Sachbericht zum Verwendungsnachweis – Teil II: Eingehende Darstellung

Projekt:	Echtzeiterkennung und Nachweis hybrider Desinformationskampagnen in Online-Medien (HybriD)
Teilvorhaben:	Detektion, Analyse und Lagebild-Integration von hybriden Desinformationskampagnen in Online-Medien
Förderkennzeichen:	16KIS1534
Zuwendungsempfänger:	complexium GmbH, Berlin
Berichtszeitraum:	01.10.2021 – 31.03.2025

1. Ursprüngliche Aufgabenstellung und wissenschaftlich-technischer Stand

Zum Projektstart war die Forschung zu Desinformation primär auf die Identifikation einzelner Akteure (z. B. Social Bots) oder die Klassifizierung isolierter Falschnachrichten („Fake News“) fokussiert. Diese Ansätze griffen jedoch zu kurz, um die wachsende Bedrohung durch breit angelegte, hybride Kampagnen zu erfassen, die auf einem koordinierten Zusammenspiel von teils automatisierten, teils menschlich gesteuerten Accounts basieren.

Das Verbundvorhaben „HybriD“ verfolgt daher einen innovativen und ganzheitlichen Ansatz. Das Ziel war nicht die binäre Klassifikation einzelner Inhalte, sondern die **gesamthafte Detektion wahrscheinlicher Desinformationskampagnen** durch die Erkennung von musterbasierten, temporalen Anomalien in der Online-Kommunikation. Das Teilvorhaben der Complexium GmbH konzentrierte sich dabei auf den Aufbau der technischen Analyse-Infrastruktur sowie die Entwicklung eines interaktiven Demonstrator-Dashboards, um diese neue Art von Bedrohungen frühzeitig und zielgenau erkennen und analysieren zu können.

2. Detaillierte Darstellung der durchgeführten Arbeiten

Chronologischer Ablauf der Arbeiten

Das Vorhaben wurde planmäßig und in enger Abstimmung mit den Verbundpartnern durchgeführt. Der Ablauf gliederte sich in mehrere Phasen:

- **Phase 1:**
  Aufbau der Infrastruktur (2021-2022)
  Zunächst wurden die grundlegenden technischen Voraussetzungen geschaffen. Dies umfasste die Konzeption der IT-Infrastruktur und einer gemeinsamen Datenarchitektur in Abstimmung mit den akademischen Partnern. Parallel wurde mit dem Aufbau der Datenerhebung begonnen, wofür spezifische Crawler für zentrale Plattformen wie Twitter, Telegram und Reddit entwickelt wurden, die auch komplexe Inhalte wie Kommentare erfassen konnten. Den Verbundpartnern wurde über eine Schnittstelle Zugang zur Datenbank geschaffen.

- **Phase 2:**
  Erweiterung der Datenquellen und Entwicklung von Analyseverfahren (2022-2023)
  Die Datenerhebungs-Infrastruktur wurde signifikant erweitert und umfasste nun auch Plattformen wie 4chan, Mastodon, Discord, Instagram und später auch BlueSky, Threads und Gettr. Dies ermöglichte eine umfassende Abdeckung des digitalen Informationsraums. Technisch wurde die Analyse von Telegram-Inhalten vertieft, sodass nun auch geteilte Medien (Videos, Bilder) bis zu ihrem Ursprung zurückverfolgt werden können. Ein zentraler Fortschritt war die Entwicklung eines neuen, KI-basierten Systems auf Basis eines Sprachmodells, das Beiträge automatisiert nach Ereignistypen kategorisiert und geographisch verortet.

- **Phase 3:**
  Konsolidierung und Visualisierung (2024-2025)
  In der letzten Phase wurden die gesammelten Daten und Analysefähigkeiten in einem interaktiven Demonstrator zusammengeführt. Es wurde eine fortschrittliche Visualisierungslösung entwickelt, die komplexe Netzwerkstrukturen und die Verbreitungsmuster (Diffusion) von

Narrativen über verschiedene Plattformen und Zeiträume hinweg darstellt. Zudem wurde ein stabiles Prozedere für den kontinuierlichen Datenaustausch mit den Forschungspartnern etabliert, um die wissenschaftliche Auswertung zu unterstützen.

Diese im Rahmen des Teilvorhabens durchgeführten Arbeiten orientierten sich dabei zudem eng an der ursprünglichen Planung, wurden jedoch agil an die dynamische Entwicklung der digitalen Bedrohungslandschaft und des technologischen Fortschritts angepasst.

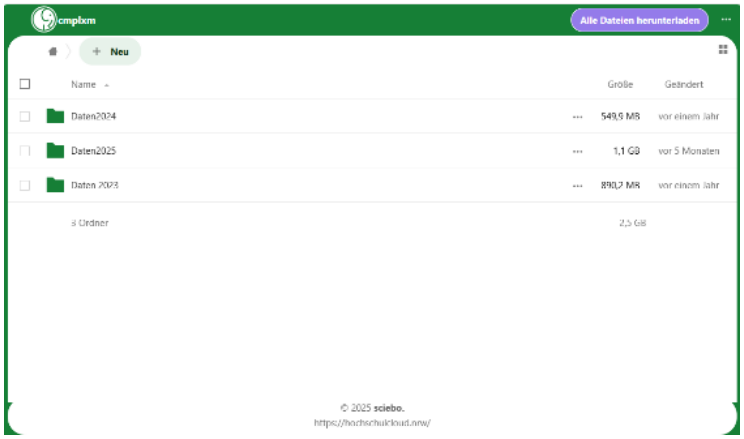
Vergleich zur ursprünglichen Vorhabenbeschreibung anhand der definierten Aufgabenpakete:

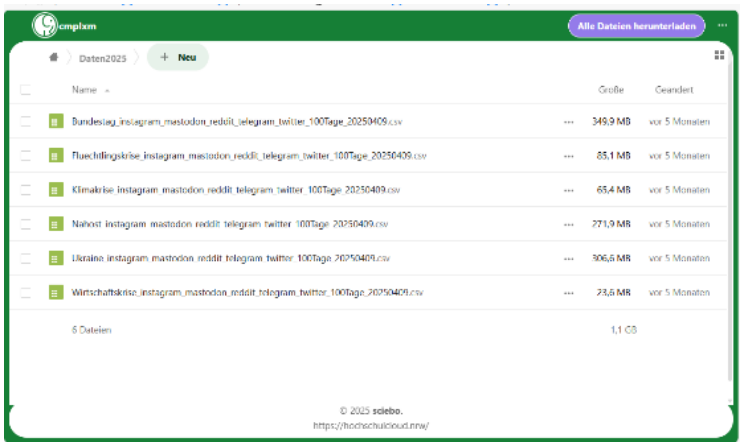
AP 1: Inhaltliche Konzeption

- **Plan:** Sichtung der Forschungsliteratur, theoretische Fundierung und Identifikation von technologischen Optionen zur Detektion von Desinformationsmustern.

- **Umsetzung:** Die Konzeptionsphase wurde wie geplant durchgeführt. Es wurde eine tiefgehende Analyse des Stands der Technik vorgenommen und in enger Abstimmung mit den Verbundpartnern eine robuste und flexible IT-Infrastruktur sowie eine gemeinsame Datenarchitektur konzipiert. Diese bildete die Grundlage für alle weiteren technischen Entwicklungen. Dabei wurde ein Austauschformat definiert, mit dem Complexium den Projektpartnern die gesammelten Daten für die weitere Analyse zur Verfügung stellen kann.

Abbildungen: **Größenvolumen der bereitgestellten Daten.**





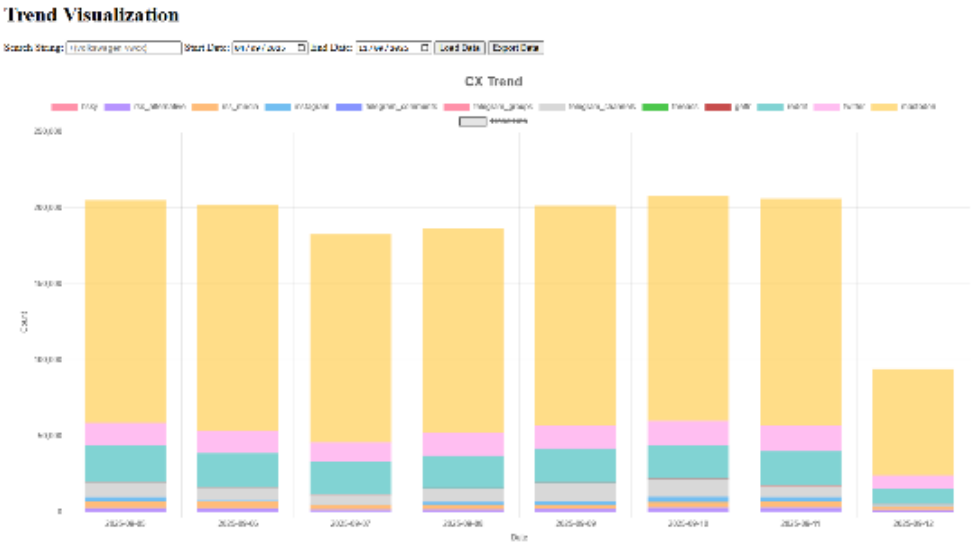
AP 2: Infrastruktur und Daten

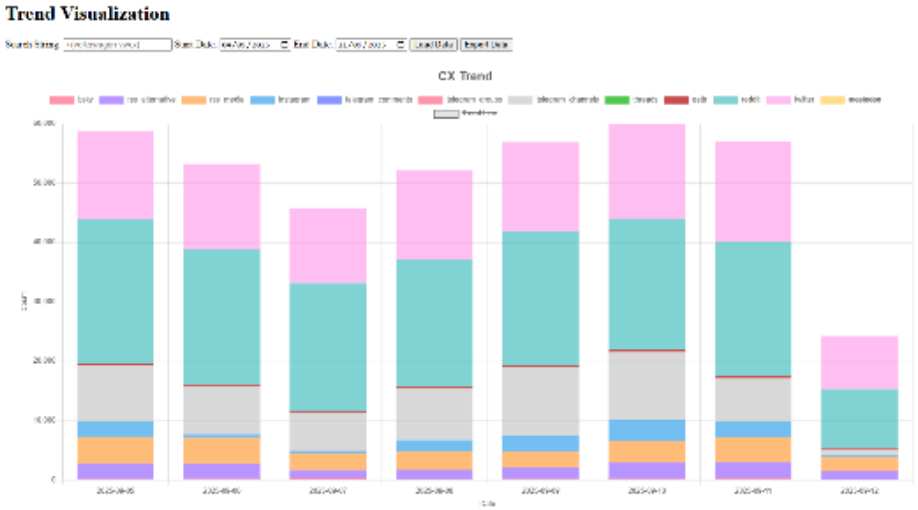
- **Plan:**
  Aufbau einer technischen Infrastruktur (Backend) zur Erfassung, Speicherung und Bereitstellung von Datenströmen aus sozialen Medien. Entwicklung von Crawlern für unterschiedliche Plattformen und Formate zur Datenbeschaffung.

- **Umsetzung:**
  Die Umsetzung übertraf die ursprüngliche Planung. Zunächst wurden wie geplant Crawler für Twitter, Telegram und Reddit entwickelt. Im Projektzeitraum und während der Implementierung der Datenerfassungstools wurde Twitter von Elon Musk übernommen. Dadurch änderte sich nicht nur der Zugang zu den Daten dieser Plattform, sondern die gesamte Landschaft der Plattformen wandelte sich grundlegend. Benutzergruppen wanderten in andere soziale Medien ab, und auf Twitter, später in X umbenannt, kam es zu einer signifikanten Themenverschiebung. Diese Entwicklungen machten eine Erweiterung der ursprünglichen Planung unabdingbar.

  Aufgrund dieser Veränderung in der Plattformlandschaf wurde die Datenerfassung im Projektverlauf signifikant auf weitere relevante Quellen wie **4chan, Mastodon, Discord, Instagram, Gettr, Threads und BlueSky** ausgeweitet. Insbesondere die Erfassung von Telegram wurde technisch vertieft, um nicht nur Textnachrichten, sondern auch die Verbreitung von Medieninhalten (Bilder, Videos) und deren Ursprung nachverfolgen zu können. Ein stabiles Datenaustauschprotokoll wurde etabliert, um den Verbundpartnern kontinuierlich relevante Datenkorpora bereitzustellen.

Abbildungen: **Abdeckungsbreite und Ausmaß der aufgenommenen Daten.**



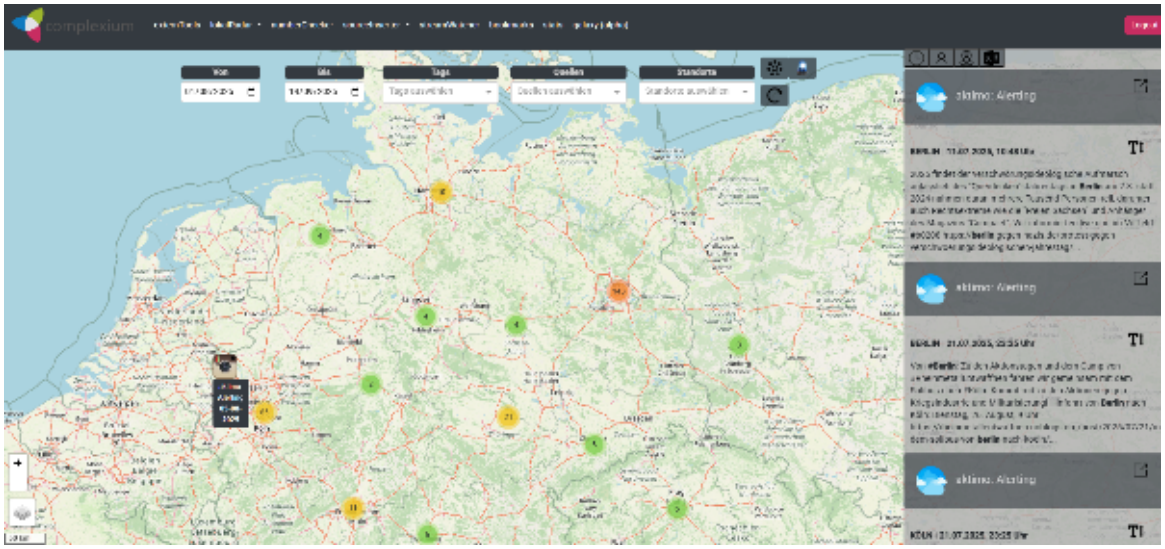


AP 3: Identifikation und Nachweis

- **Plan:**
  Entwicklung von unüberwachten und überwachten Lernverfahren zur Echtzeit-Identifikation von Kampagnenmustern (Shapes) in Datenströmen.

- **Umsetzung:**
  Die geplanten Ziele wurden erreicht und durch den Einsatz neuester KI-Technologien erweitert. Ein zentrales Ergebnis ist die Entwicklung eines neuen Analysesystems, das auf einem **elaborierten Sprachmodell (KI)** basiert. Dieses System ermöglicht eine hypothesenfreie Inhaltsanalyse und kann Beiträge automatisiert nach **Ereignistypen kategorisieren** sowie ihren potenziellen **Wirkungsbereich geographisch verorten**. Diese technologische Weiterentwicklung ergänzt die ursprünglich skizzierten Stream-Clustering-Verfahren und erlaubt eine wesentlich präzisere und tiefere Analyse der Inhalte.

Abbildung: **Geographische Verortung der aufgenommenen Beiträge.**



AP 4: Triangulation durch Experten

- **Plan:**
  Aufbau eines Expertennetzwerks und regelmäßige Durchführung von Workshops und Interviews zur Validierung der Ergebnisse und zur Integration von Praxiswissen.

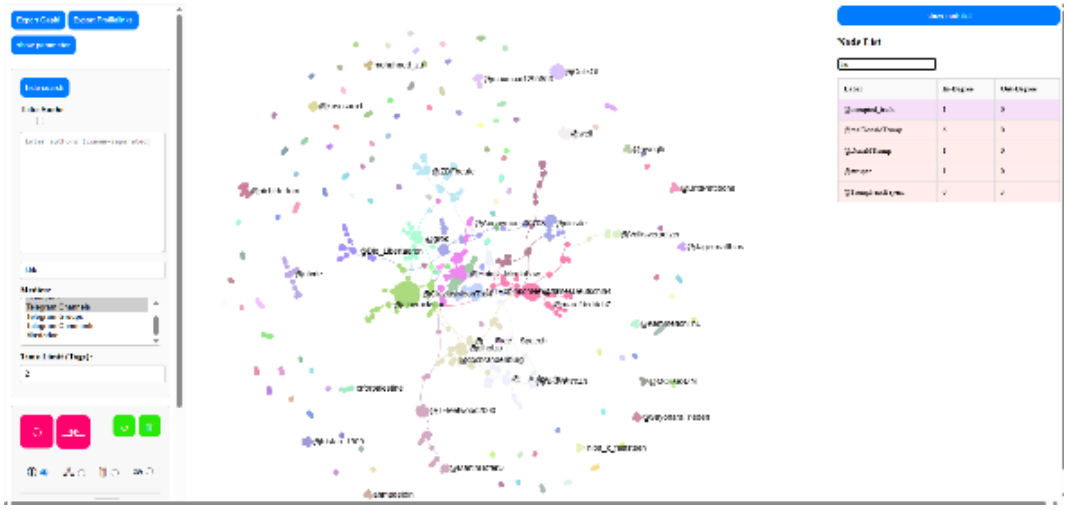
- **Umsetzung:**
  Die Interaktion mit Fachexperten wurde intensiv und erfolgreich umgesetzt. Die Projektergebnisse wurden in zahlreichen Formaten (Workshops, Vorträge, Paneldiskussionen) einem Fachpublikum aus Wirtschaft, Politik, Medien und Sicherheitsbehörden vorgestellt und zur Diskussion gestellt. Dieser Austausch stellte sicher, dass die entwickelten Verfahren und Tools eine hohe Praxisrelevanz aufweisen und an realen Bedrohungsszenarien ausgerichtet sind.

AP 5: Demonstrator eines Analyse-Dashboards

- **Plan:**
  Realisierung eines web-basierten, interaktiven Demonstrators zur visuellen Echtzeitanalyse, zur Tiefenanalyse („Deep Dive“) und zur Integration von Expertenfeedback.

- **Umsetzung:**
  Es wurde eine fortschrittliche, interaktive **Visualisierungslösung als Demonstrator** implementiert. Der Demonstrator stellt komplexe Netzwerkstrukturen dar und erlaubt die präzise Nachverfolgung von Diffusionsmustern von Narrativen über verschiedene Plattformen und Zeiträume hinweg. Damit wurden die Kernanforderungen an den Demonstrator vollständig erfüllt und eine anwenderfreundliche Schnittstelle für die Analyse der komplexen Daten geschaffen.

Abbildung: **Erschließung von Netzwerken.**



3. Wichtigste Positionen des zahlenmäßigen Nachweises

Die Bearbeitung der verschiedenen Arbeitspakete schlug sich in Personalkosten als dem mit über 95% dominierenden Kostenblock nieder. In Summe wurde eine fast perfekte Einhaltung des vorkalkulierten Budgets erreicht: Es ergab sich eine Einsparung gegenüber der Vorkalkulation von 0,236 %.

	IST	Vorkalkulation	Abweichung
0813 Material	4.270,30 €	4.600,00 €	*329,70 €*
0837 Personalkosten	468.938,09 €	463.695,86 €	*- 5.242,23 €*
0838 Reisekosten	3.500,00 €	4.000,00 €	*500,00 €*
0847 vorhabenspezifische Abschreibungen	10.396,59 €	12.705,00 €	*2.308,41 €*
0850 sonstige unmittelbare Vorhabenkosten	5.113,56 €	8.384,00 €	*3.270,44 €*
Summe	**492.218,54 €**	**493.384,86 €**	***1.166,32 €***
Einsparung ggü. Vorkalkulation			*0,236%*

4. Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Die Bedrohung durch Desinformation ist hochdynamisch. Sowohl die Taktiken der Akteure als auch die von ihnen genutzten Plattformen ändern sich kontinuierlich. Die im Projekt durchgeführten Arbeiten waren daher nicht nur angemessen, sondern zwingend notwendig, um die Projektziele zu erreichen.

Notwendigkeit der Quellenerweiterung: Die Ausweitung der Datenerfassung auf neue Plattformen war eine notwendige Reaktion auf die Migration von relevanten Diskursen und Akteuren weg von etablierten Netzwerken. Nur so konnte eine umfassende Abdeckung des Informationsraums gewährleistet werden.

Angemessenheit des KI-Einsatzes: Die sprunghafte Entwicklung im Bereich der Künstlichen Intelligenz eröffnete neue Möglichkeiten für die Datenanalyse. Die Integration eines modernen Sprachmodells war ein angemessener und entscheidender Schritt, um die Erkennungsleistung auf ein State-of-the-Art-Niveau zu heben und die Analysegenauigkeit signifikant zu verbessern.

Angemessenheit der agilen Vorgehensweise: Die flexible Anpassung der Arbeitspakete an neue technische Möglichkeiten und veränderte Rahmenbedingungen war essenziell für den Projekterfolg und stellte eine effiziente und zielgerichtete Verwendung der Fördermittel sicher.

5. Wesentliche Ergebnisse und Zusammenarbeit

Im Rahmen des Vorhabens wurden die angestrebten Ziele erreicht und eine leistungsfähige Systematik zur Analyse hybrider Bedrohungen geschaffen. Die wesentlichen Ergebnisse sind:

- **Umfassende Datenerhebungs- und Analyse-Infrastruktur:**
  Es wurde eine modulare und erweiterbare Infrastruktur aufgebaut, die Daten aus einer Vielzahl relevanter sozialer Netzwerke und Online-Foren in nahezu Echtzeit erfasst und für tiefergehende Analysen aufbereitet.
- **KI-gestützte Analyseverfahren:**
  Durch den Einsatz moderner KI-Modelle konnte die Analysepräzision deutlich erhöht werden. Das System kann Inhalte nicht nur thematisch, sondern auch nach Ereignistyp und geographischer Relevanz klassifizieren.
- **Interaktiver Visualisierungs-Demonstrator:**
  Der entwickelte Demonstrator ermöglicht es Analysten, die Entstehung und Verbreitung von Kampagnen visuell nachzuvollziehen. Komplexe Zusammenhänge in den Daten werden durch intuitive Netzwerkdarstellungen greifbar gemacht.
- **Praxisnaher Transfer und Austausch:**
  Die Methodik und Zwischenergebnisse wurden kontinuierlich mit externen Experten aus Wirtschaft, Politik und Sicherheitsbehörden diskutiert und validiert. Die Zusammenarbeit im Verbund mit der **Universität Münster** und der **HAW Hamburg** verlief durchgehend eng und produktiv, insbesondere bei der Definition der technischen Anforderungen und der Bereitstellung von Daten für die wissenschaftliche Forschung.

Die Projektergebnisse wurden zudem in zahlreichen Fachkreisen vorgestellt, unter anderem bei Workshops an der **Führungsakademie der Bundeswehr**, auf Konferenzen des **OSAC-Chapters** Germany und des **TÜV Nord** sowie in Gesprächen mit Sicherheitsverantwortlichen aus der Wirtschaft und öffentlichen Stellen.

6. Voraussichtlicher Nutzen und Verwertbarkeit des Ergebnisses

Die im Projekt entwickelten Technologien und Erkenntnisse bieten ein hohes Verwertungspotenzial, das sowohl wirtschaftliche als auch wissenschaftlich-gesellschaftliche Aspekte umfasst und im fortgeschriebenen Verwertungsplan konkretisiert wird.

- **Wirtschaftlicher Nutzen:**
  Die Projektergebnisse bilden die Grundlage für kommerzielle Dienstleistungen im Bereich der

Früherkennung digitaler Bedrohungen. Die entwickelte Datenbasis und die Analyse-Tools schaffen einen signifikanten Wettbewerbsvorteil. Bereits parallel zum Projektverlauf wurden Gespräche und Pilotierungen mit mehreren deutschen Unternehmen geführt. Konkret wurde ein **„Aktivismus-Monitor“** als marktfähige Lösung entwickelt, der sich an Verantwortungsträger für Unternehmenssicherheit, Bedrohungsmanagement und Personenschutz richtet. Die wirtschaftlichen Erfolgsaussichten werden angesichts einer steigenden Bedrohungslage als sehr positiv bewertet.

- **Wissenschaftlicher und gesellschaftlicher Nutzen:**
  Die Ergebnisse sind für öffentliche Stellen (z. B. Sicherheitsbehörden) und Aufgaben (z. B. Schutz kritischer Infrastrukturen) hervorragend geeignet. Gespräche mit der „Autobahn GmbH des Bundes“, öffentlich-rechtlichen Sendern und militärischen Stellen (z. B. Operative Kommunikation) haben das hohe Interesse an den entwickelten Lösungen bestätigt. Die modular konzipierte Systemarchitektur gewährleistet die wissenschaftliche und wirtschaftliche Anschlussfähigkeit für zukünftige Forschungs- und Entwicklungsvorhaben, insbesondere in der Integration weiterer KI-Verfahren und der Erschließung zusätzlicher Datenquellen.

7. Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Das Feld der Desinformationsforschung und der digitalen Analyse entwickelt sich kontinuierlich weiter. Während des Projektverlaufs wurde eine stetige Weiterentwicklung von Ansätzen und Methoden zur Datenauswertung bei anderen nationalen und internationalen Forschungsstellen beobachtet.

Diese Entwicklungen wurden aufmerksam verfolgt und relevante Aspekte sind in die Umsetzung der Arbeitspakete eingeflossen. Es wurden jedoch keine externen Forschungsergebnisse bekannt, die die grundlegende Zielsetzung oder den gewählten methodischen Ansatz des HybriD-Projekts in Frage gestellt hätten. Vielmehr bestätigte der allgemeine Trend die Relevanz des gewählten ganzheitlichen und plattformübergreifenden Ansatzes.

8. Erfolgte oder geplante Veröffentlichungen des Ergebnisses

Die Projektergebnisse wurden bereits während der Laufzeit aktiv in relevante Fachkreise kommuniziert, um einen breiten Transfer zu gewährleisten. Dies erfolgte primär durch:

- **Vorträge und Workshops:** Bei zahlreichen Veranstaltungen und Konferenzen wurden die Methodik und Zwischenergebnisse vorgestellt, u. a. bei der Führungsakademie der Bundeswehr, dem deutschen OSAC-Chapter, der Konrad-Adenauer-Stiftung, einer Fachkonferenz zur Flughafensicherheit, dem Risknet-Summit und dem Symposium Anlagensicherung des TÜV Nord.
- **Paneldiskussionen:** Teilnahme an Fachdiskussionen wie der Medienrechtskonferenz in Frankfurt/Oder.

Als eigenständige Fortführung dieser Arbeiten ist die durch complexium die Entwicklung eines ganzheitlichen Lagebildes für Unternehmen sowie die Konzeption eines umfassenden Lagebildes Hybrid geplant.

complexium stellt in passenden Gesprächen etwa mit Unternehmenssicherheitsbereichen stets die Möglichkeit eines „Werkstattbesuchs“ vor Ort heraus. Ebenso werden die entsprechenden Aspekte auf Konferenzen vorgetragen oder eingebracht.

Ein enger Austausch besteht zudem mit dem **VSW Verband für die Sicherheit der Wirtschaft e.V. Bundesverband**: Derzeit wird eine gemeinsame „Studie Desinformation“ erarbeitet, um weiter für die Bedrohung zu sensibilisieren. Teil dieser Studie wird das angestrebte „Lagebild Hybrid/hybride Angriffe“ werden.

Complexium war bereits auf der der gemeinsamen Sicherheitskonferenz von VSW und BfV im Jahr 2025 mit einem Ausstellungsstand vertreten. Eine Wiederholung 2026 ist vermutlich möglich.