THE Open GROUP

Calculating Reserves for Cyber Risk

Vetting Cyber Risk Models

A White Paper by:

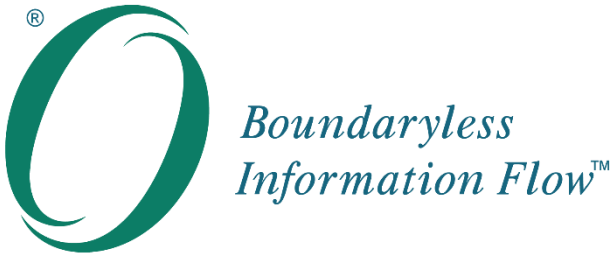
Mike Jerbic
Dr. Robert Mark

July 2022

Calculating Reserves for Cyber Risk: Vetting Cyber Risk Models
Document No.: W221

Published by The Open Group, July 2022.
Any comments relating to the material contained in this document may be submitted to:
The Open Group, Apex Plaza, Forbury Road, Reading, Berkshire, RG1 1AX, United Kingdom
or by email to:
ogpubs@opengroup.org

Table of Contents



Boundaryless Information Flow™
achieved through global interoperability
in a secure, reliable, and timely manner

Executive Summary

Measuring cyber risk in economic terms enables management to treat cyber risk as an enterprise-wide risk. Boards of directors and senior management need a tangible, practical way of performing their fiduciary and regulatory compliance duties in managing this risk. Financial Institutions (FIs) and their regulators must also have sufficient assurance in the accuracy of their operational models to rely upon them as risk estimators on which to base reserve calculations. Making that assurance case – called "model vetting" – for cyber risk models is a significant challenge for any enterprise, not just FIs.

To govern cyber risk effectively and comply with global FI regulatory capital reserve requirements, FIs must not only calculate capital reserves but also need to vet their cyber risk models. Effectively measuring and managing cyber risk is essential to building and operating a globally interoperable, secure, and reliable financial system. In this second White Paper in the *Calculating Reserves for Cyber Risk* series, we present an overview of how cyber risk models can be quantified and vetted. Model vetting requires that models are made transparent, relevant, and parsimonious. We outline here how Chief Risk Officers, Chief Information Officers, and Chief Information Security Officers can work together to meet these model vetting requirements. This supports Boundaryless Information Flow™ through the application of the Open FAIR™ Standards.

Introduction

In the first White Paper of this series, *Calculating Reserves for Cyber Risk: Integrating Cyber Risk with Financial Risk* (see References) we showed how the board, senior management, Chief Risk Officer (CRO), and Chief Information Security Officer (CISO) of a Financial Institution (FI) can work together to analyze, quantify, and communicate cyber risk in economic, Value at Risk (VaR) terms. An FI can treat cyber risk as an enterprise-wide risk and calculate necessary economic and regulatory reserve requirements. Business opportunity costs are made more transparent as reserve requirements become visible. In "freeing up" reserves by reducing risk, cybersecurity teams can escape their historical, stereotypical view as simply being a cost of doing business and instead clearly demonstrate business opportunity value.

Measuring cyber risk in economic terms similar to other financial risks such as credit and market risk enables those risks to be combined into a single integrated risk measure. This integration makes cyber risk fungible with other enterprise risks within the FI, further adding business value through an enhanced ability to avoid, transfer, mitigate, or accept that risk. For example, integrating cyber risk and financial risk enables cyber risk models to be used as a practical, real-world integrated solution of how to treat cyber risk as an operational risk within an FI.

Senior management, and the board and regulators of an FI require that models which impact reserves need to be vetted by the risk management team. In particular, models used to measure cyber risk economically for reserve calculations must undergo a rigorous review and approval process. In other words, management and the board need to have assurance that the model is relevant and accurate for the risk being modeled. By gaining that assurance, an organization can use the model to calculate reserves with the understanding that the "model risk" has been examined, made transparent, and vetted according to a formal process that is acceptable to the regulator.

This document, the second White Paper of the series, introduces core concepts that an FI's risk management team can use to build confidence in Open FAIR™ cyber risk models.

The scope of vetting a cyber risk model includes these main areas of inquiry:

- The transparency of the model so that CROs, management, regulators, and the board of directors can know with high assurance what the model is and does, its limitations, and assumptions
- The logic behind the model to show how relevant and significant it is to model cyber risk scenarios
- The accuracy of the model to show that its results are consistent with the history of the frequency of losses, and their magnitude when they occurred
- The ability of the model to represent accurately the uncertainty of the risk being modeled; for example, VaR models need to capture the uncertainty associated with the "tail" of the distribution, and pass scrutiny in how that tail is captured and that it accurately represents infrequent but extreme events
- How the model can be used in stress tests

All models are to some degree inaccurate representations of the physical world, as they:

- Make assumptions, some of which may not be true

- Rely on the quality of data used – the data used must be reliable and relevant
- Estimate an uncertain future that cannot be perfectly predicted – in other words, all models must accurately estimate outcomes, never predict them
- Represent the highly complex physical world as a simplified, but useful, representation of that world – models are inherently limited by what their designers know about the real world, which makes any model imperfect, but the goal of a model is to make estimates good enough to inform high-quality risk decisions

A model is said to have "parsimony" when the model *"…choice embodies an economy of conceptualization wherein we avoid either being unnecessarily elaborate or being too simplistic*"[1] and is therefore one that not only represents faithfully what can be known through research and scientific analysis but also faithfully represents what is unknowable[2] – such as the uncertainty of the future.

Risk modeling has matured and represents a high degree of complexity in an FI. Modelers in an FI specialize in silos such as credit risk, market risk, and elements of operational risk like Anti-Money Laundering (AML) risk. They apply their specialized knowledge to build complex, comparatively precise risk models that combine theory, history, and judgment of how the future likely differs from the past to estimate future results. Because of the knowledge and experience built over many decades in modeling risk in an FI, a parsimonious risk model is comparatively complex in relation to cyber risk modeling.

For a variety of practical reasons, cyber risk modeling is currently less mature than modeling credit risk and market risk. The theory behind a cyber risk model, similar to many operational risk models, suffers from the high uncertainty of how cyber risk Threat Agents and their tactics, methods of attack, and technology applied to cyberattacks will change. History shows they change very rapidly. There is also less of a historical record of cyberattacks and losses compared to credit and market risk events, and the record that does exist shows significant volatility. The current historical record for cyber loss is believed to be a comparatively poor indicator or predictor for future cyber events. Practices such as back-testing and stress testing that are commonplace when vetting risk models in the financial industry are, at best, emerging practices used to evaluate cyber risk models.

Cyber risk models, when compared to credit risk and market risk models in an FI, reflect a greater uncertainty about future cyberattacks and rely less heavily upon the history of prior successful attacks as a predictor of those future attacks. A risk team's effort to vet cyber risk models calls for examining what they can know and what they cannot know about the frequency and magnitude of future cyber losses. This document outlines some of those considerations.

[1] See *Parsimony – A Model Risk Paper*, by Gary Nan Tie and Dr. Bob Mark, 2020, published by PRMIA Institute; see References, where parsimony is described as "A parsimonious model choice embodies an economy of conceptualization wherein we avoid either being unnecessarily elaborate or being too simplistic. We choose a parsimonious model in terms of only that which is needed to understand our problem and robustly extrapolate. By doing so, we better understand and mitigate model risk. Parsimony is context dependent. What is complicated in one framework may be simpler in another. So, while everything should be as simple as possible and not simpler it should also be context dependent".

[2] What is unknowable today may be knowable tomorrow. Parsimony is a moving target.

How Financial Institution Risk Managers Accept and Vet Risk Models: Model Risk Analysis

Model Vetting as a Means to Evaluate and Ensure Model Quality

Decision-makers rely upon models to estimate future states in an uncertain world given what modelers know about that world. Modelers make assumptions to fill in:

- Missing knowledge[3]
- Uncertainty associated with imperfect information[4]
- Uncertainty due to random chance
- Inherent inability to predict the future in the first place

A useful model improves the decision quality when it is both accurate and sufficiently precise to add value to a decision. Models embody and reflect the uncertainty of the limited knowledge of the modeler and the uncertainty of an unpredictable future, while at the same time reflecting what the modeler knows with a high degree of confidence about the world being modeled.

All models are imperfect since they have limited ranges of utility and are based on assumptions that, if made incorrectly, will lead to errors in estimates and, therefore, decision-makers may over-rely on the accuracy and usefulness of the models.

Representing What's Known, Unknown and Knowable, and Unknowable

At its best, a model represents what is known and unknown (or uncertain) about how the future relates to the present. The concepts of known knowns, known unknowns, unknown knowns, and unknown unknowns (as shown in Figure 1) have been well articulated in risk literature and therefore are useful for applying them to the concept of model quality and the various forms of model error. A perfect model represents everything knowable – that is, everything that can be known about how the world works in predicting the future the model represents. That perfect model also represents the limit of human knowledge, the unknowable uncertainty due to random chance and humanity's limited ability to predict the future precisely and accurately. Models that do not include everything knowable represent a level of ignorance that can be reduced through additional effort, such as research and development of the model. Models that mistake what is unknowable for something known or claimed to be known commit the error of overconfidence or conceit in the modeler's knowledge.

[3] "Missing knowledge" would be anything that an analyst could research that is relevant to an estimate. Historical data that a firm has on its cyber losses over the last two or three years helps to frame practical limits on Loss Magnitude. An analyst who does not research that information is not incorporating all available knowledge into the model.

[4] Analysts use statistical distributions to reflect uncertainty. An analyst may have research showing that losses typically are distributed log-normal. Using that distribution to characterize an estimate of Loss Magnitude expresses the uncertainty to the best knowledge of the analyst.

		A Model's Claim as	
		Known	Unknown
But the Claim Actually is	Knowable	Facts, Verifiable, Objective	Ignorance
	Unknowable (Today)	Conceit, Unverifiable Claims	Uncertainty

Figure 1: Known and Unknown versus Knowable and Unknowable for Certainty

The quality of a model is measured by how much of what is known and knowable is reflected in the model as well as the uncertainty associated with the unknown and unknowable. The level of ignorance and the level of unverifiable claims a model has detract from the quality of a model to the extent that the false estimates it makes of future outcomes will change decisions. All models are flawed as representations of the real world, but models good enough to accomplish their decision-support goals are "fit for purpose".

Models that accurately express everything knowable and the uncertainty of the unknowable exhibit parsimony; in other words, they follow Albert Einstein's adage that "everything should be made as **simple as possible**, **but no simpler**". [emphasis added]

How Models Fail the Parsimony Test

Models have common failure modes in how they fail a test of parsimony.

Models are Too Complex for What is Known – or Knowable

Modelers make errors when they:

- Over-fit a model to the available historical record
- Do not challenge assumptions
- Make too many assumptions that turn out to be wrong
- Believe that they know more about how the world works than they really do or can

In describing attempts to model macroeconomies to plan a nation's output and allocation of resources efficiently and centrally, F. A. Hayek[5] famously said:

"The curious task of economics is to demonstrate to men how little they really know about what they imagine they can design. To the naive mind that can conceive of order only as the product of deliberate arrangement,

[5] The Fatal Conceit: The Errors of Socialism, by F.A Hayek, 1988; see References.

it may seem absurd that in complex conditions order, and adaptation to the unknown, can be achieved more effectively by decentralizing decisions and that a division of authority will actually extend the possibility of overall order. Yet that decentralization actually leads to more information being considered."

What Hayek described as the "Error of Socialism" is an example of the decision-maker's faith in complex economic modeling and the central control of the nation to meet human needs over a simpler system governed by the market price mechanism. Those planners, in constructing a model of collective desires, values, production, and allocation, can fool themselves into thinking they know more than is possible about the efficient production and allocation of scarce resources.

When constructing models of any complex system of human interaction, modelers must remember that simple systems (but not too simple) are frequently more useful and lead to better outcomes than complex attempts at modeling complex systems. Attempting to model complex systems through highly complex models often results in Hayek's "fatal conceit".

Models Use "Too Much" of the Historical Record

Models can be too complex when they cannot separate the "signal" from the "noise" embedded within the historical record. Models that are based upon recreating the historical record perfectly and extending that past into the future will only exactly model the past. Models that do this will not recognize the inherent uncertainty of estimating an uncertain future since they essentially convert the inherent uncertainty of the future into a false certainty. This concept is best explained through an example.

Suppose a modeler has the objective to predict the next roll of two fair, six-sided dice. The modeler uses the historical record of the previous ten rolls and develops a model that exactly replicates that historical record, then uses that model to predict which of the six faces shows up on each die upon the following roll. This model overstates what is knowable about the future of the roll of the dice. For example, the model falsely converts inherent uncertainty associated with a fair dice roll into certainty in predicting a future dice roll.[6]

A parsimonious model would reflect only what is known and what is unknown: There is an equal probability of each of the six faces of the die facing up on the next roll. It would give a probability distribution, not a prediction, of what the next roll will be.

Modelers Use Too Many Assumptions in lieu of Data or Research

As substitutes for information (facts), assumptions are taken to be true, even if they are not. Errors in making assumptions – that is, assumptions demonstrated to be untrue – necessarily make models inaccurate. Modelers and decision-makers may rely upon untrue assumptions due to:

- Insufficient research where assumptions can be verified as true
- Taking something unknowable (necessarily uncertain) and claiming by assumption that it is known within the model

[6] An equivalent problem would be predicting the next occurrence of an adverse financial event that has a probability of 2.78% of occurring within a given timeframe – that is, the same chance as a roll of two dice coming up "snake eyes", where each die shows a score of one.

Regardless of the of the reason, assumptions taken as data indicate an overconfidence in what is known about how the world really works. If this materially affects the accuracy of a model's estimates or predictions, then modelers have used too many assumptions *in lieu* of data that could be researched and therefore the model reflects more certainty and knowledge than can be intellectually justified.

Models are Too Simple

Models may be simpler than what is possible, making those models less complete than they otherwise would be. Models that are too simple may reflect insufficient research or inadequate analytic sophistication applied to their development.

To continue the analogy of modeling a roll of a six-sided die, suppose the modeler assumed the die was fair when it was actually unfair. Further suppose that the unfairness of the die is discoverable or knowable through research. Non-parsimonious modeling of the die as fair suppresses or ignores what could have been learned and represented in a model of that die.

Summary of How Models Can Fail the Parsimony Test

Simple models are not necessarily "bad", and complex models are not necessarily "better".

A high-quality model reflects what is knowable about how the world works and equally reflects what is unknown or unknowable about that world, as shown in Figure 2. A mix of hard work and intellectual humility go a long way toward developing high quality models.

		A Model's Claim	
		Known	Unknown
But the Claim Actually is	Knowable	Parsimony	Underfit, Too Simple, More Research
	Unknowable (Today)	Overfit, Overly Complex, Challenge Assumptions	Parsimony

Figure 2: Known and Unknown versus Knowable and Unknowable for Parsimony

Models are useful to the extent that they inform decision-makers sufficiently to make better decisions. Models should pass a test related to fitness for purpose. A simple model that is "good enough" to inform a relatively low-value decision is appropriate, even if it neither reflects everything knowable, nor everything about what is unknowable. Additional knowledge requires effort, and if that effort is not justified for the decision at hand, modelers should not expend it. In this case, a simple, inexpensive model is preferable to a more complex and expensive one.

On the other hand, models that claim more than they can intellectually justify about how the world works reflect a false confidence in what is known and knowable about the world, misleading decision-makers that the estimated or predicted outcome is more likely to occur than can be justified. Models that claim less than

what can be known offer less precision or accuracy in the model than could be achieved with more effort and investment into the model.

Parsimony is the Goal: Comparing Complexity between Financial Models and Open FAIR Models

Risk management in an FI has developed comparatively complex practical models. Risk professionals in an FI believe they have sufficient evidence to defend their accuracy for the precision in which those models generate and achieve parsimony through a model vetting process. Parsimonious models in an FI are back-tested against historical records. For example, if risk professionals believe that the future will look a lot like the past, then they provide a comparatively high degree of statistical complexity in order to fit to the historical record.

In the cyber world, cyber risk professionals have developed a comparatively simple Open FAIR model that relies more upon Subject Matter Expert (SME) judgment, and less on stable historical data and patterns. Nevertheless, the Open FAIR model is useful for comparative uncertainty in calculating cyber risk.

Introduction to Vetting Models Used in the Financial Sector

We can broadly think of model risk as the kind of mistakes that relate directly to the model itself and are divided into two main groups of errors:

- The model is irrelevant – there is no solid theory to support the cyber risk model or empirical evidence to verify its validity
- The model is incorrect – a model is incorrect if there are mistakes in the set of equations; for example, a model to calculate cyber risk might be based on an erroneous loss distribution assumption

Data scientists and financial engineers are constantly struggling to find the optimal compromise between complexity (to better represent reality) and tractability. An example is the use of:

- Log-normal probability distributions with constant volatility to estimate price risk[7]
- Poisson distributions to calculate cyber event frequency under the assumption that each event is independent and that the interarrival time between events is distributed as an exponential distribution[8]

Models that require the statistical estimation of a number of input parameters that are not directly observable compound the problem. This introduces forecasting errors that increase any model risk.

[7] For example, the Black-Scholes model (refer to: https://en.wikipedia.org/wiki/Black%E2%80%93Scholes_model) makes use of this assumption, even though it is well documented that volatility varies over time and that actual returns in virtually every market exhibit what the statisticians call "fat tails"; i.e., the probability of a large price change is, in reality, greater than the model allows for.

[8] We know that real cases exist which demonstrate, for example, that concentrated attacks for zero-day exploits do not satisfy this assumption.

Model Vetting

Model Vetting Overview

The purpose of model vetting is to offer assurance to the firm's management, the board, and regulators that any operational risk or model such as cyber risk is reasonable and that the proposed implementation faithfully represents the model.

Model vetting includes items such as:[9]

- *Documentation* – risk managers ask for a full documentation of the model

  The documentation must be stated in sufficient detail so that, in principle, the risk manager can independently implement a separate model that produces the same results with the same inputs.

- A mathematical statement of the model – this would include:
  - An explicit statement of all the components of the model
  - The calibration[10] procedure for the model parameters
  - A working version of the implementation
- *Soundness of the model* – the model risk team needs to check that the model is a reasonable representation of reality

  At this stage, the risk manager concentrates on how well the predictive results perform in a back-test and does not become overly focused on the mathematics.

- *Checking results and stress testing the model* – the model risk team compares the results of a benchmark model with those of the proposed model

  Stress testing the model helps to identify the range of parameter values for which the model works to be identified.

- *Checking data* – using dirty data is a key source of model risk

  The model risk team must have access to an independent database to facilitate estimation of model parameters. The data needed to calibrate a model need to be fully documented and their origin fully described.[11] The unintentional forms of data anomalies, traditionally termed "dirty data", are always present.

[9] See *The Essentials of Risk Management, 2nd Edition*, by Michel Crouhy, Dan Galal, and Robert Mark, February 2014; see References.

[10] To Open FAIR practitioners, "calibration" refers to an analyst making calibrated estimates as discussed in Douglas W. Hubbard's book *How to Measure Anything: Finding the Value of "Intangibles";* see References. Calibration here refers to the process of adjusting a model's parameters to best fit historical data. One example of this definition of calibration is the Ordinary Least Squares method of linear regression.

[11] Important data can be discovered through an anomaly detection process for a host of reasons. For example, unsupervised anomaly detection machine learning models are well designed to uncover unknown unknowns.

Calculating Reserves for Cyber Risk: Vetting Cyber Risk Models

Model Risk Vetting Scope / Questions

The scope of model vetting includes comprehensive answers to questions such as how and why the model was selected, how the risk calculated from the model compares to historical records, also known as back-testing,[12] how the model will be monitored when put into use, and how the model is likely to extrapolate (forecast) into the future.[13] Note that the scope of the model risk vetting should be adjusted based on the complexity of the model and how it is used.

Problem Formulation Questions

All questions that are answered must adhere to well-formed criteria. For each question that is answered within the scope of the model vetting process:

- Have you written down the question you are trying to answer?
- Have you identified what constitutes an acceptable answer?
- Have you determined what level of uncertainty is being dealt with?

Model Selection Questions

Risk management will ask questions on how and why a particular model was selected to model a risk. Example questions include:

- How are the model's parameters estimated?
- How has the model been assessed for model fit?
- What alternative parsimonious models could be considered?

Model Validation Questions

Models will be examined for their ability to reproduce that record using parameters appropriate for that historical record and how deviations from that record are explained. Example questions include:

- Have you cross-validated your model and documented how this was done?[14]

[12] Financial risk models such as market risk in the trading book and credit risk for loans are usually extensively back-tested or validated against a substantial historical record to help argue that those models are effective at accurately estimating future risk. Operational risk models typically have less of a record or, in the case of many cyber risk models, may have none at all. For example, suppose a financial institution is vetting a cyber risk model that is quantifying the risk associated with ransomware but has never had a successful ransomware attack. In that case, there is no historical record against which to validate the model. The process to vet cyber risk models will have to be flexible enough adapt to whatever historical record exists. That said, there is an emerging body of data from sources such as the Cyentia Institute to document and organize cyber losses across multiple industry sectors. It is likely that if the historical record is sparse now, it will grow and become more informative over time. The "relevant historical record" is that record that is useful to inform model vetting at the time models are vetted. What that relevant record is today will likely be different tomorrow.

[13] See *Parsimony: A Model Risk Paper*, by Gary Nan Tie and Dr. Bob Mark, 2020; see References.

[14] This step also checks for the reasonableness of the output from a subset of the Monte Carlo simulation calculations. For example, if we make the following purely analytical calculation:
LEF analytical mean x (sum of each PLM analytical mean + SLEF analytical mean x (sum of each analytical SLM mean))
then it should be in the same neighborhood as the overall mean generated from the MCS simulation.

- Have you back-tested the model?
- Can you identify when the model will fail?

Model Extrapolation (Forecasting) Questions

The purpose of a model is to estimate future outcomes. Risk managers in an FI call this extrapolation. Questions on how the model estimates future outcomes include:

- Is the class of models appropriate and fit for purpose?
- Can we combine different model extrapolations to make better (lower variance) forecasts?
- Which extrapolations lead to more stable forecasts, and why?

Model Monitoring Questions

When the model is approved and used, it must be monitored so that it can be continuously improved upon. Model monitoring questions include:

- Is there an unexpected emergent behavior?
- Is the behavior modeled subject to phase transition?
- Have you identified what interventions may be needed when monitoring results?

Model Implementation Questions

Every model should be checked to see whether its operation as implemented differs from its intended design. Material issues to look for include:

- Did a programming bug affect the model's output?
- Has the model been sufficiently tested, and its results verified?[15]
- Was the data input to the model of sufficient quality?[16]

Summary

The purpose of model vetting is to build the argument that the proposed, or vetted, model is of sufficient quality that it is a parsimonious fit for its intended purpose. In other words, model vetting is the quality assurance process for risk models in an FI. The example questions and scope discussed in this section give a starting point of what to expect in vetting a cyber risk model in an FI.

[15] Model vetting will include checks that the Monte Carlo simulation math works correctly. For example, if two independent identically distributed random variables are uniformly distributed, then their sum is triangularly-distributed. Sample data checks from the Monte Carlo simulation should validate that such a relationship holds.

[16] Users of Risk Information can rate (score) the quality of data based on categories such as: Accessibility, Auditability, Completeness, Extensibility, Flexibility, Integration, Integrity, and Timeliness. Refer to: *How Risky is Your Risk Information*, by Robert Mark and Dilip Krishna, Autumn 2008, and to: *Basel Committee on Banking Supervision: Principles for Effective Data Aggregation and Risk Reporting*, January 2013; see References.

Risk Model Vetting

In-house analysts or outside consultants use risk modeling tools to model the cyber risk frequency and magnitude information on behalf of the firm or their client. These analysts get their information either directly from data or indirectly by interviewing SMEs. SMEs also provide Loss Scenarios specific to the firm and perform their own research to estimate risk factors. Analysts use the output from modeling tools to describe the amount of cyber risk and present results.

The quality of these risk models depends upon:

- The quality of the risk scenario developed in the first place: analysts work with their firm's management to define what is most important to model and estimate
- The quality of the assumptions the analysts make: analysts work with their firm's SMEs and management to make assumptions about the Loss Scenario and the areas of the business likely to be impacted should it happen

  Those assumptions and estimates can be debated and challenged for accuracy and reasonableness.
- The quality of the data the analyst has available: analysts are trained to make calibrated estimates based upon incomplete information

  That information can come from SMEs within the organization, from research into the probable frequency and probable magnitude of losses as publicly reported within the same or similar industry, or from privately available information. These estimates are first tested to be accurate, at the possible expense of precision.

Cyber risk analysts through their research and documentation often need to estimate the probability of a loss occurring that has never occurred before, which is a challenge in managing any risk. If the uncertainty is high, then a useful model reflects that uncertainty with a wider range of potential outcomes.

Cyber risk analysts need to prepare themselves to defend their analytic reasoning. All assumptions, estimates, and the rationales behind them must be documented. The information going into a commercially available model needs to be checked for accuracy, and the model's calculations must be confirmed for their correctness. Cyber risk analysts should include in their documentation the logic and soundness of their risk model as adopted by a respected industry standard-setting organization. It is this combined chain of vetting that provides the assurance that their reasoning, input data, risk models, and overall analytic process are transparent. Open FAIR analyses must meet this requirement.

Stress Testing and Scenario Analysis

A "stress test" estimates how much an institution might lose in a specific, exceptional scenario. A scenario may consist of severely adverse changes to external risk factors such as environmental, economic, political, or technological elements external to the FI. Cyber risk is a Basel Pillar 1 operational risk and, therefore, should be stress-tested.

Stress tests are used to give insight into what could happen if conditions lie outside the norm. Stress tests typically posit: "What if several unlikely and extreme external shocks occurred at once?" Any single shock might be considered within normal conditions, but several occurring all at once may inadvertently be viewed as so unlikely that they should be ignored. Stress test scenarios also foresee future consequences that may result from the initial set of shocks.

As stated previously, and to re-emphasize, all models are approximations of the real world and work to simulate or estimate "normal conditions". Stress tests allow extending the validity of models to extreme or exceptional conditions.

For example, under normal business conditions, a particular cyber loss hypothetically might be best estimated by a normal distribution that is fully characterized by a mean and standard deviation of the loss. Under extreme conditions, however, when a confluence of environmental, economic, and political stresses simultaneously occur, the loss may not fit a normal distribution, so an adjustment must be made. Either the parameters of the normal distribution are changed to reflect the extreme condition or the loss distribution itself is changed (say, to a non-normal, fatter tail distribution).

Stress testing gives analysts, management, and regulators visibility into what losses are possible and extends the usability of a model beyond modeling normal business conditions to include modeling a limited set of exceptional ones. Stakeholders such as investors (shareholders) may have interests, too, in knowing whether any business has sufficient capital to withstand stress shocks. Accounting and investor disclosure rules may emerge to expand the concept of stress testing and reporting/disclosure beyond financial services.[17]

FIs use these stress test results to plan what they would do under extreme but plausible "what if" scenarios. The results are used in several ways:

- The quantified results of the stress tests inform capital calculations
- The exercise of determining stress scenarios and estimating how outcomes differ from "normal conditions" informs management and regulators on what combinations of events can cause extreme losses

When these are known, the FI has insight into what to look for as leading indicators of exceptional conditions and can prepare to mitigate extreme losses once a stress condition is observed. That preparation can be reflected in an institution's business continuity and recovery planning.

[17] The authors expect that sectors beyond the financial sector will increasingly measure expected and unexpected losses, including stress test related losses to manage their cyber risk. These techniques in other words will likely be more widely used beyond the financial services sector in the future.

Much of the purpose of stress testing is to present severely plausible "near doomsday" scenarios to help the institution to have the tools to observe and to prepare for them in advance. Extreme losses can be minimized through such a preparatory approach.

Stress Testing – Context for Cyber Risk Model Stress Testing

Basel III requirements and local regulatory requirements, such as the Dodd-Frank Act Stress Test (DFAST) in the United States, subject risk models to stress tests for purposes of calculating reserves. Cyber risk stress tests need to be developed in context with and analogous to the regulatory risk stress tests since regulators will need to see how the cyber risk stress tests are harmonized within the existing regulatory framework.

Regulators view stress testing and scenario analysis as a necessary complement to the use of internal VaR models. As we have discussed, VaR is a statistical model whose implementation requires making assumptions about the risk factors that is often dictated by the availability of data and use of conventional distributions to model uncertainty. VaR models typically work sufficiently well under "normal market conditions" and provide good insight into those normal conditions. However, VaR models do not typically work well to model adverse or extremely adverse events[18] that history shows either happened or plausibly could occur in the future. Risk managers and regulators use stress test scenarios to gain insight into plausible extreme loss exposures.

Best practices call for determining losses related to various extreme scenarios for which the underlying distribution of the risk factors deviates substantially from a standard probability distribution. Scenarios are based on an arbitrary combination of stress shocks, yet they must be consistent with the basic laws of economics. In any constructed scenario, the chain of events must make economic sense. The chain of events that may logically follow a major shock depends on the context and may be quite different from one crisis to another.

Regulators also require that FIs run scenarios that capture the specific characteristics of the company; i.e., scenarios that involve the risk factors to which their company is most sensitive. An example would be the disappearance of liquidity following a crisis leading to several well-publicized losses.[19]

A stress envelope in response to a particular stress shock[20] is a useful approach. Stress envelopes integrate a set of "stress shocks" across judiciously chosen plausible stress scenarios. Each stress scenario defines a combination of adverse or severely adverse stress shocks that could plausibly occur.

The potential number of combinations of basic stress shocks is overwhelming. In practice, only a relatively small number of scenarios can be routinely analyzed. This means that the scenarios have to be selected according to the specifics of the particular portfolio. The usefulness and accuracy of the diagnosis that emerges out of the scenario analysis depends on the judgment and experience of the analysts who design and run these scenarios. Even the best analysts rely on the past as a guide to the future.

[18] "Adverse" and "severely adverse" correspond to terms used in the DFAST; see Section "Using Open FAIR to Analyze Stress Scenarios".
[19] Regulators require FIs to include "liquidity risk" in their scenario analyses.
[20] A stress shock is an event that is plausible but not considered likely under normal conditions. Examples include an external shock such as a highly infrequent market price move such as a one-day, 500-point fall in the S&P 500 index or an operational event such as the massive scale log4j vulnerability; refer to: https://www.cisa.gov/uscert/apache-log4j-vulnerability-guidance.

In summary, as discussed, conventional distributions may be good at estimating or modeling normal conditions, but they may not be good at modeling extreme conditions. Stress testing attempts to model and capture tail events otherwise uncaptured by many conventional distributions. In other words, stress testing is an attempt to go beyond conventional models to capture estimates of extreme events that the complex real-world generates.[21]

Examples of Cyber Risk Stress Scenarios

Examples of a cyber stress test include "What would happen if…" scenarios that do not represent typical conditions but are still plausible, such as:

- A combination of failures driven by extreme weather or a natural disaster that disables multiple critical dependencies of an IT system at the same time

  IT systems depend upon multiple infrastructures, such as electric power, air conditioning to keep servers from over-heating, cyber security controls, network connectivity, and physical security, all working at the same time. Any single failure is likely planned for and is manageable. At some point, however, multiple concurrent failures are not considered as sufficiently likely to be included in ordinary risk modeling.

- A combination of socio-economic conditions such as a once-in-every-one-hundred-year global pandemic that causes business practices rapidly and unexpectedly to change outside of usual business controls, leading to new but unknown opportunities for exploitation; a pandemic-driven opportunistic increase in the hostile activity of Threat Agents

- Geopolitical tensions rise, which may increase the benefits to state-sponsored actors to conduct cyberattacks, so those actors increase the frequency of their cyber espionage and warfare attacks

- A coordinated attack on critical infrastructure is extraordinarily successful in rendering the critical infrastructure inoperable for an extended period of time, paralyzing the FI's critical service delivery

  If the chance of success of such attacks rises either through improved hacking technology or the increased desperation of Threat Agents, beneficiaries of such an attack have added incentives to increase the likelihood of widespread critical infrastructure attack.

- A deep, global recession that raises unemployment to extreme levels causes new, financially motivated Threat Agents in hard-hit emerging market economies to rapidly escalate cyberattacks against FIs, either for the Threat Agents' personal gain or for the financial gain of the state overall

[21] There are several techniques to model stress environments, with only one shown in the examples described in the following section. Another technique is to simulate jumps or changes to modeled distribution parameters within an existing Monte Carlo simulation to reflect a stress condition. In that technique, distribution parameters (min, most likely, max) of the input parameters describing Loss Event Frequency and Loss Magnitude distributions are not static within the simulation. They dynamically jump to reflect stress conditions so that the stress test appears within the single Monte Carlo simulation. At present, this approach is considered nonstandard in Open FAIR analyses and likely is not supported within commercially available modeling tools. That said, there is no reason that dynamic techniques cannot be used if deemed most appropriate for the stress being modeled.

- Rapid technological change is deployed that has unknown and undiscovered opportunities for malevolent exploitation

  Zero-day exploits are examples of rapid technology change that provide time-limited opportunities Threat Agents can take advantage of.

- One or more of the scenarios above may all occur at once

All of these scenarios represent an extreme change to "normal" conditions represented in standard risk analyses. These changes are often driven by factors external to financial services institutions and represent a change in the natural, geopolitical, technical, or economic environment in which those institutions operate.

Using the Open FAIR Taxonomy to Analyze Stress Scenarios

To align with the DFAST language, the example below analyzes an "Adverse Scenario" and a "Severely Adverse Scenario".

Stress scenarios consist of plausible but extreme combinations of rare events that go beyond anticipated typical, natural variations from a "normal environment". Stress scenarios for cyber risk that have never occurred before may have little data available to correlate external environmental, geopolitical, and economic extreme events to rising cyber risk exposure. The Open FAIR model gives us an insight into how analysts can relate stress test scenarios to cyber risk factors. Once these relationships can be plausibly linked, SMEs can estimate how the extremes indicated in stress tests change baseline risks analyzed for Loss Event Frequency (LEF) and Loss Magnitude (LM).

To analyze any stress scenario (adverse or severely adverse), an analyst would look at how one or more risk factors would change from "normal environmental conditions" under a stress test. The analytic approach answers questions as described in the structured approach below:

- How does Loss Event Frequency change – does it rise, fall, or stay the same?
- Which risk factors under the Loss Event Frequency branch in the Open FAIR tree change?
- Does the Threat Event Frequency (TEF) change? Does Vulnerability (Vuln) change? Why?
- If the Threat Event Frequency changes, is it because the Contact Frequency (CF) changed or because the Probability of Action (PoA) changed, or both?
- If the Vulnerability changes, is it because the Threat Capability (TCap) changed or because the information Asset's Resistance Strength (RS) changed, or both?

The analyst needs to next estimate the magnitude of a change in the risk factors, as follows:

- How does Loss Magnitude change – does it rise, fall, or stay the same?
- Do Primary Losses, Secondary Losses, or both change?
- Which form(s) of Primary Loss change(s) (usually productivity, response, and replacement costs)?
- Does the Secondary Loss Event Frequency (SLEF) change?

- Which form(s) of Secondary Loss change(s) (usually response, reputation, competitive advantage, and fines and judgments)?

After completing this analysis to decompose the impact of the stress scenario upon cyber risk, the analyst can estimate the impact of the scenario in economic terms.

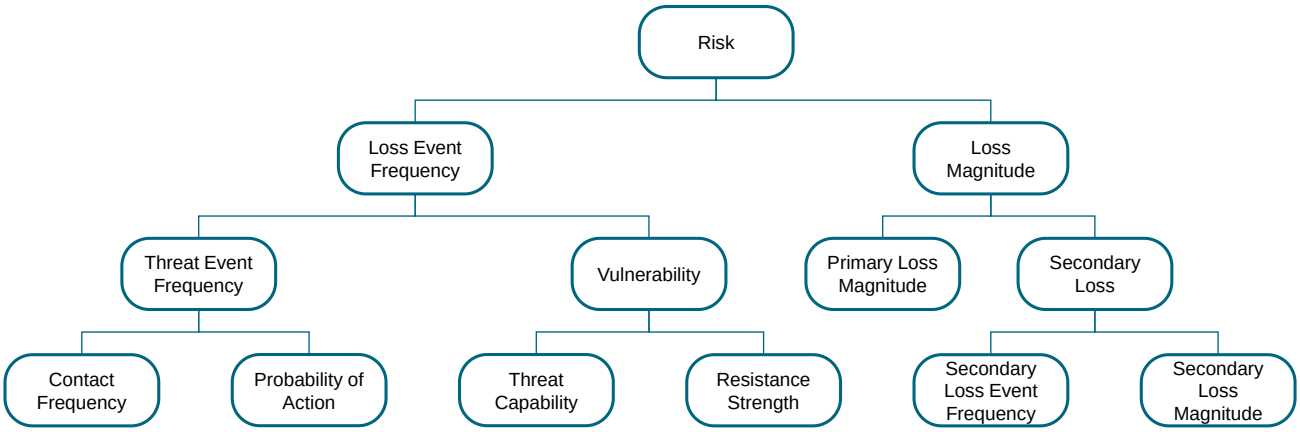


Figure 3: Decomposing Risk Factors: The O-RT Standard

Adverse Scenario Stress Test Example

The adverse stress scenario extends our example used in Section "Risk of a Portfolio that Includes Cyber Risk" from the first White Paper of this series, *Integrating Cyber Risk with Financial Risk* (see References), with the following stress scenario.

Start with a Reference Baseline: The Cyber Risk Example from the First White Paper of this Series

We modeled an unmitigated Loss Scenario involving the breach of Personal Identifiable Information (PII) confidentiality, as shown in Table 1.

Table 1: Summary of the PII Confidentiality Loss Scenario from the First White Paper of this Series, Integrating Cyber Risk with Financial Risk

		Loss Magnitude			
			Secondary Loss		
	LEF	**Primary Loss**	**SLEF**	**SLM**	
	(Events/Yr)	Response ($)	Probability	Response ($)	Fines & Judgments ($)
Minimum	0.2	30,000	0.2	15,000	1,000,000
Most Likely	0.5	100,000	0.3	25,500	1,200,000
Maximum	1.0	200,000	0.5	60,000	1,500,000

The Monte Carlo simulation distribution of losses from the scenario in *Integrating Cyber Risk with Financial Risk* is copied below in Figure 4.[22]

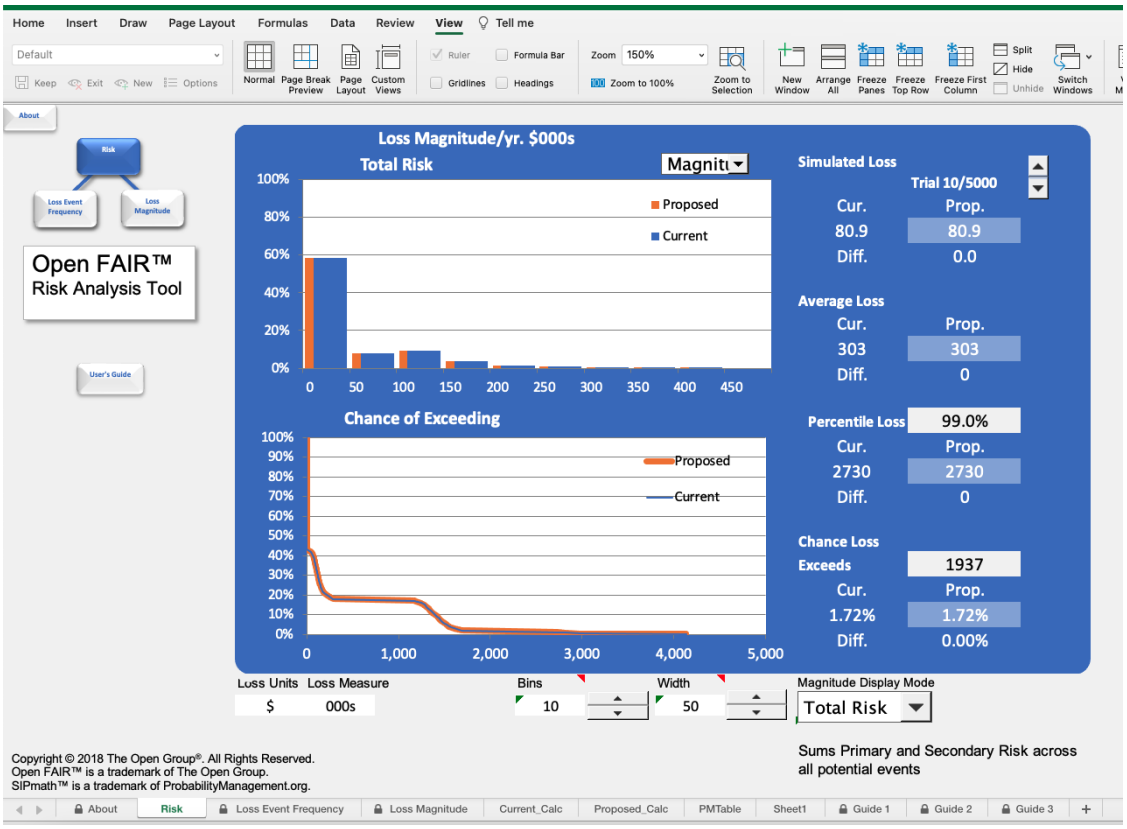


Figure 4: Monte Carlo Simulation of Loss Scenario from the First White Paper of this Series, Integrating Cyber Risk with Financial Risk

This example from the first White Paper, *Integrating Cyber Risk with Financial Risk*, and its simulated results provide a reference baseline model from which Adverse and Severely Adverse Scenarios are developed here in this section to stress test that model. Throughout the stress test scenarios that follow, all references to the reference baseline refer to this model and analysis.

Modeling the Reference Baseline Scenario Under the Adverse Scenario

Suppose that the FI's risk management team wanted to evaluate a PII disclosure Loss Scenario to determine what would happen if a hostile emerging-economy nation state was under economic stress from a global recession and decided to attack the financial infrastructure of developed economies.

[22] See Page 28 in *Integrating Cyber Risk with Financial Risk*.

The analyst would assess how Loss Event Frequency and Loss Magnitude change from the reference baseline under the stress conditions as follows:

- Loss Event Frequency Analysis

  The Threat Event Frequency would likely rise because of the increased willingness of the hostile Threat Agents to try and penetrate FIs for financial gain. Assuming that the Vulnerability of the information Assets were unchanged, the Loss Event Frequency must necessarily rise.

- Loss Magnitude Analysis

  Because the:

  — Attacks are more sophisticated than usual: the analyst and SMEs believe and argue that primary response costs will rise to determine what went wrong and how the Assets were penetrated

  — Hostile actors monetize their cyber hacking: Secondary Loss Event Frequency rises to a greater extent than under more normal conditions

The analyst and SMEs believe that secondary response costs and the exposure to fines and judgments are unchanged under this stress scenario.

Using expert judgment and any other information available to inform estimates, the analyst estimates the impact to the risk factors have been driven to be as shown in Table 2.

Table 2: Summary of PII Confidentiality Loss Scenario from the First White Paper, Integrating Cyber Risk with Financial Risk, Evaluated through an Adverse Scenario Stress Test

			Loss Magnitude			
				Secondary Loss		
	LEF		**Primary Loss**	**SLEF**	**SLM**	
	(Events/Yr)		Response ($)	Probability	Response ($)	Fines & Judgments ($)
Minimum	0.75		60,000	0.4	15,000	1,000,000
Most Likely	1		200,000	0.6	25,500	1,200,000
Maximum	1.5		500,000	1	60,000	1,500,000

The Adverse Scenario risk is shown below as the “Proposed” (orange) scenario in Figure 5. The “Current” (blue) graph is the reference baseline as described in Table 1 and Figure 4.

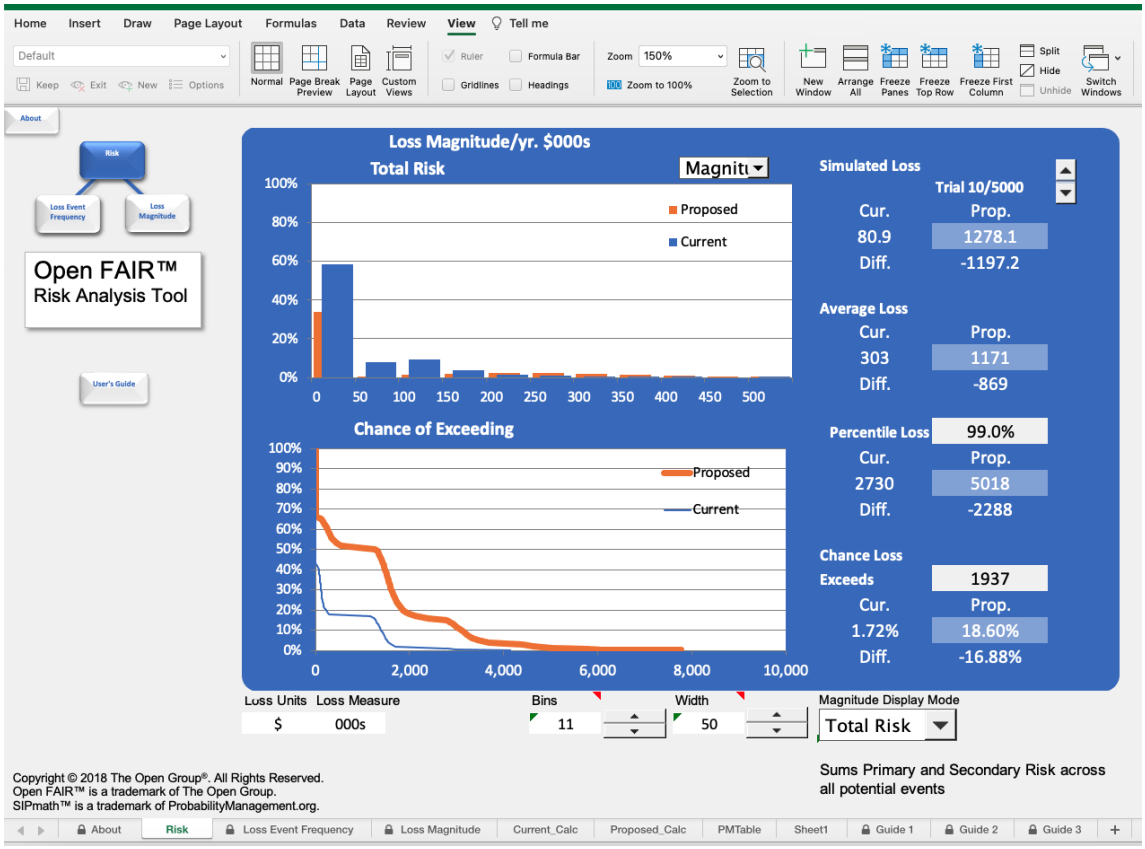


Figure 5: Adverse Stress Scenario Results (Orange, Proposed) Compared to the Reference Baseline (Baseline, Current)

The total risk exposure rises due to an increase in both the Loss Event Frequency and Loss Magnitude. The VaR and the Conditional Value at Risk (CVaR) of the stress test are respectively derived from the tail percentiles (e.g., 99%) and expected values in the tail as described in the first White Paper of this series, *Integrating Cyber Risk with Financial Risk.* Table 3 shows these measures distilled from the analysis captured in Figure 5.

Table 3: Comparison of Reference Baseline Scenario and Stress-Tested Scenario

	Current (Reference Baseline)[23] ($)	Proposed (Adverse Scenario)[24] ($)	Difference[25] ($)
Average Loss per Year:	303,000	1,171,000	-868,000
99th Percentile Loss:	2,730,000	5,018,000	-2,288,000
99th Percentile Tail Expected Value:[26]	3,130,000	6,128,000	-2,998,000

Severely Adverse Scenario Stress Test Example

A "Severely Adverse Scenario" consists of plausible scenarios that if they occurred would result in total losses more severe than the "Adverse Scenario".

Severely Adverse Scenario Illustrative Example

The Severely Adverse Scenario adds additional risk to the geopolitical and economic stresses described in the Adverse Scenario. In particular, we assume that geopolitical conditions deteriorate and cause further economic stress beyond that described in the Adverse Scenario. In this Severely Adverse Scenario, financial cybercrime is temporarily not considered a crime by the hostile nation state. Cybercrime increases under this scenario since cyber criminals face fewer risks to themselves in terms of law enforcement actions within their own nations.

Severely Adverse Scenario Loss Event Frequency Analysis

Compared to the Adverse Scenario, the Severely Adverse Scenario has a higher Threat Event Frequency due to the reduced deterrent effect from the de-criminalization of cyber hacking and cybercrime in the region. Analysts further expect the number of cyber criminals to increase. Within the Open FAIR framework, the Contact Frequency rises (since more cyber criminals enter the "business" of cybercrime against FIs) and the Probability of Action increases (from the reduction in the deterrent effect from de-criminalization). In combining these two increased risk factors, the Threat Event Frequency under this scenario is estimated to significantly rise compared to the Adverse Scenario.

Vulnerability and its associated Threat Capability and Resistance Strength risk factors are assumed not to change in this example.

[23] Current is the Reference Baseline Scenario from the first White Paper of the series, *Integrating Cyber Risk with Financial Risk*.

[24] Here, the "Proposed" state models the risk associated with the Loss Scenario under the stress test. Note that the signs of the differences are negative, indicating that the stress test risk exceeds the reference baseline of the original risk scenario.

[25] The difference reflects the change in risk measures between the stress-tested scenario and the reference baseline.

[26] *The Open FAIR™ Risk Analysis Tool (I181), January 2018 (*see References) does not calculate the expected value of a tail as part of its standard output. The authors accessed the internal Monte Carlo simulation data within the risk analysis tool and manually calculated the expected value of the tail in this analysis.

With the Threat Event Frequency significantly rising and Vulnerability staying the same (again, as compared to the Adverse Scenario), Loss Event Frequency is estimated to significantly rise.

Severely Adverse Scenario Loss Magnitude Analysis

The analyst considered that the Severely Adverse Scenario only changed incentives around the Threat Event Frequency and losses, when committed, would be as they were described in the Adverse Scenario. There is no change to the Adverse Scenario's Loss Magnitude analysis.

Severely Adverse Scenario Analysis Results

In considering the effect of the Severely Adverse Scenario, the analyst and other SMEs adjusted the Loss Event Frequency estimate as shown in Table 4 and kept the same Loss Magnitude estimates used in the Adverse Scenario.

Table 4: Summary of PII Confidentiality Loss Scenario in the First White Paper, *Integrating Cyber Risk with Financial Risk*, Evaluated Through a Severely Adverse Scenario Stress Test

			Loss Magnitude			
				Secondary Loss		
	LEF		**Primary Loss**	**SLEF**	**SLM**	
	(Events /Yr)		Response ($)	Probability	Response ($)	Fines & Judgments ($)
Minimum	1.25		60,000	0.4	15,000	1,000,000
Most Likely	1.75		200,000	0.6	25,500	1,200,000
Maximum	3		500,000	1	60,000	1,500,000

Figure 6 shows the results of the Adverse and Severely Adverse Scenarios as analyzed with the risk factors above.

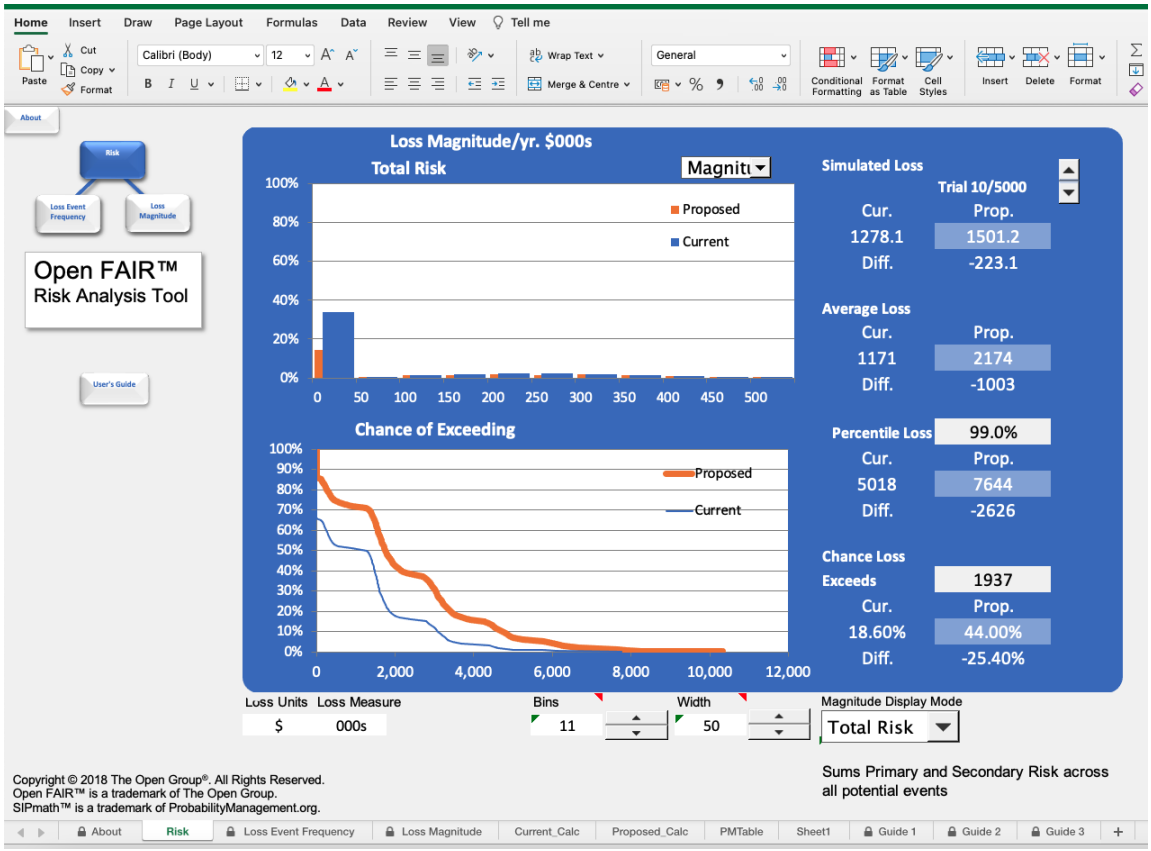


Figure 6: Adverse (in Blue, Current) and Severely Adverse (in Orange, Proposed) Scenario Stress Test Results

Putting these three scenarios together – the Reference Baseline, the Adverse Scenario, and the Severely Adverse Scenario – gives the summary figures in Table 5.

Table 5: Summary of Reference Baseline, Adverse Scenario, and Severely Adverse Scenarios

	Reference Baseline ($)	**Adverse Scenario ($)**	**Severely Adverse Scenario ($)**
Average Loss per Year:	303,000	1,171,000	2,174,000
99th Percentile Loss:	2,730,000	5,018,000	7,644,000
99th Percentile Tail Expected Value[27]:	3,130,000	6,128,000	8,596,000

[27] *The Open FAIR™ Risk Analysis Tool (I181)*, January 2018 (see References) does not calculate the expected value in the tail. These calculations were done manually.

Summary

Cyber risk analysis lends itself to stress testing in an analogous way to risk analysis performed in an FI. By linking extremes of environmental, economic, political, and technological stresses to cyber risk factors, analysts have an approach to stress test cyber risk Loss Scenarios.

Cybersecurity and risk functions can use the analytical process described here to estimate how these external stressors affect cyber risk factors, such as Contact Frequency, Probability of Action, Threat Capability, and Asset Resistance Strength as they roll up to the higher-level Loss Event Frequency factors of Threat Event Frequency and Vulnerability. CISOs and their cybersecurity teams can take an adaptive approach to loss prevention controls, those controls that avoid, deter, and resist threat actions. During times of stress, CISOs and executive management can choose to add controls to prevent losses during extreme periods if they know what to look for and have prepared for those contingencies. Stress testing helps all parties get the insight to learn what to look for and how to prepare for it.

Similarly, CROs and CISOs can incorporate external environmental factors into its impact on the magnitude of a loss. For example, in a stress test environment regarding:

- Primary Loss: do productivity, replacement, or response costs rise?
- Incidence of Secondary Losses: does the Secondary Loss Event Frequency rise?
- Secondary Loss: do response, reputation, competitive advantage costs, or exposure to legal and regulatory remedies of fines and judgments rise on a per-incident basis?

Quantitative analysis helps to frame these questions and provides CROs, CISOs, and general management insight into how detection, response, and recovery from incidents are affected by external factors.

The act of stress testing and the analysis behind it give all operational parties further insight into what to look for as events unfold. If there is a known plausible relationship between macroeconomic, geopolitical, or environmental outcomes causing an increased frequency and severity of a particular type of cyber-criminal behavior, then CISOs and risk managers who are also informed can be more prepared to take pre-emptive action. In other words, risk is better managed by anticipating a change in the Loss Event Frequency and/or Loss Magnitude of a cyberattack caused by a sudden and dramatic change in the environment, such as a market crash, elevated geo-political tension, or other systemic stresses. Stress test analysis can make everyone more risk-informed and achieve a greater risk intelligence. The result of acting on that improved intelligence improves preparation for and resiliency to unexpected events.

Finally, developing empirical relationships between abnormal circumstances and elevated risks of cyber losses is an area of research the authors believe is lacking.

Organization of the Cyber Risk Analysis Supply Chain

A specialization and division of labor has emerged to deliver cyber risk analysis goods and services, which includes:

- The Open Group that creates reliable knowledge by defining and standardizing terms, definitions, relationships, and analytic processes on cyber risk analysis
- Tool vendors who create and sell general purpose risk calculating tools based upon the standards
- Analysts and consultants who use the standards, tools, data, and their expertise to evaluate specific cyber risk scenarios
- Data providers who research and publish information about cyber events within the categories defined in the standards

Taken in combination, these actors form an ecosystem of specialists each of whom plays a part in quantifying cyber risk in economic terms.

The Open Group and its Open FAIR Standards

The Open Group is an important component of the cyber risk analysis supply chain since it is a consortium of information system hardware, software, and services providers (vendors) and Information Technology (IT) consumers (customers) whose members develop global IT standards.[28] The Open FAIR cyber risk analysis standard was published in 2006 and initiated its analyst certification program in 2013. There are two standards: The Open Group Standard for Risk Analysis (O-RA) and The Open Group Standard for Risk Taxonomy (O-RT) (see References). The Open Group Security Forum publishes and maintains additional white papers, guides, certification study guides, and *The Open FAIR Risk Analysis Tool*[29] (see References) to elaborate upon and support the two standards and the certification scheme.[30] The Open Group functions as a knowledge producer within the ecosystem.

The O-RT Standard is a taxonomy that defines terms, definitions, and their relationships (logic) that software suppliers can use to implement software tools. Software suppliers make general purpose tools that analysts use to model cyber risk Loss Scenarios. Software suppliers use their knowledge of the cyber risk domain and

[28] The Open Group is organized into interest areas called Forums. Among its Forums, The Open Group Architecture Forum is the largest interest group, and it publishes and maintains the dominant standard in Enterprise Architecture: The *TOGAF® Standard*; refer to www.opengroup.org/library/c220. The Open Group owns and licenses the *UNIX® Standard* (refer to: www.opengroup.org/library/x1201) to define UNIX branded operating systems, and The Open Group Security Forum is responsible for projects and standards in the areas of information security architecture and information risk analysis.

[29] The *Open FAIR™ Risk Analysis Tool* is a Microsoft® Excel spreadsheet that The Open Group developed as an illustrative example to demonstrate the Open FAIR method and how it could contribute to calculating cyber risk. The Tool develops a "Current" state and a "Proposed" state to analyze the present state of an information system's risk and what it would be should a proposed set of additional controls be added. "Current" refers to that current, initial state being analyzed, while the "Proposed" state is that Current state as changed by the set of additional controls. Refer to: www.opengroup.org/library/I181, or see References.

[30] Refer to: https://www.opengroup.org/certifications/openfair.

what their analysts/professional services consultants require to implement calculation tools that apply the Open FAIR terms, definitions, and relationships as defined in the O-RT Standard. That said, at the time of writing, there is no standard or certification that articulates what an Open FAIR tool must do, aside from using the Open FAIR terms accurately and implementing their relationships within the tools. Details, such as what Loss Event Frequency and Loss Magnitude distributions to use in modeling cyber risk factors, are left up to the tool vendor – and its customer base of analysts and consultants – to decide.

The O-RA Standard is a standard that defines the basics of what a cyber risk analysis contains and an approach to using the risk factors defined in the O-RT Standard. It, too, is general purpose and does not define proscriptively what an analyst must do, but it instead describes the terms, definitions, and approach to applying the Open FAIR taxonomy's terms, definitions, and relationships to cyber risk analysis.

Tool Vendors Make General Purpose Tools

Risk analysts and consultants use cyber risk analysis software tools to define and model risk scenarios, input risk factors in a structured way, calculate cyber risk results, and display those results. Because at the time of writing the market for cyber risk analysis tools and services has not yet grown sufficiently large to specialize into specific industry segments, tool vendors provide general-purpose modeling, calculation, and visualization tools that are meant to be used by knowledgeable analysts. For example, tool vendors provide a selection of distributions that analysts choose to model a particular cyber risk scenario. The survival of tool vendors depends upon market discipline. They survive or fail by how successfully their analyst and consultant customers can measure cyber risk with those tools.

There is competition in the market for these modeling tools, and there is no "single right way" to implement them. Competition has led to diversity in features, functions, and quality of tools currently available. There is no current standard that specifies what any tool must do.

Analysts and Consultants Use Those Tools

Analysts (internal to a cyber risk organization) and consultants (external professional services providers) use and depend upon the language and logic of the standards to define, analyze, and report cyber risk analyses to their clients in a way that is precise and unambiguous. Analysts use vendor-provided tools and apply them to the specific needs of their clients. Analysts use their professional expertise and judgment to apply these tools to generate specific risk scenarios that are geared to particular industries.

Analysts can disagree on how to effectively model cyber risk due to the dynamic nature of cyber risk. For example, analysts are often challenged in choosing what distributions to use and the parameters to use given what is known and unknown. Consultants perform the essential analytic problem-solving function that defines and analyzes risk scenarios for their clients through using their expertise and judgment in applying standards, selecting tools, and eliciting estimates from SMEs.

Data Providers are an Emerging Link in the Supply Chain

Analysts need data to inform their estimates and discretionary decisions in modeling cyber risk. Cyber risk data has recently become more generally available from such sources as the Cyentia Institute, IBM®, and the Ponemon Institute. As more data is collected and communicated, these sources and others will perform an increasingly important and valuable role in cyber risk analysis.

Summary

Analysts and consultants use their experience and knowledge to add value to risk analysis to make discretionary decisions throughout the analytical process. The strength of the analysis depends upon the analyst's reasoning in the context of a specific risk scenario. That strength increases when analyses are broadly consistent with an industry standard's terminology, logic, and analytic method; whose results are derived from tools that have survived market competition; have well-reasoned analytic rationales; and use all relevant information available. The combination of standards, tools, analysts and consultants, and information providers form a coherent cyber risk analysis supply chain.

How the Open FAIR Model Has Been Vetted

The model vetting process answers three questions:

- Is the model relevant?
- Is the model correct?
- Has the model been implemented correctly?

To answer these questions, each participant in the supply chain of the Open FAIR analysis provides part of the answer through documentation, examination of soundness, and all information available to inform and test models.

Members of The Open Group Security Forum, the cybersecurity interest group, accepted the two Open FAIR standards as relevant and correct in defining and describing cyber risk. Subsequently, vendors have developed Open FAIR analysis tools that combine risk factors described in those standards to calculate the probable frequency and probable magnitude of future loss of risk scenarios, and firms provide training and consulting services to use those tools to identify and quantify risk factors using the information their clients have to describe those risk factors.

This section breaks down how The Open Group, vendors, and analysts/consultants address these areas to answer the main questions listed above.

Model Relevance

The O-RT Standard presents a set of terms, definitions, and their relationships to each other to describe a vocabulary, logic, and model of cyber risk in the terms of a structured tree as shown in Figure 7 (shown earlier as Figure 3 and reproduced here for convenience).

Risk is defined as: “The probable frequency and probable magnitude of future loss”.

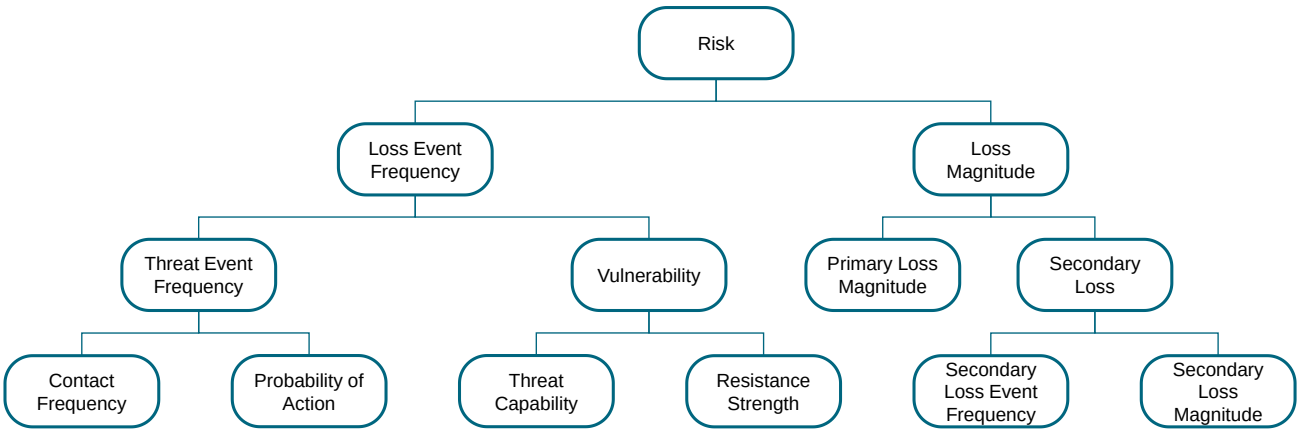


Figure 7: The Open FAIR Risk Taxonomy

The O-RA Standard applies this vocabulary, logic, and model to developing and analyzing stylized Loss Scenarios that can be decomposed and analyzed within the Open FAIR taxonomy. The model that the taxonomy describes has the following properties:

- Terms are precisely defined, which enables clear communication among stakeholders in managing cyber risk
- The taxonomy describes precise mathematical relationships between terms in the model, which brings transparency to the model and improves independent verification of tools implementing the model
- Key assumptions are defined
- All risk factors in the taxonomy are:
  - Modeled as random variables
  - Statistically independent of each other
- Key analytical interpretation is allowed, which means the model can be adapted for specific risk scenarios
- The distributions of risk factors are left to the analyst

  The Open FAIR standards do not prescribe any particular distribution for any risk factor. Instead, the analyst, considering the information available, uses the distribution that the data indicate is the best fit.

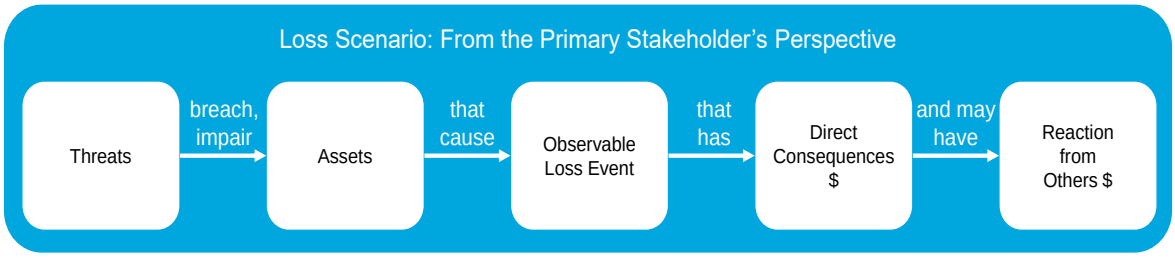


Figure 8: Stylized Loss Scenario in Open FAIR Risk Analysis

Open FAIR Tool Vetting

Vendors produce risk analysis tools that are general purpose cyber risk analysis and calculation tools, appropriate for any client who uses IT. Vendors must interpret the Open FAIR standards, select what they think their customers need to model Loss Scenarios accurately, verify the tool's mathematical correctness, and release and support it to consultants and analysts.

Vendors use traditional software architecture, engineering, development, and quality assurance processes to develop, test, and vet risk calculation tools. These techniques yield tool software quality and assurance similar to any other enterprise-class software automating business.

Tool vendors compete for analysts and consultants to use their tools: the market helps to ensure that tools are relevant and correct. Tools found to have errors either do not survive or are changed once those errors have been discovered.

Open FAIR Model Vetting

Analysts use modeling tools to model individual cyber risk scenarios on behalf of their firm or client. These analysts get their information to input into the risk modeling tool by researching available information on historical cyber incidents and their magnitude of loss, interviewing SMEs within the firm, developing Loss Scenarios specific to the firm, and performing their own research to estimate the Open FAIR risk factors.

The quality of these risk models depends upon:

- The quality of the risk scenario being developed in the first place

  Analysts work with their firm's management to define what is most important to model and estimate

- The quality of the assumptions the analysts makes

  Analysts will have to work with their firm's SMEs and management to make assumptions about the Loss Scenario and the areas of the business likely impacted should it happen. Those assumptions and estimates can be debated and challenged for accuracy and reasonableness.

- The quality of the data the analyst has available

  Open FAIR analysts are trained to make calibrated estimates based upon incomplete information. That information can come from SMEs within the organization, from research into the probable frequency and probable magnitude of losses as publicly reported within the same or similar industry, or private, subscription-based information service providers. These estimates are first tested to be accurate, at the expense of precision.

As mentioned earlier, compared to risk analyses that take place in an FI, cyber risk analyses often have less precision than financial risk analysts are used to due to the rapid change in composition of threat communities, the tactics and technology those communities use to attack systems, and the rapidly changing technology used to defend against those attacks. Cyber risk analyses are considered high quality when they represent the uncertainty in the data and assumptions used to describe the Loss Scenario being analyzed. If that uncertainty is high, then the uncertainty in the results, too, will be high.

Conforming to the standard requires the analyst to document all the assumptions, estimates, and rationales behind them. Through their research and documentation, analysts must defend their analytic reasoning. An analyst who has confidence in the accuracy of the information going into a commercially available model can trust that the calculations are correct and rely upon the logic and soundness of the O-RA and O-RT Standards as adopted by a respected industry standard-setting organization.

What The Combination of the Standards, Tools, and Analysts Brings to Cyber Risk Modeling

The combination of the standards, tools, and analysts brings a capability to model and quantify cyber risk in economic terms. The standards bring clarity of definition to cyber risk, a stylized model of how cyber losses occur, and a set of terms, definitions, and relationships that can be applied to quantifying risk.

At the same time, the authors of this document claim that the standards are not so prescriptive that the analyst loses the flexibility to improve their application in individual analyses. In other words, the analyst can, when the data warrants, improve the quality of the analysis by:

- Changing how distributions are parameterized

  The Open FAIR standards, for example, describes the parameters of all estimates as "minimum, most likely, and maximum". If the analyst chooses a distribution that is completely described by two parameters – say, average and standard deviation – the analyst has the flexibility to use those two parameters instead and comply with the spirit of the standard.

- Reducing the uncertainty captured in estimates

  The Open FAIR standards instruct analysts to use the calibrated estimating technique to estimate risk factors in the absence of a lot of data. That technique specifies a probability of accuracy of 90 percent overall in any risk factor estimate. If the analyst has data sufficient to improve that accuracy beyond 90 percent, they can (and should) estimate risk factors to that improved probability of accuracy – for example, to the 99 percent or beyond.

The Open FAIR standards were designed with the knowledge of cyber risk at the time of publication. The authors believe that the standards should be interpreted flexibly enough to incorporate new knowledge discovered after the publication of those standards while still using the language of risk; such as the standardized terms, definitions, and relationships defined within that language.

Estimating Risk Factors

Analysts face the challenge of estimating risk factors that can be back-tested against historical experience while incorporating SME opinion on how the future is likely to differ from the past. What embodies experience is all the available information the analyst can access about how many Loss Events per year have occurred and the impact of those events on the FI. That information can come from within the FI, from membership-based communities the FI has joined (e.g., Information Sharing and Analysis Centers (ISACs) and public-private partnerships usually with law enforcement), and from available published reports on cyber losses. Taken collectively, analysts gain useful insight into how often cyber losses occur and how bad they are when they do.

All risk factor estimates reflect uncertainty; that is, all estimates are expressed as random variables. The O-RA Standard states that each random variable should be expressed as a range consisting of minimum, most likely, and maximum values. The O-RA Standard does not specify a distribution for any risk factor since that choice is left to the analyst.

When the analyst models a risk factor using a distribution that needs less than three parameters to fully specify it, analysts do not have to conform to expressing the parameters of that distribution through minimum, most likely, and maximum values. For example, if an analyst chooses the normal distribution to estimate response costs, then only the mean and the standard deviation (or variance) are needed to completely describe that estimate.

Making Estimates Within an Assumed, Transparent Context

All estimates should be made within a transparent set of assumptions and context.

The Open FAIR risk model was constructed under the assumption that risk factors are statistically independent of:

- Each other

  In other words, a zero correlation between risk factors. That said, analysts can revise the model to reflect known correlations.[31]

- Geopolitical, geo-economic, and other systemic factors

  Although political and economic conditions vary and can plausibly affect Loss Event Frequency and Loss Magnitude, those influences are ignored.

These pragmatic adjustments to the Open FAIR risk analysis method conform to the intent of the Open FAIR standards. Risk analysts are expected to use all knowledge they have about any specific analysis to model it

[31] This assumption of independence may not always be true. For example, a ransomware breach that causes a high productivity loss from data being unavailable may have a commensurately higher response cost loss as the organization tries to recover and some component of response costs may scale with restoring or recovering the amount of unavailable data.

most accurately. In other words, if one or more of the standard assumptions are not true for a specific analysis, analysts should estimate those effects by adjusting risk factor estimates as needed. Making these pragmatic adjustments to the Open FAIR standard risk analysis method conforms to the intent of the Open FAIR standards.

Risk factors are always estimated within some context, and making that context transparent is necessary to establish the credibility of those estimates. Essential context (and constraint) includes, but is not limited to:

- What environment is being modeled (normal, adverse, severely adverse)?
- What is the stylized, standard Loss Scenario for the analysis (described in Figure 8)?
- Why is the analysis being done?
- Are the results useful in context of the problem being addressed?
- What time, money, and other resources are available to spend on the analysis?
- What is the availability and quality of the data used to inform estimates?

All analyses are conducted under a set of assumptions and within some context. Effectively establishing whether the analysis is fit for purpose requires that each assumption and contextual element is made transparent.

Using Information to Inform Estimates

Analysts use a variety of statistical techniques to analyze historical information and assess how the future is likely to differ from that history. SMEs can help to inform assessments to arrive at an accurate estimate that best captures the uncertainty of the cyber risk exposure the FI faces going forward.

Internally Available Information on Loss Event Frequency and Loss Magnitude

If the FI has suffered prior cyber losses or has suffered near misses, it likely has some researchable record of those events.

Information Sources to Inform Loss Event Frequency Estimates

Information security organizations frequently have Security Information and Event Monitoring (SIEM)[32] tools that can inform estimates of Loss Event Frequency and through loss detection can alert information security operations of the Loss Event as it occurs to mitigate Loss Magnitude. Based upon the logs those tools rely upon, it may be possible to use SIEM tools to give insight into the Contact Frequency and the Threat Event Frequency.

[32] Security Information and Event Management (SIEM) is defined as "Security Information and Event Management (SIEM) technology supports threat detection, compliance and security incident management through the collection and analysis (both near real time and historical) of security events, as well as a wide variety of other event and contextual data sources. The core capabilities are a broad scope of log event collection and management, the ability to analyze log events and other data across disparate sources, and operational capabilities (such as incident management, dashboards, and reporting)"; refer to: *Gartner Glossary: Security Information and Event Management (SIEM)* (see References).

SIEM tools can also give insight into the breadth of the damage of a successful attack – how much data were compromised and how long system outages lasted, informing LM estimates.

Information Sources to Inform Loss Magnitude Estimates

Analysts should examine how an FI has historically responded to an event inside the organization, such as the cost of:

- Internal people involved in the response and the estimated business impact of the event, by measuring the internal accounting cost of the person-hours required to respond to prior incidents
- External consultants such as outside cyber forensic experts, media relations, outside counsel, etc., which can be used to estimate the range of outside expertise needed to respond to a cyber event once it occurs
- Replacing destroyed equipment
- Any productivity that was impaired such as through a loss of information system availability, which can be used to estimate that impact by the measuring the internal accounting cost of the idled workforce affected

Externally Available Information on Loss Event Frequency and Loss Magnitude

Several firms and organizations specialize in researching and reporting cyber Loss Events. This information augments internal information to inform the FI analysts' estimates on Loss Event Frequency and Loss Magnitude.

These reports or information communication can be classified as:

- Published and made generally available to the public at little to no charge
- Published or made available through paid subscription
- Made available semi-formally through closed, membership-based interest groups
- Made available through "public-private partnership" organizations, usually between law enforcement and industry

Published Reports Generally Available to the Public[33]

Several firms publish annual reports on their research into data breaches. Some of the best known are:

- The Verizon™ Data Breach Investigation Report (DBIR) (see References)

  The 2020 DBIR report analyzes and summarizes 3,950 data breaches across 16 different industry segments. The report is available at https://enterprise.verizon.com/resources/reports/dbir/.

[33] What follows is a non-exhaustive selection of example reports and sources to give analysts a general idea of information that is available to those who are willing to search it out. We do not "vouch" for the quality of these but simply share them to guide analysts in what they can expect to find. Analysts must decide for themselves whether the data quality in any report is fit for analytical use.

- IBM and the Ponemon Institute collaborate to publish its annual Cost of a Data Breach Report (see References)

  In 2020, that report presents analysis of the cost of over 500 real cyber security incidents. The report is available at https://www.ibm.com/account/reg/us-en/signup?formid=urx-46542.

- The Cyentia Institute publishes its annual Information Risk Insights Study (IRIS) (see References) and measures the frequency and cost of cyber incidents using ten years of historical data

  IRIS is based upon data that Advisen™ makes available to Cyentia. In 2020, Cyentia published two analyses: the IRIS 20/20 Xtreme report that looks at the 100 largest cyber Loss Events in the previous five years and the IRIS 20/20 general report that analyzed all Loss Events, not just the extreme ones. Both reports are available at https://www.cyentia.com/iris/.

These data sets are generally available to the public and collectively give analysts insight into the Loss Event Frequency and Loss Magnitude of cyber Loss Events across multiple industry segments.

Information Available Through Paid Subscription

The Cyentia Institute IRIS reports are based upon the data that Advisen researches and analyzes. Advisen makes their data and analytic capabilities available to their clients on a fee basis and is an example of research and analysis firms who provide cyber risk data through paid subscriptions. More can be found at www.advisenltd.com.

Membership-Based Interest Groups

There are many membership-based organizations, sometimes referred to as ISACs, that share and analyze cyber risk information, especially timely, current information about cyber threats, attacks, and response tactics. Members in these organizations share confidentially, under Chatham House style rules[34] timely information on threats, novel attack scenarios, and best practices in resisting those attacks. Through that information and experience sharing, individuals can share what they see going on with other members to improve the cyber risk management practices of all. These groups may only be open to specific industry sectors such as financial services, or they may be open to all who extensively rely upon IT to conduct their business.

Example organizations are:

- The International Information Integrity Institute; refer to: https://www.i4online.com
- The Financial Services Information Sharing and Analysis Center; refer to: www.fsisac.com
- The Bank Policy Institute's BITS organization; refer to: https://bpi.com/bits/

By connecting with member-based industry organizations, analysts can obtain a clearer view into the current state of the cyber risk those members are exposed to, improving the analyst's accuracy and precision in estimating cyber risk factors.

[34] Refer to: https://en.wikipedia.org/wiki/Chatham_House_Rule.

Public-Private Partnerships

Government regulators and law enforcement can outreach to the private sector to combat problems such as cybercrime through a Public-Private Partnership and collaboration. For example, Infragard® is a public-private partnership between the US Federal Bureau of Investigation and the private sector. More can be found at https://www.infragard.org.

Information is Available if Analysts are Willing to Look

Analysts have a wide range of information on the history of Loss Events that are available, both inside and outside their FI. By drawing on this information, analysts can make informed, reasonable estimates on Loss Event Frequency and Loss Magnitude that can be compared to the FI's historical record or to other organizations of like size or scope.

Using SME Opinion to Inform Estimates

All risk factor estimates are forward-looking, and while historical data can inform those estimates, SMEs can offer additional insights into how the future is likely to differ from the past.

SME opinion can come from any of the organizational sources the analyst gets historical information from:

- Internally to the FI
- Producers of public reports
- Cyber risk and financial sector ISACs
- Public-private partnership

Experts in all of these groups likely have well-informed opinions that would enhance risk factor estimates beyond extension of historical observable data.

Using Market Prices to Estimate Risk Factors

An undeveloped but emerging information source to inform risk factor estimates can come from observing prices in various markets. For example, cyber insurance policy prices can serve as one measure of the economic value of risk since an insurance company would not rationally charge less than the expected annual loss of a risk for any policy it would underwrite. Cyber insurance is no different. Prices of insurance against such hazards as ransomware serve as one indicator of risk valuation.

An emerging, and perhaps speculative, field of inquiry is the options market as a source of cyber threat and cyber risk intelligence. A company's options being traded in an unusual way could indicate an imminent danger of attack. A cyber Threat Agent prior to commencing an attack might purchase an unusually large amount of put options[35] against the cyber target, hoping to cash in on the victim's falling stock price after the attack is complete. Combining observed options market prices and volume with other cyber threat

[35] Refer to: https://en.wikipedia.org/wiki/Put_option.

intelligence information could inform management, policy makers, law enforcement, and other groups of emerging cyber risk events. This inquiry is admittedly in its infancy, but the concept has been applied in other contexts, such as terrorism.

Capturing the Tail of Risk Factors in an Estimate

To model VaR accurately, analysts need to capture the likely tail events in their estimates. To do this, analysts need to make calibrated estimates of risk factors that also satisfy the requirements the financial sector has for capturing and estimating tail events. Ensuring these estimates satisfy an FI's risk management requirements requires an adjustment to the guidance presented in the O-RA Standard on how to do calibrated estimation.

Calibrated Estimates of the Parameters in Modeling Cyber Risk to Calculate Reserves

A 90 percent confidence (or "degree of belief") in the accuracy of a risk factor's range estimate would mean that for an example calibrated estimate of a Primary Loss response cost of between $30,000 and $200,000, 90 observations of the response cost loss will be within that range, with ten outside the range. Of those 10 losses outside the range, about five of them would exceed the upper bound of the range, and those five represent uncaptured tail events. Not capturing these tail events makes any VaR calculation blind to what is really happening in the tail. The 90 percent confidence in the accuracy of a range estimate must be adjusted when completing calibrated estimation for VaR.

Accuracy First, Precision Second

The guideline of 90 percent overall accuracy (given that it is based on a 95 percent accuracy for each of the minimum and maximum estimates) on calibrated estimates reflects a pragmatic balance between accuracy and precision, which is achievable when estimators have limited, minimal information available to inform an estimate. That balance represents what people cognitively can be trained to achieve in the context of limited information - the usual circumstance estimators face when estimating cyber risk factors.

It is believed that eliciting estimates beyond the 95 percent accuracy limit, say to 99 percent, exceeds the cognitive ability of most estimators when those estimates rely upon judgment and a sparse historical record, usually resulting in the estimate's range being so wide that is not very useful for decision-making, and likely not even accurate at the 99 percent cut-off.

In the case of making estimates to support a VaR calculation, analysts will have to use all the available information they have (including SME opinion), and then ensure that their calibrated estimates have sufficient range to capture the tail events implied in the historical record and SME opinion. Note that this may entail having higher confidence, such as a 99 percent confidence in the risk factor's estimated minimum or maximum value, assuming there are sufficient data of sufficient quality to do so. More research needs to be done to improve cyber risk modeling for tail risk analysis.

Achieving Parsimony is the Goal

In making estimates of future outcomes, the analyst's goal is to achieve parsimony: to accurately reflect both what they reasonably know about that future and what they do not. The SME calibrated estimate tradeoff is one approach to achieving parsimony, but that approach is not the only one. To meet the needs of financial managers and regulators, analysts need to concentrate on capturing accurate estimates of tail events. That goal is consistent with the intent of the Open FAIR model.

Calculating Reserves for Cyber Risk: Vetting Cyber Risk Models

Analysts can find data from many sources to inform reasonable estimates, and SMEs can enhance the historical record to forecast how the future can differ from the past. The historical record appears to show some remarkable consistency year to year in cyber losses across multiple industries. There is evidence to suggest that cyber losses have some degree of predictability.

There is also evidence that Threat Agents change, that their techniques change rapidly, and that single LM is rising with systemic threats to information infrastructure. These changes imply that SME opinion should be added to the historical record to estimate the potential for improbable but high impact events. Cyber risk analysis is still an art whose practitioners will develop new models and relationships over time between geopolitical, macroeconomic, technological, economic, and other factors and how they affect a FI's cyber risk exposure. In estimating the Open FAIR risk factors, analysts can apply these discovered relationships to better estimate Loss Event Frequency and Loss Magnitude.

In the end, analysts have to accurately assess what they know and what they do not know about the cyber risk they are modeling. What they know is embodied in the model.

What analysts do not know is embodied by the question that reflects uncertainty: "What is really going on here?" The answer to this question requires the SME to change the relationship between inputs and estimated outputs to reflect a new story of how often cyber losses are occurring and how bad they are when they do. We expect that the future of financial cyber risk modeling is a parsimonious combination of established models, such as the Open FAIR model, and additional increasingly sophisticated models that enhance our knowledge of the cyber risk exposure.[36]

One thing is certain: what is parsimonious today will be considered primitive tomorrow. Knowledge about cyber risk and how to effectively model the intersection of cyber risk with other risks, such as credit risk, will certainly grow in the coming years. What is viewed today as being an unknown unknown may be uncovered years from now through techniques such as unsupervised machine learning anomaly detection techniques. Analysts modeling cyber risk need to stay current in their knowledge of what they can reliably estimate and what they cannot in that changing risk modeling landscape. More research is needed, through which cyber risk models will become more accurate, more precise, and more verifiable through back-testing over time. Research and the ongoing exploration of how a model's estimates compare with actual future results form the foundation of a continuous improvement process that increases the model's quality over time. Model development and vetting are not "one-time" activities but processes that conform to the Shewhart-Deming Plan-Do-Check-Act[37] continuous improvement paradigm.

We say this because the arguments that cyber risk cannot be measured sufficiently to pass strict model vetting scrutiny are similar to those that financial risk experts had decades ago about credit risk and market risk. We expect cyber risk modeling to improve over time just as the quality of credit and market risk models have improved over time.

[36] Standards, too, must be parsimonious, which means they must continually improve.

[37] Refer to: https://en.wikipedia.org/wiki/PDCA.

Making Open FAIR Calculations Transparent for Model Vetting Purposes

The Open FAIR Body of Knowledge consists of two standards:

- The Open Group Standard for Risk Taxonomy (O-RT) defines the terms, definitions, and relationships that describe a Bayesian network of risk factors and how they relate probabilistically to each other
- The Open Group Standard for Risk Analysis (O-RA) describes a stylized information Loss Scenario, the major actors in that scenario, the Asset at risk, and the forms of loss the Asset's stakeholder is expected to incur

The Open FAIR standards specify:

- The definition of risk as being "the probable frequency and probable magnitude of future loss"
- A standard, stylized Loss Scenario to describe a potential future loss and how it occurs
- The definitions and probabilistic relationships behind risk factors of that future loss
- A Loss Flow that includes Primary Losses that occur directly from a cyber asset breach as well as the Secondary Losses that sometimes but do not always occur as fallout from the breach
- A standard specification to quantify risk factor estimates and their calibrated estimation process to accurately characterize the uncertainty of those factors
- The assumption that risk factors as random variables satisfy the Independently Identically Distributed (IID) assumption that is foundational to statistical analysis

Neither of the two standards, however, specifically define how calculations to quantify risk should or must be performed. At the time of writing, The Open Group does not specify conformance for how or whether a tool complies with the Open FAIR standards from an algorithmic or calculation standpoint, and the Open FAIR standards are silent on certain decisions tool implementers must make in implementing an Open FAIR risk calculating tool.

We have inserted examples in the following section in order to help make the Open FAIR calculating tool more transparent, and we describe these examples in sufficient detail to enable the reader to follow each calculation in the tables and figures that are provided below. The calculations are based on a practical example that we constructed to illustrate how *The Open FAIR Risk Analysis Tool* (see References) can be used to make cyber risk transparent. The readers of this document, the second white paper in our series, should not interpret the following discussion as the only "right" way to implement the standards; simply as the way a group of SMEs have implemented a tool consistent with what they concluded was a faithful interpretation of the risk factors defined in those standards.

How Losses Occur

The O-RA Standard defines a standard, stylized model of how losses occur, as summarized by Figure 9 (shown earlier as Figure 8 but repeated here for convenience).

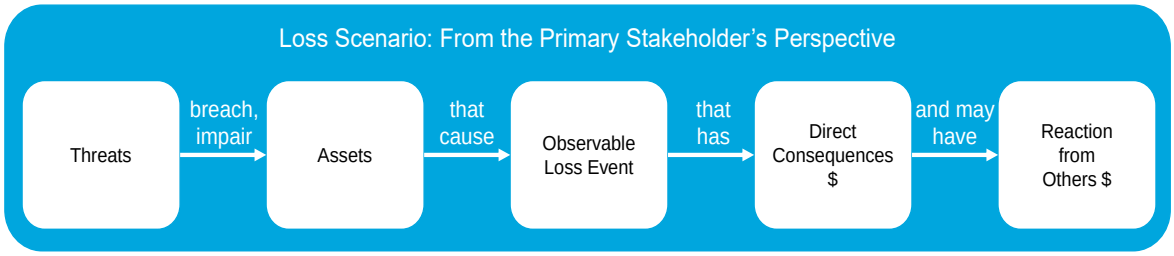


Figure 9: Stylized Loss Scenario in Open FAIR Risk Analysis

In this stylized scenario, a Primary Stakeholder has information Assets that Threat Agents can breach or impair by causing observable confidentiality, availability, or integrity losses to the information Asset. Those losses will have one or more direct economic consequences to the Asset's owner or responsible Primary Stakeholder and may have Secondary Losses to that stakeholder as a consequence of reactions by others.

The O-RT Standard decomposes these Loss Scenarios into two fundamental factors: how often they occur (the Loss Event Frequency (LEF)) and how bad they are when they occur (the Loss Magnitude (LM)). Integrating the joint Loss Event Frequency and Loss Magnitude distributions provides a risk distribution of total economic loss. A Monte Carlo simulation analysis is used to generate common risk metrics, such as average, most likely (i.e., mode), minimum, maximum, percentiles and threshold analytics, cumulative distribution, and VaR statistics in order to enhance decision-making.

Assumptions and Conventions Behind the Open FAIR Risk Model

The Assumption of IID Random Variables

Open FAIR risk factors variables are IID:

- Every Loss Event is independent of each other; Loss Events are not correlated with each other
- The magnitude of each form of loss (productivity, replacement, response, reputation, competitive advantage, and fines and judgment) is independent of another

The Open FAIR Standardization of Parameterization of Uncertainty of Estimates

Every risk factor in the Open FAIR taxonomy is a random variable characterized by three values that express the imperfect knowledge of the analyst or SME estimating that variable in an attempt to estimate the outcome of an uncertain future Loss Event. In particular, the O-RA Standard specifies that every risk factor's estimate consists of three parameters to express a range for the estimate: minimum, most likely, and maximum value.

To estimate those parameters, analysts and SMEs within the organization use their expert judgment, available historical information on similar losses within the organization, and relevant information outside of the organization such as published information breach reports within the organization's industry, available threat intelligence, and analysis of security logs, to inform a calibrated estimate.

Above all, calibrated estimates are meant to be accurate, and analysts are trained to sacrifice precision to accomplish that accuracy. Through the techniques of calibrated estimation, analysts and SMEs should strive for an estimate whose range, specified by the minimum and maximum value, will be found in the future to have been accurate: the actual future value measured should fall within the previously estimated range 90 percent of the time. To achieve that 90 percent accuracy goal, estimators, according to the O-RA Standard,

can complete calibrated estimation training to have a 95 percent confidence in the accuracy of the minimum value of the estimate and a 95 percent confidence in the accuracy of the maximum value of the estimate.

The Analyst's Discretionary Choice of Distribution in Modeling Risk Factors

While the characterization of a risk factor's estimate is standardized (minimum, most likely, and maximum), the statistical distribution of that risk factor is not; instead, being left to the risk analyst or SME to choose based upon best fit to historical data, expert judgment, and any other factors relevant to the specific Loss Scenario being analyzed.[38]

Modeling a risk factor then consists of two components: the standardized range specification of minimum, most likely, and maximum values that will be found to be accurate 90 percent of the time and the discretionary statistical distribution the analyst selects that reflects any additional information the analyst knows about the risk factor being modeled.

Modeling and Decomposing Risk: Loss Event Frequency and Loss Magnitude

This section provides an example of the type of documentation needed to satisfy the risk calculation transparency requirement for model vetting. This example is based upon how *The Open FAIR Risk Analysis Tool* (see References) combines risk factors and performs its Monte Carlo simulation. The risk scenario, estimates, and inputs to the calculations are those found in the first White Paper of this series, *Integrating Cyber Risk with Financial Risk* (see References).

This section has three major calculation sections and takes a top-down view of the risk calculation, with risk factors optionally being described and derived from their respective factors. To help the reader to see how calculations are done, the numbers used in this example were synthesized to clarify how risk factors combine. The numbers used in both the Loss Event Frequency and the Loss Magnitude calculations shown below are taken directly from the example in our first White Paper, *Integrating Cyber Risk with Financial Risk*.

The first, section "Calculating Total Risk from Loss Event Frequency and Loss Magnitude", describes how total risk measured by an annualized loss exposure can be calculated from the simulated Number of Loss Events (NLE) derived from the top-level Loss Event Frequency specification[39] and the calculated single Loss Magnitude derived from the forms of loss specification[40] to arrive at a distribution of simulated total annual loss or risk. This calculation models total risk as the distribution of Monte Carlo simulation trials of annualized total losses.

[38] Necessarily, a calculating tool must convert the standardized minimum, most likely, and maximum characterization of the risk factor's estimate into the parameters needed to specify the distribution that the analyst selects as the most relevant, accurate representation of that risk factor being modeled.

[39] The Loss Event Frequency from the first White Paper of this series, *Integrating Cyber Risk with Financial Risk*, is a triangular distribution with a Min of 0.2, Most Likely of 0.5, and Max of 1.0; see p.25 and Table 1.

[40] On p26-27 of the first White Paper of this series, *Integrating Cyber Risk with Financial Risk*, the forms of loss are specified (primary response costs, secondary response costs, secondary fines, and judgment costs, with a Secondary Loss Event Frequency.

The second, section "Modeling Loss Event Frequency from its Subfactors", shows how the number of Loss Events simulated is calculated from estimates made at any level within the Open FAIR Loss Event Frequency taxonomy. Within this section:

- Section "Modeling Loss Event Frequency Directly from a Minimum, Most Likely, and Maximum Estimate" shows how the NLE can be calculated from the distribution of Loss Event Frequency
- Section "Modeling Loss Event Frequency through its Subfactors Threat Event Frequency and Vulnerability" shows how estimates of Threat Event Frequency (TEF) and Vulnerability (Vuln) can be combined to form a calculated distribution Loss Event Frequency
- Section "Modeling Threat Event Frequency through its Subfactors Contact Frequency and Probability of Action" shows how estimates of Contact Frequency (CF) and Probability of Action (PoA) can be combined to form a calculated distribution of Threat Event Frequency
- Section "Modeling Vulnerability through its Subfactors Threat Capability and Asset Resistance Strength" shows how estimates of Threat Capability (TCap) and Resistance Strength (RS) can be combined to form a calculated distribution of Vulnerability

The numbers used in the tables were synthesized to tie to each other to aid the reader in understanding how risk factors can be combined, which introduces a necessary fiction on how triangular distributions combine. All distributions of risk factors in our example are modeled as triangular distributions for ease of explanation. In this case where two triangular distributions are sampled, such as in the section "Modeling Loss Event Frequency through its Subfactors Threat Event Frequency and Vulnerability" we can see that Threat Event Frequency and Vulnerability are combined to form Loss Event Frequency. Observe that each Monte Carlo simulation trial Threat Event Frequency and Vulnerability sample is multiplied to form a Loss Event Frequency trial Poisson mean. In this case to make the numbers tie, or equal to the Loss Event Frequency used in the first White Paper of this series, *Integrating Cyber Risk with Financial Risk*, and in the section "Calculating Total Risk from Loss Event Frequency and Loss Magnitude", the Threat Event Frequency and Vulnerability Monte Carlo simulation trial values were selected so that they equal the Loss Event Frequency when multiplied by one another. This artificial selection creates the fiction that two Monte Carlo simulation triangular distributions multiplied together produces a triangular distribution. In reality, they do not, but suspending that reality helps make the calculations clearer so that the reader can follow how risk factors within the taxonomy can combine mathematically.

The third, section "Modeling Loss Magnitude", shows how the Loss Magnitude can be calculated from its factors, forms of loss resulting in Primary Loss, Secondary Loss Event Frequency, and forms of loss resulting Secondary Loss. This section takes the same numbers and distributions used in the example in the first White Paper of this series, and elaborates how the parameters specifying distributions of those loss factors are combined to form the Loss Magnitude distribution of single total losses. That distribution is used in calculating total risk in the section "Calculating Total Risk from Loss Event Frequency and Loss Magnitude".

Calculating Total Risk from Loss Event Frequency and Loss Magnitude

Risk at its highest level of abstraction consists of:

- How often bad things happen
- How bad they are when they do happen

The Open FAIR taxonomy uses input risk factor terms such as Loss Event Frequency and Loss Magnitude. Each of these terms is a random variable, and the calculated combination is a distribution of total loss, dependent upon the two inputs.

Each trial of the Monte Carlo simulation generates the NLE within a year and the total Loss Magnitude of one simulated Loss Event. Output statistics are generated from a set of simulated trials; 5000 trials in this example.

$$Calculated\ Risk\ in\ a\ Trial = \sum_{i=1}^{NLE} Ramdomly\ Selected\ Single\ Loss\ Magnitude_i$$

Equation 1: Calculated Risk in a Trial

Suppose a 5,000 trial Monte Carlo simulation generated the NLE in a single year that looked like that shown in Table 6.

Table 6: Monte Carlo Simulation of Loss Events in One Year

Trial	NLE (Events in One Year)
1	0
2	0
3	1
4	0
…	
5,000	2

Suppose a 5,000 trial Monte Carlo simulation of the Loss Magnitude formed an "urn" of 5,000 single total losses from which to draw the Loss Magnitude of a single loss, as shown in Table 7. When drawing multiple times, this is an urn "with replacement". Every simulated outcome satisfies the IID assumption.

Table 7: Monte Carlo Simulation of the Loss Magnitude (LM) of a Single Loss Event

Trial	Total Single LM ($)
1	126,200
2	122,900
3	54,800
4	1,368,300
…	
5,000	173,700

Total Risk integrates Loss Event Frequency and Loss Magnitude via a Monte Carlo simulation to arrive at a total simulated loss.

Table 8: Monte Carlo Simulation of Total Risk where the Loss Event Frequency and the Loss Magnitude of a Single Loss are Random Variables Described by Their Respective Monte Carlo Simulation

Simulation Year (Trial)	Total Risk in Simulation Year ($)
1	0
2	0
3	122,900[41]
4	0
…	
5,000	1,161,600 (2 Loss Events, 2 drawn LMs)[42]

As can be seen in Table 8, simulation trial year 5,000 had two Loss Events simulated and, therefore, the total risk consists of the sum of two losses drawn at random from the Loss Magnitude simulation in Table 7. Every other simulated year would have its own draw from the Loss Magnitude table of the number of losses in the Loss Event Frequency table for that simulated year and then add up those losses to arrive at the total loss for the simulated year. The process is similar to performing a random draw from an urn of 5,000 single Loss Magnitudes (with replacement) of the number of Loss Events in one year and adding those drawn Loss Magnitudes together.

Risk is expressed as a distribution of these simulated year trials and statistics of that distribution. Most common representations derived from the Monte Carlo simulation are loss exceedance curves, VaR calculations, averages, medians, modes, minimums, maximums, and so on.

Modeling Loss Event Frequency from its Subfactors

According to the stylized Loss Scenario diagram in Figure 8, Threat Agents breach or impair Assets to cause an observable confidentiality, integrity, or availability loss of the information Asset. Open FAIR models those events as a sequence:

1. First the Threat Agent contacts the Asset. How often a Threat Agent contacts the Asset is the Contact Frequency.
2. After contact, a conscious Threat Agent must decide whether to try and breach or impair that Asset. The Threat Agent has a Probability of Action, the probability that a contact will result in an attempt to cause harm, or a Threat Event, on the Asset. Not all contacts are attempts to cause harm.
3. Assets can resist Threat Events through their control strength, or Resistance Strength. When the Resistance Strength exceeds the Threat Capability applied against the Asset – that is, the resources,

[41] The single Loss Magnitude of Trial 2 happened to have been drawn.

[42] The two Loss Magnitudes drawn here were from trials other than Loss Magnitude Trials 1-4 and 5,000.

skill, and strength the Threat Agent brings to bear on the Asset – then the Threat Agent is unsuccessful in penetrating the Asset. When the Threat Capability exceeds the Resistance Strength, the Threat Agent's attempt to breach or impair the Asset succeeds. Vulnerability in the Open FAIR model is the probability that the Threat Capability exceeds the Resistance Strength. Vulnerability is also the probability that a Threat Event results in a Loss Event.

Analysts have several choices in how they model how often those adverse Loss Events occur. They can model them directly by estimating the Loss Event Frequency itself, or they can break down the Loss Event Frequency into its subfactors. Figure 10 gives a graphical view of how Loss Events unfold.

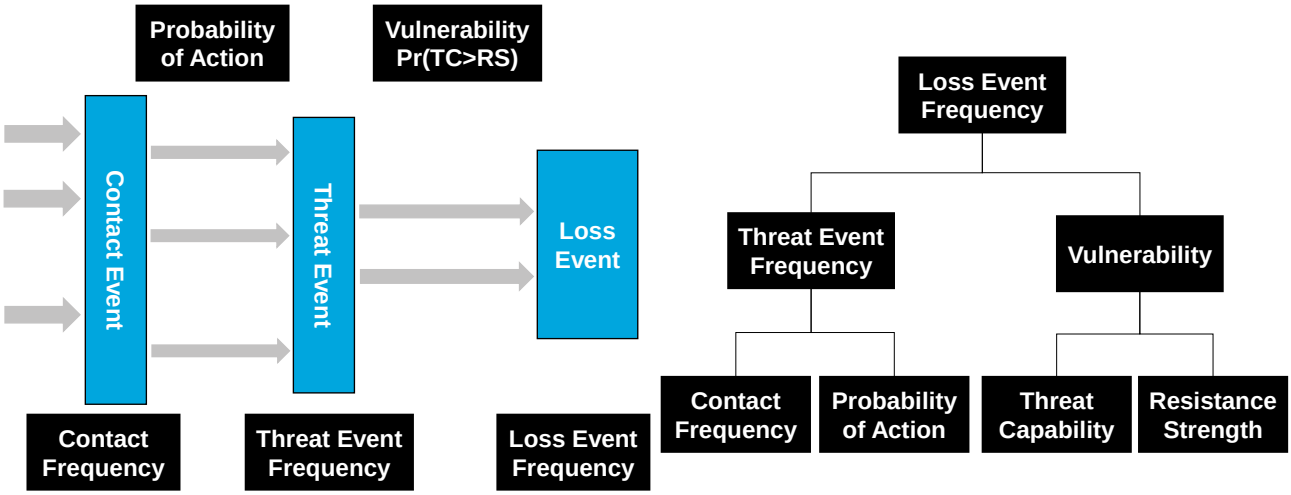


Figure 10: Open FAIR Risk Factors Preceding a Loss Event

$$Vuln = Pr(Loss\ Event\ |Threat\ Event) = Pr(TCap > RS)$$

$$PoA = Pr(Threat\ Event\ |Contact\ Event)$$

$$LEF \leq TEF \leq CF$$

Equation 2: Key Relationships that Model Loss Event Frequency (LEF)

The Loss Event Frequency can be modeled directly as a distribution, or can be derived from its subfactors, Threat Event Frequency and Vulnerability, or can be derived from the subfactors of Threat Event Frequency and/or Vulnerability.[43] In the example that follows, all estimates of risk factors are characterized by a minimum, most likely, and maximum value. All distributions shown in the example are triangular.

[43] The subfactors for Threat Event Frequency are Contact Frequency and Probability of Action. The subfactors for Vulnerability are Threat Capability and Resistance Strength.

Modeling Loss Event Frequency Directly from a Minimum, Most Likely, and Maximum Estimate

As is standard in the Open FAIR taxonomy, all directly estimated risk factors – that is, risk factors not derived from their subfactors – are specified as a range with a minimum, most likely, and maximum value. If an analyst specified Loss Event Frequency [44] directly, say from a combination of an historical record of losses and expert judgment to estimate the next year's Loss Event Frequency, the analyst might specify the Loss Event Frequency, as shown in Table 9.[45]

Table 9: Parameterizing, Specifying Loss Event Frequency (LEF)

LEF	Estimate
Minimum	0.2 (once every two years)
Most Likely	0.5 (once every other year)
Maximum	1

The calculator tool generates integral numbers of Loss Events to simulate, say, 5,000 Monte Carlo simulation trials. The analyst would have to specify a distribution that corresponds to these input values and generates an integer value for them: fractional Loss Events in a simulated Monte Carlo simulation trial do not make sense.

In *The Open FAIR Risk Analysis Tool*, the calculations take place in two steps. The initial estimate is a triangular distribution of Min and Max with its peak (mode) at the Most Likely value as shown in the probability density function in Figure 11.

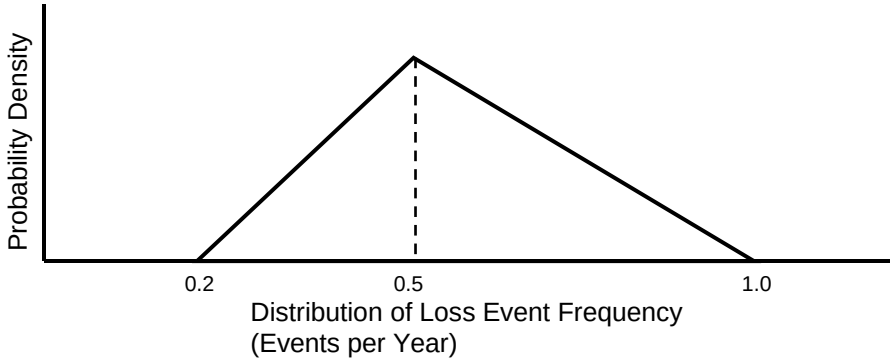


Figure 11: Probability Density Function of a Triangular Distribution: Minimum 0.2, Most Likely 0.5, Maximum 1.0

[44] The analyst also must interpret the exact meaning of Loss Event Frequency. It could the actual number of Loss Events per year, which has to be an integer, or it could be the probability of a Loss Event occurring per unit time, a real number. This document, in using the Poisson distribution, takes this later interpretation. To express the uncertainty in the probability of a Loss Event occurring within a time period, the analyst here is specifying a triangular distribution, characterized by the minimum, most likely, and maximum value of the Poisson mean λ.

[45] Note that the parameters in Table 9 are the same as those in our first White Paper, *Integrating Cyber Risk with Financial Risk*.

The Open FAIR Risk Analysis Tool utilizes the output from a 5,000 trial Monte Carlo simulation of this distribution to generate 5,000 simulated means of a Poisson Distribution. By choosing to use this tool, the analyst is aware they are expressing uncertainty in estimating the mean frequency of a Poisson Distribution, λ, through the triangular distribution shown in Table 10.

Table 10: Monte Carlo Simulation or 5,000 Trials from the Triangular Distribution Describing Loss Event Frequency

Trial	**λ**
1	0.45
2	0.6
3	0.53
4	0.32
…	
5,000	0.72

The distribution of 5,000 Poisson means is then each sampled once to arrive at a Loss Event Frequency estimate in another Monte Carlo simulation, shown in Table 11.

Table 11: Monte Carlo Simulation of the Number of Loss Events (NLE) from a Poisson Mean, Itself a Random Variable

Trial	**λ (From Prior calculation)**	**NLE (From One Trial of Poisson Distribution)**
1	0.45	0
2	0.6	0
3	0.53	1
4	0.32	0
…		
5,000	0.72	2

As discussed, the triangular distribution to characterize the Loss Event Frequency estimate and the Poisson distribution to generate integral Loss Events were used in our example but other analysts, using other tools, could have made other choices on how to best model and calculate the Loss Event Frequency from its subfactors.

Modeling Loss Event Frequency through its Subfactors Threat Event Frequency and Vulnerability

The Open FAIR taxonomy describes Loss Event Frequency as a composition of two subfactors, Threat Event Frequency and Vulnerability, as shown in Figure 12. Threat Event Frequency is the estimated number of attempts a Threat Agent makes to try and cause harm as described in the Loss Scenario. Vulnerability is the probability that the attempt to cause harm actually does; that is, the probability that a Threat Event becomes a Loss Event. Threat Event Frequency and Vulnerability are assumed to be statistically independent of each other.

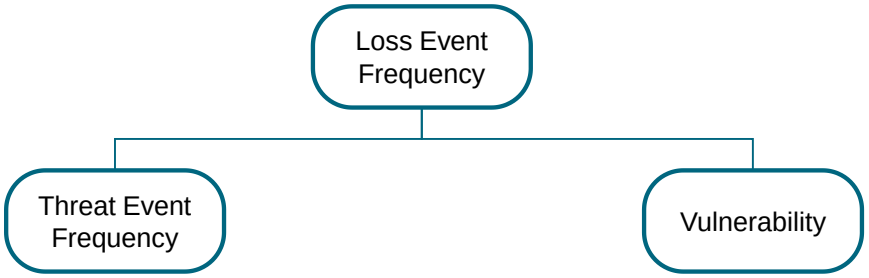


Figure 12: Loss Event Frequency Decomposed into Threat Event Frequency and Vulnerability

In *The Open FAIR Risk Analysis Tool*, an analyst can describe Threat Event Frequency as a triangular distribution of Threat Events per year specified by the parameters of minimum, most likely, and maximum. Example distributions for Threat Event Frequency and Vulnerability are shown in Figure 13.

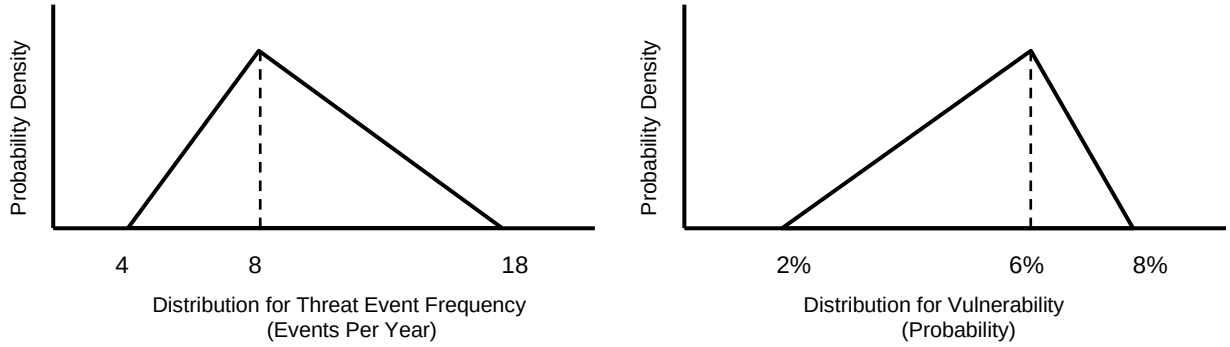


Figure 13: Loss Event Frequency Modeled as Distributions of its Subfactors

The Open FAIR Risk Analysis Tool samples each distribution 5,000 times, and multiplies together the Threat Event Frequency and Vulnerability to arrive at an Loss Event Frequency input as if it were inputted directly without the decomposition, as a calculated substitute for the Loss Event Frequency input into the model shown in Table 10 and Table 11. The calculated Loss Event Frequency is then used as a mean to the Poisson Distribution to generate an integral number of Loss Events for each of the 5,000 Monte Carlo simulation trials. See Table 12 for an example of the process.

Note: In this example, the Loss Event Frequency and Loss Event numbers are the same as in Table 11 to show how they could arrive from calculations derived from the subfactors Threat Event Frequency and Vulnerability.

Table 12: Monte Carlo Simulation of Loss Event Frequency (LEF) and Loss Events from Threat Event Frequency (TEF) and Vulnerability (Vuln)

Trial	TEF (from distribution)	Vuln (from distribution)	LEF = TEF × Vuln (Poisson λ)	NLE (From one trial of Poisson Distribution)
1	6.82	6.59%	0.45	0
2	15.38	3.90%	0.6	0
3	8.28	6.40%	0.53	1
4	4.57	7.0%	0.32	0
…				
5,000	11.07	6.50%	0.72	2

Modeling Threat Event Frequency Through its Subfactors Contact Frequency and Probability of Action

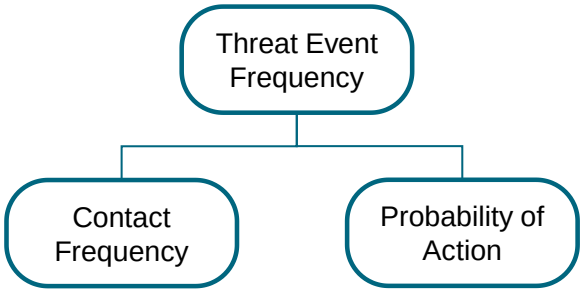


Figure 14: Threat Event Frequency Decomposed into Contact Frequency and Probability of Action

Like the relationship of Loss Event Frequency to Threat Event Frequency and Vulnerability, Threat Event Frequency can be decomposed into its subfactors Contact Frequency and Probability of Action. *The Open FAIR Risk Analysis Tool* calculates Threat Event Frequency from Contact Frequency and Probability of Action in the same way as it calculates Loss Event Frequency from Threat Event Frequency and Vulnerability as shown in the example below.

The Monte Carlo simulation samples the Contact Frequency distribution, shown in Figure 15, 5,000 times and puts those results into the Contact Frequency trial column in Table 13. Then the Probability of Action distribution in Figure 15 is sampled and put into the Probability of Action column in Table 13, and the results are multiplied together to arrive at a Threat Event Frequency distribution which is then further calculated upon to arrive at Loss Event Frequency as shown in the previous section. Contact Frequency and Probability of Action are assumed to be statistically independent of each other.

Note: In this example the Threat Event Frequency numbers are the same as in Table 12 to show how they could arrive from calculations derived from the subfactors Contact Frequency and Probability of Action.

Suppose that the Contact Frequency and the Probability of Action were modeled with the triangular distributions shown in Figure 15.

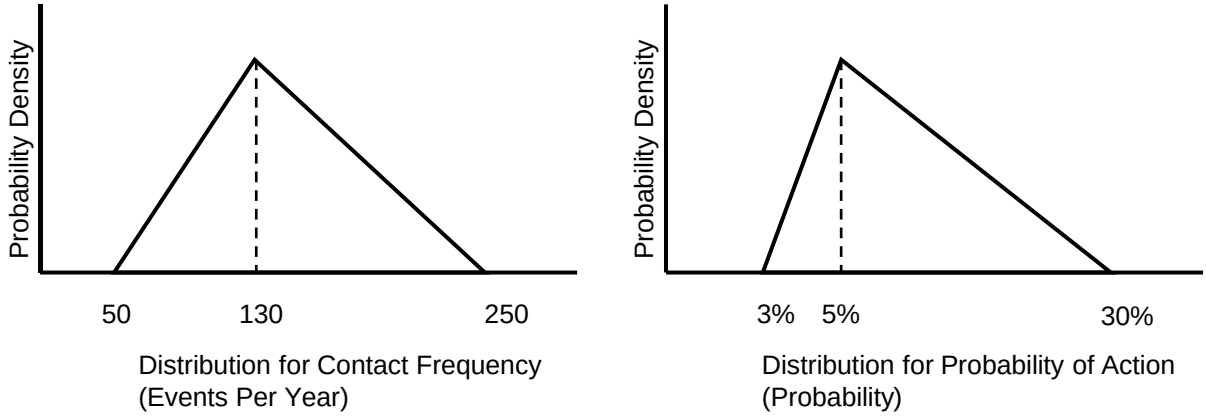


Figure 15: Example Distributions for Threat Event Frequency Subfactors Contact Frequency and Probability of Action

Table 13: Monte Carlo Simulation of Loss Event Frequency and Loss Events from Threat Event Frequency and Vulnerability

Trial	CF (from distribution)	PoA (from distribution)	TEF = CF × PoA
1	125.2	5.4%	6.82
2	107.4	14.3%	15.38
3	163.1	5.1%	8.28
4	129.3	3.5%	4.57
…			
5,000	229.7	4.8%	11.07

Modeling Vulnerability through its Subfactors Threat Capability and Asset Resistance Strength

Analysts can decompose Vulnerability into its subfactors Threat Capability and the Asset's Resistance Strength. These factors are comparatively abstract as compared to the other risk factors, so a bit of elaboration is warranted here.

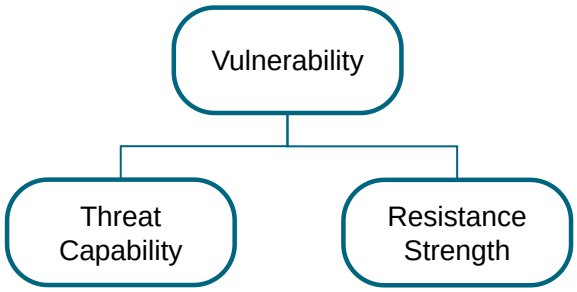


Figure 16: Vulnerability Decomposed into Threat Capability and Resistance Strength

Each Threat Agent within a broader threat community has its own Threat Capability, its own unique skill, resources, organization, etc. as compared to others within a larger Threat Community. A Threat Community consists of Threat Agents who share similar attributes, such as motive, sponsorship of or membership in an organization, personal risk appetite, and willingness to commit collateral or social damage. Each Threat Agent's Threat Capability can be ranked as a percentile of all Threat Agents within a given Threat Community. In other words, the Threat Capability used in a cyberattack varies, is represented as a random variable, and is measured as the percentile rank of the Threat Agent's Threat Capability within the broader Threat Community. The threat community is completely described by the 0-100 percentile range, and the particular capability of the Threat Agent that is most likely to try to breach or impair the Asset is described as a percentile range as subset of that community.

Similarly, the ability of the Asset to resist Threat Agent attempts to breach or impair that Asset falls on a spectrum. Not all information Assets within the scope of the Loss Scenario are protected equally, or the analyst may not have precise knowledge of the strength of those controls. Asset Resistance Strength is measured as the percentile of the threat community that the Asset's controls will successfully resist.

A Threat Agent succeeds when the Threat Capability applied in the specific Threat Event exceeds the Resistance Strength of the Asset, where TCap > RS.

Threat Capability, then, can be modeled as a distribution of the percentile of the threat community that is likely to try to breach or impair the information Asset, and Resistance Strength, similarly, can be modeled as a distribution of the percentile of the Threat Community that the Asset can resist. Vulnerability is the probability distribution of Threat Agent success, where TCap > RS. Threat Capability and Resistance Strength are assumed to be statistically independent of each other.

As an example, suppose an analyst and SMEs believe that their most likely Threat Agents are well-funded state-sponsored adversaries or well-organized criminals who are trying to commit a ransomware attack for financial gain. The analyst and SMEs believe that these Threat Agents fall within the top 80^{th} to 100^{th} percentile of the financially motivated threat community.

In estimating the Resistance Strength of their Asset, the analyst and SMEs estimate that their state-of-the-art controls are almost always effective; i.e., able to resist the Threat Capability of somewhere between the 90^{th} to 99^{th} percentile of Threat Agents; see Table 14 and Figure 17 for the distributions of Threat Capability and Resistance Strength used in this example.

The Monte Carlo simulation modeling of the Threat Capability and Resistance Strength in *The Open FAIR Risk Analysis Tool* would look like as shown in Table 14.

Table 14: Specification of Triangular Distributions for Threat Capability (TCap) and Resistance Strength (RS)

	TCap	**RS**
Minimum	80	90
Most Likely	87.5	94.5
Maximum	95	99

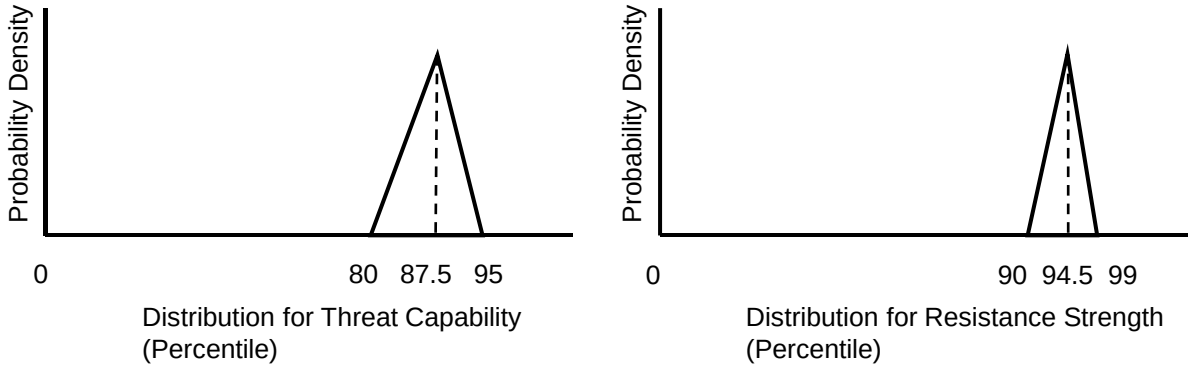


Figure 17: Example Distributions for Vulnerability Subfactors Threat Capability and Resistance Strength

Table 15: Monte Carlo Simulation Deriving Vulnerability (Vuln) from Threat Capability (TCap) and Resistance Strength (RS)

Trial	TCap (from distribution)	RS (from distribution)	TCap > RS? 1=Yes 0=No	Vuln Pr(TCap > RS) (average)
1	87	90	0	7.3%
2	89	92	0	
3	92	91	1	
4	93	97	0	
…				
5,000	94	93	1	

Note: The developers of *The Open FAIR Analysis Tool*, instead of implementing a complicated calculation and simulation to generate a distribution for Vulnerability consistent with what has been presented in Table 12, compromised to generate one number, the average Vulnerability, that is used in each entry in that Vulnerability column as a substitute for Vulnerability in this particular tool. This means that 7.3% would be the entry in every cell in the Vulnerability column in Table 12. This is an example how tool makers make choices in developing their tools for specific client audiences, industry segments, or other applications, trading off at times complexity for reduced cost of development.

Modeling Loss Magnitude

Loss Magnitude consists of four high level risk factors: Primary Loss Magnitude, Secondary Loss, Secondary Loss Event Frequency, and Secondary Loss Magnitude. See Figure 3 and Figure 18 for the relationships of these terms.

Primary Loss Magnitude describes losses that occur with each instance of the Loss Scenario and is measured in dollars. Secondary Losses sometimes occur and are conditional upon the Primary Loss occurring in the first place. The probability of a Secondary Loss occurring is the Secondary Loss Event Frequency, measured in probability. The magnitude of the Secondary Loss are those losses that can occur, measured in dollars.

Loss Magnitude is the sum of the Primary Loss Magnitude and the conditional Secondary Loss if it occurs. Because the Secondary Loss only sometimes occurs, in a Monte Carlo simulation, the Secondary Loss may be zero while the Primary Loss Magnitude will always be positive, non-zero.

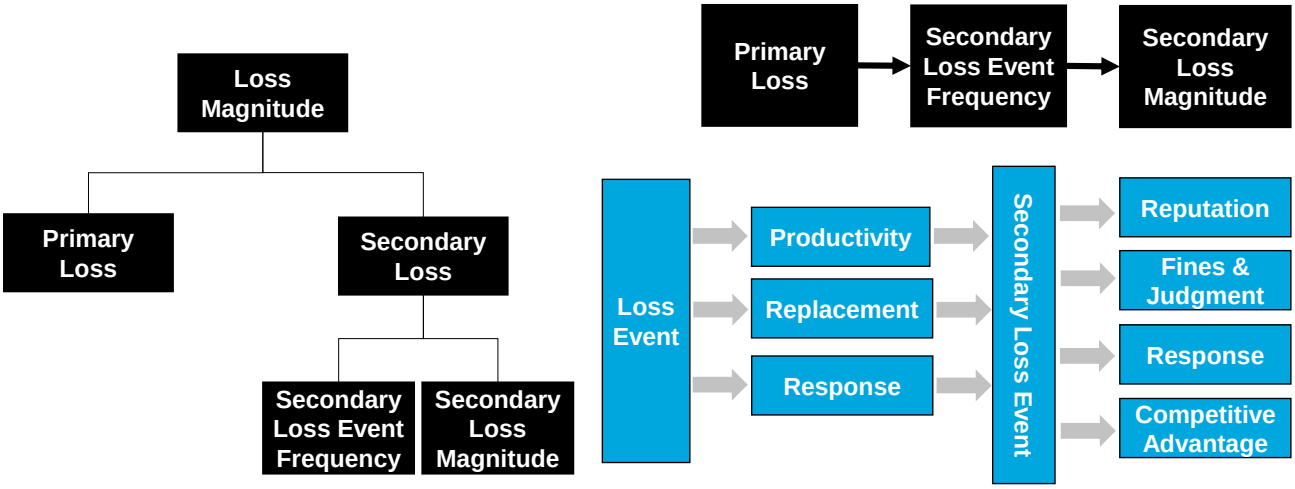


Figure 18: Open FAIR Risk Factors Modeling Loss Magnitude Expressed as Primary and Secondary Loss

Each loss (Primary and Secondary) is embodied by one or more of the six standard forms of loss as shown in List 1.

1. Productivity
2. Response
3. Replacement
4. Reputation
5. Competitive advantage
6. Fines and judgments

List 1: Forms of Loss

Essential formulas relating to these terms are:

$$LM = PLM + SL$$

$$PLM = \sum_{i=1}^{6} Primary\ Loss\ Form_i$$

$$Secondary\ Loss = \begin{Bmatrix} (0 \mid Pr(1 - SLEF)) \\ (SLM \mid Pr(SLEF)) \end{Bmatrix}$$

$$SLM = \sum_{i=1}^{6} Secondary\ Loss\ Form_i$$

$$SLEF = \Pr\left(Secondary\ Loss \mid Primary\ Loss\right)$$

($_i$ refers to forms of loss in List 1)

Equation 3: Essential Formulas

Each of the Loss Magnitude risk factors, Primary Loss Magnitude and its forms of loss, Secondary Loss Magnitude and its forms of loss, and Secondary Loss Event Frequency are all estimates expressed through a minimum, most likely, and maximum value, and a distribution. Analysts use the O-RA Standard convention to characterize these estimates with minimum, most likely, and maximum values. The distribution is not standardized and left to the analyst and tool vendor to provide in the model. Each form of loss is statistically independent of each other.

Modeling Loss Magnitude: An Example

In the example that follows, *The Open FAIR Risk Analysis Tool* models each primary and secondary form of loss as a triangular distribution specified by a minimum, a maximum, and a most likely value and a Secondary Loss Event Frequency as a triangular distribution of a Binomial Distribution mean with one trial. In this way, the tool performs a Monte Carlo simulation to generate trials of Primary and Secondary Loss Magnitudes, and the probabilities of Secondary Loss Event Frequency.

Modeling Primary Loss Magnitude

Each form of loss is specified as a range with a minimum, most likely, and maximum value. *The Open FAIR Risk Analysis Tool* used that characterization as a specification for a triangular distribution of the form of loss.

Suppose an analyst is modeling the Primary Loss of a typical PII data breach loss as consisting of only Response costs. The other forms of loss – productivity, replacement, reputation, competitive advantage, and fines and judgments – are not applicable as Primary Losses to the Loss Scenario. The analyst specifies the calibrated estimate as shown in Table 16.

Table 16: Calibrated Estimate for Primary Loss Magnitude (PLM) Forms of Loss

PLM	Response ($)
Minimum	30,000
Most Likely	100,000
Maximum	200,000

These parameters specify a triangular distribution in Figure 19 for the Response cost.

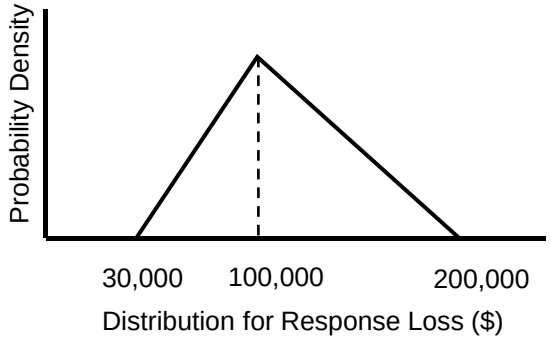


Figure 19: Triangular Distributions Describing Primary Losses for Response Costs

The Open FAIR Risk Analysis Tool generates Monte Carlo simulation Trials of this distribution to arrive at a Monte Carlo simulation of the total Primary Loss Magnitude for the Loss Scenario.

Table 17: Primary Loss Magnitude (PLM) Monte Carlo Simulation

Trial	Response Loss ($)	Total PLM ($)[46]
1	126,200	126,200
2	122,900	122,900
3	54,800	54,800
4	113,400	113,400
…		
5,000	173,700	173,700

Modeling Secondary Loss Event Frequency

The analyst now estimates Secondary Loss Event Frequency, the probability that a Primary Loss will result in a Secondary Loss. That estimate in *The Open FAIR Risk Analysis Tool* is specified as a triangular distribution as shown in Table 18 and Figure 20.

Table 18: Secondary Loss Event Frequency (SLEF) Estimate

SLEF	Probability
Minimum	0.2
Most Likely	0.3
Maximum	0.5

[46] If there were more than just Response costs to consider, then each form of loss (Productivity, Replacement, etc.) would be modeled as an independent distribution, with a Monte Carlo simulation generated as shown for the Response cost. The total would be the simple sum of each Monte Carlo simulation trial row. An example of multiple loss forms is seen in Section Modeling Secondary Loss Magnitude; see Table 21.

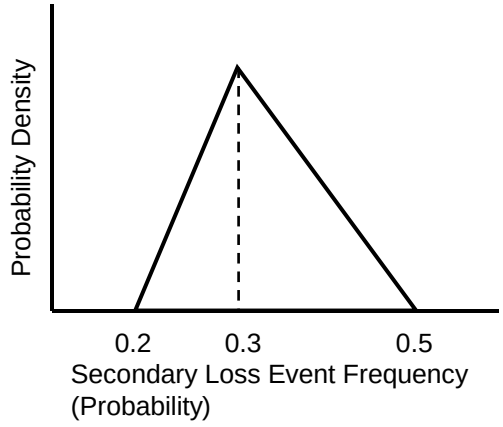


Figure 20: Secondary Loss Event Frequency Triangular Distribution Representation

Next, *The Open FAIR Risk Analysis Tool* samples the Secondary Loss Event Frequency triangular distribution for a value. That value is used as a binomial mean for one trial to simulate whether a secondary Loss Event occurred (1) or did not occur (0), as shown in Table 19.

Table 19: Monte Carlo Simulation of Secondary Loss Events

Trial	Binomial Mean Sampled from SLEF Distribution	Binomial Test: 1 Indicates a Secondary Loss Occurred in that Trial
1	0.28	0
2	0.23	0
3	0.33	0
4	0.38	1
…		
5,000	0.41	0

Only Trial 4 had a Secondary Loss whose magnitude would be added to that trial's Primary Loss.

Modeling Secondary Loss Magnitude

Using the same process as modeling Primary Loss Magnitude, the analyst parametrizes the minimum, most likely, and maximum estimates and selects a distribution to model the uncertainty of the Secondary Loss Magnitude should it occur. *The Open FAIR Risk Analysis Tool* only offers the triangular distribution, so that is the one used in the example below.

Suppose an analyst is modeling two secondary forms of loss to estimate the impact of a confidentiality loss of proprietary information:

- Response costs
- Fines and judgments

The other forms of loss do not apply to this Loss Scenario example:

- Productivity
- Replacement
- Reputation
- Competitive advantage

The analyst specifies the estimate for these two forms of loss as shown in Table 20.

Table 20: Parameters Describing the Uncertainty of the Loss Magnitude of Two Forms of Secondary Losses

SLM	Response ($)	Fines & Judgments ($)
Minimum	15,000	1,000,000
Most Likely	25,500	1,200,000
Maximum	60,000	1,500,000

These parameters specify two triangular distributions, one for each of the forms of loss. Each form of loss is IID.

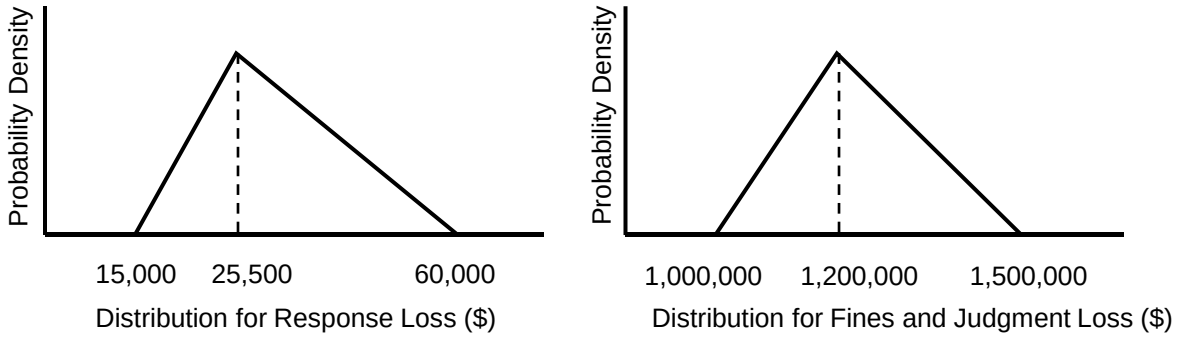


Figure 21: Triangular Distributions Describing Secondary Losses for Response and Competitive Advantage Costs

The *Open FAIR Risk Analysis Tool* provides Monte Carlo simulation Trials of each distribution and adds them together to arrive at a Monte Carlo simulation of the total Secondary Loss Magnitude for the Loss Scenario, as shown in Table 21.

Table 21: Secondary Loss Magnitude (SLM) Monte Carlo Simulation

Trial	Response Loss ($)	Fines & Judgments Loss ($)	Total SLM ($) (Response + Fines & Judgments Losses)
1	37,700	1,281,000	1,318,700
2	36,700	1,270,700	1,307,400
3	19,900	1,071,800	1,091,700

Trial	Response Loss ($)	Fines & Judgments Loss ($)	Total SLM ($) (Response + Fines & Judgments Losses)
4	31,800	1,223,100	1,254,900
…			
5,000	25,000	1,147,600	1,172,600

Modeling Total Single Loss Magnitude

The Loss Magnitude of a single Loss Event is the sum of the Primary Loss Magnitude and Secondary Loss Magnitude, if any, for a simulated loss trial. In Table 22, only Trial 4 has an SL (from the Secondary Loss Event Frequency Monte Carlo simulation, shown in Table 19, so the Secondary Loss Magnitude for Trial 4 is added to the Primary Loss Magnitude for that trial to arrive at the total Loss Magnitude for that trial). All other Loss Magnitude Trials, shown in Table 22, consist only of Primary Loss Magnitude.

Table 22: Total Single Loss Magnitude Monte Carlo Simulation

Trial	PLM ($)	Secondary Loss Event? 1=yes 0=no	SLM ($)	Total Single LM ($)
1	126,200	0	1,318,700	126,200
2	122,900	0	1,307,400	122,900
3	54,800	0	1,091,700	54,800
4	113,400	1	1,254,900	1,368,300
…				
5,000	173,700	0	1,172,600	173,700

Summary

Quantifying cyber risk using the Open FAIR standards requires the analyst to use all information available to make accurate, forward-looking estimates of the minimum, most likely, and maximum values of the selected risk factors that characterize the probable frequency and probable magnitude of future loss of an analyzed Loss Scenario. All risk factors are modeled as IID random variables.

Risk analysis tool vendors offer statistical distributions that can be characterized with these estimates, and analysts select the distribution that best reflects what the analyst knows about the risk factor being estimated as well as reflects the analyst's uncertainty. In this example, the triangular distribution was used throughout, but it did not have to be.

In implementing a risk calculating tool, tool vendors interpret the Open FAIR standards to combine risk factors mathematically and use Monte Carlo simulation to generate results that reflect the model, faithfully

represent the Loss Scenario, and decompose it into its Loss Event Frequency and Loss Magnitude risk factors, and calculate risk.

The Open FAIR Risk Analysis Tool is but one example of a set of choices a tool vendor could make. The example that we used in this document shows how Loss Event Frequency can be calculated from a triangularly-distributed estimate of Loss Events per year that is interpreted to reflect the mean of a compound Poisson distribution. The Loss Magnitude of a single Loss Event was decomposed into the Primary Loss based on a triangular distribution of losses that always occur, and a conditional Secondary Loss, also based on triangular distributions of losses that occur with the Secondary Loss Event Frequency probability.

Conclusion

Vetting cyber risk models in the financial services sector requires assurance that models are transparent and predictive. In this document, we describe how cyber risk models can be shown to be parsimonious.

We have shown how cyber risk model accuracy can be improved from a combination of estimates that are based on SME opinion and detailed historical data. Toward this end, the challenge is to reflect both what is known about the risk scenario and what is uncertain.

We described how risk analysts and cyber practitioners can work together to vet critical aspects of the Open FAIR taxonomy and show that the Open FAIR model is designed to achieve parsimony. We also described how the terms, definitions, and relationships among risk factors provide a measure of transparency in how risk and risk scenarios are defined and measured.

This document expands on the standards in the Open FAIR taxonomy and provides an example of how to document a model's calculations so that they are transparent to those charged with vetting models. All risk models in an FI require this level of transparency. We have provided a practical example to illustrate key points.

Risk models must be robust enough to be used under normal, adverse, and severely adverse conditions. This document provides an example of how to apply assumptions and data to cyber risk modeling under each of these conditions.

Back-testing cyber risk models remains a challenge and is a subject for further research. More information is being published on the Loss Event Frequency and Loss Magnitude of cyber losses in different industry segments. We have described how to achieve parsimony with better and more robust risk models by incorporating new information into cyber risk factor estimates. As time passes, we expect to improve what is known about cyber risk, the frequency and magnitude of cyber losses, and the environmental factors that influence cyber risk.

In combination, the two-paper "*Calculating Reserves for Cyber Risk*" series demonstrates how organizations can evaluate and measure cyber risk in economic terms. The implications of this are profound:

- Boards of Directors and senior management have been directed to improve their enterprise-wide cyber risk intelligence

  Measuring cyber risk in economic terms enables management to treat cyber risk as an enterprise-wide risk. Boards and senior management now have a tangible, practical way of performing their fiduciary and regulatory compliance duties in managing this risk. This series demonstrates how to make cyber risk fungible with other financial risks such as market and credit risk, giving directors and senior management a holistic view into their enterprise's risk posture.

- Cybersecurity controls provide tangible, defensible benefits to the enterprise

  The reduced cyber risk derived from those controls can be estimated and quantified as an economic benefit. The language and economics of risk management give CISOs, CROs, and senior management the language to discuss and resolve the common problem of effective cyber risk management. Cybersecurity is not an end in itself. It is the means to accomplish effective cyber risk management.

- As IT grows in strategic importance for any enterprise, the techniques discussed in this two-paper series will apply to a broad range of industries

  Stakeholders such as regulators, shareholders, customers, and suppliers across all industries, expect that any enterprise is resilient to the impact of all major risks, including cyber risk. Resiliency includes having sufficient reserves to cover losses from these risks. In the future, enterprises across all industries, not just the financial sector, should be able to calculate and hold sufficient reserves for cyber risk.

- All organizations need to sufficiently trust the cyber risk models that they are using for the decisions those models are informing; that is, all cyber risk models must be fit for purpose

  Validating cyber risk models includes examining the quality of the data, the processes used to construct and analyze risk scenarios, the correctness of the risk model being used, and the accuracy of the modeled results.

Risk managers and cyber risk practitioners are working together to achieve cyber risk model parsimony. Keeping parsimony as an objective, risk modelers and their vetters are striving to integrate historical data with SME judgment to arrive at the best, most defensible risk model parameter estimates. More research is needed on how to combine data and SME judgment when conducting cyber risk analyses.

Taken together, this two-paper series provides any enterprise, not just FIs, with a roadmap that harmonizes cyber risk with other risks in a parsimonious manner and leads to cyber risk models that are continuously improving and fit for purpose.

Appendix: Scoping a Project to Calculate Reserves for Cyber Risk

The checklist summary provided below in Table 23 and Table 24 is extracted from both of the documents in our White Paper series, and can be used as an initial scope to calculate reserves for cyber risk. This checklist is structured as a series of questions that when answered leads to the initial information required to start a project to calculate reserves for cyber risk. No doubt any organization that undertakes such a project will add to this list, and we welcome feedback on the checklist and how it might be expanded and/or revised.

The questions are organized by section titles within each of the two papers in the series. The section title above each table provides provide guidance on where in each White Paper to look for content in answering the particular questions within the table. Note that the tables are ordered the same as the sections in each White Paper. Also note that there are not questions for every section in each White Paper.

Table 23: Questions Relating to the First White Paper of this Series: Integrating Cyber Risk with Financial Risk

Value at Risk (VaR): A Core Concept of Financial Risk Management	**Complete?**
1. What is the VaR threshold being used to calculate reserves (e.g. 99 percentile)?	
2. What is being calculated: VaR, CVaR, something else?	
Risk of a Portfolio that Includes Cyber Risk	**Complete?**
3. What is the cyber risk for which the reserve is being calculated?	
4. What is the purpose of the analysis: • Assessment? • Comparison of risk mitigation alternatives?	
5. What is the formula for calculating reserves based on VaR, CVaR, or other tail risk measures?	
6. If calculating reserves for multiple risks, how are those risks correlated?	
7. What analytic perspective is being taken (who/what is the stakeholder?)	
8. Do risk modeling assumptions support the distributions chosen to model the risk?	
9. What are risk measures of the various risk mitigation alternatives?	
10. What risk modeling software tools will be used in the project?	
11. Do the software tools support the distributions chosen to model the risk scenarios?	
12. Do the software tools generate the results needed to support the reserve calculations?	
13. What data sources will inform risk factor estimates?	
14. What is the quality of the data?	
15. What SMEs will be relied upon to inform risk factor estimates?	

16. What risk factors are being modeled?	
17. What forms of loss are being modeled?	
18. What are the assumptions made in the model and analysis?	
19. How realistic are those assumptions?	
20. How have the data and SME opinions been incorporated into the model's risk factor estimates?	
Risk Adjusted Return on Capital (RAROC)	**Complete?**
21. Is RAROC relevant to inform an information security controls investment decision?	
22. What is the hurdle rate on RAROC based investment decisions?	
23. How are RAROC and the hurdle rate used as outputs of this analysis?	
Measuring Return on Security Investment (ROSI)	**Complete?**
24. How do ROSI and RAROC analyses compare to each other?	
25. Do risk mitigation proposals meet the minimally acceptable RAROC hurdle rate?	
26. Do risk mitigation proposals meet the minimally acceptable ROSI measure?	

Table 24: Questions Relating to the Second White Paper in the Series: Calculating Reserves for Cyber Risk: Vetting Cyber Risk Models

How FI Risk Managers Accept and Vet Risk Models: Model Risk Analysis	**Complete?**
27. How has the model been analyzed for parsimony?	
28. How is the model relevant to the risk(s) and conditions being modeled?	
29. What is the model's correctness with respect to the risk(s) and conditions being modeled?	
30. Is the model sufficiently transparent that it can be independently implemented?	
31. Are the results of the model (i.e., the model's calculations) tested for correctness?	
Stress Testing and Scenario Analysis	**Complete?**
32. What stress test scenarios have been developed and analyzed?	
Organization of the Cyber Risk AnalysisSupply Chain	**Complete?**
33. What risk analysis standard are relevant to the modeling?	
34. Are vendor tools sufficiently transparent so that their calculations are known and can be verified?	
35. Are the analysts and SMEs assessing the risks sufficiently trained in cyber risk analysis?	

Estimating Risk Factors	**Complete?**
36. What data were used to validate model performance?	
37. What data were omitted / not used?	
38. How are the data representative of the risk scenario modeled?	
39. What are the back-testing requirements?	
40. Is there sufficient historical data to meet the back-testing requirements?	
41. What back-testing has been performed?	
42. How has the model incorporated: • Historical data? • Future expectations? • Uncertainty?	
Making Open FAIR Calculations Transparent for Model Vetting Purposes	**Complete?**
43. Is each model calculation transparent and documented clearly?	
44. Has the model been independently implemented and verified?	
45. Has the model's implementation been verified against its design? (Model correctness)	

Glossary

Action

An act taken against an Asset by a Threat Agent. Requires first that contact occurs between the Asset and Threat Agent.

Asset

The information, information system, or information system component that is breached or impaired by the Threat Agent in a manner whereby its value is diminished, or the act introduces liability to the Primary Stakeholder.

Competitive Advantage Loss or Cost

One of the six forms of loss, Competitive Advantage Losses are losses associated with diminished competitive advantage. Competitive Advantage Loss is specifically associated with Assets that provide competitive differentiation between the organization and its competition. Examples include trade secrets, merger and acquisition plans, etc.

Conditional Value at Risk (CVaR)

Conditional VaR is an alternate risk measure that gives an indication of the magnitude of the potential losses in the tail. In particular, CVaR is the expected loss beyond VaR (i.e., the expected loss given that the loss exceeds the VaR).

Contact Event

Occurs when a Threat Agent establishes a physical or virtual (e.g., network) connection to an Asset.

Contact Frequency (CF)

The probable frequency, within a given timeframe, that a Threat Agent will come into contact with an Asset.

Control

Any person, policy, process, or technology that has the potential to reduce the Loss Event Frequency – Loss Prevention Controls – and/or Loss Magnitude – Loss Mitigation Controls.

Control Strength (CS)

The strength of a control as compared to a standard measure of force.

FAIR

Factor Analysis of Information Risk.

Fines and Judgments Losses or Costs

One of the six forms of loss, fines and judgments losses or costs, are those associated with legal or regulatory actions levied against an organization. Note that this includes bail for any organization members who are arrested.

Loss Event

Occurs when a Threat Agent's action (Threat Event) is successful in breaching or impairing an Asset.

Loss Event Frequency (LEF)

The probable frequency, within a given timeframe, that a Threat Agent will inflict harm upon an Asset.

Loss Flow

The structured decomposition of how losses materialize when a Loss Event occurs.

Loss Magnitude (LM)

The probable magnitude of loss resulting from a Loss Event. Losses are categorized into six forms of loss: productivity, replacement, response, competitive advantage, reputation, and fines and judgments – organized into Primary Loss Magnitude and Secondary Loss Magnitude.

Loss Scenario

The story of loss that forms a sentence from the perspective of the Primary Stakeholder.

Parsimony

The goal that a risk model is neither too complex nor too simple.

Primary Stakeholder

The person or organization that owns or is accountable for an Asset.

Probability of Action (PoA)

The probability that a Threat Agent will act against an Asset once contact occurs.

Productivity Loss or Cost

One of the six forms of loss, Productivity Losses are losses associated with the reduction in an organization's ability to generate its primary value proposition (e.g., income, goods, services).

Replacement Loss or Cost

One of the six forms of loss, Replacement Losses are those associated with the intrinsic value of an Asset. Typically represented as the capital expense associated with replacing lost or damaged Assets (e.g., rebuilding a facility, purchasing a replacement laptop).

Reputation Loss or Cost

One of the six forms of loss, Reputation Losses are those associated with an external perception that an organization's value proposition is reduced, or leadership is incompetent, criminal, or unethical.

Resistance Strength (RS)

The strength of a Control as compared to the probable level of force (as embodied by the time, resources, and technological capability; measured as a percentile) that a Threat Agent is capable of applying against an Asset.

Response Loss or Cost

One of the six forms of loss, Response Losses are the expenses associated with managing a Loss Event (e.g., internal or external person-hours, logistical expenses).

Risk

The probable frequency and probable magnitude of future loss.

Risk Analysis

The process to comprehend the nature of risk and determine the level of risk. [Source: ISO Guide 73:2009]

Risk Assessment

The overall process of risk identification, risk analysis, and risk evaluation. [Source: ISO Guide 73:2009]

Risk Factors

The individual components that determine risk, including Loss Event Frequency, Loss Magnitude, Threat Event Frequency, etc.

Risk Management

Coordinated activities to direct and control an organization with regard to risk. [Source: ISO Guide 73:2009]

Secondary Stakeholder

Individuals or organizations that may be affected by events that occur to Assets outside of their control. For example, consumers are Secondary Stakeholders in a scenario where their personal private information may be inappropriately disclosed or stolen.

Threat

Anything that is capable of acting in a manner resulting in harm to an Asset and/or organization; for example, acts of God (weather, geological events, etc.), malicious actors, errors, failures.

Threat Agent

Any agent (e.g., object, substance, human, etc.) that is capable of acting against an Asset in a manner that can result in harm.

Threat Capability (TCap)

The probable level of force (as embodied by the time, resources, and technological capability) that a Threat Agent is capable of applying against an Asset.

Threat Community

A subset of the overall Threat Agent population that shares key characteristics.

Threat Event

Occurs when a Threat Agent acts against an Asset.

Threat Event Frequency (TEF)

The probable frequency, within a given timeframe, that a Threat Agent will act against an Asset.

Value at Risk (VaR)

Value at Risk is defined as the worst loss that might be expected from holding a security or a portfolio over a given period of time (say a single day or 10 days) given a specified level of probability known as the confidence level.

Vulnerability (Vuln)

The probability that a Threat Event will become a Loss Event; the probability that Threat Capability is greater than Resistance Strength. (Synonym: Susceptibility)

Acronyms & Abbreviations

AML	Anti-Money Laundering
CF	Contact Frequency
CISO	Chief Information Security Officer
CRO	Chief Risk Officer
CVaR	Conditional Value at Risk
DFAST	Dodd-Frank Act Stress Test
FAIR	Factor Analysis of Information Risk
FI	Financial Institution
IID	Independently Identically Distributed
ISAC	Information Sharing and Analysis Center
IT	Information Technology
LEF	Loss Event Frequency
LM	Loss Magnitude
NLE	Number of Loss Events
O-RA	The Open Group Standard for Risk Analysis
O-RT	The Open Group Standard for Risk Taxonomy
PII	Personally Identifiable Information
PoA	Probability of Action
RAROC	Risk-Adjusted Return on Capital
RS	Resistance Strength
SICI	Systemically Important Cloud Infrastructure
SIEM	Security Information and Event Management
SIFI	Systemically Important Financial Institution
SLEF	Secondary Loss Event Frequency
SME	Subject Matter Expert
TCap	Threat Capability
TEF	Threat Event Frequency

VaR	Value at Risk
Vuln	Vulnerability

References

(Please note that the links below are good at the time of writing but cannot be guaranteed for the future.)

- *Basel Committee on Banking Supervision: Principles for Effective Data Aggregation and Risk Reporting*, January 2013; published by Bank for International Settlements; refer to: https://www.bis.org/publ/bcbs239.pdf
- *Calculating Reserves for Cyber Risk: Integrating Cyber Risk with Financial Risk, The Open Group White Paper (W215)*, November 2021, (the first White Paper of this Series), published by The Open Group; refer to www.opengroup.org/library/w215
- *Cost of a Data Breach Report 2020, The Full Version*; refer to: https://www.ibm.com/account/reg/us-en/signup?formid=urx-46542. (Note: An IBM login is required for access.)
- *Cost of a Data Breach Report 2020, The Highlights*; refer to: https://www.ibm.com/downloads/cas/QMXVZX6R
- *Gartner Glossary: Security Information and Event Management (SIEM)*; refer to: https://www.gartner.com/en/information-technology/glossary/security-information-and-event-management-siem
- *How Risky is Your Risk Information*, by Robert Mark and Dilip Krishna, Autumn 2008, Journal of Risk Management in Financial Situations (JRMFI), 1(4); refer to: https://www.ingentaconnect.com/content/hsp/jrmfi/2008/00000001/00000004/art00011
- How to Measure Anything: Finding the Value of "Intangibles" in Business, Third Edition, by Douglas W. Hubbard, April 2014, Wiley
- *IRIS 20/20 XTREME: Analyzing the 100 Largest Cyber Loss Events of the Last Five Years*; published by Cyentia Institute; refer to: https://www.cyentia.com/wp-content/uploads/IRIS2020-Xtreme.pdf
- *ISO Guide 73:2009, Risk Management – Vocabulary*, November 2009; refer to https://www.iso.org/standard/44651.html
- *National Institute of Standards and Technology (NIST): Cyber Security Framework (CSF)*; refer to: https://www.nist.gov/cyberframework
- *Parsimony – A Model Risk Paper*, by Gary Nan Tie and Dr. Bob Mark, 2020, published by PRMIA Institute; refer to: https://prmia.org/PRMIAInstitute/Resources/Papers/Parsimony_-_A_Model_Risk_Paper
- *The Essentials of Risk Management, 2nd Edition*, by Michel Crouhy, Dan Galal, and Robert Mark, February 2014, published by McGraw Hill
- *The Fatal Conceit: The Errors of Socialism*, by F.A Hayek, 1988, published by The University of Chicago Press; refer to: https://press.uchicago.edu/ucp/books/book/chicago/F/bo3643985.html
- *The Open FAIR™ Risk Analysis Tool (I181)*, January 2018, published by The Open Group; refer to: www.opengroup.org/library/i181

- *The Open Group Standard for Risk Analysis (O-RA), Version 2.0.1 (C20A)*, published by The Open Group, November 2020; refer to: www.opengroup.org/library/c20a
- *The Open Group Standard for Risk Taxonomy (O-RT), Version 3.0.1 (C20B)*, published by The Open Group, November 2020; refer to: www.opengroup.org/library/c20b
- *Verizon: Data Breach Investigations Report 2020, Executive Summary*; refer to: https://enterprise.verizon.com/resources/reports/dbir/

Acknowledgements

The Open Group Security Forum acknowledges the contribution of the following people in the refinement and publication of this document:

- Christopher Carlson, C T Carlson LLC
- Dr. Jack Freund, VisibleRisk
- Camille Kloppenburg, Security Compass
- Eva Kuiper, Invited Expert
- Altaz Valani, Security Compass
- John Linford, Forum Director, Security & OTTF, The Open Group

The Open Group and the authors of this document gratefully acknowledge the feedback received at the presentation prior to publication made to the Society of Information Risk Analysts. All errors are the responsibility of the authors.

About the Authors

Dr. Robert (Bob) Mark

Dr. Bob Mark is a Managing Partner at Black Diamond Risk Enterprises. He serves on several boards, has led Treasury/Trading activities, and was a Chief Risk Officer at Tier 1 banks. He is the Founding Executive Director of the MFE Program at UCLA, he has co-authored three books on Risk Management, and holds an Applied Math PhD. Bob was awarded Financial Risk Manager of the Year by GARP, is a co-founder of PRMIA, has published extensively in leading business and finance journals, and is an Individual Contributor in The Open Group Security Forum.

Mike Jerbic

Mike Jerbic is the Founder and Managing Director of Trusted Systems Consulting Group specializing in cyber risk management. He is a retired lecturer in the Economics Department at San Jose State University, and serves as Chair of The Open Group Security Forum. Prior experience includes product development engineering and management at Hewlett Packard, and IT project management consulting. He has authored several articles and book chapters for the American Bar Association as a cyber risk and economics contributor.

About The Open Group

The Open Group is a global consortium that enables the achievement of business objectives through technology standards. With more than 870 member organizations, we have a diverse membership that spans all sectors of the technology community – customers, systems and solutions suppliers, tool vendors, integrators and consultants, as well as academics and researchers.

The mission of The Open Group is to drive the creation of Boundaryless Information Flow™ achieved by:

- Working with customers to capture, understand, and address current and emerging requirements, establish policies, and share best practices
- Working with suppliers, consortia, and standards bodies to develop consensus and facilitate interoperability, to evolve and integrate specifications and open source technologies
- Offering a comprehensive set of services to enhance the operational efficiency of consortia
- Developing and operating the industry's premier certification service and encouraging procurement of certified products

Further information on The Open Group is available at www.opengroup.org.