Minds & Machines (2016) 26:317–321
DOI 10.1007/s11023-016-9408-z

# On the Risks of Relying on Analogies to Understand Cyber Conflicts

**Mariarosaria Taddeo**[1,2]

Received: 17 November 2016 / Accepted: 18 November 2016 / Published online: 23 November 2016


Efforts to regulate cyber conflicts—and cyber-defence postures more generally—rose to prominence almost a decade ago, when the risks for national and international security and stability arising from the cyber domain became clear.[1] As I argued elsewhere (Taddeo 2014), these efforts often rely on an *analogy-based approach*, according to which the regulatory problems concerning cyber conflicts are only apparent, insofar as these are not radically different from other forms of conflicts. Those endorsing this approach claim that the existing legal framework[2] governing armed conflicts is sufficient to regulate the cyber battlefield. All that is needed is an in-depth analysis of such laws and an adequate interpretation of the phenomena. As Schmitt stresses.

> "a thick web of international law norms suffuses cyber-space. These norms both outlaw many malevolent cyber-operations and allow states to mount robust responses" (Schmitt 2013, 177).

While the use of analogies to regulate cyber conflicts proved to be effective in the short term, and was necessary a decade ago to avoid the so-called digital Wild West;

[1] http://www.nato.int/docu/review/2013/cyber/timeline/EN/index.htm.

[2] The referred legal framework mainly encompasses the four Geneva Conventions and their first two Additional Protocols, the international customary law and general principle of law, the Convention restricting or prohibiting the use of certain conventional weapons, and judicial decisions. Arms control treaties, such as the Nuclear Non-Proliferation Treaty and the Chemical Weapons Convention, are often mentioned as providing guidance for action in the case of kinetic cyber attacks (Schmitt 2013). At the same time, coercive measures addressing economic violations are generally considered legitimate in the case of cyber attacks that do not cause physical damage (Lin 2012; O'Connell 2012).

✉ Mariarosaria Taddeo
mariarosaria.taddeo@oii.ox.ac.uk

[1] Oxford Internet Institute, University of Oxford, Oxford, UK

[2] Alan Turing Institute, London, UK

in the medium- and long-term, this approach poses serious risks to the stability of current and future information societies. For efforts based on analogies between kinetic and cyber conflicts often end with *ad hoc* solutions, fall short of political and ethical foresight, and overlook, and are limited by, the theoretical vacuum underlying them.

A relation of mutual influence exists between the way in which conflicts are waged and the societies waging them (Taddeo and Glorioso 2016a, b). For this reason the analogy-based approach risks entrapping future information societies in the past, thereby missing the opportunity to address questions concerning the impact of this new form of conflict on our societies, on their values, the rights and security of their citizens, and on national and international politics.

And not just that. If taken too far, analogies between kinetic and cyber conflicts become misleading and pose more problems than they can solve. This is, for example, the case of cyber deterrence. Estimates indicate that the cyber security market will grow to US$170 billion by 2020, posing the risk of a progressive militarisation of cyber domain ensuing a cyber arms race and competition for digital supremacy, and hence increasing the possibility of escalation (Markets and Markets 2015). In this scenario, cyber deterrence is crucial to maintain international security and foster stability.

However, deploying deterrence strategies within the cyber domain is proving to be a serious challenge, due to the global reach, the anonymity, the distributed, and the interconnected nature of the domain (Chadwick and Howard 2009). In this case, relying on analogies with nuclear deterrence aggravates, rather than helping to meet, the challenge. Consider, for example, game-theory models for nuclear deterrence, such as the famous "brinkmanship" model (Powell 2008). This model relies on a clear identification of the attacker, and on demonstrating state's capability of retaliating and commitment to retaliate (credibility) should the attacker not desist from his/her intent. While credibility and capability may not be problematic in the case of cyber deterrence, the identification of the source of the threat (attribution) in the cyber domain is one of the crucial hurdles to address. The difficulties in identifying the attacker make problematic the strategic assessment of pains and gains, and the understanding of the attacker's strategies, payoffs, and preferences, making this model inadequate to define successful deterrence strategies in the cyber domain.

This is not tantamount to saying that game theory models do not apply to the case of cyber deterrence; but it is indicative of the need to develop *new* domain-specific models able to account for the specificity of the cyber domain and of the conflicts to deter. This concerns (Taddeo 2012, 2014; Dipert 2013; Floridi and Taddeo 2014):

- The domain: ranging from the virtual to the physical;
- The agents and targets: involving artificial alongside human agents, as well as physical and virtual objects; and
- The level of violence: spanning from non-violent to potentially highly violent phenomena.

These aspects of the cyber domain and of cyber conflicts are of great relevance for they are reshaping our understanding of key concepts such as harm, violence, target, combatants, weapons, attack, state sovereignty, and territoriality, and political power (Chadwick and Howard 2009; Cornish 2016).

Clearly, understanding such conceptual changes—and identifying their impact on international relations and on military strategies—are preliminary and necessary steps to any attempt to identify legitimate strategies and to define the right policies, norms, and laws regulating cyber conflicts. For this reason, it is necessary to realise the limits of the analogy-based approach, and to move past it. As Betz and Stevens (2013) put it:

> "It is little wonder that we attempt to classify […] the unfamiliar present and unknowable future in terms of a more familiar past, but we should remain mindful of the limitations of analogical reasoning in cyber security".

Analogies can be powerful, for they inform the way in which we think and constraint ideas and reasoning within a conceptual space (Wittgenstein 2009). However, if the conceptual space is not the right one, analogies become misleading and detrimental for any attempt to develop innovative and in-depth understanding of new phenomena, as in the case of cyber deterrence, and they should be abandoned altogether. When the conceptual space is the right one, analogies are at best a step on the Wittgenstein's ladder and need to be disregarded once they have taken us to the next level of the analysis. This is the case of the analogies between kinetic and cyber conflicts.

I believe that the time has come to develop a new framework to understand and regulate this cyber conflicts. This framework relies on three methodological pillars: political ontology, *political constructionism*, and data ethics. Let me explain.

There is an ontological hiatus between the entities involved in cyber conflicts and those taken into consideration in kinetic ones (Taddeo 2014). An analysis of the nature of the environment, of the agents involved in conflicts, and of the political power exerted in the cyber domain is thus a necessary preliminary step to any attempt to regulate this new type of conflict. As Hay put, it "ontological assumptions (relating to the nature of the political reality that is the focus of our analytical attentions) are logically antecedent to the epistemological and methodological choices" (Hay 2011, 1–2).

'Political constructionism' is a neologism that refers to the design process of policies, norms, and laws. It builds on Floridi's analysis of norms as artefacts.

> "designed according to a set of requirements, which are specified on the basis of available resources, in order to implement a set of desired functions, and all this in view of achieving some ultimate purpose […]", (Floridi 2015, 1100).

Data ethics builds on political ontology and provides the purpose required by political constructionism for the design of norms. Data ethics is the.

> "branch of ethics that studies and evaluates moral problems related to data (including generation, recording, curation, processing, dissemination, sharing and use), algorithms (including artificial intelligence, artificial agents, machine

learning and robots) and corresponding practices (including responsible innovation, programming, hacking and professional codes), in order to formulate and support morally good solutions (e.g. right conducts or right values)", (Floridi and Taddeo 2016, 3).

As such, data ethics embraces the correct level of abstraction[3] to identify the key ethical principles that should underpin the regulation of cyber conflicts (Taddeo and Glorioso 2016a), and which will contribute to shape open, pluralistic, and stable information societies.

A conceptual framework built on these three pillars will deliver the adequate responses to address the threats arising from the cyber domain. The alternative is developing unsatisfactory, short-sighted approaches and facing the risk of a *cyber backlash*: a deceleration of the digitization process imposed by governments and international institutions to prevent this kind of conflicts to erode both in the trust in economy and in political institutions.

## References

Betz, D. J., & Stevens, T. (2013). Analogical reasoning and cyber security. *Security Dialogue, 44*(2), 147–164. doi:10.1177/0967010613478323.

Chadwick, A., & Howard, P. N. (Eds.). (2009). *Routledge handbook of internet politics. Routledge handbooks*. London: Routledge.

Cornish, P. (2016). *Deterrence as the basis for ethical constraints on conflict in the cyber domain. In ethics and policies for cyber warfare. Philosophical studies*. Berlin: Springer.

Dipert, R. (2013). The essential features of an ontology for cyberwarfare. In P. Yannakogeorgos, A. Lowther (Eds.) *Conflict and cooperation in cyberspace* (pp. 35–48). London: Taylor & Francis. http://www.crcnetbase.com/doi/abs/10.1201/b15253-7.

Floridi, L. (2008). The method of levels of abstraction. *Minds and Machines, 18*(3), 303–329. doi:10.1007/s11023-008-9113-7.

Floridi, L. (2015). Toleration and the design of norms. *Science and Engineering Ethics, 21*(5), 1095–1123. doi:10.1007/s11948-014-9589-x.

Floridi, L., & Taddeo, M. (2016). What is data ethics? *Philosophical Transactions A, 374*, 20160360.

Floridi, L., & Taddeo, M. (Eds.). (2014). *The ethics of information warfare*. New York: Springer.

Hay, C. (2011). Political ontology. In R. Goodin (Ed.), *The Oxford handbook of political science* (pp. 78–96). Oxford: Oxford University Press.

Hoare, C. A. R. (1972). Structured programming. In O. J. Dahl, E. W. Dijkstra, and C. A. R. Hoare (Ed.), (pp. 83–174). London: Academic Press Ltd. http://dl.acm.org/citation.cfm?id=1243380.1243382.

Lin, H. (2012). Cyber conflict and international humanitarian law. *International Review of the Red Cross, 94*(886), 515–531. doi:10.1017/S1816383112000811.

Markets and Markets. (2015). Cyber security market by solutions and services—2020. http://www.marketsandmarkets.com/Market-Reports/cyber-security-market-505.html?gclid=CNb6w7mt8MgCFQoEwwodZVQD-g.

[3] The method of abstraction is a common methodology in Computer Science (Hoare 1972) and in Philosophy and Ethics of Information (Floridi 2008). It specifies the different Levels of Abstraction LoAs at which a system can be analysed, by focusing on different aspects, called observables. The choice of the observables depends on the purpose of the analysis and determines the choice of LoA. Any given system can be analysed at different LoAs. For example, an engineer interested in maximising the aerodynamics of a car may focus upon the shape of its parts, their weight, and the materials. A customer interested in the aesthetics of the same car may focus on its colour and on the overall look and may disregard the shape, weights, and materials of the car's components.

O'Connell, M. E. (2012). Cyber security without cyber war. *Journal of Conflict and Security Law, 17*(2), 187–209. doi:10.1093/jcsl/krs017.

Powell, R. (2008). *Nuclear deterrence theory: The search for credibility. Digitally printed version. Paperback re-issue*. Cambridge: Cambridge University Press.

Schmitt, M. (2013). Cyberspace and international law: The penumbral mist of uncertainty. *Harvard, 126*(176), 176–180.

Taddeo, M. (2012). Information warfare: A philosophical perspective. *Philosophy and Technology, 25*(1), 105–120.

Taddeo, M. (2014). Just information warfare. *Topoi*, April 1–12. doi:10.1007/s11245-014-9245-8.

Taddeo, M., & Glorioso, L. (Eds.). (2016). *Ethics and policies for cyber operations. Philosophical studies*. Berlin: Springer.

Taddeo, M., & Ludovica, G. (Eds.) (2016b). Regulating cyber conflicts and shaping information societies. In *Ethics and policies for cyber operations*. Philosophical studies series. Berlin: Springer.

Wittgenstein, L. (2009). *Philosophical investigations* (4th ed.). Chichester: Wiley-Blackwell.