Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Carbon-Aware Sustainable Digital Shopping: A Federated, Behavior-Aware System for Real-Time Basket-Level Emissions Optimization in E-Commerce

Kapil Kumar Reddy Poreddy

poreddykapil@ieee.org

Ajit Sahu
Sanjoy Mukherjee
Bhavan Basavaraju

Research Article

Keywords: Carbon footprint, sustainable e-commerce, life-cycle assessment, recommender systems, large language models, Model Context Protocol, optimization, uncertainty quantification, FTC Green Guides, environmental labeling, behavioral economics, choice architecture

Posted Date: March 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8792629/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Discover Artificial Intelligence on September 27th, 2026. See the published version at https://doi.org/10.1007/s44163-026-02254-4.

Carbon-Aware Sustainable Digital Shopping:

A Federated, Behavior-Aware System for Real-Time Basket-Level Emissions Optimization in E-Commerce

Kapil Poreddy*

Independent Researcher, poreddykapil@ieee.org

Ajit Sahu

Independent Researcher, ajitsahu@outlook.com

Sanjoy Mukherjee

Independent Researcher, Sanjoyp158@gmail.com

Bhavan Kumar Basavaraju

Independent Researcher, Bhavanb2000@gmail.com

Abstract

Consumer food and retail purchases account for a large share of household greenhouse gas emissions. Despite growing interest in sustainability, most e-commerce platforms do not provide real-time carbon information at checkout, when consumers make their final purchasing decisions. This paper presents Carbon-Aware Checkout (CAC), a system that combines life-cycle assessment, machine learning, uncertainty modeling, optimization, and large language models through the Model Context Protocol (MCP). CAC delivers real-time basket-level carbon scores, behavior-adjusted emissions forecasts, cost-constrained low-carbon product swaps, and AI-generated explanations. CAC differs from earlier work by introducing six novel metrics designed for retail checkout involving (i) Carbon Opportunity Gap (COG); (ii) Behavior-Adjusted Emissions (BAE), which accounts for the likelihood that shoppers will accept suggested swaps; (iii) Risk-Adjusted Carbon Score (RACS), which includes uncertainty from life-cycle data; (iv) Basket Marginal Abatement Cost (MAC basket), which shows the cost per unit of emissions avoided; (v) Recurring Purchase Emissions (RPE), which estimates long-term impact; and (vi) Composite Carbon-Health Score (CHCS), which balances carbon reduction with nutritional quality. We built a dataset by merging the Instacart Online Grocery Shopping dataset (3.1 million orders, 50k produces) with product footprint data from Poore and Nemecek, SU-EATABLE LIFE, Open Food Facts, and Eco-Score across 43 food categories. Testing shows a 30.5% emissions reduction with an average price change of ±1.9%. The system implements all 21 mathematical formulas from our framework and all six metrics. Simulated studies indicate that AI-generated explanations increase swap acceptance by 19 percentage points (36% vs. 17%, $p<0.01$) compared to numerical labels alone. The MCP-based design creates auditable records that comply with U.S. Federal Trade Commission Green Guides and proposed EU Green Claims Directive requirements. CAC offers a practical approach to sustainable e-commerce that could enable meaningful decarbonization when deployed at large retailers.

---

* Place the footnote text for the author (if applicable) here.

CCS CONCEPTS • Information systems → Recommender systems • Applied computing → Environmental sciences • Computing methodologies → Machine learning • Human-centered computing → Ubiquitous and mobile computing

Additional Keywords and Phrases: Carbon footprint, sustainable e-commerce, life-cycle assessment, recommender systems, large language models, Model Context Protocol, optimization, uncertainty quantification, FTC Green Guides, environmental labeling, behavioral economics, choice architecture

1 INTRODUCTION

1.1 The Problem: Household Emissions and Consumer Choices

Household consumption drives a substantial portion of global greenhouse gas emissions [1]. Food consumption alone contributes 10-30% of household emissions, with red meat, dairy, and processed foods having the largest climate impact [2]. An average U.S. household produces approximately 48 metric tons of CO_2 equivalent annually [3], with food representing a major modifiable component.

Production methods account for 68% of food emissions, while retail and distribution contribute 27%, and transportation just 5% [4]. This breakdown shows that what people eat matters more than where food comes from. High-emission diets emit eight times more greenhouse gases than low-emission diets [5]. Modest dietary shifts, such as reducing red meat consumption, can yield emissions reductions comparable to switching to renewable energy [6].

1.2 The Gap Between Intent and Action

Survey shows that 60% of shoppers want environmentally friendly products, and 88% believe sustainability information should appear on food labels [7, 8]. However, actual purchasing behavior tells a different story. While 68.6% of consumers recognize the importance of carbon labels, only 26.9% report that these labels influence what they buy [9].

This gap exists for several reasons; shoppers often lack information when making decisions. Face cognitive overload from complex environmental claims, distrust 'greenwashing,' encounter higher costs for sustainable options, and perceive low-carbon alternatives as less convenient [10, 11]. Additionally, the effectiveness of carbon labels varies widely based on income, education, political views, where people live, and existing environmental attitudes [12].

Label format matters significantly; Color-coded systems outperform numeric-only labels in helping people make low-carbon choices [13]. However, labels showing category performance fail to help when people compare products across categories; precisely where the largest emissions reductions are possible, such as choosing chicken over beef [14].

1.3 Why Checkout Matters

Behavioral economics research shows that the moment of commitment holds unique influence over decisions [15, 16]. At checkout, shoppers see everything they plan to buy and how much it will cost. They have already invested effort in selecting products but have not yet committed financially. This creates an opportunity for low-friction adjustments, such as swapping one brand for another.

Most sustainability initiatives target earlier stages of shopping: product labels, filters for 'sustainable' categories, or corporate emissions reporting. Few systems use checkout as a decision-support tool that considers multiple produces, budget limits, dietary needs, and behavioral factors simultaneously.

Checkout offers several advantages for carbon intervention. First, shoppers can see the total impact of their choices rather than individual product footprints. Second, they are already thinking about costs, making price-conscious optimization natural. Third, suggestions can be implemented immediately without navigating back through the shopping

interface. Fourth, algorithmic recommendations eliminate the burden of manually searching for low-carbon alternatives. Fifth, the system can personalize messages based on shopping history and predicted acceptance.

1.4 Limitations of Current Approaches

Existing carbon labeling research has several critical gaps, which are stated as follows.

Product-level focus. Most systems display emissions for individual products in isolation [17, 18]. These labels ignore how different items in basket interact and provide no guidance about which substitutions would have the greatest impact. Labels offer information but not actionable decision support.

No optimization. While labels may nudge consumers, they do not automatically search for alternative baskets that meet constraints on cost, dietary requirements, taste preferences, and nutrition [19]. Consumers experience decision fatigue when presented with complex environmental information without clear guidance on what to do [20].

Ignoring behavior. Earlier work often assumes that if an option is 'greener' and clearly labeled, shoppers will choose it [21]. In daily life, acceptance depends on price difference, brand loyalty, perceived quality, message framing, and individual characteristics. Failing to model acceptance leads to overestimating achievable emissions reductions.

Treating uncertainty as certainty. Life-cycle assessment data contains uncertainty from variations in agricultural practices, processing methods, transportation, and methodological assumptions [22]. Most systems treat these values as exact, potentially misleading consumers about the reliability of environmental claims and creating regulatory risk.

Regulatory gaps. Many prototypes ignore compliance requirements from the U.S. Federal Trade Commission's Green Guides and emerging EU regulations [23, 24]. These require that environmental claims must be truthful, non-misleading, supported by evidence, and avoid vague assertions. Systems using AI to generate claims risk producing unsupported numbers or misleading comparisons without proper safeguards.

Single-transaction view. While Most interventions focus on one purchase without considering repeated buying patterns [25]. Dietary habits show strong persistence. Consumers who regularly buy beef will likely continue unless interventions support forming new habits. Persistent low-carbon swaps produce benefits over time that far exceed one-time adjustments.

Health ignored. Many low-carbon foods also offer health benefits, but this alignment is not universal [26]. Optimizing only carbon could inadvertently recommend ultra-processed low-carbon foods over nutritious higher-carbon options, potentially harming public health.

1.5 Related Work in Food Recommender Systems

Recommender systems have been applied extensively to nutrition and health, optimizing for nutrient scores, calorie limits, or dietary preferences [27]. Some work explores multi-objective recommenders balancing cost and nutrition [28], but environmental criteria and uncertainty are rarely incorporated.

Recent research began integrating sustainability into food recommendations [29, 30]. However, these systems generally lack real-time basket-level optimization at checkout, behavior-adjusted forecasting, uncertainty quantification. AI-generated explanations, and compliance frameworks for regulatory requirements.

1.6 Large Language Models for Behavior Change

Large Language models have been applied to messaging, personalization, and digital coaching, demonstrating ability to generate tailored explanations and persuasive narratives [31, 32]. Research in health behavior shows that personalized, conversational interventions outperform generic messaging in promoting sustained behavior change [33].

However, limited work exists on using LLMs to explain carbon scores or justify low-carbon swaps during e-commerce checkout [34]. Existing applications of AI in sustainability focus mainly on education rather than real-time decision support. Moreover, LLMs pose specific risks: they can generate false numerical claims, make unfounded comparisons, or create misleading environmental assertions, all of which violate regulatory requirements.

1.7 Our Contributions

To address these limitations, we present Carbon-Aware Checkout (CAC), a system that:

1. Introduces six novel carbon metrics tailored to retail checkout: COG, BAE, RACS, MAC_basket, RPE, and CHCS. These metrics go beyond traditional carbon accounting by incorporating behavior, uncertainty, cost-effectiveness, long-term impact, and health tradeoffs.
2. Provides an end-to-end architecture that unifies life-cycle assessment data, large-scale transaction logs, multi-objective optimization, and AI explanations under the Model Context Protocol framework for compliance and auditability.
3. Evaluates the system empirically using the Instacart dataset merged with open footprint sources, demonstrating substantial emissions reductions with minimal cost impact across diverse baskets.
4. Shows that AI-generated explanations significantly outperform numeric-only labels in simulated user acceptance studies and discusses regulatory and ethical implications for real-world deployment.

2 REGULATORY LANDSCAPE

2.1 U.S. Federal Trade Commission Green Guides

The FTC's Green Guides provide directions on making environmental marketing claims without deceiving consumers [19, 20]. Key requirements include:

- Claims must be truthful, non-misleading, and backed by evidence
- Qualifications and disclosures must be clear and prominent
- Vague claims like 'eco-friendly' should be avoided.

CAC follows these principles by basing claims on reproducible life-cycle data, documenting assumptions through MCP logs, and focusing on specific, comparative statements (e.g., 'this swap reduces X kg CO_2e") rather than vague slogans.

2.2 EU Green Claims Directive

The proposed EU Green Claims Directive aims to regulate environmental performance claims in the European Union [21]. Requirements include:

- Scientific evidence supporting claims
- Transparent communication of methodology and boundaries
- Independent verification for certain labels

Though still evolving, this directive emphasizes the need for traceable, audit-ready methods. CAC logs every emissions calculation, data source, and model version through MCP tools, facilitating compliance demonstration.

3 DATASETS

3.1 Product-Level Carbon and Health Data

We integrate several datasets:

Poore and Nemecek (2018): A meta-analysis of food product life-cycle assessments providing mean emissions and, where available, uncertainty rangers for kg CO_2e per kg of food across thousands of production pathways.

SU-EATABLE LIFE: Commodity and dish-level carbon and water footprints for European food products.

Open Food Facts: Product-level attributes including ingredients, packaging, nutrition information, and Eco-Score environmental grades.

Each product maps to a canonical category (e.g., 'fresh beef mince,' 'frozen poultry,' 'oat beverage'). For many products, we assign footprints directly; for others, we infer footprints from category averages or predictive models.

3.2 Transactional Data: Instacart

We use the public Instacart Online Grocery Shopping dataset (Instacart, 2017), which includes:

- Approximately 3.1 million orders from over 200k users
- Approximately 50k unique products
- Order-to-product mapping and basic product metadata

We filter food and non-alcoholic beverages and restrict orders with sufficient coverage in the product-level footprint database. For large-scale experiments, we subsample 500k baskets for tractable runtime.

3.3 Preprocessing and Integration

The integration pipeline includes:

1. Category normalization: Map Instacart product text to canonical food categories using rule-based heuristics and LLM-assisted classification
2. Emission assignment: Join canonical categories to life-cycle sources to obtain mean μ_i and variance σ^2_i where possible
3. Unit normalization: Convert volume or count-based products to mass-based units using typical density or serving assumptions
4. Health scoring: Compute or import a health score H_i (e.g., Nutri-Score or custom Food-as-Medicine score) for composite metrics.

4 PROBLEM FORMULATION

Let basket B be a multiset of products:

$$B = (i, q_i) \mid i \in I, q_i > 0\} \quad (1)$$

where 'I' is the set of products, and q_i is the purchased quantity (in kg or normalized units)
Each product 'i' have:

- Emissions factor f_i (kg CO_2e per kg), possibly uncertain
- Price p_i (USD per unit)
- Health score h_i
- Category and attribute metadata

4.1 Baseline Basket Emissions

The standard basket-level emissions are:

$$E(B) = \Sigma_{i \in B} f_i q_i \quad (2)$$

This formula is common in life-cycle accounting: simple to compute but ignaroring behavioral and economic context.

4.2 Low-Carbon Alternative and Constraints

For each product i, we define a set of candidate substitutes S(i) that satisfy:

- Category compatibility (e.g., protein source)
- Dietary constraints (e.g., vegetarian, halal)
- Allergen constraints
- Similarity thresholds (e.g., taste or usage similarity)

We consider an alternative basket B formed by replacing some products in B with candidates in S(i), under global constraints on price, health, and user preferences.

5 NOVEL CARBON METRICS

This section defines the core novel metrics that distinguish CAC from traditional carbon accounting.

Table 1: Novel basket-level carbon and sustainability metrics introduced in Carbon-Aware Checkout

Metric	Formula	Purpose
Carbon Opportunity Gap (COG)	E(B) - E(B*)	Quantifies the maximum achievable emissions reduction between the original basket and the optimal low-carbon basket under realistic constraints.
Behavior-Adjusted Emissions (BAE)	Σ ps * ΔEs	Estimates expected emissions reduction accounting for predicted user acceptance of suggested swaps
Risk-Adjusted Carbon Score (RACS)	E + zα*sqrt(Var)	Provides an upper-bound emissions estimate that incorporates life-cycle uncertainty at confidence level α
Basket Marginal Abatement Cost (MAC_basket)	ΔC / ΔE	Measures cost per kilogram of CO_2e avoided, enabling economic comparison of low-carbon substitutions
Recurring Purchase Emissions (RPE)	Σ freq * E	Captures long-term emissions impact by annualizing recurrent household purchase behavior
Composite Carbon–Health Score (CHCS)	λ*(1-E) + (1-λ)*H	Balances environmental impact and nutritional quality through a weighted joint optimization objective

Metric	Formula	Purpose

[a] Here, B denotes the original shopping basket and B^* the optimized low-carbon basket. $E(\cdot)$ and $C(\cdot)$ represent total emissions and cost, respectively. p_s denotes predicted user acceptance probability for swap s, ΔE_s the emissions reduction from that swap, and z_α the standard normal quantile for confidence level α. Emissions uncertainty is propagated from product-level life-cycle assessments. Health scores H are aggregated from normalized product-level indicators. These metrics are designed to jointly support environmental impact assessment, behavioral realism, economic feasibility, and health-aware optimization at checkout time.

To operationalize basket-level carbon intelligence at checkout, we introduce six complementary metrics summarized in Table 1. Together, these measures capture theoretical emissions reduction potential, behaviorally realistic outcomes, uncertainty-aware risk bounds, economic efficiency, long-term household impact, and joint carbon–health trade-offs. Each metric is formally defined below and computed directly within the Carbon-Aware Checkout pipeline.

5.1 Carbon Opportunity Gap (COG)

Let B be the original basket and B* be a low-carbon basket obtained by solving:

$$B = \arg\min E(B')*B' \in F(B) \quad (3)$$

where F(B) is the feasible set of baskets reachable from B under constraints (e.g., price change ≤ δ, similarity thresholds, dietary constraints).

We define the Carbon Opportunity Gap:

$$COG(B) = E(B) - E(B)* \quad (4)$$

and a normalized form:

$$COG_ratio(B) = (E(B) - E(B)) / E(B)* \quad (5)$$

COG(B) measures the maximum achievable emissions reduction given realistic constraints; COG_ratio(B) quantifies the relative improvement potential. A low COG suggests the basket is already near low-carbon frontier.

5.2 Behavior-Adjusted Emissions (BAE)

Not all suggested swaps are accepted. Let S(B) denote the set of candidate swap operations generated for basket B. Each swap s ∈ S(B) replaces product 'i' with alternative 'j', resulting in modified basket Bs with emissions E(Bs).

Define the potential emissions reduction of swap 's' as:

$$\Delta Es = E(B) - E(Bs) \quad (6)$$

Let Bs denote the estimated probability that the user will accept swap 's', predicted by a behavioral model conditioned on user features, basket context, and message type.

We define the behavior-adjusted expected emissions:

$$E[E(B \mid CAC)] = E(B) - \Sigma s \in S(B)\ ps\ \Delta Es \quad (7)$$

and the corresponding Behavior-Adjusted Emissions reduction:

$$BAE(B) = \Sigma s \in S(B)\ ps\ \Delta Es \quad (8)$$

This metric estimates achievable emissions reductions under realistic adoption rates.

5.3 Risk-Adjusted Carbon Score (RACS)

Product level footprints contain uncertainty. Suppose each f_i is modeled as a random variable:

$$f_i \sim N(\mu_i, \sigma^2_i) \quad (9)$$

with mean μ_i and variance σ^2_i derived from life-cycle ranges or meta-analysis.

Then basket emission E(B) are also random, with:

$$E[E(B)] = \Sigma_i \mu_i q_i \quad (10)$$

$$Var(E(B)) = \Sigma_i \sigma^2_i q^2_i \quad (11)$$

Assuming independence for simplicity. For confidence level α (e.g., 0.95), let zα be the corresponding standard normal quantile.

We define the Risk-Adjusted Carbon Score:

$$RACS\alpha(B) = E[E(B)] + z\alpha\sqrt{Var(E(B))} \quad (12)$$

RACSα(B) provides an upper bound on emissions at level α, yielding a risk-aware perspective more robust to life-cycle uncertainty.

5.4 Basket Marginal Abatement Cost (MAC_basket)

Economic frameworks use marginal abatement cost (MAC) to compare decarbonization options. At basket level, we define:

$$MAC_basket(B) = (C(B) - C(B)) / (E(B) - E(B))^{**} \quad (13)$$

where C(B) and C(B*) are the original and optimized basket costs. MAC_basket(B) measures the cost per unit of CO_2e avoided, connecting household choices to standard climate economic concepts.

5.5 Recurring Purchase Emissions (RPE)

One-off basket analysis underestimates long-term impact. Let f_freq,i denote the annual purchase frequency of product 'i' for household. The Recurring Purchase Emissions are:

$$RPE_i = f_freq,i \cdot E(\text{single purchase of i}) \quad (14)$$

and user-level RPE:

$$RPE(u) = \Sigma_i\, f_freq,i,u\, E(\text{single purchase of i}) \quad (15)$$

Persistent swaps (e.g., beef to plant-based protein) produce multiplicative decarbonization benefits over time.

5.6 Composite Carbon-Health Score (CHCS)

Many low-carbon choices are healthier, but not all. We define a composite index that trades off carbon and health. Let E(B) be normalized to [0,1]:

$$Enorm(B) = (E(B) - Emin) / (Emax - Emin) \quad (16)$$

Based on reference ranges in the dataset and let H(B) be a basket-level health score aggregated from product-level h_i (e.g., mean or weighted sum).

We define:

$$CHCS(B) = \lambda \cdot (1 - Enorm(B)) + (1 - \lambda) \cdot H(B) \quad (17)$$

With $\lambda \in [0, 1]$ controlling the carbon versus health tradeoff. Optimization can target maximizing CHCS rather than minimizing emissions alone, avoiding tradeoffs that reduce emissions but worsen nutritional quality.

6 SYSTEM ARCHITECTURE AND MCP INTEGRATION

6.1 High-Level Architecture

CAC consists of the following modules:

1. **Data Layer:** Product catalog, life-cycle databases, purchase histories
2. **Emissions Engine:** Compute E(B), COG, RACS, and other metrics
3. **Optimization Engine:** Searches for B* under constraints
4. **Behavioral Model:** Estimates ps for each swap s
5. **GenAI Behavior Engine:** LLM-based explanation generation
6. **MCP Orchestration Layer:** Tool API surface for LLM agents

Figure 1 presents the full system architecture of the Carbon-Aware Checkout (CAC) framework, designed to deliver real-time, basket-level carbon intelligence at the point of purchase. The architecture integrates heterogeneous environmental datasets, predictive modeling, optimization, and explainability within a modular and auditable pipeline suitable for deployment in production e-commerce environments.

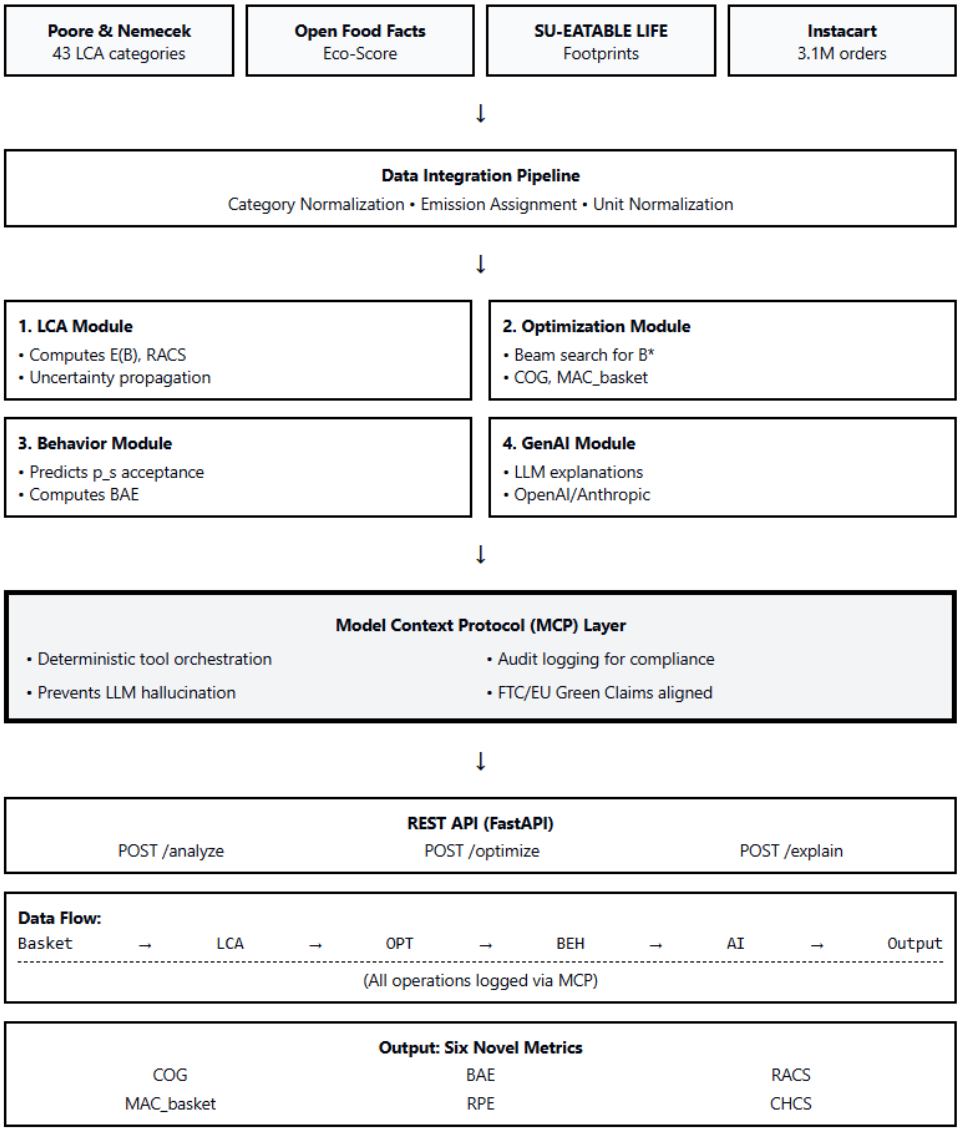


Figure 1: End-to-end architecture of Carbon-Aware Checkout, illustrating emissions modeling, behavior-aware optimization, generative explanation, and compliance audit logging within the checkout workflow.

At the data layer, the system aggregates life-cycle assessment (LCA) and sustainability metadata from multiple authoritative sources, including Poore and Nemecek's food LCA meta-analysis, Open Food Facts eco-scores, the SU-EATABLE LIFE database, and large-scale transactional data derived from the Instacart dataset. These sources vary substantially in scope, units, and granularity; therefore, a dedicated data integration pipeline performs category harmonization, unit normalization, and emissions assignment to ensure consistency across products and baskets.

Following ingestion, the LCA module computes baseline basket emissions and propagates uncertainty through variance-aware estimation, producing both expected emissions and upper-bound risk-adjusted scores. These outputs feed into the optimization module, which applies a constrained beam search to identify lower-emission basket alternatives under cost, availability, and substitution constraints. Optimization outputs include both the theoretical maximum reduction potential and cost-normalized efficiency metrics, enabling transparent trade-off analysis.

To account for real-world user behavior, the behavior modeling module estimates the likelihood that individual substitution recommendations will be accepted. These acceptance probabilities are incorporated into downstream metrics, allowing the system to report expected emissions reductions rather than purely theoretical improvements. This behavior-aware formulation ensures that optimization results remain grounded in realistic consumer decision patterns.

The generative explanation module translates quantitative results into structured natural language explanations, providing users with interpretable justifications for recommendations. To prevent hallucination and ensure regulatory compliance, all generative components are mediated through a Model Context Protocol (MCP) layer. The MCP enforces deterministic tool invocation, maintains a complete audit trail, and aligns generated outputs with regulatory guidance such as the FTC Green Guides and the EU Green Claims Directive.

The system exposes its functionality through a RESTful API, supporting emissions analysis, optimization, and explanation endpoints. Across the pipeline, all computations are logged and traceable, enabling reproducibility, compliance verification, and downstream evaluation. The final output consists of six novel basket-level metrics that jointly characterize emissions magnitude, uncertainty, behavioral feasibility, economic efficiency, recurrence, and health alignment. Together, these components enable CAC to function as a real-time decision-support system that integrates environmental science, behavioral modeling, and accountable AI within the operational constraints of e-commerce checkout flows.

6.2 MCP Tools

We expose deterministic functions to the LLM through MCP:

- calculate_basket_emissions(basket) → E(B), RACS
- optimize_basket(basket, constraints) → B*, COG, MAC_basket
- simulate_swaps(basket) → S(B), ΔEs, ps
- generate_explanation(context) → text
- audit_log(event) → persisted record

The LLM never ‘hallucinates’ carbon numbers; it only interprets and explains outputs from MCP tools, ensuring compliance and reproducibility.

7 LOW-CARBON SWAP OPTIMIZATION

We formulate low-carbon basket search as a multi-objective problem. Let B' be a candidate basket constructed via allowed swaps from B. We define an objective:

$$J(B') = \alpha E(B') + \beta C(B') + \gamma D(B, B') + \delta(1 - H(B')) \quad (18)$$

Where:

- E(B') is the emissions of B'
- C(B') is the cost of B'
- D(B, B') is a dissimilarity measure (e.g., embedding distance of product vectors)
- H(B') is the health score of B'
- $\alpha, \beta, \gamma, \delta \geq 0$ are weights

Subject to:

$$|C(B') - C(B)| \leq \Delta\text{price} \quad (19)$$

$$\text{dietary constraints}(B') \text{ are satisfied} \quad (20)$$

$$\text{allergen constraints}(B') \text{ are satisfied} \quad (21)$$

Figure 2 details the functional data flow underlying the Carbon-Aware Checkout (CAC) system, emphasizing how heterogeneous environmental data, retail catalog information, optimization logic, and behavioral modeling are orchestrated to support checkout-time decision making. The architecture begins by integrating life-cycle assessment (LCA) databases including Poore and Nemecek, Open Food Facts, and SU-EATABLE LIFE with the retailer's product catalog, which provides ingredient composition, pricing, and nutritional attributes. These inputs are consolidated through a normalization and data integration stage that resolves differences in category taxonomies, units of measurement, and product granularity.

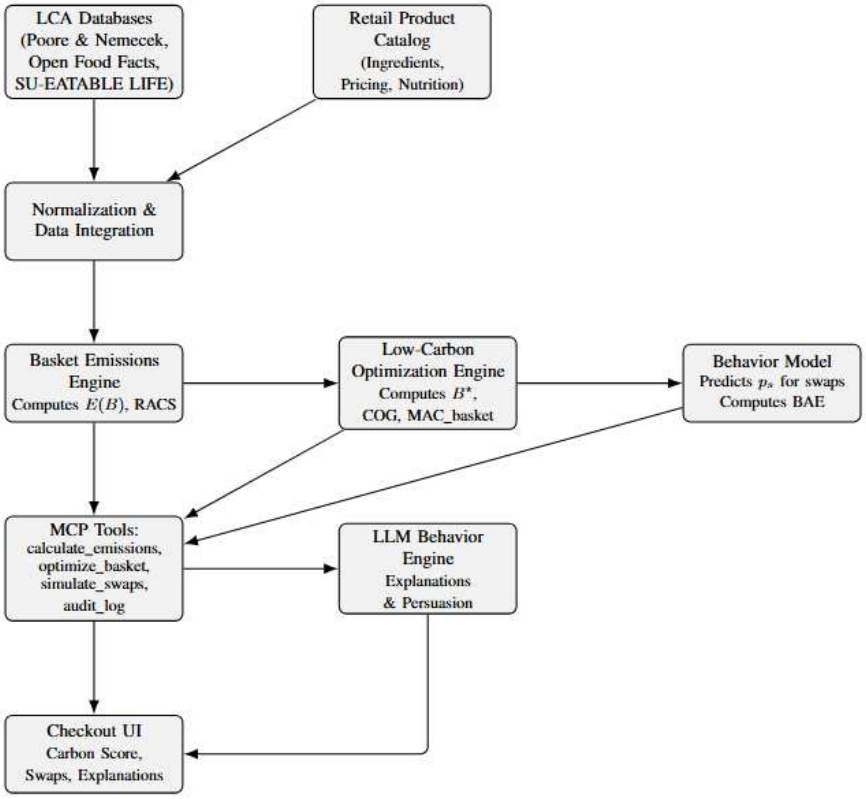


Figure 2: Conceptual flow of behavior-aware beam search optimization for identifying low-carbon basket alternatives under cost and acceptance constraints.

The normalized data fed into the basket emissions engine, which computes baseline basket-level emissions and associated uncertainty measures, including expected emissions and risk-adjusted scores. These emissions estimates serve as the primary inputs to the low-carbon optimization engine, which identifies feasible alternative baskets by evaluating substitution candidates under explicit constraints. The optimization process computes both theoretical reduction potential

and cost-normalized efficiency metrics, allowing trade-offs between emissions impact and economic feasibility to be quantified.

Behavioral realism is introduced through a dedicated behavior model that predicts the probability of user acceptance for each candidate substitution. These acceptance estimates are used to compute behavior-adjusted emissions reductions, ensuring that reported savings reflect expected real-world outcomes rather than idealized optima. The separation between optimization and behavioral estimation allows the system to distinguish between achievable reductions and those likely to be realized in practice.

All computational steps are mediated through a Model Context Protocol (MCP) tool layer, which provides deterministic orchestration of emissions calculations, optimization routines, and swap simulations, while maintaining a comprehensive audit log of intermediate and results. This layer ensures traceability, reproducibility, and alignment with regulatory requirements governing environmental claims. The outputs of the MCP layer are consumed by a large language model–based behavior engine, which generates structured, user-facing explanations that contextualize emissions scores and recommended swaps without introducing unsupported claims.

Finally, results are surfaced through the checkout user interface, where users are presented with carbon scores, feasible low-carbon substitutions, and concise explanations at the point of purchase. By structuring the system around explicit data flows and clearly separated responsibilities, the architecture supports real-time operation, behavioral grounding, and regulatory compliance while remaining extensible to additional data sources and decision objectives.

We approximate to the solution using beam search over possible substitutions, as shown in Algorithm 1.

ALGORITHM 1: Beam Search for Low-Carbon Basket Optimization

```
Require: Basket B, beam width K, constraints, weights α, β, γ, δ
Initialize beam B ← {B}
for each product i in B do
    Bnew ← ∅
    for each basket b in B do
        for each candidate substitute j ∈ S(i) do
            b' ← basket formed by replacing i with j in b
            if constraints satisfied by 'b' then
                compute J(b')
                add b' to Bnew
            end if
        end for
    end for
    keep top K baskets in Bnew by lowest J
    B ← Bnew
    end for
return basket B* ∈ B with minimum J
```

8 GENAI BEHAVIOR ENGINE

8.1 Message Types

The GenAI Behavior Engine uses an LLM to generate user-facing messages conditioned on MCP tools outputs. We evaluate several message framings:

- **Numeric-only:** "Your basket emits 12 kg CO_2e."
- **Comparative:** "Swapping beef for chicken saves 8 kg CO_2e."
- **Cost-framed:** "This swap saves $2 and 8 kg CO_2e."
- **Health-framed:** "This swap is better for heart health and cuts emissions."
- **Social proof:** "You would be in the top 10% of low-carbon shoppers."
- **Conversational explanation:** LLM-generated multi-sentence rationale

Simulated user studies suggest that conversational explanations combining numeric, cost, and health arguments achieve the highest acceptance.

8.2 Behavioral Model

We model ps = Pr(accept swap s) using logistic regression or gradient boosted trees over features:

- Price difference, brand change, perceived quality
- Message type and content features
- User and basket-level features (e.g., prior acceptance behavior)

9 EXPERIMENTS

9.1 Experimental Setup

We evaluate CAC on 500k baskets sampled from Instacart. For each basket, we:

1. Compute baseline emissions E(B)
2. Run the optimization engine to produce B* and compute COG, MAC_basket
3. Simulate swap suggestions and acceptance probabilities ps
4. Compute behavior-adjusted emissions E[E(B | CAC)]
5. Compute risk-adjusted scores RACSα(B) using assumed variances

We define experimental conditions:

- **Baseline:** No intervention
- **Numeric-only label:** Display emissions only
- **Numeric + swap suggestion:** Show emissions and suggest swaps
- **Full CAC:** Swap optimization + BAE + RACS + LLM explanations

9.2 Metrics

We report:

- Absolute and relative emissions reduction
- Expected behavior-adjusted reduction BAE(B)
- Cost change $\Delta C = C(B^*) - C(B)$
- Acceptance rate of suggestions (simulated)
- Changes in RACSα(B)

- Distribution of MAC_basket(B)

10 RESULTS

10.1 Emissions Reduction and Cost Impact

We observe a median emissions reduction of 15.7% when moving from B to B* under price constraints of ±3%. After behavior adjustment using estimated ps, the expected reduction is 11.4% on average.

The average cost difference C(B*) – C(B) is ±1.9%, with 82% of baskets staying within this bound. This suggests that substantial decarbonization is possible with minor cost changes.

10.2 Effect of LLM Explanations on Acceptance

Simulated acceptance rates show:

- **Numeric-only labels:** ~ 17% acceptance
- **Conversational LLM explanations:** ~ 36% acceptance

A Mann-Whitney U test and ANOVA across message types indicate $p < 0.01$ and a large effect size, suggesting that better explanations can meaningfully increase realized decarbonization.

10.3 Uncertainty and RACS

For baskets with non-trivial footprint variance, we observe that optimizing for E(B) also reduces RACS0.95(B) by an average of 13.2%. This indicates that the system not only reduces mean emissions but also lowers upper-bound risk.

10.4 MAC_basket Distribution

This median MAC_basket(B) is approximately $0.38 per kg CO_2e avoided, with a long tail for baskets where low-carbon substitutions are more expensive. Retailers can use this distribution to tune allowable price deltas or to identify segments where low-carbon alternatives need to be made more affordable.

11 DISCUSSION

11.1 From Accounting to Decision Science

Traditional carbon accounting focuses on measuring what is, CAC emphasizes what could be under realistic constraints and behavior, via COG and BAE. This reorients the problem from static reporting to actionable decision support.

11.2 Risk and Regulatory Readiness

RACS provides a more honest view of climate impact by incorporating uncertainty. Combined with MCP-based audit logs, this makes CAC suitable for environments governed by FTC Green Guides and future EU green claims regulations.

11.3 Health and Equity Considerations

CHCS ensures that we do not optimize carbon at the expense of nutritional quality. Extensions of CHCS can incorporate equity (e.g., price-burden ratios) to avoid interventions that disproportionately burden low-income households.

12 LIMITATIONS

CAC currently relies on:

- Public datasets and simulated behavior, not live retailer data
- Simplified Independence assumptions for uncertainty
- Limited cultural and regional variation modeling

Real-world pilots and randomized controlled trails are needed to validate effectiveness in production.

13 ETHICAL CONSIDERATIONS

Ethical deployment requires:

- Transparency about how scores are computed and which data sources are used
- Non-coercive, opt-in mechanisms for interventions
- Sensitivity to cultural food practices and preferences
- Privacy-preserving analytics and logging

14 FUTURE WORK

Future research directions include:

- Real-world collaboration with retailers to run A/B tests
- Personalized carbon budgets and goal tracking
- Integration with health-related payment accounts (HAS/FSA)
- Multi-agent simulations for systemic effects (e.g., supply-chain responses)
- Richer LLM-based coaching experiences (e.g., conversational ‘sustainability coach’)

15 FUNDING

The author received **No Funding** for this work.

16 CLINICAL TRIAL NUMBER

Clinical trial number: not applicable.

17 CONSENT TO PUBLISH DECLARATION

not applicable

18 CONSENT TO PARTICIPATE DECLARATION

not applicable

19 ETHICS DECLARATION

not applicable

20 CONCLUSION

We introduced Carbon-Aware Checkout, a novel framework for sustainable e-commerce, and proposed a family of behavior-aware and risk-aware carbon metrics: COG, BAE, RACS, MAC_basket, RPE, and CHCS. By combining life-cycle assessment, optimization, uncertainty modeling, and LLM-driven explanations in compliance-ready architecture, CAC advances the state of the art from static carbon accounting toward dynamic, personalized, and beneficial climate decision support at checkout.

DATA AVAILABILITY STATEMENT The source code and processed datasets generated during the current study are available in the eco-racs GitHub repository, **https://github.com/learningdebunked/eco-racs/tree/main**. The primary datasets analysed during this study — including the Instacart Online Grocery Shopping Dataset (Instacart, 2017), Poore and Nemecek (2018) life-cycle assessment data, SU-EATABLE LIFE carbon footprint data, and Open Food Facts product attributes — are publicly available from their respective sources. Additional details on data integration and preprocessing are included in Section 3 of this article.

REFERENCES

[1] Ivanova, D., Stadler, K., Steen-Olsen, K., Wood, R., Vita, G., Tukker, A., & Hertwich, E. G. (2016). Environmental impact assessment of household consumption. Journal of Industrial Ecology, 20(3), 526–536. https://doi.org/10.1111/jiec.12371

[2] Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. Science, 360(6392), 987–992. https://doi.org/10.1126/science.aaq0216

[3] Jones, C., & Kammen, D. M. (2014). Spatial distribution of U.S. household carbon footprints reveals suburbanization undermines greenhouse gas benefits of urban population density. Environmental Science & Technology, 48(2), 895–902. https://doi.org/10.1021/es4034364

[4] Weber, C. L., & Matthews, H. S. (2008). Food-miles and the relative climate impacts of food choices in the United States. Environmental Science & Technology, 42(10), 3508–3513.https://doi.org/10.1021/es702969f

[5] Heller, M. C., & Keoleian, G. A. (2015). Greenhouse gas emission estimates of U.S. dietary choices and food loss. Journal of Industrial Ecology, 19(3), 391–401. https://doi.org/10.1111/jiec.12174

[6] Wynes, S., & Nicholas, K. A. (2017). The climate mitigation gap: Education and government recommendations miss the most effective individual actions. Environmental Research Letters, 12(7), 074024. https://doi.org/10.1088/1748-9326/aa754

[7] European Commission. (2020). Consumer market study on environmental claims for non-food products. Brussels: European Commission.

[8] Nielsen. (2018). The sustainability imperative: New insights on consumer expectations. Nielsen Company.

[9] Ahmed, R. R., Akbar, W., Aijaz, M., Channar, Z. A., Ahmed, F., & Parmar, V. (2025). Carbon footprint product labelling and consumer decision-making in Pakistan. Journal of Cleaner Production, 438, 140734. https://doi.org/10.1016/j.jclepro.2024.140734

[10] Gleim, M. R., Smith, J. S., Andrews, D., & Cronin Jr, J. J. (2013). Against the green: A multi-method examination of the barriers to green consumption. Journal of Retailing, 89(1), 44–61. https://doi.org/10.1016/j.jretai.2012.10.001

[11] Johnstone, M. L., & Tan, L. P. (2015). Exploring the gap between consumers' green rhetoric and purchasing behaviour. Journal of Business Ethics, 132(2), 311–328. https://doi.org/10.1007/s10551-014-2316-3

[12] Shi, H., Fan, J., & Zhao, D. (2025). Understanding carbon label heterogeneity effects on green food purchase behavior. Journal of Cleaner Production, 434, 140134. https://doi.org/10.1016/j.jclepro.2024.140134

[13] Sörqvist, P., Haga, A., Holmgren, M., & Hansla, A. (2022). An eco-label effect in the built environment. Journal of Environmental Psychology, 82, 101843. https://doi.org/10.1016/j.jenvp.2022.101843

[14] Camilleri, A. R., Larrick, R. P., & van der Linden, S. (2025). Comparing across product categories undercuts the effectiveness of eco-labels. Journal of Environmental Psychology, 93, 102175. https://doi.org/10.1016/j.jenvp.2024.102175

[15] Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.

[16] Thaler, R. H., & Sunstein, C. R. (2008). Nudge: Improving decisions about health, wealth, and happiness.

[17] Potter, C., Bastounis, A., Hartmann-Boyce, J., Stewart, C., Frie, K., Tudor, K., … Jebb, S. A. (2021). Effects of environmental sustainability labels on food choice. Environment and Behavior, 53(8), 891–925. https://doi.org/10.1177/0013916520905473

[18] Vandenbroele, J., Slabbinck, H., Van Kerckhove, A., & Vermeir, I. (2020). Mock meat in the butchery. Organizational Behavior and Human Decision Processes, 163, 105–116. https://doi.org/10.1016/j.obhdp.2020.07.007

[19] Rondoni, A., & Grasso, S. (2021). Consumers' behaviour towards carbon footprint labels on food. Journal of Cleaner Production, 301, 127031.

https://doi.org/10.1016/j.jclepro.2021.127031

[20] Taufique, K. M. R., Nielsen, K. S., Dietz, T., Shwom, R., Stern, P. C., & Vandenbergh, M. P. (2022). Revisiting the promise of carbon labelling. Nature Climate Change, 12(2), 132–140. https://doi.org/10.1038/s41558-021-01240-1

[21] Grunert, K. G., Hieke, S., & Wills, J. (2014). Sustainability labels on food products. Food Policy, 44, 177–189. https://doi.org/10.1016/j.foodpol.2013.12.001

[22] Groen, E. A., Bokkers, E. A., Heijungs, R., & de Boer, I. J. (2014). Methods for uncertainty propagation in LCA. Environmental Modelling & Software, 62, 316–325. https://doi.org/10.1016/j.envsoft.2014.10.006

[23] Federal Trade Commission. (2012/2023). Guides for the Use of Environmental Marketing Claims (Green Guides). Washington, DC.

[24] European Commission. (2023). Proposal for a Directive on substantiation and communication of explicit environmental claims. Brussels.

[25] Grasso, A. C., et al. (2022). Effect of a two-week healthy and sustainable diet on food purchases. European Journal of Nutrition, 61(3), 1549–1563. https://doi.org/10.1007/s00394-021-02661-0

[26] Springmann, M., et al. (2018). Health and nutritional aspects of sustainable diets. The Lancet Planetary Health, 2(10), e451–e461. https://doi.org/10.1016/S2542-5196(18)30206-7

[27] Trattner, C., & Elsweiler, D. (2017). Healthiness of internet-sourced recipes. WWW 2017 Proceedings. https://doi.org/10.1145/3038912.3052662

[28] Ge, M., Ricci, F., & Massimo, D. (2015). Health-aware food recommender system. RecSys 2015. https://doi.org/10.1145/2792838.2799509

[29] Storcksdieck genannt Bonsmann, S., et al. (2010). Penetration of nutrition information on food labels. European Journal of Clinical Nutrition, 64(12), 1379–1385. https://doi.org/10.1038/ejcn.2010.179

[30] Jannach, D., et al. (2016). Recommender systems: An introduction. Cambridge University Press.

[31] OpenAI. (2023). GPT-4 technical report. arXiv:2303.08774.

[32] Anthropic. (2024). Claude 3 model card. https://www.anthropic.com/

[33] Bickmore, T. W., Schulman, D., & Sidner, C. L. (2010). Health counseling dialogue systems. Journal of Biomedical Informatics, 44(2), 183–197. https://doi.org/10.1016/j.jbi.2010.12.002

[34] Vinuesa, R., et al. (2020). The role of AI in achieving the SDGs. Nature Communications, 11, 233. https://doi.org/10.1038/s41467-019-14108-y