

Evaluating the Business Value of Enterprise Data Protection

A Framework for Cost, Risk, and Operational Resilience

SREEHARI PRASAD JINKA VENUGOPAL¹ (MAIL09HARI@GMAIL.COM)¹

¹MBA, ENTERPRISE PRODUCT MANAGEMENT, DATA PROTECTION, CYBER RESILIENCE, TRUSTWORTHY AI
GOVERNANCE PHOENIX, ARIZONA, USA.

ORCID: 0009-0002-5969-9357

Abstract

Enterprise data protection is usually justified in the language of fear and evaluated in the language of cost. Budget owners see licences, storage and staff time, while the benefits arrive as incidents that did not happen or that were survived with less damage, which are hard to observe and easy to dismiss. This paper proposes the Data Protection Value Framework (DPVF), a structured way to evaluate the business value of data protection along three connected dimensions: the cost of protection, the risk it reduces, and the operational resilience it provides. DPVF links assets, vulnerabilities, threats, controls and countermeasures in a common risk ontology, values assets by business importance, expresses risk as a distribution of annual losses and downtime and not as a single score, and combines the dimensions through explicit weights that reflect the organization's risk appetite. It is organized in the layers of enterprise risk management (framework, model and methodology) so that evaluation results feed governance decisions. We examine the framework in a Monte Carlo study of four protection portfolios (baseline, backup-centric, prevention-centric and integrated) over 50,000 simulated years for a mid-sized enterprise. At base assumptions, the integrated portfolio cut the 95th-percentile annual loss from about 4.0 million to 1.2 million and the 95th-percentile annual downtime from 178 to 57 hours, yet its expected net financial return was slightly negative (about minus 50 thousand per year), and it paid for itself in only 18% of simulated years. When the assumed ransomware frequency doubled, it delivered the highest net value of the four portfolios. A weighting sweep showed that the preferred portfolio shifts from baseline to backup-centric to integrated as the weight on risk and resilience increases. The results are synthetic, depend on stated assumptions, and are intended to show how the framework reasons, not to estimate returns for any real organization. We close with governance guidance and a research agenda.

Keywords: Data Protection, Business Value, Return on Security Investment, Operational Resilience, Enterprise Risk Management, Loss Distribution, Cyber Risk, Decision Support.

1. Introduction

Every organization protects data, and every organization asks, sooner or later, whether it is spending the right amount in the right way. The question is harder than it appears. The costs of data protection are visible and recurring: backup and recovery platforms, endpoint and identity controls, encryption, monitoring, staff, training and audit. The benefits are counterfactual. They consist of losses that were avoided, outages that were shorter than they would have been, regulatory penalties that never materialized, and customer trust that was never tested. When nothing happens, protection looks like an expense that produced nothing, and when something does happen, the question changes from value to blame.

This asymmetry produces predictable behavior. Budgets for data protection tend to rise after incidents and to be challenged after quiet periods. Technologies are purchased because a peer suffered a particular attack, and not because they address the largest exposure. Different functions use incompatible vocabularies: security speaks of threats and controls, finance speaks of cost and return, operations speaks of recovery time and availability, and compliance speaks of obligations. As a result, decisions are made by negotiation and not by a shared evaluation.

Three difficulties make a shared evaluation hard. First, losses from data incidents follow heavy-tailed distributions. Most years bring only small events, while rare events are very large, so that average annual loss understates the exposure that threatens the organization and any evaluation based only on expected value will undervalue protection that works mainly in the tail. Second, the value of protection is not only financial. Operational resilience, meaning the ability to continue or quickly resume critical operations, matters to customers, regulators and employees in ways that a financial model captures poorly. Third, protection measures interact. Prevention reduces how often incidents occur, recovery reduces how much they cost when they do, and the combination can be worth more or less than the sum of the parts.

This paper proposes a framework that addresses these difficulties. The primary objective is to design and examine the Data Protection Value Framework (DPVF). The specific tasks are: (i) to position data protection evaluation within the foundations of enterprise risk management; (ii) to define a risk ontology that connects assets, vulnerabilities, threats, controls and countermeasures to the valuation of assets; (iii) to specify how cost, risk and resilience are measured and combined, with explicit weights; (iv) to define governance and reporting that make the evaluation repeatable; and (v) to examine the framework's behavior in a controlled simulation, including its sensitivity to threat frequency and to the weights placed on competing objectives. The contribution is a coherent, transparent and testable approach to evaluating data protection, together with explicit statements about what the evidence does and does not show.

The remainder of the paper is organized as follows. Section II reviews the literature. Section III summarizes related work. Section IV presents the DPVF methodology. Section V details the experimental setup, Section VI reports results, Section VII discusses implications, Section VIII addresses limitations, and Section IX concludes.

2. LITERATURE REVIEW

A. Enterprise risk management foundations

Enterprise risk management (ERM) provides the vocabulary in which data protection decisions should be framed. The COSO framework integrates risk with strategy and performance, and ISO 31000 provides principles and a process for managing risk of all kinds [1], [2]. ISO/IEC 27005 and NIST Special Publication 800-30 apply the same logic to information security risk, describing how to identify assets, threats and vulnerabilities, estimate likelihood and impact, and select treatments [3], [4]. A distinction that is useful in practice, and shown in Figure 1, separates the *framework* (governance structure, roles and oversight standards), the *model* (operational execution, dashboards and workflows) and the *methodology* (assessment approach, scoring and mitigation steps). These three layers sit beneath a strategy focus on risk appetite alignment, compliance integration and data visibility.

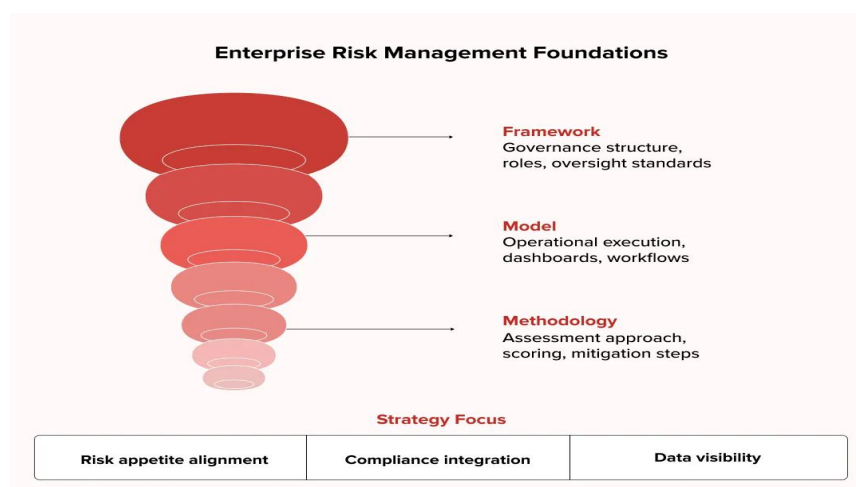


Fig. 1. Enterprise risk management foundations: framework, model and methodology layers, with a strategy focus on risk appetite alignment, compliance integration and data visibility.

The layering matters for data protection because evaluation is not a one-off calculation. A methodology that scores risk is useful only if a model turns it into recurring workflows and dashboards, and if a framework assigns accountability and oversight. Many data protection assessments fail at this transition: they produce a report, but no one is responsible for acting on it, and no mechanism repeats it when conditions change.

B. The economics of security investment

The economics of information security has long been concerned with how much to spend. Anderson argued that many security failures are better explained by incentives and economics than by technology [5], and Anderson and Moore surveyed the field's contributions on externalities, asymmetric information and investment [6]. Gordon and Loeb showed that the optimal investment in protecting an information set does not always rise with its vulnerability, and that it is generally only a fraction of the expected loss [7]. Cavusoglu and colleagues presented a model for evaluating security investments that accounts for the strategic behavior of attackers and the choice of technologies [8]. Sonnenreich and colleagues proposed return on security investment (ROSI), which relates the risk reduction obtained from a measure, net of its cost, to the cost itself [9]. ROSI is attractive because it speaks the language of finance, but its estimates depend on the quality of loss and risk-reduction inputs, which are uncertain and difficult to validate. Kaplan and Norton's balanced scorecard offers a different perspective, arguing that financial measures alone are insufficient for managing performance and should be complemented by customer, process and learning measures [10].

C. Quantifying cyber and data risk

Quantitative risk analysis represents risk as a distribution of losses. The FAIR approach decomposes risk into the frequency of loss events and the magnitude of loss per event, and expresses each as a range and not as a point estimate [11], [12]. Hubbard and Seiersen argue that even rough probabilistic estimates calibrated by experts outperform ordinal scoring such as red, amber and green, which hides uncertainty and cannot be aggregated meaningfully [13]. Loss distributions in this domain are typically heavy-tailed, so that measures of the tail are needed. The quantitative risk management literature describes value at risk and expected shortfall, and Artzner and colleagues explain why a measure of risk should have properties such as subadditivity, which value at risk lacks [14], [15], [16]. For decision purposes, these measures answer the question that average loss cannot: how bad can a bad year be?

D. Operational resilience and recovery

Resilience engineering views safety and reliability as the ability of a system to adjust its functioning to sustain operations under both expected and unexpected conditions [17]. In data protection, resilience is reflected in recovery time and recovery point objectives, and in the demonstrated ability to meet them. NIST contingency planning guidance and the Cybersecurity Framework emphasize planning, testing and recovery as core functions [18], [19]. Resilience contributes to business value in ways that are not purely a reduction of loss: it supports customer commitments, limits regulatory exposure for operational disruption, and reduces the dependence of the business on any single protective measure. Industry studies of breach costs, such as the annual IBM and Ponemon study, report that detection and response capabilities and the time taken to contain an incident are associated with lower costs, which is consistent with a value of resilience that operates through the severity of events and not their frequency [20].

3. RELATED WORK

Prior work approaches the evaluation of data protection from several directions, and few contributions integrate them.

In security economics, models of optimal investment and return on security investment provide mathematical structure for the spending decision [7], [8], [9]. They typically concentrate on expected loss and treat the protective measure as a single quantity, so that they say little about which combination of measures to buy, how the measures interact, or how to value reductions in tail risk and downtime.

In quantitative risk analysis, FAIR and related approaches offer practical methods for estimating loss frequency and magnitude [11], [12], [13]. They are strong on risk estimation, but

they are generally used to rank risks, not to compare alternative protection portfolios on cost, risk and resilience simultaneously, nor to connect the results to governance.

In risk management standards and guidance, the processes of identifying assets, vulnerabilities and threats and selecting treatments are described in detail [2], [3], [4]. These sources are comprehensive about process and less specific about how to quantify the business value of a treatment or how to trade competing objectives.

In resilience engineering and continuity planning, resilience is described and recovery is planned and tested [17], [18], [19], but resilience is rarely expressed in the same units as financial value, which leaves it outside the usual investment conversation.

In multi-criteria decision analysis, methods for combining several objectives through explicit weights are well established, as in the work of Keeney and Raiffa, and the sensitivity of rankings to weights can be studied systematically [21], [22]. These methods are seldom applied to data protection portfolios with a complete treatment of loss distributions.

The gap that DPVF addresses is therefore one of integration. No widely adopted approach (a) connects assets, vulnerabilities, threats and controls to asset valuation in a common ontology, (b) represents risk and resilience as distributions and not as single scores, (c) compares complete protection portfolios on cost, risk and resilience with explicit and testable weights, and (d) embeds the result in an ERM governance structure.

4. METHODOLOGY: THE DPVF FRAMEWORK

A. Overview and structure

DPVF is organized in the three ERM layers shown in Figure 1. At the framework layer it defines governance: who owns the evaluation, who approves risk appetite and weights, and how results are reviewed. At the model layer it defines the workflows and dashboards through which evaluation is repeated and reported. At the methodology layer it defines how assets are valued, how risk is estimated, how cost and resilience are measured, and how the three are combined. Throughout, the three strategy foci of the figure, risk appetite alignment, compliance integration and data visibility, are treated as requirements: the weights must reflect stated appetite, compliance obligations must enter the analysis as constraints and benefits, and every input must be traceable to data that the organization can inspect.

B. A risk ontology for data protection

DPVF uses the entity relationships shown in Figure 2 as its risk ontology. An asset has an identifier, name, description, type and importance level. An asset may have zero or more vulnerabilities, each with a name, description and weight. Each vulnerability may be exploited by one or more threats, which have a description and an impact. Each vulnerability can be addressed by zero or more security controls and zero or more countermeasures, each of which has a weight. In the figure, the elements are annotated with the dimensions of a capability framework labelled CSCF: access and trust and supply chain attacks relate to the threat entity, security governance relates to controls and countermeasures, asset valuation and continuity and operations relate to the asset and its importance, and defence relates to the weight of the vulnerability.

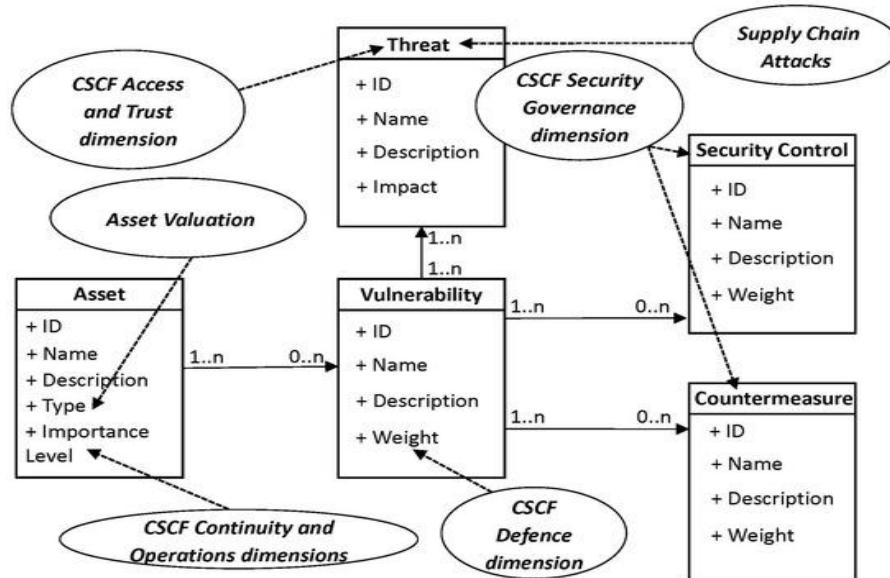


Figure 2. Entity-relationship view of the risk ontology: assets, vulnerabilities, threats, security controls and countermeasures, with the capability dimensions that annotate them.

The ontology serves two purposes. It provides a shared vocabulary that lets security, finance, operations and compliance refer to the same entities. And it makes the evaluation traceable: every claim that a measure reduces risk can be tied to the vulnerabilities it addresses and the assets those vulnerabilities expose. An asset that is highly important but exposed through many weakly addressed vulnerabilities is easy to identify, as is a control that is costly but addresses no vulnerability of any important asset.

C. Asset valuation

Asset valuation determines what an incident would cost. DPVF values data assets along four lines. Direct recovery cost includes the labor and technology needed to restore data and systems. Business interruption covers revenue lost and costs incurred while processes are unavailable. Liability and regulatory exposure covers fines, notification, legal costs and contractual penalties. Intangible loss covers reputation and customer attrition, which are uncertain but are often the largest component in severe cases. Each asset also receives an importance level reflecting its role in critical operations. Valuation is expressed as ranges, not point values, and the ranges are refined as incident data accumulate. Table 1 lists the cost and value categories used in the framework.

TABLE 1. Cost and Value Categories in DPVF

Dimension	Category	Examples
Cost	Acquisition and licensing	Backup, endpoint, identity, encryption and monitoring tools
Cost	Infrastructure and storage	Backup storage, immutable copies, network and compute
Cost	Operations and labor	Administration, monitoring, restore testing, incident handling
Cost	Compliance and training	Audits, certifications, awareness programs
Risk reduction	Lower frequency of incidents	Prevention, detection and access controls
Risk reduction	Lower severity of incidents	Recovery capability, containment, segmentation
Resilience	Shorter downtime, tighter recovery objectives	Verified restores, automation, tested runbooks
Enabling value	Compliance, assurance and trust	Lower audit effort, insurance terms, customer confidence

D. Measuring cost

Cost is measured as the total cost of ownership over a planning horizon, usually three to five years, including acquisition, infrastructure, operations and compliance. Two practices improve the measure. First, labor is included, because a tool that is cheap to buy and expensive to operate is not cheap. Second, the cost of the baseline is included, so that the value of a portfolio is measured against what the organization would otherwise have, which is usually not nothing: even without a dedicated program, native platform controls and informal practices exist.

E. Measuring risk

For each threat scenario, DPVF estimates how often it occurs in a year and how much it costs when it does, each as a range reflecting uncertainty. Scenarios are drawn from the ontology and typically include ransomware, accidental or malicious data loss, service outages, data breaches with regulatory consequences, and misconfiguration or corruption. A protection portfolio is then described by its effect on each scenario: how much it reduces the frequency of occurrence (prevention), and how much it reduces the severity of loss and the downtime when the event occurs (recovery and containment). A simulation over many years, in which event counts and severities are drawn at random, produces the full distribution of annual losses and downtime for each portfolio. From the distribution the framework reports the mean, the 95th and 99th percentiles, and the probability that a portfolio's total cost over a year, including both spending and residual losses, is lower than that of the baseline.

The use of paired simulation, in which the same random events are applied to every portfolio, makes comparisons sharper, because differences between portfolios then reflect their effects and not luck.

F. Measuring resilience

Resilience is measured with operational metrics that can be observed: the expected and tail downtime in hours per year, the recovery time and recovery point achieved in tests against the stated objectives, and the share of critical assets whose recovery has been verified. These measures are reported separately from financial loss, because they matter to the organization in ways that dollars may not capture, and because customers and regulators increasingly ask for them directly. Resilience also enters the financial analysis through the severity of downtime-related loss.

G. Value synthesis and weights

DPVF reports three views and does not collapse them prematurely. The financial view shows the expected net value of each portfolio relative to the baseline: the loss avoided, plus enabling benefits such as lower insurance cost and audit effort, minus the additional spend. The tail and resilience view shows percentile losses and downtime. The composite view combines cost and risk-and-resilience outcomes through explicit weights, expressed on an index from 0, where only cost counts, to 100, where only risk and resilience count. Outcomes are first rescaled so that the best and worst portfolios on each dimension anchor the scale, and the weighted combination then yields a composite value index for each portfolio.

The weight is a statement of risk appetite and not a technical parameter, and it must be set by those who are accountable for the risk, namely senior management and the board. DPVF therefore requires the weight to be set explicitly, documented, and accompanied by a sensitivity sweep showing how the preferred portfolio changes as the weight moves. If the preferred portfolio changes at a weight close to the stated appetite, the decision is fragile and deserves additional scrutiny. This approach follows multi-attribute decision analysis, in which explicit trade-offs are recommended over implicit ones [21].

H. Sensitivity and uncertainty analysis

Since inputs are uncertain, DPVF requires sensitivity analysis. One-at-a-time variation of the most influential inputs, especially the frequency of the largest threats, shows how robust the ranking of portfolios is. Wider global sensitivity methods can be used where the inputs are numerous [22]. The output of the analysis is not a single number but a statement of the conditions under which each portfolio is preferred, for example the ransomware frequency above which the integrated portfolio has the highest expected net value.

I. Governance and reporting

A risk owner sets the weight and approves the evaluation. A data protection owner operates the controls and the recovery capability. A finance partner validates cost and valuation inputs. A compliance owner confirms obligations and evidence. A review board meets at least annually and after significant incidents to review the inputs, the evaluation and the decisions taken. Reporting uses dashboards that show cost, tail risk, downtime and recovery verification side by side, and includes the sensitivity results. The evaluation is repeated when the threat landscape, the business or the portfolio changes materially, and incident data are used to recalibrate estimates, which closes the loop between the methodology, the model and the framework layers.

5. EXPERIMENTAL SETUP

The experiment is limited in purpose. Real organizations cannot observe how their losses would have differed under an alternative portfolio, and incident data are too sparse and too sensitive to support a controlled comparison. We therefore used a Monte Carlo study of a hypothetical mid-sized enterprise, implemented in Python with a fixed random seed. The study examines three questions: how four protection portfolios differ in expected loss, tail loss and downtime; how the economics change when the frequency of the largest threat changes; and how the preferred portfolio depends on the weight placed on cost against risk and resilience.

Threat scenarios. Five scenarios are modelled. For each, Table 2 gives the assumed annual frequency and the typical severity and downtime per event. Severity and downtime per event follow log-normal distributions, which produces the heavy tails characteristic of incident losses.

TABLE 2. Assumed Threat Scenarios (baseline portfolio)

Scenario	Events per year	Typical loss per event (median)	Typical downtime per event (median, hours)
Ransomware	0.15	900,000	96
Accidental or insider data loss	0.80	40,000	8
SaaS or cloud outage	1.50	90,000	10
Data breach with regulatory impact	0.10	2,000,000	24
Misconfiguration or corruption	0.60	120,000	14

Protection portfolios. Four portfolios are compared. P0, baseline, relies on native platform controls and informal practice. P1, backup-centric, adds immutable backup and routine restore capability, which mainly reduces the severity and duration of events. P2, prevention-centric, adds endpoint protection, data loss prevention, identity hardening and posture management, which mainly reduces how often events occur. P3, integrated, combines prevention and backup with restore validation and orchestrated response, in the spirit of the readiness approach discussed in the companion literature on recovery readiness. Table 3 gives the assumed annual cost and the qualitative effect of each portfolio; the quantitative effects on frequency, severity and downtime by scenario are inputs of the simulation.

TABLE 3. Assumed Portfolio Costs and Benefits

Portfolio	Annual cost	Enabling benefits per year	Main mechanism
P0 Baseline	150,000	0	Native controls
P1 Backup-centric	450,000	40,000	Lower severity and downtime
P2 Prevention-centric	750,000	90,000	Lower frequency
P3 Integrated	1,000,000	160,000	Lower frequency, severity and downtime

Simulation. For each of 50,000 simulated years, event counts for each scenario are drawn from a Poisson distribution, and severity and downtime per event are drawn from the log-normal distributions. The same drawn events are applied to every portfolio: prevention removes a random share of events, and recovery and containment scale the severity and downtime of those that remain. This paired design isolates the effect of each portfolio.

Metrics. We report annual cost, mean annual loss, the 95th and 99th percentile of annual loss, mean and 95th-percentile annual downtime, the expected net value of each portfolio relative to the baseline (loss avoided plus enabling benefits minus additional spend), the resulting return on the additional spend, and the share of simulated years in which the portfolio's total cost (spend plus residual loss, minus enabling benefits) was lower than the baseline's.

Sensitivity analyses. Two analyses are performed. In the first, the ransomware frequency is multiplied by 0.5, 1, 2 and 3 and the expected net values are recomputed. In the second, the weighting index is varied from 0 to 100 in steps of 5, and for each value the composite value index of each portfolio is computed from cost and from risk and resilience outcomes (the 95th-percentile loss and the mean downtime), each rescaled so that the best and worst portfolios anchor the scale.

All assumptions, including frequencies, severities, portfolio effects and costs, are inputs of the simulation and not measurements of any real organization.

6. RESULTS

A. Portfolio economics at base assumptions

Table 4 reports the results at base assumptions. The baseline portfolio had a mean annual loss of about 1.01 million, a 95th-percentile annual loss of about 4.0 million, and a 99th-percentile loss of about 10.9 million, with 58 hours of downtime in an average year and 178 hours in a bad year. Each added portfolio reduced loss and downtime. The integrated portfolio reduced mean loss to about 0.37 million, the 95th-percentile loss to 1.2 million, the 99th-percentile loss to 4.6 million, mean downtime to 19 hours and 95th-percentile downtime to 57 hours. The backup-centric portfolio, at lower cost, achieved the second-largest reductions in downtime, and the prevention-centric portfolio achieved larger reductions in loss than downtime.

TABLE 4. Simulated Annual Results by Portfolio (50,000 simulated years; amounts in thousands)

Portfolio	Annual cost	Mean loss	95th-pct loss	99th-pct loss	Mean downtime (h)	95th-pct downtime (h)	Expected net value vs. baseline	Return on added spend (%)	Years better than baseline (%)
P0 Baseline	150	1,008	4,019	10,862	58.4	178.3	0	n/a	n/a
P1 Backup-centric	450	736	2,968	8,879	33.3	96.2	+12	4	23.7
P2 Prevention-centric	750	580	2,018	6,341	39.3	118.8	-82	-14	14.1
P3 Integrated	1,000	369	1,217	4,596	19.3	56.9	-50	-6	18.0

The financial result is more nuanced than the risk result. After allowing for the additional spend and for the enabling benefits, only the backup-centric portfolio had a positive expected net value (about 12 thousand per year, a return of 4% on the added spend). The prevention-

centric portfolio was negative (about minus 82 thousand), and the integrated portfolio was slightly negative (about minus 50 thousand). Measured year by year, none of the portfolios paid for itself in most years. The backup-centric portfolio had lower total cost than the baseline in 23.7% of simulated years, the integrated portfolio in 18.0% and the prevention-centric portfolio in 14.1%. This pattern is a property of heavy-tailed losses: in most years no large event occurs, and the additional spend is not recovered, whereas in the rare bad years the protection delivers large savings.

B. Sensitivity to threat frequency

Table 5 reports the sensitivity of expected net value to the frequency of ransomware. If ransomware were half as frequent as assumed, all three portfolios would have a negative expected net value, between minus 53 and minus 155 thousand. At the base frequency the backup-centric portfolio is marginally positive. At twice the base frequency, the integrated portfolio has the highest expected net value, about 178 thousand, ahead of the backup-centric (142 thousand) and prevention-centric (74 thousand) portfolios. At three times the base frequency, the integrated portfolio's net value rises to about 390 thousand, against 281 thousand for the backup-centric portfolio and 205 thousand for the prevention-centric portfolio.

TABLE 5. Expected Net Value Relative to Baseline by Ransomware Frequency (thousands per year)

Ransomware frequency (multiple of base)	P1 Backup-centric	P2 Prevention-centric	P3 Integrated
0.5	-53	-148	-155
1	+12	-82	-50
2	+142	+74	+178
3	+281	+205	+390

The break-even is therefore informative. The integrated portfolio is worth its cost when ransomware events are expected roughly every three to four years or more often, with the other assumptions unchanged, and it is not worth its cost when they are rarer. This statement is more useful to decision makers than a single return figure, because it focuses attention on the input that matters and on the evidence needed to estimate it.

C. Sensitivity to weights

Figure 3 shows the composite value index of each portfolio as the weighting index moves from 0, where only cost counts, to 100, where only risk and resilience outcomes count. The baseline portfolio is preferred when the weight on risk and resilience is below about 45. The backup-centric portfolio is preferred in a narrow band from about 45 to 55, and the integrated portfolio is preferred from about 60 upward. The prevention-centric portfolio is never preferred: under these assumptions it is more expensive than the backup-centric portfolio and less effective at reducing tail loss and downtime than the integrated portfolio, so it is dominated on both dimensions over the whole range.

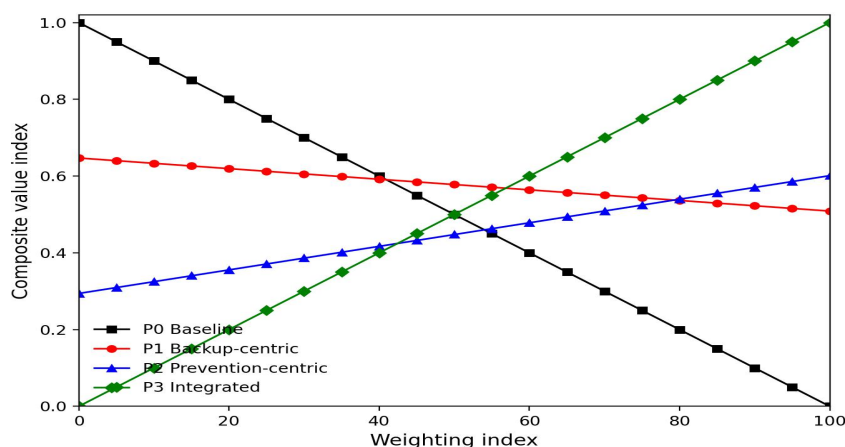


Figure 3. Composite value index of the four portfolios as the weighting index moves from cost only (0) to risk and resilience only (100).

The backup-centric portfolio holds a stable middle position, scoring between 0.51 and 0.65 across the whole range, which makes it a robust compromise when the organization is unsure of its appetite. The baseline and integrated portfolios are the extremes: each scores 1.0 at one end of the range and 0 at the other, so the choice between them is a pure statement of appetite.

D. Summary of findings

Three findings stand out. First, the value of the integrated portfolio is concentrated in the tail: it reduced 95th-percentile loss by about 70% and 95th-percentile downtime by about 68% relative to the baseline, even though its expected financial return was slightly negative and it was cheaper in total in fewer than one year in five. Second, the economic case is highly sensitive to the frequency of the largest threat, and doubling the ransomware frequency was enough to make the integrated portfolio the best in expected value. Third, the preferred portfolio depends on the weight placed on risk and resilience against cost, and one of the four portfolios, prevention-centric, was dominated under every weight.

7. DISCUSSION

The study supports the central argument of the framework: the business value of data protection cannot be captured by one number. An evaluation that looked only at expected net financial return would reject the integrated portfolio at base assumptions and would offer little reason to prefer any portfolio other than the cheapest. An evaluation that looked only at tail loss and downtime would select the integrated portfolio without regard to its cost. DPVF presents both views, together with the conditions under which they agree, and therefore turns a disagreement about conclusions into a discussion about inputs and appetite.

The simulation also illustrates why protection spending is hard to defend in quiet years. In four years out of five, even the best portfolio cost more than it saved. This is the insurance property of data protection, and it explains why organizations alternately overspend after incidents and underspend after quiet periods. Presenting the distribution makes the pattern visible and allows the board to decide whether it is willing to bear the tail or prefers to pay to reduce it.

For practitioners, five recommendations follow. Evaluate whole portfolios and not individual tools, because the effects of prevention and recovery differ in kind and combine in ways that single-tool evaluations miss. Report tail loss and downtime next to expected loss, and do not rely on average figures. Identify the one or two threat frequencies that drive the economics, and invest in evidence for those, for example through incident history, industry data and red-team exercises. State the weight between cost and risk and resilience explicitly, obtain approval from the accountable executive, and show the sensitivity sweep. And revisit the evaluation after incidents and at least annually, because both the threat landscape and the organization change.

The framework also has implications for compliance and assurance. Regulators and customers increasingly ask not only whether controls exist but whether recovery capability has been demonstrated and whether the organization understands its exposure. A documented evaluation with explicit assumptions, traceable inputs and a record of decisions provides evidence for such inquiries, and the ontology of assets, vulnerabilities, threats and controls gives auditors a structure to follow.

Finally, the framework is deliberately conservative about automation and false precision. The numbers in an evaluation are estimates with wide uncertainty, and their value lies in making assumptions explicit and in showing how conclusions depend on them. A model that produced a single confident return figure would conceal exactly the uncertainty that decision makers need to see.

8. LIMITATIONS AND FUTURE WORK

The principal limitation is that the evaluation is synthetic. Every input, including threat frequencies, severities, downtime, portfolio effects, costs and enabling benefits, is an assumption chosen to be plausible, and none was calibrated against data from a real organization. The absolute results, such as the expected net value of each portfolio, should not be expected to hold in practice, and different but equally plausible assumptions could reverse the ranking of portfolios at base conditions. The more durable conclusions are structural: the concentration of value in the tail, the sensitivity to the frequency of the largest threat and the dependence of the preferred portfolio on appetite.

Second, the simulation treats threat scenarios as independent and stationary. In reality, incidents cluster, attackers adapt to defenses, and a successful attack can trigger several loss types at once. Correlation and adaptation would typically make tail outcomes worse and the value of resilience greater, so the model may understate the case for resilient portfolios. Third, the intangible components of loss, such as reputation and customer attrition, are folded into severity and not modelled separately, although in severe incidents they may dominate. Fourth, the composite value index relies on min-max rescaling, which is sensitive to the set of portfolios compared: adding or removing a portfolio changes the scale for all. Fifth, the enabling benefits are small, fixed amounts, whereas real benefits such as improved insurance terms, faster audits and customer wins are uncertain and vary with context.

Future work should therefore include: (i) calibration of scenario frequencies and severities against incident data from real organizations, with expert elicitation methods and uncertainty ranges; (ii) models that include correlated and adaptive attackers; (iii) explicit modelling of intangible loss and of enabling benefits; (iv) comparison of weighting and aggregation methods, including approaches that do not depend on rescaling; and (v) a field pilot in which an organization applies the framework to an actual investment decision and compares its predictions with subsequent experience.

9. CONCLUSION

This paper presented DPVF, a framework for evaluating the business value of enterprise data protection along the dimensions of cost, risk and operational resilience. It is organized in the layers of enterprise risk management, uses a risk ontology that connects assets, vulnerabilities, threats, controls and countermeasures, represents risk as distributions of loss and downtime, and combines the dimensions through explicit and testable weights, with governance and sensitivity analysis built in. In a Monte Carlo study of four portfolios, the integrated portfolio reduced 95th-percentile loss from about 4.0 million to 1.2 million and 95th-percentile downtime from 178 to 57 hours, yet had a slightly negative expected return at base assumptions and paid for itself in under one year in five; it became the best portfolio in expected value when ransomware frequency doubled, and it was the preferred portfolio under the composite index when the weight on risk and resilience was 60 or more.

The study is evidence of coherence and not of field effectiveness. The next step is empirical: calibrate the inputs with real data, apply the framework to an actual investment decision, and compare its predictions with experience. If the results are consistent with the logic shown here, organizations can replace arguments from fear or from cost alone with a transparent evaluation that shows what protection costs, what it changes in the bad years, and what the organization must believe about its threats to justify it.

References

1. Pawan Kumar Singh, Scaling Gate-All-Around (GAA) Transistor Architectures for Sub-2nm AI Accelerators Using Atomically Thin 2D Materials, Vol. 33 No. 08 (2024): JOCAAA, Journal Name: Journal of Computational Analysis and Applications, Journal's ISSN: 1521-1398 (Paper),1572-9206 (Online), <https://eudoxuspress.com/index.php/pub/article/view/5713>
2. Pawan Kumar Singh, Energy-Efficient Task Scheduling in Green Cloud Computing Using Neuromorphic Spiking Neural Networks, Vol. 33 No. 08 (2024): JOCAAA, Journal Name: Journal of Computational Analysis and Applications, Journal's ISSN: 1521-1398 (Paper),1572-9206 (Online), <https://eudoxuspress.com/index.php/pub/article/view/5712>
3. A. Mahendra Vardhan and S. Sridhar, "Determining False Positive Analysis of Software Vulnerabilities with Predefined Scan Rules using Random Forest Classifier and Decision Tree Technique," 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2022, pp. 622-625, DOI: <https://doi.org/10.1109/ICAC3N56670.2022.10074458>
4. Ronakkumar shah , Cognitive Process Intelligence: Integrating Generative AI and Automated Workflow Discovery, Vol. 33 No. 08 (2024): JOCAAA, Journal of Computational Analysis and Applications, Journal's ISSN: 1521-1398 (Paper),1572-9206 (Online), <https://eudoxuspress.com/index.php/pub/article/view/5676>

5. Ronakkumar shah, research Trends in Smart Workflow Automation: A Bibliometric Analysis of Emerging Themes, Home / Archives / Vol. 33 No. 05 (2024): JOCAAA, Journal of Computational Analysis and Applications, Journal's ISSN: 1521-1398 (Paper),1572-9206 (Online), <https://eudoxuspress.com/index.php/pub/article/view/5675>
6. Dip Bharatbhai Pate, Dynamic vs. Static visualizations: Understanding their use cases in big data analysis, International Journal of Science and Research Archive, 2022, 05(02), 392-395., DOI <https://doi.org/10.30574/ijrsra.2022.5.2.0077>
7. Dip Bharatbhai Patel, Building scalable business intelligence systems in the cloud: A technical approach, International Journal of Science and Research Archive, 2021, 02(01), 212-215., Received on 03 February 2021; revised on 17 April 2021; accepted on 19 April 2021, DOI: <https://doi.org/10.30574/ijrsra.2021.2.1.0047>
8. Dip Bharatbhai Patel, Understanding the Impact of Text Analytics on Social Media Sentiment Analysis, Vol. 5 No. 5 (2019) ,September-October, DOI: <https://doi.org/10.32628/CSEIT1936413>
9. Dip Bharatbhai Patel, Real-Time Data Streaming in BI Dashboards: Opportunities and Challenges, Archives / Vol. 5 No. 4 (2019) : July-August, DOI: <https://doi.org/10.32628/CSEIT192142>
10. Dip Bharatbhai Patel, the Role of Data Visualization in Modern Intelligence, Archives / Vol. 2 No. 2 / Research Article, Published, 2020-11-28 DOI: <https://doi.org/10.32996/jcsts.2020.2.2.4>
11. A. Srivastava, "Securing PHI in cloud-based healthcare systems: Challenges, frameworks, and future directions," J. Comput. Anal. Appl., vol. 33, no. 2, Aug. 2024. [Online]. Available: <https://www.eudoxuspress.com/index.php/pub/article/view/4349/3190>
12. A. Srivastava, "AI-assisted cloud migration of healthcare claims data from legacy systems: A metadata-driven framework," Int. J. Commun. Networks Inf. Secur., vol. 14, no. 3, Nov. 2022. [Online]. Available: <https://ijcnis.org/index.php/ijcnis/article/view/8682>
13. A. Srivastava, "Enhancing provider and claims data accuracy using artificial intelligence on cloud-native data platforms," J. Comput. Anal. Appl., vol. 31, no. 4, Nov. 2023. [Online]. Available: <https://eudoxuspress.com/index.php/pub/article/view/4347/3189>
14. A. Srivastava, "A unified framework for secure cloud data management: Integrating governance, metadata intelligence, and privacy controls," J. Inf. Syst. Eng. Manage., vol. 9, no. 2, Apr. 2024. [Online]. Available: https://www.jisem-journal.com/download/45_A_Unified_Framework.pdf
15. A. Srivastava, "Optimizing large-scale data migration for cloud-native architectures," J. Comput. Anal. Appl., vol. 31, no. 2, pp. 652-662, May 2023. [Online]. Available: <https://eudoxuspress.com/index.php/pub/article/view/4603>
16. Abdul Aziz. "Intelligent Health Analytics: Mastering AI in Healthcare Data". Acta Scientiae, vol. 25, no. 4, Aug. 2024, pp. 544-55, <https://www.periodicos.ulbra.org/index.php/acta/article/view/687>. <https://doi.org/10.5281/zenodo.21699886>
17. Event-Aware Multi-Layer Storage Risk Forecasting for Oracle Database Estates Using HAPF, Raghu Gollapudi, International Journal of Computational and Experimental Science and Engineering, 10(4), 2024 ,10.22399/ijcesen.5183
18. Backup Integrity and Recovery Readiness Assessment for High-Availability Databases Raghu Gollapudi, Computer Fraud & Security, Vol. 2023, Issue 8, pp. 35-48, 10.52710/cfs.1031
19. Risk-Controlled Near-Zero-Downtime Oracle Database Migration Using GoldenGate Raghu Gollapudi, International Journal of Computational and Experimental Science and Engineering, 8(3), 2022 ,10.22399/ijcesen.5382
20. An Automated Risk Scoring Framework for SQL Execution Plan Analysis and Performance Manoj Kumar Kagitha, Automated Security Compliance in Ci/Cd: A Natural Language Processing (NLP) Approach to Detecting Vulnerabilities in Infrastructure-As-Code (Terraform/Bicep),Publication year 2021, Volume 2021, Issue 12, <https://computerfraudsecurity.com/index.php/journal/article/view/970>
21. Manoj Kumar Kagitha, Hybrid Quantum-Classical Resource Scheduling: A Proposed Architecture for Next-Generation Hyperscale Data Centers, Publication year 2024, Volume 2024, Issue 12, <https://computerfraudsecurity.com/index.php/journal/article/view/969>
22. Manoj Kumar Kagitha, Self-Healing Cloud Infrastructure: A Hybrid Reinforcement Learning Framework for Predicting Microservice Failures in Azure Kubernetes Service (AKS),Publication year 2023, Volume 2023, Issue 12,<https://computerfraudsecurity.com/index.php/journal/article/view/971>

23. Manoj Kumar Kagitha, LOAD BALANCING TECHNIQUES IN DISTRIBUTED CLOUD ENVIRONMENTS, (IJESMR - 2017), D.O.I:10.53555/jcr.v6:i1.13380, <https://jcreview.com/archives/volume-6/issue-1/13380>
24. Manoj Kumar Kagitha, AI-DRIVEN OBSERVABILITY FOR HYPERSCALE COLOCATION DATA CENTERS: A CASE STUDY OF LOG ANALYTICS AND ANOMALY DETECTION PIPELINES, Archives / Vol. 27 No. 7 (2019): JOCAAA, <https://www.eudoxuspress.com/index.php/pub/article/view/5053>
25. Manoj Kumar Kagitha, GREENOPS IN THE CLOUD: AN AI-DRIVEN TELEMETRY ANALYSIS MODEL FOR REDUCING CARBON FOOTPRINT IN HIGH-AVAILABILITY CLUSTERS, Vol. 28 No. 6 (2020): JOCAAA, <https://www.eudoxuspress.com/index.php/pub/article/view/5054>
26. Bandari, V. V., & Lekkala, H. B. (2024). AI-based operator behavior monitoring and cost optimization using digital traceability in manufacturing systems. *Power System Protection and Control*, 52(1), 38–45. <https://pspac.info/index.php/dlbh/article/view/342>
27. Venumadhav Vavilala, Shankar Balla (2024, May). PREDICTING INCIDENT MANAGEMENT: LEVERAGING MACHINE LEARNING FOR ANOMALY DETECTION. *Power System Protection and Control Scopus Q1 Journal*. PSPC. <https://pspac.info/index.php/dlbh/article/view/270>
28. Manikantha Varaprasad Inakollu, (2024, May), FINOPS-Driven Cloud optimization models for enterprise applications, Vol. 52 No. 2 (2024): April-June 2024, 99-110, *Power System Protection and Control*, ISSN-1674-3415, URL: <https://pspac.info/index.php/dlbh/article/view/283>, DOI: <https://doi.org/10.46121/pspc.52.2.10>
29. Naresh Lokiny. (2022). Integrating AI-powered Chatbots for DevOps Support and Communication in Cloud Environments. *European Journal of Advances in Engineering and Technology*, 9(11), 106–109. <https://doi.org/10.5281/zenodo.13325989>
30. Naresh Lokiny, (2021), "Disaster Recovery and Business Continuity Planning in DevOps Cloud with AI", *International Journal of Science and Research (IJSR)*, 10(3), 2024-2027. <https://dx.doi.org/10.21275/SR24724151733>, <https://www.ijsr.net/getabstract.php?paperid=SR24724151733>
31. Naresh Lokiny, & Ranganath Nandanampati. (2020). DevSecOps: Integrating Security into DevOps with AI in Cloud. *Journal of Scientific and Engineering Research*, 7(10), 239–242. <https://doi.org/10.5281/zenodo.13348695>
32. Naresh Lokiny, & Pradip Reddy. (2021). Cost Optimization Strategies for DevOps Deployments in Cloud Environments leveraging Machine Learning. *European Journal of Advances in Engineering and Technology*, 8(3), 69–72. <https://doi.org/10.5281/zenodo.13325845>
33. Jayanth Para, 6G Internet Technology Cyber Threat Notification & Alert System, Vol. 52 No. 4 (2024): October-December 2024, *Power System Protection and Control*, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/170>, DOI: <https://doi.org/10.46121/pspc.52.4.9>
34. Jayanth Para, AI Based Cloud Computation Observational Method & Process, Vol. 51 No. 4 (2023): October-December 2023, *Power System Protection and Control*, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/171>, DOI: <https://doi.org/10.46121/pspc.51.4.3>
35. Kaleshwar Aryasomayajula, PREVENTING BIAS IN AI-BASED DECISION-MAKING: ANALYZING TECHNIQUES TO REMOVE UNFAIR PREJUDICE IN ALGORITHMIC OUTCOMES, Vol. 50 No. 2 (2022): April-June 2022, *Power System Protection and Control*, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/294>
36. Kaleshwar Aryasomayajula, MODERN DEVELOPMENT PRACTICES: INVESTIGATIONS INTO SERVERLESS COMPUTING, LOW-CODE/NO-CODE PLATFORMS, AND DEVELOPER PRODUCTIVITY IN REMOTE ENVIRONMENTS, Vol. 50 No. 4 (2022): October-December 2022, *Power System Protection and Control*, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/364>
37. Kaleshwar Aryasomayajula, Micro services: "A Comparative Analysis of Monolithic vs. Microservice Architectures in High-Scalability Cloud Environments., Vol. 51 No. 2 (2023): April-June 2023, *Power System Protection and Control*, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/374>
38. Vikas Suresh Bagora, KERNEL-LEVEL PERFORMANCE OPTIMIZATION FOR CARRIER-GRADE BROADBAND AND HIGH-SPEED NETWORKING SYSTEMS, Vol. 47 No. 1 (2019), *Power System Protection and Control*, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/359>

39. Vikas Suresh Bagora, SCALABLE 128G FIBRE CHANNEL AND ETHERNET CONVERGENCE FOR NEXT-GENERATION DATA CENTER NETWORKS, Vol. 50 No. 4 (2022): October-December 2022, Power System Protection and Control, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/362>
40. Vikas Suresh Bagora, SECURE EMBEDDED NETWORKING ARCHITECTURES USING TRUSTED EXECUTION ENVIRONMENTS IN BROADBAND GATEWAYS, Vol. 52 No. 2 (2024): April-June 2024, Power System Protection and Control, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/363>
41. Santhosh Chitraju Gopal Varma, The Evolution of Cloud-Native Architectures: Exploring the Synergy between Kubernetes and Microservices, DOI: <https://doi.org/10.63282/3050-9246.IJETCSIT-V11I4P104>, Vol. 1 No. 4 (2020), <https://ijetcsit.org/index.php/ijetcsit/article/view/108>
42. Santhosh Chitraju Gopal Varma, Artificial Intelligence in Cloud Computing: Building Intelligent, Distributed, and Fault-Tolerant Systems, Vol. 3 No. 1 (2022), DOI: <https://doi.org/10.63282/3050-9416.IJAIBDCMS-V3I1P105>, <https://ijaibdcms.org/index.php/ijaibdcms/article/view/78>
43. Sunil Karmakar, AI driven Identity & Access Management with Oracle Cloud security, Vol. 30 No. 2 (2022): JOCAAAA, <https://eudoxuspress.com/index.php/pub/article/view/4118>
44. Rakshitha Kiran P., Samitha Khaiyum, Aparna R. Palandye, Apoorva S. D., Arjun R., Akshaya S. A., Annapoorna S., Leveraging LLaMA3 and LangChain for Rapid AI Application Development, Vol. 20 No. 10s (2024), <https://doi.org/10.52783/jes.5539>, <https://journal.esrgroups.org/jes/article/view/5539>.