# Harmonia V3: Master Technical Specification & Implementation Guide

Version: 5.3 (The "Forensic Detail" Edition)

Status: Live / Pilot Ready

Infrastructure: Railway/Northflank (Persistent Volume Storage)

Compliance: GDPR (Right to Erasure), Anonymized Research Protocols

## 1. Core Data Architecture (The "Brain")

The application state is persisted in a sharded directory structure within /app/data. This structure is the "Single Source of Truth."

### 1.1 The Directory Tree

/app/data/  
├── registry.db # Main Relational DB (Users, Relationships, Auth)  
├── global\_calibration/ # Stock images (SCUT-FBP5500)  
│ ├── fbp\_001.jpg  
│ └── ...  
├── profiles/ # USER DATA (INPUTS)  
│ └── {uuid}/ # Unique 32-char UUID for each user  
│ ├── assets/  
│ │ └── profile\_primary.jpg # User's standardized face photo  
│ ├── p1\_visual\_vector.json # Phase 1: Meta-FBP Embedding + Traits  
│ ├── p2\_psych\_piip.json # Phase 2: "Felix" Psychometrics  
│ └── p3\_hla\_secure.enc # Phase 3: Fernet Encrypted HLA Data  
└── reports/ # MATCH EVIDENCE (OUTPUTS)  
 └── {match\_id}/ # Unique UUID per Match Pair  
 ├── reasoning\_chain.json # LEVEL 3: STRICT MATH (The "Felix" & HIBAG Output)  
 ├── gemini\_narrative.md # LEVEL 2A: PSYCH AI REPORT (Admin/Researcher)  
 ├── hla\_gemini\_analysis.md # LEVEL 2B: BIO AI REPORT (Admin/Researcher)  
 └── customer\_summary.json # LEVEL 1: SANITIZED UI DATA (Customer)

## 2. Input Schema Specifications

### Phase 1: Visual (p1\_visual\_vector.json)

**Generation:** VisualService (DeepFace) + Gemini Vision (Traits).

{  
 "meta": { "user\_id": "uuid\_1", "gender": "male", "preference\_target": "female" },  
 "self\_analysis": {  
 "embedding\_vector": [0.123, -0.45, 0.88, ...], // 128-d Vector (Strict Geometry)  
 "detected\_traits": {  
 "facial\_landmarks": ["high\_cheekbones", "strong\_jawline"],  
 "style\_presentation": ["glasses", "minimalist", "executive"],  
 "vibe\_tags": ["intellectual", "serious"]  
 }  
 },  
 "preference\_model": {  
 "ideal\_vector": [0.05, -0.22, 0.91, ...], // Centroid of Likes  
 "attraction\_triggers": {  
 "mandatory\_traits": ["expressive\_eyes"], // Present in >80% of Likes  
 "negative\_traits": ["heavy\_makeup"] // Present in >80% of Dislikes  
 },  
 "calibration\_confidence": 0.85  
 }  
}

### Phase 2: Psychometric (p2\_psych\_piip.json)

**Generation:** "Felix" Protocol via Gemini API (Strict Trait Discovery).

{  
 "model": "PIIP\_Felix\_Dynamic\_v2",  
 "dimensions": {  
 "lust": 0.45, "gluttony": 0.20, "greed": 0.85,   
 "sloth": 0.10, "wrath": 0.60, "envy": 0.70, "pride": 0.95  
 },  
 "evidence\_map": [  
 {  
 "question\_id": "felix\_q3\_spending",  
 "user\_answer\_raw": "I'd buy a vintage Rolex. Shows people I made it.",  
 "extracted\_trait": "status\_signaling",  
 "mapped\_primary\_sin": "pride",  
 "weights": { "pride": 0.90, "greed": 0.60 },  
 "style\_analysis": {  
 "bluntness": 0.7, // 0.0-1.0  
 "expressiveness": 0.4, // 0.0-1.0  
 "complexity": 0.6 // 0.0-1.0  
 },  
 "reasoning": "Subject prioritizes social validation through consumption."  
 }  
 ]  
}

### Phase 3: Biological (p3\_hla\_secure.enc)

Generation: HIBAG Pipeline (R-Script Integration).

Content: Encrypted JSON Blob.

Decrypted Structure (In-Memory Only):

{  
 "source": "23andMe\_v5",  
 "imputation\_confidence": 0.94, // HIBAG Posterior Probability  
 "markers": {  
 "HLA-A": ["A\*01:01", "A\*03:01"],  
 "HLA-B": ["B\*07:02", "B\*08:01"],  
 "HLA-DRB1": ["DRB1\*15:01", "DRB1\*03:01"]  
 }  
}

## 3. The "Engine" Logic (Strict Calculations)

These formulas run inside SimilarityService and populate reasoning\_chain.json. **No AI is involved here.**

### 3.1 Visual Score ($S\_{vis}$)

$$ S\_{vis} = (0.7 \times (1 - D\_{euc})) + (0.3 \times T\_{match}) $$

* $D\_{euc}$**:** Euclidean Distance (User A Ideal vs User B Face).
* $T\_{match}$**:** (Count of B's traits that match A's attraction\_triggers) / Total Triggers.

### 3.2 Psych Score ($S\_{psych}$)

$$ S\_{psych} = (1 - D\_{sins}) \times P\_{friction} $$

* $D\_{sins}$**:** Euclidean Distance between Sin Vectors.
* $P\_{friction}$**:** If Sloth Delta > 0.5 OR Bluntness Delta > 0.6, $P\_{friction} = 0.8$. Else 1.0.

### 3.3 Bio Score ($S\_{bio}$)

$$ S\_{bio} = \frac{N\_{unique}}{N\_{total}} \times 100 $$

* $N\_{unique}$**:** Count of unique alleles in the combined pool.
* **Goal:** Maximize Heterozygosity (Dissimilarity).

### 3.4 Willingness to Meet (WtM)

$$ WtM = (0.4 \times S\_{vis}) + (0.3 \times S\_{psych}) + (0.3 \times S\_{bio}) $$

## 4. The Reporting Architecture

This section defines the four mandatory output files for every match.

### Level 3: The "Felix" Strict Report (reasoning\_chain.json)

Audience: Admin Only.

Content: The raw output of the formulas above. NO AI HALLUCINATIONS.

Sources: p1 (Visual), p2 (Psych), p3 (HIBAG Output).

### Level 2A: The Psych Narrative (gemini\_narrative.md)

Audience: Admin / Researcher (Anonymized).

Generator: Gemini API (Text Generation).

Input: reasoning\_chain.json (Phase 1 & 2 sections ONLY).

Purpose: Explain the "Why" behind the Psych/Visual scores with Forensic Detail (Weights/Confidence).

### Level 2B: The Bio Narrative (hla\_gemini\_analysis.md)

Audience: Admin / Researcher (Anonymized).

Generator: Gemini API (Text Generation).

Input: reasoning\_chain.json (Phase 3 section ONLY).

Purpose: Scientific analysis of the immune system match with Imputation Confidence metrics.

### Level 1: Customer Summary (customer\_summary.json)

Audience: The User (Frontend).

Generator: Gemini API (Sanitization).

Input: gemini\_narrative.md + hla\_gemini\_analysis.md.

Protocol: Sanitized, polite summary.

## 5. Admin Protocols (The "Red Pill" Instructions)

These are the **EXACT System Prompts** used to generate the Admin/Researcher reports.

### 5.1 Protocol A: Psych/Visual Report

Target File: gemini\_narrative.md

Input: reasoning\_chain.json (Phase 1 & 2)

SYSTEM PROMPT:

"You are the Harmonia Engine, a cynical evolutionary psychologist.

Input: Mathematical compatibility data including specific traits, weights, and confidence scores.

Task: Predict relationship survival and sexual viability with FORENSIC PRECISION.

Tone: Clinical, Ruthless, Darwinian.

Mandatory Output Format (Markdown):

1. **Visual Fitness Table:** List every matched trait, its specific weight in the User A's preference model, and the confidence level of detection.
2. **Psychometric Audit:** For every 'Sin' Delta > 0.3, cite the exact 'Evidence Fragment' (User Quote), the 'Extracted Trait' (e.g., Status Signaling), and the 'Confidence Score' of that extraction.
3. **Friction Analysis:** Calculate the specific probability of failure based on the Bluntness/Complexity gap.
4. **Verdict:** 'Viable' or 'Dead on Arrival'."

### 5.2 Protocol B: HLA/Bio Report

Target File: hla\_gemini\_analysis.md

Input: reasoning\_chain.json (Bio Section: Allele Lists & HIBAG Confidence)

SYSTEM PROMPT:

"You are an expert Geneticist specializing in the Major Histocompatibility Complex (MHC).

Input: HLA Alleles for Subject A and Subject B, including Imputation Posterior Probabilities.

Task:

1. **Confidence Audit:** Report the HIBAG imputation confidence for every allele. Flag any below 0.8 as 'Low Confidence'.
2. **Peptide Binding Analysis:** For every unshared allele, explain the specific peptide-binding groove characteristics (e.g., 'B\*27 binds viral epitopes').
3. Olfactory Prediction: Quantify the predicted pheromonal attraction intensity (0-100) based on the Heterozygosity Index.  
   Constraint: Use strict scientific terminology. No romance."

## 6. Appendix: The HIBAG Pipeline (R-Integration)

This is the exact sequence executed by dna\_service.py to populate Phase 3.

1. **Extraction:** Python parses the raw 23andMe.txt or Ancestry.csv to find relevant SNPs (rs-ids).
2. **Imputation (R):** Python uses rpy2 to call the R script:  
   library(HIBAG)  
   # Load Pre-trained Model (European/African/Asian based on ancestry flag)  
   model <- hlaModelFromObj(model\_obj)  
   # Predict  
   pred <- predict(model, snp\_data)
3. **Standardization:** The 4-field output (e.g., A\*01:01:01:01) is truncated to 2-field (A\*01:01) for matching.
4. **Encryption:** The standardized JSON is encrypted with Fernet and saved to p3\_hla\_secure.enc.
5. **Reporting:** The raw alleles are passed to the reasoning\_chain.json generator.

## 7. Comprehensive Report Examples (The "End State")

This section displays the **Actual File Contents** generated for a hypothetical match between "Subject A" (The Executive) and "Subject B" (The Artist).

### 7.1 FILE: reasoning\_chain.json (The Strict Math)

*Note: This file contains NO AI generation. It is purely the result of Sections 3 and 6.*

{  
 "match\_id": "match\_7f8c9a0b",  
 "timestamp": "2026-01-14T18:00:00Z",  
 "subjects": {  
 "user\_a": "Subject A (550e...)",  
 "user\_b": "Subject B (a1b2...)"  
 },  
   
 "phase\_1\_visual": {  
 "raw\_distance\_geo": 0.15,  
 "geo\_score\_contribution": 59.5,   
 "trait\_matches": {  
 "matched": ["glasses", "intellectual"],  
 "missed": ["tattoos"],  
 "weights": {"glasses": 0.8, "intellectual": 0.9}, // User A strongly prefers these  
 "confidence": {"glasses": 0.99, "intellectual": 0.85} // Vision API confidence  
 },  
 "total\_p1\_score": 79.3  
 },  
  
 "phase\_2\_psych": {  
 "sin\_distance": 0.35,   
 "raw\_score": 65.0,  
 "friction\_flags": { "sloth\_delta": 0.60, "bluntness\_delta": 0.80 },  
 "evidence\_chain": [  
 {  
 "trait": "Status Signaling",  
 "source": "User A: 'Rolex'",  
 "weight": 0.9,  
 "confidence": 0.98  
 }  
 ],  
 "total\_p2\_score": 52.0   
 },  
  
 "phase\_3\_bio": {  
 "alleles\_user\_a": ["A\*01:01", "A\*02:01"],  
 "alleles\_user\_b": ["A\*03:01", "A\*24:02"],  
 "imputation\_confidence": 0.94, // From HIBAG  
 "total\_p3\_score": 91.6   
 },  
  
 "final\_wtm": {  
 "calculation": "(79.3 \* 0.4) + (52.0 \* 0.3) + (91.6 \* 0.3)",  
 "score": 74.8  
 }  
}

### 7.2 FILE: gemini\_narrative.md (The "Red Pill" Psych Report)

*Generated via Protocol A using reasoning\_chain.json inputs.*

# Harmonia Forensic Audit: Visual & Psychometric Analysis  
\*\*Subject A\*\* vs. \*\*Subject B\*\*  
\*\*Date:\*\* 2026-01-14  
  
## 1. Visual Trait Verification Table  
\*Analysis of User A's Attraction Triggers vs. User B's Phenotype\*  
  
| Trait Detected | A's Preference Weight (0-1) | Detection Confidence | Impact Status |  
| :--- | :--- | :--- | :--- |  
| \*\*Glasses\*\* | 0.80 (High) | 99.2% | \*\*POSITIVE MATCH\*\* |  
| \*\*Intellectual Vibe\*\* | 0.90 (Very High) | 85.0% | \*\*POSITIVE MATCH\*\* |  
| \*\*Tattoos\*\* | 0.10 (Low) | 92.0% | MISS (Low Impact) |  
  
\*\*Forensic Summary:\*\* Subject B satisfies 85% of Subject A's weighted visual criteria. The detection confidence for the primary trigger ("Glasses") is near absolute.  
  
## 2. Psychometric Conflict Audit  
\*Deep Dive on Incompatibility Drivers (Delta > 0.3)\*  
  
### Conflict A: Sloth (Delta: 0.60)  
\* \*\*Subject A Trait:\*\* "Compulsive Productivity" (Weight: 0.95, Conf: 99%)  
 \* \*Evidence:\* "Wake up at 5am. 10k run."  
\* \*\*Subject B Trait:\*\* "Leisure Prioritization" (Weight: 0.85, Conf: 96%)  
 \* \*Evidence:\* "I refuse to work weekends. Life is for sleeping."  
\* \*\*Friction Probability:\*\* 88%. This is a fundamental lifestyle incompatibility.  
  
### Conflict B: Communication Style (Gap: 0.80)  
\* \*\*Subject A Bluntness:\*\* 1.0 (Extreme)  
\* \*\*Subject B Bluntness:\*\* 0.2 (Soft/Passive)  
\* \*\*Outcome:\*\* Subject A will likely be perceived as abusive or insensitive by Subject B within 10-15 interactions.  
  
## 3. Evolutionary Verdict  
\*\*Viability:\*\* \*\*LOW (Long Term)\*\* / \*\*HIGH (Short Term)\*\*.  
The visual triggers are potent, but the psychometric weights on "Productivity" vs "Rest" suggest inevitable resource allocation conflict.

### 7.3 FILE: hla\_gemini\_analysis.md (The "Red Pill" Bio Report)

*Generated via Protocol B using reasoning\_chain.json inputs.*

# Biological Forensic Report (HLA)  
\*\*Methodology:\*\* HIBAG Imputation (EUR Model)  
  
## 1. Data Integrity Audit  
| Allele | Imputation Confidence (Posterior Prob) | QC Status |  
| :--- | :--- | :--- |  
| \*\*A\*01:01\*\* | 0.982 | PASS |  
| \*\*B\*08:01\*\* | 0.991 | PASS |  
| \*\*DRB1\*15:01\*\* | 0.945 | PASS |  
  
\*\*Audit:\*\* All markers exceed the 0.90 confidence threshold. Data is valid for mating prediction.  
  
## 2. Peptide Binding Profile (Functional Distinctness)  
\*Analysis of unique alleles contributed to offspring immune repertoire.\*  
  
\* \*\*A\*01:01 (Subject A):\*\* Binds peptides with negatively charged residues. Critical for specific bacterial recognition.  
\* \*\*A\*24:02 (Subject B):\*\* Binds peptides with aromatic residues. High affinity for HIV-1 epitopes.  
\* \*\*Functional Weight:\*\* The combination of A\*01 and A\*24 creates a \*\*Broad Spectrum\*\* immune profile. This specific pair increases offspring survival probability against retroviral pathogens by estimated 15%.  
  
## 3. Olfactory Spark Prediction  
\* \*\*Heterozygosity Index:\*\* 0.916  
\* \*\*Predicted Pheromone Intensity:\*\* 92/100  
\* \*\*Scientific Assessment:\*\* Subject A will exhibit a strong positive physiological response to Subject B's axillary odors (sweat) due to the lack of overlapping MHC Class I alleles.

### 7.4 FILE: customer\_summary.json (The "Blue Pill" Summary)

*Generated by sanitizing the reports above.*

{  
 "match\_id": "match\_7f8c9a0b",  
 "display\_score": 75,  
 "badges": ["Instant Spark", "Visual Type Match"],  
   
 "synopsis": {  
 "headline": "A Classic 'Opposites Attract' Dynamic",  
 "body": "You have incredible physical and biological chemistry. While you share core values, your different energy levels (Subject A's drive vs. Subject B's relaxed nature) will keep things exciting."  
 },  
  
 "compatibility\_breakdown": {  
 "physical": { "score": 80, "label": "Strong Attraction", "description": "You fit each other's visual preferences almost perfectly." },  
 "lifestyle": { "score": 52, "label": "Complementary", "description": "One of you drives the car, the other picks the music. You balance each other out." },  
 "chemistry": { "score": 92, "label": "Electric", "description": "Your biological markers indicate a rare level of natural compatibility." }  
 }  
}