生成式人工智能系统密码应用指引

Cryptography application guide
for generative artificial intelligence systems

（版本 1.0）

中国密码学会密评联委会

2026 年 8 月

目　　次

前　言

近年来，生成式人工智能以前所未有的速度不断突破与迭代，并在政务、金融、医疗、能源等重要领域的智能化转型中发挥着核心驱动作用。然而，随着生成式人工智能深度融入生产生活，其安全风险已从单一技术环节向全生命周期渗透。基础设施的分布式架构暴露新型攻击面，模型参数的“黑箱”特性滋生信任危机，训练与推理数据的海量流动加剧隐私泄露隐患，生成内容的不可控性更可能被恶意利用引发社会风险。在此背景下，如何构建与生成式人工智能技术特征适配的安全防护体系，已成为保障数字经济健康发展的重要命题。

密码技术作为网络安全的核心基石，在生成式人工智能安全体系中承担着不可替代的角色。从基础设施层的算力集群访问控制、存储数据加密，到模型层参数防泄露与对抗样本防御，再到应用层用户身份认证、生成内容溯源，密码学的主动防御能力贯穿生成式人工智能系统全链路。但相较于传统信息系统，生成式人工智能的动态性、复杂性与开放性，对密码应用提出了更高要求。

为落实《密码法》、《数据安全法》及人工智能安全治理相关要求，破解生成式人工智能基础设施、模型安全、数据流转、应用落地中的密码应用难题，特编制本指引。本指引立足生成式人工智能系统技术特点，围绕运行环境与基础设施安全、数据安全、模型安全、人工智能模型服务安全及智能体应用安全五个方面，系统识别关键安全威胁，提供可参考的安全指引，助力构建“技术可控、风险可防、责任可溯”的生成式人工智能系统安全生态，推动人工智能与密码技术的协同发展。

本指引作为GB/T 39786-2021《信息安全技术 信息系统密码应用基本要求》重要扩展，立足生成式人工智能技术特性与安全风险，系统提出适配其全生命周期的密码应用技术指引，旨在为生成式人工智能系统开发单位、服务提供商及行业用户提供安全实践指引，推动构建“技术可控、风险可防、责任可溯”的生成式人工智能系统，切实落实《密码法》、《数据安全法》及相关法律法规要求。

本文件起草单位：商用密码检测认证中心、三未信安科技股份有限公司、深信服科技股份有限公司、深圳市网安计算机安全检测技术有限公司、中国工业互联网研究院（工业和信息化部密码应用研究中心）、北京百度网讯科技有限公司、腾讯科技（深圳）有限公司、国家信息技术安全研究中心、中国科学院信息工程研究所、中国科学院软件研究所、中国电子科技集团公司第十五研究所（信息产业信息安全测评中心）、智巡密码（上海）检测技术有限公司、永信至诚科技集团股份有限公司、北京信安世纪科技股份有限公司、阿里云计算有限公司

本文件主要起草人：李红芳、罗鹏、高志权、张立花、叶盛元、叶润国、苗子萌、付夏冰、邓诗智、杨宏志、杨辰、陈小庆、吴月升、彭石坚、魏士慧、何瑶、宋利、杨龙、刘军荣、孙明明、汪宗斌、马姚、姚杨

本文件内容将根据人工智能实际发展情况适时迭代更新。

生成式人工智能系统密码应用指引

1 范围

本文件分析了生成式人工智能系统在运行环境与基础设施安全、数据安全、模型安全、人工智能模型服务和智能体应用安全等方面的主要安全风险，提出了生成式人工智能系统的密码应用技术框架，描述了生成式人工智能系统的密码应用方法，为生成式人工智能系统开发单位、服务提供商及行业用户提供可参考的密码应用指引。

本文件适用于生成式人工智能系统的开发、建设和运营。

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 45438-2025 网络安全技术 人工智能生成合成内容标识方法

GB/T 39786-2021 信息安全技术 信息系统密码应用基本要求

GB/T 41867-2022 信息技术 人工智能 术语

GB/T 45230-2025 数据安全技术 机密计算通用框架

GB/T 45288.1-2025 人工智能 大模型 第1部分：通用要求

GB/T 45574-2025 敏感个人信息处理安全要求

GB/T 45652-2025 网络安全技术 生成式人工智能预训练和优化训练数据安全规范

GB/T 45909-2025 网络安全技术 数字水印技术实现指南

GB/T 45958-2025 网络安全技术 人工智能计算平台安全框架

GB/T 46284-2025 人工智能 联邦学习技术规范

GB/T 46800-2025 生成式人工智能技术应用社会影响评估指南

GM/T 0135-2024 多方安全计算 技术框架

GM/Z 4001-2013 密码术语

3 术语和定义

GM/Z 4001、GB/T 41867中界定的及以下术语和定义适用于本文件。

3.1

大模型 large-scale model

大规模深度学习模型 large-scale deep learning model

基于大量数据训练得到，具有复杂计算架构，能处理复杂任务，且具备一定泛化性的深度学习模型。

注：大模型的参数量由其功能和模态决定，一般不低于1亿。大模型训练使用的数据总量受参数量的影响，达到收敛的大模型的参数量的对数与其训练数据总量的对数成正比。

[来源：GB/T 45288.1-2025，3.1]

3.2

生成式人工智能 generative artificial intelligence;GenAI

生成文本、图片、音频、视频、软件代码等内容的人工智能模型及相关产品服务。

[来源：GB/T 46800-2025，3.1]

3.3

智能体 agent

智能体是具备自主感知、记忆、决策、交互与执行能力的智能系统，是人工智能产品及服务的重要形态。

注：本文件中的智能体特指基于生成式人工智能技术的智能体应用，一般为软件系统。

3.4

模型配置文件 model configuration file

是用于记录机器学习/深度学习模型全生命周期关键信息的结构化文件。

3.5

模型权重文件 model weight file

是存储机器学习/深度学习模型训练后关键参数数据的二进制或结构化文件。

3.6

提示词 prompt

使用大模型进行微调或下游任务处理时，插入到输入样本中的指令或信息对象。

[来源：GB/T 45288.1-2025，3.5]

3.7

语料 corpus

在真实语言使用环境中自然产生的、经过系统采集和加工的语言材料集合。

3.8

预训练 pre-training

使用大规模数据使生成式人工智能模型获得通用知识的训练过程。

[来源：GB/T 45652-2025，3.4]

3.9

优化训练 fine-tuning

在预训练基础上，使用特定领域数据使生成式人工智能模型获得面向领域服务能力的训练过程。

[来源：GB/T 45652-2025，3.5]

3.10

检索增强生成 retrieval augmented generation（RAG）

从预构建的知识库或数据源中检索相关信息，在特定任务上增强生成人工智能模型的技术方法。

3.11

RAG 数据库 retrieval-augmented generation database

是一种结合“外部知识检索”与“大语言模型生成”的混合 AI 技术。专为 RAG 系统设计，用于存储、管理“可供检索的结构化/非结构化知识”，并支持高效语义匹配的专用数据库。

3.12

同态加密 homomorphic encryption

对称或非对称加密，允许第三方（既不是加密者也不是解密者）以加密的形式对数据执行计算，同时确保明文数据的机密性。

[来源：GB/T 46284-2025，3.6]

3.13

差分隐私 differential privacy

在数据分析中，通过按特定概率分布向数据添加随机噪声，以严格控制隐私预算来保护个人隐私的技术。

[来源：GB/T 46284-2025，3.7]

3.14

多方安全计算 secure multi-party computation; SMPC

保证各参与方在不泄露原始输入的前提下，协同计算预期函数并得到最终结果的密码技术。

[来源：GM/T 0135-2024，3.8]

3.15

机密计算 confidential computing

在可信的硬件基础上，通过隔离、加密、证明等机制，保护使用中数据安全的计算模式。

[来源：GB/T 45230-2025，3.3]

3.16

零知识证明 zero-knowledge proof

一种证明方与验证方之间进行一次或多次交互问询的方法：证明方能够使验证方确信某个命题为真，同时不会向验证方泄露该命题真值以外的任何额外信息。

3.17

数字水印 digital watermarking

通过在数字内容中嵌入不易察觉的特定信息，标识或保护数字媒体的版权、来源或内容的信息安全技术，其所嵌入的信息对数据使用者是隐蔽且不可辨识的。

[来源：GB/T 45909-2025，3.1，有改动]

3.18

文件元数据 file metadata

按照特定编码格式嵌入到文件中的描述性数据，用于记录文件来源、属性、用途、版权等信息。

[来源：GB/T 45438-2025，3.5]

4 缩略语

下列缩略语适用于本文件。

ABE：基于属性的加密（Attribute-Based Encryption）

AI：人工智能（Artificial Intelligence）

API：应用程序编程接口（Application Programming Interface）

FPE：格式保留加密（Format-Preserving Encryption）

GPU：图形处理器（Graphics Processing Unit）

HMAC：带密钥杂凑的消息鉴别码（Keyed-Hash Message Authentication Code）

IoT：物联网（Internet of Things）

IPSec：IP 安全协议（Internet Protocol Security）

RAG：检索增强生成（Retrieval-Augmented Generation）

SMPC：多方安全计算（Secure Multi-Party Computation）

SSH：安全交互（Secure Shell）

SSL：安全套接字层（Secure Sockets Layer）

TCM：可信密码模块（Trusted Cryptography Module）

TEE：可信执行环境（Trusted Execution Environment）

TLCP：传输层密码协议（Transport Layer Cryptography Protocol）

VPN：虚拟专用网络（Virtual Private Network）

5 概述

生成式人工智能系统能够根据输入提示词自动生成文本、图像、音频、视频、代码等多模态内容的智能系统。其核心通常由大规模参数化的深度神经网络模型构成，具备强大的上下文理解、知识推理与内容创造等能力。本文件从建设及运营角度提出与密码应用相关的生成式人工智能系统参考架构，如图1所示。

图1 生成式人工智能系统参考架构

——运行环境与基础设施层包括计算资源、存储资源、网络资源等硬件资源，以及资源虚拟化/调度和人工智能运行框架等软件资源；

——数据层贯穿系统各层次，包括预训练数据、优化训练数据和使用阶段数据；

——模型层包括基础模型、定制化模型，其中，定制化模型是依据用户需求对基础模型进行微调，定制适用于具体使用环境的模型；

——应用层包括人工智能模型服务、智能体应用等，智能体包含嵌入或集成大模型能力的其他软件。

6 安全风险

6.1 运行环境与基础设施安全风险

6.1.1 网络环境安全风险

(a) 不安全端口及外部接口。人工智能分布式部署（如多节点集群、跨云边端API服务）中，未受保护的管理端口、外部交互接口因缺乏身份认证或传输加密机制，存在被攻击的风险。

(b) 网络隔离与访问控制不足。云架构下模型训练/推理组件与敏感数据区间的网络隔离或访问控制机制，因缺乏密码技术保护，易引发跨域越权访问或数据泄露。

(c) 不安全的网络协议与配置。网络配置或网络协议缺乏身份鉴别及机密性、完整性保护，存在模型交互数据与控制指令被窃听或篡改的风险。

6.1.2 计算环境安全风险

(a) 操作系统风险。操作系统未启用密码策略（如弱口令、无多因素认证）或关键操作（如特权命令执行）未进行安全保护，存在未授权访问、恶意代码植入或系统配置篡改的风险。

(b) 虚拟化技术风险。虚拟机/容器镜像分发过程未采用完整性保护，或容器间通信时未采用安全协议，存在镜像被篡改、容器通信内容被窃听的风险。

(c) 计算资源弹性配置风险。弹性扩缩容时未验证新增计算节点身份，存在恶意节点非法占用资源或劫持计算任务的风险。

(d) 运行框架风险。在启动和交互时，未对人工智能运行框架内的关键组件、配置文件等进行完整性校验、身份鉴别或通信保护，攻击者可能利用框架漏洞或篡改配置操纵训练过程、窃取梯度信息或植入后门。

(e) 第三方依赖库风险。在安装或运行前，未对第三方依赖库进行来源真实性或完整性验证，可能因使用恶意库组件导致运行过程被操控或敏感数据泄露。

(f) 日志篡改风险。系统日志未采用完整性保护，攻击者可通过篡改或伪造记录掩盖异常访问痕迹。

(g) 不安全的身份认证机制。用户或API未采用安全的身份鉴别或访问控制策略，攻击者可能越权访问模型参数、篡改训练数据或非法调用推理服务。

6.1.3 存储环境安全风险

(a) 存储资源隔离控制不足。不同业务类型或不同安全等级的存储资源未实现逻辑隔离，可能导致敏感数据跨域非法访问或混淆，引发隐私泄露或业务干扰风险。

(b) 数据分布式存储机制缺陷。分布式存储系统未采用安全的数据副本同步或容灾备份机制，可能因副本被篡改、丢失或同步错误，导致数据无法正确恢复、影响模型服务准确性等风险。

6.1.4 供应链安全风险

(a) 资源更新与分发机制风险。人工智能应用、依赖库或模型参数在更新与分发时，未进行来源真实性或完整性验证，未采用传输保护，可能导致模型功能异常或敏感信息泄露。

(b) 复杂多源供应链风险。人工智能依赖海量、多源的训练数据及复杂的开源预训练模型、框架和第三方API服务，可能存在数据投毒、模型篡改、依赖服务断等风险。

6.2 数据安全风险

6.2.1 预训练数据安全风险

(a) 数据来源管控不足。预训练阶段模型训练数据的采集、清洗、存储和使用过程中，因缺乏来源合法性验证、合规性审计，或缺少访问控制机制等，可能存在侵权、泄密、偏差传播或违规的风险。

(b) 隐私泄露。预训练语料中未脱敏处理或未经隐私影响评估的个人信息，叠加模型对高频或结构化片段的强记忆效应，使攻击者可通过提示词诱导或成员推断攻击，复现训练数据中的隐私内容导致隐私泄露。

6.2.2 优化训练数据安全风险

(a) 回流数据滥用。使用用户交互产生的回流数据（如对话记录、反馈日志）进行优化训练时，攻击者可能通过模型输出反推或窃取原始数据，泄露回流数据中的敏感信息或商业机密。

(b) 私域数据越权访问或篡改。优化训练时未采用访问控制保护的私域数据（如医疗、金融等高敏数据），可能被未授权访问或篡改，破坏数据的机密性与完整性。

(c) 标注数据污染。优化训练时，若数据标注过程遭恶意干扰（如错误标签注入、标注工具被投毒）或存储的标注数据被恶意篡改，将损害标注数据的真实性与完整性，影响模型输出的可信性。

6.2.3 使用阶段数据安全风险

(a) RAG数据污染。未受保护的外部知识库与生成模型交互过程，存在数据源可信性不足、数据污染与投毒、敏感信息泄露、越权访问、检索结果偏差等风险。

(b) 生成内容可信标识缺失。人工智能生成内容若缺乏可信标识机制，存在显式或隐式来源标识被篡改、虚假信息泛滥、溯源能力缺失、社会信任受损等风险。

(c) 智能体应用敏感数据泄露风险。智能体应用未对关键参数、多智能体协同数据、记忆数据及关键操作指令和决策指令进行防护，可能导致敏感数据、长期记忆或RAG知识库被篡改或泄露的风险。

6.3 模型安全风险

6.3.1 模型篡改与泄露风险

(a) 模型篡改。未受保护的模型存在参数篡改、后门注入、推理逻辑篡改、模型权重文件替换等风险。

(b) 模型泄漏。未受保护的模型可能因模型参数、架构、训练数据或输出行为泄露，导致模型知识泄露（如模型抽取、模型逆向、成员推断、参数泄漏、梯度泄漏、提示词泄漏）。

6.3.2 模型对抗攻击风险

(a) 反演攻击。攻击者利用对模型的白盒或黑盒访问能力，通过反复查询或逆向分析模型输出与参数，实施成员推断、属性提取或梯度反演等重构训练数据中的个体敏感样本（如人脸图像、医疗记录），导致隐私信息泄露。

(b) 提示词注入攻击。攻击者通过指令覆盖、目标劫持或上下文欺骗等，篡改或绕过模型内嵌的系统提示模板与约束条件（如安全策略），导致模型泄露系统内部指令、执行非预期恶意任务或输出失控内容。

6.4 应用安全风险

6.4.1 人工智能模型服务风险

(a) 模型服务非授权访问。人工智能模型服务采用不安全的身份鉴别或访问控制机制，导致攻击者可能绕过认证、非法调用服务、窃取核心模型参数或越权访问受保护的训练数据与输出结果。

(b) 数据传输风险。人工智能模型客户端与服务端间缺乏有效的安全传输保障，可能导致通信内容被窃听、篡改或伪造等风险。

6.4.2 智能体应用安全风险

(a) 智能体应用非授权访问。智能体应用身份鉴别机制失效或权限分配不当，可能导致身份仿冒、凭证被盗用、智能体被授予权限过高执行越权数据访问、高危系统命令滥用等风险。

(b) 多智能体应用风险。多智能体应用间通信协议不安全或访问控制策略不当，存在协同敏感数据泄露、身份仿冒接入、消息指令被篡改等风险。

(c) 运行环境安全风险。智能体应用在集成外部组件或工具时，未采用安全的访问控制与可信管控、运行环境隔离机制，或使用过度代理权或不安全的函数调用（如缺乏严格的输入验证、沙箱隔离和用户确认机制），可能导致恶意代码执行、数据泄露等风险。

(d) 规划和控制安全风险。未受保护的智能体应用无法验证任务规划合规性与指令可信性，存在高危操作滥用、指令伪造篡改与行为不可否认性缺失的风险。

(e) 监控与审计安全风险。智能体应用未采用日志完整性保护、未对高权限行为实施“人-智能体”协同审查，存在日志被篡改、高权限操作滥用、恶意行为无法追溯等风险。

7 密码应用技术框架

生成式人工智能系统的密码应用技术框架如图 2 所示。

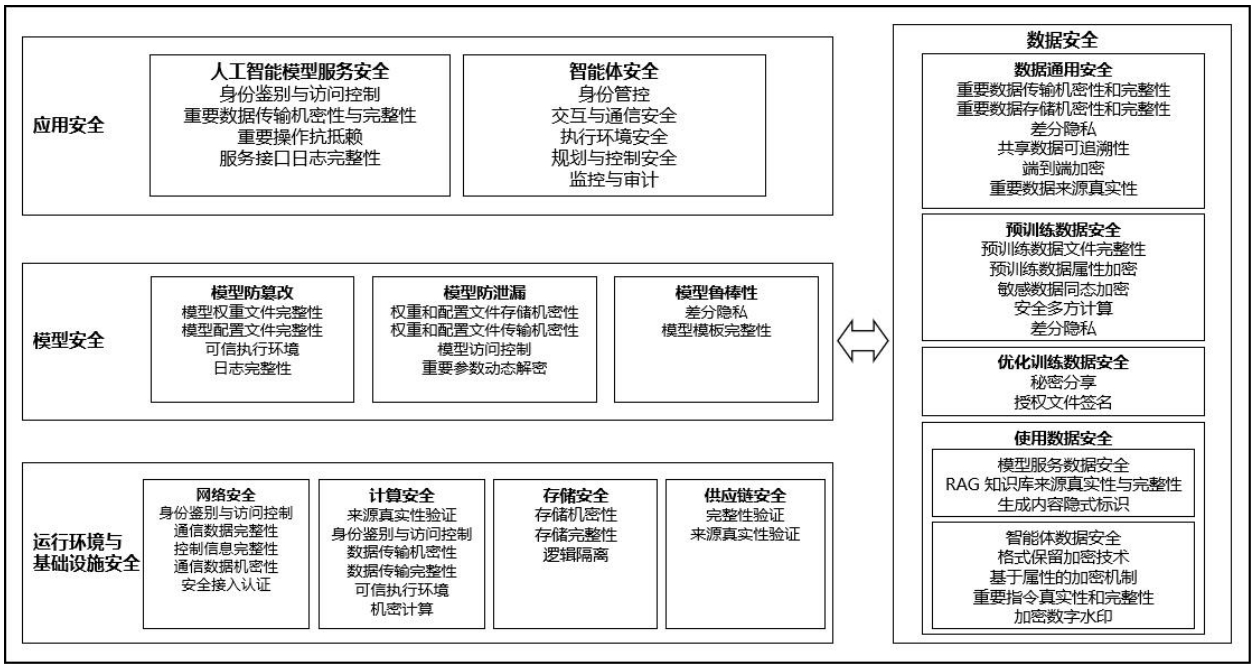

图2 生成式人工智能系统密码应用技术框架

生成式人工智能系统密码应用技术框架由运行环境与基础设施安全、数据安全、模型安全、应用安全四部分组成。

(a) 运行环境与基础设施安全包含网络安全、计算安全、存储安全、供应链安全等密码应用措施，保证模型运行环境与基础设施的安全。

(b) 数据安全包含数据通用安全、预训练数据安全、优化训练数据安全、使用数据安全等密码应用措施，保证模型及应用的重要数据及敏感个人信息安全。

(c) 模型安全包含模型防篡改、模型防泄漏、模型鲁棒性等密码应用措施，保证模型自身的安全。

(d) 应用安全包含调用人工智能模型服务的密码应用措施，以及覆盖身份鉴别与访问控制、通信协议安全、运行环境安全、规划和控制安全、监控与审计等方面的智能体应用密码应用措施。

根据生成式人工智能系统的特点和密码技术的应用现状，本文件对生成式人工智能系统密码防护措施进行了分层，如图3所示。在信息系统密码应用基本要求之上，结合业务需要构建适配生成式人工智能系统特点的分层密码防护体系，以更好地指导密码技术在生成式人工智能系统中的应用落地。

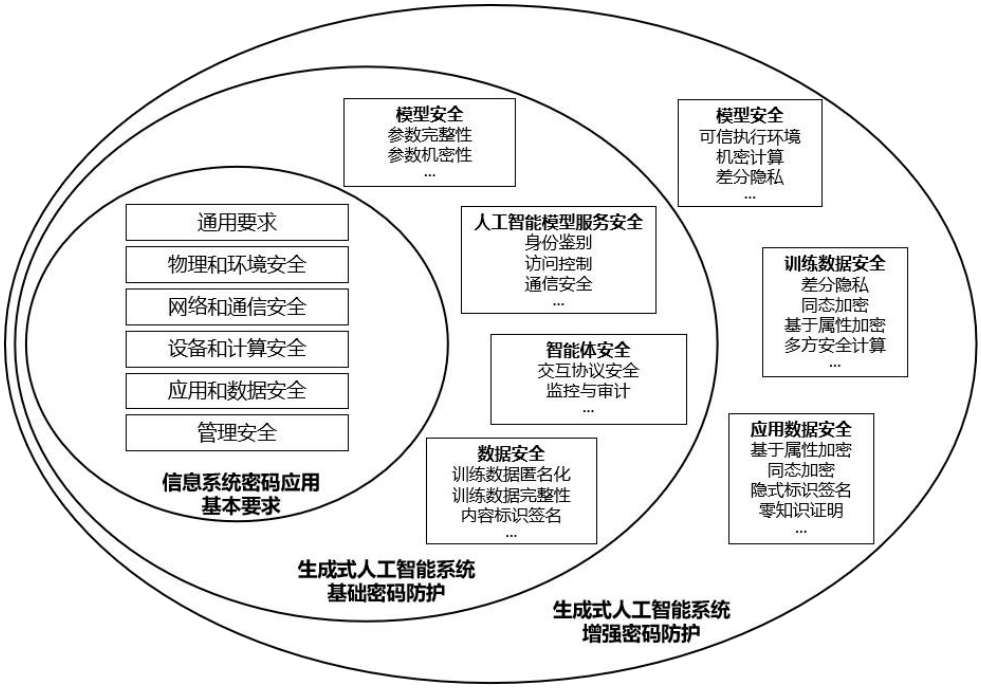

图3 生成式人工智能系统密码防护措施

为直观落地本指引的密码应用技术要求，附录A从训练系统、应用系统和智能体应用给出生成式人工智能系统的应用示例，附录B从人工智能模型的语料准备、模型训练、模型部署与推理及模型更新与退役等四个关键阶段给出了全生命周期密码应用示例。

8 密码应用指引

8.1 基本安全要求

生成式人工智能系统在物理和环境安全、网络和通信安全、设备和计算安全、管理要求等方面的密码应用基本要求参考GB/T 39786-2021，本文件8.2章节仅描述运行环境与基础设施安全保护对象，8.3至8.5章节描述生成式人工智能系统相关新增密码要求。信息系统密码应用基本要求与本文件安全措施的对应关系参考附录C。

应用本文件提出的密码措施时，需验证其实现的正确性与有效性。

8.2 运行环境与基础设施安全

8.2.1 网络安全保护对象

采用密码技术对人工智能系统客户端与服务端间通信、人工智能分布式部署节点间通信等网络通信链路的身份鉴别、机密性、完整性、安全接入认证等进行保护。

8.2.2 计算安全保护对象

采用密码技术对支撑模型运行的设备与硬件资源、操作系统、运行时依赖等来源真实性、身份鉴别、机密性、完整性等进行保护，并对用于数据处理、模型训练和模型推理等不同目的的计算资源进行安全隔离。

对于高安全需求场景，可结合机密计算技术保护模型与数据在训练和推理阶段的安全性，具体可参考 GB/T 45230-2025。

8.2.3 存储安全保护对象

采用密码技术对模型及应用存储的重要数据和敏感个人信息等的身份鉴别、机密性、完整性等进行保护；对多业务线或不同安全域的存储资源，采用独立密钥加密等密码技术实现逻辑隔离保护，防止重要数据和敏感个人信息泄露；对安全性标注数据可进行隔离存储。

8.2.4 供应链安全保护对象

采用密码技术对模型所使用的训练框架、第三方依赖库、预训练权重及推理引擎等组件的完整性及来源真实性进行保护；对模型所使用的人工智能加速处理器及加速模组、训练服务器及专用计算设备的固件的完整性及来源真实性进行保护，对组件、固件更新及升级前验证其完整性及来源真实性。

8.3 数据安全

8.3.1 通用安全

通用安全措施可采用：

(a) 对采集和生成的个人敏感数据（如姓名、身份证号等用户数据），采用加密或脱敏的方式处理，也可结合差分隐私、格式保留加密等技术进行处理，防止隐私泄露；

(b) 在共享数据中嵌入加密数字水印，通过提取和验证数字水印追溯泄露节点；

(c) 对数据提供方的数据采用数字签名技术进行保护，保障数据来源的真实性和完整性；

(d) 采用密码技术对存储的敏感数据进行机密性和完整性保护；

(e) 采用消息鉴别码、数字签名等密码技术对数据分类分级的标识进行完整性防护。

8.3.2 预训练数据安全

预训练阶段的数据安全措施可采用：

(a) 数据被模型纳入训练集之前，每个数据文件采用消息鉴别码或数字签名等密码技术进行完整性验证，保障数据使用前未被篡改；

(b) 对敏感程度不同的预训练数据（如核心商业数据、普通公开数据）可结合基于属性的加密，允许满足属性条件的用户进行解密；

(c) 数据清洗、特征提取等处理环节，对敏感数据（如用户行为特征、企业核心指标）可结合同态加密，防止处理节点被入侵导致中间数据泄露；

(d) 训练数据分布在多个机构且无法集中时，可结合多方安全计算进行联合训练，防止训练数据泄露；

(e) 在训练算法的关键步骤中可结合差分隐私技术引入噪声，保障无法推断出某个特定个体是否在训练集中，更无法重构出原始数据。

8.3.3 优化训练数据安全

优化训练阶段的数据安全措施可采用：

(a) 分布式微调场景中，可结合秘密共享机制将训练数据梯度拆分为多个份额，梯度更新时无需暴露完整梯度，防止反推原始梯度信息；

(b) 对于行业敏感数据的采集授权文件，采用数字签名等密码技术确认授权有效性。

8.3.4 使用阶段数据安全

8.3.4.1 模型服务数据安全

使用阶段的数据安全措施可采用：

(a) 采用数字签名等密码技术对接入RAG的外部知识库进行保护，保证数据来源的真实性和完整性；

(b) 采用消息鉴别码或数字签名等密码技术对内容生成日志（如生成者、时间、内容、分类等）进行完整性保护，防止日志被篡改；

(c) 采用数字签名等密码技术对人工智能模型服务生成的文本、图像、音频或视频等生成合成内容文件元数据嵌入隐式标识，保障生成内容的可追溯性；

(d) 在合成内容文件元数据嵌入隐式标识应符合 GB 45438-2025 的要求。

8.3.4.2 智能体数据安全

智能体应用使用阶段的数据安全措施可采用：

(a) 采用密码技术的加解密功能对智能体的技能包（Skills）、配置数据、长期记忆、上下文窗口及RAG数据库中敏感数据进行加密存储，确保只有获得授权的智能体实例能解密检索，以防止敏感数据泄露、内存污染或上下文污染；采用基于密码技术的访问控制及隔离机制，对智能体的长期记忆、上下文窗口及RAG数据库实施读写控制、记忆内容验证、会话隔离等操作，仅允许已验证的数据源写入，并定期清理未验证的条目记录；

(b) 采用消息鉴别码或数字签名等密码技术保护智能体的访问控制策略、权限规则及授权凭证等关键安全信息的完整性与真实性，防止其在存储或传输过程中被未授权篡改、注入或破坏；

(c) 采用数字签名等密码技术对智能体的关键操作指令、决策指令和对外发布的重要信息等进行签名，以确保其真实性和完整性；

(d) 采用密码技术对智能体处理的敏感数据集合嵌入加密数字水印。当发生数据泄露时，可通过提取水印信息追溯泄露源头（如具体是哪个工具接口或哪个协作智能体节点），明确安全责任边界；

(e) 可结合格式保留加密等密码技术，对智能体在调用外部工具或插件时传递的包含敏感数据（如个人隐私、智能体私钥、API密钥、访问令牌或调用凭证）的参数字段进行加密，实现敏感数据的“最小化披露”与“可用不可见”；

(f) 对多智能体系统中智能体间需协同完成任务的共享敏感数据，可采用基于属性的加密（ABE）或多方安全计算（SMPC）等密码技术，确保数据在访问和联合计算过程中的安全。

8.4 模型安全

8.4.1 模型防篡改

模型防篡改的安全措施可采用：

(a) 采用消息鉴别码或数字签名等密码技术对模型配置文件和权重文件进行完整性保护，在模型分发、下载、部署、启动和更新时验证模型配置文件和权重文件的完整性，防止模型配置文件和权重文件被篡改；

(b) 可结合可信执行环境，为模型运行提供硬件隔离的可信区域，保障模型推理、梯度计算等都在可信区域中进行；

(c) 采用消息鉴别码或数字签名等密码技术对模型全生命周期日志（如下载、调用、微调、导出等）进行完整性保护，防止日志被篡改。

8.4.2 模型防泄露

模型防泄露的安全措施可采用：

(a) 在模型配置文件和权重文件存储时使用密码技术的加解密功能进行机密性保护，防止模型配置文件和权重文件泄露；

(b) 在模型分发、下载时使用密码技术的加解密功能，对模型配置文件和权重文件进行传输机密性保护，防止传输过程中模型配置文件和权重文件泄露；

(c) 模型加载过程使用内存加密等密码技术，模型配置文件、权重文件不以明文形式运行在内存，防止攻击者直接复制或反编译模型文件；

(d) 对模型调用接口设置身份认证与访问控制机制，使模型操作（如下载、部署、更新）遵循最小权限原则，防止模型通过接口泄露。

8.4.3 模型鲁棒性

模型鲁棒性安全措施可采用：

(a) 采用消息鉴别码或数字签名等密码技术对模型的系统提示模板进行完整性保护，用于对模型的输入进行结构完整性检查，防御对抗性攻击；

(b) 模型训练或推理中可结合差分隐私机制，对梯度或输出中添加噪声，限制单个样本对输出的影响，防御模型反演攻击与成员推理攻击。

8.5 应用安全

8.5.1 人工智能模型服务安全

人工智能模型服务的安全措施可采用：

(a) 采用数字签名等密码技术对调用人工智能模型服务的用户、接口或系统进行身份鉴别，采用基于角色或身份的访问控制策略，遵循最小权限原则，防止非授权访问；

(b) 在人工智能客户端与服务端之间采用密码技术建立安全传输通道，保障关键交互数据、模型输出结果或系统指令等数据的机密性和完整性；

(c) 对重要的人工智能模型服务操作（如模型参数查询、大规模数据导出、管理配置变更等）采用数字签名等密码技术实现抗抵赖，确保操作行为的不可否认性；

(d) 采用密码技术对人工智能模型服务的API令牌等凭证进行存储机密性和完整性保护，防止因凭据泄露导致的安全风险；

(e) 采用消息鉴别码或数字签名等密码技术对人工智能模型服务接口的调用日志进行完整性保护，记录并监控所有访问请求和操作行为，及时发现并告警异常访问模式或潜在攻击行为。

8.5.2 智能体安全

8.5.2.1 身份鉴别与访问控制

身份鉴别与访问控制安全措施可采用：

(a) 采用基于密码技术的身份管理机制，为注册智能体应用分配唯一、可验证的身份标识，并签发数字证书作为智能体应用数字身份凭证，通过身份管理机制防止智能体应用身份仿冒与凭证盗用，并实现对智能体应用数字身份的全生命周期管理；

(b) 采用基于密码技术的身份鉴别机制，实现智能体应用与用户、智能体应用之间、智能体应用与第三方工具的双向身份鉴别，防止非授权访问，实现多智能体应用系统中通信双方的身份鉴别，防止非授权智能体应用接入系统；

(c) 对智能体应用权限执行动态管理与访问控制，采用密码技术生成具有时效性和特定范围的即时访问令牌或临时凭证，确保即使智能体应用被劫持等场景下，攻击者也只能获得受限的、可撤销的权限；具有执行高权限行为如访问敏感工具或敏感数据的智能体应用应遵循最小权限原则，防止权限滥用。

8.5.2.2 交互与通信安全

通信协议的安全措施可采用：

(a) 采用密码技术在智能体应用与用户、智能体应用之间、智能体应用与第三方工具/API间建立安全传输通道，保障传输过程中数据的机密性和完整性；

(b) 采用SSH等密码技术建立专用安全管理通道，对智能体应用运行资源（如服务器、容器、虚拟机）进行远程管理，保障管理通道的双向身份鉴别、数据机密性和完整性；

(c) 采用密码技术对MCP等智能体应用交互协议的指令消息和响应消息进行机密性和完整性保护，防止消息在交互过程中被篡改、注入或敏感信息泄露。

8.5.2.3 执行环境安全

运行环境的安全措施可采用：

(a) 采用消息鉴别码或数字签名等密码技术，对智能体应用运行所依赖的操作系统、运行框架、插件、工具、脚本、配置文件、策略文件等进行来源真实性及完整性保护，在部署、启动、更新和动态加载时进行验证，防止运行环境被篡改或植入恶意组件；

(b) 采用数字签名等密码技术对智能体应用调用的外部工具、插件、函数接口、代码执行环境等建立身份鉴别和授权校验机制，确保智能体应用仅能调用经过授权和可信验证的运行资源，防止不可信工具接入导致远程代码执行、数据泄露或系统破坏；

(c) 对智能体应用执行环境采用可信执行环境等隔离机制，限制智能体应用对主机系统、文件系统、网络资源、数据库及外部服务的访问范围，防止智能体应用越权访问运行环境中的重要资源；

(d) 可结合内存加密技术对智能体应用运行过程中产生的中间文件、缓存数据、上下文状态、任务队列和工具调用结果等进行保护，并在任务完成或权限失效后及时清理，防止运行态数据泄露、篡改或重放。

8.5.2.4 规划和控制安全

规划和控制的安全措施可采用：

(a) 智能体应用生成的任务规划（Plan）在执行前，采用数字签名等密码技术生成规划的可验证凭证。验证方（如管理节点或用户）无需了解规划细节，即可验证该规划是否符合预设的安全策略（如是否包含高危操作），确保智能体应用决策逻辑的可信执行，可结合零知识证明生成该规划的可验证凭证；

(b) 对于来自外部用户或管理端的控制指令（如任务终止、权限升降级、工具调用授权），采用数字签名等密码技术或预共享密钥建立身份鉴别机制。对高敏感度的控制指令，要求发送方进行数字签名，确保指令来源的真实性、指令内容的完整性以及操作行为的不可否认性。

8.5.2.5 监控与审计

监控与审计的安全措施可采用：

(a) 采用消息鉴别码或数字签名等密码技术对智能体应用的运行日志、工具调用日志、异常行为记录、安全告警信息和资源访问记录等进行完整性保护，记录并监控所有访问请求和操作，及时发现异常行为或违背安全策略的操作；

(b) 监控智能体应用高权限操作、执行未授权操作企图等异常行为，采用数字签名等密码技术建立人机协同控制机制；对高权限敏感操作须经人工核查确认，由授权人员对操作确认指令进行数字签名，确保操作来源真实、行为不可抵赖且全程可追溯。

附　录　A

（资料性附录）

生成式人工智能系统密码应用示例

A.1 生成式人工智能训练系统

生成式人工智能训练系统是指支撑大模型开展数据归集、样本标注、特征处理、分布式训练及参数迭代的全流程基础平台。典型系统包含训练数据仓库、GPU算力集群、参数服务器、任务调度编排、密码安全支撑及运维管控等核心模块。系统承载海量原始训练数据与模型参数，涉及数据流转、算力调度、任务运维等多环节，是人工智能大模型研发的核心底座，承载大量高敏感训练资源，整体安全防护需求严苛。

生成式人工智能训练系统安全架构采用分层防护模式，构建全方位安全屏障。密码基础设施层作为信任根，提供密钥管理、加解密等基础支撑；数据层通过加密存储与签名验签，保障训练数据安全；生成式人工智能训练资源层防护数据在算力节点间流转的机密性与完整性；调度服务层实现调度指令加密与节点可信认证，防范恶意攻击；管理运维层通过身份鉴别与操作审计，杜绝越权与恶意操作。生成式人工智能训练系统的密码应用架构如图A.1所示。

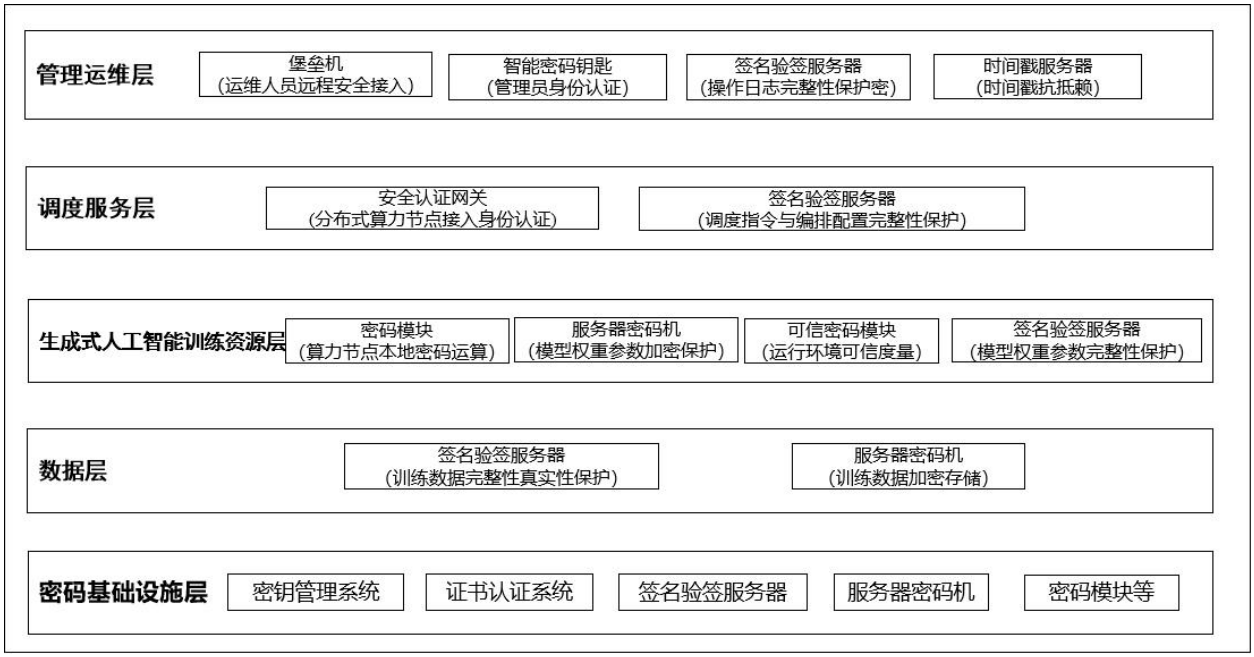

图 A.1 生成式人工智能训练系统密码应用架构

(a) 密码基础设施层：作为整个系统的信任根，部署证书认证密钥管理系统、服务器密码机、时间戳服务器、安全认证网关等密码设备，统一实现密钥全生命周期管理、加解密运算、可信时间戳、安全通信链路，为上层安全链路提供密码支撑。

(b) 数据层：通过签名验签服务器对每份训练数据文件进行来源真实性与完整性校验，防范数据投毒风险；服务器密码机负责原始语料、标注数据、特征数据的加密存储。

(c) 生成式人工智能训练资源层：签名验签服务器负责模型配置文件与权重文件的完整性校验，确保加载的模型文件未被篡改；服务器密码机负责模型权重、梯度信息的加密保护；训练节点内嵌可信密码模块，结合可信执行环境为训练任务提供硬件级隔离运行空间，保障训练数据、中间计算结果在内存运算过程中的机密性；密码模块部署于GPU训练节点，提供本地高性能密码运算能力，降低分布式训练中密码计算的网络开销。

(d) 调度服务层：通过安全认证网关实现分布式算力节点的安全接入认证，为合法节点签发身份凭证；签名验签服务器对任务调度指令、容器编排配置进行数字签名保护，防止指令被篡改或恶意任务注入。

(e) 管理运维层：通过堡垒机建立运维远程安全接入通道，管理员使用智能密码钥匙存储身份证书，结合多因素认证机制实现强身份鉴别；签名验签服务器对训练全生命周期操作日志进行完整性保护，结合可信时间戳服务器为审计记录附加可信时间，确保操作行为可追溯、不可篡改。

A.2 生成式人工智能应用系统

生成式人工智能应用系统是指面向终端用户提供智能对话、内容生成、知识问答、代码辅助等服务的完整业务系统。典型的应用系统通常包括前端用户交互界面、API网关、模型推理服务、RAG知识检索模块、以及后端数据存储等核心组件。在政务、金融、医疗、能源等关键行业中，此类系统处理大量敏感数据，对安全性要求极高。

生成式人工智能应用系统层面的密码安全需覆盖身份鉴别、通信保护、数据机密性与完整性、访问控制、操作抗抵赖以及生成内容溯源等维度。图A.2以一个典型的生成式人工智能应用系统为例，阐述密码技术在应用系统全链路中的部署实践。

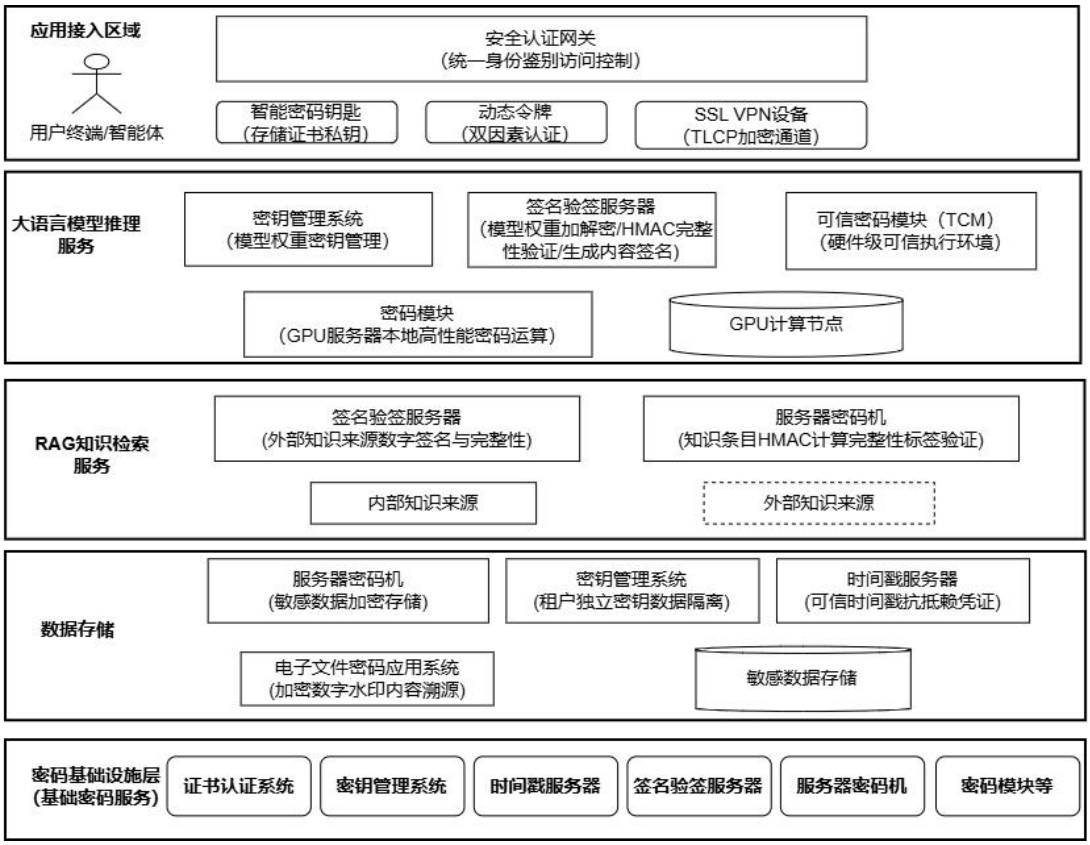

图 A.2 生成式人工智能应用系统密码应用架构

(a) 密码基础设施层：提供证书签发、密钥管理、可信时间戳、签名验签四类基础密码服务；应用系统层按功能划分为用户接入、模型推理、知识检索、数据存储五个区域，各区域部署相应密码产品实现安全防护；各区域与密码基础设施之间通过密码设备应用接口对接，形成统一的密码服务能力；采用密码技术对推理引擎、依赖库等组件的完整性及来源真实性进行验证。

(b) 数据存储：通过服务器密码机对敏感数据加密存储，密钥管理系统为多业务线或不同安全域分配独立密钥实现逻辑隔离保护。生成内容经电子文件密码应用系统嵌入加密数字水印实现来源可追溯；

重要操作由签名验签服务器附加数字签名、时间戳服务器附加可信时间戳，形成抗抵赖的操作凭证；对安全性标注数据实施隔离存储；采用密码技术对API令牌等凭证进行安全存储。

(c) RAG知识检索服务：通过签名验签服务器验证外部知识来源的数字签名，确认数据提供方身份和内容完整性；知识条目写入数据库时由服务器密码机计算完整性标签，检索结果提交模型前由服务器密码机验证标签匹配性。

(d) 大语言模型推理服务：签名验签服务器负责模型文件的完整性验证和生成内容签名；服务器密码机负责模型权重文件加密存储、解密加载以及提示词模板完整性保护，可采用差分隐私机制对模型输出添加噪声以防御模型反演攻击；加密密钥由密钥管理系统统一管理；基于国产处理器的硬件级可信执行环境，以及平台内嵌的可信密码模块（TCM）保障推理数据的安全和隐私；密码模块接入GPU服务器提供本地高性能密码运算能力，降低密码计算的网络延迟。

(e) 应用接入区域：通过安全认证网关实现统一身份鉴别和访问控制，签名验签服务器完成令牌验证和日志签名，调用时间戳服务器为审计记录附加可信时间戳；通信链路由SSL VPN设备建立TLCP加密通道，用户终端或智能体通过智能密码钥匙存储证书和私钥，结合动态令牌实现双因素认证；可采用差分隐私、格式保留加密等技术对采集的个人敏感数据进行处理。

A.3 生成式人工智能智能体应用

生成式人工智能智能体应用是指具备自主感知、规划决策、工具使用及多智能体协作能力的智能软件系统。在复杂的应用场景中，智能体不仅需要与用户进行安全交互，还需在多智能体间建立可信的通信链路，并安全地调用外部工具或API。

生成式人工智能智能体应用的密码应用架构主要由智能体交互接入层、智能体大脑与规划层、工具调用与执行层、记忆与知识管理层以及统一的密码基础设施层组成。各层级通过标准密码服务接口调用底层密码资源，实现全链路的安全防护。图A.3示例重点阐述如何利用密码技术构建智能体的身份管控体系、规划执行安全机制及工具调用防护，以应对身份仿冒、越权操作及敏感数据泄露等风险。

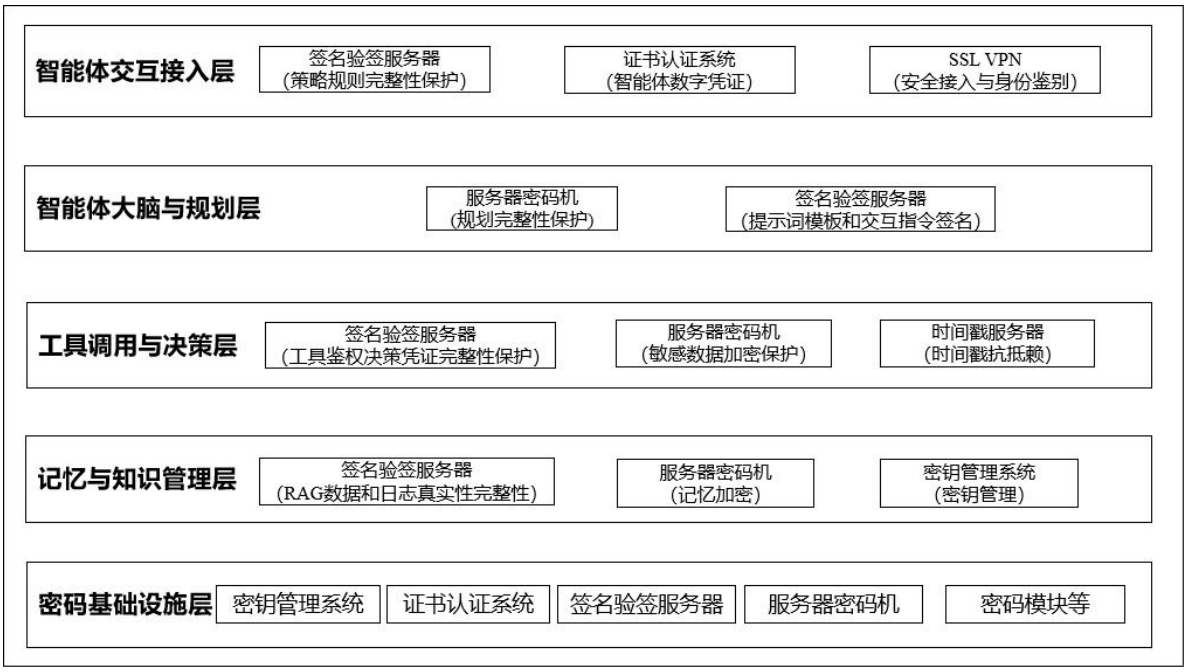

图 A.3 生成式人工智能智能体密码应用架构

(a) 智能体交互接入层：该层基于证书认证系统，为每个智能体签发唯一的SM2数字证书，建立基于数字证书的双向身份认证机制；通过SSL VPN建立TLCP通道，实现智能体与用户、智能体与智能体

之间的安全接入与身份鉴别，防止非授权实体接入系统；使用签名验签服务器保护访问控制策略及权限规则的完整性以及保护智能体的访问控制策略、权限规则及授权凭证等关键安全信息的完整性与真实性。

(b) 智能体大脑与规划层：通过签名验签服务器对系统提示词模板及生成的任务规划进行完整性保护，防止决策逻辑被篡改，并结合可信时间戳，确保决策行为的抗抵赖性与可追溯性；使用服务器密码机对智能体应用生成的任务规划进行完整性保护。

(c) 工具调用与执行层：通过签名验签服务器对外部工具、插件、函数接口进行身份鉴权与完整性校验，确保智能体仅能调用经过授权的可信运行资源；服务器密码机对工具调用参数中的敏感字段加密保护，防止调用过程中隐私泄露，采用消息鉴别码（MAC）机制对智能体与工具间的交互指令和响应消息、智能体应用的运行日志、工具调用日志、异常行为记录、安全告警信息及其他操作行为、资源访问记录等进行完整性保护；对智能体执行的敏感数据访问、工具调用等关键操作指令，采用数字签名技术进行记录，并附加可信时间戳，确保操作行为的不可否认性与可审计性。

(d) 记忆与知识管理层：通过签名验签服务器对RAG外部知识库的数据源进行数字签名验证，确认知识来源真实性与内容完整性；服务器密码机对智能体技能包、配置数据、长期记忆、上下文窗口及RAG数据库中的敏感数据进行加密存储；密钥管理系统为不同安全级别的记忆数据分配独立密钥，实现逻辑隔离与读写权限管控；使用签名验签服务器对智能体运行日志、工具调用日志采用数字签名进行完整性保护，满足安全审计与事后追溯需求。

(e) 密码基础设施层：作为整个架构的信任根，部署证书认证密钥管理系统、证书认证系统、签名验签服务器及时间戳服务器，统一提供密钥全生命周期管理、数字证书签发、数据完整性校验及可信时间源等基础密码服务；采用密码技术对训练框架、依赖库、推理引擎等组件、芯片模组及设备固件的完整性及来源真实性进行验证，组件、固件更新及升级前验证其完整性及来源真实性；在部署、启动、更新和动态加载时，对智能体应用运行所依赖的操作系统、运行框架、插件、工具、脚本、配置文件、策略文件等，采用消息鉴别码或数字签名等密码技术进行完整性及来源真实性保护。

附 录 B
（资料性附录）
人工智能模型全生命周期密码应用示例

人工智能模型的全生命周期通常包括语料准备、模型训练、模型部署与推理、模型更新与退役四个关键阶段。数据作为其核心要素，贯穿在全生命周期的各个阶段，是人工智能模型安全防护的关键对象。密码技术为数据安全和模型安全提供了有效的保障。

各阶段涉及的主要密码技术、密码产品及关键安全活动如图B.1所示。

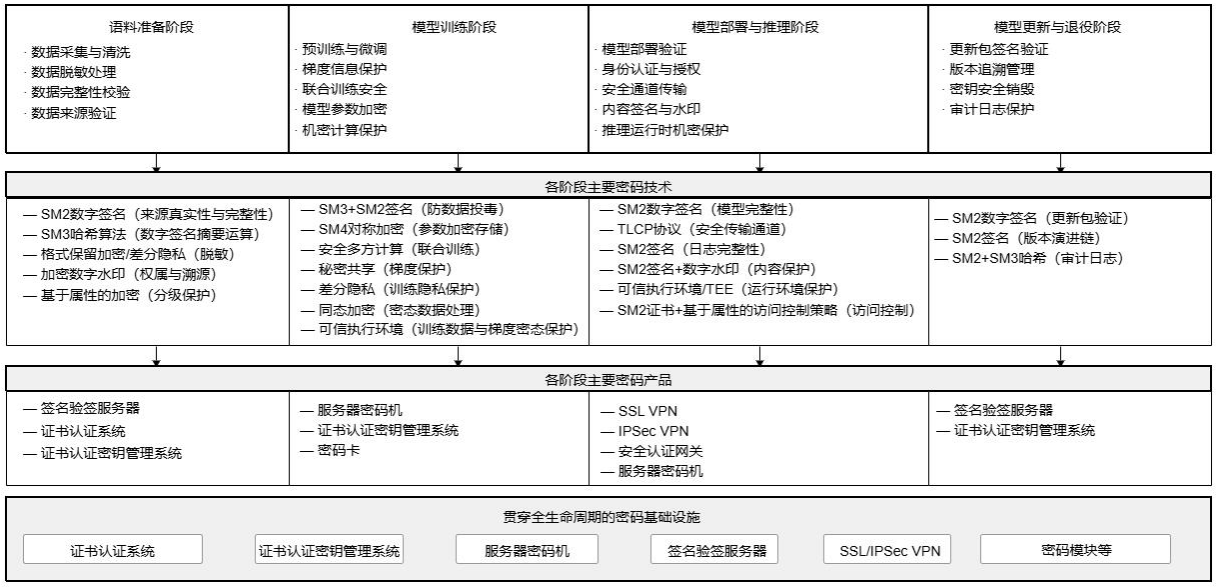

图 B.1 人工智能模型全生命周期密码应用框架

B.1 语料准备阶段

语料准备阶段的密码应用要点如下：

(a) 采用SM2数字签名技术对数据提供方的原始数据进行签名，保障数据来源的真实性；数据使用前对每个数据文件进行签名验证，确保数据在采集、传输和存储过程中未被篡改。

(b) 采用密码技术对存储的敏感语料数据进行机密性和完整性保护，对包含个人敏感信息的语料数据，可采用格式保留加密（FPE）或差分隐私技术进行脱敏处理，在保留数据统计特性的同时防止隐私泄露。

(c) 在共享数据中嵌入基于密码技术的加密数字水印，用于标识数据和追溯泄露节点。

(d) 采用消息鉴别码或数字签名等密码技术对数据分类分级的标识进行完整性防护，防止标识被篡改；对安全性标注数据采用隔离存储措施，与普通训练数据实施逻辑隔离。

(e) 对敏感程度不同的预训练数据，可采用基于属性的加密技术进行分级保护，通过密钥管理系统实现密钥的统一管理和分发。

本阶段主要使用的密码产品包括：签名验签服务器、证书认证系统、证书认证密钥管理系统等。

B.2 模型训练阶段

模型训练阶段的密码应用要点如下：

(a) 数据被纳入训练集前，采用SM2数字签名技术对每个数据文件进行完整性和来源真实性验证，防止数据投毒攻击。

(b) 分布式训练场景中，采用秘密共享机制将梯度信息拆分为多个份额，梯度更新时无需暴露完整梯度，防止通过梯度反推原始训练数据。

(c) 采用SM4对称加密对模型配置文件和权重文件进行加密存储，通过服务器密码机提供密钥保护和加解密服务。

(d) 在模型训练过程中，将用于数据处理和模型训练等不同目的的计算资源进行安全隔离，采用可信执行环境（TEE）技术为训练任务提供硬件隔离的安全计算环境，保障训练数据、梯度信息和中间计算结果在主机内存及算力资源中的机密性，防止未授权访问和侧信道攻击。

(e) 对模型训练全生命周期日志（包括数据下载、训练启动、微调、导出等操作）采用消息鉴别码或数字签名等密码技术进行完整性保护，防止日志被篡改。

(f) 数据清洗、特征提取等处理环节，对敏感数据可采用同态加密技术进行密态处理，防止处理节点被入侵导致中间数据泄露。

(g) 训练数据分布在多个机构且无法集中时，可采用多方安全计算技术进行联合训练，各参与方数据不出域，在密态下完成联合训练。

(h) 在训练算法关键步骤中可引入差分隐私机制，对梯度添加噪声，确保无法推断特定个体是否在训练集中。

本阶段主要使用的密码产品包括：服务器密码机、证书认证密钥管理系统、密码卡等。

B.3 模型部署与推理阶段

模型部署与推理阶段的密码应用要点如下：

(a) 模型部署前，采用SM2数字签名技术对模型配置文件和权重文件进行完整性和来源真实性验证，防止模型文件在分发、下载和部署过程中被篡改或替换；采用密码技术对训练框架、依赖库、推理引擎等组件、芯片模组及设备固件的完整性及来源真实性进行验证。

(b) 采用基于SM2数字证书的身份认证机制，对调用人工智能服务的用户、智能体和系统进行身份鉴别，结合基于属性的访问控制策略实现细粒度访问控制，遵循最小权限原则；采用密码技术对人工智能模型服务的API令牌等凭证进行安全存储，防止凭据泄露。

(c) 在客户端与模型服务端之间通过TLCP协议建立安全传输通道保障交互数据的机密性和完整性；模型分发、下载时使用密码技术的加解密功能，对模型配置文件和权重文件进行传输的机密性保护，防止传输过程中模型文件泄露。

(d) 对人工智能服务生成的内容采用SM2数字签名，保证生成内容在传输和存储过程中未被篡改，并在生成的文本、图像、音频或视频中嵌入加密数字水印保障内容可追溯；对模型的系统提示模板使用消息鉴别码或数字签名等密码技术进行完整性保护，防御对抗性攻击；接入RAG的外部知识库采用数字签名等技术保证数据来源的真实性和完整性。

(e) 模型运行过程中采用可信执行环境（TEE），确保模型配置文件和权重文件不以明文形式在内存中运行，防止侧信道攻击和内存提取攻击；对多业务线或不同安全域的存储资源，采用独立密钥加密等密码技术实现逻辑隔离保护，防止重要数据和敏感个人信息泄露；对安全性标注数据实施隔离存储。

(f) 采用消息鉴别码或数字签名等密码技术对操作日志进行完整性保护，记录并监控所有访问请求和操作行为，确保日志不可篡改和可审计；对重要的人工智能模型服务操作采用数字签名技术实现抗抵赖。

本阶段主要使用的密码产品包括：SSL VPN、IPSec VPN、安全认证网关、服务器密码机等。

B.4 模型更新与退役阶段

模型更新与退役阶段的密码应用要点如下：

(a) 模型更新包在分发前采用SM2数字签名进行签名，接收方在安装更新前验证签名以确认更新包的完整性和来源真实性，防止恶意更新注入；更新包传输过程中采用密码技术的加解密功能保障机密性，防止更新包在传输过程中被截获或篡改。对训练框架、依赖库、推理引擎等组件、芯片模组及设备固件更新，在升级前验证其完整性及来源真实性。

(b) 建立基于SM2数字签名的模型版本追溯机制，对每个模型版本的配置文件、权重文件和变更记录进行签名，形成可验证的版本演进链。

(c) 模型退役时，通过证书认证密钥管理系统执行密钥安全销毁流程，彻底清除模型加密密钥和关联凭证，确保退役模型的参数数据无法被恢复；对智能体应用的退役，应同步撤销其数字证书和访问令牌，清理其长期记忆、上下文窗口及RAG数据库中的关联数据；对退役模型的访问权限实施基于密码技术的访问控制，确保退役模型不再被任何主体调用或访问。

(d) 保留退役模型的操作审计日志，审计记录采用SM2数字签名和SM3哈希保护，满足合规审计和事后追溯需求。

本阶段主要使用的密码产品包括：签名验签服务器、证书认证密钥管理系统等。

注：本附录中列出的密码技术和密码产品为生成式人工智能系统各生命周期阶段的典型应用，具体选用应根据系统安全等级要求和实际业务场景确定。

附 录 C
（资料性附录）
信息系统密码应用基本要求与安全措施映射表

指标体系			安全措施
技术要求	物理和环境安全	-	8.1
	网络和通信安全	身份鉴别	8.2.1、8.5.2.2
		通信数据完整性	8.2.1、8.5.2.2
		通信过程中重要数据的机密性	8.2.1、8.5.2.2
		网络边界访问控制信息的完整性	8.2.1、8.5.2.2
		安全接入认证	8.2.1、8.5.2.2
		密码服务	8.1
		密码产品	8.1
	设备和计算安全	身份鉴别	8.2.2、8.5.2.3
		远程管理通道安全	8.2.2
		系统资源访问控制信息完整性	8.2.2、8.5.2.3
		重要信息资源安全标记完整性	8.2.2、8.5.2.3
		日志记录完整性	8.2.2、8.4.1、8.5.2.5
		重要可执行程序完整性、重要可执行程序来源真实性	8.2.2、8.2.4、8.4.1、8.5.2.3
		密码服务	8.1
		密码产品	8.1
	应用和数据安全	身份鉴别	8.4.2、8.5.1、8.5.2.1 、8.5.2.4、8.5.2.5
		访问控制信息完整性	8.4.2、8.5.1、8.5.2.1
		重要信息资源安全标记完整性	8.3.1、8.5.2.4
		重要数据传输机密性	8.3.1、8.3.3、8.3.4、8.4.2、8.5.2.4
		重要数据存储机密性	8.2.3、8.3.1、8.3.2、8.3.3、8.3.4、8.4.2、8.5.2.4
		重要数据传输完整性	8.3.1、8.3.4、8.4.2、8.5.2.4
		重要数据存储完整性	8.2.3、8.3.1、8.3.2、8.3.4、8.4.1、8.4.2、8.4.3、8.5.2.4、8.5.2.5
		不可否认性	8.3.3、8.3.4、8.5.2.4
		密码服务	8.1
		密码产品	8.1
管理要求	-	-	8.1

注：物理与环境安全和管理要求部分按照信息系统要求进行建设。