

Meine Katze macht ‘nen super Latte

Schwierigkeiten ungebundener Bedienung - wenn Geräte sich ungewollt angesprochen fühlen

Kontaktangaben

Thom Scheiner

thom.scheiner@uid.com

Tobias Limbach

tobias.limbach@uid.com

User Interface Design GmbH

Claudius-Keller-Straße 3c

81669 München

Abstract

Sprach- und Raumgestensteuerung haben miteinander gemein, dass der Nutzer sie ohne physischen Kontakt bedient und sie voll auf ihre Sensorik setzen. Noch sind diese Technologien nicht so breit im Einsatz, dass Probleme durch Verwechslungen immanent würden; aber die Verbreitung dieser Generation von Produkten hat bereits begonnen – etwa mit dem Rauchmelder von Nest.

Wenn nun mehrere Geräte in der Küche stehen und ihre Sensoren Wirkräume aufspannen, kann es sprichwörtlich eng werden für Gestikulierer. Räume, in denen ein Gerät Gesten erkennt, können sich dabei durchaus überlappen. Eine „exklusive“ Bedienung wie mit Tastatur/Maus findet eben gerade nicht statt. Sprachsteuerung ist ebenfalls ungerichtet, jedenfalls wenn sie natürlichsprachlich ist. Zwar könnte man – wie bei Smartphones – vor jedem Kaffeebezug etwas sagen wie „Jura Impressa III: Kaffee“. Aber das würde den Komfort doch arg einschränken.

Im Workshop diskutieren wir, welche Probleme sich aus heutiger Sicht und auch künftig ergeben, und suchen Lösungsansätze.

Keywords

Natürlich Geste Sprache Steuerung Voice

Einleitung

Mai 2014 erfolgte eine Rückrufaktion für die „Nest Protect“-Rauchmelder der Firma Nest Labs (siehe [1], [2], [3]). Der Grund: eine fehlerhafte Funktion namens „Nest Wave“. Das Winken, mit dem man per Gestenerkennung beispielsweise beim Braten einen Rauchmelder kurzzeitig abschalten kann, damit diese keinen Fehlalarm durch Küchendünste verursachen, könnte auch von anderen Nest-

Rauchmeldern auf sich bezogen werden. Diese würden sich dann ebenso abschalten und sich später gegebenenfalls nicht wieder aktivieren.

Wer Maus oder Tastatur kontrolliert, steuert genau das, was dran hängt. Bei ungebundener Bedienung hingegen können im Prinzip mehrere Geräte dasselbe Signal erfassen, solange es in ihrem Rezeptionsbereich auftrat – egal ob das Signal für das Gerät gedacht war oder nicht. Im schlimmsten Fall weiß hier der Nutzer womöglich nicht, dass seine Laute und Gesten überhaupt erfasst werden. Das Gerät kennt die Intention des Nutzers nicht und fehlinterpretiert das Fuchteln, mit dem man eine Fliege verscheuchen wollte.

Der Unterschied zwischen Rezeptionsbereichen und Maus, Tastatur und Joystick

Gebunden vs. ungebunden

Auf dem Amiga 500 war alles noch ganz einfach: Eine Maus (serieller Anschluss), eine fest verbaute Tastatur, und ein bis zwei Joysticks an den Sub-D-Buchsen. Wer die Peripherie kontrollierte, kontrollierte damit auch den Rechner – eine andere Form der Einwirkung gab es nicht. Wenn zwei Nutzer gemeinsam am Rechner saßen, kontrollierte dennoch eigentlich nur einer das Geschehen. Der Rechner akzeptierte nur Signale von definierten Eingabegeräten, und diese Eingabegeräte waren sehr begrenzt vorhanden. Sie waren fest an den Rechner gebunden.

Bei Computern ist der Grundaufbau heute noch recht ähnlich, auch wenn man theoretisch recht viele Mäuse und Tastaturen (Kabel oder Funkempfänger) anstöpseln und damit eingreifen kann. In der Realität sind die meisten Computer mit einer Tastatur und einer, selten zwei Mäusen und vielleicht einem Grafiktablett ausgestattet – das reicht auch für die meisten Nutzer. Selbst Funkmäuse und -tastaturen sind ja durch das Pairing dem Computer eindeutig bekanntgemacht worden, zählen hier also auch zu den gebundenen Geräten.

Die ungebundene Bedienung beschreibt hingegen eine Form der Kontrolle, die nur noch eine räumliche, jedoch keine festgelegte oder vereinbarte (soz. „monogame“) Verbindung zwischen Mensch und Gerät benötigt. Eingaben gelangen hier über Sensoren und deren Rezeptionsbereich [Wirkbereich/Rezeptorfeld] in den Rechner. Dieser Rezeptionsbereich hat keine klar erkennbaren physischen Grenzen und kein standardisiertes Eingabegerät – anders eben, als bei Standardeingabegeräten wie Tastaturen und Funkmäusen. Es kann eine (nicht standardisierte) große oder kleine, linke oder rechte Hand sein, es geht aber auch mit Katzenpfote oder fünf Bockwürsten.

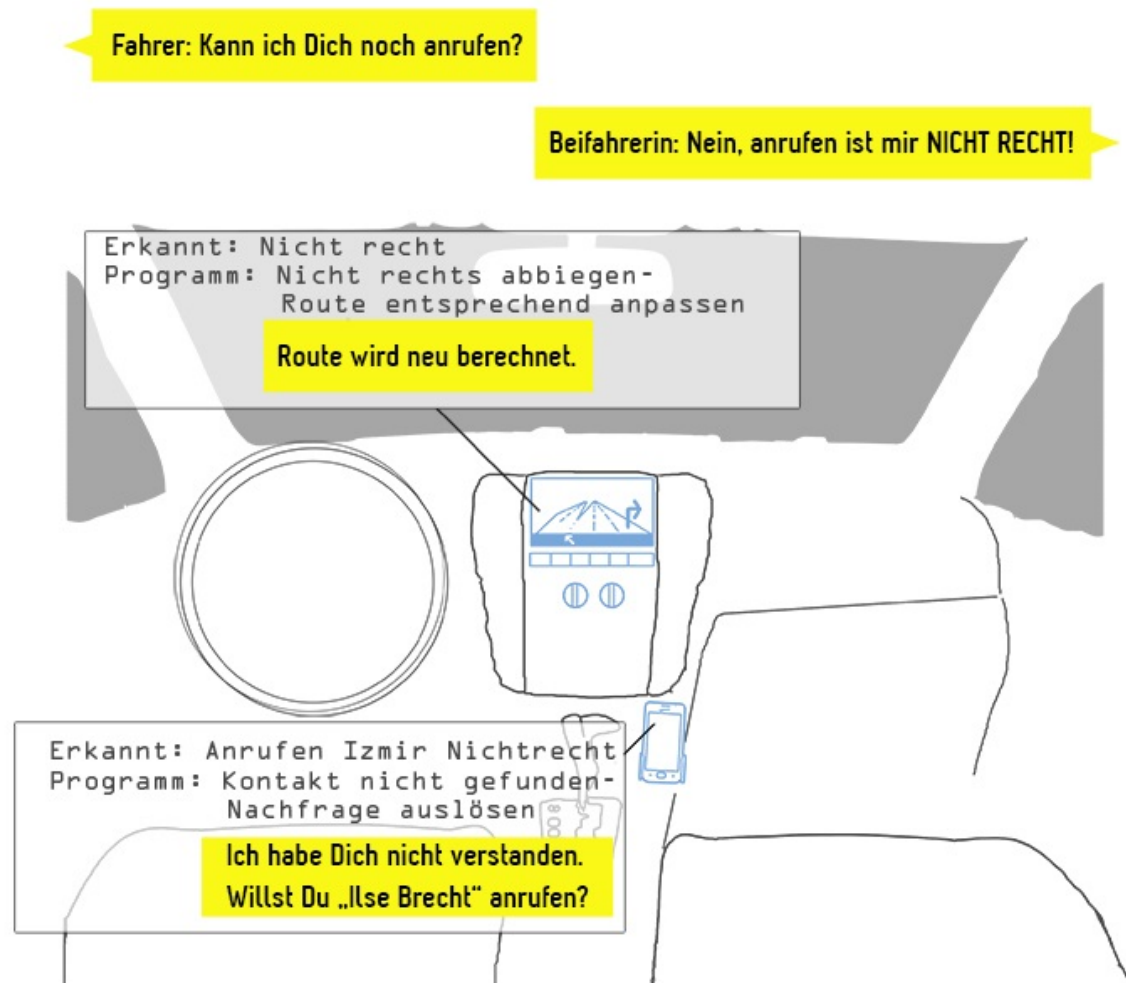


Abbildung 1: Ungebundene Bedienung per Sprachkommando: Das Auto-Bordsystem und das Smartphone hören mit, während sich Fahrer und Beifahrer unterhalten. Und beide versuchen auszuwerten, was da so gesagt wird.

Akustische vs. gestische Signale und Rezeptoren

Der Schwerpunkt bei ungebundener Bedienung liegt auf der akustischen und der gestischen Steuerung. Bei akustischer Steuerung werden Sprachkommandos oder andere Schallsignale (etwa Klatschen) ausgewertet. Bei gestischer Steuerung werden Bewegungen in einem zwei- oder dreidimensionalen Raum getrackt und ausgewertet.

Andere denkbare Rezeptoren (etwa für Temperaturen, Gerüche, Gesichtsausdrücke oder Hirnströme) werden hier nicht betrachtet.

Verbreitung von Geräten mit ungebundener Bedienung

Bekannte Beispiele sind:

- Gestische Sensorik
 - Leap Motion Controller / Myestro Controller
 - Diverse Spielkonsolen (Nintendo Wii, Sony Playstation Move, Microsoft Kinect)
 - Nest Protect
 - Kaffeemaschinen
- Akustische Sensorik

- o Navigationsgeräte
- o Kfz-Bordsysteme
- o Sprachsteuerung in Smartphones (etwa Apples „Siri“)
- o und schon seit den 1980-ern Wecker (etwa Braun Voice Control)

Die Katze und der Latte Macchiato– Fallstricke bei ungebundener Bedienung

Ungebundene Bedienung wirft eine Reihe von Schwierigkeiten auf und kann den erwarteten Nutz- oder zumindest Mehrwert eines solchen Gerätes schmälern. Die Rückrufaktion mit einhergehendem mehrmonatigem Verkaufsstopp für die „Nest Protect“ Rauchmelder belegt eindrücklich, dass schon ein kleines Einzelproblem bei der ungebundenen Bedienung starke – auch wirtschaftliche – Auswirkungen haben kann.

Stehen zwei Personen im Rezeptionsbereich eines entsprechenden Geräts und führen z.B. Gesten aus, so könnte jede der Gesten verwertbar sein. Welcher Geste soll Folge geleistet werden? Was tun bei widersprüchlichen Gesten? Was tun bei gleichen Gesten?

Rezeptionsbereiche sind unsichtbar ...

Das Feld, welches die Rezeptoren überwachen, ist unsichtbar. Bestenfalls erkennt man die Geräte als solche und weiß, dass man sich in ihrem Rezeptionsbereich mit einer ungefähr bekannten Größe und Form aufhält. Bei der LeapMotion, die etwas größer als ein USB Stick ist und am besten auf dem Tisch vor dem Bildschirm/Laptop platziert wird, ist dies noch gut möglich. Vertikal über der LeapMotion spannt sich unsichtbar ein Gestenraum auf, hier drin kann man interagieren. Wie groß genau der Raum ist, erkennt man durch Probieren. Bei anderen Gestenerkennern wie z.B. der Myestro ist dies so gut wie unmöglich. Der Myestro Gestenraum ist vom Sensor gesehen horizontal im Raum, und nur mit Versuch- und Irrtum zu finden.

Ähnlich ist es mit Sprachbedienung. Was heute noch eher kommandobasiert und schlicht herkommt („Kaffeemaschine! [Piep] „Cappuccino zwei, mit Zucker, doppelt [Piep]“) und eigentlich überhaupt keine Erleichterung ist, wird sich in Zukunft hin zu einer natürlich-sprachlichen Eingabe wandeln. Dann heißt es eher „Ich habe Lust auf ein Heißgetränk... [Kaffeemaschine und Teekoher spulen schon mal hoch, Kühlschrank ist enttäuscht]... aber ich hatte schon zwei Espresso gerade... [Tee freut sich]... also doch eine Tasse Earl Grey! [der ab jetzt schon im Kessel zieht].“

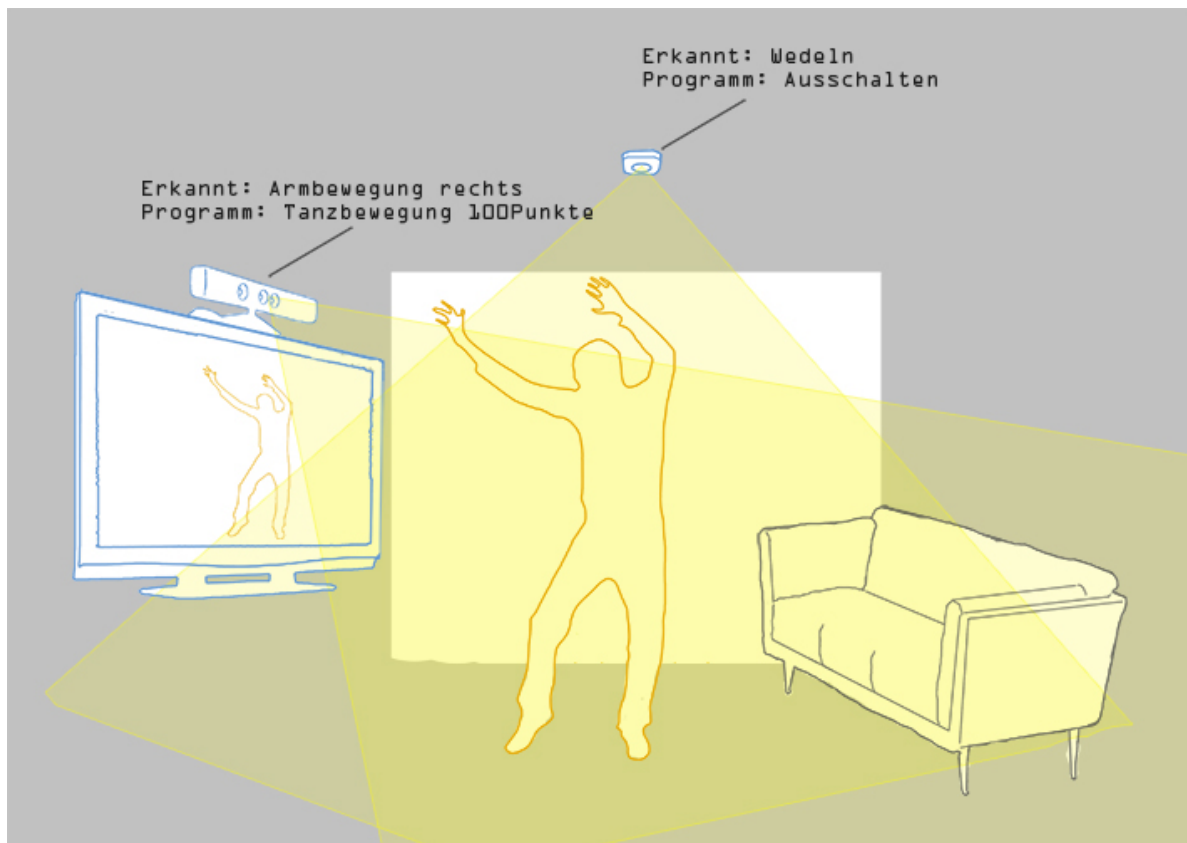


Abbildung 2: Rezeptionsbereiche sind oft unsichtbar und können sich überlappen. Im Wohnzimmer könnte man beim Spiel mit der Microsoft Kinect auch von anderen Sensoren erfasst werden – etwa vom Rauchmelder, der das Herumgefuchtel als Kommando zum Ausschalten interpretiert.

... und sollen manchmal auch verborgen sein

Beim Besuch bei Freunden oder im Büro von neuen Kunden weiß man nichts über deren Ausstattung. Heute ist das noch kein Problem, da die Geräte mit ungebundener Bedienung noch kaum verbreitet sind; bis zur nächsten Dekade dürfte sich das geändert haben. Dann spannen sich wohl an vielen Stellen unsichtbare Rezeptionsbereiche auf, die wir mit unseren Gesten und Worten reizen. Sind wir uns dessen dann bewusst? Schränken wir uns deshalb ein in dem was wir sagen oder tun, nur um nicht unbeabsichtigt etwas ein- oder auszuschalten? Werden wir als Nutzer dann zum Problemverursacher?

Sicher für Kinder ...

Vielleicht möchte mein Kind den Rauchmelder lieber abschalten, weil es fürchtet, ich könnte ihm das Zündeln verbieten; dabei sollte er grade dann seine Arbeit machen. Woher weiß das Gerät, wer es bedienen darf und wer nicht? Braucht es ein Rechte-Management mit verschiedenen Nutzerrollen und -berechtigungen?

... Katzen ...

Dass ein Gerät irgendeine Bewegung einer Katze als Aufforderung missversteht, einen Latte Macchiato zuzubereiten, ist zugegebenermaßen unwahrscheinlich. Dass ein schnöder Bewegungssensor dem Haustier für seine nächtlichen Streifzüge zuverlässig das Licht einschalten kann, ist hingegen schon längst Realität. Werden sich alle Haustiere aus der Gesteninteraktion zuverlässig herausfiltern lassen? Vielleicht habe ich ein seltenes exotisches Haustier, das so exotisch ist, dass mein Rauchmelder es gar nicht als Nicht-Mensch erkennt. Greift dann die

Hausratsversicherung vielleicht nicht, weil es mir (unberechtigterweise) den Rauchmelder abgeschaltet hat und deshalb beim Brand kein Alarm ausgelöst wurde?

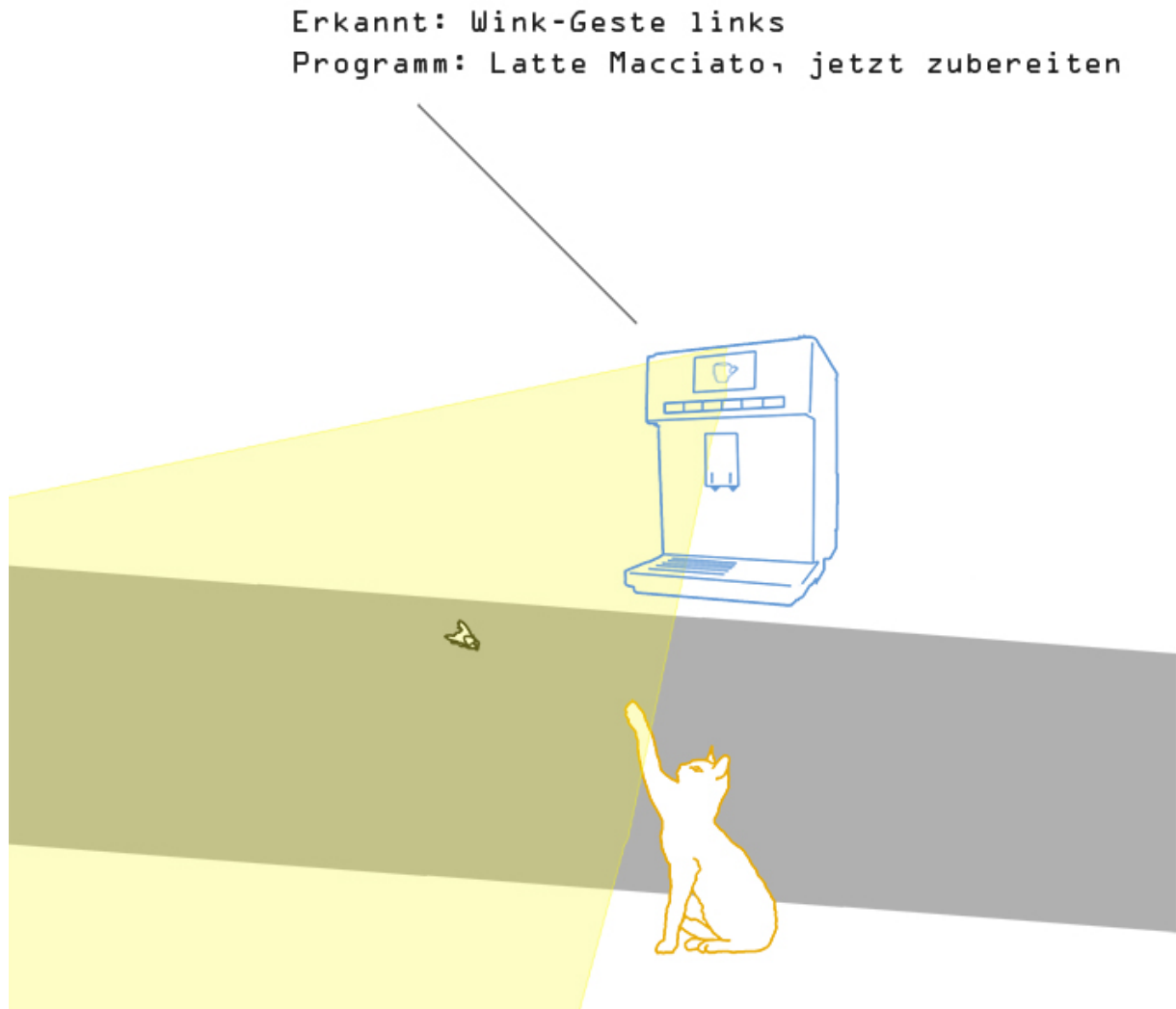


Abbildung 3: Die Katze will die Fliege verjagen. Wenn sie dabei im Rezeptionsbereich eines Geräts steht, wäre es gut wenn das Gerät sie als nicht berechtigtes Wesen erkennt und deshalb ihre Gesten ignoriert; sonst ist der Kaffeebohnenbehälter bald leer.

... und vor anderen Geräten

Auf die Spitze getrieben: Im Auto, grade raus aus Hamburg; mein Navi ist auf München programmiert, und im Radio läuft ein Hörspiel. Kurzer Halt an einem Parkplatz, um aus dem Kofferraum eine Flasche Wasser zu holen und meine Jacke reinzulegen. Hörspiel: „ blablabla gib mir bitte die Adresse von Anna!“ Nun werden im Hörspiel eine Minute lang Karteikarten durchsucht. Das Smartphone ist leider schneller und spricht die Adresse von Anna: „Kastanienweg 3, 10567 Berlin“; das Bordsystem quittiert „Die Route wird neu berechnet“. So, ich steige wieder ein, lasse den Motor an, und weiter geht's – nach Berlin. (Zu aus dem Kontext gerissenen Informationen aus Hörspielen siehe auch [4].)

Werden Sprachausgaben irgendwann so natürlich klingen, dass man maschinelle Aussprache nicht mehr von menschlicher Aussprache unterscheiden kann? Werden Geräte sich dann untereinander

darüber informieren müssen, was sie gleich auf für Menschen wahrnehmbaren Kanälen kundtun wollen – damit die anderen Geräte dies dann rausfiltern können? Welche Ausgabe-Kanäle sind davon betroffen – Sprache, Licht, ...?

Die Systeme müssen den Aktor und die Tragweite einer Aktion berücksichtigen. Die Tragweite einer Aktion lässt sich z.B. daran bemessen, wie unumkehrbar, risikoreich oder teuer die Aktion und seine Folgen sind.

- Prozesse geringer Tragweite: sofort ausführen
- Prozesse mittlerer Tragweite: Auf Abbruch-Möglichkeit hinweisen, dann nach Timeout ausführen (siehe Nest Rauchmelder)
- Prozesse hoher Tragweite: Individuell nachfragen, erst nach Antwort ausführen.

Beim Absender könnte das z.B. so aussehen:

- Katze = keine Reaktion
- Kind = nur Prozesse geringer Tragweite möglich
- Erwachsene = Prozesse großer Tragweite möglich
- Besonders autorisierte Personen = Ändern dieser Einstellungen

Herausforderung für UX

Vieles hängt natürlich von der Auflösung der Sensorik, Performanz und Robustheit des Rezeptionsbereichs ab, also von technischen Rahmenbedingungen. Doch wo ist dann das Handlungsfeld für UX Professionals? Hier gibt es einige:

- Verstehen der Nutzer und ihrer Bedürfnisse > Rückmeldung an die Produktentwicklung (Nutzungskontexte analysieren und verstehen)
- Bewusstes Gestalten der Rückmeldungen des Geräts (visuell, auditiv, haptisch, vielleicht gar olfaktorisch) für hohe Usability und positive UX
- Lösungsorientiertes Gestalten von Systemen, die sich ja in einen dynamischen Kontext befinden
 - Wie nah beziehungsweise weit muss ein Sensorfeld von der Kaffeemaschine entfernt sein, um komfortabel aber immer noch sicher bedienbar zu sein?
 - Wie kann man von vorneherein eine Überlappung von Sensorfeldern verhindern?
- Standards erarbeiten für
 - die Interaktion mit einzelnen und mehreren Geräten gleicher und unterschiedlicher Klassen (z.B. Herd = Sicherheitskritisch, daneben das Radio = nicht sicherheitskritisch)
 - die Art der Aktivierung („OK, glass: ...“) und den Umfang des Gesten- bzw. Wortschatzes (siehe auch [5])
- Klassifizierung von Interaktionen im Sinne einer Risikoanalyse
 - die Einordnung und Bewertung von Steuerkommandos in Abhängigkeit von Tragweite sowie Absender
 - Prozesse geringer Tragweite: sofort ausführen
 - Prozesse mittlerer Tragweite: Auf Abbruch-Möglichkeit hinweisen, dann nach Timeout ausführen (siehe Nest Rauchmelder)

- o Prozesse hoher Tragweite: Individuell nachfragen, erst nach Antwort ausführen.
- o Absender: Katze = keine Reaktion; Kind = nur Prozesse geringer Tragweite möglich; Erwachsene... autorisierte Personen...)

Der Workshop auf der UP14

Ziel dieses Workshops ist, das Bewusstsein für die unsichtbare Problematik zu schärfen; wir wollen künftige Produkte mitentwickeln können, die sich besser einpassen und nutzen lassen in einer Welt, die reich an Rezeptoren sein dürfte.

Im Workshop wollen wir diverse Szenarien aufbauen und diskutieren. Absurde wie künftig alltägliche Fälle dienen als Stimulus. Wir wollen ergründen, welche weiteren Probleme bei der Verwendung ungebunden bedienter Geräte auftreten könnten und wie wir als UX Professionals positive Nutzungserlebnisse gestalten können.

Literaturangaben

[1] SPIEGEL Online: <http://www.spiegel.de/netzwelt/gadgets/nest-rauchmelder-von-google-probleme-stoppen-verkauf-a-962524.html> (23. Mai 2014)

[2] SPIEGEL Online: <http://www.spiegel.de/netzwelt/gadgets/nest-rauchmelder-mit-problemen-rueckruf-fuer-440-000-geraete-a-971018.html> (22. Mai 2014)

[3] Nest Labs: <https://nest.com/letter-from-the-ceo/> (23. Mai 2014)

[4]

Wikipedia: http://de.wikipedia.org/wiki/Der_Krieg_der_Welten#Wirkung_des_H.C3.B6rspiels_von_1938 (30. Juni 2014)

[5] Apple: <https://support.apple.com/kb/HT3597> (23. Mai 2014)

Vita/Viten

Thom Scheiner (ist auch im Beitrag #96 enthalten)



Thom Scheiner ist Kommunikationspsychologe (FH) und Senior Usability Engineer in der Münchner Geschäftsstelle der User Interface Design GmbH (UID). Seit 2005 leitet und bearbeitet er vornehmlich Projekte aus der Industrie-Branche in In- und Ausland sowie Enterprise- und Web-Projekte.

Seine Aufgaben umfassen Projektmanagement, Koordination internationaler Studien, Anforderungsmanagement, Konzeptentwicklung, Evaluation und Dokumentation.

Thom Scheiner wirkt als Autor an der Neuauflage der Richtlinie VDI/VDE 3850 "Gebrauchstaugliche Gestaltung von Benutzungsschnittstellen für technische Anlagen" mit an Blatt 1 und verantwortet mit Blatt 3 die Neuauflage der Richtlinie für die Gestaltung von Touchscreen-Oberflächen.

Tobias Limbach

Tobias hat seine separat geschickt.