

# **Presentation Abstract: Addressing AI-Driven Mis- and Disinformation Risks at the Local Governance Level through Coordinated, Youth-Empowered Policy Projects with City and State Governments**

*Authors: Carolyn Wang\*, Kaitlyn Cui\*, Kayla Huang\*, Mo Akintan\*, Sofia Rodriguez, Katherine He, Sam Jung, Jennifer Wang*

Digital disinformation and misinformation pose unprecedented risks to democracy amidst the rise of generative artificial intelligence. More than ever, it is imperative for governments to establish methods to mitigate the risks imposed by information manipulation and propaganda. Citizens rely on cities and counties as a source of information for essential services, many of which require intergovernmental communication and civilian input. However, these essential services are increasingly susceptible to threats from textual, audio, and video deep fakes.

Many smaller city and county governments lack the resources necessary to deploy teams to address such risks. As information dissemination increasingly shifts to youth-dominated social media platforms, there is also an urgent need to establish a talent pipeline that engages a younger generation of tech-savvy, civically-engaged students in conversations around disinformation and democracy.

This abstract introduces the Paragon Policy Fellowship, a nationwide initiative that integrates technically-proficient, civically-engaged college students into policy projects with local government agencies to tackle emerging technology issues, including AI-powered media. We explain how two successful case studies from the fellowship, derived from the Paragon Organizational Model, serve as a proof-of-concept for addressing complex disinformation issues at the local level. These case studies, completed in partnership with the City of San José, empower internal government employees and external community members to collaboratively guard against harmful media disinformation. We argue that the project-based Paragon Fellowship illustrates an adaptable civic engagement model to successfully engage youth with policymakers in local government capacity-building surrounding disinformation harms.

## **The Paragon Model**

The Paragon Model operates on three pillars to support successful collaboration between governments and youth constituencies:

- (1) **Empowering local government through a community-driven approach:** The Paragon team matches local students from diverse academic backgrounds with their local governments for four months to conduct in-depth research and produce a policy brief advising on a technology policy issue specific to the local community, enabling governments to address AI issues they lack the capacity to address on their own.

- (2) **Increasing and maintaining diversity within the AI governance space:** Novel, comprehensive policy solutions necessitate diverse perspectives and backgrounds: 75% of Paragon fellows are from underrepresented backgrounds, 41% are first-generation students, and 34% are Pell Grant-eligible. Throughout the fellowship, fellows have access to policy directors with extensive government work experience, speaker events featuring AI policy experts, and office hours providing mentorship.
- (3) **Building a grassroots-first network for digital cooperation, literacy, and advancement with potential for global scalability:** To foster cross-cultural communication, Paragon created an online community to empower discourse, collaboration, and knowledge-sharing across diverse AI governance-focused issues.

### **Paragon Case Studies**

This abstract highlights two projects which address different channels of disinformation within the context of local governments. The City of San José, situated in the heart of Silicon Valley, is the birthplace of AI advancements that have revolutionized the creation of multimedia communications. As AI has evolved, it has also heightened concerns about the potential for constituents to fall prey to mass-generated AI content. In Spring 2024, two Paragon fellow teams worked with San José to design strategies for mitigating risks associated with disinformation originating from (1) internal intergovernmental communications and (2) the external public-input process.

- (1) **Mitigating AI-Powered Misinformation for Government Communication:** Paragon fellows collaborated with San José over four months to create a policy manual for social media staff with three key misinformation response methods: policy approaches addressing mis/disinformation, technical tools for AI detection and response, and public education and media literacy initiatives to encourage community involvement. These guidelines were presented to internal City staff to implement including comprehensive incident documentation, authenticity verification of official government communications, and public awareness via a social media campaign.
- (2) **Mitigating Harms of AI-Generated Misinformation in the Public-Input Process:** In four months, the San José Paragon Fellows created a policy manual that provides guidance aligned with San José's policies on public participation, proposing a standard operating procedure, guidelines, and educational initiative to minimize the impact of bot-generated comments on government staff and processes. This comprehensive report was presented to the city council and relevant stakeholders.

Existing AI governance strategies lack emphasis on community-centered solutions tailored to address how local governments are uniquely impacted by and well-positioned to address disinformation challenges. The Paragon Model effectively synthesizes two integral tools for the fight against synthetic media and disinformation — firstly, constructing effective models for reinforcing information integrity at the local level, and secondly, investing in the development of a younger generation of technically-minded policy professionals.

*\*Primary Authors*