Gefördert vom

bmb+f

Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie

Mensch-Technik-Interaktion

SmartKom

Auswertung der ergonomischen Evaluation des SmartKom Prototypen

–Szenario Home–

Befragungszeitraum 01. April 02 bis 31. Juni 02

Ute Kartal, Silke Steininger, Nicole Beringer, Daniel Sonntag

Ludwig Maximilians Universität München

Memo Nr. 12

Dezember 2002

Dezember2002

Auswertung der Ergebnisse der Evaluations–Aufnahmen Szenario Home

Fragebogenversion vom 01.04.02

Ute Kartal, Silke Steininger, Nicole Beringer, Daniel Sonntag

Ludwig Maximilans Universität München
Schellingstr. 3
80799 München

Tel.: (089) 2180–5751
FAX: (089) 2800362

E–Mail: {ukartal,kstein,beringer,sonntag}@phonetik.uni–muenchen.de

Dieses Memo gehört zu Teilprojekt 7: Intuitives Arbeiten mit dem Computer:
SmartKom Home/Office; AP 7.5 Ergonomische Evaluation

Das diesem Memo zugrundeliegende Forschungsvorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01 IL 905 gefördert. Die Verantwortung für den Inhalt liegt bei den Autoren.

Inhaltsverzeichnis

1. Befragungszeitraum

Die Befragung wurde während des Evaluationszeitraumes für das Szenario Home vom 01.04.02 bis einschließlich 31.06.02 durchgeführt. Diese Memo befaßt sich ausschließlich mit den Befragten, die den SmartKom Prototypen Home getestet haben.

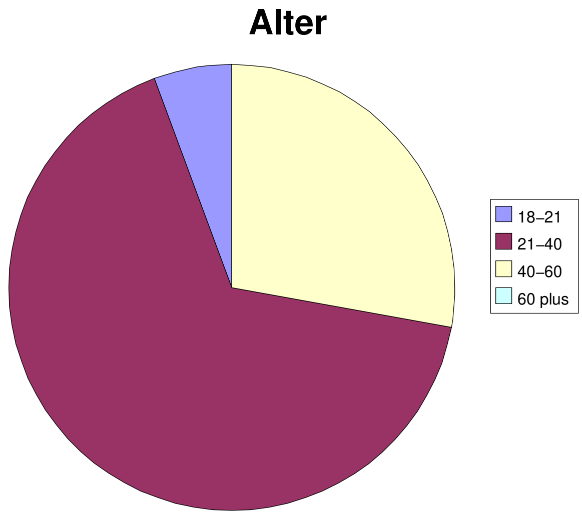
2. Anzahl und Zusammensetzung der Befragten

Insgesamt wurden 28 Personen befragt, bei einer Person existiert aufgrund technischer Schwierigkeiten kein Fragebogen. Die folgenden Daten beziehen sich damit auf 27 gültige Fragebögen zur Evaluation des Home–Prototypen.

Alter

In Anlehnung an die Alterskategorien des Statistischen Bundesamtes wurde folgende Verteilung erreicht:

Alter	***Prozent***	***Verteilung Statistisches Bundesamt*** *(in Prozent)*	***Anzahl Vpn***
14–18	33.3	5.2	9
18–21	3.7	4	1
21–40	44.4	32.7	12
40–60	18.5	31.2	5
60 plus	0	26.9	0
Summe	99.9	100	n=27



Hier wird ersichtlich, daß die erstrebte Rekrutierung der Versuchspersonen entsprechend der Altersstruktur der Bevölkerung der Bundesrepublik Deutschland nicht gelungen ist. Dies ist auf die erschwerten Bedingungen beim Aufbau des Systems SmartKom zurückzuführen, welche eine Nutzung durch komplett naive Versuchspersonen zu Beginn des Evaluationszeitraumes Home ausschloß.

Geschlecht

13 weibliche und 14 männliche Versuchspersonen teilen die Befragten ungefähr paritätisch.

Staatsangehörigkeit

Nur 2 der 27 Befragten gaben an, keinen deutschen Paß zu besitzen. Dies sind 7.4% aller Befragten dieser Welle. Damit ist die angestrebte Verteilung von ca. 8% „Inhaber eines nicht–deutschen Passes" erfüllt.

Schulabschluß

Entsprechend der Kategorien des Statistischen Bundesamtes verteilten sich die Befragten wie folgt auf die verschiedenen Schulabschlüsse:

	Prozent	***Verteilung Statistisches Bundesamt*** *(in Prozent)*[1]	***Anzahl Vpn***
Volks–/ Hauptschule	14.8	50.5	4
Realschule/ gleichwertiger Abschluß	0	28.7	0
(Fach–) Hochschulreife	66.7	20.8	18
Ohne Abschluß	18.5	–	5
Keine Angaben	0	–	0
Sonstige	0	–	0
Summe	100	100	27

[1]Verteilung beruht auf Personen mit allgemeinbildendem Schulabschluss. Personen ohne Abschluss wurden hier nicht erfaßt.

Erfragt wurde in diesem Zusammenhang der zuletzt erreichte allgemeinbildende Schulabschluss. Dies ist nicht zu verwechseln mit akademischen Graden oder Ausbildungsabschlüssen, die in dieser Verteilung nicht aufgelistet sind.

Wie oben bereits erwähnt, war eine Rekrutierung der

Versuchspersonen ausschließlich nach statistischen Werten in dieser Phase der Evaluation nicht praktikabel. Deshalb differieren die Werte erheblich zur angestrebten Verteilung entsprechend den Daten des Statistischen Bundesamtes.

Sonstige Variablen

Mit 25 Personen sind die meisten Probanden deutsche Muttersprachler. Zwei Versuchspersonen gaben an, jeweils Englisch und Japanisch als Muttersprache zu haben. Dies entspricht, in Anlehnung an die oben bereits diskutierte Nationalität, unseren Zielvorgaben.

3. Wichtige Faktoren der ergonomischen Evaluation

1. Zugang zu einem Computer

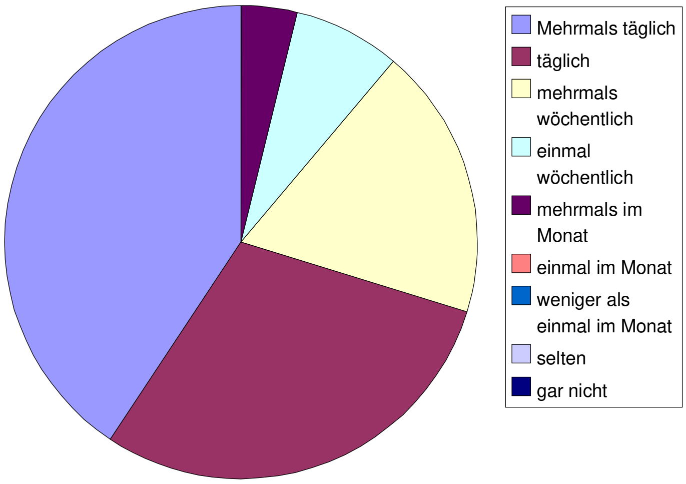
26 Benutzer gaben an, Zugang zu einem Computer zu haben. Eine Person verweigerte die Antwort.

2. Nutzungshäufigkeit eines Computers

Wie auch bei den vorherigen Fragebögen wird die Nutzungshäufigkeit eines Computers mit einer Skala von „mehrmals täglich" bis „gar nicht" erfragt. Hierbei sah die Verteilung wie folgt aus:

	Prozent	***Anzahl Vpn***
Mehrmals täglich	40.7	11
täglich	29.6	8
mehrmals wöchentlich	18.5	5
einmal wöchentlich	7.4	2
mehrmals im Monat	3.7	1
einmal im Monat	0	0
weniger als einmal im Monat	0	0
selten	0	0
gar nicht	0	0
Summe	**99.9**	**27**

Nutzungshäufigkeit eines Computers

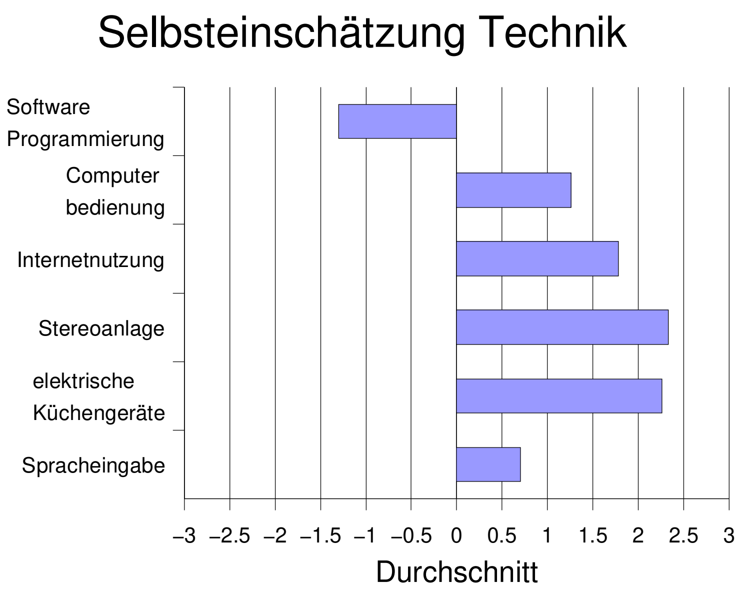


Wie auch bei den letzten Befragungswellen, ist die Mehrzahl der Befragten intensiver Computernutzer. Dies hängt sicherlich damit zusammen, daß, trotz intensiver Bemühungen bei der Rekrutierung, Personen mit Hemmungen gegenüber Eltektronik seltener zu einem Computertest zu bewegen sind.

3. Selbsteinschätzung im Bereich Technik

	Durchschnitt
Software Programmierung	−1.3
Computer bedienung	1.26
Nutzung des Internets	1.78
Bedienung der Stereoanlage	2.33
Bedienung elektrischer Küchengeräte	2.26
Bedienung eines Systems über Spracheingabe	0.7

(Skala von +3=hohe Kompetenz bis −3=niedrige Kompetenz)



Diesmal ist eine etwas bessere Selbsteinschätzung bei der Kompetenz zur Softwareprogrammierung als bei der letzten Welle zu verzeichnen. Allerdings ist dies immer noch der negativste Wert der ganzen Reihe, wie zu erwarten war.

Auch diesmal scheinen sich die Probanden nicht darüber im klaren zu sein, daß die Bedienung eines Systems über Spracheingabe eigentlich weniger technische Kompetenz erfordert als die Bedienung elektrischer Küchengeräte. Dieses immer wieder auftretende

Phänomen, daß Probanden ihre Kompetenz der Bedienung eines Systems über Spracheingabe als schlechter einschätzen als die Bedienung ihrer Küchengeräte, ist vielleicht darauf zurückzuführen, daß ersteres noch nicht weit verbreitet ist. Vermutlich ist diese schlechte Selbsteinschätzung damit nur auf Schüchternheit und Unwissenheit als auf tatsächliche Kompetenz zurückzuführen.

4. Emotionserkennung vor dem Experiment

Die Probanden wurden vor dem Experiment darüber befragt, ob Ihnen die Aufzeichnung ihrer Mimik und Gestik zur Analyse der Emotionen angenehm ist, oder ob sie Vorbehalte gegenüber dieser Art der Interaktion zeigen. Diese Frage wird in ähnlicher Form nach Ablauf des Experimentes wiederholt; daraus lassen sich dann eventuelle Einstellungsänderungen durch die Nutzung von SmartKom erkennen.

Diesmal ergab sich eine Neigung zu leicht positiver Einstellung gegenüber der Emotionserkennung.

Einstellung zu Emotionserkennung vor Experiment	***Prozent***	***Anzahl***
positiv	38.5	10
neutral	30.8	8
negativ	30.8	8
Summe	100.1	26
keine Antwort		1

Der Mittelwert dieser Verteilung(mit positiv=+1, neutral=0, negativ=–1) liegt nun bei 0.08 sehr gering auf der positiven Seite.

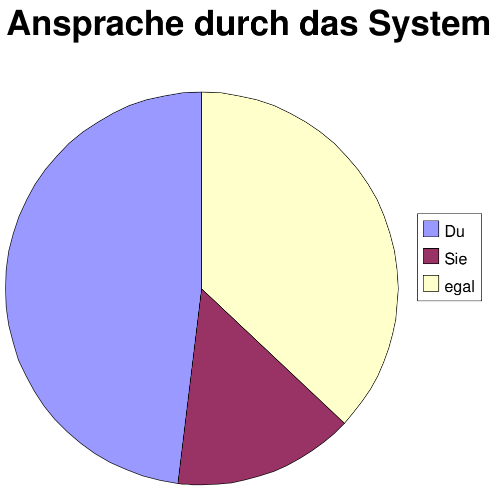
Änderungen in der Einstellung zur Emotionserkennung

In dieser Befragungswelle stellte sich heraus, daß nur eine einzige Versuchsperson angab, ihre Einstellung gegenüber Emotionserkennung geändert zu haben. Dies erfolgte von negativ zu positiv. Die überwiegende Mehrheit der Befragten (19) gaben an, keine Einstellungsänderung zu Emotionserkennung im allgemeinen erfahren zu haben. Allerdings wurden sehr häufig negative Kommentare bezüglich der Emotionserkennung des SmartKom Prototypen geäußert. So wurde der Kommentar „Ich hatte nicht das Gefühl, daß meine Emotionen erkannt wurden“ so oder in leicht abgewandelter Form allein sieben Mal geäußert. Die acht kommentarlosen „nein“ – Antworten relativieren die vorher genannte Anzahl zu rund der Hälfte aller Kommentare.

5. Ansprache durch das System

An dieser Stelle wurde abgefragt, wie der Proband von dem System angesprochen werden wollte. Hierbei zeigte sich folgende Verteilung:

	Prozent	*Anzahl*
Du	48.1	13
Sie	14.8	4
egal	37	10
Summe	99.9	27



Die Mehrheit der Befragten gab also an, mit „Du" tituliert werden zu wollen. Die zweithäufigste Antwort, daß es den Probanden „egal" ist, wie sie angesprochen werden, kontrastiert die seltene Präferenz „Sie" beträchtlich. Natürlich ist dies auch mit der Struktur der Befragten zu kontrastieren. Da der Hauptanteil der Befragten in der Altersgruppe 21–44 Jahre zu finden ist, ist diese starke Präferenz für „Du" auch verständlich.

6. Einige allgemeine Fragen zum System

Hierbei wurden Fragen zu einigen Aspekten von SmartKom gestellt. Auch szenariospezifische Detailfragen wurden implementiert.

Auf die Frage „Bitte geben Sie an, wie sehr sie folgenden Aussagen zustimmen" hatten die Probanden jeweils einen Aspekt von SmartKom von +3 (stimme zu) bis –3 (stimme nicht zu) zu bewerten.

Die Mittelwerte der jeweiligen Aspekte verteilen sich wie folgt:

Aspekt	*Mittelwert*
Die Aufgabe war einfach zu lösen	0.04
SmartKom hat mich verstanden	−1.37
SmartKom ist einfach zu verstehen	0.52
SmartKom hat zum größten Teil passende Antworten gegeben	−1.15
Die Geschwindigkeit des Systems war angemessen	−2.15
Ich wußte während des Dialogs immer was ich sagen konnte	−0.56
Die Benutzung von SmartKom wäre für mich eine Erleichterung	−0.93
Die Eingabe über Zeigen hat funktioniert	−0.59
Die Eingabe über Sprache hat funktioniert	−1
Ich würde SmartKom gerne nutzen	−0.52
SmartKom reagierte schnell auf meine Anfragen	−2.33

Auffällig hier ist die durchweg negative Bewertung des Systems SmartKom im allgemeinen. Die einzigen zwei positiven Durchschnittswerte beziehen sich auf die Aufgabe und die Verständlichkeit des Systems, wobei selbst die positive Bewertung nicht sonderlich extrem sind.

Szenariospezifische Detailfragen

	Mittelwert	*Anzahl nicht genutzt*	*Keine Antwort*
Das Telefonieren hat gut funktioniert	0	23	1
Die Kinoauskunft entprach meinen Wünschen	–0.11	22	1
Das TV–Programm lies sich einfach aussuchen	–0.93	2	0
Die Programmierung des Videorekorders war einfach	–0.3	20	1
Die Stifteingabe hat gut funktioniert	–0.04	3	0

Wie hier bei durch die Anzahl der „nicht genutzten" Features leicht zu erkennen ist, blieben die meisten Probanden bei dem Aussuchen der Fernsehsendungen stecken. Andere Aspekte von SmartKom konnten sie wohl aus Zeitgründen nicht testen. Bedingt durch die meistens sehr niedrige Anzahl Versuchspersonen, die ein Feature genutzt haben, wird an dieser Stelle darauf verzichtet, diese Ergebnisse hinsichtlich der Bedeutung der Mittelwerte zu kommentieren.

7. Semantisches Differential zur Funktionalität von SmartKom

In diesem Abschnitt wurden diverse Aspekte zur Funktionalität von SmartKom abgefragt. Hier werden nun die Mittelwerte der siebenstufigen Antwortskala widergegeben. Auch hier sind positive Aussagen mit positiven Werten verknüpft. Die Skala reicht von +3 (Neigung zur positiven Antwort) bis zu –3 (Neigung zur negativen Antwort). Unentschiedenheit kann durch ankreuzen der mittleren Antwortmöglichkeit ausgedrückt werden.

Frage	Mittelwert (Skala von positiv=+3 bis negativ=–3)
Einfachheit der Bedienung	–0.07
Informationen sind in Anzahl und Qualität angemessen	–0.37
Eingaben nur einmal nötig	–1.7
Hilfefunktion ist gut	–1.11
Display ist übersichtlich	**1.33**
Reaktion der Spracherkennung ist schnell	–1.96
Reaktion der Gestikerkennung ist schnell	–0.96
Erlaubt Barge–In	–1.37

Wieder überwiegen die negativen Bewertungen des Systems. Allerdings ist hierbei die durchaus positive Bewertung der Übersichtlichkeit des Displays (hervorgehoben) zu beachten.

8. Schwierigkeiten bei der Bedienung von SmartKom und Verbesserungsvorschläge

Diese offene Frage sollte die Versuchsperson animieren, die auffälligsten konzeptuellen Probleme von SmartKom schonungslos offen zu legen. Folgende Probleme wurden genannt (Anzahl Nennungen in Klammern; Mehrfachnennungen möglich); Zitate der Probanden werden aufgelistet:

SmartKom versteht nicht (13): SmartKom reagierte nicht auf meine Befehle.

Das System ist langsam (12): Reaktionsgeschwindigkeit zu gering.

Fehlende Funktionen (8): Zeigefunktion ging nicht; System gab falsche Antwort; Hilfe fehlte; Lautstärkeregelung nicht möglich

Schlechte/fehlende Dialogführung (12): Ich wusste nicht, wann Smartkom arbeitet und wann es auf eine Eingabe wartet; Smartakus Verhalten war nicht eindeutig zu verstehen (""Hand hinter das Ohr"" könnte dazu führen, dass die Spracheingabe wiederholt wird).

Unsicherheit der Versuchsperson (2):man wusste nicht genau, was das Programm hören wollte

System instabil (1): Hat nicht sofort reagiert, ist nach einiger (Zeit) abgestürzt

Sonstiges (1): noch nicht ausgereift und unkoordiniert, die angestrebte Software ist noch sehr verbesserungsfähig

Im Anschluß an diese Frage wurde nach verbesserungswürdigen Elementen von SmartKom gefragt. Diese wurden genannt (Anzahl Nennungen in Klammern, Mehrfachnennungen möglich); einige Zitate der Probanden werden aufgelistet:

Spracherkennung (10)

Geschwindigkeit (16)

Stifterkennung (4)

Mehr Funktionen (2): Das Angebot sollte vielseitiger sein und das System schneller reagieren, sonst kein überwiegender Vorteil gegenüber Fernsehzeitschrift

Zwischenmeldungen zu Systemzustand (10): Es wäre gut, wenn mehr Zwischenstandmeldungen kommen würden, z.B. System sucht oder System hat gefunden etc. Die Gestik des Männchens ist zwar gut gedacht, aber zu unauffällig. Bis man merkt, dass es einem etwas durch Körpersprache mitteilen will, ist es bereits zu spät und man ist leicht genervt.

Bessere Hilfe (2)

Sonstiges (4): alle Elemente sollten verbessert werden; Name Smartakus sollte geändert werden; Bedienung ist schlecht;

9. Kommunikationsmöglichkeiten mit SmartKom

Die überwiegende Mehrheit der Probanden (17) gab an, mit den Möglichkeiten der Kommunikation mit SmartKom zufrieden zu sein. Nur 6 Versuchspersonen waren unzufrieden und 4 gaben an, nicht zu wissen, ob ihnen die Kommunikationsmöglichkeiten ausreichend waren. Dies ist ein positives Bild. Die Frage nach gewünschten zusätzlichen Kommunikationsmitteln ergibt folgendes Bild:

	Anzahl Probanden mit Ja
Telefonhörer	1
Tastatur	3
Maus	6
Drucker	2
Sonstiges	2

Unter „Sonstiges" fiel zum Beispiel der Wunsch eines Probanden nach „Mehr Buttons für Eingabe via Stift". Ansonsten ist doch erstaunlich, daß sich so viele Probanden eine Maus als Kommunikationsmittel wünschen. Anscheinend ist der mitgelieferte Stift nicht Konkurrenz oder Ersatz genug für die Maus.

10.SmartKom eher Mensch oder Maschine

Als Neuerung wurde in dieser Welle die Frage nach der „Menschlichkeit" des Systems aufgenommen. Ziel war es, herauszufinden, ob die Interaktion mit SmartKom mit der eines Menschen vergleichbar ist. Unterschwelliges Ziel bei einer „vereinfachten" Bedienung eines Systems wie SmartKom ist es ja, eine Kommunikation fast wie mit einem menschlichen Gegenüber zu gewährleisten.

Mit einem Durchschnitt von –1.85, in einer Skala von +3 eher wie ein Mensch bis –3 eher wie eine Maschine, votieren die Probanden für die Maschinenähnlichkeit von SmartKom.

Nach einer Begründung für ihre Antwort gefragt, gaben die meisten Probanden an, die Qualität der Sprachausgabe wäre zu schlecht, um menschlich zu erscheinen (10). An zweiter Stelle rangieren die schlechten Reaktionen des Systems (7); z.B. „Die Antworten sind so standardisiert, dass man sofort das Gefuehl bekommt, mit einer Maschine zu komminizieren." Einige Probanden nervte der Avatar (3): „Kein kompetentes Verhalten von Smartakus (unangemessenes Feedbackverhalten: Abzählen der Finger, Hand hinters Ohr). Zu langsame Antwortzeiten. Mittelmässige Sprachsynthese". Positives Feedback bezüglich Stimme und Avatar gab es zwei mal:" Aber das Maennchen auf dem Bildschirm laesst sie zumindest etwas lebhaft erscheinen."

11.Schwierigkeiten in der Rolle zu bleiben

Erstaunlicherweise spaltet sich die Gruppe der Befragten fast zur Hälfte in Personen, die Schwierigkeiten hatten, in der Rolle zu bleiben (12 Vpn) und solche, die angaben, ohne Probleme ihre Aufgabe gelöst zu haben (15 Vpn). Dies steht im krassen Gegensatz zu den Ergebnissen der Befragung im Rahmen der WOZ–Experimente. Dort hatte die überwältigende Mehrheit der Versuchspersonen angegeben, keine Schwierigkeiten gehabt zu haben, in der Rolle zu bleiben. Der Schluß liegt nahe, daß die Versuchspersonen, durch die Bedienung des Systems verunsichert, sich genötigt sahen, Hilfe zu erbitten. Die anschließend erbetenen Ausführungen verdeutlichen dies. Hier einige Beispiele:

–"weil es einen nie verstanden hat war es schwierig in der Rolle zu bleiben";

–„Als ich das System einigermaßen begriffen hatte, war die Testzeit schon vorbei. Ich hatte das Gefühl, gar nicht so weit gekommen zu sein, die Vorteile des Programms überhaupt nutzen zu können.";

–„Weil es zu lange dauerte bis die Maschine reagierte und ich nicht wusste, ob ich denn nun überhaupt weitermachen, oder warten, oder wiederholen sollte."

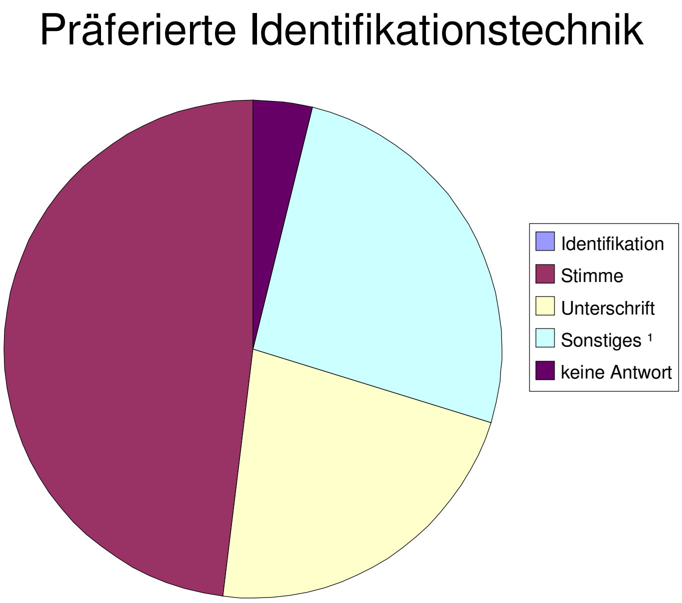
12.Identifikationstechniken

An dieser Stelle wurden die Versuchspersonen nach ihren Präferenzen zur Identifikation gefragt. Zur expliziten Auswahl standen die Anmeldung über Stimmerkennung oder per Unterschrift. Da die Frage offen gestellt wurde, sind auch andere Möglichkeiten der Identifizierung nennbar, die Fragestellung ermuntert dazu jedoch nicht.

Hier ist die Verteilung der Antworten:

Identifikation	***Prozent***	***Anzahl***
Stimme	48.1	13
Unterschrift	22.2	6
Sonstiges [1]	25.9	7
keine Antwort	3.7	1

[1]Sonstiges = undefinierbare Antworten



Wie aus der Grafik ebenfalls ersichtlich, ist auch dieses Mal die Stimmerkennung die am meisten präferierte Identifikationstechnik. Erstaunlicherweise ähneln sich die Begründungen für die Präferenz der jeweiligen Identifikationstechnik stark. Die Befürworter der Stimmidentifikation begründen dies mit der einfachen und schnellen Handhabung und dass die Unterschrift fälschbar und fehlerbehaftet sei, die Stimme dagegen schwerer. Die Befürworter der Unterschriftenidentifikation halten dagegen, daß die Stimmerkennung unzuverlässig ist, die Stimme verändert sein kann und die Spracherkennung in diesem System schlecht sei. Ein interessanter Vorschlag seitens der Probanden ist, eine Kombination aus Identifikation mit Stimme und Unterschrift anzubieten, da eine Kombination schwerer fälschbar sei.

13.Potentielle Benutzergruppe

Diese Frage beschäftigt sich mit der von den Versuchspersonen eingeschätzten Verwendungsmöglichkeit und –wahrscheinlichkeit des SmartKom–Produktes für verschiedene Benutzergruppen.

Benutzergruppe	**Mittelwert**
Leitende Angestellte	0.3
Unternehmer	**0.56**
Angestellte	0.37
Freiberufler	**0.59**
Arbeitslose	–0.56
Hausfrauen/–männer	0.04
Kinder/ Jugendliche	**1.44**
Studenten	1.04
Senioren	–0.59
Jedermann	–0.19
Für mich	0.26

Senioren und Arbeitslose liegen auch diesmal in der Einschätzung der potentiellen Benutzer weit abgeschlagen an letzter Stelle. Erfreulich ist die positive Einschätzung für Unternehmer und Freiberufler. Jugendlichen wird das größte Interesse an SmartKom zugeschrieben. Dies steht in Einklang mit Memo Nr.11, in welcher Kinder/Jugendliche ebenfalls an erster Stelle standen. Vielleicht ist dies auf das angewandte Szenario („Home") zurückzuführen. Wobei die verwendete Hardware, das Webpad, schon einen sehr „geschäftlichen" Eindruck vermittelte.

14.Einschätzung verschiedener Aspekte von SmartKom

Die Frage „Wie sehr gefallen Ihnen folgende Aspekte von SmartKom" konnte auf einer Skala von +3 (=gefiel mir sehr) bis –3 (=gefiel mir gar nicht) beantwortet werden. Dabei ergab sich folgendes Bild:

Aspekte	Mittelwert
Displayeingabe (allgemein)	0.11
Displayausgabe (allgemein)	**0.63**
Spracheingabe (allgemein)	–0.48
Sprachausgabe (allgemein)	0.22
Die Möglichkeit mit dem System „menschlich" zu interagieren	0.11
Das Design insgesamt	**0.96**
SmartKom reagiert auf meine Emotionen	–0.52

(Skala von +3=gefiel mir sehr bis –3=gefiel mir gar nicht)

Das Design insgesamt fiel sehr positiv auf, die Displayausgabe ebenso. Besonders negativ trat die Spracheingabe und die Emotionserkennung in Erscheinung. Alle weiteren erfragten Aspekte rangierten um den unspektakulären Nullpunkt.

15. Störendes an SmartKom

In der offenen Frage zu „Störendes am System Smartkom" wurde am häufigsten (16 Mal) die fehlende Geschwindigkeit genannt. Mit elf Nennungen rangiert der Kommentar, daß es nicht funktioniert habe, an zweiter Stelle. Die Persona wurde viermal kritisiert, die fehlende explizite Hilfefunktion zweimal. Gezählt wurden die Nennungen, wobei mehrere Nennungen in einer Antwort möglich sind. Hier nun einige Zitate zu den eben genannten Themen:

> **Geschwindigkeit:** „Dass er ewig lang nicht reagiert hat und ich mehrmals meine Anweisungen wiederholen musste, und dass ich nur bestimmt Sätze verwenden kann"; „Es dauert noch zu lange, man wird schnell ungeduldig"; „Es war zu langsam, und zu unpersönlich, eine Kommunikation war fast nicht möglich".
>
> **Fehlende Funktionalität:** „Er hat nicht wirklich funktioniert"; „Es reagierte nicht auf meine Befehle".
>
> **Persona:** „Langsamkeit fehlendes Verständnis die Persona nicht wirklich eine Hilfe – aufgrund der Langsamkeit des Systems eher nervend. Und wieso kaut sie Kaugummi Das ist unhöflich!"; „Die Ausgefeiltheit des animierten Männchens ist zwar ganz lustig

beim ersten Mal, aber für ein Benutzerprogramm, bei dem es letztendlich nur um Informationsbeschaffung geht, könnte es einem schnell auf die Nerven gehen"; „unpersönlich".

Fehlende Hilfefunktion: „schlechte Informationen über Bedienung".

16.Werteinschätzung des Systems in Euro

Danach gefragt, wieviel die Probanden denn für ein System dieser Art zahlen würden, wurde wieder vielfach moniert, daß das System ja gar nicht funktioniere. Der dennoch ermittelte Durchschnittswert beträgt 135,20 Euro, wobei jedoch zu berücksichtigen ist, daß vielfach Versuchspersonen nicht bereit waren, Geld für dieses System auszugeben. Viele „Null–Euro" Werte beeinflussen daher diesen Mittelwert.

17.Sehr schlecht bzw. sehr gut durchgeführte Aktionen

Häufig (7 Mal) wurde eine oder mehrere falsche Bildschirmausgaben bemängelt.

Dies hängt sicherlich mit der am zweithäufigsten schlecht durchgeführten Aktion zusammen, dem Verstehen von Befehlen. Die fehlende Zustandsanzeige lag gleich auf mit der Schnelligkeit, die ebenfalls zu wünschen übrig lies.

Die Darstellung des Fernsehprogrammes wurde als grafisch sehr gut beurteilt (11 Nennungen). Das Aussehen der Persona, die Sprachausgabe, die Begrüssung und die Aussage, dass nichts sehr gut durchgeführt wurde, lagen mit jeweils zwei Nennungen gleich auf bei den Antworten zu dieser Frage. Unerwartet viele Probanden (7) beantworteten diese Frage nicht.

18.Benotung einzelner Aspekte von SmartKom

Bei dieser Frage wurden die Probanden gebeten, ihrer Einschätzung der verschiedenen aufgelisteten Aspekte von SmartKom durch eine schulähnliche Notengebung Ausdruck zu verleihen. Die Skala reicht hierbei also von 1="sehr gut" bis 6="ungenügend".

	Inno–vation	Qualität	High–Tech	Spaß	Nützlichkeit	Gesamt
Displayeingabe	3.04	3.52	3.04	3.11	3.11	3.16
Displayausgabe	**2.85**	**2.67**	**2.81**	**2.96**	**2.63**	**2.78**
Spracheingabe	2.37	3.52	3	3.11	2.78	2.95
Sprachausgabe	2.74	3.11	3.04	3	3	2.97
Mit dem System „menschlich" zu reden	2.63	3.78	3.15	2.96	3.52	3.2
Das Design insgesamt	2.81	2.7	3.15	2.7	3.3	2.93
Versenden von emails						Nicht genutzt
Identifikation mit Unterschrift						Nicht genutzt
Identifikation über Stimmerkennung (N=7)	2.29	2.43	2.57	2.43	2	2.34
Telefonieren						Nicht genutzt
Stifteingabe	3.07	2.96	2.89	3.07	2.07	2.81
Programmieren des Videorekorders (N=2)	3.5	4.5	4	3.5	3.5	Nicht genutzt (N=25) 3.8
Auskunft über das Fernsehprogramm	**2.63**	**3.15**	**2.85**	**2.81**	**2.63**	**2.8**
Kinoauskunft						Nicht genutzt
Gesamt	**2.76**	**3.17**	2.99	2.96	2.88	2.95

Der Vollständigkeit halber wurden in dieser Tabelle alle abgefragten Aspekte aufgeführt, auch wenn diese von der Mehrzahl der Befragten oder allen nicht genutzt wurden. Die Aspekte „Identifikation über Stimmerkennung" und „Programmieren des Videorekorders" wurden von einer so geringen Anzahl Befragter benotet, daß eine ausführliche Auswertung an dieser Stelle unterbleibt.

In der Gesamtwertung der einzelnen Aspekte liegt die „Displayausgabe" mit einer Durchschnittsnote von 2.78 etwas vor der „Auskunft über das Fernsehprogramm" mit 2.80.

Gerechnet über die verschiedenen Aspekte, benoteten die Probanden die Innovation des Konzeptes SmartKom am besten. Am schlechtesten schnitt jedoch die Qualität ab. Insgesamt erscheint damit auch die Gesamtnote von 2.95 ins Bild zu passen.

In dieser Aufstellung gingen, wie oben bereits erwähnt, die nicht genutzten Aspekte sowie die nur von einer geringen Anzahl Probanden benoteten „Programmieren des Videorekorders" und die „Identifikation über Stimmerkennung" nicht in die Berechnung mit ein.

19.Automarke

Bei einer Gliederung nach Image der in dieser offenen Frage genannten Automarken zeigt sich folgende Aufteilung:

Gliederung	***Genannte Automarke***	***Anzahl Vpn***
Niedriges Preissegment	Renault Twingo,Citroen 2CV (Ente), Trabant, Fahrrad, Fiat	8
Mittleres Preissegment	Honda, Ford, Opel, VW	5
Höheres Preissegment	Audi, BMW, Mercedes	7
Sonstiges		4
Keine Antwort		3

Parallelen zu einem Trabant und sogar zu einem Fahrrad sollten zu denken geben. Allerdings wurde doch relativ häufig ein Auto der gehobenen Preisklasse zum Vergleich herangezogen. Eine eindeutige Tendenz zur einen oder anderen Preisklasse ist also nicht so ohne weiteres auszumachen.

20.Bewertung der Persona Smartakus

In diesem Kapitel werden verschiedene Aspekte der Persona näher abgefragt. Dazu zählt die wahrgenommene Kompetenz, die Natürlichkeit der Aktionen, das allgemeine Aussehen der Persona sowie generelle Fragen zu guten und schlechten Aspekten derselben. Dabei zeichnet sich folgendes Bild ab.

Die Persona wirkt mit –0.45 auf einer Skala von +3 bis –3 etwas inkompetent, jedoch verbirgt sich in diesem Mittelwert mit 8 Befragten ein hoher Anteil an Antwortverweigerungen bzw. „kann ich nicht sagen"–Antworten. Zum Teil wurden bei der Begründung auf diese Frage ganz konkretes Fehlverhalten der Persona genannt: „Missverständliches, unangemessenes Verhalten; z.B. Hand hinters Ohr könnte als Sprechaufforderung aufgefasst werden, das Zählen der Hand als Warten auf eine Reaktion des Benutzers"; „Kann ich nicht wirklich sagen, weil ich eher immer auf das grüne Licht gestarrt habe bzw ungeduldig auf eine Reaktion gewartet habe".

Die Natürlichkeit der Aktionen der Persona schnitt mit einem Mittelwert von –0.22 auf der selben Skala jedoch noch schlechter ab als in der vorherigen Frage. Diesmal verweigerten mit 2 Personen auch weitaus weniger Befragte eine Antwort.

Einen Bonuspunkt erhält die Persona bei der Frage nach dem Aussehen; hier schneidet sie mit einem Mittelwert von 0.67, und nur zwei Verweigerungen, sogar durchaus positiv ab.

Bei einer offenen Frage nach besonders guten Aspekten der Persona stand das Verhalten und Aussehen derselben mit 14 Nennungen durchaus im Vordergrund. Die Ausstrahlung des Smartakus wurde zweimal, „keine positiven Aspekte" allerdings auch viermal genannt. Es gab 8 Antwortverweigerungen.

Die Frage nach den negativen Aspekten der Persona zeigte folgendes Bild: Besonders häufig wurden unpassende Aktionen (10 Nennungen) erwähnt. Z.B. „das dumme Rumgezähle während der Wartezeit, das dauernde wieder Abwenden des Blickes, die lähmende Langsamkeit, der enorm kleine Kopf, der wenig Hirn hinter der Mütze vermuten lässt"; „Dieses ständige Finger abzählen"; „Einige der Bewegungsroutinen schon merkwürdig. Kaugummiblasen machen – was soll das" und so weiter. Das Aussehen wurde zweimal, die Langsamkeit einmal moniert. Mit 9 Nennungen wurde „nichts negatives" vergleichsweise häufig aufgelistet.

Alles in allem wurde die Persona als Repräsentation des Systems gesehen. Funktioniert dieses schlecht, wird auch die Persona schlecht bewertet. Erstaunlich ist in diesem Fall jedoch, dass trotz allgemein negativer Bewertung des Systems, die Persona dennoch häufig als „süß" bezeichnet wurde.

4. Fazit

Der SmartKom Prototyp, Szenario Home, wurde an 27 Versuchspersonen getestet. Die hier detailliert ausgewertete Benutzerbefragung ergab ein allgemein negatives Bild des Prototypen.

Allgemein schlecht wurde die lange Wartezeit zwischen

Befehlseingabe und Reaktion des Systems beanstandet. Daran angeschlossen die Tatsache, dass es keine Benutzerführung gab und die Probanden sich alleine gelassen fühlten. Als Fazit der Befragung über die präferierte Identifikationstechnik fällt auf, daß die Sicherheit der jeweiligen Technik von den Versuchspersonen als besonders relevantes Thema angesehen wird. Trotz allgemein negativer Bewertung des Systems erscheint die Persona jedoch immer noch ein gewisser Sympathieträger zu sein, wenn auch die Angemessenheit einzelner Aktionen, z.B. Kaugummiblasen, Finger zählen, häufig bezweifelt wurde.

Im Hinblick darauf gilt es zu hoffen, daß die Weiterentwicklung des Basisystems mit der Darstellung der Persona Schritt halten kann.