Fighting disinformation in a pandemic world:
the role of AI and cognitive sciences

A 2-day Advanced Research Workshop

Rome, November 16-17, 2021

Organized by Herbert Simon Society and Kyiv School of Economics (KSE)

During the COVID-19 pandemic, it soon became apparent that the viral contagion was paralleled, or even surpassed, by the corresponding spread of fake news on the ongoing crisis: as usual, part of this was just the unwanted consequence of poor epistemic practices and widespread lack of digital literacy in Internet users (misinformation); however, another portion of this harmful digital clutter was deliberately produced and distributed by malicious agents, either for profit, ideological reasons, or even a perverse idea of fun (disinformation). Moreover, the negative impact of this distortion of accurate information was not limited to the acute phase of the COVID-19 outbreak, but continued also during subsequent recovery stages: for instance, international aids sent by foreign countries were often accompanied by intense social media campaigns via unofficial agencies, aimed at leveraging a friendly act of support to promote the PR agenda of the helping country, thus potentially hiding the fist of disinformation within the glove of international cooperation.

This strong correlation between pandemic outbreaks and deliberate disinformation should not be surprising: by its very nature, a pandemic like COVID-19 creates a climate of social anxiety, personal fear, political radicalization, global unrest, and opinion polarization, all of which are excellent tools in the hands of well-trained agents to conduct so called "information operations". The global and prolonged nature of the crisis further strengthens the potential impact of such operations, and the fact that future outbreaks of contagion agents akin to COVID-19 are not only possible, but even likely, increases the chances that pandemic-related disinformation may become a key asset in digital warfare. This risk is magnified by the fact that the sanitary restrictions needed to fight against a pandemic (e.g., lockdown and social isolation) strengthens our dependence on Internet technology and social media, which are the natural breeding grounds of online disinformation: thus the same strategy that protects us from physical contagion makes us more vulnerable to digital aggression. For all these

reasons, proper response protocols and routines must be urgently devised, using the disinformation strategies individuated in the current crisis as testing grounds to face future, similar threats.

In the quest for adequate counter-measures against disinformation in a pandemic world, the role of AI and cognitive science must be central, on pain of failure. AI systems can help create safer and higher quality digital platforms for citizens to engage with a pandemic crisis, both on their own and at the collective and institutional level, whereas cognitive sciences can offer key insights on how to ensure that users actually adopt the available technology as widely and correctly as possible. Moreover, AI and cognitive sciences can collaborate to the fight against disinformation also at a deeper level: understanding the socio-cognitive determinants of our vulnerability to false information is a necessary precondition for the design of more robust and safer platform for well-informed public engagement, whereas the new digital technology created by AI applications constitutes an essential factor for our understanding of cognition in an interconnected world. Thus the definition of protocols to fight online disinformation on pandemic threats must involve AI and cognitive sciences, not separately from each other, but jointly and taking full advantage of their integration.

Agile and evidence-based government decisions on how to design information policy with the assistance of data science and AI. Innovative data collection and NLP tools can help the governments quickly identify the main myths, dominating the public discourse and help design the best counter-measures. In the long term perspective, the pandemic is doing a great job in promoting a culture of evidence-based decision making. The public is becoming more reluctant to trust the words and is closely monitoring the deeds of the policymakers thus demanding more informed decisions and AI and data science can help in that.

This workshop will bring together world-leading scholars from AI, cognitive sciences, and related disciplines to tackle these issues, with the specific mandate of formulating clear guidelines for implementation by NATO and Partner countries, as part of their ongoing response to both digital and pandemic threats and their associated risks. **The list of topics** to be explored features, among others, the following:

- AI-solutions to support better information quality: how to harness the power of data analytics and machine learning to offer clearer, more authoritative, and more persuasive information channels, and how to make such channels resilient to digital infiltration.

- AI-triggered epistemic vigilance: how to design information platforms in order to trigger the cognitive mechanisms involved in the critical assessment of news, and how to avoid promoting instead cognitive laziness.

- From groupthink to the wisdom of crowds: how to create AI social platforms that systematically leverage collective intelligence, while at the same time having built-in socio-technical mechanisms to isolate and ostracize malicious and harmful agents

- Innovative methods for critical thinking education and digital literacy as early response measures: how to start the fight against online disinformation well before any outbreak, by adequately strengthening the epistemic immune system of Internet users.

- Social nudges for responsible information consumption: how to leverage social norms to promote a virtuous collective discourse around pandemic threats, while avoiding the pitfalls of reputational mechanisms (e.g., distrust and even open hostility among citizens).

- Choice architecture against binge browsing under stress: how to make sure that the hard self-restrain required to navigate the digital ecology during a pandemic becomes as natural and easy as possible for citizens, by carefully designing how their browsing choice get made.

- Risk misperception and its remedies: how biases in risk assessment affect our reactions to pandemic threats and related (dis)information, and how to correct such biases.

- Trust in institutions and the orchestration of the public response to pandemic threats: how to maximize and leverage social capital in dealing with a pandemic, via careful design of the interaction among all institutional actors involve.

- The language of pandemic emergency: how message crafting can improve dramatically the effectiveness of risk communication and the compliance with proposed measures and guidelines, based on subtle yet easy to design pragmatic and presentational cues.

- Designing interdisciplinary collaboration in response teams: how to best mobilize the complex and broad array of scientific and technical competences needed to face a pandemic threat and its associated digital risks, avoiding the pitfalls of tunnel vision.