THE CASE FOR DESIGNING TECH FOR SOCIAL COHESION

A working group formed in 2020 to explore the need for and interest in forming a Council on Technology and Social Cohesion. This 4-page document is a summary of the full reports, which can be found on the Council's website. The reports supported the design of the Designing Tech for Social Cohesion Conference in February 2023.
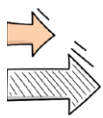
KEY TAKE AWAYS

→ The design of technology platforms is not neutral. Design affordances and algorithms can amplify toxic polarization or help to build social cohesion.

→ Digital town squares are increasingly important for information sharing and deliberation. But disinformation and other harmful content plague digital spaces.

→ Content moderation is important, but it is not keeping pace with the scale of harmful digital content and toxic polarization.

→ Tech regulation is important, but digital spaces are resistant to regulation and digital polarization spills are undermining policy solutions.

→ Pro-social tech platforms already exist, and we can learn from these tech design affordances and algorithms that support social cohesion.

→ Scaling social cohesion requires partnerships between practitioners and tech platforms to design better platforms and improve how people use tech in democratic processes.

A LANDSCAPE ANALYSIS OF TOXIC POLARIZATION ONLINE

Toxic polarization is increasing globally, contributing to violence, and hampering efforts to solve pressing public problems. While not the origin of social and political division, there is wide agreement that the design of many social media products amplifies polarization.

Tech companies are building a "Trust and Safety" teams to address online digital harms. But the challenge of content moderation is increasing. Political actors, cyber armies, and a growing for-profit disinformation industry amplify and incentivize individual producers of divisive digital propaganda aimed at polarizing societies with a "divide and conquer" strategy. In an era of publicly amplified political propaganda, content moderation itself is polarizing. In 2022, the tech sector's relatively new "Trust and Safety" infrastructure laid off 100,000 tech workers and downsized or eliminated human rights and content moderation teams due to falling tech company stock prices, Elon Musk's Twitter acquisition, and other global factors.

Toxic polarization occurs when people perceive other groups as existential threats, distrust and dehumanize others with us-vs-them narratives and justify the use of violence against others.

Technology or tech refers here to digital tools, with a particular but not exclusive focus on social media.

Social cohesion refers to the glue that keeps society together; it is the opposite of toxic polarization. The United Nations defines social cohesion as "the extent of trust in government and within society and the willingness to participate collectively toward a shared vision of sustainable peace and common development goals." Three dimensions of social cohesion include individual agency, horizontal relationships, and vertical relationships.

Bridge building and peacebuilding are types of **prosocial interventions** that support the goal of social cohesion in three ways.

1) Increasing individual agency;
2) Bridging relationships between groups; and
3) Building public trust between society and governing institutions.

PeaceTech refers to technology that both supports the analysis of polarization and bridge building or peacebuilding interventions to support social cohesion.

3 APPROACHES TO DIGITAL POLARIZATION
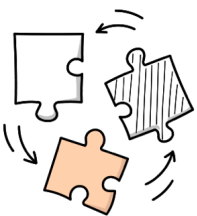
This research found three distinct but complementary narratives or approaches to thinking about polarization and social cohesion in digital spaces.

1. The **"User-Centered" Narrative** describes harmful content online as *generated by users*, with social media products and search engines acting as a *mirror* of society. Tech company Trust and Safety teams use a variety of strategies to reduce digital harm including, for example, setting community guidelines; human and algorithmic moderation to remove, demote or alter the ranking; recommending and amplification of content; and partnerships with groups helping tech companies identify harmful content.

2. The **"Tech Design Regulation" Narrative** describes harmful content as generated by the design of tech *affordances and algorithms* that are optimized for user engagement and profits. *Engagement-driven affordances* such as an ability for users to "Like" may encourage anxiety and social comparisons while the *engagement-driven algorithms* rank content to show users divisive and emotionally engaging content that pits groups against each other. Many of the most popular digital tools offer people a megaphone for shouting, much like the design of a gladiator arena. Government regulation of tech companies to date has focused on privacy and cybersecurity concerns, not the affordances and algorithms that amplify toxic polarization. Tech products optimized for user engagement, advertising, and profit incentivize the spread of false and hateful posts. Regulating algorithms can incentivize socially responsible digital spaces and sanction those that fuel toxic polarization.

> ***Affordances*** are the features of a tech product that limit what people can do in a digital space.
>
> ***Algorithms*** are the computational settings of a tech product that determine what content users can see.

3. The **"Social Cohesion by Design" Narrative** offers a third and complementary approach to content moderation and tech regulation of algorithms. This approach addresses the challenge of toxic polarization by designing digital spaces for pro-social content and social cohesion. Computer engineers working with the bridging movement and peacebuilding practitioners have already designed tech products with affordances and algorithms optimized to amplify and scale social cohesion rather than polarization.

TECH AFFORDANCES AND ALGORITHMS TO SUPPORT SOCIAL COHESION
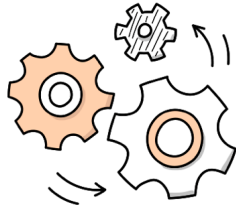
As detailed in the full report, a growing number of UN agencies, governments, universities, NGOs, civil society initiatives, and start-up tech companies are designing digital spaces or "peacetech" with affordances and algorithms that support four goals related to social cohesion:

1. Scale *individual agency*, including opportunities to participate in civic life including

→ Opportunities for meaningful online civic engagement (eg. Ushahidi, Kazm, #IamHere movement, Lithuanian Elves, Braver Angels)

→ Measures of impact so people view civic action as worth their time (eg. petition sites that report policy changes and impacts)

→ Coaching in effective communication strategies (eg. eBay, Angry Uncle Bot)

2. Scale *horizontal relationships within and between groups* including:

→ Audio and video affordances that humanize others and discourage harmful content by design (eg. Marco Polo, Gatherround, SlowTalk, Soliya)

→ Guardrails to minimize trolling and harmful content by removing the possibility of harmful personal responses in comment sections on public issues (eg. Pol.is)

→ Affordances to enable "listening at scale" to enhance "perspective taking" to better understand the views and interests of diverse groups (eg. Pol.is, Remesh, and Reddit's ChangeMyView)

→ Reality testing on perception gaps to improve accurate understanding (eg. More in Common's Perception Gap Quiz)

→ Visualization of shared interests and common ground between people (eg. Pol.is and Remesh)

3. Scale *vertical relationships between public institutions and society* including

→ Incentives to develop policy solutions on polarized public issues that take into consideration the core interests of others (Pol.is and Remesh)

→ Transparency on institutional performance such as blockchain tech products that increase confidence and public trust in governance

4. *Analyze digital polarization and information ecosystems*

→ Offer affordances to identify polarized content, trending topics, hashtags, and key influencers (eg. Phoenix and Sparrow)

Some of these platforms began with computer engineers with training in social cohesion. Others started as initiatives of the UN or NGOs in partnership with tech startups to create products that would support bridge building and peacebuilding work. And in some cases, like the case study below, a big tech company like Twitter borrowed and adapted these ideas for large-scale efforts to improve social cohesion.

CASE STUDY:

POL.IS AND TWITTER'S COMMUNITY NOTES

Inspired by insights from social cohesion efforts in nonviolent communication and attempts at collective decision-making in the Occupy Movement, Colin Megill designed the tech platform Pol.is to improve computational democracy. Experiments in Taiwan and the UK showed that Pol.is could help a divided public find areas of common ground and develop policy solutions on polarized public issues. Drawing on inspiration from Pol.is' affordances and algorithms, Twitter staff developed a program called Community Notes (formerly Birdwatch) to empower Twitter users to add helpful notes to Tweets that might be misleading. Wired Magazine calls this "one of the most exciting content moderation innovations ever to come out of not just Twitter, but any major platform."

Carl Miller. "Elon Musk Embraces Twitter's Radical Fact-Checking Experiment." *WIRED Magazine*. 28 November 2022.

A ROADMAP FOR COLLABORATION ON TECHNOLOGY AND SOCIAL COHESION
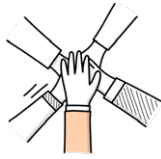
There is emerging interest in and support for a Council on Tech and Social Cohesion to create opportunities for partnership and collaboration in depolarizing societies online and offline and designing tech with affordances and algorithms to support prosocial content. While there are a variety of efforts to support "tech for good," "public-interest tech" or "tech for social impact," this initiative is distinct. This initiative creates a space for tech companies with *practitioners who use peacebuilding and bridge building to foster social cohesion*.

Interviewees for this report identified a range of potential areas for partnership. The full Roadmap with more details and examples can be found on the Council's website.

Toda Peace Institute

Search for Common Ground
Trust, Collaboration, Breakthroughs

Braver Angels

UNIVERSITY OF NOTRE DAME

Belgium
partner in development

B · KBF CANADA

More in Common

Center for Humane Technology

Alliance for Peacebuilding

CIVICHEALTH PROJECT

PSYCHOLOGY OF TECHNOLOGY INSTITUTE

* Members of the Council on Technology and Social Cohesion

1 **Institutionalize the Cohesive Tech Movement** – including mapping relevant stakeholders, creating a shared calendar of events, host cross-company forums for information sharing, and matchmaking forums to create partnerships between peacebuilding and bridging movement organizations to create partnerships with tech engineers, start-ups, and companies.

2 **Promote Public Awareness of Tech Roles in Social Cohesion** – create compelling digital content and articles with case studies and discussions of the role of technology in supporting social cohesion.

3 **Incubate Prosocial Tech** – including supporting peacetech start-ups, piloting and experimenting in different contexts with peacetech, organizing a "sharktank for peacetech" to offer feedback from seasoned tech experts, and offering coaching for peacetech startups on how to make a market case or pitch to funders.

4 **Explore Prosocial Funding Models** – including foundations, governments and international organizations, cooperative and open-source funding models, venture capital funding.

5 **Train and Build Capacity on Tech & Social Cohesion** – including offering workshops on Social Cohesion for Executive and Tech Staff to understand and prioritize elements of social cohesion and tech workshops for peacebuilding and bridging movement groups.

6 **Measure Tech's Impacts on Toxic Polarization and Social Cohesion** – including refining digital indicators and metrics of social cohesion, and forming data analysis partnerships.

7 **Improve Content Moderation** – including partnering to identify and analyze harmful content, and experimenting with citizen panels for content moderation and improving procedural justice.

8 **Protect Information Ecosystems** – including partnerships with local organizations to identify digital risks such as disinformation and hate speech, coordinate interventions to improve information ecosystems, create early warning systems to prevent violence and support public interest journalism online and offline, and protect mediation and peacebuilding efforts from digital harms.

9 **Explore Government Regulation to Incentivize Tech for Social Cohesion** – including a framework of metrics to reward tech companies by creating a market signal for the positive contributions to social cohesion, and tax or sanction companies for harmful content or information pollution.

10 **Advocate for Big Tech Adoption of Prosocial Affordances and Algorithms** – including bridging-based ranking, incentives for listening and checking perceptions, guardrails to prevent harmful content, and incentives for humanizing others.