Int. j. adv. multidisc. res. stud. 2024; 4(6):3539-3563

Received: 10-11-2024
Accepted: 20-12-2024

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

AI-Enabled Synthesis of Stakeholder and Survey Data in Operations Research: A Review of Advances and Applications

[1] Valentina Ochuko Obukadata, [2] Rosalyn Ezeako, [3] Ifeoma E Okoli, [4] Shalom Alugwe
[1] University of New Haven, Connecticut, United States
[2] Firma Divina Projects, Nigeria
[3, 4] Maersk Logistics, Canada

DOI: https://doi.org/10.62225/2583049X.2024.4.6.6931
Corresponding Author: **Valentina Ochuko Obukadata**

Abstract

Operations research depends on inputs that are not naturally numerical. Objectives are negotiated, constraints are contested, criterion weights are argued over, and the parameters that finally enter a model are usually distilled from interviews, facilitated workshops, questionnaires and free-text commentary. Converting that material into a usable model has long been the slowest, least reproducible and least auditable part of the analytical process. Between 2023 and 2024, with unusual density in the latter year, a body of work emerged that applies large language models and related generative systems to precisely this conversion problem. This paper reviews that work and situates it within the wider applied analytics literature on data-driven decision systems, procurement and supply chain optimisation, safety and risk management, healthcare operations, energy transition planning, and enterprise financial decision-making. The paper develops a conceptual model that decomposes the path from stakeholder situation to decision model into six functional layers: elicitation, qualitative coding and synthesis, synthetic respondent generation, autoformulation from natural language, preference and weight elicitation for multi-criteria methods, and explanation, verification and governance. Six propositions derived from the model are examined against the literature, and applications are then surveyed across eight domains in which stakeholder and survey data feed directly into operational models. Three findings structure the argument. First, progress has been fastest at the two ends of the pipeline, in text-to-model translation and in automated deductive coding, while the middle, where stakeholder meaning is negotiated and converted into legitimate model structure, remains comparatively underserved. Second, the evaluation base is thin relative to the enthusiasm, with most reported gains measured against convenience benchmarks rather than controlled comparison. Third, and most consequentially for operations research specifically, the errors these systems make are structured rather than random, since persona conditioning amplifies group differences and models drift toward socially acceptable answers on sensitive items, which converts a noise problem into a bias problem that optimisation then propagates into unfair allocations. The discussion characterises this pattern as an inversion, in which the capability profile of the technology is the reverse of the difficulty profile of the practice. Theoretical, methodological, practical and policy implications are drawn out, limitations of the review, the evidence base and the model are stated, and the paper closes with a research agenda organised around verification, participatory legitimacy, and the epistemic status of synthetic evidence in decision models that carry real consequences.

Keywords: Operations Research, Large Language Models, Stakeholder Engagement, Survey Research, Problem Structuring, Autoformulation, Multi-Criteria Decision Analysis, Silicon Sampling, Qualitative Synthesis, Decision Support Systems

1. Introduction

1.1 The soft edge of a hard discipline

The canonical description of the operations research process runs through problem definition, data collection, model formulation, solution, validation and implementation. Textbook presentations devote most of their attention to the middle stages, where the mathematics lives. Practitioners tend to report the opposite distribution of effort. Getting a problem defined in a way that its owners recognise, and getting reliable parameters out of the people who hold the relevant knowledge, routinely consumes more calendar time than solving the resulting model.

This asymmetry produced its own methodological tradition. Problem structuring methods, including Soft Systems Methodology (Checkland, 1981), Strategic Options Development and Analysis (Eden & Ackermann, 1998) and the broader family catalogued by Rosenhead and Mingers (2001), exist because messy, multi-actor situations resist direct formalisation. Multi-criteria decision analysis developed elaborate machinery for eliciting and aggregating preferences precisely because criterion weights are judgements rather than measurements. Facilitated modelling itself is difficult to audit, since a workshop is a live event whose reasoning is largely invisible in the artefacts it produces, as Franco and Greiffenhagen (2018) showed through detailed ethnomethodological study of how such sessions unfold. Marttunen *et al*. (2017) documented how frequently practitioners combine a structuring method with an analytic one, which is itself an admission that neither half suffices alone.

The same asymmetry appears throughout the applied literature. Studies of multi-stakeholder governance alignment in joint ventures describe the coordination problem as prior to, and harder than, the optimisation problem it eventually enables (Eyetsemitan *et al*., 2020). Work on supplier relationship management frames collaborative capability rather than contract structure as the binding constraint on supply chain performance (Ike *et al*., 2021). Safety research is more explicit still, since safety leadership influence in large temporary project organisations is a construct that exists only in perception data (Arumosoye & Obriki, 2024), and the recurrence mechanisms of unsafe behaviour are behavioural rather than physical (Obriki & Arumosoye, 2022).

1.2 What changed

Large language models altered the economics of working with text. Three capabilities matter for operations research specifically. First, models can conduct adaptive dialogue, which means an instrument can probe rather than merely record. Second, they can apply a coding framework expressed in ordinary language, without task-specific labelled training data. Third, they can translate a prose description of a decision problem into a formal model and executable solver code.

Fan *et al*. (2024) framed the third of these as one pillar of a broader artificial intelligence for operations research agenda, situating natural language to model translation alongside parameter generation and solver acceleration. Their framing is useful but incomplete for present purposes, because it treats upstream human material as a data source rather than as a site of contested meaning. The literature reviewed here sits partly inside their taxonomy and partly outside it.

A parallel shift occurred in applied decision systems. Work on advanced decision support for healthcare operations and resource planning makes the same argument in a domain where the inputs are overwhelmingly qualitative (Adelanwa *et al*., 2024a).

1.3 Scope of the review

This review covers work that uses artificial intelligence, predominantly large language models, to collect, code, synthesise, simulate or formalise stakeholder and survey data for operations research and adjacent decision-analytic purposes. It excludes work that applies artificial intelligence to solve optimisation problems whose inputs are already given, and it excludes forecasting and predict-then-optimise pipelines (Elmachtoub & Grigas, 2022) except where stakeholder judgement supplies the predicted quantity.

The contributions are fivefold. First, a conceptual model of AI-enabled stakeholder and survey synthesis, presented in Section 4 as six functional layers mapped onto stages of the operations research process rather than onto model architectures, together with six propositions that the synthesis then examines. Second, a comparative reading of the evidence base that distinguishes claims supported by controlled comparison from claims supported by demonstration. Third, a cross-domain application survey drawing on the applied analytics literature in procurement, healthcare, safety, energy, finance, marketing and agriculture, which is where stakeholder data actually enters operational models. Fourth, a set of theoretical, methodological, practical and policy implications drawn from the synthesis. Fifth, a research agenda that treats verification and participatory legitimacy as first-order technical problems rather than as ethical postscripts.

2. Methodology

2.1 Review design

This is a scoping review with narrative synthesis rather than a systematic review in the PRISMA sense. That choice reflects the state of the field. The relevant work is scattered across operations research, computational linguistics, survey methodology, human-computer interaction and applied domain journals, uses incompatible vocabulary for similar constructs, and includes a large preprint fraction that formal inclusion criteria would either over-weight or exclude entirely.

Sources were identified through database and preprint-server searching across combinations of three term groups. Method terms included large language model, generative artificial intelligence, foundation model and multi-agent system. Data terms included survey, questionnaire, open-ended response, interview, stakeholder, qualitative and preference elicitation. Domain terms included operations research, optimisation, decision analysis, multi-criteria decision analysis, problem structuring, supply chain, scheduling and resource allocation. Backward and forward citation chasing extended the initial set.

A second corpus was assembled from the applied decision-analytics literature, covering domains in which stakeholder and survey data feed directly into operational models. This corpus is treated as evidence about the demand side of the problem, showing where and why such data enters models, rather than as evidence about method performance. The emphasis falls on 2024 publications, with earlier work included where it established a line of inquiry and later work included where it supplies evaluation evidence that 2024 lacked.

2.2 Inclusion and exclusion criteria

Work was included where it satisfied three conditions. It had to concern the collection, coding, simulation or formalisation of human judgement data rather than the solution of problems whose inputs were already given. It had to involve artificial intelligence methods, predominantly but not exclusively large language models. And it had to bear on decision modelling, either directly through optimisation and multi-criteria analysis or indirectly through the construction of parameters that such models consume.

Work was excluded where the artificial intelligence component operated purely downstream of a settled model, where the study concerned language model capability in the abstract with no decision-analytic application, or where a claim of stakeholder synthesis rested on a demonstration too under-specified to assess. Predict-then-optimise pipelines were excluded except where stakeholder judgement supplied the predicted quantity.

No language restriction was imposed in principle, although the retrieved corpus is overwhelmingly English-language, which is itself a finding rather than only a limitation and is discussed in Section 9.

2.3 Corpus composition

Two corpora were assembled and are used for different evidential purposes throughout.

The methodological corpus consists of studies that report the performance of an artificial intelligence method on an elicitation, coding, simulation or formalisation task. It is treated as evidence about what these methods can currently do, and claims drawn from it are qualified by the strength of the underlying design, distinguishing controlled comparison from demonstration.

The applied corpus consists of decision-analytic and operational studies across procurement, healthcare, occupational safety, energy, finance, marketing, process improvement, education and agricultural development. It is treated as evidence about the demand side of the problem, establishing where and why judgement data enters operational models and what the consequences of misrepresenting it would be. It is not used to support claims about method performance, and no attempt is made to pool results across the two corpora.

2.4 Synthesis procedure

Synthesis proceeded in four passes. The first pass coded each methodological source to a stage of the operations research process, which produced the six-layer structure presented in Section 4. The second pass coded each applied source to the layer or layers whose data it consumes, which produced the application mapping in Section 6. The third pass extracted evaluation designs and reliability statistics where reported, which produced the assessment in Sections 7.3 and 7.4. The fourth pass identified disconfirming and critical evidence explicitly, in order to avoid the tendency of technology reviews to aggregate positive demonstrations.

Where sources conflicted, the study with the stronger design was given precedence and the disagreement was reported rather than resolved. Where a claim rested on a single demonstration, it is described as such in the text.

2.5 Quality appraisal

No formal quality appraisal instrument was applied, since none of the available instruments accommodates the mixture of computational benchmark studies, survey methodological experiments and applied conceptual frameworks that this corpus contains. Instead, three appraisal questions were applied consistently to methodological sources: whether the comparison condition was a serious baseline or a convenience one, whether the reported metric measured the construct of interest or a proxy for it, and whether the study reported performance disaggregated in any way. The answers inform how strongly each finding is stated.

The review is written from an operations research perspective, which shapes what counts as an important gap. A survey methodologist reading the same corpus would likely weight Layer 1 more heavily and Layer 4 less. This orientation is stated rather than corrected for.

3. Background: The Stakeholder Data Problem in Operations Research

3.1 Where soft data enters hard models

It is worth being concrete about how stakeholder and survey material actually becomes model input, because the abstraction obscures the difficulty.

In procurement and supply chain optimisation, supplier evaluation weights, risk scores and resilience indices are derived from expert judgement and structured questionnaire instruments. Frameworks for data-driven vendor evaluation and bid selection combine measured performance with assessed capability, and the assessed component dominates in early-stage sourcing (Ahiaeke Patrick *et al*., 2021). Resilience indices for post-pandemic supply chains are explicitly constructed from perceived exposure and perceived recovery capability rather than from realised disruption frequency (Efobi *et al*., 2023). Supplier relationship management frameworks operationalise trust and collaboration, neither of which has a natural instrument (Ike *et al*., 2021). Quantitative risk architecture for public-private partnerships layers stakeholder-assessed likelihood over technical exposure (Ike *et al*., 2024).

In safety and risk management the dependence is even more direct. Near-miss and hazard observation systems generate large volumes of short free-text reports that must be classified before they can inform any predictive model (Obogo *et al*., 2021). Behavioural safety programmes rest on observation data whose coding is a judgement (Obriki *et al*., 2023a). Safety culture maturity models are perception instruments (Arumosoye & Obriki, 2021). Predictive safety analytics for early detection of high-risk activity depends on the quality of the upstream classification far more than on the choice of predictive algorithm (Obriki & Arumosoye, 2020).

In healthcare operations, patient flow models, workforce planning models and service redesign decisions all consume staff and patient perception data. Governance-driven workforce planning frameworks require nurse-patient ratio judgements that combine measured workload with assessed acuity (Omaghomi *et al*., 2023). Health system transformation work notes that the data backbone required for such models is itself an organisational and governance achievement rather than a technical one (Ilodigwe & Adesemoye, 2021).

In energy planning, community acceptance is now treated as a modelled variable rather than an external condition. Work on turning social license into a measurable variable predicts community trust from survey and contextual data (Komi, 2024a), and related analysis models the causal link between community trust and project survival (Komi & Adeniji, 2024d). Just transition analysis combines econometric estimation with skills mapping derived from workforce survey instruments (Komi & Adamolekun, 2024b).

In agriculture and rural development, technology adoption models rest almost entirely on smallholder survey data, and the literature on socioeconomic barriers to adoption makes clear that the barriers are perceptual and institutional as much as financial (Michael & Ogunsola, 2022). Assessments of climate-smart practice contribution to household food security depend on self-reported outcome measures (Michael & Ogunsola, 2021b).

In information security, human factors research is survey research. Studies of password practice and authentication behaviour, of bring-your-own-device adoption risk, and of awareness training effectiveness are all instrument-based, and their outputs feed directly into control prioritisation decisions (Onche *et al.*, 2024).

3.2 Why the conversion is hard

Four properties of stakeholder material resist automation, and they compound.

The material is short and context-poor. Open-ended survey responses, near-miss reports and free-text feedback often consist of a few words whose meaning depends on organisational context the analyst holds and the text does not.

The material is contested. Different stakeholder groups describe the same situation using incompatible framings. A constraint to one group is a target to another. Aggregating across groups without first surfacing the disagreement produces a model that no group recognises.

The material is strategic. Respondents anticipate the use to which their answers will be put. Procurement questionnaires, safety perception surveys and staff satisfaction instruments all elicit answers shaped by expected consequences.

The material carries legitimacy. Where a model's authority derives from having consulted affected parties, the consultation is not merely a data-collection step. It is part of what makes the eventual recommendation acceptable. Multi-stakeholder governance alignment work makes this point directly (Eyetsemitan *et al.*, 2020), as does the energy transition literature on why projects survive or fail (Komi & Adeniji, 2024d).

Automation interacts differently with each property. It helps most with the first, is neutral on the third, and risks actively damaging the fourth.

4. Conceptual Model

4.1 Rationale

The synthesis requires an organising structure, and the two obvious candidates are unsatisfactory. Organising by model architecture, which is the convention in the computational literature, groups together work that solves quite different problems and separates work that solves the same one. Organising by application domain, which is the convention in the applied literature, obscures the fact that a procurement problem and a healthcare rostering problem consume stakeholder data in structurally identical ways.

The model proposed here organises instead by function within the operations research process. It decomposes the path from a stakeholder situation to a decision model into six layers, each defined by the transformation it performs on the material rather than by the technique it uses. The decomposition was derived inductively during the first synthesis pass described in Section 2.4 and then applied deductively in later passes.

4.2 The six-layer pipeline

Fig 1 presents the model. Four layers lie on the main path from situation to model: elicitation converts a situation into recorded stakeholder input, coding converts that input into structured categories, autoformulation converts a structured problem description into a formal model, and the model yields a recommendation. Two layers sit off the main path in supporting roles: synthetic respondent generation feeds pretesting, imputation and robustness analysis rather than the primary data stream, and preference elicitation supplies the weights and trade-off parameters that the formulation requires but that no amount of text coding can produce. A sixth layer wraps the others, supplying explanation, verification and governance, and it is the layer that closes the loop back to the stakeholders whose input started the process.

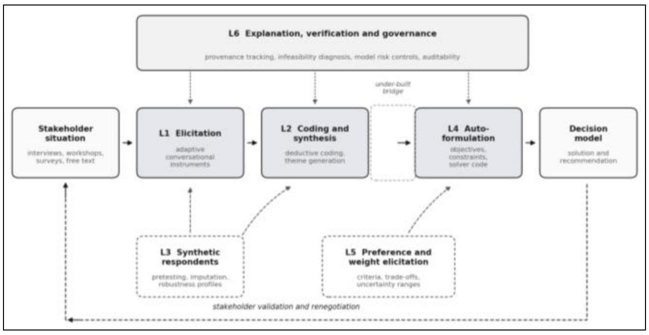

Fig 1: The six-layer stakeholder-to-model pipeline. Solid arrows show the primary data path, dashed arrows show supporting and governance relationships, and the dashed box marks the under-built bridge between coded stakeholder material and formulation-ready problem descriptions

Two features of the figure carry the argument of the paper. The first is the dashed box between Layer 2 and Layer 4, marking the transition from coded stakeholder material to a problem description clean enough for an autoformulation system to consume. Section 5.4.3 argues that this transition is where the least work has been done and where the most is needed. The second is the feedback loop running beneath the pipeline. Where a model's authority derives from consultation, that loop is not an optional refinement but the mechanism through which the model acquires its legitimacy, and every automation decision in the layers above changes what passes through it.

4.3 Layer definitions

Layer	Function	Representative work
L1 Elicitation	Collecting stakeholder input adaptively	Chopra and Haaland (2023); Wuttke *et al.* (2024); Yu *et al.* (2024)
L2 Coding and synthesis	Converting text to structured categories and themes	Chew *et al.* (2023); Mathis *et al.* (2024); De Paoli (2024)
L3 Synthetic respondents	Generating or imputing respondent-level data	Argyle *et al.* (2023); Sun *et al.* (2024); Park *et al.* (2024)
L4 Autoformulation	Translating problem descriptions into formal models	Ramamonjison *et al.* (2023); Xiao *et al.* (2024); AhmadiTeshnizi *et al.* (2024)
L5 Preference elicitation	Deriving MCDA parameters from judgement	Emerging; largely extrapolated from L1 to L3
L6 Explanation and governance	Making models legible, checked and accountable	Chen *et al.* (2024); Mostajabdaveh *et al.* (2024); Ominyi (2024a)

4.4 Propositions

The model generates six propositions, which the synthesis in Sections 5 and 6 examines and Section 7.7 revisits.

P1. Automated performance is higher where the transformation is constrained by an externally supplied framework than where the framework must itself be induced. Deductive coding should therefore outperform inductive theme generation, and translation of well-posed descriptions should outperform interpretation of contested ones.

P2. Effort in the field is distributed inversely to difficulty, concentrating at the ends of the pipeline where inputs are cleanest and thinning in the middle where meaning is negotiated.

P3. Synthetic respondent data is defensible in the supporting role shown in Fig 1, for pretesting and robustness analysis, and indefensible on the main path, where it substitutes for consultation rather than augmenting it.

P4. Error at Layers 2 and 3 is not randomly distributed across respondents, so bias introduced upstream is propagated by optimisation into systematically unequal allocations rather than averaged away. This is the proposition for which the evidence is most indirect and most in need of a dedicated test.

P5. Layer 6 determines the value of the layers beneath it, because a formulation derived by an opaque process cannot be validated by the stakeholders who supplied its inputs, and only they can detect misinterpretation of intent as distinct from error in execution.

P6. Evaluation practice measures the wrong quantity. Coding accuracy and formulation accuracy are proxies for decision equivalence and stakeholder recognition, and neither of the latter is currently measured.

4.5 Boundary conditions

The model applies to decision problems whose parameters are substantially judgemental and whose stakeholders are identifiable. It does not usefully describe problems whose inputs are wholly instrumented, such as scheduling against measured machine throughput, where Layers 1 to 3 collapse and the pipeline reduces to formulation and solution. It also assumes a single formalisation is the goal; participatory settings in which the purpose is to sustain rather than resolve disagreement fit the model only partially, and Section 9.3 returns to this.

5. Literature Synthesis

This section works through the six layers in turn. For each, it reports what has been established, what rests on demonstration alone, and what the applied corpus indicates about demand. The distribution of research effort across the layers is uneven in a way that Proposition 2 anticipates. Layer 4 has attracted the most concentrated methodological attention, with competitive benchmarks and successive system generations. Layers 2 and 3 are active but fragmented across disciplines. Layer 1 is developing quickly. Layer 5 is conspicuously thin. Layer 6 is nascent within operations research but comparatively well developed in the adjacent governance literature.

5.1 Layer 1: AI-mediated elicitation

The oldest constraint in stakeholder data collection is the trade-off between depth and scale. Structured surveys reach many people but confine them to anticipated answer spaces. Semi-structured interviews recover unanticipated content but cost too much per respondent to run at scale. Every applied instrument in the domains surveyed above sits somewhere on this frontier, and most sit at the shallow end for cost reasons.

Chopra and Haaland (2023) delegated the interviewer role to an artificial intelligence system and ran several hundred qualitative interviews on a single substantive question. The design point of interest for operations research is not the substantive finding but the demonstration that adaptive probing at survey scale is feasible, together with their observation that the factors surfaced in initial top-of-mind answers differed systematically from those surfaced by later probing. This matters directly for model construction. A first-pass constraint list gathered by questionnaire is likely to be a different object from the constraint list that emerges under sustained questioning, which has implications for how near-miss reporting systems and hazard observation programmes are designed (Obogo *et al*., 2021).

Wuttke *et al*. (2024) tested the approach against a human-interviewer control condition, randomly assigning participants to artificial or human interviewers working from an identical protocol. They report response quality broadly comparable to human-conducted interviews, with AI interviews eliciting substantially longer answers, alongside identified failure modes in guideline adherence. The controlled comparison is the methodologically important feature. Most work at this layer reports only that the system ran.

Yu *et al*. (2024) approached the same territory from a systems engineering angle, proposing a modular architecture for conversational survey agents that separates question management, response interpretation and follow-up generation. Modularity matters for operations research deployment because it allows the elicitation logic to be audited independently of the underlying model, which is the same argument made in the governance literature about separating decision logic from decision infrastructure (Ominyi & Anichukwueze, 2024a).

Several applied literatures describe elicitation problems that adaptive instruments would address directly.

Safety perception surveys suffer from a well-known ceiling effect in which respondents report high safety climate scores that do not predict incident rates. Adaptive probing could distinguish genuine confidence from acquiescence, which matters for maturity models built on such instruments (Arumosoye & Obriki, 2021; Obriki *et al*., 2023a).

Information security awareness research faces the opposite problem, in which respondents under-report insecure behaviour. Studies of password practice, device adoption and training effectiveness all rely on self-report instruments whose validity is bounded by social desirability (Onche *et al*., 2024). Virtual interviewing has historically shown some reduction in social desirability bias, though this has not been established for current conversational agents across culturally varied populations.

Public health surveillance systems require rapid collection of symptom and behaviour data from populations with limited literacy and variable language. Frameworks for national surveillance and epidemic intelligence assume a data collection layer whose cost currently constrains coverage (Oparah *et al*., 2023). Adaptive multilingual instruments are a natural fit, subject to the performance gaps discussed in Section 7.5.

Education and training evaluation depends on learner feedback instruments whose design constrains what can be learned. Conceptual frameworks for evidence-based digital learning and for real-time performance analytics both presuppose richer feedback than fixed instruments produce (Orise & Niniola, 2022).

Agricultural extension and adoption research faces cost and reach constraints that adaptive voice-based instruments could relax, particularly for the socioeconomic barrier questions where fixed-response instruments perform worst (Michael & Ogunsola, 2021b, 2022).

Public safety communication is a further case in which the instrument and the population are badly matched. Work on artificial intelligence for multilingual and accessible emergency alerting identifies comprehension across languages and disability categories as the binding constraint on alert effectiveness (Odejide, 2023), and a trust-centred model for multilingual emergency communication treats recipient trust as a design variable rather than an assumption (Odejide, 2024). These are Layer 1 problems in a setting where the cost of a failed instrument is immediate and unevenly distributed.

Administrative work supplies a quieter example. Studies of contemporary office management practice, of the influence of personal and professional attributes on organisational effectiveness, and of professional ethics and workplace experience among female secretarial staff are questionnaire-based and address constructs that resist fixed-response measurement (Onche *et al*., 2021b). Where the construct of interest is workplace experience, adaptive probing is a validity requirement rather than a convenience.

Interviewer effects are well documented in survey methodology, and an artificial interviewer has its own, currently uncharacterised, effect profile. Consent and disclosure norms for AI-conducted elicitation are unsettled, and interact with data protection obligations across jurisdictions (Annan, 2022). Whether adaptive probing improves or degrades comparability across respondents remains open, since the instrument is no longer identical for each participant, which is a problem for any downstream model that assumes exchangeability.

5.2 Layer 2: Coding, thematic analysis and synthesis

Coding free text is the classic bottleneck between qualitative stakeholder data and quantitative modelling. The recent literature attacks it from both the deductive and inductive directions, and the results differ sharply between them.

On the deductive side, Chew *et al*. (2023) established the basic pattern of supplying a codebook in natural language and having the model apply it, which removes the labelled-training-data requirement that constrained earlier machine learning classifiers. Subsequent work extended this to open-ended survey responses in non-English settings, where responses are typically short, context-poor and semantically subtle, and where most prior evidence had come from English-language and topically simple material. Zenimoto *et al*. (2024) took a different route, generating pseudo-responses to improve coding of open-ended items.

On the inductive side, Mathis *et al*. (2024) compared open-source model generated thematic analysis of healthcare interviews against traditional analysis, and De Paoli (2024) worked through inductive thematic analysis as an explicit exploration of where the approach breaks down. The pattern across these studies is consistent. Performance on deductive coding against a clear framework is substantially better than performance on inductive theme generation, and both degrade where the material requires reading tone, affect or conversational nuance. Later blinded comparisons with formal non-inferiority testing reinforced this asymmetry, supporting language models as an adjunct to rather than a replacement for human analysts.

For operations research the gap is significant in a specific way. Deductive coding is what one does when the model structure is already settled and stakeholder text needs to be mapped onto known parameters. Inductive theme generation is what one does when the model structure is still in question, which is exactly the problem-structuring phase where the discipline most needs help and where current tools are weakest.

The applied literature on human-in-the-loop machine learning is directly relevant here and is frequently overlooked by the survey methodology community. Reviews of the state of the art distinguish configurations by where human judgement enters, whether at annotation, at model correction or at decision (Ladapo *et al*., 2022), and service management research supplies the operational vocabulary for describing such hand-offs in production settings (Ladapo *et al*., 2023). Work on human-centred automated annotation with generative models addresses the specific case in which the model proposes labels and humans adjudicate, which is the dominant configuration in practical coding pipelines (Ladapo *et al*., 2024).

This literature supplies something the qualitative coding studies generally lack, which is a vocabulary for describing what the human is actually doing. Reporting that a study used human validation is uninformative unless the configuration is specified.

Related work on context-aware data curation for fine-tuning pipelines addresses the upstream question of what material a domain-adapted coding model should see (Oyesiji *et al*., 2024), and parameter-efficient multilingual adaptation is directly relevant to coding in the linguistically diverse settings where much stakeholder data originates (Oyesiji *et al*., 2023).

Reliability reporting has improved. Krippendorff's alpha with conventional thresholds is now common as an evaluation instrument, which permits comparison across studies in a way that raw accuracy percentages did not (Krippendorff, 2019). It also exposes an uncomfortable fact, which is that many reported agreements fall in the tentative rather than reliable band.

The applied domains have their own reliability traditions that the automated coding literature could usefully absorb. Incident investigation research has long distinguished between coding the event and coding the causal attribution, with markedly lower agreement on the latter (Obriki & Arumosoye, 2024). Business intelligence applications for mental health resource allocation depend on classification of presenting need, where category boundaries are contested by construction (Anioke & Atima, 2023b). Public health governance models using process optimisation face the same issue when performance metrics are derived from coded case narratives (Anioke & Atima, 2023a).

Model governance work adds a further consideration. Using a language model to judge the output of another language model is now a common evaluation shortcut, and reviews of this practice document systematic biases in the judging model that propagate into release decisions (Nwakamma *et* *al*., 2024a). Where automated coding is validated by automated means, the resulting reliability estimate is not independent of the system being evaluated.

5.3 Layer 3: Synthetic respondents and data augmentation

This is the most contested layer. Argyle *et al*. (2023) introduced the conditioning of a language model on sociodemographic backstories drawn from real respondents to produce corresponding synthetic respondents, and argued that outputs reproduced human response distributions and cross-item correlation structure closely enough to warrant the term algorithmic fidelity. Sun *et al*. (2024) extended conditioning from individual backstories to group-level demographic information, finding close approximation of real marginals on many items and the worst performance on socially sensitive topics, where models drift toward socially acceptable answers. Kim and Lee (2024) applied the machinery to imputation and retrodiction in opinion data, which is the use case closest to standard operations research data preparation. Park *et al*. (2024) scaled interview-grounded agent construction to roughly a thousand individuals. An earlier strand had already proposed treating conditioned models as economic subjects whose behaviour could substitute for experimental participants (Horton, 2023).

A critical literature developed in step. Bisbee *et al*. (2024) found that persona-conditioned synthetic responses substantially inflated estimates of affective polarisation relative to human benchmarks, attributing the distortion to stereotype amplification in the conditioning step. This is consistent with evidence that model outputs align more closely with the views of some demographic groups than others even before any persona conditioning is applied (Santurkar *et al*., 2023). Sarstedt *et al*. (2024) synthesised a large set of comparisons between synthetic and human samples and concluded that models reproduce surface-level patterns more faithfully than deeper behavioural regularities, recommending confinement to upstream activities such as pretesting and pilot design rather than main studies. Later methodological work highlighted analytic flexibility as a distinct threat, since the number of defensible prompting and conditioning choices is large enough that a determined analyst can reach many conclusions from the same underlying model.

Evaluation practice has itself become a research topic. Distributional comparison, the dominant early standard, cannot distinguish coherent respondent-level prediction from aggregate pattern matching, which has motivated stricter designs such as cross-survey transfer, in which a model receives one block of a respondent's answers and must predict a disjoint block.

Operations research has always augmented sparse data. What distinguishes synthetic respondents is the source of the augmentation and the nature of the claim being made.

Conventional practice in demand forecasting fits a model to observed history and extrapolates, with uncertainty quantified from the fit (Tonoyan *et al*., 2021). Predictive analytics for demand forecasting accuracy in e-commerce procurement follows the same logic (Akin-Oluyomi *et al*., 2023a). Attribution modelling in marketing constructs counterfactual journeys from observed touchpoints, with well-documented bias problems that arise from the modelling assumptions rather than from the data source (Sanni *et al*., 2022). Machine learning models addressing uncertainty in cross-channel campaign forecasting quantify that uncertainty explicitly (Wedraogo & Sanni, 2024). Predictive audience segmentation constructs latent groups from behavioural data and validates them against outcomes (Oziri *et al*., 2020). Frameworks for predictive and prescriptive marketing analytics distinguish the two modes precisely because the prescriptive claim requires stronger warrant (Rukh *et al*., 2023).

In every one of these cases the synthetic quantity is anchored to observed behaviour of the same population. Synthetic respondents are anchored to a training corpus whose relationship to the target population is unknown and unverifiable. That is a categorical difference, not a difference of degree, and it should govern how the outputs are used.

Interpretable machine learning work in renewable energy makes a complementary point. Early failure prediction models in distributed systems are valued partly because their reasoning can be inspected against engineering knowledge (Komi & Adamolekun, 2021b). Synthetic respondent generation offers no equivalent inspection surface, since there is no domain theory against which the generated preferences can be checked.

Predictive analytics for financial risk detection illustrates the consequence. Where synthetic augmentation is used to balance rare-event classes, the augmented cases inherit whatever correlations the generator encodes, and downstream detection thresholds are set against a distribution that does not exist (Adelanwa *et al*., 2023b).

Synthetic respondents are defensible as an instrument for exploring the sensitivity of a decision model to plausible preference structures, for pretesting an elicitation instrument before fielding it, and for stress-testing a model against hypothetical stakeholder configurations. They are not defensible as a substitute for consultation where the model's legitimacy rests on having consulted. The distinction is not primarily technical, and the marketing research literature has moved furthest toward codifying it (Sarstedt *et al*., 2024).

5.4 Layer 4: Autoformulation from natural language

This is the layer where operations research has done the most on its own terms. The NL4Opt competition (Ramamonjison *et al*., 2023) established the task of translating natural language descriptions into linear programming formulations, with subtasks for entity recognition and formulation generation, and supplied the benchmark against which subsequent work was measured. Three strategies then diverged.

Multi-agent decomposition assigns specialised roles to cooperating agents. Chain-of-Experts (Xiao *et al*., 2024) coordinates such agents over complex operations research problems, introducing the ComplexOR dataset alongside NL4Opt evaluation. Mostajabdaveh *et al*. (2024) developed a multi-agent, multi-stage framework that couples formulation with verification, which is the more interesting design choice because it treats correctness as something to be established rather than assumed.

Pipeline systems with solver integration build formulations, generate solver code and iterate on errors. OptiMUS (AhmadiTeshnizi *et al*., 2024a) and OptiMUS-0.3 (AhmadiTeshnizi *et al*., 2024b) exemplify this route, with the later release targeting scale and introducing the NLP4LP dataset. The evolution of this system across versions is one of the few places in the literature where like-for-like progress can be traced.

Fine-tuning approaches train on optimisation modelling data rather than relying on prompt engineering or agent orchestration. ORLM (Tang *et al*., 2024) uses semi-synthetic training data and LLMOPT (Jiang *et al*., 2024) pursues a related learning-based route. These are complementary to the agentic approaches rather than competing with them.

Search-based and hybrid methods also appeared, including Monte Carlo tree search over the space of plausible model structures (Astorga *et al*., 2024), which produces a set of functionally distinct candidate models rather than a single answer. That output format is arguably better suited to stakeholder settings, where the point of the exercise is often to surface alternative framings rather than to converge prematurely.

Two cautions have emerged. Workflow-based agentic designs can propagate errors through sequential stages, and at least one comparative study reports degraded performance for a pipeline system relative to a simpler standard baseline on a given dataset, which suggests that agent orchestration is not uniformly beneficial. Evaluation by final optimal value is also a coarse instrument, since a system can produce a formulation that is wrong in structure but coincidentally right in value, or can eliminate errors by deleting the constraints that caused them. Work on human-aligned evaluation for linear programming word problems responds directly to this (Xing *et al*., 2024).

Essentially all work at this layer assumes a clean, self-contained, unambiguous problem description as input. Real stakeholder material is none of these things.

The gap is visible when one compares the benchmark problem statements with the descriptions that actually circulate in applied settings. Data-driven lean supply chain optimisation in manufacturing and retail requires reconciling process descriptions from operators, supervisors and planners that do not agree (Efobi *et al*., 2022). Negotiation optimisation for procurement cost reduction depends on reservation values that no party will state directly (Akinleye & Adeyoyin, 2022a). Resilient logistics for humanitarian supply chains integrates predictive analytics with field assessments whose categories shift between assessments (Anene & Clement, 2022).

Some adjacent work does address structured formalisation from operational descriptions. Constraint satisfaction modelling for cloud network resource allocation and predictive resource scaling both formalise problems whose statements originate in service-level discussions rather than in mathematics (Ahmed *et al*., 2020).

Warehouse and fulfilment settings show the same dependence, since integrated path-planning and slotting optimisation must first convert layout and throughput requirements stated by operations staff into a combined formal model (Akanbi & Sunday, 2024). Building the bridge between coded stakeholder material and formulation-ready problem descriptions is, on the evidence of this review, the single largest structural gap in the field.

Manufacturing settings supply particularly clean examples of the same difficulty. Non-destructive evaluation of weld integrity formalises defect categories whose boundaries rest on inspector judgement (Atta *et al*., 2021).

5.5 Layer 5: Preference and weight elicitation

Multi-criteria weight elicitation is cognitively demanding and error-prone. Pairwise comparison approaches impose a burden that grows quadratically in the number of criteria, producing fatigue and inconsistency. Methodological responses within the multi-criteria community have targeted this directly, through reduced comparison sets, deck-of-cards variants and ordinal regression formulations that trade precision for tractability.

The applied literature shows how frequently the problem arises. Quantitative evaluation of environmental, social and governance adoption within supply chains requires weighting across incommensurable dimensions (Akinleye, 2023a). Integrating such metrics into infrastructure auditing compounds the problem by adding temporal weighting (Okojie *et al*., 2023b). Multi-layered risk architecture for public-private partnerships requires both likelihood and consequence weights from parties with divergent exposure (Ike *et al*., 2024). Data visualisation research on procurement decision-making finds that presentation format materially affects the weights decision-makers reveal, which is a direct challenge to the assumption that elicited weights are stable properties of the decision-maker (Babatope *et al*., 2023).

Risk-based decision-making in pharmaceutical manufacturing formalises this as a structured judgement process with defined criteria and thresholds (Aneke & Adesemoye, 2020), and regulatory readiness maturity models embed weighting decisions in the scoring rubric itself (Eze *et al*., 2022b). Maturity models for predicting airport safety audit outcomes do the same, with the weights implicit in the dimension structure rather than elicited explicitly (Adeyelu & Dagodzo, 2024). Capital allocation and financial planning models require explicit trade-off parameters between growth, risk and liquidity that are ultimately judgemental (Oduleye & Medon, 2023).

Four natural applications of language models at this layer can be identified, none of which has a substantial evaluation literature.

Conversational weight elicitation would derive implicit trade-offs from narrative discussion rather than from explicit comparisons, returning a weight vector with a stated uncertainty range. The elicitation machinery of Layer 1 supplies the mechanism; what is missing is validation against conventionally elicited weights.

Text-to-criteria extraction would induce criteria and their hierarchical structure from stakeholder documents and transcripts rather than having the analyst propose them.

Consistency repair would detect and interrogate inconsistent judgement patterns during elicitation rather than after it, converting a post-hoc correction into a conversational clarification.

Synthetic preference profiles would support robustness analysis, testing whether a recommended alternative survives across plausible weight configurations rather than a single elicited one. This is the use of Layer 3 outputs that the evidence best supports, since it requires only that generated profiles be plausible rather than that they be accurate.

The thinness of this layer is notable given that it is arguably where artificial intelligence could contribute the most decision-relevant value per unit of effort. It may reflect disciplinary separation, since the multi-criteria community and the language modelling community have limited overlap in venues and vocabulary.

5.6 Layer 6: Explanation, verification and governance

Two contributions are indicative. Chen *et al*. (2024) applied language models to diagnosing infeasible optimisation problems, which is a genuinely stakeholder-facing task, since infeasibility usually means that one party's stated constraint is incompatible with another's, and the diagnostic output is an input to renegotiation rather than a technical artefact. Mostajabdaveh *et al*. (2024) build verification into the formulation process itself.

The broader position emerging is that explanation is not a presentation-layer concern. If a formulation is derived from stakeholder text by an opaque process, the stakeholders cannot validate it, and validation by the people who supplied the inputs is the only check that catches misinterpretation of intent as opposed to error in execution.

The operations research literature at this layer is nascent, but an adjacent body of work on algorithmic accountability and model risk management is well developed and directly applicable.

Reviews of algorithmic accountability and model risk management in financial institutions establish the basic control architecture, including model inventory, validation independence and challenge functions (Ominyi & Anichukwueze, 2023). Conceptualisations of responsible governance structures for automated decision systems in regulated contexts extend this to systems whose outputs are decisions rather than predictions (Ominyi & Anichukwueze, 2024a). Systematic review of organisational readiness for the European Union Artificial Intelligence Act supplies a maturity framing that translates readily to analytical functions (Ominyi *et al*., 2024b).

Complementary work addresses specific control problems. Algorithmic accountability under trade secret protection identifies a structural obstacle to the transparency that stakeholder validation requires (Annan, 2024). Data privacy governance for cross-border platforms bears on the practicalities of elicitation across jurisdictions (Annan, 2022).

Security considerations for agentic workflows are directly relevant to Layer 4 systems. Work on securing enterprise agentic workflows against prompt injection, tool poisoning and memory manipulation describes attack surfaces that a formulation pipeline consuming untrusted stakeholder text inherits by construction (Nwakamma *et al*., 2024b).

Sector-specific governance work supplies further structure. Comparative governance of AI-driven healthcare management addresses executive oversight and regulatory interface (Afrihyia *et al*., 2024). Data integrity and governance in regulated life sciences operations codifies expectations that translate to analytical pipelines (Aneke *et al*., 2023). Platform governance and deployment tooling in regulated customer systems address the access control layer that stakeholder data collection depends on (Badmus *et al*., 2024).

The operations research community would gain more from importing this apparatus than from rebuilding it.

Data governance work supplies the control layer on which stakeholder pipelines depend. Enterprise data sensitivity classification and retention frameworks determine what stakeholder material may be processed and for how long (Aliliele *et al*., 2024b), and risk-based business intelligence architecture formalises the exposure that follows from getting that wrong (Mbonu *et al*., 2021). Data protection impact assessment models for multi-cloud financial environments bear directly on multi-jurisdiction elicitation (Mbonu *et al*., 2022b), while identity and access management integration determines who can see the material once it has been collected (Mbonu *et al*., 2020). API governance and risk prioritisation frameworks are relevant wherever a pipeline calls external model services (Aliliele *et al*., 2023b), and work on artificial intelligence enabled general controls with audit automation extends the control environment to cover the analytical machinery itself (Mbonu *et al*., 2022a).

Where stakeholder data cannot leave a device or a jurisdiction, deployment constraints become design constraints, and work on energy, reliability and latency trade-offs in green communication systems characterises the frontier that on-device processing operates against (Asiedu & Quainoo, 2024b).

6. Applications Across Domains

6.1 Procurement and supply chain

This is the densest application area, because procurement decisions consume stakeholder judgement at every stage and because the resulting models are consequential.

Supplier evaluation and bid selection combine measured performance with assessed capability, and geospatially informed evaluation frameworks illustrate how heterogeneous the input types are (Ahiaeke Patrick *et al*., 2021). Digital supply chain governance frameworks in energy and infrastructure define the decision rights within which such assessments are made (Okonkwo *et al*., 2024a), and predictive procurement planning to sustain operational uptime requires demand signals that originate in maintenance judgement rather than in transaction history (Okonkwo *et al*., 2024b). Port logistics efficiency and demurrage elimination models depend on process descriptions elicited from operators (Okonkwo *et al*., 2020).

Automated coding of supplier correspondence, tender documentation and audit findings is the most immediately practical application. End-to-end visibility frameworks improving transparency and traceability require classification of exception events that are reported in free text (Nnabueze *et al*., 2021). Automation and digital twin frameworks for reducing procurement errors depend on accurate specification capture (Akinlade *et al*., 2024).

Risk and resilience applications are more speculative but potentially more valuable. Resilience index construction is explicitly a weighting exercise over perceived exposures (Efobi *et al*., 2023). Predictive vendor risk assessment models require the same (Tonoyan *et al*., 2024a). Integrated frameworks for artificial intelligence and predictive analytics in supply chain management describe the target architecture (Rukh *et al*., 2024).

Two contributions supply useful caution. Localised supply chain solutions for community development make a related point about the mismatch between centrally specified models and locally held knowledge (Anene & Clement, 2024). Workforce optimisation frameworks addressing staffing volatility depend on availability and preference data that workers have incentives to misreport (Falemi *et al*., 2023). Order promising sits at the same boundary: a quoted lead time is a commitment made under uncertainty, and the risk threshold a firm accepts when issuing it is a managerial judgement rather than a measured quantity. Work on online lead time quotation under contingency formalises the quoting decision as a sequential problem under disruption, which makes explicit the parameters that would otherwise remain implicit in a planner's discretion (Ekun & Muriel, 2024).

6.2 Healthcare operations and public health

Healthcare generates large volumes of free-text patient and staff feedback, has multiple stakeholder groups with divergent objectives, and makes operational decisions that are resource-allocation problems. It is therefore the most active application area for automated coding. Analytics-driven work on productivity and workflow inefficiency depends on process descriptions supplied by staff (Nnaji & Akinlolu, 2023), early detection systems for non-communicable disease consume clinical narrative alongside measured indicators (Afrihyia *et al*., 2023), and governance and accountability models for public-private partnerships in health infrastructure add a negotiated allocation of risk between parties (Aminu-Ibrahim *et al*., 2024).

Reported uses include inductive thematic analysis of patient and clinician interviews to identify service-delivery constraints (Mathis *et al*., 2024), classification of patient feedback to surface recurring operational failures, and blinded comparison of automated and human analysts on clinical qualitative material. The typical pipeline runs from unstructured feedback through automated coding to a set of criteria or constraints that then enter a scheduling, capacity or prioritisation model.

The applied literature supplies the demand cases, of which governance-driven workforce planning is representative, since staffing models depend on acuity judgements rather than on measured workload alone (Omaghomi *et al*., 2023).

At system level, the data backbone required for transformation is described as a governance and architecture achievement rather than a technical one (Ilodigwe & Adesemoye, 2021), and real-world evidence and predictive analytics for decision-making sit on top of it (Ilodigwe & Adesemoye, 2024).

In public health, national real-time surveillance dashboards require classification of case narratives and symptom reports at speed (Oparah *et al*., 2023), and business intelligence for mental health resource allocation depends on classification of presenting need where category boundaries are contested (Anioke & Atima, 2023b). Drug shortage early warning models combine supply signals with reported clinical impact (Eze *et al*., 2022a).

Healthcare infrastructure planning is a distinct and underappreciated case, because the decisions are capital-intensive, long-lived and justified by projected population benefit that must be elicited rather than measured. Infrastructure-driven expansion of diagnostic access and programme management for coordinated multi-site delivery both depend on risk and dependency assessments supplied by project participants (Aminu-Ibrahim *et al*., 2021).

At facility level the pattern repeats. Risk-managed construction strategies for highly regulated healthcare environments depend on assessed rather than observed risk (Ogbete *et al*., 2021), and lifecycle performance evaluation of purpose-built diagnostic laboratories combines measured throughput with judgements about fitness for purpose (Ogbete *et al*., 2023).

Pharmacy and medicines access supply a further set of demand cases. Reviews of last-mile drug distribution barriers to underserved communities and of supply chain inefficiencies affecting medication affordability both synthesise qualitative field evidence into constraints that distribution models then consume (Asiedu & Asiedu, 2023), and conceptual work on equitable access to chronic therapies in hospital pharmacy settings makes the equity weighting explicit (Asiedu & Asiedu, 2024a).

Clinical service research contributes evaluation designs that operational models borrow.

6.3 Occupational safety and risk management

Safety is an underappreciated application area for this technology, and it is one where the input data is almost entirely textual.

Near-miss and hazard observation systems generate short free-text reports at volumes that exceed human coding capacity, which is precisely the constraint that automated deductive coding addresses (Obogo *et al*., 2021). Data-driven occupational safety risk control models consume the coded output (Obriki & Arumosoye, 2020). Behavioural safety programmes generate observation data with the same properties (Obriki *et al*., 2023a), and incident investigation approaches oriented toward prevention require causal attribution coding where inter-rater agreement is historically weakest (Obriki & Arumosoye, 2024).

Two constructs in this domain are pure perception measures and therefore candidates for adaptive elicitation. Safety culture maturity models rest on organisational learning indicators that are self-reported (Arumosoye & Obriki, 2021), and safety leadership influence in large temporary project organisations is measurable only through perception instruments (Arumosoye & Obriki, 2024). Recurrence mechanisms of unsafe behaviour are similarly behavioural constructs (Obriki & Arumosoye, 2022).

Aviation safety oversight presents a structurally similar problem with a stronger regulatory frame. Managing encroachments and unauthorised structures around aerodromes requires extracting requirements from heterogeneous regulatory and operational documentation (Adeyelu, 2021). Risk-tiered oversight models for unmanned operations in shared airspace require judgement-based tiering (Adeyelu, 2023), and maturity models for predicting audit outcomes embed weighting in the scoring structure (Adeyelu & Dagodzo, 2024).

Information security awareness sits alongside these as a behavioural risk domain where the measurement instrument is the survey and where training effectiveness evaluation depends on self-reported behaviour change (Onche *et al*., 2024).

6.4 Energy transition, power systems and asset maintenance

The energy domain has moved furthest toward treating stakeholder acceptance as a modelled variable rather than an external condition, which makes it the most interesting test case for this technology.

Predicting community trust in renewable projects converts social license into a measurable quantity (Komi, 2024a), and modelling the causal link between community trust and project survival makes the operational consequence explicit (Komi & Adeniji, 2024d). Just transition analysis combines econometric estimation with skills mapping from workforce instruments (Komi & Adamolekun, 2024b), and diversification pathway modelling extends this to economic outcomes beyond direct job replacement. Reviews of financing mechanisms that fail to reach poor households are constructed from programme evaluation evidence whose synthesis is a coding problem (Komi & Adeniji, 2024c).

Fairness in forecasting is a distinctive contribution from this domain. Work showing that accurate load forecasts are not necessarily fair forecasts identifies a failure mode with direct analogues in stakeholder data synthesis, since a model can minimise aggregate error while systematically misrepresenting subgroups (Komi, 2023). Predictive analytics for smart city emissions and infrastructure risk monitoring extends the pattern to urban systems (Okojie *et al*., 2023a).

Power systems engineering supplies the technical counterpart to the social acceptance work discussed above, and it is where judgement enters as engineering assessment rather than public opinion. Digital twin modelling for real-time monitoring and fault detection in smart grids, and AI-powered digital twins for predictive maintenance, depend on failure taxonomies that originate in maintenance records written as free text (Shittu *et al*., 2023).

Maintenance and asset management are the clearest case of free-text data feeding an operational model. Vibration-based condition monitoring of rotating machinery and sensor-based fault detection in building climate systems generate instrumented signals that must nonetheless be interpreted against maintainer judgement about severity (Sunday & Omoegun, 2022).

Investment and deployment decisions close the loop. Data-driven portfolio optimisation across solar, wind and storage assets requires scenario weights that are ultimately judgemental (Ilesanmi *et al*., 2024), and blockchain-based peer-to-peer energy trading embeds assumptions about participant behaviour that are elicited rather than observed (Shittu *et al*., 2021). Project management research on solar deployment identifies stakeholder alignment rather than technical feasibility as the dominant delivery risk (Amayo *et al*., 2024).

6.5 Financial planning, audit, financial crime and payments

Enterprise financial decision systems consume expert judgement extensively, and the audit function supplies a natural verification analogue for the pipelines described in Layer 6.

Data-driven cost management models for strategic financial planning aggregate operational signals with management judgement (Oduleye & Medon, 2021), and decision analytics models for sustainable profitability formalise the trade-offs (Lawal & Oduleye, 2023). Variance analytics for cost control depends on explanation classification, which is a coding task (Aliliele *et al*., 2024a), and comprehensive financial reporting models formalise the compliance constraints that classification must satisfy (Medon & Oduleye, 2022). Integrated predictive analytics for strategic forecasting and quantitative measurement of financial analysis impact both rest on judgemental inputs (Oduleye & Medon, 2023).

On the assurance side, integrated data management for financial forecasting and predictive risk assessment describe the control environment (Agu *et al*., 2023). Risk-based auditing in strategic financial management and data-driven compliance approaches to risk governance together sketch a validation apparatus that analytical pipelines could adopt (Akomolafe *et al*., 2023).

Financial crime detection is a large and distinct branch of the same problem, in which the label being predicted is itself the output of an analyst judgement. Anomaly detection in financial time series is framed conceptually rather than purely statistically for the same reason (Atakpa & Fobellah, 2023). Adaptive fraud risk scoring for real-time transaction monitoring must set thresholds that trade false positives against missed cases, which is a weighting decision rather than a modelling one (Fadayomi *et al*., 2024), and detection models for synthetic identity fraud face a target that is defined by convention (Elebe *et al*., 2023).

Audit and assurance research supplies the counterweight. Venture and SME risk analytics show the same dependence on assessed rather than measured inputs (Eboh & Aliliele, 2024), as does machine learning in insurance underwriting, where the historical decisions being learned from encode prior underwriter judgement (Oluwo *et al*., 2024).

Payments and cross-border operations extend the pattern into multi-jurisdiction settings, where harmonising risk governance, technology infrastructure and compliance frameworks is a coordination problem across parties holding different objectives and facing different regulators (Okafor *et al*., 2024b).

Public sector financial management supplies a final set of cases in which the constraint is institutional. Strategic finance in joint venture and cross-border structures depends on assumptions negotiated between partners rather than derived from data (Isiekwu *et al*., 2021), and lifecycle performance management models supply the operational counterpart (Adelanwa *et al*., 2023a).

6.6 Marketing, customer operations and revenue systems

Marketing research has engaged with synthetic respondents more systematically than any other applied field, and its conclusions are the most explicitly bounded.

Customer acquisition and retention frameworks depend on segmentation whose validity rests on behavioural anchoring (Ogundairo & Ayivi-Donkor, 2023). Predictive maintenance modelling for e-commerce systems and traffic pattern forecasting across owned media both consume operational signals with substantial human interpretation (Oghenemaiga *et al*., 2024). Customer relationship and workflow automation in small healthcare practices shows the same pattern at small scale (Eyetsemitan *et al*., 2024a).

6.7 Lean operations and small enterprise process improvement

Process improvement methodologies are stakeholder elicitation methodologies in all but name, which makes them a natural adoption site.

Lean Six Sigma adaptation for small enterprises depends on process mapping conducted with operators (Ambali *et al*., 2021). Data-driven process optimisation in micro-enterprises and integrated lean-digital scaling frameworks both require the same upstream elicitation (Eyetsemitan *et al*., 2024b). Cycle time reduction in professional services onboarding is a classic elicitation-then-optimise problem (Oyeleye *et al*., 2022). Algorithmic process optimisation with simulation modelling and predictive cost optimisation both depend on process parameters that originate in interview data (Tonoyan *et al*., 2022b), and real-time performance monitoring systems close the loop (Tonoyan *et al*., 2024b).

6.8 Education, workforce and inclusive service design

Education and workforce applications supply instruments and evaluation designs that operations research borrows. Evidence-based digital learning platforms and real-time educational performance analytics both formalise feedback loops from learner data (Orise & Niniola, 2022).

Special education and inclusive provision deserve separate treatment, because the decisions are individualised, legally constrained and documented in narrative form, which makes them an almost ideal test case for automated coding. Systematic review of behaviour intervention plan effectiveness in secondary settings depends on coding plan content and implementation notes (Oloto & Yeboah, 2023).

Individualised planning research supplies the operational counterpart. Data-driven development of individualised education programmes proposes conceptual models for turning progress data into plan content (Yeboah & Oloto, 2022b), monitoring implementation fidelity addresses whether the plan is executed as written (Yeboah & Oloto, 2023), and using student progress data to personalise instruction is the classroom-level instance of the same loop (Yeboah *et al*., 2022a).

Instructional and institutional research completes the picture. Integrating information and communication technology with mathematical thinking evaluates a digital intervention through learner outcome instruments (Boakye *et al*., 2020), and policy and practice research on educational technology for inclusive learning addresses the conditions of adoption (Ijiga *et al*., 2024). Work on artificial intelligence for student engagement and motivation, on hyperpersonalisation and hyperlocalisation as affordances, is the closest analogue in the applied corpus to Layer 1 adaptive elicitation, since the system adapts to the individual rather than administering a fixed sequence (Seun *et al*., 2024).

6.9 Built environment and infrastructure asset monitoring

Built environment decisions are long-lived, spatially fixed and justified by projected occupant experience, which makes them unusually dependent on elicited data.

Techno-economic evaluation of renewable-material construction for low-income housing combines cost modelling with affordability assessments drawn from household instruments (Uduokhai *et al*., 2024b), and critical review of housing policy implementation strategies synthesises programme evidence into constraints that allocation models then apply (Uduokhai *et al*., 2023c). Post-occupancy satisfaction modelling in government-sponsored housing is a direct instance of Layer 1 and Layer 2 working together, since satisfaction is measured by instrument and the resulting categories feed design standards (Uduokhai *et al*., 2023a).

Design and delivery research supplies further demand cases. Design thinking approaches to architectural practice require criteria weights across cost, durability and sustainability dimensions that participants supply rather than derive (Uduokhai *et al*., 2023d). Building information modelling research addresses the coordination of specifications originating from multiple disciplines (Uduokhai *et al*., 2023b), and system dynamics modelling of circular economy integration formalises feedback structures elicited from sector participants (Uduokhai *et al*., 2024a).

Remote asset monitoring illustrates what changes when elicitation is partly replaced by instrumentation. Integrated frameworks combining aerial survey, LiDAR and geographic information systems formalise the resulting data into asset models (Dagodzo & Ahiaeke Patrick, 2021a), and encroachment detection using geospatial techniques converts a judgement about right-of-way violation into a classification task (Dagodzo & Ahiaeke Patrick, 2022a). Integrating drone operations into enterprise geographic information systems is the organisational counterpart (Dagodzo & Ahiaeke Patrick, 2023), while deep learning for vegetation classification in power corridors shows the classification layer directly (Dagodzo *et al*., 2022b). The instructive point for this review is that instrumentation displaces elicitation only for the observable portion of the problem. Whether an encroachment matters, and how urgently, remains a judgement.

6.10 Agricultural and food systems

Agricultural development is the clearest case of operations research dependent on survey data collected under difficult conditions. Digital agriculture tool assessment, climate-smart practice contribution to household food security, socioeconomic barriers to technology adoption, quantitative agricultural economics modelling for food system efficiency, and assessment of artificial intelligence in agricultural decision-making all rest on smallholder instruments whose cost and reach constrain what can be modelled (Michael & Ogunsola, 2021b, 2022).

Agricultural and livestock research supplies the clearest example of operational modelling built on primary data collected under difficult field conditions. Phenotypic characterisation of indigenous poultry across regions depends on structured field observation protocols whose consistency across enumerators is a live concern (Ekeocha *et al*., 2021). Feeding trials evaluating farm residues, alternative energy sources and silage supplementation generate performance data that feed least-cost ration formulation, which is a classical optimisation problem with judgement-laden constraints (Tawose *et al*., 2023a). Physiological and stress indicator work under field conditions introduces welfare constraints that are assessed rather than measured (Oluwadele *et al*., 2023). Aflatoxin detection and quantification in commercial feeds establishes safety thresholds that then enter procurement decisions as hard constraints (Tawose *et al*., 2023b), and integration of artificial intelligence with biosensors for rapid pathogen detection points toward instrumented replacement of some of that judgement (Olorunkosebi *et al*., 2023).

7. Discussion

7.1 Principal findings

Four findings emerge from the synthesis.

The first concerns the shape of progress. Capability is high and improving where the transformation is constrained by an externally supplied framework, and markedly lower where the framework must itself be induced from the material. Deductive coding against a stated codebook performs substantially better than inductive theme generation, and translation of well-posed problem statements performs better than interpretation of contested ones. This is the pattern Proposition 1 anticipated, and it holds across both Layer 2 and Layer 4 despite the very different techniques involved.

The second concerns the distribution of effort. Layer 4 has benchmarks, competing architectures, successive system generations and a traceable progress record. Layer 5 has almost nothing, despite being the layer at which a single parameter choice can reverse a recommendation. Layer 1 is developing quickly but has only one controlled comparison against human interviewers in the retrieved corpus. The field has invested most heavily where the inputs are cleanest.

The third concerns the character of the errors. The critical literature does not report that these systems are simply noisy. It reports that persona conditioning amplifies group stereotypes and that models drift toward socially acceptable answers on sensitive items. These are distinct phenomena with a common consequence, which is that the error is structured rather than random.

The fourth concerns evaluation. Every layer that has an evaluation tradition measures a proxy. Formulation benchmarks measure translation of clean text. Coding studies measure agreement with a human standard that is itself variable. Synthetic respondent studies have largely measured distributional similarity, a criterion the field is now abandoning as insufficient. Nothing measures whether the resulting decision differs from the one careful conventional analysis would produce.

7.2 The inversion problem

Taken together these findings describe an inversion. The capability profile of the technology is the reverse of the difficulty profile of the practice.

In operations research, the hardest and most consequential work occurs early, when the problem is still being defined and competing framings are live. That is precisely where automated support is weakest, because the task requires inducing structure rather than applying it. The easiest and most mechanical work occurs later, once the formulation is settled and the remaining question is technical. That is where automated support is strongest.

The practical risk this creates is not that the tools fail visibly. It is that they succeed on the tractable portion of the pipeline and thereby increase the throughput of a process whose upstream quality has not improved. A pipeline that formalises ten times as many problem descriptions ten times as fast, from stakeholder material coded no more accurately than before, produces more models of the same fidelity and reduces the time available to notice when one of them has misread what the stakeholders meant.

7.3 What the benchmarks measure

Layer 4 has benchmarks and therefore has comparability, but those benchmarks measure translation of already-clean descriptions. Layer 2 has borrowed reliability statistics from content analysis, which is appropriate but only assesses agreement with a human standard that is itself variable. Layer 3 evaluation is in active dispute, with distributional fidelity increasingly regarded as insufficient. Layers 1 and 5 have essentially no shared evaluation instruments.

A field-level consequence follows. Current benchmarks reward performance on the easiest sub-problem in the pipeline and are silent on the hardest. Nothing currently measures whether a model constructed from automatically synthesised stakeholder data recommends a different decision than one constructed conventionally, which is the only quantity that matters operationally.

7.4 Reproducibility and analytic flexibility

Non-determinism in model outputs, undocumented prompt variation, model version drift and proprietary interface changes together make exact replication rare. Reporting standards have improved but remain inconsistent. The analytic flexibility argument means that under-specified reporting is not a minor omission but a mechanism by which arbitrary conclusions survive review.

The model governance literature offers relevant discipline here. Model inventories, versioning requirements and validation independence are standard expectations in regulated financial contexts (Ominyi & Anichukwueze, 2023), and there is no principled reason why analytical pipelines feeding operational decisions should be held to a lower standard.

7.5 Bias and its distributional shape

Two distinct bias phenomena are documented and should not be conflated, and a third is a plausible extension of them that the evidence does not yet settle.

Caricature and stereotype amplification occurs when persona conditioning exaggerates group differences (Bisbee *et al*., 2024). Social desirability or harmlessness drift occurs when models shift toward acceptable answers on sensitive items (Sun *et al*., 2024). Both are documented directly. The third, characteristic-correlated error, in which accuracy varies systematically with attributes of the person whose text is being processed, follows from the first two by extension rather than from direct measurement, since evidence that outputs align more closely with some demographic groups than others (Santurkar *et al*., 2023) implies unequal error rates without establishing them for coding tasks specifically.

The third is the most consequential for operations research, because it produces biased parameters rather than noisy ones, and optimisation propagates bias in parameters into systematically unfair allocations. The energy fairness literature has already articulated the general form of this problem in the forecasting context, showing that a model can minimise aggregate error while systematically misrepresenting subgroups (Komi, 2023), and universal programme design research documents how apparently neutral instruments deepen existing inequality (Komi & Adeniji, 2024c).

Performance gaps in non-Western and non-English contexts are documented and consistent with training corpus composition, which matters given that a substantial share of practice on resource-allocation problems occurs in exactly those contexts. Multilingual adaptation work offers a partial technical response (Oyesiji *et al*., 2023), but adaptation improves fluency more reliably than it improves cultural calibration.

7.6 Security and integrity of the pipeline

A formulation pipeline that consumes stakeholder text is an application that processes untrusted input. Prompt injection, tool poisoning and memory manipulation are live attack surfaces for agentic workflows (Nwakamma *et al*., 2024b), and a stakeholder consultation is an obvious injection vector where the participants have an interest in the outcome. This consideration is almost entirely absent from the operations research literature on autoformulation and deserves attention before such systems are used in contested settings such as procurement scoring or grant allocation.

The adjacent security literature makes the scale of that surface concrete. Threat actor analysis for strategic enterprise security planning addresses precisely the case in which the party supplying data to a system has an interest in its output (Dosunmu & Ogundele, 2023). Intrusion detection and prevention models describe the monitoring an analytical pipeline would sit behind (Dosunmu & Ogundele, 2020), and threat intelligence integration frameworks describe how an organisation maintains awareness of evolving attack methods (Dosunmu & Ogundele, 2022). Three contributions are especially relevant to verification. Breach and attack simulation frameworks provide continuous validation of controls rather than point-in-time assessment (Dosunmu & Ogundele, 2024a), threat-informed defence engineering measures control effectiveness against specific adversary behaviours (Dosunmu & Ogundele, 2024c), and cyber risk quantification models convert qualitative exposure into figures that support investment prioritisation, which is the same elicitation-to-parameter problem this review examines throughout (Dosunmu & Ogundele, 2024b).

Where analytical pipelines control physical processes the stakes rise further. Adversarial machine learning in critical infrastructure describes attacks that manipulate model inputs rather than breaching systems (Adebayo *et al*., 2022), and AI-augmented threat detection in industrial control systems describes the defensive counterpart (Adebayo *et al*., 2023). Security architecture for the convergence of information and operational technology in regulated energy environments addresses the boundary at which analytical systems meet physical ones (Adegbite *et al*., 2022), and zero trust architecture for operational technology proposes an access model suited to it (Ahmed *et al*., 2021). Blockchain-assisted secure data exchange for supervisory control networks addresses integrity of the data stream itself (Shittu *et al*., 2022).

A further contribution bears on the input-manipulation problem at the infrastructure layer, where deep learning for malware classification addresses adversarial content that is designed to be misread (Idika *et al*., 2021). At the organisational level, identity-centric zero trust architecture determines who may submit material to a pipeline in the first place (Ogbole *et al*., 2021), and continuous monitoring for cloud misconfiguration addresses the most common practical failure in the environments where such pipelines run (Aliliele *et al*., 2023a).

7.7 Revisiting the propositions

Propositions 1, 2 and 6 are supported by the synthesis. The framework-constrained advantage in P1 appears consistently across layers. The inverse distribution of effort in P2 is evident from the relative maturity of Layers 4 and 5. The proxy-measurement problem in P6 is visible in every evaluation tradition surveyed.

Proposition 4 is partially supported. That the errors are structured rather than random is established for stereotype amplification and for drift on sensitive items. That the structure tracks individual respondent characteristics closely enough to bias an allocation is a reasonable inference from those findings rather than a measured result, and it is the claim most in need of a dedicated study.

Proposition 3 is supported in its negative half and untested in its positive half. The evidence that synthetic respondents should not substitute for consultation is strong. The claim that they are useful for pretesting and robustness analysis is plausible and endorsed by the marketing literature, but the retrieved corpus contains no study that demonstrates decision-relevant value from such use.

Proposition 5 is argued rather than demonstrated. No study in the corpus tests whether stakeholder validation of an automatically derived formulation catches errors that verification alone misses. This is a gap worth naming precisely, because P5 carries much of the normative weight of the paper.

8. Implications

8.1 Theoretical implications

The six-layer model contributes a way of relating two literatures that have developed largely in isolation. Problem structuring methods and multi-criteria decision analysis have long theorised the upstream stages of the operations research process, while the computational literature on autoformulation has theorised the downstream translation step. Placing them on a single pipeline makes visible the transition between them, which neither literature currently owns, and identifies it as a research object rather than an implementation detail.

A second theoretical implication concerns the status of stakeholder data. The synthesis suggests that such data cannot be treated as a noisy measurement of an underlying parameter, which is the implicit assumption when synthetic augmentation is proposed. Where a model's authority derives from consultation, the data is partly constitutive of the model's legitimacy rather than merely informative about its parameters. That distinction has no representation in current formalisms and deserves one.

A third implication is that fairness in operations research needs an upstream formulation. The energy forecasting literature has already established that a model can minimise aggregate error while systematically misrepresenting subgroups. The coding evidence extends the same structure to qualitative inputs. Fairness constraints applied at the optimisation stage cannot correct for bias introduced during elicitation and coding, because by then the misrepresentation is embedded in the parameters the constraints operate on.

8.2 Methodological implications

For researchers, the most immediate implication is that reporting standards need to change in three specific ways. Subgroup breakdowns of coding and simulation accuracy should be standard rather than exceptional, given the documented correlation of error with respondent characteristics. Human-in-the-loop configurations should be specified precisely, since reporting that a study used human validation is uninformative without saying where human judgement entered and how disagreement was resolved. And prompting, conditioning and model version choices should be reported in full, because the space of defensible analytic choices is wide enough that under-specification permits arbitrary conclusions to survive review.

A second methodological implication concerns benchmark construction. The field needs a benchmark whose inputs are authentic and unclean, paired with expert-validated formulations, and whose scoring rewards flagging ambiguity rather than resolving it silently. Existing benchmarks reward the opposite behaviour.

8.3 Implications for practice

For analysts and modelling teams, the synthesis supports a specific division of labour rather than wholesale adoption or abstention.

Automated deductive coding is ready for use where a codebook exists and has been validated, which describes near-miss and hazard reporting, supplier audit findings, patient feedback classification and consultation response coding. It should be deployed with human adjudication of a sampled subset and with reliability reported, not assumed.

Adaptive elicitation is ready for pilot use where the alternative is a fixed-response instrument that is known to be inadequate, which describes safety perception surveys, adoption barrier research and multilingual needs assessment. Comparability across respondents becomes a live question once the instrument adapts, and analysts should expect to defend it.

Autoformulation is ready for use as a drafting aid on well-specified problems, with the formulation treated as a proposal requiring verification rather than as an output. Constraint provenance should be recorded so that each model element can be traced to the input that generated it.

Synthetic respondents are ready for instrument pretesting and for robustness analysis over plausible preference configurations. They are not ready to substitute for consultation, and using them that way in procurement scoring, grant allocation or service planning would be difficult to defend if challenged.

8.4 Implications for policy and governance

Three governance implications follow directly.

Organisations deploying these pipelines in regulated or publicly accountable settings should extend existing model risk management arrangements to cover them rather than treating analytical tooling as out of scope. Model inventory, validation independence and challenge functions transfer with little modification.

Procurement and public consultation processes that use automated synthesis should disclose that they do so. The legitimacy argument in Proposition 5 is not satisfied by accuracy alone; affected parties have a reasonable interest in knowing how their input was processed, and the anti-discrimination exposure created by structured rather than random coding error is a legal risk as well as an ethical one.

Finally, pipelines that ingest stakeholder text are applications processing untrusted input, and in contested settings the participants have a material interest in the outcome. Security review should be part of deployment approval, not an afterthought.

Cross-border and multi-jurisdiction settings raise the governance question in its sharpest form, and an adjacent policy literature has already mapped much of the terrain. Reviews of legal barriers to trade participation identify compliance capability rather than tariff structure as the binding constraint for smaller participants (Jones & Ominyi, 2020), and systematic review of market access outcomes from trade agreements assesses whether formal liberalisation translates into realised access (Jones & Ominyi, 2022). Compliance frameworks for technology transfer between developed and developing economies bear directly on where analytical capability is allowed to reside (Jones & Ominyi, 2023). Two contributions address artificial intelligence explicitly: applications in customs administration and trade facilitation describe automated decision-making with immediate distributional consequences (Jones & Ominyi, 2024a), and harmonising digital trade regulation across jurisdictions describes the coordination problem that follows (Jones & Ominyi, 2024b). Diplomatic and regional policy research supplies the multi-actor analogue. Policy alignment modelling for climate diplomacy addresses coordination among parties with divergent objectives and no common authority (Liadi, 2022), and linking diplomatic engagement to development outcomes attempts to make that coordination measurable (Liadi, 2023). Soft power projection through education and cultural diplomacy and cross-border security cooperation modelling both formalise relationships that exist primarily in assessment rather than in data (Liadi, 2024b). These settings are the limiting case for the legitimacy argument in Proposition 5, since no party can be compelled to accept a model and acceptance is therefore the whole of the outcome.

9. Limitations

Publication and demonstration bias is severe. A large share of contributions report that a technique worked on a dataset the authors selected, with no pre-registered hypothesis and no comparison to a strong baseline. Negative results exist but are concentrated in a handful of deliberately critical papers.

Terminological drift complicates synthesis. Synthetic data, silicon sample, simulated respondent and generative agent are used with overlapping and inconsistent referents. The same is true of autoformulation, which sometimes denotes model construction and sometimes end-to-end problem solving. Similar drift affects the applied literature, where framework, model and architecture are frequently interchangeable.

Rapid model turnover undermines cumulative comparison. Results obtained with earlier generation models and results obtained with later reasoning models are not straightforwardly comparable, and few papers separate capability effects from method effects.

A fourth limitation is scope asymmetry between the two corpora. The methodological corpus was assembled through systematic term-combination searching, while the applied corpus was drawn from a consolidated reference collection. The applied corpus is therefore not a representative sample of applied practice and should be read as illustrative of the demand structure rather than as an estimate of its distribution.

A fifth is the perspective noted in Section 2.5. The judgement that Layer 5 is under-served and that the Layer 2 to Layer 4 transition is the central gap follows from an operations research framing of what matters.

The evidence base itself constrains what can be concluded, in ways that no review design can compensate for.

Controlled comparison is rare. Across the whole methodological corpus, the number of studies that randomly assign an automated and a human condition working from an identical protocol is very small. Most reported gains are measured against convenience baselines chosen by the authors.

Effect sizes are difficult to compare across studies because model versions differ, prompting is under-reported and metrics are inconsistent. Aggregation across studies would create a false impression of cumulative knowledge.

Negative results are concentrated in a small number of deliberately critical papers rather than distributed across the literature, which suggests that the true rate of failed demonstrations is higher than the published record indicates.

Evidence from non-English and non-Western settings is sparse relative to the share of applied practice occurring in those settings, so conclusions about performance are least reliable exactly where the consequences of error fall most heavily.

The six-layer model is a simplification in three respects that users of it should keep in view.

It presents as sequential a process that is iterative in practice. Analysts move back from formulation to elicitation repeatedly, and the figure's left-to-right arrangement understates that.

It assumes convergence on a single formalisation. Participatory traditions in which the purpose of the exercise is to sustain productive disagreement rather than resolve it are only partially described by the model, and the feedback loop in Fig 1 is doing more work in those settings than the diagram suggests.

It treats the six layers as separable, which is an analytical convenience. In agentic systems that elicit, code and formulate within a single workflow, the boundaries the model draws do not correspond to observable system components, and diagnosing where an error originated becomes correspondingly harder.

10. Future Research Directions

The agenda below is ordered by the ratio of decision-relevant value to research effort rather than by tractability. Items A to C address the structural gaps identified in Section 7.2, items D to F address the evaluation and legitimacy problems raised by Propositions 3, 5 and 6, and items G and H concern reporting and governance practice.

A. Build the Layer 2 to Layer 4 bridge: The single largest structural gap is the absence of methods that take realistically messy stakeholder material and produce formulation-ready problem descriptions, with ambiguity flagged rather than silently resolved. A benchmark of authentic, unclean stakeholder inputs paired with expert-validated formulations would advance the field more than another point of accuracy on existing benchmarks.

B. Treat verification as the primary object: Formulation verification, infeasibility diagnosis and constraint provenance tracking should be designed in rather than added on. Every model element should be traceable to the stakeholder input that generated it, which is the analytical analogue of the audit trail expectations already standard in regulated domains (Aneke *et al*., 2023).

C. Develop Layer 5 seriously: Conversational preference elicitation with calibrated uncertainty, consistency repair during elicitation, and robustness analysis over synthetic preference profiles are all tractable and all under-explored. Validation against conventionally elicited weights is the obvious first study.

D. Establish a use-boundary standard for synthetic respondents: The field needs an explicit, citable convention distinguishing legitimate uses, including instrument pretesting, sensitivity analysis and robustness checking, from illegitimate ones, including substituting for consultation and generating headline findings. Marketing research has moved furthest toward this and operations research should not reinvent it.

E. Evaluate at the decision level: Coding accuracy and formulation accuracy are proxies. What matters is whether the decision recommended by the resulting model differs from the decision that careful conventional analysis would produce, and whether stakeholders recognise the model as representing what they said. Neither is currently measured.

F. Address legitimacy directly: Where a model's authority derives from having consulted affected parties, automating the consultation changes the nature of the claim. This is a design constraint on the technical system rather than an afterword to it, and it needs treatment in the operations research literature rather than by reference to general ethics guidance.

G. Report distributionally: Aggregate accuracy should be accompanied by subgroup breakdowns as standard, given the documented characteristic-correlation of error.

H. Import the governance apparatus: Model risk management, validation independence, readiness assessment and auditability standards already exist in adjacent fields (Annan, 2024; Ominyi & Anichukwueze, 2024a; Ominyi *et al*., 2024b). Adapting them is faster and more defensible than developing parallel conventions.

11. Conclusion

The recent literature demonstrates that artificial intelligence can materially accelerate the parts of the operations research process that have always resisted acceleration. Conversational elicitation makes depth affordable at scale. Automated coding makes free-text stakeholder data tractable as model input, reliably for deductive application of an existing framework and less reliably for inductive discovery. Autoformulation systems have advanced quickly on well-posed problem descriptions and now have benchmarks, competing architectures and a visible progress trajectory.

The gaps are equally clear. The field is strongest where the input is already clean and weakest where meaning is still being negotiated, which inverts the priority order of actual practice. Preference elicitation, the layer with the highest ratio of decision impact to research attention, is nearly untouched. Evaluation practice measures translation fidelity rather than decision quality. A growing critical literature shows that the errors these systems make are not randomly distributed across respondents, which turns an accuracy problem into an equity problem once the output enters an allocation model. And the security properties of pipelines that consume contested stakeholder text have barely been examined.

The applied evidence surveyed here reinforces rather than softens these conclusions. Across procurement, healthcare, safety, energy, finance, marketing, process improvement and agricultural development, the pattern is consistent. Operational models depend on judgement data, the cost of collecting that data constrains what can be modelled, and the quality of its classification bounds the value of everything downstream. That is precisely the constraint this technology addresses. It is also precisely the point at which errors become invisible, because a fluent synthesis of stakeholder input is indistinguishable in form from an accurate one.

The reasonable position is neither adoption nor abstention. AI-enabled synthesis is well suited to expanding the volume and depth of stakeholder input that an analyst can process, and poorly suited to replacing the stakeholder input itself.

The technologies extend the analyst's reach. Treating them as a substitute for the people whose problem is being modelled misunderstands both what they can do and what the model is for.

12. References

1. Adebayo A, Adegbite MP, Ahmed MO. Adversarial machine learning in critical infrastructure: A conceptual framework for threat modeling AI enabled OT systems. World Journal of Innovation and Modern Technology. 2022; 6(1):184-234. Doi: https://doi.org/10.56201/wjimt.v6.no1.2022.pg184.234
2. Adebayo A, Adegbite MP, Ahmed MO. AI augmented threat detection in industrial control systems: A systematic review of machine learning approaches for ICS anomaly detection. International Journal of Engineering and Modern Technology. 2023; 9(3):287-340. Doi: https://doi.org/10.56201/ijcsmt.v9.no3.2023.pg287.340
3. Adegbite MP, Adebayo A, Ahmed MO. A security architecture model for IT and OT convergence in regulated energy networks: Design principles and governance alignment. International Journal of Computer Science and Mathematical Theory. 2022; 8(2):81-132. Doi: https://doi.org/10.56201/ijcsmt.v8.no2.2022.pg81.132
4. Adelanwa A, Basnet A, Anene UN. Data driven digital transformation models for lifecycle performance management in infrastructure delivery. International Journal of Advanced Multidisciplinary Research and Studies. 2023a; 3(6):2646-2662. Doi: https://doi.org/10.62225/2583049X.2023.3.6.5968
5. Adelanwa A, Basnet A, Anene UN. Predictive analytics models for financial risk detection and fraud prevention in public systems. International Journal of Advanced Multidisciplinary Research and Studies. 2023b; 3(6). Doi: https://doi.org/10.62225/2583049X.2023.3.6.5969
6. Adelanwa A, Basnet A, Anene UN. Advanced AI based decision support systems for healthcare operations and resource planning. Gyanshauryam, International Scientific Refereed Research Journal. 2024a; 7(1):244-276. Doi: http://doi.org/10.32628/GISRRJ
7. Adeyelu OO. A framework for managing encroachments and unauthorized structures around airport boundaries: Regulatory evidence, enforcement, and corrective action protocols for Nigerian civil aviation. International Journal of Multidisciplinary Research and Growth Evaluation. 2021; 2(6):954-972. Doi: https://doi.org/10.54660/.IJMRGE.2021.2.6.954-972
8. Adeyelu OO. A risk-tiered oversight model for unmanned aircraft operations in shared controlled airspace over Nigerian airports. International Journal of Scientific Research in Civil Engineering. 2023; 7(4):147-184. Doi: https://doi.org/10.32628/IJSRCE237419
9. Adeyelu OO, Dagodzo D. A maturity model for predicting airport safety audit outcomes in resource-constrained regulatory environments. International Journal of Scientific Research in Civil Engineering. 2024; 8(4):132-170. Doi: https://doi.org/10.32628/IJSRCE248423
10. Afrihyia E, Akinse SG, Ojukwu PU. Development and implementation of an AI-driven early detection system for non-communicable diseases. International Journal of Advanced Multidisciplinary Research and Studies. 2023; 3(6):2680-2691.
11. Afrihyia E, Akinse SG, Ojukwu PU. Comparative governance of AI-driven healthcare management: Executive oversight, regulatory structures, and accountability in the United States and developing countries. International Journal of Advanced Multidisciplinary Research and Studies. 2024; 4(6):3087-3102. Doi: https://doi.org/10.62225/2583049X.2024.4.6.5920
12. Agu MU, Akomolafe O, Bello A. A review of integrated data management systems for enhancing financial forecasting accuracy. International Journal of Advanced Multidisciplinary Research and Studies. 2023; 3(6):2335-2344.
13. Ahiaeke Patrick MC, Okonkwo CS, Mayo W, Okeke OT. Model for data driven vendor evaluation and bid selection using geospatial intelligence. Shodhshauryam, International Scientific Refereed Research Journal. 2021; 4(4):426-443. Doi: https://doi.org/10.32628/SHISRRJ214461
14. AhmadiTeshnizi A, Gao W, Udell M. OptiMUS: Scalable optimization modeling with (MI)LP solvers and large language models. In Proceedings of the 41st International Conference on Machine Learning. PMLR, 2024a.
15. AhmadiTeshnizi A, Gao W, Brunborg H, Talaei S, Lawless C, Udell M. OptiMUS-0.3: Using large language models to model and solve optimization problems at scale. arXiv, 2024b. https://arxiv.org/abs/2407.19633
16. Ahmed KS, Odejobi OD, Oshoba TO. Predictive model for cloud resource scaling using machine learning techniques. Journal of Frontiers in Multidisciplinary Research. 2020; 1(1):173-183.
17. Ahmed MO, Adegbite MP, Adebayo A. Zero trust architecture for operational technology in North American critical infrastructure: A framework for implementation and resilience optimization. International Journal of Engineering and Modern Technology. 2021; 7(1):66-113. Doi: https://doi.org/10.56201/ijemt.vol.7.no1.2021.pg66.113
18. Akanbi O, Sunday EA. An integrated path-planning and slotting optimization model for AMR-enabled high-density warehousing. International Journal of Advanced Multidisciplinary Research and Studies. 2024; 4(6):3226-3243. Doi: https://doi.org/10.62225/2583049X.2024.4.6.6181
19. Akin-Oluyomi OT, Atima ME, Akinleye OK, Okoruwa PO. Predictive analytics approaches for improving demand forecasting accuracy in e-commerce procurement systems. International Journal of Advanced Multidisciplinary Research and Studies. 2023a; 3(6):2173-2182.
20. Akinlade OF, Filani OM, Nwachukwu PS. Automation and digital twins framework reducing procurement errors and turnaround time. International Journal of Scientific Research in Humanities and Social Sciences. 2024; 1(1):197-216.
21. Akinleye OK. Framework for quantitative evaluation of ESG adoption within SME supply chains in emerging economies. International Journal of Advanced Multidisciplinary Research and Studies. 2023a;

3(6):2317-2334.
22. Akinleye OK, Adeyoyin O. A negotiation optimization model for reducing procurement costs in manufacturing firms. Shodhshauryam, International Scientific Refereed Research Journal. 2022a; 5(5):398-435. Doi: https://doi.org/10.32628/SHISRRJ225891
23. Akomolafe O, Agu MU, Bello A. A conceptual model for implementing risk-based auditing in strategic financial management. International Journal of Advanced Multidisciplinary Research and Studies. 2023; 3(6):2274-2286. Doi: https://doi.org/10.62225/2583049X.2023.3.6.5359
24. Aliliele C, Mbonu IS, Iwuanyanwu U. A conceptual framework for continuous cloud misconfiguration monitoring and enterprise risk mitigation strategies. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2023a; 9(10):373-394. Doi: https://doi.org/10.32628/CSEIT2361071
25. Aliliele C, Mbonu IS, Iwuanyanwu U. A review of API governance and risk prioritization frameworks in modern financial institutions. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2023b; 9(10):395-433. Doi: https://doi.org/10.32628/CSEIT2361072
26. Aliliele C, Eboh EE, Oluwo K. Advances in variance analytics for cost control in manufacturing and industrial enterprises. International Journal of Advanced Multidisciplinary Research and Studies. 2024a; 4(6):3163-3185. Doi: https://doi.org/10.62225/2583049X.2024.4.6.6012
27. Aliliele C, Mbonu IS, Iwuanyanwu U. A conceptual framework for enterprise data sensitivity classification and regulatory traceability mechanisms. International Journal of Advanced Multidisciplinary Research and Studies. 2024b; 4(6):3103-3124. Doi: https://doi.org/10.62225/2583049X.2024.4.6.5991
28. Amayo EB, Owulade OA, Isi LR. Strategic project management approaches for successful solar energy deployments in emerging markets: Lessons from West Africa. Iconic Research and Engineering Journals. 2024; 7(11):783-803.
29. Ambali KB, Eyetsemitan RA, Oyeleye AO, Fadayomi O. Lean Six Sigma for small enterprises: A systematic review and Lite-DMAIC adaptation framework for resource-constrained organizations. Iconic Research and Engineering Journals. 2021; 5(5):562-583. Doi: https://doi.org/10.64388/IREV5I5-1716957
30. Aminu-Ibrahim AY, Ogbete JC, Ambali KB. Program management models for coordinated multi-site healthcare infrastructure expansion projects. International Journal of Multidisciplinary Research and Growth Evaluation. 2021; 2(6):661-678.
31. Aminu-Ibrahim AY, Ogbete JC, Ambali KB. Governance and accountability models for public private partnerships in healthcare infrastructure development. International Journal of Advanced Multidisciplinary Research and Studies. 2024; 4(6):2943-2960. Doi: https://doi.org/10.62225/2583049X.2024.4.6.5699
32. Aneke OB, Adesemoye AC. A conceptual framework for risk-based decision-making in good manufacturing practice environments. Iconic Research and Engineering Journals. 2020; 4(6):325-341.
33. Aneke OB, Sydney E, Adesemoye AC. Data integrity and governance in regulated life sciences operations: A review of failures, controls, and remediation strategies. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2023; 9(10):538-568.
34. Anene UN, Clement T. A resilient logistics framework for humanitarian supply chains: Integrating predictive analytics, IoT, and localized distribution to strengthen emergency response systems. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2022; 8(5):398-424.
35. Anene UN, Clement T. Localized supply chain solutions for sustainable community development: A strategic model for economic revitalization and regional resilience. International Journal of Scientific Research in Science and Technology. 2024; 11(5).
36. Anioke SC, Atima ME. Public health governance models using process optimization and performance metrics for regulatory oversight. International Journal of Advanced Multidisciplinary Research and Studies. 2023a; 3(6):2534-2548. Doi: https://doi.org/10.62225/2583049X.2023.3.6.5491
37. Anioke SC, Atima ME. Business intelligence applications for mental health resource allocation and public health program accountability. International Journal of Advanced Multidisciplinary Research and Studies. 2023b; 3(6):2549-2563. Doi: https://doi.org/10.62225/2583049X.2023.3.6.5492
38. Annan AO. Data privacy governance models for cross-border digital platforms. International Journal of Multidisciplinary Research and Growth Evaluation. 2022; 3(6):949-964.
39. Annan AO. Algorithmic accountability and trade secret protection in artificial intelligence. Shodhshauryam, International Scientific Refereed Research Journal. 2024; 7(5):315-347.
40. Argyle LP, Busby EC, Fulda N, Gubler JR, Rytting C, Wingate D. Out of one, many: Using language models to simulate human samples. Political Analysis. 2023; 31(3):337-351.
41. Arumosoye OM, Obriki OD. Organizational learning-based conceptual maturity model for continuous safety performance improvement. International Journal of Multidisciplinary Research and Growth Evaluation. 2021; 2(1):970-980. Doi: https://doi.org/10.54660/.IJMRGE.2021.2.1.970-980
42. Arumosoye OM, Obriki OD. Conceptual model of safety leadership influence in large temporary project organizations. International Journal of Advanced Multidisciplinary Research and Studies. 2024; 4(6):3034-3047. Doi: https://doi.org/10.62225/2583049X.2024.4.6.5895
43. Asiedu CS, Asiedu AA. Pharmaceutical supply chain inefficiencies and medication affordability in resource-constrained health systems: A narrative review. International Journal of Health and Pharmaceutical Research. 2023; 8(4):192-222. Doi: https://doi.org/10.56201/ijhpr.v8.no4.2023.pg192.222
44. Asiedu CS, Asiedu AA. Equitable access to chronic therapies in hospital pharmacy settings: A conceptual framework for continuity of care. International Journal of Medical Evaluation and Physical Report. 2024a; 8(6):274-298. Doi:

https://doi.org/10.56201/ijmepr.v8.no6.2024.pg274.298
45. Asiedu W, Quainoo R. Rethinking energy, reliability, and latency trade-offs in green communication for next generation IoT. International Journal of Multidisciplinary Research and Growth Evaluation. 2024b; 5(6):1987-1994.
46. Astorga N, Liu T, Xiao Y, Van Der Schaar M. Autoformulation of mathematical optimization models using LLMs. arXiv, 2024.
47. Atakpa MI, Fobellah AN. Anomaly detection in financial time-series data: A conceptual model for healthcare and banking applications. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2023; 9(2):967-997. Doi: https://doi.org/10.32628/CSEIT2342441
48. Atta S, Tofah P, Adenuga OM. Advances in non-destructive testing methods for weld integrity evaluation in high-capacity power infrastructure. International Journal of Engineering and Modern Technology. 2021; 7(1):20-47. Doi: https://doi.org/10.56201/ijemt.vol.7.no1.2021.pg20.47
49. Babatope OM, Akokodaripon DA, Akinleye OK, Okoruwa PO. The role of data visualization in enhancing strategic procurement decision-making effectiveness. International Journal of Advanced Multidisciplinary Research and Studies. 2023; 3(6):2205-2212.
50. Badmus O, Jooda D, Anunagba CO. A privacy-by-design framework for role-based security architecture in Salesforce environments handling sensitive personal data. International Journal of Advanced Multidisciplinary Research and Studies. 2024; 4(6):3244-3254. Doi: https://doi.org/10.62225/2583049X.2024.4.6.6243
51. Bisbee J, Clinton JD, Dorff C, Kenkel B, Larson JM. Synthetic replacements for human survey data? The perils of large language models. Political Analysis. Advance Online Publication, 2024.
52. Boakye K, Ofori SD, Ogbona CS, Yeboah TJ, Bobga MA. Integrating ICT and mathematical thinking: Exploring the effectiveness of digital tools in secondary mathematics instruction. Iconic Research and Engineering Journals. 2020; 4(4):372-398. Doi: https://doi.org/10.64388/IREV4I4-1714179
53. Checkland P. Systems thinking, systems practice. Wiley, 1981.
54. Chen H, Constante-Flores GE, Li C. Diagnosing infeasible optimization problems using large language models. INFOR: Information Systems and Operational Research. 2024; 62(4):573-587.
55. Chew R, Bollenbacher J, Wenger M, Speer J, Kim A. LLM-assisted content analysis: Using large language models to support deductive coding. arXiv, 2023. https://arxiv.org/abs/2306.14924
56. Chopra F, Haaland I. Conducting qualitative interviews with AI (CESifo Working Paper No. 10666). CESifo, 2023.
57. Dagodzo D, Ahiaeke Patrick MC. An integrated framework for UAV, LiDAR, and GIS in infrastructure corridor management. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2021a; 7(5):497-524. Doi: https://doi.org/10.32628/CSEIT217566
58. Dagodzo D, Ahiaeke Patrick MC. A review of right-of-way encroachment detection methods using geospatial technologies. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2022a; 8(1):638-667. Doi: https://doi.org/10.32628/CSEIT2281226
59. Dagodzo D, Ahiaeke Patrick MC. A framework for integrating drone operations into enterprise GIS systems for utility companies. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2023; 9(4):924-957. Doi: https://doi.org/10.32628/CSEIT23564536
60. Dagodzo D, Ahiaeke Patrick MC, Aliliele C. AI and deep learning for vegetation classification in power corridor management: A review. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2022b; 8(1):668-698. Doi: https://doi.org/10.32628/CSEIT2281227
61. De Paoli S. Performing an inductive thematic analysis of semi-structured interviews with a large language model: An exploration and provocation on the limits of the approach. Social Science Computer Review. Advance Online Publication, 2024.
62. Dosunmu AA, Ogundele PO. Intrusion detection and prevention models for enhancing organizational cyber defense effectiveness. Iconic Research and Engineering Journals. 2020; 4(6):310-324. Doi: https://doi.org/10.64388/IREV4I6-1713226
63. Dosunmu AA, Ogundele PO. Threat intelligence integration frameworks supporting proactive enterprise cybersecurity decision making. Gyanshauryam, International Scientific Refereed Research Journal. 2022; 5(3):397-416.
64. Dosunmu AA, Ogundele PO. Cyber threat actor analysis models for strategic enterprise security planning. Shodhshauryam, International Scientific Refereed Research Journal. 2023; 6(5):513-531.
65. Dosunmu AA, Ogundele PO. Breach and attack simulation frameworks for continuous validation of enterprise security controls. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2024a; 10(3):1100-1119.
66. Dosunmu AA, Ogundele PO. Cyber risk quantification models for prioritizing enterprise security investment decisions. International Journal of Multidisciplinary Research and Growth Evaluation. 2024b; 5(6):1777-1785. Doi: https://doi.org/10.54660/.IJMRGE.2024.5.6.1777-1785
67. Dosunmu AA, Ogundele PO. Threat informed defence engineering models for measuring security control effectiveness at scale. International Journal of Advanced Multidisciplinary Research and Studies. 2024c; 4(6):2847-2858.
68. Eboh EE, Aliliele C. AI-driven data analytics framework for risk assessment and detection of venture capital and private equity investment fraud in U.S. capital markets. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2024; 10(6):2624-2665. Doi: https://doi.org/10.32628/CSEIT2410787
69. Eden C, Ackermann F. Making strategy: The journey of strategic management. Sage, 1998.
70. Efobi OZ, Akinleye OK, Fasawe O. Conceptual model

for data-driven lean supply chain optimization in manufacturing and retail operations. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2022; 8(2):703-734. Doi: https://doi.org/10.32628/CSEIT2541335
71. Efobi OZ, Akinleye OK, Fasawe O. Conceptual framework for developing a resilience index for post-pandemic supply chains. Shodhshauryam, International Scientific Refereed Research Journal. 2023; 6(2):421-432.
72. Ekeocha AH, Aganga AA, Adejoro FA, Oyebanji A, Oluwadele JF, Tawose OM. Phenotypic characteristics of indigenous chickens in selected regions of Nigeria. Journal of World's Poultry Research. 2021; 11(3):352-358. Doi: https://doi.org/10.36380/jwpr.2021.42
73. Ekun K, Muriel A. Online lead time quotation under contingency. In A. Brown Greer, C. Contardo, & J.-M. Frayret (Eds.), Proceedings of the IISE Annual Conference & Expo 2024, 2024. Doi: https://doi.org/10.21872/2024IISE_6090
74. Elebe O, Hammed NI, Omoegun GO, Fadayomi O, Bello AD. An advanced machine learning model for detecting synthetic identity fraud in e-commerce platforms \[Preprint\]. ResearchGate, 2023.
75. Elmachtoub AN, Grigas P. Smart predict, then optimize. Management Science. 2022; 68(1):9-26.
76. Eyetsemitan RA, Ambali KB, Oyeleye AO, Fadayomi O. Multi-stakeholder governance alignment in joint venture operations: A conceptual framework for coordinating business processes in highly regulated environments. Iconic Research and Engineering Journals. 2020; 4(4):418-441. Doi: https://doi.org/10.64388/IREV4I4-1716955
77. Eyetsemitan RA, Oyeleye AO, Ambali KB, Fadayomi O. CRM and workflow automation in small healthcare practices: A process efficiency framework for scalable patient engagement. International Journal of Multidisciplinary Research and Growth Evaluation. 2024a; 5(6):1931-1949. Doi: https://doi.org/10.54660/IJMRGE.2024.5.6.1931-1949
78. Eyetsemitan RA, Oyeleye AO, Ambali KB, Fadayomi O. Data-driven process optimization in micro-enterprises: A conceptual framework for funnel analysis and bottleneck identification. International Journal of Multidisciplinary Research and Growth Evaluation. 2024b; 5(6):1950-1968. Doi: https://doi.org/10.54660/IJMRGE.2024.5.6.1950-1968
79. Eze FI, Anene UN, Akinleye OK. Creation of a drug shortage early warning and regulatory response model for essential medicines in global health systems. Research Journal of Pure Science and Technology. 2022a; 5(1):57-95. Doi: https://doi.org/10.56201/rjpst.v5.no1.2022.pg57.95
80. Eze FI, Anene UN, Akinleye OK. Design of an integrated regulatory readiness maturity model for pharmaceutical companies seeking global market entry. International Journal of Health and Pharmaceutical Research. 2022b; 7(1):77-112. Doi: https://doi.org/10.56201/ijhpr.v7.no1.2022.pg77.112
81. Fadayomi O, Bello AD, Elebe O, Hammed NI, Omoegun GO. An adaptive fraud risk scoring model for real-time transaction monitoring at scale. IIARD International Journal of Banking and Finance Research. 2024; 10(10):212-230.
82. Falemi A, Akhigbe R, Akin-Oluyomi OT. A supply chain workforce optimization framework for addressing staffing volatility during rapid expansion cycles. International Journal of Advanced Multidisciplinary Research and Studies. 2023; 3(1):1745-1762. Doi: https://doi.org/10.62225/2583049X.2023.3.1.5342
83. Fan Z, Ghaddar B, Wang X, Xing L, Zhang Y, Zhou Z. Artificial intelligence for operations research: Revolutionizing the operations research process. arXiv, 2024. https://arxiv.org/abs/2401.03244
84. Franco LA, Greiffenhagen C. Making OR practice visible: Using ethnomethodology to analyse facilitated modelling workshops. European Journal of Operational Research. 2018; 265(2):673-684.
85. Horton JJ. Large language models as simulated economic agents: What can we learn from homo silicus? (NBER Working Paper No. 31122). National Bureau of Economic Research, 2023.
86. Idika CN, Salami EO, Ijiga OM, Enyejo LA. Deep learning driven malware classification for cloud-native microservices in edge computing architectures. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2021; 7(4).
87. Ijiga OM, Awoyemi O, Atobatele FA, Okonkwo CA. Integrating educational technology: Policies and practices for inclusive learning. International Journal of Scientific Research in Science, Engineering and Technology. 2024; 11(5):408-420.
88. Ike PN, Ogbuefi E, Nnabueze SB, Olatunde-Thorpe J, Aifuwa SE, Oshoba TO, *et al*. Supplier relationship management strategies fostering innovation, collaboration, and resilience in global supply chain ecosystems. International Journal of Multidisciplinary Evolutionary Research. 2021; 2(2):52-62. Doi: https://doi.org/10.54660/IJMER.2021.2.2.52-62
89. Ike PN, Aifuwa SE, Nnabueze SB, Olatunde-Thorpe J, Ogbuefi E, Oshoba TO, Akokodaripon D. Quantitative risk architecture for public-private partnerships: A multi-layered model for allocating public and private risk. International Journal of Advanced Multidisciplinary Research and Studies. 2024; 4(6):2669-2682. Doi: https://doi.org/10.62225/2583049X.2024.4.6.5021
90. Ilesanmi MO, Okoh OF, Balogun SA, Ijiga OM. Data-driven portfolio optimization for utility-scale solar, wind, and battery energy storage systems (BESS): Integrating performance analytics with investor EBITDA targets. International Journal of Scientific Research and Modern Technology. 2024; 3(11):141-155.
91. Ilodigwe L, Adesemoye AC. The data backbone of health system transformation: A governance and architecture framework for interoperability, data quality, and advanced analytics at national scale. International Journal of Health and Pharmaceutical Research. 2021; 6(2):52-76. Doi: https://doi.org/10.56201/ijhpr.vol.6.no2.2021.pg52.76
92. Ilodigwe L, Adesemoye AC. Advances in real-world evidence, predictive analytics, and data-driven decision making for improving health system performance and

patient outcomes. International Journal of Health and Pharmaceutical Research. 2024; 9(5):174-197. Doi: https://doi.org/10.56201/ijhpr.v9.no5.2024.pg174.197
93. Isiekwu CP, Oluwo K, Dada T. A conceptual model for CFO-led strategic finance in joint venture and cross-border partnerships. Iconic Research and Engineering Journals. 2021; 4(7):383-400. Doi: https://doi.org/10.64388/IREV4I7-1714350
94. Jiang C, Shu X, Qian H, Lu X, Zhou J, Zhou A, *et al*. LLMOPT: Learning to define and solve general optimization problems from scratch. arXiv, 2024. https://arxiv.org/abs/2410.13213
95. Jones BC, Ominyi M. A review of legal barriers to import-export participation for Nigerian entrepreneurs. IIARD International Journal of Economics and Business Management. 2020; 6(4):51-70. Doi: https://doi.org/10.56201/ijebm.vol.6.no4.2020.pg51.70
96. Jones BC, Ominyi M. A systematic review of market access outcomes from free trade agreements and duty preference programs for locally made African goods. IIARD International Journal of Economics and Business Management. 2022; 8(6):67-95. Doi: https://doi.org/10.56201/ijebm.v8.no6.2022.pg67.95
97. Jones BC, Ominyi M. Reviewing compliance frameworks for technology transfers between developed and emerging markets in an era of export controls and sanctions. World Journal of Innovation and Modern Technology. 2023; 7(2):121-149. Doi: https://doi.org/10.56201/wjimt.v7.no2.2023.pg121.149
98. Jones BC, Ominyi M. Applications of artificial intelligence in customs administration and trade facilitation: A review. World Journal of Innovation and Modern Technology. 2024a; 8(6):181-217. Doi: https://doi.org/10.56201/wjimt.v8.no6.2024.pg181.217
99. Jones BC, Ominyi M. A conceptual framework for harmonizing digital trade regulations in cross-border e-commerce between the United States and Nigeria. International Journal of Economics and Financial Management. 2024b; 9(9):337-364. Doi: https://doi.org/10.56201/ijefm.v9.no9.2024.pg337.364
100. Kim J, Lee B. AI-augmented surveys: Leveraging large language models and surveys for opinion prediction. arXiv, 2024. https://arxiv.org/abs/2305.09620
101. Komi NM. When accurate forecasts are not fair forecasts: Deep learning load prediction and equity across districts. Shodhshauryam, International Scientific Refereed Research Journal. 2023; 6(6):476-546. Doi: https://doi.org/10.32628/SHISRRJ247102
102. Komi NM. Turning social license into a measurable variable: Predicting community trust in renewable energy development. International Journal of Engineering and Modern Technology. 2024a; 10(11):239-296. Doi: https://doi.org/10.56201/ijemt.v10.no11.2024.pg239.296
103. Komi NM, Adamolekun A. Interpretable machine learning for early failure prediction in distributed renewable energy assets. International Journal of Multidisciplinary Research and Growth Evaluation. 2021b; 2(6):1015-1038. Doi: https://doi.org/10.54660/.IJMRGE.2021.2.6.1015-1038
104. Komi NM, Adamolekun A. Quantifying a just transition: An econometric and skills-mapping analysis of worker displacement in the coal-to-renewable shift. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2024b; 10(3):1242-1299. Doi: https://doi.org/10.32628/CSEIT25113587
105. Komi NM, Adeniji IO. No single instrument reaches the poor: A systematic review of financing mechanisms for inclusive renewable energy. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2024c; 10(3):1185-1241. Doi: https://doi.org/10.32628/CSEIT25113586
106. Komi NM, Adeniji IO. Why energy projects survive or fail: Modeling the causal link between community trust and infrastructure durability. International Journal of Advanced Multidisciplinary Research and Studies. 2024d; 4(6):3263-3296. Doi: https://doi.org/10.62225/2583049X.2024.4.6.6459
107. Krippendorff K. Content analysis: An introduction to its methodology (4th ed.). Sage, 2019.
108. Ladapo OO, Dosunmu AA, Jooda D, Abolaji TO. Human-in-the-loop machine learning: A state of the art. Journal of Frontiers in Multidisciplinary Research. 2022; 3(1):656-669. Doi: https://doi.org/10.54660/.JFMR.2022.3.1.656-669
109. Ladapo OO, Dosunmu AA, Jooda D, Abolaji TO. An IT service management literature review: Challenges, benefits, opportunities, and implementation practices. International Journal of Advanced Multidisciplinary Research and Studies. 2023; 3(6):2875-2889. Doi: https://doi.org/10.62225/2583049X.2023.3.6.6205
110. Ladapo OO, Jooda D, Dosunmu AA, Abolaji TO. Keeping humans in the loop: Human-centered automated annotation with generative AI. International Journal of Multidisciplinary Futuristic Development. 2024; 5(1):81-95. Doi: https://doi.org/10.54660/IJMFD.2024.5.1.81-95
111. Lawal OA, Oduleye TE. A review of decision analytics models for sustainable profitability in technology firms. Gyanshauryam, International Scientific Refereed Research Journal. 2023; 6(6):455-475.
112. Liadi KO. A policy alignment model for Nigeria's foreign policy and global climate diplomacy goals. International Journal of Multidisciplinary Research and Growth Evaluation. 2022; 3(6):790-801. Doi: https://doi.org/10.54660/.IJMRGE.2022.3.6.790-801
113. Liadi KO. A model for linking Nigeria's diplomatic engagement to sustainable development outcomes in ECOWAS states. Gyanshauryam, International Scientific Refereed Research Journal. 2023; 6(3):392-420.
114. Liadi KO. Designing a cross-border security cooperation model: Nigeria's foreign policy response to regional terrorism. Gyanshauryam, International Scientific Refereed Research Journal. 2024b; 7(3):311-338. Doi: https://doi.org/10.32628/GISRRJ247324
115. Marttunen M, Lienert J, Belton V. Structuring problems for multi-criteria decision analysis in practice: A literature review of method combinations. European Journal of Operational Research. 2017; 263(1):1-17.
116. Mathis WS, Zhao S, Pratt N, Weleff J, De Paoli S. Inductive thematic analysis of healthcare qualitative interviews using open-source large language models: How does it compare to traditional methods? Computer Methods and Programs in Biomedicine. 2024; 255, Article 108356.

117. Mbonu IS, Aliliele C, Iwuanyanwu U, Uzoka E. A review of identity and access management integration strategies in hybrid and multi cloud environments. International Journal of Multidisciplinary Research and Growth Evaluation. 2020; 1(5):795-810. Doi: https://doi.org/10.54660/.IJMRGE.2020.1.5.795-810
118. Mbonu IS, Aliliele C, Iwuanyanwu U, Uzoka E. A conceptual framework for risk based business intelligence architecture in financial technology platforms. International Journal of Multidisciplinary Research and Growth Evaluation. 2021; 2(6):731-746. Doi: https://doi.org/10.54660/.IJMRGE.2021.2.6.731-746
119. Mbonu IS, Aliliele C, Iwuanyanwu U, Uzoka E. A conceptual framework for AI enabled IT general controls and SOX audit automation processes. Gyanshauryam, International Scientific Refereed Research Journal. 2022a; 5(5):384-414. Doi: https://doi.org/10.32628/GISRRJ2256239
120. Mbonu IS, Iwuanyanwu U, Aliliele C, Uzoka E. A review of data protection impact assessment models in multi cloud financial infrastructure systems. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2022b; 8(1):589-623. Doi: https://doi.org/10.32628/CSEIT25442
121. Medon JJ, Oduleye TE. A comprehensive financial reporting model for strengthening compliance and organizational accountability systems. International Journal of Multidisciplinary Research and Growth Evaluation. 2022; 3(6):768-777.
122. Michael ON, Ogunsola OE. Investigating the contribution of climate-smart agricultural practices to household food security and income stability. International Journal of Multidisciplinary Research and Growth Evaluation. 2021b; 1(2):209-219. Doi: https://doi.org/10.54660/IJMRGE.2020.1.2.209-219
123. Michael ON, Ogunsola OE. Examining the socioeconomic barriers to technological adoption among smallholder farmers in remote rural areas. Shodhshauryam, International Scientific Refereed Research Journal. 2022; 5(6):484-519.
124. Mostajabdaveh M, Yu TT, Ramamonjison R, Carenini G, Zhou Z, Zhang Y. Optimization modeling and verification from problem specifications using a multi-agent multi-stage LLM framework. INFOR: Information Systems and Operational Research. 2024; 62(4):599-617.
125. Nnabueze SB, Ike PN, Olatunde-Thorpe J, Aifuwa SE, Oshoba TO, Ogbuefi E, *et al*. End-to-end visibility frameworks improving transparency, compliance, and traceability across complex global supply chain operations. International Journal of Multidisciplinary Futuristic Development. 2021; 2(2):50-60. Doi: https://doi.org/10.54660/IJMFD.2021.2.2.50-60
126. Nnaji N, Akinlolu VS. Addressing productivity and workflow inefficiencies using analytics-driven performance models in healthcare operations. Shodhshauryam, International Scientific Refereed Research Journal, 2023.
127. Nwakamma S, Ojukwu J, Oyesiji SO. LLM-as-a-judge for automated model governance: A review of methods, biases, and release-gating practices. World Journal of Innovation and Modern Technology. 2024a; 8(6):185-216. Doi: https://doi.org/10.56201/wjimt.v8.no6.2024.pg185.216
128. Nwakamma S, Ojukwu J, Oyesiji SO. Securing agentic AI enterprise workflows against prompt injection, tool poisoning, memory manipulation, and excessive agency threats. World Journal of Innovation and Modern Technology. 2024b; 8(6):217-248. Doi: https://doi.org/10.56201/wjimt.v8.no6.2024.pg217.248
129. Obogo SF, Arumosoye OM, Obriki OD. Advances in proactive hazard recognition and near miss reporting systems. International Journal of Multidisciplinary Research and Growth Evaluation. 2021; 2(6):835-846. Doi: https://doi.org/10.54660/IJMRGE.2021.2.6.835-846
130. Obriki OD, Arumosoye OM. Conceptual framework for human error causation in high-risk construction and industrial activities. International Journal of Multidisciplinary Research and Growth Evaluation. 2020; 1(5):715-727. Doi: https://doi.org/10.54660/.IJMRGE.2020.1.5.715-727
131. Obriki OD, Arumosoye OM. Conceptual framework explaining recurrence mechanisms of unsafe behaviors in high-hazard worksites. Shodhshauryam, International Scientific Refereed Research Journal. 2022; 5(3):380-402. Doi: https://doi.org/10.32628/SHISRRJ225489
132. Obriki OD, Arumosoye OM. Systematic review of incident investigation approaches and prevention-oriented learning in industrial operations. Shodhshauryam, International Scientific Refereed Research Journal. 2024; 7(4):239-263. Doi: https://doi.org/10.32628/SHISRRJ247163
133. Obriki OD, Obogo SF, Arumosoye OM. Review of behavioral safety programs for risk reduction in large workforces. International Journal of Advanced Multidisciplinary Research and Studies. 2023a; 3(6):2760-2775. Doi: https://doi.org/10.62225/2583049X.2023.3.6.6076
134. Odejide GS. Artificial intelligence in multilingual and accessible emergency alerting: A scoping review of evidence, applications, and gaps in the United States. International Journal of Scientific Research in Science, Engineering and Technology. 2023; 10(6):499-512. Doi: https://doi.org/10.32628/IJSRSET26132152
135. Odejide GS. A trust-centered conceptual model for AI-driven multilingual emergency communication. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2024; 10(6):2781-2811. Doi: https://doi.org/10.32628/CSEIT2410794
136. Oduleye TE, Medon JJ. A data-driven cost management model for improving strategic financial planning and performance evaluation. International Journal of Multidisciplinary Research and Growth Evaluation. 2021; 2(6):524-537.
137. Oduleye TE, Medon JJ. A quantitative model for measuring the strategic impact of financial analysis on enterprise growth. International Journal of Advanced Multidisciplinary Research and Studies. 2023; 3(1):1663-1672. Doi: https://doi.org/10.62225/2583049X.2023.3.1.5334
138. Ogbete JC, Aminu-Ibrahim AY, Ambali KB. Risk managed construction strategies for complex and highly regulated healthcare facility developments. International Journal of Scientific Research in Computer Science,

Engineering and Information Technology. 2021; 7(5):362-392. Doi: https://doi.org/10.32628/CSEIT217561
139. Ogbete JC, Aminu-Ibrahim AY, Ambali KB. Lifecycle performance evaluation of purpose built diagnostic laboratories supporting long term healthcare delivery. International Journal of Advanced Multidisciplinary Research and Studies. 2023; 3(6):2605-2621. Doi: https://doi.org/10.62225/2583049X.2023.3.6.5698
140. Ogbole JI, Okoruwa PO, Fadayomi O, Abolaji TO, Edivri J, Akeju B. Conceptual model for identity-centric zero trust architecture in enterprise security governance. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2021; 7(5):393-415. Doi: https://doi.org/10.32628/IJSRCSEIT217562
141. Oghenemaiga E, Basnet A, Anene UN. Predictive analytics applications for forecasting traffic patterns across owned and operated media platforms. International Journal of Advanced Multidisciplinary Research and Studies. 2024; 4(6). Doi: https://doi.org/10.62225/2583049X.2024.4.6.5668
142. Ogundairo KM, Ayivi-Donkor SS. Machine learning approaches to customer acquisition and retention: A framework for predictive lifecycle management. IIARD International Journal of Economics and Business Management. 2023; 9(10):210-246. Doi: https://doi.org/10.56201/ijebm.v9.no10.2023.pg210.246
143. Okafor CM, Osuji VC, Dako OF. Harmonizing risk governance, technology infrastructure, and compliance frameworks for future-ready banking systems. International Journal of Scientific Research in Humanities and Social Sciences. 2024b; 1(1):316-337.
144. Okojie J, Ike P, Idu J, Nnabueze SB, Filani O, Ihwughwavwe S. Predictive analytics models for monitoring smart city emissions and infrastructure risk in urban ESG planning. International Journal of Multidisciplinary Futuristic Development. 2023a; 4(1):45-57. Doi: https://doi.org/10.54660/IJMFD.2023.4.1.45-57
145. Okojie JS, Filani OM, Ike PN, Okojokwu-Idu JO, Nnabueze SB, Ihwughwavwe SI. Integrating AI with ESG metrics in smart infrastructure auditing for high-impact urban development projects. International Journal of Multidisciplinary Futuristic Development. 2023b; 4(1):32-44. Doi: https://doi.org/10.54660/IJMFD.2023.4.1.32-44
146. Okonkwo CS, Agbabiaka J, Ogunwole O, Mayo W, Okeke OT. Model for demurrage elimination and port logistics efficiency in emerging economies. International Journal of Multidisciplinary Research and Growth Evaluation. 2020; 1(5):552-562. Doi: https://doi.org/10.54660/IJMRGE.2020.1.5.552-562
147. Okonkwo CS, Agbabiaka J, Mayo W, Okeke OT. Conceptual framework for digital supply chain governance in energy and infrastructure sectors. Gyanshauryam, International Scientific Refereed Research Journal. 2024a; 7(4):335-356. Doi: https://doi.org/10.32628/GISRRJ247423
148. Okonkwo CS, Agbabiaka J, Mayo W, Okeke OT. Model for predictive procurement planning to sustain operational uptime. International Journal of Scientific Research in Humanities and Social Sciences. 2024b; 1(2):909-928. Doi: https://doi.org/10.32628/IJSRSSH242772
149. Olorunkosebi MT, Ayinde MK, Mensah GK. Integrating artificial intelligence and advanced biosensors for rapid and sensitive detection of antimicrobial resistant foodborne pathogens. Journal of Advances in Microbiology Research. 2023; 4(2):175-182. Doi: https://doi.org/10.22271/micro.2023.v4.i2b.293
150. Oloto AM, Yeboah TJ. A systematic review of behavior intervention plan effectiveness in secondary special education programs. Gyanshauryam, International Scientific Refereed Research Journal. 2023; 6(6):385.
151. Oluwadele JF, Tawose OM, Ekeocha AH, Akinlabi EY, Adeitan OO, Bello KA. Physiological indicators and stress index of scavenging chickens at LAFARGE (Ewekoro) and DANGOTE (Ibese) cement factory areas of Ogun State, Nigeria. Journal of Austrian Society of Agricultural Economics. 2023; 19(4). https://www.sagepublisher.com/volume/JASAE/19/04
152. Oluwo K, Dada T, Isiekwu CP. Transforming insurance underwriting with machine learning: A review and application cases. International Journal of Advanced Multidisciplinary Research and Studies. 2024; 4(6):3008-3019.
153. Omaghomi TT, Akinlolu VS, Fapohunda M. A governance-driven framework for improving workforce planning and nurse-patient ratio optimization in public hospitals. Multidisciplinary Journal of Educational Research. 2023; 4(6):1357-1368.
154. Ominyi M, Anichukwueze CC. A review of algorithmic accountability and model risk management approaches in financial services. Journal of Accounting and Financial Management. 2023; 9(12):218-238. Doi: https://doi.org/10.56201/jafm.v9.no12.2023.pg218.238
155. Ominyi M, Anichukwueze CC. Conceptualizing responsible AI governance structures for automated decision systems in regulated industries. International Journal of Social Sciences and Management Research. 2024a; 10(11):403-430. Doi: https://doi.org/10.56201/ijssmr.v10.no11.2024.pg.403.430
156. Ominyi M, Anichukwueze CC, Uzougbo NS. A systematic review of organizational readiness for the EU Artificial Intelligence Act in multinational enterprises. World Journal of Innovation and Modern Technology. 2024b; 8(6):185-203. Doi: https://doi.org/10.56201/wjimt.v8.no6.2024.pg185.203
157. Onche VO, Adenekan ET, Sabboh GM. Influence of personal and professional attributes of secretaries on organisational success of universities in Oyo State, Nigeria. International Journal of Science Academic Research. 2021b; 2(4):1355-1359.
158. Onche VO, Ogbonna CS, Adegbite MP. Effectiveness of cybersecurity awareness training on insider-threat behavior among university administrators: An integrative academic review. International Journal of Computer Science and Mathematical Theory. 2024; 10(2):117-144. Doi: https://doi.org/10.56201/ijcsmt.v10.no2.2024.pg117.144
159. Oparah OS, Ezeh FE, Olatunji GI, Ajayi OO. Framework for designing national real-time disease surveillance dashboards for public health stakeholders. Shodhshauryam, International Scientific Refereed

Research Journal. 2023; 6(1):208-227.
160. Orise C, Niniola S. Real-time analytics for educational excellence: A conceptual framework for performance dashboards in health professional training. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2022; 8(2):802-815.
161. Oyeleye AO, Eyetsemitan RA, Ambali KB, Fadayomi O. Reducing client onboarding cycle time in small professional services firms: A Lean Six Sigma process redesign framework. Gyanshauryam, International Scientific Refereed Research Journal. 2022; 5(2):467-498. Doi: https://doi.org/10.32628/GISRRJ225357
162. Oyesiji SO, Nwakamma S, Ojukwu J. Parameter-efficient multilingual adaptation of on-device language models: A survey. International Journal of Engineering and Modern Technology. 2023; 9(3):287-320. Doi: https://doi.org/10.56201/ijcsmt.v9.no3.2023.pg287.320
163. Oyesiji SO, Nwakamma S, Ojukwu J. A conceptual framework for context-aware data curation in fine-tuning pipelines. International Journal of Engineering and Modern Technology. 2024; 10(11):197-228. Doi: https://doi.org/10.56201/ijemt.v10.no11.2024.pg197.228
164. Oziri ST, Arowogbadamu AAG, Seyi-Lande OB. Predictive analytics applications in reducing customer churn and enhancing lifecycle value in telecommunications markets. International Journal of Multidisciplinary Futuristic Development. 2020; 1(2):40-49.
165. Park JS, Zou CQ, Shaw A, Hill BM, Cai C, Morris MR, *et al*. Generative agent simulations of 1,000 people. arXiv, 2024. https://arxiv.org/abs/2411.10109
166. Ramamonjison R, Yu TT, Li R, Li H, Carenini G, Ghaddar B, *et al*. NL4Opt competition: Formulating optimization problems based on their natural language descriptions. In Proceedings of the NeurIPS 2022 Competition Track (Proceedings of Machine Learning Research, Vol. 220, pp. 189-203). PMLR, 2023.
167. Rosenhead J, Mingers J. (Eds.). Rational analysis for a problematic world revisited. Wiley, 2001.
168. Rukh S, Oziri ST, Seyi-Lande OB. Framework for enhancing marketing strategy through predictive and prescriptive analytics. Shodhshauryam, International Scientific Refereed Research Journal. 2023; 6(4):531-569.
169. Rukh S, Seyi-Lande OB, Oziri ST. An integrated framework for AI and predictive analytics in supply chain management. International Journal of Scientific Research in Humanities and Social Sciences. 2024; 1(1):451-491.
170. Sanni JO, Atima ME, Attah A. Systematic review of attribution modeling methods resolving bias in multi-touch journeys. Shodhshauryam, International Scientific Refereed Research Journal. 2022; 5(1):304-334. Doi: https://doi.org/10.32628/SHISRRJ247135
171. Santurkar S, Durmus E, Ladhak F, Lee C, Liang P, Hashimoto T. Whose opinions do language models reflect? In Proceedings of the 40th International Conference on Machine Learning. PMLR, 2023.
172. Sarstedt M, Adler SJ, Rau L, Schmitt B. Using large language models to generate silicon samples in consumer and marketing research: Challenges, opportunities, and guidelines. Psychology & Marketing, 2024.
173. Seun A, Ajayi A, Elumilade RA, Bobga MA, Yeboah TJ, Ofori SD, *et al*. PearlAI: Hyperpersonalization and hyperlocalization as new AI affordances in basic education. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2024; 10(6):2553-2568. Doi: https://doi.org/10.32628/CSEIT25113574
174. Shittu H, Shittu MA, Adeleke OJ, Adedokun OJ. Blockchain based energy trading models for peer-to-peer renewable microgrids. International Journal of Research in Electrical and Electronics Engineering. 2021; 4(3):1-19. Doi: https://doi.org/10.34218/IJREEE_04_03_001
175. Shittu MA, Adeniji IO, Shittu H, Opara IS. Blockchain-assisted secure data exchange architectures for SCADA-controlled power systems. Iconic Research and Engineering Journals. 2022; 6(3):359-379. Doi: https://doi.org/10.64388/IREV6I3-1714041
176. Shittu MA, Shittu HA, Adeleke OJ, Adedokun OJ. Digital twin modeling for real time monitoring and fault detection in smart substations. International Journal of Industrial Engineering Research and Development. 2023; 14(2):25-44. Doi: https://doi.org/10.34218/IJIERD_14_02_003
177. Sun S, Lee E, Nan D, Zhao X, Lee W, Jansen BJ, *et al*. Random silicon sampling: Simulating human sub-population opinion using a large language model based on group-level demographic information. arXiv, 2024. https://arxiv.org/abs/2402.18144
178. Sunday EA, Omoegun GO. Smart fault detection in HVAC systems using sensor-based monitoring. International Journal of Scientific Research in Science, Engineering and Technology. 2022; 7(4):380-403.
179. Tang Z, Huang C, Zheng X, Hu S, Wang Z, Ge D, *et al*. ORLM: Training large language models for optimization modeling. arXiv, 2024. https://arxiv.org/abs/2405.17743
180. Tawose OM, Ekeocha AH, Oluwadele JF. Nutritional quality and utilization of water hyacinth-cassava peels silage by West African dwarf (WAD) goats. FUOYE Journal of Agriculture and Human Ecology. 2023a; 6(1):35-45.
181. Tawose OM, Oluwadele JF, Ekeocha AH, Odumboni AA, Egbeyemi OS. Aflatoxin detection and quantification in poultry feeds available in selected areas in Ekiti State. FUOYE Journal of Agriculture and Human Ecology. 2023b; 7(2). Doi: https://doi.org/10.62923/fuojahe.v7i2.302
182. Tonoyan A, Dada O, Ayivi-Donkor SS. Advances in demand forecasting: Machine learning algorithms for revenue projection and financial planning accuracy. Gyanshauryam, International Scientific Refereed Research Journal. 2021; 4(1):282-317.
183. Tonoyan A, Dada O, Ayivi-Donkor SS. Conceptual model for predictive cost optimization: Machine learning applications in business transformation analysis. International Journal of Economics and Business Management. 2022b; 8(6):68-102. Doi: https://doi.org/10.56201/ijebm.v8.no6.2022.pg68.102
184. Tonoyan A, Dada O, Ayivi-Donkor SS. Advances in supply chain resilience: Predictive models for vendor

risk assessment and procurement cost optimization. International Journal of Social Sciences and Management Research. 2024a; 10(11):525-551. Doi: https://doi.org/10.56201/ijssmr.v10.no11.2024.pg.525.551
185. Tonoyan A, Dada O, Ayivi-Donkor SS. Real-time KPI tracking systems: A review of automated performance monitoring and data-driven decision making. World Journal of Innovation and Modern Technology. 2024b; 8(6):247-281. Doi: https://doi.org/10.56201/wjimt.v8.no6.2024.pg247.281
186. Uduokhai DO, Garba BMP, Nwafor MI, Sanusi AN. Modeling user experience and post-occupancy satisfaction in government-sponsored housing projects. Gyanshauryam, International Scientific Refereed Research Journal. 2023a; 6(2):479-497.
187. Uduokhai DO, Giloid S, Nwafor MI, Adio SA. Evaluating the role of building information modeling in enhancing project performance in Nigeria. International Journal of Advanced Multidisciplinary Research and Studies. 2023b; 3(6):2154-2161.
188. Uduokhai DO, Nwafor MI, Sanusi AN, Patrick BM. Critical review of housing policy implementation strategies in Sub-Saharan African urban economies. Shodhshauryam, International Scientific Refereed Research Journal, 2023c. Doi: https://doi.org/10.32628/SHISRRJ236927
189. Uduokhai DO, Nwafor MI, Sanusi AN, Garba BMP. Applying design thinking approaches to architectural education and innovation in Nigerian universities. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2023d; 9(4):852-870.
190. Uduokhai DO, Nwafor MI, Sanusi AN, Garba BMP. System dynamics modeling of circular economy integration within the African construction industry. International Journal of Scientific Research in Humanities and Social Sciences. 2024a; 1(2):871-887.
191. Uduokhai DO, Garba BMP, Nwafor MI, Sanusi AN. Techno-economic evaluation of renewable-material construction for low-income housing communities. International Journal of Scientific Research in Humanities and Social Sciences. 2024b; 1(2):888-908.
192. Wedraogo L, Sanni JO. Machine learning models addressing uncertainty in cross-channel campaign performance forecasting accuracy. International Journal of Advanced Multidisciplinary Research and Studies. 2024; 4(6):2875-2890. Doi: https://doi.org/10.62225/2583049X.2024.4.6.5648
193. Wuttke A, Assenmacher M, Klamm C, Lang M, Wurschinger Q, Kreuter F. AI conversational interviewing: Transforming surveys with LLMs as adaptive interviewers. arXiv, 2024. https://arxiv.org/abs/2410.01824
194. Xiao Z, Zhang D, Wu Y, Xu L, Wang YJ, Han X, *et al*. Chain-of-Experts: When LLMs meet complex operations research problems. In Proceedings of the Twelfth International Conference on Learning Representations, 2024.
195. Xing L, Wang X, Feng Y, Fan Z, Xiong J, Guo Z, *et al*. Towards human-aligned evaluation for linear programming word problems. In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation. ELRA and ICCL, 2024, 16550-16556.
196. Yeboah TJ, Oloto AM. Advances in data driven IEP development: Conceptual models for supporting diverse learners. Shodhshauryam, International Scientific Refereed Research Journal. 2022b; 5(1):335-365.
197. Yeboah TJ, Oloto AM. Conceptual models for monitoring IEP implementation fidelity and educational service delivery outcomes. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2023; 9(3):944-973.
198. Yeboah TJ, Bobga MA, Boakye K, Ogbona CS. Data-driven teaching: Using student progress data to personalize learning in special education classrooms. Journal of Frontiers in Multidisciplinary Research. 2022a; 3(2):225-240. Doi: https://doi.org/10.54660/.JFMR.2022.3.2.225-240
199. Yu J, Zhao J, Miranda-Moreno L, Korp M. Modular conversational agents for surveys and interviews. arXiv, 2024.
200. Zenimoto Y, Hasegawa R, Utsuro T, Yoshioka M, Kando N. Coding open-ended responses using pseudo response generation by large language models. In Proceedings of the 2024 NAACL Student Research Workshop. Association for Computational Linguistics, 2024, 242-254.