AIBOM在运营商大模型软件安全治理中的应用与发展研究

张侃[1]，王诗雨[2]，朱沛潼[2]，徐帅健妮[2]，殷铭[2]，何国锋[2]
（1. 中国电信集团有限公司网络和信息安全管理部，北京 100140;
2. 中国电信股份有限公司研究院，上海 200120）

摘　要：随着人工智能（artificial intelligence，AI）技术在电信行业的快速应用，大模型软件安全作为“底座安全”，治理需求日益凸显。人工智能物料清单（artificial intelligence bill of material，AIBOM）作为软件物料清单（software bill of material，SBOM）的延伸，能够系统化呈现组件、模型与数据等信息，成为保障大模型软件可信与合规的关键抓手。首先，梳理了AIBOM的定义、核心要素以及其与传统SBOM的区别，强调其在大模型软件安全治理中的战略价值。其次，通过行业案例分析，总结了AIBOM落地过程中的技术与治理难点。在此基础上，从电信运营商视角出发，探讨了大模型软件安全治理的独特挑战，并阐述了AIBOM在应对AI资产透明度、跨部门协同和合规监管等问题上的核心作用。最后，以中国电信实践为例，展示了其在标准政策、工具链与风险管控方面的探索经验，并提出了未来演进方向。

关键词：AIBOM；人工智能安全；电信运营商；软件安全治理
中图分类号：TN915.08；TP309
文献标志码：A
doi: 10.11959/j.issn.1000−0801.DXKX250538

Research on application and development of AIBOM in large-scale model software security governance for telecom operators

Zhang Kan[1], Wang Shiyu[2], Zhu Peitong[2], Xu Shuaijianni[2], Yin Ming[2], He Guofeng[2]
1. Network and Information Security Management Department, China Telecom Group Co., Ltd., Beijing 100140, China
2. Research Institute of China Telecom Co., Ltd, Shanghai 200120, China

Abstract: With the rapid adoption of artificial intelligence (AI) technologies in the telecommunications industry, the security of large-scale model software—serving as the “foundation of safety”—has become an increasingly pressing governance demand. As an extension of the software bill of material (SBOM), the artificial intelligence bill of material (AIBOM) systematically records information related to components, models, and datasets, and has emerged as a key instrument to ensure the trustworthiness and compliance of large-scale AI systems. Firstly, the definition and core elements of AIBOM were reviewed, as well as its distinctions from traditional SBOM, its strategic significance in large-model software security governance was highlighted. Then industry cases were analyzed to summarize technical

收稿日期：2025−09−08；修回日期：2026−04−01
通信作者：王诗雨，wangsy37@chinatelecom.cn

and governance challenges in AIBOM adoption. Building on this, the unique challenges faced by telecom operators in large-model software security governance were explored, and the central role of AIBOM in addressing AI asset transparency, cross-departmental coordination, and regulatory compliance was elaborated. Finally, the practices of China Telecom were drawn on to demonstrate exploratory efforts in standards, toolchains, and risk management, and future research directions for AIBOM development were outlined.

Key words: AIBOM, AI security, telecom operator, software security governance

0 引言

新一代大模型技术的迅猛发展，正在推动人工智能（artificial intelligence，AI）进入全新的阶段，并给各行业带来跨越式的变革[1]。在这一进程中，电信运营商凭借自身在数字化基础设施和业务场景上的独特优势，成为人工智能应用与落地的重要推动力量。相关研究认为，人工智能将长期深刻影响通信网络体系结构、运营模式及工业生态演进[2]。一方面，运营商长期积累了海量、多样的高质量数据和大规模云化的计算资源，为大模型训练与部署提供了坚实基础；另一方面，其庞大的用户群体与丰富的业务场景，为人工智能的商业化探索和生态构建创造了有利条件。因此，领先运营商率先洞察到AI的颠覆性趋势，积极推进数智化转型，并在技术研发和工业布局上取得了诸多实践成果，逐步形成了差异化的竞争优势[3]。

与此同时，随着人工智能系统在各行业的深入应用，其供应链结构日益复杂，由此带来的安全风险也逐步显现，尤其是在生成式内容与自动化决策等场景中，相关风险已成为人工智能安全治理的重要议题[4]。模型、数据、依赖库与运行环境之间的多维耦合，使得人工智能系统的可信性与可治理性成为学术界与工业界共同关注的议题。

国家高度重视人工智能安全，在2023年10月提出《全球人工智能治理倡议》；2024年6月，工业和信息化部、中央网络安全和信息化委员会办公室、国家发展和改革委员会、国家标准化管理委员会四部门联合印发《国家人工智能产业综合标准化体系建设指南（2024版）》，明确将安全治理标准纳入整体体系结构，推动标准体系落地[5]；2024年7月，党的二十届三中全会明确提出完善生成式人工智能发展和管理机制，建立AI安全监管制度；2025年4月，中共中央政治局指出要完善法律法规、伦理准则及风险预警机制，确保AI朝着有益、安全、公平的方向发展[6-7]；2025年6月，国家标准GB/T 45654-2025《网络安全技术 生成式人工智能服务安全基本要求》正式发布，为AI服务提供者、监管部门及第三方评估机构提供了全面的安全技术指南。这些政策明确了针对强大技术的治理与监管力度应同步加强，而中国电信“为强大的技术提供‘向善’的约束”的理念也与此不谋而合。

在人工智能安全治理体系中，内容安全和算法安全等问题侧重于模型行为与逻辑机制的风险控制，而大模型软件安全（即应用或嵌入大模型的软件系统的安全）——覆盖软件供应链可信度、模型组件追溯性、训练环境的安全性和运行合规性——构成整个治理体系的“底座安全”，为大模型的可信运行提供基础性保障，尤其对于运营商人工智能安全治理体系，是不可或缺的重要组成部分。在传统软件供应链安全领域，已有研究与实践表明，通过引入软件物料清单（software bill of material，SBOM）等清单化手段提升软件组成与依赖关系的透明度，是支撑漏洞追踪、风险识别与治理审计的重要基础，为复杂系统的安全治理提供了可行路径[8]。在此基础上，人工智能物料清单（artificial intelligence bill of

material，AIBOM）的概念应运而生，以SBOM为基础，在AI系统中增设模型、数据集、运行环境等特有字段，实现组件透明化、可追溯、可验证，也为中国电信提出的“向善”约束提供了可操作的机制。

在国家政策持续强化安全治理的同时，也明确强调人工智能“发展与安全并重”的总体原则，要求在推动技术创新和行业应用的过程中，同步构建健全的风险管控与治理体系。因此，大模型软件安全不仅是风险防控的要求，也是保障工业健康发展的必要基础，其治理路径应坚持促进创新、保障发展与有效监管的统一。

本文旨在为大模型软件安全治理提供理论支撑与实践参考，推动AI技术朝着可信、安全、可持续的方向发展。

1　AIBOM概述

作为近年来兴起的治理工具，AIBOM逐渐被视为解决人工智能供应链透明性与可信性问题的核心抓手。本节旨在系统梳理AIBOM的概念脉络、核心要素及其与传统SBOM的差异，并进一步阐释其在大模型软件安全治理中的基础性地位。通过对国内外相关研究与实践的归纳，可以发现AIBOM不仅在语义与结构上延续了SBOM的技术传统，还在模型、数据与运行环境等AI特有维度进行了关键扩展。这种扩展使其兼具工程化落地价值与治理支撑能力，从而逐步成为人工智能可信治理的重要研究前沿。

1.1　定义与核心要素

根据国际开源组织SPDX的草案规范，AIBOM被定位为一种技术性清单格式，旨在系统化记录模型构建过程中的全部组成部分，从而在语义和结构上延续并拓展SBOM的成熟经验[9]。从治理与应用视角出发，Santos等[10]将AIBOM界定为“一种面向人工智能系统的整体化机制，用于呈现其所有组成要素，类似制造业中的物料清单”，突出其在提升系统透明性与可信性方面的作用。由此可见，AIBOM的研究与标准化工作已逐步形成国际共识。

结合现有研究与工业实践，本文所提出的定义强调其工程可落地性与运营价值，即AIBOM是一种基于SBOM标准（如SPDX、CycloneDX）的扩展性清单格式，用于对人工智能系统关键组成部分（模型、数据、依赖、运行环境等）及其演化关系（lineage）进行标准化描述。其核心目标是提供可交换、可追溯、可验证的元信息基础，以支撑跨组织的合规审计、安全治理与全生命周期管理。这一视角既承继了国际标准的规范性，也契合我国“人工智能促进计划”等国家战略发展方向[11]，对于构建安全、透明与高效的AI生态具有重要意义。

在核心要素方面，AIBOM不仅为模型及相关工件提供统一标识，还涵盖谱系与来源、数据集描述、运行环境与依赖、评测与风险、控制与合规以及审计与证明等多个维度。这些要素共同作用于提升人工智能系统的透明性与可验证性：一方面支持跨组织的比对、防篡改与责任追溯，另一方面为合规监管、安全治理与生命周期管理提供数据支撑。AIBOM的核心要素见表1。

1.2　与传统SBOM的区别

与传统SBOM相比，AIBOM在语义和结构上实现了扩展。Boming等[12]指出，其关键特征在于引入了面向AI场景的专属字段，既包括模型本身，也涵盖训练与测试数据。传统SBOM虽然能够有效描述软件组件及依赖关系，但在数据、模型与运行环境等AI特有维度上存在不足，难以全面覆盖相关风险。因此，AIBOM应被视为SBOM在人工智能领域的演化与延伸，其关注点已从“组件与依赖”拓展至“模型、数据、环境与合规”的全生命周期治理。SBOM与AIBOM的对比分析见表2。

表1 AIBOM的核心要素

核心要素	定义	作用	典型实例
资产标识	唯一标识AI模型及工件	保证跨组织传递的唯一性与可验证性	SPDX标准中的SPDXID字段
谱系与来源	描述模型来源及演变关系	支持责任划分、合规检查与安全事件溯源	CycloneDX中的DependencyGraph字段
数据集描述	记录训练与推理所用数据集	确保合法合规，支持隐私保护与跨境审查	CycloneDX中的Service字段
运行环境与依赖	描述运行所依赖的软件与硬件	保障漏洞检测与环境一致性	SPDX中的Package字段
评测与风险	记录模型性能与安全评测	提供可信度量，支持准入与治理	扩展命名空间中的Metrics (accuracy, f1)
控制与合规	描述安全控制与法规映射	确保运行合规，推动“合规即代码”	CycloneDX中的Policy字段
审计与证明	提供可信证明与审计证据	确保AIBOM自身可信性，支持第三方审计	SPDX中的CreationInfo字段

表2 SBOM与AIBOM的对比分析

对比维度	SBOM	AIBOM
描述对象	软件包、依赖库、容器镜像	模型、数据集、训练/微调作业、推理服务、运行环境
生命周期覆盖	主要集中于软件交付与测试阶段	除软件本身，还覆盖训练、微调、部署、推理的大模型全生命周期
风险类型	依赖漏洞（CVE）、开源许可合规	除软件常见风险，还包含数据投毒、模型后门、偏见、越权调用、隐私泄露、环境漏洞
数据与合规	聚焦开源许可证与版权问题	除开源许可证与版权问题，还包含数据来源、采集合规、隐私保护、跨境传输、模型对齐等
变化频率	软件版本更新相对稳定	模型与数据迭代频繁，运行时表现动态影响风险
标准支持	SPDX、CycloneDX等已成熟应用	在SPDX/CycloneDX基础上扩展模型、数据、评测、合规等字段，未形成行业共识

1.3 在大模型软件安全治理中的地位

AIBOM在大模型软件安全治理中发挥着基础性作用。其核心价值不仅在于延续SBOM的透明性目标，还在于面向AI场景扩展至模型、数据与运行环境等关键要素。通过对这些组件进行结构化与标准化记录，AIBOM显著提升了大模型系统的可解释性与可追溯性。Karen等[13]指出，基于SPDX 3.0构建的AIBOM已将算法、数据集与合规信息纳入其中，从而为AI系统的透明治理奠定了重要基础。

在合规与风险治理层面，AIBOM同样具有重要意义。大模型的治理不仅涉及数据合法性、隐私保护与跨境流动，还面临大模型投毒、模型后门等新兴威胁。AIBOM通过结构化描述与可验证证据，为审计、责任划分和合规检查提供支持。Santos等[10]进一步指出，AIBOM核心目标之一是建立可验证的信任机制，从而提升AI系统的安全性与合规性。

总体而言，AIBOM在大模型软件安全治理中兼具透明性、合规性与可追溯性功能，并进一步延伸至生命周期管理、风险预警与跨组织协作，其治理理念与近年来从生命周期与风险防范双视角构建软件与数据安全技术体系的研究方向高度契合[14]。它不仅是大模型治理的重要工具，也为实现人工智能系统的全生命周期可信治理提供了坚实的基础。

2 工业界应用现状

随着AIBOM研究逐渐成型，工业界亦在积极探索其在实际业务场景中的应用路径。与学术界注重理论框架与标准定义不同，企业实践更侧重于在复杂业务链路中实现模型、数据与运行环境的可追溯性与治理。总体而言，已有探索初步展现了AIBOM在提升透明性与安全性方面的价值，但大多仍处于局部应用阶段，在标准兼容、跨组织互操作以及运行时动态治理方面尚未形成成熟能力。

2.1 行业案例

为了揭示AIBOM在工业界的应用情况，本文

选取国外厂商 Wiz 与国内厂商悬镜安全（SeerAI）作为典型案例进行分析。二者在路径选择上呈现出差异化特征：前者强调将 AIBOM 嵌入云平台安全框架，形成企业级资产视图；后者则突出“静态清单—血缘图谱—运行时探针”的一体化安全管理模式，展现出更强的场景驱动属性。

（1）国际案例：Wiz 的 AI-SPM 实践

Wiz 是一家专注于云原生安全的厂商，其提出的人工智能安全态势管理（AI security posture management，AI-SPM）方案中，AIBOM 被视为核心组成部分。不同于传统 SBOM 仅关注软件依赖透明度的做法，Wiz 尝试利用 AIBOM 覆盖模型、数据、身份权限与运行环境，以构建企业级 AI 安全资产视图。其实现路径主要包括：①基于模型谱系链路追踪，形成对基础模型、微调任务与生成模型的可视化管理；②结合数据集依赖分析，以识别敏感信息泄露或隐含的合规风险；③对身份权限与运行环境进行依赖检查，实现账号、组件漏洞与配置风险的统一映射。

尽管该方案在工程化落地方面具有启发性，但其局限亦较为明显。一方面，过度依赖云平台应用程序接口（API）与访问控制（IAM）体系，导致其在多云或跨组织场景中的适配性不足；另一方面，该方案仍偏向静态清单与准入控制，对运行时动态风险（如提示注入、推理偏差）的捕获支持有限。整体来看，Wiz 的实践主要体现为“在云安全框架内嵌入 AIBOM”，距离覆盖 AI 全生命周期治理仍存在差距。

（2）国内案例：悬镜安全的 SeerAI 实践

悬镜安全是国内较早将 AIBOM 工程化落地的厂商之一，其方案覆盖数据、开发、MLOps、运行时与运维治理的全链路，突出“静态清单+血缘图谱+运行时探针”的一体化设计。在具体实施中，悬镜通过 Notebook 与项目扫描自动生成依赖清单，并结合大模型元数据与血缘图谱，实现模型与数据集的可追溯性；同时，依托运行时探针与大语言模型（LLM）防火墙，对推理调用、配置策略与异常行为进行持续监控，形成“动静结合”的安全治理视角。

然而，该方案亦存在不足。一是过度依赖厂商自建的漏洞情报与血缘库（非开源），缺乏与国际标准的有效对齐，难以实现跨国互认与通用性；二是在多厂商模型与异构环境下，运行时探针的可移植性有限，导致部署与扩展成本偏高。因此，悬镜的实践虽凸显了场景驱动的工程化特征，但在标准化与跨组织互操作方面仍需进一步完善。

2.2　落地关键与难点

首先，基于上述案例分析，AIBOM 在工业界的应用虽已初见成效，但要实现广泛落地，仍需聚焦若干关键环节并克服相应挑战。

在关键要素方面，标准化是落地的前提条件。AIBOM 需要与 SPDX、CycloneDX 等现有标准保持兼容，同时扩展出适配 AI 场景的专属字段，从而形成跨组织、跨平台的统一表达体系。其次，元数据采集能力是实现有效治理的技术基础。覆盖模型、数据与运行环境的全生命周期自动化采集机制，能够显著提升 AIBOM 的完整性与可验证性。再次，跨域互操作能力是其发挥价值的关键。由于 AI 供应链往往涉及多云、多组织环境，只有实现不同生态间的互认与交换，才能建立可信的供应链协作网络。最后，运行时监控与动态治理不可或缺。AIBOM 不仅应关注静态清单，还需结合运行时风险探测，实现从构建到推理阶段的全链路安全覆盖。

落地难点当前主要体现在 4 个方面：一是国际标准尚未完全统一，导致不同厂商间存在割裂，限制了生态互通；二是工业界对 AIBOM 的理解存在分歧，部分实现偏重合规审计，忽视了运行时动态治理；三是大模型软件的复杂性导致风险呈现多样化与动态化特征，传统 SBOM 范式难以直接迁移。

总体而言，AIBOM 的工程化落地既需要在

标准、工具与机制层面形成统一框架，也需要在工程实践中逐步解决生态割裂、风险动态性和合规约束等现实问题。只有在学术研究与工业实践的双重推动下，AIBOM才能真正成为大模型软件安全治理的基础性工具。同时，AIBOM的完善也为大模型在关键行业的规模化应用奠定透明、可信的基础，从而促进人工智能技术的创新发展与工业价值释放。

3 运营商视角的大模型软件安全治理

AI技术的深度应用已成为电信运营商数字化转型的核心引擎[15]。在电信行业全面推进数智化转型的背景下，大模型能力已成为业务创新和服务升级的重要驱动力。因而，大模型治理体系必须在促进创新应用的同时兼顾安全可靠，实现“发展与安全并重”的整体目标。随着大模型在电信运营商业务中的广泛应用，其安全治理问题呈现出明显的行业特性。运营商不仅需要管理大规模用户数据和复杂网络环境，还要承担社会责任与监管合规的双重压力。在这一背景下，AIBOM的引入为运营商建立可控、可追溯和可审计的治理体系提供了基础支撑。

3.1 运营商的独特挑战

运营商在大模型软件安全治理中面临的挑战具有多维度和复杂性。一方面，运营商业务具有跨区域、多节点部署和多系统协作的特点。大模型可能被用于用户行为分析、网络流量预测、智能客服、故障诊断等多个业务场景，每个场景的模型依赖不同的数据集和训练策略，这使得统一的安全管理和风险监控变得异常复杂。另一方面，运营商掌握大量敏感数据，包括用户身份信息、通信内容及网络行为数据，这类数据的使用受制于严格的隐私保护和合规要求。在构建和管理AIBOM时，如何在保障治理透明度的同时，确保数据隐私和合规性，是运营商特有的挑战。

此外，大模型自身的特性也增加了治理难度。模型在微调过程中可能引入隐藏后门，运行阶段可能出现幻觉输出或异常行为，而运营商需要在不中断业务的前提下对这些风险进行监控和响应。大模型更新迭代快速，跨系统调用频繁，如何实现动态追踪、实时预警以及跨节点协同治理，是传统软件安全工具难以解决的问题。同时，大模型对底层算力平台与运行环境的稳定性高度依赖，超大规模智算集群的可靠运行已成为支撑大模型安全与治理的重要前提[16]。更为复杂的是，运营商需面对来自监管机构的审计要求，要求模型构建和使用过程具备完整的可追溯性，以证明合规性和责任边界的清晰性。

这些特有挑战决定了运营商对大模型软件安全治理提出了更高要求：治理体系不仅要覆盖模型全生命周期，还要兼顾跨系统、多区域的操作管理、敏感数据保护以及实时风险响应能力。AIBOM在此背景下或可成为运营商实现全局可控、安全可信的关键工具。

3.2 AIBOM的战略地位

在运营商的大模型治理体系中，AIBOM扮演着战略性基础设施的角色。首先，它提供了统一的治理框架，使运营商能够对分布在不同业务系统和节点的大模型进行集中管理和风险控制。通过对模型构建、微调、训练数据及接口行为的全生命周期记录，AIBOM将潜在风险转化为可量化、可操作的管理指标，使运营商能够实现预防性治理和实时监控。其次，AIBOM支撑运营商在合规与审计方面的需求。在电信行业，监管机构对用户数据安全和业务连续性有严格要求。AIBOM通过标准化的记录和可验证的治理流程，使运营商在面对审计时能够提供完整、可追溯的证据链条，从而降低合规风险，并增强公众与监管部门的信任。最后，AIBOM为运营商内部的跨部门协同提供技术手段。运营商在大模型的部署与使用过程中涉及研发、数据、网络、安全等

多个部门，AIBOM 可以通过结构化的治理信息，建立统一语言和操作接口，实现跨部门的数据共享、安全监控和事件响应。通过这种方式，运营商能够在复杂环境下实现安全治理闭环，将模型从“可用资产”转化为“可治理资产”。

总体而言，AIBOM 在运营商的大模型软件安全治理中具有不可替代的战略地位：它不仅是风险管控的工具，更是合规证明和业务可靠性的保障。

4　运营商大模型软件安全治理的实践路径——以中国电信为例

为应对运营商在大模型软件安全治理中面临的多重挑战，行业普遍探索了从政策制度、标准规范到工具体系与流程治理相结合的系统化路径。中国电信的相关实践具有一定代表性，体现了运营商在大模型软件全生命周期治理中的共性需求和可推广方法。本节以中国电信的探索为例，对运营商大模型软件安全治理的总体思路与关键实践进行归纳总结。

4.1　大模型软件安全治理政策发展与总体思路

随着生成式人工智能的广泛应用，围绕大模型的安全问题逐渐成为监管部门和行业关注的重点。特别是自 2023 年以来，国内政策和行业规范密集出台，从法律法规到企业内部指引，为大模型治理提供了制度保障。中国电信在此背景下，紧密跟随国家战略部署，结合企业业务实际，逐步建立起覆盖全链条的大模型软件安全治理政策体系。中国电信大模型软件安全治理政策发展可分为政策响应、规范与制度探索、规范试行与体系化落地 3 个演进阶段。

（1）政策响应阶段：发布《关于做好生成式 AI 信息安全管理工作的通知》，首次明确了大模型信息安全管理的组织分工和责任界定。

（2）规范与制度探索阶段：从内容安全入手，推出《生成式人工智能内容安全测评规范（试行）》，在企业内部形成了测评的初步操作性标准，随后发布《生成式人工智能安全风险视图与治理思路》，全面梳理安全风险并确立中国电信大模型软件安全治理思路。

（3）规范试行与体系化落地阶段：陆续出台《关于强化大模型一体机安全防护工作的通知》《生成式人工智能安全治理实施指引（试行）》和《生成式人工智能安全评估与评测规范（试行）》，将治理要求具体化、体系化，形成了从制度到标准再到落地工具的全链路闭环。在落地指引中，大模型软件安全也作为关键部分发挥着举足轻重的作用。

基于体系化落地实践经验，以中国电信为例的运营商持续优化大模型安全治理思路，明确以安全可信、保障发展、可控向善为核心导向，以环境安全可靠、数据治理有效、模型算法可信、内容安全合规、应用风险可控为整体目标，提出包括风险治理机制建设、内容安全加固、安全应急处置、全量测评、常态化检查与众测、自动化测评工具研发以及算法/模型安全探索七大行动方向，要求安全治理贯穿大模型的“开发–预训练–开源–微调–部署–备案–运营–废弃”全流程闭环，确保大模型在整个生命周期中均受安全约束。其中，大模型软件安全在整体治理思路中承担了“底座安全”的角色，为大模型的可信运行提供基础保障，是治理体系不可或缺的关键一环。

4.2　大模型软件安全治理实践

在总体思路的引领下，围绕“环境安全可靠、数据治理有效、模型算法可信、内容安全合规、应用风险可控”的治理目标，制定并实施了一系列针对大模型软件安全的重点举措。这些举措既有制度层面的顶层设计，也有标准与工具层面的技术落地，逐步形成了体系化、可操作的治理框架，中国电信大模型软件安全治理框架如图 1 所示。

图1　中国电信大模型软件安全治理框架

（1）标准构建：提出并推动AIBOM标准化建设

以中国电信为代表的运营商率先提出并推动AIBOM标准化建设，将大模型软件资产纳入可视化、可追溯的治理范畴。中国电信人工智能物料清单构成要求如图2所示，在继承传统SBOM文档构成与通用字段的基础上，新增了模型字段、训练数据集字段、训练环境字段三大类数据字段，全面覆盖了大模型软件的关键要素，为治理提供了统一的标准化描述语言。在此基础上，中国电信已完成国内首个人工智能物料清单相关行业标准的立项工作，标志着我国在大模型软件安全治理标准化方面实现了突破。

（2）自动化工具研发与应用：构建AIBOM管理与检测能力

在工具体系方面，通过自主研发自动化

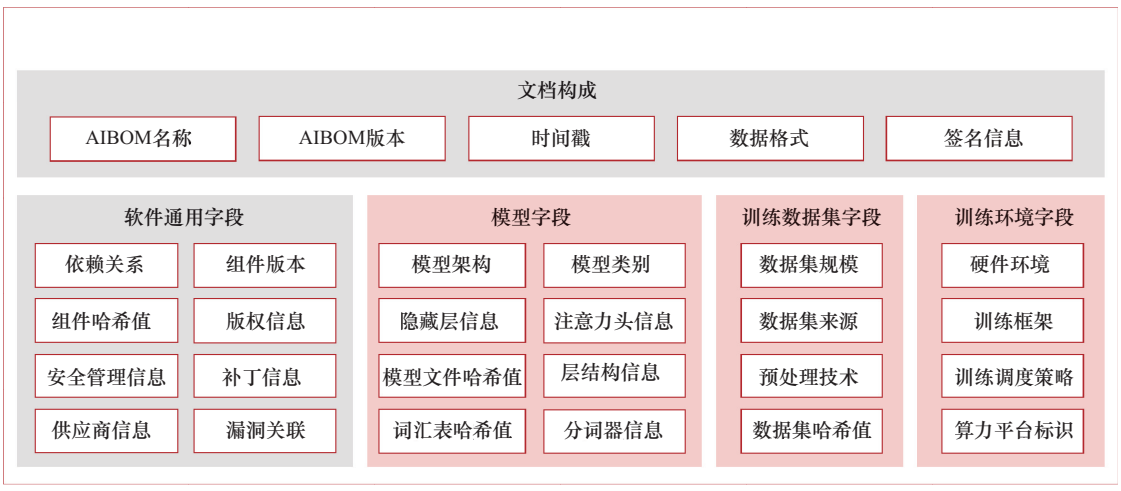

图2　中国电信人工智能物料清单构成要求

AIBOM 管理与检测工具，实现元数据采集、风险情报聚合、影子模型检测、数据集溯源等核心能力，降低大模型软件安全治理的人力成本。

① AIBOM 采集：研发支持模型引入、训练、部署等环节的全生命周期采集工具链，可自动化捕获来自 HuggingFace 等主流 AI 平台的元数据信息，确保资产登记的完整性与一致性。

② AI 风险情报采集：构建面向开源社区、软件供应商和安全机构的多源风险情报实时采集机制，动态扩展 AI 风险溯源数据库，涵盖组件漏洞、应用漏洞、框架漏洞、知识产权风险、模型投毒等类别，并提供修复与缓解的参考路径。

③ 影子模型检测：在 AIBOM 元数据架构中引入模型血缘、数据谱系和环境依赖的增强描述机制，统计与分析模型血缘图谱数据。依托 AIBOM、AI 组件指纹库及血缘图谱，可实现影子模型的识别与风险判断，从而提升责任追溯性与透明性。

④ 数据集溯源：构建覆盖报备数据字段与数据集库的双重比对机制，通过数据一致性检查与篡改检测，降低恶意数据集引入带来的安全风险。

在上述能力的基础上，可信 AI 开发组件库的构建亦为运营商提供安全可控的模型与依赖来源，为源头构建安全研发体系提供支撑。同时，电信运营商正在探索构建兼容 SPDX 与 CycloneDX 等多标准的 AIBOM 转换中间件，以实现多标准协同与互操作，为行业提供更广泛的适配与推广价值。

（3）大模型软件供应链安全管理平台构建：实现资产全量覆盖与动态纳管

在平台建设方面，构建大模型软件供应链安全管理平台，并通过建立统一的大模型软件资产台账，实现对集团范围内资产的全量覆盖与动态纳管，以支撑风险治理与审计追溯需求。在平台功能设计上，应强调差异化管理与闭环管控，形成以 AIBOM 为核心的 4 项关键能力。

① AIBOM 分级分类：针对不同业务场景与风险等级，制定差异化的 AIBOM 采集规范，对高敏感场景明确强制采集与审计字段，确保治理粒度能够与风险暴露程度相匹配。

② AIBOM 动态更新：建立覆盖模型引入、训练、部署及迭代等阶段的动态更新机制，形成可追溯的 AIBOM 全生命周期闭环管理能力。通过技术手段实现资产变更在 1 h 内同步与更新，确保资产台账的实时性与一致性，从而为风险快速感知与处置提供保障。

③ AIBOM 风险治理：在大模型软件全量资产的基础上，构建供应链攻击面分析模型，结合 AIBOM 元数据进行风险聚合、漏洞关联与攻击路径推演，从而实现由“资产可见”向“风险可治”的转变。平台进一步支持将治理结果纳入安全评估体系，形成“政策—标准—工具”的闭环联动。

④ 大模型软件资产测绘：通过跨平台与跨云环境的自动化检测能力，动态识别分布于阿里云、AWS、Azure、腾讯云等多云环境下的 AI 软件资产，并自动构建资产画像与风险图谱，实现大模型软件供应链风险的全域感知与精准定位。该功能不仅覆盖 100% 纳管资产，还能够实现对资产变更趋势与风险演化的持续监测，为战略决策与资源优化提供量化支撑。

此外，通过差异化 AIBOM 分级分类、动态更新机制、风险聚合分析及资产测绘能力，电信运营商可实现从“资产可见”向“风险可治”的系统化能力演进。

（4）安全评测与风险监测：强化实证检验

在资产全量纳管的基础上，开展针对大模型软件的安全评测与风险监测，为优化治理措施提供实证数据支撑。上线前，开展覆盖软件成分分析、许可证合规、病毒检测、漏洞扫描与渗透测试在内的 5 类 20 项安全评测，通过评测并取得安全合格证明后，方可上线使用。上线后，基于大模型软件管理平台开展定期风险监测，及时识别

相关安全漏洞与模型风险，实时上报威胁情报至综合调度系统，对涉险软件自动派发工单，将大模型软件安全治理纳入集团统一处理流程，实施分级处置与持续改进，形成闭环的安全治理机制。

（5）常态化检查与闭环治理：建立长效机制

为确保大模型软件安全治理的持续有效，在内部实践中逐步建立了常态化检查与闭环治理机制。针对内部单位各业务单元的大模型软件系统，实行定期检查与不定期抽检相结合的方式，借助集AIBOM检测工具、代码检测工具等多种工具为一体的软件安全工具箱实现技术赋能，覆盖软件供应链安全、模型安全与合规性、训练数据与环境安全等多维度风险。一旦发现漏洞或隐患，立即通过AIBOM管理工具实现责任追溯，启动整改与复测流程，并利用工具箱验证漏洞整改措施有效性，将漏洞发现、整改、复测、验证等多个环节紧密衔接起来，实现软件风险可知可治、治理过程可管可控。

4.3 大模型软件安全治理实践成效

在标准体系逐步建立的背景下，运营商在大模型软件安全治理方面已取得阶段性成效。

以中国电信为代表的行业实践显示，标准化建设正在推动大模型软件供应链治理从探索走向体系化落地。已完成首个人工智能软件供应链领域CCSA行业标准《电信与互联网 人工智能系统软件供应链安全管理要求》的立项，并牵头启动了首个面向AIBOM的CCSA行业标准《电信与互联网 人工智能物料清单（AIBOM）格式要求》的立项工作。这不仅填补了国内人工智能软件供应链治理标准方面的空白，也为大模型软件安全治理的规范化、体系化奠定了制度基础，提升了我国在相关国际规则制定中的话语权和先导力。

在政策体系的牵引下，运营商在资产治理方面也取得了显著进展。实践表明，通过集团级的治理体系建设，可以实现对大模型软件资产的全面识别、集中纳管和持续更新。例如，中国电信已完成对全网范围大模型软件资产的系统化梳理，形成覆盖信息系统名称、定级备案情况、服务对象类型、安全评测结果等多维度要素的资产台账；截至2025年5月，已对49家省级与专业公司200余个大模型建立动态更新机制，其中约60%已通过安全评测，显著提升了风险识别与治理效率。

在支撑可信开发方面，行业实践亦展现出积极成效。中国电信构建了基于AIBOM的第三方组件治理能力。通过对大模型开发常用的5 000余个组件进行收集与分析，结合AIBOM中关于来源、版本与安全属性的描述，筛选出2 000余个符合安全与可信要求的组件，纳入“可信AI开发组件库”并对外发布。可信AI开发组件库的建设路径也显示出较高的行业可复用性。通过对大模型研发组件的筛选与治理，运营商可构建安全基线，推动研发模式从“经验驱动”向“规范驱动”转变。

总体而言，该体系建设路径为运营商提升大模型软件研发、部署与运营过程的风险可视化和可控性提供了普遍意义上的方法论参考。

5 未来研究与发展重点

随着AIBOM在国内国际标准化组织和工业界的探索逐步展开，其研究与应用仍处于早期阶段，如何在保障大模型可信性、合规性与可治理性的同时兼顾可操作性与经济性，成为亟待解决的问题。尤其是在智能体、端侧推理、模型控制协议服务及大模型一体机等新型应用涌现的背景下，人工智能安全治理正加速迈入“深水区”。国际上，欧盟已围绕人工智能伦理原则、风险分级与治理责任划分等方面形成了较为系统的治理框架，标志着人工智能安全治理正逐步从原则层面走向制度化和工程化实践[17]。

上述应用复杂化趋势与治理实践进展，对AIBOM的发展提出了更高的要求。从AIBOM的工程化落地与实际应用需求出发，未来研究可重点聚焦标准化、可信性、环境适配3个方向。

（1）标准统一与互操作性：当前 AIBOM 依托 SPDX、CycloneDX 等 SBOM 扩展，但语义差异导致跨标准兼容不足。亟须推动标准融合，或建立高效映射机制，实现互操作性，以降低企业应用门槛与维护成本。

（2）AIBOM 的不可篡改性：可信性是 AIBOM 治理价值的前提。未来可探索区块链、分布式账本、数字签名及零知识证明等技术，确保清单生成、传输与存储全流程的完整性与可验证性，并建立多方信任框架以支撑跨组织协同。

（3）云环境下的同步问题：端、边、云多层部署使 AIBOM 一致性面临挑战。未来可研究轻量化清单表示以适配资源受限设备，并结合可信执行环境（TEE）、远程证明及 CI/CD 工具，实现跨域、多平台的实时同步与验证。

6　结束语

大模型正在成为推动新一轮工业变革的核心力量，但其安全风险同样不可忽视。本文围绕 AIBOM 在电信运营商大模型软件安全治理中的研究与实践展开系统探讨。首先，界定了 AIBOM 的内涵与核心要素，并对比传统 SBOM，阐明其在大模型软件安全治理中的战略价值。随后，通过行业案例分析揭示 AIBOM 落地的主要难点，并在运营商视角下强调其在应对模型复杂性、合规性与供应链风险方面的不可替代性。在此基础上，本文总结了中国电信的治理思路与实践路径。这些实践不仅为大模型软件安全治理提供了方法论借鉴，也为 AIBOM 的工程化落地积累了宝贵经验。

展望未来，AIBOM 的研究仍需在标准化框架、不可篡改性保障和跨环境协同等方面深化。随着其体系逐步成熟，AIBOM 有望成为大模型软件安全治理的核心枢纽，为电信运营商及更广泛工业提供可持续的安全保障。

参考文献：

[1] 韩炳涛, 刘涛. 大模型关键技术与应用[J]. 中兴通讯技术, 2024, 30(2): 76-88.
Han B T, Liu T. Key technologies and applications of large models[J]. ZTE Technology Journal, 2024, 30(2): 76-88.

[2] 欧阳晔, 王立磊, 杨爱东, 等. 通信人工智能的下一个十年[J]. 电信科学, 2021, 37(3): 1-36.
Ouyang Y, Wang L L, Yang A D, et al. Next decade of telecommunications artificial intelligence[J]. Telecommunications Science, 2021, 37(3): 1-36.

[3] 中国移动. 2023 电信 AI 产业发展白皮书[R]. 2023.
China Mobile. White paper on AI industry development in telecommunications 2023[R]. 2023.

[4] 乔喆. 人工智能生成内容技术在内容安全治理领域的风险和对策[J]. 电信科学, 2023, 39(10): 136-146.
Qiao Z. Risks and countermeasures of artificial intelligence generated content technology in content security governance[J]. Telecommunications Science, 2023, 39(10): 136-146.

[5] 工业和信息化部, 国家标准化管理委员会. 国家人工智能产业综合标准化体系建设指南 (2024 版)[R]. 2024.
Ministry of Industry and Information Technology, Standardization Administration of China. Guidelines for the construction of a comprehensive standardization system for the artificial intelligence industry (2024 Edition)[R]. 2024.

[6] 新华网. 建立人工智能安全监管制度[N]. 新华网, 2024-11-06.
Xinhuanet. Establishing an artificial intelligence security supervision system[N]. Xinhuanet, 2024-11-06.

[7] 中共中央政治局. 把握人工智能发展规律 构建风险预警体系[N]. 人民网, 2025-04-26.
The Political Bureau of the CPC Central Committee. Grasping the law of artificial intelligence development and building a risk early warning system[N]. People's Daily Online, 2025-04-26.

[8] 王戈, 郭新海, 刘安, 等. 基于 SBOM 的软件安全治理实践[J]. 邮电设计技术, 2023(8): 9-13.
Wang G, Guo X H, Liu A, et al. Practice of software security governance based on SBOM[J]. Designing Techniques of Posts and Telecommunications, 2023(8): 9-13.

[9] SPDX Workgroup. SPDX AI BOM draft specification 3.0[EB]. 2023.

[10] Santos O, Radanliev P. Toward trustworthy AI: an analysis of artificial intelligence (AI bill of materials) (AI BOMs)[EB]. 2023.

[11] 国家发展和改革委员会. 新一代人工智能发展规划[R]. 2017.
National Development and Reform Commission. Development

plan for a new generation of artificial intelligence[R]. 2017.

[12] Xia B, Zhang D, Liu Y, Lu Q, Xing Z, Zhu L. Trust in software supply chains: blockchain-enabled SBOM and the AIBOM future[EB]. 2023.

[13] Bennet K, Rajbahadur G K, Suriyawongkul A, et al. Implementing AI bill of materials (AI BOM) with SPDX 3.0: a comprehensive guide to creating AI and dataset bill of materials[PP]. V1. arXiv (2025-04-23)[2025-09-08]. arXiv: 2504.16743.

[14] 陈禹存, 黄科满, 杜小勇. 基于生命周期与风险防范双视角的数据流通安全技术体系[J]. 大数据, 2024, 10(6): 16-32.
Chen Y C, Huang K M, Du X Y. A data circulation security technology system based on the dual perspectives of lifecycle and risk prevention[J]. Big Data Research, 2024, 10(6): 16-32.

[15] 中国联通. 中国通信运营商AI+DevOps实践报告 (2024)[R]. 2024.
China Unicom. Report on AI+DevOps practices of Chinese telecom operators (2024)[R]. 2024.

[16] 牛红韦华, 黄永宝, 丁国强, 等. 基于数据和知识驱动的超万卡智算集群稳定性保障实践[J]. 电信科学, 2025, 41(7): 145-163.
Niu H W H, Huang Y B, Ding G Q, et al. A data and knowledge-driven practice for ensuring stability in ultra-large intelligent computing clusters[J]. Telecommunications Science, 2025, 41(7): 145-163.

[17] 曹建峰, 方龄曼. 欧盟人工智能伦理与治理的路径及启示[J]. 人工智能, 2019(4): 39-47.
Cao J F, Fang L M. The path and enlightenment of EU artificial intelligence ethics and governance[J]. AI-View, 2019(4): 39-47.

[作者简介]

张侃（1970– ），男，博士，中国电信集团有限公司网络和信息安全管理部正高级工程师、高级资深专家，主要研究方向为信息通信、网信技术。

王诗雨（1995– ），女，中国电信股份有限公司研究院工程师、软件安全能力中心研究员，主要研究方向为软件供应链安全、安全有效性验证、人工智能技术。

朱沛潼（2000– ），男，中国电信股份有限公司研究院科研助理，主要研究方向为信息安全技术。

徐帅健妮（1994– ），女，博士，中国电信股份有限公司研究院高级工程师，软件安全能力中心副总监，主要研究方向为软件供应链安全、内生安全、容器镜像安全和密码学。

殷铭（1994– ），男，中国电信股份有限公司研究院工程师、软件安全能力中心副总监，主要研究方向为软件供应链安全、移动应用安全。

何国锋（1973– ），男，博士，中国电信股份有限公司研究院正高级工程师，安全应用技术研发部主任，主要研究方向为信息通信、网信安全。