Cyber Risk Definition and Classification for Financial Risk Management

Filippo Curti, Jeffrey Gerlach, Sophia Kazinnik, Michael Lee and Atanas Mihov*

September 20th, 2021

Abstract: Cyber risk is undeniably one of the most critical emerging risks to the financial industry. However, even though cyber risk is recognized as a significant threat to financial institutions and, more generally, to financial stability, the quantification and analysis of cyber risk has not yet matured to the point where it can be consistently measured and managed against corporate risk appetites. This impedes efforts to effectively measure and manage such risk, diminishing institutions' individual and collective readiness to handle system-level cyber threats. This paper aims to address this gap by providing a preliminary cyber risk definition and classification of cyber risk for risk management purposes. As such, the proposed definition and classification would ensure that adopting institutions are utilizing common language and allowing consistent data collection and sharing. We provide a deeper dive into the reasoning behind the variables we propose to collect and demonstrate how some of the existing cybersecurity events map into our proposed scheme.

Keywords: operational risk; nonfinancial risk; cyber risk; risk measurement

* Filippo Curti, Jeffrey Gerlach, and Sophia Kazinnik are at the Federal Reserve Bank of Richmond. Michael Lee is at the Federal Reserve Bank of New York. Atanas Mihov is at the University of Kansas. Email addresses: filippo.curti@rich.frb.org; jeffrey.gerlach@rich.frb.org; sophia.kazinnik@rich.frb.org; michael.j.lee@ny.frb.org; amihov@ku.edu. We thank Tom Barkin, Steve Bishop, Luke Carrivick, Jill Cetina, Nida Davis, Michelle Gluck, Greg Gupton, Jason Healey, Marco Migueis, Keith Morales, Patricia Mosser, Nika Lazaryan, Hema Parekh, Will Robinson, Stacey Schreft, David Stabenaw, and Todd Waszkelewicz, for helpful comments and suggestions. We thank Sonia Karami, Cooper Killen, Laurel Mazur, and James Schulte for excellent research assistance. The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Richmond, the Federal Reserve Bank of New York, or the Federal Reserve System.

SECTION 1: INTRODUCTION

Cyberattacks are currently on the rise and are becoming more prevalent and sophisticated over time.[1] Cyber incidents pose a major threat to the financial system.[2] In fact, cyberattacks on traditional financial institutions and cryptocurrency exchanges alike are estimated to have resulted in the theft of billions of dollars. The hacks of major financial firms, consumer credit reporting agencies, retailers and government agencies have compromised the personal information of hundreds of millions of individuals. Data breaches of third-party service providers put intellectual property and confidential information of their serviced financial firms at major risk. Ransomware attacks have infected hundreds of thousands of computer systems globally.

A number of factors contribute to cyber risk at financial institutions, including an increasing trend in globalization, the use and early adoption of quickly evolving technology, significant dependencies and interconnections within both the financial system and information technology infrastructure, the growing sophistication of cyber criminals, and the intrinsic nature of financial institutions' business and services.[3] Awareness of the risks associated with cyber incidents has compelled supervisors and regulators across the world to take steps intended to mitigate cyber risk at financial institutions, including enhancing resiliency capabilities and implementing plans for effective response to and recovery from cyberattacks.

Even though cyber risk is recognized as a significant threat to financial stability, the measurement and analysis of cyber risk within the financial sector has not matured to a point where it can be consistently measured and managed against corporate risk appetites or viewed from a system-wide perspective by regulators and supervisors. This impedes efforts to effectively measure and manage such risk, diminishing institutions' individual and collective readiness to handle system-level cyber threats. In order to begin to classify such risk, this paper provides a preliminary cyber risk definition and classification of cyber risk for risk management purposes. As such, the proposed definition and classification would ensure that adopting institutions are utilizing common language and allowing consistent data collection and sharing.[4,5] This work can additionally support the application of

[1] Cybersecurity experts predict that in 2021, there will be a cyberattack every 11 seconds. This is nearly twice what it was in 2019, and four times the rate five years ago.

[2] See, for example, Eisenbach et al. (2020).

[3] See, for example, Healey et al. (2021) and Crosignani et al. (2020).

[4] Data on cyber losses in the financial industry are not being captured in a consistent and comprehensive way. Available data products are largely based on publicly available information. Vendors include CyberDB (https://cyberdb.co/), ORX (https://managingrisktogether.orx.org/), Advisen (https://advisenltd.com/), and Verisk (https://verisk.com/).

[5] The Federal Reserve System or other regulatory agencies might be particularly well-positioned to facilitate and coordinate data collection efforts due to their secure information technology and data warehouse infrastructures, commitment to information and data confidentiality, and non-profit business orientation. Collected data could be additionally analyzed and used by the Federal Reserve or other regulatory agencies to provide horizontal perspectives on cyber risk management and mitigation for the benefit of participating financial institutions.

modeling frameworks such as Factor Analysis of Information Risk (FAIR) to quantify and measure risk in the financial sector. The cyber loss definition and classification provided in this document, however, are intended to standardize, not necessarily replace, current bank practices. It is also important to note that while our framework incorporates certain elements related to information technology, our main focus is on the financial risk management aspects of cyber risk.

SECTION 2: OBJECTIVE OF CYBER RISK DEFINITION AND CLASSIFICATION

The objective of this paper is to formalize a cyber risk definition and classification scheme in order to support the work of regulatory agencies and private sector participants to facilitate cyber risk management in the financial sector. A cyber risk definition and classification could be useful to support work in the following areas:

Cross-sector shared recognition and identification of relevant cyber risks. A common definition and classification would foster a common understanding of cyber risks and their underlying triggers. In addition, a common set of definitions and shared understanding across the financial sector, including among authorities and private participants, could further facilitate information sharing and appropriate cooperation in cyber risk management.

Assessment and monitoring of financial stability risks. As regulatory and supervisory agencies assess and monitor financial stability risks associated with cyber incidents, this work could be supported by a common definition and classification of cyber risks. For instance, as part of their assessment of vulnerabilities in the US financial system, regulatory and supervisory agencies consider the potential for operational risks, including cyber risks, to result in shocks that could be transmitted across the financial system.

Data collection and information sharing. A definition and classification that supports a common understanding across the financial sector can help advance data collection and information sharing critical to enhancing a collective knowledge of cyber risk by offering a coherent framework for creating and managing data and enabling systematic and compatible aggregation of information.

Regulatory guidance related to cyber risk management. A common classification could enhance the work of regulatory and supervisory agencies in providing guidance related to cybersecurity and cyber resilience, including identifying effective practices and/or emerging threats. For example, utilizing common language could help foster effective regulatory approaches while reducing the risk of duplicative and potentially conflicting regulatory and supervisory requirements.

In general terms, a common definition and classification scheme will facilitate work in the areas outlined above.

SECTION 3: CYBER RISK DEFINITION

Definitions related to cyber risk exist in different contexts. In this paper, we treat cyber risk as a form of operational risk. We therefore build our definition upon one put forth by the Basel Committee on Banking Supervision.[6] Specifically, we define *cyber risk* as the risk of loss resulting from digital incidents caused by internal, external, or third parties, including theft, compromised integrity and/or damage to information and/or technology assets, internal and external fraud, and business disruption. Notably, this definition is largely consistent with known concurrent private sector efforts to define cyber risk. For example, the ORX's Cyber and Information Security Risk initiative defines cyber risk as the risk of loss (both financial and non-financial) arising from digital events caused by external or internal actors or third parties.[7]

To expand our definition of cyber risk further, we build upon the definition derived from the Financial Stability Board's *Cyber Lexicon*.[8] FSB's *Cyber Lexicon*, a widely used list of terms that are relevant to cyber risk in the financial sector, originally takes root in the NIST definition framework.[9] The *Cyber Lexicon* was specifically created to address financial sector cyber resilience and is currently consistent with the majority of industry practices.

We define *cyber event* as an observable occurrence in an information system that a) jeopardizes the cybersecurity of an information system or the information the system processes, stores, or transmits; or b) violates the security policies, security procedures, or acceptable use policies of the information system, whether or not it is a result of malicious activity.[10] We then define a *cyber incident* as a *cyber event* that has resulted in a financial loss.

In essence, both of these definitions build upon the cyber incident definition proposed in FSB's *Cyber Lexicon*.[11] However, we assume our *cyber event* to be based on FSB's cyber incident definition and create an additional partition by introducing the financial loss clause to our *cyber incident*. This is done to refine the mechanism of capturing only the most relevant occurrences related to cyber risk, both in terms of cost and impact.

A single cyber incident may have multiple loss impacts. Figure 1 illustrates this point. For example, a single cyberattack might be associated with the disruption of services at the attacked institution, a data

[6] Operational risk is defined as the risk of loss resulting from inadequate or failed internal processes, people, and systems or from external events.

[7] See Carrivick et al. (2021).

[8] The FSB published the *Cyber Lexicon* in 2018, as a limited scope lexicon that comprises approximately 50 core terms related to cybersecurity and cyber resilience in the financial sector. The goal of this initiative was to develop and propose common definitions of a core set of terms relevant to financial sector participants in both the public and private sectors.

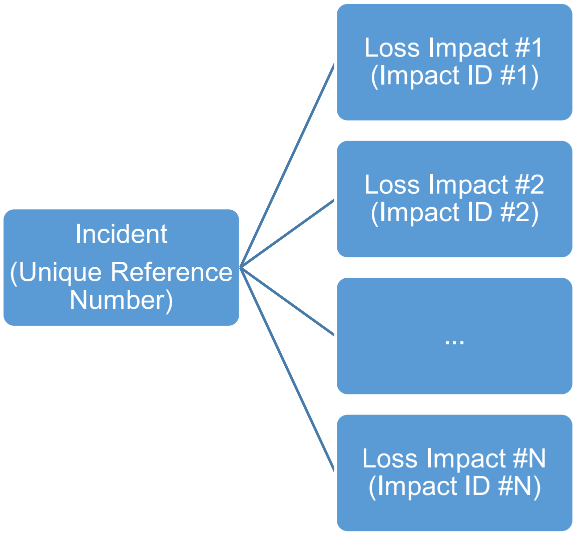
[9] NIST, Glossary of Key Information Security Terms, Revision 2 (May 2013) https://nvlpubs.nist.gov/nistpubs/ir/2013/NIST.IR.7298r2.pdf

[10] A *cyber event* is, by its very nature, a detectable occurrence that breaks through at least two layers of internal controls.

[11] We do not use FSB's cyber event definition, as it would capture too wide of an event set. The definition of FSB's cyber event is as follows. Cyber event is defined as any observable occurrence in an information system. Cyber events sometimes provide indication that a cyber incident is occurring. Source: Adapted from NIST (definition of "Event").

breach, and theft of customer funds.[12] A cyber loss impact is defined as a financial loss (excluding insurance or tax effects) resulting from a cyber incident and includes all expenses associated with a cyber incident, including an indirect cost estimate. Inherent in this definition are elements of legal risk, including privacy protection risk as applicable.[13] This definition excludes strategic and reputational risk.

Figure 1: Cyber Incident Structure



SECTION 4: CYBER LOSS CLASSIFICATION

4.1 Classification Principles

An important underlying principle of creating a cyber loss impact classification scheme is to create categories that aggregate loss impacts that are relatively similar in nature and contain similar drivers in order to facilitate actionable steps from a risk management perspective. Specifically:

a. A cyber loss impact should be uniquely identified to belong to a particular classification category.
b. A particular cyber loss classification category should have impacts with similar underlying drivers.

If a single cyber incident has multiple loss impacts, each loss impact could be plausibly assigned to a different classification category. In cases of incidents with multiple loss impacts, there should be a

[12] In this instance, the cyber loss incident has three separate impacts.

[13] Legal risk includes, but is not limited to, exposure to fines, penalties, or punitive damages resulting from supervisory actions as well as private settlements.

common identifier at the incident level (e.g., a unique reference number) to link these individual records to the same underlying incident. Figure 1 above illustrates this structure with impact identifiers assigned to the incident in chronological order.

4.2 Classification

In line with the above reasoning, we organize our proposed cyber risk classification around six main concepts listed below. The classification scheme is also summarized below in Table 1, Panels A and B.

a. ***Intent***: an indicator for whether the cyber incident was deliberate or accidental.
   - Intentional: when the cyber incident is malicious/intentional.
   - Unintentional: when the cyber incident is not intentional.

b. ***Cyber incident consequence:*** the consequence of a cyber incident.
   - Business Disruption, System and Execution Failure (BDSEF, CN01): Any type of internal or external incident that disrupts the business or causes a software/hardware/IT failure where there was no initial data, technology, or monetary loss.
   - Data Breach - PII (CN02): Any type of data loss or exposure involving Personally Identifiable Information (PII).[14]
   - Theft or Loss of Non-PII Information (CN03): Any type of theft or loss of technology, intellectual property, business proprietary information, or any other information that is not PII.
   - Theft of Funds (CN04): Any type of incident that led to an immediate and direct loss of funds and was carried out via a digital channel.

c. ***Origin***: an indicator for whether the cyber incident originated at the institution or at an external entity.
   - External Party: When the cyber incident initiated at a third party/vendor or any other external entity.
   - Non-External Party: When the cyber incident initiated at the institution or its subsidiary.

d. ***Basel event type category***[15]***:*** the Basel Event category assigned to the cyber incident.

[14] PII here is defined as any information about an individual that can be used to distinguish or trace an individuals' identity and any other information that is linked or linkable to an individual. Source: NIST SP 800-163 under Personally Identifiable Information (NIST SP 800-122).

[15] As previously discussed, cyber risk is considered a form of operational risk. In this regard, the Basel event type categorization is important from a consistency perspective of how to map cyber risk to the broader concept of operational risk. The Basel event type categorization also provides additional granularity to meaningfully differentiate cyber loss events already classified according to other classification concepts.

- Internal Fraud (ET1): Losses due to acts of a type intended to defraud, misappropriate property, circumvent regulations, the law, or company policy, excluding diversity/discrimination events, which involve at least one internal party.
- External Fraud (ET2): Losses due to acts of a type intended to defraud, misappropriate property, or circumvent the law, by a third party.
- Employment Practices, and Workplace Safety (ET3): Losses arising from acts inconsistent with employment, health or safety laws or agreements, from payment of personal injury claims, or from diversity/discrimination events.
- Clients, Products, and Business Practices (ET4): Losses arising from an unintentional or negligent failure to meet a professional obligation to specific clients (including fiduciary and suitability requirements) or from the nature or design of a product.
- Damage to Physical Assets (ET5): Losses arising from loss or damage to physical assets from a natural disaster or other events.
- Business Disruption and System Failures (ET6): Losses arising from disruption of business or system failures.
- Execution, Delivery and Process Management (ET7): Losses from failed transaction processing or process management, from relations with trade counterparties and vendors.

e. ***Cyber incident cause**:* the method through which a malicious cyberattack is carried out.[16]

- Denial-of-Service (CA01): A denial-of-service (DoS) attack floods systems, servers, or networks with traffic to exhaust resources and bandwidth. As a result, the system is unable to fulfill legitimate requests. A distributed denial-of-service (DDoS) occurs when attackers use multiple compromised devices to perform the attack.
- Man-in-the-Middle (CA02): A Man-in-the-middle (MitM) attack, also known as an eavesdropping attack, occurs when attackers insert themselves into a two-party transaction. Once the attackers interrupt the traffic, they can filter and steal data.
- Phishing (CA03): Phishing is the practice of sending fraudulent communications that appear to come from a reputable source, usually through email. The goal is to steal sensitive data like credit card and login information or to install malware on the victim's machine.
- Drive-By Attack (CA04): In a drive-by download attack, hackers look for insecure websites and plant a malicious script into HTTP or PHP code on one of the pages. This script might install malware directly onto the computer of someone who visits the site,

[16] The list itself is loosely based on the MITRE ATT&CK classification, and it expected to be expanded on a continuous basis.

or it might redirect the victim to a site controlled by the hackers. The "Watering Hole" is the most common strategy to execute this type of attack.[17]

- Password Attack (CA05): A password attack happens when an unauthorized party obtains the access to a person's password by looking around the person's desk, "sniffing" the connection to the network to acquire unencrypted passwords, using social engineering, gaining access to a password database, or outright guessing (brute force or dictionary attack)
- SQL Injection (CA06): A Structured Query Language (SQL) injection occurs when an attacker inserts malicious code into a server that uses SQL and forces the server to reveal information it normally would not.
- Cross-Site Scripting (CA07): Cross-site scripting (XSS) attacks use third-party web resources to run scripts in the victim's web browser or scriptable application.
- Birthday Attack (CA08): Birthday attacks are made against hash algorithms that are used to verify the integrity of a message, software, or digital signature.[18]
- Malware (CA09): Software designed with malicious intent containing features or capabilities that can potentially cause harm directly or indirectly to entities or their information systems.
- Zero-Day Exploit (CA10): A zero-day exploit hits after a network vulnerability which is exploited before a patch or solution is developed.
- Other (CA99): Any other type of cyberattack that is not defined. This category would serve as a "catch all" category for cyberattacks with a known type but that are not captured by another existing category.
- Unknown (CA00): When the type of cyberattack is unknown to the institution.

f. ***Asset Exploited:*** the tangible or intangible asset through which the incident was carried out.

- Network (AE01): Incident involving either network, server and/or switches, routers, cables, and other devices in the server room.
- Hardware (AE02): Incident involving hardware, such as PoS, personal computer/laptop, ATM, etc.
- Media/Data (AE03): Incident involving physical documentation containing classified information or either data or data-related vulnerabilities.
- People/Processes (AE04): Incident involving either direct user privileges, assistance from people, or processes/procedures involving people.

[17] A "Watering Hole" attack targets a victim that belongs to a particular group (organization, industry, or region). In this attack, the strategy of an attacker is to guess or observe which websites the group often uses and infects one or more of them with malware.

[18] This is a brute force type of attack where the success of the attack largely depends on the higher likelihood of collisions found between random attack attempts and a fixed degree of permutations, as described in the well-known birthday paradox problem.

- Application/Software (AE05): Incident involving software or application-related vulnerabilities.
- External Provider (AE06): Incident involving cloud or cloud-related assets.
- Other (AE98): Incident involving other assets that do not fit in any of the above categories.
- Not Applicable (AE00): Asset exploited not applicable.

Table 1: Classification Matrix: Panel A – Intentional

	Intentional			
Incident Consequence	*External Party*	*Non-External Party*	**Basel Event-Type Category**	**Incident Cause**
BDSEF	An intentional business disruption at a third-party provider causes disruption to the firm.	An intentional act causes business disruption at the firm.	ET6	CA 1-99
	Human error that leads to an intentional business disruption at a third party /external provider.	An internal human error that leads to an intentional business disruption at the firm.	ET7	CA 1-99
Data Breach - PII	An employee of a third-party provider uses their physical access to steal PII-classified data from the firm.	An employee of the firm uses their physical access to steal PII-classified data from the firm.	ET1	CA 1-99
	An external party gains physical access under the control of a third-party provider to steal PII data from the firm.	An external party gains physical access that enables him to steal PII data directly from the firm.	ET2	CA 1-99
Theft or Loss of Non-PII Information	An employee of a third-party provider steals non-PII data from the firm with remote access.	An employee of the firm steals non-PII data from the firm with remote access.	ET1	CA 1-99

	An external party steals non-PII firm data from a third-party provider with remote access.	An external party steals non-PII firm data from the firm with remote access.	ET2	CA 1-99
Theft of Funds	An employee of a third-party provider uses their access to steal money from the firm or its customers.	An external party defrauds a third party resulting in monetary loss to the firm or the firm's customers.	ET1	CA 1-99
	An employee of the firm uses their access to steal money from the firm or its customers.	An external party defrauds the firm resulting in a monetary loss to the firm or the firm's customers.	ET2	CA 1-99

Table 1: Classification Matrix: Panel B – Unintentional

	Unintentional			
Incident Consequence	*Third Party*	*Non-Third Party*	**Basel Event-Type Category**	**Incident cause**[19]
BDSEF	An unintentional business disruption at a third-party provider causes disruption to the firm.	A software or hardware failure at the firm causes business disruption.	ET6	0 - Not Applicable
Data Breach – PII	A human error allows for unintentional business disruption at a third-party provider, exposing PII data.	A human error allows for unintentional business disruption at the firm, exposing PII data.	ET7	0 - Not Applicable
Theft or Loss of Non-PII Information	A third-party provider loses non-PII firm data as a result of a hardware or software failure.	The firm loses non-PII data as a result of a hardware or software failure.	ET6	0 - Not Applicable

[19] Incident cause is inapplicable here because the table lists unintentional (non-malicious) incidents only.

	A third-party provider loses non-PII firm data as a result of a faulty process or human error.	The firm loses non-PII firm data as a result of a faulty process or human error.	ET7	0 - Not Applicable

We provide more examples in Section 6, illustrating the proposed classification scheme through real-life examples of cyber incidents.

Lastly, it is important to emphasize that the proposed classification scheme is expected to evolve and be periodically updated as new technologies, their applications in banking and finance, and associated cyber threats continue to develop and emerge.

SECTION 5: DATA COLLECTION VARIABLES

In order to gather the appropriate information needed to study cyber-related losses, we propose the data collection proposal we describe in this section. We have settled on the proposed number of variables after receiving extensive feedback from industry participants and consortiums. The decision to have two schedules was made after several rounds of discussions with industry participants.

While the larger banks might be able to afford collecting all cyber incidents, for smaller banks it might be too cost-prohibitive to collect this type of detailed data. For these banks, our proposal would be for them to report using the aggregate level schedule. While we acknowledge that the data collection schedule is not all encompassing, we attempt to take into account the consideration of costs and benefits from potential industry participants.

The two proposed schedules in our data collection schedule are as follows:

- A detailed “loss incident” schedule. This schedule would track cyber risk incidents from which financial losses were realized and would be particularly useful for financial loss modeling. We discuss the variables that compose this schedule in this section.
- An aggregated monthly schedule. This schedule would track both the cyberattacks that resulted in financial losses (incidents) and the ones which did not result in financial losses (events) at a monthly frequency. Such a schedule would be particularly useful for tracking cyber risk trends in addition to financial loss modeling.

Next, we discuss the variables proposed for collection in both the incident and aggregate levels. Tables 2 and 3 provide detailed description of these variables. Our schedule is dynamic in nature, i.e., constructed in such way that we will be able to expand it on a continuous basis. Largely, our list is consistent with the ORX CISR data collection schedule (Carrivick et al. 2020); however, there are several variables that are absent from the ORX data collection schedule and present in ours. These differences stem largely from certain aspects of how both frameworks are structured. In the below list, we provide an explanation for why we include each of the variables in our proposed data collection.

Loss Incident Level Variables:

- *Chronological Order ID*
  For incidents with multiple impacts, this variable represents a cardinal order that reflects the chronology of the different impacts. Capturing this variable would allow researchers to identify which attack type (if there were multiple) was first.
- *Occurrence Date*
  This variable captures the date that the cyber loss incident occurred or began. Capturing this variable would allow researchers to investigate the determinants of cyber losses through a time-series analysis.
- *Discovery Date*
  This variable captures the date that the cyber loss incident was first discovered by the institution. The loss incident's discovery date should not be earlier than its occurrence date. This variable would allow researchers to estimate how long each event went undiscovered – time to discover - and gain a deeper understanding into which cyber losses go undiscovered for longer.
- *Remediation Date*
  This variable captures the date that the cyber loss incident was fully remediated by the institution. The loss incident's remediation date should not be earlier than its occurrence and discovery dates. This variable would allow researchers to estimate the amount of time it takes to remediate – time to remediation - each type of cyber loss.
- *Accounting Date*
  This variable captures the date that the financial impact including the remediation cost of the cyber loss incident was first recorded on the institution's financial statements. The accounting date should be consistent with, and no later than, the date a legal reserve is established. Generally, the loss incident's accounting date should not be earlier than its occurrence date or discovery date; however, there are cases where accounting date can accurately be reflected prior to discovery date. Capturing this variable would allow researchers to better understand the financial impact of the cyber loss incident from an accounting perspective.
- *Gross Loss Amount ($USD)*
  This variable captures the total financial impact of the cyber loss incident before any recoveries as well as excluding insurance and/or tax effects. Capturing this would allow researchers to directly estimate the total financial impact of the cyber loss incident. The Gross Loss Amount would include all expenses associated with a cyber loss incident except for opportunity costs, forgone revenue, provision and provision write backs, and costs related to risk management and control enhancements implemented to prevent future cyber losses. Also, the following type of incidents would not be included in the Gross Loss Amount or the institution's completed schedule:

  - Near Misses: A cyber risk incident that did not result in an actual financial loss or gain to the institution.
  - Timing Incidents: A cyber risk incident that causes a temporary distortion of the institution's financial statements in a particular financial reporting period but that can be fully corrected when later discovered (e.g., revenue overstatement, accounting and mark-to-market errors).
  - Forgone Revenues/ Opportunity Costs: Inability to collect potential future revenues due to cyber risk related failures.
  - Gains: Situations where a cyber risk results in a financial gain for the institution.
- *Remediation Cost ($USD)*

  This variable is set to capture the direct remediation cost of the cyber loss incident before any recoveries and excluding insurance and/or tax effects. Capturing this variable would allow researchers to study which types of cyber loss incidents are more costly than others in terms of remediation costs. The Remediation Cost would be included in the Gross Loss Amount and represents all the expenses the institution bears to fully remediate the cyber incident. The costs related to risk management and control enhancements implemented to prevent future cyber losses would not be included.
- *Indirect Cost ($USD)*

  This variable is set to capture the indirect costs of the cyber loss incident. The Indirect Cost would include expenses related to foregone revenues and/or opportunity costs.[20] This variable is important to capture, even as an approximation, because indirect costs can sometimes be as substantial, if not even more so, in comparison to direct costs.
- *Recovery Amount ($USD)*

  This variable is set to capture the recovery amount following a cyber-related incident. Capturing this variable would allow researchers to capture which cyber loss events are more challenging in terms of the recovery of funds. We define recovery as an independent occurrence related to the cyber loss incident, separate in time, in which funds or outflows of economic benefits are received from a third party, excluding funds received from insurance providers.
- *Insurance Recovery ($USD)*

  This variable captures funds recouped as a result of existing insurance coverage as related to the cyber risk incident. Capturing this variable would allow researchers to learn more about which factors explain how insured amounts are recovered.
- *Cyber Incident Consequence Category*

  All loss incidents reported by the institution would be mapped to one of the four "Cyber Incident Consequence" categories. These categories are described in detail is Section 4.2.a.

[20] Here, we define forgone revenues and opportunity costs as the inability to collect potential future revenues due to failures related to cyber risk.

Capturing this would allow researchers to distinguish between different types of cyber incidents.

- *Asset Exploited*
  This variable captures the category of a tangible or intangible asset through which the incident was carried out.[21] Capturing this would allow researchers to learn more about specific vulnerabilities that allow attacks to take place.
- *Cyber Incident Cause Category*
  All loss incidents reported by the institution would be mapped to one of the twelve "Cyber Incident Cause" categories. These categories are described in detail in Section 4.2.b. Capturing this would allow researchers to learn more about the specific causes of each cyber incident.
- *Intent Indicator (Intentional vs. Unintentional)*
  This variable captures the presence or absence of intent in each incident. Unlike many other cyber data collections, we are including the unintentional incidents. When it comes to cyber risk, there are many events that start out unintentionally, but that lead to severe circumstances.[22] Capturing these events would allow researchers to differentiate between intentional and unintentional incidents and gain a deeper understanding of underlying risk drivers for both of these categories.
- *External Party Indicator*
  This variable is included to capture whether an incident transpired due to the involvement of a third party or an internal actor. By allowing this differentiation, we can learn more about different risk drivers that lead to cyber incidents for both of these categories.
- *Basel Event-Type Category: Level 1*
  All loss events reported by the institution would be mapped to one of the seven "Level 1 Event Types". These categories are described in detail is Section 4.2.e. Capturing this would allow researchers to map the collected cyber events to the existing Basel Operational Risk Management framework.
- *Basel Business Line: Level 1*
  This variable captures the business line involved, as defined by the Basel Operational Risk Management framework. Capturing this would allow researchers to study underlying risk drivers for each of the business lines.
- *Acquired or Merged Entities*
  If the loss incident being reported originated from an acquired or merged entity, then this variable would capture the name of the respective acquired or merged entity.
- *Detailed Description of Loss Incident*
  This variable would allow for the detailed description of the loss incident to be filled out by the bank. Capturing this would allow researchers to derive more details about each incident, if

[21] Specifically, the categories for the asset exploited are: Network, Hardware, Media/Data, People/Processes, Application/Software, External Provider, Other, and Not Applicable. Detailed definitions are provided in Section 4.2.f.

[22] For example, according to the UK Information Commissioner's Office (ICO), human error was the cause of approximately 90% of data breaches in 2019. This is up from 61% and 87% the previous two years.

needed. It will also allow for future enhancement of the data collection framework, as it will provide more insight into which relevant aspects of cyber incidents are not captured by the existing framework.

- *Detailed Description of Remediation Action*
  This variable would allow for a detailed description of the remediation action taken to address the cyber risk incident (including technical details for information technology fixes).
- *Threat Actor*
  Type of threat actor, either entity or person, that caused or contributed to the event. While we recognize the difficulty of capturing this variable, capturing it would provide an immense amount of insight into the incident. Capturing this would allow researchers to differentiate between the types of actors and learn more about underlying risk drivers and other relevant characteristics that contribute to each cyber incident.
- *Primary/Secondary Control Failure*
  These variables capture the codes for primary and secondary controls which were set to prevent the event from occurring. We are proposing to use the NIST framework to capture the failed controls. The reason we use NIST in order to capture failed controls is because most banks utilize NIST as their first choice for controls classification. For example, 74% of institutions with an ORX membership are relying on NIST to capture failed controls. Capturing this would allow researchers to learn more about the main determinants of failure in the prevention of cyber incidents.
- *Event Status*
  An indicator denoting that all necessary information related to the event is known and has been submitted. Capturing this would allow researchers to keep track of current and closed cyber incidents.

For the aggregate schedule, we propose to collect only select variables from the above list that would allow us to capture the picture of cyber loss in the aggregate.[23] As the definitions remain the same, we do not include a separate list within the paper.

SECTION 6: CLASSIFICATION MATRIX CASE STUDY

In this section, we discuss several real-life examples of cyber incidents and demonstrate how they would map this into our proposed data collection schedule.

[23] Specifically, we include the following variables in the aggregate schedule: Reporting Date, Incident Cause, Total Number of Cyber Events, Total Number of Cyber Incidents (Direct or indirect loss > 0), Loss Amounts: Total Gross Losses, Total Recovery Amount, Total Defense Cost.

6.1 Intentional Incidents

Description*: Human error at the third party leads to an attack and PII exposure*

In mid-May of 2017, misconfiguration of the device inspecting encrypted traffic affected certain versions of Apache Struts due to an unpatched open-source flaw. As a consequence, Equifax exposed the personal information of 145.5 million U.S. citizen. The data breach was discovered by Equifax after hackers had access to Equifax's network for 76 days.[24] Under the initial proposed settlement, Equifax agreed to pay up to $700M in order to cover the cost of the data breach including consumers' loss and identity theft services.[25]

Classification: *Intentional/BDSEF/External Party/ET7CA06*

Description*: Password protected computer is stolen from the third party, and PII information becomes compromised*

On February 13, 2020 the theft of an employee laptop from GridWorks IC, a third-party vendor of Health Share of Oregon, exposed the personal and medical information of 654,000 members. The Health Share of Oregon data breach disclosed sensitive data, including names, addresses, phone numbers, dates of birth, Social Security numbers, and Medicaid ID numbers. Following this event, Health Share offered the members involved in this incident one year of identity theft protection, credit monitoring, and fraud consulting at no cost.[26]

Classification: *Intentional/Data Breach – PII/External Party/ET2/CA99*

Description*: An insider infiltrates the firm's computer system and gains access to confidential (non PII) data*

In February 2018, SunTrust Bank disclosed that a former employee had shared information regarding 1.5 million customers with a criminal third party. According to SunTrust, PII was not exposed, and breached data mostly included name, phone number, addresses, and account balances of the 1.5 million

[24] Ng, Alfred. "How the Equifax hack happened, and what still needs to be done." cnet.com, https://www.cnet.com/tech/services-and-software/equifaxs-hack-one-year-later-a-look-back-at-how-it-happened-and-whats-changed/

[25] Schneider, Avie and Chris, Arnold. "Equifax To Pay Up To $700 Million In Data Breach Settlement." npr.org, National Public Radio, 22 July 2019, www.npr.org/2019/07/22/744050565/equifax-to-pay-up-to-700-million-in-data-breach-settlement

[26] "Identity Protection." healthshareoregon.org, Health Share of Oregon, www.healthshareoregon.org/idprotection

customers. In addition, SunTrust has not detected any unexplained or criminal activities linked to the impacted accounts.[27]

Classification*: Intentional/Theft or Loss of Non-PII Information/Non-External Party/ET1/CA99*

Description: *Criminals use malware to gain access to firm accounts and transfer funds.*

In February 2016, in an effort to hack the software that the Bangladesh central bank uses to send Swift massages, a team of hackers successfully transferred $101 million from the bank's account at the Federal Reserve Bank (FRB) of New York to a number of bank accounts in Philippines. In this attack, hackers initially requested the transfer of $951 million to their accounts, out of which $850 million was detected and flagged as suspicious transactions.[28]

Classification: *Intentional/Theft of Funds/External Party/ET2/CA09*

Description: *Firm is hit by DDoS attack that subsequently disrupts service.*

In September 2017, a DDoS attack took place in Google services. This attack lasted for almost six months and was the largest DDoS recorded at the time.[29]

Classification: *Intentional/BDSEF/External Party/ET6/CA01*

Description*: Hacker infiltrates firm's computer system and gains access to the networks, systems and data.*

In March 2020, nation-state hackers compromised a DLL file linked to a software update for the Orion platform by SolarWinds. The supply chain attack impacted up to 18,000 SolarWinds customers including six U.S Government departments such as the Department of the Treasury and private companies such as Microsoft, Intel, and Cisco. In some cases, such as the Department of the Treasury,

[27] Hufford, Austen, and Rexrode, Christina. "SunTrust Employee May Have Stolen Information About 1.5 Million Clients." Wsj.com, Wall Street Journal, 20 April 2018, www.wsj.com/articles/suntrust-employee-may-have-stolen-information-about-1-5-million-clients-1524231553

[28] Al-Mahmood, Syed Zain. "Hackers Lurked in Bangladesh Central Bank's Servers for Weeks." Wsj.com, Wall Street Journal, 22 March 2016, www.wsj.com/articles/hackers-in-bangladesh-bank-account-heist-part-of-larger-breach-1458582678

[29] Smith, Adam. "CHINA LAUNCHED THE BIGGEST DDOS ATTACK IN HISTORY AGAINST GOOGLE, COMPANY CLAIMS." Independent.co.uk, Independent, 19 October 2020, www.independent.co.uk/life-style/gadgets-and-tech/google-ddos-attack-hack-biggest-china-b1155500.html

no classified information was breached. However, some of the SolarWinds clients may never know if their data had been compromised as the attack was elaborately executed.[30]

Classification: *Intentional/Theft or Loss of Non-PII Information/External Party/ET2/CA09*

6.2 Unintentional Incidents

Description*: An external event leads to an unexpected service disruption*

On August 8, 2011, a Microsoft data center in Dublin, Ireland suffered an eleven-hour-long outage due to lightning striking one of the transformers and causing a widespread fire. The outage left many of Microsoft's customers unable to access business critical data and operations for the duration of the outage.[31]

Classification: *Unintentional/BDSEF/External Party/ET/CA0*

Description*: Employee of the firm loses a flash drive/laptop containing PII information*

In June 2018 a judge upheld the decision to fine the University of Texas MD Anderson Cancer Center $4.3 million for HIPAA violations. The cancer center suffered three data breaches between 2012 and 2013, which resulted in the loss of the health information of over 33,500 individuals. In one case an unencrypted laptop was stolen from an employee's residence. The other two breaches involved the loss of unencrypted USB devices.[32]

Classification: *Unintentional/Data Breach – PII/Non-External Party/ET7/CA0*

SECTION 7: CONCLUSION

How large are the losses of U.S. banks due to cyber risk? This question remains open, even today. This paper was, in part, motivated by the sense of urgency that this unanswered question poses. With the recent alarming developments in the realm of cyberspace, having the ability to define, classify, and measure cyber risk for financial institutions is of a pressing nature. We construct this proposed data

[30] Volz, Dustin and Robert McMillan. "Hack Suggests New Scope, Sophistication for Cyberattacks." Wsj.com, https://www.wsj.com/articles/hack-suggests-new-scope-sophistication-for-cyberattacks-11608251360

[31] Nick Gaunt, "Microsoft Data Centre in Dublin Offline Due to Lightning", https://www.bluechip.co.uk, https://www.bluechip.co.uk/blog/microsoft-data-centre-in-dublin-offline-due-to-lightning-affecting-amazons-ec2-platform/

[32] Flahive, Paul. "MD Anderson Cancer Center Fined $4.3 M for Data Breach." Tpr.org, Texas Public Radio, 19 June 2018, www.tpr.org/technology-entrepreneurship/2018-06-19/md-anderson-cancer-center-fined-4-3-m-for-data-breach

collection framework as a vehicle meant to pave its way to greater understanding and monitoring of cyber risk that U.S. banking institutions currently face.

Even though cyber risk is on the rise, the quantification and analysis of cyber risk has not yet matured to the point where it can be consistently measured and managed against corporate risk appetites. This impedes efforts to effectively measure and manage such risk, diminishing institutions' individual and collective readiness to handle system-level cyber threats. This paper sets out to provide a preliminary cyber risk definition and classification of cyber risk for risk management purposes in order to fill this existing gap.

Table 1: Cyber Loss Incident Data Collection Schedule (Detailed, $250,000 Threshold)

Field Reference	Field Name	Description	Format N: Numeric C: Character A: Alphanumeric
A	Unique Identifier	Report the unique identifier for each row of data in the institution's data submission. The unique identifier should remain constant with the specified row of data in subsequent submissions and become a permanent element of the data. The unique identifier should not include any white spaces, tabs, or special characters.	A
B	Reference Number	Report the unique institution-established identifier assigned to each loss incident. The reference number should not include any white spaces, tabs, or special characters.	A
C	Chronological Order ID	For incidents with multiple impacts, please assign a cardinal order that reflects the chronology of the different impacts starting at 1. For incidents with a unique impact, please assign 1.	N
D	Occurrence Date	Report the date that the cyber loss incident occurred or began. The Occurrence Date must be submitted in the following format: MM/DD/YYYY. For example, "January 5, 2011" should be "01/05/2011."	Date MM/DD/YYYY
E	Discovery Date	Report the date that the cyber loss incident was first discovered by the institution. The loss incident's discovery date should not be earlier than its occurrence date. The Discovery Date must be submitted in the following format: MM/DD/YYYY. For example, "January 5, 2011" should be "01/05/2011."	Date MM/DD/YYYY
F	Remediation Date	Report the date that the cyber loss incident was fully remediated by the institution. The loss incident's remediation date should not be earlier than its occurrence and discovery dates. The Remediation Date must be submitted in the following format: MM/DD/YYYY. For example, "January 5, 2011" should be "01/05/2011."	Date MM/DD/YYYY
G	Accounting Date	Report the date that the financial impact including remediation cost of the cyber loss incident was first recorded on the institution's financial statements. The accounting date should be consistent with, and no later than, the date a legal reserve is established. Generally, the loss incident's accounting date should not be earlier than its occurrence date or discovery date; however, there are cases where accounting date can accurately be reflected prior to discovery date. The Accounting Date must be submitted in the	Date MM/DD/YYYY

Field Reference	Field Name	Description	Format N: Numeric C: Character A: Alphanumeric
		following format: MM/DD/YYYY. For example, "January 5, 2011" should be "01/05/2011."	
H	Applicable Loss Data Collection Threshold	Report the institution-established loss data collection threshold that was applicable to the respective business line/ function and in effect at the time the loss incident was captured.	N
I	Gross Loss Amount ($USD)	Report the total financial impact of the cyber loss incident before any recoveries and excluding insurance and/or tax effects. The GLA should include all expenses associated with a cyber loss incident except for opportunity costs, forgone revenue, provision, and provision write backs, and costs related to risk management and control enhancements implemented to prevent future cyber losses. Also, the following type of incidents should **not** be included in the Gross Loss Amount or the institution's completed schedule: *Near Misses:* A cyber risk incident that did not result in an actual financial loss or gain to the institution. *Timing Incidents:* A cyber risk incident that causes a temporary distortion of the institution's financial statements in a particular financial reporting period but that can be fully corrected when later discovered (e.g., revenue overstatement, accounting and mark-to-market errors). *Forgone Revenues/ Opportunity Costs:* Inability to collect potential future revenues due to cyber risk related failures. *Gains*: Situations where a cyber-risk results in a financial gain for the institution. In addition, Gross Loss Amounts: Should be reported in units of one (not thousands), rounded to the nearest unit (for example, a one-million-dollar loss would be reported as 1,000,000). Must be reported in US dollars. Loss amounts recorded in foreign currency should be converted to	N

Field Reference	Field Name	Description	Format N: Numeric C: Character A: Alphanumeric
		US dollars using a foreign exchange rate as of the accounting date associated with the respective loss. Cannot be reported as a negative value, except cases where it represents a decrease in reserves.	
J	Remediation Cost ($USD)	Report the direct remediation cost of the cyber loss incident before any recoveries and excluding insurance and/or tax effects. The Remediation Cost should be included in the Gross Loss Amount and represents all the expenses the institution bears to fully remediate the cyber incident. The costs related to risk management and control enhancements implemented to prevent future cyber losses should not be included.	N
K	Indirect Cost ($USD)	Report the indirect costs of the cyber loss incident. The Indirect Cost should include expenses related foregone revenues, opportunity costs, or other foregone financial activity. *Forgone Revenues/ Opportunity Costs:* Inability to collect potential future revenues due to cyber risk related failures, even if such activity was completed after the cyber related incident ended. *Other forgone financial activity*: Inability to process financial activity other than revenue collection, such as outgoing payments or settlement of securities, even if such activity was completed after the cyber related incident ended.	N
L	Recovery Amount ($USD)	A recovery is an independent occurrence related to the cyber loss incident, separate in time, in which funds or outflows of economic benefits are received from a third party, excluding funds received from insurance providers. Recovery Amounts: • Should not be included in the Gross Loss Amount column or netted into the gross loss amount. • Should exclude provisions and provisions write backs. • Should have the same reference number as the associated loss incident. • Should be reported in units of one (not thousands), rounded to the nearest unit	N

Field Reference	Field Name	Description	Format N: Numeric C: Character A: Alphanumeric
		(for example, a one-million-dollar loss would be reported as 1,000,000). • Should be reported in US dollars. Recoveries recorded in foreign currency amounts should be converted to US dollars using a foreign exchange rate as of the accounting date associate with the respective recovery. • Cannot be reported as a negative value.	
M	Insurance Recovery ($USD)	Report funds recouped as a result of existing insurance coverage as related to the cyber risk incident.	N
N	Cyber Incident Consequence Category	All loss incidents reported by the institution must be mapped to one of the four "Cyber Incident Consequence" categories in Reference Table 3. This field must contain the respective Cyber Incident Consequence code specified in Reference Table 3 (i.e., CN01, CN02, CN03, and CN04). The exact code provided must be used (e.g., "CN01") with no additional characters or spaces added.	A
O	Cyber Incident Cause Category	All loss incidents reported by the institution must be mapped to one of the twelve "Cyber Incident Cause" categories in Reference Table 4. This field must contain the respective Cyber Incident Cause code specified in Reference Table 4 (i.e., CA01, CA02, CA03… CA99). The exact code provided must be used (e.g., "CA01") with no additional characters or spaces added.	A
P	Intent Indicator (Intentional vs. Unintentional)	For all loss incidents originally caused by an intentional act please select 1, otherwise 0.	N
Q	External Party Indicator	For all loss incidents originally caused by an external or third party failure please assign 1, otherwise 0.	N
R	External Party Name	For all loss incidents originally caused by an external or third party failure please indicate name	A
S	External Party Location	For all loss incidents originally caused by an external or third party failure please indicate headquarters city and state	A
T	External Party ID number	For all loss incidents originally caused by an external or third party failure please indicate third party ID number (if available, see attached list)	N

Field Reference	Field Name	Description	Format N: Numeric C: Character A: Alphanumeric
U	Basel Event-Type Category: Level 1	All loss events reported by the institution must be mapped to one of the seven "Level 1 Event Types" in Reference Table E.1.a. This field must contain the respective Level 1 Event-Type code specified in Reference Table E.1.a (i.e., ET1, ET2, ET3… ET7). The exact code provided must be used (e.g., "ET1") with no additional characters or spaces added.	A
V	Basel Business Line Level 1	All loss events reported by the institution must be mapped to one of the nine "Level 1 Business Lines" in Reference Table E.1.b. This field must contain the specific Level 1 Business Line code identified in Reference Table E.1.b (i.e., BL1, BL2, BL3,…,BL9) which corresponds to the Level 1 Business Line.	N
W	Acquired or Merged Entities	If the loss incident being reported originated from an acquired or merged entity, then include the name of the respective acquired or merged entity in this field. If not, then insert "NA" (not applicable). "Incidents originating from acquired or merged entities" refer to loss incidents that have a capture date prior to the acquisition/merger date. This requirement should also apply to loss incidents originating from acquired or merged entities that have capture dates after the acquisition/merger date, if those losses have not yet been integrated into the business lines/functions of the merged entity.	C
X	Detailed Description of Loss Incident (required for incidents > $250k)	For all cyber loss incidents with gross loss amounts greater than or equal to $250 thousand, include a detailed description of the loss incident. Generally, the "short-form" descriptions captured in an institution's internal loss database should suffice.	C
Y	Detailed Description of Remediation Action (required for incidents > $250k)	For all cyber loss incidents with gross loss amounts greater than or equal to $250 thousand include a detailed description of the remediation action taken to address the cyber risk incident (including technical details for information technology fixes). "Short-form" descriptions should suffice.	C
Z	Threat Actor	Type of Threat Actor, either entity or person that caused or contributed to the event.	A
AA	Primary Control Failure	Code for the primary control which was set to prevent the event from occurring (NIST).	A

Field Reference	**Field Name**	**Description**	**Format** **N: Numeric** **C: Character** **A: Alphanumeric**
AB	Secondary Control Failure	Code for the secondary control failure which was set to prevent the event from occurring (NIST).	A
AC	Event Status	An indicator denoting that all necessary information related to the event is known and has been submitted (1 for the open event, 0 for the closed).	N

Table 2: Cyber Event Data Collection Schedule (Aggregate)

Field Reference	Field Name	Description	Format N: Numeric C: Character A: Alphanumeric
A	Reporting Date	Report the last day of the month during which the cyberattacks occurred.	Date MM/DD/YYYY
B	Cyber Incidents Cause	All cyber incidents reported by the institution in this schedule must be mapped to one of twelve "Cyber Incident Cause" in Reference Table 4. This field must contain the respective Cyber Incident Cause code specified in Reference Table 4 (i.e., CA01, CA02, CA03… CA99). The exact code provided must be used (e.g., "CA01") with no additional characters or spaces added.	A
C	Number of Cyberattacks	Report the number of total attacks that the institution has been targeted with during the month of the reporting date	N
D	Number of Successful Cyberattacks	Report the number of successful attacks that the institution has been targeted with during the month prior of the reporting date	N
E	Total Gross Loss Amount	Report the total gross amount lost across all the successful cyberattacks. Should be reported in units of one (not thousands), rounded to the nearest unit (for example, a one-million-dollar loss would be reported as 1,000,000).	N
F	Total Recovery Amount	Report the total recovery amount across all the successful cyberattacks. Should be reported in units of one (not thousands), rounded to the nearest unit (for example, a one-million-dollar loss would be reported as 1,000,000).	N
G	Total Defense Cost	Report the total defense cost amount spent during the quarter preceding the current submitted schedule.	N

References

Carrivick, Luke and Bishop, Steve and Ivell, Thomas and Wong, Valerie and Farha, Ramy, *An Emergent Taxonomy for Operational Risk: Capturing the Wisdom of Crowds* (February 14, 2020). Journal of Operational Risk, Vol. 15, No. 2 (2020)

Crosignani, Matteo and Macchiavelli, Marco and Silva, André F., *Pirates without Borders: The Propagation of Cyberattacks through Firms' Supply Chains* (May 27, 2021). FRB of New York Staff Report No. 937

Eisenbach, Thomas M. and Kovner, Anna and Lee, Michael*, Cyber Risk and the U.S. Financial System: A Pre-Mortem Analysis* (January 1, 2020). FRB of New York Staff Report No. 909, January 2020, Rev. May 2021

Financial Stability Board, "Cyber Lexicon," November 12, 2018. (https://www.fsb.org/wp-content/uploads/P121118-1.pdf)

Healey, Jason & Mosser, Patricia & Rosen, Katheryn & Wortman, Alexander, 2021*. The Ties That Bind: A Framework to Assess the Linkage Between Cyber Risks and Financial Stability*, Journal of Financial Transformation, Capco Institute, vol. 53, pages 94-107.

NIST, '*NIST Special Publication* 800-53 (Rev. 4)', 2013. [Online]. Available: https://web.nvd.nist.gov/view/800-53/Rev4/home.