Differential Problem-Solving

A framework for resource allocation under accelerating technology

Siri Southwind

Working paper. Draft.

Abstract

The cost of solving things is collapsing. Sequencing a human genome, training a foundation model, predicting a protein structure, simulating a molecule, launching a kilogram, classifying an image, parsing a corpus — every one of these costs has fallen by orders of magnitude in the past two decades, and in many cases continues to fall. This paper argues that the dominant question facing allocators of resource — capital, attention, compute, talent — has shifted from *can this problem be solved* to *when should it be solved, and at what cost relative to where the technology will be in one, three, five or ten years*. I propose a framework, *Differential Problem-Solving*, that promotes the time-dependence of tractability to a first-class variable alongside the importance, tractability and neglectedness of the Effective Altruism tradition. The framework integrates real-options theory, Wright's-law cost trajectories, the Hamming "important problems" tradition, Bostrom's *Differential Technological Development*, and the scenario-planning tradition associated with Pierre Wack and Royal Dutch Shell. It produces falsifiable verdicts on which problems should be attacked now and which should be deferred, and it requires those verdicts to be made scenario-conditional rather than implicitly anchored on a single forecast. I apply it to a curated set of historical projects to test calibration, and to a list of live bets to test predictive value. The framework is designed to be revisable; it makes specific claims that can be wrong in specific ways. I close with limits, open questions and a proposed protocol for empirical refinement.

Keywords: problem selection, technology forecasting, real options, research allocation, differential development, cause prioritisation, scenario planning

1. Introduction

For most of human intellectual history, the binding constraint on consequential work was capability. Could the thing be done at all? Could you, with the resources available, do it? The question of *which* problem to attack was secondary to the question of *whether* a problem was attackable.

This priority has been quietly inverted. Across an increasing number of domains, the cost of attacking a problem of fixed difficulty is on a steep declining curve. A specific bioinformatics question that required a billion-dollar institutional programme in 2003 now requires a graduate-student weekend. A specific protein-structure prediction that occupied a postdoctoral career in 1993 is generated in seconds in 2024. A specific class of natural-language tasks that required a department of linguists and engineers in 2017 is solved by API call in 2026. The pattern is not universal — many problems remain genuinely hard for reasons that compute and AI cannot touch — but the pattern is widespread enough to change the calculus of allocation.

When capability is scarce, the right unit of analysis is the project: pick something, attack it, see if you can do it. When capability is collapsing, the right unit of analysis is the *moment*. The same problem attacked in 2024 and in 2027 may be entirely different in cost, in scope and in cascade.

Allocators who reason as if capability were still scarce systematically misallocate.

This paper proposes a framework for reasoning about that allocation problem. The contribution is not the dimensions individually — almost every dimension I name appears somewhere in the existing literatures of cause prioritisation, technology forecasting, real-options theory or operations research. The contribution is the integration: promoting the time-dependence of tractability to a first-class variable, providing a vocabulary that allocators across domains can share, and producing verdicts that are specific enough to be argued with.

The paper proceeds as follows. Section 2 surveys the prior literature and locates the contribution. Section 3 presents the framework formally. Section 4 develops mathematical models, drawing on real-options theory and optimal stopping. Section 5 applies the framework to a set of historical cases to test calibration. Section 6 makes predictions on live cases and proposes the falsifiability protocol. Section 7 discusses limits. Section 8 concludes.

2. Related work

The intellectual lineage of the framework spans several literatures that have not, to my knowledge, been integrated.

2.1 The problem-list tradition

Hilbert's 1900 lecture in Paris, listing twenty-three open problems in mathematics, is the founding act of explicit problem-allocation thinking (Hilbert, 1900; Yandell, 2002). Stephen Smale's 1998 update extended the gesture, with mixed retrospective accuracy. The Clay Millennium Prize Problems (2000) added explicit financial incentives. The tradition treats problem selection as a discipline in its own right but provides no machinery for *timing*.

Richard Hamming's 1986 talk *You and Your Research* is the spiritual ancestor of this paper. Hamming asked: "What are the most important problems in your field? Why aren't you working on them?" The framework here is, in part, Hamming's question rendered tractable for an environment in which "important" and "feasible" have come apart on different time-scales.

2.2 Cause prioritisation

The Effective Altruism tradition has produced the most developed existing framework for problem selection (Ord, 2020; MacAskill, 2015; Open Philanthropy, 2014–). The core formulation evaluates problems by *importance* (how much value would a solution produce), *tractability* (how much does an additional unit of resource move the needle) and *neglectedness* (how few people are already on it). The framework here borrows IT–N essentially wholesale and adds a fourth term, *timing*, treating tractability as a function of when the question is asked rather than as a static input.

The EA tradition has been criticised for various reasons (uncertainty quantification, moral framework choices, neglect of structural factors). The framework here is partly compatible with most of those critiques and partly orthogonal to them.

2.3 Differential Technological Development

Nick Bostrom's principle of differential technological development — that beneficial technologies should be accelerated relative to dangerous ones rather than allowing technologies to arrive in the order their researchers happen to deliver them (Bostrom, 2014; Bostrom and Ćirković, 2008) — is the most direct ancestor of the present framework. *Differential Problem-Solving* applies the same move at the level of individual problems rather than whole technologies, and adds time-trajectory analysis as a first-class concern.

2.4 Real options and optimal stopping

The mathematical machinery imported here is largely from finance and operations research. Real-options theory (Trigeorgis, 1996; Dixit and Pindyck, 1994) provides the formal vocabulary for valuing the option to wait under uncertainty. Optimal stopping theory (Wald, 1947; Chow, Robbins and Siegmund, 1971) provides the analytical framing for sequential decisions over an unknown distribution. Both bodies of theory are mature; the contribution here is to apply them explicitly to problem selection rather than to investment timing.

2.5 Cost trajectories and learning curves

Wright's 1936 paper introduced the empirical observation that the cost of producing aircraft fell by a roughly constant percentage with each doubling of cumulative production (Wright, 1936). The pattern has held for an extraordinary range of technologies since: solar panels, batteries, sequencing, semiconductors, satellite launches (Nagy et al., 2013; Lafond et al., 2018). The framework relies heavily on Wright's-law-shape forecasts as the empirical backbone of the *cost trajectory* dimension. Shannon's information-theoretic framing of communication (Shannon, 1948) is a methodologically adjacent move: both Wright and Shannon insist that quantities others treat as resistant to measurement can in fact be measured, and that the act of insisting changes what becomes possible.

2.6 Innovation and growth economics

The economics of ideas, particularly Romer's endogenous growth theory (Romer, 1990) and the more recent work on declining research productivity (Bloom et al., 2020), provides background for the framework's claim that *which* problems are attacked is itself a determinant of growth. The Carlota Perez tradition (Perez, 2002) on the installation and deployment phases of technological revolutions provides another useful lens.

2.7 Forecasting and superforecasting

The discipline of calibrated probability estimation introduced by Tetlock and the Good Judgment Project (Tetlock, 2005; Tetlock and Gardner, 2015) provides methodological support for the framework's scoring scheme.

2.8 Scenario planning and the Royal Dutch Shell tradition

Tetlock-style forecasting and scenario planning are usually treated as competing rather than complementary. They are complementary. Tetlock seeks calibrated point estimates against operationalised questions; scenario planning, in the tradition associated with Pierre Wack at Royal Dutch Shell (Wack, 1985a, 1985b), refuses point estimates entirely and instead constructs small sets of plausible, internally-consistent narratives spanning the critical uncertainties. Wack's Group Planning team is best known for the 1972 scenarios that anticipated the 1973 oil shock, repositioning Shell to act faster than its peers when the embargo arrived. The discipline was developed further by Schwartz (1991), van der Heijden (1996) and Kahane (2004), and earlier prefigured in the strategic-defence work of Kahn (1962) at the RAND Corporation.

The framework presented here is single-scenario by default — each of the sixteen dimensions, scored honestly, depends on an unstated forecast about how the world will unfold over the relevant time horizon. The scenario tradition is the natural corrective. The implication, developed in Section 4.4, is that the dimension scores should be computed across a small set of scenarios rather than against a single implicit forecast, and that the framework's portfolio shapes should be re-read in scenario language: *patient infrastructure* shares are the bets that pay off across the entire envelope of plausible futures; *moonshot* shares are the bets that pay massively in one scenario and produce useful by-products in the rest; *just-early* shares are the bets timed against a specific scenario about the cost trajectory and are therefore most fragile if the scenario is wrong.

2.9 The science of science

The recent quantitative literature on which projects produce more, which collaborations are productive, which kinds of papers anticipate breakthroughs (Wang and Barabási, 2021; Fortunato et al., 2018; Park, Leahey and Funk, 2023) is empirical where the present framework is theoretical. The two should converge.

3. The framework

The framework consists of sixteen primary dimensions clustered into four families. The companion repository maintains a fuller canonical list in its dimensions file; the sixteen presented here are those most consequential for typical allocation decisions, and the remainder are useful in specific subdomains. The dimensions are not orthogonal — several correlate substantively — and the framework does not claim to be a calculator. Its purpose is to make the relevant variables explicit so that the implicit *why now* in any allocation decision can be made articulable, criticisable and revisable.

Two dimensions deserve to be foregrounded. *Verification cost* (D1) and *cost trajectory* (D12) are the most important variables in almost every interesting allocation decision today. Cost trajectory captures the framework's central temporal insight. Verification cost captures the *generation–verification asymmetry* — the fact that AI has made it dramatically cheaper to produce hypotheses, code, designs and analyses than to check them. The bottleneck has moved to the verification side and is staying there.

3.1 Cost and difficulty

(D1) *Verification cost.* What does it take to confirm a proposed solution actually solves the problem? In 2026 this dimension is doing more work than ever, because generation cost is collapsing while verification cost is essentially flat. A problem whose verifier is itself hard is now substantially worse-positioned than the same problem ten years ago.

(D2) *Difficulty today.* What does it actually take to solve this with current tools, expressed in money, person-hours, compute and time?

(D3) *Required talent density.* How concentrated is the relevant expertise, and how transferable from adjacent fields?

(D4) *Capital intensity.* What fraction of the cost is up-front, sunk, irreversible?

(D5) *Coordination cost.* How many independent actors must agree, contribute or stay out of the way?

(D6) *Physical-resource dependency.* What does the project require from the physical world that cannot be substituted away — energy intensity, dependency on specific elements (rare earths, lithium, gallium, helium, isotopes), and permitted physical actions (siting, regulatory consent, environmental impact, dual-use export)? When cognition is cheap, physical-resource dependency is increasingly the binding constraint.

3.2 Value

(D7) *Direct value.* What does the world look like immediately after the problem is solved?

(D8) *Cascade value.* What other problems become solvable, cheaper or differently-shaped as a result?

(D9) *Demonstration value.* What is the value of removing the question of feasibility for an entire category of problems?

(D10) *Optionality value.* What future actions become possible that are not possible today?

(D11) *Decay rate of value.* How fast does the value of a solution erode as substitutes arrive?

3.3 Time and curve dynamics

(D12) *Cost trajectory.* How is the cost-to-solve changing year on year, and on what underlying curve?

(D13) *Tractability trajectory.* Is the *shape* of the problem changing because complementary technologies are arriving?

(D14) *Window.* Is there a closing horizon beyond which the problem cannot be solved at all, or cannot be solved with current evidence intact?

3.4 Strategic context

(D15) *Reusability of by-products.* If the headline goal fails, are the data, infrastructure, methodology or talent generated still valuable?

(D16) *Crowding and neglectedness.* How many capable teams are already attacking this, and what is the marginal return of one more?

A separate *dual-use weighting* applies to a small but consequential set of problems carrying catastrophic-risk asymmetries. For these, the standard cascade and demonstration readings invert: large cascade is a reason for caution rather than confidence, and demonstration value can change sign because demonstrating possibility tells malicious actors the capability exists. The dual-use modification is presented separately in the companion repository rather than as a numbered dimension because it changes how several of the standard dimensions should be read. The framework remains morally neutral on most allocation decisions; on dual-use catastrophic-risk decisions it cannot afford to be silent.

The dimensions interact. A problem with high cascade value but a fast cost-decline trajectory and no closing window is a *wait* candidate. A problem with modest direct value but a closing window and large by-product reusability is an *attack now* candidate. A problem with all the right scores but extreme crowding is an *attack a sub-problem instead* candidate. The framework's job is to make these readings explicit.

3.5 Decision rule

The framework reduces to four allocation moves at any decision moment. Attack now with brute force, with the goal of demonstrating feasibility and producing useful by-products. Attack now with elegance, where the curve is mature and the marginal team's contribution is incremental. Wait, on the bet that the cost will fall and the option will still be exercisable. Or attack a different problem entirely, on the bet that one of the alternatives dominates this one on the relevant axes.

A fifth move — *decompose* — is often the right one: attack a sub-problem now, defer the rest, on the bet that solving the linchpin sub-problem unlocks the dependency graph cheaply.

The framework is silent on which of these moves is correct in any specific case; it provides the vocabulary for arguing.

4. Mathematical formalisation

This section sketches the formal machinery. The technical content is intentionally light; the aim is to be rigorous about the *kind* of object the framework is, not to produce a full quantitative model. A more developed treatment is left for future work.

4.1 The waiting decision as a real option

Consider a problem that, if solved at time t, produces value $V(t)$ and costs $C(t)$ to solve. Both are stochastic processes. The allocator may choose to attack at any time t or to defer. The decision problem is:

$$\max_t E[\, V(t) - C(t) - r(t) \,]$$

where $r(t)$ is the opportunity cost of capital and attention deployed to this problem at time t rather than to alternatives.

If $C(t)$ is on a fast declining curve (Wright's-law-shape, with declining variance) and $V(t)$ is roughly stationary, the optimal policy is typically to wait. If $V(t)$ declines (decay of relevance) or has a hard horizon (closing window), the optimal policy is typically to attack early. If $V(t)$ has a step function — a complementary technology arrives at some unknown time and unlocks much higher value — the optimal policy depends on the prior over the arrival time and on the cost of being unprepared.

The classical real-options literature (Trigeorgis, 1996; Dixit and Pindyck, 1994) provides closed-form solutions under various distributional assumptions. The application to research-and-development problems is non-trivial because the underlying distributions are often non-Gaussian and the volatility itself is changing. A detailed treatment of this for specific problem classes is the most natural follow-up to this paper.

4.2 The wait curve

A useful informal device is the *wait curve*: a plot of *expected cost-to-solve* over time on one axis and *probability-weighted value* on the other. The right time to attack is approximately where the rate of cost decline first falls below the rate of value decline, adjusted for the expected cascade value of moving early.

In closed-form: attack at time t* such that

$$(d \log C) / (d t)\big|_{t^*} \approx (d \log V) / (d t)\big|_{t^*} - \kappa$$

where κ is a positive constant capturing the expected cascade and demonstration value of being among the first to attack rather than last.

The closed form is unhelpful in practice because C and V are uncertain. The wait-curve image, however, is useful as a heuristic: it organises the timing question around the ratio of the two underlying rates of change rather than around either curve in isolation.

4.3 The portfolio shape

For an allocator managing many bets simultaneously, the framework implies a portfolio strategy rather than a series of independent decisions. Specifically:

A *modal* portfolio share — the largest single category — should sit on problems in the *just early enough* zone, where the cost trajectory is favourable but moving early provides demonstration and cascade value.

A *moonshot* portfolio share — typically five to fifteen per cent — should sit on problems with low probability of success and asymmetric upside, on the bet that the portfolio's expected value is dominated by the rare wins. The shape is Taleb's barbell strategy applied to research allocation: heavy weight on safe positions, a small allocation to convex bets, very little in the middle (Taleb, 2007, 2012).

A *patient infrastructure* portfolio share — typically five to fifteen per cent — should sit on closing-window and cataloguing problems whose value compounds over decades and cannot be retroactively created.

A *deliberately unallocated* portfolio share — typically ten to twenty per cent — should remain available for problems that the framework cannot evaluate because they are not yet on anyone's list.

The remainder, fifty to seventy per cent, sits on the standard *attack-now-with-elegance* category that conventional allocation already handles well. The framework's main argument is that the *first* and *third* categories — *just early* and *patient infrastructure* — are systematically underweighted in most institutional portfolios.

4.4 Scenario-conditional scoring

Each dimension score, when the discipline is honest, is conditional on an unstated forecast about how the world will unfold. The cost trajectory of synthetic biology depends on whether biosecurity tightens; the closing window for indigenous-language documentation depends on whether ML-assisted transcription reaches remaining communities before generational hand-off completes; the cascade value of fusion depends on whether grid-scale storage solves intermittency on a different curve. A dimension score that ignores its own conditioning is a number masquerading as analysis.

The discipline borrowed from the scenario tradition (Wack, 1985a, 1985b; Schwartz, 1991; van der Heijden, 1996) is to score each dimension across a small set of plausible, internally-consistent scenarios — three is usually right, four when the field is genuinely contested — rather than against a single implicit forecast. The scenarios should differ on the variables that most plausibly drive the score, not on cosmetic surface features. The output is not a single sixteen-number row but a small matrix: dimensions on one axis, scenarios on the other, with a robustness reading at the foot.

The most useful reading is the *shape of robustness*. Let S denote the set of scenarios and $D(s)$ the dimension score in scenario s. A bet is *robust* if D(s) is positive across all $s \in S$; *directional* if it is positive in some scenarios and negative in others; *fragile* if it is negative across most. The framework's recommended portfolio shape (Section 4.3) re-reads cleanly in scenario language: *patient infrastructure* bets are robust by construction; *moonshot* bets are directional with antifragile by-products that produce useful capital across all scenarios; *just-early* bets are directional without the antifragile cushion and are the most exposed if the underlying scenario fails to obtain.

This formulation also gives the framework a defensible response to the criticism that it is single-scenario by default. The criticism is correct of any specific application; it is not correct of the framework when the scenario discipline is applied. Most failure modes of the framework, in practice, are failures of the user to make the implicit scenario explicit before scoring.

5. Historical calibration

The framework is testable, in part, against the historical record. This section presents brief readings of six cases, three from the *vindicated* class and three from the *misallocated* class, with the framework's verdict and the reasoning behind it.

5.1 The Human Genome Project (1990–2003)

Vindicated. Roughly three billion dollars over thirteen years for the first reference human genome.

Framework reading at decision time: high direct value (genomic medicine), enormous cascade value (the entire downstream genomics field), high cost-trajectory uncertainty, demonstration value substantial. Verdict: *attack now with brute force.*

Framework reading retrospective: vindicated. The HGP did not just produce a genome; it bent the cost curve, generated infrastructure, trained a generation of bioinformaticians and produced the reference assembly that all later sequencing aligns to. The dollar cost is large; the productivity-per-dollar is among the highest in twentieth-century biomedical investment.

5.2 ImageNet (2007–2010)

Vindicated. Fei-Fei Li and team paid for the labelling of fourteen million images via Mechanical Turk for an estimated few hundred thousand dollars.

Framework reading at decision time: low direct value (a benchmark dataset is not directly useful), high cascade value (training data for the next generation of vision systems), demonstration value uncertain, crowding low (the consensus was sceptical), cost-trajectory of compute favourable. Verdict: *attack now with brute force, expect the cascade to fire later*.

Framework reading retrospective: among the highest-leverage data-creation projects ever undertaken. AlexNet (2012) and the deep-learning revolution that followed were directly enabled by the dataset.

5.3 The Apollo programme (1961–1972)

Contested vindication. Roughly two hundred and fifty billion dollars in current money.

Framework reading at decision time: high direct value (national prestige and Cold War strategic positioning), substantial cascade value (integrated circuits, materials science, software engineering), high coordination cost, large capital intensity, no closing window in the literal sense but a politically-imposed deadline. Verdict: *attack now*, conditional on the strategic premise.

Framework reading retrospective: contested. The cascade was real but the magnitude is debated. The framework's verdict is positive on the premise that the alternative (not landing on the moon at all in the 1960s) would have produced substantially less of the cascade. Reasonable analysts disagree on this counterfactual by a factor of three or more.

5.4 Hand-tuned chess engines after Deep Blue (1999–~2010)

Misallocated. Significant academic and commercial effort continued on hand-tuned position-evaluation engines for a decade after the 1997 Kasparov match made it clear that brute-force search and increasingly hardware would dominate.

Framework reading at decision time: low marginal direct value, low cascade value, cost-trajectory of the alternative (compute) collapsing, crowding moderate-to-high, demonstration question already closed by Deep Blue. Verdict: *stop.*

Framework reading retrospective: misallocated. The community's institutional path-dependence kept the work going past its useful life. AlphaZero (2017) closed the question entirely, by which point the misallocation was essentially complete.

5.5 Enterprise rule-based NLP (2017–2022)

Misallocated. The global consulting industry spent tens of billions of dollars across thousands of engagements building hand-rolled entity extractors, sentiment analysers and document classifiers.

Framework reading at decision time: cost-trajectory of foundation-model alternatives extremely steep, cascade value of bespoke systems essentially zero, crowding extreme, demonstration value already accumulating elsewhere. Verdict: *stop, wait, or buy off-the-shelf.*

Framework reading retrospective: misallocated. Most of the systems built between 2017 and 2022 were rebuilt or scrapped between 2022 and 2024 once foundation models commoditised the underlying capability. The aggregate waste is among the largest in the recent history of corporate IT.

5.6 The Iridium satellite constellation (1991–1999)

Misallocated. Roughly five billion dollars for a global mobile-phone constellation that filed for bankruptcy nine months after service launch.

Framework reading at decision time: direct value real but heavily dependent on terrestrial cellular failing to expand, cost-trajectory of terrestrial cellular favourable, capital intensity extreme, reversibility low. Verdict: *do not attack at this scale*, possibly *wait* or *attack with smaller commitment.*

Framework reading retrospective: misallocated. Terrestrial cellular expanded much faster than the planning forecast, hollowing the addressable market. The infrastructure later found a niche under different ownership; the original allocation was a misread of the dominant input curve.

5.7 Calibration discussion

The framework's verdicts on the six cases above accord broadly with the post-hoc consensus, which is unsurprising — the framework was developed in part by examining cases like these. The harder test is on cases where the consensus is wrong: where the framework would have flagged a misallocation in advance, or vindicated a project the consensus rejected. The retrospective stupidity index (see 03_models_and_scoring.md of the broader repository) is the working tool for this exercise.

6. Predictions and falsifiability

A framework that produces no falsifiable verdicts is a mood. This paper, and the broader repository on which it draws, makes specific predictions on live cases. Three classes of prediction deserve naming.

6.1 Specific named bets

In the broader repository, the *current bets* file lists problems the framework rates as *attack now*, *attack with caveats*, *probably wait*, and *open*. The list is updated on six-month cycles. Each call is specific enough to be evaluated retrospectively five years from now.

The framework's predictive accuracy on the *attack-now* and *probably-wait* lists is one direct test. If the *attack-now* problems substantially underperform comparable controls over the next five-to-ten years, or the *probably-wait* problems substantially outperform, the framework's calibration is wrong.

6.2 Named anti-predictions

The *current 50* list in the broader repository lists projects the framework predicts will age badly. Specific named entries — sovereign frontier-model programmes, the NEOM Line, current Reality Labs spending levels, specific hyperscaling AI infrastructure bets, NASA's SLS — are each individually wrong if the project succeeds. If a substantial majority of the named entries on that list survive and thrive over the next five-to-ten years, the framework's curve-reading is wrong.

I expect to be wrong on roughly twenty per cent of that list. The point is not perfection; it is specificity.

6.3 Class-level predictions

The framework also makes structural predictions that are independent of any specific case. Three are particularly testable:

Foundation-model commoditisation will continue to compress the value of thin-wrapper businesses. This is happening in 2025; the framework predicts it accelerates rather than stabilises through 2027.

Closing-window infrastructure projects will become more strongly vindicated as their substrates are consumed by AI. The Internet Archive, the Materials Project, the Protein Data Bank, the various long-running cohorts will be valued more in 2030 than in 2024. If they are not, the framework's cascade-mapping is wrong.

Differential allocation will produce measurable advantage at the funder level. Funders who explicitly use timing-aware frameworks should outperform funders who use only conventional cause-prioritisation by a margin large enough to be visible in the data over a decade. This is the strongest and most falsifiable claim. If it is wrong, the entire enterprise is wrong.

7. Limits

The framework has named limits.

It does not price morality, or unknown unknowns, or the political economy of allocation. It cannot distinguish vindicated contrarians from doomed ones in advance. It does not model team quality. It is partially inadequate for catastrophic-risk decisions, where Bostrom's existential-risk literature remains the better tool.

It is also subject to two structural failure modes worth naming explicitly.

The first is *false precision*: producing a numeric score and treating the number as objective when the dimensions are qualitative. The framework is meant to force articulation, not to produce calculator-style verdicts.

The second is *contrarian pose*: mistaking *being against the consensus* for *being right*. The framework rewards counter-consensus reasoning where the reasoning is better than the consensus reasoning, not counter-consensus positions in general. Most counter-consensus positions are counter-consensus for a reason.

A more detailed treatment of limits and failure modes is provided in 12_limits_and_falsifiability.md of the broader repository.

8. Conclusion

I have argued that the dominant question facing allocators of resource has shifted, in many domains, from *can this problem be solved* to *when, and at what cost relative to alternatives*. The shift is consequential. Most institutional allocation systems — public science funding, philanthropic giving, corporate R&D, much of venture capital — were built when capability was the binding constraint and reason as if it still were. The result is a portfolio drift away from optimal that compounds each year as the underlying tractability landscape moves faster than the funding does.

Differential Problem-Solving is one attempt to make the timing question explicit. The framework integrates time-dependence of tractability with the existing literatures of cause prioritisation, real-options theory, Wright's-law cost trajectories, and the Hamming–Hilbert problem-list tradition. It is not a calculator. It is a vocabulary, a checklist, and a discipline that produces falsifiable verdicts on what to attack now and what to defer.

The framework will be wrong in specific ways. The wrong ways are the test of its value: a framework that cannot be wrong cannot be improved. The current bets and anti-bets in the broader repository are made specifically so that the framework can be evaluated against the world over the next five to ten years.

The world has more important problems than it has people who can recognise which to attack and when. To the extent that this framework helps any of those people argue more clearly, it has done its job.

References

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). *Are Ideas Getting Harder to Find?* American Economic Review, 110(4), 1104–1144.

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies.* Oxford University Press.

Bostrom, N., and Ćirković, M. M. (eds) (2008). *Global Catastrophic Risks.* Oxford University Press.

Box, G. E. P. (1976). *Science and Statistics.* Journal of the American Statistical Association, 71(356), 791–799.

Bush, V. (1945). *As We May Think.* The Atlantic Monthly, July.

Chow, Y. S., Robbins, H., and Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping.* Houghton Mifflin.

Deutsch, D. (2011). *The Beginning of Infinity: Explanations That Transform the World.* Allen Lane.

Dixit, A. K., and Pindyck, R. S. (1994). *Investment under Uncertainty.* Princeton University Press.

Fortunato, S., Bergstrom, C. T., Börner, K., et al. (2018). *Science of science.* Science, 359(6379), eaao0185.

Hamming, R. W. (1986). *You and Your Research.* Bell Communications Research Colloquium Seminar.

Hilbert, D. (1900). *Mathematische Probleme.* Lecture, International Congress of Mathematicians, Paris.

Kahane, A. (2004). *Solving Tough Problems: An Open Way of Talking, Listening, and Creating New Realities.* Berrett-Koehler.

Kahn, H. (1962). *Thinking About the Unthinkable.* Horizon Press.

Knight, F. H. (1921). *Risk, Uncertainty, and Profit.* Houghton Mifflin.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions.* University of Chicago Press.

Lafond, F., Bailey, A. G., Bakker, J. D., et al. (2018). *How well do experience curves predict technological progress? A method for making distributional forecasts.* Technological Forecasting and Social Change, 128, 104–117.

Lakatos, I. (1978). *The Methodology of Scientific Research Programmes.* Cambridge University Press.

MacAskill, W. (2015). *Doing Good Better.* Avery.

Nagy, B., Farmer, J. D., Bui, Q. M., and Trancik, J. E. (2013). *Statistical basis for predicting technological progress.* PLoS One, 8(2), e52669.

Open Philanthropy (2014–). *Cause Reports* and *Worldview Investigations.* Open Philanthropy Project.

Ord, T. (2020). *The Precipice: Existential Risk and the Future of Humanity.* Hachette.

Park, M., Leahey, E., and Funk, R. J. (2023). *Papers and patents are becoming less disruptive over time.* Nature, 613(7942), 138–144.

Perez, C. (2002). *Technological Revolutions and Financial Capital.* Edward Elgar.

Popper, K. R. (1963). *Conjectures and Refutations: The Growth of Scientific Knowledge.* Routledge.

Popper, K. R. (1972). *Objective Knowledge: An Evolutionary Approach.* Oxford University Press.

Romer, P. M. (1990). *Endogenous Technological Change.* Journal of Political Economy, 98(5), S71–S102.

Schwartz, P. (1991). *The Art of the Long View: Planning for the Future in an Uncertain World.* Doubleday.

Shannon, C. E. (1948). *A Mathematical Theory of Communication.* Bell System Technical Journal, 27(3), 379–423; 27(4), 623–656.

Smale, S. (1998). *Mathematical problems for the next century.* The Mathematical Intelligencer, 20(2), 7–15.

Taleb, N. N. (2007). *The Black Swan: The Impact of the Highly Improbable.* Random House.

Taleb, N. N. (2012). *Antifragile: Things That Gain from Disorder.* Random House.

Tetlock, P. E. (2005). *Expert Political Judgment.* Princeton University Press.

Tetlock, P. E., and Gardner, D. (2015). *Superforecasting: The Art and Science of Prediction.* Crown.

Trigeorgis, L. (1996). *Real Options: Managerial Flexibility and Strategy in Resource Allocation.* MIT Press.

van der Heijden, K. (1996). *Scenarios: The Art of Strategic Conversation.* John Wiley.

Wack, P. (1985a). *Scenarios: Uncharted Waters Ahead.* Harvard Business Review, 63(5), 73–89.

Wack, P. (1985b). *Scenarios: Shooting the Rapids.* Harvard Business Review, 63(6), 139–150.

Wald, A. (1947). *Sequential Analysis.* John Wiley.

Wang, D., and Barabási, A.-L. (2021). *The Science of Science.* Cambridge University Press.

Wright, T. P. (1936). *Factors affecting the cost of airplanes.* Journal of the Aeronautical Sciences, 3(4), 122–128.

Yandell, B. H. (2002). *The Honors Class: Hilbert's Problems and Their Solvers.* A K Peters.

Working paper. Comments and counter-arguments welcome. The companion Problem Timing repository contains historical examples, ranked lists, field guides, audience-specific one-pagers, the dual-use modification and a live revision protocol with dated commitments.

— Siri Southwind