The Value of Information

Why Decision-Aware Analysis Is the Right Direction of Inquiry — and What the “Collect Everything” Paradigm Misses

A Kindynos Advisory Whitepaper

Prepared by Kindynos Advisory

Public Whitepaper — Kindynos Advisory
Public release version 1.0 · 20 June 2026

Preface — the inverted question

The dominant operating assumption in data analytics, risk management, and machine learning today is that more data is better, that wider coverage produces better decisions, and that the rational response to uncertainty is to collect everything one can and apply increasingly sophisticated algorithms to the result. The Big Data paradigm of the 2010s made this assumption explicit: as Halevy, Norvig, and Pereira observed in their 2009 paper "The Unreasonable Effectiveness of Data," simple models trained on very large datasets often outperform more elaborate models trained on smaller ones. *The early Big Data intuition held that scale of data could substitute for theory and model structure.* The implied logic is that uncertainty is a data problem, and that data, eventually in sufficient quantity, dissolves it.

This whitepaper argues that the logic is inverted. Uncertainty is not, in its essential form, a data problem. It is a *decision* problem. The right question is not "what data should I collect to reduce uncertainty?" but "what decision am I trying to make, what are the consequences of being wrong, and which pieces of information would actually change the decision?" Without an explicit consequence model — without understanding what events will materially affect outcomes and what game-theoretic responses are available — the question of *which* data matters cannot even be posed, let alone answered. The collect-everything paradigm is not merely inefficient; it is structurally backward. It treats data as if it had intrinsic value, when in fact information has value only relative to a decision it might change and the cost of being wrong about that decision.

There is a precise academic name for the framework that makes this argument formally: **Value of Information** (VoI) analysis, a roughly sixty-year tradition in decision theory, health economics, and environmental management. VoI flips the question. Rather than asking "how much data should we collect," it asks "what is the maximum amount a decision-maker would rationally pay to acquire a given piece of information before making a decision?" That maximum, computed from the consequence structure of the decision itself, can be — and often is — zero. The cutting edge of machine-learning research has, in the past several years, formalized the same inversion under a different name: **Decision-Focused Learning** (DFL) and the **Predict-then-Optimize** framework, both of which train predictive models against the downstream *decision quality* they enable rather than against generic prediction accuracy. The fields converge on the same insight from independent directions: the consequence structure of the decision must come first; the data follows.

> *The argument has direct architectural consequences. A risk-management system built on the data-first paradigm collects everything it can about every holding, every market, every macroeconomic variable, runs the result through a battery of predictive models, and presents the user with a dashboard. A risk-management system built on the consequence-first paradigm starts from a model of how events propagate through a*

> *portfolio, identifies which events have material impact, and only then directs data collection toward the events whose probability or magnitude actually matters. The first system is a common default across AI and analytics platforms. The second is harder to build but is what every serious operational risk practitioner does informally, every day, with their own attention. This paper makes the case that the second approach is not a niche preference; it is the correct architectural commitment for any decision system whose output materially affects outcomes.*

The argument proceeds in five parts. Section 1 states the inversion sharply and locates it in the academic literature. Section 2 reviews Value of Information theory, with worked formulations and the historical and contemporary critique of the collect-everything paradigm. Section 3 surveys the contemporary academic field — Decision-Focused Learning, Predict-then-Optimize, decision-aware ML — and traces how serious researchers have arrived at the same conclusion from the algorithmic side. Section 4 describes how this principle is operationalized in EVA, the deterministic risk-management platform Kindynos Advisory has built: the event propagation engine and the game-theoretic decision apparatus together are designed to constitute the consequence model needed to make VoI-style evaluation mechanically testable for covered portfolios, event sets, and decision surfaces. Section 5 states the implications for enterprises currently invested in the collect-everything paradigm and proposes a different starting question for risk-management technology procurement.

1. The inversion, stated sharply

There are two ways to approach decision-making under uncertainty.

The **first way**, which dominates modern data analytics and machine learning, begins with data. The presumed pipeline is: collect as much information as possible about the system in question, train predictive models on that information, and present the predictions to a decision-maker who then uses them to choose an action. The implicit theory of value is that information has intrinsic worth — that more information is, by default, better than less. The engineering effort is directed at expanding data acquisition, improving model accuracy, and presenting outputs at scale.

The **second way**, which formal decision theory has known how to formalize for at least the 1960s, begins with the decision. The pipeline is inverted: identify the decision and its possible actions, construct an explicit model of the consequences of each action under each possible state of the world, identify which states of the world materially change the optimal action, and *only then* ask what information would help discriminate between those states. The implicit theory of value is that information has worth only insofar as it changes a decision whose consequences differ across the states the information can distinguish.

These two approaches are not interchangeable. They produce different answers to "what should I collect?" — sometimes radically different. A piece of information that the first approach would prioritize highly (e.g., a high-frequency time series of a

variable that is statistically interesting) the second approach will discard if the variable's possible values do not span the threshold at which the optimal action changes. Conversely, a piece of information the first approach would deprioritize (because it is sparse, hard to collect, or has low statistical signal in isolation) the second approach may identify as critical if it materially affects which action is optimal.

The inversion has a precise mathematical name in decision theory: the **Value of Information** (VoI) of a piece of evidence is the expected reduction in the cost of suboptimal decision-making that the evidence would produce. As the canonical formulation states it: "The expected value of a research project is the expected reduction in the probability of making the 'wrong' decision multiplied by the average consequence of being 'wrong' (the 'opportunity loss' of the decision)." If a piece of information cannot change the optimal action, its value is zero, regardless of how interesting, how voluminous, or how computationally tractable it is. If it can change the optimal action and the consequences of the wrong action are large, its value is large. The framework provides a derivable, defensible answer to "is this data worth collecting?" that depends entirely on the consequence structure of the decision, not on the statistical properties of the data itself.

This is the inversion. The collect-everything paradigm treats data as the input and decisions as the output. VoI treats decisions as the input — specifically, the consequence model of decisions — and data as the output of a derivation. The two paradigms are not points on a spectrum; they are different architectures of inquiry.

> *The clearest single illustration of why the inversion matters comes from a historical example that pre-dates modern machine learning by nearly a century. In 1936, The Literary Digest attempted to predict the outcome of the United States presidential election using what was, at the time, the largest political poll ever conducted: 2.3 million completed questionnaires. By the standards of the day — and by the standards the contemporary Big Data paradigm still implicitly endorses — this was an extraordinary data collection effort. The Digest confidently predicted that Alf Landon would defeat Franklin Roosevelt. Roosevelt won by a landslide. Meanwhile, George Gallup, working with a sample of approximately 3,000 carefully chosen individuals — three orders of magnitude smaller — predicted the actual outcome correctly. The 2.3 million responses were biased; the three thousand were chosen to reflect the electorate's actual structure. Volume is not value. Coverage is not relevance.*

The lesson generalizes. The right amount of data, and the right type of data, is determined by the structure of the decision the data is meant to inform — and the structure of the decision can be examined independently of, and prior to, the data itself.

2. The Value of Information framework

The Value of Information framework has a mathematical structure that is worth stating explicitly, because the structure is what makes the framework actionable rather than merely rhetorical. The formulation in its modern form dates to the 1960s in the operations-research and Bayesian-decision-theory literatures, was substantially developed in health economics in the 1990s and 2000s, and has been applied across environmental management, energy systems, geospatial information policy, and clinical-trial design. The framework's longevity is itself a signal: it works, it has been refined by serious researchers over decades, and it remains a durable and widely used framework for deciding whether information is worth acquiring — including the Big Data paradigm, which addresses a different (and arguably less important) question.

2.1 The formal structure

The framework rests on four components, each of which must be specified before the value of any piece of information can be computed:

1. A decision space. The set of actions available to the decision-maker. For a risk manager: hedge or do not hedge, allocate or do not allocate, hold or sell. For a clinical trial: approve or do not approve. For an environmental policy: mitigate, adapt, or accept. The decision space must be finite, named, and operationally meaningful.

2. A consequence model. A function that maps each (action, state-of-the-world) pair to an outcome the decision-maker cares about. The outcome is typically measured in a unit relevant to the domain — dollars, lives, regulatory exposure, expected return, risk-adjusted return. The consequence model is the most demanding requirement of the framework and the one the collect-everything paradigm most often skips.

3. A prior distribution over states of the world. What the decision-maker believes today, before acquiring any new information, about which states are how likely. The prior is updated by Bayesian conditioning when new information arrives.

4. A loss function. A way of valuing outcomes, including risk preferences (risk-neutral, risk-averse, risk-seeking). For most operational risk applications, expected-value calculations are appropriate, sometimes adjusted for tail-risk or regulatory capital.

Given these four components, the Value of Perfect Information (VPI) — the maximum amount a decision-maker would rationally pay to know the true state of the world before deciding — is computable directly: it is the expected value of the decision under perfect information minus the expected value of the decision under the prior alone. The Value of Sample Information (VSI) — the value of acquiring a specific piece of evidence that updates but does not perfectly resolve the prior — is computable similarly, by integrating over the conditional distributions the evidence

would induce. Both are amounts of money (or whatever unit the consequence model uses) and both are derivable from the four components above.

Four properties of VoI are worth highlighting because they constrain how the framework can be used (per the standard reference):

> *The value of information can never be less than zero, since the decision-maker can always ignore the additional information. No other information-gathering activity can be more valuable than that quantified by value of clairvoyance [perfect information]. Observing multiple new evidences yields the same gain in maximum expected utility regardless of the order of observation. The VoI of observing two new evidence variables is not additive.*

The first property is the floor: information is never harmful in the Bayesian framework, only useless (this is a theoretical floor — in practice, biased or misleading information can absolutely produce worse decisions, which is why the *Literary Digest* example matters). The second property is the ceiling: no real information system can exceed the value of perfect knowledge, which is finite and computable. The third and fourth are subtler: ordering of evidence doesn't change the final value, and evidence values do not add — they interact.

2.2 Three conditions for high VoI

The framework lets us make a sharper statement: there are three primary conditions under which a piece of information has high value, and absent any of them the information has low or zero value regardless of its statistical properties.

Condition one — the information must be capable of changing the decision. If, under the decision-maker's prior, the optimal action is "hold" and under every realization of the new evidence consistent with the prior the optimal action remains "hold," the information's value is zero. The information may be statistically interesting, may be hard to acquire, may be voluminous; but if it cannot change the action, it cannot reduce opportunity loss, and its VoI is zero. This is the most common reason data-collection efforts produce low return: the data could not have changed the decision regardless of what it showed.

Condition two — the consequences of the wrong decision must be material. Even if the information could change the action, if the two possible actions produce nearly identical outcomes, the value of resolving the uncertainty is small. The framework formalizes "material" precisely: the consequences are material when the opportunity loss of the wrong action is large in the units the decision-maker cares about.

Condition three — the prior must be uncertain enough that the information actually shifts the posterior. If the decision-maker already knows, with high confidence, which action is optimal, additional information has limited value because it is unlikely to change the conclusion. This is why VoI is high precisely when the decision-maker is most uncertain about the right answer.

The three conditions together produce a derivable test for any proposed data-collection effort. The test the collect-everything paradigm cannot apply, because it has no consequence model against which to apply it, is: "given the decisions we make and the consequences of getting them wrong, what is the maximum we would rationally pay for this data?" If the answer is small or zero, the data is not worth collecting regardless of how easy or interesting it is.

2.3 The Big Data critique

The Big Data paradigm — the intuition that scale of data can substitute for theory and model structure — has been substantively critiqued in the academic literature on grounds that align directly with the VoI framework. A peer-reviewed paper in *AI Magazine* in 2022, "Decoding human behavior with big data?", argues: "Big data analytics is by design atheoretical and does not provide process-based explanations of human behavior, making it unfit to support deliberation that is transparent and explainable… The accuracy of complex statistical and machine learning algorithms used in big data analytics in predicting human behavior is not consistently higher than that of simple rules of thumb; rather, it has been found to be lower in situations such as predicting election results, criminal profiling, and granting bail."

A second peer-reviewed critique, "Impact of Biases in Big Data" (Glauner, Valtchev, and State, 2018), opens with the *Literary Digest* example precisely because it illustrates the core failure mode: voluminous-but-unrepresentative data produces confidently wrong answers. The authors observe: "The underlying paradigm of big data-driven machine learning reflects the desire of deriving better conclusions from simply analyzing more data, without the necessity of looking at theory and models. Is having simply more data always helpful?" The implicit answer the paper supplies is: not when the data is biased relative to the decision being made.

The contemporary industrial-engineering literature reaches the same conclusion through different language. A 2021 analysis of the "more data equals better results" assumption notes that voluminous data routinely produces biased models that fail in deployment. The 2026 reanalysis in *Towards Data Science* similarly cautions that for many production machine-learning systems, additional data can produce minimal or no improvement past a relatively low threshold, and that effort directed toward better model specification — i.e., toward better understanding of what the model is *for* — produces substantially larger improvements.

> *The convergence across the literatures is striking. Decision theorists (the VoI tradition) say: data has value only relative to a decision and its consequences. ML researchers (the bias and overfitting literature) say: more data without theory can produce biased, brittle, or mis-specified models. Behavioral scientists (the AI Magazine critique) say: theory-grounded simple rules often outperform theory-free complex models. The fields use different vocabulary; they say the same thing. The Big Data paradigm, taken as a universal answer to decision-making under uncertainty, is structurally inadequate. The right approach starts with the consequence model.*

3. The academic field — Decision-Focused Learning

In the last several years, the machine-learning research community has independently arrived at the same conclusion the VoI tradition reached decades earlier, and has formalized it under two related names: **Predict-then-Optimize** (PtO) and **Decision-Focused Learning** (DFL). The literature is recent, technical, and rapidly growing. It is also unambiguous about the direction of the inversion: predictive models should be trained to maximize the quality of the decisions they enable, not the accuracy of their predictions in isolation. This is the VoI inversion stated as a training objective. The two are not identical — VoI asks whether information is worth acquiring before a decision, while DFL asks how predictive models should be trained when their outputs feed a downstream optimization — but they share the principle that prediction and data collection should be judged by downstream decision quality, not generic accuracy.

The canonical reference is **Elmachtoub and Grigas (2022)**, which introduced the "Smart Predict, then Optimize" (SPO+) framework and articulated the central problem: machine-learning models trained to minimize prediction error (e.g., mean squared error on a forecasting task) will often produce predictions whose downstream decision quality is far worse than predictions from a less accurate model that was trained against decision regret instead. The intuition is straightforward: prediction errors that are large but irrelevant to the decision do not matter; prediction errors that are small but happen to fall across the threshold at which the optimal action changes matter enormously. Generic prediction-loss training weights all errors equally, which is exactly the wrong weighting if the goal is to make good decisions.

A 2023 comprehensive review of the field, "Decision-Focused Learning: Foundations, State of the Art, Benchmark and Future Opportunities," frames the field's central contribution: "DFL is an emerging paradigm in machine learning which trains a model to optimize decisions, integrating prediction and optimization in an end-to-end system. This paradigm holds the promise to revolutionize decision-making in many real-world applications which operate under uncertainty, where the estimation of unknown parameters within these decision models often becomes a substantial roadblock." A 2024 follow-up extends the framework with bidirectional feedback between optimization and prediction stages; a 2025 paper specializes predict-then-optimize for settings such as cost-coefficient prediction in linear programs.

The unifying claim across this literature is precisely the VoI inversion. As one 2022 NeurIPS paper puts it: "Predictive models are typically learned independently of the downstream optimization task in machine learning–based decision-making systems, and recent work in the predict-then-optimize setting has shown that it is possible to achieve better task-specific performance by tailoring the predictive model to the downstream task." Tailoring the predictive model to the downstream task is, in VoI terms, structuring data collection and modeling around the decision and its consequences rather than around generic predictive accuracy.

The field also produces direct empirical evidence for the inversion. A 2025 application to ride-sourcing subsidy allocation (using real-world data from Didi Chuxing, a large ride-hailing platform) compared decision-focused training against conventional predict-then-optimize on two real-world problems. The decision-focused approach consistently produced better decision quality despite producing prediction-accuracy metrics that were, in some cases, slightly worse. A 2025 application to seaport power-logistics scheduling makes the operational case directly: “Asymmetric and nonlinear effects of forecast errors on scheduling are poorly captured by symmetric, linear statistical losses such as MSE. Decision-focused learning instead evaluates the decision quality of forecasts through their impact on the specific downstream optimization tasks.”

This is the academic ML field, in its own vocabulary and against its own benchmarks, reaching the same conclusion as the VoI tradition: the decision and its consequences come first. The data and models follow. The Big Data paradigm, with its emphasis on volume and predictive accuracy as ends in themselves, is increasingly complemented, in the research literature, by approaches that start from the decision structure.

4. The EVA architecture — VoI made operational

The framework above has direct architectural consequences for risk-management technology. This section describes how Kindynos Advisory’s EVA platform (Event Intelligence) is architected to operationalize the consequence-first inversion. The description is intentionally architectural rather than commercial: the purpose is to show what the inversion looks like when built end-to-end, not to make a product case.

4.1 The two layers EVA was built around

EVA was designed around two layers that together constitute the consequence model the VoI framework requires:

The event propagation engine. A deterministic, rules-based engine that models how a financial event — a regulatory change, a supply-chain shock, a macroeconomic dislocation, a geopolitical incident — propagates through a portfolio of holdings. For each holding, the engine computes a per-event impact estimate with explicit valuation provenance: which tiered valuation model produced the estimate (fundamental, causal, heuristic, residual, asset-specific), which probability source produced the event probability (catalog calibration, ML predictor, belief contract, composed), which causal narrative supports the impact. The engine is the *consequence model* in VoI’s vocabulary: it answers the question “if this event occurs, what changes for this holding?”

The decision-strategy layer. A game-theoretic apparatus that takes the impact analysis as input and produces actionable strategies: which holdings to hedge, which to rotate, which pairs to construct, which positions to size up or down. The layer applies an explicit decision-quality objective (risk-adjusted expected return, subject

to constraints) and produces ranked recommendations with per-leg theses, sub-scores, and provenance. The layer is the *decision-quality optimization* in DFL's vocabulary: it determines which actions are optimal given the consequence model and produces them with full traceability.

Together, the two layers are designed to constitute an end-to-end consequence model. In principle, any proposed data acquisition can be evaluated against the two layers: if a piece of data would not change the impact propagation for any portfolio, or would not change the strategy recommendations the decision layer produces, its VoI in the EVA context is zero; if it would change both, its VoI is in principle the expected reduction in decision regret it would produce. This is the architectural basis on which the framework's central question — "what is this data worth?" — can be posed; EVA is built to make this evaluation possible rather than to compute it as a turnkey figure today.

4.2 Why this matters for data-acquisition strategy

The EVA architecture inverts the standard data-acquisition workflow. The standard workflow is: identify data that exists, acquire it, run it through models, present the result. The EVA workflow is: identify the events that materially affect portfolios in the relevant ecosystems, determine which probabilities and valuation parameters are decision-sensitive, and *only then* direct data collection toward those probabilities and parameters. Data whose acquisition would not move the decision is deprioritized regardless of its statistical interest. Data whose acquisition would discriminate between actions whose consequences differ materially is prioritized regardless of its difficulty.

Consider a consequence-first system in which some events are well-calibrated and others are not yet calibrated. A naive data-acquisition strategy would respond by funding the cheapest, easiest, most-voluminous calibration effort available. A VoI-informed strategy responds differently: it identifies which uncalibrated events, if they occurred, would materially change strategy recommendations across the portfolios being analyzed, and directs calibration effort there first. Events whose probability would not change any strategy recommendation, even if it went from zero to one, are deprioritized — not because they are uninteresting, but because their calibration cannot affect the decision the system is being asked to support.

The consequence is that in a consequence-first design, the data-acquisition strategy is *derived from* the consequence model, not the other way around. Two analysts looking at the same uncalibrated ecosystem might come to different conclusions about whether to calibrate it based on their differing portfolios, differing time horizons, or differing risk tolerances — and the VoI framework gives both analysts a defensible derivation for the conclusion they reach. A regulator, investment committee, or board reviewing the analyst's data-acquisition choices can audit the derivation. The choices are not preference; they are decision-theoretic outputs of an explicit consequence model.

4.3 The contrast with the standard architecture

A risk-management system built on the Big Data architecture would, by contrast, attempt to collect comprehensive data on every variable that might conceivably affect a portfolio: economic indicators, market microstructure, sentiment data, alternative-data feeds, ESG scores, credit spreads, supply-chain indicators, weather, and so on. Sophisticated machine-learning models would then attempt to predict outcomes from this combined feature space. The user would see a dashboard of predictions, presumably with confidence intervals, and would be expected to integrate the predictions with the consequence structure mentally.

> *The architectural difference is not the data or the models. Both architectures can use the same data and similar models. The difference is the **direction of derivation**. In the Big Data architecture, the data comes first and the consequence model — to the extent one exists at all — is implicit and reconstructed by the user from the predictions. In the EVA architecture, the consequence model comes first and the data is derived as the answer to "what would help here?"*

The Big Data architecture's failure mode is the one the VoI critique names: a confidently presented prediction that is statistically defensible but cannot change the decision, because the optimal action is the same under all values the prediction might take. The EVA architecture's failure mode is different and bounded: it can produce wrong recommendations if the consequence model itself is wrong, but the consequence model is explicit, exposed, and inspectable — and a regulator, committee, or analyst who disagrees with a recommendation can trace the disagreement to a specific component of the consequence model rather than to an opaque statistical pipeline.

5. Implications

If the consequence-first inversion is correct as an architectural commitment, four implications follow for enterprises currently invested in the data-first paradigm and for the broader risk-management technology market.

First, the standard procurement question for risk-management AI is the wrong question. The standard question is some version of "what data does the system have, and how accurate are its predictions?" The right question, given the consequence-first framework, is "what is the system's consequence model, how is it constructed, and how is the data-acquisition strategy derived from it?" An honest vendor with a consequence-first architecture can answer the second question with specifics: the event propagation logic, the valuation tier composition, the decision-quality optimization objective, the strategy-generation criteria. An honest vendor without a consequence-first architecture will answer the second question by pivoting back to the first — "we have lots of data and our models are very accurate" — because they have no consequence model to describe. The pivot is itself diagnostic.

Second, data-acquisition budgets in financial institutions may be misallocated under the data-first paradigm. A recent building-energy study applying Value of Information analysis (Langtry et al., 2025) applies VoI to three concrete data-collection decisions — smart meters for heat-pump maintenance, occupancy monitoring for ventilation scheduling, and ground thermal tests for ground-source heat-pump design — and shows that each is economically justified for supporting its decision *only* after VoI analysis identifies which decisions they would inform — and that "further study of VoI in building energy systems would allow expenditure on data collection to be economised and prioritised, avoiding wastage." The same logic applies, with much higher stakes, to financial institutions. A risk function that spends tens of millions of dollars annually on alternative-data feeds, sentiment scores, satellite imagery, and similar data products is often doing so without an explicit VoI derivation that ties the data, the decision it could change, and the consequence of that change into a single auditable model. If the consequence model were explicit, the institution might discover that some portion of its data spend is buying information whose decision value, for the portfolios and actions in question, is low or zero.

Third, the architectural distinction has direct implications for regulatory defensibility. A risk report presented to a regulator must, increasingly, be accompanied by an explanation of how the risk number was derived and what data went into it. A risk function operating under the data-first paradigm faces a structural difficulty: it can describe its data and its models, but cannot describe a consequence model that would let a regulator audit whether the data and models are appropriate to the decisions being made. A risk function operating under the consequence-first paradigm has the inverse situation: the consequence model is explicit and inspectable, the data-acquisition strategy is derivable from it, and a regulator can audit both. As regulatory scrutiny of AI-derived risk outputs intensifies, the institutions whose architectures support this kind of audit will have a structural advantage.

Fourth, the academic ML field's pivot toward Decision-Focused Learning is likely to accelerate. Production ML systems in industry today are overwhelmingly built on the predict-first paradigm: models are trained against generic accuracy losses, deployed against business processes, and evaluated against prediction quality. The research literature has moved past this; the production stacks have not. The institutions that catch up first — by rebuilding their ML pipelines around explicit consequence models, decision-focused training objectives, and VoI-informed data acquisition — will have an advantage that compounds: better decisions today, smaller data budgets, more defensible regulatory positions. The institutions that do not catch up will continue to spend on data whose VoI is low and to deploy models whose accuracy is high but whose decision quality is mediocre.

5.1 When each architecture is correct

To be fair to the data-first paradigm: it is not always wrong. There are domains in which "collect everything, predict, present" is the correct architecture. The honest test is: which way does the cost asymmetry of the decision run?

If the cost of a wrong decision is high and the consequences differ materially across actions, the consequence-first architecture is correct. The data and models are subordinate to a consequence model that should be built first, exposed explicitly, and audited routinely.

If the cost of any single decision is low, decisions are made in volume, and the consequences across actions are similar, the data-first architecture is appropriate. Content recommendation, ad targeting, search ranking, generic-language summarization — these are domains where the consequences of individual wrong decisions are small, the volume of decisions is enormous, and aggregate accuracy is the right objective. The VoI framework is not violated in these domains; it is simply that the VoI of better data in these domains often *is* high and *is* tied directly to accuracy, because the consequence model is "more accurate predictions produce slightly better outcomes at scale, summed over millions of decisions."

The error is not the data-first paradigm per se. The error is applying the data-first paradigm to domains where the consequence model is explicit, the decisions are infrequent, the consequences of error are large, and the audit requirements are stringent — that is, to domains like risk management, regulatory compliance, capital allocation, and clinical decision-making. In these domains the consequence-first architecture is the correct commitment.

6. Conclusion

The dominant operating assumption in AI and analytics today is that more data is better, that broader coverage improves decisions, and that uncertainty is a data problem to be dissolved through sufficient collection. This whitepaper has argued, with citations to decades of decision-theory research and several recent years of cutting-edge machine-learning literature, that the assumption is inverted. Uncertainty is not, in its essential form, a data problem; it is a decision problem. The right question is not "what data should I collect to reduce uncertainty?" but "what decision am I trying to make, what are the consequences of being wrong, and which pieces of information would actually change the decision?" The framework that makes this question answerable is Value of Information analysis. The contemporary machine-learning research community has independently arrived at the same conclusion under the name Decision-Focused Learning. Both fields are saying: start from the decision and its consequences, derive the data from that.

The architectural consequence, for any system whose output materially affects outcomes, is that the consequence model must come first and be explicit, exposed, and inspectable. The data follows as the answer to "what would help discriminate between actions whose consequences differ?" not as an undifferentiated raw input to be processed by predictive models. The Big Data paradigm's universalist claim — that more is always better, that accuracy is always the right objective — is structurally inadequate for any decision domain where the cost of being wrong is high, the consequences differ materially across actions, and the analyst's reasoning must be defensible to a regulator, an investment committee, or a board.

EVA, the deterministic risk-management platform Kindynos Advisory has built, is architected on this inversion from its foundation. The event propagation engine and the game-theoretic decision-strategy layer are designed together to constitute an explicit consequence model, with data acquisition derived from that model rather than the other way around. The design intent is that the system can be asked, for a proposed data acquisition, what the data would be worth — an evaluation grounded in the consequence model, intended to be defensible to an auditor and traceable through the valuation, probability, source, and determinism provenance the architecture is built to carry across the values it displays.

> *The work of the field, for institutions building risk-management technology that must withstand serious scrutiny, is to construct consequence models that make Value of Information computable for the decisions the institution actually makes. The data follows. The Big Data instinct to collect first and reason later is, in these domains, exactly the wrong starting point. Reasoning first; collection second. That is what the literature has been saying for decades. It is also what the most recent ML research is saying, in its own vocabulary, today.*

Sources & References

[1] Glynn, P. D., Rhodes, C. R., Chiavacci, S. J., Helgeson, J. F., Shapiro, C. D., & Straub, C. L. (2022). Value of Information and Decision Pathways: Concepts and Case Studies. Frontiers in Environmental Science. https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2022.805214/full

[2] Jackson, C., Presanis, A., Conti, S., De Angelis, D. (2017). Value of Information: Sensitivity Analysis and Research Design in Bayesian Evidence Synthesis. arXiv 1703.08994. https://arxiv.org/pdf/1703.08994

[3] Wilson, E. C. F. (2015). A practical guide to value of information analysis. PharmacoEconomics, 33(2), 105–121. https://link.springer.com/article/10.1007/s40273-014-0219-x

[4] Value of Information. Wikipedia. [background orientation only] https://en.wikipedia.org/wiki/Value_of_information

[5] Value of Information. Opasnet (international environmental health). [background only] https://en.opasnet.org/w/Value_of_information

[6] Howard, R. A. (1966). Information Value Theory. IEEE Transactions on Systems Science and Cybernetics, 2(1), 22–26. [foundational VoI reference] https://doi.org/10.1109/TSSC.1966.300074

[7] Glauner, P., Valtchev, P., State, R. (2018). Impact of Biases in Big Data. arXiv 1803.00897. https://arxiv.org/pdf/1803.00897

[8] Katsikopoulos, K. V., et al. (2022). Decoding human behavior with big data? Critical, constructive input from the decision sciences. AI Magazine. https://onlinelibrary.wiley.com/doi/full/10.1002/aaai.12034

[9] Raj, D. (2021). Avoiding The Biggest Big Data Fallacy. LinkedIn. https://www.linkedin.com/pulse/avoiding-biggest-big-data-fallacy-where-more-means-higher-david-raj

[10] Devansh. (2024). Understanding the Big Data-AGI Fallacy. Medium / Machine Learning Made Simple. https://machine-learning-made-simple.medium.com/understanding-the-big-data-agi-fallacy-made-by-ai-companies-893ab4685bd1

[11] Does More Data Always Yield Better Performance? Towards Data Science (January 2026). https://towardsdatascience.com/does-more-data-always-yield-better-performance/

[12] Quality of Data in Machine Learning. arXiv 2112.09400. https://arxiv.org/pdf/2112.09400

[13] Elmachtoub, A. N., Grigas, P. (2022). Smart Predict, then Optimize. Management Science. (The canonical PtO reference.) https://pubsonline.informs.org/doi/10.1287/mnsc.2020.3922

[14] Mandi, J., et al. (2023). Decision-Focused Learning: Foundations, State of the Art, Benchmark and Future Opportunities. arXiv 2307.13565. https://arxiv.org/abs/2307.13565

[15] Tang, B., Khalil, E. B. (2023). CaVE: A Cone-Aligned Approach for Fast Predict-then-Optimize. arXiv 2312.07718. https://arxiv.org/pdf/2312.07718

[16] Wang, X., Du, J., Peng, Y., Ma, W. (2025). From Sequential to Recursive: Enhancing Decision-Focused Learning with Bidirectional Feedback. AAAI 2026. arXiv 2511.08035. https://arxiv.org/pdf/2511.08035

[17] Lawless, C., Zhou, A. (2022). A Note on Task-Aware Loss via Reweighing Prediction Loss by Decision-Regret. arXiv 2211.05116. https://arxiv.org/pdf/2211.05116

[18] Yang, J., et al. (2025). DFF: Decision-Focused Fine-tuning for Smarter Predict-then-Optimize with Limited Data. arXiv 2501.01874. https://arxiv.org/pdf/2501.01874

[19] Decision-Focused Learning without Differentiable Optimization. NeurIPS 2022. https://proceedings.neurips.cc/paper_files/paper/2022/file/0904c7edde20d7134a77fc7f9cd86ea2-Paper-Conference.pdf

[20] Learning Loss Functions for Predict-then-Optimize. Harvard TeamCore. https://teamcore.seas.harvard.edu/learning-loss-functions-predict-then-optimize/

[21] Predict-then-Optimize for Seaport Power-Logistics Scheduling. arXiv 2511.07938 (November 2025). https://arxiv.org/pdf/2511.07938

[22] Langtry, M., et al. (2025). Rationalising data collection for supporting decision making in building energy systems using Value of Information analysis. Journal of Building Performance Simulation; arXiv 2409.00049. https://arxiv.org/abs/2409.00049

Note on sources and currency. *Foundational Value-of-Information theory rests on Howard, R. A. (1966), "Information Value Theory"; Raiffa, H. (1968), Decision Analysis; and Pratt, Raiffa & Schlaifer, Introduction to Statistical Decision Theory. This whitepaper draws on a mix of foundational decision-theory and statistical literature (1960s through 2010s), peer-reviewed applied research in health economics and environmental science (1990s through 2020s), and the recent academic machine-learning literature on Decision-Focused Learning and Predict-then-Optimize (2022 through 2026). Citations are current as of preparation in May 2026 and reflect an actively developing field; vocabulary continues to evolve. The paper is research analysis and strategic commentary, not legal, tax, investment, accounting, or regulatory advice.*