

Sachbericht zum Verwendungsnachweis

Teil II – Eingehende Darstellung der Arbeiten

Zuwendungsempfänger: CORRECTIV – Recherchen für die Gesellschaft gemeinnützige GmbH

Verbundprojekt: noFake – KI-unterstütztes Assistenzsystem für die Crowdsourcing-basierte Erkennung von über digitale Plattformen verbreiteter Desinformation

Teilvorhaben: NoFake Community

Förderkennzeichen: 16KIS1521

Bewilligungszeitraum: 01.12.2021 - 28.02.2025

Einleitung zu Teil II – Eingehende Darstellung der Arbeiten

Das Teilvorhaben “NoFake Community” von CORRECTIV nahm innerhalb des Verbundprojekts noFake eine vermittelnde Rolle zwischen technologischer Entwicklung, journalistischer Praxis und zivilgesellschaftlicher Beteiligung ein. Während die wissenschaftlichen Partner der Technischen Universität Berlin, der Ruhr-Universität Bochum und der TU Dortmund technische, linguistische und rechtliche Grundlagen erforschten, lag der Schwerpunkt von CORRECTIV auf der praktischen Erprobung und Umsetzung eines partizipativen Ansatzes zur Bekämpfung von Desinformation.

Ziel des Teilvorhabens war es, Bürgerinnen und Bürger systematisch in Prozesse der Desinformationsaufklärung einzubeziehen und digitale Werkzeuge in die Beteiligungsplattform der Faktencheck-Community zu integrieren. Damit sollten nachhaltige Wege der Medien- und Informationskompetenzbildung eröffnet, zugleich das Faktenchecken durch Crowdsourcing und kollektive Recherche ergänzt sowie technische Lösungen für einen effizienten Faktencheck-Prozess erprobt werden. Das Vorhaben verband dafür redaktionelle Arbeitsweisen, technologische Entwicklung und medienpädagogische Konzepte zu einem Modell, das kollaboratives Arbeiten im Faktenchecken erprobt und weiterentwickelt.

Die folgenden Abschnitte stellen die im Projektzeitraum durchgeführten Arbeiten, erzielten Ergebnisse und gewonnenen Erfahrungen entlang der einzelnen Arbeitspakete dar. Sie verdeutlichen, wie Forschungsergebnisse der Verbundpartner in die Praxis übertragen und durch praktische Erkenntnisse aus Community-Arbeit, Plattformentwicklung und Medienkompetenz-Bildung ergänzt wurden. Zugleich zeigen die Ergebnisse aus den Arbeitspaketen,

unter welchen Bedingungen Community-basiertes Faktenchecken tragfähig ist und welche strukturellen Voraussetzungen für eine Verstetigung erforderlich sind.

AP 1.1 Prozesse erforschen und Methoden entwickeln für das crowd-basiertes Monitoring

CORRECTIV entwickelte und erprobte Methoden für community-basiertes Faktenchecken, die den gesamten Prozess von der Einreichung einer Behauptung bis zur Veröffentlichung einer überprüften Faktensammlung abbilden. Das Ziel war, den Faktencheck-Prozess so zu gestalten, dass Freiwillige ohne journalistische Vorbildung daran mitwirken können. Durch die Beteiligung erlangen sie nicht nur Medien- und Informationskompetenz, sie können neues Wissen direkt anwenden und ihnen wird Teilhabe am öffentlichen Diskurs durch journalistische Mittel ermöglicht. Die Veröffentlichungen wirken über die Community hinaus und sind gleichzeitig ein zentrales Incentive zur Beteiligung.

Das Team kombinierte etablierte Ansätze aus Prebunking und Debunking zu einem mehrstufigen Peer-Production-Modell. In dieser Struktur wirken Freiwillige, Community-Managerinnen und die Redaktion arbeitsteilig zusammen. Community-Mitglieder können unterschiedliche Beteiligungsformen wählen. Sie schlagen Behauptungen vor, sammeln Fakten, diskutieren Einordnungen und Rechercheergebnisse oder lesen einfach nur mit. Die Community-Recherchen veröffentlicht CORRECTIV in Form von sogenannten Faktensammlungen, ein für das Projekt entwickeltes Format, das den kollaborativen Ursprung der Recherchen widerspiegelt. Jede Faktensammlung wird vor Veröffentlichung von einem Faktencheck-Redakteur oder einer Redakteurin geprüft, um Qualitätsstandards sicherzustellen. Die redaktionellen Standards im Faktenforum orientieren sich an dem Kodex internationaler Faktencheck-Netzwerke wie dem European Fact-checking Standards Network (EFCSN) und sind auf der Faktenforum-Webseite dokumentiert: <https://www.faktenforum.org/ueber-uns/unsere-standards-und-arbeitsweisen>

Um dieses Modell empirisch zu untermauern, führte CORRECTIV zu Projektbeginn qualitative und quantitative Erhebungen durch. 326 Personen nahmen an einer Online-Befragung teil, sechs vertiefende Leitfadeninterviews ergänzten die Ergebnisse. Diese Daten bildeten eine Grundlage für das Design der späteren Plattformfunktionen und für die Entwicklung von Lernmaterialien. Im Frühjahr 2025 zum Ende der Projektlaufzeit führte CORRECTIV eine weitere Nutzerumfrage (n=790) durch, um die weitere Entwicklung des Projekts zu untermauern. Die Ergebnisse geben zudem einen Eindruck über die Projekterfolge zum Förderzeitraumende: <https://correctiv.org/in-eigener-sache/2025/07/24/community-umfrage-correctiv-faktenforum-unt-erstuetzt-im-kampf-gegen-desinformation/>

Das Arbeitspaket legte den Grundstein, Faktenchecken in einem kollaborativen, denzentralen Prozess umzusetzen. Erste Rückmeldungen aus Workshops und ein mehrmonatiger Testlauf für

Recherche- und Community-Abläufe über einen Discord-Server zeigten das Interesse an Engagement für eine faktenbasierte Debatte. Sie machten zudem deutlich, dass eine klare Struktur und Begleitung der Community-Recherchen nötig ist. Diese Erkenntnisse waren entscheidend für die spätere Gestaltung der Community-Formate, Lernmodule und dem Software-Front-End.

Ergebnis des Arbeitspakets 1.1 ist ein methodisch fundiertes Konzept für kollaboratives Faktenchecken, das technische, redaktionelle und medienpädagogische Anforderungen integriert. Die Erkenntnisse waren leitend für die folgenden Arbeitspakete, etwa für die Entwicklung der Plattform und der Schulungsmaterialien.

AP 1.2 – Entwicklung eines Software-Front-Ends für das Monitoring

Auf Grundlage der im Arbeitspaket 1.1 entwickelten Konzepte ging CORRECTIV die technische Umsetzung des CORRECTIV.Faktenforums um. Ziel war eine Plattform, die den Faktencheck-Prozess vollständig abbildet, kollaborativ zugänglich macht – von der Einreichung potenziell falscher Behauptungen bis zur Veröffentlichung geprüfter Inhalte. Gleichzeitig soll sie unterstützende Tools und Community-bildende Maßnahmen integrieren können.

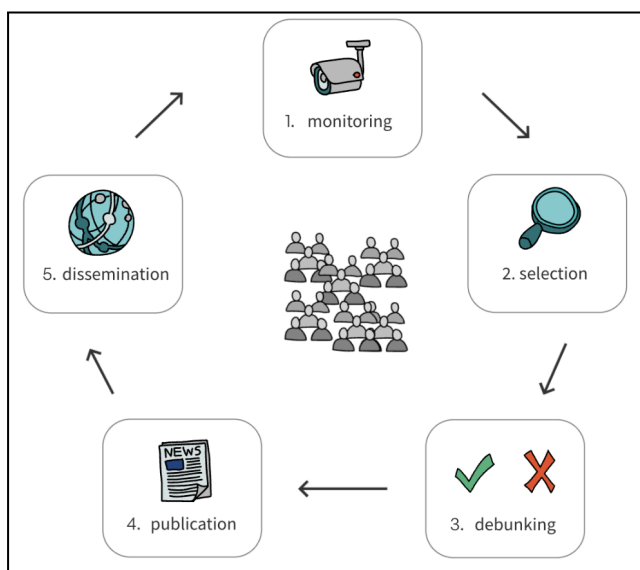


Abbildung 1: Schritte im Faktencheck-Prozess, die die Plattform kollaborativ zugänglich machen soll. Quelle: CORRECTIV

CORRECTIV initiierte innerhalb von noFake einen mehrtägigen Design-Sprint mit einer externen Agentur, um die Plattform-Entwicklung anzustoßen. Für eine Synchronisierung und ein besseres Verständnis innerhalb des Verbunds war dieser sehr hilfreich. Die Gestaltung des User-Interfaces orientierte sich letztlich aber vor allem an den praktischen Erfahrungen redaktioneller Abläufe und einem Community-Testlauf über einen Discord-Server.

Die Plattform basiert auf TypeScript und Vue.js; als Schnittstelle kommt GraphQL zum Einsatz, umgesetzt über Nuxt.js und Hasura. Die Stackwahl wurde von anderen Civic-Tech-Projekten von CORRECTIV informiert, mit denen das Faktenforum Synergien schafft, etwa über einen fachlichen Austausch von Entwicklern. Die Architektur erlaubt eine flexible und offene Struktur im Sinne des Open-Source-Ansatzes von CORRECTIV. Das Projektteam dokumentierte die Entwicklung während des Projekts fortlaufend auf GitHub. Bevor die Plattform in Betrieb genommen wurde, wurde sie einem PEN-Test von einem IT-Sicherheitsunternehmen unterzogen.

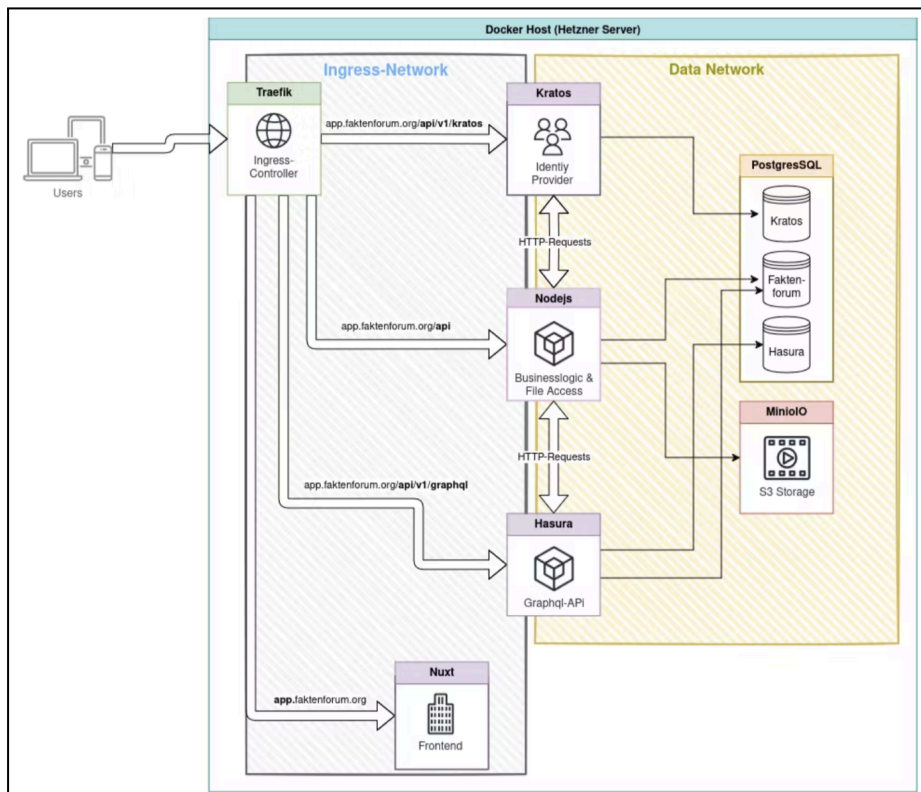


Abbildung 2: Die Logik der Faktenforum-Plattform.

Nutzerinnen und Nutzer loggen sich über die Faktenforum-Webseite ein, um in den Nutzerbereich der Plattform zu gelangen. Dort können sie Behauptungen einreichen, Dokumente und Quellen hochladen, sich an Diskussionen beteiligen und den Fortschritt einzelner Faktenchecks verfolgen. Die Oberfläche wurde bewusst niedrigschwellig gestaltet, um digitale Barrieren zu vermeiden und Teilhabe zu erleichtern. Neben dem Bereich für kollaboratives Faktenchecken umfasst die Plattform offene Themenforen und ein Handbuch, das in acht Kapiteln Lerninhalte in Text und Video zum Faktenchecken anbietet.

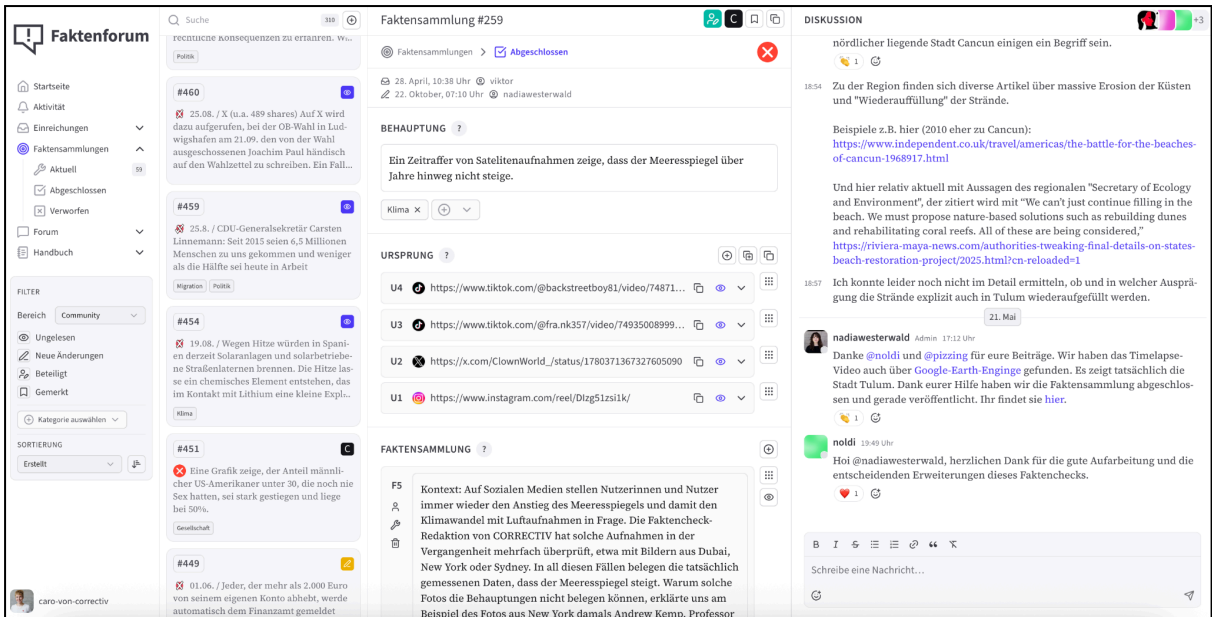


Abbildung 3: Nutzer oder Community-Manager schlagen Behauptungen zur Prüfung vor und dokumentieren deren Verbreitung über eine Linksammlung. Einzelne Fakten dienen der Einordnung und werden kollaborativ zusammengetragen. Jeder Behauptung ist ein Diskussionsforum angehängt.

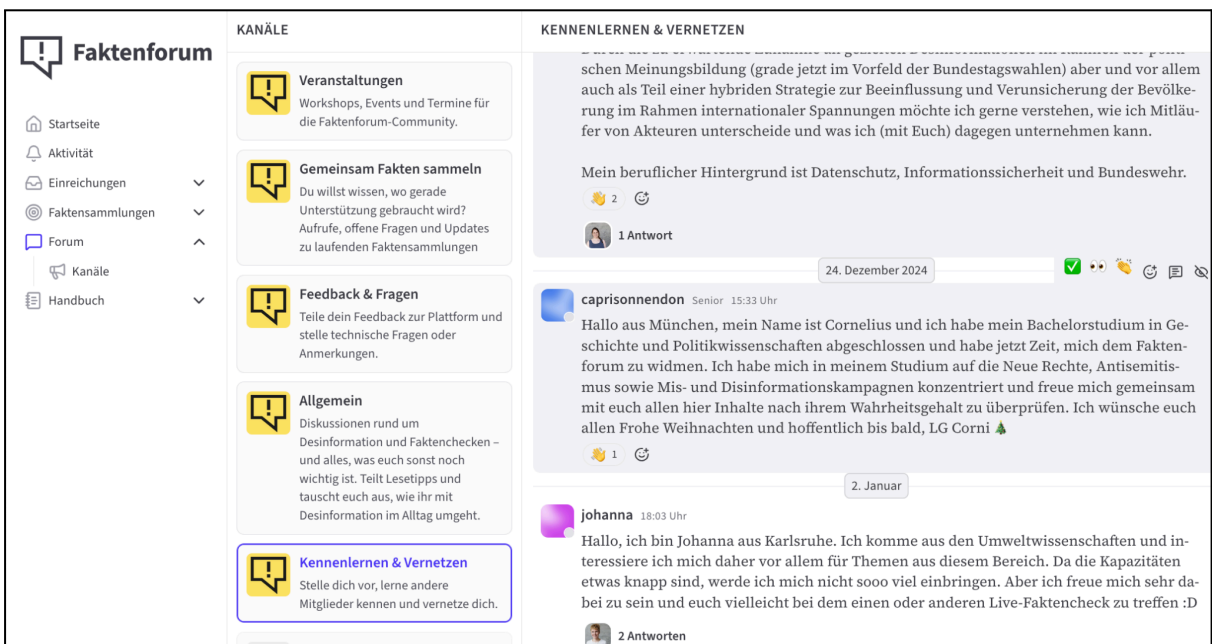


Abbildung 4: In offenen Diskussionsforen können sich Community-Mitglieder über konkrete Behauptungen hinaus beteiligen.

Faktenforum


- Startseite
- Aktivität
- Einreichungen
- Faktensammlungen
- Forum
- Handbuch

- 1 Community-Richtlinien
- 2 So funktioniert das Faktes...
- 3 Was ist Desinformation?
- 4 Was überprüfen wir?
- 5 Die Basics: Suchmaschine...
- 6 Geolocation: Faktencheck...
- 7 Deepfakes: So erkennst d...
- 8 Quellen: So prüfen und b...
- 9 Simple Parolen: So kanns...

7 Deepfakes: So erkennst du manipulierte Inhalte

Bilder, Videos und Audiospuren können täuschend echt manipuliert werden – mithilfe künstlicher Intelligenz entstehen sogenannte Deepfakes. Doch woran erkennt man sie? Hier erfahrt ihr, welche Anzeichen auf eine Fälschung hindeuten und welche Tools bei der Überprüfung helfen.

Lass dich nicht täuschen! So erkennst du...
Faktencheck
correctiv.org/faktencheck



Täuschend echt?

So erkennst du Deepfakes

Ansehen auf YouTube

Was sind Deepfakes

Deepfakes sind mithilfe von KI manipulierte Bild-, Video- und Audio-Dateien, die täuschend echt wirken können. Teilweise sind die KI-Tools, mit denen das möglich ist, in entsprechenden Apps und andere Software eingebaut. „Deep“ kommt von *Deep Learning*, sprich maschinellem Lernen, einer Technik der

Abbildung 5: Im Handbuch finden die Nutzerinnen und Nutzer Lerninhalte und Faktencheck-Tipps – hier das Kapitel zu Deepfakes.

Nach einer internen Testphase startete im Oktober 2024 ein Pilotbetrieb mit 35 Community-Mitgliedern. Das dabei gewonnene Feedback führte zu Anpassungen bei Navigation, Benachrichtigungssystemen und Kommunikationsfunktionen. Der öffentliche Launch erfolgte am 5. Dezember 2024 in Berlin. Innerhalb eines Monats registrierten sich 157 Nutzerinnen und Nutzer; bis zum Projektende waren es über 1.500, zum ersten Geburtstag der Plattform im Dezember 2025 über 2.000. Im Jahr 2025 veröffentlichte CORRECTIV mehr als 100 Community-Recherchen in Form von Faktensammlungen. Die meistgeklickte Veröffentlichung vom 3. März 2025 erzielte rund 300.000 Views.

Mit zunehmender Nutzung entsteht ein kontinuierlich wachsender, strukturierter Datensatz aus Behauptungen, Recherchematerialien und Einordnungen. Diese Datenstruktur ist in der Form einzigartig und für unterschiedliche Ansätze zum Verständnis sowie zur Eindämmung von Desinformation relevant. Über das Projekt hinaus bildet er eine Grundlage für weiterführende Analysen.

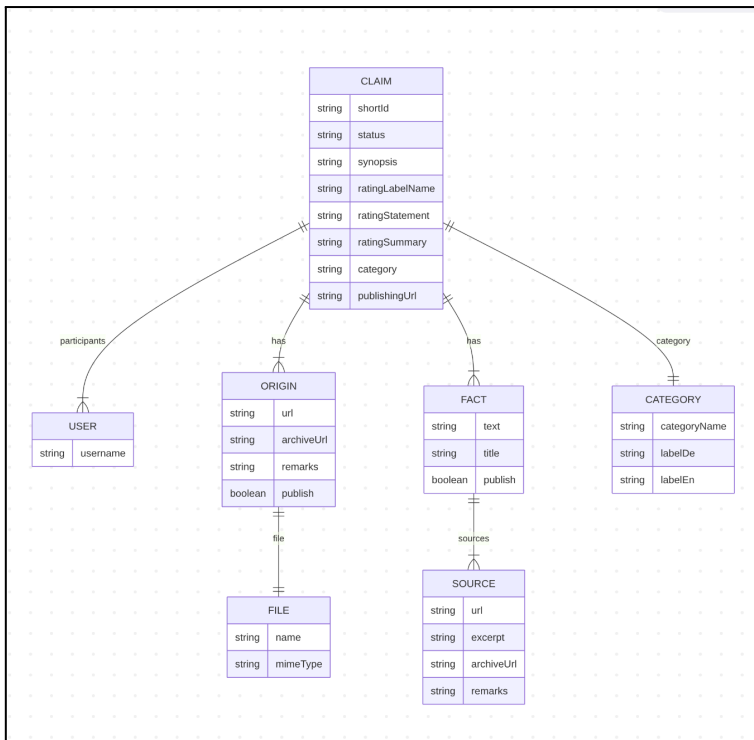


Abbildung 6: Die Struktur des wachsenden Datensatzes im Faktenforum.

Neben der technischen Entwicklung trug das Arbeitspaket wesentlich zur Vernetzung des Projekts bei. CORRECTIV präsentierte das Faktenforum über die Projektlaufzeit hinaus bei über 50 Fachveranstaltungen mit unterschiedlichen Zielgruppen. Dazu gehören der *Digitalgipfel*, die *re:publica*, das *ZEIT Forum Wissenschaft*, die Frankfurter Buchmesse, beim *Aspen Institute Dialogue*, bei internationalen Konferenzen wie mehrmals bei der *Global Fact* oder beim *JournalismAI Festival*, der Jahreskonferenz des BVDA oder von *netzwerk Recherche*, die Lokalkonferenz, die Wissenschaftskonferenz des ZAF, beim *bfuture-Festival*, beim *transform_D Summit* der DSEE und vielen mehr. Die Resonanz dieser Veranstaltungen zeigte, dass das Projekt eine Lücke zwischen Forschung, Medienpraxis und Gesellschaft schließt und einen wichtigen Beitrag zur Eindämmung von Desinformation und zur Steigerung der gesellschaftlichen Resilienz gegen Desinformation leisten kann.

Mit der Faktenforum-Plattform entstand erstmals eine skalierbare Infrastruktur, die Partizipation am Faktenchecken ermöglicht und den Einsatz von KI-Assistenz vorbereitet.

AP 4.1, 4.4 und 4.6 Kollaboration, Community-Management, KI-Assistenz und Validierung

CORRECTIV entwickelte in diesen Arbeitspaketen die organisatorischen und technischen Grundlagen für kollaboratives Faktenchecken. Ziel war, Freiwillige effizient in redaktionelle Prozesse einzubinden und zu vernetzen und gleichzeitig redaktionelle Qualität sicherzustellen.

Seit Ende 2023 wurde das Community-Management intensiviert. CORRECTIV organisierte regelmäßig Online-Workshops, in denen neue Mitglieder geschult und bestehende in spezifischen Themen weitergebildet wurden. 2024 fanden 23 Workshops statt, die Themen wie Quellenbewertung, visuelle Recherche oder KI-gestützte Tools behandelten. Diese Formate wurden zu einem festen Bestandteil der Community-Arbeit. 2025 richtete das Faktenforum über 85 Community-Workshops aus und diversifizierte die Formate. Die Themen reichten von Quellenbewertung und Monitoring bis zum reflektierten Einsatz KI-gestützter Tools.

Frühe Tests zeigten, dass Motivation innerhalb der Community weniger durch Wettbewerb als durch Sinnstiftung, Lerngelegenheiten und persönliche Rückmeldungen entsteht. Vor diesem Hintergrund verzichtete CORRECTIV nach Testphasen bewusst auf umfassende Gamification-Elemente und setzte stattdessen auf Betreuung, Feedback und transparente redaktionelle Einbindung.

Mit dem Checkbot und der „Checkworthiness“-Erkennung wurden erste KI-Tools für einen Testlauf eingebunden. Sie unterstützen Nutzerinnen bei der Recherche und Einordnung, ohne den menschlichen Prozess zu ersetzen. Die Integration dieser Systeme wurde gemeinsam mit den Partnern der TU Berlin und der Ruhr-Universität Bochum umgesetzt und auf der Plattform getestet. Es stellte sich heraus, dass der mögliche Mehrwert von KI-basierten Tools abhängig vom jeweiligen Anwendungsfeld ist. Der Checkworthiness-Algorithmus wurde beispielsweise nur von den CORRECTIV-Mitarbeitenden getestet, denn der automatisierte Selektionvorschlag für oder gegen einen Faktencheck, kann User davon abhalten, mögliche Falschbehauptungen einzureichen – entgegen der Projektziele. Denn diese verlangen schon bei Monitoring und Selektion Beteiligung.

Die Validierung des Ansatzes im Live-Betrieb bestätigte die Tragfähigkeit des Modells: Community-basiertes Faktenchecken funktioniert, wenn technische Umfeld, redaktionelle Sorgfalt und kontinuierliche Begleitung ineinandergreifen. Im Ergebnis entstand ein belastbares Modell für kollaboratives Arbeiten im Faktenchecken, das von CORRECTIV auch nach Projektende weitergeführt, skaliert und weiterentwickelt wird.

Notwendigkeit und Angemessenheit der geleisteten Arbeiten

Die Projektarbeiten von CORRECTIV waren notwendig, um den Ansatz des Community-basierten Faktencheckens technisch und organisatorisch zu etablieren. Der Aufbau der Plattform, die Begleitung der Community und die Sicherung redaktioneller Qualität erforderten hohe personelle und konzeptionelle Ressourcen.

Die Mittel wurden angemessen eingesetzt – vor allem für Personal, Entwicklung, Workshops und die Produktion von Lernmaterialien. Ein wesentlicher Teil entfiel auf die technische Entwicklung

des Frontends, auf die redaktionelle Begleitung der Community sowie auf Schulungs- und Netzwerkaktivitäten.

Das Projekt reagierte zugleich auf neue Herausforderungen durch generative KI und veränderte Informationsräume. Durch die stetige Anpassung an neue Entwicklungen – etwa die zunehmende Anzahl von Deepfakes, der Einfluss ausländischer Desinformationskampagnen etwa auf Wahlen oder die Veränderungen in Social-Media-Ökosystemen – blieb die Arbeit stets relevant. Mit dem Faktenforum konnte CORRECTIV ein praxisbewährtes Modell schaffen, das technologische Innovation und gesellschaftlichen Nutzen verbindet.

Abschließende Reflexion und Ausblick

Das CORRECTIV.Faktenforum zeigt, dass technologisch gestützte und zugleich gemeinwohlorientierte Innovationen im Journalismus realisierbar sind. Die Entwicklung des CORRECTIV.Faktenforums, die Einbindung von KI-Assistenzsystemen und der Aufbau einer aktiven Community stellen eine Bereicherung im Kampf gegen Desinformation dar.

Das Projekt hat darüber hinaus neue Formen der Zusammenarbeit im Journalismus erprobt und weiterentwickelt. Die strukturierte Verbindung einer professionellen Redaktion mit ehrenamtlicher Beteiligung zeigt, dass kollaborative Arbeitsweisen auch unter Wahrung journalistischer Qualitätsstandards tragfähig sind. Damit eröffnet das Modell Perspektiven für gemeinnützige Medienarbeit, die auf Transparenz, nachvollziehbaren Prozessen und wechselseitigem Vertrauen basiert. Die Vernetzungsmöglichkeiten setzen zudem gesellschaftlicher Polarisierung etwas entgegen.

Zugleich entstanden neue Lern- und Qualifizierungsformate: Workshops, modulare Lerninhalte und begleitete Recherchen können künftig systematisch auf der Plattform umgesetzt und in Kooperation mit externen Bildungs- und zivilgesellschaftlichen Partnern weiterentwickelt werden. Das Faktenforum fungiert damit nicht nur als Produktionsumgebung für Faktenchecks, sondern auch als praxisnaher Lernraum für Medien- und Informationskompetenz.

CORRECTIV plant, die im Projekt entwickelten Prozesse organisatorisch und technisch zu verstetigen und schrittweise weiterzuentwickeln. Dazu zählen insbesondere der Ausbau des strukturierten Datensatzes zu Falschbehauptungen und deren Einordnung, die Weiterentwicklung der Plattform als Test- und Anwendungsumfeld für KI-gestützte Assistenzsysteme im journalistischen Kontext sowie die Erschließung zusätzlicher Nutzungsmöglichkeiten der Ergebnisse für Forschung, Bildung und zivilgesellschaftliche Akteure.

Darüber hinaus ist vorgesehen, den Ansatz über bestehende Netzwerke wie das European Fact-Checking Standards Network (EFCSN) und EDMO im europäischen Kontext weiterzugeben und perspektivisch mehrsprachig auszubauen. Auf diese Weise soll das Faktenforum auch über den

Förderzeitraum hinaus als nachhaltige Infrastruktur für kollaboratives Faktenchecken und Medienkompetenzarbeit wirken. Eine mehrsprachige Ausrichtung ermöglicht nicht nur Kooperationen über Ländergrenzen hinweg; sie ist auch für Deutschland relevant , um mehrsprachige Menschen in den Faktencheckprozess zu integrieren.

Insgesamt hat das Projekt wesentlich dazu beigetragen, belastbare Strukturen zur Stärkung digitaler Öffentlichkeit, zur Erhöhung von Desinformationsresilienz und zur Förderung von Medienkompetenz zu schaffen.