What is the amount of exposure for a cloud-hosted application? What is the return on security investment (ROSI) on a previous investment? Even further—with the Equifax breach in mind—what is the financial impact in case a scenario with a similar unpatched Apache Struts application or any other unpatched application(s) arises? Imagine how this same scenario gets more tortuous in a post-EU General Data Protection Regulation (GDPR) era.

It is the general consensus that cyberrisk is surely a business risk. From a business risk standpoint, the most important question to be answered is to know the adequate cyber insurance coverage for an organization to cover its bases in case of a breach. There is no straightforward answer to this today. This entirely depends on several variables, including the risk posture of the organization and the insurance provider, who can, in most cases, is not willing to offer a package that would cover what the business anticipates, as it does not have the right tools or data to estimate the risk posture of the customer.

Cyberrisk Insurance Landscape

Cyber insurance, along with cyberrisk, has become a very common agenda item on the boardroom discussion list in recent times. Both enterprises and insurance companies are finding it difficult to quantify the controls in place and the amount of risk each of the parties is undertaking. Cyber insurance has undergone a substantial evolution from a coverage perspective as there are several new risk factors that were not witnessed or considered before (such as cyberextortion, espionage and privacy breaches).

Cyber insurance coverage is additional to the liability, property and theft insurance that has been traditionally offered. But the challenge here is twofold. Insurers do not have a set baseline or robust setup to evaluate the organization’s cyberrisk to determine insurance premiums. Today, most of this is done by leveraging basic questionnaires.
to evaluate the current state of cyberrisk. This practice may result in owning a high risk that could negatively impact the insurance company. On the other hand, if the questions are misinterpreted by the organization, this may result in higher premiums. The post-incident insurance implications are adverse if the organization overstated the controls while acquiring the policy.

Traditionally, auto or home insurance companies provide insurance based on variables such as the driver’s age, type of car driven, year a home was built, and proximity to fire and police services. This risk-aware decision-making is possible primarily because the data and metrics have been available for several decades. Similar maturity and metrics are not available for IT risk management, which implies there is a lot of uncertainty. This is where

statistics and probability can help. Figure 1 illustrates that the dearth of data triggers the vicious cycle of cyberinsurance. In fact, it is actually the inability of both the provider and consumer to mine just enough data to estimate the cyberrisk that triggers this vicious cycle.

Fitch Ratings Inc. reported that the Insurance Data Security Model Law was adopted by the US National Association of Insurance Commissioners to promote more rigorous cyberrisk management. Similar regulations are needed to promote more rigorous cyberrisk management.

Figure 1—The Vicious Circle of Cyber Insurance

Lower sales mean less access to data.

Buyers doubt coverage value, lowering take-up rate/market penetration.

Insurers have not collected a critical mass of less data on their own.

A large percentage of cyberevents go unreported.

Cyberattacks are constantly evolving, making exposure difficult to anticipate.

Insurers cautiously offer modest limits for restricted coverage.

Insufficient data undermine insurer confidence for underwriting and pricing.

Source: Deloitte University Press. Reprinted with permission.
practices. They point out that limited historical data loss, varying policy language, and terms and challenges in quantifying risk aggregations present considerable uncertainty for insurers. Any slight reduction in this considerable uncertainty would enhance the current state. Statistical and probabilistic methods are leveraged when uncertainty is involved. This article provides evidence that statistical and probabilistic risk assessments can help both parties arrive at a conclusion as to how much risk is being transferred in quantitative terms.

In lieu of the vicious cycle of cyber insurance mentioned previously, a (cyber)consumers, providers and regulators (CPR) cycle in figure 2 is proposed, and it can enable robust cybersecurity and risk practices if harmony is attained and maintained. The triangle illustrates that the cyberinsurance providers, customers and regulations such as GDPR, Payment Card Industry (PCI), US Health Insurance Portability and Accountability Act (HIPAA), and US Sarbanes-Oxley Act (SOX) are interdependent and together can contribute to improve the state of cybersecurity and insurance. Increases in the number of breaches often result in new regulations that drive insurance providers to raise the cost of coverage. This is conspicuously evident in the case of the upcoming GDPR rollout. In a different vein, new regulations also drive cyber insurance customers to adopt more stringent security controls (possibly reducing future breaches), and with insurance coverage rising, they are forced to accurately estimate potential risk. This would stabilize the coverage price and enforce providers to optimize coverage level. The US Department of Homeland Security emphasizes that a robust cybersecurity insurance market could help reduce the number of successful cyberattacks. Accurately estimating the potential cyberrisk is a good place to start for a security and risk professional. From a security program perspective, the burgundy arrows in figure 2 should be the top priority to reap the benefit of better coverage at

![Figure 2—(Cyber)Consumers, Providers and Regulators (CPR) Cycle](image-url)
optimal cost and to reduce the number of breaches in the long haul.

Due to recent data breaches, more CISOs have been hired globally in recent times, and some of these individuals have finally procured their long-craved seat at the boardroom table. This simply means that the CISO has an increased responsibility to inform the board of the current risk state and share meaningful security metrics so the board is well informed to make the right decisions. Making the right decisions has paramount importance as enterprises may be able to avert major financial risk and possible reputational damage or even prevent going out of business. This includes securing a robust cyber insurance policy that covers any cataclysmic risk. When these decisions are primarily based on risk assessments, it is critical to use methods that function and, most importantly, measure how well these risk assessment methods work. After all, one cannot manage what one cannot measure. Before all else, a baseline for common cyberrisk language needs to be established.

“Before all else, a baseline for common cyberrisk language needs to be established.”

Terminology Consensus

“Risk,” “vulnerability,” “threat” and “asset” each have a contextualized meaning and are often used interchangeably with one another. For example, malicious insiders, weak passwords, nation-state actors, cybercriminals, hacktivists and network shares are not risk. But the taxonomy in most organizations today concerning risk is that most of these are misinterpreted as a potential risk. Risk practitioners need to have a nomenclature consensus and adept understanding of the difference between a threat, threat agent, vulnerability, asset and risk. A common vocabulary harmony needs to exist not only within organizations, but also among insurance providers, law enforcement and corporations, which greatly assists in executing the cyber CPR efficiently. This is best attained by practice and training. Further guidance can be found in the Factor Analysis of Information Risk (FAIR) book.14

Quantitative Cyberrisk Assessments Today

It has been relentlessly advocated that attributing numbers to colors on a heat map will not make it a quantitative risk assessment. Figure 3 is a simpler version of the risk matrix example to explain the range compression problem with heat matrices. Two risk scenarios follow:

- **Risk A**—Likelihood is 40 percent, Impact = US $6 million
- **Risk B**—Likelihood is 80 percent, Impact = US $1.5 million

The risk is evaluated by multiplying impact and likelihood. Clearly the expected loss for Risk A, US $2.4 million, is much greater than the expected loss for Risk B, US $1.2 million.

But the risk matrix depicts otherwise. It shows Risk A to be a medium risk and Risk B to be a high-level risk, which is just the opposite of what the mathematical evaluation of the expected loss suggests.

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Impact</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;US $1M (Minor)</td>
<td>US $1M-$10M (Moderate)</td>
<td>≥US $10M (Catastrophic)</td>
</tr>
<tr>
<td>High (&gt;75%)</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Medium (&gt;25%-75%)</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Low (&lt;25%)</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
</tr>
</tbody>
</table>

Change is an unwelcome nemesis anywhere in any form. The priority of organizations, especially dealing with cybersecurity, should be to drive a change in the thought process around adopting probabilistic quantitative risk assessments and clear any misconceptions.15 “Culture eats strategy
for breakfast”16 appropriately describes how organizations blindly adopt the proposition to leverage quantitative cyberrisk measurement models based on age-old practices, myths about data availability and statistical ignorance. And sometimes, organizational politics also play a major factor. The blatant fact here is that quantitative risk assessments based on probabilistic models need to be adopted as a standard to help make better, more accurate decisions. Unfortunately, most leading frameworks and consortiums still use heat maps.

“We use probability because we lack perfect information, not in spite of it.”24 One key element addressed in the book from which this quote is taken is that statistics help in estimating rarely occurring events with minimal or just enough data sets.

Key Elements—Analysts, Data and Tools

It is well known that three elements—people, process and technology—form the crux of any successful business transformation. Similarly, for risk estimation, the three elements that are key for quantitative cyberrisk analysis are skill of analysts, having just enough data and leveraging commonly available tools.

The skill of analysts is to extract the data that matters and perform a reasonable estimation. Consider the bald tire scenario25 when explaining the interpretation of risk terms and mental calculations made by practitioners based on invalid assumptions. The point is that inaccurate assumptions will jeopardize the entire risk analysis exercise. After acquiring enough data, statistical methods can be implemented using commonly available tools, such as Microsoft Office Excel, FAIR-based software or tools like Analytica by Lumina.

The FAIR model is a widely adopted model today that utilizes Monte Carlo and Program Evaluation Review Technique (PERT) to estimate risk. Similarly, a model that utilizes some decomposition tactics along with Monte Carlo simulations and Bayes method has been suggested. Simple analyses or prototyping can be performed by leveraging Excel spreadsheets using built-in statistical functions. For complex scenarios and larger organizations, these preliminary evaluations are scalable and can be integrated into enterprise governance, risk and compliance (GRC) solutions by leveraging programming languages such as Python, R.26

Case Study: Risk Due to Loss of PHI Data Via Email

To showcase how this can be done, an example to evaluate the risk for an email misdirection or confidential data loss via email described in the FAIR Institute’s blog27 was chosen. The analysis was done

THE THREE MAIN STEPS IN FAIR INCLUDE DEFINING THE SCOPE, PERFORMING THE RESEARCH AND MAPPING IT TO THE FAIR MODEL.
The three main steps in FAIR include defining the scope, performing the research and mapping the results to the FAIR model. After these have been established, one can finally make decisions based on the result.

**Define Scope**
The key elements of any risk scenario are actor, threat type, event, asset and time. Defining the scope of a scenario is a critical step, and it comprises identifying the asset at risk, the threat actor and the effect. Sensitive or critical data in the email are the asset risk here. An internal user is the threat. The user may be a privileged user who has access to sensitive data (such as protected health information [PHI]) or a nonprivileged user who may have access to sensitive data (such as personally identifiable information [PII]). An inadvertent act or an intentional malicious act would have the same effect. Hence, whether the act is malicious or inadvertent does not matter here, unless cybercriminals are included, and they are out of scope for this discussion as it pertains to emails sent by internal users. The effect of this kind of act will be the loss of confidentiality of critical information.

Risk scenarios involved in this scope are described in **figure 4**.

**Research and Map**
Instead of mapping it to the FAIR model in this step, another model was used to perform the decomposition and analysis. The threat sources in the previous scope are the line items in the spreadsheet shown in **figure 5**. The line items and decomposition can be anything, including applications, threat sources, business units in an organization or vulnerabilities, depending on the preference. This would determine if it is a risk analysis or a risk assessment.

FAIR ontology recommends evaluating the loss event frequency and loss magnitude to evaluate risk. The following are some questions risk practitioners should pose to subject matter experts (SMEs) to evaluate loss event frequency (LEF):29

- How frequently are PHI data sent via email, and how many patient records (on average) are in one email?
- How often does an employee deliver an email to an incorrect recipient?
- Is the PHI within the emails encrypted without needing to login to an account to access the reports? If so, that is a vulnerability.

Based on the previous responses, the likelihood that the event will happen is evaluated. (See the second column in **figure 5**.) This can also be decomposed further using Bayes methods.30

Then the loss magnitude (in the FAIR methodology) can be evaluated in two steps—primary and secondary loss. Loss of productivity and replacement costs occur mostly as primary loss. Legal liabilities/fines, intellectual property loss and reputational damages occur as secondary loss. Incident response costs fall into both primary and secondary loss categories.

For this case study, primary costs are customer service time to handle the email glitch (investigating and responding to the event) and to replace the terminated employee(s) (if such a thing is part of the policy enforcement). There is no loss in productivity as there is no operational disruption. Secondary costs include offering credit monitoring to customers, fines by a regulator if personal credit and/or health information was released, and potential settlements on customer lawsuits. There is also

| Figure 4—Risk Scenarios in Scope for Data Loss in Unencrypted Email |
|----------------|------------------|------------|----------------|
| Threat Type    | Threat Actor/Agent/Source | Asset        | Threat Effect/Event |
| Inadvertent/malicious intent | Privileged insider       | Customer information | Confidentiality |
| Inadvertent/malicious intent | Nonprivileged insider | Customer information | Confidentiality |
reputational damage, especially if it is a publicly traded corporation. With this in perspective, the Excel template is leveraged to estimate the loss magnitude by decomposing it into observables: replacement cost, response cost, cost in legal liabilities and fines, and reputational cost. Plugging in the calibrated estimates for these decomposition variables using the knowledge and input from the SMEs provides the expected losses shown in figure 5. The range for legal liabilities and fines, for example, should include the 4 percent annual global turnover or US $23.7 million dollars (whichever is greater) fine that is levied if GDPR applies to the organization.

The total loss is then evaluated, and a Monte Carlo simulation of 100,000 such scenarios is run (figure 6).

A histogram is devised (figure 7) that helps plot the loss exceedance curve (LEC) in figure 8. These mathematical calculations and simulations can be performed easily by leveraging tools such as Excel or R. The difficult part is getting the estimates to be accurate, and expertise alone will not help. Getting the right estimates involves posing the right questions to the SMEs and slowly narrowing down to a final value. Posing the right questions comes from practice along with tools (e.g., RiskLens’ CyberRisk Suite) that come with preconfigured questions that will assist the risk practitioner.

Making Decisions
Once all of the math is complete, it is time to paint a picture that highlights the current risk state compared to risk appetite. The LEC shown in figure 8 depicts that there is a 30 percent chance that the loss will be greater than US $2.2 million. Similarly, there is 10 percent chance the loss will be more than US $30 million.
This offsets the belief that security is not an investment that provides a return.\textsuperscript{32}

\[
\text{Return on control percentage} = \left(\frac{\text{Reduction in Losses}}{\text{Cost of Control}} - 1\right) \times 100
\]

Figure 9 shows sample events that are categorized based on return on control percentage\textsuperscript{33} and a response to mitigate, immediately mitigate or track is suggested.

**What Is Next?**

Many organizations have already leveraged statistical risk assessments. Work is in progress to make these models widely available and increase awareness of the benefits of embracing uncertainty. All this progress would make it simpler for organizations to evaluate cyberrisk in a meaningful way rather than classifying it as a specific color or assigning it an unrealistic value. This, in turn, will help insurance providers to accurately understand the onus they are bound to undertake and embolden them to come up with better pricing and accurate coverage.

Although these statistical methods include numbers that often perplex boards of directors and CISOs,
the fact is that a small patching vulnerability in a web application could result in a breach that can cost millions or even billions of US dollars.34 There is no doubt that the magnitude will be even higher if such breaches transpire in the GDPR age. One can wait to be part of the historical data or do the actual math (numbers do not lie) upfront to mitigate the risk that really matters. Organizations may not need statistical cyber risk assessments in the future when historical data are abundant and the uncertainty becomes negligible. But, until then, the goal is to keep reducing that uncertainty.

Endnotes


