

7. Juli 2005

Automatische Überprüfung von Guidelines zur Usability Bewertung interaktiver Systeme

Nico Hamacher

Braunschweiger Verkehrskolloquium

Inhalt

- | Anforderungen und Eigenschaften von FIS
 - Gestaltungsvariationen
 - Möglichkeiten der Bewertung

- | Guideline-Bewertung von FIS
 - Eigenschaften
 - Anwendungsbeispiel

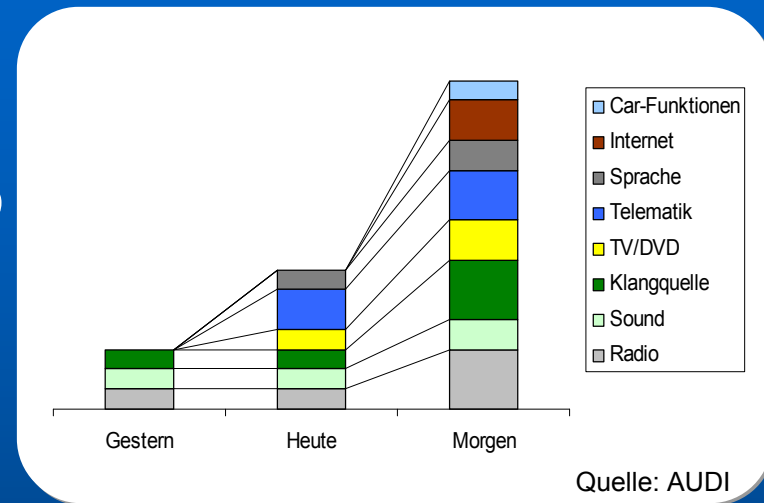
- | Bewertungs-Tool REVISER

- | Zusammenfassung

Entwicklung von Fahrer-Informationssystemen (FIS)

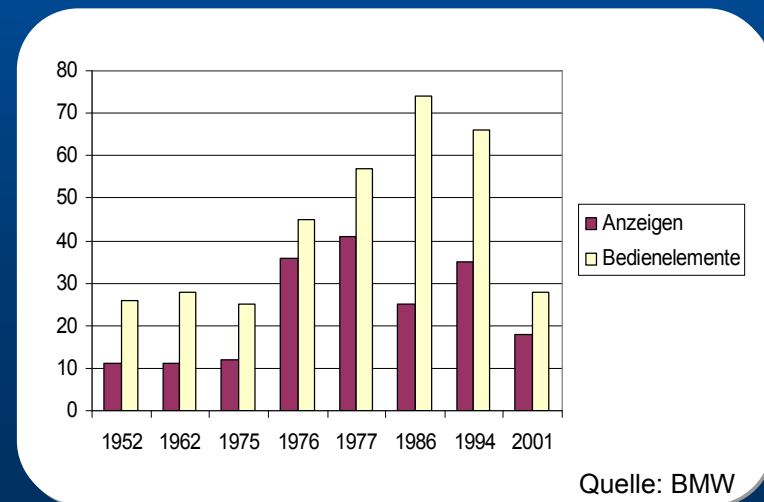
I Zunahme an Funktionen im Auto

- aktuell: bis 700
- überwiegend Komfortfunktionen



I Verringerung von Bedienelementen und Anzeigen

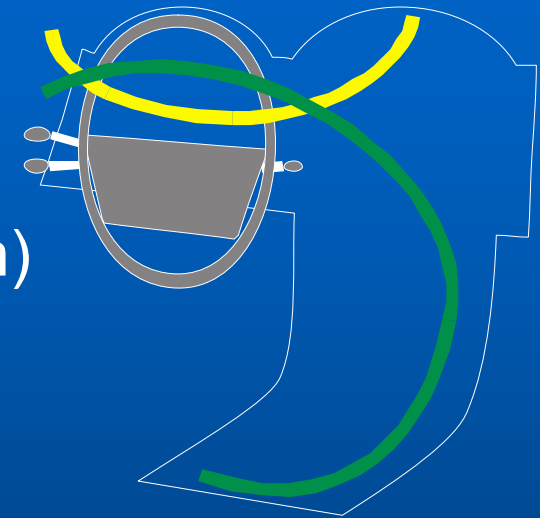
- Integration der Systeme
- ein zentrales FIS



„State of the Art“ bei FIS

Variationsmöglichkeiten bei FIS:

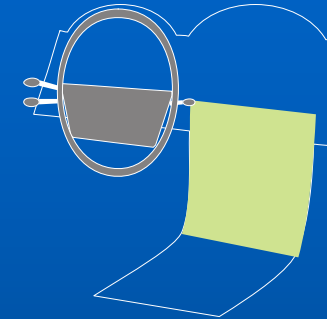
- | Lage der Anzeigen (Sichtbereich)
- | Lage der Bedienelemente (Greifbereich)
- | Art der Bedienelemente
 - Hardkeys
 - Softkeys
 - Controller
 - Touchscreen



- = Grenze des idealen Greifraums
- = Grenze des idealen Sichtbereichs

⇒ Es gibt bisher (noch) kein „bestes Konzept“

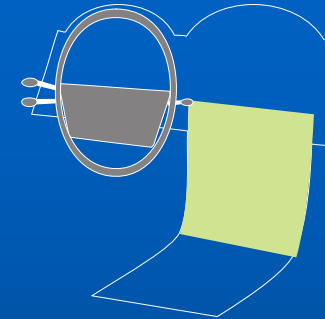
Volkswagen Phaeton



■ = Greifraum, ■ = Sichtbereich

- Hardkeys
- Softkeys
- Controller
 - Drehen
 - Drücken

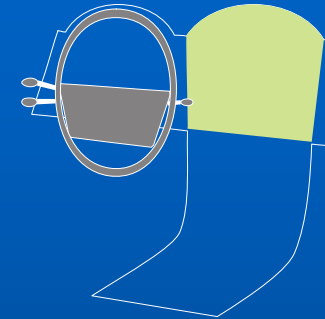
Mercedes Benz (Comand)



■ = Greifraum, ■ = Sichtbereich

- Hardkeys
- Controller
 - Drehen
 - Drücken

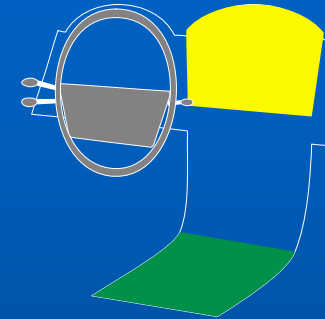
Lexus LS



■ = Greifraum, ■ = Sichtbereich

- Hardkeys
- Touchscreen

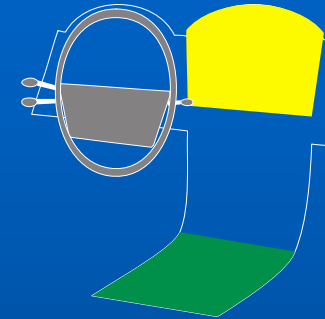
BMW (iDrive)



■ = Greifraum, ■ = Sichtbereich

- Controller
 - Drehen
 - Drücken
 - Schieben

Audi A8 (MMI)



■ = Greifraum, ■ = Sichtbereich

- Hardkeys
- Softkeys
- Controller
 - Drehen
 - Drücken

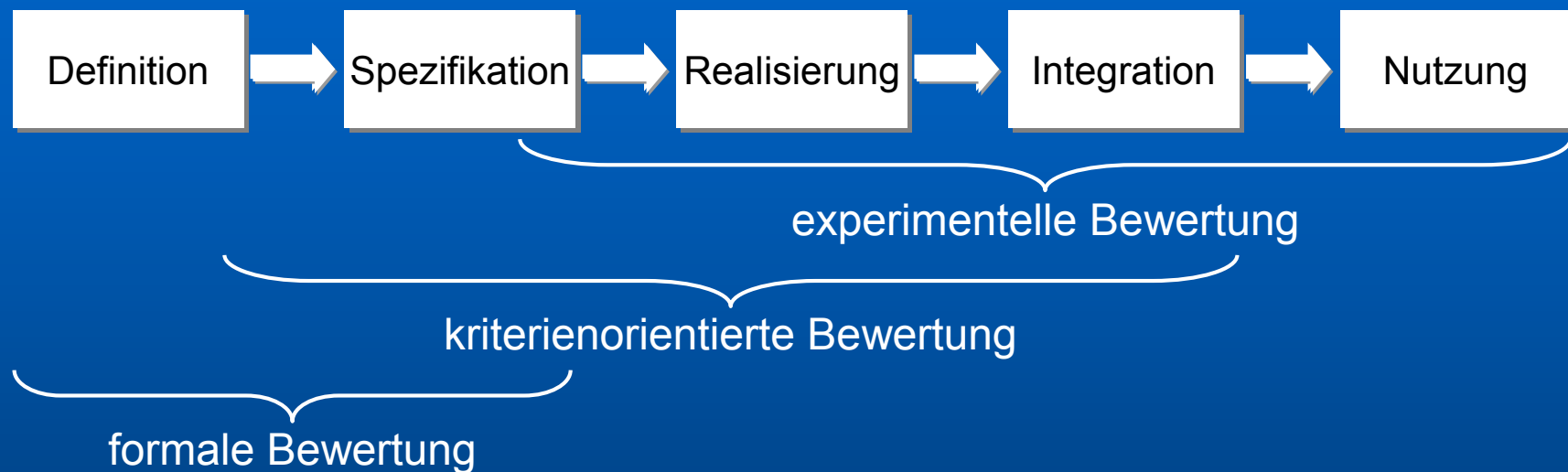
Modellierungsrelevante Eigenschaften von FIS

- | Einschränkungen bei Eingabemodalitäten
 - verteilte Elemente
 - unterschiedliche Eingabemöglichkeiten (Tasten, Controller)

- | Einschränkungen bei Displays
 - begrenzte Anzeigefläche
 - kleine Schrift
 - viele unterschiedliche Informationen auf einem Bildschirm

- | Einschränkungen bei Menüs
 - umfangreicher Menübaum
 - unterschiedliche Menüformen

Bewertung während der Systementwicklung



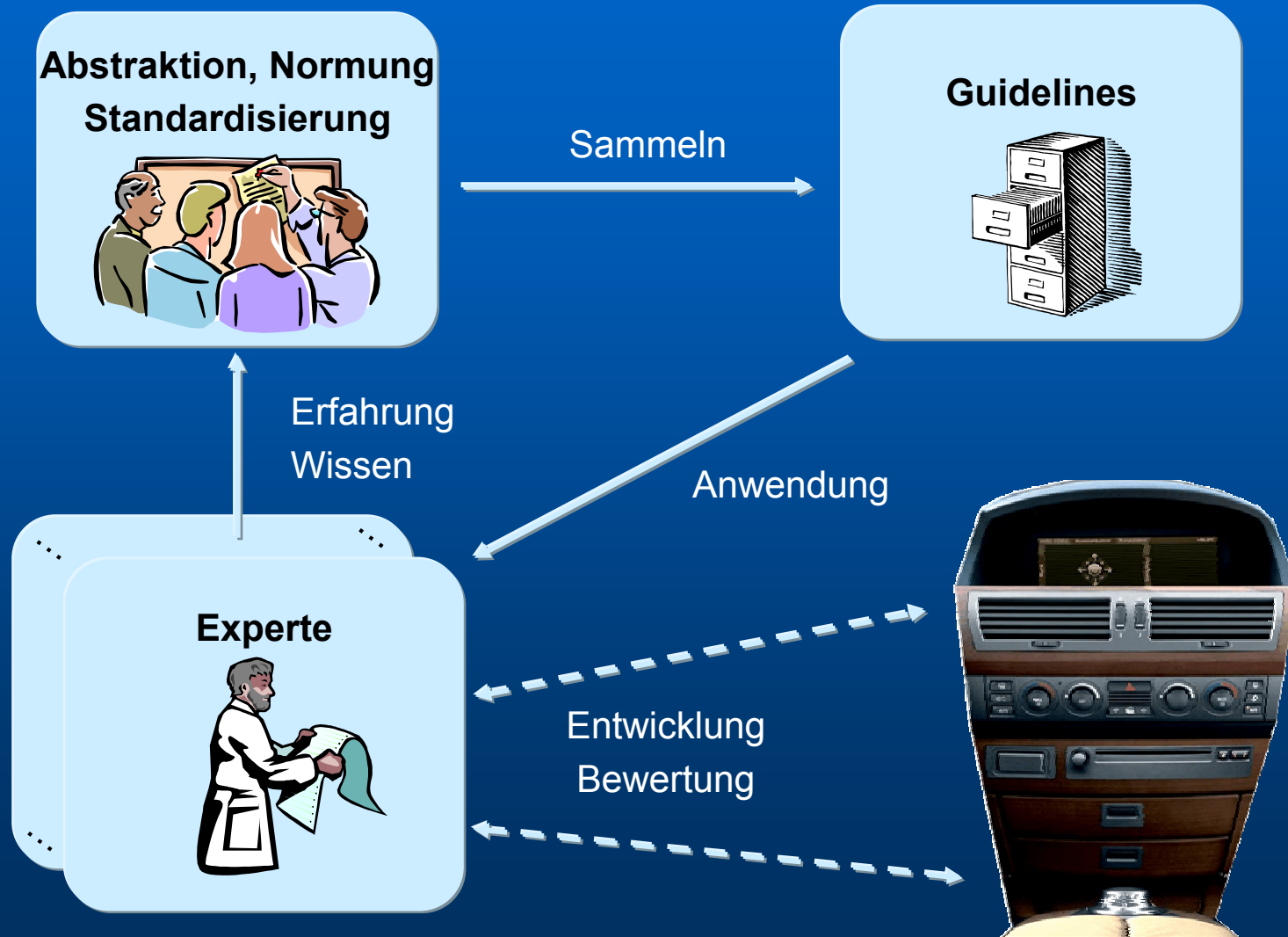
Bewertungsmethoden

- experimentell: Bewertung der Interaktion von Versuchspersonen mit einem Prototypen
- kriterienorientiert: Bewertung eines Prototypen anhand von Prüflisten oder Richtlinien durch Experten
- formal: Bewertung von Beschreibungen des Systems (Spezifikationen) oder des Verhaltens eines Operators (Benutzermodelle)

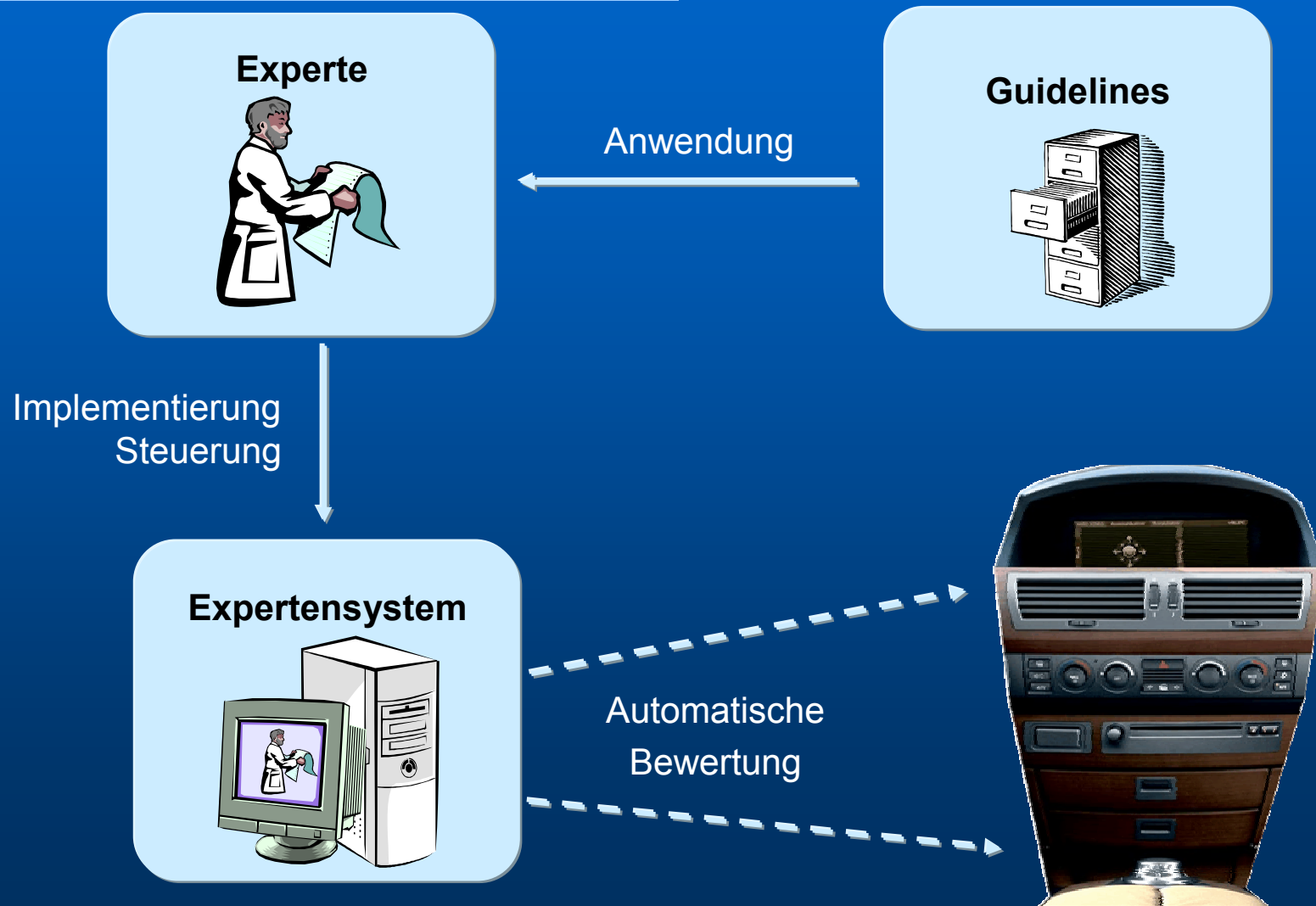
Abstraktionsgrad von Guidelines (Kriterien)



Kriterienorientierte Bewertung von FIS



Automatische kriterienorientierte Bewertung von FIS



Automatisierung nötig?

Aufwand der Implementierung, Haltbarkeit von Guidelines?

- | Untersuchung von Nielsen (2005):
 - 944 Usability Guidelines für Display-Systeme (ca. 20 Jahre alt)
 - 90% der Guidelines sind noch gültig
 - 70-78% sind gültig und nützlich

Experten-Erfahrung nicht eindeutig genug?

- | Untersuchung von Souza und Bevan (1990):
 - Interface Experten sollen UI eines Menüsystems unter Berücksichtigung von 45 Guidelines entwickeln
 - Experten machten Fehler oder hatten Probleme mit 91% der gegebenen Guidelines.
 - Entwickelte Menüsysteme verstießen im Mittel gegen 11% der gegebenen Guidelines

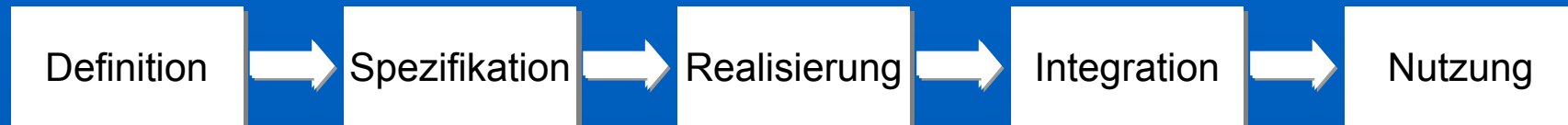
Anforderung an Guideline-Bewertung

- | Operationalisierung der Guidelines:
 - Beschreibung der Guideline als Regel
 - Festlegung des Zielelementes
 - Implementierbare Bedingung

- | Art der Bewertung:
 - Fehlerüberprüfung
 - Alarm
 - Verbesserungsvorschläge
 - Vergleich von Designalternativen

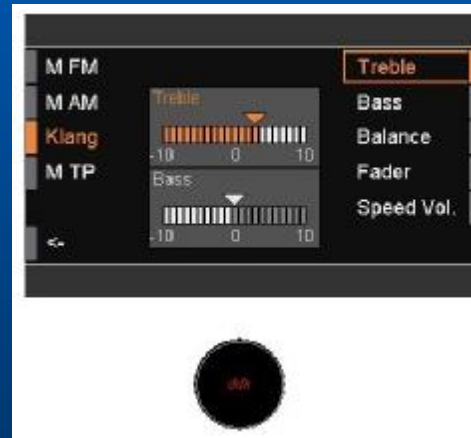
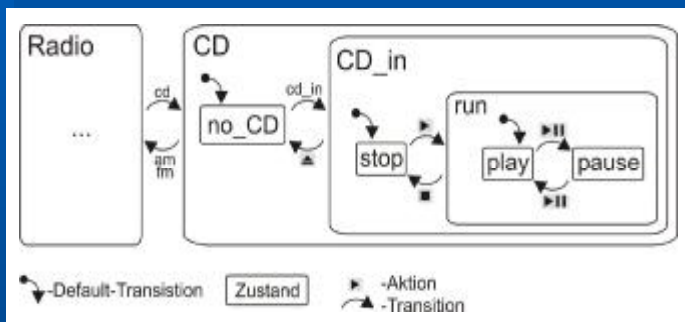
- | Durchführung der Bewertung
 - Statisch
 - Dynamisch (Agentenbewertung)

Gerätebeschreibung als Bewertungsgrundlage



Formale
Spezifikation

Prototypen-/Systembeschreibung



Beispiel einer Gerätebeschreibung



(box (xpos 444) (ypos 50) (xsize 480) (ysize 240) (thickness 1) (color #333333) (fill #333333) (isVisible 1) (isDisplay 1) (level 1))

(text (xpos 454) (ypos 261) (xsize 82) (ysize 24) (textheight 19) (content "Manuell") (fontfamily "Swis721 Ex BT") (thickness 1) (color #cc6600) (isVisible 1) (isDisplay 0) (level 20))

(line (xpos 524) (ypos 255) (xsize 321) (ysize 1) (pointsx 0 320) (pointsy 0 0) (thickness 2) (color #ff9900) (isVisible 1) (isDisplay 0) (level 21))

(circle (xpos 523) (ypos 150) (xsize 10) (ysize 10) (radius 5) (thickness 1) (color #ff9900) (fill #ff9900) (isVisible 1) (isDisplay 0) (level 31))

(text (xpos 635) (ypos 51) (xsize 102) (ysize 25) (textheight 20) (content "Radio FM") (fontfamily "Swis721 Ex BT") (thickness 1) (color #ff9900) (isVisible 1) (isDisplay 0) (level 32))

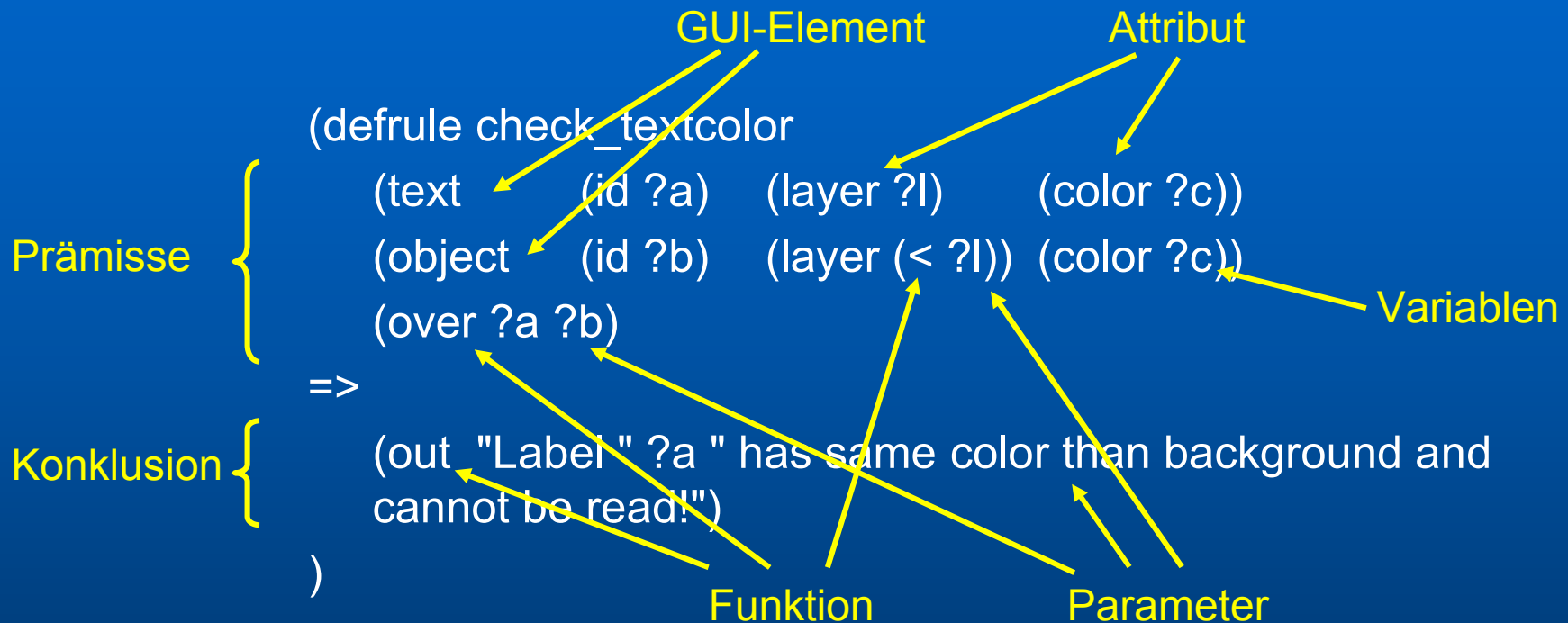
(mmi-element (xpos 644) (ypos 660) (xsize 80) (ysize 40) (textheight 10) (content "RETURN") (fontfamily "Allegro BT") (thickness 1) (color #000000) (isVisible 1) (isDisplay 0) (level 9))

(mmi-element (xpos 734) (ypos 660) (xsize 80) (ysize 40) (textheight 10) (content ">") (fontfamily "Allegro BT") (thickness 1) (color #000000) (isVisible 1) (isDisplay 0) (level 10))

(mmi-element (xpos 434) (ypos 488) (xsize 80) (ysize 40) (textheight 10) (content "RADIO") (fontfamily "Allegro BT") (thickness 1) (color #000000) (isVisible 1) (isDisplay 0) (level 11))

...

Aufbau einer Regel



Überprüfung durch Inferenz:

- GUI-Elemente (Fakten) auf Existenz
- Funktionen auf Erfüllung (true – false)
- Wenn alle Ausdrücke der Prämisse *true*, dann Konklusion ausführen

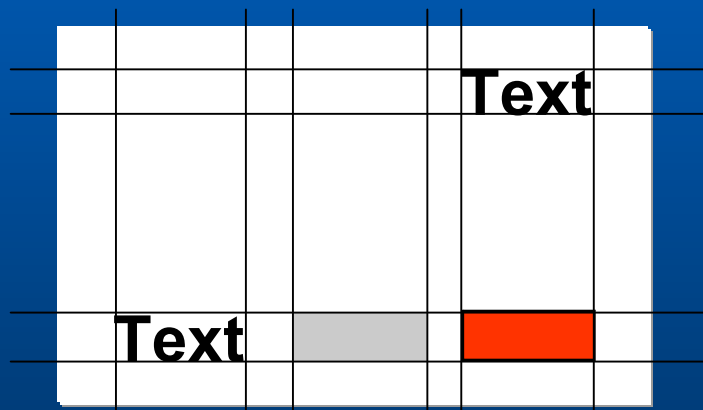
Beispiele von zu ermittelnden Werte

- | *Aspect-Ratio* Verhältnis zwischen Breite und Höhe
 - Gut: zwischen 0,5 und 0,8
- | *Gridedness* Maß der „Angeordnetheit“
 - je größer desto besser
- | *Farben* Anzahl Vordergrund-/Hintergrundfarben
- | *Margins* Abstände zum Rand
- | *Non Widget-Area* Freie Fläche
 - Gut: >30
 - 100 = hohe Ausnutzung
- | *Vertikale Balance* Oben-Unten-Ausgewogenheit
- | *Horizontale Balance* Links-Rechts-Ausgewogenheit
- | *Widget Density* Objekt-Dichte
 - Gut: <100
- | *Widgets* Anzahl der sichtbaren Objekte
- | *Text-Columns* Anzahl der Spalten mit bündigem Text

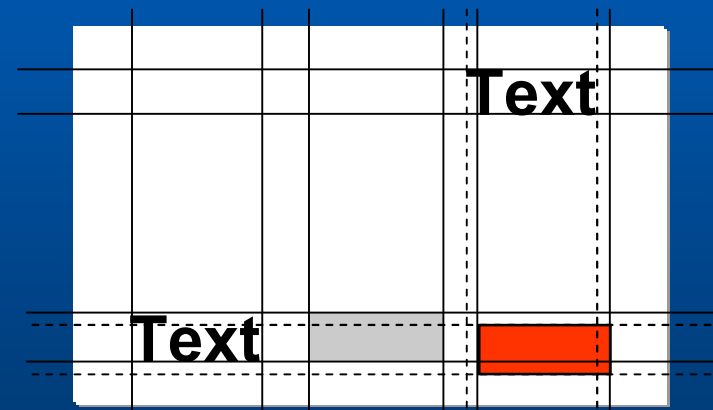
Beispiel: Gridedness

$$g = \frac{\# \text{Elemente}}{\# \text{unterschiedliche Kanten}} \cdot 100$$

Je größer desto besser!



KantenH = 6
KantenV = 4
Gridedness $g = 40$



KantenH = 8
KantenV = 6
Gridedness $g = 28,5$

(Shneiderman 91)

Beispiel: Maß für den subjektiven Farbkontrast

I ΔE Kontrastverhältnis zwischen zwei Farben

- Gut: $\Delta E > 100$

$$\bullet \Delta E(Yu'v') = \sqrt{\left(155 \cdot \left(\frac{\Delta Y}{Y_m}\right)\right)^2 + (367 \cdot \Delta u')^2 + (167 \cdot \Delta v')^2}$$

ΔY = Differenz von $Y_{\text{Vordergrund}}$ und $Y_{\text{Hintergrund}}$

$Y_m = \max(Y_{\text{Vordergrund}}, Y_{\text{Hintergrund}})$

$\Delta u'$ = Differenz von $u_{\text{Vordergrund}}$ und $u_{\text{Hintergrund}}$

$\Delta v'$ = Differenz von $v_{\text{Vordergrund}}$ und $v_{\text{Hintergrund}}$

Beispiele für das Farbkontrastmaß

I ΔE Kontrastverhältnis zwischen zwei Farben

- Gut: $\Delta E > 100$

Gutes Kontrastverhältnis:

Beispieltext $\Delta E = 188$

Beispieltext $\Delta E = 126$

Beispieltext $\Delta E = 146$

Beispieltext $\Delta E = 140$

Schlechtes Kontrastverhältnis:

Beispieltext $\Delta E = 31$

Beispieltext $\Delta E = 7$

Beispieltext $\Delta E = 47$

Beispieltext $\Delta E = 11$

Aber auch:

Beispieltext $\Delta E = 144$

Konsistenz

- | *Gleichförmigkeit, Geschlossenheit*
- | Konsistenz über vergleichbare Arbeitsaufgaben hinweg ermöglicht dem Benutzer, gemeinsame Abläufe für ausgeführte Arbeitsabläufe zu entwickeln
- | Konsistenz erhöht grundsätzlich die Vorhersehbarkeit eines Dialogs

Unterschiedliche Konsistenz:

- | **Gestaltungskonsistenz**
(z.B. Lage und Art der sichtbaren Objekte)
- | **Begriffskonsistenz**
(z.B. gleiche Begriffe für Gleiches, verständliche Oberbegriffe in Menüs)
- | **Bedienkonsistenz**
(z.B. gleiche Aktionen für Gleiches, Verwendung von Bedien-Stereotypen)

Maß zur Gestaltungskonsistenz

Konsistenzmaße nur im Zusammenhang einer Bearbeitungssequenz

Verhältnis eines Bearbeitungsschrittes zur gesamten Sequenz

- | Jeweiliges Maß bestimmen
- | Mittelwert (E) und Standardabweichung (σ) bestimmen
- | Z-Score berechnen

$$\text{Z-Score}(x) = \frac{x - E(x)}{\sigma}$$

- | Für $-1 \leq \text{Z-Score}(x) \leq 1$ gilt Maß x als konsistent

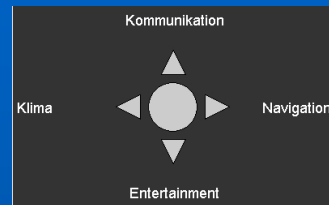
Konsistenz einer kompletten Aufgabe

- | Maß für die *mittlere Variation* (V , als Gegenteil der Konsistenz)
- | V berechnen

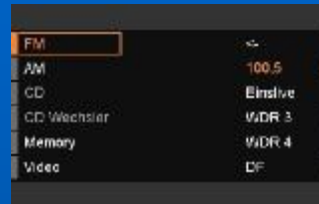
$$V(\text{Aufgabe}) = \frac{\sigma}{E(x)} \cdot 100$$

- | Je kleiner, desto konsistenter ist das jeweilige Maß x für die gesamte Aufgabe

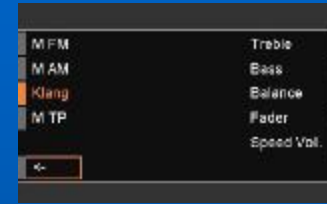
Beispiel: BMW-Klangeinstellung



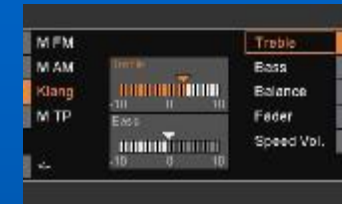
1. (1 Schritt)



2. (5 Schritte)



7. (4 Schritte)



11. (3 Schritte)

Maß	Z-Score		Z-Score		Z-Score		Z-Score	
Word Count	4	-2,67	17	0,17	14	-0,49	22	1,26
Column-All Ratio	0	-3,35	0,71	0,2	0,8	0,62	0,67	-0,04
Column Count	0	-1,59	2	-0,65	3	-0,18	7	1,7
Text Widget Count	4	-2,29	14	0,28	10	-0,75	18	1,3
Non Widget Area	94,87	-3,46	100	0,29	100	0,29	100	0,29
Gridedness	30	-2,66	50,88	-0,11	50	-0,22	63,03	1,37
Foreground Color Count	3	-1,99	7	0,26	5	-0,86	9	1,38
Margin East	6	3,46	0,5	-0,29	0,5	-0,29	0,5	-0,29
Margin West	6	3,46	0,5	-0,29	0,5	-0,29	0,5	-0,29
Margin South	8	-3,46	32,50	0,29	32,5	0,29	32,5	0,29
Margin North	7	-3,46	31,50	0,29	31,5	0,29	31,5	0,29
Vertical Area-Balance	-1,57	0,3	-18	-0,63	-9,44	-0,15	9,89	0,95
Horizontal Area-Balance	5,73	0,69	-23,99	-1,13	-0,39	0,31	4,55	0,61
Display Aspect Ratio	0,63	0	0,63	0	0,63	0	0,63	0
Widget Density	9	-1,2	29	-0,3	21	-0,66	75	1,77

Z-Score = Verhältnis eines Wertes zur Gesamtheit | Gut: $-1 \leq Z\text{-Score} \leq 1$

Anwendungsbeispiel

Aufgaben:

- | Radio
 - Den dritten Sender anwählen
- | Klangeinstellungen
 - Zwei Rasterstufen mehr Höhen
- | Navigation
 - Das letzte gespeicherte Ziel anwählen
 - Zielführung starten

Unterschiedliche FIS:

- Audi MMI: Controller + Tasten
- BMW iDrive: Controller

Ergebnisse – Anwendungsbeispiel

	<u>Audi</u>	<u>BMW</u>
Benutzte Farben (Gesamt)	10	8
ΔE-Alarme	3	3
Schritte gesamt (Klang, Sender, Navi)	18 (7,5,6)	44 (13,9,22)
Menü-Überschriften automatisch erkannt	ja	nein

V-Werte (Variation eines Maßes in einer Aufgabe - je kleiner, desto besser):

	Sender		Navigation		Klang		Über alle Aufgaben	
	Audi	BMW	Audi	BMW	Audi	BMW	Audi	BMW
Column-All Ratio	25,23	35,36	11,83	38,76	86,60	29,81	48,28	35,61
Column Count	0,00	41,42	35,36	47,14	86,60	62,98	55,49	61,81
Text Widget Count	24,49	24,38	18,00	34,68	15,28	30,12	21,59	33,31
Non Widget Area	5,20	1,62	20,14	1,07	2,71	1,37	16,63	1,30
Word Count	29,51	26,26	12,76	36,84	18,62	28,20	27,50	37,05
Gridedness	5,89	13,51	5,74	4,84	1,95	15,80	5,45	20,84
Foreground Color Count	10,53	19,58	11,11	20,00	20,38	27,24	18,71	26,21
Margin East	0,00	155,56	0,00	85,27	16,64	158,77	19,66	154,91
Margin West	0,00	155,56	0,00	111,00	0,00	158,77	0,00	133,16
Margin South	0,00	25,86	0,00	16,26	0,00	21,32	0,00	20,03
Margin North	0,00	26,76	0,00	18,22	0,00	22,04	0,00	21,53
Vertical Area-Balance	31,72	195,98	46,47	1505,70	119,15	258,37	86,68	603,68
Horizontal Area-Balance	40,71	1932,99	46,03	94,16	73,26	297,65	58,45	158,85
Widget Density	16,87	23,47	17,62	28,60	34,64	62,47	33,21	59,70
Widget Count	16,87	23,47	17,62	28,60	34,64	62,47	33,21	59,70

SirValUse-Studie (Navi/Gesamt-Schulnoten)

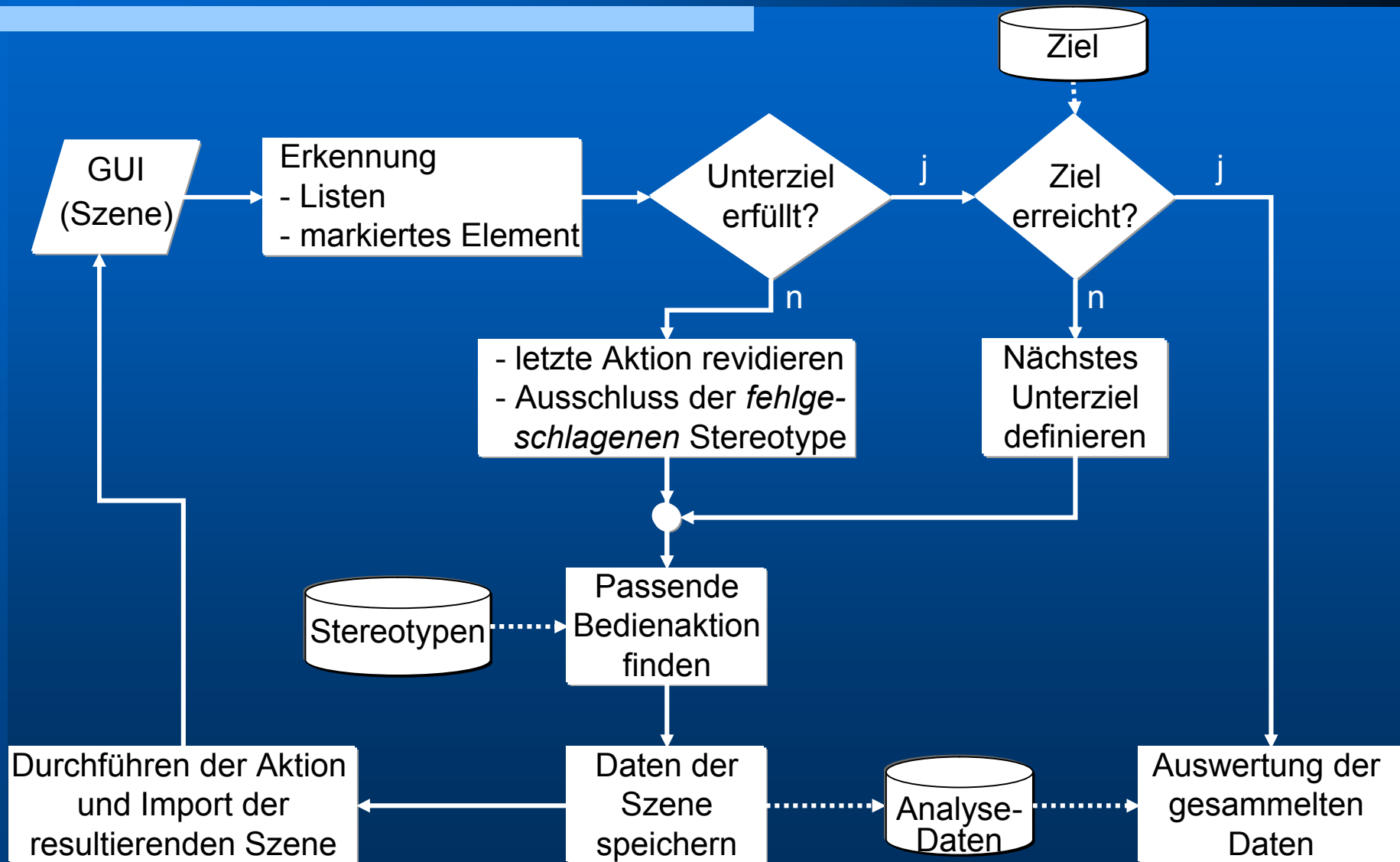
2 / 2,0

4 / 3,5

Bedienkonsistenz

- | Gleichförmige Bedienung
 - Gleiche Bedienschritte mit gleichen Bedienelementen
 - Menübedienung in der ersten/zweiten/.../folgenden Menüebene
 - Datenmanipulation
 - Unterscheidung zwischen Softkeys, Hardkeys, Controller
- | Simulierte Bedienung mit interaktivem Agent
 - Vorgaben
 - Ziel: Erreichen eines bestimmten Textes (z.B. „WDR 3“)
 - Szene: alle Bedienelemente und aktuelle GUI
 - Interaktionswissen: Stereotypen
 - mentale Modelle / Metaphern
 - Simulation
 - Erfüllung der Stereotypen mit gegebenen GUI-Daten

Bedienkonsistenz - Agent



Ergebnis einer Agentenbewertung

- I Pro Szene:
 - Gestaltungsfehler (Alarmer)
 - Gestaltungs-Kennzahlen

- I Pro Aufgabensequenz
 - Gestaltungskonsistenz (Z-Score)
 - Gedächtnisbelastung = Wissen, das zur erfolgreichen Bedienung einer Aufgabe benötigt wird

- I Über mehrer Aufgaben
 - Gleichförmigkeit der Bedienung
 - Gedächtnisbelastung zur Bedienung des Systems
 - Konsistenz-Maße (V-Werte)
 - Begriffskonsistenz

Begriffswissen - Ontologien

Text als Orientierung:

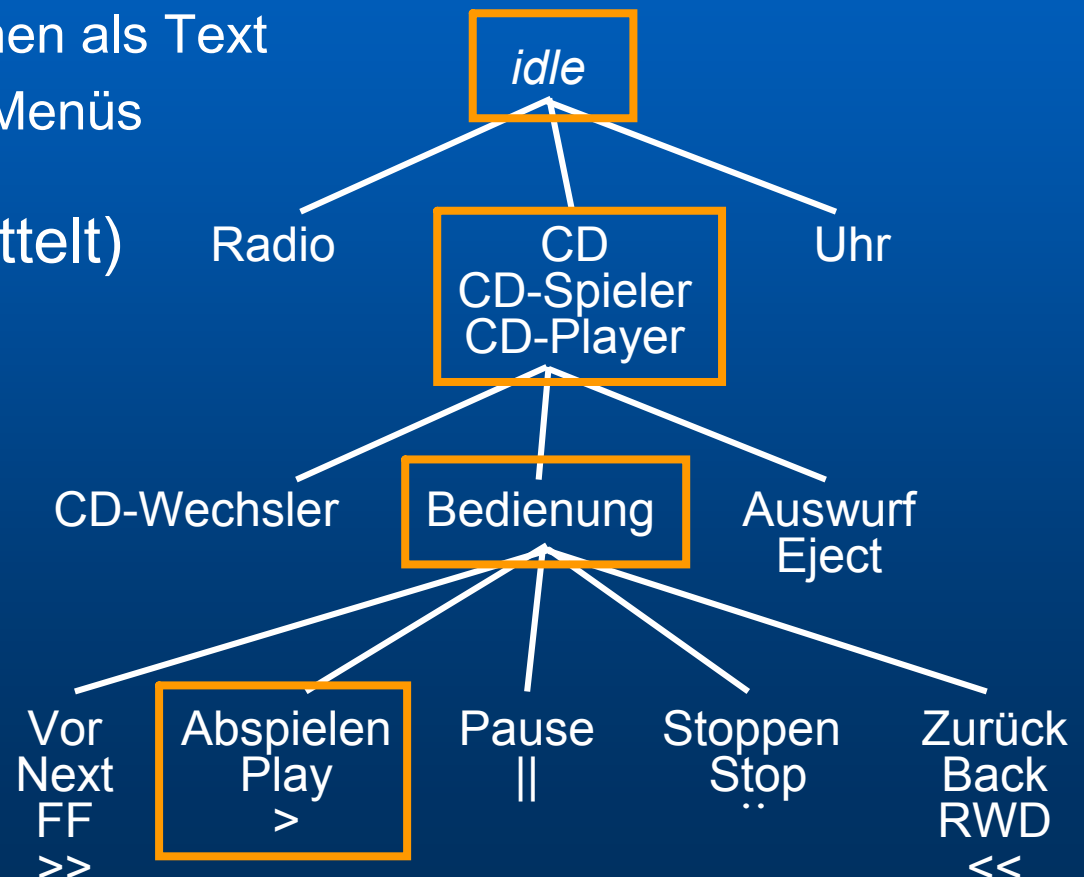
- Beschreibung von Funktionen als Text
- Darstellung von Informationen als Text
- Navigation durch textuelle Menüs

Ontologien (empirisch ermittelt)

- Begriffsklassen
- Hierarchische Anordnung

Einsatz von Ontologien

- Bildung von *Begriffspfaden* (*Zielpfad* für Agent)
- Überprüfen von Begriffszusammenhängen

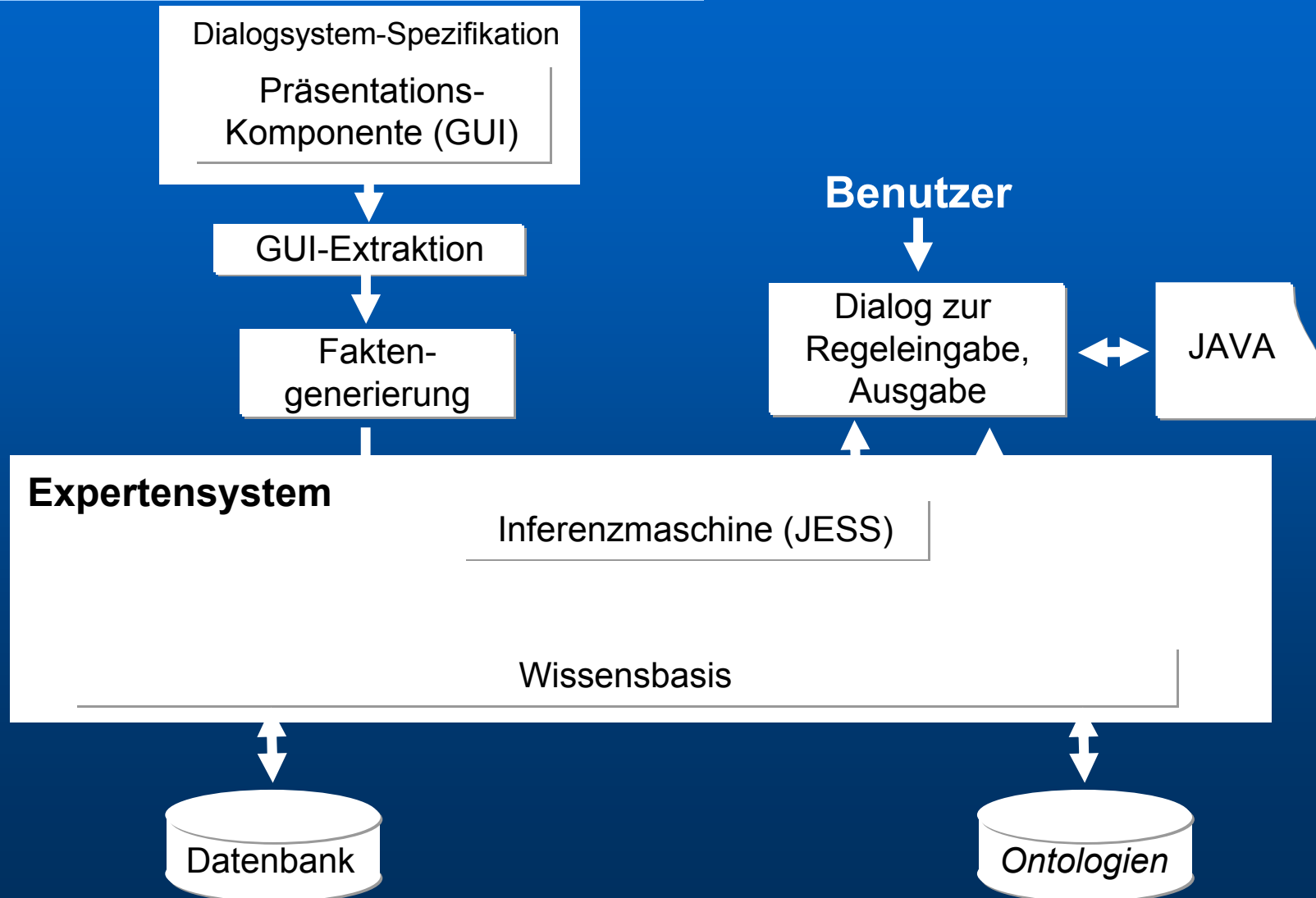


Tool zur Guideline-Bewertung (LTI)

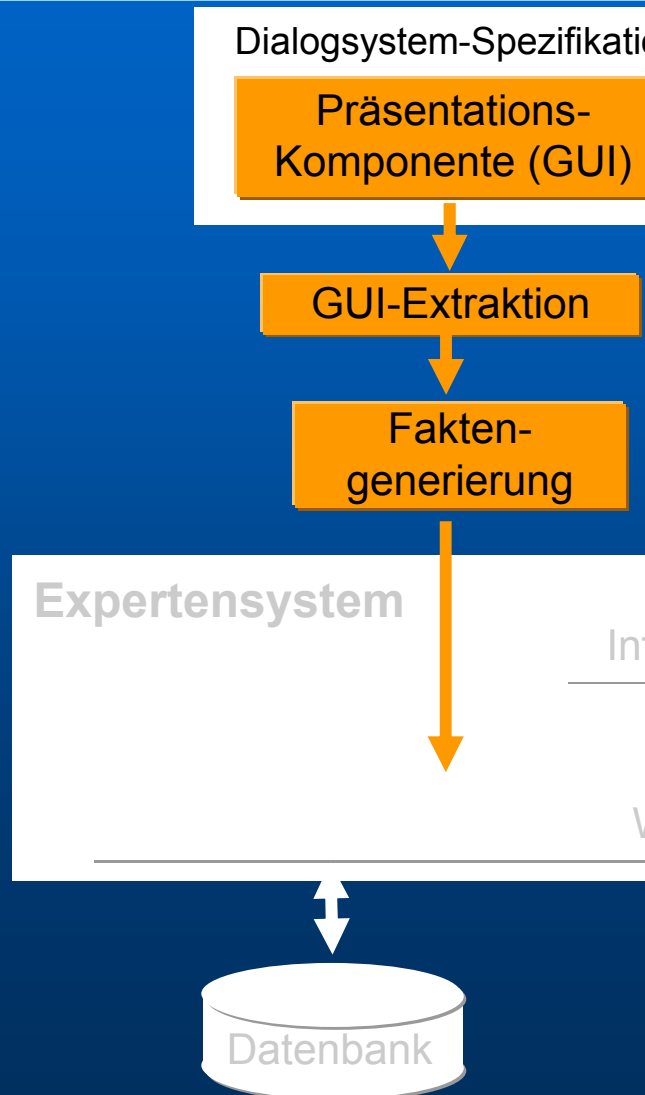
- | Editor zur schnellen Regeleingabe
- | Anbindung eines Expertensystems zur Überprüfung der Regeln
- | Konverter zur Generierung der Fakten aus technischen Spezifikationen / Prototypen
- | Einbindung von Java-Methoden
- | Einbindung von Ontologien
- | Möglichkeit der Integration eines autonomen Agenten für die Simulation der Bedienung

REVISER

Rapid Evaluation of Interactive Systems based on ExpeRt knowledge



Beschreibung des GUI



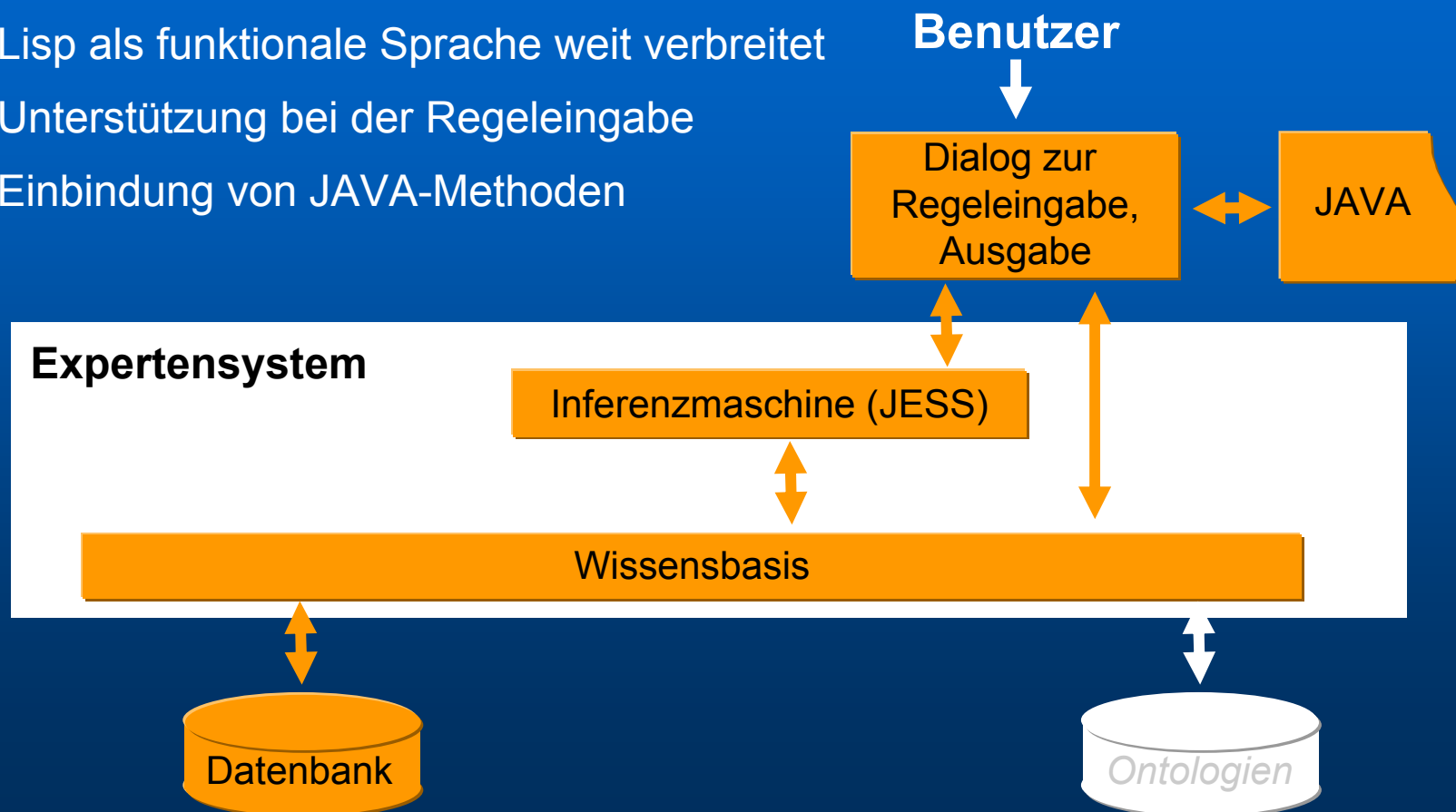
- | Beschreibung aller sichtbaren Elemente
- | Grafische Dekomposition (kompatibel zu IBM CUA-Format)
- | Beschreibung in allg. lesbaren/gültigen XML-Format

```

<But_Func type= button
  x= 43
  y= 20
  width= 40
  height= 24
  layer= 31
  label= Lab_Func
  ...
/>
<Lab_Func text= 'Func'
  font= Arial
  size= 6
  color= black
  ...
/>
  
```

JESS (Java Expert System Shell)

- | Java-Version des gängigen CLIPS
- | Inferenz mit Vorwärts-/Rückwärtsverkettung
- | Lisp als funktionale Sprache weit verbreitet
- | Unterstützung bei der Regeleingabe
- | Einbindung von JAVA-Methoden



Zusammenfassung

- | Guideline-Bewertung Standard in der Bewertung von FIS

- | Automatische Guideline-Bewertung
 - Deutlicher Zeitgewinn
 - Interpretationssicherheit
 - Robuste Vergleichsmaße durchgängig in allen Entwicklungsphasen
 - Statische Bewertung sowie dynamische Bewertung durch Agent
 - Kein Ersatz von Experten sondern Arbeitserleichterung

- | REVISER als automatisches Guideline-Bewertungs-Tool
 - Gerätespezifikationen/-prototypen als Bewertungsgrundlage
 - Guidelines als Regeln
 - Automatische Überprüfung der Regeln
 - Beliebige Erweiterbarkeit durch Java

Kontakt

Dipl.-Inform. Nico Hamacher

RWTH Aachen

Lehrstuhl für Technische Informatik

Ahornstr. 55

52074 Aachen

www.techinfo.rwth-aachen.de

0049 (0)241 80 26105

hamacher@techinfo.rwth-aachen.de