

团 体 标 准

T/XXXX XX—XXXX

科技成果数字画像系统架构

Science and technology achievements digital profile system architecture

(征求意见稿)

XXXX—XX—XX 发布

XXXX—XX—XX 实施

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 总体架构	1
5 基础设施层	2
5.1 计算设施	3
5.2 网络设施	3
5.3 存储设施	3
5.4 安全设施	3
6 数据资源层	3
6.1 科技成果数据	3
6.2 机构数据	4
6.3 技术需求数据	4
6.4 用户数据	4
7 算法服务层	4
7.1 特征提取算法	4
7.2 画像生成算法	4
7.3 匹配推荐算法	5
8 业务应用层	5
8.1 成果管理	5
8.2 需求管理	5
8.3 匹配推荐	6
8.4 态势分析	6
参考文献	8

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。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国标准化研究院提出。

本文件由中国技术市场协会归口。

本文件起草单位：中国标准化研究院、中国技术市场协会、北京信息科技大学等。

本文件主要起草人：暂略。

科技成果数字画像系统架构

1 范围

本文件规定了科技成果数字画像系统架构，包括总体要求、性能和功能要求。

本文件适用于科技成果数字画像系统架构的设计、功能开发和实施应用。

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 32399 信息技术 云计算 参考架构

GB/T 44731 科技成果评估规范

3 术语和定义

GB/T 32399、GB/T 44731界定的以及下列术语和定义适用于本文件。

3.1

科技成果 science and technology achievements

通过科学研究与技术开发产生的具有一定学术价值或应用价值的成果。

注1：科技成果通常分为基础研究成果、应用研究成果、技术开发和产业化成果三种类型。

注2：科技成果的主要技术载体，又称“主成果”或“主交付物”，通常包括但不限于(所有物质态的)硬件、软件、工艺、方法等。

注3：科技成果的辅助技术载体，又称“副成果”或“副交付物”，通常包括但不限于标准、专利、论文、报告、软著、专著、图纸、方案等。

[来源：GB/T 44731—2024，3.1]

3.2

科技成果数字画像 science and technology achievements digital profile

基于科技成果相关信息提取描述科技成果特征的标签数据集，通过自动化方式处理分析，并以可视化方式展示指定维度的科技成果特征的方法。

3.3

科技成果数字画像系统 science and technology achievements digital profile system

使用科技成果数字画像（3.2）方法，提供指定维度的科技成果特征展示、指定技术需求的科技成果匹配推荐等功能的软件系统。

3.4

架构 architecture

通过系统元素、元素间的关系，以及系统设计和进化原则体现出来的一个系统在其环境中的结构和组织方式。

[来源：GB/T 32399—2024，3.1]

4 总体要求

4.1 层次结构

科技成果数字画像系统的层次结构见图1，包括以下四个层次：

- 1) 基础设施层，提供科技成果数字画像系统建设所需的基础设施设备，包括计算设施、网络设施、存储设施、安全设施等。
- 2) 数据资源层，提供科技成果数字画像系统所需的数据以及对数据的预处理、存储、分析等资源支撑，主要数据资源包括科技成果数据、技术需求数据、机构数据和用户数据等。
- 3) 算法服务层，提供科技成果数字画像系统各业务应用所需的各类算法引擎，包括特征提取算法、画像生成算法、匹配推荐算法等。
- 4) 业务应用层，为满足科技成果数字画像的常见业务需求，提供模块化的业务功能或解决方案，包括成果管理、需求管理、匹配推荐、态势分析等。



图1 科技成果数字画像系统层次结构

4.2 画像要求

4.2.1 数据来源真实可靠。科技成果数字画像的基础数据应来源于客观数据、成果持有方或其授权的第三方提供的真实数据。

4.2.2 画像特征维度丰富。科技成果数字画像应能实现不少于3种维度的成果特征提取和画像的展示，例如区域、行业领域、机构或机构类型等。

4.2.3 多层次分类展示。科技成果数字画像应能够实现2个以上层级的成果筛选和画像的展示，例如特定区域特定领域的科技成果列表或画像。

5 基础设施层

5.1 计算设施

计算设施的基本要求如下：

a) 应提供分布式计算能力；

b) 应通过对一组计算机进行节点设置、任务调度，实现负载均衡、高可用性和高性能计算；

c) 应提供较强的计算加速能力；

d) 宜采用公有云与私有云结合的混合云解决方案，以私有云确保系统稳定、可靠运行，以公有云确保系统快速部署、高效运维。

5.2 网络设施

网络设施的基本要求如下：

a) 应具备较低的通信延迟和较高的吞吐量，支持系统的快速响应和高效处理数据；

b) 应具备高度可靠性，确保数据传输和通信的连续、可靠；

c) 应具备高度可扩展性，能够增加网络带宽和处理能力，以适应业务需求的不断变化。

5.3 存储设施

存储设施的基本要求如下：

a) 应提供高速数据读写和低延迟的数据访问能力；

b) 应提供可扩展的存储空间，并支持多种存储介质和类型；

c) 应提供冗余的存储和备份策略，避免数据因为硬件故障或其他意外事件而丢失。

5.4 安全设施

安全设施的基本要求如下：

a) 应有防火墙，对网络数据流量进行过滤和监控，以保护网络免受恶意攻击；

b) 应具有身份认证和访问控制模块，通过验证用户身份和对用户访问权限的控制来确保信息系统的安全；

c) 应具有安全日志管理模块，并配备权限管理功能，以监测和记录系统的操作和事件，识别和解决潜在的安全问题，并仅授权给特定的用户或角色（如管理员）进行访问；

d) 应具有防病毒措施，使用防病毒软件等对信息系统进行实时监测和防护，保障系统不受病毒和恶意软件的攻击。

6 数据资源层

6.1 科技成果数据

6.1.1 科技成果基本信息，应包括成果名称、所属领域、成果简介、关键词等。

6.1.2 科技成果知识产权信息，应包括涉及的专利、软件著作权等知识产权的名称、编号、授权时间、

持有方、法律状态等。

6.1.3 科技成果评价信息，应包括技术成熟度、成果所处阶段、成果经济社会价值等。

6.1.4 科技成果转化相关信息，应包括拟转化方式、拟转化金额、所属地区等。

6.1.5 其他信息，应包括相关成果名称、依托项目名称等。

6.2 机构数据

6.2.1 机构基本信息，应包括机构名称、机构类型、所属地区、成立时间、机构规模、行业领域等。

6.2.2 机构成果信息，应包括成果数量、成果类型、成果领域等。

6.2.3 机构需求信息，应包括需求数量、需求类型、拟合作方式、拟投入金额。

6.2.4 其他信息，应包括关联机构名称、关联关系等。

6.3 技术需求数据

6.3.1 需求基本信息，应包括需求名称、所属领域、需求简介、关键词等。

6.3.2 知识产权要求，应包括对授权时间、持有方、法律状态等的要求。

6.3.3 对成果评价信息的要求，应包括技术成熟度、成果所属阶段、成果经济或社会价值等的要求。

6.3.4 其他需求，应包括拟合作方式，拟投入金额以及合作地点等。

6.4 用户数据

6.4.1 用户基本信息，应包括用户名、密码、用户类别、年龄、性别、联系方式、所属地区等。

6.4.2 用户收藏信息，应包括成果列表、需求列表、机构列表等。

6.4.3 用户浏览记录，应包括浏览的成果条目、需求条目和机构等。

6.4.4 用户检索记录，应包括检索的关键词、检索式、检索语句等。

7 算法服务层

7.1 特征提取算法

特征提取算法包括一系列的数据工作，用于将分散在各处的科技成果、技术需求、机构用户等数据资源进行汇聚，通过清洗、转换、去重、融合等操作完成数据的标准化处理，并根据数据特点和业务需求，提取科技成果数据的典型特征，为上层的业务应用提供数据支撑。基本要求如下：

a) 应具备对知识进行抽取的能力，包括对数据资源层数据进行知识抽取，将其中关键的业务本体信息进行抽取；

b) 应具备对数据进行生产的能力，包括对数据进行汇聚、接入、处理、组织的功能，支撑数据资源层各项数据资源的生成；

c) 应具备对数据进行分析的能力，包括对数据进行可视化分析模型构建，按照业务应用组织相关业务逻辑模型。

7.2 画像生成算法

画像生成算法基于知识图谱和可视化技术，针对科技成果评价、科技成果转化、技术需求匹配等业务场景中出现的实体、关系、事件等对象构建点边关系网络，提供面向整个关系网络的重点实体、重点关系、重点网络子图的查询、展现、挖掘、分析等能力的可视化分析工具。基本要求如下：

a) 应具备知识可视化能力，包括利用可视化图表组件对相关数据资源进行可视化大屏的构建；

b) 应具备知识表示的能力，包括通过本体设计工具对标准相关的实体、关系、事件进行表示；

c)应具备知识融合的能力,包括对抽取的业务本体信息按照知识表示方式进行实时知识融合,形成相关知识领域的知识图谱;

d)应具备知识推理的能力,包括对知识信息进行过滤筛选、整合扩展、布局、子图查询等功能;

e)应具备知识服务的能力,包括对查询、推理、挖掘等知识数据的访问API进行管理。

7.3 匹配推荐算法

匹配推荐算法基于自然语言处理、深度学习等技术,针对科技成果或技术需求的所属领域、关键词、技术文本等信息,通过相似度计算对候选项排序,返回最符合需求的匹配成果列表。基本要求如下:

a)应具备关键词检索匹配能力,包括从文本中提取技术关键词和领域关键词,生成关键词列表等;

b)应具备词性句法分析的能力,包括对词性的筛选、句子的拆解分析、同义词替换等;

c)应具备匹配相似度计算能力,包括关键词共现频数、余弦相似度等常见指标的计算;

d)应具备定制扩展匹配的能力,例如考虑地理邻近性、语义相似度等多种因素的匹配推荐能力;

8 业务应用层

8.1 成果管理

8.1.1 成果检索

成果检索提供对科技成果相关信息的一键查询和高级查询功能等。基本要求如下:

a)应具备一键查询能力,包括从基于成果名称、关键词等基本信息的检索;

b)应具备高级查询能力,包括同时对所有基本信息同时检索的能力等;

c)应具备动态筛选能力,包括基于指定的一个或多个字段筛选显示的成果列表等。

8.1.2 成果画像

成果画像提供对科技成果相关信息的可视化展示能力。基本要求如下:

a)应具备成果信息可视化能力,包括利用可视化图表组件对科技成果相关信息进行可视化大屏构建;

d)应具备动态调整能力,包括对成果信息进行过滤筛选、整合扩展、布局、子图查询等功能。

8.1.3 成果标签管理

成果标签管理提供对科技成果特征标签的查询、新增、修改、删除等功能。基本要求如下:

a)应支持角色与权限管理,不同角色仅在权限范围内进行查询、修改、删除、审批等操作;

b)宜支持多级标签功能以及标签体系导入导出功能。

8.1.4 成果数据管理

成果数据管理提供对科技成果基本信息、评价信息等各种信息的查询、新增、修改、删除等功能。基本要求如下:

a)应支持角色与权限管理,不同角色仅在权限范围内进行查询、修改、删除、审批等操作;

b)应支持数据导入导出功能,可以按照指定模板导入或导出一个或一组成果数据;

c)宜具备数据分析与汇总功能,能根据不同字段、标签等对成果数据进行描述性统计、聚类分析等操作。

8.2 需求管理

8.2.1 需求检索

需求检索提供对技术需求相关信息的一键查询和高级查询功能等。基本要求如下:

- a) 应具备一键查询能力，包括从基于需求名称、关键词等基本信息的检索；
- b) 应具备高级查询能力，包括同时对所有基本信息同时检索的能力等；
- c) 应具备动态筛选能力，包括基于指定的一个或多个字段筛选显示的需求列表等。

8.2.2 需求画像

需求画像提供对技术需求相关信息的可视化展示能力。基本要求如下：

- a) 应具备需求信息可视化能力，包括利用可视化图表组件对技术需求相关信息进行可视化大屏构建；
- d) 应具备动态调整能力，包括对需求信息进行过滤筛选、整合扩展、布局、子图查询等功能。

8.2.3 需求标签管理

需求标签管理提供对技术需求特征标签的查询、新增、修改、删除等功能。基本要求如下：

- a) 应支持角色与权限管理，不同角色仅在权限范围内进行查询、修改、删除、审批等操作；
- b) 宜支持多级标签功能以及标签体系导入导出功能。

8.2.4 需求数据管理

需求数据管理提供对技术需求基本信息等各种信息的查询、新增、修改、删除等功能。基本要求如下：

- a) 应支持角色与权限管理，不同角色仅在权限范围内进行查询、修改、删除、审批等操作；
- b) 应支持数据导入导出功能，可以按照指定模板导入或导出一个或一组成果数据；
- c) 宜具备数据分析与汇总功能，能根据不同字段、标签等对需求数据进行描述性统计、聚类分析等操作。

8.3 匹配推荐

8.3.1 成果推荐

成果推荐功能提供对指定技术需求输入的相似成果推荐，一般以列表形式列出排名靠前的一部分科技成果的名称等基本信息。

8.3.2 机构推荐

机构推荐功能提供对当前科技成果或机构有类似科技成果的机构列表。

8.3.3 类似需求

类似需求功能提供与当前技术需求类似的需求基本信息列表，或提供与当前机构具有类似技术需求的机构列表。

8.3.4 智能问答

智能问答功能提供针对当前技术需求输入的模糊搜索，并匹配相应的成果列表、潜在合作机构列表等信息。

8.4 态势分析

8.4.1 领域态势

领域态势功能提供对特定领域的科技成果或技术需求数据分析的可视化展示。

8.4.2 地区态势

地区态势功能提供对特定地区的科技成果或技术需求数据分析的可视化展示。

8.4.3 机构态势

机构态势功能提供对特定机构的科技成果或技术需求数据分析的可视化展示。

8.4.4 时间态势

时间态势功能提供对特定时间的科技成果或技术需求数据分析的可视化展示。

参 考 文 献

- [1] T/CESA 1279—2023 标准信息服务平台 参考架构
 - [2] T/CMSS 0007—2023 科技成果智能化与标准化评价规范
 - [3] T/TMAC 002.F—2021 科技成果评价
-