AI for Science Strategic Compass: Aligning Discovery Tensions with Core AI Functions

Ran Liu
UniBridgeAI
sophie.liu@unibridgeai.com

Zhibin Lin
Durham University
zhibin.lin@durham.ac.uk

Xiaowei Huang
University of Liverpool
xiaowei.huang@liverpool.ac.uk

Abstract

AI is transforming scientific discovery, yet researchers face a fragmented, fast-moving field of AI that lacks stable, strategy-level guidance for method selection and integration. In this study, we introduce the AI for Science Strategic Compass (AFSC), a compact decision framework that aligns four cross-domain scientific-discovery tensions (Complexity, Constraint, Scarcity, Explosion) with six core AI functions (Represent; Reason & Infer; Optimize & Control; Simulate & Emulate; Generate & Create; Autonomize & Orchestrate) via a 6×4 Strategy Matrix. We adopt a function-based typology that is domain-agnostic and comparatively stable under ongoing methodological change, enabling direct alignment with these tensions and yielding decision-relevant guidance. Each cell is labeled with a keyword that captures the shared mitigation logic and lists three strategic pathways linked to representative method families. Pathways are anchored to a function-internal atomic triad, stabilizing the vocabulary as techniques change. Automated corpus audits validate the framework's scope: the four tensions collectively cover all sampled abstracts across six natural science domains, and the six functions account for 98.9% of capabilities reported in recent AI papers. AFSC shifts selection from tool-driven browsing to strategy-first planning, lowering cognitive load and remaining portable across domains. We illustrate its use with an exoplanet spectral retrieval case study that demonstrates systematic integration of complementary AI approaches across functions to address multiple research tensions.

1 Introduction

Artificial intelligence is reshaping scientific research by accelerating discovery, extracting structure from complex data, and extending the frontier of testable hypotheses and designs [Wang et al., 2023a, Boiko et al., 2023a, Reddy and Shojaee, 2025, Canty et al., 2025, Rapp et al., 2024]. However, the AI knowledge base evolves faster than disciplinary curricula, terminology is fragmented across subfields, and many laboratories, particularly those without formal AI training, lack a strategy-level guide that links specific scientific problems to appropriate AI capabilities.

Generic AI surveys synthesize broad method families by learning paradigm, modality, or architecture and have established a shared vocabulary for the field [Gui et al., 2024, Zha et al., 2025, Xu et al., 2023]. Yet their AI-centric vantage point is often either too abstract to inform concrete choices in a laboratory or so technical that it raises cognitive load rather than lowering it. Domain-specific reviews translate techniques into a single scientific context and improve local relevance [Ma et al., 2024, Hasselgren and Oprea, 2024, Smith and Geach, 2023], but they narrow methodological coverage and embed assumptions about data, resources, and metrics that hinder transfer across fields. Procedural frameworks and evaluation methodologies add rigor through phases, roles, and metrics (e.g.,

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: The Reach and Limits of AI for Scientific Discovery.

[Tekinerdogan, 2024, Cappello et al., 2025]), yet they typically presuppose specialized infrastructure and address bounded scenarios, offering little cross-domain strategic guidance. Autonomous and closed-loop systems demonstrate impressive end-to-end capability and throughput [Szymanski et al., 2023a, Koscher et al., 2023, Wang et al., 2025], but they showcase solutions rather than provide general criteria for prioritizing and integrating AI under local constraints. In short, the literature remains fragmented and cognitively demanding; decision science suggests that complex dynamic settings require simplified but principled frames for strategy [Simon, 1955, Gigerenzer and Selten, 2002].

We present the AI for Science Strategic Compass (AFSC), a compact, function-based framework that aligns what science needs with what AI can do. AFSC organizes the AI landscape into six core functions and aligns each with four universal scientific-discovery tensions, instantiating a 6×4 Strategy Matrix in which every cell names a shared mitigation logic and offers three strategic pathways linked to representative method families. To establish scope and coverage, we validate the tensions and functions with automated corpus audits: the four tensions collectively cover all sampled abstracts across six natural-science domains, and the six functions account for nearly all capabilities reported in recent AI papers. By abstracting from algorithms to functions and anchoring each pathway in an atomic layer of three minimal, mutually exclusive and collectively exhaustive (MECE) categories per fuction, AFSC lowers cognitive load [Sweller, 2011] while preserving theoretical rigor, yielding guidance that remains stable as techniques evolve.

2 Four universal research tensions

We treat four system-intrinsic barriers to scientific discovery as the problem descriptors to which the Compass aligns AI functions. For clarity we use their full names and adopt short labels for later reference. *System Complexity* (*Complexity*) is the intrinsic structural intricacy that makes modeling, explanation, and generalization difficult even when data are abundant; it encompasses high dimensionality, tightly coupled variables, nonlinear or chaotic interactions, emergence, non-stationary shifting, and multiscale or multimodal signals. *Experimental Constraint* (*Constraint*) is the set of limits on running empirical trials that slow or cap evidence acquisition; typical causes include high per-trial cost or long cycle times, safety or irreversibility that undermines repeatability, physical inaccessibility, and low throughput or limited parallelism. *Data Scarcity* (*Scarcity*) is a shortfall of sufficiently informative and reliable evidence relative to problem difficulty; it includes few-shot regimes, rare or inaccessible phenomena, weak or missing labels, noisy or biased curation across heterogeneous sources, and incomplete or inconsistent records. *Combinatorial Explosion* (*Explosion*) is the exponential growth of design, parameter, configuration, or solution spaces that render exhaustive search infeasible in discrete, continuous, or mixed settings. The four tensions are orthogonal in intent and collectively exhaustive at a coarse granularity.

To test cross-domain coverage, we sampled 3,000 abstracts from top journals across six domains in 2021–2025 via the Crossref REST API (DOI-deduplicated) [Crossref, 2025]. For each abstract, an LLM via the OpenAI API (model: `gpt-5-mini`) generated 1–3 bottleneck hypotheses [OpenAI, 2025]; sentence-level evidence was retrieved with Okapi BM25 [Robertson and Zaragoza, 2009], and hypothesis–evidence entailment was scored by a DeBERTa-v3 cross-encoder (`cross-encoder/nli-deberta-v3-base`) using MNLI-style templates [He et al., 2021, Reimers and Gurevych, 2019, Wolf et al., 2020, Williams et al., 2018]. We used a no-abstention Top-2 policy to emphasize coverage and logged (evidence, hypothesis, probability) per label for auditability. Top-2 coverage was 100% (OTHER=0), and the most common co-occurrence is *Complexity* + *Scarcity* (65.9%). These results are consistent with the tensions being domain-general and collectively exhaustive at coarse granularity.

3 Six core AI functions and their dependencies

We structure the Compass around six domain-agnostic AI functions at an intermediate level of abstraction because this is the only granularity that can be aligned with research tensions while remaining stable as individual AI methods evolve. *Representation* encodes raw, heterogeneous inputs into structured or latent states; *Reason & Infer* operates on those states to produce explicit constraints, causal–probabilistic relations, and calibrated beliefs; *Optimize & Control* selects actions or designs under stated objectives and constraints, either open-loop or in closed-loop policy control; *Simulate*

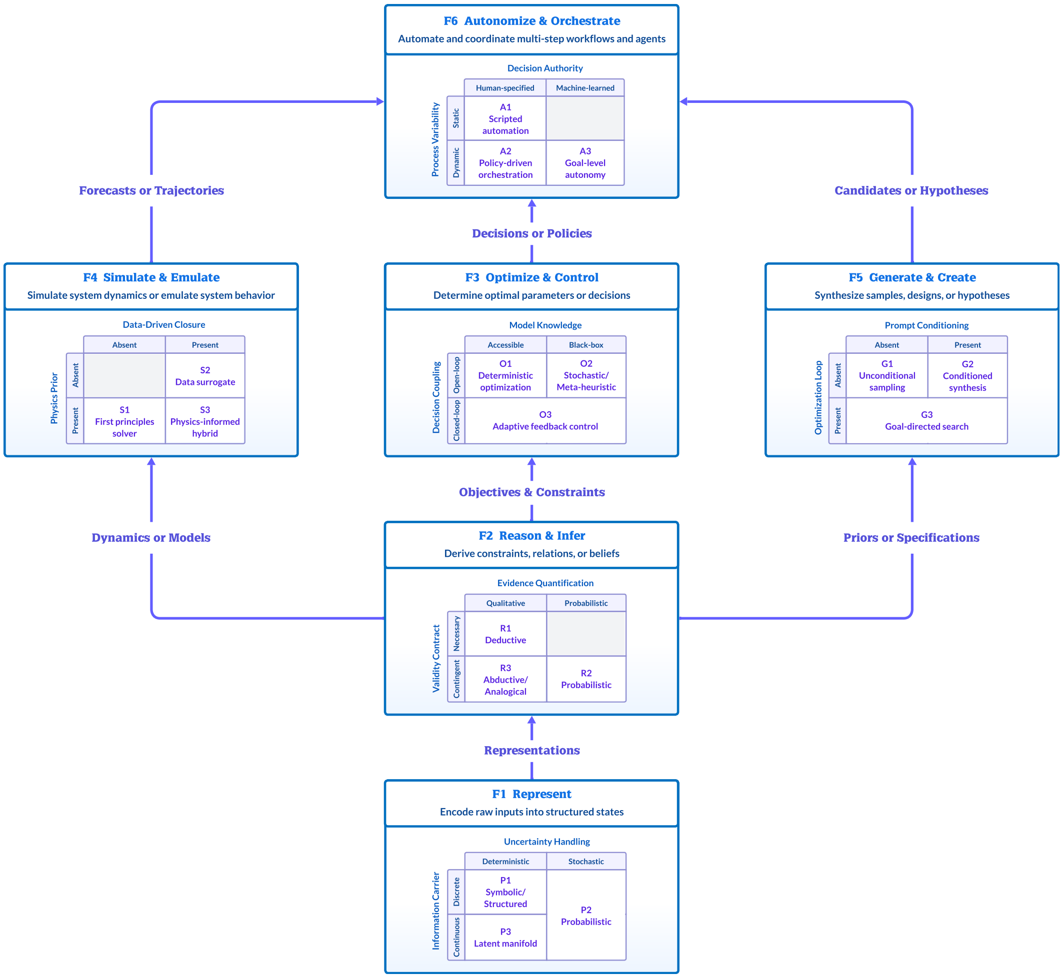

Figure 1: AI core function ontology: a two-level capability framework. The upper level lists the functions; the lower level shows, for each function, its three atomic categories obtained by crossing two intrinsic binary axes. Minimal prerequisite structure: Representation → Reason & Infer → {Optimize & Control, Simulate & Emulate, Generate & Create} → Autonomize & Orchestrate. Optional lateral compositions among Optimize, Simulate, and Generate are omitted; dependencies are functional rather than temporal. High-resolution, citable version: `https://doi.org/10.5281/zenodo.17669687`

& Emulate reproduces dynamics to forecast, test counterfactuals, and run virtual experiments using first-principles solvers, data surrogates, or hybrids; *Generate & Create* synthesizes candidate data, artefacts, or designs conditioned on prompts or goals; *Autonomize & Orchestrate* composes and supervises these capabilities in end-to-end workflows. Each function further decomposes into a triad of atomic capability categories obtained by crossing two intrinsic binary axes; these atomic triads give each function a minimal, MECE internal structure and later serve as anchors for strategic pathways.

The functions and their atomic triads form a minimal prerequisite chain (Fig. 1). Raw signals must be encoded before they can support inference; separating representation from reasoning is standard in cognitive and information-processing theory [Marr, 1982, Ackoff, 1989]. Reasoning then supplies what downstream modules require (explicit objectives and constraints for optimization and closed-loop control, and governing equations or learned dynamics for simulation), so *Reason & Infer* is a prerequisite for *Optimize & Control* and *Simulate & Emulate* [Åström and Murray, 2008, Raissi et al., 2019]. Similarly, generative synthesis depends on targets or priors made explicit by reasoning, making *Generate & Create* downstream of *Reason & Infer*. Workflow autonomy is meaningful only once decisions or policies, forecasts or trajectories, and candidate artefacts or hypotheses exist to be scheduled and supervised; accordingly, *Autonomize & Orchestrate* depends on *Optimize & Control*, *Simulate & Emulate*, and *Generate & Create*. Lateral exchanges among *Optimize*, *Simulate*, and

Generate are common in practice but are optional compositions rather than logical prerequisites; when they occur, *Reason & Infer* typically mediates scoring, constraint checking, and calibration.

To assess coverage of the six-function taxonomy on recent AI papers, we stratified arXiv (2019–2025) by subfield and year, downloaded PDFs, and mined Methods/Contributions/Evaluation passages. For each function, we retrieved candidate passages with BM25 (curated lexicon) [Robertson and Zaragoza, 2009] and scored function–passage entailment using a DeBERTa-v3 cross-encoder (`cross-encoder/nli-deberta-v3-base`) via Sentence-Transformers/HuggingFace Transformers with MNLI-style hypotheses [He et al., 2021, Reimers and Gurevych, 2019, Wolf et al., 2020, Williams et al., 2018]. A deterministic policy with fixed thresholds accepted a function only when supported by strong entailment on retrieved evidence; if the top two label scores fell within a small margin, we retained both labels (Top-2), and OTHER was used only when no function met acceptance. Each decision stores the supporting passage and retrieval/entailment scores for audit. Coverage was 98.9% on $N = 628$ papers; spot audits of OTHER cases indicate they stem primarily from weak evidence extraction rather than a missing seventh function. This supports the taxonomy as providing near-complete, auditable coverage of capabilities reported in recent AI research. The compact backbone provides a stable, domain-agnostic scaffold for method selection, lowers cognitive load, and supports more impartial, context-aware decisions as techniques evolve.

4 Strategy matrix

4.1 Derivation overview

We constructed the Compass via a theory-guided procedure, complemented by two auditable empirical checks; the resulting 6×4 Strategy Matrix is shown in Fig. 2.

(i) We first identified four universal discovery tensions from cross-domain pain points and validated collective exhaustiveness by an automated corpus audit on 3000 abstracts across six natural-science domains.
(ii) We then elicited a minimal, mutually exclusive set of six core AI functions defined by epistemic role rather than technique, ensuring method- and domain-agnostic scope.
(iii) For each function we fixed two intrinsic binary axes (first-principles motivated, empirically recurrent), crossed them, and applied a void/merge test to remove logically empty quadrants and merge operationally indistinct ones, yielding a triad of atomic categories—minimal, non-decomposable classes along those axes that render each function internally MECE [Birkhoff, 1940, Davey and Priestley, 2002]. These atomic triads form the second level of the ontology in Fig. 1.
(iv) For every tension–function pairing, we distilled three strategic pathways articulated at the mitigation-mechanism level and transferrable across domains; a candidate was retained only if it expressed a distinct mechanism, admitted a minimal atomic signature, and transferred across at least two scientific domains. After identification, we attach the minimal atomic signature as a post-hoc anchor to intrinsic properties rather than transient techniques, which keeps pathway-level revisions infrequent.

Each cell then receives a single keyword naming the shared mitigation logic, and its pathways are linked to representative method families so the strategy is actionable without prescribing a single model. $\star$ marks non-exclusive high-leverage entry points for that tension.

Pathways may evolve by split (one label conflates separable mechanisms), merge (two labels are mechanistically interchangeable), retag (primary relief lies in another function or tension), or addition (a genuinely new mechanism recurs across domains with a verifiable minimal signature). If a proposed pathway cannot be anchored to any minimal atomic signature within its function, we reassign it to a more appropriate function or drop it as ill-posed. Atomic triads are revised only when robust cross-domain evidence shows that the current two axes fail to span the function's variability. Anchoring pathways to atomic categories thus provides a stable, falsifiable, technique-agnostic basis.

4.2 Reading and using the matrix

The Compass is organized into three layers, each with a distinct role. The function layer specifies six domain-agnostic capabilities and their prerequisite relations, together delimiting the scope of what an AI system can do. The atomic layer fixes, for each function, two intrinsic axes and induces a triad of

Function (F) \ Tension (T)	T1 Complexity	T2 Constraints	T3 Scarcity	T4 Explosion
F1 Represent Uncertainty Handling: Deterministic / Stochastic Information Carrier: Discrete / Continuous P1 Symbolic/ Structured (Discrete, Deterministic) P2 Probabilistic (Stochastic) P3 Latent manifold (Continuous, Deterministic)	**Factorize** ★ • Latent factorization (P3) disentangled representation learning • Hierarchical abstraction (P3) hierarchical VAEs or transformers • Structured graph extraction (P1) typed-graph mining; schema induction	**Robustify** • Physics-aware embedding (P3+P1) physics-informed neural encoders • Noise-robust encoding (P3) denoising autoencoder; contrastive learning • Domain-invariant mapping (P3+P2) domain adversarial networks	**Amplify** • Signal-boost encoding (P3) contrastive/self-supervised pretraining • Cross-source fusion (P3) multimodal joint embeddings • Uncertainty-aware imputation (P2+P3) VAE/diffusion-based imputers	**Compress** • Regularity Compression (P1+P3) symmetry-aware/equivariant encoders • Multi-resolution abstraction (P3) pyramidal/multiscale encoders • Compact latent indexing (P1) deep hashing; product quantization
F2 Reason & Infer Evidence Quantification: Qualitative / Probabilistic Validity Contract: Necessary / Contingent R1 Deductive (Necessary, Qualitative) R3 Abductive/ Analogical (Contingent, Qualitative) R2 Probabilistic (Contingent, Probabilistic)	**Causalize** ★ • Causalize system dynamics (R3+R2) causal structure learning; causal GNNs • Probabilistic dependency modeling (R2) Bayesian networks & factor graphs • Deductive causal invariants (R1) differentiable constraint layers	**Prequalify** • Deductive feasibility prequalification (R1) neuro-rule engines; constraint programming • Probabilistic feasibility estimation (R2) probabilistic graphical models • Abductive constraint induction (R3) neuro-symbolic inductive logic programming	**Generalize** • Rule-driven extrapolation (R1) neuro-symbolic regression • Bayesian prior integration (R2) hierarchical Bayesian models • Few-shot hypothesis induction (R3) meta learning; prompt-tuned LLMs	**Prune** ★ • Deductive constraint propagation (R1) differentiable CSP networks • Probabilistic branch ranking (R2) neural posterior-guided search • Abductive pathway trimming (R3) neuro-symbolic weighted abduction
F3 Optimize & Control Model Knowledge: Accessible / Black-box Decision Coupling: Open-loop / Closed-loop O1 Deterministic optimization (Open-loop, Accessible) O2 Stochastic/ Meta-heuristic (Open-loop, Black-box) O3 Adaptive feedback control (Closed-loop)	**Navigate** • Gradient-based surrogate navigation (O1) neural operator surrogates • Heuristic black-box search (O2) Bayesian optimization (BO) • Self-adaptive feedback control (O3) meta-reinforcement learning (RL) adaptation	**Satisfy** ★ • Learned safety certificates (O1) barrier/trust-region loss surrogates • Constrained acquisition search (O2) safe BO; constrained evolutionary search • Risk-sensitive adaptive policies (O3) risk-sensitive model predictive control	**Prioritize** • Initialization-based optimization (O1) meta-learning initialization • Uncertainty-driven sampling (O2) Bayesian active learning • Curriculum-adaptive control (O3) meta-RL curriculum scheduler	**Guided-Search** ★ • Multi-fidelity surrogate screening (O1) hierarchical neural operator cascades • Structure-guided search (O2) GNN-guided branching & pruning • Hierarchical policy refinement (O3) hierarchical RL/model-based refinement
F4 Simulate & Emulate Data-Driven Closure: Absent / Present Physics Prior: Absent / Present S2 Data surrogate (Physics Prior Absent, Closure Present) S1 First principles solver (Physics Prior Present, Closure Absent) S3 Physics-informed hybrid (Physics Prior Present, Closure Present)	**Approximate** • Coarse-grain approximation (S1+S2) reduced-order models with learned closure • Stochastic scenario sampling (S2) variational/diffusion simulators • Residual-hybrid acceleration (S3) physics-informed residual correction	**Virtualize** ★ • Virtual lab emulation (S1+S2) digital-twin emulators (composite) • Rule-constrained simulation (S1) constraint-enforcing numerical solvers • Safe exploration loops (S3) safety-constrained model-based RL	**Synthesize** ★ • Mechanistic data synthesis (S1) equation-/physics-based simulators • Surrogate extrapolation (S2) diffusion or GAN emulators • Physics-informed augmentation (S3) physics-informed generative emulators	**Accelerate** • Physics-based prefiltering (S1) reduced-order modeling screening • Structure-guided pruning (S2) graph/symbolic-heuristics emulation • Adaptive multi-fidelity screening (S3) staged solver-emulator loops
F5 Generate & Create Prompt Conditioning: Absent / Present Optimization Loop: Absent / Present G1 Unconditional sampling (Loop Absent, Conditioning Absent) G2 Conditioned synthesis (Loop Absent, Conditioning Present) G3 Goal-directed search (Loop Present)	**Probe** • Counterfactual probing (G3) counterfactual VAEs; causal GANs • Edge-case exploration (G2+G3) active tail exploration • Latent subspace probing (G1) latent space traversal/interpolation	**Prototype** • Unconstrained prototype drafting (G1) unconditional GANs; latent mixing • Rule-conditioned prototyping (G2) rule-/constraint-conditioned diffusion • Constraint-loop prototyping (G3) constraint-aware BO for generators	**Augment** ★ • Data augmentation (G1) classical or generative augmentation • Weak label expansion (G2) LLM pseudo labeling; conditional diffusion • Utility-guided augmentation (G3) acquisition-guided generative augmentation	**Seed** • Diversity-maximized sampling (G1) determinantal point process samplers • Constraint-aware seed search (G2) grammar-constrained samplers • Hierarchical assembly (G3) multi-stage RL assembly
F6 Autonomize & Orchestrate Decision Authority: Human-specified / Machine-learned Process Variability: Static / Dynamic A1 Scripted automation (Static, Human-specified) A2 Policy-driven orchestration (Dynamic, Human-specified) A3 Goal-level autonomy (Dynamic, Machine-learned)	**Auto-Compose** • Scripted multimodal coordination (A1) workflow DSLs; rule-based tool-routing • Policy-driven scaling & routing (A2) policy-based workflow schedulers • Closed-loop pipeline auto-tuning (A3) Bayesian/RL pipeline tuning	**Auto-Enforce** • Validation & workflow codification (A1) protocol DSLs; provenance graphs • Policy-driven guardrails & feedback (A2) feedback controllers for labs • Self-lab orchestration (A3) safe-RL controllers; digital-twin planners	**Auto-Curate** • Scripted acquisition & integration (A1) schema-aware ETL pipelines • Policy-driven auto-labeling (A2) active learning labeling schedulers • Autonomous quality refinement (A3) data-cleaning/noise-filtering agents	**Auto-Screen** • Batch high-throughput screening (A1) workflow DAG pipeline frameworks • Policy-driven triage & scheduling (A2) multi-armed bandit schedulers • Closed-loop active screening (A3) active learning acquisition controllers

Figure 2: AFSC strategy matrix. Rows: six core AI functions. Columns: four universal scientific-discovery tensions. The left column shows, for each function, its two intrinsic binary axes and the resulting three atomic categories (which render the function internally MECE). Each cell is labeled with a keyword naming the shared mitigation logic and lists three distinct strategic pathways with representative method families (illustrative, not prescriptive). $\star$ marks high-leverage cells, typical entry points for that tension (non-exclusive). Full pathway definitions and method-family citations appear in Appendix C. High-resolution, citable version: `https://doi.org/10.5281/zenodo.17672434`

atomic categories—minimal, non-overlapping mechanisms along those axes that render the function internally MECE and provide intrinsic anchors independent of particular algorithms. The strategy layer is the 6×4 Matrix: each cell is labeled with a single keyword capturing its shared mitigation logic, lists three strategic pathways that realize that logic, and cites representative method families.

A typical workflow proceeds as follows. Identify the dominant tension(s) in the scientific problem; consult the starred cells ($\star$) as high-leverage, non-prescriptive entry points; select one or more strategic pathways within the chosen cell that fit your data, expertise, computational budget, and experimental constraints; then instantiate the pathway with a suitable method family or an equivalent alternative. A row-wise scan shows how a single function changes stance across tensions; a column-wise scan contrasts mechanisms across functions for a fixed tension. Because each pathway is anchored to a fixed atomic layer, the conceptual vocabulary remains stable even as specific algorithms evolve, enabling consistent comparison and incremental updates without revising the scaffold.

5 Case study: exoplanet spectral retrieval

Problem and dominant tensions. Retrieving atmospheric parameters from exoplanet spectra is an ill-posed inverse problem. Observations mix multiple molecules, overlapping lines, and cloud opacity; many parameter vectors produce near-identical spectra, creating degeneracy [Madhusudhan, 2019, Welbanks and Madhusudhan, 2019]. These features induce *Complexity* via multiscale, entangled structure and non-identifiability, and open an interpretability gap where black-box fits cannot attribute spectral segments to physical causes. The workflow also faces *Constraints*: forward radiative-transfer evaluations with sampling-based Bayesian retrievals (MCMC or nested sampling) are computationally expensive, and repeated space-based transits are scarce [Vasist et al., 2023].

Compass-guided selection. Following the Matrix's starred cues, we address *Complexity* via *Representation* $\rightarrow$ *Complexity* (Factorize) and *Reason & Infer* $\rightarrow$ *Complexity* (Causalize), and address *Constraint*s via *Optimize & Control* $\rightarrow$ *Constraint*s (Satisfy) and *Simulate & Emulate* $\rightarrow$ *Constraint*s (Virtualize).

Instantiated pathways.

Factorize. Learn a compact latent spectra embedding to accelerate convergence and mitigate overfitting (latent factorization, P3); augment it with a multiscale "skeleton" that captures coarse-to-fine topology (hierarchical abstraction, P3); and construct a typed spectral graph encoding relations among bands, molecules, and cloud or continuum components (structured graph extraction, P1).

Causalize. Expose parameter–wavelength links by coupling the skeleton and parameter vector with cross-attention and an attention–skeleton alignment loss (causalize system dynamics, R3+R2). Irreducible degeneracy is modeled with a mixture-density posterior (probabilistic dependency modeling, R2). Hard physics (e.g., monotonicities, equilibrium chemistry) enters via conditioning vectors and regularizers (deductive causal invariants, R1).

Satisfy. Conduct gradient-guided parameter search by backpropagating through a differentiable surrogate and shaping the objective with physics terms (learned safety certificates, O1), together with a two-phase scheduler that suppresses early noise drift and adapts later refinement (risk-sensitive adaptive policies, O3).

Virtualize. A first-principles RT forward model generates 22 000 synthetic spectra for training (rule-constrained simulation, S1). A latent-diffusion surrogate conditioned on the skeleton and learned parameters provides fast emulation for both reconstruction and retrieval (virtual-lab emulation, S1+S2).

Outcome. These pathways preserve physical interpretability, expose parameter attributions, and shorten the retrieval loop, while remaining aligned with the Compass's strategy layer rather than ad-hoc model choices.

6 Practical value and scope

The AFSC serves as a compass for a fragmented, fast-evolving AI landscape. It aligns a problem's dominant tension with the relevant function and atom-anchored pathways, turning unconstrained, tool-driven browsing into targeted, strategy-first exploration. This panoramic view of capability reduces availability bias and tool myopia and supports defensible, context-aware decisions about integrating AI into a research workflow. In practice, users proceed from (i) tension identification to (ii) starred high-leverage functions, (iii) appropriate strategic pathways, and (iv) method families, preserving methodological freedom while sharply narrowing the design space. Because the vocabulary is anchored at the atomic layer, it remains stable as techniques evolve: most new methods instantiate combinations of existing atoms, so the scaffold rarely requires structural revision. This stability provides a time-robust basis for decision-making across domains and enables cumulative learning without reframing the map.

7 Limitations and outlook

AFSC is a decision aid, not a performance guarantee. The pathways are representative rather than exhaustive, and some domains may require additional variants. The current Matrix reflects a theory-guided design with automated corpus checks; broader evaluation remains open. We will pursue multi-lab user studies, ablations of pathway choices, and longitudinal tracking of downstream impact. We are also building an open Matrix browser with per-cell exemplars, links to implementations, and a community contribution workflow.

8 Conclusion

AFSC aligns four universal tensions (*Complexity*, *Constraint*, *Scarcity*, *Explosion*) with six core AI functions (*Representation*; *Reason & Infer*; *Optimize & Control*; *Simulate & Emulate*; *Generate & Create*; *Autonomize & Orchestrate*) in a single Strategy Matrix. By lifting the focus from algorithms to functions and anchoring pathways in a first-principles, MECE atomic layer, the Compass provides compact, cognitively tractable guidance that remains stable amid rapid technical change and supports disciplined, cross-domain practice.

Acknowledgements

We thank colleagues in the AI for Science community for discussions on early versions of the AFSC strategy matrix, and the NeurIPS AI for Science Workshop reviewers for their constructive feedback. Their comments helped refine the presentation of the Compass and motivated our ongoing effort to develop an engineered, auto-updating digital platform version of the framework as a shared resource for the community.

References

Joshua Achiam, David Held, Aviv Tamar, and Pieter Abbeel. Constrained policy optimization. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 22–31. PMLR, 06–11 Aug 2017. URL `https://proceedings.mlr.press/v70/achiam17a.html`.

Russell L. Ackoff. From data to wisdom. *Journal of Applied Systems Analysis*, 16:3–9, 1989.

Akshay Agrawal, Brandon Amos, Shane Barratt, Stephen Boyd, Steven Diamond, and J. Zico Kolter. Differentiable convex optimization layers. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL `https://proceedings.neurips.cc/paper_files/paper/2019/file/9ce3c52fc54362e22053399d3181c638-Paper.pdf`.

Brandon Amos, Ivan Jimenez, Jacob Sacks, Byron Boots, and J. Zico Kolter. Differentiable mpc for end-to-end planning and control. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL `https://proceedings.neurips.cc/paper_files/paper/2018/file/ba6d843eb4251a4526ce65d1807a9309-Paper.pdf`.

Dirk V. Arnold and Nikolaus Hansen. A (1+1)-cma-es for constrained optimisation. In *Proceedings of the 14th Annual Conference on Genetic and Evolutionary Computation*, GECCO '12, page 297–304, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450311779. doi: 10.1145/2330163.2330207. URL `https://doi.org/10.1145/2330163.2330207`.

Jordan T. Ash, Chicheng Zhang, Akshay Krishnamurthy, John Langford, and Alekh Agarwal. Deep batch active learning by diverse, uncertain gradient lower bounds. In *International Conference on Learning Representations*, 2020. URL `https://openreview.net/forum?id=ryghZJBKPS`.

Karl Johan Åström and Richard M. Murray. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, Princeton, NJ, 2008. ISBN 978-0-691-13576-2. doi: 10.1515/9781400828739.

Alexei Baevski, Yuhao Zhou, Abdelrahman Mohamed, and Michael Auli. wav2vec 2.0: A framework for self-supervised learning of speech representations. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 12449–12460. Curran Associates, Inc., 2020. URL `https://proceedings.neurips.cc/paper_files/paper/2020/file/92d1e1eb1cd6f9fba3227870bb6d7f07-Paper.pdf`.

Jiaxin Bai, Yicheng Wang, Tianshi Zheng, Yue Guo, Xin Liu, and Yangqiu Song. Advancing abductive reasoning in knowledge graphs through complex logical hypothesis generation. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1312–1329, 2024. doi: 10.18653/v1/2024.acl-long.72. URL `https://aclanthology.org/2024.acl-long.72.pdf`.

Hangbo Bao, Wenhui Wang, Li Dong, Qiang Liu, Owais Khan Mohammed, Kriti Aggarwal, Subhojit Som, Songhao Piao, and Furu Wei. Vlmo: Unified vision-language pre-training with mixture-of-modality-experts. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 32897–32912. Curran Associates, Inc., 2022. URL `https://proceedings.neurips.cc/paper_files/paper/2022/file/d46662aa53e78a62afd980a29e0c37ed-Paper-Conference.pdf`.

Rémi Bardenet, Subhroshekhar Ghosh, and Meixia Lin. Determinantal point processes based on orthogonal polynomials for sampling minibatches in sgd. In *Proceedings of the 35th International Conference on Neural Information Processing Systems*, NIPS '21, Red Hook, NY, USA, 2021. Curran Associates Inc. ISBN 9781713845393.

Denis Baylor, Eric Breck, Heng-Tze Cheng, Noah Fiedel, Chuan Yu Foo, Zakaria Haque, Salem Haykal, Mustafa Ispir, Vihan Jain, Levent Koc, Chiu Yuen Koo, Lukasz Lew, Clemens Mewald, Akshay Naresh Modi, Neoklis Polyzotis, Sukriti Ramesh, Sudip Roy, Steven Euijong Whang, Martin Wicke, Jarek Wilkiewicz, Xin Zhang, and Martin Zinkevich. Tfx: A tensorflow-based production-scale machine learning platform. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '17, page 1387–1395, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450348874. doi: 10.1145/3097983.3098021. URL `https://doi.org/10.1145/3097983.3098021`.

Yoshua Bengio, Salem Lahlou, Tristan Deleu, Edward J. Hu, Mo Tiwari, and Emmanuel Bengio. Gflownet foundations. *Journal of Machine Learning Research*, 24(210):1–55, 2023. URL `http://jmlr.org/papers/v24/22-0364.html`.

Felix Berkenkamp, Andreas Krause, and Angela P. Schoellig. Bayesian optimization with safety constraints: safe and automatic parameter tuning in robotics. *Machine Learning*, 112:3713 – 3747, 2016. URL `https://api.semanticscholar.org/CorpusID:3682030`.

Jules Berman and Benjamin Peherstorfer. Randomized sparse neural galerkin schemes for solving evolution equations with deep networks. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, volume 36, pages 4097–4114. Curran Associates, Inc., 2023. URL `https://proceedings.neurips.cc/paper_files/paper/2023/file/0cb310ed8121549488fea8e8c2056096-Paper-Conference.pdf`.

Tom Beucler, Michael Pritchard, Stephan Rasp, Jordan Ott, Pierre Baldi, and Pierre Gentine. Enforcing analytic constraints in neural networks emulating physical systems. *Phys. Rev. Lett.*, 126:098302, Mar 2021. doi: 10.1103/PhysRevLett.126.098302. URL `https://link.aps.org/doi/10.1103/PhysRevLett.126.098302`.

Garrett Birkhoff. *Lattice Theory*. American Mathematical Society, December 1940. ISBN 9781470431730. doi: 10.1090/coll/025. URL `http://dx.doi.org/10.1090/coll/025`.

Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer, New York, 2006. ISBN 9780387310732. doi: 10.1007/978-0-387-45528-0.

Daniil A. Boiko, Robert MacKnight, Ben Kline, and Gabe Gomes. Autonomous chemical research with large language models. *Nature*, 624(7992):570–578, December 2023a. ISSN 1476-4687. doi: 10.1038/s41586-023-06792-0. URL `http://dx.doi.org/10.1038/s41586-023-06792-0`.

Daniil A Boiko, Robert MacKnight, Ben Kline, and Gabe Gomes. Autonomous chemical research with large language models. *Nature*, 624(7992):570–578, 2023b.

Joan Bruna and Stephane Mallat. Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1872–1886, 2013. doi: 10.1109/TPAMI.2012.230.

Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15): 3932–3937, 2016. doi: 10.1073/pnas.1517384113. URL `https://www.pnas.org/doi/abs/10.1073/pnas.1517384113`.

Richard B. Canty, Jeffrey A. Bennett, Keith A. Brown, Tonio Buonassisi, Sergei V. Kalinin, John R. Kitchin, Benji Maruyama, Robert G. Moore, Joshua Schrier, Martin Seifrid, Shijing Sun, Tejs Vegge, and Milad Abolhasani. Science acceleration and accessibility with self-driving labs. *Nature Communications*, 16(1), April 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-59231-1. URL `http://dx.doi.org/10.1038/s41467-025-59231-1`.

Franck Cappello, Sandeep Madireddy, Robert Underwood, Neil Getty, Nicholas Lee-Ping Chia, Nesar Ramachandra, Josh Nguyen, Murat Keceli, Tanwi Mallick, Zilinghan Li, Marieme Ngom, Chenhui Zhang, Angel Yanguas-Gil, Evan Antoniuk, Bhavya Kailkhura, Minyang Tian, Yufeng Du, Yuan-Sen Ting, Azton Wells, Bogdan Nicolae, Avinash Maurya, M. Mustafa Rafique, Eliu Huerta, Bo Li, Ian Foster, and Rick Stevens. EAIRA: Establishing a methodology for evaluating AI models as scientific research assistants. *arXiv preprint arXiv:2502.20309*, 2025. doi: 10.48550/arXiv.2502.20309. URL `https://arxiv.org/abs/2502.20309`.

Lu Chen, Yan Li, Yanjie Ma, Lin Gao, and Liang Yu. Multiscale graph equivariant diffusion model for 3d molecule design. *Science Advances*, 11(16):eadv0778, 2025. doi: 10.1126/sciadv.adv0778. URL `https://www.science.org/doi/abs/10.1126/sciadv.adv0778`.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 1597–1607. PMLR, 13–18 Jul 2020. URL `https://proceedings.mlr.press/v119/chen20j.html`.

Evgenii E Chzhen, Christophe Giraud, Zhen LI, and Gilles Stoltz. Small total-cost constraints in contextual bandits with knapsacks, with application to fairness. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. URL `https://openreview.net/forum?id=uZvG0HLkOB`.

Taco Cohen and Max Welling. Group equivariant convolutional networks. In Maria Florina Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 2990–2999, New York, New York, USA, 20–22 Jun 2016. PMLR. URL `https://proceedings.mlr.press/v48/cohenc16.html`.

Taco S. Cohen and Max Welling. Steerable CNNs. In *International Conference on Learning Representations*, 2017. URL `https://openreview.net/forum?id=rJQKYt5ll`.

Andrew R. Conn, Katya Scheinberg, and Luis N. Vicente. *Introduction to Derivative-Free Optimization*. MPS-SIAM Series on Optimization. SIAM, Philadelphia, 2009. ISBN 978-0-89871-668-9. doi: 10.1137/1.9780898718768.

Crossref. Crossref rest api documentation. `https://www.crossref.org/documentation/retrieve-metadata/rest-api/`, 2025. Accessed 2025-06-23.

Michael R. Crusoe, Sanne Abeln, Alexandru Iosup, Peter Amstutz, John Chilton, Nebojša Tijanić, Hervé Ménager, Stian Soiland-Reyes, Bogdan Gavrilović, Carole Goble, and The CWL Community. Methods included: standardizing computational reuse and portability with the common workflow language. *Commun. ACM*, 65(6):54–63, May 2022. ISSN 0001-0782. doi: 10.1145/3486897. URL `https://doi.org/10.1145/3486897`.

Ekin D. Cubuk, Barret Zoph, Dandelion Mané, Vijay Vasudevan, and Quoc V. Le. Autoaugment: Learning augmentation strategies from data. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 113–123, 2019. doi: 10.1109/CVPR.2019.00020.

Wang-Zhou Dai, Qiuling Xu, Yang Yu, and Zhi-Hua Zhou. Bridging machine learning and logical reasoning by abductive learning. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL `https://proceedings.neurips.cc/paper_files/paper/2019/file/9c19a2aa1d84e04b0bd4bc888792bd1e-Paper.pdf`.

B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, April 2002. ISBN 9780511809088. doi: 10.1017/cbo9780511809088. URL `http://dx.doi.org/10.1017/CBO9780511809088`.

Filipe de Avila Belbute-Peres, Kevin Smith, Kelsey Allen, Josh Tenenbaum, and J. Zico Kolter. End-to-end differentiable physics for learning and control. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL `https://proceedings.neurips.cc/paper_files/paper/2018/file/842424a1d0595b76ec4fa03c46e8d755-Paper.pdf`.

Ewa Deelman, Karan Vahi, Gideon Juve, Mats Rynge, Scott Callaghan, Philip J. Maechling, Rajiv Mayani, Weiwei Chen, Rafael Ferreira da Silva, Miron Livny, and Kent Wenger. Pegasus, a workflow management system for science automation. *Future Gener. Comput. Syst.*, 46(C):17–35, May 2015. ISSN 0167-739X. doi: 10.1016/j.future.2014.10.008. URL `https://doi.org/10.1016/j.future.2014.10.008`.

Ewa Deelman, Tom Peterka, Ilkay Altintas, Christopher D. Carothers, Kerstin Kleese van Dam, Kenneth Moreland, Manish Parashar, Lavanya Ramakrishnan, Michela Taufer, and Jeffrey Vetter. The future of scientific workflows. *The International Journal of High Performance Computing Applications*, 32(1):159–175, 2018. doi: 10.1177/1094342017704893.

Igor Douven. *The Art of Abduction*. MIT Press, Cambridge, MA, 2022. doi: 10.7551/mitpress/14179.001.0001.

Richard Evans and Edward Grefenstette. Learning explanatory rules from noisy data. *J. Artif. Intell. Res.*, 61: 1–64, 2018. doi: 10.1613/JAIR.5714. URL `https://doi.org/10.1613/jair.5714`.

Stefan Falkner, Aaron Klein, and Frank Hutter. BOHB: Robust and efficient hyperparameter optimization at scale. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 1437–1446. PMLR, 10–15 Jul 2018. URL `https://proceedings.mlr.press/v80/falkner18a.html`.

Matthias Feurer, Jost Tobias Springenberg, and Frank Hutter. Using meta-learning to initialize bayesian optimization of hyperparameters. In *Proceedings of the 2014 International Conference on Meta-Learning and Algorithm Selection - Volume 1201*, MLAS'14, page 3–10, Aachen, DEU, 2014. CEUR-WS.org. ISBN 16130073.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 1126–1135. PMLR, 06–11 Aug 2017. URL `https://proceedings.mlr.press/v70/finn17a.html`.

Carlos Florensa, David Held, Markus Wulfmeier, Michael Zhang, and Pieter Abbeel. Reverse curriculum generation for reinforcement learning. In Sergey Levine, Vincent Vanhoucke, and Ken Goldberg, editors, *Proceedings of the 1st Annual Conference on Robot Learning*, volume 78 of *Proceedings of Machine Learning Research*, pages 482–495. PMLR, 13–15 Nov 2017. URL `https://proceedings.mlr.press/v78/florensa17a.html`.

Fabian Fuchs, Daniel Worrall, Volker Fischer, and Max Welling. Se(3)-transformers: 3d roto-translation equivariant attention networks. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 1970–1981. Curran Associates, Inc., 2020. URL `https://proceedings.neurips.cc/paper_files/paper/2020/file/15231a7ce4ba789d13b722cc5c955834-Paper.pdf`.

Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario March, and Victor Lempitsky. Domain-adversarial training of neural networks. *Journal of Machine Learning Research*, 17(59):1–35, 2016. URL `http://jmlr.org/papers/v17/15-239.html`.

Hongyang Gao and Shuiwang Ji. Graph u-nets. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 2083–2092. PMLR, 09–15 Jun 2019. URL `https://proceedings.mlr.press/v97/gao19a.html`.

Jacob R. Gardner, Matt J. Kusner, Zhixiang Xu, Kilian Q. Weinberger, and John P. Cunningham. Bayesian optimization with inequality constraints. In *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32*, ICML'14, page II–937–II–945. JMLR.org, 2014.

Maxime Gasse, Didier Chetelat, Nicola Ferroni, Laurent Charlin, and Andrea Lodi. Exact combinatorial optimization with graph convolutional neural networks. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL `https://proceedings.neurips.cc/paper_files/paper/2019/file/d14c2267d848abeb81fd590f371d39bd-Paper.pdf`.

Gerd Gigerenzer and Reinhard Selten, editors. *Bounded Rationality: The Adaptive Toolbox*. MIT Press, Cambridge, MA, 2002. ISBN 9780262571647.

Nate Gruver, Samuel Stanton, Nathan Frey, Tim G. J. Rudner, Isidro Hotzel, Julien Lafrance-Vanasse, Arvind Rajpal, Kyunghyun Cho, and Andrew Gordon Wilson. Protein design with guided discrete diffusion. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, Red Hook, NY, USA, 2023. Curran Associates Inc.

Jie Gui, Tuo Chen, Jing Zhang, Qiong Cao, Zhenan Sun, Hao Luo, and Dacheng Tao. A survey on self-supervised learning: Algorithms, applications, and future trends. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12):9052–9071, December 2024. ISSN 1939-3539. doi: 10.1109/tpami.2024.3415112. URL `http://dx.doi.org/10.1109/TPAMI.2024.3415112`.

Nico Gürtler, Dieter Büchler, and Georg Martius. Hierarchical reinforcement learning with timed subgoals. In *Proceedings of the 35th International Conference on Neural Information Processing Systems*, NIPS '21, Red Hook, NY, USA, 2021. Curran Associates Inc. ISBN 9781713845393.

Danijar Hafner, Timothy P Lillicrap, Mohammad Norouzi, and Jimmy Ba. Mastering atari with discrete world models. In *International Conference on Learning Representations*, 2021. URL `https://openreview.net/forum?id=0oabwyZbOu`.

Erik Härkönen, Aaron Hertzmann, Jaakko Lehtinen, and Sylvain Paris. Ganspace: discovering interpretable gan controls. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA, 2020. Curran Associates Inc. ISBN 9781713829546.

Florian Häse, Loïc M. Roch, and Alán Aspuru-Guzik. Next-generation experimentation with self-driving laboratories. *Trends in Chemistry*, 1(3):282–291, 2019. doi: 10.1016/j.trechm.2019.02.007.

Catrin Hasselgren and Tudor I. Oprea. Artificial intelligence for drug discovery: Are we there yet? *Annual Review of Pharmacology and Toxicology*, 64(1):527â€“550, January 2024. ISSN 1545-4304. doi: 10.1146/annurev-pharmtox-040323-040828. URL `http://dx.doi.org/10.1146/annurev-pharmtox-040323-040828`.

Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. Masked autoencoders are scalable vision learners. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 15979–15988, 2022. doi: 10.1109/CVPR52688.2022.01553.

Pengcheng He, Jianfeng Gao, and Weizhu Chen. DeBERTaV3: Improving DeBERTa using ELECTRA-style pre-training with gradient-disentangled embedding sharing, 2021. URL https://arxiv.org/abs/2111.09543.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 6840–6851. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf.

Jonathan Ho, Chitwan Saharia, William Chan, David J. Fleet, Mohammad Norouzi, and Tim Salimans. Cascaded diffusion models for high fidelity image generation. *J. Mach. Learn. Res.*, 23(1), January 2022. ISSN 1532-4435.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=nZeVKeeFYf9.

Yuanming Hu, Luke Anderson, Tzu-Mao Li, Qi Sun, Nathan Carr, Jonathan Ragan-Kelley, and Frédo Durand. Difftaichi: Differentiable programming for physical simulation. *ICLR*, 2020.

Daolang Huang, Yujia Guo, Luigi Acerbi, and Samuel Kaski. Amortized bayesian experimental design for decision-making. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024a. URL https://openreview.net/forum?id=zBG7WogAvm.

Kaixuan Huang, Yukang Yang, Kaidi Fu, Yanyi Chu, Le Cong, and Mengdi Wang. Latent diffusion models for controllable rna sequence generation. *arXiv preprint arXiv:2409.09828*, 2024b.

Yu-Xuan Huang, Wang-Zhou Dai, Le-Wen Cai, Stephen H Muggleton, and Yuan Jiang. Fast abductive learning by similarity-based consistency optimization. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 26574–26584. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper_files/paper/2021/file/df7e148cabfd9b608090fa5ee3348bfe-Paper.pdf.

Zizhou Huang, Davi Colli Tozoni, Arvi Gjoka, Zachary Ferguson, Teseo Schneider, Daniele Panozzo, and Denis Zorin. Differentiable solver for time-dependent deformation problems with contact. *ACM Trans. Graph.*, 43 (3), May 2024c. ISSN 0730-0301. doi: 10.1145/3657648. URL https://doi.org/10.1145/3657648.

G. E. Hughes and M. J. Cresswell. A new introduction to modal logic. *Studia Logica*, 62(3):439–441, 1996. doi: 10.4324/9780203028100.

Florian Häse, Matteo Aldeghi, Riley J Hickman, Loïc M Roch, Melodie Christensen, Elena Liles, Jason E Hein, and Alán Aspuru-Guzik. Olympus: a benchmarking framework for noisy optimization and experiment planning. *Machine Learning: Science and Technology*, 2(3):035021, jul 2021. doi: 10.1088/2632-2153/abedc8. URL https://dx.doi.org/10.1088/2632-2153/abedc8.

Michael Janner, Yilun Du, Joshua Tenenbaum, and Sergey Levine. Planning with diffusion for flexible behavior synthesis. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 9902–9915. PMLR, 17–23 Jul 2022. URL https://proceedings.mlr.press/v162/janner22a.html.

Ashish K Jayant and Shalabh Bhatnagar. Model-based safe deep reinforcement learning via a constrained proximal policy optimization algorithm. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 24432–24445. Curran Associates, Inc., 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/9a8eb202c060b7d81f5889631cbcd47e-Paper-Conference.pdf.

Nan Jiang, Maosen Zhang, Willem-Jan van Hoeve, and Yexiang Xue. Constraint reasoning embedded structured prediction. *Journal of Machine Learning Research*, 23(345):1–40, 2022. URL http://jmlr.org/papers/v23/21-1484.html.

Herve Jégou, Matthijs Douze, and Cordelia Schmid. Product quantization for nearest neighbor search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(1):117–128, 2011. doi: 10.1109/TPAMI.2010.57.

Kirthevasan Kandasamy, Gautam Dasarathy, Junier Oliva, Jeff Schneider, and Barnabás Póczos. Gaussian process bandit optimisation with multi-fidelity evaluations. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'16, page 1000–1008, Red Hook, NY, USA, 2016. Curran Associates Inc. ISBN 9781510838819.

George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021. doi: 10.1038/s42254-021-00314-5. URL https://www.nature.com/articles/s42254-021-00314-5.

Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4396–4405, 2019. doi: 10.1109/CVPR.2019.00453.

Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and Improving the Image Quality of StyleGAN . In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8107–8116, Los Alamitos, CA, USA, June 2020. IEEE Computer Society. doi: 10.1109/CVPR42600.2020.00813. URL https://doi.ieeecomputersociety.org/10.1109/CVPR42600.2020.00813.

Gul Rukh Khattak, Sofia Vallecorsa, Federico Carminati, and Gul Muhammad Khan. Fast simulation of a high granularity calorimeter by generative adversarial networks. *The European Physical Journal C*, 82(4):386, 2022.

Hyunjik Kim and Andriy Mnih. Disentangling by factorising. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 2649–2658. PMLR, 10–15 Jul 2018. URL https://proceedings.mlr.press/v80/kim18b.html.

Hyunjik Kim, Andriy Mnih, Jonathan Schwarz, Marta Garnelo, Ali Eslami, Dan Rosenbaum, Oriol Vinyals, and Yee Whye Teh. Attentive neural processes. In *International Conference on Learning Representations*, 2019a. URL https://openreview.net/forum?id=SkE6PjC9KX.

Kwanyoung Kim, Dongwon Park, Kwang In Kim, and Se Young Chun. Task-aware variational adversarial active learning. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8162–8171, 2020. URL https://api.semanticscholar.org/CorpusID:211082697.

Yoon Kim, Chris Dyer, and Alexander Rush. Compound probabilistic context-free grammars for grammar induction. In Anna Korhonen, David Traum, and Lluís Màrquez, editors, *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2369–2385, Florence, Italy, July 2019b. Association for Computational Linguistics. doi: 10.18653/v1/P19-1228. URL https://aclanthology.org/P19-1228/.

Andreas Kirsch, Joost van Amersfoort, and Yarin Gal. Batchbald: Efficient and diverse batch acquisition for deep bayesian active learning. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://papers.neurips.cc/paper/8925-batchbald-efficient-and-diverse-batch-acquisition-for-deep-bayesian-active-learning.pdf.

Pascal Klink, Carlo D' Eramo, Jan R Peters, and Joni Pajarinen. Self-paced deep reinforcement learning. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 9216–9227. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/68a9750337a418a86fe06c1991a1d64c-Paper.pdf.

Brent A. Koscher, Richard B. Canty, Matthew A. McDonald, Kevin P. Greenman, Charles J. McGill, Camille L. Bilodeau, Wengong Jin, Haoyang Wu, Florence H. Vermeire, Brooke Jin, Travis Hart, Timothy Kulesza, Shih-Cheng Li, Tommi S. Jaakkola, Regina Barzilay, Rafael GÃ³mez-Bombarelli, William H. Green, and Klavs F. Jensen. Autonomous, multiproperty-driven molecular discovery: From predictions to measurements and back. *Science*, 382(6677), December 2023. ISSN 1095-9203. doi: 10.1126/science.adi1407. URL http://dx.doi.org/10.1126/science.adi1407.

Claudius Krause and David Shih. Fast and accurate simulations of calorimeter showers with normalizing flows. *Phys. Rev. D*, 107:113003, Jun 2023. doi: 10.1103/PhysRevD.107.113003. URL https://link.aps.org/doi/10.1103/PhysRevD.107.113003.

Yufei Kuang, Jie Wang, Haoyang Liu, Fangzhou Zhu, Xijun Li, Jia Zeng, Jianye HAO, Bin Li, and Feng Wu. Rethinking branching on exact combinatorial optimization solver: The first deep symbolic discovery framework. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=jKhNBulNMh.

Matt J. Kusner, Brooks Paige, and José Miguel Hernández-Lobato. Grammar variational autoencoder. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 1945–1954. PMLR, 06–11 Aug 2017. URL https://proceedings.mlr.press/v70/kusner17a.html.

Jeffrey N. Law, Shubham Pandey, Prashun Gorai, and Peter C. St John. Upper-bound energy minimization to search for stable functional materials with graph neural networks. *JACS Au*, 3(1):113–123, 2023. doi: 10.1021/jacsau.2c00540.

Kookjin Lee and Kevin T. Carlberg. Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders. *J. Comput. Phys.*, 404(C), March 2020. ISSN 0021-9991. doi: 10.1016/j.jcp.2019.108973. URL `https://doi.org/10.1016/j.jcp.2019.108973`.

Junnan Li, Richard Socher, and Steven C.H. Hoi. Dividemix: Learning with noisy labels as semi-supervised learning. In *International Conference on Learning Representations*, 2020a. URL `https://openreview.net/forum?id=HJgExaVtwr`.

Rongjie Li, Songyang Zhang, and Xuming He. Sgtr: End-to-end scene graph generation with transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 19486–19496, June 2022a.

Shibo Li, Wei Xing, Robert M. Kirby, and Shandian Zhe. Multi-fidelity bayesian optimization via deep neural networks. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA, 2020b. Curran Associates Inc. ISBN 9781713829546.

Shibo Li, Zheng Wang, Robert Kirby, and Shandian Zhe. Deep multi-fidelity active learning of high-dimensional outputs. In Gustau Camps-Valls, Francisco J. R. Ruiz, and Isabel Valera, editors, *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pages 1694–1711. PMLR, 28–30 Mar 2022b. URL `https://proceedings.mlr.press/v151/li22b.html`.

Zongyi Li, Hongkai Zheng, Nikola Kovachki, David Jin, Haoxuan Chen, Burigede Liu, Kamyar Azizzadenesheli, and Anima Anandkumar. Physics-informed neural operator for learning partial differential equations. *ACM / IMS J. Data Sci.*, 1(3), May 2024. doi: 10.1145/3648506. URL `https://doi.org/10.1145/3648506`.

Haomiao Liu, Ruiping Wang, Shiguang Shan, and Xilin Chen. Deep supervised hashing for fast image retrieval. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2064–2072, 2016. doi: 10.1109/CVPR.2016.227.

Yongshuai Liu, Jiaxin Ding, and Xin Liu. Ipo: Interior-point policy optimization under constraints. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 4940–4947, 2020.

Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 9992–10002, 2021. doi: 10.1109/ICCV48922.2021.00986.

Andre KY Low, Flore Mekki-Berrada, Abhishek Gupta, Aleksandr Ostudin, Jiaxun Xie, Eleonore Vissol-Gaudin, Yee-Fun Lim, Qianxiao Li, Yew Soon Ong, Saif A Khan, et al. Evolution-guided bayesian optimization for constrained multi-objective optimization in self-driving labs. *npj Computational Materials*, 10(1):104, 2024.

Lu Lu, Pengzhan Jin, and George Em Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, 2021. doi: 10.1038/s42256-021-00302-5.

Andreas Lugmayr, Martin Danelljan, Andres Romero, Fisher Yu, Radu Timofte, and Luc Van Gool. Repaint: Inpainting using denoising diffusion probabilistic models. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11451–11461, 2022. doi: 10.1109/CVPR52688.2022.01117.

Yingying Ma, Pengcheng Xu, Minjie Li, Xiaobo Ji, Wenyue Zhao, and Wencong Lu. The mastery of details in the workflow of materials machine learning. *npj Computational Materials*, 10(1), July 2024. ISSN 2057-3960. doi: 10.1038/s41524-024-01331-5. URL `http://dx.doi.org/10.1038/s41524-024-01331-5`.

Nikku Madhusudhan. Exoplanetary atmospheres: Key insights, challenges, and prospects. *Annual Review of Astronomy and Astrophysics*, 57(1):617–663, 2019. doi: 10.1146/annurev-astro-081817-051846.

Robin Manhaeve, Sebastijan Dumancic, Angelika Kimmig, Thomas Demeester, and Luc De Raedt. Deepproblog: Neural probabilistic logic programming. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL `https://proceedings.neurips.cc/paper_files/paper/2018/file/dc5d637ed5e62c36ecb73b654b05ba2a-Paper.pdf`.

Zhiping Mao and Xuhui Meng. Physics-informed neural networks with residual/gradient-based adaptive sampling methods for solving partial differential equations with sharp solutions. *Applied Mathematics and Mechanics (English Edition)*, 44(7):1069–1084, 2023. doi: 10.1007/s10483-023-2994-7.

David Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman, 1982.

Pierre-Alexandre Mattei and Jes Frellsen. MIWAE: Deep generative modelling and imputation of incomplete data sets. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 4413–4423. PMLR, 09–15 Jun 2019. URL https://proceedings.mlr.press/v97/mattei19a.html.

Natalie Maus, Haydn Jones, Juston Moore, Matt J Kusner, John Bradshaw, and Jacob Gardner. Local latent space bayesian optimization over structured inputs. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 34505–34518. Curran Associates, Inc., 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/ded98d28f82342a39f371c013dfb3058-Paper-Conference.pdf.

S. Hessam M. Mehr, Matthew Craven, Artem I. Leonov, Graham Keenan, and Leroy Cronin. A universal system for digitization and automatic execution of the chemical synthesis literature. *Science*, 370(6512):101–108, 2020. doi: 10.1126/science.abc2986. URL https://www.science.org/doi/abs/10.1126/science.abc2986.

John Mern, Anil Yildiz, Zachary Sunberg, Tapan Mukerji, and Mykel J. Kochenderfer. Bayesian optimized monte carlo planning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(13):11880–11887, May 2021. doi: 10.1609/aaai.v35i13.17411. URL https://ojs.aaai.org/index.php/AAAI/article/view/17411.

Philipp Moritz, Robert Nishihara, Stephanie Wang, Alexey Tumanov, Richard Liaw, Eric Liang, Melih Elibol, Zongheng Yang, William Paul, Michael I. Jordan, and Ion Stoica. Ray: a distributed framework for emerging ai applications. In *Proceedings of the 13th USENIX Conference on Operating Systems Design and Implementation*, OSDI'18, page 561–577, USA, 2018. USENIX Association. ISBN 9781931971478.

Byeonghu Na, Yeongmin Kim, Minsang Park, Donghyeok Shin, Wanmo Kang, and Il-Chul Moon. Diffusion rejection sampling. In *Proceedings of the 41st International Conference on Machine Learning*, ICML'24. JMLR.org, 2024.

Ofir Nachum, Shixiang Gu, Honglak Lee, and Sergey Levine. Data-efficient hierarchical reinforcement learning. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 3307–3317, Red Hook, NY, USA, 2018. Curran Associates Inc.

OpenAI. Openai api platform documentation. https://platform.openai.com/docs/api-reference, 2025. Accessed 2025-08-15.

Kushagra Pandey, Jaideep Pathak, Yilun Xu, Stephan Mandt, Michael Pritchard, Arash Vahdat, and Morteza Mardani. Heavy-tailed diffusion models. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=tozlOEN4qp.

George Papamakarios, Theo Pavlakou, and Iain Murray. Masked autoregressive flow for density estimation. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/file/6c1da886822c67822bcf3679d04369fa-Paper.pdf.

George Papamakarios, David Sterratt, and Iain Murray. Sequential neural likelihood: Fast likelihood-free inference with autoregressive flows. In Kamalika Chaudhuri and Masashi Sugiyama, editors, *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pages 837–848. PMLR, 16–18 Apr 2019. URL https://proceedings.mlr.press/v89/papamakarios19a.html.

Jinwoo Park, Jaehyeong Park, Youngmok Jung, Hwijoon Lim, Hyunho Yeo, and Dongsu Han. Topfull: An adaptive top-down overload control for slo-oriented microservices. In *Proceedings of the ACM SIGCOMM 2024 Conference*, ACM SIGCOMM '24, page 876–890, New York, NY, USA, 2024a. Association for Computing Machinery. ISBN 9798400706141. doi: 10.1145/3651890.3672253. URL https://doi.org/10.1145/3651890.3672253.

Kanghee Park, Jiayu Wang, Taylor Berg-Kirkpatrick, Nadia Polikarpova, and Loris D' Antoni. Grammar-aligned decoding. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, editors, *Advances in Neural Information Processing Systems*, volume 37, pages 24547–24568. Curran Associates, Inc., 2024b. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/2bdc2267c3d7d01523e2e17ac0a754f3-Paper-Conference.pdf.

Amin Parvaneh, Ehsan Abbasnejad, Damien Teney, Reza Haffari, Anton Van Den Hengel, and Javen Qinfeng Shi. Active learning by feature mixing. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12227–12236, 2022. doi: 10.1109/CVPR52688.2022.01192.

Max B. Paulus and Andreas Krause. Learning to dive in branch and bound. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, Red Hook, NY, USA, 2023. Curran Associates Inc.

Nick Pawlowski, Daniel Coelho de Castro, and Ben Glocker. Deep structural causal models for tractable counterfactual inference. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 857–869. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/0987b8b338d6c90bbedd8631bc499221-Paper.pdf.

Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, September 2009. ISBN 9780521749190. doi: 10.1017/cbo9780511803161. URL http://dx.doi.org/10.1017/CBO9780511803161.

Karl Pertsch, Oleh Rybkin, Frederik Ebert, Shenghao Zhou, Dinesh Jayaraman, Chelsea Finn, and Sergey Levine. Long-horizon visual planning with goal-conditioned hierarchical predictors. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 17321–17333. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/c8d3a760ebab631565f8509d84b3b3f1-Paper.pdf.

Brenden Petersen, Larma K., Mundhenk Mikel Landajuela, Santiago T. Nathan, P. Claudio, Soo Kim, Kim K., and T. Joanne. Deep symbolic regression: Recovering mathematical expressions from data via risk-seeking policy gradients. *Arxiv:1912.04871 Cs, Stat*, 2021.

Ben Poole, Ajay Jain, Jonathan T. Barron, and Ben Mildenhall. Dreamfusion: Text-to-3d using 2d diffusion. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=FjNys5c7VyY.

Shrimai Prabhumoye, Alan W Black, and Ruslan Salakhutdinov. Exploring controllable text generation techniques. In Donia Scott, Nuria Bel, and Chengqing Zong, editors, *Proceedings of the 28th International Conference on Computational Linguistics*, pages 1–14, Barcelona, Spain (Online), December 2020. International Committee on Computational Linguistics. doi: 10.18653/v1/2020.coling-main.1. URL https://aclanthology.org/2020.coling-main.1/.

Yunzhe Qi, Yikun Ban, Tianxin Wei, Jiaru Zou, Huaxiu Yao, and Jingrui He. Meta-learning with neural bandit scheduler. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, volume 36, pages 63994–64028. Curran Associates, Inc., 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/c9e6ac15e689e06139d7b39e1667b165-Paper-Conference.pdf.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 8748–8763. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/radford21a.html.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140):1–67, 2020. URL http://jmlr.org/papers/v21/20-074.html.

Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

Jacob T. Rapp, Bennett J. Bremer, and Philip A. Romero. Self-driving laboratories to autonomously navigate the protein fitness landscape. *Nature Chemical Engineering*, 1(1):97–107, January 2024. ISSN 2948-1198. doi: 10.1038/s44286-023-00002-4. URL http://dx.doi.org/10.1038/s44286-023-00002-4.

Ali Razavi, Aaron van den Oord, and Oriol Vinyals. Generating diverse high-fidelity images with vq-vae-2. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/file/5f8e2fa1718d1bbcadf1cd9c7a54fb8c-Paper.pdf.

Chandan K Reddy and Parshin Shojaee. Towards scientific discovery with generative ai: Progress, opportunities, and challenges. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(27):28601–28609, April 2025. ISSN 2159-5399. doi: 10.1609/aaai.v39i27.35084. URL http://dx.doi.org/10.1609/aaai.v39i27.35084.

Nils Reimers and Iryna Gurevych. Sentence-BERT: Sentence embeddings using siamese BERT-networks. In *Proceedings of EMNLP*, 2019. URL https://arxiv.org/abs/1908.10084.

Stephen Robertson and Hugo Zaragoza. The probabilistic relevance framework: BM25 and beyond. *Foundations and Trends in Information Retrieval*, 3(4):333–389, 2009. doi: 10.1561/1500000019.

Tim Rocktäschel and Sebastian Riedel. End-to-end differentiable proving. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/file/b2ab001909a8a6f04b51920306046ce5-Paper.pdf.

Salva Rühling Cachay, Benedikt Boecking, and Artur Dubrawski. End-to-end weak supervision. *Advances in Neural Information Processing Systems*, 34:1845–1857, 2021.

Krzysztof Rzadca, Pawel Findeisen, Jacek Swiderski, Przemyslaw Zych, Przemyslaw Broniek, Jarek Kusmierek, Pawel Nowak, Beata Strack, Piotr Witusowski, Steven Hand, and John Wilkes. Autopilot: workload autoscaling at google. In *Proceedings of the Fifteenth European Conference on Computer Systems*, EuroSys '20, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450368827. doi: 10.1145/3342195.3387524. URL https://doi.org/10.1145/3342195.3387524.

Axel Sauer and Andreas Geiger. Counterfactual generative networks. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=BXewfAYMmJw.

Samuel S Schoenholz and Ekin D Cubuk. Jax, m.d. a framework for differentiable physics*. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(12):124016, dec 2021. doi: 10.1088/1742-5468/ac3ae9. URL https://dx.doi.org/10.1088/1742-5468/ac3ae9.

John Schulman, Sergey Levine, Pieter Abbeel, Michael Jordan, and Philipp Moritz. Trust region policy optimization. In Francis Bach and David Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 1889–1897, Lille, France, 07–09 Jul 2015. PMLR. URL https://proceedings.mlr.press/v37/schulman15.html.

Daniel Selsam, Matthew Lamm, Benedikt Bünz, Percy Liang, Leonardo de Moura, and David L. Dill. Learning a SAT solver from single-bit supervision. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019. URL https://openreview.net/forum?id=HJMC_iA5tm.

Shreya Shankar, Labib Fawaz, Karl Gyllstrom, and Aditya Parameswaran. Automatic and precise data validation for machine learning. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, CIKM '23, page 2198–2207, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9798400701245. doi: 10.1145/3583780.3614786. URL https://doi.org/10.1145/3583780.3614786.

C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, July 1948. ISSN 0005-8580. doi: 10.1002/j.1538-7305.1948.tb01338.x. URL http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x.

Yujun Shen and Bolei Zhou. Closed-form factorization of latent semantics in gans. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1532–1540, 2021. doi: 10.1109/CVPR46437.2021.00158.

Xiaoming Shi, Siqiao Xue, Kangrui Wang, Fan Zhou, James Y. Zhang, Jun Zhou, Chenhao Tan, and Hongyuan Mei. Language models can improve event prediction by few-shot abductive reasoning. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, Red Hook, NY, USA, 2023. Curran Associates Inc.

Ashish Shrivastava, Tomas Pfister, Oncel Tuzel, Joshua Susskind, Wenda Wang, and Russell Webb. Learning from simulated and unsupervised images through adversarial training. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2107–2116, 2017.

David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, Timothy Lillicrap, Karen Simonyan, and Demis Hassabis. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 362(6419):1140–1144, 2018. doi: 10.1126/science.aar6404. URL https://www.science.org/doi/abs/10.1126/science.aar6404.

Herbert A. Simon. A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118, 1955. doi: 10.2307/1884852.

Michael J. Smith and James E. Geach. Astronomia ex machina: a history, primer and outlook on neural networks in astronomy. *Royal Society Open Science*, 10(5), May 2023. ISSN 2054-5703. doi: 10.1098/rsos.221454. URL `http://dx.doi.org/10.1098/rsos.221454`.

Ryan Smith, Jason A. Fries, Braden Hancock, and Stephen H. Bach. Language models in the loop: Incorporating prompting into weak supervision. *ACM / IMS J. Data Sci.*, 1(2), April 2024. doi: 10.1145/3617130. URL `https://doi.org/10.1145/3617130`.

Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012. URL `https://proceedings.neurips.cc/paper_files/paper/2012/file/05311655a15b75fab86956663e1819cd-Paper.pdf`.

Kihyuk Sohn, David Berthelot, Chun-Liang Li, Zizhao Zhang, Nicholas Carlini, Ekin D. Cubuk, Alex Kurakin, Han Zhang, and Colin Raffel. Fixmatch: simplifying semi-supervised learning with consistency and confidence. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA, 2020. Curran Associates Inc. ISBN 9781713829546.

Stian Soiland-Reyes, Peter Sefton, Mercè Crosas, Leyla Jael Castro, Frederik Coppens, José M. Fernández, Daniel Garijo, Björn Grüning, Marco La Rosa, Simone Leo, Eoghan Ó Carragáin, Marc Portier, Ana Trisovic, RO-Crate Community, Paul Groth, and Carole Goble. Packaging research artefacts with ro-crate. *Data Science*, 5(2):97–138, 2022. doi: 10.3233/DS-210053. URL `https://doi.org/10.3233/DS-210053`.

Xingyou Song, Sagi Perel, Chansoo Lee, Greg Kochanski, and Daniel Golovin. Open source vizier: Distributed infrastructure and api for reliable and flexible blackbox optimization. In Isabelle Guyon, Marius Lindauer, Mihaela van der Schaar, Frank Hutter, and Roman Garnett, editors, *Proceedings of the First International Conference on Automated Machine Learning*, volume 188 of *Proceedings of Machine Learning Research*, pages 8/1–17. PMLR, 25–27 Jul 2022. URL `https://proceedings.mlr.press/v188/song22a.html`.

Baochen Sun and Kate Saenko. Deep coral: Correlation alignment for deep domain adaptation. In Gang Hua and Hervé Jégou, editors, *Computer Vision – ECCV 2016 Workshops*, pages 443–450, Cham, 2016. Springer International Publishing. ISBN 978-3-319-49409-8.

Tianxiang Sun, Yunfan Shao, Hong Qian, Xuanjing Huang, and Xipeng Qiu. Black-box tuning for language-model-as-a-service. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 20841–20855. PMLR, 17–23 Jul 2022. URL `https://proceedings.mlr.press/v162/sun22e.html`.

John Sweller. *Cognitive Load Theory*, pages 37–76. Elsevier, 2011. doi: 10.1016/b978-0-12-387691-1.00002-8. URL `http://dx.doi.org/10.1016/B978-0-12-387691-1.00002-8`.

Nathan J. Szymanski, Bernardus Rendy, Yuxing Fei, Rishi E. Kumar, Tanjin He, David Milsted, Matthew J. McDermott, Max Gallant, Ekin Dogus Cubuk, Amil Merchant, Haegyeom Kim, Anubhav Jain, Christopher J. Bartel, Kristin Persson, Yan Zeng, and Gerbrand Ceder. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature*, 624(7990):86–91, November 2023a. ISSN 1476-4687. doi: 10.1038/s41586-023-06734-w. URL `http://dx.doi.org/10.1038/s41586-023-06734-w`.

Nathan J Szymanski, Bernardus Rendy, Yuxing Fei, Rishi E Kumar, Tanjin He, David Milsted, Matthew J McDermott, Max Gallant, Ekin Dogus Cubuk, Amil Merchant, et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature*, 624(7990):86–91, 2023b.

Aviv Tamar, Yinlam Chow, Mohammad Ghavamzadeh, and Shie Mannor. Policy gradient for coherent risk measures. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015. URL `https://proceedings.neurips.cc/paper_files/paper/2015/file/024d7f84fff11dd7e8d9c510137a2381-Paper.pdf`.

Yusuke Tashiro, Jiaming Song, Yang Song, and Stefano Ermon. Csdi: conditional score-based diffusion models for probabilistic time series imputation. In *Proceedings of the 35th International Conference on Neural Information Processing Systems*, NIPS '21, Red Hook, NY, USA, 2021. Curran Associates Inc. ISBN 9781713845393.

Bedir Tekinerdogan. Ai-sos: A strategic framework for integrating artificial intelligence in system of systems. In *2024 19th Annual System of Systems Engineering Conference (SoSE)*, pages 57–64. IEEE, June 2024. doi: 10.1109/sose62659.2024.10620943. URL `http://dx.doi.org/10.1109/SOSE62659.2024.10620943`.

Gerald Tesauro, V T Rajan, and Richard Segal. Bayesian inference in monte-carlo tree search. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, UAI'10, page 580–588, Arlington, Virginia, USA, 2010. AUAI Press. ISBN 9780974903965.

Megh Thakkar, Tolga Bolukbasi, Sriram Ganapathy, Shikhar Vashishth, Sarath Chandar, and Partha Talukdar. Self-influence guided data reweighting for language model pre-training. In *The 2023 Conference on Empirical Methods in Natural Language Processing*, 2023. URL https://openreview.net/forum?id=rXn9WO4M2p.

Gary Tom, Stefan P. Schmid, Sterling G. Baird, Yang Cao, et al. Self-driving laboratories for chemistry and materials science. *Chemical Reviews*, 124(16):9633–9732, 2024. doi: 10.1021/acs.chemrev.4c00055.

Brandon Trabucco, Kyle Doherty, Max A Gurinas, and Ruslan Salakhutdinov. Effective data augmentation with diffusion models. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=ZWzUA9zeAg.

Arash Vahdat and Jan Kautz. Nvae: A deep hierarchical variational autoencoder. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 19667–19679. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/e3b21256183cf7c2c7a66be163579d37-Paper.pdf.

Aaron van den Oord, Oriol Vinyals, and koray kavukcuoglu. Neural discrete representation learning. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/file/7a98af17e63a0ac09ce2e96d03992fbc-Paper.pdf.

Malavika Vasist, François Rozet, Olivier Absil, Paul Mollière, Egor Nasedkin, and Gilles Louppe. Neural posterior estimation for exoplanetary atmospheric retrieval. *Astronomy & Astrophysics*, 672:A147, 2023. doi: 10.1051/0004-6361/202245263.

Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th International Conference on Machine Learning*, ICML '08, page 1096–1103, New York, NY, USA, 2008. Association for Computing Machinery. ISBN 9781605582054. doi: 10.1145/1390156.1390294. URL https://doi.org/10.1145/1390156.1390294.

Martin J. Wainwright and Michael I. Jordan. 2008. doi: 10.1561/2200000001.

Chengshi Wang, Yeon-Ju Kim, Aikaterini Vriza, Rohit Batra, Arun Baskaran, Naisong Shan, Nan Li, Pierre Darancet, Logan Ward, Yuzi Liu, Maria K. Y. Chan, Subramanian K.R.S. Sankaranarayanan, H. Christopher Fry, C. Suzanne Miller, Henry Chan, and Jie Xu. Autonomous platform for solution processing of electronic polymers. *Nature Communications*, 16(1), February 2025. ISSN 2041-1723. doi: 10.1038/s41467-024-55655-3. URL http://dx.doi.org/10.1038/s41467-024-55655-3.

Hanchen Wang, Tianfan Fu, Yuanqi Du, Wenhao Gao, Kexin Huang, Ziming Liu, Payal Chandak, Shengchao Liu, Peter Van Katwyk, Andreea Deac, Anima Anandkumar, Karianne Bergen, Carla P. Gomes, Shirley Ho, Pushmeet Kohli, Joan Lasenby, Jure Leskovec, Tie-Yan Liu, Arjun Manrai, Debora Marks, Bharath Ramsundar, Le Song, Jimeng Sun, Jian Tang, Petar Veličković, Max Welling, Linfeng Zhang, Connor W. Coley, Yoshua Bengio, and Marinka Zitnik. Scientific discovery in the age of artificial intelligence. *Nature*, 620(7972):47–60, August 2023a. ISSN 1476-4687. doi: 10.1038/s41586-023-06221-2. URL http://dx.doi.org/10.1038/s41586-023-06221-2.

Ke Alexander Wang, Geoff Pleiss, Jacob R. Gardner, Stephen Tyree, Kilian Q. Weinberger, and Andrew Gordon Wilson. Exact gaussian processes on a million data points. In *Advances in Neural Information Processing Systems*, volume 32, 2019a. URL https://proceedings.neurips.cc/paper/9606-exact-gaussian-processes-on-a-million-data-points.pdf.

Po-Wei Wang, Priya Donti, Bryan Wilder, and Zico Kolter. SATNet: Bridging deep learning and logical reasoning using a differentiable satisfiability solver. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 6545–6554. PMLR, 09–15 Jun 2019b. URL https://proceedings.mlr.press/v97/wang19e.html.

Weiran Wang, Xinchen Yan, Honglak Lee, and Karen Livescu. Deep variational canonical correlation analysis, 2017. URL https://openreview.net/forum?id=H1Heentlx.

Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. Self-instruct: Aligning language models with self-generated instructions. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 13484–13508, Toronto, Canada, July 2023b. Association for Computational Linguistics. doi: 10.18653/v1/2023.acl-long.754. URL https://aclanthology.org/2023.acl-long.754/.

Zi Wang and Stefanie Jegelka. Max-value entropy search for efficient Bayesian optimization. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 3627–3635. PMLR, 06–11 Aug 2017. URL https://proceedings.mlr.press/v70/wang17e.html.

Luis Welbanks and Nikku Madhusudhan. On degeneracies in retrievals of exoplanetary transmission spectra. *The Astronomical Journal*, 157(5):206, 2019. doi: 10.3847/1538-3881/ab14de.

Adina Williams, Nikita Nangia, and Samuel R. Bowman. A broad-coverage challenge corpus for sentence understanding through inference. In *Proceedings of NAACL-HLT*, 2018. URL https://aclanthology.org/N18-1101/.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowitz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. Transformers: State-of-the-art natural language processing. In *Proceedings of EMNLP (Demos)*, pages 38–45, 2020. URL https://aclanthology.org/2020.emnlp-demos.6/.

Keyi Wu, Thomas O'Leary-Roseberry, Peng Chen, and Omar Ghattas. Large-scale bayesian optimal experimental design with derivative-informed projected neural network. *J. Sci. Comput.*, 95(1), March 2023. ISSN 0885-7474. doi: 10.1007/s10915-023-02145-1. URL https://doi.org/10.1007/s10915-023-02145-1.

Tailin Wu, Takashi Maruyama, and Jure Leskovec. Learning to accelerate partial differential equations via latent global evolution. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 2240–2253. Curran Associates, Inc., 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/0f817dcbad81afb21fb695f1b2e55e44-Paper-Conference.pdf.

Wei Xiao, Tsun-Hsuan Wang, Ramin Hasani, Makram Chahine, Alexander Amini, Xiao Li, and Daniela Rus. Barriernet: Differentiable control barrier functions for learning of safe robot control. *IEEE Transactions on Robotics*, 39(3):2289–2307, 2023. doi: 10.1109/TRO.2023.3249564.

W.W. Xing, A.A. Shah, P. Wang, S. Zhe, Q. Fu, and R.M. Kirby. Residual gaussian process: A tractable nonparametric bayesian emulator for multi-fidelity simulations. *Applied Mathematical Modelling*, 97: 36–56, 2021. ISSN 0307-904X. doi: https://doi.org/10.1016/j.apm.2021.03.041. URL https://www.sciencedirect.com/science/article/pii/S0307904X21001724.

Peng Xu, Xiatian Zhu, and David A. Clifton. Multimodal learning with transformers: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(10):12113–12132, October 2023. ISSN 1939-3539. doi: 10.1109/tpami.2023.3275156. URL http://dx.doi.org/10.1109/TPAMI.2023.3275156.

Zongheng Yang, Zhanghao Wu, Michael Luo, Wei-Lin Chiang, Romil Bhardwaj, Woosuk Kwon, Siyuan Zhuang, Frank Sifei Luan, Gautam Mittal, Scott Shenker, and Ion Stoica. SkyPilot: An intercloud broker for sky computing. In *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*, pages 437–455, Boston, MA, April 2023. USENIX Association. ISBN 978-1-939133-33-5. URL https://www.usenix.org/conference/nsdi23/presentation/yang-zongheng.

Daochen Zha, Zaid Pervaiz Bhat, Kwei-Herng Lai, Fan Yang, Zhimeng Jiang, Shaochen Zhong, and Xia Hu. Data-centric artificial intelligence: A survey. *ACM Computing Surveys*, 57(5):1–42, January 2025. ISSN 1557-7341. doi: 10.1145/3711118. URL http://dx.doi.org/10.1145/3711118.

Chao Zhang, Fangbo Tao, Xiusi Chen, Jiaming Shen, Meng Jiang, Brian Sadler, Michelle Vanni, and Jiawei Han. Taxogen: Unsupervised topic taxonomy construction by adaptive term embedding and clustering. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '18, page 2701–2709, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450355520. doi: 10.1145/3219819.3220064. URL https://doi.org/10.1145/3219819.3220064.

Zhen Zhang, Mohammed Haroon Dupty, Fan Wu, Javen Qinfeng Shi, and Wee Sun Lee. Factor graph neural networks. *Journal of Machine Learning Research*, 24(181):1–54, 2023. URL http://jmlr.org/papers/v24/21-0434.html.

Yuxuan Zhao and Madeleine Udell. Matrix completion with quantified uncertainty through low rank gaussian copula. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 20977–20988. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/f076073b2082f8741a9cd07b789c77a0-Paper.pdf.

Xun Zheng, Bryon Aragam, Pradeep K Ravikumar, and Eric P Xing. Dags with no tears: Continuous optimization for structure learning. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL https://proceedings.neurips.cc/paper_files/paper/2018/file/e347c51419ffb23ca3fd5050202f9c3d-Paper.pdf.

A Automated audits for coverage

We complement the theory-guided construction with two auditable corpus audits: a tension-coverage audit over natural science literature and a function-coverage audit over recent AI papers. Both audits store per-item evidence and scores for inspection.

A.1 Tension coverage (natural-science abstracts)

We sampled $N = 3{,}000$ journal-article abstracts (2021–2025) across six domains (Physics, Chemistry, Materials, Ecology, Astronomy, Genomics) via the Crossref REST API [Crossref, 2025], targeting 500 per domain with DOI-based deduplication. We relied on deposited abstracts; when a top venue lacked abstracts in Crossref, we expanded the venue list within the domain until the quota was met. For each abstract, an LLM via the OpenAI API (model: `gpt-5-mini`) generated 1–3 latent "discovery bottleneck" hypotheses at the scientific-content layer (excluding workflow frictions); responses were cached for reproducibility [OpenAI, 2025]. We split abstracts into sentences, retrieved top evidence sentences for each hypothesis using Okapi BM25 [Robertson and Zaragoza, 2009], and computed textual entailment with a cross-encoder NLI implemented as `cross-encoder/nli-deberta-v3-base` [He et al., 2021], via Sentence-Transformers and HuggingFace Transformers [Reimers and Gurevych, 2019, Wolf et al., 2020], using MNLI-style templates [Williams et al., 2018]. Decisions were evidence-aware: BM25 provided gating and evidence-driven promotion; NLI scores dominated fusion. To emphasize coverage we used a no-abstention Top-2 policy, retaining up to two tension labels per abstract. Top-2 coverage by the four tensions is 100% (OTHER = 0). The dominant co-occurrence is *Complexity* + *Scarcity*; *Explosion* appears primarily as secondary. This supports the collective exhaustiveness (at coarse granularity) of the four scientific-discovery tensions across domains.

A.2 Function coverage (recent AI papers)

We stratified arXiv (2019–2025) by subfield $\times$ year and sampled $N = 628$ papers. PDFs were fetched via canonical arXiv URLs with retries and checksums, then parsed to plain text. We mined candidate passages primarily from *Methods/Contributions/Evaluation/Ablations*, guided by high-yield heuristics (section headers; "we propose/introduce ... to ..."; pipeline/agent/tool-use phrases; benchmark mentions). For each of six AI functions (Represent; Reason & Infer; Optimize & Control; Simulate & Emulate; Generate & Create; Autonomize & Orchestrate), we maintained a curated, high-precision lexicon and retrieved top-k passages per function with Okapi BM25 [Robertson and Zaragoza, 2009]. We then scored *function–passage* entailment using a DeBERTa-v3 cross-encoder (`cross-encoder/nli-deberta-v3-base`) implemented with Sentence-Transformers and HuggingFace Transformers [He et al., 2021, Reimers and Gurevych, 2019, Wolf et al., 2020], under MNLI-style hypotheses [Williams et al., 2018]. A deterministic, evidence-aware rule accepted a function if entailment $\geq \theta$ with adequate BM25 support; promoted strong-evidence cases (high BM25 with a clear NLI margin gap); and emitted Top-2 when the top labels were within a small margin. *OTHER* was reserved for genuine coverage failures (no function passed gates). Each assignment logged the winning passage, BM25/NLI scores, and gate/override flags. The Coverage is 98.9% (OTHER = 1.11%). Frequent Top-2 pairs combine *Autonomize & Orchestrate* with a core algorithmic function (e.g., *Reason*, *Simulate*, or *Optimize*), reflecting typical paper structure (pipeline plus capability). This indicates a near-complete coverage of the reported capabilities by the six-function taxonomy. Retrieval noise can cause OTHER; proportions reflect the mining policy and are not population frequencies. The thresholds were fixed a priori and stable under small perturbations.

B Intrinsic axes and atomic triads

Our compass treats each core AI function as a two-dimensional conceptual space spanned by two intrinsic binary axes. An axis is adopted only if it (i) follows from first principles, (ii) recurs across disciplines, and (iii) captures the observed variability of the function. Crossing the axes yields four theoretical quadrants. In every case, one quadrant is either logically void or operationally redundant; removing the void cell or merging indistinguishable cells produces a triad of mutually exclusive and collectively exhaustive classes. We call these atomic categories in the lattice-theoretic sense [Birkhoff, 1940, Davey and Priestley, 2002]: minimal under the chosen axes, and any higher-level construct (a problem-centered strategy or an algorithmic innovation) can be expressed as a join of these atoms. Because the axes are intrinsic and the atoms are minimal, the taxonomy is stable as techniques evolve and provides the scaffold for the strategy layer.

Representation. Axes: information carrier (discrete symbols vs continuous vectors) and uncertainty handling (deterministic vs stochastic). In common learning pipelines, discrete tokens are embedded into continuous logits before probabilistic handling, making the discrete–stochastic and continuous–stochastic quadrants operationally indistinguishable; we thus merge them [Shannon, 1948, Bishop, 2006]. Triad: P1 Symbolic/Structured, P2 Probabilistic, P3 Latent-manifold.

Reason & Infer. Axes: validity contract (necessary vs contingent) and evidence quantification (qualitative vs probabilistic). "Necessary $\times$ probabilistic" is void—probabilistic claims presuppose contingency, leaving three well-studied calculi [Hughes and Cresswell, 1996, Douven, 2022, Pearl, 2009]. Triad: R1 Deductive, R2 Probabilistic, R3 Abductive/Analogical.

Optimize & Control. Axes: decision coupling (open vs closed loop) and model knowledge (accessible vs black box). In closed loop, policies driven by analytic versus surrogate/finite-difference gradients behave identically once the update law is fixed, so those quadrants merge [Åström and Murray, 2008]; black-box plants motivate derivative-free or meta-heuristic search [Conn et al., 2009]. Triad: O1 Deterministic optimization, O2 Stochastic/meta-heuristic search, O3 Adaptive feedback control.

Simulate & Emulate. Axes: physics prior (present vs absent) and data-driven closure (present vs absent). Models lacking both prior and closure are uninformative and discarded; the survivors match standard practice in physics-informed ML [Raissi et al., 2019, Karniadakis et al., 2021]. Triad: S1 First-principles solver, S2 Data surrogate, S3 Physics-informed hybrid.

Generate & Create. Axes: optimization loop (present vs absent) and prompt conditioning (present vs absent). A loop without a target is incoherent; with an inner loop, the goal can be internalized and updated during generation (e.g., diffusion-based planning), effectively collapsing the prompt-present pair in practice [Prabhumoye et al., 2020, Janner et al., 2022](Prabhumoye et al., 2020; Janner et al., 2022). Triad: G1 Unconditional sampling, G2 Conditioned synthesis, G3 Goal-directed search.

Autonomize & Orchestrate. Axes: process variability (static vs dynamic) and decision authority (human vs machine). A strictly static script cannot host a learned policy (void), yielding three workflow regimes supported by evidence from scientific workflow systems and self-driving laboratories [Deelman et al., 2018, Häse et al., 2019, Tom et al., 2024]. Triad: A1 Scripted automation, A2 Policy-driven orchestration, A3 Goal-level autonomy.

C Pathway identification, anchoring, and representative method families

Pathways are identified top-down by mitigation mechanisms: for each function–tension pairing, we formulate three mechanism-level strategies. Once a pathway is named, we attach its minimal atomic signature—the smallest sufficient set of atoms (a single atom or a combination)—as a post hoc anchor to intrinsic properties rather than transient techniques; pathway-level revisions are therefore infrequent. We retain a pathway if its mechanism is distinct within the cell, transfers across domains, and admits a minimal signature; we merge interchangeable labels, split when one label conflates separable mechanisms, re-tag when the primary relief lies in another function, and add only for genuinely new mechanisms recurring across domains. We claim MECE at the function and atomic levels; the strategy layer is illustrative rather than exhaustive, and non-redundant within a cell.

We assemble the strategic pathways into the 6×4 Strategy Matrix (Fig. 2). Rows correspond to the six core functions; the left-hand column for each row represents the two intrinsic axes and the resulting atomic triad; columns are the four discovery tensions. Each cell is labeled with a single keyword that names the shared mitigation logic and lists three strategic pathways with their atomic signatures; starred cells ($\star$) indicate high-leverage pairings for the corresponding tension.

For each pathway, we cite representative method families—broad and literature-anchored categories rather than individual models—so that the strategy is actionable without prescribing a specific algorithm. Families were selected by (i) coverage (used across multiple domains), (ii) maturity (canonical surveys or benchmarks), and (iii) explanatory fit to the pathway's mechanism; they are illustrative, not exhaustive. Mapping a method family to a pathway is illustrative rather than prescriptive: alternative families realizing the same mechanism are acceptable and do not alter the pathway's atomic signature. The six function-wise tables below instantiate the matrix: for each cell (function × tension), we present the cell's keyword and three strategic pathways, each with a one-sentence mechanism-level definition and an atomic signature, and for each pathway we list the representative method families with in-table citations that are illustrative rather than exhaustive.

Table C1: **Representation** — P1 *Symbolic/Structured*, P2 *Probabilistic*, P3 *Latent-manifold*

Tension *(Keyword)*	**Strategic Pathways**	**Method Families (illustrative)**
Complexity *(Factorize)*	• **Latent factorization (P3):** Learn a lower-dimensional latent coordinate system with weakly coupled factors, reducing intrinsic dimensionality.	• Disentangled / factorized latent encoders [Kim and Mnih, 2018] • Low-rank and tensor-factorization encoders [Hu et al., 2022]
	• **Hierarchical abstraction (P3):** Build a multi-level representation where higher levels summarize and organize lower levels, enabling scale-appropriate computation.	• Hierarchical VAEs / multi-scale latent encoders [Vahdat and Kautz, 2020] • Hierarchical or segmented transformers [Liu et al., 2021]
	• **Structured graph extraction (P1):** Map observations into typed symbolic structures (entities, relations, rules) to expose constraints and sparsity for combinatorial pruning.	• Typed-graph mining & ontology induction [Zhang et al., 2018] • Programmatic schema/grammar induction [Kim et al., 2019b] • Scene/semantic graph parsers [Li et al., 2022a]
Constraint *(Robustify)*	• **Physics-aware embedding (P3+P1):** Encode invariants and constraints (symmetries, conservation, rule structure) via continuous fields plus discrete entities/relations to preserve validity.	• Equivariant encoders [Cohen and Welling, 2016] • Physics-informed encoders [Raissi et al., 2019] • Neural fields with constraint features [Beucler et al., 2021]
	• **Noise-robust encoding (P3):** Learn latent representations that attenuate measurement noise and artefacts while preserving signal, with implicit or explicit noise modeling.	• Denoising autoencoders [Vincent et al., 2008] • Diffusion-based denoisers [Ho et al., 2020] • Consistency-regularized contrastive encoders [Chen et al., 2020]
	• **Domain-invariant mapping (P3+P2):** Separate domain factors and align distributions so task features transfer with quantified uncertainty.	• Domain adversarial encoders [Ganin et al., 2016] • Moment/marginal alignment encoders [Sun and Saenko, 2016]

continued on next page

Tension *(Keyword)*	**Strategic Pathways**	**Method Families (illustrative)**
Scarcity *(Amplify)*	• **Signal-boost encoding (P3):** Use self-supervised pretraining to harvest structure from unlabeled data, improving sample efficiency for downstream tasks.	• Masked Autoencoders [He et al., 2022] • Contrastive pretraining [Chen et al., 2020] • Predictive coding encoders [Baevski et al., 2020]
	• **Cross-source fusion (P3):** Align multiple sources or modalities into a shared latent space to transfer supervision and fill coverage gaps.	• Joint multimodal embeddings [Radford et al., 2021] • Deep CCA / multi-view alignment [Wang et al., 2017] • Mixture-of-encoders multimodal co-embedding [Bao et al., 2022]
	• **Uncertainty-aware imputation (P2+P3):** Probabilistic completion of missing data with calibrated uncertainty.	• Variational imputers [Mattei and Frellsen, 2019] • Diffusion/score-based imputers [Tashiro et al., 2021] • Probabilistic graphical imputers [Zhao and Udell, 2020]
Explosion *(Compress)*	• **Regularity compression (P1+P3):** Encode symmetries and invariants as discrete indices with equivariant continuous features to eliminate redundant search.	• Group-equivariant CNNs / transformers [Cohen and Welling, 2016] • Steerable / Lie-group encoders [Cohen and Welling, 2017] • Symmetry-aware encoders [Fuchs et al., 2020]
	• **Multi-resolution abstraction (P3):** Use hierarchical indices and multiscale latents for coarse-to-fine navigation and inference.	• Multiscale graph encoder–decoder [Gao and Ji, 2019] • Wavelet / scattering feature pyramid encoders [Bruna and Mallat, 2013] • Hierarchical latent pyramid models [Razavi et al., 2019]
	• **Compact latent indexing (P1):** Quantize or hash embeddings into compact discrete codes, enabling sub-linear retrieval and pruning.	• Deep hashing families [Liu et al., 2016] • Product quantization indexing [Jégou et al., 2011] • Vector-quantized autoencoders [van den Oord et al., 2017]

Table C2: **Reason & Infer** — R1 *Deductive*, R2 *Probabilistic*, R3 *Abductive/Analogical*

Tension *(Keyword)*	**Strategic Pathways**	**Method Families (illustrative)**
Complexity *(Causalize)*	• **Causalize system dynamics (R3+R2):** Learn causal structure and effect strengths to separate drivers from correlates, enabling intervention-aware simplification of model search.	• Causal discovery (score-/constraint-/invariance-based; nonlinear variants) [Zheng et al., 2018] • Causal graphical modeling with neural parameterization [Pawlowski et al., 2020]
	• **Probabilistic dependency modeling (R2):** Build calibrated graphical or conditional-density models that capture uncertainty in dependencies under partial information.	• Bayesian networks & factor graphs [Zhang et al., 2023] • Deep conditional density estimators (autoregressive/flow models) [Papamakarios et al., 2017]
	• **Deductive causal invariants (R1):** Establish and enforce invariants/constraints that must hold under interventions, pruning hypotheses inconsistent with theory.	• Differentiable theorem proving [Rocktäschel and Riedel, 2017] • Neuro-symbolic constraint layers [Wang et al., 2019b]
Constraint *(Prequalify)*	• **Deductive feasibility prequalification (R1):** Use known rules/constraints to pre-screen candidate designs or experiments, eliminating impossible or non-compliant options before optimization.	• Neuro-rule engines / differentiable logic layers [Manhaeve et al., 2018] • Constraint programming / SAT–SMT with neural guidance [Selsam et al., 2019]
	• **Probabilistic feasibility estimation (R2):** Estimate the probability of constraint satisfaction under data/model uncertainty.	• Probabilistic graphical models [Wainwright and Jordan, 2008] • Simulation-based inference [Papamakarios et al., 2019]
	• **Abductive constraint induction (R3):** From observed passes/failures, infer latent rules/guards that best explain feasibility patterns and generalize them.	• Differentiable ILP [Evans and Grefenstette, 2018] • Neuro-guided abduction [Dai et al., 2019]
Scarcity *(Generalize)*	• **Rule-driven extrapolation (R1):** Apply mechanistic/symbolic relations to extend predictions beyond the training regime with logical validity.	• Neuro-symbolic regression [Petersen et al., 2021] • Sparse system-identification [Brunton et al., 2016]
	• **Bayesian prior integration (R2):** Combine informative priors with limited data to produce calibrated posteriors that generalize.	• Gaussian process regression [Wang et al., 2019a] • Hierarchical Bayesian models [Kim et al., 2019a]

continued on next page

Tension *(Keyword)*	**Strategic Pathways**	**Method Families (illustrative)**
	• **Few-shot hypothesis induction (R3):** Generate and refine candidate hypotheses from few examples via analogical or meta-level reasoning.	• Gradient-based meta learning [Finn et al., 2017] • Prompted LLMs abductive reasoning frameworks [Shi et al., 2023]
Explosion *(Prune)*	• **Deductive constraint propagation (R1):** Propagate hard constraints to shrink the search space by eliminating inconsistent branches early.	• Differentiable CSP networks [Jiang et al., 2022] • Neural-guided SAT/SMT/CP solvers [Selsam et al., 2019]
	• **Probabilistic branch ranking (R2):** Score and select branches by success probability or expected value to focus search effort.	• Bayesian value estimation for branch-and-bound [Mern et al., 2021] • Posterior-guided heuristic search [Tesauro et al., 2010]
	• **Abductive pathway trimming (R3):** Prefer explanations with minimal assumed causes, dropping branches not required by the best explanation.	• Neuro-symbolic weighted abduction [Huang et al., 2021] • RL-guided abduction [Bai et al., 2024]

Table C3: **Optimize & Control** — O1 *Deterministic optimization*, O2 *Stochastic/meta-heuristic search*, O3 *Adaptive feedback control*

Tension *(Keyword)*	**Strategic Pathways**	**Method Families (illustrative)**
Complexity *(Navigate)*	• **Gradient-based surrogate navigation (O1):** Use differentiable surrogates to obtain gradients or adjoints and optimize in a reduced space, streamlining search.	• Neural operator surrogates [Lu et al., 2021] • Differentiable physics surrogates [de Avila Belbute-Peres et al., 2018]
	• **Heuristic black-box search (O2):** Explore the objective with uncertainty- or heuristic-driven proposals when gradients are unavailable or unreliable.	• Bayesian optimization [Snoek et al., 2012] • Evolutionary strategies [Sun et al., 2022]
	• **Self-adaptive feedback control (O3):** Maintain closed-loop policies that update online from rollouts or streaming data to track changing dynamics.	• Meta RL adaptation [Finn et al., 2017] • Adaptive model predictive control [Amos et al., 2018]
Constraint *(Satisfy)*	• **Learned safety certificates (O1):** Train differentiable barrier or Lyapunov certificates and embed them in deterministic objectives to enforce feasibility at low cost.	• Barrier/Lyapunov networks [Xiao et al., 2023] • Differentiable penalty/projection layers [Agrawal et al., 2019]

continued on next page

Tension *(Keyword)*	**Strategic Pathways**	**Method Families (illustrative)**
	• **Constrained acquisition search (O2):** Optimize acquisition functions that couple utility with feasibility or safety under uncertainty.	• Safe/constrained Bayesian optimization [Gardner et al., 2014] • Constrained evolutionary search [Arnold and Hansen, 2012]
	• **Risk-sensitive adaptive policies (O3):** Learn closed-loop controllers that satisfy constraints or risk budgets during execution.	• Risk-aware RL [Tamar et al., 2015] • Constrained policy optimization [Achiam et al., 2017]
Scarcity *(Prioritize)*	• **Initialization-based optimization (O1):** Use meta-learned initializations or learned optimizers to reduce steps to convergence on new tasks.	• MAML initialization [Finn et al., 2017] • Transfer-initialized solvers [Feurer et al., 2014]
	• **Uncertainty-driven sampling (O2):** Query-efficient evaluations by selecting points that maximize information gain or value of information.	• Information-theoretic acquisition [Wang and Jegelka, 2017] • Bayesian active learning for expensive evaluations [Kirsch et al., 2019]
	• **Curriculum-adaptive control (O3):** Adapt task or domain difficulty online to accelerate policy learning with minimal data.	• Self-paced RL [Klink et al., 2020] • Teacher–student curriculum generation [Florensa et al., 2017]
Explosion *(Guided-Search)*	• **Multi-fidelity surrogate screening (O1):** Coarse-to-fine evaluation using cheap surrogates to prune candidates before expensive solves.	• Multi-fidelity Bayesian optimization [Kandasamy et al., 2016] • Hierarchical surrogate cascades [Falkner et al., 2018]
	• **Structure-guided search (O2):** Learn problem structure-guided heuristics to prune or branch effectively in combinatorial spaces.	• GNN-guided branching & pruning [Gasse et al., 2019] • Neural heuristic guidance for tree search or combinatorial optimization [Silver et al., 2018]
	• **Hierarchical policy refinement (O3):** Plan at a coarse level and refine to fine-grained actions via hierarchical control.	• Hierarchical RL [Nachum et al., 2018] • Model-based RL with hierarchical planning [Pertsch et al., 2020]

Table C4: **Simulate & Emulate** — S1 *First-principles solver*, S2 *Data surrogate*, S3 *Physics-informed hybrid*

Tension *(Keyword)*	**Strategic Pathways**	**Method Families (illustrative)**
Complexity *(Approximate)*	• **Coarse-grain approximation (S1+S2):** Combine physics-based coarse models with learned surrogates to capture fine-scale effects at reduced resolution and cost.	• Projection-based or reduced-order models [Berman and Peherstorfer, 2023] • Latent dynamical-system emulators [Wu et al., 2022]
	• **Stochastic scenario sampling (S2):** Train data-driven stochastic simulators that sample plausible trajectories or outcomes to approximate distributional futures.	• Variationally trained stochastic simulators [Hafner et al., 2021] • Diffusion-based simulators [Janner et al., 2022]
	• **Residual-hybrid acceleration (S3):** Attach a learned residual or corrective policy to a first-principles solver to reduce error and iteration count while preserving governing structure.	• Residual physics-informed neural networks [Mao and Meng, 2023] • Gaussian-process residual models [Xing et al., 2021]
Constraint *(Virtualize)*	• **Virtual lab emulation (S1+S2):** Build executable virtual models of experimental workflows to rehearse procedures and test feasibility under controllable parameters.	• Differentiable physics engines [Schoenholz and Cubuk, 2021] • Agent-based or rule-based laboratory emulators [Häse et al., 2021]
	• **Rule-constrained simulation (S1):** Enforce hard rules and constraints inside the simulator so generated trajectories remain admissible.	• Constraint-enforcing numerical solvers (barrier/penalty/projection families) [Huang et al., 2024c] • Trust-region simulators [Schulman et al., 2015]
	• **Safe exploration loops (S3):** Close the loop with constraint-aware design-of-experiments, selecting next trials within certified risk or validity bounds.	• Bayesian experimental-design controllers [Wu et al., 2023] • Safe RL controllers [Liu et al., 2020]
Scarcity *(Synthesize)*	• **Mechanistic data synthesis (S1):** Use governing-equation solvers to generate labeled data across parameter regimes when measurements are unavailable.	• Physics-constrained neural solvers [Li et al., 2024] • Symbolic or numerical equation-driven simulators [Hu et al., 2020]
	• **Surrogate extrapolation (S2):** Learn empirical surrogates that extrapolate beyond observed regimes with quantified uncertainty.	• Flow-based or autoregressive emulators [Krause and Shih, 2023] • Adversarial generative emulators [Khattak et al., 2022]

continued on next page

Tension *(Keyword)*	Strategic Pathways	Method Families (illustrative)
	• **Physics-informed augmentation (S3):** Generate synthetic samples conditioned on physical invariants or constraints, and select under-covered regions iteratively.	• Physics-conditioned generative models [Chen et al., 2025] • Simulator–generator joint training loops [Shrivastava et al., 2017]
Explosion *(Accelerate)*	• **Physics-based prefiltering (S1):** Apply fast analytic or coarse-physics filters to reject infeasible candidates before high-fidelity simulation.	• Reduced-order physics screening [Lee and Carlberg, 2020] • Analytic bounding and approximation models [Law et al., 2023]
	• **Structure-guided pruning (S2):** Learn structure-aware heuristics that approximate solver decisions and prune branches or candidates early.	• Graph-based surrogate heuristics [Paulus and Krause, 2023] • Symbolic rule-learning for pruning [Kuang et al., 2024]
	• **Adaptive multi-fidelity screening (S3):** Allocate simulation budget across fidelity levels with closed-loop policies that update using uncertainty and cost.	• Multi-fidelity Bayesian optimization [Li et al., 2020b] • Active learning with adaptive fidelity selection [Li et al., 2022b]

Table C5: **Generate & Create** — G1 *Unconditional sampling*, G2 *Conditioned synthesis*, G3 *Goal-directed search*

Tension *(Keyword)*	Strategic Pathways	Method Families (illustrative)
Complexity *(Probe)*	• **Counterfactual probing (G3):** Generate plausible alternatives under explicit intervention targets, searching inputs or latents so that specified factors change while others are held fixed.	• Counterfactual generative modeling [Sauer and Geiger, 2021] • Goal-directed diffusion generators [Poole et al., 2023]
	• **Edge-case exploration (G2+G3):** Bias generation toward tail regions using conditioning and adaptive guidance to surface rare or brittle behaviors.	• Tail-focused diffusion samplers [Pandey et al., 2025] • Importance-weighted or rejection-guided samplers [Na et al., 2024]
	• **Latent subspace probing (G1):** Explore the generator's intrinsic manifold by traversing or interpolating latent directions to reveal controllable factors.	• Latent traversal and interpolation [Härkönen et al., 2020] • Geodesic or spectral manifold probes [Shen and Zhou, 2021]
Constraint *(Prototype)*	• **Unconstrained prototype drafting (G1):** Use unconditional sampling and mixing to sketch broad candidate prototypes without validity constraints.	• Adversarial generative models [Karras et al., 2020] • Latent mixing or style-mixing methods [Karras et al., 2019]

continued on next page

Tension *(Keyword)*	**Strategic Pathways**	**Method Families (illustrative)**
	• **Rule-conditioned prototyping (G2):** Generate candidates conditioned on rules/grammars/masks/property descriptors to ensure compliance at draw time.	• Grammar-constrained decoders [Kusner et al., 2017] • Mask-conditioned diffusion generators [Lugmayr et al., 2022]
	• **Constraint-loop prototyping (G3):** Evaluate–generate–refine in a closed loop under stated constraints to steer prototypes toward admissible regions.	• RL fine-tuning with constraint penalties [Jayant and Bhatnagar, 2022] • Constrained Bayesian search over generator controls [Maus et al., 2022]
Scarcity *(Augment)*	• **Data augmentation (G1):** Create training signal via transformations or unconditional synthesis to expand coverage without labels.	• Geometric, photometric, or spectral transformations [Cubuk et al., 2019] • Unconditional generative augmentation [Trabucco et al., 2024]
	• **Weak label expansion (G2):** Create labels or label-like signals using teacher models or constraints, attaching noisy but useful annotations to existing or generated data.	• LLM labeling frameworks [Wang et al., 2023b] • Self-training and consistency-based pseudo-labeling [Sohn et al., 2020]
	• **Utility-guided augmentation (G3):** Choose what to generate next by maximizing downstream utility or information gain with a generator-in-the-loop.	• Active learning with generator proposals [Kim et al., 2020] • Bayesian acquisition-guided generation [Gruver et al., 2023]
Explosion *(Seed)*	• **Diversity-maximized sampling (G1):** Select seeds to maximize coverage and diversity in latent or feature space before downstream search.	• Determinantal point process samplers [Bardenet et al., 2021] • Maximum-entropy generators [Bengio et al., 2023]
	• **Constraint-aware seed search (G2):** Generate-and-prune under constraints using grammars, masks, or property predicates to keep only admissible seeds.	• Grammar-driven samplers with rule-based pruning [Park et al., 2024b] • Constraint-aware property-guided decoders [Huang et al., 2024b]
	• **Hierarchical assembly (G3):** Compose complex artifacts from parts via multi-stage plans where generation and selection alternate across a hierarchy.	• Hierarchical generative models [Ho et al., 2022] • RL assembly policies [Gürtler et al., 2021]

Table C6: **Automate & Orchestrate** — A1 *Scripted automation*, A2 *Policy-driven orchestration*, A3 *Goal-level autonomy*

Tension *(Keyword)*	**Strategic Pathways**	**Method Families (illustrative)**
Complexity *(Auto-Compose)*	• **Scripted multimodal coordination (A1):** Compose fixed, human-authored pipelines that coordinate heterogeneous tools and data via explicit handoffs.	• Workflow domain-specific languages [Crusoe et al., 2022] • Static pipeline orchestrators and schedulers [Deelman et al., 2015]
	• **Policy-driven scaling & routing (A2):** Use human-specified policies to place, scale, and route tasks dynamically at run time.	• Policy-based workflow schedulers [Yang et al., 2023] • Container-orchestration autoscaling frameworks [Rzadca et al., 2020]
	• **Closed-loop pipeline auto-tuning (A3):** Learn controllers that select or adjust components, hyperparameters, and resource allocations online to optimize end-to-end objectives.	• RL pipeline controllers [Park et al., 2024a] • Meta-controllers for configuration search [Song et al., 2022]
Constraint *(Auto-Enforce)*	• **Validation & workflow codification (A1):** Formalize protocols, validations, and provenance as executable steps before experiments.	• Protocol domain-specific languages [Mehr et al., 2020] • Provenance and lineage graphs [Soiland-Reyes et al., 2022]
	• **Policy-driven guardrails & feedback (A2):** Install dynamic guardrails (halt, rollback, review) triggered by human-authored risk or quality policies during operation.	• Safety and compliance guardrails [Berkenkamp et al., 2016] • Active learning lab schedulers with policy thresholds [Low et al., 2024]
	• **Self-lab orchestration (A3):** An autonomous planner–executor that plans, executes, measures, and adapts experiments under constraints using learned policies.	• Robotic experiment platforms [Szymanski et al., 2023b] • LLM planners for lab tasks [Boiko et al., 2023b]
Scarcity *(Auto-Curate)*	• **Scripted acquisition & integration (A1):** Compose fixed, human-authored pipelines to acquire data and integrate schemas.	• Schema-aware ETL pipelines [Shankar et al., 2023] • Web and API crawlers [Raffel et al., 2020]
	• **Policy-driven auto-labeling (A2):** Apply rule-/policy-guided labeling with model assistance to generate or refine annotations at scale.	• Weak-supervision labelling frameworks [Rühling Cachay et al., 2021] • LLM labeling bots under policies [Smith et al., 2024] • Active learning labeling schedulers [Ash et al., 2020]

continued on next page

Tension *(Keyword)*	Strategic Pathways	Method Families (illustrative)
	• **Autonomous quality refinement (A3):** Learn to detect, correct, and reweight noisy, duplicate, or low-quality data in closed loop.	• Noise filtering agents [Li et al., 2020a] • Learned deduplication & outlier detection [Thakkar et al., 2023]
Explosion *(Auto-Screen)*	• **Batch high-throughput screening (A1):** Run fixed batch pipelines that evaluate large candidate sets in parallel through scripted stages.	• Workflow DAG pipeline frameworks [Baylor et al., 2017] • Parallel batch execution frameworks [Moritz et al., 2018]
	• **Policy-driven triage & scheduling (A2):** Use human-specified scoring/eligibility rules to prioritize and schedule candidates over time.	• Policy-driven triage with learned scoring [Chzhen et al., 2023] • Multi-armed bandit schedulers [Qi et al., 2023]
	• **Closed-loop active screening (A3):** Select the next candidates iteratively using value and uncertainty models to maximize discoveries under budget.	• Active learning acquisition controllers [Parvaneh et al., 2022] • Adaptive experimental-design controllers [Huang et al., 2024a]