Cybersecurity Metrics

Yavor Valentinov Papazov
Business Park Sofia
Building 11-B, Floor 1
Mladost 4
1766 Sofia
BULGARIA

yavor@esicenter.bg

ABSTRACT

This paper aims to summarize the development in and current state of the field of cybersecurity and IT security metrics, an interdisciplinary domain, sufficiently dynamic and often misunderstood. As such, the paper discusses various topics, related to the usage of cybersecurity metrics in management, covering the wide array of definitions for metrics, the established frameworks for choosing metrics, taxonomies, an overview of a proper metrics program and, finally, an attempted analysis at currently open problems in this domain. Notably, however, we will discuss very few metrics per se, keeping our focus at a higher level - metrics programs, instead of metrics themselves.

1.0 INTRODUCTION

Being identified as the fifth domain of war, cyberspace has steadily and significantly grown in importance, both in our daily lives and in multiple aspects of society. However, this rapid, almost explosive growth has led to less-than-perfect security stance of the devices, networks, applications, operating systems and so on.

Today's governments and enterprises need to build and maintain IT infrastructure that is both flexible to innovations and secured against capable attackers, while the very building blocks of this infrastructure are often heterogeneous and spanning over widely different technical platforms on all levels – from the hardware and OS to the application stack. The task of securing such an environment is becoming increasingly more difficult, both due to the aforementioned diversity and the increasing offensive capabilities of attackers. In order to maintain the security of this environment one must also assess it and compare it to an acceptable standard for security – 'You can't manage what you can't measure' as the overused adage goes.

This article attempts to give a high-level overview of the most important developments in the field of cybersecurity metrics and to clearly define the open questions and directions that require further research.

Section 2 (Section 1 being this introduction) focuses on the definition of metric in the domain – definitions, what can and what can't be a metric, how to properly identify and choose metrics. Several practical frameworks for selection of metrics are also briefly covered. The section concludes with a very brief discussion of classifications of metrics.

Section 3 covers what a 'metrics program' consists of and how it could operate – the life cycle of a metrics program, selecting initial metrics, data collection processes, identifying and selecting meaningful metrics for monitoring, reacting based on metrics feedback loop and pivoting the process itself to improve metrics collection and management.

Section 4 consists of an attempted overview of initiatives in the industry and government sectors in the field of cybersecurity, IT security and computer security metrics.

Section 5 discusses the 'open' challenges, related to cybersecurity metrics programs and their deployment, and prominent attempts at solving them – the issues of automating data collection and metrics calculation, "big data" and its challenges in this context, ensuring proper metrics are being utilized and the issue of determining threshold and target values for the selected metrics and a few other issues.

2.0 METRICS IN CYBERSECURITY

A logical start of any discussion of metrics is the definition of the word 'metric'. The term has uses in many disciplines, in which its definitions vary both in rigor and essence. In mathematics, a metric is a non-negative function that maps a pair of points in a space to a number – their "distance" (or a suitable equivalent in the respective space). This definition is the easiest in that it is the most rigorous – there are four conditions for a function to be a metric (we shall remain firmly uninterested in these axioms) and inadmissible candidates can easily be rooted out. In business practice, a metric is defined much more loosely. The online "Business Dictionary" defines metrics as "Standards of measurement by which efficiency, performance, progress, or quality of a plan, process, or product can be assessed" [53]. However, this definition is harder to unpack, as it is much wider. Ironically enough, the meaning behind the term 'metric' in this context is intuitively much closer to another related mathematical term - 'norm'.

These two definitions obviously differ significantly. In the cybersecurity domain, a definition of a metric is bound to be somewhere in between – preferably, it should have sufficient rigor to permit rooting out non-metrics and allowing a more formal approach to both modelling and assessment, while being sufficiently wide and abstract to include the variety of meanings 'cybersecurity' itself carries. As Andrew Jaquith notes, "anything that quantifies a problem space and results in a value could be considered a metric" [1]. Certain authors also draw a distinction between 'measure' and 'metric' - defining 'measure' as 'concrete, objective attribute', while relaxing those requirements for 'metric', allowing some degree of abstractness and subjectivity in a metric [15]. This level of ambiguity in a definition serves to underscore the degree to which 'cybersecurity metrics' as a field of science is still in its infancy. Furthermore, many sources list and discuss a huge number of metrics, highlighting the issue of selecting the appropriate metrics [20].

Ultimately, most authors in the field agree that any result of a measurement, that can influence management decisions, can be considered a metric. However, it is immediately obvious that many 'metrics' that fit the latter definition fail miserably at bolstering improvement and identifying gaps in capabilities. Therefore, we must define also what makes a metric 'good', or in other words, fit for usage in organizations that actively seek to improve their cybersecurity posture. Ref. [17] gives an example of reasonable requirements for privacy metrics, but some of them are applicable for security metrics in general. In particular, these metrics must be orthogonal to utility and cost metrics (valid for strictly speaking 'security metrics', but not necessarily true for measurements of e.g. security efforts cost efficiency, which may be useful as KPIs), their value domain must be well defined, ordered, and "not too coarse".

It should be noted that so far only the pure mathematical definition rules out qualitative value domains. Indeed, while using quantitative metrics is often regarded as a superior choice, Hayden points out that not all knowledge can be quantified and, in particular, measurements, relating to people, are rather difficult to pin down on a numerical scale [54]. He further warns against attempts to convert a qualitative scale to a quantitative one. As a typical example of such transgression, let us consider the traffic light colors – with their easily recognizable meaning. It is easy to map the values 'Green', 'Yellow' and 'Red' to the numbers 1 to 3 on a 'threat' or 'risk' scale, however the value 2.5, which is inside the numerical interval (and therefore a valid result for a quantitative scale), has no meaning in the original value domain [54]. Whether or not qualitative measurements should be considered 'metrics' and whether they could be 'good metrics' for the purpose of cybersecurity is the subject of a vivid debate [15]. In this article, we will assume a neutral position on the topic.

It is deserving of a mention that while many authors discuss what makes a metric 'good', most of them agree that a metric on its own is not necessarily good or bad, but rather it can be good or bad in the context of an organization. Due to that, we will abstain from listing and discussing particular metrics whenever possible, and will instead focus on the available approaches to selecting the appropriate metrics for an organization. In particular, noting the immaturity and volatility of the field, it is often suggested that at the present, comparing the values of the same metric between different organizations may hold little value [1]. Indeed, we will discuss the issue of (lack of) universal cybersecurity metrics and standards further with regards to the open problems in the field. Interested readers can find a good summary of different authors' requirements for 'good' IT security metrics in Ref. [18].

Having spent a significant amount of effort defining 'metrics' in the context of cybersecurity, we will not waste further energy on the tangential (but still important) issue of defining and drawing a boundary between the terms cybersecurity, information security and computer security. Instead, we postulate that this article deals to some degree with metrics in *all* of these fields. While the differences between these terms are of great importance, widely-accepted, rigorous and unchallenged definitions of these terms are not yet available, as can be evidenced by the availability of contradictory interpretations in literature. Therefore, when using the term 'cybersecurity metrics', we will mean metrics in cybersecurity, information security or computer security (sometimes even information assurance metrics), *without* necessarily implying that cybersecurity as a domain contains the other 3. Such a decision is taken from purely pragmatic concerns – a detailed and complete discussion of the differences is well out of scope for this article.

Let us instead, for now, turn our attention to the available frameworks and paradigms for choosing the appropriate metrics.

2.1 GQM

The GQM (Goal-Question-Metric) paradigm is one of the simplest (at first glance) frameworks for choosing the metrics that drive management decisions. The GQM approach can be summarized as follows [3]:

- Define the Goals of the organization.
- Formulate Questions for each of the goals, whose answers accurately reflect the progress, achieved towards the achievement of those goals.
- Choose Metrics that answer the questions posed.

It is important to note that the relations goal-question and question-metric are one-to-many and that a question may be shared by more than one goal, as well a metric can be needed to answer more than one question [54].

This approach has the obvious upside of being closely aligned with the organization's strategy and vision, thereby removing irrelevant metrics out of the picture. Another immediate strength is the perceived lack of complexity in this approach, which could surely facilitate 'bringing everybody on board' with such a metrics program [54].

Originally developed in order to help foster a good software quality assurance program for NASA during the 70s, the method is mostly associated with Victor R. Basili, who along with G. Caldiera and H. Dieter Rombach 'officially' publishes it in 1994 [3], for the relatively narrow field of software engineering. Despite that, the paradigm has received much attention (and respectively adoption) in many industries, sufficient to bolster the development of an extension, "GQM+Strategies" [4], which attempts to improve the original strengths of the GQM approach, namely the organization-wide introduction of high-level goals that drive the conversation on the topic of metrics. In that spirit, the "GQM+Strategies" paradigm was further brought closer to the general public with the book "Aligning Organizations Through Measurement" [5].

Unsurprisingly, the GQM method has found many supporters and has spurred a lot of research in the field of intersection – software security measurement [6], [7].

Within the (IT) security sector, the method was heavily promoted by Hayden in "IT Security Metrics: A Practical Framework", where he suggests that the approach could be used as the core during the deployment of a metrics program.

It should be noted that due to its origins, GQM is still a project-oriented approach. While this can be a strong side when pressure is high to achieve short-term objectives, it also means the 'bigger picture' – a more holistic approach to cybersecurity metrics, one that goes above projects, may be more elusive. Hayden suggests this effect can be negated by combining metrics in catalogues and even proposes an entirely new paradigm – the Security Process Management (SPM) framework – that is designed to exist on a higher level than GQM and ensure integration of security measurement activities at the strategic level for the organization. We will discuss SPM further later.

2.2 PRAGMATIC

While the GQM paradigm assumes a strongly top-down view of identifying the relevant metrics in an organization, the PRAGMATIC [1] framework takes a much leaner approach. Rather than demand top-down focus for metric selection, the authors instead focus their efforts on solving a more general problem – defining metrics for metrics (metametrics). The result can be summarized as 9 requirements for a metric to be acceptable for the organization:

- **P**redictive – metrics should not only accurately reflect the present state, but also the future; allowing for *preventive* instead of just *corrective* measures.
- **R**elevant – while obvious in hindsight, metrics need to have relevance to the security posture of the organization, i.e. improvement in the metric values should indicate *actual* improvement in the organization.
- **A**ctionable – as noted earlier, a metric is a decision support tool, therefore its value should help decision makers, and conversely, it is logical that decision makers have influence over the metric and the ability to take corrective measures, should it be required. Furthermore, it is suggested that such a metric should not only allow for action, but even compel management to act.
- **G**enuine – this quality relates to a certain degree to objectivity, in the sense that the same metric should have similar values in the same context, even when the measurement is performed by a different individual. The authors, however, quickly note that it is sufficient for a metric to be *reasonably* objective, even suggesting a different moniker for Genuine – 'Good enough'.
- **M**eaningful – particularly in the context of their desired recipient, e.g. is metrics designed for upper-level management should not be overly technical. If the metric fails to convey useful information to the decision maker, it is of very little use.
- **A**ccurate – this relates to the most obvious requirement for any metric – to reflect reality sufficiently accurate. While more accuracy is always better, this may come at a price and measuring beyond the required levels of precision is essentially useless, while still taking up resources.
- **T**imely – a metric that produces indication of security improvement or deterioration too late is of little use. Instead, it is highly preferable to have a rapid feedback loop, allowing for a quick action in response.
- **I**ndependent – understood here mostly as 'resistant to manipulation', that is the metric reflects the current status of security objectives of the organization and cannot be skewed to encourage behavior that improves the metric, without being beneficial to the organization.
- **C**ost – while gathering the data and calculating a metric is virtually never free, the cost of computing the metric should be low enough to make such a process worthwhile in business terms.

It is worth mentioning that the authors admit that while PRAGMATIC works very well as a mnemonic, they may have not enumerated all possible desirable qualities a metric could possess, and encourage readers to try out alternative criteria and adapt the approach on their own [25].

2.3 SMART

The SMART method was originally introduced as a framework to help specify goals and objectives, formally introduced by G. T. Doran in 1981 [12].

Like PRAGMATIC it is a mnemonic, listing desirable qualities. While initially the SMART paradigm was focused on goals and objectives, multiple articles have since discussed how the SMART criteria relate to metrics. The criteria themselves are [11]:

- **S**pecific
- **M**easurable
- **A**ctionable
- **R**elevant
- **T**imely

It should be noted that the SMART paradigm is quite loaded as a mnemonic, since many of the letters have multiple possible meanings – e.g. the M can stand for Measurable, Manageable, Meaningful, the A – for Achievable, Appropriate, Attainable, etc. [10].

The authors of the PRAGMATIC approach take the time to discuss its relationship with SMART, pointing out that even though the two mnemonics share some of their requirements, the PRAGMATIC metametric system is much more concrete and rigid than SMART, especially when taking into account the different meanings of the term. Furthermore, they point out that highly PRAGMATIC metrics are also likely to be SMART metrics, but the converse is not necessarily true. [9]

QuERIES is another framework, focused on quantifying the operational cybersecurity risk, suggested by Hughes and Cybenko in Ref. [22]. Unlike the previously discussed frameworks, QuERIES is focused not on *selecting* metrics, but rather as a method in *calculating* the value of a single metric – a performance indicator for the operational risk, associated with a particular information system.

While we have covered the three most common frameworks for selecting metrics, it should be noted that there is research regarding the role of metrics in relation to other planning and management frameworks, e.g. metrics in the context of the Balanced Scorecard [13], [14].

We will very briefly discuss the classifications and taxonomy of cybersecurity metrics. There are many attributes by which to classify metrics discussed in the literature. As some examples, we can consider immunity versus resilience metrics, vulnerability versus security control metrics [16], top-down versus bottom-up metrics, organisational versus operational versus technical metrics, implementation versus effectiveness/efficiency versus impact metrics, etc. A concise, yet detailed discussion of different classifications of metrics can be found in Ref. [15].

3.0 IMPLEMENTATION OF A METRICS PROGRAM

There are multiple suggestions on workflows for implementing a metrics program in the literature. However, a common element can easily be noticed in virtually all such approaches – the cyclical nature of such a program. The reader should forgive the overused (but still very true) Bruce Schneier quote, “Security is a process, not a product” [23], and as such, it is only logical to assume that monitoring and managing security

must be processes, as well. The suggestions for this process have been increasing in both complexity and robustness during the development of the discipline.

3.1 The NIST SP 800-55 Approach

First discussed in 2003, this approach has been documented in NIST SP 800-55. The ITL (Information Technology Laboratory) at NIST also published a bulletin, where in less than 3 pages, the approach is described in broad strokes [24]. This approach is extended in NIST SP 800-55 Rev. 1 [42] and summarized concisely in Chapter 7 of NIST SP 800-100 [43]. The following section further summarizes NIST SP 800-100.

The NIST-suggested approach consists of two high-level processes – the Metrics Development Process (used to select the initial metric set and selecting an appropriate subset at any given time) and the Metrics Implementation Process.

The Metrics Development Process consists of two core activities:

- "Identifying and defining the current information security program".
- "Developing and selecting specific metrics to measure implementation, efficiency, effectiveness, and the impact of the security controls".

These two activities are not necessarily sequential. The process can be further broken down into seven phases. The first four of them fall within the first activity, while the latter three fall within the second defined activity. Figure 1 shows the relationship between those phases and the two activities.

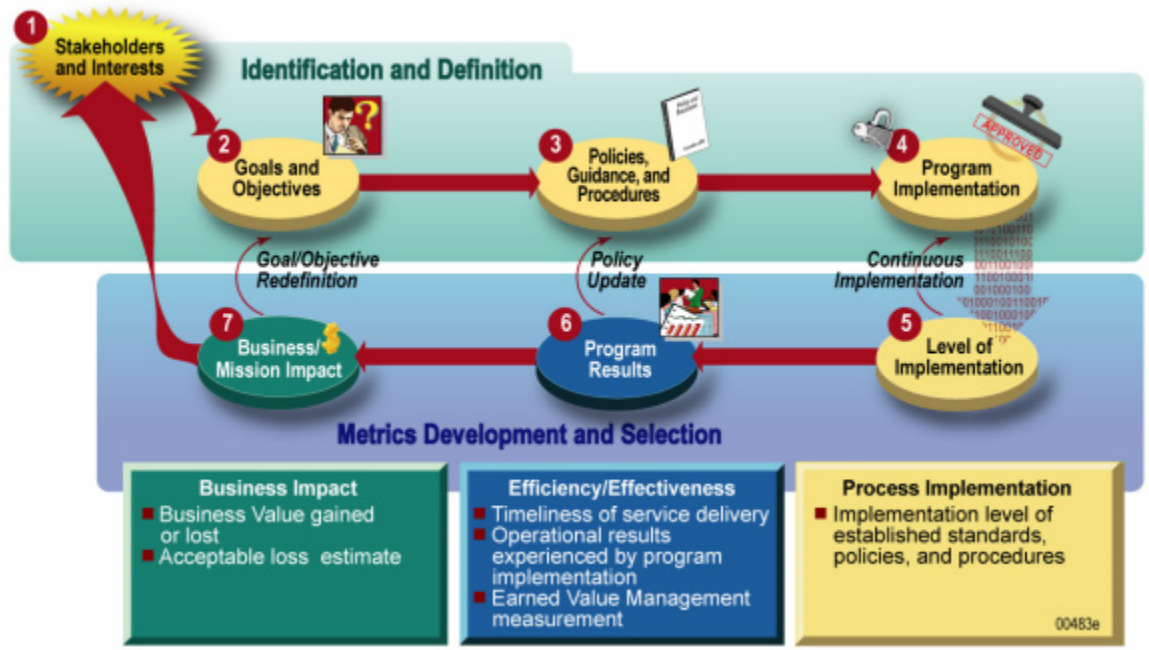

Figure 1: The NIST Information Security Metrics Development Process (Courtesy: NIST).

The other process, the Metrics Implementation Process, consists of six phases (visualized above in Figure 2):

1) Prepare for Data Collection – this phase includes identification, definition, development and selection of the metrics that are to be used.

2) Collect Data and Analyze the Results – in the second phase, the actual data is collected, converted in an appropriate format for analysis, analyzed, and the areas in need of improvement are identified.

3) Identify Corrective Actions – in this phase the appropriate actions to improve security performance are identified and they are prioritized.

4) Develop Business Case – the fourth phase is essentially focused on 'translating' the IT security issues that are identified in business terms. Suggested activities include cost assessment, sensitivity analysis and budgeting activities for the corrective actions proposal.

5) Obtain Resources – at this point it is assumed that resources have been allocated to the corrective efforts and they must now be prioritized and further assigned to tasks and activities.

6) Apply Corrective Actions – the final phase is concerned with the actual implementation of the corrective measures, identified and agreed upon in previous phases. The authors take the time to remind that the process is cyclical in nature and point out that it restarts at this point.

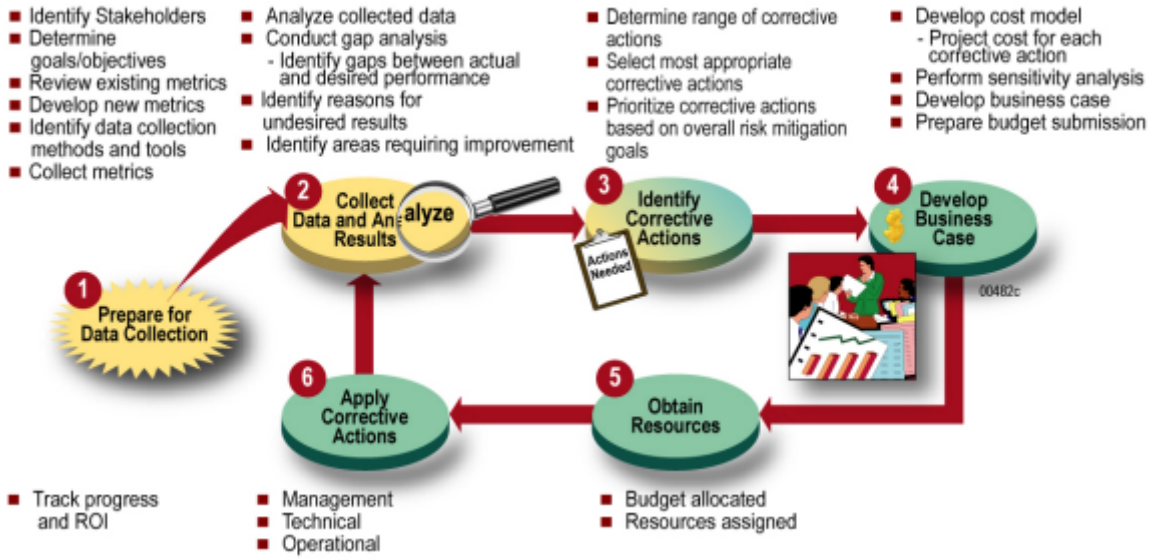

Figure 2: The NIST Information Security Metrics Program Implementation Process (Courtesy: NIST).

3.2 The "Information Security Measurement System Lifecycle" (PRAGMATIC)

Another approach, to structuring an IT security and cybersecurity metrics process, named the "Information Security Measurement System Lifecycle" is proposed in "PRAGMATIC Security Metrics" ([25], Chapter 8). The lifecycle consists of 8 phases:

1) Determine and specify metrics requirements – consists of determining the needs and objectives of the organization with relation to information security. The authors note that, like software development and many other engineering disciplines, this phase is the most crucial towards achieving good results, yet is often the most misunderstood.

2) Prepare a business case – the product of this phase – the business case – should contain strong arguments in favor of the metrics program, i.e. "justifying the investment". It should also document the scope of the program and its expected positive outcomes.

3) Design the measurement system, select security metrics – the purpose of the phase is to select a set of metrics that describe the entire measurement system with sufficient completeness. The suggested workflow by the authors is to select candidate metrics for each important measurement category, then score them by the PRAGMATIC metametrics set and shortlist the winners.

4) Develop the metric(s) – this step mostly considers the necessary steps to transform data into metrics. This includes concerns, such as whether data is collected and analyzed locally or at a central point, whether the metric is calculated at an interval, or is triggered by an event and other important decisions. The authors take care to point out the necessity to automate the calculation of metrics, that cannot be easily computed by hand, or need to be continuously monitored.
5) Test the metric(s) – at this point we have created a set of metrics that we believe can give a sufficiently holistic picture of the cybersecurity posture of the organization. The next logical step is to test that assumption – using e.g. trial runs, prototypes and pilot studies. The authors underscore the importance of assessing the reliability of the set of metrics.
6) Launch the measurement system – this phase is likened to all other change management activities – it is suggested to ensure that deployment does not conflict critical business processes.
7) Use, manage, and maintain the measurement system – at that point, the information security measurement program is (hopefully) running at full speed, and auxiliary activities are the focus – monitoring the process, performing minor tweaks and changes, etc.
8) Mature the measurement system – the final stage of the process; it is suggested that at this point, a 'review process' of sorts is established, albeit at a large interval (authors suggest every 1-2 years). During this process, every step of the process is reconsidered in light of the new circumstances, context and data available, so that the iterative nature of the workflow is preserved.

3.3 SPM

The final approach to designing a metrics program that we will discuss has already been mentioned earlier – the Security Process Management (SPM) framework ([25], Chapter 4).

The SPM framework starts where GQM ends – we have already noted the tendency for project-orientation of the GQM framework. As such, it is important to have a higher-level strategy for integrating the metrics process in the organization at a level *above* projects. That is exactly the issue that SPM is designed to address.

An attempt to summarize the SPM framework follows:

1) Develop metrics – it is suggested that this is done with the GQM framework. The set of these metrics is somewhat project-specific, however later phases take care of this issue.
2) Collect and analyze metrics data in the form of Security Measurement Projects (SMP). As projects, these must have well-defined goals, start and end dates and their results must be stored.
3) Commit to a Security Improvement Program (SIP) that is a layer above SMPs and coordinate SMPs around this program. Do not treat SMPs are individual and isolated projects, cross-reference and examine their results to gain further knowledge and insight into their correlation and their relationship with the organization security needs.
4) Analyze and improve the Security Improvement Program itself, using the gained insight from the previous SMPs and their outcome, issues, 'lessons learned', etc. Do this continually, so that you can manage security as a process, just like any other management aspect.

It is worth noting that unlike the previous discussed approaches, the above list does not consist of steps that are to be taken in that order, but rather documents activities at different levels – from operational to strategic – and as such, these processes coexist at the same time. Also, unlike the other methods, the SPM framework is much more abstract and open to interpretation.

Finally, it is worth a mention that there is a multitude of other approaches to this issue that we will not enumerate and discuss due to the limited scope of the article. An example alternative flow is discussed in Ref. [55].

Other problems that we will not discuss here include metrics dashboarding and visualization.

4.0 GOVERNMENT AND INDUSTRY EFFORTS

This section generally follows the structure of and attempts to summarize the IATAC State-of-the-Art Report (SOAR), titled "Measuring Cyber Security and Information Assurance" [2]. As such, it includes newer information, wherever available, but is also much less detailed and incomplete. The interested reader is urged to consult the original report.

We have divided efforts in 2 broad categories: Regulations (which includes government standards, policies, etc.) and Initiatives – including both government and industry (following a similar classification, present in [2]). Let us now examine each of those groups.

4.1 Regulations

4.1.1 FISMA

The first and foremost government regulation that is related to information security is the Federal Information Security Modernization Act (FISMA) of 2014, which supersedes the previous FISMA (Federal Information Security Management Act of 2002) [26]. The act revises a number of provisions and some general trends can be observed: more responsibility is placed upon the Department of Homeland Security (DHS) and an effort has been made in simplifying existing FISMA reporting. The DHS publishes yearly metrics, further discussed in section 4.2.1.3.

4.1.2 NIST SP 800-55 Rev. 1

The NIST SP 800-55 Rev 1, named "Performance Measurement Guide for Technical Security" was published in 2008 and supersedes NIST SP 800-55. Since we have already discussed the NIST approach to managing a metrics program in 3.1, we will only list it here.

4.1.3 ISO/IEC 27004

The ISO/IEC 27004:2009 "Information technology — Security techniques — Information security management — Measurement" standard was first published in 2009. At the time of writing of this article, the newer revision of the standard (ISO/IEC 27004:2016) is under publication. The standard "provides guidance on the development and use of measures and measurement" in relation with ISO 27001 – Information Security Management [29].

The process consists of four phases:

1) Developing measures.
2) Operating a measurement program.
3) Analyzing and reporting the results.
4) Evaluating and improving the measurement program itself.

4.1.4 ISO/IEC 15408 / Common Criteria

The ISO/IEC 15408 (Common Criteria for Information Technology Security Evaluation) standard consists of three parts – Introduction and general model, Security functional components and Security assurance components. The current revisions of part 2 and 3 were published in 2008, while for part 1 it was published in 2009 [30]. The standard is more commonly known as the 'Common Criteria' Standard. While the standard does mention metrics mostly in Part 2 - "Security functional requirements" [31], since the standard defines 7

Evaluation Assurance Levels (EALs) [32], the assurance level of a software product or system can *itself* be considered a metric – albeit a very coarse one.

4.2 Initiatives

4.2.1 Government Initiatives

4.2.1.1 ICS-CERT

Industrial Control Systems (ICS) and Supervisory control and data acquisition (SCADA) systems have been gaining more and more attention as the information systems which could have the highest impact if compromised, while also often remaining unpatched for years and in some cases being *unpatchable* [39].

In June 2009, the Industrial Control Systems CERT (ICS-CERT) has published the "Primer Control Systems Cyber Security Framework and Technical Metrics". The document establishes the Control Systems Cyber Security Framework (CSCSF) and defines a set of ten metrics and seven dimensions of security for control systems [39].

4.2.1.2 SAMATE

The NIST Software Assurance Metrics And Tool Evaluation (SAMATE) project has its goal at establishing a methodology for assessing software assurance tools. Among the classes of tools that are in scope are Source code analyzers, Web vulnerability scanners and Binary code scanners [40].

4.2.1.3 FISMA Metrics

The DHS publishes on a yearly basis a set of metrics that each government agency is to report to the Office of Management and Budget (OMB). For 2016, three sets of metrics were released, concerning different individuals within agencies – metrics for the Senior Agency Official for Privacy (SAOP), the Inspector General (IG) and the Chief Information Officer (CIO). Currently, only the FY 2017 (Financial Year 2017) CIO metrics have been published. Since 2016, these metrics "are organized around the … Framework for Improving Critical Infrastructure Cybersecurity" [36].

4.2.1.4 NIS WG1

Unlike the rest of the initiatives describes, this one does not originate in the US, but in the EU. The Network and Information Security (NIS) Public-Private Platform was declared as a step forward in the EU Cybersecurity Strategy and the NIS Directive [33]. Within the NIS Platform efforts, three working groups have been created, with working group 1 (WG1) focused on "risk management, including information assurance, risks metrics and awareness raising". So far, the only documents, produced by the NIS Platform with significant relevance to cybersecurity metrics are the "Survey of Risk Management Methods, Frameworks and Capability Maturity Models for the EU Network Information Security Platform" [34] and the State-of-the-Art report on Secure ICT Landscape, including a section, titled "Metrics in cybersecurity" [35].

*4.2.1.5 Metrics and CMM/CMMI **

The Capability Maturity Model (CMM for Software) "provides software organizations with guidance on how to gain control of their processes for developing and maintaining software and how to evolve toward a culture of software engineering and management excellence" [56]. It was developed in the Software Engineering Institute (SEI) at Carnegie-Mellon University (CMU) with assistance from the Mitre Corporation and its first official release was published in 1991. The model went through some revisions and spawned related capability maturity models for different areas and functions – for example Systems

Engineering CMM (SE-CMM), however over time these models grew a certain degree of overlap and even contradicted each other. To solve this problem the CMM Integration (CMMI, originally CMM Improved) model was created, which superseded most other available CMMs and is applicable outside of the original intended domain of CMM – the software development process [50].

The success of the SW-CMM and the need for a similarly structured model for "security services, products or operations" lead to the development of the Systems Security Engineering CMM (SSE-CMM) by the International Systems Security Engineering Association [51]. It defines multiple process areas and five maturity levels, similar to the original CMM:

- Level 1 – Performed Informally.
- Level 2 – Planned and Tracked.
- Level 3 – Well Defined.
- Level 4 – Quantitatively Controlled.
- Level 5 – Continuously Improving.

The process was used as the base for as an ISO standard – ISO 21827:2002 (superseded by ISO 21827:2008) [52].

It is worth noting that while the *request* for these initiatives was from government, they were mostly industry or academia-driven projects – for example, CMM and CMMI were requested by DoD, but implemented by SEI, which is why an asterisk has been included in the title, indicating a deviation from the convention used elsewhere in this paper.

4.2.2 Industry Initiatives

4.2.2.1 CIS

The Center for Internet Security (CIS) is a not-for-profit organization, "dedicated to enhancing the cybersecurity readiness and response among public and private sector entities" [57]. Among its numerous activities in the domain, CIS has released a list of 28 metrics for 7 business functions in a document, titled "The CIS Security Metrics" [28]. Those metrics are very strictly defined and allow for very little room in interpretation with the purpose of allowing cross-organizational sharing and benchmarking. The current release of the CIS Security Metrics is version 1.1.0, published at November 1st, 2010.

4.2.2.2 OWASP

The Open Web Application Security Project is a worldwide not-for-profit organization, focused on improving web application security. Among its most famous projects are the OWASP Top 10 and the Application Security Verification Standard (ASVS). While these projects' contribution to the advancement of information and cyber security is undeniable, the OWASP project that is most related to metrics, the "Application Security Metrics Project" is inactive and has failed to produce even a first release [27].

4.2.2.3 BSA EU Cybersecurity Dashboard

The Software Alliance (BSA) has developed a 'dashboard', describing the current status of cybersecurity legislation, capabilities, education efforts and public-private partnerships of each EU member [37]. It is worth noting that most of the metrics considered are binary in nature or concern the year in which a certain entity was created in the respective country.

4.2.2.4 SIRA NIST CSF Metrics

The Society of Information Risk Analysts has started an open, CC BY-SA 3.0-licensed effort to design metrics, that accurately measure the "Functions, Categories, and where possible Sub-Categories" of the NIST Cybersecurity Framework [38].

5.0 OPEN CHALLENGES

5.1 Information Sharing and Universal Metrics

In the beginning of the article, we pointed out that, at present, many authors suggest an approach to metrics that is strongly focused on the context of the organization. While this does provide many benefits (like ensuring the relevance of the metrics collected), it may also represent a barrier towards the adoption of universal cybersecurity metrics.

The widely accepted answer to this issue is cross-organizational information sharing, which is expected to improve performance in security in many ways, including faster sharing of new attack vectors, vulnerabilities, and others between interested parties.

In the USA, a potential solution to this problem has been proposed in the form of the Cybersecurity Information Sharing Act (CISA). Among other things, the CISA provides a mechanism for private entities to share cybersecurity-related information between each other and with the federal government [58] and blanket liability protection when sharing that information [44], [45].

It is worth noting that the CISA was met with serious criticism from certain advocates, particularly from civil liberties groups [61]. The criticism was mainly focused on the act's perceived negligence towards personal privacy and the perceived possibility of militarization of the Internet as a result of overly broad definitions of 'cybersecurity threat'. Another point of criticism is the perceived inefficiencies of current information sharing initiatives and the potential for failure within the government itself [59].

A good overview of current information sharing approaches in EU can be found in Ref. [60].

An entirely different approach has been suggested by Lie and Satyanarayanan in Ref. [62], where they suggest open challenges as an additional mechanism in assessing security. While they note that some issues have been observed consistently with this approach, they suggest that this is due to such 'challenges' being operated by security vendors with an implicit interest in the failure of the challengers and are often operated under unrealistic, arbitrary or too narrow rules. Instead, it is suggested that there are three main requirements for open challenges:

- Fairness – allowing comparisons between products.
- Sustainability – ensuring sufficient financial interest for all parties included to create a healthy self-sustaining ecosystem.
- Flexibility – sufficient to allow adoption in both academic and business context.

They proceed to define an evaluation framework that answers these requirements [62].

5.2 Automation

We have noted the importance of having metrics that can be calculated cheaply. The easiest way to achieve this by far is to automate the process. Indeed, having an automatic process can bring great value to the program – it allows for, among other things, unscheduled calculations of the metric, reduce workload for employees, and, in general, allow for near-continuous metrics feedback loop.

However, that utility comes at a price. The heterogeneous devices that make up modern enterprise and government networks increase the complexity of an automatic solution. The collected data, used for metric calculation, often comes in wildly different formats, is in need of normalization, validation and in general is subject to many of the issues faced in data mining. Currently, some cross-platform and cross-vendor formats have been developed and widely deployed, like the CVSS vulnerability metric, SCAP, OVAL and others, however many other barriers to further interoperability remain [47].

5.3 Metrics in the Context of 'Big Data'

In the last paragraph, we noted that some of the challenges faced in metrics are similar to challenges in the 'big data' and data mining fields [48]. Indeed, the similarities go further: quoting Andrew Jaquith – "if I can collect enough data on enough things, perhaps I can put everything in a blender and tease out the relationships via correlation" [1]. The 'blender' in this scenario has striking similarities to the abstract (maybe even 'buzzword') usage of the term 'big data'.

While many authors warn that gathering a huge amount of metrics data and stacking it for its own sake does not improve the metrics program in the slightest, but only serves to decrease its efficiency, it is worth noting that in some cases aggregating huge sets of data is impossible to avoid. As an example, for a large enterprise, it is perfectly possible to have thousands of PCs, stationary phones and mobile devices, hundreds of servers and applications. Calculating even a standard, off-the-shelf metric like "Percent of Systems Without Known Severe Vulnerabilities" (defined in CIS Security metrics, [28]) on a regular basis and having basic historical querying capabilities could result in hundreds of gigabytes of data. As for more technically challenging metrics that require, for example, parsing logs, usage of 'big data' tools and approaches may be the only possible way.

In general, 'big data' techniques are 'hyped' as the next-generation tools in cybersecurity [49] and such a trend cannot pass metrics by. It remains to be seen what effect will the 'transformative' capacity of 'big data' have on cybersecurity metrics programs.

5.4 'Responsibility to Act' as Deterrent Factor

One of the not-so-often mentioned implications of a metrics program is that once such a program is in place, the management of the organization can no longer claim ignorance as a defence. While this may seem like a positive aspect (and it often is), it can have an observable effect when taking the decision whether or not to implement such a program.

As an example, in many organizations at lower maturity levels, management may be unofficially aware of the existence of severe security issues, however no official records can be produced, proving that awareness. At this point, management will often react negatively to a security metrics initiative, as the results will prove beyond reasonable doubt that the organization is in dire need of improvement in the domain of cybersecurity. Since starting a cybersecurity metrics program will generate data and therefore trigger the 'responsibility to act', management would instead silently deflect any attempt to introduce metrics in the organization.

This issue has been discussed by multiple authors, for example in "PRAGMATIC Security Metrics" [25], the authors discuss it under the name "Implausible Deniability". A warning in Ref. [54] addresses a similar concern – collected data that is not acted upon is not an asset, but instead a liability when subjected to e-discovery.

While this issue is mostly present in cases of malign management and is therefore not an inherent issue of cybersecurity metrics, an attempt at addressing the issue has been made through the Cybersecurity Information Sharing Act (CISA) [44], where protection from liability is provided to private entities for the purpose of monitoring of information systems. While the CISA failed to pass, the Consolidated

Appropriations Act, 2016 was passed as law and introduced the liability protection clause [46]. However, most other countries have failed to implement similar measures.

5.5 Metrics Gaps and 'Black Swan' Events

This issue is somewhat intrinsic to any measurement system – the more mature any metric program is, the easier it is to assume the program is 'good enough' and to miss issues in the metric collection program *itself*. Brotby and Hinson warn about this in Ref. [25], calling the issue "Metrics Gaps".

While this can be counteracted by a careful and complete review process of the metrics collection program, even a well-designed and implemented process fails to address the issue of the 'Black swan' events.

A 'black swan' event is an extremely unlikely event with huge impact, that can easily be predicted retrospectively, but not prospectively [63]. Informally, 'black swan' events are in the 'unknown unknowns' category and as such, the expectation is that they cannot be predicted or mitigated in advance.

In the context of metrics, the issue of 'black swan' events is not that they cannot be predicted – that is true by definition. The problem is, rather, that a well-running metrics program, giving consistently good results in terms of metric values, may create the illusion of perfect security and therefore convince management that the 'unknown unknowns' do not exist and require no resources to identify and manage.

While some research in this direction is already present [64], it remains to be seen whether and how a systematic approach can systematically identify gaps in capabilities, that are yet to be exploited by attackers and are generally not part of the current threat model.

6.0 REFERENCES

[1] Jaquith, A., *Security Metrics: Replacing Fear, Uncertainty, and Doubt*, 1st Edition, Addison Wesley, 2007, ISBN-13: 978-0321349989.

[2] *Measuring Cyber Security and Information Assurance*, State-Of-The-Art Report, Information Assurance Technology Analysis Center (IATAC), May 8, 2009.

[3] Basili, V., Caldiera, G., Rombach, H., *Goal Question Metric Paradigm*, Encyclopedia of Software Engineering – 2 Volume Set, 1994, ISBN 9780471028956.

[4] Basili, V., et al., *GQM+ Strategies – Aligning Business Strategies with Software Measurement*, Proceedings of the DASMA Software Metric Congress (MetriKon 2007): Magdeburger Schriften zum Empirischen Software Engineering, pages 253-266, Kaiserslautern, Germany, November 15-16 2007.

[5] Basili, V., et al., *Aligning Organizations Through Measurement*, Springer International Publishing, 2014, (e-)ISBN 978-3-319-05047-8.

[6] Islam, S., Falcarin, P., *Measuring Security Requirements for Software Security*, Cybernetic Intelligent Systems (CIS), 2011 IEEE 10th International Conference on, September, 1-2, 2011, DOI: 10.1109/CIS.2011.6169137, (e-)ISBN 978-1-4673-0688-1.

[7] Abdulrazeg, A., et al., *Security Measurement Based On GQM To Improve Application Security During Requirements Stage*, International Journal of Cyber-Security and Digital Forensics (IJCSDF) 1(3): 211-220, 2012, ISSN: 2305-0012.

[8] Jonsson, E., Pirzadeh, L., *A Framework for Security Metrics Based on Operational System Attributes*, SECUREWARE 2013, 2013.

[9] Brotby, W., Hinson, G., *Shouldn't metrics be SMART?*, accessed on October 9, 2016, http://www.securitymetametrics.com/html/faq.html#SMART.

[10] Haustein, J., *Successful Metrics*, June 25, 2012, accessed on October 9, 2016, https://confluence.cornell.edu/display/metrics/Successful+Metrics.

[11] *Using SMART metrics to Drive Action*, Tenable Security, Whitepaper, downloaded from https://www.tenable.com/whitepapers/using-smart-security-metrics-to-drive-action.

[12] Doran, G. T., *There's a S.M.A.R.T. way to write management's goals and objectives*, 1981, Management Review. AMA FORUM. 70 (11): 35–36.

[13] *Security Metrics and the Balanced Scorecard*, Intel Security, October 12, 2011, accessed on October 9, 2016, https://blogs.mcafee.com/business/security-connected/security-metrics-and-the-balanced-scorecard/.

[14] Volchkov, A., *How to Measure Security From a Governance Perspective*, ISACA Journal Volume 5, 2013, accessed on October 9, 2016, http://www.isaca.org/JOURNAL/ARCHIVES/2013/VOLUME-5/Pages/How-to-Measure-Security-From-a-Governance-Perspective.aspx.

[15] *Resilience Metrics and Measurements: Technical Report*, ENISA, February 1, 2011, accessed on October 3, 2016, downloaded from https://www.enisa.europa.eu/publications/metrics-tech-report.

[16] Patriciu, V., Priescu, I., Nicolaescu, S., *Security Metrics for Enterprise Information Systems*, Journal of Applied Quantitative Methods, Volume 1, Issue 2, 2006, ISSN 18424562, accessed on October 1, 2016, http://www.jaqm.ro/issues/volume-1,issue-2/pdfs/patriciu_priescu_nicolaescu.pdf.

[17] Wagner, I., Eckhoff, D., *Technical Privacy Metrics: a Systematic Survey*, accessed on October 6, 2016, https://arxiv.org/pdf/1512.00327.pdf.

[18] Yasasin, E., Schryen, G., *Requirements for IT Security Metrics – An Argumentation Theory Based Approach*, 2015, ECIS 2015 Completed Research Papers. Paper 208.

[19] Pfleeger, S., *Useful Cybersecurity Metrics*, IT Professional, July/August 2009, pp. 38-45.

[20] Pfleeger, S., Cunningham, R., *Why Measuring Security Is Hard*, IEEE Security & Privacy, Volume 8, Issue 4, July/August 2010, pp. 46-54, ISSN: 1540-7993.

[21] Black, P., Scarfone, K., Souppaya, M., *Cyber Security Metrics And Measures*, Handbook of Science and Technology for Homeland Security, Vol 5, 2008.

[22] Hughes, J., Cybenko, G., *Quantitative Metrics and Risk Assessment: The Three Tenets Model of Cybersecurity*, Technology Innovation Management Review, August 2013, ISSN: 1927-0321, accessed on October 6, 2016, http://timreview.ca/article/712.

[23] Schneier, B., *The Process of Security*, Information Security, April 2000, accessed on October 9, 2016, https://www.schneier.com/essays/archives/2000/04/the_process_of_secur.html.

[24] Lennon, E., *IT Security Metrics*, ITL Bulletin, NIST, August 2003, accessed on October 6, 2016, http://csrc.nist.gov/publications/nistbul/bulletin08-03.pdf.

[25] Brotby, W., Hinson, G., *PRAGMATIC Security Metrics Applying Metametrics to Information Security*, Auerbach/CRC Press, 2013, ISBN: 978-1439881521.

[26] *Federal Information Security Modernization Act of 2014*, Public Law No: 113-283 (12/18/2014), 113th Congress of the United States of America, accessed on October 9, 2016, https://www.congress.gov/bill/113th-congress/senate-bill/2521/text.

[27] *OWASP Application Security Metrics Project*, OWASP Foundation, August 22, 2006, accessed on October 6, 2016, https://www.owasp.org/index.php/Category:OWASP_Application_Security_Metrics_Project.

[28] *The CIS Security Metrics*, The Center for Internet Security (CIS), November 1st, 2010, accessed on October 9, 2016, https://benchmarks.cisecurity.org/tools2/metrics/CIS_Security_Metrics_v1.1.0.pdf.

[29] *Information technology -- Security techniques -- Information security management -- Monitoring, measurement, analysis and evaluation*, ISO/IEC 27004:2016, ISO/IEC JTC 1/SC 27, accessed on October 9, 2016, status information at http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=64120.

[30] *Information technology -- Security techniques -- Evaluation criteria for IT security -- Part 1: Introduction and general model*, ISO/IEC 15408-1:2009, ISO/IEC JTC 1/SC 27, accessed on October 9, 2016, downloaded from http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html.

[31] *Information technology -- Security techniques -- Evaluation criteria for IT security -- Part 2: Security functional components*, ISO/IEC 15408-2:2009, ISO/IEC JTC 1/SC 27, accessed on October 9, 2016, downloaded from http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html.

[32] *Information technology -- Security techniques -- Evaluation criteria for IT security -- Part 3: Security assurance components*, ISO/IEC 15408-3:2009, ISO/IEC JTC 1/SC 27, accessed on October 9, 2016, downloaded from http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html.

[33] *NIS Platform*, ENISA, accessed on October 6, 2016, available at https://resilience.enisa.europa.eu/nis-platform.

[34] *Survey of Risk Management Methods, Frameworks and Capability Maturity Models for the EU Network Information Security Platform*, NISP WG1, ENISA, accessed on October 6, 2016, https://resilience.enisa.europa.eu/nis-platform/shared-documents/survey-of-risk-management-methods-frameworks-and-capability-maturity-models-for-the-eu-network-information-security-platform.

[35] *State-Of-The-Art of Secure ICT Landscape*, NIS Platform Working Group 3, July 2014, accessed on October 6, 2016, https://resilience.enisa.europa.eu/nis-platform/shared-documents/wg3-documents/state-of-the-art-of-the-secure-ict-landscape/at_download/file.

[36] *FY 2017 CIO FISMA Metrics*, DHS, October 1, 2016, accessed on October 9, 2016, https://www.dhs.gov/sites/default/files/publications/FY%202017%20CIO%20FISMA%20Metrics-%20508%20Compliant.pdf.

[37] *EU Cybersecurity Maturity Dashboard 2015*, BSA | The Software Alliance, accessed on October 9, 2016, http://cybersecurity.bsa.org/.

[38] *SIRA NIST CSF Metrics Project*, Society of Information Risk Analysts (SIRA), accessed on October 9, 2016, http://nistcsf.societyinforisk.org/doku.php.

[39] *Primer Control Systems Cyber Security Framework and Technical Metrics*, June 2009, Control Systems Security Program, DHS, accessed on October 9, 2016, https://ics-cert.us-cert.gov/sites/default/files/documents/Metrics_Primer_7-13-09_FINAL.pdf.

[40] *Introduction to SAMATE*, NIST, 2004, accessed on October 9, 2016, https://samate.nist.gov/index.php/Introduction_to_SAMATE.html.

[41] Collier, Z., et al., *Security Metrics in Industrial Control Systems*, "Cyber Security of Industrial Control Systems, Including SCADA Systems", Springer, NY, 2016, accessed on October 11, 2016, https://arxiv.org/ftp/arxiv/papers/1512/1512.08515.pdf.

[42] Chew, E., et al., *Performance Measurement Guide for Information Security*, NIST Special Publication 800-55 Revision 1, NIST, July 2008, accessed on October 9, 2016, http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-55r1.pdf.

[43] Bowen, P., Hash, J., Wilson, M., *Information Security Handbook: A Guide for Managers*, NIST Special Publication 800-100, NIST, October 2006, accessed on October 9, 2016, http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-100.pdf.

[44] *Cybersecurity Information Sharing Act of 2015*, 144th Congress of the United States of America, Held at the desk, accessed on October 9, 2016, https://www.congress.gov/bill/114th-congress/senate-bill/754/text.

[45] Karp, B., *Federal Guidance on the Cybersecurity Information Sharing Act of 2015*, Harvard Law School Forum on Corporate Governance and Financial Regulation, March 3, 2016, accessed October 9, 2016, https://corpgov.law.harvard.edu/2016/03/03/federal-guidance-on-the-cybersecurity-information-sharing-act-of-2015/.

[46] *Consolidated Appropriations Act, 2016*, Public Law No: 114-113, (12/18/2015), 114th Congress of the United States of America, accessed on October 9, 2016, https://www.congress.gov/bill/114th-congress/house-bill/2029/text.

[47] Kott, A., Arnold, C., *Towards Approaches to Continuous Assessment of Cyber Risk in Security of Computer Networks*, A version of this paper appeared in IEEE Security and Privacy, December 25, 2015, accessed on October 9, 2016, https://arxiv.org/ftp/arxiv/papers/1512/1512.07937.pdf.

[48] Vaarandi, R., Pihelgas, M., *Using Security Logs for Collecting and Reporting Technical Security Metrics*, Military Communications Conference (MILCOM), 2014 IEEE, October 6-8, 2014, DOI: 10.1109/MILCOM.2014.53.

[49] Morgan, S., *Cybersecurity is the killer app for big data analytics*, June 30, 2015, CSO Online, accessed on October 6, 2016, http://www.csoonline.com/article/2942083/big-data-security/cybersecurity-is-the-killer-app-for-big-data-analytics.html.

[50] CMMI Product Team, *CMMI® for Development, Version 1.3*, Software Engineering Institute (SEI), CMU, November 2010, accessed on October 9, 2016, http://resources.sei.cmu.edu/asset_files/TechnicalReport/2010_005_001_15287.pdf.

[51] Kormos, C. et al., *Using Security Metrics to Assess Risk Management Capabilities*, 22nd National Information Systems Security Conference, NIST, NCSC, accessed on October 6, 2016, http://csrc.nist.gov/nissc/1999/proceeding/papers/p29.pdf.

[52] Davis, N., *Secure Software Development Life Cycle Processes*, Software Engineering Institute (SEI), CMU, July 2013, https://resources.sei.cmu.edu/asset_files/WhitePaper/2013_019_001_297287.pdf.

[53] *metrics*, BusinessDictionary.com. WebFinance, Inc., accessed on October 9, 2016, http://www.businessdictionary.com/definition/metrics.html.

[54] Hayden, L., *IT Security Metrics: A Practical Framework for Measuring Security & Protecting Data*, McGraw-Hill Education, 2010, ISBN: 0071713409.

[55] Payne, S., *A Guide to Security Metrics*, SANS Institute, June 19, 2006, accessed on October 9, 2016, https://www.sans.org/reading-room/whitepapers/auditing/guide-security-metrics-55.

[56] Paulk, M., et al., *Capability Maturity Model for Software, Version 1.1*, SEI, CMU, February 1993, accessed on October 9, 2016, https://www.sei.cmu.edu/reports/93tr024.pdf.

[57] *Who We Are*, Center for Internet Security, accessed on October 6, 2016, https://www.cisecurity.org/about/.

[58] *Guidance to Assist Non-Federal Entities to Share Cyber Threat Indicators and Defensive Measures with Federal Entities under the Cybersecurity Information Sharing Act of 2015*, The Department of Homeland Security, The Department of Justice, June 15, 2016, accessed on October 9, 2016, https://www.us-cert.gov/sites/default/files/ais_files/Non-Federal_Entity_Sharing_Guidance_%28Sec%20105%28a%29%29.pdf.

[59] Dourado, E., O'Sullivan, A., *"Information Sharing": No Panacea for American Cybersecurity Challenges*, Mercatus Center, George Mason University, June 22, 2015, accessed on October 9, 2016, https://www.mercatus.org/publication/information-sharing-no-panacea-american-cybersecurity-challenges.

[60] *Cyber Security Information Sharing: An Overview of Regulatory and Non-regulatory Approaches*, ENISA, December 16, 2015, accessed on October 9, 2016 https://www.enisa.europa.eu/publications/cybersecurity-information-sharing.

[61] Tien, L., *A Zombie Bill Comes Back to Life: A Look at The Senate's Cybersecurity Information Sharing Act of 2014*, June 29, 2014, accessed on October 9, 2016, https://www.eff.org/deeplinks/2014/06/zombie-bill-comes-back-look-senates-cybersecurity-information-sharing-act-2014.

[62] Lie, D., Satyanaran, M., *Quantifying the Strength of Security Systems*, Proceeding HOTSEC'07 Proceedings of the 2nd USENIX workshop on Hot topics in security, 2007.

[63] Taleb, N., *The Black Swan: The Impact of the Highly Improbable*, Random House, 2007, ISBN 978-1400063512.

[64] Reagan, JR, Raghavan, A., Thomas, A., *Quantifying risk: What can cyber risk management learn from the financial services industry?*, Deloitte University Press, accessed on October 9, 2016, http://dupress.deloitte.com/dup-us-en/deloitte-review/issue-19/quantifying-risk-lessons-from-financial-services-industry.html.