

Liana S

Bard Center for the Study of Hate

Internship with the Institute for Strategic Dialogue

Summer 2025

### The Architecture of Hate: Investigating Propaganda, Polarization, and Online Harm at ISD

For the past year and a half, I focused on documenting human rights violations: verifying videos of police brutality, archiving forced deportations, and analyzing state-authorized violence. As human rights violations researchers, we often get caught up in the technical process of verification and classification, before contextualizing these events to convey urgency and impact. But it is crucial to take a step back from time to time to ask deeper, structural questions about the nature of the hate we observe. What forces enable it to take root? How do harmful narratives become normalized, obscured, or even celebrated? Why does it feel like the presence and social tolerance of hate have grown in recent decades - and what political, technological, or economic factors might explain that shift?

Part of the answer lies in the extraordinary speed at which narratives and moods shift in the current political and social landscape. The post-Cold War international order has been marked by a paradigm shift: the threats to human rights and security no longer come exclusively from traditional state actors or clearly identifiable militant organizations. Instead, we are increasingly confronted with a post-organizational reality, where non-state actors, loosely connected online communities, and even individuals can generate and sustain extremist ideologies. In this environment, hateful propaganda can directly incite or inspire violence, erasing the clear boundary that once separated messaging from harmful action. The systems that spread such hate continually evolve to adapt and reach new audiences.

To explore these dynamics, I joined the Institute for Strategic Dialogue (ISD) through the Bard Center for the Study of Hate's internship program. My placement was in ISD's Information Operations (InfoOps)

team, where I helped investigate how sanctioned pro-Kremlin media outlets in the EU circumvent content bans, and how Russian state-aligned narratives about the deportation of Ukrainian children operate across digital platforms. Alongside this work, I engaged with ISD's broader approach: how it conceptualizes hate, what kinds of solutions it prioritizes, how it measures success, and what ethical tensions emerge from operating as both an analyst and an actor within the disinformation space.

## **Introduction to Institute for Strategic Dialogue**

ISD operates as a global “think-and-do tank,” combining research, policy advising, and community programming to counter hate, extremism, and disinformation. From my very first week, it was clear that ISD does not treat hate as a moral failing or emotional outburst - it understands hate as a structured phenomenon. According to ISD, hate is central to extremism, and extremism is the advocacy of political and social changes in line with a system of belief that claims the superiority and dominance of one identity-based “in-group” over an “out-group”. Hate, in this context, involves beliefs and practices that malign, delegitimize, or exclude entire classes of people on the basis of immutable characteristics such as ethnicity, religion, gender, sexual orientation, or disability.<sup>1</sup>

In digital environments, this cultivation happens with extraordinary ease and reach. Social media platforms reward extremity and confrontation. Comment sections become echo chambers of cruelty, no longer bound by shame or anonymity. What once required effort to conceal is now proudly declared. Bot networks and coordinated campaigns further distort public discourse, giving extreme narratives an appearance of mass legitimacy. All of this reinforces a perception that hate is widespread and inescapable - even if in some cases that perception itself is a product of manipulation.

A recurring theme in my conversations with people at ISD was polarization - how the “us versus them” narrative is both a consequence and a tool of hate. It is cultivated through narrative repetition, historical

---

<sup>1</sup> Institute for Strategic Dialogue, *Extremism*, ISD Global, accessed July 31, 2025, <https://www.isdglobal.org/extremism/>.

grievance, and moral dissonance. But it is also incentivized structurally: platforms promote engagement, engagement thrives on outrage, and outrage becomes currency. In this landscape, hate is not just expressed - it is gamified, monetized, and often made invisible by the very systems we rely on to detect it.

ISD works to challenge these dynamics through three core approaches: policy, education, and digital analysis. In the policy space, the organization collaborates with the EU and national governments to shape platform accountability laws and public communications strategies. Its research has fed directly into the development of the Digital Services Act, where ISD's frameworks help distinguish between misinformation and coordinated state manipulation. These tools are not politically neutral; they rest on the normative assumption that hate and propaganda are not inevitable byproducts of free speech, but strategically deployed weapons that require targeted responses.

Complementing this top-down engagement are ISD's community-based programs, particularly the Strong Cities and Young Cities networks. These initiatives work with municipalities and youth leaders to counter polarization through civic education. Participants are not merely taught to identify disinformation - they are encouraged to build emotional literacy and solidarity. This personal approach is the opposite to policy work, but it is just as crucial and effective. This reframing of hate as both an emotional and infrastructural process gave me a new appreciation for just how layered counter-strategy must be.

## **My Involvement**

My role in the InfoOps team provided a close look at how digital methods, such as scraping, narrative mapping, and traffic analysis, can be used to understand how hate and propaganda circulate. One of the main projects I was involved in was mapping access to banned Russian media outlets post-sanctions.

Despite the EU restrictions, users continued accessing Russian state-aligned media through mirror sites, VPNs, and search-engine manipulation.<sup>2</sup> This raised questions not just about policy enforcement, but

---

<sup>2</sup> Institute for Strategic Dialogue, *Holding the Line: Auditing the EU's Ban of Russian State Media, 3 Years On*, August 2025, [https://www.isdglobal.org/digital\\_dispatches/investigation-holding-the-line-auditing-the-eus-ban-of-russian-state-media-3-years-on/](https://www.isdglobal.org/digital_dispatches/investigation-holding-the-line-auditing-the-eus-ban-of-russian-state-media-3-years-on/).

about the symbolic power of media visibility. Russia could continue to position itself as a censored truth-teller - subverting the moral authority of democratic regulation while maintaining reach. I also confirmed yet again that there is a pressing necessity for securitization of the hybrid warfare threat posed by Russia - the current sanctions on the media outlets that are spreading propaganda which seeks to undermine democratic processes in EU member states are not sufficient to fully counter the threat and conduct damage control.

The second project brought together my past and present research in a very personal and pressing way. I had previously investigated the deportation of Ukrainian children by Russian authorities from a human rights perspective: documenting violations, tracing legal implications, and gathering testimony. At ISD, I returned to this issue, but from a different angle: narrative warfare. My supervisor, also Ukrainian, and I initiated the project from scratch, developing a framework to identify and analyze harmful disinformation narratives around Ukraine - particularly claims framing the country as a hub for child trafficking.

These narratives didn't emerge in isolation. They often intersected with larger conspiratorial ecosystems - stories about organ harvesting, satanic rituals, or elite child abuse rings. What was especially alarming, however, was not just the content, but the mode of distribution. Our initial hypothesis was that much of this amplification might be coordinated or inauthentic, possibly driven by bots or troll farms. But the data told a different story. We found no clear evidence of coordinated inauthentic behavior. Instead, what we saw was organic sharing by users. Real people were posting, resharing, and commenting with conviction and moral outrage. The scale was disturbing not because it was manipulated, but because it was *sincerely believed*.

This absence of artificial amplification made the threat feel more out of control. Disinformation spread through inauthentic means can, at least in theory, be flagged and removed. But when the engine of spread is authentic belief - fueled by emotion, trauma, or ideological framing - the challenge becomes social and cultural, not merely technical. The idea that Ukraine is a place of moral rot or criminality, especially

around children, didn't need to be imposed - it resonated. And it distracted from the real issue - Ukrainian children being illegally transferred to Russia-controlled territories with few, if any, opportunities to return.

That's what made this project particularly important: these trafficking narratives didn't just mislead - they displaced attention. They created moral fog. While pro-Kremlin networks generated outrage about fictional abuse, the actual abuses committed by Russian forces - forced deportations, illegal adoptions, re-education programs, cultural erasure - went underreported or were even reinterpreted as benevolent. The goal of disinformation here wasn't just to deny Ukrainian victimhood - it was to overwrite it entirely. To take the language of harm and reassign it. In doing so, the machinery of propaganda didn't just justify state violence - it masked it with moral urgency. And by tracing that dynamic, I came to see how essential narrative analysis is - not as a substitute for documentation, but as its necessary partner.

### **Balance of Urgency and Skepticism**

One of the more nuanced challenges I encountered at ISD was how to balance urgency with skepticism. The InfoOps space is filled with daily reports of new threats: extremists using AI to generate recruitment content, disinformation networks mimicking civic organizations, foreign campaigns targeting minority groups with emotionally charged messages. The pace is at times overwhelming. It can be easy to internalize a sense of crisis. But what impressed me about ISD was its refusal to indulge in panic. The team constantly emphasized empirical grounding - cross-checking assumptions, resisting hype, and contextualizing findings. Sometimes, alarmist narratives about technological threats are themselves part of the disinformation strategy. Knowing when *not* to escalate became just as important as identifying genuine risk.

At the same time, no one pretended that the digital domain is the whole battlefield. Everyone I worked with understood that online hate has offline consequences - fueling violence, legitimizing exclusion, hardening state repression. The platform is not the problem; it's the medium. The problem is systemic:

political incentives, identity formation, narrative infrastructures. And that's what ISD tries to address - not just the expressions of hate, but the conditions that make them effective and thriving.

This commitment to nuance is perhaps most visible in how ISD handles its own role as a narrator of hate. There's a constant tension between exposing disinformation and amplifying it. Every time an organization like ISD publishes a report on an extremist network or state-aligned campaign, there's a risk: visibility is a resource, and these actors thrive on reaction. ISD researchers are aware of this. Rather than sensationalizing findings, reports are structured to limit exposure of raw content, provide context, and minimize dramatization.

### **Emotional Cost of Work**

Of course, this work is not without emotional cost. Many team members spoke openly about desensitization and emotional exhaustion. Looking at hate every day changes how you relate to it. ISD attempts to mitigate this through ethical briefings, mental health resources, and regular internal conversations that re-center the human impact behind the data. But emotional fatigue remains a structural condition of the field. When discussing the issue of desensitization with one of my colleagues, they brought up the necessity to zoom out every now and then to remember that behind data that we see on the screen are people who are being affected daily by it in different forms. For human rights work, desensitization is both a curse and a miracle: it may get easier to work with explicit content after being exposed to it for some time, and yet it changes our perception of the levels of violence and hate we consider "normal." What once felt shocking begins to feel routine. The danger is not only emotional numbness, but a gradual recalibration of our moral threshold. We risk losing the ability to recognize the extraordinariness of harm - to see humanity in every instance of abuse rather than reducing it to a datapoint in a broader trendline.

And yet, without some degree of desensitization, the work would be impossible. If every image or testimony elicited the same visceral reaction as the first time we encountered it, burnout would be

inevitable. The paradox, then, is learning how to build resilience without slipping into indifference. Some colleagues describe this as cultivating a kind of professional distance - not apathy, but the ability to compartmentalize in order to remain effective. Still, even with such strategies, the risk remains that the constant exposure to systemic hate and violence reshapes how we interpret the world beyond our screens.

In this sense, zooming out becomes more than a coping mechanism - it is an ethical imperative. To pause, to recontextualize, and to remind ourselves that each dataset corresponds to lived realities helps counteract the numbing effect of repetition. It keeps the work tethered to its purpose: not simply cataloging hate, but understanding and challenging the systems that produce it. Without this periodic reorientation, we risk turning into technicians of atrocity rather than advocates for justice.

### **Possible Collaborations**

One area where ISD could meaningfully expand its work is through deeper, more sustained collaboration with academic institutions and scholars. While ISD's strength lies in its independence and ability to respond quickly to emerging threats, that very agility often limits the capacity for deeper theoretical engagement. Integrating structured partnerships with universities through research fellowships, rotating academic residencies, or collaborative projects could add critical conceptual depth to ISD's already-strong methodological rigor.

Such partnerships would be especially valuable in areas where the ethical stakes are high: content moderation, digital surveillance, disinformation ethics, and platform regulation. Scholars in feminist theory, critical race studies, surveillance studies, and digital humanities could offer frameworks that challenge assumptions and expose blind spots. Collaborations with universities could also strengthen training, evaluation, and reflective practice within ISD. For instance, co-designed research with universities could explore longer-term outcomes of resilience programming, or trace how policy shifts influenced by ISD manifest on the ground. Even occasional joint publications or peer-reviewed spin-offs

from ISD reports could bridge the gap between practice and scholarship, enhancing both legitimacy and reach.

## **Final Reflections**

One of the most important things I've taken away from this experience is a more systematic understanding of how hate operates. Before working at ISD, my focus was on documenting direct harms, such as deportations, abuses, legal violations. That kind of work is essential. It taught me how to handle evidence responsibly, how to build a case based on verification, and how to remain sensitive to the dignity of the people affected. But during my time at ISD, I began to look at the structure around those harms: the justifications, the narratives, and the mechanisms through which violence is presented as something else: order, protection, even care. I started to understand that combating hate isn't just about exposing violence, but also about challenging the systems that allow it to make sense to others.

This experience made me more critical and more cautious - not cynical, but more aware of how easily manipulative content can blend into the everyday. It also changed how I think about intervention. Not everything that seems threatening in theory turns out to be impactful in practice, and not everything that appears insignificant at first is harmless. ISD taught me to interrogate evidence not just for its content, but for its resonance: how it circulates, who engages with it, and what unintended effects exposure might produce. That connects directly to ethical questions I've encountered before in human rights investigations, especially around the "Do No Harm" principle<sup>3</sup> and Berkley protocol. Originally designed for medical and humanitarian contexts, and adapted by research institutions like UC Berkeley's Human Rights Center, the principle urges investigators to carefully assess whether publishing, exposing, or analyzing a piece of information could contribute to harm, even indirectly.<sup>4</sup>

---

<sup>3</sup> CDA Collaborative Learning Projects, *Do No Harm: A Brief Introduction from CDA* (2018), <https://www.edacollaborative.org/wp-content/uploads/2018/01/Do-No-Harm-A-Brief-Introduction-from-CDA.pdf>.

<sup>4</sup> United Nations Office of the High Commissioner for Human Rights (OHCHR), Human Rights Center at UC Berkeley School of Law, *Berkeley Protocol on Digital Open Source Investigations: Second Edition* (2022), [https://www.ohchr.org/sites/default/files/2024-01/OHCHR\\_BerkeleyProtocol.pdf](https://www.ohchr.org/sites/default/files/2024-01/OHCHR_BerkeleyProtocol.pdf).

At ISD, these concerns didn't disappear; they simply shifted form. Instead of asking whether a survivor's testimony might place them at risk, I was asking whether platform research or naming specific actors might accidentally amplify their reach. The ethical questions were no less urgent, even if they weren't always straightforward. There were moments where we debated how much detail to include, whether naming certain channels or influencers was justified, and how to avoid making hate content more visible while still documenting its presence and strategy. These dilemmas don't have easy answers, but being in an environment that openly discussed them was reassuring.

At the same time, I saw clearly how the tools and values of human rights research and digital analysis can strengthen one another. The former brings a long tradition of centering affected communities, building ethical safeguards, and resisting flattening narratives. The latter adds capacity for scale, pattern recognition, and structural mapping. Rather than feeling like two opposing approaches, they became, for me, two sides of the same work. My previous experience helped me read behind the data to see people, politics, and power in what might otherwise look like trendlines. My work at ISD taught me to place those stories into a wider ecosystem to analyze how they're distorted, challenged, or erased altogether.

Practically, I leave this internship with sharpened research and writing skills. I'm more comfortable with social media analysis, disinformation typologies, and how to ask questions which can be answered with data. But I also leave with a deeper appreciation for the responsibility that comes with doing this kind of work, as well as for the people who are working at ISD. They are not only brilliant analysts but also deeply resilient people, managing the heavy reality of this work with grace and humanity. What I admired most was how they struck a balance that felt rare: never so cynical that the work seemed hopeless, but never so idealistic that they ignored its difficulties.

Even in the face of exhausting subject matter, they made space for care: for each other and for the communities we were studying. Despite the seriousness of the work, the atmosphere on the team was always supportive and genuinely curious. People looked out for one another, shared insights freely, and

never lost their sense of humor, even on tough days. It might sound simple, but being part of such a thoughtful and committed group gave me a lot of hope. Watching how they balanced sharp analysis with empathy - and how they managed to stay neither too cynical nor overly idealistic - showed me that resilience in this field isn't about being hardened, but about staying human together.

## Works Cited

1. Institute for Strategic Dialogue. *Extremism*. ISD Global. Accessed July 31, 2025.  
<https://www.isdglobal.org/extremism/>.
2. Institute for Strategic Dialogue. *Holding the Line: Auditing the EU's Ban of Russian State Media, 3 Years On*. August 2025. [https://www.isdglobal.org/digital\\_dispatches/investigation-holding-the-line-auditing-the-eus-ban-of-russian-state-media-3-years-on/](https://www.isdglobal.org/digital_dispatches/investigation-holding-the-line-auditing-the-eus-ban-of-russian-state-media-3-years-on/).
3. CDA Collaborative Learning Projects. *Do No Harm: A Brief Introduction from CDA*. Cambridge, MA: CDA Collaborative, 2018.  
<https://www.cdacollaborative.org/wpcontent/uploads/2018/01/Do-No-Harm-A-Brief-Introduction-from-CDA.pdf>.
4. United Nations Office of the High Commissioner for Human Rights (OHCHR), Human Rights Center at UC Berkeley School of Law. *Berkeley Protocol on Digital Open Source Investigations: Second Edition*, 2022.  
[https://www.ohchr.org/sites/default/files/202401/OHCHR\\_BerkeleyProtocol.pdf](https://www.ohchr.org/sites/default/files/202401/OHCHR_BerkeleyProtocol.pdf).