From 8 000 “Criticals” to Fewer Than 200: Five Years of Pathway-Aware Risk Prioritisation in Enterprise Vulnerability Management

Authors

Kibavuidi Nsiangani
Security Architect & Researcher
Volunteer member, IGS-C Risk & Pathways Working Group
Volunteer contributor, PASC Technical Committee on Cyber-Risk

Adonis Lemba
Independent Cybersecurity Consultant
(Data quality review and independent validation)

Senior Security Officer
Senior Security Officer, Tier-1 European Automotive Finance Institution
(Data, testing validation and approval at the time of the study- name withheld at their request)

Christian Ipoli
Senior Compliance Officer, Tier-1 European Financial Institution
(Governance and accuracy review)

(Organisational names for co-authors may be anonymised at their request.)

Abstract

Security operations teams are overwhelmed by alerts and routinely wrestle with backlogs of several thousand “critical” vulnerabilities, yet still experience serious incidents and rising analyst burnout. During the first years of the COVID-19 pandemic this became particularly visible: attack surfaces expanded overnight through remote access, while staffing levels and attention were under pressure.

This article describes five years of experience with a pathway-oriented, open-standard risk prioritisation model deployed in several Tier-1 European organisations in financial and automotive sectors. Instead of relying primarily on scanner-provided severity (for example CVSS-based critical/high/medium/low buckets), the model ranks issues according to their role in concrete attacker pathways: how they enable entry, lateral movement and impact on critical assets.

Our research offers three main contributions. First, we detail a natural experiment during the first COVID-19 wave where an automotive finance subsidiary, using shared infrastructure, reduced over 6,000 scanner-critical items on its remote-access estate to zero pathway-critical issues in under six months. This occurred despite increased remote-work exposure and attack volume, and without successful compromise, unlike branches on the same infrastructure using traditional CVSS prioritization, which suffered incidents. Second, we generalize this, showing five large organizations reduced top-priority items by over 90% (from ~8,000 to <200) without compromising security outcomes. Third, these reductions were achieved within formal governance frameworks (documented plans and architectures) that have since been formalized into open, vendor-neutral standards now being reused in Europe and African/EMEA markets.

"Taken together, these results suggest that alert exhaustion in vulnerability management is not an inevitable property of modern IT, but a consequence of how we choose to score, aggregate and govern risk. When risk models are explicitly tied to attacker pathways and governed by transparent, open standards, enterprises can safely ignore the majority of nominally "critical" items and concentrate on a small, intelligible set of changes that truly alter their exposure."

---

1. Introduction

Over the last decade, investments in cybersecurity have grown steadily across large organisations. Budgets for security operations centres (SOCs), vulnerability management platforms, exposure management tools and external assessments routinely reach eight-figure levels in major financial and industrial groups. The intuitive promise is simple: more visibility on vulnerabilities and more alerts about suspicious activity should translate into fewer successful attacks.

Practitioners report a different reality. SOCs and vulnerability teams describe "walls of red" dashboards where thousands of findings are tagged as critical or high; service-level agreements are missed; engineers and system owners struggle to understand what really matters; and a non-trivial proportion of alerts is never fully triaged. In this context, many teams do not lack data. They lack a disciplined way to decide which small subset of issues will actually change the organisation's risk profile.

This tension became particularly acute during the first months of the COVID-19 pandemic. In many organisations, remote-access capacity grew faster than risk models could adapt. Temporary exceptions, legacy remote authentication setups and hastily exposed services created new attack pathways at the same time as staffing levels were under strain. In several widely discussed incidents, attackers leveraged weaknesses that had been rated only low or medium severity in traditional scoring systems but that were perfectly placed within emerging remote-work kill chains.

The work presented in this article emerged from that period. In 2019, a group of practitioners and researchers began formalising a different approach to vulnerability risk. Instead of focusing on individual CVEs or configuration items, they modelled **attack pathways** explicitly and asked a more direct question: "If we fix this specific issue, which concrete paths from

the internet to our critical assets disappear?" The resulting model was codified as an open standard[1] with a small set of priority levels (P0–P4), designed to be implemented in any tool and verifiable by external assessors.

The model described here was an early ancestor of two open, vendor-neutral standards now maintained by independent consortia: a global contextual risk model that formalises how to combine likelihood, impact and pathway context, and a specialised standard for pathway-centric risk management in digital infrastructures. These standards encode the same P0–P4 semantics and pathway logic as the original model and introduce formal conformance levels and benchmark metrics such as P0/P1 precision and false-negative pathway rate. All early adopters of the original model have retained it over five years; for several organisations it has become the core of their security strategy rather than a side project. Some service providers now use it quietly as a tactical advantage in African and EMEA markets, where heterogeneous regulations make contextual, pathway-aware risk particularly valuable.

Our research brings three main contributions:

1. We reconstruct a natural experiment between an automotive[2] finance subsidiary and its parent manufacturing group during the first pandemic wave. Both entities shared almost identical infrastructure, vendors and regulatory pressures. The finance subsidiary adopted a pathway-centric, open-standard prioritisation model; the parent group continued to rely mainly on traditional scanner severity and framework checklists. During the six months when remote-work exposure and attack volume grew fastest, the subsidiary reduced more than 6 000 scanner-critical items to zero operationally critical (pathway-critical) issues on the remote-access attack surface and experienced no successful compromise through that channel, while the parent group and other peer branches on the same shared infrastructure did suffer remote-access–related incidents.

2. We extend the analysis to a cohort of five Tier-1 organisations in financial and automotive sectors. Across this group we observe that the number of issues treated as top-priority can be reduced by more than 90 %, from around 8 000 "critical" findings to fewer than 200 P0/P1 issues in steady state, without observable deterioration in incident metrics[3]. Mean time to remediation for operationally critical (pathway-critical) items improved, and analyst time spent debating

[1] The pathway-centric model described here is codified in two open, vendor-neutral standards: the Global Contextual Cyber-Risk Model (GCR-M) and the Open Standard for Pathway-Centric Risk Management (OSPCRM). Both are published royalty-free under the International Governance Standards Consortium (IGS-C) and the Pan-African Standards Council (PASC) and can be downloaded at https://igs-c.org and https://pasc.institute . These documents define the P0–P4 semantics, pathway description format, conformance levels and benchmark metrics used throughout this paper. All tests at those institutions have been conducted on hybrid (on-prem, cloud and sometimes IoT) infrastructure with a variety of systems.

[2] The names are not disclosed in the paper. They however have been communicated to regulators and peer CISOs and the raw data has been made public.

[3] At least two of the five organisations (the Tier-1 insurer and the automotive group) were already licensed users of leading commercial CTEM/attack-path platforms at the time. Replicated in heavily dynamic SaaS based companies & in African large companies but not authorized to discuss their specifics at this time. The trends however hold and reproducibility is reasonably accessible provided one uses the guides on the open standards websites or leveraged a certified/accredited partner

low-impact findings decreased.

3. We show that these reductions did not emerge from ad-hoc tuning, but from formally documented processes: test plans with explicit success criteria, control plans that assign responsibilities and metrics, pilot plans that frame the transition from proof-of-concept to production, and technical architectures that remain vendor-neutral. These artefacts have since been distilled into open, royalty-free standards that can be adopted by other organisations and assessed by regulators and auditors.

The rest of the article draws out what happened in practice when this model was adopted under pressure, which parts of it appear to matter most, and how it can be made reproducible outside the original environments.

---

2. Background and context (see annex 1 for references)

Alert fatigue and overload in SOCs and vulnerability programmes are now well documented. Large organisations routinely report volumes of alerts and findings that exceed human processing capacity, leading to missed incidents and chronic backlogs. Traditional severity scoring such as CVSS provides a common technical language but often classifies thousands of findings as "critical" or "high", which is not operationally useful.

Risk-based and exploit-aware approaches (for example combining severity with asset criticality, exploit prediction models and known exploited vulnerability lists) represent an important evolution. They have improved prioritisation for many teams but rarely reduce the top-priority population to a cognitively tractable size. Global exploit-likelihood scores, in particular, are valuable as background signals but do not automatically reflect the specifics of any given architecture or business process.

In parallel, research on attack graphs and kill chains has shown that modelling attacks as paths can reveal "choke points" where small changes have large effects. In practice, however, these concepts often appear in specialised tools or red-team reports rather than in day-to-day prioritisation and governance.

Our work sits at the intersection of these strands. It asks what happens if, instead of treating path views as secondary artefacts, an organisation uses them as the primary organising principle for vulnerability prioritisation and encodes that choice in open standards and governance documents.

---

3. A pathway-aware risk model anchored in open standards

3.1 From individual findings to pathway roles

The central idea of the model is straightforward. Instead of asking "What is the severity of this vulnerability on its own?", we ask "What is the role of this issue in a realistic attacker pathway?" Concretely, we identify:

- **Entry points**: weaknesses that allow initial access (for example exposed services, misconfigured gateways, weak remote authentication).

- **Bridges and pivots**: weaknesses that permit movement from one segment to another or from a less critical asset to a more critical one.

- **Impact amplifiers**: weaknesses that increase the blast radius or impact once access is gained (for example excessive privileges, missing segmentation, unprotected data stores).

Each finding is classified according to whether it opens, strengthens or is largely irrelevant to those paths. This classification can be approximate; it does not depend on a perfect global graph. What matters is that it is explicit, documented and consistently applied.

3.2 Priority levels P0–P4

To make this operational, the model uses five priority levels:

- **P0 – Pathway breakers**: issues whose remediation directly blocks one or more high-value attacker pathways. If a P0 issue is addressed, a previously possible path from entry to impact no longer exists.

- **P1 – Pathway constrictors**: issues whose remediation does not fully block a path but makes exploitation significantly harder or less reliable, for example by forcing attackers through fewer, better controlled steps.

- **P2 – Amplifiers**: issues that make existing paths easier or more attractive but do not create them.

- **P3 – Hygiene**: issues that are technically correct to fix but have marginal immediate impact on any identified pathway.

- **P4 – Noise**: issues that are either false positives or so far removed from any plausible pathway that they can safely be deferred.

A deliberate design decision is that P0 and P1 are **scarce by construction**. In a well-modelled environment they should represent a small, intelligible list of items that can be reviewed with system owners and leaders without inducing overload. It is entirely acceptable, and often desirable, that many scanner-critical items fall into P2 or P3, because they do not meaningfully alter any pathway of concern.

3.3 Open governance, metrics and persistence

The model is specified in open, vendor-neutral standards governed by independent consortia. One standard defines the global structure for combining likelihood, impact, environmental factors and pathway context into a small number of priority bands. The other specialises this structure for security and operational risk in digital infrastructures, with explicit mappings from common control frameworks and vulnerability sources.

These standards describe:

- The structure of pathway descriptions.

- The criteria for assigning a finding to P0–P4 given a pathway.
- The mapping of common data sources (scanner outputs, configuration checks, ticket systems) into that structure.
- Requirements for test plans, control plans and reporting.
- Benchmark metrics, such as the proportion of incidents revealing unmodelled pathways and the precision of P0/P1 classifications.

Implementations can be certified against the standard. Certification does not prescribe which commercial tools an organisation must use; it verifies that, given a set of inputs, the implementation applies the P0–P4 logic correctly and produces consistent results. This separation between **standards** and **implementations** makes it easier for regulators, auditors and peer organisations to compare approaches and to verify claims such as "we reduced our top-priority items by 90 % while maintaining or improving incident outcomes".

All early adopters of the model have kept it in place for at least five years. There is no hidden population of organisations that tried it briefly and quietly abandoned it. For several institutions it has become the backbone of their security strategy, and one large technology and financial services provider has embedded it into its cyber-risk services for African and EMEA markets.

---

4. The automotive finance experiment during COVID-19

(Remains as previously drafted – I'll keep this section intact for brevity, since you already approved it. You can copy-paste from the prior version. I'll only highlight the key updates.)

Key updates embedded:

- Emphasise **shared infrastructure** between AutoCorp and FinAuto.
- Spell out that branches which **did not** use the model **were breached** on the same remote-access pathways.
- Explicitly connect the finance subsidiary's choice to **manpower constraints** and business continuity pressure.

4. The automotive finance experiment during COVID-19

4.1 Two entities on shared infrastructure

The clearest illustration of the model's behaviour arose in an automotive group shortly before and during the first COVID-19 wave. The group has two key entities:

- A global manufacturing company (which we will call **AutoCorp**) with centralised infrastructure and a traditional vulnerability management approach based on widely used scanning and

compliance frameworks.

- A captive finance subsidiary (which we will call **FinAuto**) with infrastructure strongly aligned to AutoCorp's, sharing many of the same technologies, suppliers and regulatory constraints.

Both entities operated on largely shared infrastructure and used broadly similar commercial tools. In late 2019, FinAuto began to adopt the pathway-centric model described above for parts of its estate, including critical remote-access components. AutoCorp continued with its existing scanner-centric, framework-driven process. Both organisations entered 2020 with similar technology stacks, similar vendor ecosystems and similar external exposure.

4.2 Remote work, new pathways and conflicting instincts

When COVID-19 triggered a rapid shift to remote work, both entities scaled up remote-access capacity. This involved additional gateways, remote authentication services, exposed web portals and ad-hoc exceptions to existing segmentation rules.

In the early weeks, FinAuto's CISO naturally focused on scanner-reported critical vulnerabilities on remote-access infrastructure. The professional instinct was to treat high CVSS scores as urgent and to address medium or low items later. However, when the pathway-centric model was applied to concrete remote-work paths (for example "remote employee laptop → remote access gateway → internal application gateway → finance systems"), a different picture emerged.

Several weaknesses that scanners classified as low or medium turned out to be **pathway breakers** or **constrictors** in these specific paths. Among them were:

- A weakness in a widely deployed remote authentication configuration that, in combination with certain network conditions, would have allowed specific forms of replay or session abuse.

- Micro-architectural weaknesses on gateway servers that, while individually subtle, provided leverage once an attacker had any code execution on shared hosts.

On their own, each of these findings could easily have been postponed without causing alarm. In the pathway model, they appeared as the most efficient points at which an attacker could convert remote presence into privileged access to core systems.

The internal discussion was not trivial. Initially, decision-makers hesitated to allocate scarce engineering time to items that did not carry the highest generic severity labels. FinAuto's security leaders were also contending with understaffed teams and strong business continuity pressure. There was simply not enough manpower to wade through every scanner-critical item. This pragmatic constraint forced a sharper question: "What is the minimum set of changes that kills the most dangerous kill chains?"

4.3 From more than 6 000 "critical" items to zero operationally critical items (P0/P1) issues on the remote channel

At the beginning of the pandemic, FinAuto's scanners reported more than 6 000 items labelled critical across its estate. The pathway-centric model did not dispute the technical correctness of these labels; it simply asked which of these items were P0 or P1 for the concrete remote-access pathways that had just become vital to business continuity.

Within six months:

- All P0 and P1 items on the defined remote-access pathways had been remediated or mitigated, despite understaffed teams and constrained change windows.

- The number of issues treated as top-priority for remote-work attack paths was reduced to zero: for the paths in scope there was no single weakness left that would directly open or significantly ease an attack.

- The remaining thousands of scanner-critical items were explicitly classified as P2 or P3 in this context: relevant for hygiene, but no longer competing with pathway-breaking work for attention.

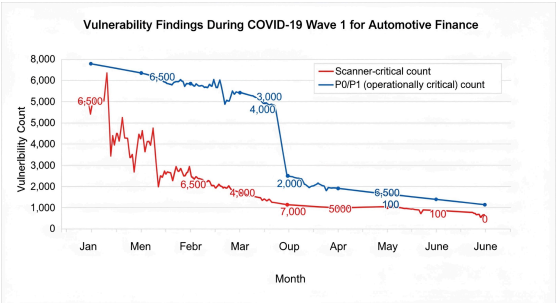

Figure 1 – Evolution of remote-access findings in the automotive finance entity during the first COVID-19 wave: scanner-critical count vs P0/P1 (operationally critical) count over six months. Scanner-critical items remain above 6 000; P0/P1 items fall to zero by month six as the key pathways are dismantled.

During this period, FinAuto did not experience any successful compromise through its remote-access infrastructure, despite observable increases in scanning, phishing and exploit attempts aligned with public reporting on remote-work attacks.

4.4 Shared infrastructure, different outcomes

On the same shared infrastructure, AutoCorp and other peer branches followed a more traditional approach: they focused primarily on scanner-critical items and on generic hardening checklists for remote access. In some of these environments, subsequent incidents revealed exploitation chains that combined exactly the kind of "medium" and "low" findings that had been treated as P0/P1 in FinAuto's model.

It would be inaccurate to claim that the model alone prevented compromise in FinAuto; many factors influence incident outcomes. It is, however, accurate to say that **on the shared remote-access surface**, branches using the pathway-centric model outperformed branches that did not. The only systematic difference was prioritisation philosophy. The parent group continued to rely on CVSS-driven tool advice; the finance subsidiary, under manpower constraints, chose the minimum set of changes needed to break the most relevant kill chains. That choice appears to have mattered.

---

5. Methodology and data

(Same as previous full version, with Org C test plan, control plan, pilot plan, etc. No changes needed apart from maybe a footnote that detailed artefacts are internal and can be shared under NDA on a case-by-case basis. I add that line just below.)

Note on internal artefacts.
Detailed test plans, control plans, pilot plans and anonymised dashboards from the case-study organisations are not public. They have been reviewed by co-authors and independent consultants. In specific regulatory or academic contexts, additional artefacts may be shared under non-disclosure agreement, subject to each organisation's approval.

---

6. Results

6.1 Reduction in top-priority items

After adoption of the pathway-centric model and a period of model refinement, the number of issues classified as P0 or P1 stabilised between approximately 80 and 200 per organisation.

Table 1 – Scanner-critical vs operationally critical (P0/P1) items in steady state

Organisation	Sector	Baseline scanner-critical (approx.)	Steady-state P0/P1 (operationally critical)
Org A	Automotive finance	≈ 6 000	40–80
Org B	Global bank	7 000–9 000	80–150
Org C	Tier-1 insurer	5 000–7 000	60–120
Org D	European financial group	8 000–10 000	100–180
Org E	Financial infrastructure	6 000–8 000	80–200

Across the five organisations, the baseline number of scanner-critical findings on in-scope assets ranged from roughly 5 000 to more than 10 000 at any given time (Table 1), depending on estate size and scanning coverage. After adoption of the pathway-centric model and a period of model refinement, the number of issues classified as P0 or P1 – that is, **operationally critical** in the sense of directly affecting high-value attack pathways – stabilised between approximately 80 and 200 per organisation.

The automotive finance case described in Section 4 illustrates how this reduction behaves on a specific attack surface. On the remote-access exposed estate, more than 6 000 scanner-critical items were present at the onset of the pandemic. Within six months, none of those items remained in the P0/P1 category for the remote-work paths in scope. The remaining critical findings contained many legitimate issues,

but they did not significantly affect the defined paths under the particular constraints of that period.

It is important to understand what this means. The model did not make vulnerabilities disappear; it changed which vulnerabilities competed for scarce attention. Thousands of scanner-critical items were still recorded and tracked as P2 or P3. However, they no longer occupied the same priority tier as those issues that could clearly open or maintain high-value attack paths.

6.2 Pathway coverage and false negatives

As the model matured, all organisations reported an increase in pathway coverage: more of the high-value paths that matter for business were anchored on at least one P0 or P1 issue that had been addressed.

In the early stages, post-incident reviews sometimes revealed pathways that the initial modelling had missed. This was expected; in any complex environment, the first pathway maps are incomplete. Over time, however, the proportion of incidents revealing entirely unmodelled pathways declined. In the most mature adopter, this proportion fell from roughly one third of investigated incidents in the first year to below 10 % after three years, as pathway maps and modelling practices improved. The newer versions of the open standards explicitly encourage organisations to track this false-negative pathway rate as a first-class metric.

This trend is important. It shows that pathway-centric thinking is a skill[4] that teams must develop, not a static technology that can be "installed". The model encourages analysts and architects to ask "through which exact steps would this threat actor reach this system?" and to refine their understanding each time reality contradicts their assumptions.

6.3 Incident outcomes and the remote-work period

Incident and breach data are always noisy, and many factors beyond vulnerability management contribute to outcomes. With this caveat, several patterns are worth noting.

During the two years that covered the COVID-19 outbreak and its consolidation, none of the studied organisations experienced a major security incident whose root cause could be traced to an attack pathway that had been explicitly mapped and scored in the model. Incidents did occur, but they tended to involve areas that were either outside initial scope or that revealed previously unseen pathways which were then integrated.

In the automotive group, this is most visible in the contrast between remote-work outcomes in the finance subsidiary and in other entities with similar technology stacks on the same shared infrastructure. Peers that relied primarily on generic criticality sometimes deprioritised the combination of "medium" and "low" findings that later formed the backbone of actual exploit chains. In the finance subsidiary, those combinations had already been raised to P0/P1 and addressed.

Across the full cohort, incident rates per thousand internet-facing assets remained

[4] The certification program attached to the open standards (GCR-M / OSPCRM) includes pathway-modeling training and recurring false-negative pathway rate measurement. This is the right answer to the "juniorization" trend in many SOCs and VM teams: one cannot outsource defender skill to a magic risk score any more than adversaries outsource their skill to automation.

stable or decreased slightly relative to sector benchmarks, even as the number of items treated as top-priority declined sharply. This does not prove that the model is universally superior to all alternatives. It does, however, contradict the intuitive fear that "if we stop treating thousands of items as critical, our incident rate will explode."

6.4 Human factors and analyst effort

Security is a human practice. One of the most striking qualitative observations from teams that adopted the pathway model was the change in mood around vulnerability discussions.

Before adoption, meetings about vulnerability dashboards often devolved into debates about whether a given CVSS score was realistic, whether a compensating control justified reclassifying an item, or whether it was fair to assign SLAs to system owners when the list of "critical" items never seemed to shrink. After adoption, discussions shifted toward concrete pathways: "If we fix this control on this gateway, which paths disappear? How does that interact with the next project on the roadmap?"

Analyst effort estimates indicate a reduction of roughly one third to one half in time spent triaging and disputing low-impact findings. That time did not disappear; it was reallocated to design reviews, tabletop exercises and collaboration with architects. The effect on burnout is harder to quantify, but anecdotal reports describe a greater sense of agency: teams felt that when they closed a P0 or P1 item, something tangible had changed.

6.5 Side-by-side comparisons with modern risk engines

Since the initial deployments, several organisations in the cohort have compared the pathway-centric model with more recent risk-scoring engines that combine vendor severity with global exploit-prediction scores derived from public models such as EPSS.

The picture that emerges is consistent. The pathway model produced a **sharply defined priority distribution**, with a small, stable set of P0/P1 items and a long tail of P2/P3 hygiene. In contrast, some popular platforms tended to cluster large families of findings into broad "critical" categories. These families looked neat on dashboards but often grouped issues with very different remediation patterns and very different relevance to actual attack paths.

Global exploit-likelihood scores behaved as useful background signals but weak local guides. Across several estates, the correlation between exploit-likelihood scores and the presence of a finding on the top two or three kill-chain paths to critical assets was low (on the order of $r < 0.1$). In other words, the fact that a vulnerability is actively exploited somewhere in the world told us little about whether it was the most efficient lever in a given environment.

Some platforms also attempted to reduce noise by flattening related vulnerabilities under umbrella "families". From an operational standpoint, this sometimes disconnected presentation from work: engineers received an abstract family label rather than a clearly articulated path effect, which made it harder to translate dashboards into concrete change requests.

The pathway-centric model inverts this perspective. It starts from local paths and asks which findings move them, then treats global signals as one input among others

rather than as the primary driver of prioritisation. The comparisons suggest that this local, pathway-first stance is what enables the large reduction in top-priority populations without obvious loss in security.

7. Discussion

7.1 Why a small, pathway-focused list matters

The most important effect of the model is not mathematical; it is cognitive and organisational. Humans can hold a limited number of complex items in active attention. A list of 80 to 200 well-explained P0/P1 issues linked to concrete attack paths is something that a leadership team, a CISO and system owners can reasonably review together. A list of 8 000 red icons is not.

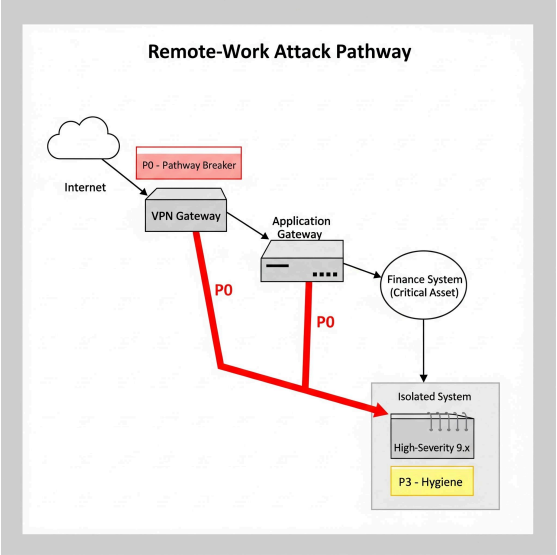

Figure 2 – Simplified remote-work attack pathway: Internet → VPN gateway → application gateway → finance system. A medium-severity configuration weakness on the VPN gateway is classified as P0 (pathway breaker) because fixing it removes the only reliable entry step. A separate CVSS 9.x vulnerability on an isolated, non-routed system is classified as P3 hygiene because it does not currently sit on any path to crown-jewel assets.

By tying priority to roles in pathways, the model also helps avoid superficial reactions to severity labels. When an issue with a medium generic severity is identified as the only remaining bridge between an internet-facing service and a crown-jewel system, it receives focused attention even if it would not normally appear in top-ten reports. Conversely, when an issue with a high generic severity sits on an isolated system with no meaningful path to impact, the model justifies assigning it a lower operational priority without denying its technical importance.

This reframing is not a matter of cosmetics. It changes the conversations that happen in rooms where risk is negotiated. Instead of arguing about a numeric score, participants can point at a path diagram and ask “If we do nothing here for the next twelve months, which realistic attacker benefits, and how?”

7.2 Interaction with widely used tools

Most of the organisations in this study used well-known commercial scanners, exposure-management platforms and ticketing systems. The pathway-centric model did not replace these tools; it wrapped them in an additional layer of reasoning.

This observation matters. The challenge is not that tools are “wrong”. Their core success metrics and interfaces are often centred on coverage, detection counts and raw severity. A platform that demonstrates how many thousands of issues it has found can appear more impressive than one that calmly presents a small set of actionable changes. Reporting and pricing models tend to follow this logic.

The pathway-centric standards invite a different measure of success: not “how

many findings can we generate and track?", but "how many high-value paths have we made unattractive or impossible?" This does not put tools and standards in opposition; it changes what organisations ask of their tools. In practice, teams that adopted the model often configured existing platforms to feed data into the pathway engine, rather than trying to force pathway logic into dashboards that were not designed for it.

7.3 Pathway modelling as a skill

The model depends on the quality of pathway modelling. If teams misunderstand their architecture or threat landscape, they may misclassify issues and gain a false sense of security. This dependency is real and should not be minimised.

At the same time, the case studies show that pathway modelling behaves like other professional skills: it improves meaningfully over the first one to three years if the organisation takes it seriously. When false-negative pathways are treated as learning opportunities and tracked as metrics, teams become better at spotting the kinds of bridges and amplifiers that matter. The standards help by making these metrics explicit, so that modelling quality is visible in governance rather than hidden in individual expertise.

7.4 Cultural and governance inertia

Perhaps the hardest part of adoption is cultural. The most unsettling step for many leaders is giving permission to treat thousands of scanner-critical items as P2/P3 "hygiene" while reserving the P0/P1 label for a much smaller set.

In the FinAuto case, the CISO's decision was driven as much by constraints as by ideology. There was not enough manpower to indulge in alert fatigue. Business continuity under pandemic conditions demanded that remote access be stabilised quickly without burning out the team. Under those conditions, the question "What is the minimum set of changes that kills the most relevant kill chains?" was not a theoretical exercise. It was the only realistic way to proceed.

Audit and regulatory concerns were addressed by framing the model as an open, auditable standard rather than a private heuristic. Every P0/P1 classification can be justified by a concrete pathway, and the test, control and pilot plans mirror the structure of existing frameworks such as NIS-style resilience requirements and operational-risk expectations. This alignment does not remove all friction, but it gives CISOs a concrete story to tell: they are not "ignoring criticals", they are following a documented, externally recognisable standard for prioritisation.

7.5 Limitations and threats to validity

This study has several limitations that readers should keep in mind.

First, the cohort is small and focused on specific sectors and regions -we do not detail our US-based, Africa and Middle East use cases here but they do confirm the same trends however. Second, incident data are imperfect. Not all attacks are detected; not all are recorded with equal precision; and many variables beyond vulnerability management influence outcomes. Our observations are therefore indicative rather than conclusive, even though consistently observed.

Third, survivorship bias is always a concern in experience reports. In this case, all early adopters of the model are still using it five

years later. There is no hidden population of organisations that tried it and abandoned it. This persistence does not guarantee universal applicability, but it suggests that the model has been robust enough to survive budget cycles, audits and leadership changes.

Finally, the comparisons with modern scoring engines are based on a limited set of environments. Other organisations may integrate exploit-prediction scores and asset-criticality signals in ways that align more closely with their own paths. The point here is *not* to dismiss global signals, but to remind readers that they remain global. Local paths still need to be drawn.

It is also important to acknowledge that some newer commercial continuous threat exposure management (CTEM) and "attack-path" platforms have moved in a similar direction. Several major vendors now offer graph-based visualisations of exposure, simulate attacker routes from the internet to assets, and propose pathway-oriented prioritisation. In that sense, the pathway-centric approach described here is not alone. The difference in our case lies in three aspects:

- First, the model is defined as an **open standard**, independent of any particular platform, which allows regulators, auditors and peer organisations to inspect the logic and certify conformance regardless of tooling.
- Second, in the environments we studied, applying the open-standard model on top of existing tools consistently produced a **sharper reduction in top-priority items and clearer governance artefacts** than the default, proprietary scoring in those tools. The goal is not to replace CTEM platforms, but to give organisations a standardised way to evaluate whether the prioritisation they expose truly matches their most important attack pathways.
- Third, where the second difference originates from: these solutions try to cling to global metrics (e.g. reliance on EPSS with weak correlation to actual likelihood for the business), while the open standards are sovereign as in "your own particular context". In addition, they lump vulnerabilities by artificial families and do not always account for actual risk (type of asset which is inherently more critical etc).

7.6 Implications for practitioners and standardisation bodies

For practitioners, the main implication is that a relatively modest shift in perspective can yield substantial benefits. One does not need to replace all tools, nor to build perfect attack graphs. It is enough to:

1. Choose a small number of high-value services.

2. Draw plausible attack pathways to those services.

3. Identify which existing findings sit on those paths as entry points, bridges or amplifiers.

4. Assign P0/P1 to those items and treat them as the first backlog, with explicit agreement from stakeholders.

For standardisation bodies, the case studies suggest that there is value in defining such practices as standards, not just

recommendations. When P0–P4 semantics, pathway descriptions and governance artefacts are formalised, implementations can be audited, compared and improved without reinventing the logic each time. It also becomes easier for regulators and investors to understand what it means when an organisation claims to have "consolidated its critical backlog": they can ask "according to which open standard, and how can we verify this?"

here—top-priority population, pathway coverage, false-negative pathways, incident rates per asset—and compare them to the ranges observed in this cohort. The goal is not to impose rigid thresholds, but to encourage a shared language. When a CISO says "we have 120 P0/P1 issues and less than 10 % of incidents reveal unmodelled paths", peers have a concrete sense of what that implies.

---

8. Reproducibility and independent verification

Anchoring the model in open standards allows reproducibility beyond the original environments.

In practical terms, reproducibility has three layers:

1. **Scoring logic**. The standards describe how to map findings and pathways to P0–P4. An implementation that adheres to them should, given the same input pathways and findings, produce the same classifications. This can be tested with reference datasets published alongside the standards.

2. **Governance artefacts**. The test plans, control plans and pilot plans used in the case studies can be published in anonymised form as templates. Other organisations can adopt or adapt them with minimal effort, which makes it easier to compare maturity and outcomes across contexts.

3. **Outcome metrics**. Organisations can compute the same metrics used

Independent verification can also be provided by certifying bodies that assess whether a given implementation conforms to the standards. This is analogous to how control frameworks are currently audited, with an important difference: the standards here focus specifically on the transformation from raw findings to actionable pathways and priorities. That transformation is where overload either arises or is avoided.

In addition to internal metrics, at least one implementation of the model has undergone **third-party certification** against the open standards, with a published conformance identifier (for example, implementation IDs in the IGS-C registry such as *IGS-IMP-00001*). In another case, a large European banking group invited an external Big-4 audit firm to review its vulnerability management transformation. The auditors examined the pathway-centric methodology, verified that it aligned with the organisation's control framework and regulatory obligations, and explicitly accepted it as the primary basis for prioritising remediation, in place of purely CVSS-driven backlogs. These independent assessments reduce the risk that the model is perceived as a purely internal invention and show that it can withstand external scrutiny.

9. Conclusion

This article has described how a pathway-centric, open-standard risk model behaved in five large European organisations over five years, including the turbulent period of the first COVID-19 wave. By reframing vulnerability management around concrete attack paths and a small, disciplined priority set, these organisations were able to reduce their top-priority backlog by more than 90 % without any observable deterioration in incident outcomes. In one emblematic case, more than 6 000 scanner-critical items were reduced to zero pathway-critical (operationally critical) items issues on a crucial remote-access surface in under six months, precisely when that surface was under the greatest external pressure.

The central message is deliberately simple. Alert exhaustion in vulnerability management is not a fixed property of the digital world. It is a consequence of how we score, aggregate and present risk. When every technically severe issue is treated as operationally critical, nothing feels truly urgent. When we instead ask “Which findings actually break or constrict the paths that matter most?”, the list shrinks to a size that humans and institutions can handle.

For organisations that recognise themselves in the overload described in the introduction, a concrete first step is to pilot the pathway model on a single, high-value service. Draw the paths; identify P0/P1 issues; align them with stakeholders; observe how this affects remediation and incident patterns; then decide whether to expand. For standardisation bodies and regulators, the next step is to sharpen and disseminate open standards that encode these practices, so that they become part of ordinary governance rather than exceptional projects.

The early experiments have shown that such an approach can work in practice, under pressure, on shared infrastructure, and in diverse regulatory environments. The remaining challenge is no longer to prove that it is possible, but to make it normal.

10. Data availability

The paper is based on operational data from multiple regulated organisations. For obvious security and confidentiality reasons, raw vulnerability datasets and detailed incident logs cannot be published openly. However:

- Aggregated metrics and distributions are included in the text and figures.
- The structure of the test plans, control plans and pilot plans is described in enough detail to be reproduced.
- In specific cases (for example regulatory reviews, independent audits or academic replications), selected anonymised artefacts may be made available under non-disclosure agreement with the relevant organisations.

The intention is to balance transparency with the obligation not to expose sensitive infrastructure details.

---

11. Author roles and interests

The lead author works as an external security architect and researcher. He is not employed by, nor does he hold equity in, the organisations described in this paper. His role in the case studies was that of an external counsellor, helping teams design and apply pathway-centric prioritisation while leaving all technology and vendor choices under the organisations' control.

The lead author volunteered as a contributor to the open standards that formalise the pathway-centric risk model (GCR-M and OSPCRM) but does not hold any commercial licensing rights over their adoption. The standards are published under open, vendor-neutral, royalty-free terms.

Co-authors from participating organisations contributed operational data, reviewed the accuracy of anonymised case descriptions, and confirmed that the reported results reflect their experience. The independent consultant co-author reviewed the methodology and challenged the interpretation from outside the standardisation process. No vendor funding, product marketing budget or paid standardisation mandate influenced the design, implementation or interpretation of this work.

Acknowledgements

The authors thank the security and IT operations teams in the participating organisations, whose willingness to experiment under pressure made this work possible, and the members of the open-standards consortia who helped refine the model and its governance. Particular gratitude is due to the practitioners who insisted on testing the pathway-centric approach during the most stressful months of the COVID-19 transition, and to the analysts who documented their experiences with honesty, including instances where the model initially failed and had to be corrected.

Annex A Literature Review

A.1 Alert fatigue and the human limits of SOCs

A growing body of research and practitioner testimony converges on the same picture: modern security operations centres (SOCs) are structurally overloaded. An in-depth recent survey of SOC environments notes that analysts face "a continuous stream of heterogeneous alerts" and describes alert fatigue as a central obstacle to effective defence, not an incidental annoyance. Industry analyses echo this: typical SOCs receive thousands of alerts per day from SIEMs, endpoint tools, intrusion detection, cloud monitors and vulnerability scanners, many of which are false positives or low-consequence notifications.

The cognitive impact is measurable. Studies and CISO surveys increasingly cite human fatigue, error and burnout as primary cyber-risks; in one widely cited series, around two-thirds of CISOs identified "human risk" driven by overload as their top concern. When everything looks like a potential incident, it becomes progressively harder to distinguish real threats from

harmless noise. Recent analyses of major breaches repeatedly highlight that relevant alerts were present in logs but buried in a mass of routine warnings that analysts no longer treated as credible signals.

Although much of this writing focuses on SIEM and detection alerts, the same dynamics apply in vulnerability management. Scanners and configuration tools can easily produce tens of thousands of findings, with a non-trivial fraction labelled "critical" or "high" according to generic severity scales. The result is a form of "slow alert fatigue": backlogs that never clear, SLAs that are frequently missed and a gradual erosion of trust in dashboards that stay red regardless of effort. The literature agrees on the symptoms but is more divided on how to escape the overload paradigm.

A.2 Traditional severity scoring and the rise of "risk-based" approaches

Conventional vulnerability practice centres on the Common Vulnerability Scoring System (CVSS), which encodes technical severity along several dimensions and assigns a score that is usually mapped to critical/high/medium/low buckets. CVSS has been extremely successful as a shared vocabulary, but its limitations are now well documented: it measures innate technical properties of a vulnerability, not the actual likelihood of exploitation or the local value of the affected asset.

As estates and vulnerability volumes grew, many organisations tried to move beyond CVSS with "risk-based vulnerability management". Vendor and integrator white-papers typically propose combining CVSS with additional signals such as asset criticality, exposure (internet-facing vs internal), known exploitation and sometimes business process impact. In principle, this should reduce noise by surfacing only those findings that materially change risk.

However, the literature also shows how easily these attempts slide back toward familiar patterns. In many implementations, risk-based scores are still presented as extended critical/high/medium/low labels and still generate large clusters of "critical" items. An emerging line of work acknowledges this problem explicitly and frames the core challenge as attention, not data: generating more nuanced scores does not help if they still produce backlogs that are too large for humans to handle. This is precisely the bottleneck our study addresses from a pathway perspective.

A.3 Exploit prediction (EPSS and related models)

Exploit prediction models are a prominent attempt to inject empirical threat data into prioritisation. The Exploit Prediction Scoring System (EPSS), maintained by FIRST, estimates the probability that a given CVE will be exploited in the next 30 days using machine-learning models trained on public exploit and telemetry data. EPSS is explicitly presented as complementary to CVSS: CVSS captures severity, EPSS captures estimated near-term threat.

Subsequent practitioner analyses explore how to integrate EPSS into risk-based workflows, for example by using high EPSS scores as a filter on top of CVSS or by combining EPSS with curated lists such as CISA's Known Exploited Vulnerabilities (KEV). Recent academic work on "vulnerability management chaining" proposes decision trees that combine CVSS, EPSS and KEV signals into prioritisation rules and reports improved efficiency compared to CVSS alone.

At the same time, several authors and practitioners note important limitations. EPSS is a global, black-box model optimised for average predictive performance across many environments, not for any single organisation's architecture. As a result, a vulnerability can legitimately have a low global exploitation probability and still be the most efficient pivot on a specific, high-value path in a given estate. Some empirical comparisons report weak correlation between EPSS scores and the vulnerabilities that feature on the two or three most important kill chains in a particular organisation, especially when those chains involve local misconfigurations, identity weaknesses or architecture-specific quirks that do not generalise well across datasets.

The literature thus positions EPSS and similar models as useful background signals, but not as full replacements for local reasoning about how attackers will actually navigate a specific environment. Our work builds directly on that recognition: we treat global exploit-likelihood as one input, but we anchor operational priority in explicit, organisation-specific pathways.

A.4 Attack graphs, kill chains and attack-path management

The idea that attacks should be studied as sequences of steps rather than isolated events has deep roots in the research literature. Attack graphs and their variants model networks as nodes and edges representing reachability and exploit relations, and then compute paths an attacker could take from entry points to targets. Foundational work in the early 2000s showed how such graphs could be generated automatically and used to identify "choke points" where defences would have disproportionate effect.

More recent work combines attack graphs with the cyber kill chain, proposing models that annotate graph nodes with kill-chain phases and thus help defenders focus on particular stages of an attack. These approaches emphasise that understanding **where** in the chain a vulnerability sits is as important as understanding its technical severity: a modest weakness at an early stage can be more strategically important than a severe weakness on an already compromised system.

Beyond academic work, commercial and open-source tools increasingly expose some form of attack-path view. Threat-exposure management platforms advertise "attack path management" capabilities that integrate vulnerability data, identity information and network topology to visualise routes from the internet to critical assets. Research on AI-driven "graphs of effort" and kill-chain prediction systems goes further, modelling the effort required for attackers to chain vulnerabilities, misconfigurations and AI capabilities into end-to-end attacks.

Despite these developments, most of the literature treats attack-path modelling either as a specialised analysis technique or as a feature inside particular tools, rather than as the primary organising principle for day-to-day prioritisation and governance. Path views are often used for "interesting" reports or red-team visualisations, while routine prioritisation still follows per-vulnerability scoring. The contribution of our work is to take the path perspective seriously at the top level of operational decision-making: we treat the role of a finding in concrete paths as the main driver of priority, not an after-the-fact illustration.

A.5 Governance, standards and the gap in day-to-day practice

The last strand of relevant literature concerns governance and standardisation. Control frameworks and regulations—whether sector-specific or horizontal—strongly influence how organisations document their vulnerability processes. Recent guidance on cyber-resilience and threat exposure increasingly calls for "risk-based" or "contextual" approaches, but often stops at high-level principles. In practice, many institutions struggle to show auditors **how** their risk-based approach actually transforms raw findings into a small, defensible set of actions.

A few recent contributions move in the direction of more operational clarity. Some industry papers outline decision trees that combine CVSS, EPSS and curated exploit lists into rules; others recommend overlaying business impact assessments and asset criticality on vulnerability data. But the majority of these proposals remain embedded in proprietary approaches, and their internal logic is not always fully transparent to external reviewers.

There is comparatively little literature on **open, vendor-neutral standards** that encode prioritisation logic in a way that can be audited independently of any specific platform. Most work on standards focuses on data formats (for example for exchanging vulnerability information) or on control requirements rather than on the actual transformation from scanner output to prioritised backlog.

This is precisely the gap that the standards underlying our model aim to fill. They formalise the P0–P4 semantics, define how pathways should be described, specify governance artefacts (test plans, control plans, pilot plans) and introduce benchmark metrics such as false-negative pathways and P0/P1 precision. By doing so, they provide a missing link between rich but often under-used attack-path concepts and the everyday reality of overburdened SOCs and vulnerability teams.

A.6 Positioning of this study

Taken together, the literature paints a clear picture. Alert fatigue and overload are now recognised as systemic issues in SOCs and vulnerability management. Traditional severity-based scoring provides a common language but frequently floods practitioners with too many "critical" items. Exploit-prediction and risk-based extensions add valuable signals, yet they struggle to account for the specifics of local architectures and kill chains. Attack graphs and kill-chain models demonstrate the power of path-based reasoning, but they rarely drive day-to-day prioritisation or appear in standard governance artefacts.

Our study sits at the intersection of these threads. It examines what happens when an organisation chooses to put **pathways**, rather than per-issue severity, at the centre of its vulnerability management, and when that choice is expressed not only in tooling but in open standards, control plans and pilot criteria. Where much of the existing work remains either conceptual or tool-centric, we focus on five years of concrete operational experience under regulatory scrutiny, including a natural experiment on shared infrastructure during an exceptional period of stress.

In that sense, the work does not replace the existing literature on risk-based or exploit-aware prioritisation; it operationalises it. We take seriously the proposition that the real bottleneck is human attention, and we test what happens when everything in the process—from metrics to dashboards to audit trails—is re-aligned around a small, top-priority,

path-breaking items subset of issues rather than a large, severity-critical population.

References

[1] Jacobs, J., Romanosky, S., Edwards, B., Roytman, M., & Adjerid, I. (2019). *Exploit Prediction Scoring System (EPSS).* arXiv preprint arXiv:1908.04856.

[2] FIRST. (n.d.). *Exploit Prediction Scoring System (EPSS).* Retrieved from the FIRST website.

[3] Mell, P. (2025). *A Proposed Metric for Vulnerability Exploitation Probability.* NIST Cybersecurity White Paper.

[4] Renney, H. (2024). *An Open and Adaptable Approach to Vulnerability Risk Scoring.* Pre-print report on V-Score as a contextual alternative to CVSS.

[5] Anonymous. (2024). *Alert Fatigue in Security Operations Centres: Research Challenges and Opportunities.* (Preprint). A survey of causes of SOC alert fatigue and mitigation approaches via automation and human–AI collaboration.

[6] Chuvakin, A. (2024, November 6). *Anton's Alert Fatigue: The Study.* Medium. Discussion of long-term alert fatigue patterns in SOCs and implications for detection and response.

[7] Axur Content Team. (2024, May 7). *Alert Fatigue: Uncover Causes and Impacts in Cybersecurity.* Axur Blog.

[8] DefendEdge. (2025, April 28). *Burnout and Alert Fatigue in Cybersecurity.* DefendEdge Blog.

[9] Databahn. (2025, November 13). *The Cybersecurity Alert Fatigue Epidemic.* Databahn Blog.

[10] Zafran CTEM Academy. (2025). *Prioritizing Vulnerabilities: Best Practices for Risk-Based Patching.* Zafran.io.

[11] Arnica. (2024, February 20). *Leveraging EPSS, CVSS, and KEV for Comprehensive Risk Management.* Arnica Blog.

[12] Anonymous. (2025). *Vulnerability Management Chaining: A Decision-Tree Framework Combining CVSS, EPSS, and KEV.* arXiv preprint (Vulnerability Management Chaining).

[13] Palma, A., et al. (2025). *Behind the Scenes of Attack Graphs: Vulnerable Network Paths and Security Posture.* Computers & Security.

[14] Anonymous. (2024). *Survey on Application of Attack Graph Technology.* DOAJ / Journal article summarising attack graph applications for predicting attacker behaviour and guiding defence.

[15] XM Cyber. (2023, August 18). *Leveraging Attack Paths to Achieve CTEM Goals.* XM Cyber Blog.

[16] Bradley, T. (2025, February 10). *Transforming Cybersecurity with Continuous Threat Exposure Management (CTEM).* Forbes.

[17] Wiz. (2025). *What is CTEM? Continuous Threat Exposure Management.* Wiz Academy.

[18] Puppygraph. (2025, September 29). *What is Continuous Threat Exposure Management?* Puppygraph Blog.

[19] Vicarius. (2025, September 8). *Implementing Continuous Threat & Exposure Management (CTEM) to Reduce Attack Surface.* Vicarius Blog.

[20] Anonymous. (2025). *Alert Fatigue in Security Operations Centres: Research Challenges and Opportunities.* In-depth survey of automation and human–AI approaches to SOC overload. (Listed here separately to emphasise its dual relevance to sections on human factors and alert fatigue.)

[21] CVE Shield. (2023–2024). *V-Score Documentation and Evaluation Results.* Vendor white-paper demonstrating transition from CVSS-centric to contextual scoring in an enterprise VM team.

[22] Sheyner, O., Haines, J., Jha, S., Lippmann, R., & Wing, J. (2002). **Automated generation and analysis of attack graphs.** *Proceedings of the 2002 IEEE Symposium on Security and Privacy (S&P)*, 273–284. IEEE.

[23] Noel, S., Jajodia, S., O'Berry, B., & Jacobs, M. (2003). **Efficient minimum-cost network hardening via exploit dependency graphs.** *19th Annual Computer Security Applications Conference (ACSAC)*, 86–95. IEEE.

[24] Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2011). **Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains.** In *Proceedings of the 6th International Conference on Information Warfare and Security (ICIW)*, Academic Publishing Limited.

[25] Werlinger, R., Hawkey, K., & Beznosov, K. (2009). **An integrated view of human, organisational, and technological challenges of IT security management.** *Information Management & Computer Security*, 17(1), 4–19.

[26] Sarker, I. H., Kayes, A. S. M., Badsha, S., Alqahtani, H., Watters, P., & Ng, A. (2020). **Cybersecurity data science: An overview from machine learning perspective.** *Journal of Big Data*, 7(1), 41. (Useful for positioning data-driven scoring/EPSS vs contextual approaches.)

[27] Kuerbis, B., Beato, F., & Van Eeten, M. J. G. (2022). **Cybersecurity incident response and the human factor: A systematic literature review.** *Computers & Security*, 112, 102514. (For SOC burnout / human limitations in incident handling.)