



中华人民共和国国家标准

GB/T XXXXX—XXXX

心理咨询服务 第 4 部分：人工智能技术辅助应用指南

Psychological counseling—Part 4: Guidelines for supporting with Artificial Intelligence technology

(点击此处添加与国际标准一致性程度的标识)

(征求意见稿)

(本草案完成时间：2026.3)

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX - XX - XX 发布

XXXX - XX - XX 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前 言	II
引 言	III
1 范围	4
2 规范性引用文件	4
3 术语和定义	4
4 总则	5
5 人员与角色要求	6
6 服务流程	6
7 安全与隐私管理	8
8 服务评估与改进	9
附 录 A (资料性) 人工智能在心理咨询中的典型风险清单	10
附 录 B (资料性) 知情同意书要素示例	11

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件是GB/T 30446的第4部分。GB/T 30446已经发布了以下部分：

- 第1部分：基本术语；
- 第2部分：服务流程；
- 第3部分：咨询信息管理；

本文件由全国服务标准化技术委员会提出并归口。

本文件起草单位：中国标准化研究院、国家心理健康与精神卫生防治中心、郑州师范学院、北京师范大学、济南中科泛在智能计算研究院、清华大学。

本文件主要起草人：郑娟尔、罗婉莹、付强、晋争、侯非、王娜娜、杨曦、林丹华、姜怀臣、毕丹丹、曹俐莉、丁哲、张登辉、孙昊明、曹琴。

征求意见稿

引 言

心理咨询服务在促进个体和群体心理健康、增强个体社会适应能力、维护社会和谐稳定方面具有重要作用。标准化有助于提升心理咨询服务的专业性、规范性和有效性，是心理咨询行业健康发展的基石。

GB/T 30446系列标准旨在确立心理咨询服务的基本工作准则，由以下部分构成。

- 第 1 部分：基本术语。目的在于规范心理咨询服务的基本术语。
- 第 2 部分：服务流程。目的在于规范心理咨询服务的基本原则、通用流程及要求。
- 第 3 部分：咨询信息管理。目的在于规定心理咨询服务中所涉及的心理咨询信息记录、归档、使用和保存等各环节的具体要求。
- 第 4 部分：人工智能技术辅助应用指南。目的在于为人工智能技术在心理咨询服务中的辅助应用提供指导。

目前，人工智能技术在心理咨询领域的应用迅速发展，为评估支持、过程管理与风险识别提供了新工具，同时也带来伦理、隐私与安全等新挑战。本文件规范人工智能技术在心理咨询服务中的辅助性应用，明确人类咨询师的主导地位与责任边界，保障服务质量与安全，推动行业规范健康发展。

心理咨询服务 第4部分：人工智能技术辅助应用指南

1 范围

本文件给出了利用人工智能技术辅助开展心理咨询服务的总体原则、人员与角色要求、服务流程控制、伦理与隐私保护、安全管理以及服务评估与改进要求。

本文件适用于提供心理咨询服务的组织在其服务过程中应用人工智能辅助系统进行辅助操作的情形。

本文件不适用于以下情形：

- a) 人工智能系统在无心理咨询师有效监督与介入下，自主提供心理咨询服务或实质性承担应由心理咨询师实施的核心干预环节的情形；
- b) 涉及精神障碍诊断、治疗等医疗服务；
- c) 人工智能底层技术研发活动（如算法训练框架研发、通用大模型研发等）。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 30446.1 心理咨询服务 第1部分：基本术语

GB/T 30446.2 心理咨询服务 第2部分：服务流程

GB/T 30446.3 心理咨询服务 第3部分：咨询信息管理

GB/T 35273 信息安全技术 个人信息安全规范

GB/T 41867 信息技术 人工智能 术语

GB/T 42582 信息安全技术 移动互联网应用程序(App)个人信息安全测评规范

3 术语和定义

GB/T 30446.1 和 GB/T 41867 界定的，以及下列术语和定义适用于本文件。

3.1

心理咨询 **psychological counseling**

由专业人员运用心理学理论和技术，通过语言与非语言交流，给予个人或群体帮助、启发和教育，使其改变认识、情感、态度和行为，解决其在生活、学习、工作等方面出现的问题，促进其心理健康和社会适应能力改善的专业过程。

[来源：GB/T 30446.1，2.1，有修改]

3.2

心理咨询服务 **psychological counseling service**

由心理咨询服务提供者提供的满足个人或群体**心理咨询(3.1)**需求的专业服务行为。

[来源：GB/T 30446.1，2.2，有修改]

3.3

心理咨询服务提供者 psychological counseling service provider

依规范取得服务资格或资质，依法提供**心理咨询服务（3.2）**的主体。

[来源：GB/T 30446.1, 2.3, 有修改]

3.4

来访者 visitor

接受**心理咨询服务（3.2）**的个人或群体。

[来源：GB/T 30446.1, 2.5]

3.5

人工智能辅助系统 AI-assisted system

针对心理咨询服务的既定目标，运用机器学习、自然语言处理、数据分析等人工智能技术，为心理咨询师提供信息整理、辅助记录、内容生成、风险提示、辅助决策支持等输出的一类工程系统。

[来源：GB/T 41867-2022, 3.1.8, 有修改]

3.6

危机信号 crisis indicator

指来访者表达的自伤/自杀意图、危害他人意图、严重情绪失控、现实解体等可能危及自身或他人安全的言语、文本、行为或其他可获得线索。

4 总则

4.1 辅助性原则

人工智能辅助系统的设计与应用宜以心理咨询专业理论、伦理规范和实践为基础，仅作为心理咨询服务的辅助手段，不得替代心理咨询师的专业判断和核心服务。当系统输出涉及高风险情形、关键决策或明显异常时，心理咨询师应立即介入并承担最终责任。

4.2 透明与知情同意原则

心理咨询服务提供者应在启用人工智能辅助系统前，以清晰、易理解的方式向来访者告知：人工智能辅助系统的使用目的、使用范围、主要功能边界、潜在风险类别、数据采集处理存储方式、来访者权利与无人工智能辅助系统的替代方案，并取得来访者明示同意并留存记录。来访者拒绝使用的，心理咨询服务提供者应提供不依赖该技术的替代路径，不得因此降低服务可及性。

4.3 最小必要与目的限定原则

心理咨询服务提供者应仅收集、使用为提供心理咨询服务所必需的数据，不得用于营销、顾客画像、信用评估等与心理咨询服务无关的目的。确因系统优化、技术迭代等必要场景使用数据的，应采取匿名化、去标识化等安全处理措施，且不得侵害来访者隐私及合法权益。

4.4 可控性与人工可接管原则

心理咨询服务提供者应确保人工智能辅助系统具备人工干预、降级与退出机制；当监测到来访者存在自伤、自杀、危害他人等危机信号、系统异常、输出明显错误或涉及关键决策时，应自动触发预警并切换为人工优先接管。

4.5 公平与包容原则

人工智能辅助系统的设计、训练和应用应充分考虑文化差异与多样性，避免算法偏见，保障不同性别、年龄、民族、文化和社会背景来访者的平等服务质量；应建立算法偏见监测与纠偏机制，保障所有来访者获得平等、公正的心理咨询服务支持。

4.6 持续改进原则

心理咨询服务提供者应建立人工智能辅助应用的制度体系与问责机制，明确角色职责、审核放行、风险事件处置、纠偏闭环与定期评审要求，并与 GB/T 30446.3 的记录管理相衔接。

4.7 安全性原则

人工智能辅助系统应具备完善的安全保护能力，符合网络安全等级保护要求，保障数据收集、传输、存储、处理过程中的安全性、保密性、可用性。系统运行应稳定可靠，避免因技术故障对心理咨询服务造成不良影响。

5 人员与角色要求

5.1 角色与授权

心理咨询服务提供者应识别并书面授权与人工智能辅助应用相关的岗位角色，至少包括：心理咨询师、督导/伦理监督人员、人工智能管理负责人、数据保护负责人、系统运行管理人员、技术支持人员、组织管理者；涉及外部供应商或第三方机构的，应通过合同与制度明确数据边界、服务质量要求、安全管理标准与权责接口。

5.2 胜任力与培训

相关人员应接受人工智能伦理、数据与网络安全、隐私保护、系统操作、危机识别与应急处置培训；培训频次应不低于每年一次，并留存记录。关键岗位宜开展胜任力评估与复训。

5.3 职责分工与升级机制

心理咨询服务提供者应建立风险升级与协同机制，至少满足：

- 识别到危机信号、系统异常或输出明显错误时，应由心理咨询师或督导立即接管；
- 涉及数据共享、委托处理、对外提供的决策，应由数据保护负责人牵头会签；
- 涉及模型配置、规则策略或功能重大变更，应由人工智能管理负责人组织评审、验证与放行；
- 需要降级或切换人工模式时，技术支持人员应按预案执行并同步通知相关人员。

5.4 记录与可追溯

心理咨询服务提供者应建立并留存：人员授权清单、培训记录、系统清单与版本、模型说明文档及其更新记录、变更与放行记录、系统操作与数据访问审计日志、个案抽审/伦理会审记录、风险事件与整改闭环记录；保存期限不短于 GB/T 30446.3 的要求。

6 服务流程

人工智能辅助系统仅可辅助 GB/T 30446.2 规定的心理咨询服务流程，包括但不限于：来访接待与匹配、初步评估、方案制定、服务实施、效果评估、回访与反馈、服务结束等环节。心理咨询师应对人工智能输出进行专业验证并承担最终决策责任。

6.1 来访接待与匹配

人工智能辅助系统可用于咨询预约排期、信息结构化整理、与心理咨询师匹配的辅助推荐；不得用于自动拒绝服务、自动标签化结论、自动转介决定。

6.2 风险告知与初步评估

6.2.1 风险告知

心理咨询师应：

- 书面说明人工智能辅助的目的、范围与主要潜在风险类别（含算法偏差/误判、信息不准确或更新滞后、情感理解局限、数据安全与隐私风险等）；
- 告知来访者拒绝权与替代方案；
- 取得来访者明示同意并留存（含电子签署）。

6.2.2 初步评估

人工智能辅助系统可对来访者信息进行结构化处理与辅助分析并生成供心理咨询师参考的初步评估报告。该报告至少应包含：

- 来访者基本情况与数据来源说明；
 - 系统分析得出的来访者心理相关主题与问题线索；
 - 模型适用范围、算法与理解局限性、结果参考性的提示语；
- 心理咨询师应结合面谈、行为观察等进行专业验证并记录采信依据。

6.3 方案制定与验证

心理咨询师可借助人工智能辅助系统提供的初步评估报告，结合专业理论、咨询经验与初步评估结论综合生成个性化心理咨询服务方案建议，包括咨询目标、可能的咨询方法、技巧以及预期目标，并与来访者进行沟通。人工智能辅助系统的建议仅供参考，不得单独形成正式方案。

6.4 服务实施

人工智能辅助系统可用于文本与数据分析、信息整合、辅助记录与提醒；心理咨询服务流程核心环节应由心理咨询师主导。核心环节包括但不限于：首次访谈建立关系、危机状态下的即时干预、情感反馈与共情回应、咨询目标协商与调整、结束阶段的总结与分离处理。

心理咨询师可利用人工智能辅助系统帮助整理心理咨询记录，为后续服务优化和效果评估提供支持。

服务过程中识别到来访者存在危机信号的，心理咨询师应立即启动人工接管，并按照应急预案采取相应处置措施。人工智能辅助系统不得替代心理咨询师对危机状态的专业识别、判断和处置。

6.5 效果评估

心理咨询效果宜结合量化数据与来访者主观体验，通过专业的心理学评估方法进行综合判定。心理咨询师可利用人工智能系统辅助分析心理咨询过程中的对话内容、情感变化及反馈数据，帮助评估咨询效果。心理咨询师不可完全依赖人工智能技术提供的评估结果。

6.6 回访与反馈

心理咨询服务期间，人工智能系统可辅助心理咨询服务提供者和心理咨询师对来访者进行定期回访，通过分析来访者反馈和对话数据，辅助识别关键问题和来访者的情感倾向，帮助心理咨询师更准确地把

握来访者需求，并提供必要的支持和建议。人工智能系统不得以自动化方式向来访者输出建议替代心理咨询师沟通。

6.7 服务结束

服务终止决策必须由心理咨询师基于专业判断作出；人工智能辅助系统的结果不得作为独立依据。

7 安全与隐私管理

7.1 技术安全

心理咨询服务提供者应建立系统安全管理机制，包括身份鉴别、分级访问控制、漏洞扫描与修复、系统版本全链路追溯、功能变更回归测试、操作行为日志审计与备份恢复，防止系统运行异常、非授权访问及滥用风险。

7.2 隐私保护

心理咨询服务提供者应：

- 依照 GB/T 35273、GB/T 42582 落实来访者个人信息安全要求；
- 严格遵循最小必要原则，仅收集、处理和使用为提供心理咨询服务所必需的个人数据，采集范围和数量与服务目的相匹配；
- 实施严格的数据访问控制机制。仅限合法授权的心理咨询师及相关工作人员能够访问心理咨询敏感数据。权限分配需严格对应其工作职责，防止来访者个人隐私泄露；
- 在不影响咨询服务质量、后续服务衔接和危机干预的前提下，对咨询记录进行去标识化或最小化展示处理；
- 采用有效加密技术与算法，对心理咨询过程中产生的个人数据进行加密存储与传输，防止数据在传输、存储过程中被非法截获、篡改或窃取；
- 明确群体数据与个体数据的使用边界，禁止将敏感数据用于非咨询目的或商业化再利用。

7.3 数据用于模型训练、评估与改进的控制

心理咨询服务提供者如需将来访者数据用于人工智能辅助系统的模型训练、微调、评估或功能改进，应满足：

- 另行开展专项隐私与安全风险评估并形成书面评估报告；
- 取得来访者针对该用途的单独明示同意（不得以默认勾选、捆绑授权替代）；
- 允许来访者随时撤回数据的使用同意；
- 采取去标识化、隔离存储、访问权限控制与用途审计措施；
- 不得将数据提供给未被授权的外部系统或用于与心理咨询服务无关的模型通用能力训练；
- 使用第三方平台、外部大模型或云端能力开展处理的，心理咨询服务提供者应明确数据传输方式、处理边界、留存规则、安全责任和退出机制，并确保相关处理活动可核验、可追溯；
- 建立训练数据集定期清理更新机制，移除过时、无效或存在隐私风险的数据。

7.4 数据质量与模型说明

心理咨询服务提供者宜关注人工智能辅助系统所使用数据的相关性、准确性、完整性、时效性和适用性，避免因数据质量缺陷导致输出明显失真、误导或不当提示。

心理咨询服务提供者宜要求技术提供方提供与系统相关的模型说明文档。该文档宜包括以下内容：

- 模型或系统的基本信息；
- 预期用途、适用场景和预期使用对象；
- 主要功能及输出类型；
- 性能指标及其适用条件；
- 训练数据和评估数据的基本概况；
- 已知局限、潜在偏差和主要风险；
- 版本信息及重要更新说明。

上述模型说明文档宜供心理咨询师、督导或伦理监督人员、组织管理者在必要时查阅，并作为系统选用、人员培训、风险评估、使用管理和问题追溯的参考依据。

7.5 应急管理

心理咨询服务提供者应在人工智能辅助系统运行过程中针对系统故障、数据泄露或篡改、伦理风险、隐私泄露等情形制定应急预案，明确第一响应人、应急处理流程、降级与人工接管流程、通报时限与报告路径；每年至少演练一次，并留存记录与整改闭环。

8 服务评估与改进

8.1 服务评估

心理咨询服务提供者宜定期评估人工智能技术在实践中的应用效果与影响，包括服务效率、输出适切性、来访者满意度及服务结果等。

宜定期征求来访者的反馈意见，了解其在接受人工智能辅助心理咨询过程中的真实感受、需求满足情况及改进建议。

8.2 伦理审查与持续改进

心理咨询服务提供者应设立伦理审查机制，并可根据应用风险引入第三方评估反馈机制，定期审查人工智能辅助系统的更新、数据使用合规性、风险控制有效性、伦理符合性。应根据评估结果改进流程、更新指导方针或暂停不符合伦理的应用。

8.3 行业协作与知识更新

鼓励心理咨询服务提供者共享经验，参与行业交流与研究，及时更新与人工智能应用相关的培训内容和行为守则。

附录 A

(资料性)

人工智能在心理咨询中的典型风险清单

- A.1 输出风险：事实性错误输出、无依据信息生成（人工智能幻觉）、过度自信表述、非专业的拟专业化误导
- A.2 越权操作风险：自动生成诊断/贴标签、自动化关键决策、自动化危机处置
- A.3 算法偏差风险：群体差异导致的误判、歧视性建议或提示、特殊群体心理分析的系统性偏差
- A.4 数据安全和隐私风险：未经来访者同意收集/再利用、第三方调用数据泄露、数据使用超出约定范围的用途漂移、未按要求开展数据匿名化/去标识化处理
- A.5 系统安全风险：系统故障中断、系统被非法攻击或内容篡改、日志缺失/不完整导致不可追溯、权限管控失效引发的非授权操作
- A.6 伦理与法律风险：责任界定不清、知情同意不充分、对弱势群体（儿童、青少年、老年人）保护不足、侵犯使用者文化习俗。

征求意见稿

附 录 B
(资料性)
知情同意书要素示例

- B.1 人工智能辅助系统的使用目的、适用范围与不适用范围
- B.2 主要风险类别与局限性说明（含误判、偏差、更新滞后、数据安全与隐私风险等）
- B.3 来访者权利：拒绝权、撤回同意方式、替代服务路径
- B.4 数据处理说明：收集项与依据、用途、存储期限与安全、共享/委托处理、删除与更正渠道
- B.5 心理咨询师责任边界：心理咨询师对关键决策、服务方案定制与服务结果承担最终责任
- B.6 单独同意项（如适用）：数据用于模型改进、评估或效果评估的单独明示同意选项

征求意见稿

参 考 文 献

- [1] 《中华人民共和国个人信息保护法》
 - [2] American Counseling Association (2014). ACA Code of Ethics. Alexandria, VA: Author.
 - [3] 国务院《新一代人工智能发展规划》
 - [4] 《中国心理学会临床与咨询心理学工作伦理守则(第二版)》
-

征求意见稿