Project Star Forge: A Planetary-Scale Foresight Engine for Global Risk Simulation

THOMAS HARDING, Independent Researcher, United States

Project Star Forge is a proposed planetary-scale foresight engine designed to simulate and analyze cross-domain global risks in an integrated manner. The system architecture consists of a central orchestration core that coordinates an ensemble of specialist simulation modules (SSMs) across domains (e.g. climate, economy, geopolitical conflict, infrastructure, etc.), all underpinned by a unified cross-domain knowledge graph. A scenario generation and simulation component, termed the Fork Cascade Engine, allows exploration of branching "what-if" scenarios and cascading effects across systems. An emotional-cognitive heuristics layer is incorporated to model human-like decision-making biases and responses within simulations, improving realism when simulating social or institutional behavior. Ethical oversight mechanisms and transparency are built into the design to align with humanitarian goals. We describe the architecture and methodology of Project Star Forge in detail, including plans for validation of its predictive capabilities. The system is positioned as a modular, extensible research platform to support policymakers and researchers in stress-testing strategies against plausible futures. The paper concludes with a discussion of how Project Star Forge's technical blueprint supports its mission as a global public good and how it can be evolved through collaborative efforts.

CCS Concepts: • **Social and professional topics → Codes of ethics**; • **Computing methodologies →** *Cognitive science*; • **Applied computing → Forecasting**.

Additional Key Words and Phrases: global risk, existential risk, NOAA, global catastrophic risk, autonomous decision-making, self-improving systems, AI agents, GPT, Grok, Claude, flood forecast, early warning system, natural disasters

ACM Reference Format:
Thomas Harding. 2025. Project Star Forge: A Planetary-Scale Foresight Engine for Global Risk Simulation. 1, 1 (August 2025), 22 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Public Manifesto – A Mission for Global Foresight and Resilience

Mission and Values: Project Star Forge is a nonprofit humanitarian initiative dedicated to building a planetary-scale foresight engine that can simulate complex, cross-domain risks and inform better global decision-making. In an age of unprecedented technological and ecological upheaval, our mission is to provide a transparent, ethical, and public good platform for anticipating and navigating 21st-century challenges. It is designed as a global public good, meaning its insights are open to all humankind rather than locked behind corporate or national silos. By operating under a nonprofit model, we prioritize ethical use, transparency, and broad collaboration over profit. This ensures that the system's development and outputs remain accountable to the public interest and not driven by private agendas.

Author's Contact Information: Thomas Harding, Independent Researcher, Dallas, United States, calmingchronicles85@gmail.com.

Why a Global Foresight Engine Now:

The 21st century confronts us with deeply interconnected risks—climate change, pandemics, biodiversity loss, geopolitical conflicts, financial crises, and even emerging AI or biotechnology threats. These challenges do not occur in isolation; shocks can cascade across domains, multiplying their effects in ways that strain our institutions and imagination [8]. Researchers warn that concurrent crises can lead to catastrophic "poly-crises" beyond historical experience, underscoring the need for new tools to craft and study future risk scenarios as a responsibility for global governance [8]. In other words, humanity faces a foresight gap: traditional planning struggles with the complexity and scale of modern problems. Project Star Forge is our answer to this gap – a means to systematically explore "what-if" scenarios across climate, geopolitics, economics, technology, and society, helping policymakers and communities prepare for a range of possible futures. By integrating data and expertise from many domains, the platform will illuminate how a change in one system (like a climate shock) can ripple into others (economic downturns, migrations, conflict) in order to inform preventive action and resilience strategies.

Transparency, Ethics, and Public Trust

Trust is the cornerstone of a global foresight system. Project Star Forge is committed to a development process that is as open and transparent as possible without compromising safety. The project adheres to emerging best practices in ethical AI: human rights, fairness, and human oversight are baked into our design principles from day one [9]. We will publish our methodologies and key assumptions, invite external audits, and enable independent experts to scrutinize and validate our simulations. This open approach is not just ethical but pragmatic: studies show that open-source modeling and transparency build wider trust and improve validity, as others can test, validate, and improve the models [1]. By defensively publishing our core designs and algorithms as a public white paper, we ensure the concept remains in the public domain – preempting any attempts at proprietary appropriation of what should fundamentally remain a public, global commons of knowledge. We recognize the potential misuse of any powerful technology, so ethical safeguards and an oversight board (detailed below) will guide what scenarios are explored and how results are applied. Project Star Forge's nonprofit governance and open ethos send a clear message: this initiative exists to benefit humanity at large, with no single government or corporation in control.

The Urgency of Preparedness

Events of recent years (from global pandemics to financial contagions and climate disasters) have made it clear that "thinking ahead and planning for a wide range of contingencies" is not a luxury but a necessity [2]. We have learned that uncertainty is no excuse for inaction [2]. A foresight engine like Project Star Forge can empower decision-makers with projections and early warnings, acting as a form of "societal radar" for looming risks. For example, climate change is widely recognized as a "threat multiplier" that exacerbates instability, resource conflicts, and disasters [2]. A tool that simulates climate-security scenarios (the focus of our Phase 0 pilot) can help governments anticipate hot-spots where droughts or floods might spark humanitarian crises or conflicts. Likewise, cross-domain simulations could illuminate how a financial crisis combined with a pandemic could stress societies, or how rapid AI advancements might interplay with geopolitical tensions. By openly sharing these foresight analyses, Project Star Forge aims to democratize access to strategic insights, so that not only major powers but all nations, NGOs, and citizens groups can better navigate the uncertain future. In short, Project Star Forge is about collective intelligence for collective survival – pooling our best data, models, and ethical thinking to chart safer paths forward in an increasingly complex world.

In the following sections, we present two deeper layers of documentation: first, an academic-style technical blueprint of the Project Star Forge system, and second, a nonprofit road map & governance plan. This structure serves readers of different backgrounds – whether you're a scientist interested in the architecture, or a policy/finance stakeholder interested in how to support and govern this initiative.

Academic Technical Blueprint – Architecture and Methodology

1 System Architecture

Modern global challenges are multi-faceted and interconnected, demanding analytical tools that break out of single-domain silos. Climate phenomena, for instance, can trigger political and economic shocks; likewise, technological disruptions can have societal and environmental ripple effects. Traditional modeling approaches – while sophisticated within domains like climate science or economics – often struggle to capture these cross-domain interdependencies.

There is growing recognition in the research community that we need integrated frameworks for data and simulation. For example, recent work on cross-domain knowledge integration highlights that solving global problems requires linking diverse datasets and breaking down data silos [11]. Similarly, scenario researchers have started combining narrative and quantitative methods to explore how different risk factors (AI, climate, biotech, etc.) might superimpose and cascade over decades [8]. Project Star Forge is designed to fill this gap: it aims to be a unifying foresight platform where various specialized models can plug in, share data via a common knowledge graph, and jointly simulate complex scenarios that no single model could represent in isolation.

From a systems-engineering perspective, Project Star Forge draws inspiration from advancements in distributed AI and simulation. The architecture is modular by design. This modularity allows each domain (e.g. climatology, epidemiology, macroeconomics, conflict simulation) to be handled by an SSM – essentially an AI or algorithmic model specialized in that area – which can be developed and improved semi-independently, yet interoperates through a common orchestration layer. The orchestration core manages the scheduling, data exchanges, and coherence of these modules, ensuring that, for example, a climate module's output (such as a drought event) can dynamically influence an economic module or a political stability module. All modules read and write to a Cross-Domain Knowledge Graph, which serves as a living memory of the simulated world state (entities, relationships, and facts spanning all domains). Knowledge graph approaches have proven effective in bridging diverse data sources in environmental and social applications, by enforcing a shared ontology and enabling queries that span traditionally separate datasets [8, 11].

Crucially, Project Star Forge is not merely a numeric simulator but also accounts for human and institutional behavior nuances. Real-world outcomes often diverge from idealized model predictions due to human factors: political will, cognitive biases, emotional reactions, misinformation, etc. To address this, we include an emotional-cognitive heuristics layer that introduces bounded rationality and behavioral economics aspects into decision-making processes within the simulation.

Incorporating such factors is increasingly common in advanced agent-based modeling – modern simulations can integrate emotional states and cognitive biases of agents to reflect more realistic human behavior, yielding insights into social phenomena that purely rational models might miss [6]. By combining data-driven modeling with behavioral heuristics, Project Star Forge aims to simulate not only "physics and economics," but also the psychology, sociology, and politics that drive real-world responses to crises.

The following sections detail the Project Star Forge architecture, component by component, and then outline the methodology of how simulations are orchestrated. We also discuss how we plan to validate and verify the system's outputs, given the complexity inherent in forecasting. By design, certain low-level implementation details are kept abstract in this paper (denoted as proprietary or to-be-developed elements) – our goal is to communicate the architecture and approach while maintaining security and focusing on high-level concepts.

1.1 Orchestration Core (Central Coordination Engine)

The Orchestration Core is the central brain of Project Star Forge. It is responsible for coordinating all the specialist modules (SSMs), timing the simulation steps, and maintaining a consistent global state across domains.

Conceptually, it functions like an operating system for the simulation: it allocates computational resources to modules, triggers them in the correct sequence or in parallel as needed, and handles the messaging between modules. This component is meant to “orchestrate” a symphony of models so that they work in concert.

Key responsibilities of the Orchestration Core include:

- **State Management:** It interfaces with the cross-domain knowledge graph (described below) to fetch the current world state that each module needs as input, and to commit the outputs of modules back into the shared state. This ensures that every specialist AI works from a single source of truth about the simulated world.
- **Scheduling and Synchronization:** Some processes (like climate or population dynamics) evolve continuously, while others (like discrete political events or market shocks) might be event-driven. The core manages a hybrid time-stepped and event-driven schedule, allowing, for example, daily timesteps for slow-changing variables and instantaneous propagation for sudden events. It synchronizes these to simulate cascading effects accurately – e.g., a drought event flagged by the climate module at time T can immediately trigger the agriculture and economy modules to update food prices, which then influence unrest levels in a governance module.
- **Integrity and Constraints:** The core enforces any global constraints or invariants. For instance, conservation laws (mass/energy in climate systems) or accounting identities (global supply equals sum of demands) are checked at this level to catch any physically or logically inconsistent outputs from modules. It effectively monitors “sanity” of the combined simulation.
- **Multi-Modal Data Fusion:** While not described in proprietary detail here, the core also integrates data streams that feed into simulations. It can incorporate real-world data updates (for near-term forecasting modes) via proprietary multi-modal data fusion pipelines. These pipelines might ingest earth observation data, economic indicators, news feeds, etc., and update the knowledge graph, but all this occurs under the core’s oversight. We simply note that the system is designed to handle large-scale data ingestion continuously, without describing the vendor-specific or cryptographic elements of how data integrity is ensured.

In summary, the Orchestration Core is what makes Project Star Forge a cohesive engine rather than a set of disparate models. It allows the platform to behave like a single, unified simulation of Earth systems, effectively implementing the “laws of interaction” between domains.

1.2 Specialist Simulation Modules (SSMs)

Project Star Forge’s Specialist Simulation Modules (SSMs) are the domain-specific engines—modular AI models, algorithms, or simulation subsystems each focused on a particular risk domain. Each SSM can be thought of as an expert in a particular slice of the world:

- A climate and earth systems model that simulates weather patterns, climate change trajectories, and natural disasters.

- A geopolitical simulation model, perhaps using game-theoretic or agent-based approaches to model international relations, conflicts, and cooperation dynamics.
- A public health SSM could simulate pandemic spread and health system responses.
- An economic SSM might be a global macroeconomic model that can simulate trade, markets, and financial crises.
- Other modules could include agriculture/food security, infrastructure and supply chain, demographics and migration, technology adoption, etc. – the architecture is extensible, so new modules can be added as needed for additional domains of risk.

All SSMs plug into the orchestration core via standardized interfaces. They receive inputs from the knowledge graph (the part of the world state relevant to their domain) and after performing their domain-specific computation for a time step or scenario branch, they output results back to the knowledge graph.

For instance, the climate SSM might output regional temperature and precipitation anomalies, which the agriculture SSM then picks up to adjust crop yield projections, which the economic SSM then uses to update food prices and GDP, and the social unrest module uses to update instability metrics, and so on.

A guiding design principle for SSMs is modularity and replaceability. Different institutions or research groups could contribute improved models for a domain, and as long as they conform to the interface and data schema of the knowledge graph, the orchestration core can incorporate them. This fosters a plug-and-play ecosystem of the best available models. It also means we can run multiple alternative models for the same domain to compare outcomes (for example, two climate models with different assumptions) if desired.

Each SSM can utilize whatever AI/algorithm is most appropriate for its domain:

- Some may be physics-based simulations (e.g., a general circulation model for climate).
- Others might be machine learning models (e.g., using deep learning to predict conflict likelihood given certain inputs).
- Others could be agent-based models (particularly for social systems, simulating many individual agents or actors).
- Many will be hybrids, possibly augmented by statistical forecasting or rules-based components.

While their internal workings can be complex, from Project Star Forge's top-level perspective they are black boxes with defined input-output behavior. We treat any intricate details like reward shaping, training routines, or proprietary data sources within a module as out of scope for this paper. For example, if an SSM uses a reinforcement learning agent to simulate geopolitical actors, the specific reward functions tuned for that agent would be considered part of the proprietary know-how of that module's developers and are not detailed here.

In essence, the SSM layer organizes specialized intelligence. By federating multiple "narrow AIs" under a common framework, Project Star Forge embodies a collective intelligence greater than the sum of its parts – a necessity for reflecting the real world's interconnected complexities. As one LinkedIn analyst phrased it, this kind of cross-domain scenario modeling connects technological, economic, social, and environmental factors in integrated simulations, yielding a truly holistic view [7].

1.3 Cross-Domain Knowledge Graph (Global Context Memory)

At the heart of Project Star Forge's data layer is a Cross-Domain Knowledge Graph, which serves as the central repository of facts and state in the simulation. A knowledge graph is essentially a network of entities (nodes) and relationships (edges), which can flexibly represent information. For a planetary-scale simulation, this is ideal because it can encode heterogeneous data – from physical parameters (like CO_2 concentration, sea level at location X) to socioeconomic indicators (population, GDP, conflict intensity in region Y) to relationships (trade links between countries, treaty alliances, supply chain networks, etc.).

The knowledge graph approach ensures that all modules are interoperating on consistent data structures. For instance, "Nigeria" or "Amazon Rainforest" or "Pacific Ocean" would be an entity node, and various properties (attributes) and relationships (edges) connect to it (e.g., Nigeria has a population of N, trades with X, Y; Amazon has deforestation rate D, affects climate region Z). The graph is cross-domain because it doesn't silo information by category; instead, everything from climate to economic to health data can be linked if relevant. This enables insights that siloed databases would miss – much like recent projects such as the KnowWhereGraph, which integrates human and environmental data to uncover hidden insights across traditionally separate datasets [6, 11].

Function in the simulation. The knowledge graph is updated in real-time during a simulation run. When SSMs produce new outputs (say, the economy module computes a recession in country A), that update is written to the graph. Downstream modules querying the graph will then see the updated state. Because knowledge graphs can handle temporal data, we also record state changes over simulation time, enabling analysis of temporal trajectories. The graph can be thought of as the memory of the simulation – at any point, it holds the "state of the world" as simulated, and its history holds the narrative of how that state evolved.

We also leverage the graph for consistency checks and inference. Some simple inferences are encoded (e.g., if entity X is a subset of Y, and Y has property P, perhaps propagate some effect to X). More complex reasoning (like causal inference) can be done by specialized algorithms reading the graph, although those are part of the analytics layer rather than core simulation loop.

An advantage of this approach is that it aligns with the FAIR data principles (Findable, Accessible, Interoperable, Reusable) – all data in the simulation is structured and linkable, facilitating transparency [8]. External stakeholders could query the knowledge graph after a simulation to understand why certain outcomes happened, tracing causal links.

The knowledge graph is built to integrate not just simulation variables but also real-world data layers when available. For example, it may incorporate geospatial layers (via coordinates on nodes/edges) for physical systems, or networks for trade/finance. By doing so, Project Star Forge can be constantly calibrated against reality: feeding in up-to-date observations ensures the simulation doesn't drift into fantasy.

Researchers have stressed the urgent need for new ways of sharing and integrating such diverse data for global challenges – the knowledge graph is our answer to that need within the Project Star Forge architecture [11].

1.4 Scenario Simulation Engine ("Fork Cascade Engine")

One of Project Star Forge's most powerful capabilities is its scenario simulation engine, internally dubbed the Fork Cascade Engine. This component handles the branching logic of "what-if" scenarios and the propagation of cascading effects through those branches. The name reflects two core functions: **Forking** (splitting into multiple futures or policy alternatives) and **Cascading** (letting events in one part of the system trigger chain reactions across the system).

How it works: The Fork Cascade Engine allows the system to create a tree of scenario outcomes. For instance, starting from a common initial state (e.g., the world in 2025), the engine can fork into multiple branches: one where a mild climate policy is adopted versus one where an aggressive climate policy is adopted. Each branch then evolves as the simulation runs. At certain points, the engine can introduce stochastic events or policy interventions, causing further forks.

This resembles techniques in scenario planning and Monte Carlo simulation, but here done with rich, dynamic models. Crucially, as each scenario plays out, we monitor for cascade triggers – these are conditions where an event in one domain causes a nonlinear response in another (for example, a banking crisis triggers trade collapse which then exacerbates a pandemic due to reduced medical supply). The engine ensures such cascade effects are captured by rapidly communicating trigger events across modules and, if needed, branching the simulation to explore alternate outcomes (e.g., if a pandemic remains local vs. becomes global due to those economic differences).

This systematic exploration of scenario space addresses the concern that many risk models underplay compound risks and their interactions. A recent quantitative scenario study explicitly modeled cascading risks across AI, climate, biotech, and markets over 50-year timelines [8]. It showed that combinations of catastrophes can lead to dramatically worse outcomes than each in isolation, including possibilities of long-lasting global depressions or even collapse under certain assumptions [4].

These findings reinforce why a tool like the Fork Cascade Engine is needed – to explore not just single disasters, but the full ensemble of plausible futures, including low-probability but high-impact combinations. By doing so, Project Star Forge can help identify tipping points and intervention opportunities. For example, if in 1000 simulated futures a certain chain of events leads to especially dire outcomes, we can highlight what early warning signals precede that chain and what interventions could break it.

Technically, the engine uses both deterministic branching (policy choices or scenario assumptions set by users) and probabilistic branching (random draws for uncertain events, akin to Monte Carlo simulations). It manages the computational load by pruning unlikely branches and merging equivalent outcomes where possible, so that the explosion of possibilities remains tractable.

The output is a rich set of scenarios complete with cross-domain narratives and data – essentially stress-tests for our world-model under different assumptions. Policymakers could examine these to see, for example, how a combination of a severe El Niño, a concurrent cyber-attack on infrastructure, and a financial panic might play out, versus a baseline world, and use those insights to craft robust strategies that perform well across many futures – the idea of robust decision-making under deep uncertainty.

1.5 Emotional-Cognitive Heuristics Layer

A distinctive aspect of Project Star Forge is the inclusion of an Emotional-Cognitive Heuristics Layer. This layer imbues simulated agents and decision nodes within the system with more human-like decision-making patterns, including biases and emotional responses. The rationale for this component is straightforward: when modeling complex social systems, purely rational actor models often fall short of reality. Humans (and institutions composed of humans) do not always optimize objectives in a vacuum; they follow heuristics, are influenced by fear, trust, short-term thinking, etc. Incorporating these factors can lead to more realistic simulations of, say, public response to crises, political decision-making, or market behavior under stress.

In practice, this layer is implemented as a set of modifiers or sub-models that can override or influence the decisions within certain SSMs:

- For example, a governance module might have a rational baseline for how a government should respond to a spreading pandemic (e.g., instituting early lockdowns if cases rise). The emotional-cognitive layer can introduce delays or biases (perhaps simulating denial or political polarization) to see how behavioral hesitancy affects outcomes, such as a slower response leading to a worse outbreak.

- In an economic model, agents might normally respond to information efficiently. The heuristics layer can introduce phenomena like herding behavior or panic selling in financial markets, reflecting observed behaviors in real crises.
- In conflict simulations, factors like fear, honor, or misperception might be introduced to deviate from a coldly rational game-theoretic path, thus exploring scenarios like accidental escalations or unexpectedly conciliatory moves.

There is a growing body of research supporting such approaches. Modern agent-based models can account for emotional states and cognitive biases, and doing so has yielded unprecedented insights into how individual psychological processes scale up to social outcomes [6]. For instance, models that include biases can reproduce the spread of misinformation or the persistence of harmful behaviors in a population in ways that rational models cannot, thereby highlighting intervention points for real-world policy.

By toggling this layer on or off, Project Star Forge can simulate best-case "rational world" scenarios and more realistic "biased world" scenarios, giving decision-makers a range of outcomes. Notably, the inclusion of emotional-cognitive factors is not to make the system whimsical or less rigorous – it is in fact adding scientific realism from psychology and behavioral economics. We treat these heuristics with the same empirical grounding as physical parameters, calibrating them to literature or data (for example, using known distributions of risk perception in populations, or documented biases like loss aversion in economics).

Ultimately, this layer aligns with Project Star Forge's ethos of reflecting humanity as it is, not as a simplified rational agent. It acknowledges the role of trust, fear, and culture in crises – elements that have tangible effects (as seen in the COVID-19 pandemic, where public emotion and trust levels impacted outcomes). By doing so, it strengthens the foresight engine's ability to highlight governance and communication strategies in addition to technical fixes. It essentially asks: "If we know the right answer, will people follow it? And if not, what then?"

2 Ethical Oversight Principles (Built-in Safeguards)

Even at the architecture level, Project Star Forge is designed with ethical guardrails. The system's immense analytical power comes with the responsibility to use it wisely and avoid misuse. We outline here the principles and mechanisms, noting that governance details are expanded in the later section. Key ethical oversight features include:

Transparent Scenario Parameters. Every simulation experiment in Project Star Forge is logged with a "scenario dossier" detailing its initial conditions, assumptions, and any interventions applied. This practice ensures traceability – both for internal debugging and for external review. It should always be auditable what went into a scenario. If a policy decision is later influenced by a Project Star Forge simulation, one can inspect exactly how that conclusion was reached [9].

Human Oversight and Control. No simulation runs autonomously without a human in the loop approving the parameters and the general intent. The platform provides tools for experts (e.g., domain scientists or ethicists on the oversight board) to review scenario setups, particularly those that might be sensitive (such as simulations of conflict or economic shocks that could be misinterpreted). This aligns with global AI principles that stress the importance of human oversight for high-impact AI systems [9].

Avoidance of Harmful Uses. The Project Star Forge project adopts a usage policy aligned with principles of beneficence and non-maleficence. We explicitly prohibit and technically guard against using the system to, for example, strategize offensive warfare, enable repression, or other unethical ends. The focus is on humanitarian and preventive scenarios. Our ethical AI framework requires an impact assessment for new use-cases, as recommended by UNESCO and OECD guidelines, to ensure we avoid conflicts with human rights or environmental well-being [9].

Bias and Fairness Checks. As with any AI-driven system, there is a risk of bias in simulations (data or model bias). We incorporate bias detection routines – essentially meta-audits – that examine simulation outcomes for anomalous patterns. For example, if the model consistently underestimates impacts on certain vulnerable populations due to data gaps, that is flagged for model improvement. Diversity in data and global inclusion in model development are part of the ethical design (to avoid a narrow set of values dominating the foresight engine).

Open Review and Publication. We will invite the broader scientific community to review the methodologies and even replicate elements of the system. Peer review and open science help catch issues early. In line with this, major simulation findings (especially those that might influence public policy) will be published openly along with sufficient detail to be critically evaluated. This not only builds trust but also leverages collective intelligence to spot errors or ethical blind spots.

By embedding these principles, Project Star Forge aims to set a standard for responsible AI-for-good. The ethical oversight is not a separate module but a pervasive set of rules and practices influencing system design, from how data is handled to how results are communicated.

In practical terms, an **Ethical Oversight Council** (described later) will have direct input into the development process, functioning somewhat analogously to an Institutional Review Board for an experiment. Their guidance translates into design choices – for instance, perhaps requiring a "second pair of eyes" approval step before any high-stakes simulation can be publicly released. As a result, the technical blueprint is intertwined with an ethical blueprint, ensuring the tool remains aligned with its humanitarian purpose.

3 Methodology

Having outlined the components, we describe how Project Star Forge operates as an integrated system. The methodology spans how scenarios are set up, how simulations run, and how outputs are analyzed and validated:

(1) **Scenario Initialization:** Each simulation begins with defining an initial state and parameters. This could be the present day world state (assembled from real data feeds into the knowledge graph), or a hypothetical future baseline (for long-term scenarios). Users (or researchers) can configure specific parameters: e.g., assuming a certain CO_2 emission pathway, or a particular geopolitical alliance configuration. The knowledge graph is populated accordingly. The Ethical Oversight principles apply here – scenarios that are too sensitive or speculative might require approval. Once initialized, the orchestration core takes over.

(2) **Simulation Loop (Orchestration):** The Orchestration Core advances the simulation in increments (time-steps or event-steps). In each cycle, it does the following:
- Queries the knowledge graph for the current global state.
- Triggers each relevant SSM in turn (or in parallel where feasible), providing it the needed slice of state. For instance, at a daily time step: the climate module updates weather variables; the agriculture module then

uses new weather to update crop yields; the economy module then updates commodity markets, etc. The sequence can be configured or auto-optimized.

- Integrates each module's output back into the knowledge graph, resolving any conflicts or inconsistencies.
- Checks for any special events or thresholds: if conditions for a cascade trigger are met (as defined by the Fork Cascade Engine's logic), then it will create scenario forks. For example, if a model indicates a 5% probability that a pandemic could emerge this year, the engine may branch one scenario with the pandemic outbreak and another without, to explore both.
- The emotional-cognitive layer continuously watches decisions within modules: e.g., if a government agent is about to respond to a crisis, a probabilistic bias might be applied here to mimic human delay or error. This is done by adjusting the input or output of that SSM subtly, according to calibrated bias parameters.

(3) **Branching and Parallel Simulation:** When scenario forks occur, the Orchestration Core essentially clones the simulation state into two (or more) parallel states and tracks them separately. From that point on, the SSM computations for each branch proceed independently (usually requiring parallel computing resources). For manageability, the number of concurrent branches is limited by pruning logic. We might simulate, say, 100 futures for a given set of assumptions, not millions – focusing on plausible diverse scenarios rather than an exhaustive sweep. This approach is akin to running a large ensemble of model runs, a practice common in climate forecasting to capture uncertainty.

(4) **Data Logging and Analysis:** As the simulation runs (which could be virtual years or decades, simulated over hours or days of real time depending on complexity), a wealth of data is produced. The knowledge graph at each time step constitutes a multi-dimensional dataset. This is logged to a data store. Once runs are complete, the analysis subsystem (not detailed earlier) kicks in. We generate visualizations and reports: e.g., graphs of global temperature, economic output, conflict incidents over time for each scenario. We also use the knowledge graph's structure to perform causal tracing – identifying which events led to particular outcomes (this is important for learning policy insights, like "what were the early indicators of the worst outcomes?").

(5) **Human-in-the-loop Review:** Intermediate results can be reviewed by analysts or the oversight team, especially in long simulations. They might decide to adjust certain parameters or explore a branch in more detail. For example, if a scenario is showing an unexpected conflict arising, an analyst might inject a hypothetical peace treaty at a point to see how it changes the trajectory – effectively an interactive simulation. The methodology supports iterative exploration in this way, combining automated simulation with expert steering.

(6) **Completion and Synthesis:** In the end, Project Star Forge produces a set of scenario outcomes. These are synthesized into a coherent foresight narrative for communication: for instance, a report might say "In 80% of simulated futures where X policy was not adopted, we observe Y risk escalating within 10 years, whereas with the policy, that risk was mitigated in most runs." The credibility of these insights comes not only from the models used but from cross-validation and stress-testing done internally.

We emphasize that the Project Star Forge methodology is as much about exploration and learning as it is about prediction. It is a foresight engine, not a crystal ball. Thus, a lot of the value comes from comparing across scenarios and understanding why different decisions yield different outcomes. This comparative approach aligns with methods like "wind tunnel testing" of policies in strategic foresight, where policies are tested against many scenarios to identify robust choices [3]. Our engine effectively automates and enriches that process with AI.

Technical Notes: On the technical side, running such complex simulations demands significant computational strategies: we employ high-performance computing techniques, optimize communication between modules (since coupling many models can be slow), and likely need to run on GPUs/TPUs for AI-heavy modules. The multi-modal nature of the platform (mixing qualitative branching with quantitative models) means we also devote effort to ensuring reproducibility – any run should be replayable with the same random seeds, and differences explainable.

4 Validation Plans

Validating a system as complex as Project Star Forge is itself a complex task. We outline a multi-pronged validation strategy to ensure the system's credibility and continuous improvement:

Benchmarking Against Historical Scenarios. One approach is to test whether Project Star Forge could have "predicted" or retrodicted known historical complex events when given data up to a point. For example, we might input data up to 2019 and see if any scenario in a large ensemble produces a global pandemic in 2020 and its ensuing economic shock. We would not expect an exact prediction, but we can assess whether the system's simulated pandemic scenarios qualitatively and quantitatively resemble COVID-19 outcomes (in infection spread, mortality, economic impact patterns, etc.). Similarly, we could test if the system can replicate known climate–conflict linkages from past decades, or the 2008 financial crisis dynamics, etc. This form of hindcasting provides a sanity check that the SSMs and their integration capture real-world cause-effect relations to a reasonable degree.

Modular Validation. Each Specialist Simulation Module, on its own, should be validated against domain standards. Climate modules must be validated against climate observations and existing climate model benchmarks. Economic modules should be back-tested on past economic data. We leverage the fact that many SSMs are adaptations of well-studied models (for example, using established climate models or economic models as a starting point). Their performance in isolation gives a baseline. Then we validate the coupling – e.g., ensure that when the climate module sends a drought signal, the agriculture module responds in a way consistent with historical drought impacts on crop yields. We will compile a set of cross-domain validation cases (perhaps using events like "the impact of the 1997–98 El Niño on global food markets and conflicts" which is documented in literature) to test the integrated system.

Extreme Scenario Face Validation. For scenarios that have no historical precedent (e.g., AI-driven risk cascades, or multiple simultaneous catastrophes), we rely on expert review for validation. This means subject matter experts (climate scientists, geopolitical analysts, etc.) will review the plausibility of the simulated outcomes. If the model shows an implausible cascade (like a very minor climate event causing disproportionate upheaval), experts can flag it, leading us to investigate if a module is over- or under-sensitive. This process will refine the system. Our Advisory Council (described later) will play a key role here, providing an interdisciplinary panel for reviewing results.

Continuous Learning and Calibration. Project Star Forge is intended to be a living system that improves over time. As new data comes in or when real events unfold, we will calibrate the models. For instance, the COVID-19 pandemic has taught many lessons about pandemic modeling and societal response; those insights feed back in (updating the public health module or the heuristics layer around human compliance to health measures). We also anticipate using techniques like Bayesian updating for model parameters as we get more observations – effectively tuning the system to reduce its uncertainty where possible. The knowledge graph's ability to incorporate new empirical data will help here, making calibration a semi-automated process in some cases.

Validation Metrics. We will track specific metrics for performance. These include:

- **Predictive accuracy (when appropriate):** though for long-term foresight this is tricky, for near-term forecasts we can measure how often events simulated actually materialize.
- **Resolution and Recall of Risks:** Does the system tend to forecast too many false alarms or does it miss known risks? We will adjust to calibrate this balance, aiming for a system that "broadly covers the space of plausible futures without overtriggering."
- **Robustness and Stability:** Small changes in input assumptions shouldn't lead to wildly divergent trivial outcomes (unless those represent true tipping points). We test stability by perturbation analysis.
- **Stakeholder usefulness:** Ultimately, a practical validation is whether policymakers and researchers find the outputs useful and insightful. During pilot deployments (like Phase 0), we will gather user feedback on whether the foresight provided actually informed their planning or thinking in a valuable way.

It's important to note we do not promise exact predictions of specific events. Instead, success is defined by offering a rigorous exploration of possibilities. The measure of that is partly qualitative – e.g., did Project Star Forge help uncover a non-obvious vulnerability or opportunity in a system that domain experts agree is plausible and important? Over a 5-year timeline, we expect a few early "wins" where the system highlights a risk scenario later echoed by real events or by other independent analyses, which will build confidence in its utility.

Finally, to bolster credibility, we will publish results of validation studies in academic forums, inviting external critique. This openness to scrutiny ensures that Project Star Forge's insights can be trusted as coming from a thoroughly vetted process, not a black box. In the words of one initiative advocating open modeling, transparency and the ability for others to test models builds the trust and rigor needed for decisions – a principle we take seriously in validating Project Star Forge [1].

5 Conclusion

In this technical overview, we have presented Project Star Forge as an ambitious yet grounded approach to simulating our complex world. Its architecture combines a central orchestration brain with modular domain intelligences, all linked by a rich knowledge graph and enhanced by behavioral realism and ethical safeguards. Technologically, it stands on the shoulders of progress in AI, data integration, and complex systems simulation. Ethically, it is guided by the imperative that such technology serves the global public interest in navigating uncertainty.

Project Star Forge is inherently a collaborative platform – no single team can perfectly model the world, so it is built to integrate contributions from many experts and to remain adaptable. In deploying it as a nonprofit endeavor, we ensure the focus stays on long-term human welfare over short-term gains. The technical blueprint we've described avoids proprietary specifics, but it lays the foundation for an international effort: researchers can envision where they might plug in their expertise (be it improving an SSM or adding a new data stream), and decision-makers can see how the engine's outputs would emerge and be vetted.

To transition from design to reality, strong partnerships and governance are as critical as algorithms. Up next, we outline the practical roadmap for developing Project Star Forge in phases, along with how we intend to resource, govern, and license the project for maximum global benefit. The success of Project Star Forge will not only be measured in the innovations in AI and modeling it produces, but in how it fosters a shared capacity for foresight – empowering humanity to face the future with eyes open and hands steady on the helm.

6 Nonprofit Road map and Governance

To achieve the Project Star Forge vision, we propose a phased development road map over approximately five years. Each phase builds capabilities and broadens the system's scope, while delivering tangible outputs to stakeholders early:

Phase 0 (Year 1): MVP – Climate-Security Demonstrator

Objective: Develop a minimum viable product focusing on the intersection of climate change and security (our initial use-case), to demonstrate the platform's value on a pressing real-world problem.

Scope: This phase implements the core architecture (orchestration engine + knowledge graph + a few key SSMs) at small scale. Specifically, it will include:

- A Climate SSM capable of simulating near-term climate impacts (e.g., droughts, extreme weather frequency in different regions under various emissions scenarios).
- A Security/Conflict SSM that simulates risk of instability or conflict, influenced by factors like resource scarcity, migration, or disaster occurrence.
- The Orchestration Core linking these, and a simplified knowledge graph encoding relevant data (populations, resources, climate indices, etc.).
- A basic Scenario Engine to explore a handful of scenarios (e.g., what if we have a strong El Niño + poor harvest + weak governance in Region X – does conflict erupt?). Emotional-cognitive factors can be rudimentary at this stage (perhaps a simple parameter for government effectiveness or public panic under stress).

Deliverable: A demonstrator report and tool that can simulate, for example, how climate change acts as a "threat multiplier" for conflict [2]. We'll create a scenario such as "2030 with moderate emissions vs 2030 with Paris Agreement goals met" and show differences in security outcomes. This phase validates the end-to-end pipeline on a contained problem and provides a showcase to attract partners and funding. It also allows us to get early feedback from climate and security experts.

Timeline: ~12 months. By the end of Year 1, we expect to have a working prototype running on cloud infrastructure, with results shared in a white paper or pilot study. This MVP will be relatively coarse in resolution but sufficient to demonstrate integrated foresight (e.g., highlighting regions where climate stress could ignite conflicts, which helps policymakers prioritize interventions).

Phase 1 (Years 2–3): Global Knowledge Graph & Domain Expansion

Objective: Scale up the platform into a global knowledge graph core with more domains integrated, transitioning from a narrow demo to a broad decision support system.

Scope: Key activities in this phase:

- **Global Knowledge Graph Build-out:** We will collate and integrate large-scale datasets to form the initial version of the global cross-domain knowledge graph. This includes sourcing data for demographics, economics, infrastructure, environment, health, etc., at global scale. We anticipate using open data from UN agencies, World Bank, Copernicus (for Earth observation), and others. The result will be a graph that maps the world's entities and baseline relationships – effectively a digital twin of Earth's systems (aligned in spirit with efforts like the EU's Destination Earth, which also seeks a digital replica of the planet) [3].
- **Enhance Orchestration & Data Pipeline:** The core will be hardened for scale – ability to handle orders of magnitude more data points and more simultaneous SSMs. We'll incorporate streaming data pipelines for

real-time updates (e.g., daily satellite climate data, news feeds for conflict events), labeled broadly as proprietary multi-modal data fusion systems in our documentation to encompass the necessary but implementation-specific work of connecting to APIs, ensuring data quality, etc.

- **Add Modular Domains:** Introduce additional SSMs beyond climate and security. Candidates: an Economic SSM (global macroeconomy/trade and possibly finance), a Public Health SSM (pandemics and health systems), and a Technology Disruption SSM (e.g., models for AI uptake or cyber risk). We'll prioritize domains that stakeholders (partners, funders) deem most urgent. Each new module will be integrated and tested within the orchestrated simulation. We will also refine the existing climate and security modules based on Phase 0 lessons.
- **User Interface & Visualization:** Develop a basic user interface and dashboard so that scenario outputs can be easily visualized (maps, graphs, etc.) by non-developers. This may include geospatial mapping of results and a scenario configuration UI for researchers.
- **Case Studies:** Use the expanded system for a set of in-depth case studies – for example, a Global Food Security Stress Test that combines climate, economy, and conflict modules, or a Pandemic & Recession Co-impact scenario. These studies will demonstrate multi-domain integration, e.g., how a pandemic could cascade into economic and political domains, and vice versa, aligning with research on cascading risks that involve multiple sectors [8].

Deliverable: By end of Phase 1 (~24–30 months in), we aim for Project Star Forge v1.0 – a functional global foresight platform with a web-based tool or application that partner organizations (e.g., a UN agency or a research institute) can use on pilot basis. The "Global Knowledge Graph" and documentation will be released, likely as open data, providing a resource on its own. We will also produce an academic publication or series of reports detailing the architecture and results of initial integrated scenarios (to maintain scientific credibility and peer review).

Timeline: 2 years for full implementation, overlapping with partnership building. By Year 3, the platform should handle global scale data and multiple domain models, effectively demonstrating the promise of cross-domain foresight. We expect this phase to confirm the feasibility and identify performance bottlenecks to address before full integration.

Phase 2 (Years 4–5): Multi-Domain Integration & Refinement

Objective: Achieve full multi-domain integration with higher fidelity and establish Project Star Forge as a robust, widely-used tool. Essentially, this is scaling from a prototype to an operational system akin to a "World Risk and Resilience Simulator."

Scope:

- **Additional Domains and Detail:** Incorporate any remaining critical domains. For example, Astrophysical Risks (space weather, asteroid threats) could be an SSM if funders like space agencies are involved; Environmental Biodiversity collapse could be another domain. We will also increase the spatial and temporal resolution of simulations as computing allows (e.g., go from country-level averages to subnational detail where data permits, especially for climate and population). Integration of modules becomes more complex here (ensuring consistency across many interlinked systems). By the end of Phase 2, Project Star Forge should cover all major categories of global catastrophic and systemic risk.
- **Performance and HPC Deployment:** As the model complexity grows, we plan to deploy Project Star Forge on dedicated high-performance computing (HPC) infrastructure or cloud supercomputing. We will optimize code for GPU acceleration, distributed computing, etc. It's likely we'll pursue collaborations with national supercomputing centers or cloud providers. For perspective, initiatives like Destination Earth anticipate needing

exascale computing; Project Star Forge, covering even broader domains, will similarly require brute computing power and advanced infrastructure [5]. We expect to leverage modern parallel computing frameworks (possibly using containerized microservices for each SSM orchestrated by a Kubernetes-like system for scaling).

- **Robustness, Uncertainty Quantification:** Develop rigorous methods to quantify uncertainty of outcomes (sensitivity analysis across models, ensemble runs with varied parameters) and to communicate confidence levels to users. This would be akin to how the IPCC reports present ranges and confidence intervals for climate projections, but extended to cross-domain scenarios.
- **Fork Cascade Engine Maturation:** Refine the scenario engine for more complex branching, maybe allowing user-defined branch logics (like exploring specific policy alternatives). We will also develop libraries of "standard scenarios" (best-case, worst-case, business-as-usual in various domains) that can be combined, to help users navigate scenario setup.
- **User Community and Training:** By this phase, we envision a growing community of practice around Project Star Forge. Part of Phase 2 is setting up workshops, training programs, and documentation for external users – scientists, analysts, perhaps government foresight teams – to run their own scenarios or contribute improvements. This transitions Project Star Forge from a project to a platform.
- **Pilot Policy Applications:** Work with partner governments or international bodies on applying Project Star Forge insights to real policy questions. For example, assisting a coalition of countries in scenario planning for climate adaptation investments, or working with a global health body to simulate responses to future pandemics. These real-world pilots will test the system's utility and help refine outputs to be decision-useful.

Deliverable: Project Star Forge v2.0 (end of Year 5) – a full-fledged, multi-domain simulation platform backed by a stable organization (the Project Star Forge Foundation, see governance) and used by a network of collaborators. Deliverables include the platform (software + knowledge graph database + documentation), a series of flagship foresight reports co-produced with partners (for instance, an annual "Global Systemic Risk Outlook" that is analogous to WEF's Global Risks Report but powered by Project Star Forge simulations), and open-source release of significant portions of the code and models. At this stage, we anticipate Project Star Forge could be entrusted as an open, shared resource for anticipatory governance by international coalitions.

Timeline: By Year 4, basic multi-domain integration achieved; by Year 5, refinement and consolidation, moving to continuous improvement mode beyond initial build-out. It's expected that beyond Year 5, the project will transition into an ongoing operation (akin to how weather prediction models are continuously run and improved) with periodic updates.

Throughout all phases, we maintain adaptability. If a major unforeseen risk emerges (say a novel AI scenario or rapid climate tipping not originally prioritized), we have the flexibility to adjust the roadmap to incorporate it. The phased approach ensures we deliver incremental value and learn at each step, rather than attempting a "big bang" world simulation all at once.

Preliminary Resource and Cost Estimates

Building and running Project Star Forge will require substantial resources, but we outline here back-of-the-envelope estimates to guide planning and fundraising. We consider three main cost drivers: compute (cloud/HPC), staffing, and data infrastructure.

Compute (Cloud GPU/CPU hours): In Phase 0, using cloud services for an MVP, we might run on the order of 10^4–10^5 GPU-hours for model training/tuning and scenario runs. For example, if using 100 NVIDIA A100 GPU instances for some months, costs could be a few hundred thousand USD. By Phase 2, with full integration and multiple concurrent scenarios, the needs ramp up dramatically. We anticipate possibly needing on the order of 10^6–10^7 GPU-hours per year for routine operations and experimentation as the platform scales. To put it in context, training a single large AI model like GPT-3 has been estimated at around 3×10^5 GPU-hours (costing several million dollars)—Project Star Forge will encompass multiple models and continuous simulations, although not all as large as GPT-style networks.

We might offset costs by using academic supercomputers; however, if fully on commercial cloud, Phase 2 compute expenses could be on the order of $5–10 million per year or more to maintain an adequately powerful cluster (this assumes negotiated rates or grants for compute, but it's a ballpark). Additionally, high-memory CPUs for knowledge graph queries and multi-node clusters for coupling models are needed. We will also explore use of national labs or international HPC initiatives (EuroHPC, etc.) given the public good nature—for example, Europe's €1 billion EuroHPC project is building exascale systems [5], which might be tapped via partnerships.

Staffing: Human capital is crucial. We foresee a lean core team in Phase 0 (perhaps 5–7 people: AI/ML engineers, a climate scientist, a conflict analyst, etc.) costing around $1–1.5M for the year (fully loaded salaries, assuming nonprofit rates with some in-kind academic contributors). In Phase 1, the team would expand (maybe 15–20 people) including data engineers for the knowledge graph, UI/UX for visualization, additional domain scientists, etc. Annual staffing budget might rise to $3–5M. By Phase 2, we might have ~30 full-time equivalents, including support staff, community managers, etc., with an annual budget in the $5–8M range for personnel.

Over five years, cumulative staffing might thus be on the order of $20M. We anticipate supplementing core staff with partner institutions' contributions (e.g., a university lends a postdoc to develop a module, a government think-tank contributes some analysts). A global advisory network (unpaid or stipend-based) will also amplify expertise without huge cost (more on governance).

Data and Storage: Hosting the global knowledge graph and associated datasets will involve large storage and data egress costs if on cloud. We expect to accumulate petabytes of data (satellite imagery, detailed socio-economic data, simulation output history). A rough estimate: by Year 5 perhaps 5–10 PB of storage needed. Cloud storage at scale might be $100k/PB/year, so maybe $0.5–1M per year for data by later phases. If using on-premises storage in an academic setting, costs might reduce but then maintenance needs to be funded.

We'll also allocate budget for data acquisition (some datasets might have licensing fees, though we'll prioritize open data). Initial phases will use mostly free public data.

Miscellaneous/Facilities: We'll need funds for collaboration infrastructure (communication, servers for development, possibly office space if not fully remote). If a small compute cluster is built in-house for prototyping (Phase 1), that might be a capital expense of a few hundred thousand. Outreach, workshops, and travel for engagement also factor in—maybe $100k/year in early phases, scaling to $200–300k when doing global engagement in Phase 2.

In summary, a very rough cost breakdown could be:

- Phase 0: ~$1–2M (small team, modest cloud usage, mostly development).
- Phase 1: ~$5M/year for 2 years = $10M (larger team, more compute, data integration).
- Phase 2: ~$10M/year for 2 years = $20M (full team and heavy compute).
- **Total:** ~$30M over 5 years, excluding any major in-kind contributions or HPC donations which we will aggressively pursue to lower direct costs.

These figures are provisional. We will refine estimates in a detailed budget plan as the project launches, identifying where partnerships (with government research grants, or cloud credits from tech companies, etc.) can defray costs.

Compared to analogous efforts, this is not unreasonable: for instance, the EU's Destination Earth program, focused mainly on climate digital twins, is funded on the order of hundreds of millions of euros over a decade [3]. Project Star Forge's scope is broader, but our nonprofit status may allow tapping philanthropic and multi-sector support to raise the needed funds.

Return on investment is measured not in profit but in risk mitigation: even slightly better preparedness for catastrophes could save countless lives and billions in economic damage. In that light, a few tens of millions for a global foresight instrument is a prudent and impactful investment.

Governance and Organizational Structure

Project Star Forge will be established with a governance model that ensures accountability, inclusivity, and transparency. We outline the envisioned structure:

Founding Foundation and Board of Trustees. We will incorporate the Project Star Forge Foundation as an independent nonprofit entity (likely a 501(c)(3) in the US or an international equivalent). The Foundation will hold the project's assets (IP, funding) in trust for the public benefit. A Board of Trustees (or Directors) will govern the foundation's fiduciary and strategic decisions. The Board will likely include a mix of experts and stakeholders: e.g., a climate scientist, a representative from an international organization (UN or similar), a technologist, a public policy expert, and a member from the global south to ensure diverse representation.

The Board's duties include overseeing the executive team, ensuring alignment with mission, approving major partnerships and budgets, and upholding ethical standards. We will implement policies (bylaws) that cement the mission—for instance, making it hard to deviate from nonprofit goals or to be taken over by for-profit interests.

Executive Team. The day-to-day operations will be led by an executive director (project lead) and a small leadership team (CTO for technical, Chief Scientist, etc.). They report to the Board. The executive team ensures the roadmap is executed, manages staff, and interfaces with partners. Importantly, they also will work closely with the advisory structures for guidance.

Global Advisory Council. Given the breadth of domains Project Star Forge touches, a Global Advisory Council will be formed, comprised of subject matter experts and stakeholder representatives worldwide. This council is not involved in daily management but provides guidance, review, and advocacy. It may include experts in climate, health, economics, security, ethics, as well as representatives from international bodies (e.g., UNDRR, WHO), civil society, and even community leaders from vulnerable regions.

The role of the Advisory Council is multi-fold:

- **Expert Review:** They review project directions and outputs, giving input on priorities (for instance, advising that certain emerging risk needs more attention, or that scenario outputs should be framed in a certain way for policy uptake).
- **Oversight:** Some members may form sub-committees such as the Ethical Oversight Committee mentioned earlier. They ensure that the project's uses align with humanitarian goals and flag any ethical issues.
- **Outreach and Trust-Building:** Advisors as respected voices can help communicate Project Star Forge's purpose and findings to broader audiences, lending credibility. For a public good project, having Nobel laureates or renowned experts on the council signals trustworthiness, akin to a scientific advisory board for public policy.

- **Membership and Rotation:** We plan for a rotating membership to keep ideas fresh and involve multiple regions. For example, we could have 20 council members with half rotating every 2–3 years. Council positions are likely voluntary or with honoraria; their independence is key.

Tiered Membership and Collaboration Network. In addition to the formal governance bodies, Project Star Forge will cultivate a network of collaborators through a tiered membership model. This is not "membership" in the sense of a club for individuals, but rather partnership tiers for institutions:

- **Core Partners:** Organizations providing significant contributions (funding, data, expertise). They could get a seat in a Partners Forum that meets regularly to discuss project progress and provide input. For example, a space agency contributing satellite data or a foundation providing multi-million funding would be core partners.
- **Contributing Members:** Entities that contribute modules, research, or moderate funding. They might be universities or companies that donate expertise. They get early access to results, and recognition in governance (possibly a say in specific committees).
- **General Members / Community:** Eventually, an open membership where any interested researcher or practitioner can join a Project Star Forge community platform. This community can participate in workshops, suggest scenarios, and use the public versions of the tool. Think of this like a "user group" which ensures we hear feedback from those applying the tool on the ground.

This tiered approach ensures broad inclusion but also clear roles. A similar structure is seen in some open-source foundations (e.g., Linux Foundation has Platinum/Gold/Silver members from companies with corresponding influence, while still remaining community-driven). We will clarify that membership influence is on advisory and collaborative aspects, not on core ethical guardrails which remain non-negotiable.

Transparency Mechanisms

Governance will institute several practices to maintain transparency:

- **Regular public reports on finances, progress, and setbacks.** Being open about challenges helps build trust and invites solutions.
- **Open meetings or published minutes** for certain council or board meetings, when appropriate (except confidential HR or security matters). For instance, an annual meeting summary could be published.
- **Audit and Evaluation:** External auditors will assess the project's finances, and ideally, an external ethics audit could be done periodically to ensure we adhere to stated principles (similar to how some AI projects invite third-party ethics assessments).
- **A commitment to publish source code and data** (with necessary exceptions for security) to the community. This allows independent verification and contributions, reinforcing the notion of Project Star Forge as a shared public endeavor.

Internal Policies

The Foundation will have policies to:

- **Prevent conflicts of interest** (board members will disclose ties, recuse where needed),
- **Ensure diversity and inclusion** in hiring and decision-making (global risks require global perspectives),
- **Enforce data governance** (respecting privacy and not misusing data).

In essence, we align with good governance norms of nonprofits and the specific AI ethics context.

Preventing Capture

An important note is ensuring no single country or corporation can dominate Project Star Forge. We may embed in our charter that the majority of board seats must be independent and not representing for-profit interests. We might also utilize legal instruments (like a requirement that any change to nonprofit status or licensing needs supermajority approval or consent of certain founders) to avoid mission drift.

Institutional Philosophy

Overall, this governance structure is designed to balance expert guidance, stakeholder input, and operational agility. The Foundation provides stability and accountability, the Advisory Council and members provide legitimacy and knowledge, and transparency keeps the project honest in the eyes of the public.

We draw lessons from existing global collaborations – e.g., the governance of large science projects like the IPCC (thousands of contributors but robust review processes), or the way the Global Risk Assessment Framework (GRAF) by UNDRR aimed to network experts to enhance understanding of systemic risk [10]. Project Star Forge's governance aspires to institutionalize a similar culture of collaboration in service of a safer future.

Open-Source Licensing and Collaboration Model

To fulfill our mission as a public good and to encourage worldwide collaboration, Project Star Forge will adopt an open licensing strategy for its outputs and software, with appropriate safeguards:

Software Licensing

We intend to release the majority of Project Star Forge's software under an open-source license. The exact license will be chosen to balance openness with preventing misuse. A strong copyleft license (like GNU AGPLv3) could ensure that improvements stay open and that if someone deploys the system, they can't turn it into a closed proprietary product. However, we are also considering a custom "non-commercial public license" variant – essentially, code is free for non-commercial, research, or public sector use, but if a company wants to commercialize it, they'd need a separate agreement ensuring alignment with our ethics (or contributing back).

This approach is similar to some "Open / Non-Commercial" dual licensing models used by other social-good tech projects. The goal is to maximize accessibility (so that developing country governments, for instance, can use the tool without barriers) while preventing exploitation (we wouldn't want a private military contractor to use it for harmful purposes, as an extreme example).

We will also leverage the concept of *defensive publication*: by publishing our methods openly (as we do with this white paper and subsequent technical releases), we create prior art that discourages patenting of core concepts by others. Our stance is that methodologies for global risk simulation belong in the commons.

That said, any truly sensitive components (e.g., certain security features to prevent tampering, or specific data pipelines under NDA) might remain closed or under a Responsible Disclosure regime to avoid enabling bad actors. But these will be minimal and ring-fenced.

Data and Knowledge Graph Licensing

The integrated global knowledge graph we compile has immense value. We plan to release the knowledge graph data (except where data sources impose restrictions) under an open data license, likely Creative Commons BY (attribution) or BY-SA (share-alike). If some datasets are proprietary or sensitive, we will mark those and try to find open equivalents.

Emphasizing open data aligns with global calls to treat certain data (like climate and disaster risk information) as a public good. It also encourages external researchers to use and improve the data. By open-sourcing the knowledge graph, we foster transparency – anyone can inspect the assumptions and baseline facts feeding the simulations.

Indeed, the KnowWhereGraph project and others have shown the benefit of open, shared knowledge bases for environmental intelligence [11].

Simulation Outputs

Similarly, the results of scenarios and analyses we produce (the foresight reports, etc.) will be published openly. We might use a CC BY license for reports. Visualizations and tools could be made freely accessible via a public web portal. This ensures the widest dissemination and impact – for instance, a small NGO in a developing country could access Project Star Forge scenario insights relevant to their region without cost.

We consider this akin to open science: just as the Human Genome Project made its data public domain daily, Project Star Forge will share global risk “insights” freely.

Collaborative Development Model

We will create an online repository (e.g., on GitHub or a similar platform) for code, and invite contributions from the community. Clear contribution guidelines and a code of conduct will govern this community. This is important for scaling up – domain experts might contribute new SSMs as plugins, or technologists might help optimize code.

An open modular architecture encourages an ecosystem to form, much like how the OpenAI Gym enabled contributions of new environments or how open-source scientific software (like OpenClimateGIS or epidemiological models) grew via community input.

Our licensing (e.g., AGPL) will ensure that if, say, a national lab improves the climate module, they can merge it back to the public codebase, benefiting all.

Non-Commercial Clause Consideration

If we do opt for a non-commercial clause in the license, we will accompany it with an offer for commercial entities to engage via partnership agreements. The reason to consider non-commercial is to safeguard against a scenario where a private firm packages Project Star Forge and sells it exclusively, undermining the public-good intent.

However, truly open source (permissive or copyleft) is more straightforward and fosters community. We’ll likely lean on advice from licensing experts and perhaps model after licenses like the Community Data License Agreement – Permissive/Non-Commercial for data, or the Creative Commons licenses.

Intellectual Property (IP) Ownership

All significant IP created under the project will be held by the Project Star Forge Foundation (for the trust of the public). Contributors and employees would assign IP to the foundation with the understanding it’s released under the open license. This central ownership helps enforce the open license and prevents fragmentation. It also simplifies handling if we integrate any third-party open-source components (we ensure license compatibility and document the provenance of all included code).

Preventing Misuse

An open tool can be used by anyone, including potentially malicious actors. While we can't eliminate this risk entirely (just as open-source encryption or AI algorithms can be dual-use), our approach is to mitigate it. We will publish a usage policy and perhaps implement usage tracking in the public platform (for example, requiring API keys for heavy use) that could be revoked if abuse is detected.

The open license could include a clause that use of the software must not violate international human rights (this is tricky to enforce legally, but at least it states intent). We will coordinate with our ethics advisors to monitor any concerning use and respond (for instance, if a government tried to use it for oppressive surveillance, the community could voice concerns). In extreme cases, because we remain the original developers, we could withhold technical support or updates from known bad actors, while keeping the main project free.

Community and Ecosystem Benefits

By being open-source, Project Star Forge can become a platform for education and research. Universities could use it in coursework on systems thinking or risk analysis. Other nonprofits can fork or adapt it for specific focuses (e.g., a regional version focusing on one continent's issues). This open ecosystem can lead to innovations we didn't anticipate – much like open-source software often does.

And crucially, openness and collaboration build trust. As one nonprofit in healthcare modeling noted, making models open allows others to test and validate them, thereby building trust in the results [1]. We expect the same trust dividend in policy contexts: governments will trust Project Star Forge outputs more if they know the model assumptions are transparent and the code is open to scrutiny.

Final Note

In conclusion, our licensing and collaboration approach is straightforward: **when in doubt, open it up.** We will protect the project's integrity, but not by enclosure – rather by community stewardship. Project Star Forge's very purpose is to be a shared foresight resource for humanity, so open-source and open-data principles are natural choices to ensure it truly belongs to everyone.

This strategy also preempts intellectual property being used to block progress: by openly staking out the ideas and tools, we signal that this knowledge is part of the global commons.

Author & point-of-contact: Thomas Harding `<calmingchronicles85@gmail.com>`

Collaborative Proposal Disclaimer. This document is a **concept-level proposal** intended to inspire open collaboration. It does **not** represent a finished system, commercial offer, or fundraising prospectus. All technical descriptions are aspirational and will require multi-disciplinary input, peer review, and independent validation.

Defensive Publication. Pursuant to 35 U.S.C. § 102(b)(1)(B), the material disclosed herein is deliberately published to establish prior art. Any patents that attempt to claim the architectural flow, orchestration logic, or scenario-forking mechanism herein described after the date of first public release may be subject to invalidation.

References

[1] AJMC Staff. 2020. *Open-Source Modeling Builds Trust, Validity, and Rapid-Cycle Learning in Value Assessment.* The American Journal of Managed Care. https://www.ajmc.com/view/open-source-modeling-builds-trust-validity-and-rapid-cycle-learning-in-value-assessment News/press article summarizing IVI brief; not a peer-reviewed journal article.

[2] John D. Banusiewicz. 2014. *Hagel to Address 'Threat Multiplier' of Climate Change.* U.S. Department of Defense. https://www.defense.gov/News/News-Stories/Article/Article/603440/hagel-to-address-threat-multiplier-of-climate-change/

[3] Anni Jakobsson. 2025. *Destination Earth and Digital Twins – A European Opportunity for HPC.* LUMI Consortium. https://www.lumi-supercomputer.eu/destination-earth-and-digital-twins-a-european-opportunity-for-hpc/

[4] Prabin Maharjan. 2025. Policy Brief No. 174: Leveraging Strategic Foresight to Mitigate Artificial-Intelligence Risks. UN DESA Policy Brief. https://desapublications.un.org/file/21159/download 5 pages.

[5] Sebastian Moss. 2020. *The Earth's Twin.* DatacenterDynamics (DCD). https://www.datacenterdynamics.com/en/analysis/earths-twin/

[6] Alexander De Ridder. 2024. *Agent-Based Modeling in Psychology.* SmythOS. https://smythos.com/managers/healthcare/agent-based-modeling-in-psychology/ Author shown on page; publication date not explicitly listed.

[7] Andre Ripla. 2025. *Leveraging AI for Corporate Foresight and Scenario Planning.* LinkedIn. https://www.linkedin.com/pulse/leveraging-ai-corporate-foresight-scenario-planning-andre-0sc3e/

[8] Trond Arne Undheim and Taimur Ahmad. 2024. Quantitative Scenarios for Cascading Risks in AI, Climate, Synthetic Bio, and Financial Markets by 2075. *Frontiers in Complex Systems* 2 (2024), 20 pages. https://doi.org/10.3389/fcpxs.2024.1323321

[9] UNESCO. 2021. Recommendation on the Ethics of Artificial Intelligence. UNESCO Recommendation. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

[10] United Nations Office for Disaster Risk Reduction. 2022. *Global Risk Assessment Framework (GRAF).* UNDRR. https://www.preventionweb.net/understanding-disaster-risk/graf

[11] Rui Zhu, Cogan Shimizu, Shirly Stephen, Colby K. Fisher, Thomas Thelen, Kitty Currier, Krzysztof Janowicz, Pascal Hitzler, Mark Schildhauer, Wenwen Li, Dean Rehberger, Adrita Barua, Antrea Christou, Ling Cai, Abhilekha Dalal, Anthony D'Onofrio, Andrew Eells, Mitchell Faulk, Zilong Liu, Gengchen Mai, Mohammad Saeid Mahdavinejad, Bryce Mecum, Sanaz Saki Norouzi, Meilin Shi, Yuanyuan Tian, Sizhe Wang, Zhangyu Wang, and Joseph Zalewski. 2025. The KnowWhereGraph: A Large-Scale Geo-Knowledge Graph for Interdisciplinary Knowledge Discovery and Geo-Enrichment. arXiv preprint. https://doi.org/10.48550/arXiv.2502.13874 arXiv:2502.13874 [cs.DB] Preprint; page count not listed on abstract page.