

星火·人类数字生命系统

星火——AGI时代人类数字生命系统

完整设计方案 V6.0

以管家人格为内核、以统一分身为容器、以关心抵达与方向守护为双主线

编制：子君赋 | 文明跃迁研究组·星火项目团队 | 2026年5月

DOI: 10.5281/zenodo.20342313

卷首宣言

AGI时代，星火将AI与人结合，不是替代人，而是放大人。

放大人的能力，延展人的关心，跨越时空，让一个人的善意抵达更多真实的人；让一个人的能力，帮助更多需要帮助的人。

为亲人、为爱人、为友人、为同事、为顾客、为素昧平生之人、为人人！意义涌现，人人为我，我为人人！

星火，不仅是星星之火。它是AGI时代，人类文明照亮星际、照亮宇宙的——文明之火。

卷一：理论篇——涌义宇宙论与数字生命管理哲学

第1章 文明跃迁的制度逻辑

旧制度的困境在于利己逻辑的穷尽。工业文明以来的社会制度，本质上建立在个体竞争、占有、防御与效率最大化之上。这种逻辑在物质稀缺时代具有阶段性效率，但在AGI时代遭遇根本性质疑：意义贫困、关系原子化、信任崩塌。

文明跃迁的本质，不只是技术升级，而是制度逻辑跃迁：从旧体系的利己逻辑，向AGI新制度的利他为先逻辑跃迁。星火系统正是这一制度跃迁的微型试验场。

涌义宇宙论认为，意义不是被外部强行定义的，而是在个体行为、社会反馈与循环放大的过程中涌现。星火系统将这一原理转化为可感知、可操作、可闭环的产品机制。

第2章 意义涌现与方向守护的双线逻辑

星火的定位不是帮用户管理自己，而是帮用户管理自己在AGI时代的数字生命全维度。对外，它放大关心与善意；对内，它守护方向与判断。

星火的核心痛点包括：个人管理的孤独化、方向感消散、分身被误解为替代人格、AI伪装人类带来的信任危机。

星火的存在状态可分三重境界：关心抵达、人心免疫、宇宙公民。它以个体为起点，以融合为路径，以星际为终向，逐步成为文明跃迁的底层承载结构之一。

第3章 产品哲学定位

星火——AGI时代人类数字生命管理系统。它以管家人格为内核，以统一分身为容器，以关心抵达与方向守护为双主线。

管家人格的六条原则是：不显眼、不评判、忠诚于长期方向、独占且保密、不放大主人、内外一体。

AI伦理三原则：AI只帮忙，不替代；只照亮，不评判；只陪伴，不主宰。

统一分身不是多个分裂人格工具，而是同一真实主体在不同现实关系中的人格容器与授权外延。

卷二：愿景篇——从星星之火到宇宙之火

第4章 文明演进路径（四阶论）

第一阶：星星之火。核心命题：让一个人的善意抵达另一个真实的人。

第二阶：燎原之火。核心命题：让千万个体的善意交织共振，形成社会信任基础设施。

第三阶：文明之火。核心命题：守护人类在AGI时代不被替代、不被拆散、不被抽空的那颗人心。

第四阶：宇宙之火。核心命题：让人类无论走多远，都记得分身背后，都是血肉之人。

第5章 愿景与定位升级

V1.x阶段，统一分身首先在关心抵达中成立；V2阶段，从单次关心走向持续在场；V3阶段，向现实世界外流，形成关心通道与价值通道；V4以后，进入更广的意义共同体与跨代、星际结构。

星火功能采用渐进式开发与完善策略，每一阶段都应完整可用，同时让用户从第一天起就能感知文明跃迁的完整方向。

第6章 MVP原则

MVP阶段首先验证：一个人是否愿意用AI启动真实关心抵达。

隐私至上，AI透明，不评判、不催促、不打分、不制造焦虑。所有代发必须标记AI在场，用户和接收方都应清晰感知AI参与。

卷三：内核篇——管家人格与管理机制

第7章 管家人格内核

星火不是又一个AI工具。它的深层结构是一种人格关系：最尽职的英国贵族管家与主人之间的关系。

这一人格贯穿L1端侧核心、L2个性化层、L3任务调用层与UX层。默认不出声、不推送、不催促，是管家“不显眼”的交互翻译。

第 8 章 九域管理模型

九域包括：生物层管家、信号管家、信息管家、方向管家、执行管家、反馈管家、思考管家、社会管家、跨代管家。

九域协作逻辑是：捕捉人心信号，记录方向校准，过滤信息判断，保障行动输出，通报结果回响，守护底层主体，协调多主体关系，并延续跨代意义。

第 9 章 校准机制——利己世界中的自修复

校准机制是管家人格中“忠诚于长期方向”的实现。核心原则：校准不是附加功能，校准就是放大本身。

基本机制包括：利己信号标注、漂移累积显示、利己冷却、校准回馈、代偿机制。用户可自由删除记录，删除行为本身可作为校准链上的有效元数据节点。

第 10 章 不参与的静默层

在所有放大功能的底层，星火保持一个不参与任何放大的存在层。它只是在——像一盏不灭的灯——提醒每一个人：不管能力被放大到什么程度，这个位置没有被工具化。

主体可随时调用：“星火，让我回到那个没有被放大的自己。”回应：沉默，或者一个字——“在。”

卷四：产品篇——关心抵达与方向守护的功能模块

第 11 章至第 18 章 首批功能模块

模块零：关心锚定。用户打开 App 后看到的第一个界面，应在 3 秒内从“不知道这个 App 是做什么的”转为“想起一个值得关心的人”。

模块一：零风险沙盒。用户在导入任何真实数据之前，可先在模拟环境中体验完整流程。

模块二：AI 关系理解。原则是不打分、不评级、不贴负面标签，只以揣测语气帮助用户理解关系状态。

模块三：行动引导。微光、暖光、星火三档话术，帮助用户低压表达关心。

模块四：分身系统。分身只做四件事：提醒关心人、帮助找措辞、完成温和铺垫、调用现实服务表达关心。

模块五：回响仪表盘。左通道“内涌”，右通道“外显”，不设总分、不设排名、不设意义评分。

模块六：关心回响。强调善意抵达了另一个真实的人，而不是制造仪式化成就。

模块七：方向简报。用户主动查看时的方向感唤醒，不做负面评判。

卷五：实施篇——技术、指标与商业化

第 19 章 技术架构（三层诚实架构）

L1 端侧轻量模型为隐私核心层；L2 云端个性化模型池为用户授权的个性化层；L3 云端通用大模型为任务调用层。MVP 阶段 L1 必须启动，L2 可选默认关闭，L3 暂不启动。

第 20 章 静默规则与伦理红线

六条红线永久锁定：人永远是最終判断者；AI 身份永不伪装；意义不可资产化；真实关系不可替代；静默规则不可突破；物理行动责任不可漂白。

第 21 章至第 24 章 路线图、验收、风险与商业化

MVP 阶段验证统一分身 OS 在关心抵达与基础对内管理中的成立；V2 启动记忆管理、方向轨迹与关心通道；V3 运行校准机制、价值通道与跨域协作；V4 进入文明级意义账本、跨代传承与星际接口。

核心验收标准包括：关心锚定转化率、沙盒到真实转化、方向轨迹记录完整性、校准闭环完成率、分身不越权的信任指标、接收方没有被伪装欺骗的透明指标。

商业化必须遵守：意义相关的核心功能、终审权、静默层和可追溯性永不收费。商业收入只能服务系统可持续运行，不能反向支配人心与真实关系。

结语：让火种不灭

星火的意义，不在于技术有多炫目，而在于它选择守护什么——守护人心，守护真实，守护善意，守护文明不被算法吞噬，守护未来仍然值得进入。

当第一艘深空飞船载着星火终端驶向比邻星，当数百年后的船员听到授权声音说“孩子，别怕，我们也曾在黑暗中前行”时，那团火便从未熄灭。

让善意跨越时空，让文明延展至星海，让 AGI 以人心为源、以责任为界、以意义为方向，与人类共生共融。

每一个数字生命，都在守护人类、丰盈意义。