ESG Energy & Emissions Optimization Agent

AI Decision-Support System for Portfolio Retrofit Prioritization

Author: Taash Chikosi

Target roles: AI Consultant / AI Product Manager / AI Solutions Architect

Target employers: Top-tier consulting firms and large enterprises adopting AI at scale

Value proposition: Design and deliver end-to-end AI agent systems that solve real business problems, quantify ROI, and integrate human-in-the-loop decision-making

Strengths: Business problem framing, AI workflow design, ML-driven predictions, RAG, executive dashboards

Outcome focus: Measurable efficiency gains, cost reduction, or revenue impact

Table of Contents

1. Executive Summary 2
2. Decision Context & Objectives 2
   Decision context 3
   Business question 3
   Intended users 3
   Success criteria 3
3. How the System Works (High Level) 4
   Inputs (pre-intervention only) 4
   Outputs 4
4. What the Results Show 4
5. Why the Results Are Trustworthy 5
   Data strategy and realism 5
   Synthetic data generation and ML readiness 6
   Feature engineering and unit of analysis 6
   Training, validation, and performance criteria 6
   Governance, confidence, and human oversight 7
6. What Is Live & What Comes Next 7
   What is live today 7
   What the deployment proves 8
   Scope boundaries 8
   What comes next 9

1. Executive Summary

This engagement delivers a **live AI decision-support system** designed to help sustainability and asset leaders **prioritize energy-efficiency investments** across commercial building portfolios. The system estimates **realistic post-intervention energy savings**, translates those estimates into **cost and emissions impact**, and provides **evidence-backed explanations with explicit confidence and governance controls**. It is intentionally designed for **early-stage screening and prioritization of retrofit actions**, enabling faster and more disciplined capital allocation without overstating certainty or replacing expert judgment.

The solution is **deployed, inspectable, and governed**, combining:

- a machine-learning model that estimates realized savings outcomes, and
- a retrieval-augmented reasoning layer that explains results, surfaces assumptions, and supports ESG-aligned decision-making.

2. Decision Context & Objectives

Decision context

Portfolio owners face a recurring challenge: **many plausible retrofit opportunities, limited capital, and increasing ESG scrutiny**. Traditional engineering studies are accurate but slow and costly, while high-level benchmarks are fast but often misleading.

The core decision is not “what is technically optimal,” but:
Which interventions are most worth investigating further right now?

Business question

Which retrofit actions should be prioritized across a commercial building portfolio over the next 12–24 months to reduce energy cost and Scope 2 emissions, while maintaining capital discipline and ESG credibility?

Intended users

- **ESG / Sustainability leaders** — portfolio prioritization and disclosure support
- **Facilities and Energy teams** — feasibility screening and sequencing
- **Finance** — capital allocation and ROI framing
- **Risk and Governance** — defensibility of ESG claims

Success criteria

- Quantified savings estimates suitable for **screening-level decisions**
- Clear translation to **kWh, AUD, and tCO_2e**
- Explicit confidence and limitation signaling
- Transparent, auditable logic that supports governance review

3. How the System Works (High Level)

The system follows a **simple, defensible decision workflow**:

Input → Predict → Translate → Explain → Decide

Inputs (pre-intervention only)

Users provide a deliberately small set of **high-signal, measurable** building characteristics:

- Floor area
- Building age
- HVAC efficiency (relative score)
- Envelope / insulation quality (relative score)
- Occupancy utilization
- Baseline annual energy consumption

These inputs reflect **known structural and operational drivers of achievable savings** and are available **before capital is committed**, enabling portfolio screening without requiring audit-grade data.

Outputs

For each scenario, the system produces:

- Predicted energy-savings potential (%)
- Translated annual impact (kWh, AUD, tCO_2e)
- Confidence signal
- A clear screening recommendation

4. What the Results Show

For a representative commercial building scenario, the system produces:

- A **quantified savings estimate** (e.g., ~20% reduction)
- Translated **cost and emissions impact**
- A **confidence signal** indicating the reliability of the estimate

The outputs are intentionally framed to answer a single, decision-relevant question:
Is this intervention worth investigating further?

The system does **not** recommend final approval, guarantee outcomes, or substitute for engineering studies. Its role is to **focus attention and capital on the most promising opportunities**.

5. Why the Results Are Trustworthy

This system was designed from the outset to be **decision-grade**, not merely predictive. Trustworthiness is established through **data realism, modeling discipline, validation design, and governance controls.**

Data strategy and realism

Because real portfolio data is rarely accessible, the dataset was deliberately engineered to **behave like real-world building portfolios**, not idealized engineering models.
Key design principles included:

- Separation of **static asset characteristics**, **operational time-series behavior**, and **intervention outcomes**, mirroring enterprise data architectures
- Explicit modeling of **seasonality, operational volatility, and noise**
- Systematic divergence between **expected** and **realized** savings
- Inclusion of **under-performing and failed interventions**

This ensures the system learns from **realistic operational uncertainty**, avoiding best-case bias.

Synthetic data generation and ML readiness

Synthetic data was generated using a reproducible, code-driven pipeline with explicit assumptions. Causal relationships were intentionally preserved (e.g., size, usage, weather sensitivity), while randomness and measurement error were injected to reflect reality.
The dataset was designed to be:

- Statistically sufficient for supervised learning
- Representative of portfolio-level decision contexts
- Suitable for feature engineering without post-hoc leakage

Feature engineering and unit of analysis

The unit of analysis is **one intervention applied to one building**. Features are derived exclusively from **pre-intervention behavior**, using a fixed historical window to summarize operational patterns.

This ensures:

- Deployment realism
- No look-ahead bias
- Alignment with how decisions are actually made

Training, validation, and performance criteria

Models were trained and evaluated using a **chronological split** based on intervention timing: earlier interventions for training, later interventions for testing. This mirrors real deployment, where models are always applied to future decisions.

Performance evaluation prioritized:

- **Absolute error** over academic fit metrics
- Stability across scenarios
- Robustness under noisy inputs

The final deployed model (Random Forest) was selected not for maximal theoretical accuracy, but for **consistent, explainable, and stable performance** under real-world conditions.

Governance, confidence, and human oversight

Predicted impact and confidence are treated separately. Confidence bands reflect data coverage and operational risk rather than inflating savings estimates.

Explicit governance controls include:

- Deterministic recommendation logic
- Mandatory human review under low confidence or high implied exposure
- Logged inputs, outputs, and overrides
- Conservative language enforcement

Together, these mechanisms ensure the system **supports judgment rather than replacing it**.

6. What Is Live & What Comes Next

What is live today

- Deployed Stream lit web application
- Version-controlled ML inference artifact
- Persisted vector store for evidence-backed explanations
- Secure secrets management and logging

Users can interact with the system in real time, adjust building scenarios, and observe how predicted impact and confidence respond.

What the deployment proves

- End-to-end operability

- Stable inference
- Transparent explanation
- Practical usability by non-technical stakeholders

Scope boundaries

The system is designed for:

- Portfolio-level screening
- Capital prioritization
- ESG impact framing
- Decision support

It is explicitly **not** designed for:

- Certified ratings
- Guaranteed savings
- Regulatory submissions
- Replacement of engineering audits

What comes next

Future enhancements could include:

- Integration with real portfolio data
- Expanded intervention typologies
- Longitudinal performance monitoring
- Deeper integration with capital planning workflows

In summary, this engagement demonstrates how AI can be applied pragmatically to improve energy-efficiency and ESG investment decisions at a portfolio scale. By combining predictive impact estimation with realistic data design, disciplined validation, explicit governance, and evidence-backed explanation, the solution enables faster, more credible prioritization while standing up to senior executive and risk scrutiny.