CTSi//circle.responsibleComputing · Science Insight · 9. November 2026

Sich geschützt fühlen, geschützt sein, und warum das nicht dasselbe sein muss

Sicherheit, Safety und Privatsphäre, wie Menschen sie erleben

Rafael Vrecar

ÖAW DOC Fellow · CTSi und Human-Computer Interaction Group, TU Wien

Über mich

Mein Blickwinkel: Sicherheit, Safety und Privatsphäre als etwas, das Menschen erleben, nicht nur als Eigenschaft von Systemen



Rafael Vrecar

rafael.vrecar@tuwien.ac.at
vrecar.eu/rafael

HCI-Forscher und Doktorand, TU Wien

Software & Information Engineering (BSc) und Media & Human-Centered Computing (MSc), beide TU Wien

Dissertation zu Perceived Security

Sicherheit, Safety und Privatsphäre, wie Menschen sie erleben

Wie sicher sich Menschen bei Technologien fühlen, und warum das von der tatsächlichen Sicherheit abweicht. Gefördert durch ein ÖAW DOC Fellowship.

Lehre in HCI, Design und Security

Design Thinking, Artifact-based Design, Usable Security, Quantified Self: als Lehrender oder in Gastvorträgen an TU Wien, FH Wiener Neustadt (HWN), TU Graz, Montanuniversität Leoben und PLUS Salzburg. Seit 2022 bei ESSE.

Mixed Methods, und Prototypen gemeinsam mit Studierenden

Online-Studien mit Vignetten, Workshops, Interviews. Prototypen aus Abschlussarbeiten, etwa ein einheitliches Consent-Fenster im Browser.

Ablauf

00 **Einstieg**
2 Entscheidungen aus dem Alltag und 4 Felder

01 **Gefühl und Wirklichkeit**
Projekt, Begriffe, Einflussgrößen

02 **Eigene Studien**
Cookie-Banner, Consent-Fenster, B2B-Plattform, Fabrik

03 **Generative KI**
Täuschung, Einwilligung, 3 Ansatzpunkte

04 **Diskussion**
3 Fragen zum Einstieg

2 Entscheidungen aus dem Alltag

Beide treffen wir nach Gefühl.



Der Passwortmanager

„Alle Passwörter an einem Ort? Das ist doch riskant.“

Fühlt sich unsicher an. Ist in der Regel sicherer als dasselbe Passwort auf 20 Seiten.

Gefühl: ungeschützt · **Wirklichkeit: geschützt**



Die offizielle E-Mail

Logo, korrekte Anrede, kein Tippfehler, 1 Link zur „Bestätigung“.

Fühlt sich sicher an. Ist Phishing.

Gefühl: geschützt · **Wirklichkeit: ungeschützt**

In beiden Fällen entscheidet das Gefühl. Einmal zu vorsichtig, einmal zu vertrauensvoll.

Geschützt fühlen und geschützt sein

2 Fragen, 4 Felder.

	IST NICHT GESCHÜTZT	IST GESCHÜTZT
FÜHLT SICH GESCHÜTZT	**Illusion von Schutz** Die offizielle E-Mail. Der Inkognito-Modus. Führt zu riskantem Verhalten.	**Verdientes Vertrauen** Die Freigabe im Online-Banking per App. Das Update im Hintergrund.
FÜHLT SICH NICHT GESCHÜTZT	**Berechtigte Sorge** Dasselbe Passwort überall.	**Unbegründete Sorge** Der Passwortmanager. Verhindert, dass Schutz genutzt wird.

Oben links und unten rechts sind die Forschungsfragen. Aufklärung allein reicht dort kaum.

Kapitel 1

Gefühl und Wirklichkeit

Projekt, Begriffe, Einflussgrößen

Das Projekt

Seit 2025 gefördert durch ein DOC-Stipendium der ÖAW: „Perception vs. Reality: The Illusion of Security“.

Die Frage

Warum weichen wahrgenommene und tatsächliche Sicherheit voneinander ab, gerade bei Technologien, die KI nutzen?

Anders als Usable Security

Usable Security: Wie wird Technik sicher und benutzbar? Perceived Security: Fühlt sie sich auch so an, wenn sie es ist?

Das Ziel

Menschen nutzen sichere Technik, statt sie zu meiden, und erkennen unsichere, statt ihr zu vertrauen.

Perceived Security: In meiner Arbeit und in diesem Vortrag zähle ich Security, Safety und Privatsphäre dazu. Das ist meine Lesart, keine allgemein anerkannte Definition.

Projekt „Perception vs. Reality: The Illusion of Security“, DOC-Stipendium der Österreichischen Akademie der Wissenschaften (DOC/27292).

Security, Safety, Privatsphäre

Technisch 3 verschiedene Dinge. Im Alltag 1 Gefühl.

Security

Schutz vor absichtlichem Schaden: der gehackte Server, gestohlene Zugangsdaten. Messbar: Tests, Audits, Zertifikate.

Safety

Schutz vor unabsichtlichem Schaden: der Server, der nach einem Kurzschluss ausfällt; der Unfall an der Maschine. Messbar: Normen, Vorfälle, Risikoanalysen.

Privatsphäre

Kontrolle darüber, wer was über mich erfährt und was damit geschieht. Messbar: Datenflüsse, Rechtsgrundlagen, Einstellungen.

IM ALLTAG

„Ich fühle mich nicht sicher“ kann alles davon bedeuten.

Perception of Security: wie sicher sich eine Person fühlt, auch wenn technisch Safety oder Privatsphäre betroffen ist.

Abgrenzung von Security und Safety nach Jore (2019). Privatsphäre und Security als verschiedene, oft verwechselte Konstrukte: Bansal (2011).

Einflussgrößen

Vom KI-Assistenten bis zum Cookie-Banner.



Die Technologie

Vertraute Muster, sichtbarer Aufwand, Fehlermeldungen, ein Schloss-Symbol. Signale, nicht Eigenschaften.



Die Person

Wissen, Erfahrung, Technikaffinität, Alter. Wissen wirkt schwächer als erwartet.



Der Kontext

Was auf dem Spiel steht. Firmengerät oder privat, Zeitdruck, Vertrauen in den Anbieter.

Das Gefühl stützt sich auf sichtbare Signale. Die technische Sicherheit sieht man nicht.

Kapitel 2

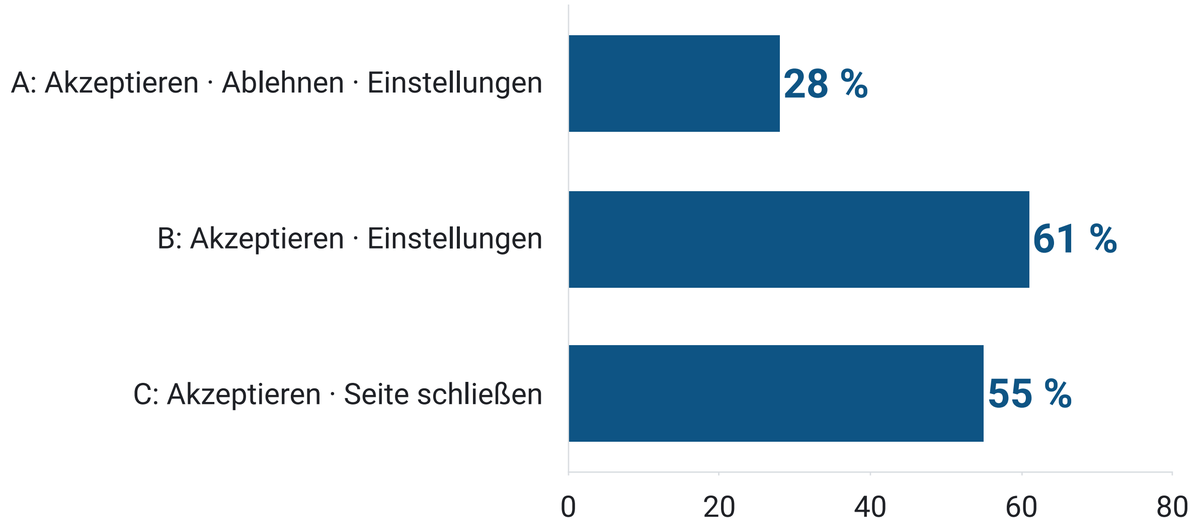
Eigene Studien

Privatsphäre, Security, Safety

Cookie-Banner

Online-Vignettenstudie mit 4 Website-Mock-ups. 136 Teilnehmende, 544 Entscheidungen (CHIRA 2025).

Anteil „Alle akzeptieren“ je Optionsset



47,8 %

aller Entscheidungen fielen auf „Alle akzeptieren“.

p < 0,001

Effekt des Optionssets (χ^2 = 44,1).

p = 0,22

für den Website-Typ. Cookie-Wissen und Alter: kein messbarer Einfluss.

Wissen schützt nicht, Design schon. Wo Ablehnen 1 Klick kostet, wird abgelehnt.

Vrecar, R., Geld, B., Dobrosovestnova, A., Čolakovová, M. G., & Weiss, A. (2026). Do cookies taste different? CHIRA 2025, CCIS 2834, 56–77. Springer. · Statische Mock-ups, Stichprobe nicht repräsentativ: die Ergebnisse zeigen eine Richtung.

Ein Fenster statt 1.000 Banner

Prototyp aus der Bachelorarbeit von Andreas Schlößl (TU Wien), betreut von Astrid Weiss und mir.

Die Idee

Der Browser stellt 1 Fenster bereit, das überall gleich aussieht. Präferenzen werden 1-mal gesetzt und gelten auf jeder Seite.

72

von 100 Punkten auf der System Usability Scale.

80 %

bevorzugten das Fenster gegenüber dem klassischen Banner.

10

Teilnehmer, 18 bis 30 Jahre, alle männlich, die Hälfte aus der Informatik. Nicht repräsentativ.

Offen: Verspricht das Fenster mehr Schutz, als es bietet?

Vrecar, R. (2026). Why AI shouldn’t decide what we share with AI and how we can do better. Workshop at CHI ’26. · Vrecar, R. (2026). Consent cannot be delegated. CTSi White Paper No. 01.

Vertrauen in eine B2B-Plattform

Workshops mit 3 Stakeholder-Gruppen und Expertenreview einer Plattform für Elektronikfertigung (CHIRA 2024).

Vertrauen heißt hier: Institution

Verträge, NDAs, Ruf. Ein Leak schadet, bevor ein Gericht entscheidet.

Transparenz und Kontrolle

Was wird gespeichert, wie lange? Jeden Schritt zurücknehmen können. Erklärbarkeit der KI-Auswahl war weniger wichtig.

Kleinigkeiten kippen das Gefühl

Fehlermeldung „Pflichtfeld“ vor der Eingabe. Kein Rückgängig beim Löschen. Uneinheitliche Meldungen. Keine Sicherheitslücke, aber weniger Vertrauen.

Nebenbefund: Auch diese Fachleute lehnen Cookies nur ab, wenn es leicht geht.

Vrecar, R., Weiss, A., Lepuschitz, W., Wedral, A., & Čolakovová, M. G. (2025). Why do(n’t) you trust us? CHIRA 2024, CCIS 2370, 177–192. Springer.

Safety in der Fabrik

In-situ-Projektionen gegen Gefahren: 3 Anwendungsfälle, getestet mit 15 Beschäftigten aus der Zielgruppe (INTERACT 2023).

Montagehilfe

Zeigt den nächsten Arbeitsschritt direkt am Werkstück. Bessere Leistung, geringe kognitive Belastung.

Ergonomisches Heben

Erinnert an die richtige Hebetechnik. Regel nicht öfter befolgt.

Sperrbereiche

Markiert den gefährlichen Bereich. Die meisten ignorierten die Regel nach eigener Einschätzung.

Menschen bewerten die Gefahr selbst. Wo Gefühl und Regel voneinander abweichen, gewinnt das Gefühl.

Wedral, A., Vrecar, R., Ebenhofer, G., Pönitz, T., Wührer, P. H., Weiss, A., & Stübl, G. (2023). Spatial augmented reality in the factory: Can in-situ projections be used to communicate dangers and health risks? INTERACT 2023, 574–594.

Kapitel 3

Generative KI

Täuschung und Einwilligung

Täuschung wird billig

Vertrauenssignale lassen sich in beliebiger Menge herstellen.

SIGNAL	BISHER	MIT GENERATIVER KI
Fehlerfreie Sprache, in meiner Sprache	Mühsam, oft auffällig	Jedes Modell, jede Sprache, in Sekunden
Persönlicher Bezug: Projekt, Kollegin, Termin	Recherche von Hand	Automatisch aus öffentlichen Quellen
Vertraute Stimme oder vertrautes Gesicht	Kaum möglich	Wenige Sekunden Audio reichen
Seriöses Design, Logo, Domain	Vorlage kopieren	Website in Minuten, für jedes Ziel neu

54 %

klickten auf vollautomatisch erstellte Spear-Phishing-Mails. Gleich viele wie bei menschlichen Experten. Kontrollgruppe: 12 %.

50 ×

profitabler werden große Kampagnen durch die Automatisierung.

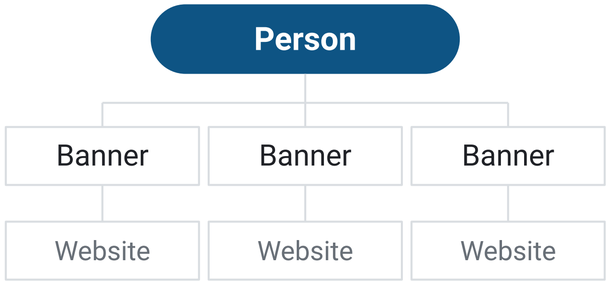
Die Oberfläche ist jetzt billig. Was bleibt: ein zweiter Kanal, ein bekannter Standard, ein System, das zeigt, was es tut.

Heiding, F., Lermen, S., Kao, A., Mayrink Verdun, C., Schneier, B., & Vishwanath, A. (2026). Evaluating large language models' ability to automate spear phishing. Expert Systems with Applications, 314, 131546. N = 101.

KI soll entscheiden, was wir teilen

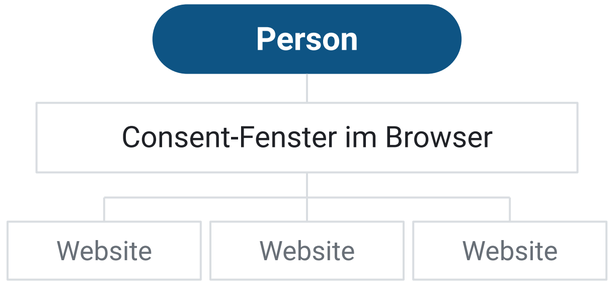
Browser-Agenten klicken Banner weg, Einstellungen werden „optimiert“.

BANNER PRO WEBSITE



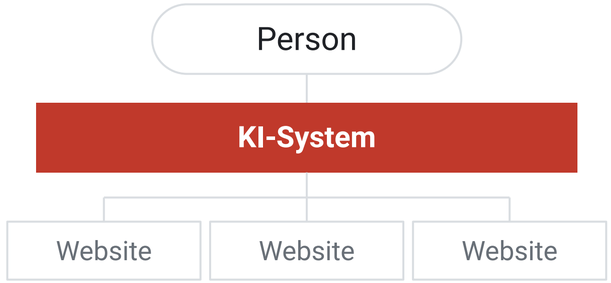
Optionen je Seite anders, Ablehnen oft nicht 1 Klick

1 FENSTER IM BROWSER



Person setzt Präferenzen 1-mal, überall gleich

KI ENTSCHEIDET



Die Entscheidung kommt nicht von der Person

Art. 4 Nr. 11 DSGVO: Einwilligung ist eine „freiwillig für den bestimmten Fall, in informierter Weise und unmissverständlich abgegebene Willensbekundung“ der betroffenen Person. Erwägungsgrund 32: auch durch die Auswahl technischer Einstellungen. Eigene Voreinstellungen: ja. Eine KI, die stellvertretend wählt: nein.

Einwilligung lässt sich nicht delegieren. KI darf erklären, entscheiden muss die Person.

Vrecar, R. (2026). Consent cannot be delegated. CTSi White Paper No. 01. · Verordnung (EU) 2016/679 (DSGVO). Abbildung nach dem White Paper.

3 Ansatzpunkte

Nicht Oberflächen, nicht Wissen allein.

Standards, die man kennt

Gleiche Optionen an gleicher Stelle, auf jeder Website. 1 Fenster statt 1.000 Banner.

Systeme, die zeigen, was sie tun

Was wird gespeichert, wie lange, wer sieht es. Jeden Schritt zurücknehmen können. Fehlermeldungen, die stimmen.

Entscheidung bei der Person

KI erklärt Begriffe und Richtlinien, trifft aber keine Wahl.

Wissen allein reicht nicht. Und Design darf nicht mehr Schutz versprechen, als dahintersteht.

Zusammenfassung

5 Punkte

01 **Geschützt fühlen ist nicht geschützt sein**

Beide Richtungen schaden: Die Illusion führt zu Risiko, die unbegründete Sorge verhindert Schutz.

02 **Das Gefühl folgt Signalen, nicht Eigenschaften**

Technologie, Person und Kontext bestimmen es.

03 **Design schlägt Wissen**

Nur die Optionen im Banner hatten einen Effekt.

04 **Details prägen das Vertrauen**

Fehlermeldungen, Rückgängig, Transparenz.

05 **Generative KI verschiebt die Frage**

Täuschung wird billig. KI darf erklären, nicht entscheiden.

Offene Diskussion

3 Fragen zum Einstieg.

1. Darf sich ein System sicherer anfühlen, als es ist, wenn das die Nutzung einer sicheren Technologie fördert?
2. Welche Signale für Schutz wären im Alltag verlässlich, und für alle lesbar, nicht nur für Fachleute?
3. Wo soll KI beim Schutz helfen dürfen, wo nicht? Und wer trägt die Verantwortung, wenn sie falsch liegt?

Oder Ihre Frage.

Danke.

Fragen: rafael.vrecar@tuwien.ac.at · vrecar.eu/rafael

CTSi//circle.responsibleComputing · Science Insight · 9. November 2026 · Sich geschützt fühlen, geschützt sein

Quellen

Bansal, G. (2011). Understanding the security in privacy-security concerns: A theoretical and empirical examination. AMCIS 2011.

Heiding, F., Lermen, S., Kao, A., Mayrink Verdun, C., Schneier, B., & Vishwanath, A. (2026). Evaluating large language models' ability to automate spear phishing. Expert Systems with Applications, 314, 131546.

Jore, S. H. (2019). The conceptual and scientific demarcation of security in contrast to safety. European Journal for Security Research, 4(1), 157–174.

Verordnung (EU) 2016/679 (Datenschutz-Grundverordnung). Amtsblatt der Europäischen Union, L 119, 1–88. Art. 4 Nr. 11 und Erwägungsgrund 32.

Vrecar, R. (2026). Consent cannot be delegated: Findings on cookie banners, the GDPR and the role of AI in consent. CTSi White Paper No. 01. TU Wien.

Vrecar, R. (2026). Why AI shouldn't decide what we share with AI and how we can do better. Position paper, workshop at CHI '26, Barcelona. hdl.handle.net/20.500.12708/228428

Vrecar, R., Geld, B., Dobrosovestnova, A., Čolakovová, M. G., & Weiss, A. (2026). Do cookies taste different? The impact of options and website type when setting cookie preferences. CHIRA 2025, CCIS 2834, 56–77. Springer.

Vrecar, R., Weiss, A., Lepuschitz, W., Wedral, A., & Čolakovová, M. G. (2025). Why do(n't) you trust us? Highlighting the importance of trust and transparency for designing B2B platforms in electronics manufacturing. CHIRA 2024, CCIS 2370, 177–192. Springer.

Wedral, A., Vrecar, R., Ebenhofer, G., Pönitz, T., Wührer, P. H., Weiss, A., & Stübl, G. (2023). Spatial augmented reality in the factory: Can in-situ projections be used to communicate dangers and health risks? INTERACT 2023, 574–594. Springer.

Wu, Y., Gupta, P., Wei, M., Acar, Y., Fahl, S., & Ur, B. (2018). Your secrets are safe: How browsers' explanations impact misconceptions about private browsing mode. WWW '18.

Der Prototyp des Consent-Fensters stammt aus der Bachelorarbeit von Andreas Schlößl (TU Wien). Icons: Lucide (ISC-Lizenz). Porträt © Matthias Heisler.