Sachbericht Teil II – Darstellung der Arbeitspakete und durchgeführte Arbeiten

Verbundname: Dynamiken der Desinformation erkennen und bekämpfen

Akronym: DYNAMO

Teilvorhaben: Teilprojekt Medienpsychologie

AP 1 Exploration
AP 1.2.1 Befragung
Ziel des Arbeitspakets war es, einen ersten empirischen Überblick darüber zu gewinnen, wie deutsche Internetnutzer*innen Nachrichten in Messengerdiensten wahrnehmen, wenn der Verdacht einer Desinformation besteht und welche Faktoren ihr Weiterleitungsverhalten beeinflussen. Anders als in der bisherigen Forschung, die primär soziale Medien wie Twitter oder Facebook untersucht hat, lag der Fokus gezielt auf Messengerdiensten. Geplant war eine bevölkerungsrepräsentative Online-Befragung mit rund 1.000 Teilnehmenden, um zu erfassen, welche Inhalte weitergeleitet werden, welche Motive dahinterstehen und welche Rolle der Verdacht auf Falschmeldungen spielt. Zusätzlich sollten Persönlichkeitsmerkmale wie Need for Cognition sowie Einstellungen zur COVID-19-Pandemie erhoben werden.

Zu Beginn wurde eine umfassende Literaturrecherche durchgeführt und ein theoretisches Fundament erarbeitet. Parallel dazu wurden Projektstrukturen aufgebaut, Hilfskräfte eingestellt und technische Ressourcen geschaffen. Für die Datenerhebung wurde ein professionelles Online-Panel eingekauft. Die repräsentative Befragung (N = 1.004) erfasste, welche Arten von Falschnachrichten über Messenger empfangen wurden und welche psychologischen Mechanismen das Weiterleiten beeinflussen. Der Schwerpunkt lag auf Emotionen: Der Erhalt von Falschnachrichten erhöhte die Wahrscheinlichkeit negativer Emotionen. Aversion (z. B. Ekel) begünstigte das Weiterleiten, um Gefühlen Ausdruck zu verleihen, während Angst vor allem das Einholen zusätzlicher Informationen motivierte. Die Ergebnisse legen nahe, dass unterschiedliche Typen von Falschnachrichten spezifische Interventionsstrategien erfordern. Für dieses Arbeitspaket wurden insbesondere Mittel für den die Studie durchführenden Mitarbeiter, die Hilfskräfte, Panelkosten und technische Infrastruktur eingesetzt. Diese waren notwendig, um eine qualitativ hochwertige und repräsentative Datenerhebung sicherzustellen. Die Ergebnisse sind sowohl wissenschaftlich als auch praktisch nutzbar, da sie ein tieferes Verständnis der emotionalen Dynamiken hinter dem Teilen von Falschnachrichten liefern und konkrete Ansatzpunkte für Gegenmaßnahmen eröffnen. Die Ergebnisse wurden in einem Manuskript für die Veröffentlichung in internationalen Journals aufbereitet und bei einer Online-Präkonferenz der 9. European Communication Conference (ECREA) und im Rahmen eines interdisziplinären Workshops des Graduiertenkollegs Privacy & Trust der TU Darmstadt präsentiert.

AP 2 Erfassen und Erkennen
AP 2.2.1 Bedeutung der Herkunfts- und Weiterleitungsquellen
Ziel dieses Arbeitspakets war es, den Einfluss der Herkunfts- und Weiterleitungsquelle einer Falschnachricht auf deren Glaubwürdigkeit sowie auf Weiterleitungs- und Verifizierungstendenzen zu untersuchen. Während in Messengerdiensten Nachrichten häufig nicht von journalistischen Quellen, sondern von Privatpersonen aus dem Freundes- oder Bekanntenkreis stammen, sollte in einem 2x2-between-subjects-Laborexperiment (N = 200) geprüft werden, wie sich die Urheberquelle (journalistisch vs. privat) in Kombination mit der sendenden Person auf die Rezeption und das Weiterleitungsverhalten auswirkt.

Im Berichtszeitraum wurde das Arbeitspaket konzeptionell vorbereitet und gemeinsam mit AP 2.2.2 detailliert geplant. Die Durchführung verzögerte sich jedoch aufgrund des aufwendigen experimentellen Designs, eines Hackerangriffs auf die Universität Duisburg-Essen Ende 2022, der die IT-Infrastruktur einschränkte, sowie eines Personalwechsels im Sommer 2023. Das Laborexperiment startete Ende Oktober 2023 und umfasste ein 2 (Quelle) x 2 (Bewertung) between-subjects-Design mit insgesamt N = 228 Teilnehmenden. Neben Quelle und Bewertung der Falschnachricht (keine Bewertung vs. negative Bewertung) wurde auch die Beziehungsstärke zwischen Sender und Empfänger erfasst. Die Teilnehmenden wurden paarweise eingeladen, bearbeiteten Kommunikationsaufgaben über den Messenger-Dienst Discord und erhielten dabei einen manipulierten Screenshot einer Falschnachricht, der vermeintlich von der jeweils anderen Person stammte. Erfasst wurden Weiterleitungsintention, Verifizierungstendenz und Verarbeitungstiefe. Die Datenerhebung gestaltete sich aufgrund der Laborumgebung, der gleichzeitigen Teilnahme zweier Personen sowie erschwerter Rekrutierung von männlichen Teilnehmenden sehr aufwändig. Sie konnte erst im März 2024 abgeschlossen werden. Für dieses Arbeitspaket wurden Mittel für die die Studie durchführenden Mitarbeiterin, zwei studentische Hilfskräfte und die Incentivierung der Teilnehmenden eingesetzt. Die Ergebnisse des Laborexperiments zeigten, dass weder Quelle noch Senderbewertung direkte oder indirekte Effekte auf die Verarbeitungstiefe, Weiterleitungs- oder Verifizierungsintention hatten. Stattdessen erwies sich die Verarbeitungstiefe als zentraler Prädiktor: Je intensiver sich Teilnehmende mit der Nachricht auseinandersetzten, desto höher waren sowohl ihre Bereitschaft, die Information weiterzuleiten, als auch ihre Motivation, sie zu überprüfen. Zudem zeigte sich, dass der Erhalt der Falschnachricht mit negativen Emotionen verbunden war, die die Weiterleitungsintention wiederum signifikant beeinflussen. Erste Ergebnisse wurden im September 2024 auf dem Kongress der Deutschen Gesellschaft für Psychologie (DGPs) online präsentiert. Finale Ergebnisse wurden in einem wissenschaftlichen Manuskript niedergeschrieben und befinden sich im Veröffentlichungsprozess.

AP 2.2.2 Faktoren zur Weiterleitung
Ziel dieses Arbeitspakets war es, aufbauend auf den Ergebnissen aus AP 2.2.1 weitere Faktoren zu identifizieren, die die Weiterleitung von Falschnachrichten begünstigen. Geplant waren zwei experimentelle Online-Studien: Eine sollte den Einfluss von Voreinstellungen sowie der Fähigkeit und Motivation zur Auseinandersetzung mit Informationen untersuchen, die zweite den Einfluss von Merkmalsausprägungen der Nachricht, insbesondere ihrer Emotionalität.

Aufgrund eines Personalwechsels im Projektablauf wurden die Arbeiten angepasst und eine der geplanten Online-Studien aus AP 2.2.2 vorgezogen. In der ersten Studie (N = 770) wurde untersucht, ob ein Zusammenhang zwischen der Weiterleitungsintention, der Nachrichtenquelle und der vorbestehenden Meinung besteht. Hierfür wurde ein 2 (Quelle: KI-basierter Chatbot vs. journalistisch tätige Person) x 2 (Meinung: Pro-Argumente vs. Kontra-Argumente) Design umgesetzt. Die Teilnehmenden erhielten eine Messenger-Nachricht zum polarisierenden Thema Gender-Sprache, die entweder von einem Chatbot oder einer journalistischen Person stammte. Der Inhalt der Nachricht enthielt entweder Argumente für oder gegen die Nutzung geschlechtergerechter Sprache. Die Analyse zeigte keinen signifikanten Zusammenhang zwischen der Weiterleitungsintention und den Faktoren Nachrichtenquelle oder bestehende Meinung.
Die zweite Studie wurde Ende 2023 inhaltlich vorbereitet, im Frühjahr 2024 durchgeführt und im Mai 2024 abgeschlossen. In diesem 2 (Emotionalität: neutral vs. negativ) x 3 (Motivation: akkurat handeln vs. sich verteidigen vs. keine Motivation) between-subjects-Experiment (N = 200) wurde untersucht, wie Emotion und Motivation die Weiterleitungsintention beeinflussen. Die Ergebnisse zeigen, dass die Motivation, sich zu verteidigen, die Intention zur Weiterleitung signifikant erhöht. Dieser Effekt verringerte sich jedoch, wenn die Falschinformation mit den Einstellungen der Teilnehmenden übereinstimmte. Die Emotionalität der Nachricht hatte keinen signifikanten Einfluss. Die geplante Induktion einer

genauen, akkuraten Verarbeitungsmotivation war nicht erfolgreich, sodass Hypothesen zu diesem Aspekt nicht geprüft werden konnten.

Für die Umsetzung des Arbeitspaketes wurden vor allem Personalmittel und Teilnehmendenhonorare für die Durchführung der Online-Erhebung eingesetzt. Diese waren notwendig, um die komplexen Designs umzusetzen und eine hinreichend große Stichprobe zu erreichen. Die Ergebnisse liefern wichtige Hinweise darauf, dass motivationsbezogene Faktoren wie Verteidigungsmotive entscheidend für das Teilen von Falschinformationen sein können. Die Resultate wurden in einem wissenschaftlichen Manuskript aufbereitet und befinden sich derzeit im internationalen Begutachtungsprozess.

AP 3 Bekämpfen

AP 3.2.1 Vertiefte Betrachtung spezifischer technischer Interventionen

Ziel dieses Arbeitspakets war es, auf Basis der im Konsortium entwickelten Konzepte zu technischen Interventionen deren Wirksamkeit im Hinblick auf die Bekämpfung von Desinformation zu testen. Geplant war eine experimentalpsychologische Online-Studie (N = 200) im within-subjects-Design, in der typische Falschnachrichten und wahre Nachrichten (als Kontrollvariable) mit verschiedenen technischen Interventionen versehen und deren Einfluss auf das Weiterleitungsverhalten untersucht werden sollten. Die ursprüngliche Planung sah vor, dass technische Vorarbeiten der informatischen Projektpartner als Grundlage für die Interventionen dienen. Da diese in der geplanten Form nicht vorlagen, wurden stattdessen aus der Literatur und der Theorie abgeleitete Interventionen (psychologische Inokulation/Prebunking) ausgewählt und getestet.

Parallel zu AP 2.2.2 wurde zu Jahresbeginn 2024 eine Studie aus AP 3.2.1 vorbereitet und im April 2024 abgeschlossen. In dieser Online-Studie (N = 180) wurde im between-subjects-Design die Wirksamkeit von psychologischer Inokulation (Prebunking) auf den Glauben an und das Teilen von wahren und falschen Informationen untersucht. Zusätzlich wurde geprüft, ob individuelle Eigenschaften wie kritisches Denken, Need for Cognition oder die Nutzungshäufigkeit von Messenger-Diensten die Wirkung der Intervention moderieren. Die Teilnehmenden wurden zufällig einer von zwei Experimentalgruppen oder einer Kontrollgruppe zugeteilt. In allen Gruppen wurden identische wahre und falsche Informationen im Rahmen eines simulierten Messenger-Gruppenchats präsentiert. Die Experimentalgruppen unterschieden sich hinsichtlich der Formulierung der Inokulationsnachricht: thematisch spezifisch (refutational-same) oder unspezifisch (refutational-different). Die Ergebnisse zeigen, dass Inokulationsnachrichten weder die Glaubwürdigkeit noch die Weiterleitungsintention für wahre oder falsche Informationen insgesamt reduzierten. Eine Ausnahme ergab sich für falsche Informationen zum Ukrainekrieg: Hier führte die Präsentation einer thematisch spezifischen Inokulationsnachricht zu einer niedrigeren Glaubwürdigkeit. Es konnten keine signifikanten Moderatoreffekte individueller Eigenschaften nachgewiesen werden.

Für die Umsetzung wurden vor allem Personalmittel und Teilnehmendenhonorare für die Online-Erhebung eingesetzt. Die gewonnenen Erkenntnisse sind relevant für die Bewertung von Prebunking-Maßnahmen, da sie aufzeigen, dass deren Wirksamkeit stark themen- und kontextabhängig sein kann. Die Ergebnisse wurden bei der Abschlussveranstaltung des Projekts online präsentiert und sind über YouTube öffentlich zugänglich. Zudem wurden sie in einem wissenschaftlichen Manuskript aufbereitet und befinden sich derzeit im Veröffentlichungsprozess.

AP 3.2.2 Vertiefte Betrachtung spezifischer technischer Interventionen

Ziel dieses Arbeitspakets war es, aufbauend auf den Ergebnissen aus AP 3.2.1 ausgewählte technische Interventionen in einem zweifaktoriellen Laborexperiment (N = 160) zu prüfen. Neben der Variation der technischen Interventionen sollte auch die Motivation der Teilnehmenden (z. B. accuracy motivation) experimentell beeinflusst werden, um die

psychologischen Wirkmechanismen besser nachvollziehen zu können. Zusätzlich war die Erhebung moderierender Personvariablen wie Need for Cognition und themenspezifischer Einstellungen (voraussichtlich im Kontext der Corona-Pandemie) vorgesehen.

Auf Basis der Befunde aus AP 2.2.2 und AP 3.2.1 wurde das ursprüngliche Studiendesign angepasst und die Mittelverwendung entsprechend umgewidmet. Da die bis dahin untersuchten Variablen keine eindeutige Erklärung dafür lieferten, warum Personen an falsche Informationen glauben oder diese weiterleiten, wurde beschlossen, zusätzliche kognitive Faktoren und Wirkmechanismen in den Blick zu nehmen.

Zur Jahresmitte wurde eine großangelegte Online-Studie (N = 1.000) im within-subjects-Design vorbereitet und umgesetzt. Sie bestand aus zwei Teilen: Zunächst wurden die Teilnehmenden hinsichtlich ihrer kritischen Denkfähigkeit getestet und befragt. Anschließend bearbeiteten sie ein Paradigma, das experimentell prüft, ob ihre Glaubwürdigkeitseinschätzungen und Urteilsbildung durch politische Präferenzen beeinflusst werden. Die Untersuchung zeigte, dass Teilnehmende zwar präferenzkonsistente (wahre und falsche) Informationen nicht glaubwürdiger bewerteten als präferenzinkonsistente, ihre Urteilsbildung bei den Themen Migration und Klimawandel jedoch in Richtung ihrer politischen Präferenzen verschoben. Eine höhere kritische Denkfähigkeit hing signifikant mit einer akkuraten Unterscheidung zwischen wahren und falschen Informationen zusammen, nicht jedoch mit einer weniger verzerrten Urteilsbildung. Die Ergebnisse legen nahe, dass Glaubwürdigkeitsbewertungen und Urteilsbildung auf unterschiedlichen kognitiven Prozessen beruhen. Kritische Denkfähigkeit könnte daher ein relevanter Ansatzpunkt zur Bekämpfung von Falschinformationen sein, auch wenn sie Verzerrungen in der Urteilsbildung nicht vollständig verhindert. Diese Erkenntnisse tragen dazu bei, besser zu verstehen, warum Menschen falsche Informationen teilen oder von ihnen beeinflusst werden, selbst wenn ihnen deren Unwahrheit bewusst ist. Für die Umsetzung wurden vor allem Personalmittel und Teilnehmendenhonorare für die Online-Erhebung eingesetzt. Die Ergebnisse wurden in einem wissenschaftlichen Manuskript aufbereitet, das sich derzeit im internationalen Begutachtungsprozess befindet und werden im Sommer 2025 bei zwei deutschen und einer internationalen Konferenz präsentiert.

AP 4 Evaluation

AP 4.2.1 Laborstudie zur Evaluation der Maßnahmen

Ziel dieses Arbeitspakets war es, die in AP 3 entwickelten technischen Interventionsmaßnahmen unter möglichst realistischen Bedingungen zu evaluieren. Geplant war eine groß angelegte Laborstudie (N = 100) im kombinierten between- und within-subjects-Design, in der falsche und wahre Nachrichten präsentiert werden. Ein Teil der Falschnachrichten sollte für eine Experimentalgruppe basierend auf den vorherigen Projektergebnissen als Falschinformation gekennzeichnet werden. Der Einfluss auf die Weiterleitungsintention sollte in Abhängigkeit von Faktoren wie Voreinstellung, Vorwissen und Need for Cognition untersucht werden.

Auf Basis der Ergebnisse aus AP 2.2.2 und AP 3.2.1 wurde das ursprüngliche Studiendesign jedoch angepasst. Da sich zeigte, dass die bislang untersuchten Variablen allein nicht ausreichend erklären, warum Personen falsche Informationen glauben oder weiterleiten, sollte AP 4.2.1 enger mit AP 3.2.2 verzahnt werden. Beide Arbeitspakete wurden methodisch zusammengeführt, um in einer einzigen, realistisch gestalteten Studie sowohl spezifische technische Interventionen als auch deren Wirksamkeit in verschiedenen motivationalen Kontexten zu prüfen.

Da eine Umsetzung innerhalb realer Messenger-Dienste wie WhatsApp aufgrund der Ende-zu-Ende-Verschlüsselung nicht möglich war, wurde zum Jahresende 2024 mit der Entwicklung eines Chat-Simulators begonnen. Dieses Tool soll eine realistisch wirkende Gruppenchat-Umgebung schaffen, in der Versuchspersonen aktiv teilnehmen können, indem

sie beispielsweise Nachrichten direkt in den Chat eingeben. So lassen sich in einem kontrollierten Experiment sowohl breit angelegte Prebunking-Maßnahmen als auch themenspezifische Inokulationsstrategien direkt vergleichen. Der Chat-Simulator ist so konzipiert, dass er die in AP 3.2.1 gewonnenen Erkenntnisse zur Wirkung von Inokulationsmaßnahmen aufgreift und zugleich die offenen Fragen aus AP 3.2.2 zu kognitiven und motivationalen Einflussfaktoren adressiert. Ziel ist es, die Wirksamkeit der Maßnahmen nicht nur auf die Glaubwürdigkeitseinschätzung, sondern auch auf das Weiterleitungsintention zu prüfen. Die Erhebung läuft zum aktuellen Zeitpunkt, so dass auch dieses Arbeitspaket – wenngleich mit Verspätung – abgeschlossen werden konnte.

AP 5 Verbreitung und Dissemination
AP 5.2.1 Bürgerbezogene Dissemination
Ziel dieses Arbeitspakets war es, die Projektergebnisse in verständlicher Form an Politik, Zivilgesellschaft und interessierte Bürger*innen zu kommunizieren sowie Handlungsempfehlungen zur Bekämpfung von Desinformation zu formulieren. Das medienpsychologische Teilprojekt trug hierbei federführend die psychologischen Inhalte bei.

Mitte 2024 wurde ein gemeinsames Policy Paper verfasst, das Ende 2024 öffentlich zugänglich gemacht und im Frühjahr 2025 auch in englische Sprache übersetzt wurde. Das Policy Paper wurde im Anschluss gezielt an politische Entscheidungsträger*innen auf Bundes- und EU-Ebene disseminiert. Es beschreibt die besonderen Herausforderungen der Desinformationsverbreitung in Messenger-Diensten wie Telegram und WhatsApp, analysiert bestehende rechtliche Rahmenbedingungen mit Fokus auf den Digital Services Act (DSA) der EU und formuliert konkrete Vorschläge zu dessen Ergänzung, um grundrechtskonforme und wirksame Maßnahmen zu ermöglichen. Im Mittelpunkt steht eine Analyse des Präventionsansatzes Prebunking, das vor der eigentlichen Verbreitung von Desinformation ansetzt. Dabei werden sowohl breit angelegte (Broad-Spectrum) als auch inhaltlich spezifische (Narrow-Spectrum) Strategien dargestellt und unter rechtlichen, technischen und psychologischen Gesichtspunkten bewertet. Während Broad-Spectrum-Ansätze als technisch einfach umsetzbar und grundrechtsschonend gelten, deuten psychologische Studien darauf hin, dass Narrow-Spectrum-Strategien wirksamer sein könnten, weshalb weiterer empirischer Forschungsbedarf besteht.

Die Ergebnisse des medienpsychologischen Teilprojekts wurden zudem bei der vom informatischen Partner organisierten Abschlussveranstaltung vorgestellt und online auf YouTube veröffentlicht. Medienpsychologische Inhalte des Policy Papers wurden in einem Podcast-Interview erläutert und sollen der breiten Zivilgesellschaft im Spätsommer 2025 öffentlich zugänglich gemacht werden. Anstelle der ursprünglich geplanten Bürgerinnen-Veranstaltung in Kooperation mit der VHS Duisburg nahm das Teilprojekt am Essener Wissenschaftssommer teil. Gemeinsam mit der Universitätsbibliothek wurde hierfür ein Quiz zum Erkennen von „Fake News“ entwickelt, das im Rahmen einer Marktplatzveranstaltung gespielt werden konnte. Die Aktion ermöglichte einen direkten Bürger*innen-Dialog und bot die Gelegenheit, praxisnahe Projektergebnisse einer breiten Öffentlichkeit zugänglich zu machen.