World Agent

Public Specification

A Spatial Intelligence Layer for AI Systems ·

Author	Dave Lorenzini, World Agent LLC
Version	1.1 · Public
Date	2026
Document	Public specification (this document)
Companion	World Agent v9. Formal Specification (available on request)

This is the public specification for World Agent, a spatial intelligence platform that gives AI systems a real understanding of physical space, not a map, not a knowledge graph, but a living layer of place identity, memory, prediction, and trust.

It is intended for technical reviewers, researchers, journalists, and prospective partners who want to understand the architecture and its commitments before requesting the full formal treatment.

What this document is, and what it is not

This document is the public specification. It covers the platform's posture, primitives, the fifteen canonical questions, the agent hierarchy, the trust and provenance model, and an outline of the formal foundations. It states the architecture's central theorems and the intuition behind each, without reproducing the proofs. A working example threads through the document so the reader sees the system applied to one vertical, real estate, in concrete terms. The full mathematical treatment, complete proofs, and detailed implementation specifications live in World Agent v9, available on request to qualified researchers and partners.

Abstract

Modern large language models reason about the physical world through brittle composition of static maps, scraped reviews, and unverified text, a regime that produces fluent guesses indistinguishable, to the model itself, from grounded knowledge.

World Agent is a spatial intelligence layer in which every place, from a continent to a parking spot , is an addressable agent with identity, memory, live state, and a duty to answer a fixed canonical question set under explicit lineage.

The contribution is mathematical as well as architectural: the platform formalizes (i) a cohort-conditioned belief calculus that treats experiential facts as conditional probability measures rather than point estimates, (ii) a non-distributive trust lattice on the four-tuple (authenticity, reputation, corroboration, recency) under which any composite scalar trust score is provably lossy, (iii) an origin-type soundness result that rules out generated-to-observed laundering by construction, (iv) a coupled-ODE model of reflexivity with a closed-form gap between observed and unobserved place quality, and (v) calibration and Sybil-resistance bounds on the consensus engine that promotes Claims to Facts.

These formalisms compose into a tractable runtime: hierarchical query complexity independent of system size, a three-mode execution ladder respecting a Pareto frontier between latency and lineage, and a per-agent privacy gate that enforces consent as a precondition for execution.

Together they constitute a defensible technical surface for grounding AI agents in physical space.

1. Posture and Mission

World Agent is the answer layer for any AI operating in physical space. Three posture rules, index don't collect, report belief not truth, local-first and consent-bound , constrain every primitive, every endpoint, and every roll-up. They are not slogans; they are structural commitments the rest of the system instantiates.

Principle	What it means in practice
Index, don't collect	the world's information stays at its source; we hold identity, memory, and links , not the originals
Report belief, not truth	every answer carries lineage, confidence, and contestation; the system never claims to know
Experience depends on the observer	experiential facts are always cohort-conditioned; one place can be many things to many people
Failure is explicit	every failure surfaces structurally to the caller, never silently substituted; the system distinguishes between absence of evidence, stale evidence, contested evidence, and judgments it is unwilling to make on its own
Local-first, consent-bound	personal facts and live observations stay on the device that produced them by default; sharing requires explicit, scoped, revocable consent

The core claim
Most spatial systems describe places. World Agent gives places the ability to answer. Every place, a continent, a building, a queue, a balcony, is an agent with identity, memory, current state, near-term prediction, and a cohort-conditioned answer to any question a user, an AI, or an operations system can ask. The substrate underneath the next decade of spatial computing.

2. The Two-Layer Truth Model

Every fact in the system is either operational or experiential. The distinction is not stylistic; the two require different pipelines, different decay, different aggregation, and different trust models. Mixing them is what produced the failure mode that broke every review-aggregation platform of the last two decades.

	Operational facts	Experiential facts
Example (general)	is the door open; how many people inside; ambient noise dB; current price	is this place welcoming; is it prestigious; does it fit a quiet evening

	Operational facts	Experiential facts
Example (real estate)	price $4,200/mo; 720 sqft; year built 1928; PS 321 school zone	quiet block; good light; family-friendly; signals brownstone Brooklyn
Truth model	converges to a single value under sufficient evidence	irreducibly plural, varies by audience
Default shape	point value in some value space	probability distribution conditioned on cohort
Aggregates as	sum, max, mean, cleanly	mixture under cohort weights, never collapsed
Decay	seconds (occupancy) to years (built year)	weeks to years (norms shift slowly)
Question class	Q1-Q11 (operational endpoints)	Q12-Q15 (experiential endpoints)

2.1 Why this distinction is the moat

Every incumbent collapses experiential facts to a point estimate, a 4.6-star rating, a walkability 92, a quiet-block 5/5. The system below their UI cannot ever recover, by any post-processing step, that 22 percent of reporters in some cohort experienced the place as the opposite of what the centroid reports. This is not a UI choice. It is a structural property of the data they store. World Agent stores the cohort-conditioned distribution as the canonical fact; the unconditional centroid, if anyone wants it, is a derived view.

Theorem 1. Information loss of unconditional aggregation
Let the cohort-conditioned distributions for an experiential fact be a family of probability measures, one per cohort. The projection that returns the unconditional weighted average across cohorts is provably non-injective: distinct families of cohort-conditioned distributions project to the same scalar. The cohort lineage cannot be reconstructed from the aggregate, no matter how many users contribute. Intuition: when you average across audiences who fundamentally disagree, the answer that comes out is one no audience actually holds, and the algebra that produced it is not invertible. (Full proof in v9.)

3. Five Primitives

The whole system rests on five types. Three subtypes of Fact handle the experiential layer. Every endpoint, every rollup, every signal, every policy is built from these primitives.

Primitive	What it is
Agent	every addressable place, object, actor, or activity in the system; has identity, geometry, an endpoint, a memory of facts, capabilities, and a duty to answer the fifteen questions about its scope
Fact	anything known about an Agent, time-stamped, sourced, trust-scored, origin-tagged; every Fact carries provenance and the four trust scores documented below
ExperientialFact (subtype)	how a cohort experiences the place; a probability distribution over outcomes, not a single value; always carries an audience cohort or is explicitly marked unconditional
CausalClaim (subtype)	what the place tends to cause in people who interact with it; every CausalClaim must include a counterfactual and an evidence class, or it is rejected at write time
NormFact (subtype)	an unwritten rule with enforcement and violation cost; norms govern affordances ("laptops welcome until 4 pm", "no strollers on the upper level")
Relation	typed, time-bounded link between two Agents; the predicate algebra covers containment, control, scheduling, contestation, derivation, and other classes; full predicate set in v9
Frame	the temporal lens on every read; covers current state, historical points, future expectations, and hypothetical scenarios; full enumeration in v9
Claim	input awaiting consensus before becoming a Fact; every Claim is graded through the consensus engine before promotion

Rules that hold at the primitive level

An ExperientialFact without an audience cohort is a bug. A CausalClaim without a counterfactual is a bug. A read without a Frame defaults to "now" but is never silent. A Fact without an origin tag cannot be written. These are not coding conventions; they are type-system constraints that the runtime enforces by refusing malformed inputs.

4. The Agent Hierarchy

Agents nest. Every Agent has at most one parent and any number of children. The hierarchy is the substrate over which roll-up and drill-down compose. A query against a parent Agent returns a summary of its children with explicit drill-down; a query against a child Agent returns its own state, with inherited context from its parents.

Agents nest. Every Agent has at most one parent and any number of children. The hierarchy spans multiple levels of scope, from global down to local, with a bounded depth that preserves

query complexity at scale. A query against a parent Agent returns a summary of its children with explicit drill-down; a query against a child Agent returns its own state, with inherited context from its parents. The level scheme, addressing convention, and granularity rules are specified in v9.

The depth is bounded. Roll-up traverses the tree; query latency is independent of the system's total size. The architecture is correct for any number of Agents; the engineering targets assume a working point near 10^9 Agents at steady state.

Theorem 6. Hierarchical query complexity
Any roll-up query against an Agent in a hierarchy of bounded depth and bounded branching can be answered in time proportional to the product of the branching factor and the depth, independent of the total system size. Intuition: the work happens at write time, not read time, children update their parents incrementally; reads touch at most one cached aggregate per level. (Full proof in v9.)

5. The Canonical Question Set

Every query an AI can ask of the spatial layer reduces to one of a small fixed set of canonical questions. The set is divided into operational and experiential questions; both kinds carry lineage; neither claims truth.

Kind	What the questions cover
Operational	current state of a place; its history and changes over time; spatial proximity to other places; legal and control structure; affordances available to a visitor; the graph of relations to other agents; predicted state over near-term horizons; discovery across the system; explicit absence of evidence; appropriate channels for contact
Experiential	fit between a place and an asker conditioned on cohort; the meaning of being at a place and what visiting signals; causal effects the place tends to produce in people like the one asking; contestation between cohorts who experience the place differently

The full set, with canonical wording, endpoint shapes, and return-type schemas, is specified in v9. Vertical packs (brokerage workflows, robot navigation, tour-guide narration, civic ops dashboards) are synthesis on top of this fixed surface, not parallel APIs. The contract is the same regardless of who is asking; what differs is which cohort the experiential answers are conditioned on and which signals the operational answers draw from.

6. A Worked Example. One Brooklyn Listing

The architecture is most concrete when applied to a vertical. Real estate is a useful first wedge because it exercises every load-bearing piece of the platform: vertical geocoding (parcel → building → unit → balcony), multi-source consensus across MLS, Zillow, Redfin, Compass, county records, and FEMA feeds; operational facts (price, square footage) and experiential facts (quiet block, signals investment-grade); causal claims with money attached (appreciation, days-on-market); and a privacy posture that survives a mortgage-lender security review. The example below is one listing at one moment in time; the same fifteen questions apply to every other vertical the platform is instantiated against.

The listing
"Sunny 2BR in pre-war Park Slope, 4:00 pm Tuesday, asking $4,200 per month." The listing has a canonical address as a child of the building, which is a child of the parcel, which is a child of the Park Slope neighborhood agent, and so on up to the root. Every question below queries that single canonical place.

6.1 What the canonical questions return

The illustration below shows the shape of the answers an AI agent receives from the spatial layer. The specific calibration values, confidence intervals, and contestation thresholds shown are illustrative; the actual numbers a deployment produces depend on the calibration history of the contributing reporters and predictors, and are specified in v9.

Question kind	Shape of the return for this listing
State	current asking price, recent price changes, market status, time on market
History	prior rent levels, prior tenancy patterns, vacancy duration across cycles
Control	owner of record, property manager, related holdings on the block
Connected	school zone, transit proximity, flood-zone status, easements, school district
Prediction	probability distribution over rent outcomes by week, with a stated confidence and calibration history; price-cut risk over a defined horizon
Gaps	explicit absence of evidence — missing floor plan, no recent air-quality scan, no daytime light measurement — returned rather than silently substituted
Fit	cohort-conditioned match scores for the asker's cohort, with per-cohort confidence intervals; the system never collapses the cohorts into a single rating
Meaning	what visiting the place signals, conditioned on the cohort interpreting the signal; honest about which cohorts read the signal which way

Question kind	Shape of the return for this listing
Causal effect	what the place tends to cause in people like the one asking, with a required counterfactual comparison set; absence of a defensible counterfactual returns "insufficient evidence"
Contestation	where cohorts disagree about an experiential axis, returned as a structured disagreement with both sides preserved — never resolved by the platform

What Q14 refuses to do

The causal-effect endpoint refuses to fabricate. If no CausalClaim has a defensible counterfactual and a non-trivial evidence class, the system returns "insufficient evidence." It does not synthesize a number. The discipline keeps the experiential layer from collapsing into broker marketing copy, and is what makes the system trustworthy for downstream AI agents that have to compose forty-step plans without hallucinating any one step.

6.2 Why this is unreachable from the incumbents' data

Today, an AI agent shopping for a Park Slope rental does fifteen ad-hoc string parses across six websites and gets fifteen incompatible answers. The Zillow surface returns a single Zestimate; the Redfin surface returns a single estimated rent; the StreetEasy surface returns a 4.6-star neighborhood score. None of them carries cohort lineage; none of them refuses to estimate when the evidence is insufficient; none of them surfaces contestation. The canonical question set is what lets an AI agent compose a multi-step plan without any one step hallucinating. This is the contract the platforms above World Agent want the underlying layer to honor, and none of them currently can.

7. Trust, Provenance, and Consensus

Every Fact carries four trust scores forming a 4-tuple: authenticity (was the report from who it says), reputation (how the source has performed on this kind of claim historically), corroboration (how many independent sources concur), and recency (how fresh is the evidence). The 4-tuple is exposed at the API; any single composite scalar derived from it is structurally lossy.

Theorem 2. Composite trust is provably lossy
The set of trust 4-tuples forms a non-distributive lattice. The map that compresses the four into a single weighted scalar is monotone but provably non-injective, distinct trust tuples collapse to the same scalar. Intuition: "signed source plus corroboration" and "fresh report by a reputable actor" are not the same kind of belief, and downstream decisions should not treat them as interchangeable. A 4.6-star score that collapses these into one number has destroyed the information that would let a calling AI agent decide whether to trust the 4.6. World Agent therefore exposes the four-tuple as authoritative; the composite scalar is offered as a convenience, never as authority. (Full proof in v9.)

7.1 Promotion from Claim to Fact

Inputs enter the system as Claims, not Facts. The consensus engine grades each Claim against the existing Facts about the same field, weights it by the reporter's trust scores, accounts for the reporter's independence from prior reporters, and promotes or rejects accordingly. The engine's false-promotion rate is bounded by its expected calibration error plus a vanishing term in the number of historical Claims observed.

Sybil resistance is structural, not heuristic
Three colluding accounts that share an autonomous system number, posting cadence, and claim sequence form a clique in the reporter-similarity matrix. The platform bounds the effective independent count via the algebraic connectivity of that matrix, in the limit of perfect collusion, the clique counts as one reporter regardless of its size. Collusion is therefore a structural cost paid in lost weight, not an attack to be detected after the fact. (Full bound and proof in v9.)

8. Origin-Type Discipline

Every Fact carries an origin tag: observed (sensors or humans reported it), derived (computed deterministically from other Facts), generated (produced by a model), or synthetic (a hypothetical or scenario). The platform refuses, by construction, to let generated or synthetic

content flow into the present-tense answer of any Q1 or Q5 query. This is enforced by the type system, not by policy.

Origin tag	Example in real estate	Where it may appear
Observed	sensor reading; broker-confirmed photo; FEMA flood zone designation	any frame, any question
Derived	Walk Score computed from public roads and amenities	any frame, any question
Generated	diffusion-staged listing photo; AI-described "vibe" of a neighborhood	scenario frame only; never present-tense ground truth
Synthetic	what-if analysis under a proposed zoning change	scenario frame only; never present-tense ground truth

Theorem 3. Origin-type soundness
Origin classes are types, not metadata. The admissible Fact-producing operations form a typed category whose only morphisms with codomain in the present-tense frame have domain in the observed or derived classes. There exists no finite composition of admissible operations from the generated or synthetic classes that produces an observed-typed Fact, and no Fact tagged generated or synthetic is admissible into the present-tense answer of Q1, Q5, or Q14. Intuition: a metadata field can be erased or ignored at any layer; a type cannot, a function whose signature requires an observed-typed argument will not accept a generated-typed value, full stop. (Full proof in v9.)

The property survives a sophisticated adversarial review. Diffusion-staging of listing photos is commercially deployed today across the real-estate industry; most platforms label it inconsistently or not at all. The first lawsuit over hallucinated fixtures will turn on whether the platform structurally distinguished observed from generated, and most will lose that argument. World Agent's discipline survives the argument because the discipline is in the type system.

9. Reflexivity. The Failure Mode Incumbents Have Lost

Every successful platform that surfaces places has the same problem: surfacing a place shapes the place. A café flagged "great for studying" becomes crowded with studiers, which makes it worse for studying. A Brooklyn block flagged "up and coming" attracts buyers and accelerates the displacement that breaks what made it interesting. Yelp, TripAdvisor, Google Maps, and Zillow have all fought this and lost, not because they didn't see it, but because their architectures had no concept of the gap between the place they were reporting on and the place that existed before they reported on it.

Theorem 4. The reflexivity gap exists and is characterizable
Under a coupled-ODE model of place quality and platform exposure, the system has a unique stable fixed point at which observed quality is strictly less than the unobserved baseline. The gap between them, the reflexivity gap, is non-zero whenever the platform both suppresses through exposure and surfaces good places (which is to say, always). Intuition: the place the platform reports on is not the place the platform was first told about , the platform has changed it, and the gap is measurable. World Agent estimates the gap online and uses it to throttle discovery so the system does not eat what it surfaces. (Full ODE, fixed-point derivation, and stability proof in v9.)

The platform estimates the gap online and uses it operationally: places that the platform is actively shaping receive less discovery weight; places the platform has not yet touched receive more. The system actively prefers not eating what it surfaces. The mechanism for estimating the gap, the per-axis governance rules, and the discovery-throttling policy are specified in v9. The discipline distinguishes a sustainable spatial intelligence layer from one that destroys its own ground truth.

10. Cohort-Conditioned Belief

Experiential facts are conditioned on cohort. A cohort is a governed first-class entity: a label, a membership predicate, a scope, a parent cohort, and a reputation history. Cohorts are not invented per query; they are curated, versioned, and subject to anti-discrimination review.

Cohorts are co-developed with domain research teams during deployment. The registry content is vertical-specific (real estate distinguishes different cohorts than destination operations or robotics); the architectural primitives are the same across verticals. Specific cohort definitions, membership predicates, scopes, and anti-discrimination governance rules are specified in v9 and in the relevant vertical-pack documentation.

The contestation index, in plain language
When two cohorts experience a place incompatibly, the platform measures the disagreement with an information-theoretic divergence between their distributions. If the divergence exceeds a field-class threshold, the place is contested on that axis. The contestation endpoint returns the cohort pair in maximum disagreement plus the full divergence structure. The platform never resolves the contestation; it surfaces it. The exact divergence measure, threshold rules, and return-shape are specified in v9.

11. The Signals Network

Facts descend from Signals. A signal is a source of evidence about places, a sensor, a data feed, a human reporter, a perception worker running over a representation. The platform classifies signals into four classes by access posture (public, premium, private, domain) and supports multiple trigger modes including continuous and event-driven activation. Every signal's contribution to consensus passes through the same trust pipeline.

Signal class	What it is
Public	open, broad-coverage signals; freely available; establish baseline coverage
Premium	licensed or partner signals; improve confidence, prediction, and coverage on high-value places
Private	client-controlled signals; never leave the tenant; enable operational deployments where data confidentiality is a requirement
Domain	vertical-specific signal packs for specialized deployments

Signals do not become truth automatically
A signal becomes evidence; evidence is graded; consensus promotes evidence to Facts. The promotion is calibrated, audited, and reversible. A platform that auto-promotes signals , as the incumbents do for most of their data, cannot honestly say where its beliefs came from when something goes wrong. World Agent always can.

12. The Representation Layer

Splats, NeRFs, meshes, photogrammetry scans, BIM models, and generative scenes are all representations of places. World Agent treats every representation as a first-class Agent with explicit origin tag (observed scan, derived mesh, generated infill, synthetic scenario), fidelity, coverage, and lineage. Perception workers running over representations emit Claims with origin tag "derived" and full method metadata. The same trust pipeline that handles human reports handles model-emitted Claims; the same origin-type discipline fences generative content into the scenario frame.

This composition gives the platform two abilities most spatial systems lack. First, it can see for itself, a perception worker running over imagery emits Facts about a place the same way a sensor or a human would, with comparable trust treatment. Second, it can compose with generative spatial AI safely, a diffusion model that fills missing parts of a NeRF cannot, by any chain of admissible operations, produce a Fact that appears as observed ground truth at Q1. The generative layer is fenced into counterfactual reasoning, where it belongs.

13. The Prediction Substrate

World Agent does not ship its own world model. It is the substrate that grades them. Predictive systems, latent world models in the JEPA family, domain simulators (hydrology, traffic, real-estate pricing), operator forecasters, and the system's own near-term rollups, plug in as a signal class, emit Claims about the future, and are calibrated against observed outcomes through the same feedback pipeline that grades every other reporter.

> **Other systems predict the world. The substrate grades the predictions.**
> Multiple predictors will disagree about the same future. The platform's expected-frame return is a cohort-weighted, calibration-weighted ensemble, never a single model's output presented as truth. Predictors that calibrate badly get throttled out. Calibration history per (place class × horizon) is itself a moat, the data accrues only with time and cannot be reconstructed from outside.

Predictions are time-boxed. A prediction whose outcome window has closed becomes a calibration datum, not a Fact about the world. Predictions never feed Q1 (state) or Q5 (control); origin-type discipline holds at the prediction surface as strictly as anywhere else. A prediction is generated evidence about the future; it is never authority about the present.

14. Execution Modes and the Latency-Lineage Frontier

Different consumers of the platform want different tradeoffs between latency and lineage. A robotics control loop wants an answer in single-digit milliseconds; an AI agent composing a booking workflow wants the answer in well under a second; a sophisticated reasoner doing a counterfactual analysis can wait seconds. The platform exposes multiple execution modes along a single Pareto frontier, with each mode appropriate to a class of caller and with structured fallback between them. Mode names, exact latency targets, and the lineage available at each tier are specified in v9.

Caller class	What they get
Robotics-grade latency	minimal lineage, cached state appropriate for sub-perception-cycle decisions; offline-capable
AI-assistant latency	moderate lineage including trust scores, cohort summary, predicted state; the default tier for most calling agents
Decision-grade depth	complete lineage including full provenance, cohort distribution, contestation structure, and scenario frames; appropriate for legal review, regulatory review, and counterfactual analysis

Theorem 7. Mode-correct degradation
Every answer carries a tier label that is correctly attached. The lineage available at a lower tier is a strict subset of the lineage available at a higher tier. The answer is never silently degraded, every fallback step records its trigger as an auditable fact about the query. Intuition: a robotics agent that requested decision-grade depth but received a faster tier sees that fact, and can decide to retry, escalate, or act anyway. The posture against silent failure holds across the entire latency frontier. (Full proof in v9.)

15. Privacy and Consent

Privacy is structural, not procedural. The platform holds public belief about places , what is here, who controls it, what it means to whom. It does not, by default, hold private belief about people. Personal facts and live observations stay on the device that produced them; sharing requires explicit, scoped, revocable consent; every read of personal data is audited and visible to the actor whose data it is.

Property	Meaning
Scoped	consent is for a specific purpose, agent, and duration, never a blanket grant
Revocable	the actor can withdraw any consent at any time; pending tasks honoring that consent are cancelled
Visible	every read of personal data is logged in a place the actor can inspect; consent state is itself queryable as a fact about the actor
Non-derived	an agent that received data under one scope cannot use it to derive a fact under a different scope
Cross-jurisdictional	consent grants re-evaluated at every regional boundary; stricter of (origin, destination) policy applies by default

Why this matters for AI agents that complete transactions
An AI agent that books a rental, files a mortgage application, or completes a destination purchase is acting under principal authority. The privacy posture has to be defensible to regulators (GDPR, CCPA, COPPA, sectoral rules) and auditable to the principal. World Agent's consent-as-precondition discipline is exactly the structure that survives regulatory review, not because the system promises to behave well, but because the runtime refuses to act without scoped consent in the first place.

16. Failure Modes. Explicit, Not Silent

Every answer either succeeds or fails in a structured way. The platform never silently substitutes a default, never hallucinates a plausible-looking answer, and never promotes low-confidence evidence to authoritative state. Calling AI agents see the failure class structurally and can decide what to do about it.

The failure taxonomy distinguishes between absence of evidence, evidence that is stale, evidence that is contested between sources, evidence that is below the confidence floor for its field class, and questions whose answer requires a judgment the system is structurally unwilling to make on its own (legal, medical, safety-critical). The full enumeration, with the exact return shape for each class, is specified in v9.

Silent failure is the worst failure
A spatial platform that returns a confident-looking wrong answer is more dangerous than one that returns no answer at all. Every incumbent does the former on at least some fields; World Agent does the latter, by construction, on every field where evidence is insufficient. The property makes the system safe for AI agents to compose multi-step plans against, the agent can always tell what the system does not know.

17. Roll-Up. What Compresses and What Doesn't

Roll-up turns children into summaries at their parents, propagating up the hierarchy. The promise that regional queries return in milliseconds rather than fanning out to millions of children depends on what is allowed to compress as state moves upward, and, equally, on what is forbidden from compressing because compressing it would destroy the property the system exists to preserve.

The protocol defines per-field-class rules. Some kinds of field aggregate losslessly across children; some summarize with bounded loss and always retain a drill-down path to the underlying detail; some are structurally forbidden from rolling up at all (provenance chains, personal facts, claims that have not promoted to facts, synthetic and scenario content). The specific rules per field class are specified in v9.

Confidence floors travel up the hierarchy
A parent's confidence in a lossy-aggregated field is bounded by the lowest-confidence child that contributed. Confidence cannot be manufactured by aggregation. The hierarchy is real because what flows through it is disciplined; if everything flowed, the hierarchy would be a lie that hides its own stale aggregates.

18. HearHere. A Working Deployment

HearHere is a live product running on the World Agent protocol, currently deployed in Valletta, Malta. It uses 56 spatial agents across six hierarchy levels (continent, country, city, district, structure, microspace), 12 geofence-triggered audio cues, interest-aware routing conditioned on visitor cohort, and a spatially-grounded chat AI that receives the user's coordinates plus the nearby agents from the spatial graph and answers from them rather than from training data.

HearHere is not a prototype. It is a live product running on the same canonical question set, the same trust pipeline, the same cohort registry, the same posture rules documented in this specification. It is the empirical proof point that the architecture works in production, not in theory. The spatial-narrative thread is the smallest viable version of what the full platform enables across destination operations, real estate, robotics, civic operations, and the other verticals the architecture is designed for.

19. What is Built, What is Next

The platform is at Phase 1 of a four-phase roadmap. Phase 1 deploys the architecture against one vertical (real estate) in one neighborhood (Brooklyn) with three named brokers, plus the parallel HearHere deployment in Valletta. Phase 2 extends to additional verticals and additional geographies, with the calibration history compounding across domains. Phase 3 deepens vertical integration into operator stacks (brokerage CRMs, destination operations centers, robotics fleets). Phase 4 is the platform's mature form, the substrate every spatially-aware AI agent reaches for by default.

Phase	Scope	Time horizon
Phase 1	One vertical, one geography, one parallel deployment	now. Q3 2026
Phase 2	Multiple verticals; geographic and partnership expansion	Q4 2026, mid-2027
Phase 3	Deep operator integration; cross-vertical calibration accrues	mid-2027. 2028
Phase 4	Default substrate for spatially-aware AI agents	2028, ongoing

Why now
AI agents are crossing the threshold from "chat assistants" to "actors that complete transactions." The current spatial substrate they sit on, scraped listings, static maps, point-estimate review scores, was built for human eyes, not for agents that compose forty-step plans. The data quality, lineage, and cohort posture that a human can paper over with judgment, an agent surfaces as a contract breach. There is a six-to-eighteen-month window to be the canonical answer layer agents reach for before either the incumbents retrofit

(badly) or every PropTech, robotics, and destination company builds its own incompatible spatial substrate.

20. Falsification

Architectures that look elegant on paper are usually wrong somewhere. The discipline is to find out where, fast. The smallest viable falsification plan for the platform: pick one vertical with high cohort variance and tractable data (real estate), pick one geography dense enough for substitution and contestation to matter (Brooklyn, three zip codes), try to answer every canonical question for every listing in scope with currently-collectable data, and catalog every failure, which cohorts are missing, which causal claims have no defensible counterfactual, which contestation goes unsurfaced.

What the architecture is steered by
Same discipline as the rest of the system: report belief, not truth, including beliefs about the architecture itself. The architecture is steered by what bends first under real load, not by what looks most elegant on paper. The papers and pilots exist to surface the bending; the platform is updated where it bends.

21. Closing

Fifteen questions. One canonical id per place. Three posture rules. Belief, not truth. AI should not guess about the world. It should ask it. World Agent makes that possible , and the mathematics summarized in this document, fully developed in the v9 specification, is what makes the asking trustworthy.

Where to go next
If this document has earned a serious read and you want the full formal treatment, the complete mathematical foundations, the proofs of the theorems summarized above, the detailed implementation specifications, the agent-consumable manifest, request access to World Agent v9 at the Research page on worldagent.exe.xyz. Researchers, technical partners, and qualified reviewers are welcome. Commercial partnership inquiries go through the partner channel.

About the author

Dave Lorenzini leads World Agent LLC, creator of the World Agent and 4D-ID spatial intelligence platforms focused on helping AI systems understand, reason about, and interact with the physical world.

A veteran technology executive and product strategist, Dave has contributed to several foundational spatial and immersive technologies, including Keyhole (which became Google Earth), Google Glass at Google X, global satellite imaging systems at Space Imaging, and advanced XR and immersive platforms through Draw & Code and ARc.

His work centers on spatial computing, AI, XR, mapping, and real-time world understanding. Through World Agent and 4D-ID, he is developing machine-readable spatial intelligence systems designed for AI agents, robotics, autonomous systems, and next-generation spatial computing platforms.

He is based in the Bay Area and continues to support the broader XR and spatial computing ecosystem through XR Guild and global AWE XR events.

How to cite

```
@techreport{lorenzini2026worldagent_public,
 author = {Lorenzini, Dave},
 title = {World Agent: A Spatial Intelligence Layer for AI Systems},
 subtitle = {Public Specification},
 year = {2026},
 number = {v1.0-public},
 institution = {World Agent LLC},
 note = {The full formal specification (v9) is available on request.}
}
```