

Fact Sheet: DSOSA and Russia's Influence Operations on Social Media

As experts discuss and track the role social media is playing in Russia's invasion of Ukraine, they have highlighted several challenges that could be addressed through legislation, including provisions in the newly released *Digital Services Oversight and Safety Act (DSOSA)*.

DSOSA is centered around systemic risks, one of which is “Any malfunctioning or intentional manipulation of a hosting service, including by means of inauthentic use or coordinated, automated, or other exploitation of the service...[that] has an actual or foreseeable negative effect on the protection of...civic discourse, electoral processes, public security, or the safety of vulnerable and marginalized communities.” [Sec 4. (a) (3) (C)]

These types of systemic risks are sometimes referred to as *coordinated influence operations* or “coordinated campaigns by one organization, party, or state to affect one or more specific aspects of politics in domestic or another state, through social media, by producing content designed to appear indigenous to the target audience or state.”¹

Combatting coordinated influence operations is something every major platform has to consider—in the face of international conflict, upcoming elections, or otherwise. Beyond these considerations, platforms are left to choose how much to invest in keeping their information ecosystem safe, often receiving ad hoc guidance from government leaders.

DSOSA offers a more sustainable approach.

DSOSA establishes a Bureau at the Federal Trade Commission staffed with experts dedicated to understanding hosting services, content moderation processes and platform design. The full bill is [described here](#).

¹ Alizadeh, M., Shapiro, J. N., Buntain, C., & Tucker, J. A. (2020). Content-based features predict social media influence operations. *Science advances*, 6(30), eabb5824.

Regarding coordinated influence operations specifically, DSOSA would incentivize platforms to go beyond “considerations” in the following ways:

1. **Mandating Risk Assessments and Risk Mitigation Reports.** Under DSOSA, large covered platforms (including Facebook, YouTube, Twitter) would be asked to assess the potential for their platform to be used for coordinated influence operations and to document the processes they have in place to mitigate those attacks. Mitigation techniques most salient to coordinated influence operations include:
 - Integrating threat modeling and red teaming processes to guard against systemic risks in the early stages of product design, and to test potential mitigations prior to the release of a product. [Sec 7 (a) (3) (A)]
 - Crisis protocols, including a description of what constitutes the specific extraordinary circumstance the crisis protocol seeks to address and the objectives the crisis protocol pursues. [Sec 7 (a) (3) (F)]
 - Targeted measures aimed at limiting the display of advertisements and adapting the content moderation or recommender systems.

2. **Voluntary Guidance.** The *Bureau of Digital Services Oversight and Safety* would be tasked with publishing guidance on an ongoing basis that platforms can choose to reference when deploying processes to address systemic risks, influence operations or otherwise. The guidance would be informed by a skilled staff armed with information gleaned from DSOSA’s many disclosure requirements, research published by certified researchers, and Advisory Committees aimed to solicit views from stakeholders, including communities most impacted by the systemic risks, as well as content moderators and employees at covered platforms focused on mitigating such risks.
 - While guidance would be voluntary, it would provide a central source for companies to reference as they set up their own systems. This process is not too dissimilar to CDC guidance.

3. **Facilitates access to critical data.** One reason we have an understanding of influence operations is because academic researchers and civil society groups have been granted access to *some* data made available by platforms. DSOSA takes this tenuous voluntary data sharing by a handful of platforms to the next level by creating an *Office of Independent Research Facilitation*. This office would certify host organizations and researchers and set data security and data management rules to facilitate research to “gain understanding and measure the impacts of the content moderation, product design decisions, and algorithms of covered platforms on society, politics...” [Sec. 10 (b) (2)]

The data made available may include: digital advertisements, high reach content, information related to content moderation decisions, content removal requests, information related to engagement and exposure of certain content, classification of information sources, archives of removed content and accounts, and anything else deemed appropriate by the Commission. [Sec. 10 (b) (3)]

The Commission may choose to make the data available through a Federally Funded Research and Development Center that would be able to “collaborate with international research organizations with a similar mission.” [Sec 10 (c) (7)]

DSOSA is designed to create capacity to better understand online platforms and systemic risks and to hold platforms accountable to the promises they make to consumers and advertisers. Foreign influence operations are only one systemic risk platforms face but it is worth discussing the ways in which comprehensive transparency legislation can address these risks.

DSOSA is [supported by a wide range of outside groups](#).

A fact sheet on the bill can be accessed [HERE](#).

A section-by-section summary can be accessed [HERE](#).

The text of the bill as introduced can be accessed [HERE](#).