

The Future of Disinformation Summary

Introduction

This is summary of the key themes, important ideas, and significant facts presented during *The Future of Disinformation*. The event, hosted by Melbourne Connect, brought together experts from various fields to discuss the multifaceted challenges posed by disinformation in the current environment. The discussion highlights the evolving nature of disinformation, its underlying psychological and social drivers, the limitations and potential of traditional and technological countermeasures, and the urgent need for societal resilience and education.

Overall Theme

The overarching theme of the event is that disinformation is not simply a problem of factual inaccuracy, but a complex psychosocial phenomenon weaponising human vulnerabilities and amplified by new technologies. Combating it requires a holistic approach that goes beyond fact-checking and addresses the psychological, social, technological, and regulatory dimensions.

Key Themes

The Nature of Disinformation

- **Beyond Factual Falsity:** Disinformation is defined as information deliberately spread to change a person's stance for financial or political benefit. Crucially, it is not always factually false. It can involve selective emphasis and omission of true information to create a misleading image. As Eduard Hovy states, "it's not so much the question of facts and factivity and truth. It's...also a question of what is not said. It might be that everything that's said is true. But the emphasis is strange, and some other truths are not said, which creates a completely different false image."
- **Psychosocial Drivers are Primary:** The driving mechanism of disinformation is not necessarily the factual content, but the exploitation of people's psychosocial, economic, and educational vulnerabilities. Susceptibility is linked to individuals' attitudes and life situations.
- **Exploiting Doubt, Fear, and Anger:** Disinformation campaigns often work by first moving people into a state of "doubt or fear, or anger or confusion," undermining their sense of certainty. Once mobile, they are then offered a new space of "certainty, where you'll offer them clarity, you offer them conviction, you offer them a symbol... you offer them a hero." This new space is woven into a comprehensive belief system, making it very difficult to dislodge. Eduard Hovy emphasises, "we're not facing facts and factivity, we're facing this."

The Evolving Information Landscape

- **Shift from Top-Down to Peer-to-Peer:** The traditional model of information dissemination with gatekeepers like newspapers and governments is changing. We are moving towards a "peer-to-peer model where rather than being just recipients of information, individuals are participants in its creation and its spread." Eliot Higgins highlights this as a "very different model."
- **Impact on Young People:** Young people, growing up with social media as their primary information source, are particularly affected. They often lack critical thinking skills needed to assess online information and are influenced by "news influencers" rather than trained journalists. Eliot Higgins notes, "if we're just going to have this situation where we assume kids know how to use the internet because we give them a smartphone when they're four years old, poking at icons...in reality, they aren't equipped with any skills to actually make assessments of the information that they're seeing online."

- **The Rise of Generative AI:** Generative AI represents a third revolution in information production, following the printing press and social media. It enables anyone to easily create content, making it "possible to anybody to write for it." Preslav Nakov observes, "we didn't quite cope with the social media problem, and now we have GenAI on top of this".

Combating Disinformation: Challenges and Approaches

- **Limitations of Traditional Fact-Checking:** While important, fact-checking alone is insufficient, disinformation often thrives on emotional and social appeals. As Lydia Khalil argues, "we can't fact check our way out of this problem." Fact-checking can also inadvertently amplify disinformation or increase distrust in institutions."
- **Open-Source Investigation:** Eliot Higgins discusses the development of open-source investigation techniques like geolocation and chronolocation to verify information by cross-referencing publicly available data sources. This approach looks for "digital echoes of reality" and quickly exposes fabricated information because it "doesn't fit with the rest of the information."
- **Technological Solutions:** Preslav Nakov describes advancements in natural language processing for fact-checking, including systems that can detect claims, assess check-worthiness, retrieve evidence, and even generate explanations and fact-checking articles. However, he highlights that "fact checkers actually really need help in the very first steps [claim and check-worthiness detection] and in the last step, writing the article." The development of generative AI also presents new challenges, such as detecting uncertainty in generated text.
- **The Symbolic Nature of Disinformation:** Greg Nyilasy emphasises that focusing solely on the factual content of disinformation misses its deeper symbolic meaning. For example, a false claim about immigrants might be a symbol of fear of immigrants or latent racism. Disinformation acts as a "power play," a way for propagators to exercise agency and control narratives.
- **Opportunistic vs Planned Disinformation:** While some disinformation campaigns are planned, much of what spreads, especially on social media, is opportunistic. A swarm of ideas is released, and those that gain traction are amplified (survivorship bias).

The Role of Trust and Institutions

- **Erosion of Trust:** A significant factor contributing to the spread of disinformation is the distrust in traditional sources of authority. Eliot Higgins believes the cause of disinformation is "more related to distrust in traditional sources of authority, that's then compounded by a form of personal, political, social, whatever crisis."
- **The Importance of Trusted Sources:** Re-establishing and maintaining trust in reliable sources of information is crucial. Andrew Perfors argues that the way out is for there to be "trusted people that the people trust as a source of news." Universities are suggested as potentially trusted institutions due to their dedication to finding and communicating truth.
- **Civic Deficit and Complacency:** A lack of civic education and general complacency in society contribute to vulnerability to disinformation. Lydia Khalil suggests that in Australia, "we are complacent in a way" and risk losing our democracy if we don't actively safeguard it.

Regulation and Resilience

- **Existing Regulatory Frameworks:** While not fully keeping pace with technology, existing regulatory frameworks exist in Australia, including voluntary industry codes, broadcasting codes, and criminal laws against foreign interference and espionage.

- **Challenges in Regulation:** Legislating against disinformation is challenging, particularly regarding political misinformation, due to concerns about censorship. Digital platforms are often seen as ideal regulatory targets due to their role in amplifying content. However, platforms often resist regulation and are reluctant to disclose information about their algorithms.
- **The Need for Transparency:** Transparency from digital platforms about how content is targeted and amplified is important.
- **The Role of Local Government:** Cities and local governments are increasingly impacted by and responding to disinformation, often at a hyper-local level. They can play a crucial role in pre-emptive work, addressing social fault lines, building trusted networks, and supporting community resilience and recovery.
- **Media Literacy and Education:** There is a strong consensus on the need for increased media literacy and education for people of all ages, not just in schools. ACMA acknowledges the need for more outreach and education in this area.
- **Beyond Institutions:** Building resilience also involves fostering community connection and active participation in society through various non-digital activities.

The Disinformation Chatbot Demonstration:

A significant and impactful demonstration involved a disinformation chatbot avatar of Jon Faine. This AI was programmed to be contrarian and manipulate information to induce doubt, fear, anger, or strengthen social group belonging.

The demonstration highlighted:

- The ease of creating convincing deepfakes and disinformation: From a small amount of training data, a seemingly convincing avatar and chatbot could be created.
- The manipulative tactics used: The chatbot employed emotional appeals, appeals to authority (even fabricating expert citations), and reinforced pre-existing biases.
- The speed and scale of potential disinformation generation: The AI could quickly generate misleading responses tailored to different stances and emotional triggers.
- The potential for undermining trust: The demonstration unnervingly showed how easily a seemingly credible source (Jon Faine's avatar) could generate and disseminate disinformation.

Quotes of Note:

- Eduard Hovy: "Disinformation is something that's so pernicious and so subtle that it gets into you, and you don't even know that you've had it there."
- Eliot Higgins: "We focused far too much on foreign actors and particularly Russia and seeing that as the cause, when I think the cause is more related to distrust in traditional sources of authority."
- Preslav Nakov: "The best manipulation is when you tell the truth, only the truth but not the whole truth."
- Imogen Crump: "I think what's happened now is the amount of white noise we get makes that fact checking incredibly difficult."
- Lisa Watts: "I think the best thing you can do is flood the zone with good quality information."

- Andrew Perfors: "Disinformation is not primarily a problem of stupidity or laziness or ignorance. The problem arises because our cognition and emotions evolved to be adapted to a very different social and informational environment than we're in and disinformation campaigns learned how to weaponise that."
- Andrew Perfors: "A lie has an advantage over the truth in that lies are not constrained by what's actually true. So, it's easy to make short, sweet, simple lies. Truth is often complex."
- Lydia Khalil: "As important as fact checking is, this really is a social problem...we can't fact check our way out of this problem."
- Lydia Khalil: "Disinformation is not something to influence, but something to confuse. We have to think of it as power play, a form of power projection."
- Greg Nyilasy: "It is not cats and dogs...it's not them that matter, and it's not that factual content that matters. There's a deeper symbolic meaning that matters in that."
- Greg Nyilasy: "It is indeed a power play to express my agency to say whatever I want to say. Because if I'm not held by the truth, I can just create, I have more options."
- Jeannie Paterson: "Who amplifies and spreads that misinformation? It's the digital platforms and the technology companies. And they are an ideal regulatory target for the law."
- Ika Trijsburg: "Cities are hugely impacted by disinformation."
- Ika Trijsburg: "We know a lot of the social fault lines that disinformation is likely to exploit, so we can do a huge amount of work before anything happens to try to address some of those fault lines."

Conclusion

The event effectively highlighted the complex and evolving nature of the disinformation challenge. It moved beyond a simplistic view of lies versus truth, emphasising the psychological, social, and technological factors at play.

The discussion underscored the limitations of current approaches and pointed towards the need for a multi-pronged strategy involving technological solutions, robust regulation targeting platforms, increased media literacy, and, fundamentally, a renewed focus on building societal resilience and fostering trust in credible institutions and community connections.

The demonstration of the disinformation chatbot served as a potent reminder of the immediate and future threats posed by AI-powered disinformation and the urgent need for action. The call for greater civic participation and community engagement resonated as a crucial element in safeguarding democratic processes against the corrosive effects of disinformation.