

Modul:

02 KI-bezogene UX

Gefördert vom:



Bundesministerium
für Bildung, Familie, Senioren,
Frauen und Jugend



UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR MULTIMEDIALE
UND INTERAKTIVE SYSTEME

1 Einleitung in KI-bezogene UX

Kursübersicht > KI-bezogene UX

In diesem Modul erläutert Tim Schrills konkretere UX-Aspekte, die beim Design und der Implementierung von KI-Systemen beachtet werden sollten.

KI-Systeme greifen tief in unsere Informationsverarbeitung ein, was zusätzliche Aspekte für die User Experience (UX) relevant macht. Diese neuen Dimensionen sind essentiell, da sie direkt beeinflussen, wie Menschen mit KI-Systemen interagieren und diese wahrnehmen. Mit der zunehmenden Integration von KI in alltägliche Systeme ist es entscheidend, KI-bezogene Aspekte in der UX zu berücksichtigen. Dazu werden in diesem Modul fünf wichtige Aspekte genannt und deren Relevanz für das Design und die Entwicklung von KI-Systemen erklärt.

Im folgenden Video wird ein Überblick über die unterschiedlichen Aspekte bei der Identifikation gegeben, auf die in den folgenden Kapiteln näher eingegangen wird.



<https://youtu.be/dbBIEqBMudI>

In der vorigen Lektion haben wir etablierte UX-Konstrukte betrachtet, die allgemeine Nutzererfahrungen beschreiben. KI-Systeme hingegen greifen tief in unsere Informationsverarbeitung ein, was zusätzliche Aspekte für die User Experience (UX) relevant macht. Diese neuen Dimensionen sind essentiell, da sie direkt beeinflussen, wie Menschen mit KI-Systemen interagieren und diese wahrnehmen.

Wichtige Aspekte der KI-bezogenen UX

1

Wahrgenommene Autonomie

Dieser Aspekt beschreibt, wie sehr Nutzende das Gefühl haben, selbstständig zu handeln und Entscheidungen zu treffen, während sie mit einem KI-System interagieren.

2

Wahrgenommenes Situationsbewusstsein

Dies bezieht sich auf das Verständnis der Nutzende über ihre Umgebung und die Änderungen, die durch das KI-System verursacht werden.

3

Wahrgenommene Mentale Belastung

Dieser Aspekt umfasst den kognitiven Aufwand, der erforderlich ist, um Informationen zu verarbeiten und Entscheidungen zu treffen, und die potenzielle Überlastung durch zu viele Informationen.

4

Wahrgenommene Vertrauenswürdigkeit

Damit ist das Vertrauen gemeint, das Nutzende in ein KI-System haben, basierend auf dessen Handlungen.

5

Wahrgenommene Diagnostizität

Dies beschreibt das Vertrauen der Nutzende in die Diagnosen oder Vorschläge des KI-Systems und wie gut diese die gewünschten Ergebnisse liefern.

Warum sind diese Aspekte wichtig für die Automation-Related UX?

Mit der zunehmenden Integration von KI in alltägliche Systeme ist es entscheidend, diese Aspekte in der UX zu berücksichtigen. KI-Systeme übernehmen Aufgaben, die früher Menschen vorbehalten waren, und erfordern daher, dass Nutzer ein Gefühl von Kontrolle, Verständnis und Vertrauen behalten. Die Gestaltung von KI-Systemen muss sicherstellen, dass Nutzer die Systeme nicht nur effektiv nutzen, sondern sich auch sicher und autonom fühlen können.

Beispiel: KI-gestütztes Verkehrsmanagementsystem

Stellen Sie sich ein KI-gestütztes Verkehrsmanagementsystem vor. Dieses System analysiert Verkehrsdaten und schlägt Optimierungen für Ampelphasen vor. Nutzer wie Verkehrsingenieure müssen das System verstehen, seine Vorschläge bewerten und gegebenenfalls anpassen können. Hier ist es wichtig, dass das System Transparenz bietet, um das Situationsbewusstsein zu fördern, eine angenehme Menge an Informationen bereitstellt, um Überlastung zu vermeiden, und Vertrauen in die Automatisierung aufbaut.

Was können Sie von dieser Lektion erwarten?

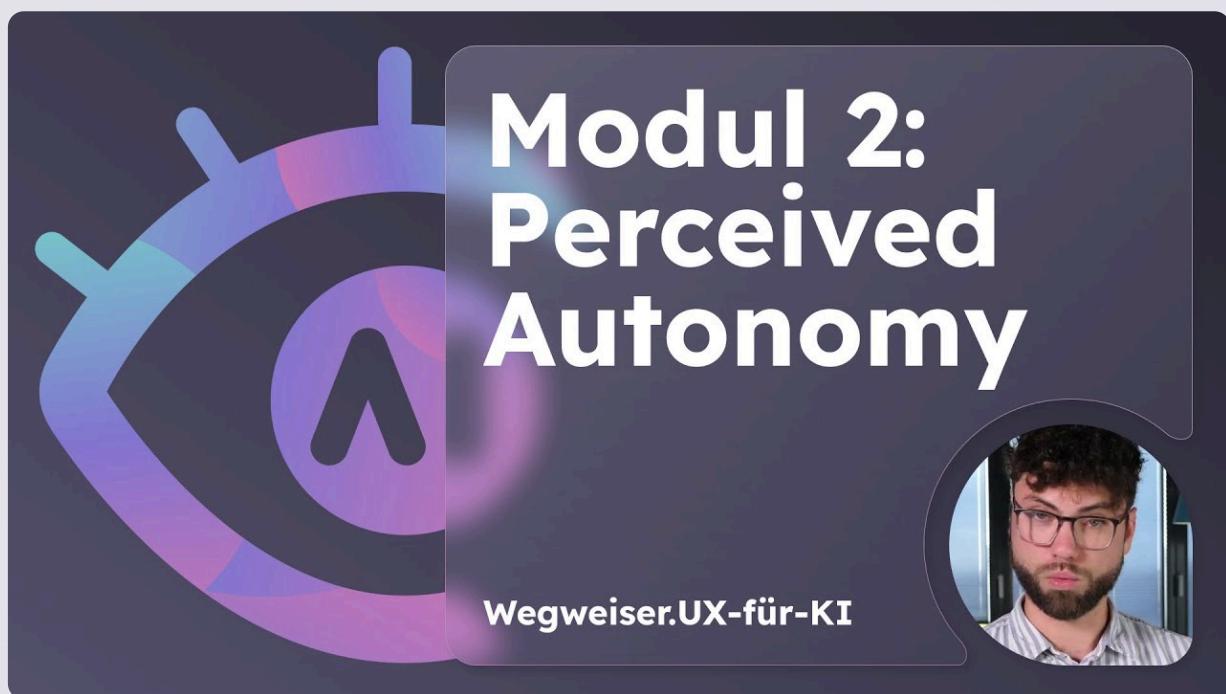
In den folgenden Kapiteln werden wir detailliert untersuchen, wie diese fünf Aspekte – Wahrgenommene Autonomie, Situationsbewusstsein, Mentale Arbeitsbelastung, Vertrauenswürdigkeit und Confidence/Diagnosticity – die Gestaltung und Nutzung von KI-Systemen beeinflussen. Wir werden sehen, wie diese Aspekte in der Praxis umgesetzt werden können und welche Designrichtlinien helfen, eine positive und effektive Mensch-KI-Interaktion zu fördern. Freuen Sie sich auf eine tiefere Auseinandersetzung mit den Herausforderungen und Chancen der KI-bezogenen UX!

Wahrgenommene Autonomie

Kursübersicht > KI-bezogene UX

Dieser Aspekt beschreibt, wie sehr Nutzende das Gefühl haben, selbstständig zu handeln und Entscheidungen zu treffen, während sie mit einem KI-System interagieren.

Im folgenden Video wird ein Überblick über den Begriff Wahrgenomme Autonomie gegeben.



<https://youtu.be/7QDwxEpQNHk>

1. Definition wahrgenommener Autonomie

Die **wahrgenommene Autonomie** beschreibt, inwiefern Nutzende eines KI-Systems das Gefühl haben, selbstständig handeln und Entscheidungen treffen zu können. Diese Wahrnehmung beeinflusst das Vertrauen in das System, die Interaktion damit und die Zufriedenheit der Nutzenden. Wesentlich ist, ob das System als transparent und unterstützend wahrgenommen wird oder ob es lediglich als ausführendes Werkzeug dient.

2. Relevante Konzepte und Modelle

Automatisierungsstufen (Levels of Automation)

Die **Automatisierungsstufen** (Levels of Automation, LOA) wurden entwickelt, um Abstufungen oder Kategorien der **Autonomie** zu veranschaulichen. Diese Struktur hilft zu verstehen, **wie Menschen mit automatisierten Systemen interagieren** und beschreibt, welche Aufgaben entweder vom Menschen oder von der Maschine übernommen werden.

Das LOA-Modell nach Parasuraman et al. (2000) umfasst **zehn Stufen der Aufgabenteilung** und Verantwortlichkeit zwischen Mensch und Maschine. Je nach Modell und Anwendungsbereich können die

Automatisierungsstufen jedoch variieren. So definiert beispielsweise die Society of Automotive Engineers (SAE) **fünf Stufen** der Automatisierung im Bereich **autonomer Fahrzeuge** (Hopkins & Schwanen, 2021).

Einfluss der Automatisierung auf psychologische Variablen

Die Einführung von Automatisierung hat einen signifikanten Einfluss auf psychologische Variablen wie **Arbeitsbelastung, Fähigkeiten, Vertrauen und Situationsbewusstsein** der Nutzer (Parasuraman et al., 2000).

Daher ist es entscheidend, das geeignete Automatisierungsniveau je nach Aufgabe auszuwählen, um unerwünschte Effekte zu vermeiden.

Vier Stufen der Informationsverarbeitung

Parasuraman, Sheridan und Wickens (2000) verknüpften die Automatisierungsstufen mit vier grundlegenden Funktionen, die auf einem Modell der menschlichen Informationsverarbeitung basieren und in einem Mensch-Maschine-System unterstützt werden sollen:

- 1.** Informationsbeschaffung,
- 2.** Informationsanalyse,
- 3.** Entscheidungsfindung
- 4.** Handlungsausführung.

Dieses Modell bietet eine strukturierte Herangehensweise zur Klassifizierung von Aufgaben, bei denen Automatisierung den Menschen unterstützen kann.

Fehler eines Systems in späteren Stadien können störender wirken als in Systemen, bei denen die Automatisierung höchstens bis zur Phase der Informationsanalyse eingesetzt wird. Akzeptanz, Vertrauen und Leistung können abnehmen, wenn in den späteren Phasen der Informationsverarbeitung zu viel Automatisierung vorhanden ist (Onnasch et al., 2014).

Einfluss der Automatisierung auf die menschliche Leistung

Parasuraman et al. (2000) haben vier zentrale Faktoren identifiziert, die beeinflussen, wie Automatisierung die menschliche Leistung beeinflussen kann:

- 1. Situationsbewusstsein:** Das Verständnis der aktuellen Umgebung und Situation durch den Menschen, welches durch Automatisierung entweder gefördert oder beeinträchtigt werden kann.
- 2. Vertrauen:** Das Vertrauen der Nutzenden in das System ist entscheidend für eine erfolgreiche Interaktion und hängt stark von der wahrgenommenen Autonomie ab.
- 3. Abbau von Fähigkeiten:** Hohe Automatisierung kann zu einem Rückgang menschlicher Fähigkeiten führen, da weniger manuelle Eingriffe und Entscheidungen nötig sind.
- 4. Arbeitsbelastung:** Automatisierung kann die Arbeitsbelastung entweder reduzieren oder erhöhen, abhängig davon, wie gut sie an die Bedürfnisse der Nutzenden angepasst ist.

Anpassbare vs. selbstanpassende Automatisierung

Die Wahl des Automatisierungs niveaus hat einen signifikanten Einfluss auf die Mensch-Automation-Interaktion. Es gibt zwei grundlegende Ansätze: **anpassbare Automatisierung** und **selbstanpassende Automatisierung**. In der anpassbaren Automatisierung wählt der Benutzer das Automatisierungs niveau manuell basierend auf seinen eigenen Bedürfnissen und Vorlieben. Im Gegensatz dazu überwacht die selbstanpassende Automatisierung den Zustand des Benutzers, wie Arbeitsbelastung oder Wachsamkeit, und passt das Automatisierungs niveau automatisch an.

1. Anpassbare Automatisierung

- Der Benutzer hat die Kontrolle über die Auswahl des Automatisierungs niveaus.
- Bietet Flexibilität und Anpassung an individuelle Präferenzen.
- Vorteilhaft, wenn Benutzer die Systeme nach ihren eigenen Bedürfnissen steuern wollen.

2. Selbstanpassende Automatisierung

- Das System überwacht den Benutzer (z. B. Arbeitsbelastung) und passt das Automatisierungs niveau automatisch an.
- Hilfreich in dynamischen Umgebungen, da es auf Veränderungen reagiert, ohne dass der Benutzer eingreifen muss.
- Kann die Arbeitsbelastung reduzieren, birgt jedoch das Risiko, dass der Benutzer die Kontrolle verliert.

3. Studien zur User Experience und KI

Kaber & Endsley (2004): The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task.

Untersuchten die Auswirkungen von adaptiver Automatisierung auf die Leistung des Menschen, das Situationsbewusstsein und die Arbeitsbelastung in dynamischen Umgebungen.

Niedriges Automatisierungsniveau: Verbessert die Leistung, da der Benutzer stark eingebunden bleibt (Kaber & Endsley, 2004). **Mittleres Automatisierungsniveau:** Führt zu verbessertem Situationsbewusstsein,

was entscheidend für komplexe und dynamische Aufgaben ist. Allerdings führen mittlere Automatisierungsstufen nicht immer zu besserer Leistung oder geringerer Arbeitsbelastung, obwohl sie das Situationsbewusstsein verbessern.

Rieger et al. (2022): Challenging presumed technological superiority when working with (artificial) colleagues.

Diese Studie untersucht, wie Menschen klassische Automatisierung und KI-basierte Systeme häufig als **Black Boxes** wahrnehmen, ohne signifikante Unterschiede in ihrer Wahrnehmung beider Technologien. Dies führt zu einem unvollständigen Automatisierungsschema, da Transparenz fehlt. Interessanterweise verändert sich die Präferenz der Menschen zwischen der Interaktion als Ratsuchender und als Bewerteter.

Während beim gemeinsamen Arbeiten menschliche Faktoren wie **Intuitivität** und **Fachwissen** bevorzugt werden, können bei der Bewertung durch Maschinen deren **Objektivität** und **Konsistenz** als vorteilhaft angesehen werden.

Deci & Ryan (1985): Self-determination theory (SDT)

Die Selbstbestimmungstheorie (Self-Determination Theory) definiert drei universelle grundlegende psychologische Bedürfnisse (Basic Psychological Needs, BPNs):

- 1. Autonomie:** das Gefühl, Kontrolle über eigene Entscheidungen und Handlungen zu haben
- 2. Kompetenz:** das Erleben von Wirksamkeit und Beherrschung einer Aufgaben
- 3. soziale Verbundenheit:** sich um andere zu kümmern und im Gegenzug Fürsorge zu erfahren

4. Operationalisierung: Fragebögen und Messinstrumente

Zoubir (2024): Preference for Automation Types Scale (PATS)

Ein Fragebogen zur Erfassung von Präferenzen der Nutzer hinsichtlich Automatisierungsaufgaben, basierend auf den Modellen von Parasuraman et al. (2000). Dieser misst, inwieweit Nutzer Automatisierung in verschiedenen Phasen der Informationsverarbeitung bevorzugen.

Moradbakhti et. al (2024): Basic Psychological Need Satisfaction for Technology Use (BPN-TU)

Die BPN-TU ist eine Skala zur Messung der Befriedigung grundlegender psychologischer Bedürfnisse bei der Nutzung von Technologie. Gemäß der Selbstbestimmungstheorie ist die Befriedigung der grundlegenden psychologischen Bedürfnisse nach Autonomie, Kompetenz und Verbundenheit entscheidend für das Wohlbefinden und die autonome Motivation.

5. Design-Guidelines für eine gute UX

1. Adaptive Automatisierung ermöglichen

Beispiel: Verkehrsmanagementsysteme für städtischen Verkehr

Ein städtisches Verkehrsmanagementsystem bietet verschiedene Automatisierungsstufen, wie z. B. die automatische Steuerung von Ampeln oder die manuelle Steuerung durch Verkehrsingenieure. Über eine benutzerfreundliche Oberfläche können die Verantwortlichen je nach Verkehrsaufkommen und speziellen Ereignissen die Automatisierungsstufe flexibel anpassen.

Diese Anpassungsmöglichkeit erlaubt eine präzise und flexible Steuerung des Verkehrsflusses, reduziert Staus und priorisiert den öffentlichen Verkehr. Die Nutzenden behalten dabei die Kontrolle und können den Automatisierungsgrad individuell anpassen, was zu einem effizienteren und reibungsloseren Verkehrserlebnis führt.

2. Situationsbewusstsein unterstützen

Beispiel: Notfallmanagementsysteme in Städten

Ein Notfallmanagementsystem liefert Echtzeitdaten zu städtischen Notfällen wie Verkehrsunfällen, Bränden oder Überschwemmungen. Es bietet automatisierte Empfehlungen für Evakuierungs Routen und Einsatzplanungen, die Einsatzleiter bei Bedarf manuell anpassen können.

Dank der Echtzeitinformationen können Einsatzkräfte schnell und präzise Entscheidungen treffen. Die Kombination aus automatisierten Vorschlägen und menschlichem Eingriff auf mittleren Automatisierungsstufen sorgt für ein optimales Gleichgewicht zwischen Effizienz und Sicherheit.

3. Flexibilität bei der Informationsverarbeitung

Beispiel: Umweltüberwachungssysteme in Städten

Ein Umweltüberwachungssystem erlaubt es den Nutzenden, die Art und Menge der überwachten Daten individuell festzulegen – etwa zur Überwachung der Luftqualität, Lärmelastung oder Wasserverschmutzung. Nutzende können den Fokus je nach Dringlichkeit und Prioritäten anpassen und die Detailtiefe der Analysen steuern.

Diese Flexibilität ermöglicht es den Nutzenden, auf spezifische Umweltfaktoren einzugehen und die Überwachung an aktuelle Bedürfnisse anzupassen. Dadurch wird eine maßgeschneiderte Umweltpolitik möglich, die effektiver auf akute Herausforderungen reagiert.

4. Transparenz sicherstellen

Beispiel: Medizinisches Diagnosetools

Ein KI-basiertes Diagnosetool für Ärzte zeigt nicht nur die Diagnoseergebnisse, sondern auch die zugrunde liegenden Daten und die Logik hinter der Entscheidung an. Die Entscheidungswege werden visualisiert, und das System erklärt, warum bestimmte Diagnosen vorgeschlagen wurden.

Diese Transparenz stärkt das Vertrauen der Ärzte in die KI, da sie genau nachvollziehen können, wie die Empfehlungen zustande kommen. Dies fördert einen effizienteren und informierten Entscheidungsprozess.

Beispiel: Überwachungssysteme für den öffentlichen Verkehr

Ein städtisches Verkehrssystem analysiert den Verkehrsfluss und erklärt transparent, wie Ampelschaltungen optimiert oder bestimmte Routen priorisiert werden.

Durch die klare Kommunikation der Algorithmen und Entscheidungsprozesse wird das Vertrauen der Öffentlichkeit gestärkt. Nutzende können nachvollziehen, wie Entscheidungen getroffen wurden, was ihre Akzeptanz und das Gefühl der Autonomie im Umgang mit dem System verbessert.

5. Nutzerzentrierte Anpassung

Beispiel: Smart City Mobilitätsplattformen

Eine Smart City Mobilitätsplattform ermöglicht es den Bürgern, ihre persönlichen Verkehrspräferenzen festzulegen - von bevorzugten Verkehrsmitteln über favorisierte Routen bis hin zu umweltbewussten Zielen wie der Reduzierung des CO2-Fußabdrucks. Die Plattform generiert daraufhin maßgeschneiderte Vorschläge, z. B. alternative Verkehrsmittel oder Fahrgemeinschaften, die den individuellen Präferenzen der Nutzenden entsprechen.

Diese nutzerzentrierte Anpassung gibt den Bürgern das Gefühl der Kontrolle über ihre Mobilitätsentscheidungen. Das Ergebnis ist eine höhere Zufriedenheit, da die Plattform auf persönliche Vorlieben eingeht. Gleichzeitig unterstützt die Lösung städtische Ziele zur Förderung nachhaltiger Mobilität.

6. Fazit

1

Die **wahrgenommene Autonomie** in der Interaktion mit KI-Systemen beeinflusst **Nutzerzufriedenheit** und **Vertrauen**.

2

Automatisierungsstufen und **Anpassungsfähigkeit** haben Einfluss auf die psychologische Wahrnehmung.

3

Transparenz, Flexibilität und **Anpassung der Automatisierungsgrade** stärken die wahrgenommene Autonomie.

4

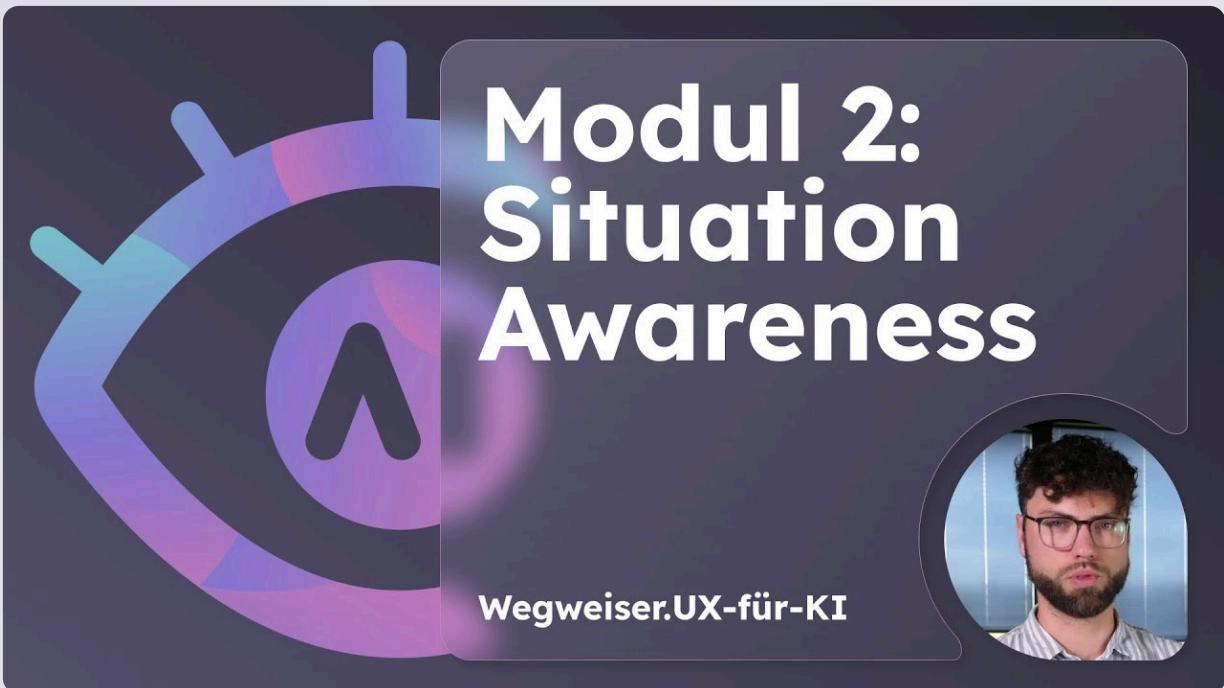
Adaptive Automatisierung und **transparente Entscheidungsprozesse** fördern **Kontrolle** und **Selbstbestimmung**, was zu effektiverer und zufriedenstellender Nutzung führt.

Wahrgenommenes Situationsbewusst- sein

Kursübersicht > KI-bezogene UX

Dies bezieht sich auf das Verständnis der Nutzende über ihre Umgebung und die Änderungen, die durch das KI-System verursacht werden.

Im folgenden Video wird ein Überblick über das wahrgenommene Situationsbewusstsein gegeben.



<https://youtu.be/mhUXTuMQ5mY>

1. Definition **Situationsbewusstsein**

Situationsbewusstsein (Situation Awareness, SA) bedeutet, die Umgebung und ihre wichtigsten Details zu verstehen. Es umfasst auch das Erkennen von Veränderungen, die im Laufe der Zeit oder durch äußere Einflüsse auftreten können. Diese Fähigkeit ist wichtig, um gute Entscheidungen in verschiedenen Situationen zu treffen.

2. Relevante Konzepte und Modelle

Drei Ebenen des Situationsbewusstsein laut Endsley (2000)

- 1. Wahrnehmung:** Die Erfassung der relevanten Informationen in einer Umgebung.
- 2. Verständnis:** Die Interpretation und das Verstehen der Bedeutung dieser Informationen.
- 3. Projektion:** Die Fähigkeit, zukünftige Zustände oder Entwicklungen einer Situation vorherzusagen, basierend auf dem aktuellen Verständnis und der Wahrnehmung der Situation.

Das **HASO-Modell** von Endsley betont ebenfalls diese drei wesentlichen Eigenschaften von Automatisierungssystemen, da sie direkt mit der Situationswahrnehmung der Nutzenden verknüpft sind.

- 1. Transparenz:** Das System muss klar kommunizieren, wie es zu seinen Entscheidungen gelangt.
- 2. Verständlichkeit:** Die Darstellung der Informationen muss leicht nachvollziehbar sein.
- 3. Vorhersehbarkeit:** Nutzende müssen einschätzen können, wie das System sich unter bestimmten Bedingungen verhalten wird.

3. Studien zur User Experience und KI

Studien wie die von **Edgar et al. (2018)** untersuchen das Situationsbewusstsein in der Mensch-Maschine-Interaktion. Sie berechnen das wahrgenommene Situationsbewusstsein auf Basis einer Vertrauensbewertung durch die Bewertung wahr/falsch-Aussagen, wobei hier ebenfalls keine signifikanten Zusammenhang mit dem Verhalten festgestellt werden können. Es ist argumentierbar, dass die Gleichsetzung von wahrgenommenem Situationsbewusstsein und Vertrauen kritisch betrachtet werden kann, da sich das Situationsbewusstsein über drei verschiedene Stufen entwickelt.

Endsley et al. (1995) betont, dass Transparenz, Verständlichkeit und Vorhersehbarkeit entscheidend für die Aufrechterhaltung des Situationsbewusstseins und des Vertrauens in automatisierte Systeme sind. Während Endsley (1998) einerseits betont, dass das wahrgenommene Situationsbewusstsein entscheidend für die Handlungsregulation ist, stellt sie auch fest, dass die Korrelation zwischen wahrgenommenem (oder subjektivem) Situationsbewusstsein oft gering ist.

In der Studie zur User Experience in Digital Contact Tracing (DCT) von **Schrills et al. (2024)** wurde gezeigt, dass das subjektive Situationsbewusstsein der Nutzenden stärker mit der wahrgenommenen Nützlichkeit zusammenhängt als das faktische Situationsbewusstsein.

4. Operationalisierung: Fragebögen und Messinstrumente

Schrills & Franke (2023): SIPA (Subjective Information Processing Awareness)

Die SIPA-Skala ist ein Werkzeug, mit dem man beurteilen kann, wie Erklärungen in der erklärbaren Künstlichen Intelligenz (XAI) auf die Nutzenden wirken. Sie basiert auf den drei Ebenen des Situationsbewusstseins: Transparenz, Verständlichkeit und Vorhersehbarkeit. Mit der SIPA-Skala lässt sich analysieren, wie gut ein System die Nutzenden dabei unterstützt, das Verhalten und die Informationsverarbeitung des Systems nachzuvollziehen.

R. M. Taylor (2017): Situation Awareness Rating Technique (SART)

Die „Situation Awareness Rating Technique“ ist eine Methode zur Bewertung des wahrgenommenen Situationsbewusstseins, die 1990 veröffentlicht wurde. Diese umfasst weitere Konstrukte wie die Arbeitsbelastung und unterscheidet sich daher von Endsleys Konzept des Situationsbewusstseins.

5. Design-Guidelines zur Förderung des Situationsbewusstseins

1. Transparenz sicherstellen

Systeme sollten alle relevanten Elemente der Informationsverarbeitung offenlegen und den Nutzenden zugänglich machen.

Beispiel: Ein KI-gestütztes Dashboard für Ärzte zeigt visuell, welche Daten zur Diagnosestellung verwendet wurden.

2. Verständlichkeit fördern

Intuitive Benutzeroberflächen und kontextbezogene Hilfen sollten die Nutzung von Systemen erleichtern.

Beispiel: Medizinische Diagnose-Tools heben die wichtigsten Informationen hervor und bieten Hilfetexte zur Erklärung komplexer Funktionen.

3. Vorhersehbarkeit verbessern

Systeme sollten Rückmeldungen geben, die die Auswirkungen von Handlungen aufzeigen, z. B. durch Simulationen oder Vorschauen.

Beispiel: Ein System zur Verkehrssteuerung könnte simulieren, wie sich geänderte Ampelphasen auf den Verkehr auswirken, bevor sie tatsächlich implementiert werden.

6. Fazit

1

Situationsbewusstseins (SA) umfasst die **Wahrnehmung**, das **Verständnis** und die **Projektion** relevanter Informationen.

2

Endsley und die SIPA-Facetten betonen die Notwendigkeit von **Transparenz**, **Verständlichkeit** und **Vorhersehbarkeit** für effektives SA.

Wahrgenommene Mentale Belastung

Kursübersicht > KI-bezogene UX

Dieser Aspekt umfasst den kognitiven Aufwand, der erforderlich ist, um Informationen zu verarbeiten und Entscheidungen zu treffen, und die potenzielle Überlastung durch zu viele Informationen.

Was ist die wahrgenommene Mentale Belastung? In diesem Video wird ein Überblick gegeben.



<https://youtu.be/0GpnGeD7BCM>

1. Definition Mentale Belastung

Mentale Arbeitsbelastung (Mental Workload) beschreibt den kognitiven Aufwand, der nötig ist, um eine Aufgabe zu erledigen. Sie umfasst die geistigen Anstrengungen, die notwendig sind, um Informationen zu verarbeiten, Entscheidungen zu treffen und Aktionen durchzuführen.

2. Relevante Konzepte und Modelle

Mentale Arbeitsbelastung ist ein zentrales Konstrukt im Bereich der Automatisierung. Für erklärbare KI (XAI) spielt sie eine besonders wichtige Rolle, da Erklärungen einen paradoxen Effekt haben können: Während KI durch Automatisierung eigentlich eine effizientere Informationsverarbeitung ermöglichen soll, können Erklärungen die mentale Arbeitsbelastung wieder erhöhen. Sie ist daher eine wichtige Metrik für die Bewertung von XAI.

Wenn Menschen durch kognitive Zwänge bewusster über Entscheidungen nachdenken sollen, kann das zwar positive Effekte haben, aber die höhere mentale Anstrengung könnte dazu führen, dass sie diese Methode weniger gerne nutzen - besonders, wenn eine einfachere direkte Empfehlung verfügbar ist.

3. Studien zur User Experience und KI

Ergebnisse mehrerer Studien z. B. von **Sewnath und Crijnen (2021)** und **Tsai et al. (2021)** sowie eine Studie zur Automatisierung von Insulinverabreichungssystemen (AID-Systeme) von **Schrills und Franke (2023)** zeigten, dass der Einsatz von Erklärungen zu einer Informationsüberlastung führen könnte. Zu viele oder zu detaillierte Erklärungen in diesen Systemen beeinträchtigen die Entscheidungsfindung und erhöhen die kognitive Belastung. Insbesondere zeigte sich, dass Systeme mit hoher Informationsoffnenlegung nicht immer zu besseren Ergebnissen führten, sondern manchmal genau das Gegenteil bewirken.

4. Operationalisierung: Fragebögen und Messinstrumente

Hart (2006): NASA Task Load Index (NASA-TLX)

Eine weit verbreitete Methode zur Messung der mentale Arbeitsbelastung ist der **NASA Task Load Index (NASA-TLX)**. Dieser Fragebogen bewertet verschiedene Dimensionen der Arbeitsbelastung, darunter:

- Mentale Anforderungen
- Physische Anforderungen
- Zeitliche Anforderungen
- Leistung
- Anstrengung
- Frustration

Der NASA-TLX ist eine hilfreiche Methode, um die kognitive Belastung in verschiedenen Arbeitsumgebungen, einschließlich der Nutzung von KI-Systemen, zu erfassen.

5. Design-Guidelines zur Reduktion der kognitiven Belastung

1. Vereinfachung von Informationen

Informationen sollten minimalistisch dargestellt und nur schrittweise offengelegt werden, um die kognitive Belastung zu reduzieren.

Beispiel: Ein KI-gestütztes Dashboard zeigt nur die wesentlichen Informationen und blendet zusätzliche Details bei Bedarf ein.

2. Anpassbarkeit der Informationsmenge

Systeme sollten es den Nutzenden ermöglichen, die Menge an Informationen anzupassen, die sie benötigen, um Entscheidungen zu treffen.

Beispiel: Personalisierte Einstellungen, die es den Nutzenden erlauben, festzulegen, wie viele Details sie sehen möchten.

3. Fokussierung auf kritische Informationen

Das System sollte kontextbezogen die wichtigsten Informationen priorisieren.

Beispiel: In einem Verkehrsleitsystem werden in einer Notfallsituation nur die kritischsten Daten hervorgehoben, wie z. B. gesperrte Straßen oder gefährliche Wetterbedingungen.

4. Reduktion der kognitiven Belastung durch Automatisierung

Routinetätigkeiten sollten automatisiert werden, um die geistige Anstrengung der Nutzenden zu minimieren.

Beispiel: Ein KI-System im Verwaltungsbereich könnte die automatische Überprüfung von Anträgen übernehmen, sodass sich Mitarbeitende auf Ausnahmen und komplexe Fälle konzentrieren können.

6. Fazit

1

Mentale Arbeitsbelastung ist in der **erklärbaren KI (XAI)** entscheidend.

2

Automatisierung erleichtert die Verarbeitung, **Erklärungen** können jedoch **Informationsüberlastung** verursachen.

3

Wichtige Systemmerkmale: **Reduzierung, Anpassbarkeit und Priorisierung von Informationen**, um Nutzer effektiv zu entlasten.

4

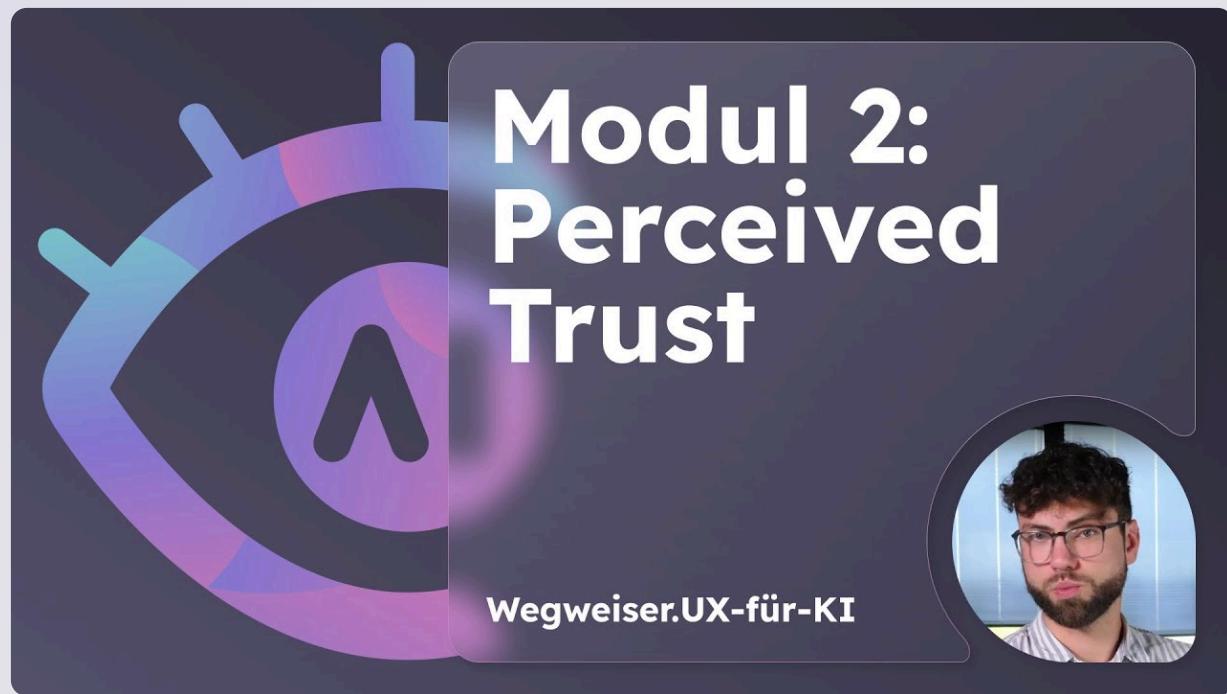
Der **NASA-TLX** hilft bei der Bewertung der Arbeitsbelastung und Verbesserung der Benutzerfreundlichkeit.

Wahrgenommene Vertrauenswürdig- keit

Kursübersicht > KI-bezogene UX

Damit ist das Vertrauen gemeint, das Nutzende in ein KI-System haben, basierend auf dessen Handlungen.

Dieses Video zeigt wie man einem System vertrauen kann und welche Formen es gibt.



1. Definition Wahrgenommene Vertrauenswürdigkeit

Im Kontext von KI-Systemen beschreibt die wahrgenommene Vertrauenswürdigkeit (perceived trustworthiness) das Vertrauen, das Nutzende einem KI-System entgegenbringen, basierend auf ihrer Wahrnehmung von dessen Zuverlässigkeit, Verständlichkeit und ethischen Prinzipien. Vertrauen wird als die Bereitschaft definiert, sich auf die Entscheidungen eines Systems zu verlassen, ohne es direkt überwachen oder kontrollieren zu können.

Vertrauen ist zentral für die Akzeptanz und Nutzung von KI-Systemen, insbesondere im Bereich der erklärbaren Künstlichen Intelligenz (XAI), wo ethische und soziale Aspekte besonders relevant sind.

2. Relevante Konzepte und Modelle

Drei Schlüsselfaktoren des Vertrauens in Systeme nach Mayer et al. (1995):

- 1. Fähigkeit:** Die Kompetenz und Fertigkeiten des Systems, seine Aufgaben korrekt auszuführen.
- 2. Wohlwollen:** Das Maß, in dem das System im besten Interesse des Nutzenden handelt.
- 3. Integrität:** Das Vertrauen, dass das System nach ethischen Prinzipien und Regeln handelt.

Vertrauen als eine Einstellung nach Lee & See (2004)

Vertrauen wird als die Einstellung beschrieben, dass ein Agent/System dazu beitragen wird, die Ziele einer Person in einer Situation zu erreichen, die durch Unsicherheit und Verwundbarkeit gekennzeichnet ist. Es ist wichtig zu beachten, dass Vertrauen als eine Einstellung und nicht als ein Verhalten konzeptualisiert werden sollte. Dabei sind drei Faktoren entscheidend:

- 1. Leistung:** Die Genauigkeit und Zuverlässigkeit des Systems.
- 2. Zweck:** Die Ziele und Absichten, die das System verfolgt.
- 3. Prozess:** Die Methoden und Verfahren, die das System verwendet.

Kognitives vs. affektives Vertrauen nach Madsen und Gregor (2000):

Sie beschreiben Vertrauen als „das Ausmaß, in dem ein Nutzer Vertrauen in die Empfehlungen, Handlungen und Entscheidungen einer künstlichen Entscheidungsunterstützung hat und bereit ist, auf deren Basis zu handeln.“

Sie unterscheiden zwischen:

- 1. Kognitives Vertrauen:** Basierend auf den wahrgenommenen Eigenschaften des Systems, wie Zuverlässigkeit und Verständlichkeit. Wenn das System beispielsweise transparent ist und nachvollziehbare Entscheidungen trifft, steigt das Vertrauen.
- 2. Affektives Vertrauen:** Emotionale Bindungen oder persönliche Erfahrungen mit dem System fördern das Vertrauen, besonders bei sprachbasierten Assistenzsystemen oder Robotern.

Drei Dimensionen des Vertrauens in der Automatisierung nach Hoff und Bashir (2015):

- 1. Dispositionelles Vertrauen:** bezieht sich auf die generelle Tendenz eines Nutzers, Automatisierung zu vertrauen, basierend auf Persönlichkeit und bisherigen Erfahrungen.
- 2. Situatives Vertrauen:** wird durch den spezifischen Kontext beeinflusst, in dem die Automatisierung verwendet wird, einschließlich Aufgabenmerkmalen und Umweltfaktoren.
- 3. Erlerntes Vertrauen:** entwickelt sich über die Zeit durch Interaktionen mit dem System, wobei positive Erfahrungen das Vertrauen stärken und negative es schwächen.

Dieses Modell ist besonders wertvoll, da es die dynamische Natur von Vertrauen und die Bedeutung der Nutzererfahrungen im Zeitverlauf hervorhebt.

Vertrauenswürdigkeits-Hinweisen nach Schlicker et al. (2022):

Vertrauenswürdigkeits-Hinweise sind entscheidend für die wahrgenommene Vertrauenswürdigkeit eines Systems. Hinweise zur Informationsverarbeitung zeigen nicht nur, wie zuverlässig das System ist, sondern helfen den Nutzern auch, dessen Funktionsweise besser zu verstehen und seine Fähigkeit zur Aufgabenbewältigung einzuschätzen. Hinweise, die nicht mit den Merkmalen der Informationsverarbeitung des Systems zusammenhängen, z. B. die Reputation des Herstellers oder soziale Hinweise wie das Nutzungsverhalten von Personen im eigenen Umfeld können ebenfalls relevant sein.

3. Studien zur User Experience und KI

Schrills (2024): Einfluss der Systemzuverlässigkeit auf Nutzervertrauen in KI-Systeme

Die Studie zeigt, dass die Selbsteinschätzung des Vertrauens in ein KI-System keinen direkten Einfluss auf das tatsächliche Vertrauen in dessen Empfehlungen hat. Stattdessen erwies sich die angegebene Zuverlässigkeit des Systems als der stärkste Einflussfaktor auf das Verhalten der Nutzenden. Das bedeutet, dass Nutzende ihr Vertrauen eher auf die wahrgenommene Zuverlässigkeit des Systems stützen als auf ihr eigenes, subjektives Vertrauensempfinden. Dieses Ergebnis legt nahe, dass die Kommunikation und Darstellung der Zuverlässigkeit eines KI-Systems einen größeren Einfluss auf die Nutzung hat als traditionelle Methoden zur Messung von Vertrauen.

Vereschak et al. (2024): Menschliche Einflüsse und Stakeholder-spezifische Anforderungen

- 1. Vertrauensanforderungen:** Die Teilnehmenden identifizierten wesentliche Elemente für Vertrauen und unterschieden es von Konzepten wie Vertrauenswürdigkeit, Verlass und Befolgung. Positive Erwartungen und wahrgenommenes Risiko waren dabei entscheidende Faktoren, wobei die Komplexität der Aufgabe als zusätzliche Voraussetzung für Vertrauen in KI-gestützte Entscheidungsprozesse hervorgehoben wurde.
- 2. Menschlicher Einfluss auf Vertrauen:** Vertrauen in KI-Systeme wurde stark von menschlichen Akteuren beeinflusst, etwa von den Personen, die das System entwickeln und einsetzen, und nicht nur von den technischen Merkmalen des Systems.
- 3. Stakeholderspezifische Vertrauensfaktoren:** Die Faktoren, die das Vertrauen zwischen Mensch und KI beeinflussen, variieren je nach Stakeholder. Zum Beispiel legen Entscheidungsträger, die das System nutzen, und diejenigen, die von den Entscheidungen betroffen sind (z. B. Patienten im medizinischen Kontext), auf unterschiedliche Vertrauensaspekte Wert.

4. Operationalisierung: Fragebögen und Messinstrumente

Jian et al. (2001): Trust in Automation (TiA)-Skala

Die Skala ist eine der am häufigsten verwendeten Skalen zur Selbsteinschätzung von Vertrauen. Sie bewertet, wie sehr Nutzer die Leistungsfähigkeit und Zuverlässigkeit automatisierter Systeme einschätzen und unterscheidet dabei zwischen Mensch-Mensch- und Mensch-Maschine-Vertrauen. Die Skala erfasst zentrale Vertrauensfaktoren und hilft, das Vertrauen der Nutzer in Automation und dessen Einfluss auf die Nutzung zu verstehen.

5. Design-Guidelines zur Förderung der Vertrauenswürdigkeit

1. Förderung von Transparenz und Verständlichkeit

Erklärbare KI (XAI) ermöglicht, dass die Entscheidungen des Systems klar und nachvollziehbar sind, was das Vertrauen der Nutzer stärkt.

Beispiel: Ein Finanzplanungssystem, das seine Berechnungen offenlegt und erklärt, welche Parameter berücksichtigt wurden, hilft Nutzern, die Entscheidungsprozesse nachzuvollziehen.

2. Sicherstellung von Systemleistung und Zuverlässigkeit

Systeme sollten verlässlich arbeiten und klare Rückmeldungen geben, um Vertrauen aufzubauen und Unsicherheiten zu reduzieren.

Beispiel: Ein medizinisches KI-System, das bei der Diagnose Fehlerindikatoren und Bestätigungen anzeigt, schafft Vertrauen in die Genauigkeit und Zuverlässigkeit der Ergebnisse.

3. Stärkung des erlernten Vertrauens

Einführungen und Schulungen können dispositionelles Vertrauen fördern, indem sie das Verständnis und die Vertrautheit mit dem System verbessern.

Beispiel: Nutzerfreundliche Tutorials für ein neues KI-gestütztes Verwaltungssystem fördern das grundlegende Vertrauen der Nutzer in das System.

4. Förderung von affektivem Vertrauen

Eine personalisierte Nutzererfahrung und positive User Experience (UX) tragen zur Entwicklung von affektivem Vertrauen bei.

Beispiel: Ein Sprachassistentensystem, das personalisierte Präferenzen berücksichtigt, stärkt die emotionale Bindung und das Vertrauen der Nutzer in das System.

5. Kontextbezogene Erklärungen für Empfehlungen geben

Um situatives Vertrauen zu fördern, sollten Systeme kontextabhängige Erklärungen liefern, die auf die spezifische Entscheidungssituation der Nutzer eingehen.

Beispiel: Ein Personalplanungssystem für das Gesundheitswesen erklärt, warum bestimmte Dienstpläne vorgeschlagen werden, z. B. basierend auf Arbeitslasten oder Personalverfügbarkeit.

6. Fazit

1

Vertrauen ist entscheidend: Vertrauenswürdigkeit fördert die Akzeptanz von KI-Systemen.

2

Unterschiedliche Vertrauensfaktoren: Fähigkeit, Integrität und Wohlwollen sind grundlegende Faktoren für Vertrauen.

3

Menschliche Einflüsse berücksichtigen: Vertrauen wird nicht nur durch das System, sondern auch durch menschliche Interaktionen und den sozialen Kontext beeinflusst.

4

Zuverlässigkeit sichtbar machen: Nutzervertrauen basiert stark auf der wahrgenommenen Zuverlässigkeit des Systems.

Wahrgenommene Diagnostizität

Kursübersicht > KI-bezogene UX

Dies beschreibt das Vertrauen der Nutzende in die Diagnosen oder Vorschläge des KI-Systems und wie gut diese die gewünschten Ergebnisse liefern.

Im folgenden Video wird ein Überblick über den Begriff Diagnostizität gegeben.



<https://youtu.be/jdsnWIIGT7c>

1. Definition wahrgenommene Diagnostizität

Die wahrgenommene Diagnostizität (Perceived Confidence/Diagnosticity) beschreibt, wie nützlich ein System für Nutzende ist, um fundierte Entscheidungen zu treffen. Ein System mit hoher Diagnostizität unterstützt die Bewertung verschiedener Optionen, sodass Nutzende auf Basis der vom System bereitgestellten Informationen die beste Entscheidung treffen können.

Je präziser und hilfreicher die Hinweise sind, desto mehr steigt die diagnostische Qualität des Systems und stärkt damit das Vertrauen der Nutzenden.

2. Relevante Konzepte und Modelle

Auswahl von Handlungsmöglichkeiten

Aus psychologischer Sicht ist Diagnostizität entscheidend, weil sie die Auswahl einer Handlung erleichtert. Informationen mit hoher Diagnostizität unterstützen Menschen dabei, gezielte Entscheidungen zu treffen. So kann ein diagnostisches System im Gesundheitswesen dazu beitragen, zwischen wenigen klaren Hypothesen zu unterscheiden, was letztlich Unsicherheit reduziert und die Entscheidungsfindung erleichtert.

Ein Beispiel dafür ist ein KI-System im Gesundheitswesen, das bei Symptomen wie Fieber die Wahrscheinlichkeiten für mögliche Erkrankungen wie Grippe oder Erkältung analysiert. Durch das Bereitstellen spezifischer Informationen hilft das System, die Entscheidung über den nächsten Schritt gezielt zu erleichtern.

Unterschied zwischen wahrgenommener und tatsächlicher Diagnostizität

Manchmal schätzen Nutzende bestimmte Informationen als sehr hilfreich für eine Entscheidung ein, obwohl diese Infos gar nicht wirklich dabei helfen, zwischen verschiedenen Möglichkeiten zu unterscheiden. Dies kann dazu führen, dass unwichtige Details überbewertet werden und dadurch Fehlentscheidungen entstehen (Nelson, 2005)

Beispielsweise hat das Symptom „Fieber“ bei der Unterscheidung zwischen COVID-19 und Grippe nur einen geringen diagnostischen Wert, auch wenn es allgemein zur Krankheitsdiagnose beiträgt.

Dieser Effekt führt zu sogenannten „Pseudo-Diagnosen“, bei denen Menschen diagnostisch wenig hilfreiche Daten auswählen und diese für ihre Entscheidung nutzen (Kern & Doherty, 1982). Ähnlich wie bei der „Illusion of Explanatory Depth“ (Chromik et al., 2021) können dadurch Fehleinschätzungen der eigenen Leistung und eine falsche Bewertung des Systems entstehen.

Unterschied zwischen Informationswert und Diagnostizität

Der Begriff „Informationswert“ bezieht sich darauf, wie stark eine Information generell Unsicherheit verringert. „Diagnostizität“ hingegen ist spezifischer: Sie hilft, zwischen bestimmten Möglichkeiten (z. B. Grippe vs. Erkältung) zu unterscheiden. Wenn es nur wenige klare Optionen gibt, ist Diagnostizität wichtiger. Wenn es viele mögliche Optionen gibt, hilft der allgemeine Informationswert mehr.

3. Studien zur User Experience und KI

Schrills und Franke (2023): Einfluss der Diagnostizität auf Vertrauen und Nutzungsabsicht

Schrills und Franke (2023) betonen, dass die wahrgenommene Diagnostizität entscheidend für die Vertrauenswürdigkeit und Nutzung von Systemen ist. Bei Digital Contact Tracing (DCT)-Apps zur Pandemiekontaktverfolgung bevorzugen Nutzende detaillierte und klare Informationen, besonders diagnostische Details wie das Tragen einer Maske während einer Pandemie. Dies stärkt ihr Vertrauen in die Genaugigkeit und Nützlichkeit der App und erhöht somit auch Zufriedenheit. Das Fehlen diagnostischer Hinweise beeinflusst das Verhalten der Nutzende und bringt sie dazu, eher allgemeine, weniger gezielte Informationen zu nutzen.

Bartlett und McCarley (2017): Suboptimale Entscheidungen durch fehlende Diagnostizität

Wenn eine KI keine spezifisch diagnostischen Informationen liefert, neigen Nutzenden oft dazu, alternative Strategien anzuwenden, die nicht immer optimal sind. Die Studie beschreibt, dass Menschen in solchen Situationen oft eine Strategie namens „probability matching“ verwenden. Dabei passen sie ihre Entscheidungen an die allgemeine Zuverlässigkeit der KI an, anstatt auf gezielte, diagnostische Hinweise zu achten.

4. Operationalisierung: Fragebögen und Messinstrumente

Wahrgenommene Ergebnisdagnostik

Hier wird erfasst, inwieweit die Nutzer der Meinung sind, dass die bereitgestellten Informationen fundiert sind und zu einer besseren Entscheidungsfindung beitragen.

Fragebögen zur Messung von Vertrauenswürdigkeit und Diagnostizität:

Diese Instrumente bewerten, wie präzise und klar die Informationen eines Systems wahrgenommen werden und wie stark sie das Vertrauen der Nutzer beeinflussen.

5. Design-Guidelines

Um Nutzer das Verständnis und Vertrauen in die Entscheidungen einer Künstlichen Intelligenz (KI) zu erleichtern, gibt es verschiedene Techniken in der Erklärbaren KI (XAI). Diese Techniken helfen, Entscheidungen der KI verständlicher und transparenter zu machen.

1. Kontrafaktische Erklärungen – „Was wäre wenn“-Szenarien

Kontrafaktische Erklärungen zeigen, wie eine kleine Änderung an den Eingabedaten zu einer anderen Entscheidung der KI führen könnte. Zum Beispiel: Wenn eine Kreditbewilligung abgelehnt wird, könnte die KI erklären, dass eine Erhöhung des Einkommens um einen bestimmten Betrag zur Bewilligung geführt hätte. Solche „Was wäre wenn“-Erklärungen helfen den Nutzenden, die Entscheidungsgrenzen der KI zu verstehen (Warren et al., 2022).

2. Semantische Anreicherung und Heatmaps – Verständlichere Visualisierungen

Durch semantische Anreicherung, also durch das Hinzufügen von zusätzlichen Erklärungen, werden Heatmaps (visuelle Darstellungen) verständlicher. Zum Beispiel in der medizinischen Bildgebung können Heatmaps aufzeigen, welche Bildbereiche für eine Diagnose besonders wichtig waren. Diese zusätzlichen Details helfen den Nutzenden, die Entscheidungslogik der KI besser nachzuvollziehen und das Vertrauen zu stärken (Gianfagna & Di Cecco, 2021; Tonekaboni et al., 2019).

3. Konfidenzbewertungen – Vertrauen durch Unsicherheitsangaben

Konfidenzbewertungen zeigen, wie sicher oder unsicher die KI bei ihren Vorhersagen ist. Diese Informationen sind besonders wichtig in Bereichen wie Gesundheit und Finanzen, wo Entscheidungen große Auswirkungen haben können. Wenn die KI ihre Unsicherheit angibt, können Nutzende besser einschätzen, ob sie der Entscheidung vertrauen möchten oder nicht (Gianfagna and Di Cecco, 2021; T. Le et al., 2023).

4. Lokale Feature-Relevanz – Bedeutung einzelner Merkmale

Die lokale Feature-Relevanz zeigt an, welche spezifischen Eingabemerkmale zu einer bestimmten Entscheidung der KI geführt haben. Beispielsweise könnte in einem System zur Betrugserkennung hervorgehoben werden, dass ungewöhnliche Transaktionsbeträge oder -orte eine große Rolle gespielt haben. Diese Detailinformationen helfen den Nutzenden, die Entscheidungen der KI mit ihrem eigenen Wissen zu vergleichen und zu validieren (Doshi-Velez and Kim, 2017; Lundberg et al., 2019).

6. Fazit

1

Die wahrgenommene Diagnostizität eines KI-Systems ist entscheidend für fundierte und präzise Entscheidungen.

2

Systeme, die klare diagnostische Informationen bereitstellen, stärken das Vertrauen der Nutzenden sowie die Akzeptanz und effektive Nutzung der Technologie.

3

Besonders wichtig ist dies in kritischen Bereichen wie dem Gesundheitswesen.

4

Die diagnostische Qualität eines Systems hat direkten Einfluss auf das Wohl der Gemeinschaft.

Zusammenfassung und Ausblick

Kursübersicht > KI-bezogene UX

Zum Ende des Moduls wird eine kurze Zusammenfassung über die Inhalte des Moduls gegeben.



<https://youtu.be/Qmg1FIFCSMw>

In dieser Lektion haben wir uns eingehend mit den spezifischen Aspekten der User Experience (UX) auseinandergesetzt, die im Kontext von KI-Systemen eine besondere Rolle spielen.

Die 5 UX-Kernaspekte und deren Designrichtlinien

1. Adaptive Automatisierung für wahrgenommene Autonomie

Wie sehr fühlen sich Nutzende in der Lage, selbstständig Entscheidungen zu treffen und zu handeln, während sie mit einem KI-System interagieren? Nutzende sollten die Kontrolle darüber haben, wie stark sie das System automatisieren oder manuell bedienen möchten. Dies fördert das Gefühl der Autonomie und gibt ihnen die Flexibilität, sich bei Bedarf stärker auf ihre eigenen Entscheidungen zu verlassen.

Bieten Sie Nutzenden die Möglichkeit, zwischen verschiedenen Automatisierungsstufen zu wechseln, sodass sie je nach Präferenz oder Aufgabenanforderung selbst entscheiden können, wie viel Kontrolle sie dem System überlassen.

2. Transparenz zur Unterstützung des Situationsbewusstseins

Situationsbewusstsein ist das Verständnis und die Wahrnehmung der Nutzenden über die aktuelle Umgebung und die Auswirkungen der KI auf diese. Ein hohes Maß an Situationsbewusstsein erfordert, dass Nutzende jederzeit relevante Informationen über das System und dessen Entscheidungen erhalten. Transparenz fördert das Verständnis für die Funktionsweise und Entscheidungsgrundlagen des Systems.

Stellen Sie sicher, dass das System die Datenquellen, Prozesse und Faktoren, die eine Entscheidung beeinflussen, klar kommuniziert. Nutzen Sie visuelle Darstellungen oder erklärende Hinweise, um komplexe Abläufe verständlicher zu machen.

3. Flexibilität bei der Informationsverarbeitung zur Reduzierung der mentalen Belastung

Mentale Belastung beschreibt den kognitiven Aufwand, der durch die Verarbeitung von Informationen entsteht und die potenzielle Überlastung durch zu viele Daten. Da zu viele oder unstrukturierte Informationen zu kognitiver Überlastung führen können, ist es entscheidend, Nutzende die Kontrolle über die Art und Menge der angezeigten Informationen zu geben. Dies hilft, die wahrgenommene mentale Belastung zu reduzieren.

Integrieren Sie Funktionen, die es Nutzende ermöglichen, die Anzeige von Informationen nach Bedarf zu filtern, zu kategorisieren oder zu priorisieren. Zum Beispiel kann eine Zusammenfassungsansicht für weniger erfahrene Nutzer und eine Detailansicht für Experten angeboten werden.

4. Zuverlässigkeit und klare Darstellung zur Förderung der wahrgenommenen Vertrauenswürdigkeit

Vertrauen entsteht, wenn das System nicht nur zuverlässig und präzise arbeitet, sondern diese Eigenschaften auch klar vermittelt. Nutzende müssen darauf vertrauen können, dass das System korrekt und ethisch agiert.

Verwenden Sie visuelle Indikatoren, die die Zuverlässigkeit und Erfolgsquote des Systems darstellen, und bieten Sie Erklärungen, die das ethische und technische Verhalten des Systems untermauern.

5. Nutzerzentrierte Anpassung zur Unterstützung der wahrgenommenen Diagnostizität

Diagnostizität ist die Fähigkeit eines Systems, Nutzende mit präzisen Informationen bei fundierten Entscheidungen zu unterstützen. Nutzende sollten das System an ihre individuellen Präferenzen und Informationsbedürfnisse anpassen können, damit es sie optimal bei der Entscheidungsfindung unterstützt. Dies stärkt die Fähigkeit des Systems, genaue und hilfreiche Informationen bereitzustellen.

Ermöglichen Sie personalisierte Einstellungen, durch die Nutzenden entscheiden können, welche Art von Informationen angezeigt werden und wie detailliert diese sein sollen. Ein personalisiertes Dashboard oder konfigurierbare Berichte können helfen, das System effizienter zu nutzen.

Ausblick auf das Modul: "Gestaltungsziele für menschzentrierte KI"

Im kommenden Modul werden wir uns mit spezifischen Eigenschaften von KI-Systemen beschäftigen, die besonders wichtig für die UX sind: Erklärbarkeit, Nachvollziehbarkeit, Vertrauenswürdigkeit, Kontrollierbarkeit und Transparenz. Wir werden untersuchen, wie diese Eigenschaften die Nutzererfahrung beeinflussen und welche Prinzipien bei der Gestaltung von KI-Systemen berücksichtigt werden sollten, um eine positive und vertrauensvolle Interaktion zu gewährleisten.

Wir hoffen, dass diese Lektion Ihnen wertvolle Einblicke in die Relevanz von UX-Konstrukten bei der Gestaltung von KI-Systemen gegeben hat. Diese Konzepte sind nicht nur für die Entwicklung von KI-Systemen entscheidend, sondern auch für die Bewertung, wie gut diese Systeme die Bedürfnisse und Erwartungen der Nutzer erfüllen. Bereiten Sie sich darauf vor, tiefer in die Dimensionen einzutauchen, die das Nutzererlebnis mit KI-Systemen weiter prägen werden.

08

Quellen

Kursübersicht > KI-bezogene UX

Literatur zu Wahrgenommener Autonomie

- Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model for types and levels of human interaction with automation. **IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans**, 30(3), 286-297. <https://doi.org/10.1109/3468.844354>
- Hopkins, D., & Schwanen, T. (2021). Talking about automated vehicles: What do levels of automation do? **Technology in Society**, 64, 101488. <https://doi.org/10.1016/j.techsoc.2020.101488>
- Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human performance consequences of stages and levels of automation: An integrated meta-analysis. **Human Factors: The Journal of the Human Factors and Ergonomics Society**, 56(3), 476-488. <https://doi.org/10.1177/0018720813501549>
- Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. **Theoretical Issues in Ergonomics Science**, 5(2), 113-153. <https://doi.org/10.1080/1463922021000054335>

- Rieger, T., Roesler, E., & Manzey, D. (2022). Challenging presumed technological superiority when working with (artificial) colleagues. **Scientific Reports**, **12**, 3768. <https://doi.org/10.1038/s41598-022-07808-x>
- Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. **Psychological Inquiry**, **11**(4), 227–268.
- Zoubir, M. (2024). **Preference for Automation Types Scale (PATS)**. <https://doi.org/10.13140/RG.2.2.22149.97769>
- Moradbakhti, L., Leichtmann, B., & Mara, M. (2024). Development and validation of a basic psychological needs scale for technology use. **Psychological Test Adaptation and Development**, **5**(1), 26–45. <https://doi.org/10.1027/2698-1866/a000062>

Literatur zu Wahrgenommenem Situationsbewusstsein (SA)

- Edgar, G. K., Catherwood, D., Baker, S., Sallis, G., Bertels, M., Edgar, H. E., Nikolla, D., Buckle, S., Goodwin, C. & Whelan, A. (2018). Quantitative Analysis of Situation Awareness (QASA): Modelling and measuring situation awareness using signal detection theory. **Ergonomics**, **61**(6), 762–777. <https://doi.org/10.1080/00140139.2017.1420238>
- Endsley, M., Sollenberger, R. & Stein, E. (2000). Situation awareness: A comparison of measures. In **Proceedings of the Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium**, Savannah, GA.

- Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. **Human Factors**, 37(1), 32–64.
<https://doi.org/10.1518/001872095779049543>
- Endsley, M. R. (2017). From Here to Autonomy: Lessons Learned From Human-Automation Research. **Human Factors: The Journal of the Human Factors and Ergonomics Society**, 59(1), 5–27.
<https://doi.org/10.1177/0018720816681350>
- Endsley, M. R., Selcon, S. J., Hardiman, T. D. & Croft, D. G. (1998). A Comparative Analysis of Sagat and Sart for Evaluations of Situation Awareness. In **Proceedings of the Human Factors and Ergonomics Society Annual Meeting**, 42(1), 82–86.
<https://doi.org/10.1177/154193129804200119>
- Schrills, T., & Franke, T. (2023). Wie erleben Nutzer die Nachvollziehbarkeit von KI-Systemen? Untersuchung des subjektiven Informationsverarbeitungsbewusstseins in automatisierten Insulinabgabesystemen (AID-Systemen). *ACM Transactions on Interactive Intelligent Systems*, 13(4), 25:1–25:34.
<https://doi.org/10.1145/3588594>
- Schrills, T., Kojan, L., Gruner, M., Calero Valdez, A. & Franke, T. (2024). Effects of User Experience in Automated Information Processing on Perceived Usefulness of Digital Contact-Tracing Apps: Cross-Sectional Survey Study. **JMIR Human Factors**, 11, e53940.
<https://doi.org/10.2196/53940>
- Taylor, R. M. (2017). Situational Awareness Rating Technique (Sart): The Development of a Tool for Aircrew Systems Design. In E. Salas (Hrsg.), **Situational Awareness** (1. Aufl., S. 111–128). Routledge.
<https://doi.org/10.4324/9781315087924-8>

Literatur zu Wahrgenommene Mentale Arbeitsbelastung

- Miller, T. (2023). Explainable AI is Dead, Long Live Explainable AI!: Hypothesis-driven Decision Support using Evaluative AI. **2023 ACM Conference on Fairness, Accountability, and Transparency**, 333–342.
<https://doi.org/10.1145/3593013.3594001>
- Vidulich, M. A., & Tsang, P. S. (2012). Mental Workload and Situation Awareness. In G. Salvendy (Hrsg.), **Handbook of Human Factors and Ergonomics** (1. Aufl., S. 243–273). Wiley.
<https://doi.org/10.1002/9781118131350.ch8>
- Longo, L., Wickens, C. D., Hancock, G., & Hancock, P. A. (2022). Human Mental Workload: A Survey and a Novel Inclusive Definition. **Frontiers in Psychology**, 13, 883321.
<https://doi.org/10.3389/fpsyg.2022.883321>
- Sewnath, G., & Crijnen, J. (2021). **How much is too much? Levels of AI Explainability within Decision Support Systems' User Interfaces for improved decision-making performance.**
- Tsai, C.-H., You, Y., Gui, X., Kou, Y., & Carroll, J. M. (2021). Exploring and Promoting Diagnostic Transparency and Explainability in Online Symptom Checkers. **Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems**, 1–17.
<https://doi.org/10.1145/3411764.3445101>
- Schrills, T., & Franke, T. (2023). How Do Users Experience Traceability of AI Systems? Examining Subjective Information Processing Awareness in Automated Insulin Delivery (AID) Systems. **ACM Transactions on Interactive Intelligent Systems**, 13(4), 1–34.
<https://doi.org/10.1145/3588594>

- Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 Years Later. **Proceedings of the Human Factors and Ergonomics Society Annual Meeting**, 50 (9), 904–908.
<https://doi.org/10.1177/154193120605000909>

Literatur zu Wahrgenommene Vertrauenswürdigkeit

- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. **The Academy of Management Review**, 20(3), 709. <https://doi.org/10.2307/258792>
- Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. **Human Factors: The Journal of the Human Factors and Ergonomics Society** , 46(1), 50–80.
https://doi.org/10.1518/hfes.46.1.50_30392
- Madsen, M., & Gregor, S. (2000). Measuring human-computer trust. **11th australasian conference on information systems** , 53, 6–8.
- Hoff, K. A., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. **Human Factors: The Journal of the Human Factors and Ergonomics Society** , 57(3), 407–434. <https://doi.org/10.1177/0018720814547570>
- Schlicker, N., Baum, K., Uhde, A., Sterz, S., Hirsch, M. C., & Langer, M. (2022). **How Do We Assess the Trustworthiness of AI? Introducing the Trustworthiness Assessment Model (TrAM)**.<https://doi.org/10.31234/osf.io/qhwvx>

- Schrills, T., & Franke, T. (2023). How Do Users Experience Traceability of AI Systems? Examining Subjective Information Processing Awareness in Automated Insulin Delivery (AID) Systems. **ACM Transactions on Interactive Intelligent Systems**, 13(4), 1–34.
<https://doi.org/10.1145/3588594>
- Vereschak, O., Alizadeh, F., Bailly, G., & Caramiaux, B. (2024). Trust in AI-assisted Decision Making: Perspectives from Those Behind the System and Those for Whom the Decision is Made. **Proceedings of the CHI Conference on Human Factors in Computing Systems**, 1–14.<https://doi.org/10.1145/3613904.3642018>
- Kohn, S. C., de Visser, E. J., Wiese, E., Lee, Y.-C., & Shaw, T. H. (2021). Measurement of trust in automation: A narrative review and reference guide. **Frontiers in psychology**, 12, 604977.
<https://doi.org/10.3389/fpsyg.2021.604977>

Literatur zu Wahrgenommene Confidence / Diagnosticity

- Wickens, C. D., & Scott, B. D. (1983). A comparison of verbal and graphical information presentation in a complex information integration decision task. In Tech. Rep. EPL-83-1/ONR-83-1. Engineering-Psychology Research Laboratory, University of Illinois Urbana.
- Nelson, J. D. (2005). Finding Useful Questions: On Bayesian Diagnosticity, Probability, Impact, and Information Gain. **Psychological Review**, 112(4), 979–999.
<https://doi.org/10.1037/0033-295X.112.4.979>

- Kern, L., & Doherty, M. E. (1982). ‘Pseudodiagnosticity’ in an idealized medical problem-solving environment. **Academic Medicine**, **57**(2), 100–104. <https://doi.org/10.1097/00001888-198202000-00004>
- Chromik, M., Eiband, M., Buchner, F., Krüger, A., & Butz, A. (2021). I Think I Get Your Point, AI! The Illusion of Explanatory Depth in Explainable AI. **26th International Conference on Intelligent User Interfaces**, 307–317. <https://doi.org/10.1145/3397481.3450644>
- Schrills, T., & Franke, T. (2023). How Do Users Experience Traceability of AI Systems? Examining Subjective Information Processing Awareness in Automated Insulin Delivery (AID) Systems. **ACM Transactions on Interactive Intelligent Systems**, **13**(4), 1–34. <https://doi.org/10.1145/3588594>
- Bartlett, M. L., & McCarley, J. S. (2017). Benchmarking Aided Decision Making in a Signal Detection Task. **Human Factors: The Journal of the Human Factors and Ergonomics Society**, **59**(6), 881–900. <https://doi.org/10.1177/0018720817700258>
- Warren, G., Smyth, B., & Keane, M. T. (2022). “Better” Counterfactuals, Ones People Can Understand: Psychologically-Plausible Case-Based Counterfactuals Using Categorical Features for Explainable AI (XAI). In M. T. Keane & N. Wiratunga (Eds.), **Case-Based Reasoning Research and Development** (Vol. 13405, pp. 63–78). Springer International Publishing. https://doi.org/10.1007/978-3-031-14923-8_5
- Gianfagna, L., & Di Cecco, A. (2021). Explainable AI: Needs, Opportunities, and Challenges. In L. Gianfagna & A. Di Cecco, **Explainable AI with Python** (pp. 27–46). Springer International Publishing. https://doi.org/10.1007/978-3-030-68640-6_2

- Tonekaboni, S., Joshi, S., McCradden, M. D., & Goldenberg, A. (2019). What clinicians want: Contextualizing explainable machine learning for clinical end use. In F. Doshi-Velez, J. Fackler, K. Jung, D. Kale, R. Ranganath, B. Wallace, & J. Wiens (Eds.), **Proceedings of the 4th Machine Learning for Healthcare Conference** (Vol. 106, pp. 359–380). PMLR.

<https://proceedings.mlr.press/v106/tonekaboni19a.html>

- Le, T., Miller, T., Singh, R., & Sonenberg, L. (2023). Explaining Model Confidence Using Counterfactuals. **Proceedings of the AAAI Conference on Artificial Intelligence**, 37(10), 11856–11864.

<https://doi.org/10.1609/aaai.v37i10.26399>

- Doshi-Velez, F., & Kim, B. (2017). **Towards A Rigorous Science of Interpretable Machine Learning** (No. arXiv:1702.08608). arXiv.

<http://arxiv.org/abs/1702.08608>

- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2019). **Explainable AI for Trees: From Local Explanations to Global Understanding** (No. arXiv:1905.04610). arXiv. <http://arxiv.org/abs/1905.04610>