**智信Trusthink人工智能创新中心**

**建设方案**

**（2025年8月-2030年7月）**

**山东烟台**

**二零二五年七月**

目 录

[一、前言 1](#_Toc204594756)

[二、建设背景 1](#_Toc204594757)

[1.全球人工智能技术面临巨大的发展瓶颈 1](#_Toc204594758)

[2.可置信AI技术是未来人工智能大模型的发展方向 2](#_Toc204594759)

[三、建设目标 3](#_Toc204594760)

[四、重点建设内容 3](#_Toc204594761)

[1.建设三个核心实验室 3](#_Toc204594762)

[2.建立智信AI开发者生态 4](#_Toc204594763)

[3.建设智信AI人工智能国际合作平台 5](#_Toc204594764)

[4.实施智信人工智能应用工程 6](#_Toc204594765)

[五、实施计划 7](#_Toc204594766)

[1.第一阶段（2025年-2026年） 7](#_Toc204594767)

[2.第二阶段（2026年-2028年） 7](#_Toc204594768)

[3.第三阶段（2028年-2030年） 8](#_Toc204594769)

[六、高水平科研团队建设 8](#_Toc204594770)

[七、创业主体及机构设置 9](#_Toc204594771)

[1.创业主体 9](#_Toc204594772)

[2.办公场地 9](#_Toc204594773)

[3.组织架构 10](#_Toc204594774)

[八、建设经费 10](#_Toc204594775)

[九、预期成效 11](#_Toc204594776)

[1.直接成效 11](#_Toc204594777)

[2.经济效益 11](#_Toc204594778)

[3.社会效益 12](#_Toc204594779)

## 一、前言

人工智能作为引领新一轮科技革命和产业变革的新兴技术，正在深刻改变人类生产生活方式，对经济社会发展和人类文明进步产生深远影响，给整个世界带来了巨大发展机遇。近年来中国在人工智能大模型领域迎来了井喷式发展，各种语言大模型、多模态大模型和具身智能日新月异，推动人工智能向更高效、更智能的方向快速发展，已经成为经济增长的新引擎。目前我国在人工智能领域的产业规模处于世界首位，但人工智能在高速发展同时，也面临着大模型幻觉、缺乏知识约束等问题，会出现难以预知的风险和复杂挑战，亟需新技术解决此类问题。

**智信人工智能创新中心**（拟命名）将在国家人工智能发展规划的指引下，通过引进美国约翰霍普金斯大学人工智能研发团队，打造全球首个可置信AI技术研发中心和科技成果产业化基地，促进我国人工智能产业从跟随到引领的转变发展，建设成为国际领先的人工智能创新中心。

## 二、建设背景

### 2.1 全球人工智能技术面临巨大的发展瓶颈

目前人工智能为“弱人工智能”，仅在图像识别、语音翻译等特定任务中表现出色，缺乏人类般的跨领域学习、推理和自主意识。主要表现为可解释性差、鲁棒性不足容易受数据干扰（如对抗性样本），在复杂、动态环境中稳定性差，例如自动驾驶遇突发情况可能误判。许多 AI 模型（如深度学习）是 “黑箱”，其决策过程难以追溯，在医疗、司法等需要明确逻辑的领域应用受限。

其次存在数据有限性导致的AI偏见，包括算法偏见/预测偏见等，现有AI算法会复制甚至放大这些问题。同时数据不完整、不准确也会导致模型失效。在数据隐私与安全方面存在大量收集用户数据可能引发隐私泄露风险。

全球人工智能大模型从GPT-4开始，基于Transformer的大模型已经使用了整个人类的文字知识进行训练，在发展过程中取得了不错的应用效果。但随着大模型在各个行业的普及应用，即使是在特定方向上，性能最为优异的衍生大模型也触及Transformer架构能力得天花板。更为严峻的是，大模型的幻觉问题是Transformer架构的先天缺陷，缺乏事实支持和逻辑验证，还会存在推理不稳定、知识更新困难、关键领域应用受损等问题，严重制约了人工智能在产业中的实际应用和可持续发展。例如在医疗领域，一个错误的诊断建议可能危及生命；在金融领域，一个不准确的风险评估可能造成巨大损失；在制造业，一个错误的决策可能导致整条生产线瘫痪。这些都是当前AI技术无法逾越的鸿沟，基于现有模型架构无法解决的难题。

### 2.2 可置信AI技术是未来人工智能大模型的发展方向

可置信AI大模型重新设计底层逻辑，引入因果建模与实验校验机制，建立语言生成与知识验证闭环，推动从生成能力向认知能力转型。将可置信技术融入到多模型矩阵，实现多模态数据融合，可以有效解决当前世界大模型面临的发展瓶颈。本团队经过多年研发，在全球首次实现了真正可置信的大模型技术突破，可以有效解决当前大模型幻觉问题、可解释性差、鲁棒性不足、数据泄露以及数据不完整等问题。可置信大模型不是对现有技术的继承改良，而是创新性的从根本上解决了目前AI大模型的信任危机，因此可置信AI技术是未来人工智能发展的重要技术之一。

## 三、建设目标

**3.1 总体目标**

智信人工智能创新中心将打造全球首个可置信AI技术创新高地，研发超感知具身智能机器人，建立可置信人工智能技术体系，打造"可置信AI技术-产品-生态"闭环的人工智能生态，培养一批国际级人工智能专业人才，推动数千家企业实现AI化转型，创造万亿级产业集群，成为未来人工智能大模型的引领者。

**3.2 技术优越性（真对下表，有针对性的补充具体技术实现路线，比如具体优越性如何实现的？）**

|  |  |
| --- | --- |
| **现有AI局限性** | **创新中心对应研发课题** |
| **当前AI能力高度细分，只能解决某个特定技术问题，无法应对新颖、未预见问题或环境** | **设计实现全新框架可置信AI大模型，基于本模型，对接生成业务可信AI处理过程** |
| **AI本身没有自我意识和内省能力。功能实现依赖于算法和数据，而非自身动机或知识本质。** | **通过自主实现AICEO + TrustDev + Spark矩阵技术，实现基于自身动机或知识本质的实现机制，令决策和业务处理革命性提高（99%以上）** |
| **在业务细分领域，基于Transformer的AI大模型幻觉成为最大制约瓶颈** | **可以根据业务场景，通过所提供的可信AI平台，实现现有模型向可置信AI系统转换。**  **相较于DeepSeek，幻觉率可降低40%。** |
| **现有机器人功能需要严格分化，非类人思考模式** | **具有超感知技术和可信AI适配的具身机器人，可以实现软硬一体的类人超智能任务实现，在国际上首次实现国际领先的类人机器人。** |
| **现有世界模型均未进入实用阶段，没有实现业务处理** | **基于是新AI框架的世界泛用模型，可以高效、准确接入工业业务处理、语言处理以及多场景决策** |

## 四、重点建设内容

### 4.1 建设三个核心实验室

①智信AI大模型实验室

实验室将重点研发具有自主知识产权的可置信大模型，建立全新的算法和底层逻辑，引入因果建模与实验校验机制，建立语言生成与知识验证闭环，推动从生成能力向认知能力转型，让大模型具备复杂推理能力、自我学习能力。融合现有开源大模型，破解大模型应用困局和制约因素，让智信AI大模型成为世界独一无二的人工智能大模型。

②超感知具身机器人实验室

超感知具身智能是机器人通过物理实体与环境交互，实现环境感知与理解、信息处理、自主决策与行动执行的智能系统。人形机器人作为其代表形态，市场前景广阔。2024年，全球人形机器人市场规模达203亿美元，预计2029年将增长至1325亿美元，年均复合增长率45.5%。

实验室将建设机器人设计与制造平台、人机协作测试场景平台、行业应用适配平台等多学科交叉平台，目前已研发的具身机器人，成功突破了机器人领域长期存在的"不可能三角”即无法同时实现人形尺寸、人类级自由度和人类力量的技术难题，拥有超高自由度和卓越性能，通用性强，无需改造基础设施，可直接应用在现有人类环境中工作，完美的与可置信人工智能大模型结合，适配各类环境，可逐步实现在工业、医疗、自动驾驶、农业、教育等领域进行商业化、产业化应用。

③智信矩阵技术实验室

智信人工智能大模型技术可以融合多个不同功能的 AI 模型或模块组成的协同系统，通过数据交互、任务分工形成一个 “矩阵” 式结构，共同完成复杂任务，打造可智信技术矩阵，提升整体效率。创新中心将打造可智信矩阵实验室，研发TrustDev、SPARK、AICEO等工具，实现文本、语音、音频、图形、代码等多模型多模态处理，提升可置信大模型研发效率，生成效率以及应用效率。

### 4.2 建立智信AI开发者生态

①TrustDev开放平台

团队已完成TrustDev平台的开发工作，TrustDev可以面向全球开发人员提供可置信AI智能体开发工具和框架，进行低代码开发，构建可置信智能体。平台将秉承开放共享精神，通过举办开发者培训和黑客马拉松，建立开发者社区和技术支持体系，让更多同道中人加入可置信AI开发中，不断升级迭代，共同探索人工智能技术，推动可置信人工智能技术发展。

②SPARK数据平台

团队已完成SPARK知识库的开发工作(不同于Apache Spark)，构建可置信人工智能行业知识库资源池，提供知识库开发工具、数据和服务，构建数据图谱，融合多源知识，支持人工智能产业互联网知识共享，同时加强训练数据采集、模型生成等环节数据安全管理。

③AICEO应用展示平台

独立研发的AICEO作为我们的核心产品，是企业智能化转型核心引擎，实现了真正的企业AI化：无界管理体验区，展示如何实现零级汇报、扁平化管理；10分钟上手培训中心，让企业管理者快速掌握AI管理工具；全景洞察指挥中心，演示AI如何提供超越人类极限的企业洞察力，让可智信AI应用成为现实。

### 4.3 建设智信AI人工智能国际合作平台

创新中心将打造可置信人工智能国际合作平台，整合行业优势资源，共建联合实验室、共建安全测试互认平台、共建数据集和语料库，开展国际人工智能交流活动，组织人工智能教育培训等，推动人工智能技术的发展和应用。建立可置信AI三方评测认证实验室、国际人工智能NQI机构，为医疗、教育、自动驾驶、法律、产业等不同置信等级要求的行业，制定相应的安全应用标准。建立完整的可置信度测试体系，为各类AI应用提供权威认证，解决大模型问题，推动人工智能在全行业的深度应用。

### 4.4 实施智信人工智能应用工程

①AI工业智能体孵化器

通过TrustDev和SPARK技术，帮助企业开发工业智能体，并与独自开发的具身机器人无缝对接，实现自主决策系统开发、自适应生产线设计、智能工厂大脑构建、无人化运营方案，打造工业智能体孵化器，真正实现人工智能+开发应用场景。

②AI智能制造示范工厂

选择2-3家代表性制造企业，通过全面部署可置信AI系统，部署预测性维护和质量智能控制，生产线全面应用我们研发的具身机器人，实现企业智能制造升级和工业4.0实践，打造真正的智能制造示范工厂。

③AI智慧医疗示范平台

与医疗机构合作，通过AI辅助诊断系统（可置信度99.9%），应用手术机器人及智能药物管理系统，展示可置信AI在医疗领域的安全应用，为未来医疗提供示范提供全面的解决方案。

④AI智慧法律、教育示范平台

开发中厚智法、LearnAny.AI两个平台，利用可置信AI技术，进行法律、教育行业AI赋能。（充实下两个平台性能优越性的简洁介绍）

## 五、实施计划

### 5.1 第一阶段（2025年-2026年）

创新中心建设初期，主要开展基础建设与试点工作，完成创新中心基础设施建设，组建核心技术团队（100人），在3-5家企业开展AICEO试点，发布可置信AI技术白皮书，完善LearnAny和中厚智法两个基于可信框架的实验性线上服务平台，开展具身智能机器人原型机研发工作。

目前LearnAny.AI与中厚智法智能体开发完成已经进入测试阶段，可进行商业化应用。

建设完成AICEO的企业决策场景示范平台建设和推广。

完成针对医疗以及1-2个特定工业生产场景的高灵巧超感知机器人的产业化应用。

### 5.2 第二阶段（2026年-2028年）

创新中心建设中期，主要工作是可置信人工智能大模型技术突破与推广，完成可置信大模型2.0研发，建立可置信大模型标准体系。持续完善TrustDev开放平台、SPARK数据平台、AICEO应用展示平台等可置信人工智能开发者生态。实现具身智能机器人批量生产，AICEO覆盖100家企业，制定首批行业标准3项，培养专业人才1000人，具备自我造血能力，实现盈利。

完成针对医疗以及不少于3个行业门类（机械、化工、汽车等）的工业生产场景的高灵巧超感知机器人的批量供货；全国范围内企业AICEO导入案例，不少于1000家客户企业。

### 5.3 第三阶段（2028年-2030年）

创新中心后期，主要将中心研发成果进行规模应用与未来布局，完成人工智能国际合作中心建设，形成完整的可置信AI产业链，服务企业达到1000家，产业规模突破千亿，成为国际可置信AI创新中心，启动基于世界模型的下一代AI研发，通过可置信AI和具身机器人的大规模应用积累数据，与各行业合作伙伴共同探索世界模型构建路径。

医疗机器人全国导入数量不少于2000台/年；基于可信AI的超感知具身机器人导入数量不少于5000台/年；AICEO的导入数量客户企业不少于1000家。

## 六、高水平科研团队建设

创新中心核心技术团队来自于约翰霍普金斯大学李浩东博士以及从事国际前沿多维度可信AI方向研究的创业团队成员，同时在国内招聘工程应用创新技术团队，辅助推进技术成果产业实施。

到2030年底，创新中心人员总数达到200人，其中固定研究人员120人，柔性引进人员80人。面向可置信人工智能、具身机器人、AI开发者生态平台、未来AI研究、AI工业化应用等多个研究方向，打造由国内外一流专家领衔的核心团队，形成多个由领军人才牵头的、结构合理、高水平科研团队。

## 七、创业主体及机构设置

### 7.1 创业团队

创业团队介绍

团队核心成员简介、工作基础、技术成果（专利、论文、标准、奖励等），**（可参考下表或直接修改下表。）**

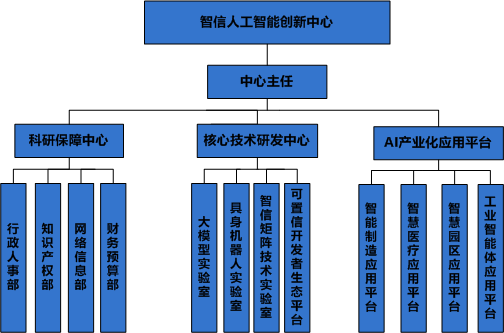
|  |  |  |
| --- | --- | --- |
| 团队名称 | 团队人数 | 团队任务 |
| 技术团队 | 李浩东博士团队共5人 | 可信AI核心框架研发；超感知具身机器人研发与制造；推动核心技术成果在企业AI决策；具身机器人、AI医疗等领域产业化；  开发升级LearnAny与中厚智法2个专业线上平台。 |
| 资金保障与法务 | 1人 | 保障前期6个月资金不少于3000万元，前2年不少于2亿元，与风投/商业伙伴的法律保障 |
| 工程应用团队 | 7人 | 负责核心技术成果产业化适配研发/产业项目定制/维保等；机器人产品的产业适配、应用对接 |
| 营销团队 | 2人 | 客户拓展/合作业务对接等 |

### 7.2 办公场地

创新中心将建设可置信大模型实验室（500㎡）、产业化应用工程技术研发中心（2000㎡）、具身机器人原型工厂实验室（1000㎡），可置信AI开发者开放平台（1500㎡）以及配套的综合会议室、小会议室、科研保障中心、健身服务中心、餐厅等共计6000㎡。

### 7.3 组织架构

创新中心的组织架构见下图，实行中心主任负责制，下设核心技术研发中心、科研保障中心、AI产业化应用平台。



组织架构图

## 八、建设经费

创新中心建设经费主要由三部分组成，主要由自有资金、政策资金、股权融资、社会资本等构成。

前期启动资金2亿元，通过广泛推介和商务洽谈，目前已经达成初步风投意向的国内外投资总额超过2亿元，其中国外资金投资意向超过1.2亿元，国内超过0.8亿元，完成自筹资金2000万元。主要用于高端人才引进、实验室建设以及大模型开发；

建设中期建设将采取社会融资、股权投资以及政策资金支持等方式募集资金20亿元，主要用于大模型推广应用以及具身机器人产业化等；

建设后期将在资本市场上市，募集资金100亿元，主要用于未来可置信人工智能开发，产业化大规模应用、智慧产业园区推广及全球化。

## 九、预期成效

### 9.1直接成效

创新中心的建成，将成为全球首个探索从当前AI路线到未来通用可信人工智能的专用技术；形成拥有完全自主知识产权的可置信AI技术体系；建成年产万台规模的具身机器人生产基地；培育可置信AI相关企业超过500家；创造高端就业岗位5万个。

### 9.2 经济效益

创新中心的建成预计将在5年内带动相关产业产值100亿元，形成1000亿元产业规模，企业平均生产效率提升50%，管理成本降低50%。

### 9.3 社会效益

创新中心的建成，将助力中国成为世界主要人工智能创新中心，为社会培养一批人工智能专业人才，确保AI技术在关键领域的安全应用，推动传统产业智能化转型，提升国家AI技术自主可控能力。

2025年7月28日