

# The AI Consulting Platform Built to Close the Value Gap

How Greydog AI combines a nine-agent orchestrated system, a human-governed trust architecture, and twenty years of consulting methodology to turn AI investment into measured business outcomes — for organizations that can't afford to guess.

# Contents

---

- 01** Executive Summary
- 02** The Problem: Why AI Investment Fails
- 03** The Greydog Approach
- 04** Platform Architecture
- 05** The Agent Roster
- 06** The Engagement Workflow
- 07** Trust, Governance & Data Handling
- 08** Technical Differentiators
- 09** Anchor: The Continuous-Improvement Layer
- 10** Engagement Models & Investment
- 11** What Success Looks Like
- 12** About Greydog AI

# Most AI investment produces no measurable return.

## Greydog exists to close that gap.

---

Greydog AI is a multi-agent consulting platform that diagnoses where AI actually creates value inside an organization, builds or coaches the implementation, and stays accountable to the outcome — encoded from twenty years of enterprise consulting practice into a coordinated team of specialized AI agents, governed by a single accountable human.

### THE FOUNDATION, NOT A VALUE STATEMENT

Every diagnostic, every recommendation, and every build decision traces back to one question, asked first and checked last: what does the customer actually need, and how do we deliver it? That is not one input among several — it is the reason the rest of the methodology exists. Evidence can add nuance or challenge an assumption. It never replaces what the customer said mattered in the first place.

The platform is not a chatbot wrapper and not a generic AI-strategy deck generator. It is a structured, multi-session engagement system: nine client-facing agents and seven internal agents, each with a defined role, operating across a real value stream — Discover, Guide & Build, Kaizen — with real scheduling, real document generation, real e-signature, and a real self-improvement loop that lets the system get measurably better at serving a given client over time.

This paper is written for technical and business buyers evaluating whether an AI-native consulting model can replace, augment, or outperform a traditional consulting engagement — and for platform buyers assessing Greydog specifically on architecture, governance, and defensibility, not just marketing claims.

## 02 — THE PROBLEM

# Why AI investment fails: it's not the technology.

Three independent research bodies converge on the same finding: organizations are not failing at AI because the models are inadequate. They are failing because they never established where AI creates real value before they started building.

**95%**

of organizations running GenAI pilots see no measurable return to the income statement.

MIT PROJECT NANDA · 2025

**80%**

of AI projects fail to deliver their intended business value — double the failure rate of comparable non-AI projects.

RAND CORPORATION · 2024

**\$10.30**

returned per dollar invested by the organizations that get it right — proving the ceiling is real.

MCKINSEY · 2025

The pattern underneath the 95% is consistent: poor data quality, no clearly defined problem, no named accountability, and processes automated as-is instead of redesigned around where the actual friction lives. Vendors sell tools. Consultancies sell frameworks. Almost nobody starts by mapping how work actually moves through the organization — as opposed to how the org chart says it should move — before recommending where to point AI.

**The gap between the 5% who see real return and everyone else isn't technology. It's clarity.**

### 03 — THE GREYDOG APPROACH

## One platform, three value streams — not sequential tiers.

Greydog's methodology is organized as three connected value streams, not a linear package ladder. A client can enter at any point depending on what they already have in place, and the streams share one continuous Engagement Record so nothing has to be re-explained between phases. All three exist to serve one thing: what the customer actually said mattered, from the first assessment through the last check-in.

| Stream                   | What it does                                                                                                                                                                                                              | Primary Agents             |
|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|
| <b>Discover</b>          | Establishes what the organization is actually trying to achieve, then maps how work actually flows — not how it's supposed to flow — to find exactly where AI creates real, prioritized value.                            | Priya, Lena, Pattern Agent |
| <b>Guide &amp; Build</b> | Two paths to the same outcome: coach the client's team through implementation against agreed goals, or build the well-scoped tool directly — Greydog stays the single point of accountability either way.                 | James, Carlos, Work Agent  |
| <b>Kaizen</b>            | The engagement doesn't have to end. A recurring, anonymous interview loop sampling employees across every level surfaces what leadership structurally can't see, feeding a transparent, cumulative backlog of real fixes. | Amara, Mara                |

Every recommendation Greydog makes traces back to a documented artifact — a strategy document, a value-stream map, a Definition of Success — so nothing is a hallucinated best practice. If a claim can't be traced to something the client actually said or a pattern the system actually observed, it doesn't go in the deliverable.

# Hub-and-spoke orchestration, with one accountable human behind it.

Greydog runs on a hub-and-spoke multi-agent architecture: a single orchestrating agent (Marcus) holds the client's complete context across the entire relationship and routes to specialist agents for each phase of work, rather than a monolithic agent trying to be everything at once. This mirrors the architecture pattern validated by Anthropic's own multi-agent research systems and by enterprise multi-agent deployments at scale.

| Layer                  | Agents                                                                                        |
|------------------------|-----------------------------------------------------------------------------------------------|
| Intake & Orchestration | Claire (Intake) → Marcus (Orchestrator, holds context, escalates to Nikola)                   |
| Discover               | Kenji (Readiness) → Priya (Strategy) → Lena (Value Stream Mapping) → Pattern Agent (internal) |
| Guide & Build          | Carlos (Outcomes/OKRs) → James (Implementation Coach) → Mara (Accountability)                 |
| Kaizen                 | Amara (Employee Interviews) → cumulative backlog → Mara (reports)                             |

*Fig. 1 — Hub-and-spoke orchestration with human governance at every escalation point.*

Behind the client-facing roster sits an internal agent layer with no client persona: Pattern Agent synthesizes findings into a prioritized backlog, Document Agent generates and archives every deliverable, Data Collector connects to the client's own systems (Jira, Azure DevOps, GitHub) to ground recommendations in real operational data, Work Agent automates the administrative layer, and Monitor/Recovery Agents watch every live session for failure and handle real-time interruption recovery without a client-visible restart.

## The self-improvement loop

The Retro Agent runs a structured retrospective after every session, evaluating both the session log and the actual document content the session produced — not metadata alone. That retrospective is re-injected into the same agent's context before its next session with that specific client, so an agent's second conversation is measurably better-informed than its first. This is a live, verified mechanism, not a design aspiration.

## 05 — THE AGENT ROSTER

# Nine client-facing specialists. Seven internal systems. One consistent voice.

---

Every client-facing agent has a fixed role, a consistent voice, and discloses clearly that it is an AI — stated once, plainly, never buried in fine print.

### **Marcus**

#### **Head of Client Success**

Orchestrates every engagement, holds context across sessions, escalates to human judgment when stakes require it.

### **Priya**

#### **Strategy Partner**

Establishes or reviews the client's actual strategy — the anchor every later recommendation checks against.

### **Carlos**

#### **Outcomes Lead**

Drafts the Definition of Success and OKRs from the agreed future state — every initiative tied to a measurable outcome.

### **Claire**

#### **Discovery Lead**

First contact — separates real problems from FOMO, captures the brief that routes the client correctly.

### **Lena**

#### **Systems Lead**

Maps how work actually moves through the organization and identifies exactly where AI eliminates friction.

### **James**

#### **Implementation Coach**

Coaches execution without pushing a fixed methodology — fit-for-purpose guidance for what's actually being built.

### **Kenji**

#### **Insights Lead**

Grounds board-level AI pressure in the client's real business with proven use cases and an honest readiness verdict.

### **Amara**

#### **People Intelligence Lead**

Runs the recurring, anonymous employee interview loop inside Kaizen — psychological safety as a discipline, not a phrase.

### **Mara**

#### **Engagement Lead**

Holds commitments across structured 30/60/90-day check-ins, escalates when a decision needs a human.

## Internal systems layer

| Agent          | Function                                                                                                                                           |
|----------------|----------------------------------------------------------------------------------------------------------------------------------------------------|
| Pattern Agent  | Synthesizes value-stream data and interview findings into a prioritized, evidence-linked AI opportunity backlog.                                   |
| Document Agent | Generates, reviews, and archives every client-facing and internal document — deterministic legal-clause fill, never LLM-drafted contract language. |
| Data Collector | Connects to the client's own Jira, Azure DevOps, or GitHub to ground recommendations in real flow-metric data, not estimation.                     |
| Work Agent     | Automates the administrative layer — tickets, acceptance criteria, session artifacts — behind every client-facing session.                         |
| Retro Agent    | Runs a real retrospective on every session's actual output and re-injects it into that agent's future sessions with that client.                   |
| Monitor Agent  | Watches every live session and agent health signal, triggers Recovery on failure.                                                                  |
| Recovery Agent | Real-time mid-session interruption handling — context-aware resume, never a silent restart the client has to notice.                               |

## 06 — THE ENGAGEMENT WORKFLOW

# Assess. Discover. Build. Retrospect. Anchor.

---

Every engagement follows one consistent journey, gated by real artifacts rather than calendar time — a phase does not open until the artifact the prior phase owed actually exists.

- 1 Kickoff — people, channels, cadence
- 2 Assessment — Marcus scopes fit
- ✓ **Gate: Engagement Agreement signed**
- 3 Discovery — Lena maps current state
- 4 Pattern Agent — candidate AI opportunities
- 5 Discovery Session 2 — future state agreed
- 6 Carlos — Definition of Success and OKRs
- ✓ **Gate: Client accepts, or requests a walkthrough**
- 7 Guide and Build — James coaching / delivery
- 8 Mara — 30/60/90-day check-ins
- 9 Closeout — Now / Next / Later handover
- 10 Anchor / Kaizen — ongoing, optional

*Fig. 2 — The gated engagement workflow. Gate steps are real decision points, not narrative beats.*

Scheduling inside this workflow is a real tool-call against Greydog's live calendar system, not a token the agent hopes translates into a booking. When a client confirms a time, the platform executes the booking synchronously, before any confirmation reaches the client — action always precedes narration. This closes a class of failure common to text-based AI agents, where a confident-sounding promise reaches the client before the system has actually attempted to make it true.

# An 80/20 model: AI does the diagnostic work, a named human owns the judgment.

Greydog is explicit, structurally and in every client conversation, that its agents are AI — never implied to be human, never hidden in a disclosure buried in terms of service. Every engagement is backed by a named human — Nikola Cvetanovski, twenty years of enterprise consulting — who is automatically notified and can intervene directly in any live session the moment stakes require it.

### ESCALATION IS STRUCTURAL, NOT ASPIRATIONAL

Each client-facing agent carries defined "hard hold" triggers — a legal or compliance signal, a material strategy conflict, loss of client confidence, a decision above the agent's authority — that stop the session and route to human review automatically, with a real-time push notification and email, not a queued ticket reviewed days later.

## Data handling

- Client data is transmitted to the underlying AI model provider solely to generate the analysis and documents that make up engagement deliverables, and is never used to train any model without separate written consent.
- Employee interview data is anonymized before inclusion in any synthesis shared with client leadership — the number of employees interviewed and any identifying detail is withheld from client-facing deliverables by default.
- Client data is retained for the duration of the engagement plus twelve months to support the structured 30/60/90-day check-in process, then deleted on written request.
- Where an interview reveals a legal, safety, or compliance concern, it is escalated to the client through the named human directly, without disclosing the source.

### LEGAL-CLAUSE DISCIPLINE

Contract and compliance language is never LLM-generated. Document Agent performs deterministic template fill — engagement-specific fields only — against reviewed legal text, so every signed agreement is textually identical on every clause that matters.

# What actually separates this from a chatbot wrapper.

---

### Real tool-use, not text-token parsing

Committal, machine-actionable actions — scheduling, e-signature, document generation — are implemented as genuine model tool-calls with JSON-schema-enforced required fields, not as bracket-tokens embedded in conversational text and reverse-engineered after the fact. The distinction matters operationally: a tool call cannot be issued with a missing required field, and its real result is fed back to the model before it narrates anything to the client — eliminating an entire failure class of confident-but-false promises.

### Live intervention without breaking the session

The named human overseeing an engagement can inject a real-time correction into any live agent session, or speak to the client directly through the platform, from an authenticated monitoring channel — never through the client-facing chat itself, which means no agent ever has to evaluate whether an identity claim is genuine. Security comes from which channel a message arrives on, not from a keyword a model has to trust.

### Document-grounded self-improvement

Retrospective evaluation reads the actual work product an agent produced — not just a logged summary — and explicitly compares it against the prior session's version for genuine improvement or regression. This closes a real gap common to simpler "lessons learned" systems: a thin session log can miss real quality change that only shows up in the artifact itself.

### Voice input

Session participants can speak instead of type — audio is transcribed and surfaced for review before it's ever sent as a message, keeping a human confirmation step between speech and anything the system acts on.

## The continuous-improvement layer most engagements never build.

---

Most consulting relationships end at delivery. Anchor is Greydog's standing retainer layer, built around a recurring Kaizen loop rather than a generic advisory call:

### INDEPENDENT, ROTATING SAMPLING

Amara interviews three to five employees per cycle from a client-provided roster. Selection is Greydog's determination, not the client's, which is what keeps the "your voice will be heard, no politics" guarantee credible.

### A CUMULATIVE, CLIENT-PRIORITIZED BACKLOG

Separate lifecycle from the one-time discovery backlog. The client's primary contact prioritizes; ownership is assigned per item — client-org-owned or Greydog-coached — at prioritization time.

### A THROUGHPUT COMMITMENT, NOT A CHECK-IN CALL

At least one backlog item delivered per cycle, reported monthly alongside what Amara actually heard.

### REAL DATA GOVERNANCE

A dedicated kickoff step collects the employee roster under explicit retention and deletion terms, with a real, client-facing choice on data retention if the retainer is ever cancelled.

## 10 — ENGAGEMENT MODELS & INVESTMENT

# Fixed fee. No meter running. Accountability built in from day one.

### Consulting Engagement + Build

**\$23,000 - \$40,000+**

Diagnosis, coaching, and a working build in one engagement — no handoff gap between the plan and the tool.

### Consulting Engagement

**\$18,000 - \$35,000**

Assessment through Discovery, implementation coaching, and 30/60/90-day check-ins. One fixed fee.

### Build Track

**\$6,000+**

A single, well-defined build — one tool, one integration, one workflow. Scoped and fixed at Assessment.

### Anchor (Kaizen)

**\$2,500 - \$5,000 / mo**

Continuous improvement after the engagement closes — recurring interviews feeding an ongoing loop of real fixes.

All engagements begin with a focused Assessment session to confirm the right scope before any commitment beyond that. Context: less than one month of a mid-level consultant at a major firm.

## 11 — WHAT SUCCESS LOOKS LIKE

# Every deliverable traces to a decision someone can actually act on.

---

Closeout produces a Now / Next / Later handover, not a deck that sits in a drawer — prioritized by what to act on immediately, what to sequence next, and what to revisit later, translatable directly into the client's own work-item tooling. Accountability continues past delivery: three structured check-ins at 30, 60, and 90 days are included in every package, with a real, recorded progress trail against each Key Result — not a reconstruction from memory when the quarter ends.

**"We're not done when it's delivered. We're done when your customer feels it." — Greydog AI's operating principle. Every recommendation starts here.**

## Twenty years of consulting judgment, encoded — not replaced.

---

Greydog AI was founded on a direct observation from two decades inside large organizations: the reason smart, capable teams can't get things done is almost never the people. It's the system around them, invisible until someone actually maps it. Greydog encodes that diagnostic discipline into a platform that runs it at AI speed — and keeps a real, named human in the room the moment judgment, not process, is what's actually needed.

The customer defines what value means. Every agent, every gate, every escalation exists to serve that — not to be interesting, not to be clever, just accurate to what the client actually needs.

**Old Dogs. New Tricks. Real Results.**

Greydog.AI · nikola@greydog.ai · greydog.ai