzelhandelskauffrau durch die Einführung von Robotern für Alltagswelten gefragt werden.

## Literaturverzeichnis

- Beck, S. (2009). Grundlegende Fragen zum rechtlichen Umgang mit der Robotik. *Juristische Rundschau*, 2009(6), 225-230.
- Berger, A.; Totzauer, S.; Lefevre, K.; Storz, M.; Kurze, A.; Bischof, A. (2017). Wicked, Open, Collaborative: Why Research through Design Matters for HCI Research. *i-com*, 16. 2017, 2, 131–142.
- Bischof, A. (2017). *Soziale Maschinen bauen: Epistemische Praktiken der Sozialrobotik*. Bielefeld: transcript.
- Courtney, J. F. (2008). Decision Making and Knowledge Management in Inquiring Organisations: toward a new decision-making paradigm for DSS. *Decision Support Systems - Knowledge management support of decision making*, 31(1): 17–38.
- Dautenhahn, K. (2007). Socially intelligent robots: dimensions of human-robot interaction. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362(1480), 679–704.
- DeGrace, P.; Stahl, L. (1990). *Wicked Problems, Righteous Solutions: A Catalog of Modern Engineering Paradigms*. Yourdon Press.
- Deutsche Forschungsgemeinschaft. (2013). *Sicherung guter wissenschaftlicher Praxis*. <http://doi.org/10.1002/9783527679188.oth1> [29.03.2018]

- Deutsches Institut für Normung e.V. (2014). *Roboter und Robotikgeräte – Sicherheitsanforderungen für persönliche Assistenzroboter (ISO 13482:2014)*; Deutsche Fassung EN ISO 13482:2014.
- Eriksson, M.; Niitamo, V.P.; Kulkki, S. (2005). *State-of-the-art in utilizing Living Labs approach to user-centric ICT innovation-a European approach*. Lulea: Center for Distance-spanning Technology. Lulea University of Technology Sweden: Lulea.
- Følstad, A. (2008). Living Labs for Innovation and Development of Information and Communication Technology: A Literature Review. *eJOV 10*, August (2008), 99–131.
- Friedman, B., Kahn, P., & Bornig, A. (2002). *Value sensitive design: Theory and methods*. University of Washington technical report, 02–12.
- Goodrich, M.A.; Schultz, A. C. (2007). Human-robot interaction: a survey. *Foundations and trends in human-computer interaction*, 1(3), 203–275.
- Herrmanny, K.; Schwarz, M.; Paldán, K.; Beckmann, N.; Sell, J.; Wagner, N.F., Dogangün, A. (2017). Designing a Holistic Behavior Change Support System for Healthy Aging. *i-com*, 16(2), 99–111.
- Knorr Cetina, Karin. (2002). *Die Fabrikation von Erkenntnis. Zur Anthropologie der Naturwissenschaft*. Frankfurt a.M.: Suhrkamp.
- Lindemann, G. (2016). Social interaction with robots: three questions, *AI & Society*, 31(4): 573–575.
- Manzeschke, A. (2015). Angewandte Ethik organisieren: MEESTAR–ein Modell zur ethischen Deliberation in sozio-technischen Arrangements. *Vom Praktisch-Werden der Ethik in interdisziplinärer Sicht: Ansätze und Beispiele der Institutionalisierung, Konkretisierung und Implementierung der Ethik*, 7.
- Meister, M. (2013). When is a robot really social? An outline of the robot sociologicus. *Science, Technology & Innovation Studies*, 10(1), 107–134.
- Mulvenna, M.; Martin, S.; McDade, D.; Beamish, E.; De Oliveira, A.; Kivilehto, A. (2011). *TRAIL Living Labs Survey: A Survey of the ENOLL Living Labs*. University of Ulster, 2011.
- Nourbakhsh, I. R.; Kunz, C.; Willeke, T. (2003). The mobot museum robot installations: A five year experiment. *Intelligent Robots and Systems*, 2003, 3636–3641.
- Pallot, M.; Trousse, B.; Senach, B.; Scapin, D. (2010). *Living lab research landscape: From user centred design and user experience towards user cocreation*. First European Summer School „Living Labs“.
- Rittel, H. W.; Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy sciences*, 4(2), 155–169.
- Rosenthal-von der Pütten, A.; Krämer, N.; Hoffmann, L.; Sobieraj, S.; Eimler, S. C. (2013). An experimental study on emotional reactions towards a robot. *International Journal of Social Robotics*, 5(1), 17–34.
- Sabanovic, S. (2010). Robots in Society, Society in Robots Mutual Shaping of Society and Technology as a Framework for Social Robot Design. *International Journal of Social Robotics*, 2(4), 439–450.
- Sanders, E. B. N. (2003). From user-centered to participatory design approaches. *Design and the social sciences*, 18–25.
- Takayama, L.; Pantofaru, C. (2009). Influences on proxemic behaviors in human-robot interaction. *Intelligent robots and systems*, 2009, 5495–5502.

- Weber, K. (2015). MEESTAR: Ein Modell zur ethischen Evaluierung sozio-technischer Arrangements in der Pflege- und Gesundheitsversorgung. *Technisierung des Alltags*, 247.
- Weiss, A.; Bernhaupt, R.; Tscheligi, M. Wollherr, D.; Kuhnen, K.; Buss, M. (2008). A methodological variation for acceptance evaluation of human-robot interaction in public places. *RO-MAN 2008*, 713–718.
- Wille, M.; Brandl, C.; Nelles, J.; Mertens, A.; Schlick, C. M. (2016). AMICAI: Ein methodisches Vorgehen zur quantitativen Analyse von ethischen, rechtlichen und sozialen Auswirkungen anwendungsnaher Forschungsprojekte. *Arbeit in komplexen Systemen—Digital, vernetzt, human*.
- Zwart, H.; Landeweerd, L.; van Rooij, A. (2014). Adapt or perish? Assessing the recent shift in the European research funding arena from ‘ELSA’ to ‘RRI’. *Life Sciences, Society and Policy*, 10(1), 11.