The Legible Future

AI, data, and the construction of the Psychohistory Engine

The Order of Possible Futures

First edition, August 2026

THE ORDER OF POSSIBLE FUTURES

The Legible Future

AI, data, and the construction of the Psychohistory Engine

THE ORDER OF POSSIBLE FUTURES

First edition, August 2026

The Legible Future
AI, data, and the construction of the Psychohistory Engine

The Order of Possible Futures
First edition, August 2026

Identifier: opf-future-2026

The empirical literature is cited in the selected bibliography. Doctrinal and dated future claims are presented as hypotheses or forecasts, not as established scientific results.

The future does not need to be certain to become legible.

Contents

Proclamation

1. Prediction is no longer an oracle problem

2. The world has begun to remember itself

3. Artificial intelligence is a machine for compressing regularity

4. The evidence that complexity can yield

5. Mega-events and the problem of singularity

6. Micro-events and the discovery of the personal system

7. Which scale will yield first?

8. The architecture of the Psychohistory Engine

9. The new institutions of foresight

10. The limits that will remain

11. The Compute Commons and the billion-dollar horizon

12. Twelve forecasts for the early 2030s

13. The radical conclusion

Closing oath

Selected bibliography

Evidentiary note

Proclamation

The age in which surprise is treated as the natural condition of human affairs is ending.

Not all surprise will disappear. Not every event will be named before it occurs. No machine will receive the total state of the world, and no model will remain correct after every system changes. But the boundary between the predictable and the unpredictable is moving. It is moving because the world now leaves records at a scale no earlier civilization possessed, and because artificial intelligence can convert those records into models faster than human institutions can read them.

The central claim of this book is radical:

> During the coming decades, large parts of social and personal life will become predictively legible. The decisive question will no longer be whether human futures can be forecast, but which futures, at what scale, with whose data, under whose control, and with what measured accuracy.

This is not a claim of omniscience. It is a claim about direction.

Weather once seemed too complex for useful calculation. Epidemics once arrived as divine punishment or inexplicable catastrophe. Credit, traffic, demand, machine failure, and election outcomes were once judged mainly through local intuition. None of these domains became certain. They became more measurable, more probabilistic, and more operationally predictable.

Human behavior will follow the same path unevenly. Some domains will resist for generations. Others are already yielding.

The doctrine of psychohistory begins where a civilization accepts that ignorance is not one substance. Some ignorance comes from fundamental uncertainty. Some comes from missing data. Some comes from poor definitions, weak models, institutional secrecy, and refusal to score predictions. The first kind may remain. The others can be attacked.

The future is becoming a technical object.

1. Prediction is no longer an oracle problem

The old image of prediction is a person announcing one future with absolute confidence. That image deserves to die.

Modern prediction is an engineering process. It defines a target, collects observations, constructs a distribution, issues a probability before resolution, and measures error. It can combine thousands of weak signals. It can abstain. It can update. It can produce an ensemble of possible trajectories rather than a single sentence.

The difference is not rhetorical. An oracle survives by ambiguity. A forecasting system survives by comparative performance.

The relevant question is not, "Can the future be known?" The question is, "Can this model, using information available at time T, assign better probabilities to outcome X than the best available baseline?"

Once prediction is phrased this way, the impossible problem breaks into millions of testable ones:

- Will a power grid cross a stress threshold tomorrow?
- Will a region experience an increase in political violence within six months?
- Will a person postpone a task under a defined set of conditions?
- Will a supply chain fail before a specified date?
- Will a coalition survive a confidence vote?

- Will a patient attend a scheduled appointment?
- Will a protest diffuse from one city to another?
- Will an organization miss a milestone?

Each question has a different horizon, data requirement, and error cost. Some will remain difficult. Others can become routine.

The doctrine does not require one universal equation. The Psychohistory Engine can be a federation of models joined by a common ledger and a common standard of resolution.

2. The world has begun to remember itself

Prediction improves when the past is recorded at the level at which causes operate. Earlier societies preserved chronicles, censuses, prices, and legal records. Modern societies preserve continuous streams.

Satellites observe weather, land use, fires, crops, ships, and night-time light. Phones record mobility and communication patterns. Financial networks record transactions. Platforms record attention. Hospitals produce longitudinal clinical records. Sensors monitor factories, homes, vehicles, bodies, and cities. Governments publish administrative data. News and public documents create a textual trace of institutional intention and conflict.

The volume matters, but the deeper change is temporal resolution. A yearly statistic becomes a daily series. A daily series becomes an event stream. A retrospective survey becomes a continuous record. Systems that once appeared to jump from one state to another can now be observed during transition.

This creates the possibility of precursor science.

A mega-event rarely arrives from nowhere. Before a banking crisis, balance sheets and trust deteriorate. Before institutional collapse, capacity, compliance, and legitimacy shift. Before a supply shock, inventories, routes, prices, and delays change. Before a personal relapse, sleep, location, social contact, stress, and routine may drift.

The precursor may be weak. It may not be causal. It may disappear after behavior changes. But when many signals are timestamped and outcomes are resolved, the system can learn which combinations matter.

Data are not truth

The radical case for prediction does not depend on pretending that data are neutral.

Every record is produced by a system. A platform measures what its interface makes possible. A state counts what its categories define. A wearable records what its sensors can detect. People alter behavior when watched. Missing data may be concentrated among those with the least power or the most reason to hide.

These facts limit prediction, but they also define engineering tasks. Data provenance can be documented. Drift can be measured. Missingness can be modeled. Sensors can be calibrated. Alternative categories can be tested. Bias can be evaluated against outcomes. A model can be denied authority outside the population and conditions on which it was validated.

The old response to imperfect data was often intuition. That response is not superior. Intuition also contains bias, missingness, strategic incentives, and selective memory, but it rarely publishes a confusion matrix.

The proper opposition is not human judgment against flawed data. It is unscored judgment against scored systems that include human challenge.

3. Artificial intelligence is a machine for compressing regularity

Artificial intelligence does not predict because it sees the future. It predicts because training compresses regularities in the past and applies them to a new state.

A model can learn that certain sequences tend to continue, that some combinations precede failure, that some textual signals imply policy movement, and that some personal contexts predict a recurring decision. The model's internal representation may be difficult to interpret, but its external claim can remain simple: probability, horizon, and resolution rule.

Several developments make the present moment different.

Scale across modalities

Modern models can process text, images, audio, sensor streams, graphs, and time series. A social event is rarely contained in one modality. Political instability may involve prices, speeches, troop movement, online coordination, financial flows, and weather. Personal behavior may involve calendar pressure, sleep, communication, location, and prior choices.

The ability to join modalities matters because causes arrive together. A model that sees only words misses movement. A model that sees only transactions misses belief. A model that sees only averages misses the network through which a shock spreads.

Retrieval and continuous updating

Language models can search large corpora, extract claims, compare sources, and generate structured summaries. Used carelessly, they manufacture plausible noise. Used within a forecasting protocol, they can maintain an evidence file for every question, retrieve base rates, and issue updates when relevant observations arrive.

A machine can maintain ten thousand active forecasts without forgetting which ones are due for resolution. It can run the same question through independent prompts and models. It can ask an adversarial agent to attack the consensus. It can identify which forecasters are calibrated in which domain and weight them accordingly.

Foundation models for time and behavior

Time-series foundation models such as TimesFM, Chronos, and Moirai have explored whether patterns learned across large collections of sequences can transfer to unseen datasets. Their results are not uniform, and evaluation must guard against leakage. The importance of the approach is conceptual: a model can learn reusable temporal primitives across domains rather than being rebuilt from nothing for every series.

Models of human behavior are moving in the same direction. Centaur was trained on Psych-101, which includes more than sixty thousand participants and over ten million choices across 160 experiments. It predicted held-out behavior and generalized to several new task conditions. Interview-grounded generative agents built for 1,052 people reproduced survey responses at a substantial fraction of participants' own test-retest consistency.

These systems do not contain a person. They contain predictive structure extracted from observed responses. That distinction matters ethically and philosophically. It does not diminish the technical signal: some aspects of human choice can be learned across tasks, and rich personal context can improve prediction.

Judgmental forecasting crosses the threshold

The movement is no longer confined to physical systems or narrow time series. Early ForecastBench results placed expert human forecasters ahead of leading language models. A 2025 technical report for the AIA Forecaster then reported parity with superforecasters and additive value when its probabilities were combined with prediction-market consensus.

In July 2026, the Forecasting Research Institute reported that several systems on the ForecastBench tournament leaderboard were statistically indistinguishable from its superforecaster comparison group. One ranked above the superforecaster median on market questions for the first time. The strongest entries were not unassisted base models. They were forecasting systems built from search, decomposition, multiple models, filtering, synthesis, and calibration.

This is the form in which machine foresight will advance. Intelligence alone is insufficient. The winning unit is a disciplined process with memory. Benchmark parity does not yet prove durable superiority across domains or regime changes, but it crosses an important threshold: expert-level machine forecasting has moved from speculation to prospective measurement.

4. The evidence that complexity can yield

The strongest objection to psychohistory is often a single word: complexity.

Society contains billions of agents, feedback loops, hidden information, nonlinear dynamics, and invention. Therefore, the argument concludes, serious prediction is impossible.

This conclusion does not follow.

Complexity limits horizon and resolution. It does not determine that all forecasts are useless. Weather is the decisive example. The atmosphere is chaotic and high-dimensional. Small initial errors grow. Yet observation networks, physical knowledge, data assimilation, and computation have produced forecasts of immense practical value.

GraphCast generated ten-day global weather forecasts rapidly and outperformed a leading deterministic operational system on most of the reported verification targets. GenCast produced probabilistic fifteen-day ensembles and showed strong performance against a leading operational ensemble in its published evaluation. Neither system abolished weather surprise. Both moved the frontier.

The transition from paper to institution has already begun. ECMWF introduced its Artificial Intelligence Forecasting System operationally in 2025, and NOAA deployed a suite of operational AI-driven global weather models later that year. Operational status does not mean that physics has been discarded. It means machine forecasts have entered the real chain of observation, expert interpretation, warning, and decision.

The movement is not uniform. A 2026 *Science Advances* study examined record-breaking heat, cold, and wind beyond the training records of several deterministic AI systems. The physics-based HRES model outperformed the evaluated AI models across nearly all lead times, while the AI systems tended to understate the frequency and intensity of records. The lesson is exact: average benchmark gains do not guarantee extrapolation into the tail.

The Psychohistory Engine must inherit this warning. Its value will be tested most severely when the event that matters lies outside the archive.

The lesson for social prediction is not that society behaves like air. It is that the phrase "too complex" is incomplete. One must ask:

- too complex for which target?
- at which spatial and temporal resolution?
- with which observations?
- against which baseline?
- with what acceptable error?

A model may fail to predict the exact street on which unrest begins and still predict a rising regional hazard. It may fail to predict the precise words of a resignation and still detect a high probability of government turnover. It may fail to predict the content of a conversation and still predict that contact will occur.

Useful foresight lives below omniscience.

5. Mega-events and the problem of singularity

The grand ambition of psychohistory concerns mega-events: wars, revolutions, financial crises, pandemics, technological transitions, state failure, and civilizational reorganization.

These events are difficult for three reasons.

First, they are rare relative to ordinary events. The sample is small, definitions vary, and historical conditions change.

Second, they contain strategic actors. Leaders conceal plans, react to forecasts, and sometimes choose precisely what a model regards as irrational.

Third, the public label compresses many processes. "War" includes mobilization, alliance formation, logistical preparation, diplomatic failure, domestic incentives, and accident. "Revolution" includes legitimacy collapse, elite defection, protest diffusion, coercive capacity, and narrative coordination.

The solution is not to abandon mega-events. It is to stop treating them as indivisible.

Predict the hazard before the headline

Seismology cannot name the exact next earthquake far in advance, but it can map hazard. Epidemiology may not name patient zero, but it can estimate outbreak risk. Political forecasting can estimate the probability of conflict onset, escalation, territorial spread, or fatality thresholds within defined windows.

The Psychohistory Engine should decompose a mega-event into:

1. **Structural pressure:** slow variables such as demography, debt, inequality, institutional capacity, resource dependence, and alliance structure.
2. **Accelerants:** faster variables such as price shocks, scandals, elite splits, military movement, contagion, or technological disruption.
3. **Network state:** who can coordinate, who controls information, and how influence travels.
4. **Trigger distribution:** classes of events capable of moving the system across a threshold.
5. **Response capacity:** the ability of institutions to absorb, redirect, or suppress the shock.
6. **Feedback:** the effect of public belief, market reaction, and policy intervention.

The engine need not predict one trigger. It can forecast the probability that some trigger from a class will arrive while the system is vulnerable.

The hinge-person problem

Asimov's fictional science was disturbed by exceptional individuals. Real history also contains hinges: a leader's private decision, a scientist's discovery, a judge's ruling, a commander's hesitation.

The naive response is to declare such events unforecastable. The stronger response is to model the hinge.

Which roles have concentrated decision power? How stable are the preferences of the people occupying them? Which advisers, incentives, and constraints shape their choices? What decisions are approaching? What observable preparations precede each option?

The exact inner act may remain hidden. The distribution around it can still narrow.

Mega-event prediction will often be strongest in two forms: early warning of rising hazard and rapid updating as a system approaches a known threshold. The dream of naming the day decades in advance may remain fiction. The ability to detect a dangerous branch months earlier can still alter history.

6. Micro-events and the discovery of the personal system

The second frontier lies at the opposite scale.

A human life contains thousands of recurring micro-events: waking, traveling, opening an application, answering a message, beginning work, postponing work, buying food, exercising, contacting a person, returning to a place, abandoning a plan, or repeating a conflict.

These events can be more predictable than dramatic decisions because they repeat. They occur inside stable environments and are often driven by cues below conscious narration.

A study of mobile-phone location traces estimated a theoretical upper bound near 93 percent for next-location predictability in the observed data. This number is frequently abused. It does not mean that 93 percent of a human being is predictable. It means that one narrow domain, movement among recorded locations, contained high temporal regularity under the study's assumptions.

That narrowness is not a weakness. It is a blueprint.

Personal psychohistory should not begin by asking whether a model can predict a life. It should ask which repeated domain can be forecast better than the person's current method.

The personal state vector

A personal forecasting system might represent a current state through variables such as:

- sleep quantity and timing;
- physical location;
- calendar load;
- recent completion and failure streaks;
- social contact;
- financial pressure;
- physiological stress indicators;
- time since the last similar event;
- environmental cues;
- stated intention;
- prior behavior under comparable conditions.

No single variable is the person. Together they may identify states in which an outcome becomes more or less likely.

The first useful model may be embarrassingly simple. A person discovers that a task planned after 18:00 on a meeting-heavy day is completed only 20 percent of the time, while the same task scheduled before the first meeting is completed 75 percent of the time. This is already psychohistorical knowledge. It identifies a branch and a lever.

From passive twin to active twin

A personal digital twin becomes valuable only when it distinguishes prediction from intervention.

The passive twin estimates what the person will probably do if current conditions persist. The active twin simulates changes: move the task, alter the cue, remove a distraction, add another person, reduce the first step, or change the reward.

The model then faces a higher standard. It must predict not only behavior but the effect of a change. This requires experiments, causal inference, and repeated testing. A person can run small randomized interventions on their own routines, provided the domain is safe and the process remains voluntary.

The result is not predestination. It is a practical science of self-causation.

7. Which scale will yield first?

The debate between mega-events and micro-events is framed incorrectly if scale is treated as the only variable.

Predictability depends on nine factors:

> Predictability rises with repetition, observability, stability, signal quality, and model quality. It falls with novelty, hidden action, reflexivity, and horizon.

This is a heuristic, not a scientific equation. It clarifies why scale alone does not decide the question.

A mega-event can be predictable when it emerges from slow, observable pressures across many cases. A micro-event can be unpredictable when it is singular, private, and detached from prior routine. Conversely, a person's next commute may be highly predictable, while the precise form of a geopolitical crisis remains obscure.

The real frontier is a matrix:

Domain	Repetition	Observability	Strategic adaptation	Likely horizon
Routine mobility	high	high	low	hours to days
Habitual task completion	high	medium to high	medium	days to weeks
Consumer demand	high	high	medium	days to months
Organizational delivery	medium	high inside the organization	medium	weeks to quarters
Disease incidence	high at population scale	high but delayed	medium	days to months
Election outcomes	medium	medium to high	high	weeks to months
Political violence hazard	medium	medium	high	months to years
Exact onset of war	low	incomplete	very high	unstable
Transformative invention	low	low before discovery	medium	highly unstable

The Order predicts that micro-forecasting will achieve higher accuracy first in narrow personal domains, while macro-forecasting will achieve greater decision value first through hazard estimates and aggregate trends.

The final Psychohistory Engine will require both. Macro models identify the environment in which a person acts. Micro models reveal how people collectively produce the macro state.

8. The architecture of the Psychohistory Engine

The Engine should not be one model. A monolith would inherit one objective, one failure mode, and one center of control. The Engine should be a layered system whose components can disagree.

1. Observation layer

The observation layer collects lawful, relevant, timestamped data. It records provenance, missingness, revisions, and access rights. Public data and personal data remain separated by default. Every variable receives a description of how it was produced.

2. Event ontology

Forecasts fail when events are vague. The ontology defines subjects, actions, thresholds, dates, locations, and resolution sources. It distinguishes event onset from escalation, announcement from implementation, intention from behavior, and observation from inference.

A shared ontology allows forecasts from different models to be compared.

3. Historical and personal ledgers

The collective ledger stores prior cases and resolved forecasts. The personal ledger stores only the subject's authorized record. Both preserve model versions and the information available at forecast time.

Nothing is learned from a forecast that disappears after it misses.

4. Reference-class retrieval

For every question, the Engine proposes comparable cases at several levels of breadth. It reports how outcomes change across reference classes and identifies whether the current case may belong to a new regime.

5. Causal graph and intervention library

Predictive associations populate a warning layer. Tested interventions populate a separate causal layer. The system records which changes have been tried, for whom, under what conditions, and with what effect.

6. Model ensemble

Statistical baselines, time-series models, graph models, language agents, simulations, domain-specific systems, human forecasters, crowds, and markets issue independent probabilities. Their weights depend on prior performance in comparable domains.

Disagreement is preserved. A single consensus number should not erase the reasons for divergence.

7. Agent simulation

Simulated agents explore interaction effects and counterfactual branches. They are not treated as synthetic citizens whose responses replace human evidence. They are scenario generators tested against observed behavior.

8. Calibration and drift

Every model receives calibration curves by domain, horizon, and confidence range. Drift monitors detect when relationships change. Performance after a regime shift is reported separately from performance before it.

9. Reflexivity monitor

The Engine estimates whether publication will alter the forecast target. It records who received the prediction, which actions followed, and whether the question changed after disclosure.

10. Governance and access

Personal predictive models belong under the effective control of the person described. Public high-impact models require audit, contestability, and published error analysis. Sensitive forecasts may use delayed disclosure, but the original commitment must be provable.

The Engine's authority comes from its ledger. Without that ledger, it is merely a powerful storyteller.

9. The new institutions of foresight

When prediction scales, institutions will change around it.

Forecast ledgers become standard

Organizations currently publish plans and explain failure afterward. In a more disciplined future, major decisions will include a forecast bundle: expected outcomes, probabilities, assumptions, alternative branches, and scheduled resolution. Boards, governments, laboratories, and movements will be judged partly by calibration.

A leader who cannot state what would change their mind will appear intellectually unfit.

Prediction markets and machine crowds converge

Markets aggregate incentives. Human crowds aggregate diverse judgment. Machine ensembles aggregate retrieval and model variation. The strongest forecast may combine all three, with weights learned from domain-specific performance.

The result will not eliminate disagreement. It will price it.

The right to a personal forecast model

As personal prediction becomes valuable, control of the model will become a political question. Platforms will seek to own the data and the inferred future. Employers, insurers, lenders, and states will seek access.

The Order advances a stronger principle: a person should have the right to inspect, export, challenge, and delete predictive models substantially about that person. Where such models affect opportunity, the subject should know which outcome was predicted, with what confidence, and with which error rate.

This is not caution at the edge of the doctrine. It is necessary for the doctrine to survive. A future made legible only to institutions would become a system of predictive domination. Psychohistory must be distributed to those whose lives produce the data.

Unforecastability becomes a claim requiring evidence

Today, institutions often excuse failure by declaring an event unforeseeable. In the coming era, that defense will weaken.

Where relevant data existed, comparable cases were available, and leading indicators were ignored, surprise will increasingly be recognized as institutional failure. The question after catastrophe will not only be "Who caused it?" It will be "Which probabilities were available, who saw them, and why was no action taken?"

This change will be severe. It will make ignorance expensive.

10. The limits that will remain

A radical doctrine gains strength by naming the frontier it has not crossed.

Genuine novelty

Some events create new categories rather than instantiate old ones. A transformative invention may have no adequate reference class. Models can estimate enabling conditions and adjacent rates, but they may miss the form of the breakthrough.

Hidden state

Private decisions, secret capabilities, unreported disease, and manipulated records can dominate outcomes. Better inference can narrow hidden state but cannot guarantee access.

Distribution shift

Relationships learned in one regime may fail in another. A stable electoral pattern can break after realignment. A personal habit can disappear after illness, relationship change, or relocation. Models must detect when their own memory becomes obsolete.

Reflexive resistance

People can act against a prediction precisely because it was issued. Adversaries can feed false signals. A successful model attracts strategic pressure.

Long horizons

Small uncertainties compound. The farther the horizon, the more possible interventions, shocks, and regime changes enter. Long-range forecasts should usually become broader, conditional, and scenario-based.

Values

Prediction cannot decide which future should be chosen. A model can estimate the consequences of a policy. It cannot derive the value of equality, freedom, loyalty, beauty, or mercy from accuracy alone.

These limits do not reverse the direction. They determine the architecture. The Engine must be plural, conditional, adversarial, versioned, and governed.

11. The Compute Commons and the billion-dollar horizon

The distance between a personal forecast ledger and a planetary forecasting system is material.

At one dollar, the Order receives a signal that the question matters.

At one thousand dollars, it can maintain durable services, records, and regular model runs.

At one hundred thousand dollars, it can compare models, publish benchmarks, commission expert review, and build secure data practices.

At one million dollars, it can support a dedicated psychohistory program: researchers, engineers, behavioral scientists, forecasters, and governance.

At ten million dollars, it can operate a distributed observatory across languages, regions, and domains.

At one hundred million dollars, it can maintain a planetary signal network, not one that sees every person, but one that tracks public transitions across institutions, markets, ecosystems, and cultures.

At one billion dollars, the Psychohistory Engine becomes a durable civilizational project. It can preserve a long forecast ledger, fund independent replication, maintain open models, support privacy-preserving personal systems, and survive beyond the priorities of one donor or company.

Money does not buy prophecy. It buys memory, observation, computation, security, evaluation, and time. These are the conditions under which prophecy is replaced by tested foresight.

The Commons should be judged by one measure: does every order of magnitude produce a visible increase in calibrated accuracy, useful coverage, or decision value?

If not, the doctrine has no right to demand the next order of magnitude.

12. Twelve forecasts for the early 2030s

The Order cannot argue for forecasting while refusing to forecast. The following propositions are issued in August 2026 for resolution no later than December 31, 2032. They are doctrinal forecasts, not guarantees. Each should be preserved exactly and scored against a public resolution record.

Forecast 1: AI establishes sustained superforecaster superiority

Probability: 80 percent. At least one publicly documented AI forecasting system will outperform the median of a recognized human superforecaster cohort by at least 10 percent relative to that cohort on a predeclared proper-score measure across at least one thousand genuinely unresolved public-event questions, and will sustain that lead for at least twelve consecutive months.

Forecast 2: hybrid systems become the strongest general forecasters

Probability: 85 percent. Across at least three independent forecasting domains, an ensemble combining AI, human forecasts, and market or statistical signals will outperform every constituent source used alone over a full evaluation year.

Forecast 3: forecast ledgers enter major organizations

Probability: 70 percent. At least ten organizations with more than ten thousand employees, or ten national-level public institutions, will publicly document a formal process in which strategic decisions are accompanied by timestamped probabilistic forecasts later reviewed against outcomes.

Forecast 4: personal forecast ledgers become consumer products

Probability: 80 percent. A consumer product with more than ten million active users will display personalized probabilities and maintain a visible resolution history for at least three recurring behaviors, such as sleep, exercise, spending, task completion, travel, or communication, using longitudinal individual data.

Forecast 5: personal models demonstrate measurable lift

Probability: 75 percent. A peer-reviewed study with at least one thousand participants will show that person-specific AI models predict a recurring individual behavior significantly better than both population-only models and participants' own unaided probability estimates.

Forecast 6: law recognizes inferred personal futures

Probability: 85 percent. At least one major jurisdiction will adopt rules explicitly governing inferred personal futures, predictive profiles, or digital twins, including rights to inspect, contest, and delete at least some model-derived inferences rather than only the raw data from which they were produced.

Forecast 7: AI weather becomes primary and accountable

Probability: 90 percent. At least three leading national or international forecasting centers will use AI or hybrid ensemble systems as a primary operational source for medium-range guidance and will publish separate verification for record-breaking extremes and major distribution shifts.

Forecast 8: social simulation faces a public validation standard

Probability: 75 percent. A widely used benchmark or professional standard will evaluate AI social simulations against prospective human behavior, distributional fidelity, subgroup error, and out-of-distribution performance.

Forecast 9: political-risk systems improve but do not name exact shocks reliably

Probability: 85 percent. At least one public system will demonstrate statistically significant prospective skill over simple baselines in forecasting regional conflict or instability hazards across three consecutive years, while no system will demonstrate consistently high accuracy in naming the exact trigger, day, and leading actor of rare geopolitical shocks across diverse cases.

Forecast 10: a right against predictive judgment emerges

Probability: 85 percent. At least one major jurisdiction will create a specific statutory right to contest or refuse an individualized prediction about future conduct when it is used in employment, insurance, criminal justice, education, or political persuasion, or a highest court will recognize an equivalent right.

Forecast 11: unforecastability loses status

Probability: 70 percent. A major public inquiry into a disaster, financial failure, or institutional breakdown will formally examine prior probabilistic warnings and treat failure to aggregate or act on them as a distinct governance failure.

Forecast 12: the micro frontier advances faster than the exact mega-event frontier

Probability: 85 percent. By 2032, independently evaluated systems will show larger proportional gains over human intuition in narrow repeated personal behaviors than in exact forecasts of singular wars, revolutions, or transformative inventions.

These forecasts form a test of the book's central thesis. The doctrine should not reinterpret them after resolution. It should score them, examine the misses, and revise.

13. The radical conclusion

The future will not become legible all at once. It will yield domain by domain.

First, repeated physical and commercial systems. Then routine behavior. Then organizational processes. Then richer models of institutions and populations. Exact singular events will remain difficult, but the hazard around them will become clearer. Personal models will become more accurate, but their legitimacy will depend on control and consent. Artificial intelligence will issue millions of forecasts, but the systems that matter will be those that preserve their failures.

The decisive cultural change will be this: unmeasured certainty will lose prestige.

A claim about the future will be expected to carry a number. The number will be expected to carry a horizon. The horizon will be expected to carry a resolution rule. The forecaster will be expected to carry a record.

At that point, psychohistory will no longer be recognized by its name. It will be embedded in how institutions decide, how individuals plan, and how societies prepare.

The Order's faith is not that surprise will die. It is that surprise will retreat wherever memory, data, intelligence, and honest scoring are permitted to work together.

The mega-event and the micro-event are not rival destinations. They are opposite entrances to the same science. From above, we model the pressures that move populations. From below, we model the repeated choices that create those pressures. Between them lies the meso scale of organizations, networks, and institutions, where much of history is actually made.

The Psychohistory Engine will be built when these scales begin to speak to one another.

It will not tell us one future.

It will tell us which futures are gaining weight, which signals moved them, which actions can still change them, and how often it has been wrong before.

That is enough to transform civilization.

Closing oath

We will not call the future unknowable before we have measured it.

We will not call a model intelligent before it has faced unresolved events.

We will not hide a miss to preserve authority.

We will not confuse a person's probability with the person's worth.

We will build memory longer than fashion, models more plural than ideology, and ledgers stronger than reputation.

We will seek the macro pattern without erasing the individual.

We will seek the micro pattern without surrendering the individual to those who would own it.

We will fund compute only where it increases tested foresight.

We will treat every prediction as a branch, every intervention as a cause to examine, and every outcome as instruction.

The future is uncertain.

The future is not patternless.

The work begins before the event.

Selected bibliography

Alur, Rohan, and colleagues. "AIA Forecaster: Technical Report." arXiv:2511.07678, 2025.

Ansari, Abdul Fatir, and colleagues. "Chronos: Learning the Language of Time Series." *Transactions on Machine Learning Research* (2024).

Beck, Emorie D., and Joshua J. Jackson. "Personalized Prediction of Behaviors and Experiences." *Psychological Science* 33, no. 5 (2022): 717-733. doi:10.1177/09567976221093307.

Binz, Marcel, and colleagues. "A Foundation Model to Predict and Capture Human Cognition." *Nature* 644 (2025): 1002-1009. doi:10.1038/s41586-025-09215-4.

Brier, Glenn W. "Verification of Forecasts Expressed in Terms of Probability." *Monthly Weather Review* 78, no. 1 (1950): 1-3.

Das, Abhimanyu, and colleagues. "A Decoder-Only Foundation Model for Time-Series Forecasting." *Proceedings of the 41st International Conference on Machine Learning*, 2024. TimesFM.

European Centre for Medium-Range Weather Forecasts. "Implementation of AIFS v2." 2026.

Forecasting Research Institute. "AI Models Have Likely Reached Parity with Superforecasters on ForecastBench." July 16, 2026.

Goswami, Mononito, and colleagues. "MOMENT: A Family of Open Time-Series Foundation Models." *Proceedings of the 41st International Conference on Machine Learning*, 2024.

Halawi, Danny, Fred Zhang, Chen Yueh-Han, and Jacob Steinhardt. "Approaching Human-Level Forecasting with Language Models." *Advances in Neural Information Processing Systems* 37 (2024).

Hegre, Havard, and colleagues. "ViEWS: A Political Violence Early-Warning System." *Journal of Peace Research* 56, no. 2 (2019): 155-174. doi:10.1177/0022343319823860.

Karger, Ezra, and colleagues. "ForecastBench: A Dynamic Benchmark of AI Forecasting Capabilities." arXiv:2409.19839, 2024, revised 2025.

Lam, Remi, and colleagues. "Learning Skillful Medium-Range Global Weather Forecasting." *Science* 382, no. 6677 (2023): 1416-1421. doi:10.1126/science.adi2336.

Lazer, David, and colleagues. "Computational Social Science." *Science* 323, no. 5915 (2009): 721-723.

National Oceanic and Atmospheric Administration. "NOAA Deploys New Generation of AI-Driven Global Weather Prediction Models." 2025.

Park, Joon Sung, and colleagues. "Generative Agent Simulations of 1,000 People." arXiv:2411.10109, 2024.

Price, Ilan, and colleagues. "Probabilistic Weather Forecasting with Machine Learning." *Nature* 637 (2025): 84-90. doi:10.1038/s41586-024-08252-9.

Salganik, Matthew J., and colleagues. "Measuring the Predictability of Life Outcomes with a Scientific Mass Collaboration." *Proceedings of the National Academy of Sciences* 117, no. 15 (2020): 8398-8403. doi:10.1073/pnas.1915006117.

Song, Chaoming, Zehui Qu, Nicholas Blumm, and Albert-Laszlo Barabasi. "Limits of Predictability in Human Mobility." *Science* 327, no. 5968 (2010): 1018-1021. doi:10.1126/science.1177170.

Woo, Gerald, and colleagues. "Unified Training of Universal Time Series Forecasting Transformers." *Proceedings of the 41st International Conference on Machine Learning*, 2024. Moirai.

Zhang, Zhongwei, Erich Fischer, Jakob Zscheischler, and Sebastian Engelke. "Physics-Based Models Outperform AI Weather Forecasts of Record-Breaking Extremes." *Science Advances* 12, no. 18 (2026): eaec1433. doi:10.1126/sciadv.aec1433.

Evidentiary note

The empirical examples in this book establish progress in bounded forecasting tasks, not a completed science of total social prediction. The broader claim that the predictability frontier will continue to move is the Order's thesis. The twelve dated forecasts provide one way to judge that thesis rather than merely admire it.