



Jan Rau

Plattform-Badges für zivile Kommunikation

Stärkung der demokratischen Öffentlichkeit im digitalen Raum

Auf einen Blick

Digitale Plattformen sind zentrale Arenen des öffentlichen Diskurses, auf denen jedoch Desinformation, extreme Positionen und polarisierende Inhalte durch die algorithmische Verbreitung oft überproportionale Sichtbarkeit erlangen. Plattform-Badges sind eine Governance-by-Design-Maßnahme zur Minderung solcher systemischen Risiken gemäß Art. 34/35 DSA: Wer sich zu Sorgfaltspflichten und deliberativer Kommunikation verpflichtet, erhält algorithmische Sichtbarkeitsvorteile. So werden Anreize für zivilen Austausch geschaffen, strukturelle Verzerrungen korrigiert und Grundrechte gewahrt, um die Resilienz demokratischer Öffentlichkeiten zu stärken.

Einleitung

Digitale Plattformen sind zu zentralen Infrastrukturen politischer Kommunikation und gesellschaftlicher Aushandlungsprozesse geworden und werden in dieser Rolle immer

wieder kritisch diskutiert. Sie strukturieren Informationsflüsse und bestimmen maßgeblich, welche Themen, Akteur:innen und Narrative Aufmerksamkeit und Sichtbarkeit im öffentlichen Raum erlangen – in der digitalen Aufmerksamkeitsökonomie eine der wichtigsten Ressourcen erfolgreicher politischer Kommunikation (Schroeder 2018; Webster 2014). Die Plattformen stehen im Kontext dieser Aufmerksamkeitsökonomie selbst im Wettbewerb miteinander und versuchen, die Zeit und Aufmerksamkeit der Nutzenden auf ihren jeweiligen Angeboten zu maximieren. Ihre Architekturen sind dementsprechend häufig auf Interaktion optimiert und belohnen Inhalte, die viele Reaktionen bei Nutzer:innen hervorrufen (Brady et al. 2023).

Was in nicht politischen Kontexten verhältnismäßig unproblematisch erscheinen mag (als klassisches Beispiel: die algorithmische Verstärkung der Sichtbarkeit emotional aufgeladener Katzen-Videos), führt in politischen Kontexten häufig zur Priorisierung von Kommunikation, die sich durch hohe Konfliktorientierung, Emotionalisierung oder Radikalität auszeichnet. Auch Desinformationen erfüllen oft diese Kriterien und erhalten dadurch wiederholt einen hohen Verbreitungsgrad (Bakir/McStay 2018; McLoughlin et al. 2024),

Plattform-Badges für zivile Kommunikation

Die Plattform



- Bietet die Badges als Opt-in-Funktion an
- Erhöht die Sichtbarkeit der Kommunikation von Badge-User:innen
- Überprüft regelmäßig die Einhaltung der Vorgaben durch die User:innen

Die User:innen



- Nehmen freiwillig am Badge-Programm teil
- Verpflichten sich zur Einhaltung von Normen ziviler Kommunikation
- Integrieren diese Normen in ihre Kommunikation auf der Plattform

Normen ziviler Kommunikation

- Sorgfaltspflichten zur Überprüfung des Wahrheitsgehalts verbreiteter Informationen
- Implementierung Grundsätze deliberativer Kommunikation (z. B. Rationalität, Reziprozität, Respekt und Konstruktivität)



Quelle: eigene Darstellung.

wovon sowohl in- als auch ausländische manipulationsorientierte Kommunikator:innen profitieren können.

Da diese konfliktorientierten, emotionalisierenden und radikalen Kommunikationsformen durch algorithmische Bevorzugung besonders erfolgreich werden, entstehen Anreize, Konflikte um des Konfliktes willen zu führen – oft genug ohne substanzielle inhaltliche Deckung (Spektakel schlägt Inhalt). Für demokratische Öffentlichkeiten resultiert daraus eine strukturelle Benachteiligung ziviler, wahrheitsorientierter oder verbindender Kommunikation, während destruktive Kommunikationsformen an Gewicht gewinnen. Gegenüber Informations- und Manipulationskampagnen ausländischer Kommunikator:innen entstehen systemische Vulnerabilitäten, ebenso gegenüber sogenannten Polarisierungsunternehmer:innen (Mau et al. 2024), die systematisch versuchen, gesellschaftliche Spaltung für ihren eigenen politischen – mitunter auch ökonomischen – Erfolg zu verstärken.

Die Europäische Union hat mit dem Digital Services Act (DSA) einen regulatorischen Rahmen geschaffen, der Plattformen erstmals umfassend in die Verantwortung für solche gesellschaftlichen Risiken nimmt. Neben verschiedenen anderen Maßnahmen verpflichtet der DSA sehr große Onlineplattformen (Very Large Online Platforms – VLOPs) dazu, systemische Risiken für demokratische Prozesse im Kontext ihrer Plattformen zu identifizieren (Art. 34 DSA) und Gegenmaßnahmen zu implementieren (Art. 35 DSA).

In diesem Kontext positioniert sich das Konzept der Plattform-Badges. Das Badge-System kann als Instrument verstanden werden, mit dem Plattformen ihrer Verpflichtung zur Risikominderung nachkommen, indem sie positive Anreize für Nutzende schaffen, bestimmte kommunikative Normen einzuhalten. Verpflichten sich Nutzende besagten kommunikativen Normen, wird im Gegenzug die Sichtbarkeit ihrer Inhalte auf den Plattformen algorithmisch erhöht. Das Badge-System versucht so, eine neue Logik der Aufmerksamkeitsverteilung auf digitalen Plattformen zu etablieren und die strukturellen Bedingungen digitaler Informationsflüsse zugunsten ziviler Kommunikation zu verbessern.

Ausgestaltung der Plattform-Badges für zivile Kommunikation

Die Grundidee des Badges folgt einem Governance-by-Design-Ansatz: Die Architektur digitaler Plattformen beeinflusst, wie sich Nutzende auf diesen Plattformen verhalten, was sie sehen und wie sie mit den Plattformen interagieren – nicht deterministisch, aber substanziell (Grüning et al. 2024; Yeung 2017). Entscheidend ist dabei, dass solche Design-Entscheidungen von Plattform-Architekturen nicht nicht wirken können: Jede Entscheidung über die Ausgestaltung eines bestimmten Plattform-Features beeinflusst automatisch, welcher Möglichkeitsraum sich für das Erleben und Handeln von Nutzenden auf den jeweiligen Plattformen ergibt. Mit Blick auf die gewachsene Bedeutung digitaler

Plattformen für den öffentlichen Diskurs ergibt sich dadurch die Dringlichkeit einer genauen Überprüfung der Effekte solcher Plattform-Design-Entscheidungen auf den öffentlichen Diskurs – und inwiefern Plattformen anders gebaut werden können, um wünschenswertere Effekte zu erzielen.

Das Badge-System knüpft hier an und soll im Wesentlichen als Mechanismus dienen, mit dem die Aufmerksamkeitslogiken digitaler Plattformen gezielt in Richtung ziviler Kommunikation verschoben werden können (siehe Abbildung 1). Der Kern des Ansatzes liegt darin, dass Nutzer:innen, die sich zu bestimmten kommunikativen Verhaltensnormen verpflichten, ein Badge, also eine Art Gütesiegel, erhalten. Diese Auszeichnung ist nicht nur ein symbolisches Signal, sondern entfaltet ihre Wirkung vor allem dadurch, dass sie algorithmisch berücksichtigt wird: Inhalte von Nutzer:innen mit Badge erhalten mehr Reichweite, werden in Empfehlungsmechanismen prominenter platziert und erscheinen sichtbarer in Feeds oder Kommentarspalten.

Kommunikationsnormen für eine demokratische digitale Öffentlichkeit

Die Vergabe eines Badges knüpft an die Einhaltung demokratisch wünschenswerter Kommunikationsnormen von ziviler Kommunikation an. Diese können prinzipiell unterschiedlich ausgestaltet werden. Beispiele solcher Normen könnten unter anderem Sorgfaltspflichten zur Vermeidung der Verbreitung von Des- und Misinformationen sein, wie sie auch der deutsche Presskodex vorsieht (Presserat (German Press Council) 2017). Hier könnten Nutzende beispielsweise darauf verpflichtet werden, Informationen auf ihren Wahrheitsgehalt zu überprüfen, eine Verfälschung desselben zu vermeiden, symbolische Fotos bei Verwendung klar als solche zu benennen etc. Eine Anpassung und Konkretisierung der Vorgaben des Presskodexes für den spezifischen Kontext sind zu empfehlen.

Eine weitere Norm für die Vergabe eines Badges könnte die Implementierung deliberativer Kommunikationsformen in die eigene Kommunikation sein. Es gibt unterschiedliche Verständnisse und Definitionen, was genau deliberative Kommunikation ausmacht, als entscheidende Komponenten werden aber oft Aspekte wie Rationalität (Verwendung von unterstützenden Argumenten für Aussagen), Reziprozität (aufeinander Bezug nehmen), Respekt (Zivilität und Vermeidung von übermäßigem Antagonismus) und Konstruktivität (Problemlösungsorientierung, Kompromissfähigkeit) genannt (Esau et al. 2021). Eine Implementation solcher Kommunikationsnormen könnte einen entscheidenden Beitrag zur Steigerung der gesellschaftlichen Verständigungsfähigkeit leisten.

Das Badge-System sollte nicht als starres Entweder-oder-Modell konzipiert werden, sondern einem gestuften Ansatz folgen: Nutzer:innen können sich einzelnen Normen verpflichten, müssen jedoch nicht unbedingt alle umfassend implementieren. Wer sich etwa nur für die Sorgfaltspflichten zur Vermeidung, Desinformation zu verbreiten, ent-

scheidet, erhält eine bestimmte Form der Sicherheitserhöhung; das Gleiche gilt für die Implementierung deliberativer Kommunikationsnormen. Wer sich zu beidem verpflichtet, profitiert von einer kombinierten Sichtbarkeitssteigerung. Hierdurch wird den Nutzer:innen eine gewisse Flexibilität für ihre jeweiligen Kommunikationsbedürfnisse eingeräumt; gleichzeitig erlaubt es unterschiedlichen Akteur:innen, sich in angemessener Weise am Badge-System zu beteiligen.

Zugang, Compliance und Ausschluss zum und vom Badge-System

Beim Erwerb des Badges sollten Nutzende über die zugrunde liegenden Normen informiert werden, beispielsweise durch kurze Texte oder Erklärvideos. Mit ihnen ist zu verdeutlichen, wie Desinformation vermieden werden kann und wie deliberative Kommunikation im Plattformkontext aussieht. Dos und Don'ts werden klar benannt, ebenso die Konsequenzen bei Verstößen. Der Erwerb des Badges ist damit nicht nur ein technischer Vorgang, sondern auch eine Form der Nutzerbildung, die kommunikative Standards in den Vordergrund rückt. Weitere Hilfestellung zur Einhaltung der Normen könnten auch auf generativer KI basierte Chatbots und andere Tools bieten, die Nutzende bei der Formulierung von Inhalten unterstützen (Gelovani et al. 2025).

Um Missbrauch vorzubeugen, sollte der Zugang an bestimmte Kriterien gebunden sein. So könnte eine Verifizierung der Identität als reale Person oder Institution sinnvoll sein (siehe entsprechende Verifizierungsmechanismen auf anderen Plattformen wie Facebook, TikTok oder ehemals Twitter). Denkbar ist auch eine plattforminterne Prüfung, die das bisherige Kommunikationsverhalten oder auch andere plattformexterne Faktoren in die Bewertung einbezieht. Je stärker diese Prüfmechanismen ausgestaltet sind, desto robuster ist das Badge-System gegenüber Manipulation; allerdings steigt auch das Risiko unzulässiger Eingriffe in die individuellen Rechte abgelehnter Nutzer:innen.

Die Einhaltung der Kommunikationsnormen sollte regelmäßig überprüft werden. Bestehende Verfahren der Inhaltsmoderation könnten dazu aufgebaut und um zusätzliche Komponenten erweitert werden. Für den Bereich Desinformation existieren bereits automatisierte Erkennungssysteme und Modelle wie Trusted Flagger oder Notice-and-Takedown (Saurwein/Spencer-Smith 2020). Diese können genutzt werden, um Verstöße gegen die Normen wie beispielsweise die ungenügende Überprüfung des Wahrheitsgehalts verbreiteter Informationen zu identifizieren. Im Bereich deliberativer Kommunikation bieten sich algorithmische Verfahren an, die auf empirisch erprobten Indikatoren zur Messung deliberativer Qualitäten der Kommunikation basieren (Esau et al. 2021). Sinnvoll wäre es, wenn die Bewertung sich nicht auf einzelne Posts fokussiert, sondern das Kommunikationsverhalten über längere Zeiträume hinweg erfasst. So entsteht ein umfassenderes Bild des jeweiligen Accounts.

Bei Verstößen gegen die festgelegten Normen sollten diese in angemessener Intensität innerhalb eines abgestuften Sanktionssystems geahndet werden – von Warnungen über eine zeitlich begrenzte Suspendierung bis hin zum vollständigen und bleibenden Entzug des Badges und der damit verbundenen Vorteile. Den Vorgaben des DSA entsprechend sollte das Badge-System in ein angemessenes internes und externes Beschwerdesystem eingebunden werden, das transparente Nutzungsbedingungen und institutionalisierte Widerspruchsmöglichkeiten beinhaltet.

Integration des Badge-Systems als Risikominderungsmaßnahme nach Art. 34/35 DSA

Der Digital Services Act stellt einen regulatorischen Rahmen dar, um den in der Einleitung geschilderten Herausforderungen der digitalen Aufmerksamkeitsökonomie und der damit einhergehenden strukturellen Benachteiligung ziviler Kommunikation zu begegnen: Art. 34 verpflichtet sehr große Onlineplattformen dazu, systemische Risiken zu identifizieren, die aus ihren Funktionsweisen entstehen oder begünstigt werden. Art. 35 verpflichtet die Plattformen, wirksame Maßnahmen zur Risikominderung zu entwickeln und umzusetzen.

Das Badge-System stellt eine konkrete Maßnahme zur Risikominderung im Sinne von Art. 34 und 35 dar: Es versucht, eine neue Logik der Aufmerksamkeitsverteilung auf den Plattformen zu etablieren und somit die strukturellen Bedingungen für zivile Kommunikation zu verbessern. Es adressiert die aufgeführten Verzerrungen nicht durch Löschung oder Sperrung von Inhalten und Accounts, sondern durch eine gezielte Umsteuerung der Sichtbarkeitsverteilung. Wer sich verpflichtet, keine Des- oder Misinformation zu verbreiten und/oder sich auf deliberative Kommunikationsgrundsätze einzulassen, erhält einen strukturellen Vorteil: Inhalte solcher Nutzer:innen werden algorithmisch verstärkt. Damit soll ein neues Gleichgewicht geschaffen werden. Akteur:innen, die destruktive Kommunikationsstrategien verfolgen, behalten zwar grundsätzlich Zugang zur Plattform, verlieren aber ihren aktuellen strukturellen Vorteil im Wettbewerb um Aufmerksamkeit. Akteur:innen, die sich an den Normen ziviler Kommunikation orientieren, gewinnen an Reichweite. Hierdurch werden für Kommunikator:innen Anreize geschaffen, diesen Normen entsprechend zu kommunizieren; als Ergebnis könnte die Sichtbarkeit ziviler Kommunikation auf digitalen Plattformen insgesamt steigen.

Anders als repressive Eingriffe wie Löschung und Sperrung von Inhalten oder Accounts stellt das Badge-System eine grundrechtsschonende Intervention dar, die sich nicht unmittelbar negativ auf Grundrechte wie Informations- und Meinungsfreiheit auswirkt. Beachtet werden muss das Diskriminierungsverbot, wobei es hier zu betonen gilt, dass das Badge-System grundsätzlich allen Nutzenden (gegebenfalls mit wenigen Ausnahmen – siehe entsprechende Diskussion oben) offensteht. Entscheidend ist auch, dass das Badge-System nicht inhaltlich einem bestimmten poli-

tischen Lager zuzuordnen ist: Egal ob progressiv, liberal, konservativ, links, mittig oder rechts – das Badge-System fokussiert sich auf die Art und Weise, wie politisch kommuniziert wird und nicht auf die inhaltlich-ideologische Verordnungsung des Kommunizierten. Auch stellt sich die Frage, inwiefern dem Diskriminierungsverbot im Kontext der aktuellen Plattformarchitekturen, deren Aufmerksamkeitsökonomien dysfunktionale Kommunikation belohnen, überhaupt Genüge getan wird – oder ob eine kommunikative Chancengleichheit zwischen unzivilen und zivilen Beiträgen erst durch Mechanismen wie dem Badge-System zustande kommt.

Beispiele für eine solche Setzung und Förderung kommunikativer Standards und Normen bei Vermeidung inhaltlicher Diskriminierung finden sich zahlreich in der Geschichte und Gegenwart von Medien(selbst)regulierung, festgehalten in verschiedenen Kontexten wie dem Presserecht und dem Pressekodex sowie im Medienstaatsvertrag (ehemals Rundfunkstaatsvertrag).

Implementierung des Badge-Systems durch Plattformen: Zwischen Freiwilligkeit und regulatorischem Druck

Art. 35 Abs. 1 DSA stellt klar, dass es im Ausgangspunkt den Plattformbetreibern obliegt, angemessene, verhältnismäßige und wirksame Risikominderungsmaßnahmen auszuwählen, die auf die jeweils eigenen systemischen Risiken im Sinne der Art. 34 und 35 zugeschnitten sind. Das Badge-System könnte von den Plattformen zunächst freiwillig als eine solche Maßnahme erwogen werden. Eine unmittelbare Befugnis der EU-Kommission, die Einführung eines Badge-Systems verbindlich einzufordern, besteht hingegen nicht ohne Weiteres.

Allerdings nennt der DSA an mehreren Stellen Mechanismen, über die das Ermessen der Plattformbetreiber eingeschränkt wird und mittels derer die Kommission Einfluss auf die Auswahl oder Ausgestaltung von Maßnahmen wie dem Badge-System nehmen kann. So könnte die Kommission im Rahmen von Leitlinien nach Art. 35 Abs. 3 DSA die Einführung eines Badge-Systems empfehlen.

Darüber hinaus haben die Betreiber nach Art. 75 Abs. 2 DSA im Falle eines unzureichenden Risikomanagements einen Aktionsplan mit Risikominderungsmaßnahmen auszuarbeiten. Das Effizienzpotenzial dieser Maßnahmen – einschließlich der möglichen Implementierung eines Badge-Systems – wird wiederum von der Kommission bewertet. Wird ein Badge-System tatsächlich integriert, unterliegt auch dessen konkrete Ausgestaltung und Funktionalität einer Bewertung durch die Kommission im Hinblick auf dessen Risikominderungspotenzial. Sollten die Plattformen ihren Verpflichtungen im Risikomanagement nicht oder nur ungenügend nachkommen, kann die Kommission Strafen in Milliardenhöhe verhängen – allerdings immer in Bezug auf das gesamte Risikomanagement, nicht aber in Bezug auf eine einzelne Maßnahme wie dem Badge-System.

In der Regulierungspraxis bestehen somit jenseits einer rein freiwilligen Implementierung verschiedene Möglichkeiten, über Leitlinien, Evaluationsverfahren und Aufsichtsmechanismen Einfluss auf die Einführung eines Badge-Systems zu nehmen. Als einzelne Maßnahme können aber nicht verpflichtend vorgeschrieben werden.

Herausforderungen vom Badge-System

Auch wenn, wie oben dargelegt, grundsätzlich positive Effekte des Badge-Systems zu erwarten sind, lassen sich verschiedene Herausforderungen identifizieren, die es zu beachten gilt.

Zum einen sollte betont werden, dass politische Konflikte ein grundlegender Bestandteil pluralistischer Demokratien sind (Mouffe 2000). Sie machen Machtverhältnisse sichtbar und eröffnen die Möglichkeit, bestehende Ungerechtigkeiten herauszufordern. Die Forschung zeigt, dass marginalisierte Gruppen oft starke moralische Positionen und emotionale Ausdrucksformen nutzen, um Aufmerksamkeit für ihre Anliegen zu gewinnen und Gemeinschaften zu bilden. Polarisierung ist daher nicht grundsätzlich schädlich, sondern kann Ausdruck tiefer gesellschaftlicher Konflikte sein. Versuche, sie zu reduzieren, ohne ihre Ursachen anzugehen, bergen das Risiko, Probleme zu verdecken, anstatt sie zu lösen (Kreiss/McGregor 2024). Deshalb ist hervorzuheben, dass der Badge-Mechanismus politischen Konflikt nicht zu unterdrücken versucht, sondern vielmehr einen Rahmen schaffen will, in dem vielfältige Stimmen sichtbar bleiben, ohne dass extreme Polarisierung die demokratische Öffentlichkeit destabilisiert. Um gesellschaftliche Veränderungsprozesse zu ermöglichen, braucht es neben offenen Konflikt- und Konfrontationsphasen auch Möglichkeitsräume des Zusammenkommens und Zusammenführens. Digitale Öffentlichkeiten geben in ihrer aktuellen Verfassung fast ausschließlich Raum für Ersteres, Zweiteres findet dagegen im viel geringeren Maße statt oder gewinnt kaum Sichtbarkeit.

Ein weiterer zentraler Kritikpunkt ist, dass die Einführung eines Badge-Mechanismus Fragen nach der Definitionsmacht aufwirft. Wer bestimmt, was als Desinformation oder Beleidigung gilt, kann erheblichen Einfluss auf den politischen Diskurs ausüben. Diese Macht ist anfällig für Missbrauch und kann bestehende gesellschaftliche Machtverhältnisse und Bias fortschreiben (Herzog 2024; Kathirgalingam et al. 2024). Um der politischen und sozialen Kontextabhängigkeit dieser Definitionsfragen gerecht zu werden, muss das Badge-System in angemessene Checks and Balances eingebettet werden. Wie bereits oben dargelegt und vom DSA gefordert, sollte das Badge-System neben transparenten und nachvollziehbaren Nutzungsbedingungen ein angemessenes Widerspruchssystem beinhalten, um offenen Machtmissbrauch zu verhindern. Divers zusammengesetzte Entscheidungsgremien bezüglich der Ausgestaltung des Badges und eine dezentrale Struktur können dazu beitragen, Voreingenommenheit und Diskriminierung zu reduzieren, und sollten daher dringend in die

Entwicklung und Implementierung des Badges involviert werden (Herzog 2024; Kettemann/Schulz 2023).

Drittens ist darauf hinzuweisen, dass extremistische oder autoritäre Akteur:innen strategisch versuchen könnten, das Badge-System zu missbrauchen. Sie könnten sich an die formalen Regeln halten, um Sichtbarkeit zu erlangen, und gleichzeitig ihre Ideologien in subtiler Form weitertragen. Auch dieses Risiko ist nicht komplett auszuschließen. Allerdings zeigen verschiedene Studien immer wieder, dass gerade solche Akteur:innen unter anderem deswegen besonders erfolgreich im digitalen Raum sind, weil sie beispielsweise stark emotionalisierte und konfliktorientierte Sprache verwenden (Freistein et al. 2022). Diese Erfolgsstrategie würde abgeschwächt werden, wenn sie sich den deliberativen Grundsätzen des Badges verpflichten und infolge der mit ihrer destruktiven Kommunikationsstrategie verbundene Reichweitengewinn ausbleiben würde. Darüber hinaus könnten zusätzliche Prüfverfahren besonders problematische Akteur:innen im Vorfeld des Badges ausschließen (hierbei müssen mögliche Konflikte mit den Grundrechten dieser Nutzer:innen beachtet werden).

Nicht-Regulierung ist keine Alternative

Abschließend muss festgehalten werden, dass ein Vorschlag wie das Badge-System zwangsläufig auf Abwehr in verschiedenen Teilen des politischen Spektrums stoßen wird. Während es grundsätzlich gilt, die entsprechende Kritik ernst zu nehmen, ihren substanziellen Gehalt zu prüfen und, wo möglich, geeignete Gegenmaßnahmen zu implementieren (siehe oben), lohnt sich auch ein Blick auf die mutmaßliche Motivlage.

Eine wesentliche Strömung der Kritik wird sich aus einer grundsätzlichen Ablehnung der Hierarchisierung von Kommunikation und der damit potenziell einhergehenden Verfestigung von Machtstrukturen speisen. Auch denkbar ist eine grundsätzliche Skepsis gegenüber der Lenkung oder Governance des menschlichen Verhaltens mit Blick auf mögliche Einbußen an der individuellen Selbstbestimmung der Bürger:innen. Hierarchisierung und Lenkung finden allerdings ohnehin statt – gegenwärtig jedoch nach den Logiken der Aufmerksamkeitsökonomie gewinnorientierter Konzerne und ohne jegliche demokratische Legitimierung. Nichtstun schreibt also lediglich diesen Status quo der Machtstrukturen und der Einflussnahme fort – mit den genannten problematischen Implikationen für den öffentlichen Diskurs und die Demokratie. Auch gilt es hervorzuheben, dass funktionale epistemische Infrastrukturen zur Erzeugung, Überprüfung und Verteilung von Wissen eine notwendige Voraussetzung kollektiver Urteils- und Entscheidungsfähigkeit sind und ein zentrales Fundament demokratischer Systeme darstellen (Herzog 2024). Sie beinhalten zwangsläufig eine Ordnung und Hierarchisierung von Information und Kommunikation nach bestimmten Regeln. Medien(selbst)regulierung wie der Presserat, der Rundfunk- und nun der Medienstaatsvertrag haben genau diesem Zweck gedient (und dienen ihm weiter). Nun gilt es,

vergleichbare Lösungen für einen Kontext zu finden, der sich durch eine substanziell gewachsene Zahl an Kommunikator:innen, einen digitalen Informationsüberfluss sowie nachgewiesene Einflussversuche ausländischer und inländischer demokratiefeindlicher Akteur:innen auszeichnet.

Eine weitere wesentliche Gruppe in Opposition zu einem Vorschlag wie dem Badge-Mechanismus werden insbesondere jene Akteur:innen sein, deren politisches (und mitunter auch ökonomisches) Geschäftsmodell im Wesentlichen im Erzeugen von Konflikt, Wut und Angst besteht und die dieses gefährdet sehen. Kritik aus dieser Richtung ist nicht grundsätzlich abzulehnen – wie bereits ausgeführt, erfüllt politischer Konflikt eine wesentliche Funktion in demokratischen Systemen und darf nicht per se unterdrückt werden. Es gilt jedoch erneut zu betonen: Das Badge-System schafft einen Rahmen, in dem vielfältige Stimmen sichtbar bleiben, aber extreme Polarisierung die demokratische Öffentlichkeit nicht gefährdet. Gleichzeitig gilt es, nicht aus den Augen zu verlieren, dass solche „Polarisierungsunternehmer:innen“ (Mau et al. 2024) nicht selten substanzielle autoritäre Tendenzen aufweisen (oder sogar offen demokratiefeindlich agieren) und ihnen ihr eigenes (politisches) Wohlergehen im Zweifelsfall wichtiger ist als das der Menschen, die sie zu vertreten vorgeben – oder gar als das Wohl der demokratischen Systeme insgesamt, in denen sie aktiv sind. Eine wehrhafte Demokratie sollte sich diesem Druck nicht beugen, sondern ihre epistemischen Infrastrukturen verteidigen und weiterentwickeln – unter Beachtung der oben aufgeführten Kritikpunkte und der Implementierung angemessener Checks and Balances.

Fazit

- Plattform-Badges können ein wirksames Instrument im Sinne von Art. 34, 35 DSA sein, um systemische Risiken wie die Verbreitung von Desinformationen und die Überpräsentation von emotionalen und konfliktorientierten Inhalten im Kontext von VLOPs einzudämmen.
- Diese Risiken stellen, auch vor dem Hintergrund ausländischer Informations- und Manipulationskampagnen sowie strategischer Polarisierungsunternehmer:innen, eine erhebliche Belastung für die Stabilität und die innere wie äußere Sicherheit europäischer Demokratien dar.
- Die Badges haben das Ziel, eine neue Logik der Aufmerksamkeitsverteilung auf digitalen Plattformen zu etablieren mit der Absicht, die strukturelle Bevorzugung problematischer Onlinekommunikation aufzubrechen und die strukturelle Benachteiligung wünschenswerter Onlinekommunikation abzuschwächen. Hierdurch kann die Resilienz europäischer Demokratien ausgebaut und gestärkt werden.
- Eine Zurückhaltung in der Umsetzung des DSA und der damit verbundenen Regulierung von VLOPs aufgrund von politischen Erwägungen in der Handels- oder Sicherheitspolitik ist vor dem Hintergrund der geschilderten

Herausforderungen als problematisch anzusehen und sollte vermieden werden.

- Plattform-Badges stellen eine grundrechtsfreundliche Art der Plattform-Governance dar. Informations- und Meinungsfreiheit werden geschützt, es findet keine Diskriminierung von inhaltlich-ideologischen Positionen statt und Missbrauch wird durch eine Einbettung der Badges in angemessene Checks and Balances vorgebeugt.
- Um die Effektivität von solchen und vergleichbaren Maßnahmen besser zu verstehen, sollte entsprechende Forschung ausgebaut und gefördert werden. Eine mögliche Einführung der Plattform-Badges sollte wissenschaftlich begleitet werden.

Anmerkung

Eine Langversion dieses Beitrages findet sich [hier](#).

Autor

Jan Rau ist wissenschaftlicher Mitarbeiter am Forschungsinstitut Gesellschaftlicher Zusammenhalt und Leibniz-Institut für Medienforschung/Hans-Bredow-Institut (HBI). Er beschäftigt sich mit der Transformation von Politik und Gesellschaft durch digitale Medien und der Frage, wie dieser Prozess aktiv gestaltet werden kann. Seine Schwerpunkte umfassen unter anderem Polarisierung, Rechtsextremismus, Desinformationen und Plattform-Governance. Er hat am Oxford Internet Institute studiert und Forschungsaufenthalte an der Princeton University, der Vrije Universiteit Amsterdam und der New York University absolviert.

Literatur

- Bakir, V.; McStay, A. (2018):** Fake News and The Economy of Emotions: Problems, Causes, Solutions, in: Digital Journalism 6 (2), S. 154–175, <https://doi.org/10.1080/21670811.2017.1345645> (24.1.2026).
- Brady, W. J.; Jackson, J. C.; Lindström, B.; Crockett, M. J. (2023):** Algorithm-Mediated Social Learning in Online Social Networks, in: Trends in Cognitive Sciences 27 (10), S. 947–960, <https://doi.org/10.1016/j.tics.2023.06.008> (24.1.2026).
- Esau, K.; Fleuß, D.; Nienhaus, S. (2021):** Different Arenas, Different Deliberative Quality? Using a Systemic Framework to Evaluate Online Deliberation on Immigration Policy in Germany, in: Policy & Internet 13 (1), S. 86–112, <https://doi.org/10.1002/poi3.232> (24.1.2026).
- Freistein, K.; Gadinger, F.; Unrau, C. (2022):** It Just Feels Right: Visuality and Emotion Norms in Right-Wing Populist Storytelling, in: International Political Sociology 16 (4), olac017. <https://doi.org/10.1093/ips/olac017> (24.1.2026).
- Gelovani, S.; Müller, P.; Meyer, H.; Stoll, A.; Pröschel, L.; Ziegele, M.; Wesseler, H. (2025):** Enhancing Democratic Listening in Online Discussions: The Role of Reaction Buttons: Discussion Moderation, and Large Language Models, 75th Annual Conference of the International Communication Association (ICA).

Grüning, D. J.; Kamin, J.; Panizza, F.; Katsaros, M.; Lorenz-Spreen, P. (2024): A Framework for Promoting Online Prosocial Behavior via Digital Interventions, in: Communications Psychology 2 (1), S. 1–10, <https://doi.org/10.1038/s44271-023-00052-7> (24.1.2026).

Herzog, L. (2024): Citizen Knowledge: Markets, Experts, and the Infrastructure of Democracy, Oxford, <https://doi.org/10.1093/oso/9780197681718.001.0001> (24.1.2026).

Kathirgalingam, A.; Lind, F.; Bernhard, J.; Boomgaarden, H. G. (2024): Agree to Disagree? Human and LLM Coder Bias for Constructs of Marginalization, <https://doi.org/10.31235/osf.io/agpyr> (24.1.2026).

Kettemann, M. C.; Schulz, W. (2023): Platform://Democracy: Perspectives on Platform Power, Public Values and the Potential of Social Media Councils, <https://doi.org/10.21241/SSOAR.86524> (24.1.2026).

Kreiss, D.; McGregor, S. C. (2024): A Review and Provocation: On Polarization and Platforms, in: New Media & Society 26 (1), S. 556–579, <https://doi.org/10.1177/14614448231161880> (24.1.2026).

Lorenz-Spreen, P.; Oswald, L.; Lewandowsky, S.; Hertwig, R. (2023): A Systematic Review of Worldwide Causal and Correlational Evidence on Digital Media and Democracy, in: Nature Human Behaviour 7 (1), S. 74–101, <https://doi.org/10.1038/s41562-022-01460-1> (24.1.2026).

Mau, S.; Westheuser, L.; Lux, T. (2024): Triggerpunkte: Konsens und Konflikt in der Gegenwartsgesellschaft (7. Auflage), Berlin.

McLoughlin, K. L.; Brady, W. J.; Goolsbee, A.; Kaiser, B.; Klonick, K.; Crockett, M. J. (2024): Misinformation Exploits Outrage to Spread Online, in: Science 386 (6725), S. 991–996, <https://doi.org/10.1126/science.adl2829> (24.1.2026).

Mouffe, C. (2000): The Democratic Paradox, London/New York.

Presserat (German Press Council) (2017): German Press Code, <https://www.presserat.de/files/presserat/dokumente/download/Press%20Code.pdf> (24.1.2026).

Saurwein, F.; Spencer-Smith, C. (2020): Combating Disinformation on Social Media: Multilevel Governance and Distributed Accountability in Europe, in: Digital Journalism 8 (6), S. 820–841, <https://doi.org/10.1080/21670811.2020.1765401> (24.1.2026).

Schroeder, R. (2018): Social Theory after the Internet: Media, Technology, and Globalization, UCL Press, <http://www.knowledge-unlatched.org/ku-select-2017-books/> (24.1.2026).

Webster, J. G. (2014): The Marketplace of Attention: How Audiences Take Shape in a Digital Age, Cambridge.

Yeung, K. (2017): „Hypernudge“: Big Data as a Mode of Regulation by Design, in: Information, Communication & Society 20 (1), S. 118–136, <https://doi.org/10.1080/1369118X.2016.1186713> (24.1.2026).

Impressum

Herausgeberin

Friedrich-Ebert-Stiftung e.V.
Godesberger Allee 149
53175 Bonn
info@fes.de

Herausgebende Abteilung

Abteilung Analyse, Planung und Beratung
www.fes.de/apb

Kontakt

Dr. Johannes Crückeberg
johannes.crueckeberg@fes.de

Jonathan Menge
jonathan.menge@fes.de

Bildnachweis

Seite 1 oben: picture alliance / Ikon Images | Stuart Kinlough

Die in dieser Publikation zum Ausdruck gebrachten Ansichten sind nicht notwendigerweise die der Friedrich-Ebert-Stiftung e.V. (FES). Eine gewerbliche Nutzung der von der FES herausgegebenen Medien ist ohne schriftliche Zustimmung durch die FES nicht gestattet. Publikationen der FES dürfen nicht für Wahlkampfzwecke verwendet werden.

Februar 2026
© Friedrich-Ebert-Stiftung e.V.

Weitere Publikationen der Friedrich-Ebert-Stiftung finden Sie hier:
www.fes.de/publikationen

