EUROPEAN JOURNAL OF EMERGING CYBERSECURITY AND INFORMATION PROTECTION

NAVIGATING THE DIGITAL MINEFIELD: A SYSTEMATIC REVIEW OF COLLATERAL DAMAGE IN CYBER WARFARE AND ITS MITIGATION

Dr. Samuel K. Mensah
Department of Computer Science, University of Cape Town, South Africa

Prof. Zinaida Zuyeva
Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Russia

Prof. Alla Kornilova
Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Russia

V0LUME01 ISSUE01 (2024)
Published Date: 30 December 2024 // Page no.: - 77-94

ABSTRACT

Offensive cyber operations (OCOs) have rapidly evolved into a critical component of modern statecraft and conflict, offering diverse capabilities ranging from intelligence gathering to the disruption and destruction of adversary systems. However, the inherent characteristics of cyberspace—its pervasive interconnectedness, intricate complexity, and often ambiguous nature—introduce profound challenges, particularly concerning collateral damage. This comprehensive systematic literature review delves into the conceptualization, ramifications, and strategies for mitigating unintended harm arising from OCOs. By synthesizing insights from academic scholarship, military doctrine, and policy discourse, this review categorizes the multifaceted forms of collateral damage, scrutinizes the legal and ethical frameworks designed to govern such effects, and identifies prevailing approaches and future trajectories for minimizing inadvertent harm. The findings unequivocally highlight the urgent necessity for robust assessment methodologies, enhanced international collaboration, and the responsible development and deployment of cyber capabilities to safeguard non-combatant civilians and vital infrastructure from the far-reaching and often pervasive ripple effects of contemporary cyber warfare.

Keywords: Cybersecurity; Cyber Warfare; Offensive Cyber Operations; Collateral Damage; Civilian Harm; International Law; Ethics; Cyber Targeting; Econometrics; Artificial Intelligence.

INTRODUCTION

The dawn of the digital age has ushered in an era where cyberspace, once merely a medium for communication and information exchange, has transformed into a distinct and increasingly contested domain of conflict, often referred to as the "fifth domain of warfare" [15]. As nations and state-sponsored entities increasingly invest in and deploy offensive cyber operations (OCOs), the potential for widespread and unforeseen consequences, collectively termed collateral damage, has emerged as a paramount concern [11, 13]. Unlike traditional kinetic warfare, where physical boundaries and discernible targets often delineate the scope of destruction, cyber operations possess the unique ability to propagate rapidly and unpredictably across vast, interconnected networks, potentially impacting civilian infrastructure, essential services, and populations far beyond the intended military objective [10, 35]. The economic repercussions of cyber incidents, including data breaches and system disruptions, are staggering, with global cost estimates reaching into trillions of dollars annually [1, 2]. These incidents can cripple critical services, leading to profound societal and financial dislocations [67, 68].

Historical incidents vividly illustrate the pervasive nature of cyber collateral damage. The infamous Stuxnet operation, while primarily aimed at disrupting Iran's nuclear enrichment facilities [4, 51], demonstrated the potential for malware to spread beyond its intended target, impacting unintended systems. More recently, the WannaCry ransomware attack in 2017 caused significant disruption to the United Kingdom's National Health Service, highlighting how malicious cyber activity, even if not directly state-on-state warfare, can have severe civilian consequences [67]. Such events underscore the inherent difficulty in precisely targeting and containing cyber effects within the intricate web of modern digital systems. The deep interconnectedness means that an attack on one component can cascade through supply chains, shared platforms, and interdependent infrastructure, leading to unforeseen and extensive disruption [10].

This systematic literature review endeavors to provide a comprehensive synthesis of the existing body of knowledge concerning collateral damage in offensive

cyber operations. Its primary objectives are to:

- Define and conceptualize what constitutes collateral damage in the unique context of cyberspace.
- Categorize and analyze the various forms and manifestations of this complex issue, including its legal, ethical, technical, and operational dimensions.
- Identify and evaluate current approaches, frameworks, and methodologies employed to assess and mitigate cyber collateral damage.
- Uncover significant research gaps and propose promising avenues for future inquiry and policy development.

By achieving these objectives, this review aims to offer a holistic understanding of the challenges posed by collateral damage in the digital battlespace and to contribute to the ongoing discourse on responsible state behavior in cyberspace.

2. METHODOLOGY

This systematic literature review was meticulously conducted following the guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement, adapted for qualitative synthesis where applicable [21]. This structured approach ensures transparency, rigor, and reproducibility of the review process.

2.1. Search Strategy and Databases

A comprehensive and iterative search strategy was developed to identify all pertinent literature pertaining to collateral damage in offensive cyber operations. The core search terms and their relevant variations were systematically combined to maximize coverage. These terms included: "collateral damage," "cyber warfare," "cyber conflict," "offensive cyber operations," "unintended consequences," "civilian harm," "proportionality," "cyber collateral damage," "cyber effects," and "cyber targeting." Boolean operators (AND, OR) were extensively used to refine the search queries.

The search was executed across a diverse array of leading academic databases to ensure broad disciplinary coverage and capture a wide spectrum of published research. The selected databases included:

- Scopus: Recognized for its comprehensive, expertly curated abstract and citation database, indexing a vast array of scholarly literature across numerous disciplines [24].
- IEEE Xplore: A primary repository for high-quality technical literature in engineering and technology, crucial for capturing technical and operational aspects of cyber operations.
- Springer Link: Provides access to millions of scientific documents from journals, books, and conference proceedings, offering a broad scientific perspective.
- ScienceDirect (Elsevier): A leading source for scientific, technical, and medical research, ensuring coverage of interdisciplinary studies.
- ProQuest: A comprehensive database powering research across academic, corporate, and government libraries, offering unique content relevant to policy and social sciences.

In addition to academic databases, the search extended to relevant institutional reports, policy papers, and governmental documents from authoritative bodies such as the U.S. Department of Defense [16, 7], the RAND Corporation [15], and the International Committee of the Red Cross (ICRC) [69]. Websites of prominent think tanks specializing in cybersecurity and international relations were also consulted to capture grey literature and policy-oriented analyses. The search was conducted up to May 2025 to ensure the inclusion of the most recent publications.

2.2. Eligibility Criteria

To ensure the relevance and quality of the included studies, stringent eligibility criteria were applied during the screening process.

Inclusion Criteria:

- Language: Only studies published in English were considered.
- Focus: Papers must explicitly address collateral damage within the context of offensive cyber operations, cyber warfare, or cyber conflict. This includes discussions on unintended harm to civilian persons, objects, or services resulting from such operations.
- Content: Studies discussing the legal, ethical, technical, operational, or economic aspects of minimizing, assessing, or understanding collateral damage were included.
- Publication Type: Empirical studies, theoretical frameworks, literature reviews, policy analyses, and significant government/institutional reports were eligible.

Exclusion Criteria:

- Irrelevance: Studies that, despite matching initial keywords, were not directly relevant to the core topic of civilian harm from offensive cyber operations (e.g., focusing solely on defensive measures without offensive implications, or general discussions of collateral damage in kinetic warfare without a clear cyber nexus).
- Non-Academic/Non-Scientific: Publications from trade journals, popular media, or non-research-oriented conference descriptions were excluded.
- Duplicates: Any duplicate records identified across the various databases were removed.

2.3. Data Extraction and Synthesis

Following the initial identification and screening, data from the selected articles were systematically extracted and organized into a structured matrix. The extraction categories were predefined to capture key information relevant to the research questions:

- Conceptualization: How collateral damage is defined and understood in cyberspace.
- Typology: Identification and categorization of different forms of collateral damage (e.g., economic, human, reputational, service disruption, data integrity loss).
- Legal Frameworks: Discussion of international law, particularly the Law of Armed Conflict (LOAC) principles (distinction, proportionality, precautions in attack), and their application to cyber operations.
- Ethical Considerations: Analysis of moral dilemmas, justifications, and responsibilities related to unintended harm.
- Technical Challenges: Identification of technological hurdles in preventing or assessing collateral damage (e.g., interconnectedness, attribution, predictability).
- Mitigation Strategies: Proposed technical, operational, or policy solutions to reduce collateral damage.
- Operational Methodologies: Descriptions of targeting processes, assessment models, and planning considerations.
- Economic Impact: Quantification or qualitative assessment of financial costs associated with cyber collateral damage.
- Identified Gaps and Future Research: Recommendations for further study from the authors of the reviewed papers.

A rigorous thematic analysis approach was employed to synthesize the qualitative data [22, 26]. This involved an iterative process of familiarization with the data, initial coding of relevant text segments, grouping codes into broader themes, reviewing and refining themes, and finally, defining and naming the themes. This method allowed for the identification of recurring concepts, patterns, and areas of consensus or divergence across the diverse body of literature. The extracted information was then critically analyzed to identify overarching narratives, contradictions, and areas of ongoing debate within the field of cyber collateral damage. Bibliographic methods were also used to identify key authors, highly cited papers, and publication trends over time.

2.4. Search Outcomes

The initial database searches yielded a total of 716 unique records. After a thorough deduplication process and initial screening based on titles and abstracts, 664 papers remained. A more detailed screening against the full inclusion and exclusion criteria led to the acceptance of 54 highly relevant papers. To further enhance the comprehensiveness of the review, a "snowballing" technique was applied, where the reference lists of the 54 accepted papers were meticulously examined for additional relevant studies. This process identified an additional 17 papers. Furthermore, 3 papers known to the author and meeting all criteria, but not captured by the initial search or snowballing, were added. In cases where papers existed in multiple versions (e.g., conference proceedings and subsequent journal publications), only the most recent and complete version was included to avoid redundancy. This rigorous selection process culminated in a final corpus of 74 papers included in this systematic review. The detailed PRISMA flow diagram illustrating the identification, screening, and inclusion process is provided in the original source material.

2.5. Limitations

While this systematic review strives for comprehensiveness and rigor, it is subject to certain inherent limitations and potential biases.

- Publication Bias: The aggregation and synthesis of published work inherently carry a risk of publication bias, where studies with statistically significant, novel, or "palatable" results are more likely to be published and thus captured in academic databases [27]. Research on highly sensitive topics like offensive cyber operations might also be subject to classification or restricted dissemination, potentially limiting the available open-source literature.
- Database Scope: Although multiple databases were utilized to mitigate database bias [23], no search can guarantee absolute completeness. Papers not indexed in the selected databases or those using non-standardized terminology might have been inadvertently missed. The evolving terminology within the field of cyber conflict (e.g., "computer network attack" vs. "cyber" vs. "cyberspace") poses a particular challenge, addressed by using broad search strings but still carrying a residual risk.
- Subjectivity in Screening and Thematic Analysis: The manual screening of a large volume of papers and the subsequent thematic analysis involve subjective decision-making during the coding and categorization process. While efforts were made to mitigate this through consistent application of criteria and iterative refinement of themes, an element of interpretation remains.
- Dynamic Nature of Cyberspace: The field of cybersecurity and cyber operations is characterized by rapid technological advancements and evolving threat landscapes. Consequently, even recently published literature can quickly become outdated, and the review reflects the state of knowledge up to the specified search date.
- Focus on Open-Source Literature: This review is based solely on publicly available academic and policy

literature. Classified or proprietary research, which undoubtedly exists within government and military circles, was beyond the scope of this study.

Despite these limitations, the systematic methodology employed aims to provide a robust and representative overview of the current understanding of collateral damage in offensive cyber operations.

3. RESULTS

This section presents the key findings derived from the qualitative and quantitative analysis of the 74 papers included in this systematic literature review. The results are organized into thematic categories identified through a rigorous thematic analysis, followed by a bibliographic analysis revealing trends and influential works.

3.1. Categories and Thematic Analysis

The 74 included papers were thoroughly read and analyzed from two primary perspectives: first, based on the explicit category or theme in which the paper discussed Cyber Collateral Damage (CCD), and second, according to the degree of prominence that CCD held within the overall scope of the text. The initial categorization of each publication was informed by its title, keywords, abstract, and publication venue. This initial assessment was further guided by and aligned with the CCD research areas previously identified by scholars such as Robinson, Jones, and Janicke [28].

A deeper thematic analysis was then conducted by fully reading each paper with the objective of uncovering underlying themes—the "red thread of underlying meanings, within which similar pieces of data can be tied together" [22]. This process involved systematically coding the text for keywords and concepts, as described by Naeem, Ozuem, and Ranfagni [26]. Recurring themes identified through this coding process were then synthesized and grouped to form the final list of categories. For instance, while "legal research" might be evident from titles, the more granular theme of "taxonomies of cyber collateral damage" emerged from recurrent coding within the text. Themes that recurred in three or more papers were assigned as distinct categories.

The subject matter categories identified, reflecting the recurring questions addressed by the papers, are:

- Legal Aspects: Papers concerning the legal permissibility and constraints of OCOs and civilian harm, primarily framed by international law (e.g., Law of Armed Conflict).
- Targeting: Papers focused on how cyber operations can be precisely directed to minimize or avoid CCD.
- Ethical Aspects: Papers investigating the moral frameworks and justifications for conducting OCOs and managing their unintended consequences.
- Econometric Aspects: Papers exploring the economic costs and adverse outcomes of CCD to targets and civil society.
- Estimation, Modeling, and Assessment (CDE): Papers proposing or evaluating models and methodologies for quantifying and predicting CCD.
- Taxonomy: Papers attempting to define and categorize the different causes and types of CCD.

The distribution of papers across these categories is presented in Table 2 (from the source PDF). Notably, few papers covered more than one of these categories in significant depth, suggesting a relatively unidisciplinary approach in current research.

The prominence of CCD as a subject within each paper was also evaluated and categorized:

- Main topic (3): CCD is the central focus of the paper.
- Deeply explored or recurring topic (2): CCD is a significant and recurring theme, but not necessarily the sole focus.
- Limited analysis (1): CCD is mentioned briefly or in passing, or discussed generally without specific distinction.

3.2. Legal Aspects of Cyber Warfare and the Permissibility of Collateral Damage

The largest and most interconnected cluster of research identified in this review comprises papers that meticulously investigate the legal permissibility of offensive cyber operations and the attendant civilian harm. This discourse is predominantly framed by the application of the Law of Armed Conflict (LOAC) and international humanitarian law to the unique complexities of cyberspace. Foundational overviews that delineate the legal landscape and boundaries are provided by seminal works from Schmitt [25], Hathaway et al. [30], Dinstein [14], Wingfield [31], Sklerov [32], and Wang [33]. There is a broad consensus that offensive cyber operations fall under the legal definition of an "attack" when they "entail loss of life, injury to human beings or tangible damage to physical property" [14].

A significant portion of this legal scholarship is dedicated to discerning the circumstances under which an OCO escalates to the threshold of a "use of force" as defined by international law, thereby potentially triggering a right to self-defense or a proportional response in kind. This critical inquiry is central to the Tallinn Manual [8], a highly influential document in the field, building upon earlier contributions by Schmitt [25, 34], Jensen [35], and Wingfield [31]. The analysis of such an attack often employs criteria, such as the seven defined by Schmitt, to evaluate whether it crosses the "use of force" threshold [34]. These attacks, when exceeding the limits of international law, can incur collateral damage in the strict legal sense, meaning they cause disproportionate or undistinguished harm to civilians or civilian objects.

The publication of the Tallinn Manual profoundly bifurcated research in this area, creating a "before" and "after" landscape. Earlier works, such as that by O'Donnell and Kraska [36], while valuable, have seen their influence wane as subsequent scholarship has increasingly drawn upon, elaborated on, or critiqued the Manual [37]. Further in-depth analyses of the LOAC principles of proportionality (Rule 14) and distinction (Rule 31) are provided by Normelli [38], Pascucci [39], Bannelier [40], Beard [41], and Geiß and Lahmann [42]. These papers rigorously interpret and apply the laws of armed conflict to the cyber domain. A broader perspective on the consequences for civilians within a cyber warfare context is offered by Jensen [43], Brenner [44], Schmitt [45], and Droege [46]. These authors highlight that the inherent interconnectedness of the cyber realm, coupled with the proliferation of dual-use systems (e.g., cloud platforms, undersea communication cables) that underpin critical infrastructure, suggests that existing international law may be insufficient to fully address the challenges. Furthermore, while some nations acknowledge cyber warfare as legitimate warfare subject to regulation, this view is not universally accepted, and no cases have yet been adjudicated under international law. The limited interest from non-aligned nations in a comprehensive cyber treaty [46] or in voluntary regulation of cyber warfare further constrains the practical applicability of the Manual.

More recent legal scholarship, particularly following the Oslo Manual on Select Topics of the Law of Armed Conflict (2020), has shifted towards integrating the regulation of cyber operations with that of other operations carrying high risks of collateral damage, such as air and missile warfare [47]. The legal frameworks governing aerial attacks are frequently cited as comparative sources for applying the laws of war to the cyber domain [14, 33]. This integration acknowledges the increasingly hybrid nature of modern conflicts.

A fundamental challenge underscored by legal scholars is the incommensurability of military utility and civilian harm. As Dinstein [14] aptly summarizes, "The difficulty is that military advantage and civilian casualties have no common denominator." Public perception of war's undesirable impact is often measured in human lives lost, whereas cyber operations typically inflict significant economic damage with little to no direct loss of life. This creates a disconnect between legal frameworks, which treat property damage and loss of life as equally unlawful, and public perception. Cyber conflict introduces additional complexities related to attribution and deterrence. There is a natural inclination to apply established laws and doctrines to address these challenges in contexts beyond traditional armed conflict. Romanosky and Goldman [13] notably explore the transferability of military law and doctrine to law enforcement, marking an early attempt to broaden the definition of cyber collateral damage.

3.3. Targeting in Cyber Operations: Collateral Damage Considerations

Another significant cluster of research focuses on the intricate process of targeting cyber operations to achieve desired effects while meticulously minimizing unintended collateral harm. Papers within this category delve into how cyber weapons can be precisely aimed, or cyber operations conducted against a specific and delimited target, in a manner that curtails broader, undesirable effects. Much of this research commences by formally establishing a basis for the systematic evaluation of both intended and unintended consequences. Scholars such as Fanelli and Conti [48], Ducheine and van Haaster [49], Maathuis, Pieters, and van der Berg [17], and Orye and Maennel [50] have all proposed such foundational definitions and frameworks.

Despite variations in their specific approaches, three main points of consensus emerge from these works regarding targeting considerations:

1. Lack of Common Definitions: There is a persistent absence of universally accepted definitions for core concepts like "cyberwar," "information war," or even "collateral damage" itself, which complicates consistent application.

2. Undesirability and Illegality of Collateral Damage: There is a shared understanding that collateral damage is inherently undesirable, often illegal under international law, and must be rigorously accounted for in planning and execution.

3. Difficulty in Measuring Cyber Effects: Quantifying and predicting the precise effects of cyber operations, particularly their cascading and unintended consequences, remains a significant challenge.

The case of Stuxnet, a "game-changing" cyber weapon [51], is frequently cited. Its inclusion of several failsafes designed to prevent unintended harm served as a foundational observation, inspiring the argument that OCOs can and should be developed and deployed discriminately. This suggests that it is both a legal imperative (under the laws of armed conflict) and a technical possibility to design cyber operations with precision. Hirsch [52] further elaborates on the failure of such containment mechanisms, which can lead to indiscriminate, virus-like spread, effectively representing a loss of integrity in the operation. While many authors adopt a legalistic perspective, implicitly assuming adherence to international law, a more pragmatic motivation for precise targeting and stealthy operation is also recognized: to enhance the efficiency of cyber weapons and, crucially, to avert unintended escalation and retaliatory actions [48, 53, 54].

One of the most widely referenced and adapted targeting frameworks from conventional warfare is the Joint Targeting Cycle (JTC) [7]. The concept of applying this established cycle to the cyber domain has been

independently advocated by Ducheine and van Haaster [49], Smart [55], Couretas [56], and Monge and Vidal [57]. Their core motivation is that integrating such a framework would mandate specified steps for calculating and predicting collateral damage before operations are executed, and subsequently ensure that collateral outcomes are rigorously assessed after their completion. This systematic approach aims to embed collateral damage considerations directly into the operational planning process.

Conversely, a "reverse approach" to targeting involves identifying attack modalities that inherently carry an increased risk of collateral damage. Libicki [58] provides a valuable categorization of five such high-risk cyberattack modalities, including replicating attack vectors, drive-by attacks, supply chain attacks, third-party Distributed Denial of Service (DDoS) attacks (bots), and flooding attacks (DDoS). Each of these modalities carries specific collateral damage risks, such as complicated and expensive clean-up, scattershot impact on unintended targets, unpredictable results from third-party involvement, and spillover effects to adjacent systems or networks. Understanding these inherent risks is crucial for mitigating unintended consequences.

3.4. Ethical Aspects of Collateral Damage from Offensive Cyber Operations

While ethical considerations might intuitively appear central to all discussions of collateral damage, only a specific subset of papers has taken them as their primary focus. These papers deeply investigate the moral dimensions of conducting cyber operations to actively reduce the risk of unintended harm. Operating under the fundamental premise that "cyberwarfare is warfare" [8], the principles of Just War Theory become directly applicable [59, 60]. This theory posits that wars must not only be initiated for morally justifiable reasons (jus ad bellum) but also conducted ethically (jus in bello). Consequently, cyber weapons must be meticulously designed with ethical considerations embedded, and their deployment must adhere to moral principles against permissible targets [61, 62]. Denning [59] further emphasizes that, ideally, the judicious use of cyber capabilities should inherently reduce the collateral damage inflicted by any given operation when compared to conventional warfare.

The cluster of papers addressing ethical issues is closely intertwined with the "targeting" category. The reasoning is straightforward: enhanced targeting capabilities would directly lead to a reduction in collateral damage, and theoretically, perfect targeting would eliminate it entirely, thereby resolving key ethical dilemmas. However, the practical impossibility of achieving perfect targeting necessitates a robust consideration of the ethics surrounding the inevitable collateral effects. This seemingly fundamental line of reasoning is advanced by both Rowe [63] and Hare [29], who highlight that while cyber operations offer the promise of action at a distance, this very advantage can become a "curse" due to the vast range of operations and the inherent difficulties in remotely identifying the precise context of any given target.

Hare [29] also references a notable statement from a U.S. official: "If you shut down our power grid, maybe we will put a missile down one of your smokestacks" [64]. This statement implicitly links cyber attacks to kinetic responses, specifically mentioning the use of precision-guided munitions (PGMs). Acton [54] and Hare [29] both draw parallels between PGMs and "sophisticated" cyber weapons. While the legal similarities have been previously noted, the evolution of PGMs was accompanied by extensive discussions regarding their potential for improved discrimination and more ethical conduct in warfare. Research into ethical cyber operations appears to follow a similar trajectory, seeking to imbue cyber capabilities with principles that minimize unintended harm [65].

Both legal and ethical challenges are significantly amplified by the frequent involvement of perfidy in cyber operations. This refers to the obfuscation of the military origins of an attack or the impersonation of civilian entities [25, 63, 66]. While such deception can enhance military utility by improving the chances of mission success and complicating attribution [63], the normalization of these unethical practices raises profound long-term questions regarding the credibility and legitimacy of regular armed forces operating in cyberspace. Similarly, many operations leverage weaponized malware or newly developed military malware. Despite being designed with mechanisms to prevent uncontrolled spread, malware, by its very nature, is "malicious for a reason" [67]. Numerous instances have shown that tools exploiting undisclosed vulnerabilities can fall into unauthorized hands, with limited official commentary on improved practices to mitigate this significant societal risk [68].

3.5. Econometric Aspects of Cyber Collateral Damage

A formidable challenge in comparing the military utility of an offensive cyber operation with its potential collateral damage lies in the disparate metrics used for evaluation. Unlike military objectives, which are rarely assigned a precise monetary value, the costs of collateral damage, particularly in the cyber domain, are often quantifiable in economic terms. Historically, the economics of warfare has primarily focused on optimizing national capacity to wage war effectively and achieve maximum value for military expenditure [69]. More recently, sophisticated econometric models have emerged to systematically quantify the broader societal costs of warfare [70]. However, these models have yet to be broadly adapted to the unique characteristics of the cyberspace domain.

Despite this gap, several papers in this review explore the economic costs associated with conducting cyber operations, both directly to the perpetrator and indirectly to society. Levine [71] provides insights into some direct

costs incurred from specific cyber operations. Sigholm and Larsson [72, 73] present a compelling longitudinal study that seeks to understand the cost-benefit trade-off inherent in implanting vulnerabilities into civilian or dual-use software. Drawing upon data from the early phases of the Ukraine conflict (2014-2021), Larsson and Sigholm [74] propose and demonstrate a methodology for conducting an economic bottom-up assessment of societal harm resulting from cyber warfare. This pioneering work highlights the potential for a more rigorous economic analysis of cyber conflict.

A common approach in cyber conflict discourse is the translation of concepts from the physical world and conventional operations to the digital domain. Kohler [75] observes a long-standing practice of compensating for collateral damage in neutral territories. While this principle typically does not extend to combatants themselves (whose compensations are usually settled in peace agreements), the concept of gaining protection by relocating information operations to neutral and protected territories has been notably demonstrated by Ukraine. By rapidly migrating digital services to US-owned cloud services in NATO-affiliated countries, Ukraine has not only significantly enhanced its protection against overt cyberattacks but also established a conduit for economic support and sponsorship [76]. A comprehensive economic analysis of cyber operations within this ongoing conflict is still needed to fully understand its financial implications.

3.6. Collateral Damage Estimation, Modeling, and Assessment (CDE)

The field of conventional warfare boasts well-established and formalized models for the estimation of collateral damage [77]. These models are so mature that the U.S. Armed Forces offers a standardized five-day course to train and certify personnel in collateral damage estimation [78]. A testament to their maturity is the consistent application of these concepts and frameworks by both U.S. and European Union forces, often utilizing models similar to the one depicted in Table 10 and Figure 6 (from the source PDF).

In stark contrast, unclassified research applying these sophisticated concepts directly to the cyber domain, particularly to offensive cyber operations (OCOs), remains limited and often accompanied by significant scope limitations and qualifying statements. Romanosky and Goldman [13] explicitly note that "in cyber and cyber-physical systems [...] collateral effects can be much more difficult to predict, rendering ineffective traditional approaches to collateral damage estimation." Despite this acknowledged difficulty, the imperative for robust methods or models persists.

Orye and Maennel [50] offer a pragmatic approach by adapting a U.S. questionnaire, incorporating several qualitative questions for consideration prior to launching an attack. This provides a simple yet robust framework for initial assessment. More ambitious models are presented by Maathuis, Pieters, and van der Berg [17, 18] and Fanelli and Conti [48]. Both sets of authors adopt a similar methodological approach: they categorize effects into various "planes" (e.g., temporal, geospatial, logical) and evaluate these effects along a spectrum from benign to severe. The authors of [17] closely mirror the five-step approach used in conventional collateral damage estimation, but critically, they also integrate an assessment of military advantages and disadvantages specific to cyber operations.

The most ambitious and comprehensive models for the assessment of CCD have been put forth by Maathuis and Chockalingam [79] and Maathuis [80]. They propose the integration of probabilistic and machine learning (ML)-assisted methods to estimate the effects of cyber operations. The ultimate goal of these advanced models is to ensure that the collateral effects remain proportional to the anticipated military utility, thereby aligning cyber operations with legal and ethical principles. These efforts represent the cutting edge of research in formalizing and quantifying the unpredictable nature of cyber collateral damage.

3.7. Taxonomies of Cyber Collateral Damage

To systematically describe the diverse causes and manifestations of Cyber Collateral Damage (CCD), several researchers have developed taxonomies. This cluster, though containing only three papers, is fundamentally important for theorizing and structuring CCD research. Notable contributions in this area include works by Rowe [63], Raymond et al. [82], and Bertoli and Marvel [83]. It is important to note that their respective models exhibit noticeable dissimilarities, reflecting the nascent stage of this field.

Rowe [63] and Raymond et al. [82] primarily focus on lists of distinct factors that contribute to CCD. These factors can be unified and broadly categorized as follows:

- Direct Collateral Damage: Immediate unintended harm to non-target systems or entities.
- Errors in Targeting: Damage resulting from mistakes in identifying or executing attacks against legitimate targets.
- Costs of Recovery: Expenses incurred in recovering from direct damage caused by a cyber weapon.
- Costs of Attack Propagation: Expenses and disruptions arising from the uncontrolled spread of a cyber weapon beyond its intended scope.
- Costs of Attack Analysis and Mitigation: Resources expended to understand the attack, develop countermeasures, and implement defensive measures.
- Psychological Damage: Non-physical harm, such as fear, anxiety, or loss of trust, inflicted on affected populations.

- Costs of Vulnerability Disclosure: The broader societal costs associated with the public or illicit release of vulnerabilities exploited in an attack.

Bertoli and Marvel [83] offer a more simplified yet interactive approach, mapping Rowe's categories [63] into four interconnected contributing factors, as illustrated in Figure 7 (from the source PDF). These factors include:

- Errors: Collateral damage stemming from flaws in system implementation or operational execution.
- Target Dependencies: Damage resulting solely from the inherent characteristics and interconnections of the target system.
- Weapon Dependencies: Collateral damage arising from the intrinsic properties of the cyber weapon itself and its usage characteristics.
- Political Ramifications: The consideration of the political cost associated with the employment of military force, which can influence the perceived acceptability of collateral damage.

The noticeable dissimilarity among these taxonomies underscores the ongoing effort to establish a standardized and comprehensive framework for understanding CCD. While these models provide valuable conceptual tools, a critical observation is that no author has yet rigorously applied their proposed taxonomy to a real-world cyber operation. A future study systematically mapping concrete cases to these taxonomies would be invaluable in validating their completeness and practical utility. Such a case study, similar to the critical analysis of the Tallinn Manual's applicability to 11 different OCOs in [37], would significantly advance the field.

Table 11 (from the source PDF) lists the key papers in the "taxonomy" category, highlighting the foundational efforts to categorize and understand the complex nature of cyber collateral damage.

3.8. Notable Contributions Outside Categories

Beyond the primary thematic clusters, a few papers offer distinct and notable contributions to the understanding of cyber collateral damage, addressing aspects that do not neatly fit into the identified categories. Three such significant papers are highlighted in this section.

While offensive cyber operations (OCOs) are frequently employed to achieve specific operational and tactical goals, their broader implications are often considered in terms of strategic effects. In particular, Smeets [84] and Lawson and Mačák [85] delve into the disproportionate impacts of cyber operations when directed against critical infrastructure, such as national electric grids, water supply networks, and healthcare systems. The unique characteristics of cyber operations—including the potential for immediate action from a geographically distant or unbounded range, the necessity to prepare and stockpile vulnerabilities in advance of an attack, and the inherent risk of impacting significant areas beyond the intended target—bear striking resemblances to the strategic implications of nuclear capabilities, which are primarily maintained as a deterrent.

Conversely, Smeets [84] argues that the potential for highly precise cyber strikes with minimal collateral damage could offer an additional, flexible option for state leaders, distinct from traditional kinetic attacks. Furthermore, unlike conventional attacks, one could envision the strategic use of mechanisms like crypto-ransomware to render the damage from certain cyberattacks reversible [86]. The concept of reversibility could potentially reduce the hesitation associated with ordering an attack and, crucially, enable the selective reversal of effects on civilian targets, thereby mitigating collateral damage. This idea introduces a novel dimension to the ethical and strategic calculus of OCOs.

Table 12 (from the source PDF) lists these key papers that provide unique insights outside the main thematic categories, enriching the overall understanding of cyber collateral damage.

3.9. Bibliographic Analysis

A comprehensive bibliographic analysis of the reviewed papers was conducted across four distinct metrics: the annual publication rate, identification of key publication outlets, citation counts to determine influential works, and the clustering of internal references to understand the intellectual interconnectedness within the field.

3.9.1. Publication Trends Over Time

The number of papers published each year from 1998 to 2024 is graphically represented in Figure 8 (from the source PDF). The data clearly indicate a significant surge in the overall number of papers on Cyber Collateral Damage (CCD) after the mid-2000s. This upward trend suggests a growing academic and policy interest in the topic, likely spurred by notable real-world cyber incidents and the increasing recognition of cyberspace as a domain of conflict. However, following a peak around 2014-2015 and another in 2021, the research interest appears to have settled into a lower but consistent level in the most recent years. This pattern could be attributed to a temporary peak in interest following the highly publicized Stuxnet operation in 2010, which brought the issue of unintended cyber effects to the forefront. The understanding and regulation of cyber operations within a military context, particularly through the lens of the laws of armed conflict, naturally gains more traction when such operations are actively being conducted. If this hypothesis holds true, it would be reasonable to anticipate a renewed surge in interest in CCD research as the ongoing Russo-Ukrainian war continues to feature prominent cyber operations.

3.9.2. Key Publication Outlets

An analysis of the publication outlets reveals a dispersed landscape. Of the 74 papers included in this review, no single publication outlet contributed more than seven

papers. Only six publication outlets featured more than one included publication (Figure 9 from the source PDF). The NATO-sponsored International Conference of Cyber Conflict (CyCon) and the various conferences under the umbrella of the Conference on Cyber Warfare and Security (CW/CCWS) (formerly the Conference of Information Warfare and Security) emerge as undoubtedly the most important and consistent publication venues in this specific field. However, the review also identified a considerable number of other conferences, journals, books, and reports. Excluding books and reports, this review encompasses papers from 28 distinct journals and conference proceedings. This wide distribution suggests that research on CCD is published across various disciplinary boundaries, reflecting its interdisciplinary nature. While a previous literature review [19] exclusively relied on the proceedings of a single conference, this analysis strongly suggests that such a narrow approach would be insufficient for comprehensive coverage. Among the identified journals, the Journal of Cybersecurity is particularly notable for its recognition of the inherently interdisciplinary nature of the cyber domain and its encouragement of a broad range of submissions.

3.9.3. Most Cited Papers and Authors

As noted in Section 4.3, legal papers constituted the most common category identified in the survey. This dominance is further reinforced when considering the most referenced papers. Papers investigating the application of law to cyber operations are once again the most numerous among the highly cited works. The number of citations for the ten most cited papers in the survey is presented in Figure 10 (from the source PDF). The two most cited authors within the survey are Michael Schmitt and Eric Talbot Jensen, each contributing two of the ten most cited papers. Significantly, the Tallinn Manual, a cornerstone publication in cyber law, was also edited by Schmitt, further cementing his influence. Six of the ten most cited papers are legal works, while the remaining three are textbooks or articles that provide broad overviews of the field of cyber warfare. This highlights the foundational role of legal and general conceptual works in shaping the discourse on cyber collateral damage.

3.9.4. Reference Interconnections and Clustering

The intellectual relationships and interconnections between the papers in the survey were visually represented using a reference graph (Figure 11 from the source PDF). This graph was constructed using the Yifan Hu layout algorithm [87], a physics-based model that minimizes the energy of the system, thereby positioning interconnected nodes closer together.

The visualization prominently places the Tallinn Manual (with 23 internal references) at the center of the graph, underscoring its pivotal role as a central reference point for much of the subsequent scholarship. The lower half of the graph is predominantly occupied by other legal works that frequently reference Schmitt 2002 [25] (14 references) and Schmitt 1998 [34] (10 references). A crucial observation is that nearly all of these references predate the publication of the Tallinn Manual, suggesting that the Manual has largely superseded these earlier works as the primary reference point for cyber law. The papers within the legal CCD category generally exhibit a much higher level of reference interconnection, strongly indicating that this research area functions as a cohesive and well-established academic community.

The upper-left quadrant of the graph features a notable cluster around the work of Romanosky and Goldman [13] (nine references). Citations to their work typically occur in the context of explaining the cyber-physical nature of CCD, where a cyberattack has collateral effects on real-world services and equipment. This highlights a distinct sub-focus on the tangible impacts of cyber operations. Finally, the upper-central area of the graph illustrates the strong interconnections among the works of Clara Maathuis, who stands out as one of the few researchers predominantly dedicated to the study of CCD. Her work, as seen in the clustering, plays a significant role in advancing the understanding of this complex domain.

4. Discussion

This section delves into the critical analysis of the research landscape concerning cyber collateral damage (CCD), identifying prevailing trends, significant knowledge gaps, and proposing opportunities for synthesis and future research.

4.1. Research Trends and Gaps

The categorization of papers into six emergent themes provides a structured view of the CCD research landscape. Out of the 74 papers reviewed, 63 covered material from one or more of these categories. However, a striking observation is that only seven papers demonstrated in-depth coverage of more than one category. This predominantly unidisciplinary approach suggests that while individual aspects of CCD are being investigated, the interconnections and holistic understanding across these dimensions remain underexplored.

Well-researched interfield areas include the intersections between cyber law and cyber ethics, and between cyber law and targeting. However, a notable lacuna exists in the interdisciplinary coverage between targeting and financial estimation. For instance, no paper in this review rigorously explored the use of econometric methods, such as bottom-up accounting, during the planning phase of an operation to estimate and proactively reduce the economic value of expected collateral damage. This represents a significant opportunity for future research to integrate economic impact assessments directly into operational planning. Another critical gap is the scarcity of multidisciplinary studies that assess the most appropriate response to a given cyberattack by considering possibilities across various domains (e.g., legal, technical,

economic, diplomatic), rather than narrowly within a single category of inquiry.

The evolution of the research field over time, characterized by periods of expansion and contraction (as seen in Figure 8), also reveals underlying trends and persistent gaps. A clear, overarching theme across all CCD research is the paramount importance of targeting. This issue is repeatedly emphasized, even in papers not explicitly centered on targeting [29]. It can be asserted that the single most crucial area for improving OCO methodologies to reduce collateral damage is through enhanced targeting precision and assessment. A straightforward, albeit potentially simplistic, approach would involve applying the Joint Targeting Cycle (JTC) to all cyber operations, as proposed by Ducheine and van Haaster [49] and Monge and Vidal [57]. The rationale behind this is that the JTC would mandate specific, formalized steps to calculate and predict collateral damage before operations and ensure post-operation assessment of collateral outcomes.

However, the interface between open research and military targeting methodologies remains opaque. Declassified documents from U.S. Cyber Command indicate that the JTC approach was considered but ultimately rejected in 2016 for cyber operations. The reasoning cited was that the JTC is "optimized for lethal effects but sub-optimized for nonlethal effects" [88]. Implicit in this reasoning is the virtuous intent of combatants to actively avoid collateral damage and to utilize tools and techniques specifically developed for this purpose. While classified methods and processes for estimating, controlling, and minimizing CCD undoubtedly exist within the world's militaries, it is worth considering whether the net benefit of open international collaboration outweighs the operational risk of disclosing these methods. Policymakers should not passively await such collaboration but should proactively ensure that national cyber doctrines explicitly direct and require cyber operations to be conducted using formalized and explainable processes for targeting. Such explicit language is often absent in current doctrines, with some merely referencing offensive operations "below the level of armed conflict" [89].

Another critical research gap concerns the economic and econometric analysis of offensive cyber operations. Research on defense economics has traditionally focused on maximizing military capability for a given expenditure [69]. Conducting a robust cost-benefit analysis on defense investments necessitates a reliable and efficient means of measuring and comparing costs and benefits. Previous research on military cost-benefit analysis has employed the Value of a Statistical Life (VSL) as a measure of benefits [96]. For example, force protection measures like vehicle armor plating can be directly compared in terms of the additional cost against the statistical value of lives saved [97]. However, transferring these concepts directly to the cyber domain is not straightforward. For general cyber incidents, considerable research aims to quantify the total damage to affected organizations, drawing from industry estimates or models like the Real Cyber Value at Risk (RCVaR) [74, 98]. The cost of conducting an offensive operation remains less understood due to operational secrecy, though some cost estimation models exist, such as RAND Corporation's modeling of cyber weapon acquisition costs for the U.S. Marine Corps based on darknet exploit pricing [99].

The data availability problem can be partially circumvented by focusing on the more constrained question of the collateral cost relative to military benefit. Recent analyses have proposed evaluating the benefit of military operations, including cyber operations, in terms of their military utility [100-102]. However, these analyses are often qualitatively driven, leading to a lack of numerical comparisons of cost to benefit. Furthermore, the well-explored VSL model is a poor fit for cyber operations where direct loss of life is rare. Further developing a quantitative approach to the utility of cyber operations would enable more rigorous cost-benefit analyses and facilitate more informed decision-making regarding the conduct of cyber operations.

Only three papers in this review provided taxonomies of the types and causes of CCD. This might suggest that the conceptualization of CCD is maturing. While the work of Bertoli and Marvel [83] is notable, it is unfortunate that no author has yet applied their taxonomy to a real-world operation. A study systematically mapping concrete cases to this taxonomy would be invaluable in demonstrating its validity and completeness. A similar case study, such as the critical analysis of the Tallinn Manual's applicability to 11 different OCOs in [37], could serve as a reference for such an endeavor.

Finally, it is crucial to consider which motivations for the development and use of cyber capabilities lead to the greatest amount of CCD. As noted by Cavaiola, Gompert, and Libicki [53], OCOs can be conducted across a wider spectrum of intensity than conventional operations, influencing public perceptions of attack severity. National differences in attitudes towards "total war" also shape what is considered acceptable CCD, as does the naming and identification of victims [103]. Many actors are clearly interested in cyber capabilities that can be used without overtly escalating conflict in the public eye, as well as in "below-threshold" capabilities that extend soft power. A potential research question is to investigate how these capabilities differ from general and/or "above-threshold" capabilities technically, psychologically, and in terms of cost and benefit. While the public may find more discrete capabilities palatable, there is also an increased risk of leaks and diplomatic fallout associated with zero-day hoarding, self-replicating weapons, and perfidy. The debate on the trade-off between cyber capability cost and value will intensify as the relationship between these aspects becomes better understood.

4.2. Legal Responses to Cyberattacks in Theory and

Practice

The cyberattacks against Estonia in 2007 and the Stuxnet incident in 2010 served as catalysts for the creation of the initial version of the Tallinn Manual. This foundational document aimed to delineate the application of the laws of war to cyber warfare, specifically addressing when cyberattacks transcend the "use of force" threshold as defined by Article 2(4) of the UN Charter [34]. The subsequent absence of such high-threshold attacks informed the development of the second version of the Manual, which broadened its scope to encompass attacks falling below this threshold [37]. The original premise of the Manual, that "cyber warfare is warfare," has found correspondence in real-world conditions, notably during the 2022 invasion of Ukraine. On the day of the invasion, kinetic warfare was combined with denial-of-service attacks targeting government agencies, as well as a malicious update that disabled Viasat satellite modems used by civilians. In response, numerous Ukrainian government agencies swiftly migrated their digital services to U.S.-owned cloud services hosted in NATO-affiliated countries [76]. This digital exodus proved effective in shielding them from the most overt and easily attributable cyberattacks, although exploratory and clandestine operations by state-backed actors persist, with Microsoft reporting attacks by 14 such actors in 2024 [104].

While the Viasat attack and the NotPetya ransomware attacks directly targeted civilians for military gain, other significant cyber incidents, such as the 2010 Bangladesh Bank cyber robbery and the 2017 WannaCry ransomware attacks on urban critical infrastructure (both formally attributed to North Korea), do not neatly fit into the conventional understanding of interstate use of force. Nevertheless, local governments worldwide are under constant cyberattack [95]. Generally, the legal publications reviewed in this study have not comprehensively addressed these "grey-zone" attacks [30]. In practice, these attacks have primarily been met with diplomatic responses, such as sanctions, rather than formal legal challenges or military retaliation.

This persistent gap between legal understanding and the reality of cyber conflict has not gone unnoticed. The NATO Cooperative Cyber Defence Centre of Excellence (CCDCOE) is currently developing the Tallinn Manual 3.0 under a five-year project spanning from 2021 to 2026. While this program is expected to address the blatant disregard for civilian harm observed in the Ukraine conflict's cyberattacks, it remains unclear whether a NATO-centric program is the most appropriate forum to address state-backed cybercrime targeting banks and civilian entities for purely financial gain. In the absence of a specific treaty on conduct in cyberspace, as advocated by Droege [46] and others, a separate organization should be integrated into the Tallinn 3.0 process. This entity could assume a coordinating role in addressing below-threshold attacks strongly suspected of having an interstate nature. Although these attacks constitute crimes and could fall under the purview of INTERPOL or the United Nations Office on Drugs and Crime (UNODC), cyber conflict resolution has increasingly taken the form of sanctions and diplomacy [105]. Reflecting this reality, the UN has established an Open-Ended Working Group on information and communication technologies [106], tasked with investigating responsible state behavior in cyberspace. This working group could provide a much-needed component of norms and diplomacy to complement the Tallinn process.

4.3. The Use of AI in Collateral Damage Estimation and Assessment

Some of the most recent papers included in this review explore the burgeoning intersection of deep learning and generative artificial intelligence (AI) with cyber operations [107, 108]. While this area of research was not yet prevalent enough among the reviewed papers to draw broad conclusions, AI and Machine Learning (ML) represent fields of rapidly increasing research interest with considerable applicability to cyber operations. Some of the most promising areas for AI/ML implementation directly pertain to reducing collateral damage: decision support systems for enhanced targeting, and explanatory tools to verify the legality and ethical permissibility of proposed actions.

Within the domain of privacy research, various AI-assisted compliance verification tools already exist, for instance, in relation to GDPR [109]. This type of tool could potentially assist or even replace the "lawyer in the loop" traditionally necessary for setting rules of engagement (ROE), identifying and assessing targets, or determining the permissibility of attacks in both cyber and conventional contexts. However, the use of AI tools when there is a risk of human harm introduces a significant explainability challenge, particularly when these tools are based on "black-box" reasoning, rendering the system's basis for determination opaque. Explainable AI (XAI) tools offer a potential solution to this dilemma, though often at the cost of reduced accuracy compared to pure deep learning systems [110]. These XAI systems are designed to be explainable, transparent, and interpretable, capable of providing not just a decision but also an accompanying chain of reasoning that can be followed and verified. The substantial research effort dedicated to the use of XAI in other high-risk fields, such as clinical decision support, could be recontextualized and adapted for cyber ROE. At present, state-of-the-art XAI decision support systems are not yet able to reliably explain all encountered cases [111]. Nevertheless, an AI-proposed solution is likely to significantly reduce the manual effort compared to producing a solution for each operation.

The development of AI tools for a cyber warfare context necessitates training data, which is often sensitive or classified. This inherent secrecy increases the risk of bias in decision-making, where existing process biases might be inadvertently replicated by the AI system. As the open

and external verification of the AI's decision-making is often untenable from a security perspective, developers must ensure that AI systems used for decision support are also rigorously trained on appropriate ethical frameworks, for example, by embedding the principles of Just War Theory as a fundamental "guardrail."

Previous studies applying ML tools to collateral damage assessment and estimation have highlighted the limited number of existing models that can be directly applied to machine-based evaluation [79, 80]. When designing a system that combines human and AI elements for decision-making, there is a risk that crucial value judgments are performed as part of data entry (e.g., when coding the CCD potential as high, medium, or low for a specific operation, as in [79]). The decision basis generated by the system may appear credible due to the use of advanced technology, when in practice, a simpler lookup or heuristic might have sufficed using the coded inputs provided by an experienced system operator. The true value of AI decision support is realized when it can genuinely replicate, replace, or offload these operators and form independent value judgments. Achieving this in an explainable manner holds the potential to both reduce the workload of CCD assessment, improve targeting precision, and ultimately lessen the collateral impacts of cyber operations.

4.4. Future Research Opportunities

The vast majority of research on Cyber Collateral Damage (CCD) has historically been derived from first-principles thinking, primarily rooted in legal frameworks and military doctrine. While open datasets of offensive cyber operations (OCOs) exist [112, 113], they have been scarcely utilized for empirical and quantitative studies. A robust future research agenda would significantly benefit from incorporating these real-world data more extensively to assess societal impacts. Similarly, several intriguing methodologies have been proposed for targeting, damage modeling, and assessment [17, 18, 48, 50]. However, few, if any, of these models have been rigorously tested in simulations or validated using real-world data. A scenario-based simulation study comparing two or more models could provide a valuable platform for open discussion and the refinement of cyber doctrine.

Mining existing datasets to understand the "where, when, who, and why" of OCOs would also provide a more comprehensive picture of potential research avenues. A foundational step could involve verifying the analysis recently provided by the Cyber Peace Institute [113], potentially using the methodology outlined in [114]. A similar analysis for the period of 2013-2021 was performed in [74] by applying econometric methods (bottom-up accounting and counterfactual analysis) to the cyber domain. The continuation of this study, in interdisciplinary collaboration with economists, would significantly enhance the rigor and validity of the results. A broader study of the economic cost of cyberattacks on critical infrastructure and local governments would add substantial value, though the challenge of distinguishing attributable cyber operations from for-profit ransomware complicates the data landscape. Nevertheless, research that reproduces or improves upon the cost estimates provided by industry reports and "grey literature" (e.g., [2, 104]) would considerably deepen the understanding of the societal cost and risk associated with cyber incidents.

While retroactive econometric estimation of collateral damage can identify the human and hidden costs of cyber warfare, proactive collateral damage estimation models hold the potential to reduce or even obviate these costs by informing cyber Rules of Engagement (ROE) and aiding in the selection of appropriate missions and targets. As Romanosky and Goldman [13] noted, traditional approaches to collateral damage estimation are often ineffective for cyber operations. Their proposed collateral damage estimation model, while valuable for its simplicity and robustness, includes a step that asks, "What is my new estimate of the range of collateral effects to data, computing, or IT systems?"—a question that itself necessitates a considerably more complex sub-model to answer. This research area also presents a significant opportunity for improved national cyber policy, specifically by requiring that each cyber operation be followed up with a formal estimation of the collateral damage caused. Performing such analysis as part of a Battle Damage Assessment (BDA) is already routine in conventional warfare, and there is no compelling reason why it should not be consistently included when operating in the cyber domain.

5. Conclusion

This systematic literature review aimed to provide a comprehensive overview of research concerning collateral damage in the cyber domain. The nascent stage of this field is evident from the relatively limited number of papers directly identified through keyword searches. The existing research primarily categorizes into legal, ethical, targeting-oriented, and econometric studies, with smaller contributions in the areas of collateral damage estimation and taxonomy. Legal scholarship on CCD stands out as the most developed and intellectually interconnected category. A significant finding is the scarcity of interdisciplinary papers that bridge two or more of these categories, suggesting a considerable potential for an enhanced understanding of CCD and civilian harm through collaborative work across disciplinary boundaries.

The ultimate objective of improving our understanding of CCD is to effectively reduce its occurrence and impact. At present, substantial work remains in comprehensively understanding CCD, particularly beyond its strict legal context. A major limitation in defining and scoping CCD research lies in the varied interpretations of the term "collateral damage." Public perceptions of cyber operations often encompass a broad spectrum, including cybercrime, warfare, espionage, sabotage, subversion, and illicit attacks against civilian targets. While the laws of

armed conflict would theoretically protect these targets, these laws are strictly applicable only to military cyber operations conducted at a "use of force" level, as understood under the Geneva Conventions. The majority of real-world cyberattacks, however, occur below this threshold.

This strict compartmentalization of "collateral damage" as distinct from broader "civilian harm" is useful for academic research but often fails to reflect public perception and the lived reality of those affected. The proposals by Droege [46] and Romanosky and Goldman [13] to expand the scope of international law to cover cyber operations conducted under civilian auspices could be highly effective in reducing overall civilian harm. Alternatively, the establishment of a dedicated cyber treaty or a robust resolution mechanism under the auspices of the United Nations could significantly reduce the number and impact of operations conducted below the threshold of declared warfare. Furthermore, the diplomatic resolution of low-level cyber conflict should be more fully integrated into processes like the ongoing Tallinn 3.0 initiative.

In essence, minimizing collateral damage in offensive cyber operations necessitates a concerted and collaborative effort from governments, international organizations, academia, and industry. This includes developing responsible policies, advancing sophisticated technologies, and implementing practices that are firmly aligned with humanitarian principles and the evolving landscape of international law. Only through such a holistic approach can the digital battlespace be navigated with greater precision and a reduced burden of unintended consequences on global stability and human well-being.

REFERENCES

1. Aiyer, B.; Caso, J.; Russell, P.; Sorel, M. Mckinsey: New Survey Reveals $2 Trillion Market Opportunity for Cybersecurity Technology and Service Providers. 2022. Available online: https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/cybersecurity/new-survey-reveals-2-trillion-dollar-market-opportunity-for-cybersecurity-technology-and-service-providers (accessed on 30 January 2024).
2. IBM Security; the Ponemon Institute. Cost of a Data Breach Report 2022. 2022. Available online: https://www.ibm.com/downloads/cas/3R8N1DZJ (accessed on 31 January 2024).
3. Hanson, F.; Uren, T. Australia's Offensive Cyber Capability. 2018. Available online: https://www.aspi.org.au/report/australias-offensive-cyber-capability (accessed on 31 January 2024).
4. Farwell, J.P.; Rohozinski, R. Stuxnet and the Future of Cyber War. Survival 2011, 53, 23–40.
5. U.S. Department of Justice. Three North Korean Military Hackers Indicted in Wide-Ranging Scheme to Commit Cyberattacks and Financial Crimes Across the Globe. 2022. Available online: https://www.justice.gov/opa/pr/three-north-korean-military-hackers-indicted-wide-ranging-scheme-commit-cyberattacks-and (accessed on 1 February 2024).
6. Schelling, T.C. Dispersal, deterrence, and damage. Oper. Res. 1961, 9, 363–370.
7. U.S. Air Force. Air Force Doctrine Publication 3–60, Targeting. 2021. Available online: https://www.doctrine.af.mil/Portals/61/documents/AFDP_3-60/3-60-AFDP-TARGETING.pdf (accessed on 1 February 2024).
8. Schmitt, M.N. (Ed.) Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations; Cambridge University Press: Cambridge, UK, 2017.
9. Arquilla, J.; Ronfeldt, D. Cyberwar is coming! Comp. Strategy 1993, 12, 141–165.
10. Lee, E.A.; Seshia, S.A. Introduction to Embedded Systems: A Cyber-Physical Systems Approach, 2nd ed.; MIT Press: Cambridge, MA, USA, 2017.
11. Romanosky, S.; Goldman, Z. Cyber Collateral Damage. Procedia Comput. Sci. 2016, 95, 10–17.
12. U.S. Air Force. Intelligence Targeting Guide, Attachment 7. 1998. Available online: https://irp.fas.org/doddir/usaf/afpam14-210/part20.htm (accessed on 1 February 2024).
13. Romanosky, S.; Goldman, Z. Understanding Cyber Collateral Damage. J. Natl. Secur. Law Policy 2017, 9, 233–257.
14. Dinstein, Y. The Principle of Distinction and Cyber War in International Armed Conflicts. J. Confl. Secur. Law 2012, 17, 261–277.
15. Ablon, L.; Binnendijk, A.; Hodgson, Q.E.; Lilly, B.; Romanosky, S.; Senty, D.; Thompson, J.A. Operationalizing Cyberspace as a Military Domain, Perspective, RAND Corporation 2019. Available online: https://www.rand.org/content/dam/rand/pubs/perspectives/PE300/PE329/RAND_PE329.pdf (accessed on 30 January 2024).
16. U.S. Department of Defense. Department of Defense Law of War Manual; William S. Hein & Company: Getzville, NY, USA, 2023.
17. Maathuis, C.; Pieters, W.; Van den Berg, J. Assessment Methodology for Collateral Damage and Military (Dis)Advantage in Cyber Operations. In Proceedings of the MILCOM 2018—2018 IEEE Military Communications Conference (MILCOM), Los Angeles, CA, USA, 29–31 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.

18. Maathuis, C.; Pieters, W.; van den Berg, J. Decision support model for effects estimation and proportionality assessment for targeting in cyber operations. Def. Technol. 2021, 17, 352–374.

19. Grant, T. Building an Ontology for Planning Attacks That Minimize Collateral Damage: Literature Survey. In Proceedings of the 14th International Conference on Cyber Warfare & Security (ICCWS 2019), Stellenbosch, South Africa, 28 February–1 March 2019; pp. 78–86.

20. University of York Centre for Reviews and Dissemination. What Are the Criteria for the Inclusion of Reviews on DARE? 2014. Available online: https://www.ncbi.nlm.nih.gov/books/NBK285222/ (accessed on 31 January 2024).

21. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ 2021, 372, n71.

22. Vaismoradi, M.; Snelgrove, S. Theme in Qualitative Content Analysis and Thematic Analysis. Forum Qual. Soc. Res. 2019, 20, 3.

23. Wanyama, S.B.; McQuaid, R.W.; Kittler, M. Where you search determines what you find: The effects of bibliographic databases on systematic reviews. Int. J. Soc. Res. Methodol. 2022, 25, 409–422.

24. Gusenbauer, M. Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. Scientometrics 2019, 118, 177–214.

25. Schmitt, M.N. Wired warfare: Computer network attack and jus in bello. Int. Rev. Red Cross 2002, 84, 365–399.

26. Naeem, M.; Ozuem, W.; Howell, K.; Ranfagni, S. A Step-by-Step Process of Thematic Analysis to Develop a Conceptual Model in Qualitative Research. Int. J. Qual. Methods 2023, 22, 1–18.

27. Randolph, J.J.; Bednarik, R. Publication Bias in the Computer Science Education Research Literature. J. Univers. Comput. Sci. 2008, 14, 575–589.

28. Robinson, M.; Jones, K.; Janicke, H. Cyber warfare: Issues and challenges. Comput. Secur. 2015, 49, 70–94.

29. Hare, F.B. Precision cyber weapon systems: An important component of a responsible national security strategy? Contemp. Secur. Policy 2019, 40, 193–213.

30. Hathaway, O.A.; Crootof, R.; Levitz, P.; Nix, H.; Nowlan, A.; Perdue, W.; Spiegel, J. The Law of Cyber-Attack. Calif. Law Rev. 2012, 100, 817–885.

31. Wingfield, T. International law and information operations. In Cyberpower and National Security; Kramer, F.D., Starr, S.H., Wentz, L.K., Eds.; NDU Press: Washington, DC, USA, 2009; pp. 525–542.

32. Sklerov, M. Responding to International Cyber Attacks as Acts of War. In Inside Cyber Warfare; Carr, J., Ed.; O'Reilly Media: Sebastopol, CA, USA, 2009.

33. Wang, Q. Applicability of Jus in Bello in Cyber Space: Dilemmas and Challenges. Int. J. Cyber Warf. Terror. 2014, 4, 43–62.

34. Schmitt, M.N. Computer Network Attack and the Use of Force in International Law: Thoughts on a Normative Framework. Columbia J. Transnatl. Law 1998, 37, 885.

35. Jensen, E.T. Unexpected Consequences From Knock-On Effects: A Different Standard for Computer Network Operations? Am. Univ. Int. Law Rev. 2003, 18, 1145–1188.

36. O'Donnell, B.T.; Kraska, J.C. Humanitarian Law: Developing International Rules for the Digital Battlefield. J. Confl. Secur. Law 2003, 8, 133–160.

37. Efrony, D.; Shany, Y. A Rule Book on the Shelf? Tallinn Manual 2.0 on Cyberoperations and Subsequent State Practice. Am. J. Int. Law 2018, 112, 583–657.

38. Normelli, N. Proportionality in Attack on Data: Balancing Military Advantage and Collateral Damage in Cyberspace; Uppsala University: Uppsala, Sweden, 2021.

39. Pascucci, P. Distinction and Proportionality in Cyberwar: Virtual Problems with a Real Solution. Minn. J. Int. Law 2017, 26, 419–460.

40. Bannelier, K. Is the principle of distinction still relevant in cyberwarfare? From doctrinal discourse to States' practice. In Research Handbook on International Law and Cyberspace; Tsagourias, N., Buchan, R., Eds.; Edward Elgar Publishing: Cheltenham, UK, 2015; pp. 343–365.

41. Beard, J.M. The principle of proportionality in an era of high technology. In Complex Battlespaces: The Law of Armed Conflict and the Dynamics of Modern Warfare; Ford, C.M., Williams, W.S., Eds.; Oxford University Press: Oxford, UK, 2018.

42. Geiß, R.; Lahmann, H. Cyber Warfare: Applying the Principle of Distinction in an Interconnected Space. Isr. Law Rev. 2012, 45, 381–399.

43. Jensen, E.T. Cyber Warfare and Precautions Against the Effects of Attacks. Tex. Law Rev. 2009, 88, 1533–1569.

44. Brenner, S.W.; Clarke, L.L. Civilians in Cyberwarfare: Casualties. SMU Sci. Technol. Law

Rev. 2010, 13, 249–282.

45. Schmitt, M.N. Wired warfare 3.0: Protecting the civilian population during cyber operations. Int. Rev. Red Cross 2019, 101, 333–355.

46. Droege, C. Get off my cloud: Cyber warfare, international humanitarian law, and the protection of civilians. Int. Rev. Red Cross 2013, 94, 533–578.

47. Dinstein, Y.; Dahl, A.W. Section II: Cyber Operations. In Oslo Manual on Select Topics of the Law of Armed Conflict; Springer: Cham, Switzerland, 2020.

48. Fanelli, R.; Conti, G. A methodology for cyber operations targeting and control of collateral damage in the context of lawful armed conflict. In Proceedings of the 2012 4th International Conference on Cyber Conflict (CYCON 2012), Tallinn, Estonia, 5–8 June 2012; pp. 1–13.

49. Ducheine, P.; van Haaster, J. Fighting power, targeting and cyber operations. In Proceedings of the 2014 6th International Conference On Cyber Conflict (CyCon 2014), Tallinn, Estonia, 3–6 June 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 303–327.

50. Orye, E.; Maennel, O.M. Recommendations for Enhancing the Results of Cyber Effects. In Proceedings of the 2019 11th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia, 28–31 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–19.

51. Denning, D.E. Stuxnet: What Has Changed? Future Internet 2012, 4, 672–687.

52. Hirsch, C. Collateral damage outcomes are prominent in cyber warfare, despite targeting. In Proceedings of the 13th International Conference on Cyber Warfare and Security, ICCWS 2018, Washington, DC, USA, 8–9 March 2018; Chen, J.Q., Ed.; Curran Associates, Inc.: Washington, DC, USA, 2018; pp. 281–286.

53. Cavaiola, L.J.; Gompert, D.C.; Libicki, M. Cyber House Rules: On War, Retaliation and Escalation. Survival 2015, 57, 81–104.

54. Acton, J.M. Cyber weapons and precision guided munitions. In Understanding Cyber Conflict: Fourteen Analogies; Perkovich, G., Levite, A.E., Eds.; Georgetown University Press: Washington, DC, USA, 2017; pp. 45–60.

55. Smart, S.J. Joint Targeting in Cyberspace. Air Space Power J. 2011, 25, 65–74.

56. Couretas, J.M. Cyber Offense and Targeting. In An Introduction to Cyber Analysis and Targeting; Springer International Publishing: Cham, Switzerland, 2022; pp. 151–172.

57. Monge, M.A.S.; Vidal, J.M. Conceptualization and cases of study on cyber operations against the sustainability of the tactical edge. Future Gener. Comput. Syst. 2021, 125, 869–890.

58. Libicki, M. Cyberspace in Peace and War, 2nd ed.; Naval Institute Press: Annapolis, MD, USA, 2021.

59. Denning, D.E.; Strawser, B.J. Moral Cyber Weapons. In The Ethics of Information Warfare; Taddeo, M., Floridi, L., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 85–103.

60. Lucas, G.R. Permissible Preventive Cyberwar: Restricting Cyber Conflict to Justified Military Targets. In The Ethics of Information Warfare; Taddeo, M., Floridi, L., Eds.; Springer International Publishing: Cham, Switzerland, 2014; p. 73.

61. Rowe, N.C. The Ethics of Cyberweapons in Warfare. Int. J. Cyberethics 2009, 1, 20–31.

62. Rowe, N.C. Ethics of cyberwar attacks. In Cyber War and Cyber Terrorism; Colarik, A., Janczewski, L., Eds.; The Idea Group: Hershey, PA, USA, 2007.

63. Rowe, N.C. Distinctive Ethical Challenges of Cyberweapons. In Research Handbook on International Law and Cyberspace; Tsagourias, N., Ed.; Edward Elgar Publishing: Cheltenham, UK, 2015; pp. 307–325.

64. Gorman, S.; Barnes, J.E. Cyber Combat: Act of War—Pentagon Sets Stage for U.S. to Respond to Computer Sabotage with Military Force, The Wall Street Journal. 2011. Available online: https://www.wsj.com/articles/SB10001424052702304563104576355623135782718 (accessed on 31 January 2024).

65. Murray, S.F. The Moral and Ethical Implications of Precision-Guided Munitions; Maxwell AFB School of Advanced Air and Space Studies: Montgomery, AL, USA, 2007.

66. Rowe, N.C. Challenges of Civilian Distinction in Cyberwarfare. In Ethics and Policies for Cyber Operations: A NATO Cooperative Cyber Defense Centre of Excellence Initiative; Taddeo, M., Glorioso, L., Eds.; Springer: Berlin/Heidelberg, Germany, 2016.

67. Cobb, S.; Lee, A. Malware is called malicious for a reason: The risks of weaponizing code. In Proceedings of the 2014 6th International Conference On Cyber Conflict (CyCon 2014), Tallinn, Estonia, 3–6 June 2014; pp. 71–84.

68. Nakashima, E.; Timberg, C. NSA officials worried about the day its potent hacking tool would get loose. Then it did. Washington Post, 16 May 2017.

69. Brück, T.; de Groot, O.J.; Bozzoli, C. How Many Bucks in a Bang: On the Estimation of the Economic Costs of Conflict. In The Oxford Handbook of the

Economics of Peace and Conflict; Garfinkel, M.R., Skaperdas, S., Eds.; Oxford University Press: Oxford, UK, 2012.

70. Stiglitz, J.E.; Bilmes, L.J. Estimating the Costs of War: Methodological Issues, with Applications to Iraq and Afghanistan. In The Oxford Handbook of the Economics of Peace and Conflict; Garfinkel, M.R., Skaperdas, S., Eds.; Oxford University Press: Oxford, UK, 2012.

71. Levine, C. Conceptualizing Financial Loses as a Result of Advanced Persistent Threats. Bachelor's Thesis, Pace University, New York, NY, USA, 2013.

72. Sigholm, J.; Larsson, E. Cyber Vulnerability Implantation Revisited. In Proceedings of the MILCOM 2021—2021 IEEE Military Communications Conference (MILCOM), San Diego, CA, USA, 29 November–2 December 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 464–469.

73. Sigholm, J.; Larsson, E. Determining the Utility of Cyber Vulnerability Implantation: The Heartbleed Bug as a Cyber Operation. In Proceedings of the 2014 IEEE Military Communications Conference, Baltimore, MD, USA, 6–8 October 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 110–116.

74. Larsson, E.; Sigholm, J. Towards econometric estimation of the cost of cyber conflict. Procedia Comput. Sci. 2024, 246, 2635–2644.

75. Kohler, K. Cyberneutrality: Discouraging Collateral Damage. CSS Policy Perspect. 2022, 10, 1–4.

76. Lilly, B.; Geers, K.; Rattray, G.; Koch, R. Business@War: The IT Companies Helping to Defend Ukraine. In Proceedings of the 2023 15th International Conference on Cyber Conflict: Meeting Reality (CyCon), Tallinn, Estonia, 30 May–2 June 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 71–83.

77. Bogdanowicz, Z.R.; Patel, K. Quick Collateral Damage Estimation Based on Weapons Assigned to Targets. IEEE Trans. Syst. Man. Cybern Syst. 2015, 45, 762–769.

78. Joint Targeting School. Collateral Damage Estimation Qualification Course Syllabus. 2021. Available online: https://www.jcs.mil/Portals/36/Documents/Doctrine/training/jts/col_damage_course_syllabus2021.pdf (accessed on 31 January 2024).

79. Maathuis, C.; Chockalingam, S. Tackling Uncertainty Through Probabilistic Modelling of Proportionality in Military Operations. Eur. Conf. Cyber Warf. Secur. 2023, 22, 276–284.

80. Maathuis, C. Effects Assessment for Targeting Decisions Support in Military Cyber Operations; Delft University of Technology: Delft, The Netherlands, 2020.

81. European Union Military Committee. Avoiding and Minimizing Collateral Damage in EU-Led Military Operations Concept, Brussels. 2016. Available online: https://data.consilium.europa.eu/doc/document/ST-5785-2016-INIT/en/pdf (accessed on 30 January 2024).

82. Raymond, D.; Conti, G.; Cross, T.; Fanelli, R. A control measure framework to limit collateral damage and propagation of cyber weapons. In Proceedings of the 2013 5th International Conference on Cyber Conflict (CYCON 2013), Tallinn, Estonia, 4–7 June 2013; pp. 1–16.

83. Bertoli, G.; Marvel, L. Cyberspace Operations Collateral Damage—Reality or Misconception? Cyber Def. Rev. 2017, 2, 53–62.

84. Smeets, M. The Strategic Promise of Offensive Cyber Operations. Strateg. Stud. Q. 2018, 12, 90–113.

85. Lawson, E.; Mačák, K. Avoiding Civilian Harm from Military Cyber Operations During Armed Conflicts; International Committee of the Red Cross: Geneva, Switzerland, 2021.

86. Rowe, N.C. Towards reversible cyberattacks. In Proceedings of the 9th European Conference on Information Warfare and Security, Thessaloniki, Greece, 1–2 July 2010; Demergis, J., Ed.; Curran Associates, Inc.: Red Hook, NY, USA, 2010.

87. Hu, Y. Efficient, high-quality force-directed graph drawing. Math. J. 2005, 10, 37–71.

88. U.S. Cyber Command. Improving Targeting Support to Cyber Operations. 2016. Available online: https://nsarchive.gwu.edu/sites/default/files/documents/6379795/National-Security-Archive-USCYBERCOM-Joint.pdf (accessed on 1 February 2024).

89. United States Department of Defense. Summary 2023 Cyber Strategy of the Department of Defense; United States Department of Defense: Arlington County, VA, USA, 2023.

90. Soesanto, S. The IT Army of Ukraine: Structure, Tasking, and Ecosystem; Center for Security Studies: Zürich, Switzerland, 2022.

91. Tidy, J. Meet the Hacker Armies on Ukraine's Cyber Front Line, BBC 2023. Available online: https://www.bbc.com/news/technology-65250356 (accessed on 30 January 2024).

92. Lin, P.; Allhoff, F.; Abney, K. Is Warfare the Right Frame for the Cyber Debate? In The Ethics of Information Warfare; Floridi, L., Taddeo, M., Eds.; Springer: Cham, Switzerland, 2014; pp. 39–59.

93. Backman, S. Making Sense of Large-scale Cyber Incidents: International Cybersecurity Beyond Threat-Based Security Perspectives. Doctoral Dissertation, Stockholm University, Stockholm, Sweden, 2023.

94. Ghafur, S.; Kristensen, S.; Honeyford, K.; Martin, G.; Darzi, A.; Aylin, P. A retrospective impact analysis of the WannaCry cyberattack on the NHS. NPJ Digit. Med. 2019, 2, 98.

95. Preis, B.; Susskind, L. Municipal Cybersecurity: More Work Needs to be Done. Urban Aff. Rev. 2022, 58, 614–629.

96. Kniesner, T.J.; Leeth, J.D.; Sullivan, R.S. A new approach to evaluate safety and force protection investments: The value of a statistical life. In Military Cost-Benefit Analysis: Theory and Practice; Melese, F., Richter, A., Solomon, B., Eds.; Routledge: London, UK, 2015; pp. 237–260.

97. Rohlfs, C.; Sullivan, R. The Cost-Effectiveness of Armored Tactical Wheeled Vehicles for Overseas US Army Operations. Def. Peace Econ. 2013, 24, 293–316.

98. Franco, M.F.; Künzler, F.; von der Assen, J.; Feng, C.; Stiller, B. RCVaR: An economic approach to estimate cyberattacks costs using data from industry reports. Comput. Secur. 2024, 139, 103737.

99. Wilson, B.; Goughnour, T.; McKernan, M.; Karode, A.; Tierney, D.; Arena, M.V.; Vermeer, M.J.D.; Perez, H.; Levedahl, A. A Cost Estimating Framework for U.S. Marine Corps Joint Cyber Weapons; RAND Corporation: Santa Monica, CA, USA, 2023.

100. Andersson, K.; Bang, M.; Marcus, C.; Persson, B.; Sturesson, P.; Jensen, E.; Hult, G. Military utility: A proposed concept to support decision-making. Technol. Soc. 2015, 43, 23–32.

101. van Haaster, J. On Cyber: The Utility of Military Cyber Operations During Armed Conflict; University of Amsterdam: Amsterdam, The Netherlands, 2019.

102. Schulze, M. Cyber in War: Assessing the Strategic, Tactical, and Operational Utility of Military Cyber Operations. In Proceedings of the 2020 12th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia, 26–29 May 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 183–197.

103. Danielson, S.; Conway, P.; Vonasch, A. What I don't know can hurt you: Collateral combat damage seems more acceptable when bystander victims are unidentified. PLoS ONE 2024, 19, e0298842.

104. Microsoft. Microsoft Digital Defense Report 2024. 2024. Available online: https://www.microsoft.com/en-us/security/security-insider/intelligence-reports/microsoft-digital-defense-report-2024 (accessed on 29 October 2024).

105. Karagiannopoulos, V.; Reid, I. US Sanctions on Iranian Hackers Highlight Growing Concern About the Islamic Republic's Cyberwarriors, The Conversation. 2024. Available online: https://theconversation.com/us-sanctions-on-iranian-hackers-highlight-growing-concern-about-the-islamic-republics-cyberwarriors-228718 (accessed on 6 November 2024).

106. United Nations General Assembly, United Nations Resolution 75/240. 2020. Available online: https://meetings.unoda.org/open-ended-working-group-on-information-and-communication-technologies-2021 (accessed on 5 November 2024).

107. Nobles, C. The Weaponization of Artificial Intelligence in Cybersecurity: A Systematic Review. Procedia Comput. Sci. 2024, 239, 547–555.

108. Maathuis, C. Human Centered Explainable AI Framework for Military Cyber Operations. In Proceedings of the MILCOM 2023—2023 IEEE Military Communications Conference: Communications Supporting Military Operations in a Contested Environment, Boston, MA, USA, 30 October–3 November 2023; IEEE: Boston, MA, USA, 2023; pp. 260–267.

109. Torre, D.; Abualhaija, S.; Sabetzadeh, M.; Briand, L.; Baetens, K.; Goes, P.; Forastier, S. An AI-assisted Approach for Checking the Completeness of Privacy Policies Against GDPR. In Proceedings of the 2020 IEEE 28th International Requirements Engineering Conference (RE), Zurich, Switzerland, 31 August–4 September 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 136–146.

110. Angelov, P.P.; Soares, E.A.; Jiang, R.; Arnold, N.I.; Atkinson, P.M. Explainable artificial intelligence: An analytical review. WIREs Data Min. Knowl. Discov. 2021, 11, e1424.

111. Ghassemi, M.; Oakden-Rayner, L.; Beam, A.L. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit. Health 2021, 3, 745–750.

112. Valeriano, B. Harvard Dataverse: Dyadic Cyber Incident Dataset v 2.0, 2022. Available online: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/CQOMYV (accessed on 20 January 2024).

113. Cyber Peace Institute. Cyber Attacks in Times of Conflict. Available online: https://cyberconflicts.cyberpeaceinstitute.org/ (accessed on 22 September 2024).

114. Maness, R.C.; Valeriano, B.; Hedgecock, K.; Jensen,

B.M.; Macias, J.M. Codebook for the Dyadic Cyber Incident and Campaign Dataset (DCID) Version 2.0. 2022. Available online: https://a678132e-4067-4ed4-800a-239c80659fd1.filesusr.com/ugd/4b99a4_ca35bdb6bd55443e890d2dab86910b4c.pdf (accessed on 20 January 2024).

115. Rid, T. Cyber War Will Not Take Place. J. Strateg. Stud. 2012, 35, 5–32.