JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

FB02

Professur für Digitalisierung, E-Business
und Operations Management

Informationen zum Proseminar WiSe 2022/2023

Virtual Reality (VR) Privacy & Trustworthy AI

an der Professur für Digitalisierung, E-Business und Operations Management
(Inh. Prof. Dr. Jella Pfeiffer)

Allgemeine Hinweise:

- *Den Ablauf des Seminars entnehmen Sie bitte den Veranstaltungen auf StudIP.*
- *Die Seminararbeit kann auf Deutsch oder Englisch verfasst werden.*
- *Aufgrund der großen Anzahl an zugeordneten Teilnehmer*innen erweitern wir den inhaltlichen Fokus des Seminars. Der Inhalt des Seminars ist nun zweigeteilt:*

Erster Themenblock: Virtual Reality (VR) Privacy

Virtual Reality hat das Potenzial, die Art und Weise zu verändern, wie wir leben, miteinander umgehen und zusammenarbeiten. Das liegt an der einzigartigen Fähigkeit der VR: Dem Eintauchen des Nutzers in eine virtuelle, aber sichere Welt. Die Einsatzmöglichkeiten von Virtual Reality sind vielfältig und nehmen ständig zu. Hierzu gehören Bereiche wie Gaming (z.B. Beats Saber), Profisport (z.B. Trainingsprogramme für verschiedene Sportarten), virtuelle Reisen (z.B. National Geographic Explore VR), Gesundheitswesen (z.B. Programme zur Behandlung von PTSD), Film & Fernsehen (z.B. 360° Videos).

Virtual Reality ist in der Lage, Nutzerdaten in einem Umfang und Ausmaß zu sammeln, wie es bei anderen Verbrauchergeräten (z.B. Smartphones und -watches) nicht der Fall ist. Darüber hinaus sind viele dieser Daten, einschließlich potenziell sensibler biometrischer Inforamtionen, für die Kernfunktionen von VR-Geräten erforderlich. Da VR in erheblichem Maße auf biometrische Daten angewiesen ist, ergeben sich einzigartige Herausforderungen in Bezug auf den Schutz der Privatsphäre und die Wahlmöglichkeiten der Nutzer bei der Erhebung und Verarbeitung personenbezogener Daten. Schließlich sind die derzeitigen Gesetze und Vorschriften in Europa nicht in der Lage, die mit dieser multimodalen Datenerfassung verbundenen Risiken für den Datenschutz zu bewältigen. Dies wirft die Frage auf, wie die Datenerhebung, die für die virtuelle Realität unerlässlich ist, geregelt werden kann. Die Herausforderung besteht darin, die Datenerhebung zu regeln und die Privatsphäre ausreichend zu schützen, ohne wesentliche VR-Funktionen einzuschränken und VR-Innovationen zu behindern.

In diesem Seminar wird auf die einzigartigen und schwerwiegenden Herausforderungen für den Datenschutz eingegangen, die die immersive Technologie mit sich bringt. Dabei werden die zwei großen Themen „Datensammlung und -auswertung“ und „Privatsphäre(schutz) und VR“ behandelt, die mögliche Fragestellungen für die eigene Proseminararbeit umfassen:

- **Datensammlung und -Auswertung:**

    - **Thema 1:** Welche wissenschaftliche Literatur hat (biometrische) Daten in VR erhoben und was genau wurde erhoben und untersucht? Welche Schwierigkeiten gibt es bei der Datenerhebung in VR?

        - Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the association for information systems*, *37*(1), 9.

    - **Thema 2:** Aus welchen Bestandteilen besteht ein VR-System? Wie und welche Arten von (z.B. biometrischen) Daten können und werden erhoben? Wie können diese Daten hinsichtlich verschiedener Dimensionen kategorisiert werden?

        - Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the association for information systems*, *37*(1), 9.

- **Privatsphäre(schutz) und VR:**

    - **Thema 3:** Welche Risiken birgt der Einsatz von Virtual-Reality-Technologien für die Nutzerprivatsphäre, und wie stehen sie im Vergleich zu dem potenziellen Nutzen dieser Technologie für Unternehmen und für die eigene Person? Und was bedeutet in diesem Zusammenhang der Begriff Privacy Paradox (engl. Privatsphäre-Paradoxon)?

        - Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the association for information systems*, *37*(1), 9.

    - **Thema 4:** Strukturierter Literaturüberblick über die De-Anonymisierung(smöglichkeiten) der gesammelten (biometrischen) Daten während der Benutzung von VR-Technik

        - Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, 26(2).

- **Thema 5:** Welche Herausforderungen stellt die immersive VR-Technologie für die Regulierungsbehörden dar? Ist die Datenschutz-Grundverordnung in ihrer derzeitigen Form in der Lage, die vorliegende Technologie zu regulieren, ohne ihre Innovation zu behindern?

    - Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the association for information systems*, *37*(1), 9.

- **Thema 6:** Strukturierter Literaturüberblick zum Thema Privacy in XR (Extended Reality oder eher Cross Reality)

    - Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, 26(2).

Gerne können Sie eigene Themenvorschläge einbringen.

Vorbereitende Literatur:

Dem BWL XII-Team ist bewusst, dass dieser Themenbereich wahrscheinlich ein bisher unbekanntes Gebiet für Studierende des Fachbereichs 02 darstellt. Um Sie auf die Inhalte vorzubereiten, empfehlen wir Ihnen folgende Literatur (bei Zugangsproblemen bitte melden!):

- Rauschnabel, P. A., Felix, R., Hinsch, C., Shahab, H., & Alt, F. (2022b). What is XR? Towards a framework for Augmented and Virtual Reality. Computers in Human Behavior, in press. https://doi.org/10.1016/j.chb.2022.107289
- Herz, M., & Rauschnabel, P. A. (2019). Understanding the diffusion of virtual reality glasses: The role of media, fashion and technology. Technological Forecasting and Social Change, 138, 228-242.
- Slater, M., Gonzalez-Liencres, C., Haggard, P., Vinkers, C., Gregory-Clarke, R., Jelley, S., ... & Silver, J. (2020). The ethics of realism in virtual and augmented reality. Frontiers in Virtual Reality, 1, 1.
- Giaretta, A. (2022). Security and Privacy in Virtual Reality - A Literature Survey. arXiv preprint arXiv:2205.00208.
- Pfeiffer, J., Pfeiffer, T., Meißner, M., & Weiß, E. (2020). Eye-tracking-based classification of information search behavior using machine learning: evidence from experiments in physical shops and virtual reality shopping environments. Information Systems Research.

- Chuah, S. H. W. (2018). Why and who will adopt extended reality technology? Literature review, synthesis, and future research agenda. Literature Review, Synthesis, and Future Research Agenda (December 13, 2018).
- Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. Journal of consumer affairs, 41(1), 100-126.
- Barth, S., & De Jong, M. D. (2017). The privacy paradox–Investigating discrepancies between expressed privacy concerns and actual online behavior–A systematic literature review. *Telematics and informatics*, *34*(7), 1038-1058.
- Kröger, J. L., Lutz, O. H. M., & Müller, F. (2019, August). What does your gaze reveal about you? On the privacy implications of eye tracking. In IFIP International Summer School on Privacy and Identity Management (pp. 226-241). Springer, Cham.

Zweiter Themenblock: Trustworthy AI

Im Jahr 2021 gab es einen Entwurf für eine Verordnung zur Regulierung der Nutzung Künstlicher Intelligenz der EU, der Artificial Intelligence Act. Es geht um Vorschriften für die Entwicklung, das Inverkehrbringen und die Nutzung von KI-Systemen. Er wird koordiniert durch den „Europäischen Ausschuss für Künstliche Intelligenz“. Dieser Vorschlag zielt darauf ab, einen Rechtsrahmen für eine vertrauenswürdige KI zu schaffen. Es werden sieben Anforderungen an die Vertrauenswürdigkeit eines KI-Systems gestellt: Human agency and oversight, technical robustness and safety, Privacy and data governance, Transparency, Diversity, non-discrimination and fairness, Societal and environmental wellbeing, Accountability.

Vorschriften zu Produktsicherheit und -haftung sind in diesem Gesetzentwurf nicht abgebildet. Diese wurden nun gesondert formuliert: https://ec.europa.eu/commission/presscorner/detail/de/ip_22_5807

In dem Themenblock beschäftigen wir uns mit unterschiedlichen Facetten von vertrauenswürdiger KI.

Thema 1: Explainable Artificial Intelligence (AI) for Transparency

Der Begriff „explainable AI" ist eine neue Bezeichnung für ein sehr altes Bestreben in der Wissenschaft, Antworten auf die Frage nach dem Warum zu geben. Ziel ist es, dass menschliche Experten verstehen, warum ein KI-System die Entscheidung so gefällt hat. Dies ist von großer Bedeutung für das kausale Verständnis und ermöglicht somit eine ethisch verantwortungsvolle KI und transparentes, überprüfbares maschinelles Lernen zur Entscheidungsunterstützung.

In Holzinger et al. (2022) werden mehrere Methoden der explainable AI kurz erläutert. Sie sollen dieses Paper und weitere Paper zu Explainable AI lesen und die verschiedenen Ansätze verstehen und in der Arbeit zusammenfassend darstellen. Neben der strukturierten Darstellung der Ansätze ist das Ziel der Arbeit, drei für Unternehmen relevante Probleme herauszusuchen bei denen typischerweise KI-Methoden eingesetzt werden (z.B. Credit Scoring, Maintenance Prediction, Identify good candidates for job, etc.). Hierzu könnten Sie auch Paper suchen, die den spezifischen Einsatz eines KI-Verfahren für den Fall beschreibt, also z.B: ein Paper, welches einen spezifischen Algorithmus zum Credit Scoring vorschlägt (ohne dabei explainable AI anzuwenden). Sie sollen dann bewerten, welche der Methoden für den jeweiligen Fall am geeignetsten sein könnte.

- Holzinger, A., Saranti, A., Molnar, C., Biecek, P., & Samek, W. (2022). Explainable AI methods-a brief overview. In *International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers* (pp. 13-38). Springer, Cham.

Thema 2: Neueste Ansätze im Bereich Explainable AI

Dieses Thema ist mit Thema 1 verbunden. Sie sollen 2-3 der allerneuesten Ansätze im Bereich Explainable AI detailliert verstehen und erläutern. Im besten Fall verwenden Sie einen der Ansätze beispelhaft an einem Datensatz, sollten die entsprechenden Bibliotheken zur Programmierung von den Autoren vorhanden sein (was meist der Fall ist). Startpunkt ist hier wieder der Übersichtsartikel von Holzinger et al. (2022)

Thema 3: Strukturierter Literaturüberblick zu Explainable AI in Information Systems

Entsprechend dem Ansatz zu strukturierten Literaturüberblick von Webster and Watson (2002) sollen Sie einen Überblick über das Feld Explainable AI in Information Systems erstellen. Sie konzentrieren sich dabei auf Veröffentlichungen im Bereich Information Systems und nehmen hier den Basket of Eight des AIS als Grundlage: https://aisnet.org/page/SeniorScholarBasket und zudem die beiden Konferenzen: International Conference on Information Systems und European Conference on Information Systems. Zugang zu den Zeitschriften können Sie teilweise über die Professur erhalten, sollten Sie Probleme haben. Ein Beispiel für solch einen strukturierten Literaturüberblick finden Sie hier: Kordzadeh und Ghasemaghaei (2022).

- Kordzadeh, N., & Ghasemaghaei, M. (2022). Algorithmic bias: review, synthesis, and future research directions. *European Journal of Information Systems*, *31*(3), 388-409.
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, 26(2).

Thema 4: Antecedents of Fairness Perceptions

In einem Literaturüberblick zu Fairness (in dem Fall algorithmic bias) listen Kordzadeh und Ghasemaghaei (2022) einige Arbeiten auf, die sich mit den Einflussfaktoren auf die empfundene Fairness befassen (siehe Abschlitt zu antecedents of fairness perceptions). Dies können sein: der technische Prozess im Algorithmus, verschiedene Erklärungsstile des Algorithmus (siehe explainable AI weiter oben), individuelle Charakteristiken der User oder auch die Aufgabe, die mit Hilfe des KI-Systems bewältigt wird. Sie sollen sich in solche Einflussfaktoren einlesen und einige Hypothesen erarbeiten, die aufzeigen, wie und warum welche Einflussfaktoren eine Auswirkung auf das Empfinden von Fairness haben könnten. Startpunkt sind die in besagtem Abschnitt genannten Quellen, wie Dodge et al. 2019 und andere.

- Dodge, J., Liao, Q. V., Zhang, Y., Bellamy, R. K., & Dugan, C. (2019). Explaining models: An empirical study of how explanations impact fairness judgment. *Proceedings of the 24th international conference on intelligent user interfaces*. Marina del Ray, CA, USA.

Thema 5: Strukturierter Literaturüberblick zu empirischen Arbeiten im Bereich Algorithmic Fairness

Entsprechend dem Ansatz eines narrativen Literaturüberblicks von vom Brocket et al. (2015) sollen Sie einen Überblick über alle empirischen Arbeiten im Bereich Algorithmic Fairness, Algorithmic Bias geben. Startpunkt ist die Arbeit von Kordzadeh und Ghasemaghaei (2022). Wichtig ist, dass Sie sich nur auf Arbeiten beziehen, die das Thema Fairness von AI-Algorithmen empirisch untersucht haben. Sie beschränken sich hier nicht nur auf den Bereich Information Systems, sondern betrachten auch human-computer-interaction, Psychologie und andere angrenzende Bereiche, da es nicht allzu viele empirische Papiere zu dem Thema geben wird.

- Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the association for information systems*, *37*(1), 9.