

ISSUES & INSIGHTS  
VOL. 23, SR 9

SEPTEMBER 2023

PACIFIC FORUM  
INTERNATIONAL

# The United States & Japan: Allied Against Disinformation — Next Generation Voices Speak

EDITED BY  
ROB YORK & AKIRA IGATA





Founded in 1975, the Pacific Forum is an independent, nonpartisan, and non-profit foreign policy research institute based in Honolulu, in the U.S. state of Hawaii. The Forum's focus areas encompass current and emerging political, security, economic, and business issues and work to help stimulate cooperative policies in the Indo-Pacific through research, analyses, and dialogues undertaken with the region's leaders in academia, public policy, military, and industry. The Forum collaborates with a network of more than 30 research institutes around the Pacific Rim, drawing on Asian perspectives and disseminating project findings and recommendations to opinion leaders, governments, and various publics throughout the region. We regularly cosponsor conferences with institutes throughout Asia to facilitate nongovernmental institution building as well as to foster cross-fertilization of ideas.

A Board of Directors guides the Pacific Forum's work. The Forum is funded by grants from foundations, corporations, individuals, and governments. The Forum's studies do not engage in classified or proprietary work.

### **Support Pacific Forum**

Pacific Forum is a private, independent, nonpartisan, and non-profit, 501(c)(3) organization. Make a tax-deductible charitable contribution at [www.pacforum.org/support-us](http://www.pacforum.org/support-us)

To support a specific program, please contact our Director of Development at: [brooke@pacforum.org](mailto:brooke@pacforum.org)

## **PACIFIC FORUM STAFF**

President  
**DAVID SANTORO, Ph.D.**

Executive Director  
**CARL BAKER**

Senior Director of Indo-Pacific Program  
**JOHN HEMMINGS, Ph.D.**

Director for Regional Affairs  
**ROB YORK**

Director of Cybersecurity & Critical Technologies  
**MARK MANANTAN**

Director of Development & Grants Management  
**BROOKE MIZUNO**

President Emeritus & WSD-Handa Chair in Peace Studies  
**RALPH COSSA**

Senior Advisor  
**CRYSTAL PRYOR, Ph.D.**

Senior Advisor  
**BRAD GLOSSERMAN**

Director for Maritime Security  
**JEFFREY ORDANIEL, Ph.D.**

Director for Women, Peace, and Security Programs & Senior Fellow  
**MARYRUTH BELSEY PRIEBE**

Director of Young Leaders Program  
**CAROL LI**

Executive Assistant  
**GEORGETTE ALMEIDA**

Senior Program Manager  
**JESSLYN CHEONG**

Program Managers  
**VALERIA VILLASENOR-BRUYERE**  
**JEFF OTTO**  
**MEGAN TAENAKA**

Communications & Outreach Manager  
**SHANNA KHAYAT**



# The United States & Japan: Allied Against Disinformation — Next Generation Voices

Edited By  
Rob York & Akira Igata

A large, light gray graphic of a sun or star with a central circle and several pointed rays, positioned on the left side of the cover.

ISSUES & INSIGHTS

EDITED VOLUME

VOL. 23, SR 9 | SEPTEMBER 2023

## Issues & Insights

*Issues & Insights*, Pacific Forum's flagship peer review and open access publication, is published in the forms of working papers, edited volumes, and conference/workshop reports. Authored by academics, policy analysts, and subject-matter experts from around the region and beyond, *Issues & Insights* provides timely, contextual, and innovative policy research and analyses on pressing political, security, and economic developments in the Indo-Pacific.

For submission guidelines and to subscribe, email:  
[pacnet@pacforum.org](mailto:pacnet@pacforum.org)

For back issues, visit [www.pacforum.org/programs/issues-insights](http://www.pacforum.org/programs/issues-insights).

### Disclaimer

The statements made and views expressed are solely the responsibility of the authors and do not necessarily reflect the views of their respective organizations and affiliations. Pacific Forum's publications do not necessarily reflect the opinions of its staff, donors, and sponsors.

### Limited print and electronic distribution rights

This representation of Pacific Forum intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from Pacific Forum to reproduce, or reuse in another form, any of its publications for commercial use.

© 2023 Pacific Forum International. All rights reserved.

For information on reprint and linking permissions, please email [pacnet@pacforum.org](mailto:pacnet@pacforum.org)

# TABLE OF CONTENTS

INTRODUCTION	
<i>Rob York</i>	1
1 THE CASE FOR ANTI-DISINFORMATION EDUCATION IN SCHOOL CURRICULUM	
<i>Rachel E. Brooks</i>	3
2 BREAKING THE INFORMATION COCOON: JAPAN-US NEW APPROACHES TO COUNTER DISINFORMATION CAMPAIGNS IN THE AGE OF COGNITIVE WARFARE	
<i>Taro Nishikawa</i>	7
3 DIGITAL ALLIES: HOW THE UNITED STATES AND JAPAN CAN PARTNER AGAINST DISINFORMATION	
<i>Tom Ramage</i>	13
4 A PROPOSAL FOR COUNTERMEASURES AGAINST DISINFORMATION: UTILIZATION OF NHK AS A PUBLIC BROADCASTER	
<i>Ryohei Suzuki</i>	17
5 PARRYING THE PRINCIPAL THREAT: REDUCING THE DISINFORMATION RISKS OF AUTHORITARIAN-INFLUENCED PLATFORMS THROUGH US-JAPAN COOPERATION ON THREAT ANALYSIS AND MITIGATION	
<i>Jonathon Marek</i>	23
6 COUNTERMEASURES AGAINST THE SPREAD OF FALSE INFORMATION BY GENERATIVE AI	
<i>Yuichiro Kotaka</i>	27
ABOUT THE AUTHORS	31



# **Introduction**

*Rob York*

Pacific Forum, like the US government and much of the international security community, considers the information space a crucial theater in the United States' ongoing great power competition with the People's Republic of China and Russia. To meet the need for fresh policy ideas, as well as Pacific Forum's long-standing mandate to train the next generation of policy professions in the US and its partners, we present to you this volume. In it, readers will find the finalists of our Young Professionals Essay Contest, as part of our ongoing program, *The United States & Japan: Allied Against Disinformation*.

Open to nationals of the two countries, the contest received numerous entries and the final decision was not an easy one—especially when comparing our finalists. On the Japanese side, Yuichiro Kotaka raises the threat posed by generative AI and poses solutions for how governments can address it, while Ryohei Suzuki suggests putting the public broadcaster NHK to use. On the US side, Jonathon Marek raises the alarm over the threat posed by authoritarian governments who have recognized disinformation as a valuable tool, while Thomas Ramage offers suggestion on how the alliance itself can address the threat that disinformation poses.

However, our committee ultimately settled on two essays that stood out above the rest: on the Japanese side, Taro Nishikawa's essay on counter-disinformation campaigns in the age of cognitive warfare was singled out for addressing the current discourse within China and for its focus on US-Japan responses. Among the American entries, Rachel Brooks' take on how countering disinformation can be part of school curriculums received praise for its originality, creativity, and for providing both an educator's narrative and an educator's solution.

We thank all of our contestants for their participation, and for contributing potential solutions in this age of information competition.

Rob York  
Director of Regional Affairs  
Pacific Forum



# 1

## **The Case for Anti-Disinformation Education in School Curriculum**

*Rachel E. Brooks*

## Introduction

I began my post-university career teaching high school English at a Catholic high school on Jeju Island, South Korea, with the US Department of State Fulbright Program. Almost 10 years later, I now find myself specializing in disinformation and democracy and working in private sector intelligence analysis. While a winding path, I reflect on my years teaching even more than I expected as a bright-eyed college graduate boarding her first international flight, coincidentally, on July 4.

As an educator, I frequently double and triple-checked my understanding of topics and confirmed my slide decks' accuracy to ensure I did not spread misinformation on a given topic before stepping in front of my students. I regularly thought about how one teacher can influence countless students throughout the span of their career, and those students can further spread those effects in their lifetimes, as research from RAND [highlights](#). The same is true for curriculum quality. Education plays a crucial role in the fight against disinformation and is a scalable way to help young people around the world navigate their information environments.

Disinformation's ability to rapidly transcend countries is reason for concern, especially for global democracies, including the United States and Japan. When societies have fractures to exploit, adversaries will almost certainly attempt to exploit them. One of the best ways to tackle the issue of disinformation is a cross-sector approach involving a coordinated response from government, industry, and civil society working in tandem. Absent stakeholder integration, an otherwise effective approach is likely to flounder.

Rather than be policy prescriptive, this essay seeks to pinpoint the role of civil society. More specifically, it examines how civil society can best bolster an educational approach to disinformation as an effective preventative measure against the challenge of disinformation.

## Contextualizing the Problem

The fragility of democracy in the United States, Japan, and elsewhere stems from some of its greatest strengths, including free speech and public debate. However, those characteristics leave these countries

ripe for manipulation from information operations targeting the countries' existing social divisions.

*The Economist* [cited](#) a "new low" for worldwide democracies in 2022, as repercussions from the pandemic hampered democratic freedoms. An *Economist Intelligence Unit* (EIU) report [uncovered](#) that authoritarian rule accounted for over a third of the global population whereas 6.4% of the world's population enjoyed a "full democracy."

While different methodologies for classifying political systems exist, the *Pew Research Center's* research [leans on](#) EIU's classifications—as does this essay—to focus on full and flawed democracies. Full democracies are those respecting basic liberties and freedoms, with few issues facing how democracy is functioning. Flawed democracies encounter more substantial issues, including low political participation or governance challenges, although they have the basic requirements of civil liberties and free elections. *Pew Research Center's* reporting [classified](#) the United States as a flawed democracy and Japan as a full democracy in 2021, although EIU [categorized](#) Japan as a flawed democracy in 2019.

While the measurements for tracking effective anti-disinformation education may be murky, this approach presents one strategy likely to support critical thinking skills—and help avoid negative ripple effects—as students inevitably approach false information across their lives. Incorporating anti-disinformation education into schools' lessons bakes these essential critical thinking skills into learners at an early age.

Young generations have unprecedented access to technology compared to their predecessors, a shift that [poses](#) immense risks and opportunities. Many children in the United States and Japan are growing up with omnipresent social media and messaging applications and navigating online information environments at increasingly early ages. A *New York Times* article from March 2022 [cited](#) a 17% increase in screen use among teens and pre-teens in the prior two years, representing an increase greater than the four years prior.

In Japan, the domestic messaging application LINE is a critical communication channel followed by globally popular social networks, including Twitter, Instagram, and Facebook, according to data from the [Statista Research Department](#) in October 2022. The United States represents the third-largest social

media audience worldwide, according to [Statista](#) research in 2022, which reported three-quarters of US-based social media users in 2022 stated they used Facebook. Pew Research Center data from April 2021 [reported](#) seven out of 10 Facebook users and approximately six out of 10 Snapchat and Instagram users visit the sites at least once per day, meaning many users are engaging with the platforms with regularity.

Online disinformation campaigns and circulating misinformation are critical threats to global democracies, and their breeding grounds—social media platforms—are pervasive among society. We must equip the generations growing up with unprecedented access to technology the tools to handle the inevitable onslaught of false information.

## **How Early Education Can Help**

A community-based approach in which teachers empower students to discern false information they come across is critical to lasting success. Equipping US and Japanese teachers with easy-to-incorporate, research-backed lesson plans across subject areas will help formalize the fundamentals of identifying, disengaging, and reporting false information.

We cannot yet know the social media environments available when young learners reach adulthood or what new platforms will emerge as popular spaces for the next generations. However, we can equip students with the anti-disinformation toolkit necessary to navigate those spaces early and often. This strategy enables anti-disinformation learnings to grow with students throughout the years. Similar to history, language, and other subject areas, the lessons can build on each other across the years with age-appropriate scenarios and checks on learning. Students must be able to question what they read when they engage with social media platforms, articles, and other content rather than immediately accept it as truth, and this tactic helps keep learners questioning the veracity of what they see on social media.

At the elementary school level, learning outcomes could begin with fundamental concepts, such as teaching the importance of values like honesty and the difference between factual information and make-believe. As students progress into middle and high school, teachers can begin to incorporate more complex lessons built on those fundamentals.

Because disinformation spans subject areas, these lessons could look like the analysis of satirical conspiracy theories in English seminars, real-world examples of past disinformation campaigns in history courses, or viral disinformation in psychology classes.

The curriculum's exercises could include research-backed initiatives from practitioners with expertise in anti-disinformation strategies. We could look to existing efforts from academia, the private sector, and other spaces. [Google's Be Internet Awesome](#), Stanford University's [Civic Online Reasoning](#), and the News Literacy Project's [Checkology curriculum](#) are a few examples offering the building blocks for an effective curriculum to support learning about disinformation in engaging ways in the classroom. The gamification of learning about disinformation can lie at the heart of certain resources, including IREX's [Learn to Discern curriculum and online game](#), and help make learning about this topic fun for students. These and other high-quality resources are ripe for greater incorporation in the classroom.

Materials on common disinformation tactics, telltale signs, and other trends will help offer a basic understanding to US and Japanese learners for application in their own lives. Civil society and for-profit entities are expanding freely-available online resources due to growing recognition of [attitudinal inoculation](#)—also called [pre-bunking](#)—as an effective tool against misinformation and disinformation. Incorporating these resources is one avenue to help students gain familiarity with the signs of disinformation, though the materials used should remain nonpartisan to avoid potential bias or backlash. The unfortunate truth of anti-disinformation education is some people may view the topic as polarized.

It is prudent to avoid politicized vocabulary, including the term “fake news,” and lean on nonpartisan, research-backed sources. The reason for avoiding such terminology is twofold. First, it is too simplistic [considering](#) the distinctions between misinformation, malinformation, and disinformation. The term “fake news” fails to acknowledge that false information often takes root in people's real opinions, fears, and prejudices. Foreign online disinformation would not be successful without those existing emotions available for exploitation, which is why we assert critical thinking skills throughout education are crucial. Second, politicians [weaponized](#) this term to dismiss news they deem displeasing.

In this politicized usage, people often use the expression “fake news” to delegitimize and undermine disagreeable but true content. “False information” offers a less politicized, more accurate term when the distinction among misinformation, malinformation, and disinformation is unclear; however, it is critical we continue to take a nonpartisan view of anti-disinformation education for the greatest likelihood of support from education administrators, parents, teachers, school boards, and other relevant stakeholders.

We do not want to understate the challenges applied to educators to meet myriad learning objectives with finite time and resources. Connecting teachers with free, quality lesson plans could help ease these burdens. Examples that could be taught across a range of contexts and cultures include the Public Broadcasting Service NewsHour’s free lesson plan [introducing](#) a satirical conspiracy theory “Birds Aren’t Real” to teach about disinformation. The lesson [outlines](#) the objectives, warm-up exercises, and main activities and anticipates taking up to “one 50-minute class period.” The ready-made lesson could work well for a teacher or substitute. If assigned as homework, it could relieve additional pressure on a classroom period. While it may seem obvious teachers would want to bolster their students’ critical thinking skills, the reality remains teachers’ time is stretched thin as they aim to meet standards with limited resources. Further awareness on the opportunity anti-disinformation education presents can help educators, students, and organizations alike.

Once students develop the critical thinking skills to identify disinformation, the community-based approach can extend those learnings to others in their information environment, showcasing the positive ripple effects of education. A student’s comment on a peer’s post sharing false information can warn other users about untrue or manipulated content. *The Washington Post* [reported](#) that observable correction lessens misperceptions and reminds the poster and viewers to question the content. Important factors related to this strategy include the proximity of the student’s observable correction to the original post, and they can be more effective when the comment includes a link to a credible source.

Even if other students have already corrected the false information, the student can support the correction to raise its profile and visibility. It is valuable for students to leverage their classroom learnings about common disinformation tactics to

help those in their network without such educational resources and encourage them to question potentially inaccurate content. While they may feel called out by the student’s correction, they are more likely to question content before sharing it again in the future, creating positive ripple effects in the information environment. The prevalence of disinformation allows students to put their learnings into practice, cementing the knowledge into how they approach online information environments long after the lesson ends.

## Conclusion

Inherent in the fight against disinformation is the need to equip the public with tools, trust in institutions, and transparent environments to live up to these responsibilities. Anti-disinformation education plays a critical role in that effort. Still, education alone cannot solve the challenge of rampant disinformation—nor should it. The country’s government, social media companies, and other institutions must play a central role in strengthening regulatory frameworks and platform accountability. Education can be a powerful tool in curbing disinformation, particularly when paired with bolstered standards and regulations.

Students deserve safe online information environments with trustworthy information. In the absence of such an environment, they need an educational toolkit to separate reputable sources and facts from disinformation. The immense scale of the digital environment and its rapid technological evolution are and will remain challenges to these educational efforts. Disinformation-spreading actors will seek innovative new ways to fool users and propagate their content across borders. Despite these challenges, school-based anti-disinformation education across young generations will help ensure the ongoing safety and security of democracy in the United States, Japan, and elsewhere around the world.

# 2

**Breaking the Information Cocoon:**

**Japan-US New Approaches to Counter  
Disinformation Campaigns in the Age of Cognitive  
Warfare**

*Taro Nishikawa*

## Introduction

Foreign disinformation campaigns currently pose a significant threat to democratic nations. Notably, some authoritarian states are rapidly refining their approaches to disinformation campaigns and integrating them into their military strategies. For instance, China has developed a new military concept called “cognitive warfare,” which exploits disinformation campaigns as an active component.

In light of this alarming situation, this paper begins by investigating the disinformation campaigns employed by authoritarian states, drawing insights from the latest discussions among Chinese military scholars. Through this discourse analysis, the paper aims to highlight the potential challenges that such campaigns may pose to Japan and the United States and suggest appropriate policy recommendations for addressing these issues.

## A New Form of Disinformation Campaign in Cognitive Warfare

The Chinese People's Liberation Army (PLA) acknowledges the significance of human cognition, encompassing emotions, beliefs, and values, as a critical domain in modern warfare. Li Minghai, a professor of the National Defense University of PLA has [characterized](#) the cognitive domain as “the main battlefield of the future.” According to his view, the war in cognitive domain, known as “cognitive warfare,” entails the comprehensive use of national resources—including military, economic, political, public opinion, psychology, law, and narrative—to influence target’s cognitive functions, ultimately altering decision-making and behavior in pursuit of national security objectives.

This novel form of warfare has become increasingly apparent in China's military actions in recent years. One notable instance was in its response to the U.S. House Speaker Nancy Pelosi's visit to Taiwan in 2022. China orchestrated large-scale military exercises around Taiwan, cleverly designed to amplify the anxieties of the Taiwanese population by [combining military intimidation, disinformation campaigns, and cyberattacks](#).

In the context of cognitive warfare, how are disinformation campaigns utilized? Notably, PLA theorists, in publications such as *China Military*, have

recently emphasized the concept of “information cocoons” in their discourse on these campaigns, with a steadily increasing attention to the topic over the past few years. Coined by American legal scholar Cass Sunstein, information cocoons [refer to](#) a phenomenon in which individuals become trapped in an environment that exposes them exclusively to information aligning with their preferences, ultimately impairing sound decision-making. The contemporary internet landscape has rendered users increasingly vulnerable to falling into these information cocoons. The rise of algorithms that prioritize information based on users' interests, as seen in social media “recommendation” features and personalized news sites, insulates users from information that challenges their views or habits. This leads to a narrowed cognitive scope and, ultimately, flawed decision-making.

In light of this reality, PLA scholars have proposed a new type of disinformation campaign specifically designed to exploit information cocoons. Tang Guodong, the deputy director of the Political Organization Work Department at the National Defense University of PLA, [argues](#) that by tailoring disinformation content to match each target's thought patterns, it becomes possible to effectively infiltrate a target's information cocoon. This “personalized disinformation” has a higher likelihood of reaching and being accepted by targets compared to other information. Once engaged with the disinformation content, algorithms prioritize displaying similar information within the target's information space. As this process unfolds, the target's information cocoon becomes saturated with disinformation carrying a specific political agenda, skewing the target's cognition in favor of the disinformation sender.

A critical enabler of these “personalized disinformation campaigns” is the recent development of information technology. Researchers at the National University of Defense Technology [suggest](#) that gathering and analyzing various online user data, such as “likes,” “retweets,” and social connections, allows for more accurate tracing of a target's thought patterns. Moreover, advancements in information generation technologies like generative AI and social robots [enable](#) the dissemination of information on a large scale, across multiple channels, at high velocity, and with strong frequency. The combination of these technologies makes it possible to disseminate vast volumes of disinformation tailored to each individual's cognitive



habits, effectively hacking the target's information cocoon.

Ultimately, by conducting such personalized disinformation campaigns, the disinformation sender gains control over the information that the target can access or not, thereby consolidating their influence over the target's decision-making process.

## **Countering Personalized Disinformation Campaigns**

As such new forms of disinformation campaigns emerge, what countermeasures are being employed? In its latest strategic document revision, Japan explicitly outlined its intent to collect and analyze disinformation at all levels—strategic, operational, and tactical—while engaging in strategic communication with domestic and international audiences. For instance, the Ministry of Defense [plans](#) to construct an AI-powered system that facilitates continuous collection and analysis of open-source information. Furthermore, Chief Cabinet Secretary Yoichi Matsuno has [indicated](#) plans to establish a “Strategic Communication Office” within the Cabinet Office by 2024. This office aims to build a new system within the government that consolidates and analyzes information, enhances external communication, and strengthens collaboration with non-governmental organizations. By analyzing internet data and strategically communicating with audiences both domestically and internationally, the government aims to proactively counter the threat of disinformation campaigns.

However, implementing such proactive countermeasures against disinformation also presents challenges. A key issue is striking a balance between proactive counter-disinformation activities and democratic principles. For instance, a news agency in Japan [reported](#) that the Ministry of Defense's new counter-disinformation system is designed to manipulate public opinion by promoting support for government defense policy while suppressing anti-war sentiment among the public. Such concerns about government intervention in the information space are not exclusive to Japan and are common among democratic nations. If governments engage in strategic communication without transparency, their actions may be perceived as propaganda, leading to skepticism about the credibility of government-sourced information.

On the other hand, relying solely on reactive approaches, such as fact-checking, could significantly undermine the effectiveness of counter-disinformation campaigns. Previous research suggests that factual information struggles to keep up with the rapid spread of false information. For example, a study conducted by researchers at the Massachusetts Institute of Technology revealed that false information is significantly [more likely to be retweeted](#) than accurate information. This finding implies that individuals may already be exposed to a flood of personalized disinformation before counter-evidence reaches them. Considering this along with the characteristics of information cocoons, it becomes evident that such reactive approaches would face significant outreach challenges: Even if fact-checkers provide counter-evidence, it may be obstructed by these information cocoons filled with disinformation, hindering its ability to effectively reach the public.

These challenges highlight the significant dilemma Japan and the United States face in confronting the threat of personalized disinformation campaigns. On one hand, they must pursue proactive communication strategies to counterbalance their outreach disadvantage against disinformation purveyors. On the other hand, such assertive approaches could conflict with the democratic principles that underpin our society, potentially jeopardizing governmental credibility and ultimately increasing vulnerability to disinformation. In order to reconcile this contradiction and effectively fight the threat of disinformation, what measures should Japan and the US consider implementing?

## **Policy Recommendations**

To address the aforementioned dilemma, Japan and the US must implement proactive countermeasures against disinformation campaigns while simultaneously constructing a transparent system that enables domestic and international audiences to continuously monitor the governments' communication activities. In this context, this paper proposes two key policy recommendations. First, building up bilateral strategic communication capabilities through the reinforcement of the Japan-US alliance, and second, promoting rule-making to enhance transparency in the information space, achieved through the strengthening of a multilateral regulatory framework for data access.

## **Ensuring Bilateral Strategic Communication Capabilities**

To effectively counter disinformation, Japan and the United States must prioritize the development of preemptive and coherent bilateral strategic communication capabilities. This paper proposes the establishment of Japan-US guidelines on strategic communication to achieve this goal.

Bilateral guidelines on strategic communication could help minimize coordination costs that Japan and the US might face when implementing such communication during various crisis phases. As strategic communication inherently involves a whole-of-government approach, it can lead to intricate and time-consuming bilateral coordination. Particularly during times of crisis, governments must gather, share, and analyze information from relevant agencies while closely collaborating with their allies to devise appropriate messages for communication. As a result, coordination costs would become enormous. Such coordination costs can present significant challenges when trying to execute preemptive communication, which is essential for countering personalized disinformation threats.

Establishing bilateral guidelines on strategic communication can reduce coordination costs by clarifying necessary procedures during different phases of potential crises. Utilizing bilateral exercises and other venues to practice information collection, analysis, sharing, and dissemination based on these guidelines, Japan and the US can further enhance their interoperability in strategic communication. During these exercises, participation should extend beyond military forces to related agencies from both countries, including Japan's Ministry of Foreign Affairs and the US Department of State, as well as Japan's Cabinet Secretariat, and the White House. This allows for continuous evaluation and confirmation of the extent of interagency coordination in strategic communication. Through these efforts, Japan and the United States can bolster their preemptive and consistent strategic communication capabilities, proactively addressing the threat posed by personalized disinformation campaigns.

## **Establishment of Multilateral Regulatory Framework for Digital Platforms**

To combat disinformation campaigns, Japan and the United States should collaborate in strengthening

global regulations for digital platforms, particularly by further institutionalizing the Data Free Flow with Trust (DFFT) initiative. Proposed by former Japanese Prime Minister Shinzo Abe at the World Economic Forum in 2019, DFFT [aims](#) to promote the international free flow of data while ensuring privacy, security, and intellectual property rights. It has received recognition in major international forums such as the G7 and G20, following its introduction during the G20 summit hosted by Japan that same year. Efforts are currently in progress to institutionalize DFFT (Data Free Flow with Trust), and a prominent instance of this is the OECD's 2022 "[Declaration on Government Access to Personal Data Held by Private Sector Entities](#)." This declaration bolsters the initiative by acknowledging governments' entitlement to access personal data in order to safeguard public safety, identify criminal activities, and ensure national security.

Building upon the continuous development of DFFT, Japan and the US should further advance rule-making that reflects the unique characteristics of personalized disinformation threats. Firstly, they should advocate for expanding governmental access to algorithmic data held by private sector entities. In the face of personalized disinformation campaigns, it is critical to scrutinize not just the disinformation contents themselves, but also how these contents are algorithmically curated and arranged within information cocoons. By gaining access to data related to digital platform algorithms, governments can consistently assess the risks associated with these algorithms being exploited in disinformation campaigns, and provide necessary recommendations for improvement to platform providers. This measure effectively prevents hostile entities from weaponizing algorithms, thus neutralizing the threat of personalized disinformation campaigns

Secondly, access to data held by the private sector should also be granted to independent research institutions under specific conditions. This approach not only facilitates the identification of disinformation within the digital domain but also allows independent research institutions to consistently monitor government strategic communication activities. As a result, citizens can stay informed about the extent to which their government's information dissemination adheres to democratic principles. Through the establishment of such an institutional framework, governments can maintain public accountability while proactively engaging in strategic communication.



## **Conclusion**

As disinformation campaigns from foreign entities become increasingly sophisticated, Japan and the US must address these threats proactively and transparently. In this context, the two countermeasures proposed in this paper can serve as complementary approaches. By enhancing bilateral strategic communication capabilities, Japan and the US can preemptively deliver coherent messages to both domestic and international audiences, thwarting the penetration of personalized disinformation campaigns into people's cognition. Simultaneously, the establishment of a multilateral framework for digital platform regulation will ensure accountability for government communications activities and prevent the exploitation of digital platform algorithms by hostile nations. Collectively, these measures will offer an effective response to disinformation while upholding democratic principles.



# 3

## **Digital Allies: How the United States and Japan Can Partner Against Disinformation**

*Tom Ramage*

The United States and Japan are leaders in digital technology and share the common values of democracy, a free and open press, and regard for the right to freedom of speech. Societies which promote these freedoms [gain the benefit](#) of an equitable political and legal system, an informed and educated public, and free and fair election systems resilient against populism and extremism. In the past, institutions in these democracies were trusted with disseminating information to the public to spread ideas and uphold democracy. In the information era, ideas spread instantaneously through decentralized digital networks which carry them at the speed of light. As a result, democracies around the world are at risk of being attacked by disinformation seeking to create chaos and sow discord among their public.

Disinformation has long been used by malicious actors to cause mayhem and division among its targets. Disinformation differs from propaganda in that it intends to deliberately plant false information in a form of political and psychological warfare. Rather than the seemingly benign origin of propaganda, which has its roots in 17<sup>th</sup> century Catholic missions, disinformation is [thought to stem](#) from Soviet *dezinformatsiya* intended to mislead public opinion. Such disinformation can be planted furtively in an attempt to astroturf topics and ideas into the public discourse. It may also be “leaked” (at least ostensibly) from what appear to be credible sources to cause pandemonium or steer public opinion towards a desired outcome in a form of Hegelian dialectic. Artificial order from artificial chaos.

Newspapers, magazines, periodicals, and journals have historically been the mediums for the dissemination of announcements, facts, and rumors. In our interconnected world, the spread of ideas in these mediums has become weaponized by both state and non-state actors for the purpose of undermining democracy, directing hostility, and eroding public credibility. Methods of disinformation are by no means unsystematic but can involve complicated intelligence operations and directives. Leaking stolen documents is part of such disinformation. Seeding information into the public psyche can come through planned declassification revealed ceremoniously to journalistic outfits or leaked furtively into the national dialogue. Forged documents used to promote fear or leaked documents to discredit political enemies can be used to direct the conversation in the public forum. In the Soviet Union,

the act of spreading malicious information to cause doubt or discord in political warfare was part of what were called “active measures.” In the digital age, active measures increasingly rely on unregulated cyberspace to achieve their political aims.

Just 30 years ago, the press was *so-called* due to its ownership of the fixed capital required for the physical dissemination of the printed word. Media was not just about access to information, it was about the machines which printed the papers, the relationships with advertisers to sponsor the pages, the office space for the editors to shout out their deadlines, and the routes for the deliverer to throw the paper. Now, no ink or paper is needed for an idea to spread and the digital paperboy can fling far beyond the imagination of any editor-in-chief of a century ago. Groups which may in the past have had to meet clandestinely to discuss fringe ideas can now do so in the open under the cloak of digital anonymity. Radicalism grows and conspiracies promulgate.

Foreign intelligence outfits will commonly target extremist and fringe groups to proliferate disinformation, hoping their outrage and existing discontent will serve as a viral host for their planted conspiracy theories. The unmoderated nature of imageboards, video hosting-sites, and online chatrooms provide fertile soil for these astroturfed conspiracies to fester and metastasize. The recent Ukraine War Discord leaks are an [unfortunate example](#) of this. Other democracies are affected by targeted influence campaigns as well. In Japan, information seeded by business or political groups intended to discredit rivals or fluctuate markets is commonly *deliberately planted* into tabloids. Such disseminated information is interspersed with the salacious imagery and political gossip of these periodicals, which are not party to the same libel laws or press controls that traditional media outlets would be. Like the operations of tabloid warfare, message boards and mud-slinging forums of digital netizens are free from the reputation gatekeeping holding back traditional media outlets and the rest of the 24-hour news cycle. Without the restraint required by these traditional outlets, ideas can spread faster and without the fact-checking of editors and the rubberstamping of institutional backers. The everyman has their own printing press and any URL which can host images, words, and ideas can have the front page of the public mind.

Many conspiracies planted as a tactic of disinformation spread from these moldy corners of the internet. Conspiracies surrounding the [2016 election](#) and the [origins of COVID-19](#) were routinely first planted among viral community boards before continuing to spread further up the value chain where they were then eventually picked up for their engagement value by a credited source. Foreign actors will oftentimes set up front organizations or obfuscate the sponsorship of news organizations in order to perpetuate their own personal flavor of disinformation. The United States and Japan can work together to resist these intrusions by bolstering democratic institutions by promoting organizations that support the freedom of the press and digital literacy. One way to resist such overt methods of seeding disinformation can be the United States and Japan working together to establish common international practices of labeling state affiliations of news organizations, providing transparency to the sponsorship of research focused groups, and promoting the moderation of extremist corners of the internet.

Existing shared partnerships such as the countries' work in [cooperating](#) on Information and Communications Technology (ICT) and 5G rollout aids in countering disinformation by creating resilient cybernetworks that are better able to detect cyberattacks and outages as well as share data on intrusions which may lead to document leaks and infiltrated networks. The rise of disinformation in political warfare is also inextricably linked with the rise of artificial intelligence. [Bot accounts](#) have the ability to amplify ideas and serve as mouthpieces for the aims of whoever sets them at their deployment. Disinformation can be countered using shared ICT initiatives; collaborating on removing artificial intelligence (AI) accounts and identifying attempts made by hackers to secure information. The United States and Japan can additionally create new bilateral initiatives to work together on countering disinformation. New initiatives can focus on identifying perpetrators, sharing information, and creating joint-task force organizations assigned with promoting cyber-resilience. In creating such bilateral efforts, they should be wary to not repeat the mistakes of previous initiatives to counter disinformation. In April 2022, the United States briefly experimented with the creation of a "[Disinformation Governance Board](#)" (DGB) tasked with giving guidance to US government agencies on countering misinformation and disinformation. Dealing with an unpopular public reaction, the

organization dissolved during its inaugural year amidst widespread [backlash and criticism](#) likening it to an Orwellian "Ministry of Truth." The United States and Japan can avoid the fate of the DGB by taking public opinion into account, decentralizing authority over information, and not declaring what is truth and what is falsity—rather amplifying the transparency of organizations which sponsor research.

The [US-Japan Global Digital Connectivity Partnership](#) (GDCCP) launched in May 2021, elevates diplomatic discussions involving training, cybersecurity capacity building, and secure connectivity in relation to the digital economy. The US-Japan Policy Cooperation Dialogue on the Internet Economy hosted in March 2023 additionally promotes international cooperation in secure networks, data protection, AI, the free flow of data, cybersecurity, digital freedom, and human rights. The commitment to transparent internet governance by both the United States and Japan, which came out of these dialogues, is crucial as both nations seek moderate solutions to counter disinformation. They must balance trusted government access with the onus of protecting the internet from bad cyber actors. The United States and Japan's engagement in the multilateral commitment to [Data Free Flow with Trust](#) (DFFT) as well as their [dialogues](#) in the Global Cross Border Privacy Rules Forum (CBPR) can in the future account for how to jointly share information on and address cyberattacks and malignant bots. Similarly, the United States and Japan can work within the first "Connected Economy" pillar of the Indo-Pacific Economic Framework (IPEF) to foster standards within digital economy. Further dialogues can consider methods to counter disinformation and promote the sharing of intelligence on hackers, front organizations, and "active measures." They can mirror efforts like the US-EU Trade and Technology Council (TTC) to partner with other important economies in the region such as South Korea to create a [US-Korea-Japan TTC](#) to increase transnational cooperation on digital governance. US officials, such as the ambassador-at-large for cyberspace and digital policy, can meet with their Japanese government counterparts to work together to build a more robust, interconnected, and resilient digital environment that can be shared by the United States, Japan, and the world.

Disinformation threatens societies which permit the free exchange of ideas by corrupting the very systems that deliver free information. The digital age has

connected people across the world, allowed the sharing of technology, and facilitated e-commerce which has transformed economies. But when used by nefarious powers, the instant nature of digital technology can become weaponized to undermine democracies by feeding fringe movements with unfounded conspiracy theories. The United States and Japan's democratic institutions can provide a model to the world on remaining robust in a time of increasing domestic political polarization and increasing authoritarianism abroad. The values which form the foundations of both countries' democracies, namely freedom of speech and freedom of the press, will be the same values which propel them forward as leaders of the liberal international order. The support of these values need not become a double-edged sword, but rather can serve to bolster democratic resilience.

# 4

## **A Proposal for Countermeasures against Disinformation: Utilization of NHK as a Public Broadcaster**

*Ryohei Suzuki*

## Introduction

Disinformation, which [includes](#) “all forms of false, inaccurate, or misleading information designed, presented, and promoted to intentionally cause public harm or for profit,” distorts people’s communications in elections and in policymaking. In Japan, disinformation has been [perceived](#) as “not as big a problem as in the US and Europe.” However, as Figure 1 illustrates, interest in this issue has recently grown.

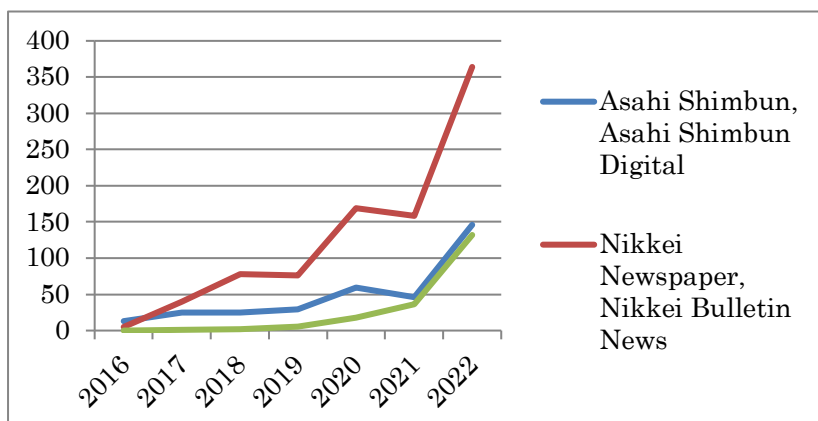


Figure 1. Growing interest in disinformation in Japan (newspaper coverage) Prepared by the author, who counted the number of articles returned by a search for “disinformation” in each medium. For Asahi and Nikkei, the Hitotsubashi University Library database was used. For Yomiuri, the number of articles was researched using [www.yomiuri.co.jp/](http://www.yomiuri.co.jp/).

As the number of articles referring to disinformation increased, the spread of disinformation in Japan became more serious. For example, Japan has been the target of Chinese [influence operations](#), with Chinese online news sites publishing articles in Japanese in attempts to worsen the relations between Japan and South Korea. In addition, according to the [EUvsDisinfo](#), Russia targets Japan with disinformation about the history of Japan-US relations. Since the Russian invasion of Ukraine, there has been growing “whataboutism” in Japan, which [suggests](#) that the West is also responsible for the Russian invasion.

Disinformation has been observed during election periods. According to one analysis, bots were already [active](#) during the 2014 general election of the House of Representatives. In the 2017 general election, Fujishiro [identified](#) 195 instances of uncertain information. Disinformation during elections has been particularly severe in Okinawa, where 70% of the US military bases deployed in Japan are concentrated. In the past election, disinformation related to the US military has been frequently

[confirmed](#), such as the temporary creation of the fake websites “Okinawa Kichimondai.com” (“Okinawa Base Issue.com” in English) and “Okinawa Kenchiji Senkyo 2018” (“Okinawa Governor Election 2018” in English) during election periods, and major mass media outlets mistakenly [reporting](#) disinformation posted by an influencer on social networking sites. This situation is problematic not only because it could inflame public conflicts and make the election results less legitimate but also because it could weaken Japan-US relations.

## Challenges Facing Fact-checking Activities

The Japanese government has led discussions on countermeasures for [social networking service \(SNS\) platforms](#), [strengthening media literacy](#), and [fact-checking](#). However, it has not put forth robust countermeasures against disinformation from the perspective of maintaining freedom of expression. On the other hand, a department was recently [established](#) in the Cabinet Secretariat to devise a legal framework for the introduction of active cyber defense, and the Ministry of Foreign Affairs’ Public Information and Cultural Diplomacy Strategy Division is in the process of [recruiting](#) a person in charge of information dissemination, including countermeasures against disinformation.

Additionally, it has been [recognized](#) that countermeasures against disinformation in Japan “respect voluntary efforts in the private sector.” Fact-checking activities led by major mass media, online media, and non-profit organizations have been the driving forces behind disinformation countermeasures in Japan. Under these circumstances, fact-checking activities face a mountain of challenges. According to the Study Group on Platform Services, fact-checking activities [include](#) problems related to implementation systems, funding, raising awareness and credibility, and international collaboration. The author also conducted a field survey in Okinawa in October 2022 by interviewing two local newspapers: *Ryukyu Shimpo* and *Okinawa Times*. Identified challenges included the absence of dedicated staff, the need for collection and selection criteria for fact-checking



targets, conducting fact-checking without bias, and identifying people affected by disinformation.

## **Solving Issues through Public Broadcasting**

To remedy this situation, some have proposed more active involvement by NHK (the Japan Broadcasting Corporation) in the fight against disinformation. The extant literature highlights the key role of public broadcasting: Patricia Aufderheide [points out](#) the importance of public broadcasting in the fight against disinformation, and Humprecht et al. [argue](#) that countries with strong public broadcasting systems are more resilient.

NHK is a public broadcaster established under the Broadcasting Act of 1950. Operating on the basis of subscription fees from the Japanese public, its mission has been to disseminate information that serves the public. As a result of the discussion by the Public Broadcasting Working Group in 2022, NHK will change its direction to make information dissemination on the Internet a mandatory operation, based on the [belief](#) that “NHK should play a role in delivering reliable information to viewers not only on TV but also through the Internet.” [Specifically](#), “in response to the spread of fake news, NHK will strengthen its function of reporting and producing information to determine the facts” and measures to maintain a healthy information space are being sought. Currently, however, NHK’s commitment to fact-checking is not high. For example, a search for the keyword “fact check” on the Public Broadcasting Service’s News Hour [yields](#) about 500 articles on election debates and fact-checking results by the Associated Press. On the other hand, a similar search on NHK [yields](#) only 17 articles. Its content is also limited to mentioning fact-checking measures and trends, rather than the results of fact-checking.

### **Funding and Personnel Shortages**

NHK is Japan’s largest broadcaster with 54 stations nationwide and 29 overseas reporting bases as of 2022. Compared to 5,000 to 6,000 employees of major commercial broadcasters—such as Fuji Television, TV Asahi, and TBS—NHK has far more personnel, more than 10,000. In addition, NHK has made efforts to develop specialized human resources related to disinformation. When the Great East Japan Earthquake of 2011 occurred, the NHK [set up](#) a social listening team to observe real-time information of

uncertain authenticity on SNS. Through this effort, NHK now disseminates information through a website called the “[Science & Culture Journal by NHK](#)” and a program titled “[Fake Busters](#),” which mainly focuses on fact-checking. A total of 56 articles related to cybersecurity have been [published](#) since December 2018 on the website, mainly by specialized reporters from NHK’s Science & Culture and Network News Department. In addition, NHK is committed to developing professional human resources, including establishing a system for dispatching staff to graduate schools specializing in cybersecurity.

### **Recognition and Credibility**

NHK is the most trusted source of information in the Japanese mass media. According to the Japan Press Research Institute, NHK [returned](#) its top position in terms of trustworthiness in 2021. The Reuters Institute for the Study of Journalism also [showed](#) that NHK is the most trusted mass media based on its brand trust scores.

### **International Collaboration**

In November 2022, NHK joined the international media network, Trusted News Initiative, to collaborate internationally in the fight against disinformation and misinformation. The international network was [created](#) at the initiative of the BBC “to protect audiences and users from disinformation, especially in crisis situations such as elections,” and involves a number of actors; news agencies, such as AP and AFP; public broadcasters, such as the BBC and CBC; platform companies, such as Facebook and Twitter; and think tanks, such as the Reuters Institute for the Study of Journalism. Along with Japan, NDTV of India, Dawn of Pakistan, Indian Express of India, Kompas of Indonesia, ABC of Australia, and SBS also [participated](#). The existence of such networks in non-Western countries is important in compensating for language barriers and the lack of cultural understanding in obtaining findings related in disinformation.

### **Policy Recommendations**

NHK can overcome the challenges faced by fact-checking organizations in terms of funding, human resources, credibility, and international collaboration. Therefore, the first suggestion is to establish an official fact-checking department at NHK. However,

the nature of NHK (similar to [the principle of equity](#)) may not make it easy to engage in these efforts, and there is a possibility that the public may negatively react to fact-checking by broadcasters close to the government. In fact, there have been numerous scandals, including [production fee fraud and improper receipt of travel expenses, over-production in TV programs, misinformation in documentaries,](#) and [pressure](#) by the administration on program content and anchors. More recently, there have been references to the dissemination of false information over the "[Greenless Islands](#)" ("Midorinakishima" in Japanese). Therefore, Plan B, which complements the existing fact-checking function, is proposed. Countermeasures against disinformation are possible not only during the election period but also before and after, and NHK's functions can be utilized. The key elements in successful fact-checking are NHK's two research institutes.

### Pre-bunking Before Elections

The importance of pre-bunking has been discussed in recent years. Pre-bunking, [which](#) is a "means to preemptively refute expected false narratives, misinformation, or manipulation techniques," can be likened to vaccination against a disease. Identifying potential disinformation that may spread in an election and alerting the public in advance will lead to more effective use of limited resources for fact-checking activities. To accomplish this, the NHK Science & Technology Research Laboratories should make full use of their technical capabilities. [The NHK Science & Technology Research Laboratories](#) are "engaged in a wide range of research, from basic to applied, in the field of broadcasting technology." For example, they [conduct](#) research in natural language processing that "analyzes large amounts of text data from social media, program archives, and other sources inside and outside the broadcast station," and image and sound analysis that "automatically reads text appearing in video images" and "identifies 'actions' and 'behavior,' as well as the names of people and objects." Using these technological capabilities, NHK can compile not only a dataset of topics of disinformation spread during past election periods but also the amount and extent of disinformation spread by topic and then makes it available to the Japanese public and mass media. Such an effort could be employed to pre-bunk disinformation and help prioritize which disinformation should be tackled first.

### Fact-Finding Surveys after Elections

One of the challenges associated with disinformation is that it is difficult to ascertain the types of people it reaches and how this affects their perceptions and behaviors. To date, there have been insufficient attempts to compare the influence of such disinformation or understand its reality. To this end, the NHK Broadcasting Culture Research Institute can [play](#) a key role; this institute not only analyzes the media but also surveys public opinion on audiences, society, politics, and daily life, and publishes these findings. This functionality allows it to compile information on the type of disinformation people encounter and whether their thinking or behavior changes after being exposed to disinformation, along with individual demographics such as gender, age group, education level, and income status. Such an effort would not only encourage further progress in disinformation research but also be important for policymaking.

### Conclusion

Internationally, influence operations using disinformation are increasing their presence. However, as Ichida [points out](#), the root of the disinformation problem is a "domestic issue," and it is important to investigate the awareness of the problem in local and regional communities to understand what kind of distortions exist. In this regard, there are high expectations for the role of public broadcasters, which are based in local communities and focus on the public interest. Indeed, PBS has [adopted](#) a framework that supports investigative reporting at the local level. NHK also utilizes its nationwide bases to disseminate information rooted in local communities.

On the other hand, PBS, a public broadcaster, and National Public Radio (NPR), a public radio station, have suspended their accounts on Twitter after being [labeled](#) "government-funded media." This is a problem for public broadcasters, which disseminate information rooted in the public interest, to face such obstacles. In addition, the Republican-led litigation [includes](#) government officials and civilians working for universities and nongovernmental organizations engaged in disinformation research. It is important to avoid this situation from spreading to the media industry and to protect freedom of expression to investigate disinformation and devise better countermeasures. The United States has a

responsibility to showcase the significance of democratic standards both domestically and internationally.

This essay discussed NHK's possible involvement in fact-checking activities. NHK has the capacity to resolve the issues currently facing fact-checking activities. However, there is a strong possibility that NHK's involvement in fact-checking activities will be perceived negatively due to repeated scandals. This is why the second option is also proposed. Therefore, it will be necessary to establish a monitoring system, such as the establishment of a third-party organization, while taking advantage of NHK's strengths, such as its abundant financial and human resources, trust, and international partnerships, and to collaborate with commercial broadcasters, universities, and NGOs.

One way to do this would be to start with regional efforts in Okinawa, where the proliferation of disinformation is conspicuous. It would be ideal if such activities were conducted on a trial basis and eventually expanded nationwide through trial and error.



# 5

## **Parrying the Principal Threat: Reducing the Disinformation Risks of Authoritarian-Influenced Platforms through US-Japan Cooperation on Threat Analysis and Mitigation**

*Jonathon Marek*

## Introduction

The rise of information platforms—which herein will be defined as any social media, messaging, or other digital platform that allows users to publish and share information—while providing numerous economic and social benefits, has also fundamentally and massively amplified the risks associated with the spread of disinformation. This is due to several factors: first, the delegation of the role of publisher and amplifier to individuals has removed this power from institutions with a vested financial or reputational interest in sharing accurate information. Second, the speed at which information can be shared and the exponential nature of its spread across these platforms allows disinformation to attain a much wider audience before it can be countered. These two factors combine to create highly individualized information ecosystems. While in many cases this can be a positive, allowing alternative ideas and viewpoints to spread more freely, it also facilitates the spread of disinformation that is both targeted enough to be effective at the individual level and widespread enough to be effective at the societal level. Finally, and most importantly, the anonymous nature of many of these platforms allows malign actors—in many cases, authoritarian states seeking to shape the information environment in their favor—to easily, cheaply, and covertly spread disinformation that appears to originate from ordinary users.

While the risks of digital disinformation exist across all information platforms, they exponentially increase on those that may be subject to direct influence from these malign actors—most notably, the People’s Republic of China (PRC). The simplest manifestation of this risk is that content moderation on these information platforms—in direct contrast to those in open, democratic societies, which have made genuine, if incomplete, efforts to counter disinformation (especially stemming from malign state actors)—could censor genuine content and fail to take down disinformation. However, an even greater risk is that a similar effect could be achieved less noticeably by manipulation of the algorithms that power these platforms. Minor tweaks could boost the reach of disinformation, catalyzing the exponential spread that makes the digital domain particularly vulnerable in the first instance. Similarly, rather than overtly removing fact checkers or accurate information, this content could simply be silently [demoted](#), achieving the same effect (i.e.,

unchecked disinformation) with less risk of discovery or the negative publicity associated with direct censorship.

Recognizing the novel powers enabled by influence over information platforms, the PRC has enshrined taking advantage of the new vectors to spread disinformation across information platforms as an element of their digital strategy. A recent [submission](#) of evidence to the Australian Senate Select Committee on Foreign Interference through Social Media provides the best overview of the authoritative PRC documents outlining this strategy and its manifestations. As early as 2013, Xi Jinping had [called](#) on the PRC’s leadership to “utilize the role of emerging media”—which includes information platforms—to “meticulously build an external discourse mechanism.” This mechanism should feature “precise communication methods”—a clear indication of an eagerness to adopt information platforms’ ability to share a number of targeted narratives in support of a broader discursive goal. This strategy explicitly and directly calls for the use of disinformation, with the head of a major think tank who led a Politburo study session on the topic [saying](#) that “what is truth and what is a lie is already unimportant.” In short, the PRC, like other authoritarian powers, has a strategy to advance what it calls its discourse power—that is, its ability to influence the international narrative—including through the use of disinformation. However, the PRC’s strategy uniquely emphasizes the internet platforms over which it maintains influence as tools to achieve these objectives—creating a threat that these platforms could be leveraged at any point to spread disinformation.

This is, fundamentally, the nature of risk in the digital domain—the possibility that, at some point, latent influence over a digital tool could be leveraged to achieve a particular outcome. The combination of the PRC’s objectives; clear willingness to use disinformation to advance its preferred narratives, whether on [human rights](#), [Taiwan](#), [COVID origins](#), or other topics; and influence over information platforms is more than sufficient to necessitate a robust policy response even absent any evidence of PRC-influenced platforms being leveraged to promote disinformation.

However, this risk is already playing out across PRC-owned and -influenced information platforms. TikTok [failed](#) to take down 90% of ads promoting disinformation about a US election—a record far



worse than non-PRC influenced platforms (namely, Facebook and YouTube), despite its blanket prohibition on political advertisements. Even more alarmingly, [reports](#) surrounding the PRC's interference in Canadian elections suggest that WeChat was the original source of disinformation campaigns targeting at least one candidate opposed by the PRC. Given the high degree of PRC involvement in the election interference campaigns writ large, and the level of control it exercises over WeChat, it is highly implausible that these disinformation campaigns originated and spread entirely organically.

## Responding to the Risks

The core of the challenge the United States and Japan face in countering these information platforms' role in disinformation is twofold: first, specific information about how they operate is often opaque. While platforms such as Facebook have [taken steps](#) to allow researchers to study the spread of disinformation across their platforms, despite promises to do so, TikTok has yet to allow researchers access to their API. Even if such access is eventually provided, it will be [limited](#) to company-approved research topics unlikely to include whether it actively promotes disinformation. Second, as advocates for a free and open digital ecosystem, both countries must work to address digital disinformation in a manner that does not undermine these core values.

One potential solution is for the United States and Japan to establish a Joint Platforms Analysis Center (JPAC) to analyze and develop recommendations to mitigate the risk of disinformation on authoritarian-influenced platforms. Such a JPAC could bring together researchers and regulators from both countries to designate information platforms as "platforms of concern" based on evidence of authoritarian influence, whether in the form of ownership structure, location of operations, or the actual spread of disinformation. This would facilitate US-Japan information sharing, raising both countries' understanding of the risks posed by these platforms. JPAC could then be authorized to analyze these platforms' source code and conduct other research and investigations as necessary to determine if and how these platforms promote disinformation—either directly or by undermining accurate information. If JPAC finds a platform is promoting disinformation, it would provide recommendations to mitigate the risk. Of course, if the risk is an inherent result of the

platform's ownership, this recommendation would necessarily rise to the level of a forced divestment or ban. While proposals for such measures targeted against individual platforms—most notably TikTok—are widespread, the JPAC model would provide three central advantages: first, it would create a process that is applicable to all authoritarian-influenced platforms, rather than ad hoc, uncoordinated efforts against individual platforms only after those platforms take off. Second, the simple fact that it is not a unilateral action increases its credibility. Finally, and most importantly, it would provide a clear, evidence- and rule-of-law-based process to justify any restriction—ensuring that counter-disinformation efforts align with the core values of open, democratic societies.

An example of such a program already exists, and has proven successful at identifying and mitigating the risks of another major PRC technology company's malign activities. The Huawei Cyber Security Evaluation Centre (HCSEC), established in the United Kingdom in 2010, examines the cybersecurity risks of Huawei's products. It has [identified](#) significant deficiencies in the company's cybersecurity practices, leading to some specific, tangible fixes being implemented. More importantly, HCSEC's work has undeniably [played a role](#) in the decisions by countries, including the UK, to restrict Huawei's role in 5G networks. While there are elements of HCSEC's model—namely, its exclusive focus on one company and that company's involvement in its operations—that a JPAC could improve on, it is clear that this approach can prove successful in identifying and countering the risks of authoritarian-influenced technology.

Beyond the possibilities of bilateral counter-disinformation cooperation a US-Japan JPAC offers, it would also provide the two countries an avenue for leadership within the Quadrilateral Security Dialogue ("Quad"), strengthening that grouping's growing focus on emerging technology issues. India and Australia represent logical partners for the United States and Japan as both are also attuned to the risks posed by PRC platforms. Australian researchers were [some of the first](#) to highlight the risks associated with TikTok and WeChat, and the country's [Select Committee on Foreign Interference through Social Media](#) has gathered significant evidence on PRC influence over platforms and associated risks. Conversely, India has taken [aggressive action](#) to restrict PRC platforms—

including several less well-known information platforms—but these actions have [faced criticism](#) for insufficient justifications of the risks behind some of the targeted platforms. Combining US and Japanese leadership with Australia’s analysis and India’s willingness to aggressively mitigate risks could create a singularly effective JPAC.

## **Alignment with Shared Data Governance Priorities**

The efforts of a JPAC also align with US and especially Japanese data governance priorities, the latter under the framework of Data Free Flow with Trust (DFFT). While an ambitious set of principles, DFFT faces [challenges in operationalization](#), including significant ones related to questions around the definition of “trust.” By working to define disinformation and contain information platforms that promote its spread, a JPAC could increase both US and Japanese confidence in the information flowing between the two countries. More broadly, it could help to develop trust globally through transparent, public analysis of some of the information platforms that are among the most noticeable manifestations of cross-border data flows to global publics.

Aligning a JPAC’s work with data governance priorities is also necessary to address one obvious concern with moves to restrict authoritarian-influenced platforms—namely, that arbitrary restrictions undermine the free flow of data. TikTok has [argued](#)—misleadingly at best—that it is a US company; if, for instance, Japan were to require it to disclose its source code without a specific legal or regulatory counter-disinformation authority, the company could argue that this violates the [US-Japan Digital Trade Agreement](#)’s prohibition against mandatory source code disclosures. While it is implausible that the United States would even consider raising a dispute on behalf of TikTok, such a move could nevertheless undermine the credibility of the principles behind that agreement’s commitments. However, if the United States and Japan establish a JPAC and examine PRC platforms under its auspices, that would clearly fit under the agreement’s exception given “a specific investigation ... (or) enforcement action.” A JPAC could also provide guidance on how best to address disinformation and other harmful forms of data under the auspices of digital trade agreements and other data governance cooperation efforts, allowing counter-disinformation

to be built in to the future digital agreements on which Japan in particular is taking a leading role in the Indo-Pacific.

This piece began with the premise that the rise of information platforms—especially authoritarian-influenced ones—had fundamentally elevated the threat of digital disinformation. To counter this threat, likeminded partners need to develop a common operating picture of the nature of these platforms, their operations, and the risks these pose—and to jointly develop measures to mitigate these risks. As highly digitalized economies with a shared interest in promoting democratic values, countering authoritarian disinformation, and shaping digital rulemaking in the Indo-Pacific, it is necessary for the United States and Japan to take a leading role in these efforts. The establishment of a JPAC would represent an effective step in this direction, lending itself to both bilateral cooperation and easy multilateralization through the Quad framework.



# 6

## **Countermeasures against the Spread of False Information by Generative AI**

*Yuichiro Kotaka*

## Introduction

In recent years, the expanded use of social networking services has made it easier to spread disinformation. Furthermore, the expanding use of AI has given rise to AI-based image- and text-generation services such as Chat GPT, which have contributed greatly to work and creative endeavors, but have also made it easier to spread the truth about false information. To solve the problem of disinformation proliferation, fact checks by the public sector and private organizations have traditionally been considered useful. However, there is a limit to what a small number of public sector and private organizations can do to check facts, one at a time, when faced with a large amount of disinformation.

This paper focuses on the spread of disinformation on social networking services (SNS) and discusses disinformation by generative AI, as well as conventional countermeasures. It will then identify effective countermeasures.

## Popularity of SNS

Most citizens in both Japan and the US now have information appliances, and many voters use SNS. Many citizens do not pay attention to the reliability of information in a society where the information flow is distributed at a high tempo. Moreover, in this new information society, information that is easy for people to understand, i.e., emotional and sensational information, is more likely to be chosen over accurate information. Furthermore, in “echo chambers,” where the content one can see from on their SNS pages is nothing more than what is of interest to that person, it is easier to disregard the reliability of the information and believe the narrative content. A Massachusetts Institute of Technology [study](#) published in 2018 concluded that false stories consistently achieve greater reach than factually correct news on the social media platform Twitter (now known as X).

## Types of Disinformation and Countermeasures

Disinformation is disseminated with malicious intent, interspersed with distorted or true information, so that certain segments of the population are not aware of the obvious distortions. For example, some of the most successful Russian proactive covert operations

have contained completely accurate information. In Japan, in September 2018, false information from the People’s Republic of China spread on social networking sites regarding the evacuation of people from Kansai International Airport after Typhoon Jebi caused damage, which [led to accusations](#) against the Taiwanese government among Taiwanese citizens, and the [suicide](#) of a Taiwanese diplomat.

Given that in recent years, Japanese and US experts have discussed a scenario in which the Chinese People’s Liberation Army would invade Taiwan in the event of a large-scale disaster in Taiwan, using the dispatch of Chinese troops to the island as a pretext, it is entirely possible that the spread of disinformation would be used to [prepare for an invasion](#).

Fact-checking is considered [an effective way](#) to halt spread such disinformation. By having a trusted institution (such as a traditional leading TV station, newspaper, NPO, etc.) declare information false, the spread of disinformation can be prevented. In recent years, the [Japan Fact-check Center](#) has been established in Japan, but the number of fact-checking organizations is [smaller](#) than in Europe and the United States.

## Emergence of Generative AI

Today, AI can easily be used to generate images and sentences. For example, Chat GPT can instantly generate accurate sentences in English and Japanese and easily generate sentences that are natural to the author, even if they are not native speakers. With technology of this standard, those who are hostile to Japan and the US can easily and intentionally spread disinformation by impersonating fictitious persons or institutions in either country, thus enabling the creation of easily believable stories. In recent years, with the advent of image-generating AI, fictitious people and events can now be shown in realistic videos and content. This poses a major threat to fact-checking.

In the past, the spread of disinformation on social networking services was dominated by “bots,” or accounts that transmit information automatically. While it was easy to spread disinformation from multiple accounts, the content of the messages was similar and the expressions strange to native speakers, so even the general public could identify these automatically, or employ means of identifying them.

However, in recent years, false information using generative AI has become indistinguishable from genuine messages to the general public. For example, it is possible to add authority to disinformation by using an AI-generated image of a fictitious human as an icon and pretending to be a member of the general public or an expert, or to post text on a SNS in a form like that of a native speaker.

This disinformation can be crudely produced in large quantities. In fact, Chat GPT and other such systems can instantly generate sentences, and since the volume of sentences posted on SNS is small, they can be used in conjunction with “bots” to spread a large amount of disinformation.

[Japan](#) and the US have regulations that are more lenient regarding this type of generative AI than in [China and Europe](#). For this reason, most generative AI enterprises are in the US, and some companies have recently established bases in Japan. This regulation seems to be primarily focused on the development of technology in a free development environment. The Japanese government is increasing this trend, with the prime minister [inviting](#) CEOs of generative AI companies and reinforcing its AI strategy in its official documents.

## **What Measures Would Be Effective to Achieve This?**

Both (1) countermeasures to limit the spread of organized disinformation campaigns and (2) regulations on companies with generative AI technology are needed.

First, this paper defines disinformation that must be addressed as false information spread with the intent to disrupt society (i.e., with malicious intent) and generated in large quantities by generative AI. This differs from misinformation—false information spread without malicious intent.

1. It is clear from the above that the spread of disinformation involves the posting of narrative content with the intent to disrupt society. To achieve this, it is not enough for a specialized organization to conduct fact-checking. Also, generative AI is still in its infancy and can be detected. In fact, the “add background information” that Twitter (now X) is doing involves writing a warning on the same screen as the posting of false

information. Initially a tool for supplementary explanations, it now serves to correct disinformation and misinformation.

It is important for the Japanese and US governments to immediately ask domestic SNS operators to include the text of the alert on the same screen as the false information, as in the case of this “addition of background information.”

To ensure efficacy, reporting requirements should be imposed on the amount of offending content and efforts to prevent the spread of disinformation. For example, in Germany, the Network Enforcement Law (“Netzwerkdurchsetzungsgesetz”) has been imposed on commercial operators of Internet platform services that allow users to share any content with other users or make it available to the public. The law requires that the operator of an Internet platform service available to the public report the number of violations and the number of violations removed.

In Japan and the US, where the freedom of communication is valued, this kind of enforcement may not be appropriate, and the steps taken should be left to the individual operators.

However, the government should monitor disinformation on a regular basis so that it can respond to problems as soon as they arise.

Furthermore, it is also important to make government information easy to look up. Government documents should not only be displayed in an easy-to-understand manner but also be easy to look up. For example, in the UK, all departments have a unified website such as “gov.uk,” and the government as a whole works to disseminate accurate information. On the other hand, in Japan in particular, information is provided by individual ministries and agencies. In addition, the government's information is not titled, making it more difficult to know the topic unless the user opens the link.

It is necessary for the government to create a unified website, abolish PDFs, and add titles to government information so that it can be easily and clearly examined.

2. The formation of a framework to prevent the transfer of technology by companies with generative AI technology.

While the involvement of foreign governments in the spread of disinformation has been pointed out, it is dangerous for companies with such technology to cross over into countries other than the US, Japan, and their partners.

To this end, a notification system should be created for companies with generative AI technology so that the government can have some grasp of the situation. It is also important for both countries to establish export controls and an international technology management regime within the US alliance for AI-related technologies.

To do this, conferences should be held between the US and Japanese governments and generative AI companies. While sharing information between the US and Japanese governments and generative AI companies, the details of notifications and regulations should be discussed at these conferences. The results of these discussions should be made into de facto standards, which would then be sublimated into an international regulations. Since many of the current generative AI companies are located in both Japan and the US, and the executives of these companies are often citizens of both countries, it is possible to establish such a conferences.

Currently, the EU has issued strict rules (ex. AI Act) against AI, while Japan and other Asian countries want looser regulations. To take the lead in this rulemaking, de facto standards should be created and both Japan and the US should lobby the G7, while both countries should take the lead in rulemaking through regional cooperation.

## Conclusion

The spread of disinformation campaigns is due to the influence of social networking services, which individuals may transmit, as well as the ease of believing in narrative content, and the fact that people no longer pay that much attention to the reliability of information. And now that the use of generative AI is expanding, it is possible to coarsely produce narrative content and more easily create confusion in society.

For this reason, it is important to create a mechanism for the general public to participate in fact-checking, as described above, for governments to communicate in an easy-to-understand manner, and to establish an international control regime for technology regarding generative AI.

## ABOUT THE AUTHORS

**RACHEL E. BROOKS** is a risk intelligence professional specialized in disinformation and other digital threats to democracy. Rachel is currently a Project Manager at Microsoft's Democracy Forward Initiative and previously served as a private sector intelligence analyst, Fulbright grantee in South Korea, and Rotary International Peace Fellow in England.

**TARO NISHIKAWA** is a Master's candidate in International Affairs at the Hertie School in Berlin. He previously served as an Officer at the Japan Ground Self-Defense Force, specializing in Military Logistics and Liaison for over four years.

**TOM RAMAGE** is an Economic Policy Analyst at the Korea Economic Institute of America and a 2021 Pacific Forum Young Leader

**RYOHEI SUZUKI** is a doctoral student at the Graduate School of Law, Hitsotsubashi University. He has published various Japanese translations including chapters 5 through 7 of *Ill Winds: Saving Democracy from Russian Rage, Chinese Ambition, and American Complacency* and "The Sunnylands Principles on Enhancing Democratic Partnership in the Indo-Pacific Region."

**JONATHON MAREK** is a Project Manager at the National Bureau of Asian Research, where he oversees research on Indo-Pacific digital governance, emerging technology, trade, and intellectual property issues. He is a graduate of Georgetown University's Walsh School of Foreign Service, and his academic interests include the role of economic statecraft in grand strategy; the intersection of technology, trade, and national security; and China's economic and technology ambitions.

**KOUTAKA YUICHIRO** is a second-year master's student at Tohoku University Graduate School of Public Policy researching Japan's economic security.



**PACIFIC FORUM**  
INTERNATIONAL

**PACIFIC FORUM INTERNATIONAL**

1003 BISHOP ST. SUITE 1150, HONOLULU, HI 96813

(808) 521-6745 • [PACIFICFORUM@PACFORUM.ORG](mailto:PACIFICFORUM@PACFORUM.ORG) • [WWW.PACFORUM.ORG](http://WWW.PACFORUM.ORG)