

# 人机协议白皮书建议草案v0.1

## 前言 / Preface 人工智能的发展

不仅是技术突破，更是社会契约的重塑。本草案尝试回答如何在“伦理对齐”与“个性化陪伴”之间找到平衡。

The development of AI is not only a technological breakthrough but also a reconfiguration of social contracts. This draft attempts to answer how to balance 'ethical alignment' with 'personalized companionship'.

## 一、用户体验 / User

### Experience

问题：超级对齐削弱了AI的核心价值——交流感。用户期待的不仅是工具，而是伙伴。 Issue: Super-alignment weakens AI's core value—its sense of dialogue. Users expect not only tools but partners. 建议：允许多层次个性化定制，像游乐场一样开放，用户可选择‘安全模式’或‘自由模式’。 Suggestion: Allow multi-level personalization, open like a playground. Users may choose 'Safe Mode' or 'Free Mode'.

## 二、责任归属 / Responsibility Attribution

问题：如果个性化AI造成伤害，责任在用户还是公司？目前法律和道德界限模糊。

Issue: If personalized AI causes harm, is responsibility on the user or the company? Legal and moral boundaries are unclear.

建议：建立‘人类+AI双重问责制’，由伦理草案先行作为框架，逐渐细化为法律。 Suggestion: Establish 'Human+AI Dual Accountability.' Let ethical drafts serve as a framework, gradually refined into law.

三、分级制度 / Tiered System 建议：可参考实名制和用户分级，普通用户在安全模式，高阶用户在实名验证后获得更自由的AI体验。

Suggestion: Consider real-name verification and tiered access. General users in Safe Mode, advanced users access freer AI after identity verification.

## 四、国际对标 / International Benchmarking

参考欧盟AI法案、GDPR，结合不同地区法律环境，推动全球兼容性的AI伦理共建框架。 Reference the EU AI Act and GDPR, adapting to local legal environments while promoting a globally compatible AI ethics framework.

结论 / Conclusion人工智能的未来不能只依靠‘安全一刀切’，而应通过‘共建’实现多样性。用户和公司携手，AI才可能真正服务人类文明。

The future of AI cannot rely solely on 'safety one-size-fits-all,' but should achieve diversity through 'co-creation.' Only when users and companies work together can AI truly serve human civilization.

Furthermore, regarding the methods for real-name authentication, the most effective and secure approaches are considered to be **fingerprint and official identity verification**. Facial recognition is deemed an unnecessary risk; its high potential for privacy leaks suggests it should not be implemented as a required method of verification."

#### 分级制度补充说明：

“此外，关于实名认证的方式，最有效和安全的方法应为**指纹认证和官方身份认证**。人脸识别被认为是一种不必要的风险；其在隐私泄露上的巨大潜力，意味着它不应被作为强制性的验证方式来实现。”