

# Sozial empathische Systeme coram publico

Anke Tallig

Fakultät für Informatik  
Technische Universität Chemnitz  
Straße der Nationen 62  
09111 Chemnitz  
anke.tallig@informatik.tu-chemnitz.de

**Abstract:** Dieses Paper beschreibt die Weiterentwicklung von öffentlichen interaktiven Systemen zu sozial empathischen Systemen. Als öffentliche interaktive Systeme werden in diesem Zusammenhang allgemein alle zugänglichen Informationssysteme coram publico verstanden, mit denen der Nutzer auf verschiedene Weise interagieren kann. Im Rahmen des interdisziplinären Graduiertenkollegs „*CrossWorlds – Connecting Virtual and Real Social Worlds*“ wird, durch die Erkennung der sozialen Signale der Nutzer, ein sozial empathisches System entwickelt. Dieses System ist ein interaktiver Public Screen mit einer *Edutainment*-Oberfläche. Das Paper zeigt Punkte auf, die aus der Mensch-Mensch-Interaktion für eine soziale Empathie softwaretechnisch umgesetzt werden können. Ebenso wird ein Vorgehensmodell der Softwareentwicklung vorgestellt, welches während der Zusammenarbeit von Ingenieurin und Sozialwissenschaftler entwickelt wurde und an die Bedingungen und Bedürfnisse der interdisziplinären Arbeit angepasst wurde. Ziel der Entwicklung sozial empathischer Systeme ist der leichtere Umgang mit diesen Systemen, damit sich nicht nur der Mensch auf das System einstellen muss, sondern das System sich auch auf seine(n) Menschen einstellen bzw. sein Angebot anpassen kann, was gleichzeitig die Wirtschaftlichkeit und die soziale Akzeptanz der Systeme erhöht.

## 1 Die Idee

Von der Idee Systeme zu entwickeln, die die Nutzer verstehen und nicht nur vom Nutzer verstanden werden müssen, träumen die Menschen seitdem sie die Entwicklung vorantreiben. Maschinen zu entwickeln, die nicht nur ihren Zweck erfüllen, wenn der Nutzer den richtigen Knopf drückt, sondern Maschinen, die angepasst an den jeweiligen Nutzer nur bestimmte Knöpfe zur Verfügung stellen, um die Nutzung einfacher und schneller zu gestalten. Interaktive Systeme, die ihre Arbeit aufnehmen, bevor der Nutzer den Prozess aktivieren muss. „Bitte berühren Sie den Bildschirm“ – warum? Damit das System aus dem Ruhemodus erweckt wird und zur Arbeit zur Verfügung steht? Warum sieht das System den Nutzer nicht kommen und bootet aus dem Ruhemodus, damit der Nutzer keine Wartezeiten hat oder womöglich denkt das System sei defekt. Warum lädt das System den Nutzer nicht ein, sich mit den Inhalten zu beschäftigen? Warum muss der Nutzer das System „bitten“ seine Informationen preiszugeben? Interaktive Informationssysteme sollen das Leben erleichtern. Aber Interaktion fordert nicht nur

vom Menschen, dem Nutzer, sich auf das System einzustellen und zu verstehen, was in diesem vorgeht, auch das System sollte sich auf seinen Interaktionspartner einstellen können, um einen reibungslosen Austausch von Informationen zu gewährleisten.

Für eine softwaretechnische Lösung von Nutzerproblemen im Umgang mit technischen Systemen muss versucht werden, eine, in den meisten Fällen, funktionierende *Mensch-Mensch-Interaktion* zu analysieren und die auf die *Mensch-Computer-Interaktion* übertragbaren Faktoren für das System zu implementieren.

Initiiert durch die *menschliche Interaktion*, die die sozialen Signale, die der Mensch aussendet, nutzt, um eine angepasste und soziale Kommunikation einzuleiten [ArM13], begann die Entwicklungsarbeit an einem sozial empathischen interaktiven System. Dieses prototypische System entsteht im Rahmen des Graduiertenkollegs „*CrossWorlds – Connecting Virtual and Real Social Worlds*“. Genau, wie sich Menschen in ihrer Kommunikation auf die *non-verbale Signale* ihres Gegenübers verlassen, sollte dies ein intelligentes interaktives System auch können. Ziel des Projektes ist es, zu beginnen, diese Fähigkeiten zu implementieren.

Da die Erkennung von sozialen Signalen in erster Linie keine technische Fragestellung ist, muss die Entwicklung einer erfolgreichen Erkennungssoftware (*Punkt 2*) in interdisziplinärer Zusammenarbeit erfolgen. Die Zusammenarbeit von Ingenieurwissenschaft und Sozialwissenschaft, wie auch in [PeM94, StJ98] beschrieben, gestaltet sich aufgrund der unterschiedlichen Vorgehensweisen der Wissenschaften problematisch. Die Entwicklung eines interdisziplinären Vorgehensmodells (*Punkt 2.1*) für die spezielle Zusammenarbeit ist für das Erreichen der Ziele unumgänglich. Denn erst die Anpassung der Softwareentwicklung an den realen Nutzer lässt den Transfer der Erkenntnisse aus der Wissenschaft in die Wirtschaft zu.

## 2 Die Grundlagen

Über die Grundlagen des nonverbalen Ausdrucks [ArM13] lassen sich die Vorteile der Anwendung in interaktiven Systemen ableiten. Die menschliche Wahrnehmung des Gegenübers findet in drei Phasen statt: Die *erste* Wahrnehmungsphase erfolgt in der Entfernung. Der mögliche Interaktionspartner wird beobachtet und es wird eingeschätzt, was seine Intension an diesem Ort ist bzw. sein könnte. Die *zweite* Phase beschreibt das Näherkommen, in dem das Bild, das sich der Mensch von seinem möglichen Interaktionspartner macht, verfeinert wird. Die *dritte* Phase ist das Gespräch oder die Interaktion an sich.

Ingenieure und Informatiker entwickeln technische Systeme mit immer umfangreicheren Funktionen und nutzerfreundlicheren Interfaces, die den Nutzer unterstützen und ihm weiterhelfen sollen. Aber das System nimmt den Menschen nicht als soziales Wesen wahr. Dem Nutzer werden verschiedene interaktive Systeme, z. B. Großbildschirme [DMi09], im öffentlichen Raum angeboten. Diese Public Screens spiegeln den Nutzer, damit er sich auf dem Bildschirm und damit dessen Interaktivität erkennt. Einige Systeme geben ab und an Geräusche von sich, um auf sich aufmerksam zu machen.

Durch z. B. Blickerkennung und -verfolgung versuchen die Hersteller von Mobilgeräten [AI13], den Nutzern den Umgang zu erleichtern.

Aber all diese Implementierungen betreffen die *dritte* Wahrnehmungsphase in der menschlichen Interaktion. Diese softwaretechnischen Umsetzungen überspringen die ersten beiden Schritte der Interaktion und können damit nicht ausreichend auf ihre Nutzer eingehen. Sie verschenken die Möglichkeit, den Nutzer besser kennenzulernen bevor die eigentliche Interaktion stattfindet. Die (soziale) Einschätzung des Partners ist ein wichtiger Punkt für Interaktion von Menschen und kann im Umgang mit interaktiven Systemen eine herausragende Rolle spielen, was nicht nur die allgemeine Akzeptanz der Systeme, sondern durch umfangreichere und gezieltere Nutzung auch deren Wirtschaftlichkeit erhöht.

## 2.1 Nonverbale Signale

Nonverbale Signale sind unter anderem die Körperhaltung, das Raumverhalten, Körperbewegung und der Gesichtsausdruck [ArM13]. Werden diese Signale wahrgenommen, kann eine softwareseitige Verarbeitung vorgenommen werden, die der des Menschen gleicht. Die Verarbeitung dieser nonverbalen Signale ist beim Menschen weitaus komplexer als es zurzeit softwaretechnisch realisierbar ist. Die Kommunikation ist ein Kreislauf, der keinen Anfang und kein Ende kennt, jede Aktion ist Reaktion und Reaktion eine Aktion [WaP11]. Sich diesen Kreislauf zum Nutzen zu machen und den fortwährend kommunizierenden Menschen, den möglichen Nutzer, wahrzunehmen und ihm so zu helfen, ist ein erreichbares Ziel. Der erste Schritt ist es, die menschlichen Wahrnehmungsphasen in ihren Grundzügen zu analysieren. Im Weiteren müssen die vorerst ausgewählten Signale extrahiert und für diese eine oder mehrere Deutung/en hinterlegt werden, um diese dem System Stück für Stück zuzuführen.

Es gibt diverse Untersuchungen, die sich mit der Analyse von menschlichem Verhalten und zwischenmenschlichem Verhalten im sozialen Raum beschäftigen. Die F-Formations (facing formations) [MaP11] sind zum Beispiel Ergebnis solcher Untersuchungen. Hier wird dargestellt, in welcher Entfernung und in welchem Winkel zwei oder mehr Menschen zueinanderstehen. Dies lässt wiederum Rückschlüsse auf deren Beziehung zueinander und unter Umständen ihre Offenheit für andere mögliche Interaktionspartner zu. Zusätzlich wurden diese F-Formations für den Kontakt mit öffentlichen und mobilen Systemen erweitert [DiE13, MaN12]. Darüber hinaus wurden Schemen für die Annäherung von möglichen Nutzern an öffentliche Bildschirme entwickelt [MiD11]. Diese Annäherungsphasen sind für die Entwicklung einer Erkennungssoftware von großem Nutzen. In Verbindung mit den Erkenntnissen, die menschliche Körpersprache betreffend, ist die Erkennungssoftware theoretisch realisierbar und der praktischen Umsetzung stehen, momentan noch, zum einen die Rechenkapazität und zum anderen die präzise Deutung der menschlichen sozialen Signale im Weg.

Auf Basis der Forschungen und auf ersten Erfahrungen mit dem Prototyp ist es möglich eine Software zu entwickeln, die die User-Konstellationen und deren sozialen Kontext wahrnimmt. Das Problem, welches oft auch Fehlverhalten in der Mensch-Mensch-

Interaktion hervorruft, ist die Interpretation der menschlichen sozialen Signale. An dieser Stelle ist die Zusammenarbeit mit Sozialwissenschaftlern unerlässlich, die den sozialwissenschaftlich notwendigen Hintergrund beleuchten. Durch diverse Nutzerstudien, die die Motivation und die Wahrnehmung der Probanden betreffen, können Fehler in der Implementierung der Wahrnehmung und Interpretation vermieden werden. Oft wird von der informatischen Seite zu Testzwecken ein Testszenario entworfen, welches auf lückenhaftem Wissen in dieser Wissenschaftsrichtung beruht. Im Rahmen der Zusammenarbeit werden die Ergebnisse der sozialwissenschaftlichen Studien einer technischen Machbarkeitsprüfung durch die Ingenieure unterzogen und somit kann die Umsetzung der sozialen Wahrnehmung eingeschränkt oder angepasst werden und eine bessere Deutung der nonverbalen Kommunikation stattfinden.

Das Modell in Abbildung 1 stellt die Phasen der Zusammenarbeit von Ingenieurwissenschaft und Sozialwissenschaft dar und ist angelehnt an das V-Modell von Boehm [BoB79].

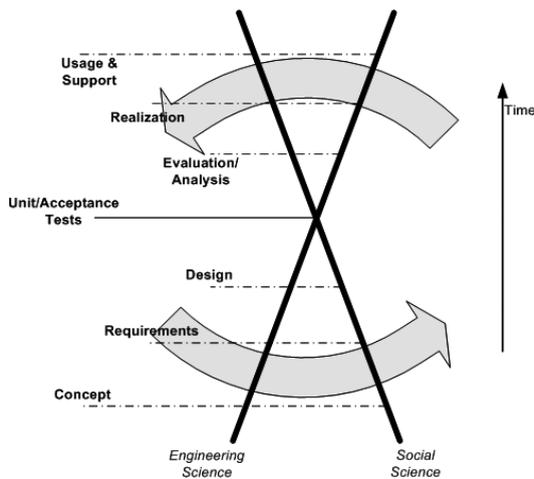


Abbildung 1: Interdisziplinäres CHI-Modell

Die zwei sich kreuzende Stränge stellen die beiden Wissenschaften dar. Vertikal ist die Zeit und horizontal sind die sieben Phasen der *gemeinsamen* Arbeit dargestellt, die parallel von beiden Wissenschaften ausgeführt werden. Die sich kreuzenden Stränge zeigen, dass die Wissenschaften von unterschiedlichen Ausgangssituationen ausgehen und sich auch die Endpunkte der wissenschaftlichen Arbeit unterschiedlich gestalten. Eine Zusammenarbeit soll keine „Hybridwissenschaft“ erschaffen, beide Wissenschaften sollen ihre Entwicklung in Reinform betreiben. Dazu gehört, dass beide Wissenschaften sollen ihre Basis behalten: Die Ingenieurwissenschaft sieht die Technik und diverse Möglichkeiten und Szenarien, die realisiert werden können, die Sozialwissenschaft sieht den Menschen im Mittelpunkt und die Möglichkeiten, ihm den Umgang mit der Technik zu erleichtern. Durch die Zusammenführung der Erkenntnisse können die nächsten Phasen geplant und abgestimmt werden und so zu einem optimalen Ergebnis führen. Ohne eine längerfristige gemeinsame interdisziplinäre Zusammenarbeit kann eine Wahrnehmung und Interpretation der sozialen Signale nicht eindeutig realisiert werden.

## 2.2 Softwaretechnische Umsetzung

Der Prototyp der Erkennungskomponente arbeitet mit einer HD-Kamera und Infrarotsensoren (IR-Sensoren). Um den Platz der Systeme nicht zu strapazieren, habe ich mich für die raumsparende Umsetzung mit einer Kamera, einem IR-Projektor und einer IR-Kamera entschieden. Wobei eine Variante welche wenigsten zwei Seiten überwacht mehr und zuverlässigere Daten geliefert hätte. Da ein solcher Aufbau aber selten in öffentlichen Räumen realisierbar ist, entschieden wir uns dagegen. Der erste Schritt ist die Festlegung des Aktionsbereiches: Der überwachte Interaktionsbereich von 1 m bis 3,5 m folgt den Beobachtungen HALLS [HE82], der, aufgrund seiner Studien, diesen Abstand zum möglichen Gesprächspartner als Interaktionsbereich festlegt. Im zweiten Schritt der Bilderkennung folgt die Wahrnehmung des Menschen an sich. Wird kein Mensch bzw. keine menschliche Statur erkannt, arbeitet die Bilderkennung nicht weiter, weil kein potenzieller Interaktionspartner in Reichweite ist. Erkennt das System einen Interaktionspartner, wird in der aktuellen Version des Prototypen, eine verbale Meldung an den möglichen User gerichtet. Diese Variante wurde gewählt, um zum einen die sozialwissenschaftlichen Studien zu realisieren und zum anderen die Aufmerksamkeit der Nutzer auf das System zu lenken. Daraus ergeben sich Vorteile, wie z. B.: Die Probanden können direkt befragt werden, wie „diese Art System“ auf sie wirkt und was sie denken, was es verfolgt (quantitative Sozialforschung), zusätzlich zur Beobachtung der Probanden (qualitative Sozialforschung). Der Vorteil der Aufmerksamkeitslenkung ist, der potenzielle Nutzer steht frontal zum System und bietet somit der Bilderkennung eine bessere Chance, den Nutzer weiter zu analysieren.

Für diese Analyse wird der Körper und seine Position im Raum bzw. die Position seiner Gliedmaßen zum Körper und der Gliedmaßen zueinander vermessen. Diese berechneten Abstände und Winkel werden mit hinterlegten Pattern verglichen. Die Interpretation, die das System vornimmt und die entsprechenden Ausgaben, erfolgen auf Grundlage der Ergebnisse der sozialwissenschaftlichen Studien. Wie weit die Erkennung mit diesem einfachen System und die Interpretation von sozialen Signalen gehen kann und wie hoch die Akzeptanz in den unterschiedlichen Erkennungsstadien ist, wird zurzeit noch getestet, ebenso wird an der Robustheit der Erkennung gearbeitet.

## 3 Die Zusammenfassung und Aussicht

Sozial empathische Systeme sollen Menschen aktiv ansprechen, die als mögliche Nutzer identifiziert werden. Damit soll das System eine höhere Nutzungsfrequenz erreichen. Außerdem wird durch die Erkennung der sozialen Signale ein angepasster und dadurch individueller Informationsaustausch ermöglicht. Durch das Einbeziehen der Informationen, die der Nutzer durch seine sozialen Signale vermittelt und ein dadurch erreichtes einfacheres Handling sowie eine bessere Datenbeschaffung kann der Nutzungsdurchsatz erhöht werden und damit die Wirtschaftlichkeit. Das aktive Einbeziehen der Nutzer verringert auch die beim Menschen vorhanden Hemmschwellen und öffnet somit auch technisch nicht versierten Menschen den Zugang zu interaktiven Systemen. Ziel der Entwicklung sind aufmerksame interaktive Systeme, die den potenziellen Nutzer aus der Ferne als diesen identifizieren und analysieren, um eine

individuelle Interaktion zu initiieren, den Nutzungsdurchsatz und die Wirtschaftlichkeit zu erhöhen.

„*Wissenschaft meets Wirtschaft*“ an dem Punkt, an dem die Analyse der sozialen Signale der Nutzer erfolgt ist und diese prototypisch implementiert sind, sowie umfangreiche Tests absolviert worden sind, deren Ergebnisse in die Weiterentwicklung eingegangen sind. Ist eine robuste Erkennung implementiert worden, kann der Schritt in die Wirtschaft erfolgen. Nach der gemeinsamen Anpassung an die Situation und die Bedürfnisse kann das System im Realraum eingesetzt werden. Die Inbetriebnahme und die ersten Erfahrungen mit dem System und der Erkennungssoftware bieten wiederum Ergebnisse, die in der Wissenschaft für die Weiterentwicklung und Neuentwicklung Verwendung finden können. Die Akzeptanz der Nutzer und die Wirtschaftlichkeit sind Ziele, die mit der Realisierung der Erkennung von sozialen Signalen erreicht werden können und dadurch für beide Seiten einen Vorteil bieten. Durch die sukzessive Einführung der neuen Eigenschaft ist es möglich, eine längerfristige Zusammenarbeit von Wissenschaft und Wirtschaft zu realisieren. Somit kann eine Interaktion entstehen, die durch einen ausgeglichenen „Wissensstand“ von Mensch und Maschine von beiden Seiten individuell gestaltet wird.

## Literaturverzeichnis

- [AI13] Champbell, M: AppleInsider; <http://appleinsider.com/articles/13/05/30/apple-shows-renewed-interest-in-gaze-detection-possibly-in-response-to-samsungs-smart-scroll> (last access: 2013-10-04).
- [ArM13] Argyle, M.: Körpersprache & Kommunikation, Nonverbaler Ausdruck und Soziale Interaktion. Paderborn: Junfermann Verlag 2013.
- [BoB79] Boehm, B. W.: Guidelines for verifying and validating software requirements and design specifications. Euro IFIP 1979. North Holland 1979.
- [DiE13] Dim, E. & Kuflik, T.: Social F-formation in Blended Reality. LAMBDa 2013, March 19-22, Santa Monica, CA, USA.
- [HE56] Hall, E. T.: The Hidden Dimension. Massachusetts Institute of Technology, 1956.
- [MaN12] Marquardt, N.: Cross-Device Interaction via Micro-mobility and F-Formations. UIST 2012, October 7-12, Cambridge, Massachusetts, USA.
- [MaP11] Marshall, P.: Using F-Formations to Analyse Spatial Patterns of Interaction in Physical Environments. Proceedings of the ACM 2011, New York.
- [MiD11] Michelis, D. & Müller, J.: The Audience Funnel: Observations of Gesture Based Interaction With Multiple Large Displays in a City Center. In: International Journal of Human-Computer Interaction. Vol. 27, Nr. 1, 2011.
- [MiD09] Michelis, D.: Interaktive Großbildschirme im öffentlichen Raum, Springer, 2009.
- [PeM94] Paetau, M; Rohde, M. & Wulf, V.: Das Maschinenmodell wird zum Auslaufmodell. In: Wechselwirkung, Zeitschrift für Technik, Naturwissenschaft, Gesellschaft, Jg.16, Nr. 69.
- [StJ98] Strübing, J.: Vom Nutzen der Mavericks, Zur Zusammenarbeit von Informatik und Soziologie auf dem Gebiet der Verteilten KI. KI-Tagung Bremen 1998.
- [TA13] Tallig, A.: A Robot Companion as mobile Edutainer. In: International Summerworkshop Computer Science 2013 Proceedings of International Summerworkshop 17.7. – 19.7.2013. Chemnitzer Science-Report CSR-13-04, July 2013.
- [TA13a] Tallig, A.: Creating an interactive Mixed Reality. HCI International 2014 (in progress).
- [WaP11] Watzlawick, P.: Menschliche Kommunikation. Bern: Hans Huber Verlag 2011.