THE CENTER FOR ADVANCED RED TEAMING
UNIVERSITY AT ALBANY
State University of New York

Towards a Definition of Red Teaming

Version 1.0 *October 2019*

CART's proposed definition of red teaming is:

Any activities involving the simulation of adversary decisions or behaviors, where outputs are measured and utilized for the purpose of informing or improving defensive capabilities.

Red teaming can overlap with, but does not completely encompass or fall within, various other types of activities, such as wargaming, penetration testing, crisis simulation and scenario building.

Why Bother With a Definition?

Red-teaming is simple in concept, yet many refer to different things when they speak about "red teaming," hindering our ability to advance its practice. A generalized definition can help ensure that we all start out on the same page. To be useful, however, any definition must identify the boundaries of the concept that it is describing. In other words it must balance three criteria:

a) **Breadth** - Be sufficiently general so as to accommodate existing usage and the broad range of practices;

b) **Essence** - Incorporate essential characteristics without which a given practice would not be considered red teaming; and

c) **Exclusivity** - Exclude any features that are not necessary for a practice to be considered part of red teaming.

How Did You Come Up With the Above Definition?

We arrived at our definition through a survey of twenty-seven existing definitions or descriptions of red teaming in published sources. This was followed by a structured analytical process to identify and evaluate the most salient conceptual components, using the above three criteria as a guide with regards to relevance.

What Do You Regard as Essential Components of "Red Teaming"?

Across all surveyed sources, we identified three essential components of red teaming, which were incorporated into the above definition:

1. **Simulation of Adversaries** [17][1]: This is the essence of the "red" in red teaming; if there is no involvement of an adversary, a practice is better described using another term. The activity may or may not include the simulation of other actors, but must include the simulation of at least one adversary. Note that an adversary is defined in relation to the user of the red teaming. For example, if the user is an intelligence agency, adversaries might include the agents of another state, terrorists, or transnational criminal organizations, while the adversaries of a corporation might include competitor firms and disgruntled employees, as well as terrorists, hackers and certain governments.

2. **Defined Output (Decisions or Behavior)** [4]: A simulation must be intended to provide a specified, recordable set of outputs. Simulating adversaries merely for entertainment purposes is not red teaming. For example, children playing "cops and robbers" or a Mission Impossible movie do not by

[1] The number of published sources that refer to a particular component is given in brackets. More detail on these sources and definitions can be obtained from CART upon request. The frequency of mentions in the literature was only one element in the decision to include or exclude a component, with the extent to which a component met the three criteria listed above playing a prominent role as well.

themselves constitute red teaming. In almost all cases the output will include one or more adversary decisions or actions, since focusing only on the adversary's imaginings tends to have less practical value.

3. **Assisting With Defense** [7]: At least one of the objectives of the simulation must be to assist in some way with defending against the adversary or adversaries being simulated. For example, if one were to simulate an adversary merely to explore how to increase the efficacy of its attack capabilities but with no desire to inform defense against threats, we believe this should not be regarded as red teaming.

Why Did You Not Include Other Elements in Your Definition?

We identified twenty-six different concepts across all the definitions surveyed, but found that the three elements above uniquely met all of our criteria. The reasons for not including several of the more commonly cited concepts are given below:

- **"Devil's Advocate" (Alternative Perspective)** [13]: While arguably necessary, we believe that this concept is subsumed by the requirement in the above definition that an adversary must be simulated. An adversary is by definition not the same as the defender and therefore will automatically generate an alternative perspective when simulated.
- **Reducing Risks** [5] / **Reviewing Plans or Concepts** [3] / **Challenging Status Quo** [7] / **Identifying Adversary Innovation** [3] / **Identifying Gaps and Exposing Vulnerabilities** [11]: While these are often the functions or objectives of red teaming, they are not essential components because activities usually regarded as red teaming can have other goals or functions (e.g., red teaming used for training defenders or red teaming used to raise awareness of a risk without necessarily identifying specific vulnerabilities).
- **Team (>1 player)** [5] / **Expert Participants** [3] / **Security or Intelligence Context** [5] / **Iterative** [2] / **Interactive** [2]: These are all common features of red teaming but are not strictly necessary. In other words, we could come up with one or more examples of activities for each of these that are widely regarded as red teaming, but do not involve these criteria.
- **Independent** [9] / **Structured** [5] / **Involving Critical Thinking** [3] / **Realistic** [1]: Although all of these elements are arguably features of high-quality red teaming and many are possibly best practices, none of them are conceptually necessary for an activity to be regarded as red teaming.

How Does Your Definition of Red Teaming Fit In With "Alternative Analysis" and "Wargaming"?

According to the above definition, **all red teaming provides alternative analysis, but not all alternative analysis is red teaming**. For example, an alternative analysis that only considers the decisions of "Blue" (i.e., the defender), would not be regarded as red teaming.[2]

Similarly, **all wargaming is red teaming, but not all red teaming is wargaming**. For example, focusing on adversary decisions in a computational simulation would constitute red teaming, but the absence of a "Blue" force means that it would not be considered to be a wargame.

What If I Don't Agree With Your Definition?

That is perfectly fine. Definitions are notoriously tricky to find consensus around and we recognize that our attempt at a generalized definition of red teaming will not be accepted by everyone. Some people might think that it is too vague and includes too many different activities, while others might think that it is too narrow and excludes practices that that they believe fall within the domain of red teaming. In developing the above definition, we tried to balance breadth and exclusivity, while striving for a basic consensus across existing definitions. Even if you do not agree with the above definition, it can at the very least serve as a starting point for further discussion by practitioners and researchers.

If I Want to Use the Above Definition, How Should I Cite It?

If our definition of red teaming turns out something that you would like to use in your own work, that is great – we are glad that you find it useful. Please use the following citation: Gary A. Ackerman and Douglas Clifford, "Towards a Definition of Red Teaming," *Center for Advanced Red Teaming* (Albany, NY: University at Albany, 2019).

[2] There are broader definitions of red teaming that do not require an adversary and instead focus on the alternative analysis aspect. The latter is practiced, among others, by the U.S. Army.