**Attribution of Cyber Attacks: Technological and Legal Dimensions**

**November 11, 2018**

**The Hebrew University of Jerusalem, Israel**

A joint workshop on the attribution of cyber-attacks was hosted by the Hebrew University Cyber Law Program and King's College, with the support of the British Council. This multidisciplinary workshop brought together academic and government experts from diverse scientific, legal, and public-policy fields, as well as from various institutions, including the Hebrew University of Jerusalem, King's College London, the Israel Defense Force, NATO and GCHQ.

The first part of the workshop focused on the technological dimension of attributing cyber-attacks, i.e. ascertaining the facts relating to their occurrence and origins. The challenges of attribution and possible solutions thereto were examined with respect to a wide range of cyber-attacks, such as hacking, implanting malware, disseminating "fake news," and disrupting the global cyber-infrastructure. Key technical features of cyber-space were shown to complicate the collection, analysis, and assessment of evidence for the purpose of attribution – while also making cyber-attacks easier to mount – with multi-stage attacks posing the most serious problem. The notion of trust was discussed in different contexts – e.g., credibility of sources of information in investigating cyber-attacks, reliability of individual users of social networks and other online platforms, and mutual trust among like-minded States – as participants considered how trust could be bolstered in order to enhance attribution capabilities. Specifically, feeding trust into a formal model for evaluating forensic evidence could assist investigators of cyber-attacks with conflicting information. Similarly, open-sourced, collectively-owned, and self-policing crypto-networks would allow basing many cyber-activities on mathematical rather than interpersonal trust.

Another major topic on the technological side was the growing sophistication and rapid proliferation of means and methods for carrying out cyber-attacks – from different kinds of malware to the ecosystems facilitating the spread of false or slanted information. Presentations explored the prospects of identifying different types of patterns in employing diverse means and methods of attack, and then using such

patterns, via natural-language processing and machine-learning, to unmask offenders. This data-driven approach may foil attempts to mislead investigations by mimicking code and technologies associated with other cyber-attacks. A separate discussion concerned the inherent insecurity of the global cyber-infrastructure's three core components, i.e. the systems used to locate websites (DNS), route user-traffic (BGP), and synchronize time (NTP). The second component's vulnerabilities, for example, could be exploited to launch potentially devastating cyber-attacks, which would be as difficult to detect as they are simple to plausibly deny, especially in real-time. Replacing this core component with a more secure system remains possible, but progress is slow due to the high costs and lack of financial incentives. Alternative solutions include global regulation and constant monitoring of internet traffic, which might adversely affect online freedoms.

A crosscutting theme of the foregoing discussions was that there are no silver bullets in the search of technological solutions to the challenges of attributing cyber-attacks. Thus, algorithms for assessing the reliability of evidence can only help, not replace human analysts; pattern-detection can only deny certain means and methods from attackers, not prevent them from developing more advanced ones; and in any case, relevant technologies can only lead investigators to the computer used to conduct a cyber-attack, not expose the individual or entity behind it. But for any solution to be effective, participants stressed, it is vital to maintain close cooperation between scientists and 'techies' on the one hand, and legal and policy experts on the other hand.

The second part of the workshop focused on the legal dimension of attributing cyber-attacks, namely identifying their human perpetrators, specifically for the purpose of assigning State responsibility to such attacks as wrongful acts under international law, to which victim States may respond with countermeasures. The applicability and adequacy of the existing legal framework for attributing cyber-attacks was one of the key issues discussed. It was observed that most difficulties in this respect are in fact not unique, as they exist with respect to espionage and to using non-State actors as proxies. However, cyber-attacks do pose special challenges to making an attribution determination with the required degree of certainty, and this problem is compounded by the vagueness of international law on this so-called burden of proof. Participants

then considered how these well-settled rules can be supplemented, for instance by using rebuttable presumptions, as well as whether they should and could be revised.

Another central discussion on the legal side concerned governmental and inter-governmental practicalities of attributing cyber-attacks. Topics here included the policy aims of attribution determinations and their implications for analyzing information and evidentiary standards; the competing interests of States with respect to attribution, e.g. preserving their own ability to take covert action while preventing other States from doing so against them; and challenges of collaboration in attributing cyber-attacks, considering the national character of such determinations, the possible disincentives to share technological knowledge and intelligence, but also the need for States to be more explicit in international cyber-matters. A related discussion explored the feasibility of establishing an international mechanism for attributing cyber-attacks. The suitability of preexisting models from other areas of international relations (e.g., arms control) was assessed, alongside recent, cyber-specific proposals put forward by different think-tanks and tech companies. While each of these options has valuable elements, it is unclear how any such mechanism can ensure compliance with the binding international rules for State conduct in cyberspace, as core legal concepts – from "sovereignty" to "due diligence" – remain under-defined. It was therefore argued that like-minded States, especially those who possess cyber capabilities, ought to initiate an inter-governmental process for further elaborating the substantive legal framework and establishing an attribution mechanism. This process should involve independent experts from industry and academia, and it must be able to make decisions by majority-vote, rather than strictly by consensus.

Throughout these legal discussions, participants paid much attention to the policy and strategic aspects of attribution, as well as to the unique situation of States in the cyber-age. Hence, one crosscutting themes of the foregoing discussions was the extreme asymmetries between States and non-state actors, from multinational corporations to criminal or terrorist organizations, as well as asymmetries among States. Another theme was attribution's potential contribution to deterrence and reducing aggression, bearing in mind the integral risk of rapid escalation by virtue of irresponsible attribution. Lastly,

presenters and participants reflected on the possible trajectories of cyber-space, in terms of security and stability, inter-governmental cooperation, and public transparency.

Main topics which remained open for further discussions and research were: 1. The issue of State and non-state actors' responsibility and direct and indirect responsibility; 2. The need to differentiate between individuals with different motivations (criminal or others) and States; 3. Is 'good enough governance' a sufficient standard for coordination and cooperation?; 4. The question whether territorial aspects still relevant and technologically applicable; 5. Does strict legal rules provide the best mechanism for addressing the challenge or should more flexible standards be developed? Are cyber-attacks more akin to Espionage than war?; 6. Should attribution be addressed through an inter-State mechanism, or other private or public framework?; 7. What is the role of AI in attribution and response (e.g., AI-governed hack backs)?; 8. Can Trust based system and certifications be developed to help attribution efforts; 10. And what is the role of ethics by design – are there suitable technological solutions that refer to ethics, in the aspects of evidence collection and attribution.