

WHITE PAPER 10 / TRANSLATIONAL HORIZON

From Image Analysis to Scientific Intelligence at Scale

How a scientific distillation method developed in 2007 became the foundation for governing population-scale data, cross-resolution evidence, and AI reasoning without allowing speed to outrun scientific meaning.

DR. BRIAN T. BENNETT

Co-Founder and CSO

PAPER 10 OF 10

AUGUST 31, 2026

NATIVE EVIDENCE / DETERMINISTIC MEASUREMENT / GOVERNED INTERPRETATION / DEFENSIBLE PUBLICATION

CONTENTS

Paper structure

The concluding paper: how manually developed scientific distillation and cross-resolution practice become governed computation capable of increasing scientific scale without allowing speed, complexity, or AI reasoning to outrun evidence.

1. Paper 9 ended with a bounded record. Paper 10 asks what happens when many of them move together.	4
2. The scientific problem came before BENNETT.	6
3. In 2007, the problem was already scientific distillation.	7
4. The mathematics matters because every reduction answers a different question.	9
5. The bottleneck moved downstream: from acquisition to scientific continuity.	10
6. BENNETT changes the scale, not the source of authority.	11
7. High-throughput science is only useful if governance scales with it.	13
8. Population response is often more informative than the representative image.	14
9. Scale requires experimental hierarchy, not merely more cells.	15
10. Dose and time turn snapshots into response surfaces.	16
11. Drug discovery is an obvious application — not the limit of the architecture.	17
12. Resolution and modality can refine a response without becoming a truth tier.	18
13. Distillation can produce a focused scientific output without erasing the evidence beneath it.	19
14. Scientific intelligence is a hierarchy of meaning.	20
15. QUESTION / ANSWER	22
16. The future is faster governed science.	23
17. What the ten papers now add up to.	25
Conclusion	26
References	28

All numbered citations in the paper link to the corresponding reference. Linked URLs in the reference list open the cited source when a public URL is available.

ABSTRACT

Abstract

Paper 9 ended with a bounded scientific record: source-linked evidence, methods, measurements, scientist actions, validation, limitations, provenance, and claim boundaries leave the workflow together without allowing the Publisher to become a new source of scientific authority.[1] It then asked whether collections of such records could support higher-order reasoning without converting association into mechanism or exceeding assay readiness. Paper 10 places that question inside a broader one: what becomes possible when governed records can be related without reconstructing what each experiment measured, which state was current, or how far each result was allowed to speak?

The answer begins much earlier than BENNETT. In 2007, I was treating a fluorescence experiment as a multidimensional scientific distillation problem rather than an image-reading problem. The practical objective was to preserve differences within a population of drug-treated cells while reducing individual-cell measurements through population state, time, drug, and dose into a focused scientific distillate.[10,11] The data could already be generated faster than I could responsibly process, compare, and reduce them by hand.

I continued working through that problem manually through approximately 2012.[17] From roughly 2012 through 2020, expanding super-resolution capability added a second burden: each experiment contained more information, distributed across different acquisition and analysis environments. When Identifyn began in 2023, relating evidence from widefield, structured illumination, confocal, Airyscan, Lattice SIM, and single-molecule localization still required the scientific continuity among those tiers to be carried manually.[2-4,17]

BENNETT therefore did not begin with artificial intelligence and search for a problem to automate. It grew from a manually developed scientific method: complexity could be reduced, but the hierarchy that gave the reduction meaning could not be skipped. When operations, comparisons, and validation obligations are explicit, governed computation may examine more perturbations, doses, timepoints, modalities, and experiments without granting AI authority to truncate evidence, collapse intermediate state, convert association into mechanism, or promote research-use evidence beyond the authority the experiment earned.

Central thesis: The controlled-distillation principle developed to make large drug-response datasets scientifically useful became the governing principle for BENNETT. Large evidence spaces cannot simply be handed to an AI model and reduced for speed. Measurements, populations, experimental conditions, comparisons, interpretations, and conclusions must move through an explicit hierarchy. BENNETT applies that principle computationally: reduce complexity without allowing AI to skip the scientific structure that gives the result meaning.

SECTION 01

1. Paper 9 ended with a bounded record. Paper 10 asks what happens when many of them move together.

The first nine papers built toward an object that can sound less consequential than it is: a scientific record whose meaning survives the workflow that produced it.

Paper 9 established that publication must not create new authority. A report should preserve what evidence was used, what representation was measured, what methods acted, what the scientist changed or accepted, what validation applied, what limitations remained, and how far the final claim was allowed to go.^[1] The Publisher therefore does not hand Paper 10 a pile of attractive images, a table of disconnected numbers, or a collection of fluent summaries. It hands the final paper something far more useful: bounded scientific records that can be compared without erasing the route by which their authority was earned.

Paper 9's unresolved responsibility was whether collections of those records could support higher-order reasoning about biological and disease state without converting association into mechanism or exceeding assay readiness. Disease-oriented reasoning is therefore not discarded by the broader title of Paper 10. It is one consequential use of the larger problem addressed here: scientific intelligence at scale.

That changes the unit of reasoning.

Until now, much of the collection has asked what one experiment can support. Paper 10 asks what happens when the experiment becomes part of a larger evidence space. A control can be compared with a perturbation. One dose can be compared with another. A response can be followed across time. A population can be divided into responding and nonresponding cells. A broad phenotype can be tested at a finer resolving scale. Two drugs can be compared. A combination can be compared with its components. A repeated result can be challenged in another sample or biological context.

The scientific opportunity grows quickly. So does the possibility of error.

A large evidence space can create false confidence just as easily as a compelling image. A repeated association can begin to feel causal. A strong dose response can begin to sound therapeutic. A phenotype can begin to sound diagnostic. A statistically stable pattern can begin to sound mechanistic. A computationally convenient summary can detach from the cells, measurements, controls, and conditions that gave it meaning.

Paper 10 does not relax the discipline established in Papers 1-9. It extends that discipline into a much larger evidence space.

The final question is not, "Can AI find patterns across enough scientific data?" Of course it can find patterns.

The harder question is: can BENNETT relate large numbers of governed scientific records efficiently enough to change scientific cycle time while keeping every important conclusion accountable to the measurements and experimental conditions that produced it?

That is the horizon of scientific intelligence at scale.

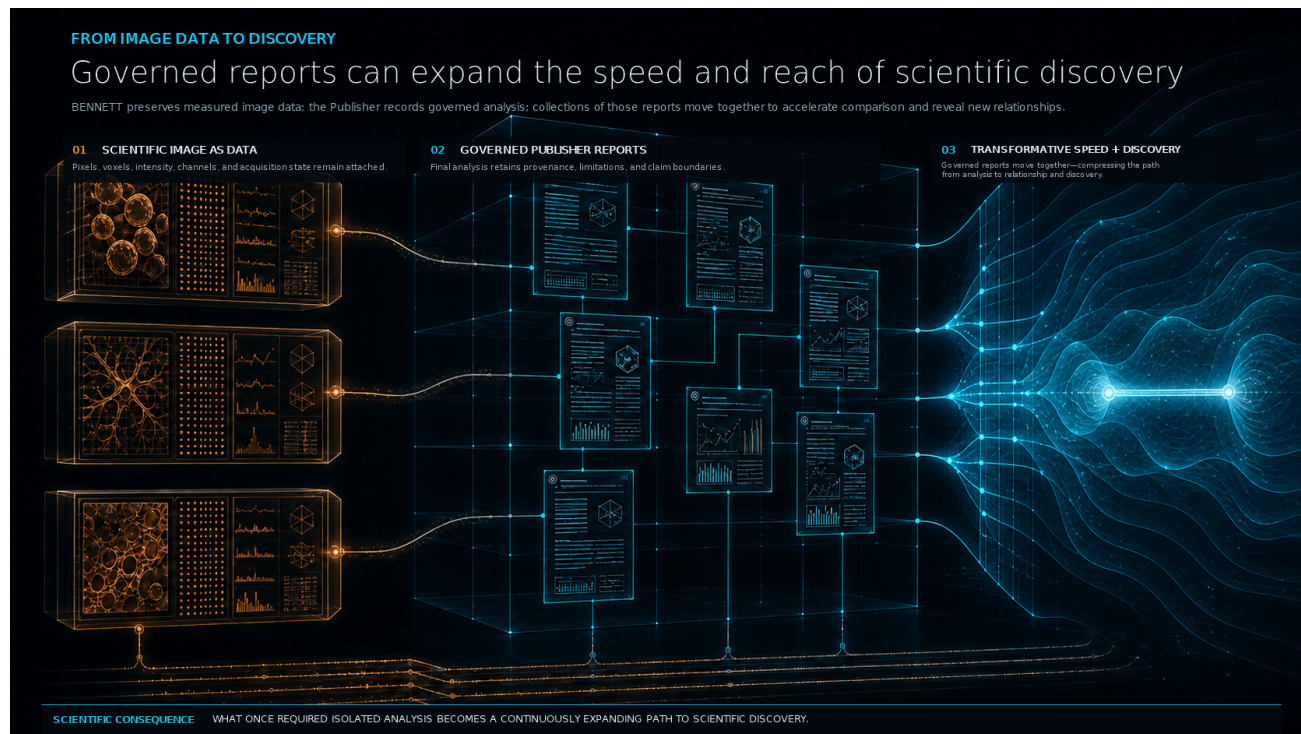


Figure 1. Conceptual research-use horizon. Governed scientific image data retain pixels, voxels, intensity, channels, acquisition state, and measurement context. The Publisher records the resulting governed analysis with provenance, limitations, and claim boundaries. Collections of those reports may accelerate comparison and reveal relationships across larger evidence spaces. This figure describes the intended scientific architecture; it does not claim that every population-scale, cross-experiment, cross-modality, or cross-domain route is currently implemented or independently validated.

SECTION 02

2. The scientific problem came before BENNETT.

It is tempting to tell the story of scientific AI as though modern models created the scientific opportunity. They did not. The scientific problems came first.

Long before today's AI systems, scientists were already confronting large data spaces, multidimensional measurements, population heterogeneity, dose and time dependence, changing evidence environments, and the need to reduce complexity without losing scientific meaning. Modern AI changes what can be done at scale; it does not create the scientific logic required to do it correctly.

For anyone building scientific AI, that order matters. A model can accelerate operations. It cannot retroactively define the experiment, decide which measurements were scientifically valid, repair missing controls, make incompatible evidence comparable, or determine how far a claim is allowed to go. Scientific AI has to inherit a scientific method. It cannot substitute for one.

In 2007, I was trying to answer a practical biological question that an individual image could not answer. If cells are exposed to a drug, what happens to the population? How many cells remain? How many show a defined molecular response? How strong is that response? Does it appear quickly or slowly? Does it persist? Does it resolve? Does the population change as dose increases? Does cell loss follow the molecular response? Does one treatment produce a different temporal or dose-dependent pattern than another?

Those questions immediately exceed the representative image.

A compelling field of cells can illustrate a response. It cannot, by itself, describe the response of the population. A mean can compress the population but hide whether the change occurred broadly or in a subset. A single endpoint can show that something happened while hiding whether the biology peaked earlier, resolved later, or never separated meaningfully from another dose. A viability readout can show population loss while saying little about the molecular response that preceded it.

The solution I was developing was to preserve the individual cell as the measurement object and then deliberately reduce the data in stages.[\[10,11\]](#)

But the data problem did not stop with that assay. I continued working through the population-scale processing and distillation problem manually through approximately 2012.[\[17\]](#) Then, from roughly 2012 through 2020, the super-resolution advances described in the earlier papers increased the information content of each experiment while high-throughput cellular work continued to increase the number of cells, conditions, doses, and timepoints that could be measured.[\[2-4,10,11\]](#) One direction increased experimental breadth. The other increased evidentiary depth. Combining both at meaningful scale was becoming almost impossible to imagine handling manually.

That distinction matters today because the core scientific idea was never simply to let a machine look at more pictures. The problem was how to preserve scientific meaning as both the number of experiments and the amount of information within each experiment increased.

I was not using what we now call modern AI. What I wanted was higher-order software capable of carrying the measurements, experimental state, comparisons, and reductions across an evidence space much larger than I could manage manually. The data were arriving faster than I could turn them into disciplined scientific answers. The reasoning task was already recognizable: preserve low-level measurements, carry their context forward, reduce them only through explicit operations, and use the resulting structure to decide what the experiment supports and what should be tested next.

That is why this history belongs in the final paper. BENNETT is not a model-first attempt to retrofit artificial intelligence onto microscopy. It is the continuation of a scientific operating problem that had already been defined manually.

The scientific method existed before the computational scale did.

SECTION 03

3. In 2007, the problem was already scientific distillation.

The word I used was "distillation." It remains the right word.[\[11\]](#)

I began with large populations of cells imaged with two established DNA-damage markers. For every included cell, I retained the molecular measurements while also retaining the population itself. That distinction was critical. I was not only asking whether the molecular response changed. I was also asking what was happening to the population as treatment took effect. Cell number, population loss, and treatment-associated cell death or apoptosis were part of the biological response, not background variables.[\[10,11\]](#)

That created a much richer scientific object than a single plate-level fluorescence value. I could ask how strongly the measured population responded, how much of the population met a defined response criterion, whether the responding population changed in strength, and whether the total population was stable or declining.

I then followed those relationships across time, drug, and dose. Molecular response could rise or fall. The fraction of responding cells could change. The response within that population could strengthen or weaken. At the same time, the total number of cells could decline as treatment-associated cell death increased. Looking at those variables together made it possible to distinguish a molecular response from a change in the population itself and to understand how both evolved across the experiment.[\[10,11\]](#)

I organized that transformation in six levels:

Level 0 was raw data collection. Level 1 was individual-cell profiling. Level 2 was cellular data integration, including molecular response and population state. Level 3 extracted time dependence. Level 4 extracted drug-and-dose dependence. Level 5 was the distillate.[11]

The final level was deliberately smaller than the data space that produced it.

That is the important point.

Scientific distillation is not data deletion. It is controlled reduction: a high-dimensional experimental record becomes a more focused scientific result while the route back to the underlying evidence remains recoverable.

I was therefore already asking the question Paper 10 now asks at a much larger scale: how do we move from thousands of measurements to one scientifically useful statement without allowing the statement to outrun the measurements?

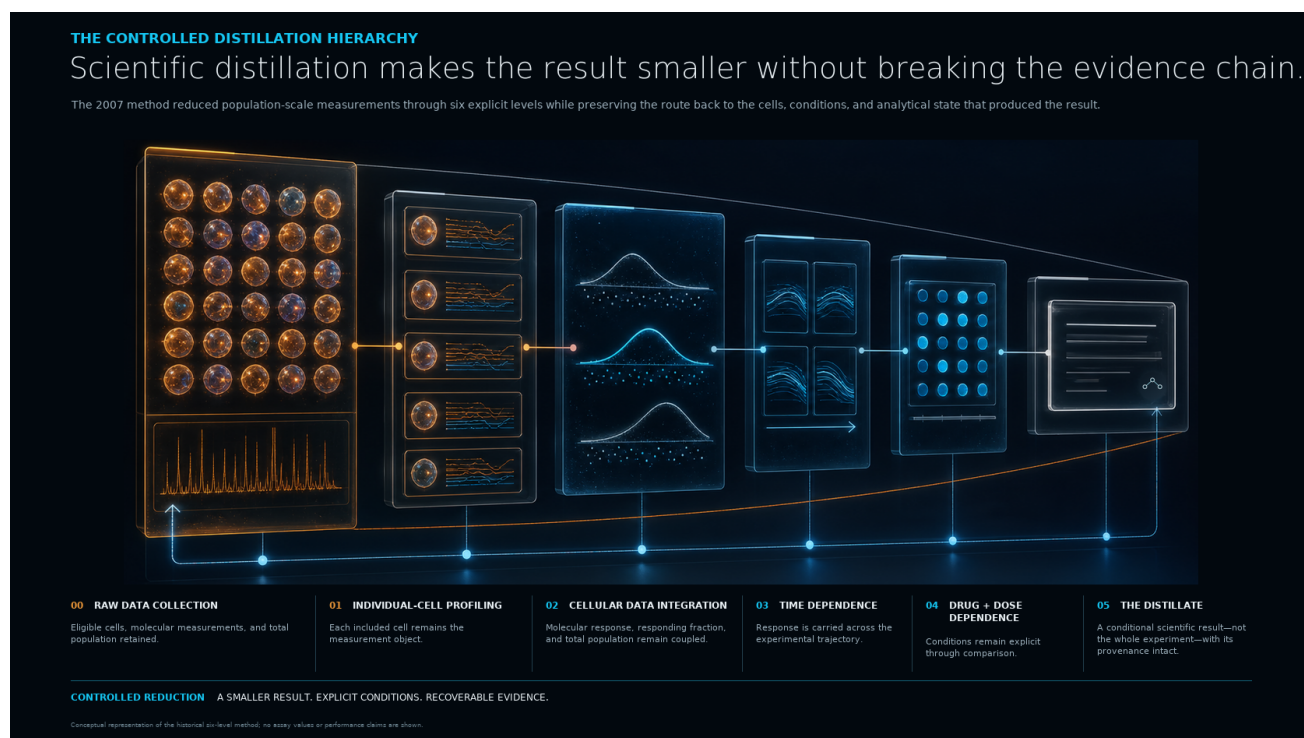


Figure 2. Controlled scientific distillation. The historical six-level method reduces eligible cells, molecular measurements, population state, time, drug, and dose into a smaller conditional scientific result while preserving the route back to the experiment. This is a conceptual representation; no assay values or performance claims are shown.

SECTION 04

4. The mathematics matters because every reduction answers a different question.

The arithmetic was never the difficult part. The scientific problem was deciding what could be reduced, under which conditions, and without losing the biological structure that gave the numbers meaning.

At the cell level, every included object carried measured signal. At the population level, those measurements could be reduced in different ways: an average across the population, the number or fraction of cells meeting a defined response criterion, the behavior of a selected response population, and the change in total population itself. Those quantities answer different questions. They are related, but they are not interchangeable.[\[10,11\]](#)

To express the governing logic conceptually—not to reproduce the original 2007 analytical formulation—consider:

$$D_j(t,p,x) = [\sum w_c(t,p,x) S_j(c,t,p,x)] / [\sum w_c(t,p,x)]$$

Here, D_j represents a conceptual distilled quantitative response for measurement j under time t , perturbation p , and dose or condition x . S_j represents the underlying cell-level measurement, and w_c represents an inclusion or weighting term defined by the active analytical state. This modern notation is illustrative rather than operational and is not presented as the equation used in the 2007 assay. The exact historical selection, thresholding, and reduction logic is not reproduced here.

The point is not the equation by itself. The point is that the distillate remains conditional on which cells were eligible, how the population was defined, which treatment, time, dose, and experimental state were active, and whether the same analytical rule was scientifically applicable across the conditions being compared.

A treatment can increase a population average because many cells change modestly or because a smaller subpopulation changes dramatically. A molecular response can strengthen while the total cell population declines. Two doses can produce similar endpoint values but very different temporal trajectories. A response that looks strong at one timepoint can disappear when the full time course is examined.

That is why the original work did not stop at a single plate-level value. The mathematics was a way of preserving biological structure inside the reduction.

A further distillation can then ask a narrower question. At what time does a meaningful separation appear? At that time, how does the response vary with dose? Which dose range produces useful separation before saturation? Does cell loss follow the molecular response or occur independently? Which condition should be carried into the next experiment?

The value of the hierarchy is that the answer at the top remains interpretable because the questions beneath it were explicit.

This is precisely the distinction that governed scientific AI must preserve. A computational system can perform arithmetic quickly. Scientific intelligence lies in knowing which reduction is valid, which population is active, which dimensions are comparable, which result is current, which limitation applies, and what the final distillate is actually allowed to mean.

SECTION 05

5. The bottleneck moved downstream: from acquisition to scientific continuity.

The 2007 assay was difficult not because addition, division, or curve extraction were difficult calculations. It was difficult because science is conditional.

Every dataset carried decisions. Were the cells identified correctly? Was the field representative? Did the analytical criterion separate a biologically meaningful population or simply divide a continuous distribution? Was a higher population response caused by more responding cells or stronger response within the same cells? Did a decrease in cell number represent treatment effect, plating variation, image-selection bias, or another experimental state? Was the selected timepoint actually the most informative point in the trajectory? Did an unexpected result represent biology, assay failure, or a condition that required another experiment?

I continued applying this kind of reasoning manually through approximately 2012.[\[17\]](#)

The first bottleneck was volume. High-throughput hardware was improving the rate at which we could generate data, but the analytical side was not keeping pace. We could acquire more than I could responsibly process, compare, and reduce by hand. I even remember overheating computers trying to push the analysis through. The anecdote is almost funny now, but the scientific problem was not: data generation was scaling faster than our ability to turn the data into disciplined scientific understanding.

Then a second problem grew alongside it.

From roughly 2012 through 2020, super-resolution microscopy expanded rapidly. New modalities and applications produced increasingly detailed and increasingly large datasets. The problem was no longer only how much data existed. The data were also becoming fragmented across instruments, modalities, reconstruction paths, and vendor software environments.[\[2-4,17\]](#)

When Identifyn began in 2023, that silo problem became an operational scientific problem for us. We could move from Zeiss widefield to widefield structured illumination, confocal, Airyscan, and Lattice SIM within the broader ZEN environment. ZEN was powerful within those modalities, but it did not perform the scientific Resolution Stepdown for us. It did not

decide what remained comparable across the resolution scale, when an apparent object had changed meaning, or how a claim should change when the evidence environment changed. Our single-molecule work on the Vutara platform created another highly detailed evidence tier. My experience with Vutara went back to beta-development work, but as a single-molecule analysis environment it still did not solve the cross-modality problem.[\[4,17\]](#)

Each system could produce exceptional data. What was missing was a common scientific language capable of relating those data without erasing the differences among them.

That meant the resolution walkdown still had to happen manually. I had to carry the scientific context from one modality to the next, decide what remained comparable, determine when the measurement object had changed, and decide what the higher-resolution evidence actually allowed us to say. At Identifyn, I performed that Resolution Stepdown manually across the supported systems. It worked, but it was slow, difficult to scale, and difficult to separate from the scientist who knew how the evidence fit together.[\[3,4,17\]](#)

By then the problem had three dimensions: more data than manual analysis could comfortably process, more information within each experiment, and scientific evidence trapped in computational silos that did not share a governed language.

That is the problem BENNETT is designed to change.

SECTION 06

6. BENNETT changes the scale, not the source of authority.

BENNETT was therefore not designed simply to calculate faster. It was designed to preserve scientific continuity across data that had never been built to speak to one another while also carrying repeatable analytical work at computational scale.

The objective is not to flatten unlike evidence into one artificial representation. Widefield, confocal, Airyscan, structured illumination, Lattice SIM, and single-molecule localization do not become equivalent because they enter the same platform. The objective is to give them a common governed language while preserving the authority, limitations, and scientific meaning of each evidence tier.

Papers 5-9 established the architecture required for that transition. The native acquisition record remains authoritative for what the instrument recorded.[\[6\]](#) Expert practice is decomposed into evidence requirements, states, applicability conditions, gates, negative paths, and validation rather than imitated from a fluent answer.[\[7\]](#) The scientist can take direct control when legitimate judgment is required without leaving the evidence route.[\[8\]](#) Analytical responsibility is divided into bounded components rather than concentrated in one opaque model.[\[9\]](#) And the Publisher preserves the resulting state without converting reporting into a new source of truth.[\[1\]](#)

The same controlled-distillation mindset now applies to AI itself. A model should not be allowed to jump from raw evidence to a persuasive conclusion simply because doing so is fast. The intermediate scientific state has to remain explicit: what was measured, which population was active, which comparison was valid, which evidence tier was being used, what changed, what remained uncertain, and how far the conclusion was allowed to go.

That is the governing difference between automation and governed scientific computation.

Automation asks whether the process can run without a person performing every step.

Governed scientific computation asks whether the process can run at scale without losing the scientific conditions that make the result defensible.

The status boundary must remain explicit. The assay-distillation method began as manual work in 2007, and I continued working through that population-scale processing problem manually through approximately 2012.[\[10,11,17\]](#) From roughly 2012 through 2020, the rapid expansion of super-resolution modalities added a second problem: richer datasets remained separated by instrument, reconstruction path, and software environment.[\[2-4,17\]](#) After Identifyn began in 2023, Resolution Stepdown remained a manual scientific task across those evidence tiers.[\[3,4,17\]](#) BENNETT currently operates supported native-data workflows internally across multiple microscope brands and file types, with deterministic analysis, governed evidence state, and controlled output behavior within its present internal validation scope. UX completion remains before public release, and independent academic and industry testing remains a planned external validation step. These are the current operating and development states established by the preceding papers.[\[2,6-9\]](#)

By contrast, the high-throughput, cross-experiment, population-, dose-, time-, and cross-modality synthesis described in this final paper is the scientific and architectural horizon. It should not be read as a claim that every route, assay family, comparative model, or cross-disciplinary use has already been implemented, independently tested, or released. Clinical diagnosis, treatment selection, patient-level prediction, and clinical utility are not claimed.

Within a route that has been implemented and validated for its intended use, computation may count included cells, measure signal, retain population state, calculate defined response quantities, construct temporal or dose-dependent relationships, and compare scientifically compatible conditions. A governed implementation must also preserve which source, sample, treatment, dose, timepoint, channel, region, object population, measurement method, resolution tier, and validation state produced each result.

But speed does not authorize the result.

The authority still comes from the experiment, the evidence, the method, the relevant validation, and the qualified scientist where judgment remains necessary.

BENNETT is designed around that boundary.

SECTION 07

7. High-throughput science is only useful if governance scales with it.

High throughput is often treated as a capacity problem: process more wells, more images, more objects, more compounds, or more samples per unit time.

For BENNETT, throughput has a second dimension.

The evidence obligations have to scale too.

If one thousand cells become one million cells, the system should not become less precise about which sample they came from. If a few doses become a full concentration series, dose identity and units should not become easier to lose. If one timepoint becomes a time course, temporal order should not become a presentation convenience. If one modality becomes several, the system should not silently collapse the resolution tiers into a single undifferentiated evidence class. If one analytical route becomes hundreds or thousands of repeated routes, method version, eligibility state, exclusions, scientist interventions, and validation should not disappear into aggregate statistics.

Speed without those controls can accelerate the wrong thing.

A poorly defined population can be classified incorrectly faster. An invalid comparison can be repeated across more samples faster. A hidden calibration or acquisition problem can contaminate more measurements faster. A persuasive but unsupported biological interpretation can propagate across more reports faster. A model can also omit intermediate evidence simply because omission makes the path to an answer shorter.

That is why the central future claim should not be "BENNETT makes science faster" without qualification.

The stronger claim is that BENNETT is designed to make governed science faster.

The system should accelerate repeatable work precisely because it preserves the conditions that allow a scientist to trust, challenge, reproduce, rerun, or reject the result. Once the evidence route is explicit, the scientist does not have to reconstruct the same context manually for every well, image, sample, modality, or comparison. The system can carry forward what remains current, expose what is missing, block what is unsupported, and return the scientist to the points where judgment is scientifically necessary.

The result is not fewer scientific decisions.

It is less time wasted reconstructing the state required to make them.

SECTION 08

8. Population response is often more informative than the representative image.

The 2007 work taught me a lesson that remains central to high-content biology: the population is not reducible to its representative cell.

A single image can make a phenotype look universal. Population measurement can reveal that only a subset of cells responds. The reverse can also matter: a molecular response can appear strong in the cells that remain while the total population is declining because of treatment-associated cell death.

That is why I preserved both the molecular measurements and the population itself. I tracked the number of included cells, the fraction meeting a defined response criterion, the behavior within that response population, and changes in total cell number across time and treatment.[\[10,11\]](#) Population loss was not a nuisance variable. It could be part of the biological response.

This distinction can be biologically important. Heterogeneity may reflect cell-cycle state, genetic or epigenetic variation, treatment exposure, microenvironment, stochastic signaling, resistant subpopulations, or another condition that deserves investigation. A mean can conceal that structure. A population fraction can expose it. A distribution can expose it more clearly still.

Modern image-based profiling has expanded this principle dramatically. Cell Painting and related methods can extract hundreds or thousands of cellular features and compare phenotypic states across chemical and genetic perturbations.[\[12-15\]](#) Those approaches demonstrate the amount of biological information that can be represented in image-derived measurements. They also reinforce the need for restraint: phenotypic similarity can be highly informative without establishing that two perturbations share the same mechanism.[\[14,15\]](#)

BENNETT's contribution is not simply to add another feature vector.

It is to keep the response connected to the complete experiment.

The important population question may be simple or complex. How many cells met the defined response criterion? How strongly? What happened to the total population? In which compartment? With what object burden? With what temporal profile? In which subpopulation? At which dose? Under which modality? After which scientific exclusions? Compared with what control? Under which accepted analytical state?

Once those dimensions are explicit, population response becomes more than a descriptive statistic.

It becomes a governed scientific object that may be compared across experiments when the experimental units, measurement definitions, controls, and relevant technical states are scientifically compatible.

SECTION 09

9. Scale requires experimental hierarchy, not merely more cells.

Single-cell measurement increases resolution within a population. It does not, by itself, increase the number of independent experiments.

Cells may be nested within fields, fields within wells, wells within plates, plates within acquisition runs or batches, and observations within specimens, donors, preparations, or biological replicates. Those levels do not carry interchangeable statistical authority. Ten thousand cells measured from one well do not automatically provide the same evidence as independent biological replication across specimens or experiments.

This distinction becomes more important as throughput increases. A system can calculate cell-level distributions with great precision while remaining uncertain about whether the observed difference generalizes beyond one well, plate, preparation, batch, or biological source. Treating every measured cell as an independent replicate can therefore produce apparent confidence that the experimental design did not earn.

A governed population record must preserve the experimental hierarchy beneath the aggregate. Depending on the intended comparison, that record may include:

- the biological unit and technical unit of replication; - specimen, donor, preparation, well, plate, field, and acquisition-run identity where applicable; - the number of eligible and excluded units at each level; - the rule and reason for exclusion; - missing conditions, failed acquisitions, segmentation failures, and non-evaluable states; - plate position, batch, instrument, acquisition, reagent-lot, and processing differences that may affect comparability; - normalization or correction methods and the evidence that authorizes them; - the statistical unit used for estimation and inference; and - the relationship between cell-level variation and variation among independent biological replicates.

Missing data—and the reasons they are missing—must remain visible. A condition with no eligible cells is not automatically a zero response. A well excluded because of an acquisition or segmentation failure is not equivalent to a biological negative. A field removed under a prespecified quality rule is different from a field removed after its result became inconvenient. A batch correction may improve comparability, but it does not erase the batch or make incompatible experiments equivalent by declaration.

Scale also increases the number of analytical opportunities. Hundreds or thousands of features, many doses, timepoints, populations, and perturbations can produce many apparently interesting relationships. Exploratory pattern detection is scientifically legitimate,

but the resulting multiplicity, model selection, threshold selection, and overfitting risk must remain attached to the claim. A pattern used to define an assay candidate should be challenged in evidence not used to select it and, where the intended use requires it, in independent samples, laboratories, or studies.

The governing principle is therefore not that more cells create more authority.

More cells can reveal population structure. More independent and appropriately controlled experiments establish whether that structure persists beyond the conditions in which it was first observed.

SECTION 10

10. Dose and time turn snapshots into response surfaces.

A single endpoint can be scientifically useful, but it is also dangerous because it can make a dynamic process look static.

I treated time as a first-class variable.[\[10,11\]](#)

That decision changed the question from "Did the treatment cause a response?" to "How did the response develop?"

A molecular marker may increase quickly and then resolve. Another may accumulate slowly. A population may show strong signaling before cell number changes. A high dose may saturate a marker so completely that it becomes less informative than a lower dose. Two treatments may produce similar endpoint values while following very different trajectories. A combination may shift the timing or magnitude of response without producing a simple additive effect.

Dose adds another dimension.

At one dose, the assay may remain close to baseline. At another, the response may separate cleanly. At a higher dose, the readout may saturate or the measured cell count may decline, making the selected measurement less informative for the scientific question. The useful dose is therefore not automatically the highest dose and the useful timepoint is not automatically the last timepoint.

I deliberately searched these relationships. Time dependence was extracted before drug-and-dose dependence was distilled.[\[11\]](#)

That ordering remains scientifically important.

A computational system can evaluate a large response surface quickly, but the scientific objective is not to find the largest number. It is to identify the region of the experimental space that best resolves the biological question under the relevant constraints.

This is where high-throughput governed analysis can become scientifically consequential.

Instead of reducing an experiment prematurely to one endpoint, the system can retain the trajectory, compare the response across the condition space, and then distill only after the relevant scientific structure is visible.

That can make follow-up experiments smarter because the next experiment begins from a richer understanding of where the response exists, where it saturates, where it diverges, and where the evidence is still ambiguous.

SECTION 11

11. Drug discovery is an obvious application — not the limit of the architecture.

The 2007 assay makes drug discovery the clearest example because the scientific variables are explicit: treatment, compound, dose, time, molecular response, population response, cell number, and comparison.[\[10,11\]](#)

The potential value is easy to see.

A drug-discovery program does not need only a list of compounds that changed an endpoint. It needs to understand how they changed it. Did the effect appear early or late? Was it dose dependent? Was the molecular response separated from a decline in measured cell count? Did only a defined subpopulation meet the response criterion? Did two markers agree or diverge? Did a low dose produce a clean biological separation before a high dose saturated the assay? Did two compounds produce similar population behavior through different temporal routes? Did a combination produce a response that merits a focused follow-up experiment?

These are scientific questions, not merely screening questions.

A governed platform capable of carrying them across large numbers of conditions can change where human attention is spent. Instead of manually reducing every field and condition before a scientist can see the pattern, computation can perform the repeatable measurements and preserve the response structure. The scientist can then spend more time on interpretation, anomaly review, experimental design, competing hypotheses, and the next experiment.

The important future statement is therefore not that BENNETT will "discover drugs" autonomously, nor that it will discover across every discipline by declaration.

That would give the system authority that the evidence does not support.

The more defensible and more consequential statement is that BENNETT is designed to expand the amount of governed experimental space a scientific team can traverse in a given period of time.

More compounds can be compared. More doses and timepoints can remain visible. More combinations can be examined. More cellular subpopulations can be retained. More

experiments can be related before a scientist has to decide where to invest the next round of bench work.

That can accelerate discovery without pretending that the computational result is the discovery itself.

And the same pattern can extend well beyond drug discovery.

Any research domain with controlled experimental variation and measurable responses may be able to use a similar architecture once its own evidence requirements, methods, failure states, and validation obligations are defined. Cell biology, pharmacology, toxicology, pathway biology, regenerative biology, infectious-disease models, environmental response, materials and imaging sciences, and other quantitative disciplines all contain versions of the same problem: a system is placed into a defined state or perturbation, a response is measured, and the scientist must decide what the pattern supports.

The generality comes from the governing method, not from pretending every discipline shares the same assay, measurement, or claim boundary.

SECTION 12

12. Resolution and modality can refine a response without becoming a truth tier.

Paper 3 established that increasing resolution does not simply reveal a truer version of the same image.[\[4\]](#)

That lesson becomes especially important when high-throughput response patterns are used to guide deeper investigation.

A broad screen may identify a population-level response at conventional resolution. That response can be entirely valid for the question being asked. A subset of conditions may then be examined with optical sectioning, Airyscan, structured illumination, SIM², or single-molecule localization to determine whether the apparent object class, spatial relationship, compartmental organization, or localization behavior changes at a finer evidence tier.

The higher-resolution experiment does not automatically invalidate the lower-resolution one. It may confirm the broader response. It may refine it. It may show that one apparent object is several. It may reveal axial separation hidden in projection. It may change the measurement object entirely. Or it may show that the cross-tier comparison is not scientifically evaluable.[\[4\]](#)

The silo problem described earlier is exactly why this cannot be left to file compatibility alone. A common computational environment is useful only if it preserves the scientific differences among the modalities. The goal is not to make a widefield image, an Airyscan image, a Lattice SIM reconstruction, and a single-molecule localization dataset look like equivalent inputs. The goal is to preserve what each evidence tier can support and govern the transition from one tier to the next.

For Paper 10, this creates an important future workflow.

High-throughput measurement can identify where the biology changes. Resolution Stepdown can then help determine what the change actually means at another evidence scale.

This is a more efficient use of demanding microscopy because the most information-rich modality does not need to become the first instrument used for every condition. The broader experiment can identify where higher-resolution evidence is scientifically valuable, and the finer evidence can return a more specific statement without being allowed to rewrite the entire population record by visual authority alone.

The sequence is not "screen, then prove."

It is "screen, measure, identify, escalate, test, and revise the claim according to what the new evidence actually supports."

That is Resolution Stepdown operating inside a larger discovery process.

SECTION 13

13. Distillation can produce a focused scientific output without erasing the evidence beneath it.

One of the most important words in the 2007 work is still the word that should govern the future: distillate.[\[11\]](#)

A distillate is not the whole experiment.

It is the smaller result that has survived a defined reduction of the experiment.

That distinction is also central to AI control. A general-purpose model can produce a concise answer by omitting intermediate material. Scientific distillation cannot work that way. The reduction has to be controlled. Each step must preserve the evidence, state, and dependencies required to understand how the smaller result was earned. The model does not get to decide that a piece of evidence is unimportant merely because skipping it makes the answer faster or cleaner.

Assay development is one natural destination for governed scientific distillation, but it is not the only one. A distillate can also be a bounded response signature, a prioritized condition, a comparison worth repeating, a narrowed hypothesis, a candidate measurement panel, or another defined scientific object that is more useful precisely because its relationship to the larger experiment remains recoverable.

If a response repeatedly survives across appropriate controls, replicates, samples, conditions, doses, timepoints, and analytical states, the full discovery workflow may eventually be narrowed into a more focused research-use assay or other repeatable scientific output. The narrowed form may require fewer variables because prior work has established which

measurement, time window, dose range, marker combination, population definition, spatial relationship, or other feature carries the relevant information.

This narrowing is not a loss of rigor if the lineage remains preserved.

It is the point of distillation.

A mature distilled output should be able to explain what it measures, which population or experimental state it applies to, which controls and conditions it requires, which analytical definition is fixed, what performance has been established for the intended use, what limitations remain, and what downstream claim the result is allowed to support.[\[16\]](#)

That is very different from allowing an AI system to discover a correlation and immediately promote it to a biomarker, mechanism, diagnostic, therapeutic predictor, materials property, or other higher-order claim.

A research-use assay candidate is one example of a bounded scientific object produced by distillation. Its transition into a biomarker, predictive tool, diagnostic, clinical decision support system, regulated product, or any other higher-authority use would require additional evidence appropriate to that intended use. Paper 10 does not claim those transitions have already occurred.

The important architectural point is that BENNETT is designed to preserve the route by which a focused scientific output was distilled.

The final output does not have to forget the thousands of measurements, cells or objects, conditions, comparisons, exclusions, evidence tiers, and validation states that made the narrowing scientifically defensible.

That is where speed and governance reinforce one another. Once a validated experimental pattern can be distilled into a focused form, subsequent studies can test it efficiently while the evidence chain remains available for challenge, revision, or expansion.

SECTION 14

14. Scientific intelligence is a hierarchy of meaning.

Scientific intelligence is useful only if the relationships among evidence, observation, measurement, interpretation, conclusion, hypothesis, and claim remain explicit.

A numerical output is not automatically a valid measurement merely because a calculation produced it.

An observation is not a measurement unless a defined method establishes what was measured and under which conditions.

A measurement is not an interpretation merely because its numerical value appears persuasive.

A repeated observation or measured association is not a causal mechanism merely because it is biologically or physically plausible.

A hypothesis is not a validated assay or model simply because it is testable.

A reproducible output does not automatically inherit authority for a broader use.

The entire BENNETT collection has been building the separations required to prevent those transitions from occurring silently.

At the base is the experiment: the defined scientific state, perturbation or comparison, sample or specimen, relevant materials, preparation, acquisition, and native evidence.[\[2-6\]](#)

That evidence can support source-linked observations. Defined and eligible methods can produce bounded measurements from the evidence. Observations and measurements may support an interpretation. Repeated observations and measurements across appropriate independent experimental units may support a bounded conclusion. A conclusion can motivate a hypothesis. The hypothesis then returns to experiment.

The hierarchy matters because controlled distillation is not permission to skip levels. BENNETT should not move directly from raw evidence to mechanism, from a measured phenotype to disease causality, or from an interesting response pattern to a validated assay simply because a model can generate the intermediate language. Each transition has to be earned by the evidence and methods appropriate to that level.

That return to experiment is critical.

Scientific intelligence is not complete when the system generates the most persuasive explanation. It is useful when the system helps identify what is supported now, what remains uncertain, what competing explanation remains plausible, and what next experiment would discriminate among them.

The cycle therefore remains scientific:

measure, compare, distill, interpret, hypothesize, test, and revise.

BENNETT is designed to reduce the time consumed by repeatable, evidence-preserving work because the architecture can retain more of the state required to repeat the cycle. It does not make the cycle unnecessary, and the scale of that benefit remains subject to validation in each implemented use.

SECTION 15

15. QUESTION / ANSWER

QUESTION Does accelerating this process simply automate scientific bias and error at a larger scale?

ANSWER It can—unless speed is subordinate to the same evidence, applicability, provenance, validation, and scientist-control obligations that govern the slower manual process.

The risk is real. A biased threshold can be applied to thousands of images faster than a scientist could apply it manually. A flawed assumption can propagate across an entire screen. A language model can produce a coherent explanation for a spurious pattern. High throughput magnifies both good methodology and bad methodology.

There is another risk that matters just as much: truncation. A general-purpose model can move quickly toward an answer without preserving every scientific dependency that made the answer defensible. In ordinary conversation that compression may be useful. In science, silent compression can erase the very state that determines whether the conclusion is valid.

That is why BENNETT's objective is not autonomous scientific throughput at any cost.

Stable scientific logic belongs in explicit governed structures rather than in model improvisation.^[7-9] Measurement remains deterministic where the method is deterministic. Applicability must be resolved before execution is treated as valid. Negative paths remain legitimate scientific outcomes. Scientist judgment remains explicit where legitimate discretion is required. Changes that alter analytical state must invalidate or revalidate dependent results rather than allowing stale authority to continue. Publisher output carries the limits forward rather than hiding them.^[1,7-9]

The safeguard is architectural.

Speed is valuable only when the system can still answer: which evidence entered, which rule or method acted, what state changed, which result depends on that state, what the scientist did, what validation applies, and how far the conclusion is allowed to go.

BENNETT should make those questions easier to answer at scale, not easier to forget.

SECTION 16

16. The future is faster governed science.

The most important future statement in this paper is not about a single assay, disease area, instrument, or scientific discipline.

It is about scientific cycle time.

The chronology matters. In 2007, the population-distillation method showed that the scientific hierarchy already worked: individual measurement, population, time, dose or condition, and distillate.[10,11] I continued working through that problem manually through approximately 2012.[17] From roughly 2012 through 2020, the burden expanded again as super-resolution modalities generated richer datasets that remained separated by instrument, reconstruction path, and software environment.[2-4,17] After Identifyn began in 2023, Resolution Stepdown provided the scientific bridge across those tiers, but that bridge was still largely manual.[3,4,17]

BENNETT is intended to change that relationship.

When repeatable scientific work is represented explicitly and its route is validated for the intended use, computation can operate over more evidence than a scientist can reasonably process by hand. The governed system can preserve experimental identity, execute bounded measurements, maintain population and temporal state, retain modality and resolution identity, compare defined conditions, identify where a result becomes stale, expose negative routes, and assemble the current scientific record for review.

The scientist can then operate at a higher level without becoming detached from the evidence.

This is not simply labor savings.

It changes what experiments become practical.

A program that once had to choose a small number of conditions because analysis was expensive may be able to preserve a richer experimental matrix. A study that once used one endpoint because manual reduction was slow may be able to retain a more informative time course. A screen that once reported an average may be able to preserve subpopulation behavior and population loss. A discovery program may be able to compare more prior experiments and request higher-resolution or orthogonal evidence only where it is most informative.

The potential scientific gain is not that every answer arrives instantly.

It is that less scientific time can be consumed by repeatable analytical work and by manually reconstructing relationships among disconnected evidence environments before the next scientific question can be asked.

That difference compounds.

Earlier measurement enables earlier comparison. Earlier comparison enables earlier hypothesis refinement. Earlier hypothesis refinement enables better targeted follow-up. Negative results can stop weak directions sooner. Strong patterns can be challenged sooner. Higher-resolution or orthogonal evidence can be requested where it is most informative. Focused assays, response signatures, or other distilled outputs can emerge from a larger and more thoroughly explored evidence space.

This is how BENNETT could contribute to drug discovery and other scientific disciplines without claiming to replace the science. The extent of that contribution must be established within each implemented and validated domain.

The objective is not faster answers.

It is faster governed science.

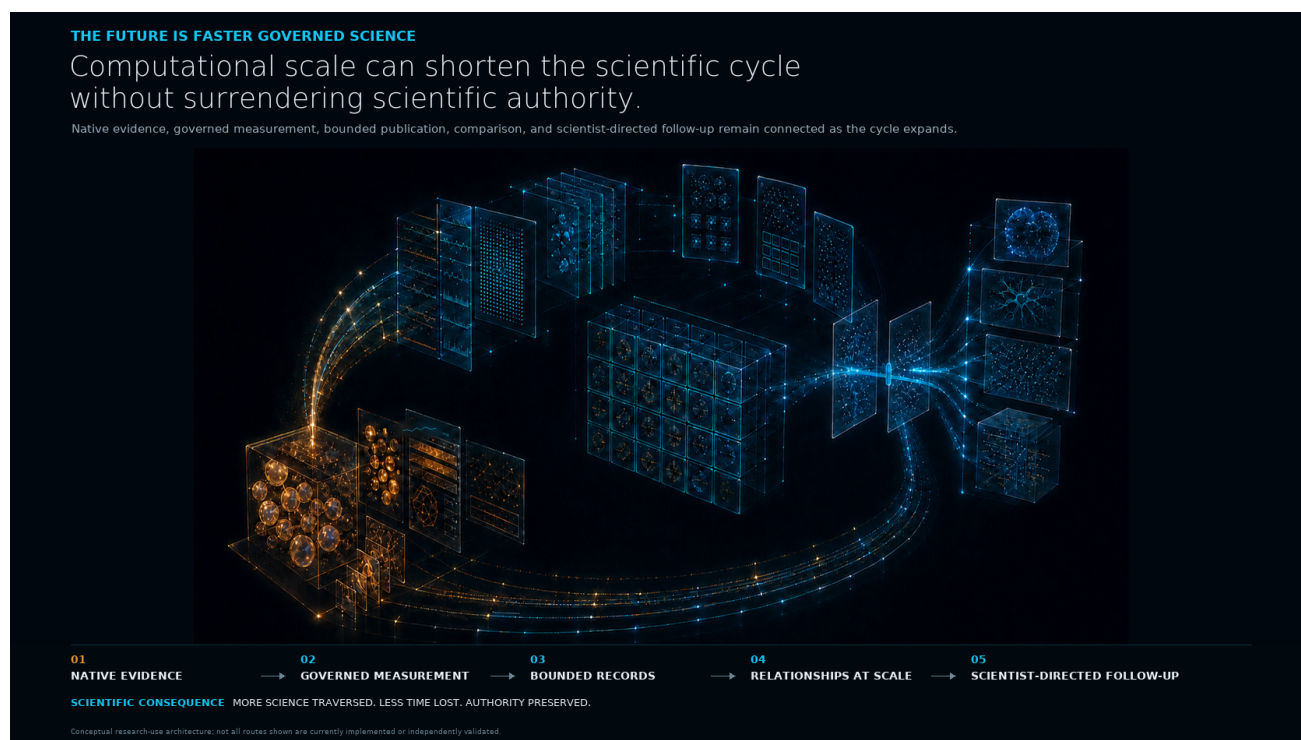


Figure 3. Faster governed science. Native evidence, governed measurement, bounded records, relationships at scale, and scientist-directed follow-up remain connected as computational scale shortens the scientific cycle. This is a conceptual research-use architecture; not all routes shown are currently implemented or independently validated.

SECTION 17

17. What the ten papers now add up to.

The collection began with an optical problem and ends with a problem of scientific scale.

The early 4Pi work showed that seeing more can change what a scientist is entitled to say.[\[2\]](#) Higher resolution separated fluorescence distributions that lower-resolution evidence had merged. That created a burden: if the evidence environment could change the interpretation, then the experiment had to remain connected to the claim.

Paper 2 expanded the unit of science from the image to the complete experiment.[\[3\]](#)

Paper 3 showed that an observation must survive a controlled change in evidence environment before its meaning can be generalized across resolution.[\[4\]](#)

Paper 4 moved upstream and established that the material system creating the fluorescence must be scientifically competent for the intended use.[\[5\]](#)

Paper 5 established the native acquisition record as the computational authority for what the instrument recorded without confusing acquisition authority with absolute biological truth.[\[6\]](#)

Paper 6 converted expert practice from tacit experience into explicit scientific state, applicability, gates, negative paths, and validation.[\[7\]](#)

Paper 7 preserved scientist authority inside the system rather than treating human intervention as an exception.[\[8\]](#)

Paper 8 decomposed analytical responsibility so that scientific capability could expand without concentrating authority in one monolithic model.[\[9\]](#)

Paper 9 carried the resulting route into a defensible scientific record without allowing publication to create new authority.[\[1\]](#)

Paper 10 reveals why all of those controls matter at scale.

The final problem is not how to make one image more intelligent.

It is how to let science move through more cells, more images, more measurements, more perturbations, more conditions, more doses, more timepoints, more resolution tiers, more experiments, and eventually more disciplines without losing the evidence chain that makes each conclusion trustworthy.

The chronology now reads as one continuous problem. The 4Pi work established the burden of seeing more.[\[2\]](#) The 2007 assay established the burden of measuring and distilling more, and I continued working through that problem manually through approximately 2012.[\[10,11,17\]](#) From roughly 2012 through 2020, the expansion of super-resolution modalities made the data richer while leaving the scientific relationship among many evidence tiers fragmented across instruments and software environments.[\[3,4,17\]](#) When Identifyn began in 2023, that

fragmentation became an operational problem: Resolution Stepdown could connect the tiers scientifically, but it still depended on substantial expert effort. BENNETT is the attempt to make that discipline computationally scalable without giving the computational system permission to erase the distinctions that made the discipline scientific.

That is the continuity from the first paper to the last.

The burden of seeing more became the burden of measuring more. The burden of measuring more became the burden of relating evidence across experiments, instruments, and scales. That burden became the need for governance. Governance is what makes scale scientifically plausible.

And scale changes the future.

CONCLUSION

Conclusion

The final proposition is not that artificial intelligence can analyze an image. Computation already does that.

The more consequential proposition is that governed computation may expand the amount of repeatable scientific work that can be performed without forcing the scientist to surrender the evidence chain that makes the work defensible.

The scientific logic for that proposition did not arrive with modern AI. The collection has shown how experimental scale and scientific continuity led from manual distillation and Resolution Stepdown to the governed architecture BENNETT now applies computationally.[1-11,17] Paper 10 adds the obligation of scale: cell-level measurements, population heterogeneity, missing data and their causes, independent replication, dose, time, modality, resolution tier, and comparative state must survive reduction. Higher-order reasoning may reveal patterns and support hypotheses, but it may not convert association into mechanism, an assay candidate into a validated test, or research-use evidence into clinical authority.

If implemented and validated routes shorten that evidence-preserving work, more experimental space can be examined before the next bench decision. Weak directions can be stopped earlier, strong patterns challenged sooner, and scientists can spend less time reconstructing prior analytical state and more time designing the experiment that answers the next question.

The scientific problem came first.

The method came first.

The mathematics came first.

The manual work came first.

Only then did computation become capable of carrying that discipline at a different scale.

That sequence matters because artificial intelligence can now move from data to explanation with extraordinary speed, while questions about what it says, what it omits, and what authority people assign to its answers remain unresolved. In science, speed can conceal the most important questions.

Where did the evidence come from?

What was actually measured?

Which assumptions entered the analysis?

What information was excluded?

What remains uncertain?

When should the system stop rather than answer?

And when must control return to the scientist?

These are not secondary questions. They may determine whether artificial intelligence accelerates scientific understanding or merely accelerates confidence.

What if scientific AI could move quickly without abandoning those questions?

What if it could preserve the route from native evidence to every measurement, comparison, interpretation, and claim? What if it could recognize when two experiments were not comparable, when a result had become stale, when an association was not a mechanism, or when the evidence was simply not strong enough to continue? What if greater computational power produced not only faster analysis, but better scientific restraint?

That possibility is larger than image analysis. It is larger than microscopy, drug discovery, or any single scientific discipline.

BENNETT may represent an early practical step toward that broader idea of governed AI: artificial intelligence that does not acquire scientific authority merely because it can generate an answer, but operates within an architecture that preserves evidence, exposes limitations, records human judgment, and keeps the scientist in control of meaning.

The promise is not science without scientists.

It is science in which computational scale no longer requires the surrender of scientific discipline.

If that becomes possible, the most important measure of scientific AI may not be how many answers it can produce or how quickly it can produce them. It may be how much complex scientific work it can help us traverse without losing the evidence, uncertainty, and human responsibility that make the result worthy of trust.

Paper 1 began with the burden created by seeing more.

Paper 10 ends with a question:

What becomes possible when science can measure more, relate more, compare more, and understand more—without leaving the evidence behind?

Scientific instruments produce data.

BENNETT produces understanding by preserving and governing the route from measurement to meaning.

The series ends here.

The larger work of governed scientific intelligence is only beginning.

REFERENCES

References

1. Bennett BT. The BENNETT Scientific Publisher. BENNETT White Paper 09. Authoritative locked PDF amendment v1.5. 2026. https://486b4e8f-f0c1-4b95-8447-81ae78bc5540.filesusr.com/ugd/175885_a785be5ad5ad496a9d7af9c146cb4ef9.pdf
2. Bennett BT. From 4Pi Microscopy to Governed Scientific Intelligence. BENNETT White Paper 01. Canonical locked v1.3. 2026. https://486b4e8f-f0c1-4b95-8447-81ae78bc5540.usrfiles.com/ugd/175885_e1d23cf6bee342279b8386aee3f0df94.pdf
3. Bennett BT. The Complete Experiment. BENNETT White Paper 02. Canonical locked v1.3. 2026. https://486b4e8f-f0c1-4b95-8447-81ae78bc5540.usrfiles.com/ugd/175885_e10f44f16e9b4a98ba19c5596f490715.pdf
4. Bennett BT. Resolution Stepdown. BENNETT White Paper 03. Canonical locked v1.1. 2026. https://486b4e8f-f0c1-4b95-8447-81ae78bc5540.filesusr.com/ugd/175885_e1b203fc28f54cf5a66be6adb01c4b3f.pdf
5. Bennett BT. Antibodies You Can Trust. Data You Can Defend. BENNETT White Paper 04. Canonical locked v1.1. 2026. https://486b4e8f-f0c1-4b95-8447-81ae78bc5540.filesusr.com/ugd/175885_0ca91b9a2ade459188a663e57fd8f8f7.pdf
6. Bennett BT. The Native File as the Acquisition GroundTruth™. BENNETT White Paper 05. Canonical locked v1.0. 2026. https://486b4e8f-f0c1-4b95-8447-81ae78bc5540.usrfiles.com/ugd/175885_c0be36ea34504c20bc50839fdc6738ec.pdf
7. Bennett BT. From Expert Practice to Governed Scientific Intelligence. BENNETT White Paper 06. Canonical locked v1.0. 2026. https://486b4e8f-f0c1-4b95-8447-81ae78bc5540.filesusr.com/ugd/175885_0cf9c0fd6a0142318b3340915e08a85a.pdf
8. Bennett BT. Scientist-in-Control AI. BENNETT White Paper 07. Canonical locked v1.0. 2026. https://486b4e8f-f0c1-4b95-8447-81ae78bc5540.filesusr.com/ugd/175885_1cc2eb114dd54f499fa2792214d6644a.pdf
9. Bennett BT. Micro-Engines, Not a Monolith. BENNETT White Paper 08. Canonical locked v1.0. 2026. https://486b4e8f-f0c1-4b95-8447-81ae78bc5540.usrfiles.com/ugd/175885_ed012b87892443b49ea92bca4d2ae47c.pdf
10. Bennett BT. Data Processing. Author-supplied assay-development record associated with the 2007 chemo-sensitivity work. Historical scientific record.
11. Bennett BT. A Chemo-Sensitivity Assay: Personalized Cancer Therapy and Drug Discovery. Author-supplied presentation associated with the 2007 assay-development work. Historical scientific record.
12. Bray MA, Singh S, Han H, et al. Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nature Protocols*. 2016;11:1757-1774. doi:10.1038/nprot.2016.105.

-
13. Caicedo JC, Cooper S, Heigwer F, et al. Data-analysis strategies for image-based cell profiling. *Nature Methods*. 2017;14:849-863. doi:10.1038/nmeth.4397.
 14. Chandrasekaran SN, Cimini BA, Goodale A, et al. Three million images and morphological profiles of cells treated with matched chemical and genetic perturbations. *Nature Methods*. 2024;21:1114-1121. doi:10.1038/s41592-024-02241-6.
 15. Chandrasekaran SN, Ceulemans H, Boyd JD, et al. Image-based profiling for drug discovery: due for a machine-learning upgrade? *Nature Reviews Drug Discovery*. 2021;20:145-159. doi:10.1038/s41573-020-00117-w.
 16. FDA-NIH Biomarker Working Group. **BEST (Biomarkers, EndpointS, and other Tools) Resource: Validation**. U.S. Food and Drug Administration and National Institutes of Health. 2016-. Updated 2020.
 17. Bennett BT. Founder chronology supplied for BENNETT White Paper 10, 2026. Source for the founder-reported continuation of the 2007 assay-distillation method through approximately 2012; the rapid expansion of super-resolution modalities and cross-platform data silos from roughly 2012 through 2020; the beginning of Identifyn in 2023 and the subsequent manual Resolution Stepdown work across supported microscopy tiers; and the historical relationship between those manual methods and the later BENNETT scientific architecture. This record establishes founder chronology; it is not presented as independent validation or a platform-performance measurement.