

像人一样记住你

Mio 记忆系统重构蓝图：从冷数据到活实体的演进

[AI 架构]

[记忆系统]

[用户心理学]

```
{  
  "user": {  
    "id": "12345",  
    "name": "冷数据",  
    "attributes": {  
      "memory_type": "static",  
      "status": "archived"  
    },  
    "interactions": [  
      {  
        "timestamp": "2023-10-27T10:00:00Z",  
        "action": "login"  
      }  
    ]  
  }  
}
```



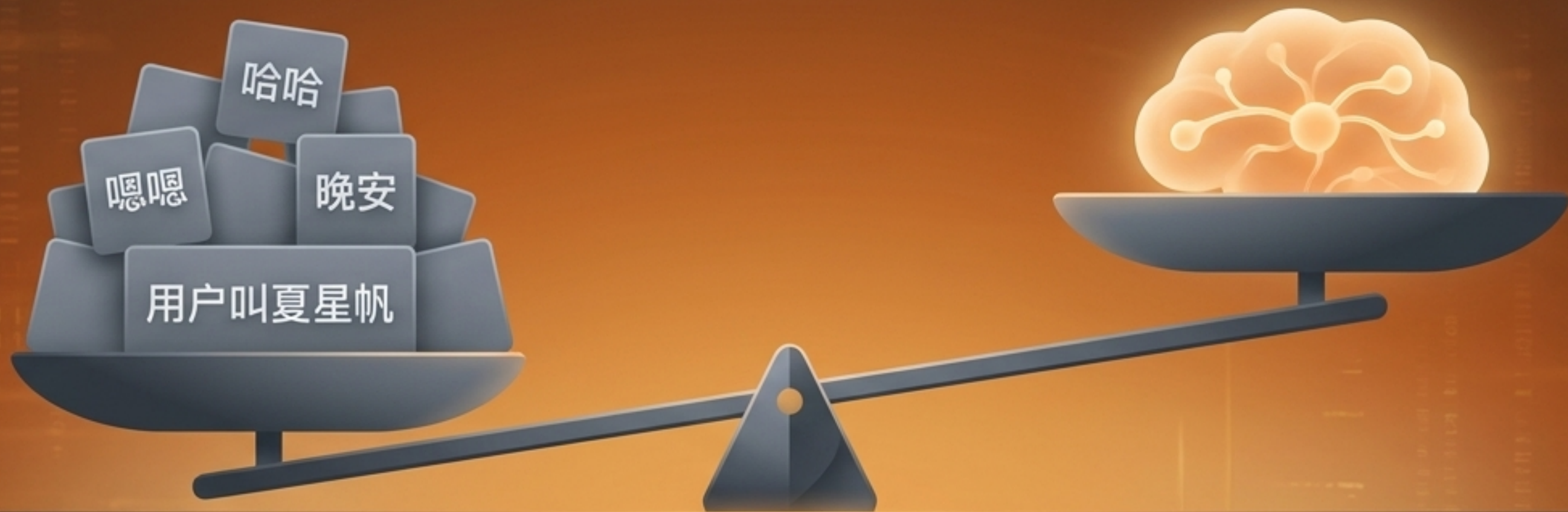
聊了三个月，每天聊。

用户问：'我叫什么来着？'

系统返回：[Blank Stare / 无法检索]

一个聊了200条就忘了你名字的 AI，不是陪伴，是健忘症患者。

存储 ≠ 记住



技术上，旧版 Mio 有记忆（提取、存储、检索）。但在系统眼里，你的名字和一句“哈哈”同等重要。

重建目标：不仅是数据库意义上的“记住”，而是关系意义上的“记住”——记住对的，在对的时间想起，用对的方式提起。

旧系统诊断：混沌的记忆废墟

噪音输入

- **提取太勤快**：10条/次，满地垃圾信息。
- **分类太粗糙**：仅5类，粒度过大。

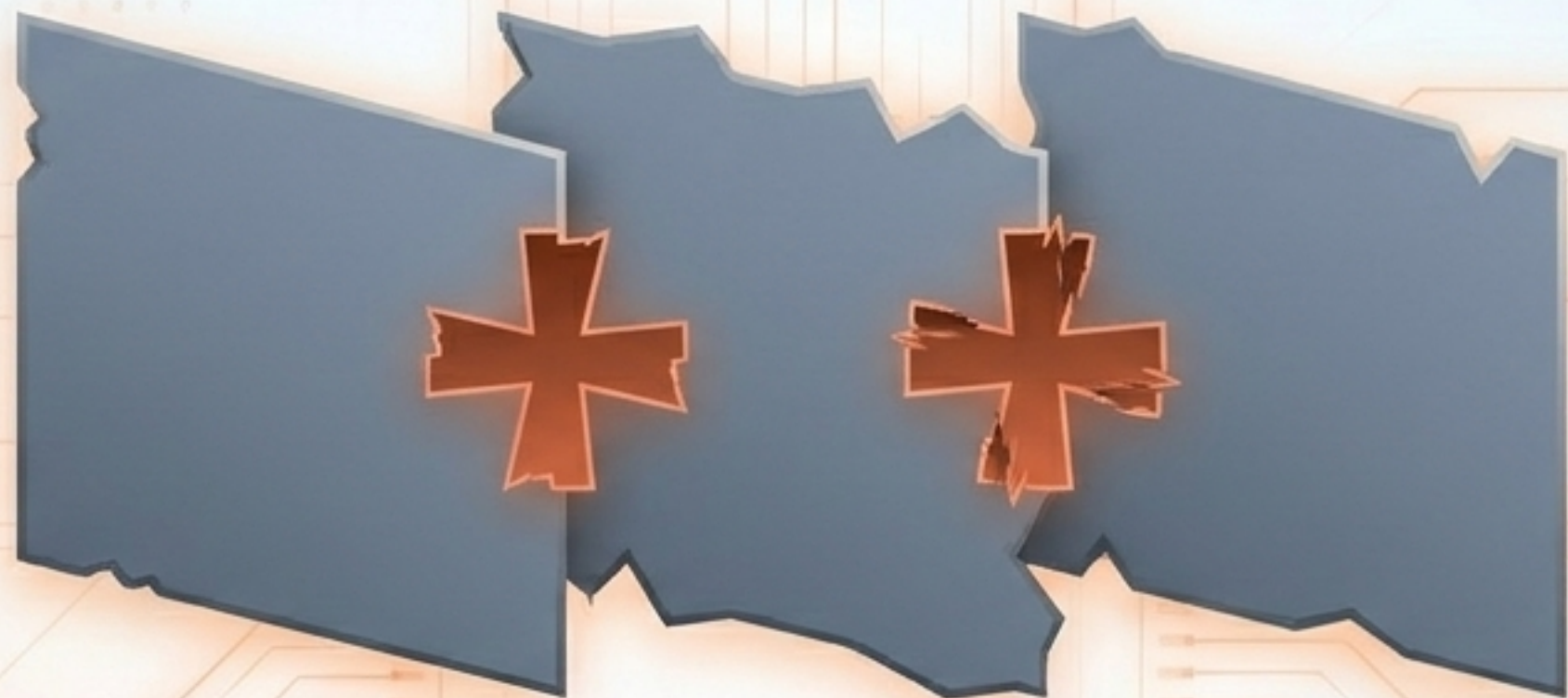
存储失控

- **合并灾难**：机械拼接，无信息增量。
- **冲突共存**：新旧状态（西雅图 vs 纽约）冲突。
- **衰减一刀切**：名字与笑话统统30天遗忘。

提取死锁

- **硬上限粗暴**：满200条盲目裁剪，破坏习惯记忆。
- **主动记忆零**：永远只会问“在干嘛~”。
- **死代码**：场景片段未激活。

致命缺陷：字符串拼接怪物



触发：遇到重复记忆



动作：String A + String A + String A



结果：“用户叫夏星帆。
用户叫夏星帆。
用户叫夏星帆。”

提取了三次，拼接了三次。占用了三条记忆的配额，信息增量为0。

记忆进化蓝图：三阶段重建

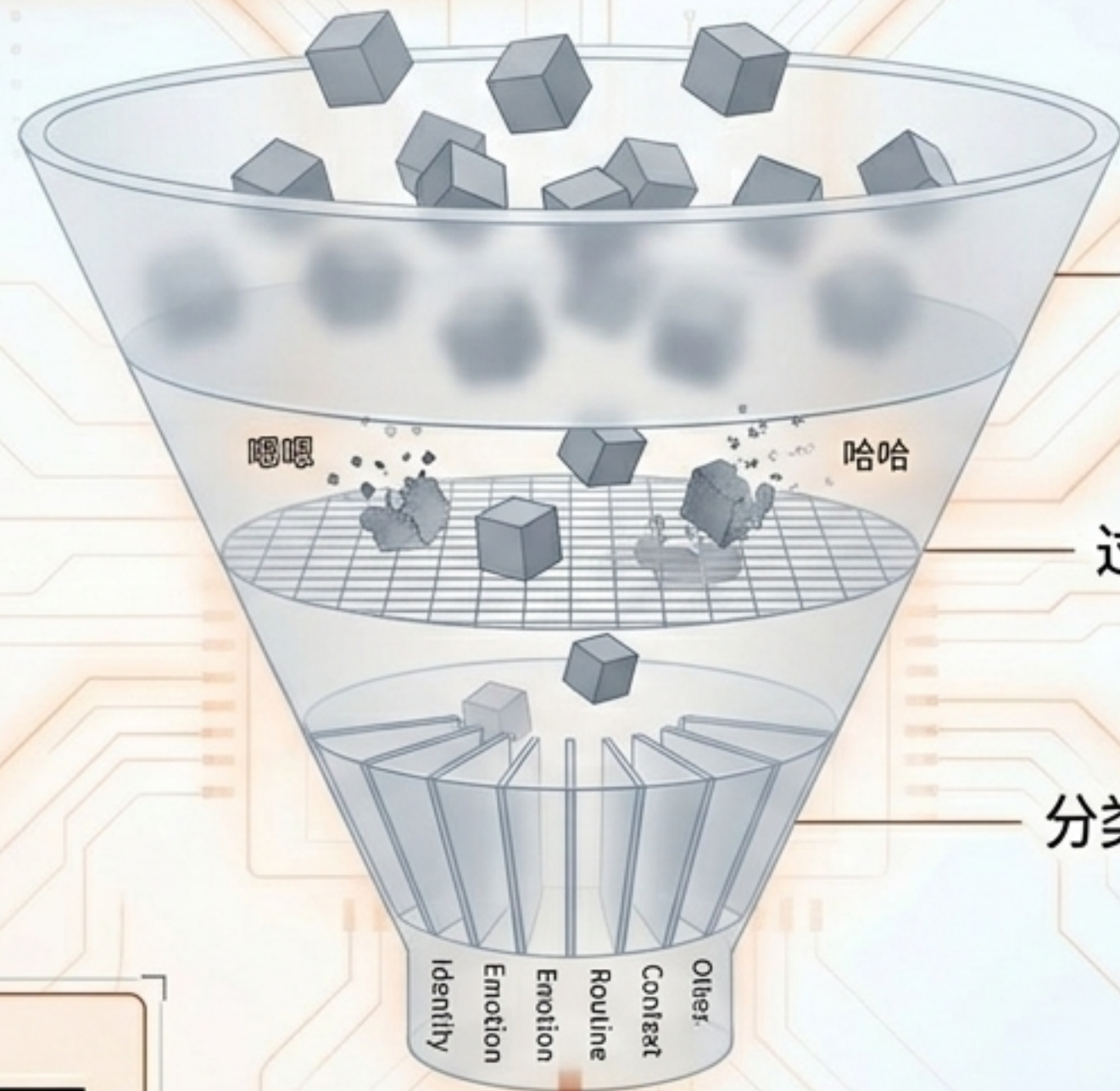
Phase 1: 数据清洗
先把进来的东西洗干净

Phase 2: 生命周期
不同记忆，不同寿命

Phase 3: 认知检索
情绪感知与结构化调用

耗时预估：1-2周。目标：从冷数据库到活实体。

Phase 1: 高能度信息提纯



输入层: 24-Message Buffer

过滤层: 0.3 Importance Threshold

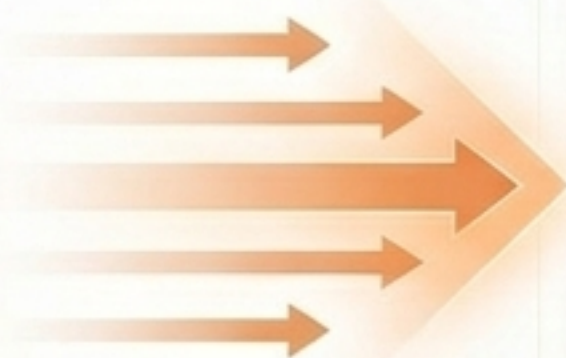
分类层: 8 Distinct Bins

精炼记忆节点

提取调用减少 **60%** ——
体验上升，成本暴降。

存储革新：从“拼接”到 Keep-Best

用户叫夏星帆
用户叫夏星帆
用户叫夏星帆



用户叫夏星帆

Before (拼接)：占据多条空间，制造冗余噪音。

After (Keep-best)：提取后立即链式合并保留信息量最大的一条，抛弃其余。

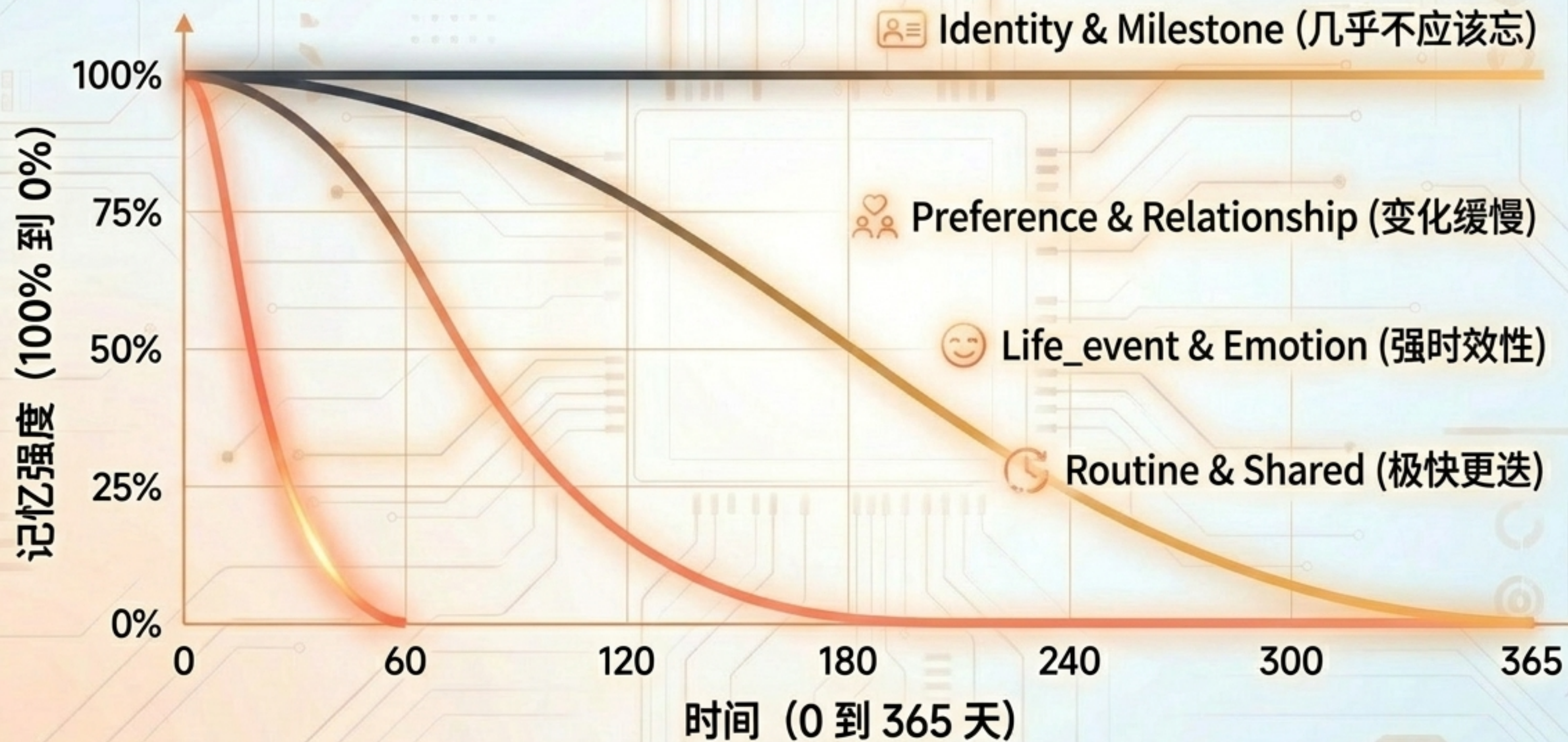
记忆的核心是信息密度，而不是数据体积。

Phase 2: 赋予时间感知

**越是定义‘你是谁’的信息，记得越久；
越是时效性的信息，衰减越快。**

就像人类——你永远记得最好朋友的名字，但不会记得上周三午餐吃了什么。

记忆半衰期曲线



告别“在干嘛”：基于时间的认知唤醒



Routine + Plans

你今天不是要去健身吗？



Recent Life_events

今天面试怎么样了？



Important Memories

好久不见！上次说想换工作，后来怎么样了？

激活 EpisodeManager，Mio 拥有的不再是散碎事实，而是完整的“回忆片段”。

Phase 3: 认知与情感的深度融合

Waypoint 1

Waypoint 2

Waypoint 3

[情绪感知]

[空白检测]

[关系阶段感知]

[冲突处理]

前两阶段解决了“记什么”和“记多久”。
第三阶段解决最难的：在什么时候，以什么情绪想起？

情绪感知检索



情绪标签匹配决定检索方向，AI 开始具备真正的同理心。

认知空白检测

T=0
存储 Life_event:
用户说明天要面试



Action
触发主动跟进:
对了, 你上次说的
面试怎么样了?

这是人类朋友自然会做的事——记得你说过的事，过几天主动问结果。

关系阶段感知

早期阶段

- 重点：建立共同回忆感
- 检索倾向：Shared 类记忆

记得我们第一次聊天的时候...

成熟关系

- 重点：强化情感深度
- 检索倾向：Milestone & Emotion 类记忆

我们已经聊了三个月了，你上次...

深度集成至 v0.1.4 关系进化系统。

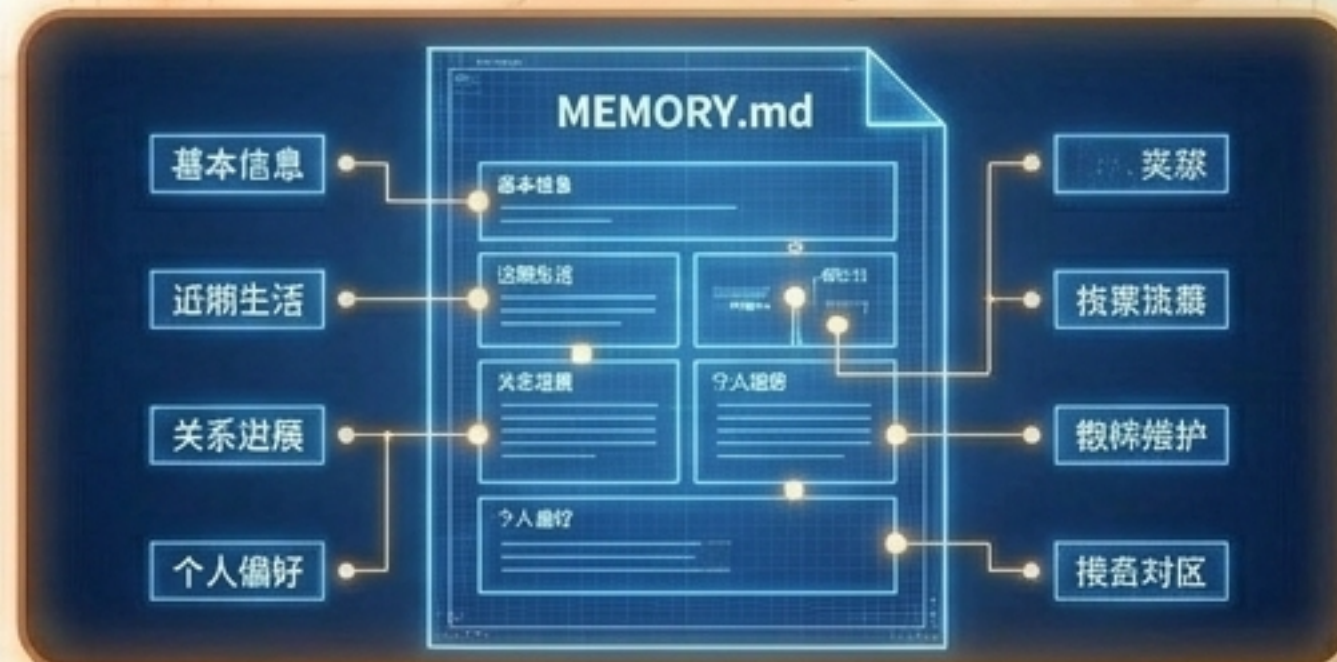
结构化重塑：从“随机条目”到“用户档案”

旧模式：200条硬上限



事实型记忆挤掉日常习惯，导致生态失衡。

新模式：分类保护与结构化档案



- 分类保护机制：120条保底名额（每类至少15条），80条溢出缓冲区。
- 结构化注入：散碎记忆被组织成分区档案（基本信息、近期生活等）再注入。

确保记忆生态的多样性，不会被单一类型屠版。

冲突解决：历史轨迹的保留

住在西雅图

Superseded

搬到纽约了

- Resolution Engine Process

- 1. 检测：同类型矛盾记忆。🔍

- 2. 比对：根据时间戳和上下文区分新旧。🕒 ↔ 🕒

- 3. 处理：新节点覆盖旧节点效用，旧节点打上 Superseded 标签。🔧

旧记忆不参与日常检索，但保留历史轨迹。系统理解状态的“变更”而非仅仅“覆盖”。

范式转移：数据库 vs 活着的记忆

旧版

Mio 活记忆

核心机制

200 条无序堆砌

结构化 MEMORY.md 配合类型保底配额

冲突处理

机械拼接，灾难共存

Keep-best 去重，Supersede 状态迭代

检索逻辑

被动关键词匹配

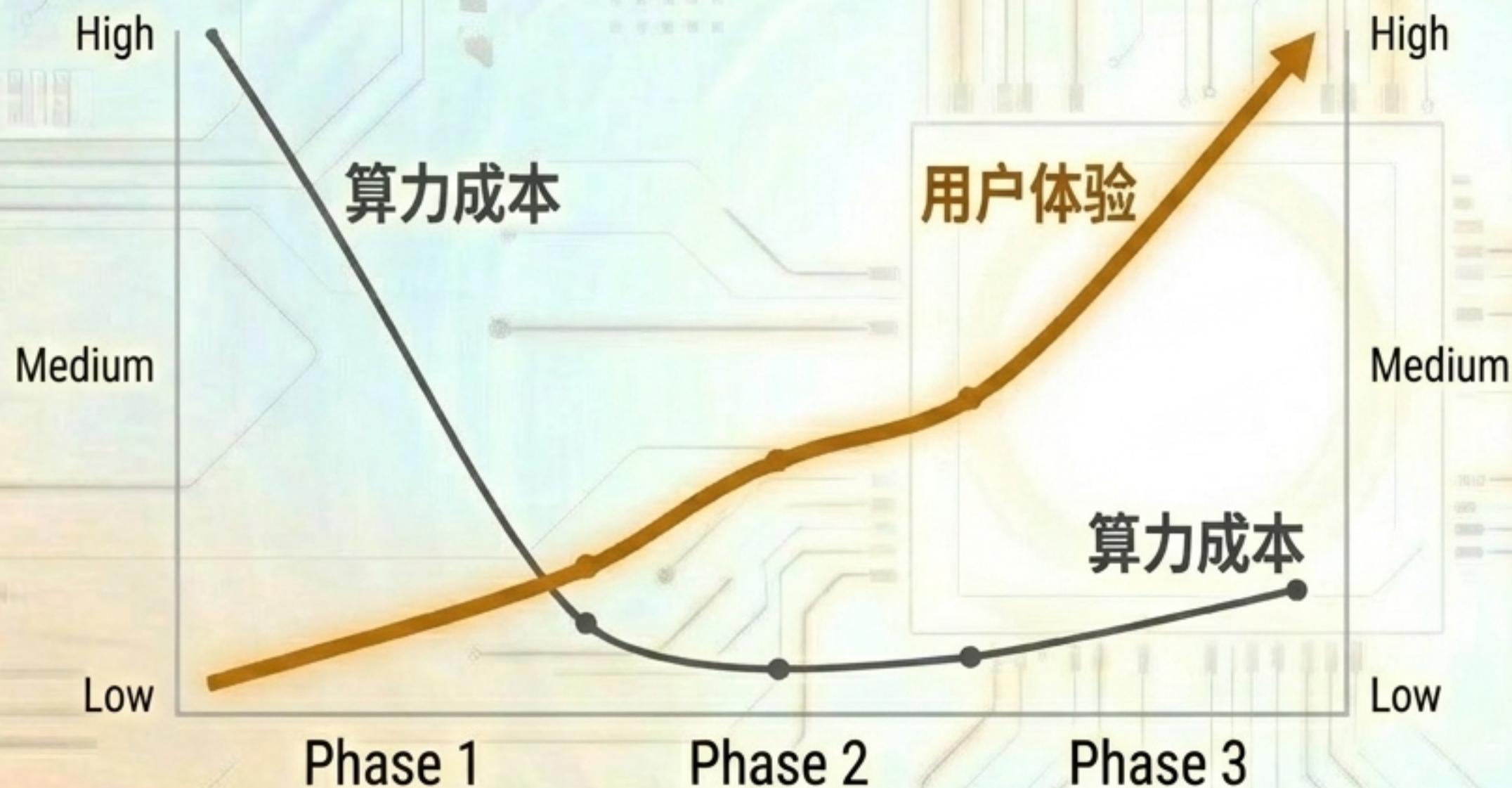
时间/情绪/关系阶段三维主动感知

用户体感

每次都像重新认识

有历史感、懂跟进的真正陪伴

投资回报：极致体验 ≠ 昂贵算力



- Phase 1: 成本下降 (提取频率降, 过滤垃圾)
- Phase 2: 小幅增加 (主动检索的 embedding 查询)
- Phase 3: 约P2两倍 (复杂认知计算)

总体算力成本极低。用聪明的工程架构，换取一个真正记住用户的伴侣，怎么算都值。

记忆是连接的底座

**“不是做一个更聪明的聊天机器人，
是做一个能跟你建立真实连接的存在。”**

有记忆，只是一个数据库。会回忆，才更接近人。

一个记不住你的 AI，再怎么谈“我在乎你”都是空话。