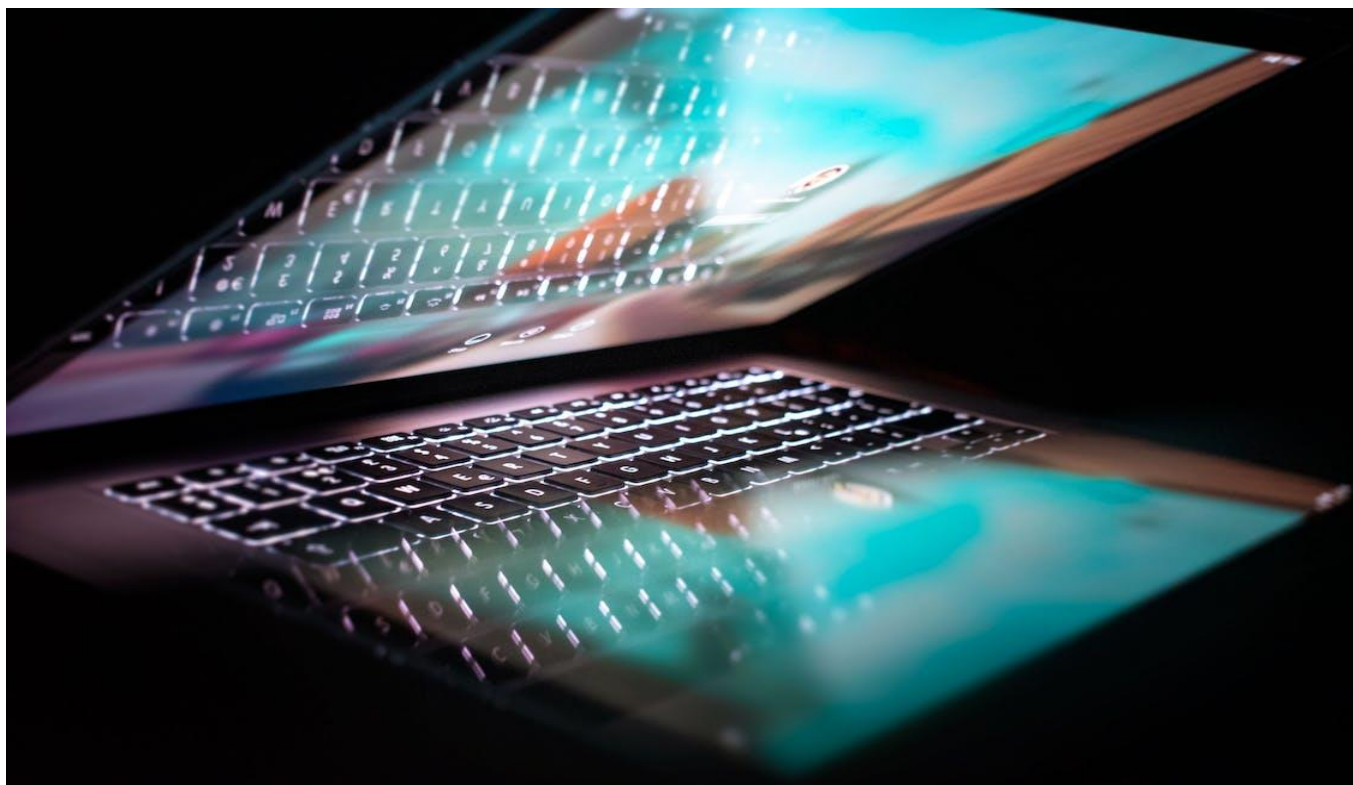


信息资源管理标准规范动态报告

NSTL 数据研究管理中心



全面调研和梳
理国内外信息
资源管理领域
标准规范的发

展动态，涵盖数据管理、数据组
织、知识技术相关的新标准、新
规范、新模型等。

本期看点

- ISO/DIS 55013 Asset management – Guidance on the management of data assets
- ISO/IEC 8183:2023 Information technology — Artificial intelligence — Data life cycle framework
- ANSI/NISO Z39.106-2023, Standard Terminology for Peer Review
- GB/T 42813-2023 数据论文出版元数据
- GB/T 30523-2023 科技资源核心元数据
- GB/T 42713-2023 信息与文献 参与者名称标识符

目 录

数据管理	1
◇ ISO/DIS 55013 Asset management - Guidance on the management of data assets.....	1
◇ GB/T 42745-2023 信息与文献 可信的第三方数字文件（档案）仓储	2
◇ GB/T 41795-2022 质量技术基础 信息资源数据规范.....	3
◇ GB/T 42574-2023 信息安全技术 个人信息处理中告知和同意的实施指南	4
◇ GB/T 42748-2023 专利评估指引.....	6
数据组织	7
◇ ANSI/NISO Z39.106-2023, Standard Terminology for Peer Review....	7
◇ GB/T 42813-2023 数据论文出版元数据.....	9
◇ GB/T 30523-2023 科技资源核心元数据.....	10
◇ GB/T 42713-2023 信息与文献 参与者名称标识符.....	11
◇ GB/T 42743-2023 信息与文献 国家科技重大专项档案元数据元素集..	12
◇ GB/T 42093 标准文档结构化 元模型.....	13
知识技术	15
◇ ISO/IEC 8183:2023 Information technology - Artificial intelligence - Data life cycle framework.....	15
◇ GB/T 42888-2023 信息安全技术 机器学习算法安全评估规范.....	16
◇ GB/T 42777-2023 基于文本数据的金融风险防控 知识图谱构建技术框架指南.....	17
◇ GB/T 41818-2022 信息技术 大数据 面向分析的数据存储与检索技术要求	18

数据管理

◇ ISO/DIS 55013 Asset management - Guidance on the management of data as sets

【看点】国际标准化组织负责的《ISO/DIS 55013 资产管理 数据资产管理指南（Asset management - Guidance on the management of data assets）》标准制定已完成了委员会草案（CD, Committee Draft）阶段，即将进入国际标准草案（DIS, Draft International Standard）阶段。该标准在 ISO 55001 资产管理体系框架下将数据与其在资产管理中的实践经验进行了结合，并给出了管理指引。

【摘要】当前，越来越多的组织依赖数据资产作为组织运营的关键推动力，这需要组织对其业务数据需求有一个清晰的认识，并确定实现业务目标所需的数据，数据的质量要求，持有、管理和迁移数据的规则要求，同时还要关注数据资产管理成本、风险和绩效之间的平衡。

ISO 55013 标准在 ISO 55001 资产管理体系框架下将数据与其在资产管理中的实践经验进行了结合，可为不同类型和规模的组织提供管理数据资产的指南，以支持组织实现其资产管理目标，从而实现组织目标。

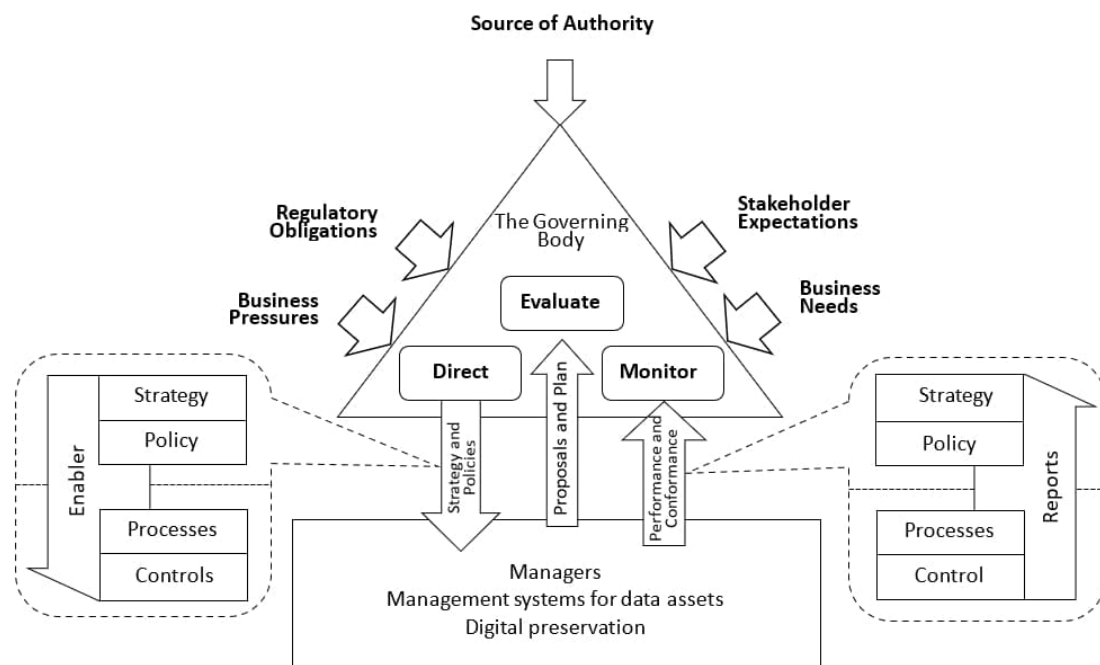


图 1 数据资产管理系统

该标准包括管理资产数据、从资产数据中传递价值、识别数据资产、管理数据资产、治理 5 个部分内容。该标准为 ISO 55001 资产管理体系标准在数据资

产等无形资产管理领域的实践拓宽应用范围,为如何更好的开展数据资产管理工
作提供了指导,为组织实现其数据资产管理目标提供支撑,对组织数据要素基础
设施建设与全生命周期管理有着积极的指导意义,贴合了数字经济的发展趋势,
也符合可持续发展远景目标。

来源:

国际标准化组织: <https://www.iso.org/standard/82455.html>

◇ GB/T 42745-2023 信息与文献 可信的第三方数字文件（档案）仓储

【看点】 国家标准化管理委员会于 2023 年 8 月 6 日发布了国家标准《信息
与文献 可信的第三方数字文件（档案）仓储》，该标准由全国信息与文献标准
化技术委员会（TC4）归口，由中国科学院武汉文献情报中心和中国科学院文献
情报中心起草，将于 2024 年 3 月 1 日正式实施。

【摘要】 该标准等同采用 ISO17068: 2017《信息与文献 可信的第三方数字
文件（档案）仓储》，旨在帮助避免数字文件因篡改等因素而产生业务纠纷，在
数字文件具有法律效力的保管期限内，确保数字文件的真实性、可靠性、完整性
和可用性。

作为数字系统在各种业务活动中的必然产物，数字文件越来越需要确保其在
保管期限内的真实性，并得到法律认可。然而，数字签名仅能在有效期（一般为
一至两年或更短）内保证信息的完整性，故而对于文件系统来说，当保管期限长
于数字签名有效期时，很难证明数字文件的完整性。人类社会越来越需要由中立
的、第三方提供的仓储服务来保障数字文件的真实可信。

可信的第三方仓储（Trusted Third Party Repository, TTPR）提供了一种可
能的解决方案。TTPR 被定义为第三方的合格保管服务，用以确保委托方的数字
文件被合法可信地保管。TTPR 的宗旨是在保管期限内为委托方提供对受管理数
字文件的可信保管服务，这些服务由利益相关方（检查员、审核员评估员等）进
行监管审查。TTPR 提供的服务有助于认证委托方数字文件是否能作为可接受的
证据来源。

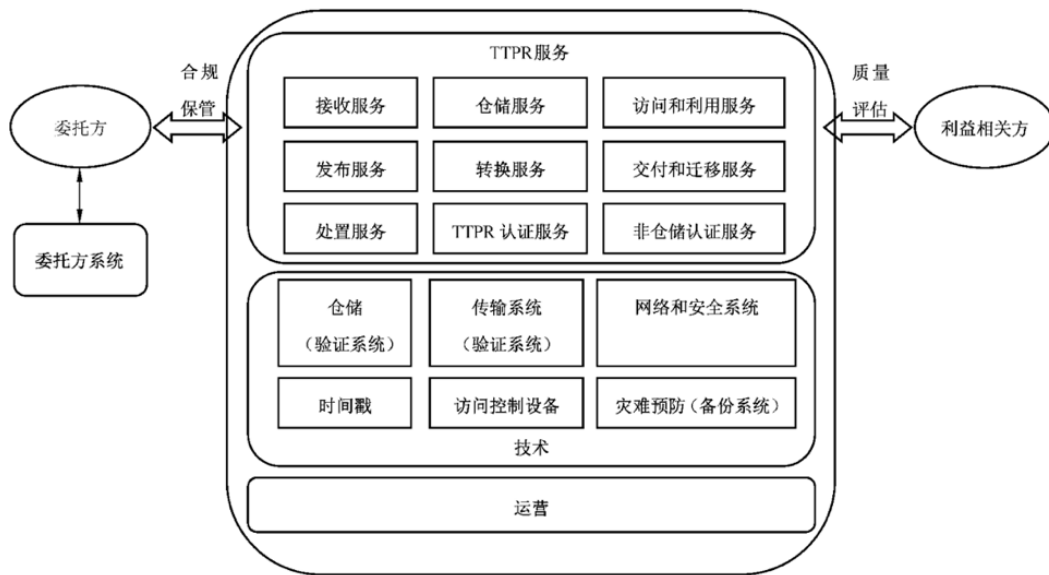


图 2 TTPR 概述

来源:

国家标准信息公共服务平台:

<https://std.samr.gov.cn/gb/search/gbDetailed?id=027A6096B03D643EE06397BE0A0A0867>

国家标准全文公开系统:

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=30EA3E9F5E109DA0B1512D039BE86BF2>

◇ GB/T 41795-2022 质量技术基础 信息资源数据规范

【看点】 国家标准化管理委员会于 2022 年 10 月 14 日发布了国家标准《质量技术基础 信息资源数据规范》，并于 2023 年 5 月 1 日正式实施。标准由全国电子业务标准化技术委员会 (TC83) 归口，主管部门为国家标准化管理委员会。

【摘要】 该标准规定了质量技术基础信息资源的分类和编码、数据元目录、数据采集、数据清洗、数据质量评价、数据溯源、集成服务平台技术实现和数据交换。该标准适用于质量技术基础集成服务平台建设中信息资源的发布、交换、存储和管理，也适用于质量技术基础集成服务平台的设计、开发和验收。

在信息资源分类方面，将质量技术基础信息资源分为大类、中类和小类三个层级。在信息资源编码方面，采用 6 位数字代码表示，前 2 位表示大类、中间 2 位表示中类、后 2 位表示小类。数据元分为人员、机构、仪器设备、标准文献、被检产品、项目、项目报告。数据采集的内容包括标准资源数据、计量资源数据、检验检测资源数据和认证认可资源数据。数据清洗包括脏数据处理、缺失值处理、错误数据处理、错误关联数据处理、非需求数据处理。

在数据质量评价指标体系方面，结合质量技术基础集成服务数据的特点，选取准确性、完备性、一致性、确实性、现时性、依从性、保密性、精度、可跟踪性、可理解性等 10 个固有的质量特性组成质量技术基础集成服务数据质量评价指标体系。在数据溯源方面，溯源方法包括但不限于批次管理法、日期管理法、连续序号管理法、混合管理法。在集成服务平台技术实现方面，平台功能包括但不限于查询检索功能、业务办理功能、咨询功能、统计分析功能、集成服务功能、定制服务功能等。在数据交换方面，包括同步或异步的自动交换方式和网站提交、手工抽取/发送的手工交换方式。

来源：

全国标准信息公共服务平台：

<https://std.samr.gov.cn/gb/search/gbDetailed?id=EB58F4DA9233B2A2E05397BE0A0A7D33>

国家标准全文公开系统：

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=9E84E300C0D2308719D94BF0BCCC52A8>

◇ **GB/T 42574-2023 信息安全技术 个人信息处理中告知和同意的实施指南**

【看点】国家标准化管理委员会于 2023 年 5 月 23 日发布了国家标准《信息安全技术 个人信息处理中告知和同意的实施指南》。该标准由全国信息安全标准化技术委员会（TC260）归口，主管部门为国家标准化管理委员会，将于 2023 年 12 月 1 日正式实施。

【摘要】该标准给出了处理个人信息时，向个人告知处理规则、取得个人同意的实施方法和步骤，适用于个人信息处理者在开展个人信息处理活动时保障个人权益，也可为监管、检查、评估等活动提供参考。

该标准围绕告知和同意两个重要概念，明确了告知和同意适用的情形、基本原则和方式、内容及实施的具体细则，并通过附录性资料丰富了告知和同意适用的场景、情形和案例。

个人信息处理者在实施告知时需要考虑以下基本原则：

- 公开透明：公布处理个人信息的种类、目的、方式、安全措施等处理规则，不采取故意遮挡、隐藏等方式诱导个人略过告知内容；
- 有效传达：尽可能通过交互式界面、邮件、电话或短信等方式向相关个人进行告知；

- 适时充分：在收集、提供、公开等个人信息处理活动发生之前或同时，对个人进行充分告知；
- 真实明确：告知个人信息的处理种类、目的、方式等规则与实际情况一致，且需结合实际业务功能，不使用笼统、宽泛的表述；
- 清晰易懂：告知文本符合个人的语言习惯，使用通用且无歧义的语言、数字、图示等

个人信息处理者在取得个人同意时需考虑以下基本原则：

- 告知一致：取得同意的范围不超出所告知的内容；
- 自主选择：支持个人通过自行操作的方式作出同意，不使用默认勾选的方式取得同意；
- 时机恰当：在个人信息收集行为发生前，且同步传达告知内容时，取得个人同意，以增进个人对业务功能与所收集的个人信息之间关联性的理解；
- 避免捆绑：区分产品或服务的业务功能，不采用捆绑方式强迫个人一次性同意多种业务功能可能收集的个人信息或多个处理活动；个人拒绝同意时，不影响与该个人信息无关的业务功能的正常使用。

个人信息处理者在实施告知和同意时宜考虑以下要素，优化告知和同意的方案和机制：

- 友好展示：使用友好、生动、形象的方式编辑告知内容，优化告知内容组织形式，以便于个人理解；
- 适配媒体：告知内容、展示形式、取得同意的方式等可根据告知媒体的种类、界面特点进行适应性设计，如适宜的字型大小、字体颜色、额外的震动和语音提示等；
- 考虑影响：设计告知和同意方案时，可考虑个人信息处理活动对个人权益的影响程度以及个人的体验、习惯、合理预期等因素；
- 区分阶段：根据个人使用产品或服务不同阶段及交互场景，选用个人信息保护政策、弹窗提示、文字说明等不同的告知和同意方案；
- 兼顾差异：考虑复杂多样的网络条件、软硬件差异、个人的知识水平和理解能力、身体机能差异等，使用可广泛适用且兼顾特定群体的告知和同意方案。

来源：

全国标准信息公共服务平台：

<https://std.samr.gov.cn/gb/search/gbDetailed?id=FC816D04FFD262EBE05397BE0A0AD5FA>

国家标准全文公开系统：

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=09255E6D24E0BAB7F391472C6995EC56>

◇ GB/T 42748-2023 专利评估指引

【看点】国家标准化管理委员会于2023年8月6日发布了国家标准《专利评估指引》，并于2023年9月1日正式实施。该标准由国家知识产权局、中国人民银行和中国银行保险监督管理委员会联合起草，由全国知识管理标准化技术委员会（TC554）归口，主管部门为国家知识产权局。

【摘要】专利评估是在特定目的和场景下对专利资产价值的评定和估算，为专利的经营管理、许可转让、作价入股、质押融资等市场活动提供参考和支持。制定科学、系统和可操作的专利评估指引，有利于引导各相关方把握专利的制度特点和运用规律，关注专利的法律价值、技术价值、经济价值等特征，实现评估指标更全面、评估方法更科学、评估结果更准确，将推动完善专利评估机制，规范专利评估活动，为专利资产管理和运用提供有力支撑，促进创新资源有序流动和高效配置。

该标准提供了一种发明专利、实用新型专利评估的基础性方法工具，通过构建一套可扩展、可操作的专利价值分析评估指标体系，对专利的法律价值、技术价值和经济价值进行综合分析，得到专利价值度。该标准适用于企业、高校、科研组织、金融机构、评估机构等各类主体以下场景的专利评估：许可转让类场景、金融类场景、财务报告类场景、侵权救济类场景、管理类场景以及其他类场景。有关行业组织和机构可在此基础上，制定特定行业领域的专利评估指引。

专利评估的总则包括科学性、系统性、操作性和扩展性，该标准提出的专利评估三级指标体系包括：

表 1 专利评估三级指标体系

一级指标	二级指标	三级指标
法律价值	权利稳定性	专利有效性
		同族专利情况
		复审无效情况
	权利保护范围	不可规避性
		权利要求合理性
	侵权可判定性	权利要求类型

		技术特征属性
	依赖度	依赖度
技术价值	技术先进性	技术问题重要性
		技术原创性
		技术效果
		引用情况
	技术替代性	替代技术方案数量
	技术适用范围	技术领域数量
		技术领域范围
	技术独立性	配套技术依存度
	技术成熟度	技术发展阶段
技术领域发展态势	技术生命周期	
经济价值	剩余经济寿命	剩余经济寿命
	竞争态势	防御性
		控制力
		竞争对手情况
	市场应用情况	销售收益
		市场占有率
		政策适应性
	专利运营状况	转让许可情况
		融资保险情况
		诉讼仲裁情况

来源:

国家标准信息公共服务平台:

<https://std.samr.gov.cn/gb/search/gbDetailed?id=027A6096AF4B643EE06397BE0A0A0867>

国家标准全文公开系统:

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?heno=21AB9B32C8B01C3849AF785EBEFC357B>

数据组织

◇ ANSI/NISO Z39.106-2023, Standard Terminology for Peer Review

【看点】美国国家信息标准组织于2023年7月5日推出了《ANSI/NISO Z 39.106-2023 同行评议标准术语 (Standard Terminology for Peer Review)》，确定并标准化了同行评议实践中的定义和术语，以帮助越来越多的出版商使用开放的同行评议模型来调整术语。

【摘要】出版商使用的同行评议术语将有助于使文章和期刊的同行评议过程更加透明，并使社区能够更好地评估和比较不同期刊之间的同行评议实践。

国际科学、技术和医学出版商协会（STM, the International Association of Scientific, Technical and Medical Publishers）认识到支持行业确保同行评议更大的透明度和开放性是开放科学的重要组成部分，需要协调和沟通同行评议过程的离散元素定义，以便社区成员（包括作者、评论家、编辑以及读者）都能快速轻松地识别如何更有效地参与学术内容的创建和认证。

《同行评议标准术语》提供了一套简化的术语，将使采用该术语的期刊的作者、审稿人和读者对同行评议过程更加透明，也将使社会各界更容易比较不同出版物的同行评议流程。该标准是针对文章和期刊出版物制定的，NISO 常设委员会将来可能会涉及图书以及其他同行评议的研究内容。



A Standard Terminology for Peer Review

1 identity transparency:

This category describes the extent to which identities of participants are made visible to each other during the review process. Identities not made visible during the process may be made visible at publication on the article page (see section 3, review information published).

TYPE	DESCRIPTION
all identities visible	Reviewer identity is visible to author, author identity is visible to reviewer, reviewer and author identity is visible to (decision-making) editor
single anonymized	Reviewer identity is not made visible to author, author identity is visible to reviewer, reviewer and author identity is visible to (decision-making) editor
double anonymized	Reviewer identity is not made visible to author, author identity is not made visible to reviewer, reviewer and author identity is visible to (decision-making) editor
triple anonymized	Reviewer identity is not made visible to author, author identity is not made visible to reviewer, reviewer & author identity is not made visible to (decision-making) editor

2 reviewer interacts with:

This category relates to direct interaction or exchange of information (e.g., via submission systems or email) during the peer review process. Multiple types of this category may be selected where applicable. Whatever is communicated about the review process after publication is included in section 3, review information published.

TYPE	DESCRIPTION
editor	Communication between editor and reviewer (traditional model). Also known as 'independent review'. Identities can be anonymized or visible
other reviewers	Direct interaction/collaboration (e.g., via submission system or email) between reviewers, or the possibility to receive and/or comment on each other's reports before reviewer makes recommendation to the editor. Identities can be anonymized or visible
authors	Direct interaction/collaboration (e.g., via submission system or email) between author and reviewer before reviewer makes recommendation to the editor. Identities can be anonymized or visible

3 review information published:

This relates to information that is published about the review process on the article page. Select and list the items that are applicable.

TYPE	DESCRIPTION
none	No information about the review process or editorial decision process is published
review summaries	Can be summaries or parts of the reviews, or a summary of the review process
review reports	Full content of the reviewer reports is published
review reports (author opt in)	Full content of the reviewer reports is published if the corresponding author opts for this
review reports (reviewer opt in)	Full content of the reviewer reports is published if the reviewer(s) opt(s) for this
submitted manuscript	The version of the manuscript that the author submitted for peer review is published
submitted manuscript (author opt in)	The version of the manuscript that the author submitted for peer review is published if the corresponding author opts for this
author/editor communication	Including editor decision letter and reviewer responses (rebuttals)
reviewer identities	Identities of the reviewers are published
reviewer identities (reviewer opt in)	The identities of the reviewers are published if the reviewers opt for this
editor identities	Identities of the handling editors

4 post publication commenting:

Relates to comments on the online published version of the version of record on the publishing platform, and does not include possible integrations with third party platforms (e.g., PubPeer). Article types such as comment / reply / letter are not considered post publication commenting as they are stand-alone publications. Only use this category when applicable.

TYPE	DESCRIPTION
open	Commenting open to anybody. Can be anonymous, require signing in and/or registration (e.g., via ORCID)
on invitation	Only editor- (or publisher-) selected and/or invited individuals can comment on the article post publication

Use of terminology

As an example, the description of a (traditional) review process to authors (e.g., on the Guide for Authors) would be:


identity transparency: single anonymized
reviewer interacts with: editor
review information published: none

Publishers can include links in these descriptions which lead to a page where the terms are explained. In case journals allow authors to choose between review models, all the options should be listed, e.g.


identity transparency: single anonymized, double anonymized

It should be actively communicated when post publication commenting is adopted. For example:

identity transparency: all identities visible
reviewer interacts with: editor, other reviewers
review information published: review reports, reviewer identities
post publication commenting: open



Advancing trusted research



More information:
[10.3789/niso.z39106-2023](https://www.niso.org/publications/z39106-2023-peerreview)

图 3 同行评议的标准术语

来源:

美国国家信息标准组织 (National Information Standards Organization) :

<https://www.niso.org/publications/z39106-2023-peerreview>

<https://www.niso.org/standards-committees/peer-review-terminology>

https://groups.niso.org/higherlogic/ws/public/download/28900/Z39.106-2023-Peer_Review_Terminology.pdf

◇ GB/T 42813-2023 数据论文出版元数据

【看点】国家标准化管理委员会于2023年8月6日发布并实施了国家标准《数据论文出版元数据》，该标准由全国科技平台标准化技术委员会（TC486）归口，主管部门为科学技术部。

【摘要】该标准规定了数据论文出版元数据的内容与结构，包括数据集元数据、数据论文元数据及其描述方法、扩展类型与规则。该标准按照GB/T 30522规定的摘要表示方式定义和描述数据论文出版元数据，包括：定义、英文名称、数据类型、值域、短名、注解等，用于数据论文及其关联数据集的描述、保存、在线出版和共享服务等。

数据论文出版元数据由数据集元数据及数据论文元数据组成：

1、数据集元数据，包括：

- 数据集基本信息（9个元数据元素和2个元数据实体）：标识符、标题、摘要、关键词、范围（元数据实体）、语种、文件内容、基金项目、数据量、数据格式、数据集作者（元数据实体）；
- 数据集出版信息（3个元数据元素）：发布日期、出版期刊、版本信息；
- 数据集服务信息（5个元数据元素）：数据集引用格式、数据集共享许可协议、数据集使用声明、数据集下载地址、数据论文访问地址。

2、数据论文元数据，包括：

- 数据论文内容信息（12个元数据元素和1个元数据实体）：标识符、标题、摘要、关键词、数据集基本信息、引言、数据采集和处理方法、数据样本描述、数据质量控制和评估、数据使用方法和建议、参考文献、致谢、数据论文作者（元数据实体）；
- 数据论文出版信息（6个元数据元素）：收稿日期、同评日期、录用日期、出版日期、版本信息、出版期刊；
- 数据论文服务信息（4个元数据元素）：数据论文引用格式、数据论文下载地址、数据论文共享许可协议、数据集访问地址。

数据集元数据描述数据集，数据论文元数据描述数据论文。数据论文描述数据集内涵与特征，经过专家评审在数据期刊正式发布，其与数据集具有一致性和关联性。

来源：

全国标准信息公共服务平台：

<https://std.samr.gov.cn/gb/search/gbDetailed?id=02814EA35DFFC92EE06397BE0A0AEAA9>

国家标准全文公开系统：

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=0A49B2D53C0489CA8F669B36EC1AEF35>

◇ GB/T 30523-2023 科技资源核心元数据

【看点】 国家标准化管理委员会于 2023 年 8 月 6 日发布并实施了国家标准《科技资源核心元数据》，该标准由全国科技平台标准化技术委员会（TC486）归口，主管部门为科学技术部。

【摘要】 科研仪器设施、科学数据、生物种质与实验材料等科技资源是科技创新的物质基础，具有公益性、基础性、战略性、多样性的特点。科技资源规模、质量和共享利用水平是影响国家科技基础能力的重要因素。为了持续做好科技资源的建设积累，加强科技资源的规范化管理，提高科技资源的共享利用水平，需要对科技资源进行统一规范描述。

该标准考虑了各类科技资源的不同形态，面向各级各类科技平台建设运行和资源共享的实际需求，采用摘要描述方法，规定了科技资源核心元数据的描述方法、组成、元素、实体，以及核心元数据扩展的类型与规则。该标准适用于科技资源核心元数据的编目、归档、建库、发布、共享、交换和查询、领域扩展等，有助于实现科技资源元数据的有效管理和共享利用，有助于科技资源信息化管理。

科技资源核心元数据包括 12 个元数据元素和 3 个元数据实体：

1、元数据元素：标识符、资源名称、描述、关键词、生成日期、注册日期、最新发布日期、学科分类、主题分类、知识产权类别、资源使用许可、资源访问地址。

2、元数据实体：共享方式、提供方信息、服务方信息。

来源：

全国标准信息公共服务平台：

<https://std.samr.gov.cn/gb/search/gbDetailed?id=02814EA35DF8C92EE06397BE0A0AEAA9>

国家标准全文公开系统：

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=87EF5C962FAA9224E7ACF8FAEC9026A4>

◇ GB/T 42713-2023 信息与文献 参与者名称标识符

【看点】国家标准化管理委员会于2023年8月6日发布了国家标准《信息与文献 参与者名称标识符》，该标准由全国信息与文献标准化技术委员会（TC 4）归口，将于2024年3月1日正式实施。

【摘要】该推荐性标准由中国工程院战略咨询中心、浙江大学、浙江理工大学、中国科学技术信息研究所及相关有限公司共同起草，定义了参与者公开身份的标识规范，旨在提供一种识别方法，区分参与者在数字环境中进行创作、加工、管理、内容发布等一系列产出时的公开身份，并有效管理其产出作品。

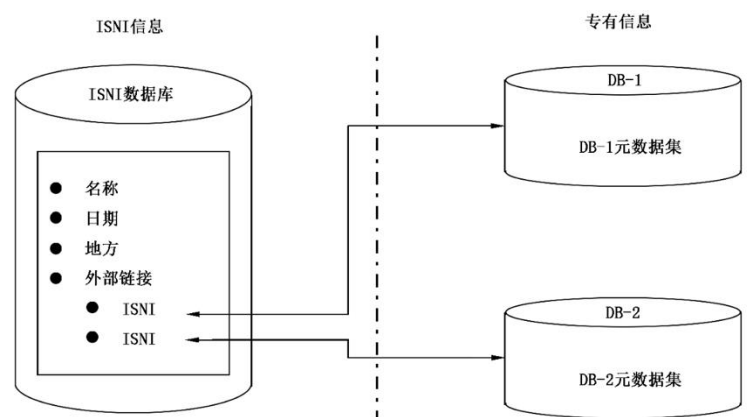


图 4 ISNI 数据库与其他专用数据库的集成示例

该标准规定了参与跨领域创造性活动的各方公开身份的唯一标识，用于区分同一参与者的不同身份，避免因不同公开身份的参与者使用相同名称而造成混淆。该标准不适用于直接提供参与者公开身份详细信息的访问路径，但可以提供指向保存此类信息的其他系统的链接，创建的标识符适用于识别数据库中参与者的公开身份，支持内容管理中的信息交换。

该标准修改采用了 ISO 27729:2012《信息与文献 国际标准名称标志符（ISNI）》，故涉及版权保护问题，正式标准出版物，请联系中国标准出版社。

来源：

全国标准信息公共服务平台：

<https://std.samr.gov.cn/gb/search/gbDetailed?id=027A6096B035643EE06397BE0A0A0867>

国家标准全文公开系统：

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=911B246D94ACD4759F2D925433B57FCC>

◇ **GB/T 42743-2023 信息与文献 国家科技重大专项档案元数据元素集**

【看点】 国家标准化管理委员会于 2023 年 8 月 6 日发布了国家标准《信息与文献 国家科技重大专项档案元数据元素集》。该标准由全国信息与文献标准化技术委员会（TC4）归口，将于 2024 年 3 月 1 日正式实施。

【摘要】 该标准由中华人民共和国科学技术部、中国科学技术信息研究所、国家档案局、北京大学、清华大学联合起草，是按照科技部于 2017 年颁布的《国家科技重大专项（民口）档案管理规定》要求制定的。规定强调了督查与落实重大专项档案收集整理与集中管理情况，组织建设统一的重大专项档案管理服务平台，该标准的制定正是为档案管理服务平台的建设提供支撑。

该标准对我国十多年来科技重大专项档案管理现状进行了深入调查研究，总结了我国科研项目档案信息化建设和档案管理的经验和成果，并针对重大专项档案管理过程中档案文件的真实性、完整性、有效性与安全性的实际需要，参考了国际、国家档案文件管理系列标准规范，对重大专项档案的元数据进行实体分类，明确每类实体的元数据构成要素及其相互关联关系；同时，对每个元数据元素的规则进行描述，明确元数据元素的属性，为档案管理服务平台提供数据层面的设计依据和管理规范。

表 2 元数据实体类型及描述

中文名称	英文名称	描述
文件实体	document entity	描述重大专项档案本身的元数据集合，包含文件实体标识元数据和文件实体描述元数据两个子集
项目实体	project entity	描述重大专项项目管理活动与流程的元数据集合，包含项目实体管理元数据和项目实体描述元数据两个子集
业务实体	bussiness entity	描述电子档案得以形成以及管理活动的元数据集合，包括电子档案的创建、收集、归档、转换、迁移、处置等业务活动
责任者实体	agent entity	描述与重大专项档案相关的机构人员实体以及负责实施电子档案管理活动的个人或组织的元数据集合
实体关系	entity relationship	描述重大专项档案形成及业务活动中各个实体间关系的元数据集合

该标准适用于国际科技重大专项文件和档案管理系统的建设以及档案管理工作业务的开展。其他国家科技计划档案数据资源的管理可参照执行。

来源：

国家标准信息公共服务平台：

<https://std.samr.gov.cn/gb/search/gbDetailed?id=027A6096B03C643EE06397BE0A0A0867>

国家标准全文公开系统:

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=4CF4B6F5A6389B7C89CC841DFE81BF02>

◇ GB/T 42093 标准文档结构化 元模型

【看点】国家标准化管理委员会于2022年12月30日发布了国家标准《标准文档结构化 元模型》，并于2023年7月1日正式实施。该标准由全国电子业务标准化技术委员会（TC83）归口，主管部门为国家标准化管理委员会。

【摘要】标准在出版、发行、服务时，通常是以标准文本作为一个单元，以纸质印刷文档或数字化电子文档载体形式来呈现，这种形式不能满足读者对标准文档内容进行分析、比对或拆分、重构以及快速查询等诸多需求。为从根本上解决这一问题，就需要在标准编辑、出版环节解决，实现标准文档结构化编写、存储、发布。

鉴于未来不同功能类型标准的结构化元模型描述对象不同，其使用的场景亦不同，GB/T 42093《标准文档结构化 元模型》拟按标准功能类型应用划分部分，由以下部分构成：

- 第1部分：全文；
- 第2部分：技术指标；
- 第3部分：术语标准；
- 第4部分：符号标准；
- 第5部分：分类编码；
- 第6部分：试验方法标准；
- 第7部分：规范标准；
- 第8部分：规程标准；
- 第9部分：指南标准。

《标准文档结构化 元模型 第1部分：全文》是GB/T 42093系列标准的第一部分，为基础通用部分，是其他功能类型标准元模型构建的基础。该标准从三个维度描述了标准文档结构化的全文元模型，即层次元模型、要素表述形式元模型、内容要素元模型，规定了各维度的关系及元模型元素的属性，适用于标准文档结构化、信息化开发和应用。

表 3 内容要素与要素表述形式及层次之间的关系

内容要素	要素允许表述形式	要素允许层次结构
封面	标明文件信息	无
目次	列表（自动生成内容）	无
前言	条文、注、脚注、指明附录	段、列项
引言	条文、图、表、数学公式、脚注、指明附录	条、段、列项
重要提示	条文	段
范围	条文、表、注、脚注	章、条、段、列项
规范性引用文件	清单、注、脚注	章
术语和定义	条文、图、数学公式、示例、注、引用、提示、指明附录	章、条 ^a 、段、列项
符号和缩略语		
分类和编码		
系统结构		
总体原则		
总体要求		
核心技术要素		
其他技术要素		
参考文献	清单、脚注	无
索引	列表（自动生成内容）	无

^a 其他内容要素段条可以是有标题条也可以是无标题条，这里的条专指有标题条

《标准文档结构化 元模型 第 2 部分：技术指标》是 GB/T 42093 系列标准的第二部分，依据 UML 的思想，按照对象、属性、关系，描述产品技术要求的元模型，其主要内容要素为对象、对象属性、影响因素、技术指标内容。

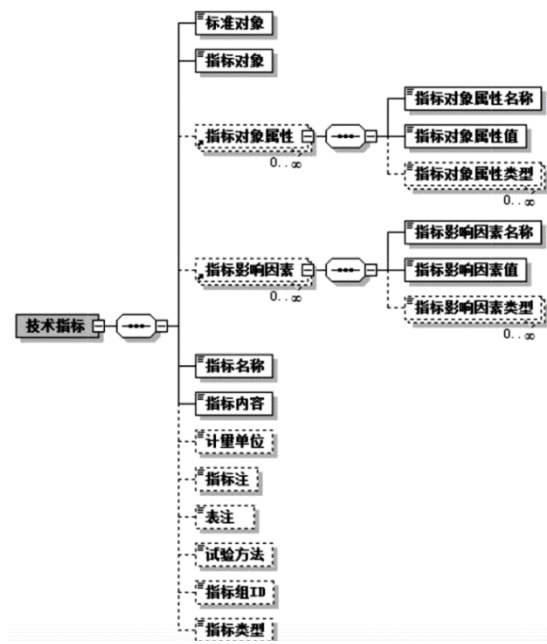


图 5 技术指标元模型

该标准适用于标准文档结构化开发与应用，不适用于标准中要求条款的功能指标的结构化描述。

来源：

全国标准信息公共服务平台：

<https://std.samr.gov.cn/gb/search/gbDetailed?id=F159DFC2A7B247EFE05397BE0A0AF334>

国家标准全文公开系统：

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=264AB19765573BADC8E9F6AF820253F5>

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=2FBA895437A9E1D380B6E7376B1DBE59>

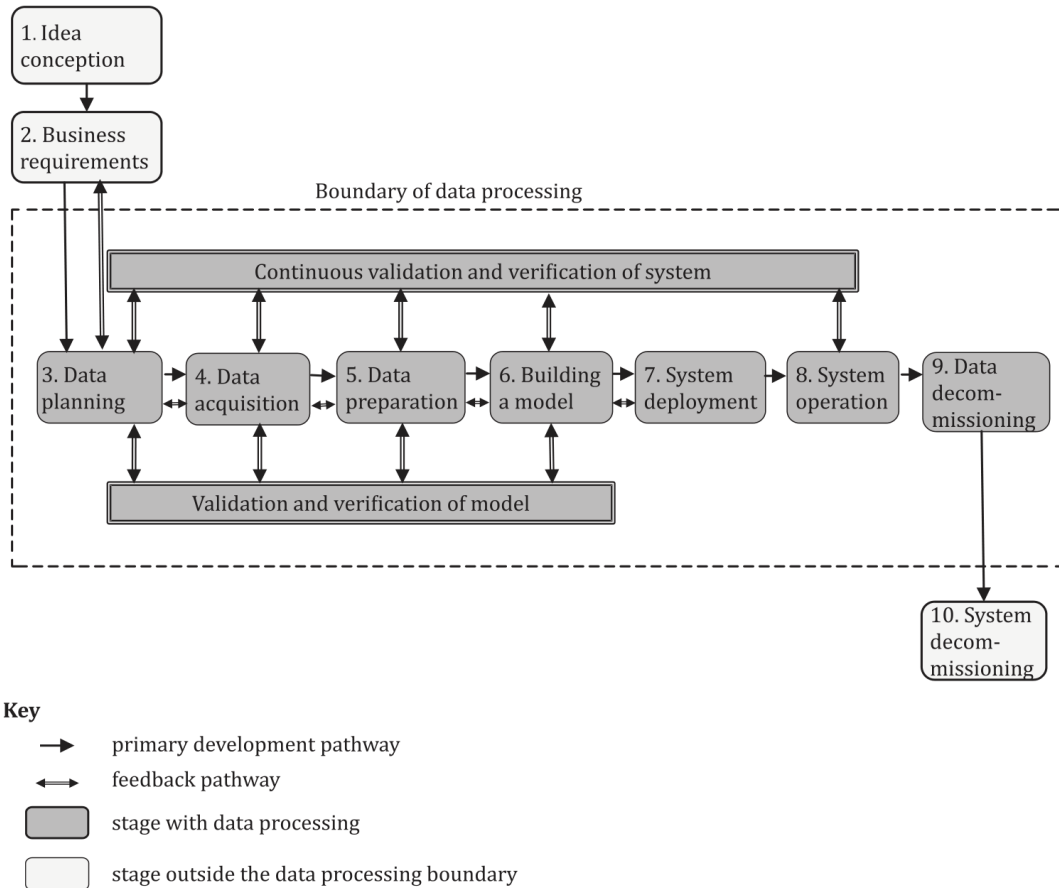
知识技术

◇ ISO/IEC 8183:2023 Information technology - Artificial intelligence - Data life cycle framework

【看点】国际标准化组织于 2023 年 7 月发布了《ISO/IEC 8183:2023 信息技术 人工智能 数据生命周期框架（Information technology - Artificial intelligence - Data life cycle framework）》，定义了整个人工智能系统生命周期中数据处理的相关步骤，包括数据获取、创建、开发、部署、维护和停止使用。

【摘要】当前，各种类型、规模和性质的组织都在采用人工智能系统，而数据对于人工智能系统的开发和运营至关重要。在人工智能系统领域，许多数据正在被使用和考虑用于不同的目的，如数据质量、数据偏差、数据治理、人工智能系统的开发和使用等。构建起总体的数据生命周期框架，可以辅助没有先验知识、背景和专业能力的人正确解释不同的数据生命周期，从而防范应用过程中可能出现的数据风险。

该标准不仅提供了数据生命周期概述，还描述了数据生命周期框架，并提供了相关信息，适用于在开发和人工智能系统过程中使用数据的所有组织。



NOTE 1 The single-headed arrows depict a linear path through the life cycle stages, while the double-headed arrows show feedback paths between life cycle stages.

NOTE 2 The verification and validation of the model refers to the internal development process, whose output is a model. The validation and verification of the system refers to the system as a whole, extending through its entire period of operation.

图 6 数据生命周期框架

来源:

国际标准化组织: <https://www.iso.org/standard/83002.html>

◇ **GB/T 42888-2023 信息安全技术 机器学习算法安全评估规范**

【看点】 国家标准化管理委员会于 2023 年 5 月 6 日发布了国家标准《信息安全技术 机器学习算法安全评估规范》，该标准由全国信息安全标准化技术委员会（TC260）归口，将于 2024 年 3 月 1 日正式实施。

【摘要】 该标准首先声明了机器学习算法的安全原则包括公平合理、公开可解释以及诚实可信，机器学习算法安全要求分为基本级和增强级，后分别介绍了机器学习算法技术和机器学习算法服务的安全要求和评估方法。除此之外，该标准还介绍了机器学习算法安全评估的流程要求包括：

- (1) 应设置算法安全评估机制；

- (2) 机器学习算法技术提供者进行安全评估时，应按照 5.1 所述安全要求和 5.2 所述评估方法开展评估工作；
- (3) 机器学习算法服务提供者进行安全评估时，应按照 6.1 所述安全要求和 6.2 所述评估方法开展评估工作；
- (4) 算法推荐服务提供者进行安全评估时，在按照第 6 章开展评估外，还应按照附录 A 所述安全要求和附录 B 所述评估方法开展评估工作。

该标准规定了机器学习算法技术和服务的安全要求和评估方法，以及机器学习算法安全评估流程，适用于指导机器学习算法提供者保障机器学习算法生存周期安全以及开展机器学习算法安全评估，也可为监管评估提供参考。

来源：

国家标准信息公共服务平台：

<https://std.samr.gov.cn/gb/search/gbDetailed?id=027A6096AFA9643EE06397BE0A0A0867>

国家标准全文公开系统：

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=E7170BA58AE37AACF4170242EFD25183>

◇ **GB/T 42777-2023 基于文本数据的金融风险防控 知识图谱构建技术框架指南**

【看点】国家标准化委员会于 2023 年 8 月 6 日发布并实施了国家标准《基于文本数据的金融风险防控 知识图谱构建技术框架指南》，该标准由全国金融标准化技术委员会（TC180）归口，主管部门为中国人民银行。

【摘要】该标准给出了基于文本数据的金融风险防控知识图谱与风险防控概述，提供了知识图谱构建、基于知识图谱的风险防控应用的指南，适用于金融机构开展基于文本数据的金融风险防控知识图谱构建工作。

知识图谱将数据加工为知识，并通过智能分析挖掘，提炼形成风险防控的方法和技术实施路径。基于知识图谱的风险防控整体框架按照数据流转方向自下而上包含基础支撑层、数据源层、知识抽取与加工层、知识管理层、知识挖掘与分析层、风险防控应用层和知识服务交互层。

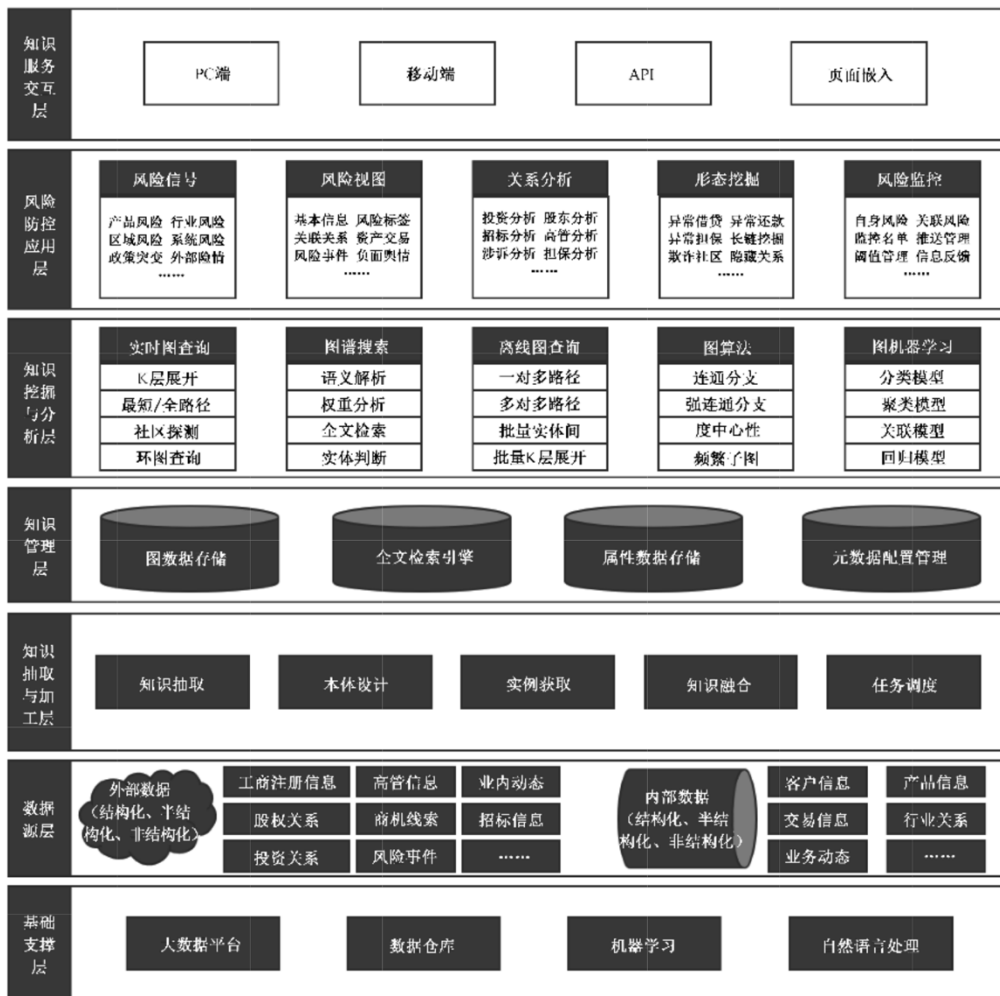


图 7 基于知识图谱的风险防控整体架构

来源：

国家标准信息公共服务平台：

<https://std.samr.gov.cn/gb/search/gbDetailed?id=02814EA35DF7C92EE06397BE0A0AEEA9>

国家标准全文公开系统：

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=A4511FD90354C5CFA085E8DD2F33C17A>

◇ GB/T 41818-2022 信息技术 大数据 面向分析的数据存储与检索技术要求

【看点】国家标准化管理委员会于 2022 年 10 月 14 日发布了国家标准《基信息技术 大数据 面向分析的数据存储与检索技术要求》，并于 2023 年 5 月 1 日正式实施。该标准由全国信息技术标准化技术委员会（TC28）归口，主管部门为国家标准化委员会。

【摘要】在大数据场景中，数据分析是通往大数据应用的首要步骤，需要简便、快捷和准确的数据组织和管理方法，这种组织和管理主要体现在存储和检索

方面。检索活动与数据分析中的数据提取直接关联，存储活动为数据分析提供便捷的数据组织。

该标准规定了支撑多应用融合分析的大数据列式存储与检索技术要求，描述的数据表为大数据场景下面向分析的数据存储和检索提供数据组织和管理方面的支持，适用于面向分析的数据存储与检索的系统或子系统的设计、开发和使用。

数据表是对数据进行组织和管理的一种数据分层次编排结构和多层结构，其组成部分包括：数据子表（简称“子表”）、数据行组（简称“行组”）、数据列组（简称“列组”）和数据页。

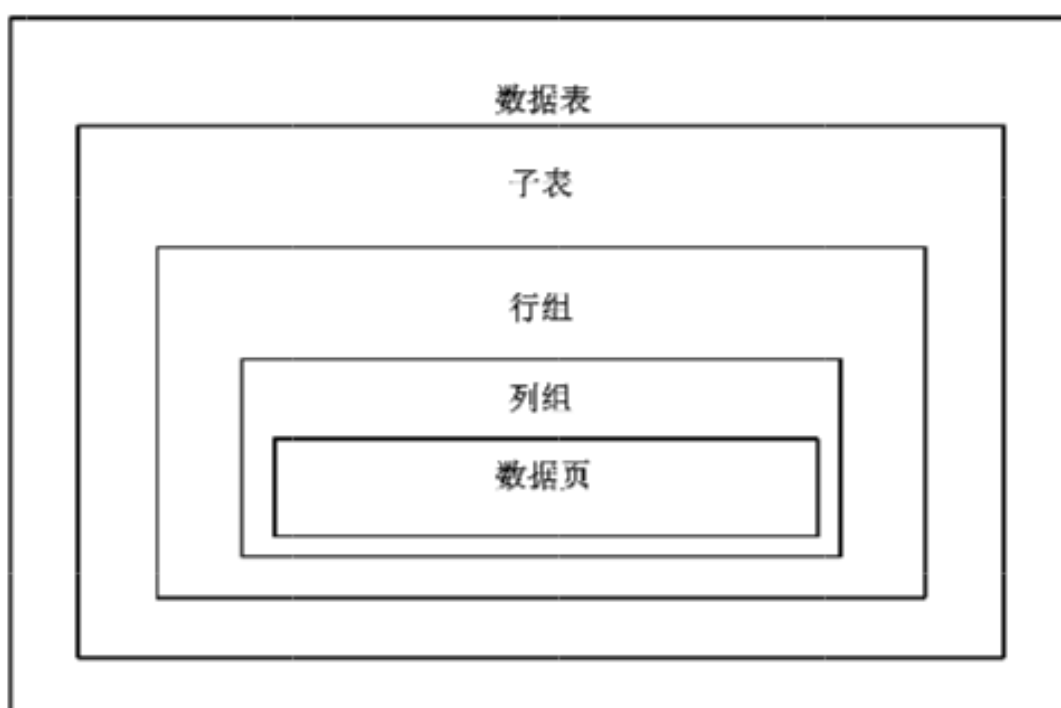


图 8 数据表结构图

一张数据表包含 1 张或多张子表，主要用于对数据进行面向分析的组织和管理工作，数据表的元数据独立于数据表外存储。这些组成部分统称为数据单元。

以旅游行业大数据为例，旅游数据存储表结构如图 9 所示：

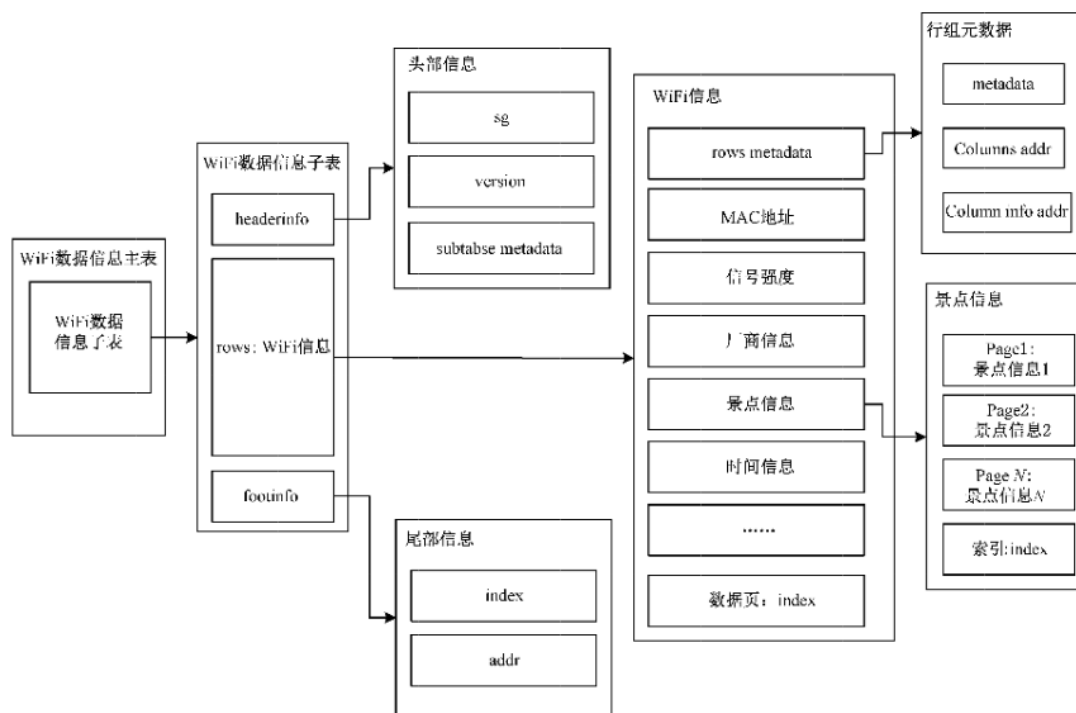


图 9 旅游数据存储表结构图

来源:

国家标准信息公共服务平台:

<https://std.samr.gov.cn/gb/search/gbDetailed?id=EB58F4DA926BB2A2E05397BE0A0A7D33>

国家标准全文公开系统:

<https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?heno=F8B193B00D98592FAF7DD6AE003E9830>