PhytoIntelligence Open-Source Documentation (v1.1)

Marie Seshat Landry

March 2025

Contents

1 Overview

1.1 Mission Statement

PhytoIntelligence aims to democratize the development of diagnostic-specific nutraceuticals by leveraging an AI-driven, open-source framework. It is designed to be modular, data-driven, and aligned with clinical, pharmacokinetic, and regulatory standards for safe and effective phytochemical formulations.

1.2 What is PhytoIntelligence?

PhytoIntelligence is a scientific methodology that employs artificial intelligence, natural language processing, molecular docking, pharmacokinetic simulation, and synergy analysis to formulate evidence-based nutraceuticals. It is diagnostic-agnostic, scalable, and grounded in peer-reviewed literature and global regulatory standards.

1.3 Key Benefits

- Universally applicable to any diagnostic condition.
- Integrates AI-driven literature mining with computational biology.
- Includes synergy scoring, bioavailability enhancement, and regulatory filters.
- Released under Creative Commons for global research collaboration.

2 Core Framework Architecture

2.1 High-Level Workflow

1. Identify diagnostic challenges and limitations of current nutraceuticals.
2. Perform AI-assisted literature mining to extract bioactive compound data.
3. Generate hypothesis and design multi-compound formulation models.
4. Use computational models to evaluate efficacy, safety, and regulatory alignment.
5. Optimize formulation for multi-target activity, bioavailability, and synergy.
6. Generate reproducible reports for experimental validation.

2.2 Mathematical Framework

The core composite score for formulation efficacy:

$$C_x = \sum_{i=1}^{n} (M_i \cdot V_i \cdot P_i \cdot B_i \cdot S_i \cdot R_i \cdot D_i)$$

Each coefficient reflects a domain of scientific or regulatory validation:

- M_i: Molecule Identification Score from literature and bioinformatics databases.
- V_i: Clinical Validation Score from in vitro, in vivo, and clinical trials.
- P_i: Pharmacokinetic Model score (ADME).
- B_i: Bioavailability enhancement coefficient.
- S_i: Synergy analysis score across biological pathways.
- R_i: Regulatory compliance factor from FDA, EFSA, WHO, USDA.
- D_i: Dosage safety score based on NOAEL and toxicology.

2.3 Diagnostic Agnosticism

The framework allows for flexible adaptation to any diagnostic through input of diagnostic-specific protein targets, literature corpora, and pharmacodynamic objectives. It supports metabolic syndromes, cancer, neurological, cardiovascular, autoimmune, and infectious conditions.

3 Modules and Methods

3.1 AI-Driven Literature Mining

The framework uses NLP and machine learning to identify and rank candidate phytochemicals. The algorithm:

- Retrieves data from PubMed, ClinicalTrials.gov, patents, and chemical libraries.
- Scores documents by quality (L_s) and efficacy (E_s).
- Computes $M_i = \sum_{s=1}^{m} (L_s \cdot E_s)$.

3.2 Molecular Docking & Simulation

Utilizes software such as AutoDock Vina to:

- Identify protein targets based on diagnostic pathology.
- Simulate ligand-protein binding and compute affinities.
- Integrate docking scores into efficacy and synergy modules.

3.3 Clinical Validation Metrics

Scores compounds based on evidence:

$$V_i = (C_{in-vitro} + C_{in-vivo} + C_{clinical}) \cdot W$$

with W giving higher weight to human clinical studies.

3.4 Pharmacokinetics & Bioavailability Optimization

$$P_i \cdot B_i = (A_i \cdot D_i' \cdot M_i' \cdot E_i) \cdot B_i$$

where:

- A_i: Absorption efficiency.
- D_i': Distribution including blood-brain barrier permeability.
- M_i': Metabolic stability.
- E_i: Excretion rate.
- B_i: Enhancement from strategies like piperine or nanoformulation.

3.5 Synergy Analysis Engine

$$S_i = \frac{\sum_{j=1}^{n}(M_i \cdot M_j)}{T}$$

where T is the number of distinct biological pathways targeted. Predicts if a compound pair has cumulative or antagonistic effects.

3.6 Regulatory Compliance & Dosage Safety

$$R_i \cdot D_i = (R_{\text{FDA}} \cdot R_{\text{EFSA}} \cdot R_{\text{WHO}} \cdot R_{\text{Organic}}) \cdot S_{\text{NOAEL}}$$

Incorporates:

- International safety data and toxicological thresholds.
- Organic and clean label compliance scoring.

4 Case Study: LC-Phyto (Lung Cancer)

4.1 Formulation

Compound	Daily Dose (mg)	Mechanism

Curcumin	500	Apoptosis, anti-inflammatory
EGCG	300	Antioxidant, angiogenesis inhibition
Resveratrol	250	DNA repair, apoptosis
Berberine	200	Cell cycle arrest, AMPK activation
Sulforaphane	100	NRF2 activation, detoxification
Quercetin	200	Antiproliferative, anti-migration
Apigenin	100	Antioxidant, anti-metastatic
Lycopene	30	Antioxidant, DNA protection
Piperine	10	Bioavailability enhancer
Beta-glucans	300	Immune system modulation

4.2 Predicted Effects

- Multi-pathway targeting for tumor suppression.
- Immune modulation and inflammation control.
- Bioavailability-enhanced delivery for systemic effect.

5 User Guide and Community Access

5.1 Portal Access

PhytoIntelligence Portal

5.2 Usage

1. Select diagnostic and upload target literature.
2. Review auto-suggested compounds and adjust parameters.
3. Generate formulation and export detailed report.

6 Licensing and Collaboration

6.1 Creative Commons Attribution 4.0 (CC BY 4.0)

Freely share and adapt with attribution. `https://creativecommons.org/licenses/by/4.0/`

6.2 Citing this Framework

Landry, M. (2025). *PhytoIntelligence: An AI-Driven Framework for Diagnostic-Specific Nutraceutical Design.*

7 Appendices

Glossary

- **Mi** - Molecule Identification
- **Vi** - Clinical Validation
- **Pi** - Pharmacokinetics
- **Bi** - Bioavailability
- **Si** - Synergy
- **Ri** - Regulatory Compliance
- **Di** - Dosage Safety
- **NOAEL** - No Observed Adverse Effect Level