

*CONVERGENCE COLLABORATIVE ON DIGITAL
DISCOURSE FOR A THRIVING DEMOCRACY AND
RESILIENT COMMUNITIES*

Discovery Report

JANUARY 2023

CONVERGENCE



TABLE OF CONTENTS

- 
- 
- 01.** **Introduction**
The Stakes
 - 02.** **Overview of the Landscape**
The Context
The Challenges
 - 03.** **Convergence Discovery and Design**
The Approach
The Issues
The Frame
 - 04.** **Opportunities for Impact**
Potential Pathways to Consensus Solutions
Conclusion
 - 05.** **About Convergence**
 - 06.** **Acknowledgements**

The Stakes

The internet, social media, and digital technology are fixtures in our daily reality; influencing how we work, play, communicate, relate, consume, produce, and disseminate information. Online discourse and information are increasingly central to our social and civic life. They have exponentially expanded access to knowledge, culture, and communication (as of January 2023, the World Wide Web contains at least [3.24 billion pages](#)). The internet has in many ways been an inherently democratizing medium for digital discourse with the potential to facilitate a knowledgeable public, robust debate, vocal dissent, and connected communities.

However, as with any industrial or technological innovation, there is a dark side. This same internet has been leveraged into a tool for disinformation and misinformation, psychological and behavioral manipulation, polarization, radicalization, surveillance, and addiction. We experience these effects daily as feelings of isolation, inability to focus, depreciated social skills, and susceptibility to false or manipulated information. We have seen how these digital harms can [rip families apart, intensify social conflict, or distract the public from important issues so they remain unresolved](#). This is especially dangerous during critical periods such as elections and public health crises, when the spread of false or discredited science can [negatively impact people's health behaviors](#).

Within a system designed to capture and monetize our attention and against the backdrop of an increasingly polarized nation, the way we use information and communication technologies has resulted in cascading harms. While experts believe that digital life will continue to expand people's boundaries and opportunities in the next decade, [nearly a third](#) believe that it will do more harm than good to people's wellbeing. [Almost 90%](#) of Americans agree that social media enables the spread of misinformation, extremism, harassment, and other types of harmful content. Within that group, Black Americans and women expressed higher concern about hurtful content compared to white Americans and men, respectively. Meanwhile, [71% of Americans](#) think that the internet does more to divide us than bring us together, triggering distrust and isolation along racial, religious, socioeconomic, partisan, and geographic lines, and [49% of experts](#) believe digital disruption will hurt democracy, compared to 33% that say it will strengthen it and 18% that think there will be no significant change.

Some believe that the central problem with digital discourse today is the threat of censorship, both formal and informal. While public shaming has existed for much of human history, “cancel culture” has become a deeply debated idea in the country’s political discourse – some see it as a call for accountability, others see it as censorship or unjust punishment to instill fear, and still others claim it does not exist at all. There are partisan and ideological differences in how Americans define “cancel culture;” conservative Republicans were more likely than other groups to see it as a form of censorship than more moderate or liberal Republicans and all Democrats.

The stakes are high. Digital platforms increasingly mediate discourse, from helping users join social groups to stay up to date on current events. The spread of false, divisive, and hurtful digital information on one hand, and over-moderation, censorship, and cancel culture on the other negatively affect our individual wellbeing, tear at our community and social fabric, skew markets, and threaten our democracy by impeding our ability to stay informed, debate, and work together to solve important problems.



Almost 90% of Americans agree that social media enables the spread of misinformation, extremism, harassment, and other types of harmful content.

The Knight Foundation



71% of Americans think that the internet does more to divide us than bring us together, triggering distrust and isolation along racial, religious, socioeconomic, partisan, and geographic lines.

The Knight Foundation

Digital Discourse: It comprises many things – useful, entertaining, and inane information; disinformation, misinformation, and other forms of media manipulation; debate and dissent; satire and parody; oddities, ephemera, GIFs, and memes.

Disinformation: False information created to intentionally mislead people.

Misinformation: False information that is mistakenly or unknowingly shared to the same effect.

Polarization: The splitting of society into distinct subgroups at different ends of a spectrum, for example in terms of their ideological beliefs or social circumstances. These groups have a positive bias towards their in-group and a negative one towards their out-group.

Censorship: The suppression of information and speech (words, images, ideas, artistic expression).

Cancel Culture: A mass exertion of social pressure to remove status or esteem from a person, place, or thing.

The Context

It is important to consider several contextual elements when thinking about how to best frame problems related to toxic digital discourse.

Institutional and social trust is low

No institution is perceived as honest by more than [25% of Americans](#), indicating that institutions might benefit from some introspection. Less than [40% of Americans](#) feel most people can be trusted. Although Americans distrust internet companies to make the right decisions related to harmful content online, [55% of them still favor the technology platforms to set their own policies over government regulation](#). Government decision-making often [centers advice from elites in academia and the corporate sector at the expense of community expertise](#).

Rooted in systemic and historical socioeconomic issues

Propaganda and public shaming predate the internet, and modern disinformation and cancel culture are [symptomatic outgrowths of a series of socioeconomic problems](#) that play out along demographic divides, such as race, gender, ideology, geography, and class, especially where there is poorer access to broadband. Digital spaces and digital life will continue to be shaped by systemic issues, such as socioeconomic inequality, racism, declining trust in institutions, and partisan culture wars. These forces hamper citizens' ability to engage in productive dialogue in digital spaces. Even if we cannot address these causes within the scope of this project, it is important to increase understanding of the issues as being systemic, rather than individual.

Disproportionate effect on marginalized communities

It is important to underscore that [historically marginalized groups](#) are more likely to be both disproportionately impacted by disinformation and censored by efforts to moderate disinformation. Some groups brought up by interviewees include Black and Latinx communities, the LGBTQ+ community, religious minorities, activists, and sex workers.

Top-down and bottom-up sources

There are both [top-down and bottom-up sources of disinformation and radicalized discourse](#). On one hand, [a lot of false, inflammatory, and sensational speech comes from politicians and their weaponization of social media platforms for political gain](#); on the other, many workers on the frontline of addressing disinformation and extremism have observed the danger of grassroots, organic spread driven by individuals and online groups.

Technology can be used for good or ill

We must be wary of grounding digital discourse challenges in technological determinism – **technology itself is not inherently good or evil, but a tool that can be used for positive or negative outcomes**. We need to work not only to prevent the detrimental impacts of digital harms, but also to understand and articulate a positive vision for how digital and networked technologies can be used to promote positive social discourse, empathy, and connection. [Wikipedia](#), for example, offers a unique view into a collaborative space where people across the globe follow a shared set of norms as to what constitutes truth.

Dis- and misinformation often center on "wedge" issues

In the U.S., disinformation and misinformation often center on **"wedge" issues**, such as elections, vaccines, and the January 6 attack on the Capitol.

Cognitive, social, and affective forces

People's consumption, dissemination, and production of information, and associated decision-making, often support more divisive, false, or sensational content due to cognitive, social, and affective forces. The human brain has a finite capacity for processing and retaining information, so it often makes decisions about what to remember driven by confirmation bias and motivated reasoning. [Socio-affective factors include whether information comes from an elite, in-group, or out-group, evokes an emotion, or aligns with your existing views or partisan identity](#). Cognitive factors include lack of analytical thinking and deliberation. Furthermore, when engaging with others digitally, people often experience [online disinhibition](#) – they self-disclose and engage more frequently or intensely than they would in-person.

Attitudes are often linked to a partisan binary

Americans' attitudes towards online content are commonly linked to a traditional partisan binary; [though political leaning matters, their views on this issue are more strongly connected to their views on free expression and privacy](#).

Social media accelerates reach and scale of disinformation in a unique way

It should be noted that [television remains the most popular platform for the consumption of news and information at 49%](#), but social media and online news outlets are close at 43%. Some studies point to [for-profit](#) television news and radio as major contributors to disinformation. Though media-driven dissemination of falsehoods and manipulations pre-date the internet – what social media uniquely contributes is the acceleration, reach, and scale of inflammatory and false information.

This challenge is global

Though this Discovery Report primarily focuses on digital discourse challenges in the context of the U.S., they are not unique to the country and manifest around the globe.

The Challenges

Increasing use of the internet continues to bring threats along with benefits, making efforts to mitigate digital harms a daunting and ever-evolving task. Several challenges make this a difficult problem to tackle.

01

Risk of restricting free speech or eroding the social benefits of access to information and network connectivity.

In addressing digital discourse challenges, we must be wary of unduly restricting free speech or eroding the social benefits of access to information and network connectivity. At the same time, though the First Amendment protects people against government limits on freedom of expression, it does not prevent

Free reach: Algorithmic or other amplification of information and speech.

a private company from setting its own rules. Even though the internet is primarily privately-owned, we see and treat it as a digital public square. This means we contend with concerns and questions related to freedom of speech governed not by the First Amendment, but by moderation standards set by technology platforms. There is also debate about differentiating **free speech** and **free reach** – some experts hold that while Americans have a right to free speech, they do not necessarily have that same right to limitless amplification of that speech.

02

Defining and scoping the issue is difficult

Defining and scoping the issue is difficult because (1) it is not an exclusively digital phenomenon; people's information ecosystems are being shaped by the decline of local news, rise of news deserts and information vacuums, decades of hyper-partisan talk radio and cable news, and **growth of partisan messaging masquerading as local news**; (2) it is hard to consistently establish intentionality, causality, or a threshold for harm when it comes to harmful information, and people have different views on what is the harm most needed to be stopped based on our politics and because of First Amendment protections; (3) there is crossover in **how different types of content are defined** (for example, disinformation and misinformation, satire and parody, misleading and decontextualized content, manipulated content such as deepfakes, ephemeral content, conspiracy theories, extremist content, hate speech); (4) it is not binary (fact vs fiction, right vs wrong); rather, all information, including disinformation and misinformation, is contextual. All these factors make it **difficult to identify, classify, and counteract toxic digital discourse**.

The background of the page features a dark, moody image of blue fiber optic cables. The cables are bundled together and run horizontally across the lower half of the page, with some individual strands visible. The lighting is dim, highlighting the texture and color of the cables against a dark background.

03

Digital technologies enable disinformation to thrive

At the same time, digital technologies create an environment that allows disinformation to (1) **spread further and faster** across interconnected channels and platforms; (2) **reach more people**; and (3) be **segmented and micro-targeted** to narrow groups and profiles.

04

Epistemological crisis

American society is confronting an epistemic crisis around how to define “truth.” We are increasingly disagreeing about epistemology -- not necessarily just about what is true, but about how we know whether something is true.

05

There is a fundamental disagreement between parties

Although there seems to be some bipartisan convergence around the need for platform regulation, it is for differing reasons. There is a fundamental disagreement between parties on what should and should not be allowed on social media, and varying levels of concern about content moderation versus censorship. The [collapse of the disinformation board proposed by the U.S. Department of Homeland Security](#) is just one example that illustrates how deeply partisan discourse and disinformation problems have become.

06

Disagreement on whether false and harmful content is fundamentally an issue of supply or demand

There is disagreement across sectors and ideologies as to whether false and harmful content is fundamentally an issue of **supply or demand**. Some blame social media companies, hyper-partisan media, politicians, and disinformation actors for polluting the information environment with conspiratorial, divisive, and false content that lacks credibility and trustworthiness. Others see this as a fundamentally human problem, that no matter how much you increase the supply of accurate information, there is a deep hunger amongst people for content that confirms their biases, stokes their anger, or offers a scapegoat for their fears.

The Approach

In the Discovery and Design phase, Convergence explored how our digital information environment and digital phenomena (such as disinformation and misinformation) affect the production and consumption of information, and in turn, people and communities.

The following questions drove our Discovery and Design process:

- ➔ What are the **core problems and challenges** related to digital discourse?
- ➔ What are **high-potential solutions** that could mitigate digital harms without infringing on free speech?
- ➔ What are the **key barriers to consensus and action** to address digital discourse challenges?
- ➔ Who are the **key stakeholders** in the same that need to be brought together to build consensus and catalyze action?

Our Discovery and Design process consisted of over 200 stakeholder interviews, accompanied by a literature review, the tracking of almost 600 pieces of state and federal legislation, three multistakeholder workshops convening a subset of experts and practitioners for discussion, and mapping existing research and activities in the space, including but not limited to the [final report](#) from Aspen Institute's Commission on Information Disorder, the [solutions summary](#) from MIT's Social Media Summit, the Heritage Foundations [roadmap](#), and the [insights report](#) and [synthesized list](#) of immediate bipartisan fixes from the German Marshall Fund.

We interviewed experts, practitioners, stakeholders, and individuals representing diverse viewpoints and experiences related to digital discourse. This includes representatives from technology platforms, advertising networks, government and policy think tanks, national and local media organizations, academia and education, civil society, science and healthcare communities, and individuals personally affected by disinformation and censorship. This allowed us to identify and map the essential challenges and assemble a diverse table.

Each of these experts offer their views in a personal capacity and not on behalf of their organizational affiliations.

CONVERGENCE DISCOVERY AND DESIGN

Following this process and leveraging the insights in this Discovery Report, Convergence mobilized a uniquely diverse and powerful model to convene a multistakeholder dialogue. Members are collectively forging and testing breakthrough consensus solutions that foster digital discourse for a thriving democracy and resilient communities. Convergence will work with participants to drive implementation of their consensus solutions and impart meaningful change.

The Issues

To add value to the national conversation and existing efforts, a multi-stakeholder dialogue on digital discourse could go in several directions. There are diverse and, at times, conflicting ideas about how to tackle issues around digital consumption of information. We have heard stakeholders mention the need to address the following specific issues within the broader landscape of digital discourse challenges:

01

Content moderation and curation

This means drawing lines that define what content is allowed, and what is not. Technology platforms, media platforms, and news outlets each have their own content moderation policies and there are bipartisan calls for reforming [Section 230](#), albeit with different underlying agendas. There is agreement among experts that the Section 230 debate has been mostly performative, unproductive, and riddled with unachievable or [harmful proposals](#). Technology platforms can provide more resources and support to moderators, and they can engage librarians to build a content curation model that does not rely so heavily on reactionary moderation. At the same time, there are concerns that focusing on content moderation solutions will incentivize censorship and will be difficult to scale, especially in languages other than English.

02

Technical design and algorithmic amplification

Platforms are designed to privilege content that fosters engagement, gets users to stay on the platform, and makes a profit—this leads to the promotion of posts that generate higher levels of outrage, polarization, and disinformation. Yet digital spaces could be programmed to promote healthier outcomes. Tech platforms can create recommendation systems that open new spaces rather than closing users into their existing preferences and biases. They can use sentiment analysis to deemphasize rage-inducing content and encourage civil discourse, or present posts chronologically. They can also offer users more transparency

and control over their time, privacy, data, and content feeds, and offer researchers more data about how disinformation spreads. [Middleware solutions](#) could unbundle platforms from recommendation algorithms and provide third party options for content prioritization. Policymakers have proposed legislation that regulates product design features, such as the previously proposed [Filter Bubble Transparency Act](#) and [Justice Against Malicious Algorithms Act](#). However, because algorithms are written by humans, some stakeholders worry that regulating algorithms could violate free speech rights.

03

Harmful business models

Technology platforms' business models are ad-based – they rely upon tracking, coupled with socio-psychometric profiling algorithms that engage and elicit an emotional response or action, such as buying into a system of beliefs or buying a product or service. Companies elevate information that optimizes time spent on the platform and segments users for advertisers, at a speed and volume that overwhelms people's cognitive defenses. This model can exacerbate inflammatory responses and be used in discriminatory ways.

Section 230:

Distinguishes interactive websites from publishers, establishing that sites cannot be held legally responsible for problematic user-generated content and cannot be sued for good-faith moderation of user-generated content.

Justice Against Malicious Algorithms Act:

Would amend Section 230 to limit liability protections when websites know their algorithms recklessly make personalized recommendations.

General Data Protection Regulation:

A European Union (EU) privacy and security law that applies to any organization that collects data related to people in the EU, whether that organization is based in the EU or not.

Filter Bubble Transparency Act:

Would require internet platforms to give users the option to engage with their site without content curated by algorithms.

Honest Ads Act

Would enhance transparency and accountability for online political advertisements by requiring purchasers and publishers to publicly disclose information about them.

Digital Services Act:

A European Union law that sets out obligations for internet and online platforms that promote transparency, innovation, and user protections.

California Consumer Privacy Act:

A 2019 California law that established privacy protections for consumers and requires websites to provide users with more transparency and control over their data.

At the same time, the models can be used in positive ways, for convenience, and for connectedness. For example, they can be leveraged during campaigns to encourage voting, to urge people to get vaccinated, or simply to provide personalized desirable movie recommendations. Technology platforms can take steps to validate and add transparency to their ads, while policies like the [Honest Ads Act](#) and proposed Section 230 amendments could hold them more accountable for the content or purpose of paid ads. Business models also shape online communities in how they manage content and behavior, and this is dependent on what they can afford to do – and sometimes what is healthiest for people and society is not the most profitable.

04

Extractive data collection

Data privacy may offer a more precise solution. By limiting access to the information that enables personalized ad targeting and polarization loops, data privacy laws could render disinformation a weapon without a target. Policy solutions can build on existing, if imperfect, privacy legislation such as the [California Consumer Privacy Act](#), and the European Union's [General Data Protection Regulation](#) and [Digital Services Act](#). However, if not implemented thoughtfully, this could impede the provision of personalized products and services that people want and pay for.

05

Lack of meaningful competition

Consumers lack meaningful options to find, send, and receive information online. A competition policy agenda includes stronger oversight of mergers and acquisitions, antitrust reform, and data portability and interoperability between services. Skeptics say true monopolies rarely occur or are transient, and that competition policies could impede innovation or spread disinformation and other false or divisive content across a greater number of sites.

Data portability: The ability to move content from one website or service to another.

Data interoperability: The ability for two separate websites or services to share data with one another.

06

Declining local news

Local newspapers [produce half of original reporting in the U.S., despite only accounting for 25% of media outlets](#). They are crucial in moments of crisis –[readership increased during the COVID-19 pandemic](#), illustrating how valuable local news was to stay informed about community responses. They are critical community infrastructure, like a library, and a core democracy-supporting force. However, local publishers are facing financial collapse and there is a divide in American journalism between [for-profit and nonprofit business models](#). Business model innovations, legislation, and a mix of commercial, philanthropic, and hybrid investment could restore the supply of local news. [Local news drives business outcomes](#)

for advertisers by building incremental audiences, providing more connected and personalized experiences, and driving direct response to advertisements. Technology platforms can also continue taking steps to prioritize local and trusted sources in their rankings. At the same time, there is a disconnect between local newsrooms and right-leaning audiences that needs to be addressed. Similarly, there are regions in the U.S. where even local news has become more polarized.

Local News Lab: Serving as a hub for creative experiments in journalism sustainability, Local News Lab has a portfolio of innovative ideas and business models that support new kinds of funding and collaboration between journalists, newsrooms, industry, and communities.

07

Lack of effective consumer education

Users continue to share false narratives, whether passively or for partisan purposes. Readers also confuse news and opinion sections of media coverage, a line the media itself has helped blur. This makes a case for a more mature digital and media literacy curriculum that better teaches consumers the architecture, dynamics, and logic of the online information ecosystem and builds their capacities to critically engage with and share content. An educated consumer can also take on a more involved role in lobbying for their digital rights or joining the next generation of policymakers, technologists, and business professionals who design and determine how technologies are used. There is an opportunity to integrate psychological strengthening approaches grounded in our understanding of the cognitive mechanisms and social determinants of information and misinformation processing.

08

Threats to national security

This extends beyond cybersecurity to include foreign influence operations, distrust in elections, rise of extremism, and threats to public health and safety. Though there is some evidence of foreign interference, this is too narrow a focus – there are many actors domestically who leverage technology platforms to disrupt and undermine the integrity of digital discourse for various reasons.

09

Intergroup dynamics of disinformation

Disinformation builds value- and identity-based narratives that pit communities against one another. People often feel a strong pull toward ingroup solidarity and outgroup conflict when consuming, producing, and distributing information. People are also highly resistant to information that they perceive as threatening their identity. These dynamics influence how people assess content.

The Frame

Our Discovery process distilled the multifaceted and complex landscape of issues outlined above into five specific problem statements that could help us determine the best frame or on-ramp, for our consensus-building Collaborative. Each slice of the problem carries with it underlying assumptions, focuses on different facets of digital discourse challenges, and implies a different set of solutions.

Problem Statements

01

Digital discourse can be framed as an information problem. The focus is on the nature, type, and quality of information itself, on increasing the supply of trustworthy content, and on decreasing the supply of disinformation, misinformation, and other misleading or harmful content. From this lens, solutions encompass fact-checking, labelling, content moderation, and take-downs, providing knowledge products and hubs, and rebuilding local media. However, perceived credibility is important – people generally are not open to information that they personally do not believe to be credible.

02

Digital discourse can be framed as a data and technology problem. The focus is on the data-driven and technological dimensions of harmful digital information, such as platforms and algorithms, as both drivers and fixes of the problems. Solutions include technological innovations and adaptations, as well as underlying business models and pertaining regulatory frameworks. The limitation here is that algorithms and the social media companies behind them are not exclusively to blame for digital harms – for example, disinformation was found to spread virally on encrypted messaging applications such as WhatsApp. Additionally, propaganda and polarization pre-date the internet.

03

Digital discourse can be framed as an education and skills problem. The focus is on building the skills of individuals, civil society, journalists, and other stakeholders (as producers, disseminators, and consumers of information) so they can navigate and engage in the online information ecosystem in responsible, safe, and ethical ways. Critics of slicing the problem this way say it over-individualizes the problem and does not consider the various ways in which different communities, people, and disciplines construct knowledge. However, more nuanced approaches to education and skill-building can focus on raising awareness of our psycho-social and cognitive vulnerabilities, such as our tendency to form us versus them groups and cognitive biases such as

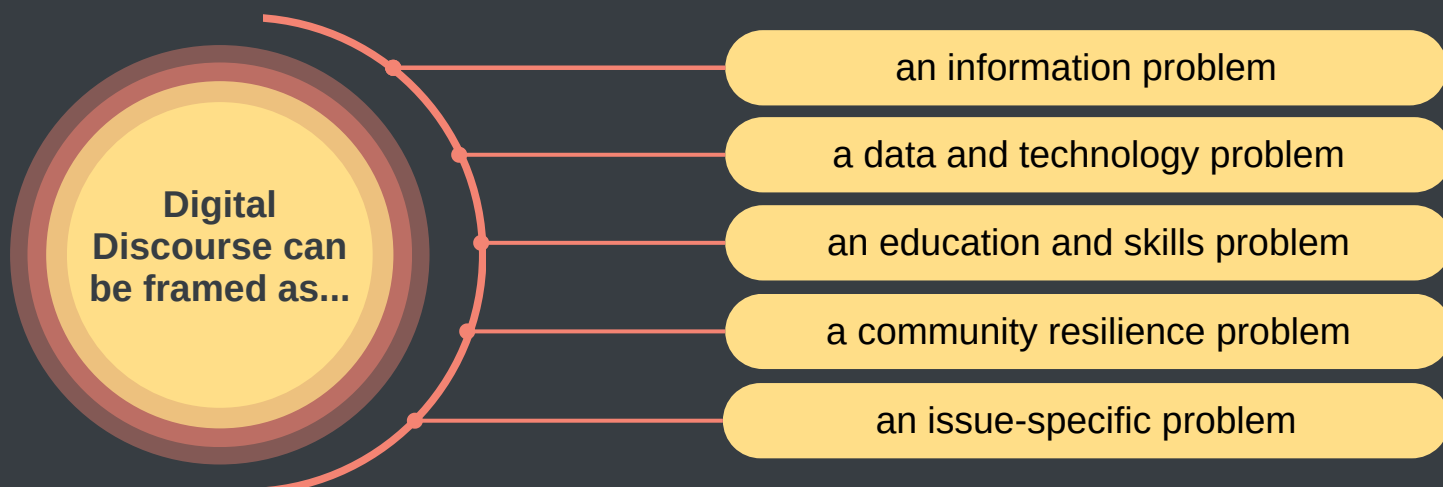
confirmation bias. There is a growing body of evidence on cognitive resilience, illustrating how pre-bunking (pre-emptive) and de-bunking (reactive) interventions can be impactful at different points in time, and how minds can be inoculated against misinformation, climate denial, and conspiracy theories.

04

Digital discourse can be framed as a community resilience problem. 70% of people across all demographic groups use social media platforms primarily to connect with friends, family, and community, and people who share their interest or views. More and more individuals are interacting with various communities through digital platforms. This requires thinking about the relational infrastructure of our information ecosystem, in addition to the technological infrastructure – namely, do our online spaces promote shared identity, mutual goals, equitable distribution of power, and frequent interactions amongst community members. Looking at toxic digital discourse through this lens puts an emphasis on the psycho-social elements of problem, and the broader trends of increasing polarization and societal divides, declining social and institutional trust, and eroding social cohesion. It also requires acknowledging that all information is contextual – for example, false narratives manifest differently in different communities. Solutions from this perspective could involve bridge-building within and between online and offline communities, working with trusted messengers, partnering with community-based organizations, and de-polarizing news, politics, and social media. There is also a crisis of loneliness and isolation, and other issues related to digital wellbeing and mental health. Lastly, in addition to deprogramming individuals who are already entrenched in extremist communities, we can also intervene to stop others from going down that path.

05

Digital discourse can be framed as an issue-specific problem. This frame focuses on the thematic dimensions of the most divisive digital discourse, such as elections, public health, or climate change. Solutions encompass specific counter-messaging strategies and partnerships with relevant stakeholders in the field.



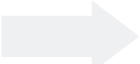
From these conceptualizations of the problem, stakeholders coalesced around a **multifaceted consumer/community-oriented approach** to digital discourse challenges that:



Considers the **affective, psychological, and social drivers** of digital disinformation, fear-driven hate, and polarization

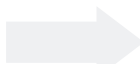


Focuses on **reimagining education and skill-building** of all stakeholders and **individual inoculation and community resilience** to digital harms

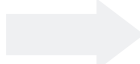


Has strong **cross-links to policy and industry interventions** that help as the consumer/community approaches scale, grapples with the challenging realities of content moderation, architects better online (and offline) communities, and that address issues of institutional accountability and power

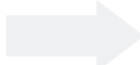
There is potential to explore a set of solutions that focus on fostering trust and belonging in digital and place-based communities. Emerging areas for exploration of consensus recommendations include:



Online communities: How do we engage communities and stakeholders in building online spaces that encourage healthy digital discourse? How do we ensure these ecosystems meet their needs for trustworthy information, belonging, and digital wellbeing?



Education and skills: Beyond just increasing interactions with trustworthy information, how do we promote more constructive online interactions and cooperative discourse? How do we use our understanding of human cognition to strengthen participation capacities of online users? How do we bridge insights from psychology, bridging and conflict resolution, social and emotional learning, and civic education to foster resilience and accountability online?



News and information: In the information era, how do we make it easier for users to access reliable and trustworthy content online? Inherently less scalable than national media, how can we sustainably fund local news organizations? How can local newsrooms better serve their communities, and provide more direct impact and value?

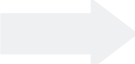
This frame articulates a distinct and necessary alternative to adequately addressing digital discourse challenges and harms, that is also achievable within the current political and industry landscape. These focus areas consider both the supply and demand side of the problem and bring in certain elements that have been missing from mainstream public policy and industry debates, such as a focus on the intersection of digital discourse and community connectedness, insights from behavioral science and bridging work, and an emphasis on the local. Our Discovery and Design process indicates that this frame will serve as a strong entry point for collaboration across ideologies and sectors to address digital harms and to build more resilience communities and a more connective democracy.

The above frame is not designed to be all-encompassing or mutually exclusive and has strengths and limitations. Any framing we choose will inevitably leave some important issues aside, including broader systemic ones. However, we believe it is important to zoom in to some extent. A manageable set of challenges with a targeted frame allows us to better determine the right stakeholders to engage and work toward a solution set that can ultimately achieve meaningful impact. Collaborative members will continue to define, adjust, and refine the frame along the way.

Potential Pathways to Consensus Solutions

Within our frame of consumer and community-oriented solutions, we have teased out possible **pathways to collaborative solutions** that mitigate harms related to digital discourse without violating free speech and other rights. These pathways triangulate consumer, community, policy, and industry action for a whole-of-society approach.

Consumer- and community-oriented pathways include calls to:




Reimagine and expand digital and media literacy to include an understanding of the architecture behind how information is presented and spread, understanding of our psycho-social vulnerabilities, cognitive strengthening, skills like overcoming motivated reasoning, conflict resolution, and dialogue. The goal is to have individuals who can engage competently, critically, and positively in digital spaces. The **Mental Immunity Project** aspires to do just that. Creative ways to reach users include gamifying pre-bunking interventions, such as games that put you in the shoes of a disinformation spreader. This approach to literacy must be created in a way that resonates with communities typically disconnected from these curricula, such as conservative, religious, and minority communities.

Mental Immunity Project: This project, run by the Cognitive Immunology Research Collaborative, empowers educators, organizational heads, community leaders, and individuals to build mental immunity in themselves and others.

Inoculation games: Educational games such as **Bad News**, **Harmony Square**, and **Go Viral!** build users' psychological resilience to false and sensationalist content online.

Christians and the Vaccine: A **partnership** by Redeeming Babel, the National Association of Evangelicals, the Ad Council, and others to equip pastors and Christian leaders to address vaccine hesitance.

#TrustedInfo2022: A **public education effort** ran by the National Association of Secretaries of State to promote state and local election officials as the trusted sources of election information.



➡ Explore creative ways to **fund, strengthen, and amplify** reliable media, especially local, investigative, and long-form journalism.

➡ **Identify, replicate, and scale community programming** to build resilience, effectively anticipate and counter disinformation and other divisive or harmful content, especially around elections and COVID-19, and build inclusive social connections that promote social cohesion on and offline.

➡ Work with non-traditional and local partners to **strengthen reliable intermediaries and institutions that engage in truth-telling and de-polarization**, including civic groups, bridgebuilding organizations, professional associations, religious institutions, community organizations, labor unions, and businesses.

➡ **Young people** are critical to resolving this challenge – we not only need to focus on the digital education and skills of children and young adults, but also look to them for entrepreneurial solutions to digital harms, and to chart the future of digital community and information spaces.

Policy-oriented pathways include calls to:

➡ **Examine existing federal policy proposals as well as explore state and local options** for nuanced and thoughtful legal interventions focused on design, advertising, transparency, and privacy that do not violate free speech. For example, the [Social Media NUDGE Act](#) promotes content agnostic interventions to reduce the harms of algorithmic amplification and social media addiction.

➡ Craft legislation that supports **meaningful digital, information, and media literacy**. Though this would primarily center on K-12 education, people of all ages would benefit from lifelong learning opportunities on this topic. An example includes New Jersey's recently passed legislation on information literacy in K-12 schools.

➡ Propose policies that **support local news**.

➡ Pass legislation that supports **bridgebuilding initiatives**, such as the [Building Civic Bridges Act](#).

Social Media NUDGE Act: This bill proposes a research and implementation process for content-neutral and evidence-based interventions to drive better experiences on social media. Examples could include screen time alerts, read, or review requirements, prompts to identify targeted advertisements, and limits on account creation and content sharing.

New Jersey Student Learning Standards in Information Literacy: The state established the nation's first K-12 program to develop information literacy, engaging teachers, libraries and other experts in the development and enforcement of these learning standards.

Building Civic Bridges Act: This bill would establish an Office of Civic Bridgebuilding within the Corporation for National and Community Service and new grants for civic bridgebuilding programs, and support training for AmeriCorps members in and research on civic bridging.

Industry-oriented pathways include calls to:

Use technologies and incentives to **support data access and academic research** into the spread of harmful content online, to bolster the efforts of technology companies already looking at these problems internally.

Continue experimenting with design approaches that add friction, disrupt virality, amplify pluralism, increase the volume and diversity of local civic information, and promote positive social discourse. This includes accrediting and elevating more trustworthy content, but beyond that, promoting more constructive, intentional and thoughtful digital interactions with news, media, and each other.

Create **robust, transparent systems** for auditing their ad platform and vetting, labelling, and archiving ads.

Partner with community organizations, local newsrooms, educators, and others to support community information ecosystems. This could include collaborations between mainstream media outlets and smaller, community-oriented, and ethnic media.

Continue diversifying and training the workforce in terms of interdisciplinary methods, ethics, human-centered design, different histories and contexts, and languages, especially to better address the realities of moderation and design in the technology space.

Create a technology pipeline that aligns with the public interest and enhances people's wellbeing.

There are countless existing efforts to tackle digital harms, but the work is disjointed. There is potential for the Convergence Collaborative to coordinate this diverse and distributed array of solutions and help them mutually reinforce each other.

Front Porch Forum: A community-building social network that most Vermont households are subscribed to. It is actively moderated, with every post reviewed by a human before publication. It centers on low-stake interactions such as borrowing, buying, and selling. The resulting daily flow of neighborly interactions leads to increased community and social capital and trust, which makes civil conversations on more difficult topics possible on the platform.

Concert Local: Vox Media partnered with Google News Initiative to launch Concert Local, a trusted advertising marketplace that provides a one-stop-shop for reaching local audiences at national scale with dynamic localized messaging.

ChatGPT: Chat Generative Pre-Trained Transformer, or ChatGPT, online AI chatbot launched by OpenAI. It is a language model that can generate human-like text. Experts see affordances for ChatGPT to make investigative journalism more productive and original journalism more valuable, but also to create or exacerbate disinformation.

There is a duality to the internet. On one side, its capacities for information exchange, communication, and connection have been used to expand humanity's boundaries and opportunities; on the other, the dark side of our modern online information ecosystem illustrates the costs of toxic digital discourse, and the threat to freedom of speech of poorly conceived solutions on the other. But this warrants neither utopianism nor despair. Despite the pain points, the opportunities above are ripe for impact, even within our current social, political, and economic landscape.

Convergence will shepherd the Collaborative group in exploring concrete questions within the frame above, identifying recommendations and solutions around which to build consensus, and charting pathways to implementation. These insights will be captured in a Consensus Blueprint report around which Convergence, the Collaborative stakeholders, and a broader network of influencers and practitioners will collaborate to bring the consensus recommendations and solutions into reality. These solutions will inform legislative, organization, or community-based policy, practices, priorities, or programmatic options.

We expect this Discovery Report and the subsequent Consensus Blueprint that will result from the Collaborative and our ongoing dialogue will have implications for leaders, experts, and practitioners across fields, as well as for information consumers. Our hope is that by bringing this uniquely diverse group together to build trust across differences, pool knowledge and resources, and champion action, we will catalyze a movement to develop collaborative solutions that lead to the betterment of ourselves, our online and physical communities, and our democracy.

About Convergence

Convergence is the leading organization bridging divides to solve critical issues through collaborative problem solving across ideological, political, and cultural lines. For more than a decade, Convergence has brought together leaders, doers, and experts to build trusting relationships, identify breakthrough solutions, and form unlikely alliances for constructive change on seemingly intractable issues. Our process is improving the lives of Americans and strengthening democracy for a more resilient and collaborative future.

ConvergencePolicy.org

ACKNOWLEDGEMENTS

The Convergence Collaborative on Digital Discourse for a Thriving Democracy and Resilient Communities is generously supported by [The John S. and James L. Knight Foundation](#), which promotes informed and engaged communities and believes that providing people with information is essential so that they and their communities can make the best choices and [New Pluralists Collaborative](#), a sponsored project of Rockefeller Philanthropy Advisors and a collaborative of diverse funders and field leaders (practitioners, storytellers, researchers, and innovators) working together to catalyze a culture of pluralism, belonging and respect in America.

This report would not have been possible without the tireless work of Convergence staff (Director, Monika Glowacki; Associate, Dylan Fabris; Associate, Emma Leyland) and the insights and time of those interviewed.

We thank you for your continued support in our efforts to find collaborative solutions to America's intractable issues.

For any questions or feedback on the report, please contact: monika@convergencepolicy.org.

Be among the first to receive the Consensus Blueprint when it's ready by signing up for Convergence emails at ConvergencePolicy.org or email us at communications@convergencepolicy.org.

CONVERGENCE





Contact

Convergence Center for Policy Resolution

1775 Eye Street NW, Suite 1150-287
Washington, DC 20006

(202) 830-2310

ConvergencePolicy.org

[@ConvergenceCtr](https://twitter.com/ConvergenceCtr)