

同济大学知识图谱项目

构建平台用户手册 V1.0

同济大学

All Rights Reserved

受控文档妥善保管

文档标识

文档名称	同济大学知识图谱项目构建平台用户手册
版本号	<V1.0>
文件状态	[√]正式发布

文档修订历史

版本	日期	描述	文档所有者
V1.0	2025/1/20	创建	
V1.1	2025/6/10	修改	

本次版本文档主要修订更新内容说明

版本	V1.0
1、编制构建平台用户手册	

目录

一、产品介绍	4
二、普通用户使用手册	4
1. 快速上手	4
1.1 创建知识空间	4
1.2 进入知识空间，设计本体	4
1.3 下载知识映射模板	6
1.4 将数据处理进模板中	7
1.5 上传模板文件，进行知识映射	7
1.6 查看图谱	9
2. 产品功能	10
2.1 登录	10
2.2 知识空间	11
2.3 资源管理	12
2.4 本体设计	14
2.5 知识处理	18
2.6 数据管理	21
2.7 图谱可视化	26
2.8 统计分析	31
2.9 抽取算法库	32
2.10 空间配置	33
2.11 日志记录	34

一、产品介绍

知识图谱构建平台是一个专门用于设计、创建、管理知识图谱的工具。平台支持从多种非结构化数据（如 txt、word、pdf、图片、ppt、音视频及网页内容等）和结构化数据（如 mysql 数据库、excel、csv 本地表格）中采集知识。通过领域专家深度参与设计本体、知识抽取、知识映射、知识编辑等步骤完成知识图谱构建。处理后的知识图谱可转化为开放的知识服务供外部应用调用、用户检索查看以实现知识共享。

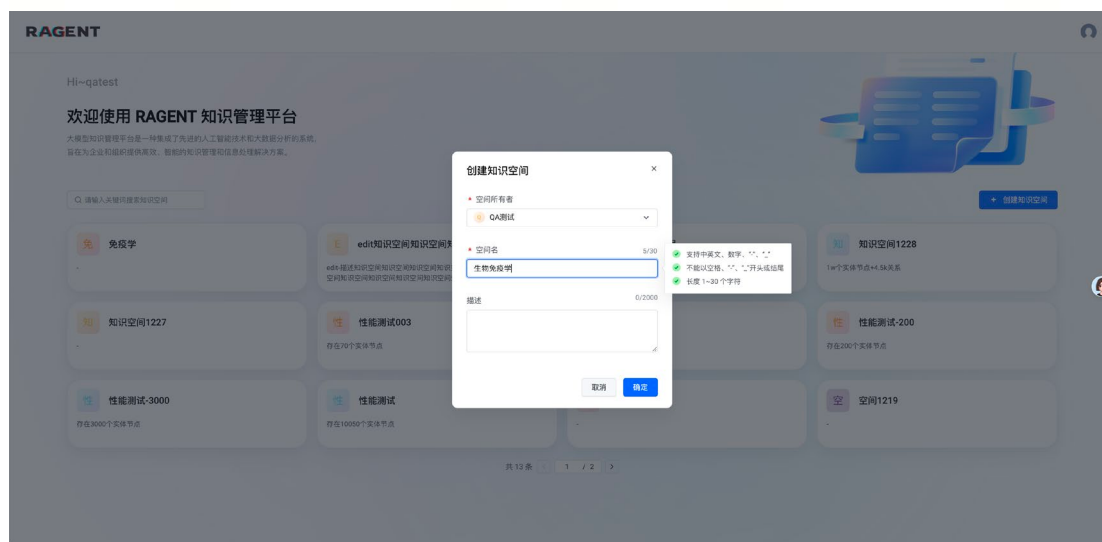
二、普通用户使用手册

1. 快速上手

1.1 创建知识空间

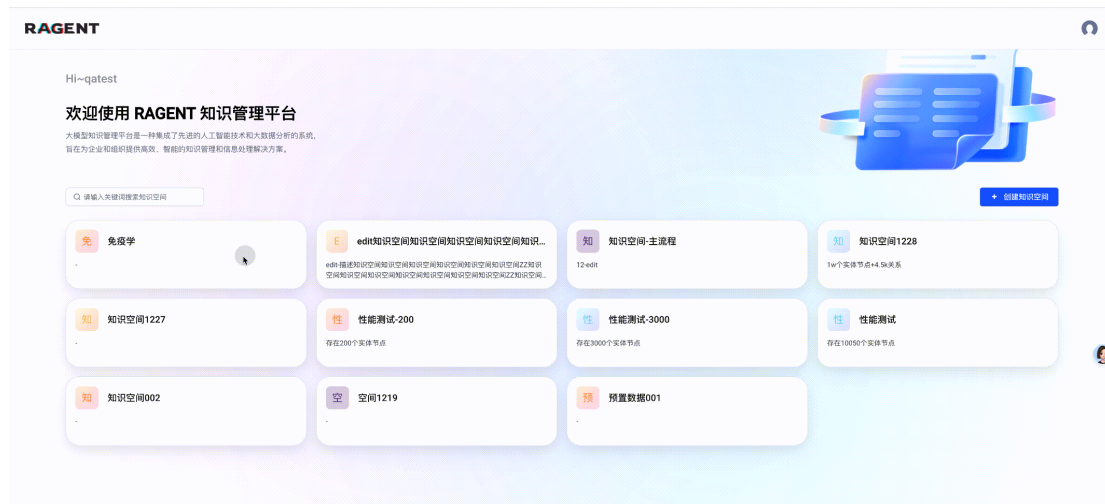
设置空间所有者、空间名、描述等信息进行知识空间创建，知识空间之间数据与权限相互隔离。

- 空间所有者：拥有空间完整权限，可以进一步构建图谱或添加空间协调成员；
- 空间名：一般为知识图谱所在的业务、领域、学科、专业相关内容，如财务、金融、土木工程等；
- 描述：描述当前知识空间包含的内容或者介绍图谱。



1.2 进入知识空间，设计本体

点击空间卡片进入知识空间进行本体设计



先添加实体类型，如：学科；设置该实体类型包含的属性，如：学科介绍。

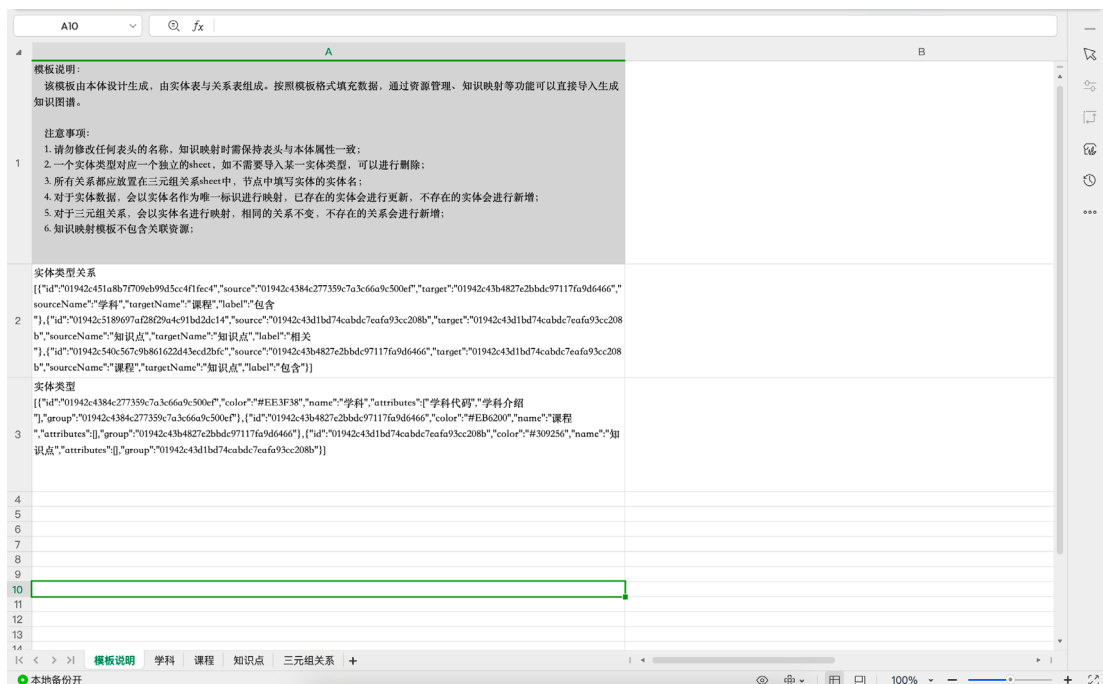
然后添加本体关系，可以通过底部的下拉框添加关系，也可以直接单击实体类型的圆点进行拖拉关系线。



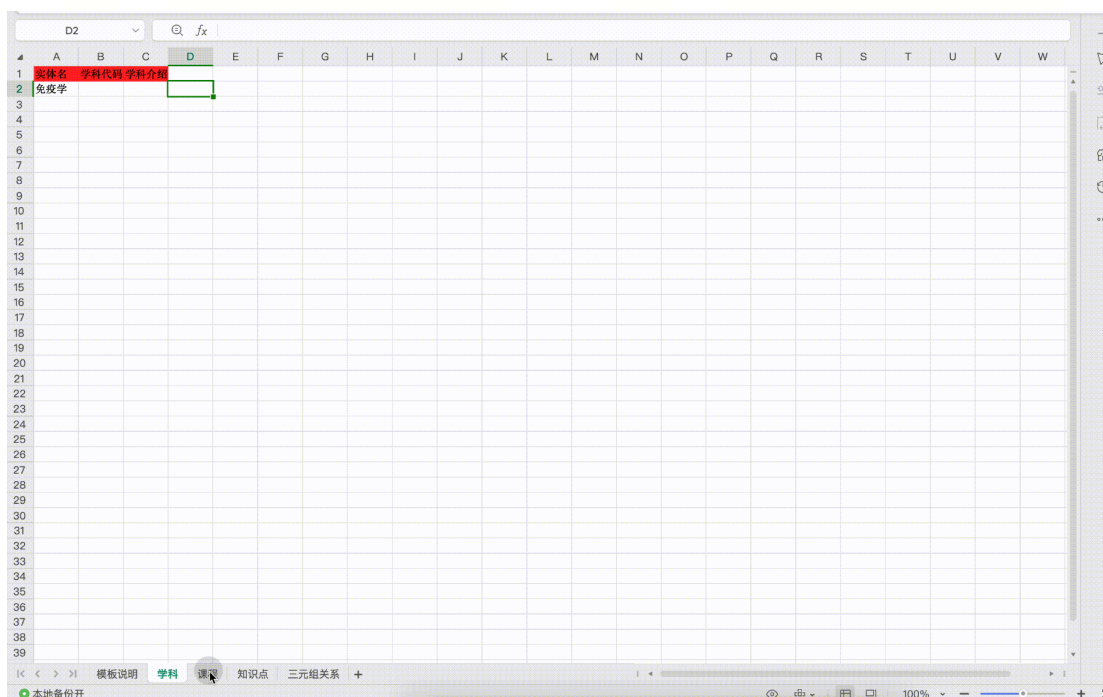
1.3 下载知识映射模板

完成本体设计后，下载知识模板。



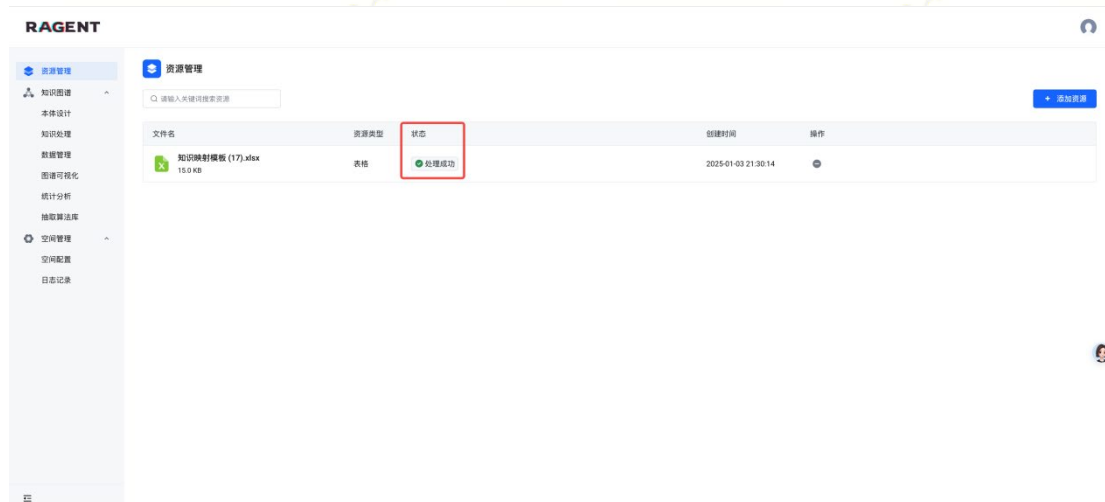
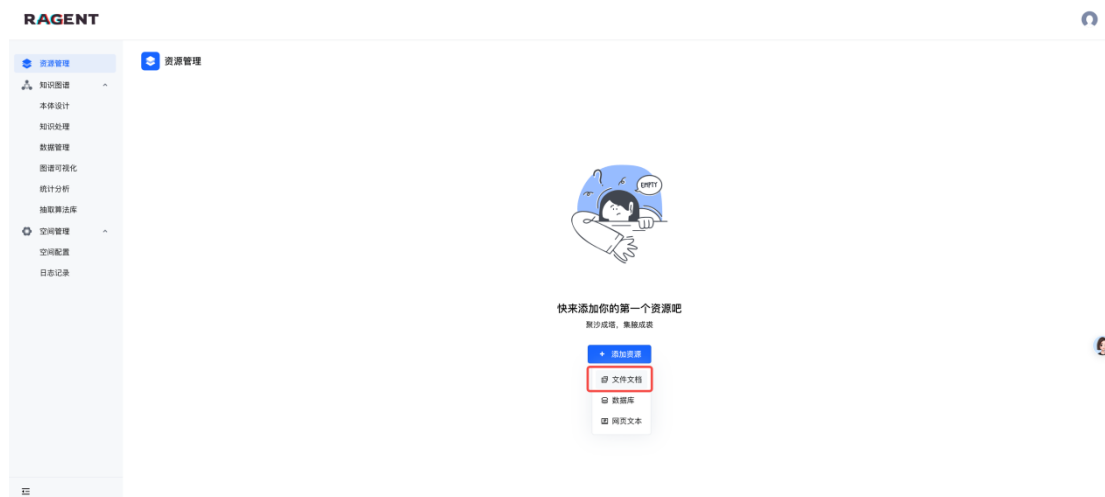


1.4 将数据处理进模板中



1.5 上传模板文件，进行知识映射

将填写好数据的模板文件上传至资源管理中，等待资源处理完成。



点击知识图谱-知识处理，选择通过模板映射



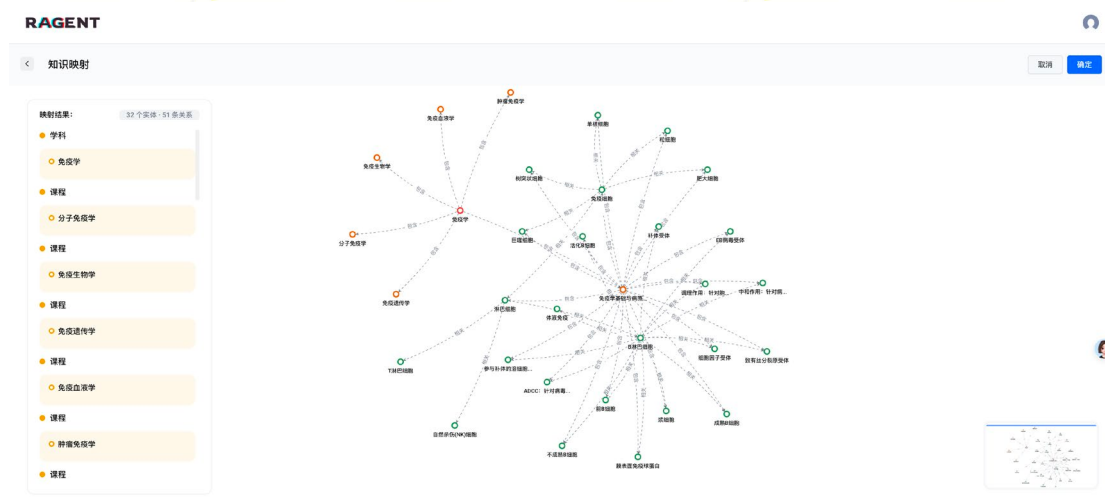
选择刚才上传的模板文件



进行自动化映射，查看映射的结果无误后，点击确定即可以正式将数据写入图谱。

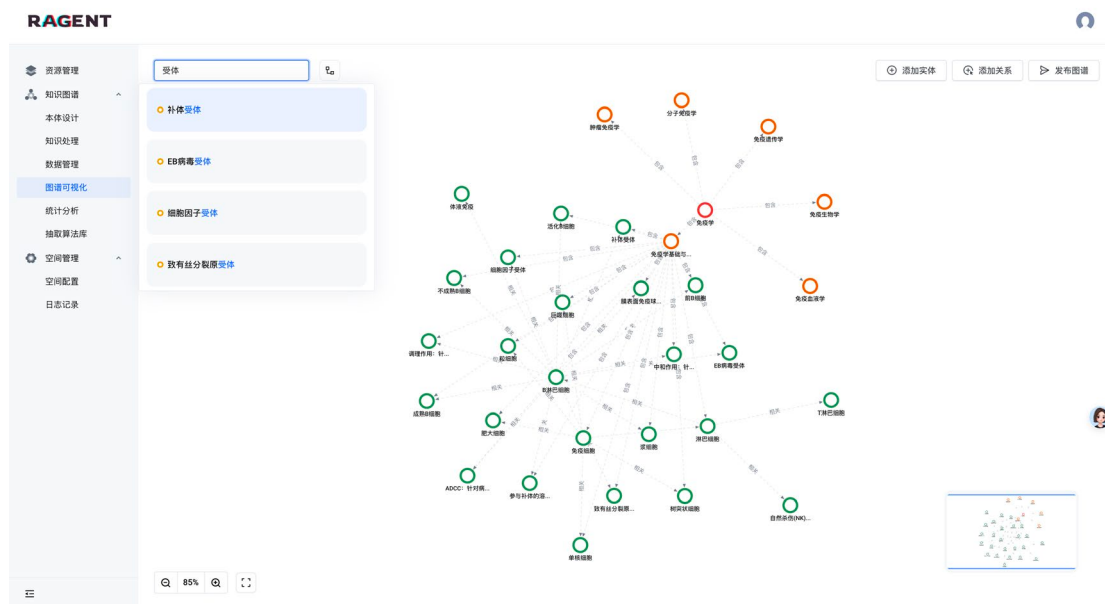
特别注意：

- 英文单引号' 是图数据库底层语法，所以如果数据中存在单引号将会报错，请替换成中文引号'。
- 网络超时是由于 HTTP 接口存在 30s 通信上限，而数据量大的情况下批量写入数据需要较长时间，但大部分情况下服务端已经收到了完整数据，可以等待一会去数据管理直接查看是否导入成功；也可以将一个数量大的 excel 拆分成多次映射。



1.6 查看图谱

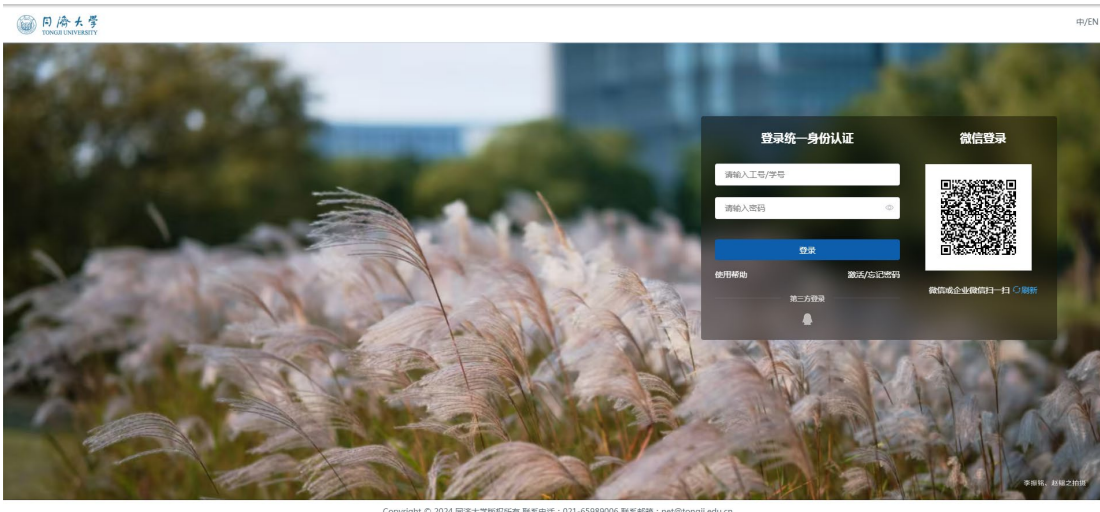
点击知识图谱-图谱可视化功能，进行图谱查看，支持搜索实体、查看实体详情等功能。



2. 产品功能

2.1 统一身份认证登录

浏览器地址栏输入网址：<https://kgc.tongji.edu.cn/>，自动跳转至统一认证登录页面进行登录。



2.2 知识空间

2.2.1 知识空间列表

显示当前登录用户有权限的空间列表，按照空间创建时间倒序排列。如果搜索、查看不到某一个知识空间，可能是由于不在空间成员中，可以联系知识空间所有者、平台负责人或管理员。



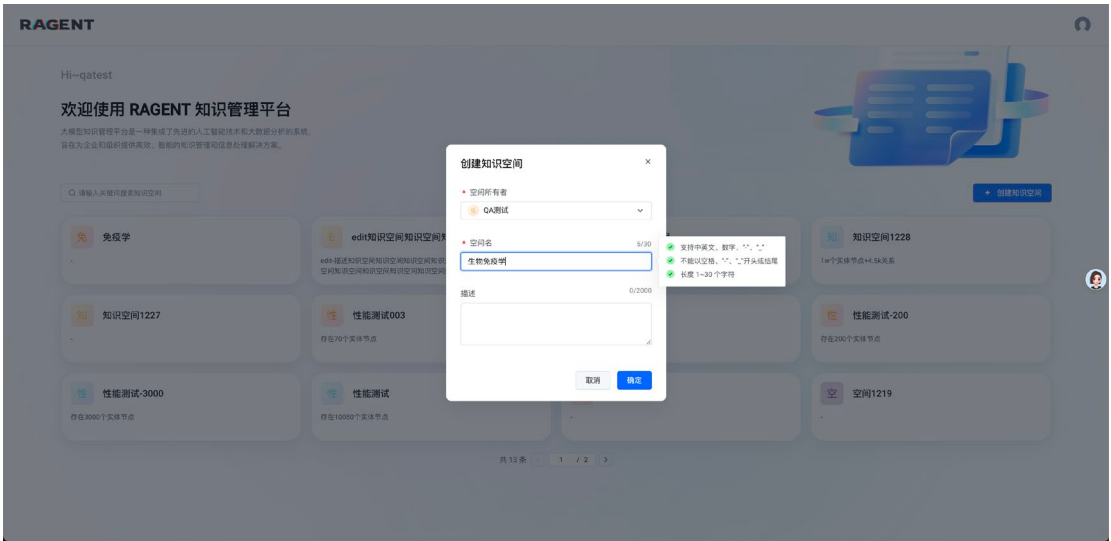
2.2.2 创建知识空间

设置空间所有者、空间名、描述等信息进行知识空间创建，知识空间之间数据与权限相互隔离。

- 空间所有者：拥有空间完整权限，可以进一步构建图谱或添加空间协调成员；
- 空间名：一般为知识图谱所在的业务、领域、学科、专业相关内容，如财务、金

融、土木工程等；

- 描述：描述当前知识空间包含的内容或者介绍图谱。

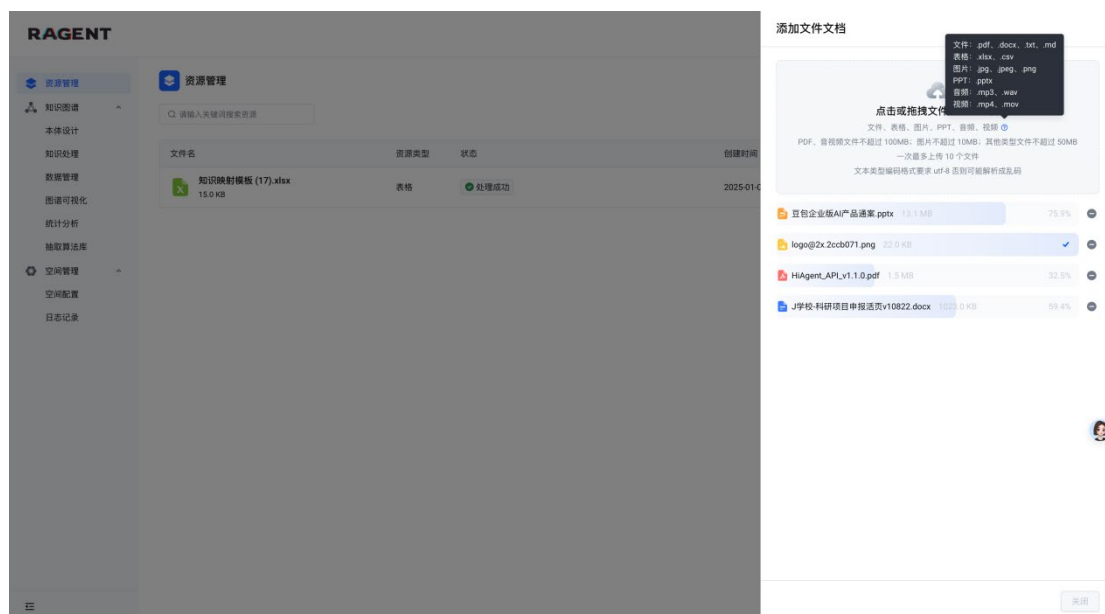


2.3 资源管理

资源类型	范围	文件格式
文本	word 文档、PDF、txt	docx、pdf、txt
表格	excel 文档	xlsx、csv
多媒体	图片、ppt、视频、音频	jpg/jpeg/png、pptx、mp4/mov、mp3/WAV
数据源	数据库	MySQL
网页资源	网页地址	-

2.3.1 添加资源-文件文档

支持上传多种类型的文件文档



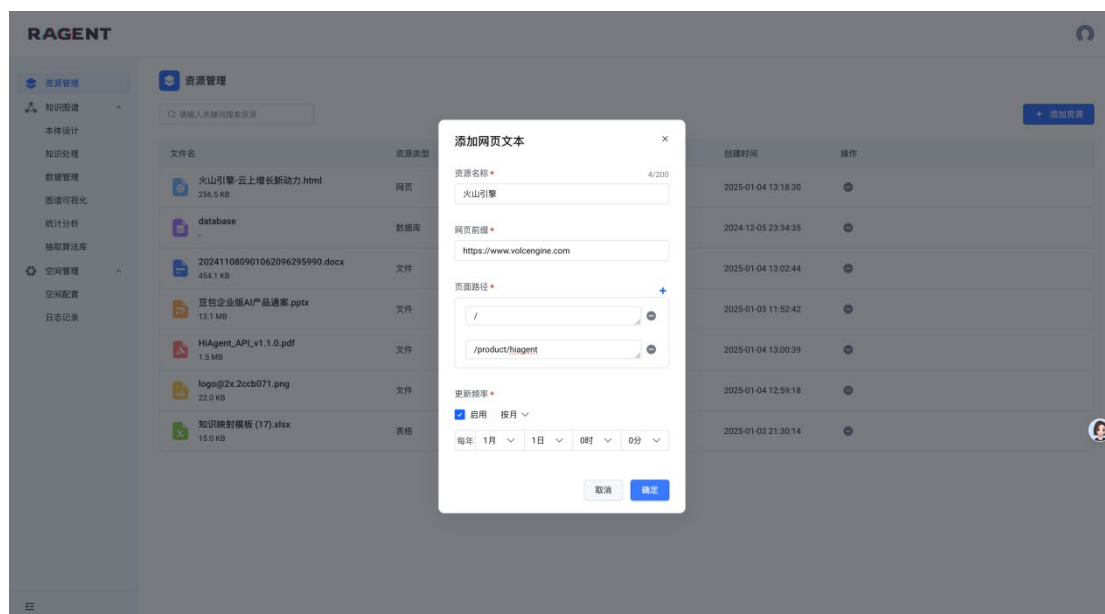
2.3.2 添加资源-数据库

支持添加 MySQL 数据库连接，需要配置 JDBC 连接所需要的 database、数据库地址、用户名、密码，验证连通性后方可点击确定完成数据库资源的添加。



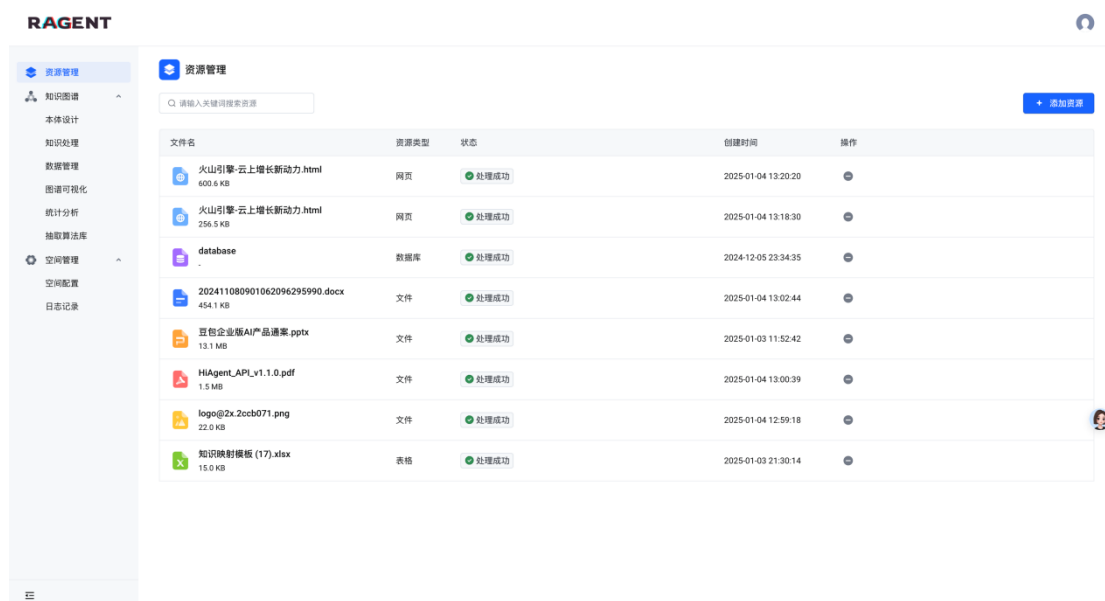
2.3.3 添加资源-网页文本

支持添加无需登录鉴权的在线网页资源，并支持自动定时更新。平台将获取网页中的文本内容并进行存储，以供后续资源关联或者知识抽取。



2.3.4 资源列表

资源列表将显示当前空间的资源信息，只有状态为处理成功的资源，可以被后续的知识抽取或知识映射使用；删除资源可能会影响已关联的知识抽取任务，请谨慎操作。



2.4 本体设计

2.4.1 添加实体类型

先添加实体类型，如：学科；设置该实体类型包含的属性，如：学科介绍。

添加实体类型

实体类型名 *

2/30

学科

颜色 *

属性

学科代码

文本

学科介绍

文本

取消

确定

2.4.2 添加关系

添加本体关系，可以通过底部的下拉框添加关系，也可以直接单击实体类型的圆点进行拖拉关系线。

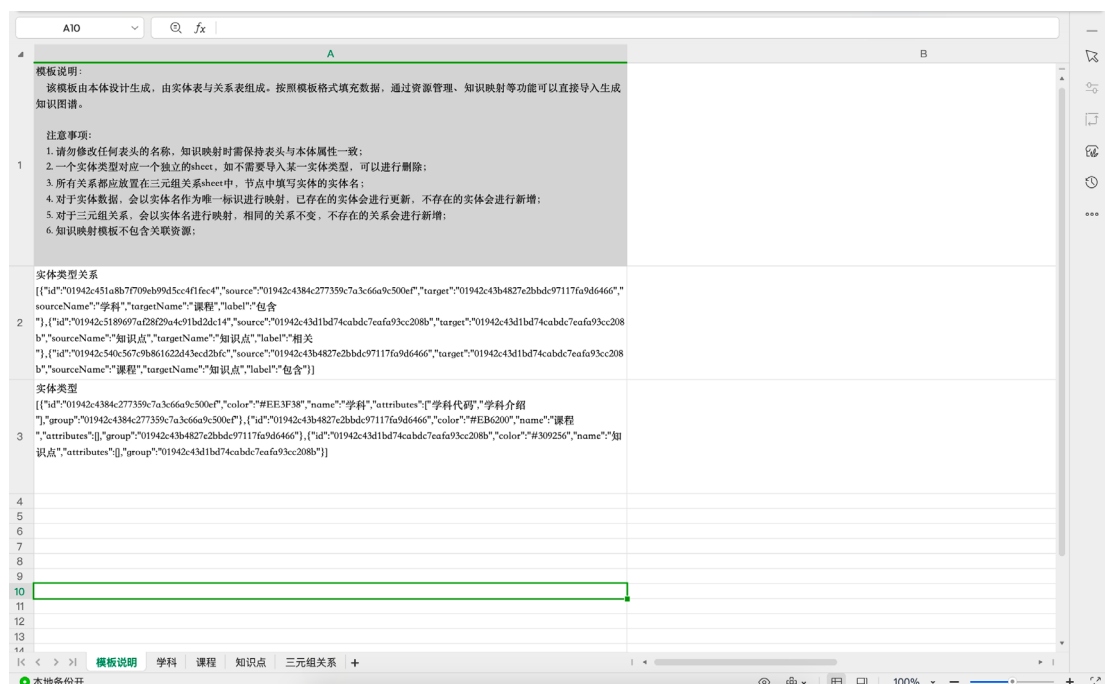




2.4.3 下载本体模板

完成本体设计后，支持下载知识模板。





模板说明：

该模板由本体设计生成，由实体表与关系表组成。按照模板格式填充数据，通过资源管理、知识映射等功能可以直接导入生成知识图谱。

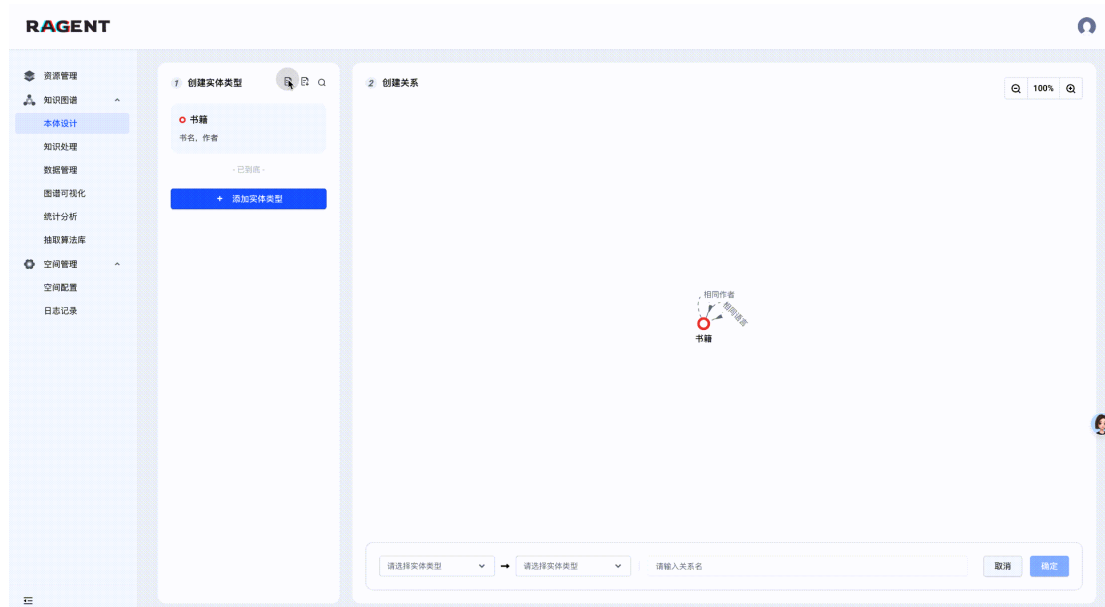
注意事项：

1. 请勿修改任何表头的名称，知识映射时需保持表头与本体属性一致；
2. 一个实体类型对应一个独立的 **sheet**，如不需要导入某一实体类型，可以进行删除；
3. 所有关系都应放置在三元组关系 **sheet** 中，节点中填写实体的实体名；
4. 对于实体数据，会以实体名作为唯一标识进行映射，已存在的实体会进行更新，不存在的实体会进行新增；
5. 对于三元组关系，会以实体名进行映射，相同的不变，不存在的进行新增；
6. 知识映射模板不包含关联资源；

2.4.4 复用本体

通过上传本体模板的功能能够实现复用本体设计。

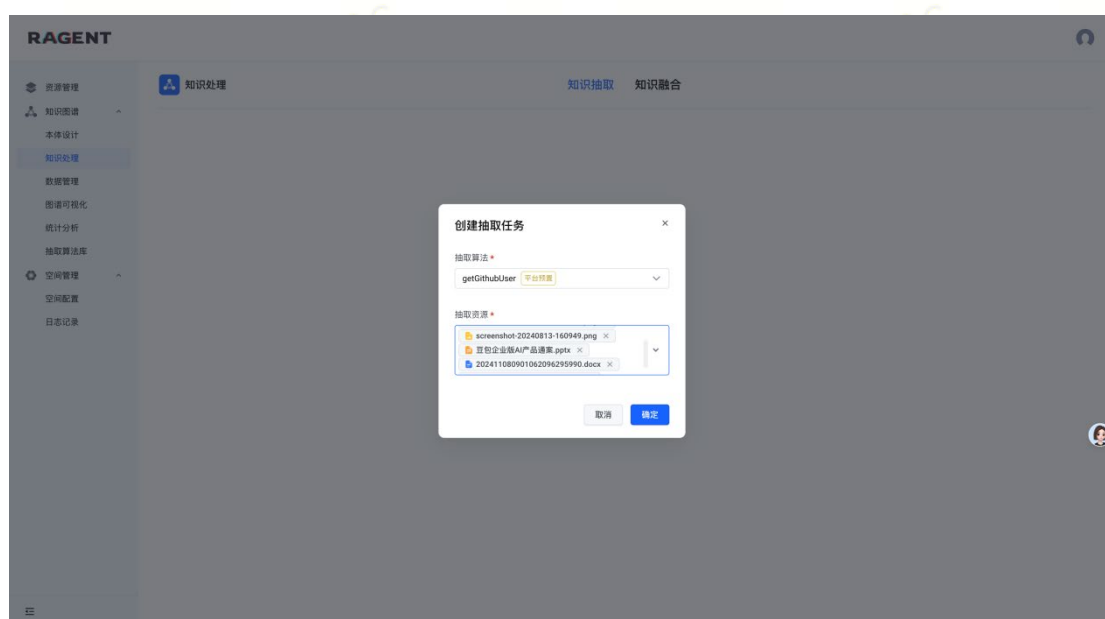
请注意不要擅自修改模板中的原始信息，以免影响模板复用与知识映射。



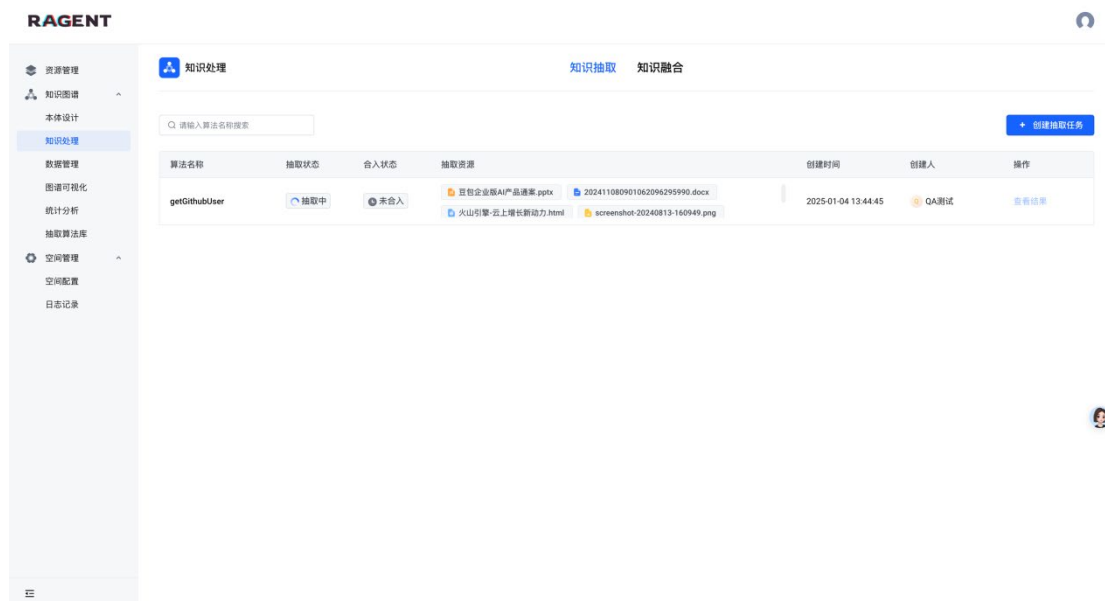
2.5 知识处理

2.5.1 知识抽取

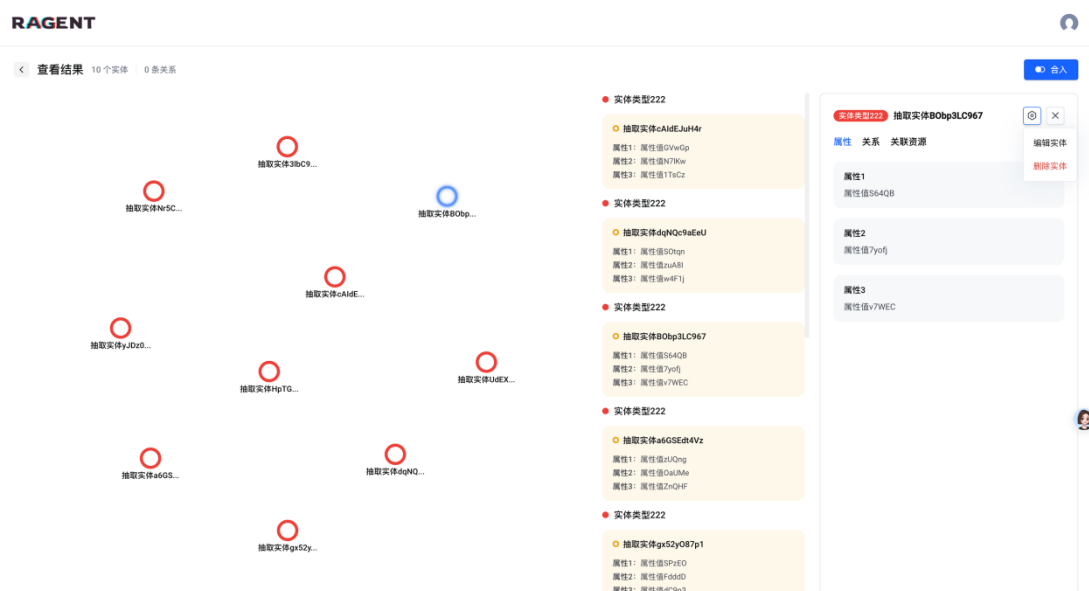
支持创建抽取任务，选择所需的抽取算法与需要抽取的资源，即可发起抽取任务。



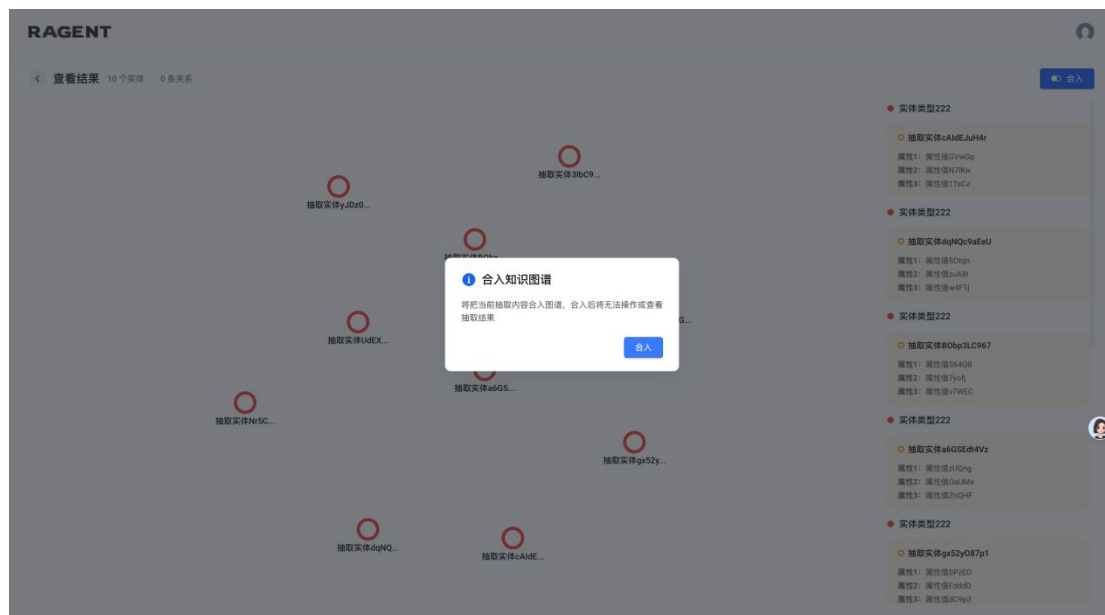
抽取状态将由外部算法服务实际的抽取情况决定，外部算法写回抽取结果后，抽取状态将变更为已抽取。



抽取完成后可查看结果，对于抽取结果不满意也可以直接编辑或删除个别实体。

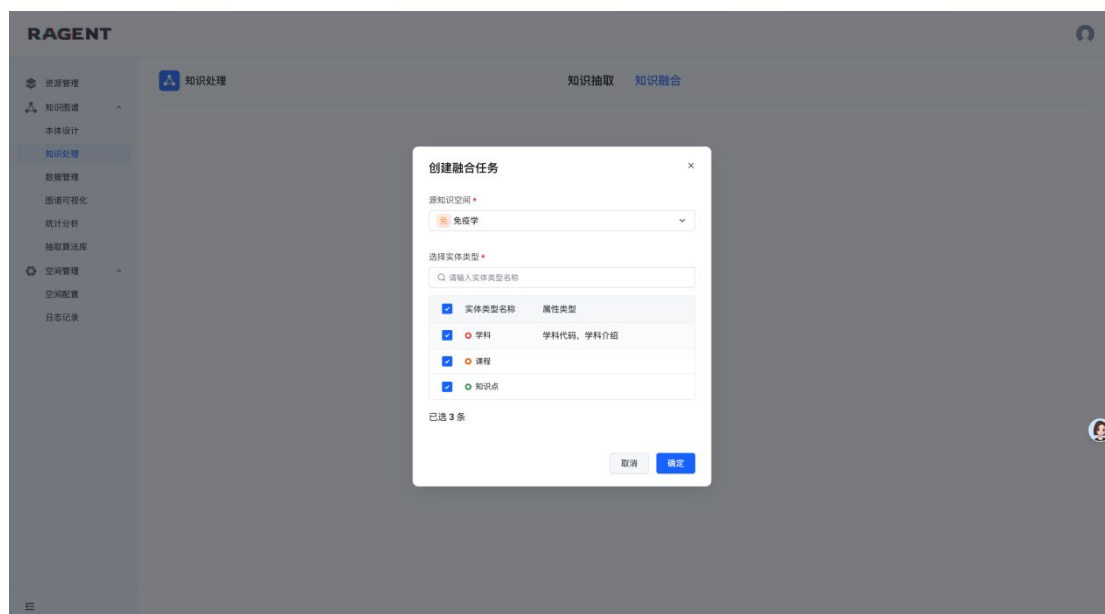


确定抽取结果无误后可以，点击合入，将当前抽取内容合入图谱，合入后将无法操作或查看抽取结果。

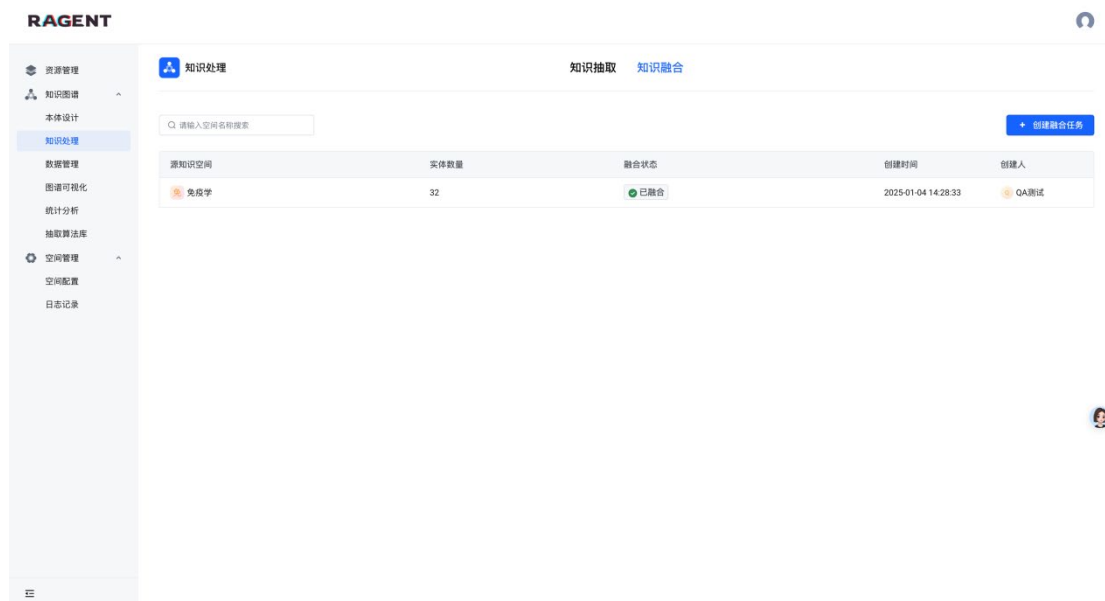


2.5.2 知识融合

支持将有权限的知识空间图谱数据融合进当前空间，仅支持融合实体类型名称、关系名称与当前空间本体设计一致的数据。



融合时将进行简单的消歧与合并，同名实体将融合合并。



2.5.3 知识映射

特别注意：

- 英文单引号' 是图数据库底层语法，所以如果数据中存在单引号将会报错，请替换成中文引号'。
- 网络超时是由于 HTTP 接口存在 30s 通信上限，而数据量大的情况下批量写入数据需要较长时间，但大部分情况下服务端已经收到了完整数据，可以等待一会去数据管理直接查看是否导入成功；也可以将一个数量大的 excel 拆分成多次映射。

(1) 通过模板映射

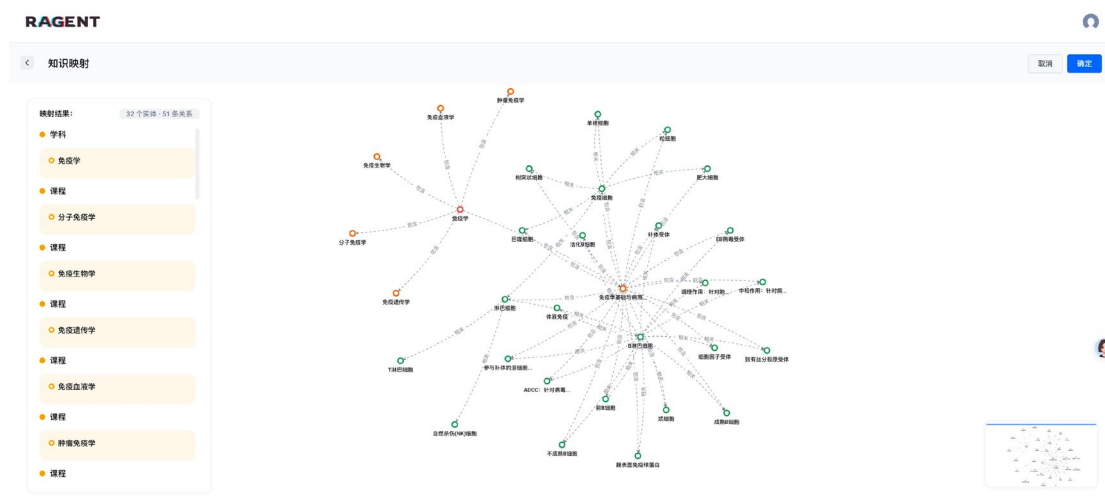
点击知识图谱-知识处理，选择通过模板映射；



选择资源管理中已经上传的模板文件；



进行自动化映射，查看映射的结果无误后，点击确定即可以正式将数据写入图谱。

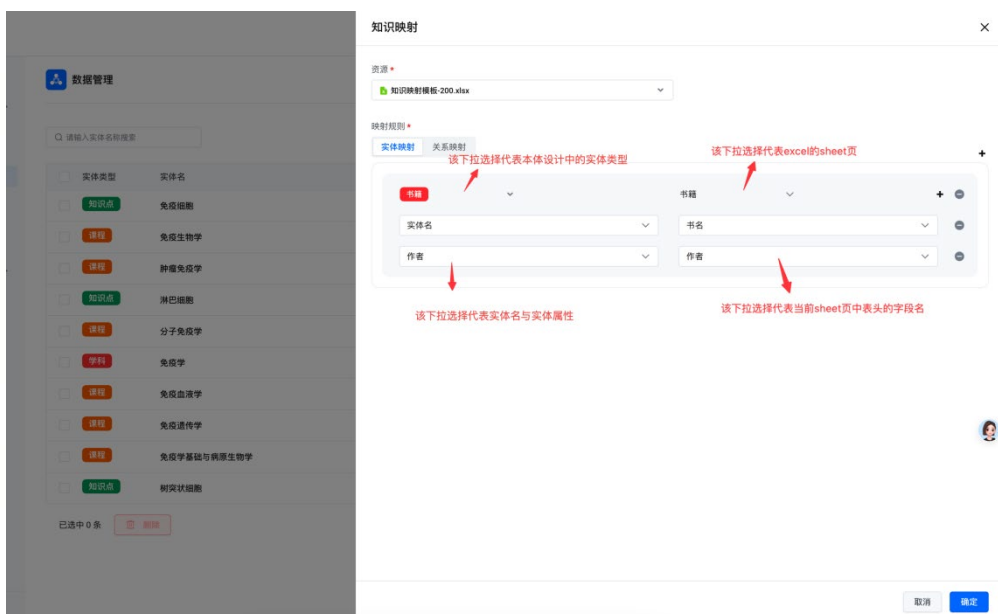


(2) 通过手动配置映射

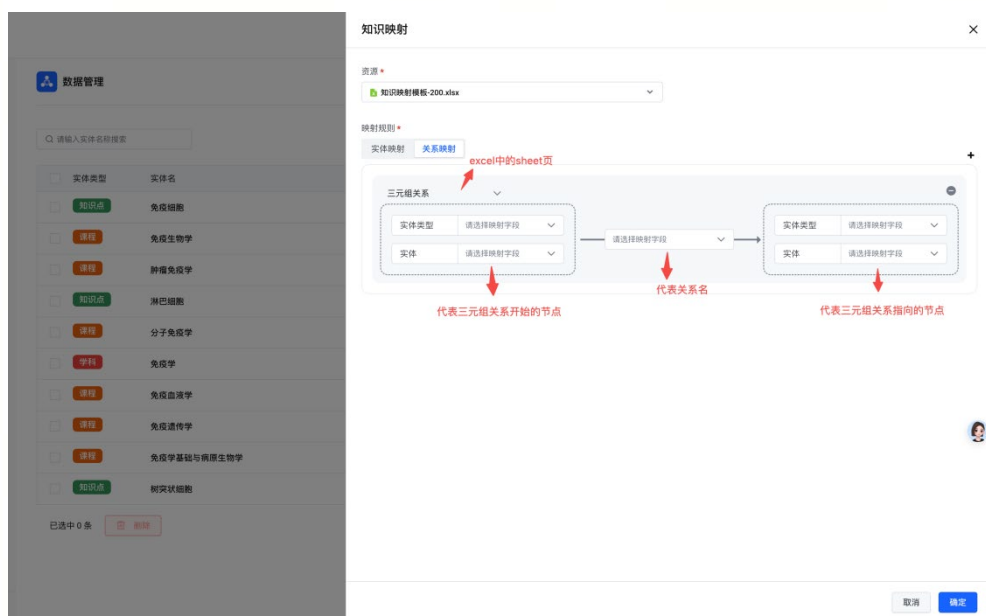
点击知识图谱-知识处理，选择通过手动配置映射；



手动配置 excel 中的表头字段与实体属性的映射规则；



手动配置 excel 中的表头字段与关系的映射规则；



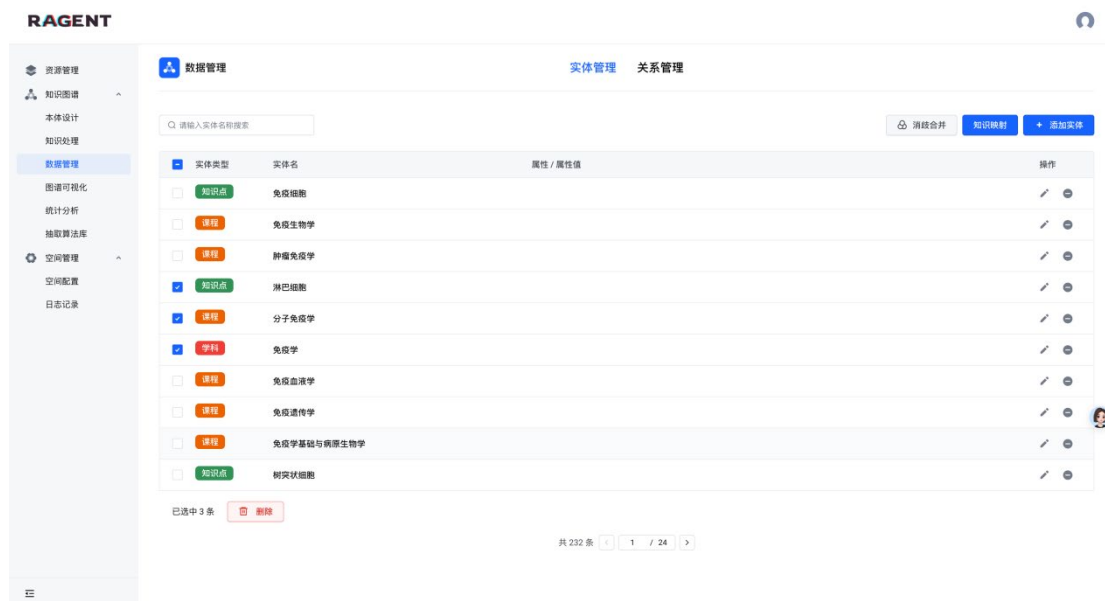
点击确定后展示映射结果，同“通过模板映射”。

2.6 数据管理

2.6.1 实体管理

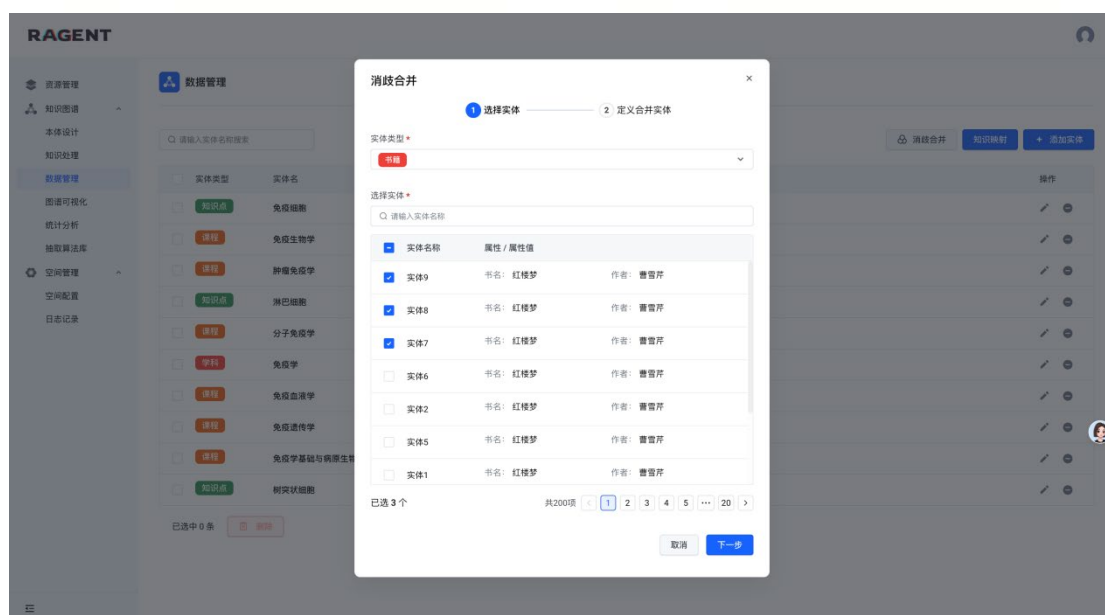
2.6.1.1 实体列表

支持在数据管理中以列表的形式查看、搜索实体数据，并对数据进行编辑、删除、批量删除的操作。



2.6.1.2 消歧合并

支持将同一个实体类型的多个实体进行手动消歧合并，可定义合并后实体的信息。





2.6.1.3 添加实体

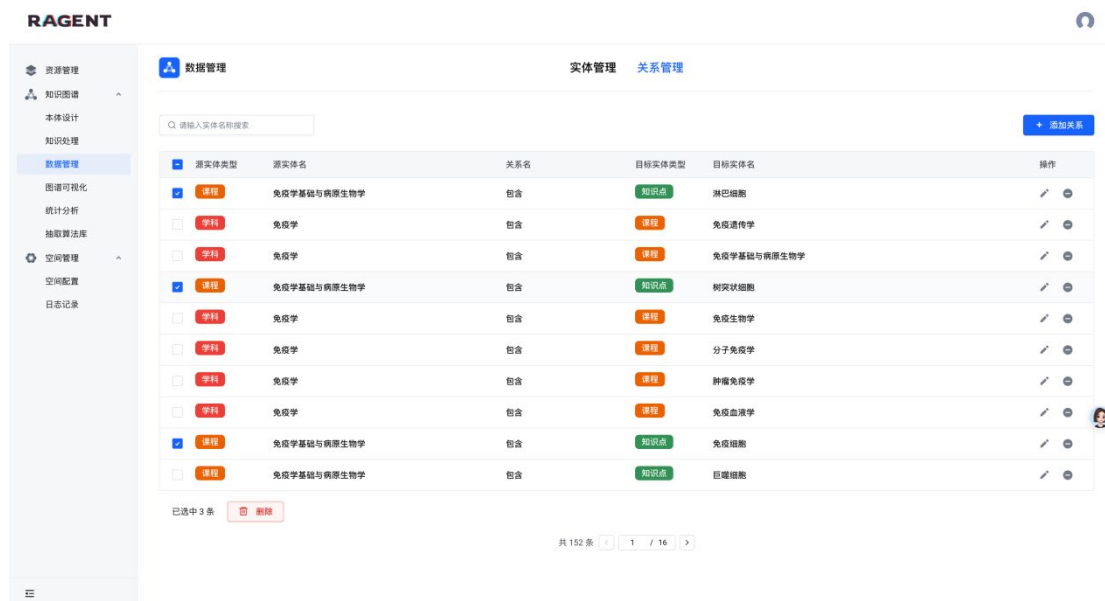
支持手动添加实体，填写实体名与实体属性即可创建。



2.6.2 关系管理

2.6.2.1 关系列表

支持在数据管理中以列表的形式查看、搜索关系数据，并对数据进行编辑、删除、批量删除的操作。



2.6.2.2 添加关系

支持手动添加关系，设置源实体、目标实体、关系即可创建。

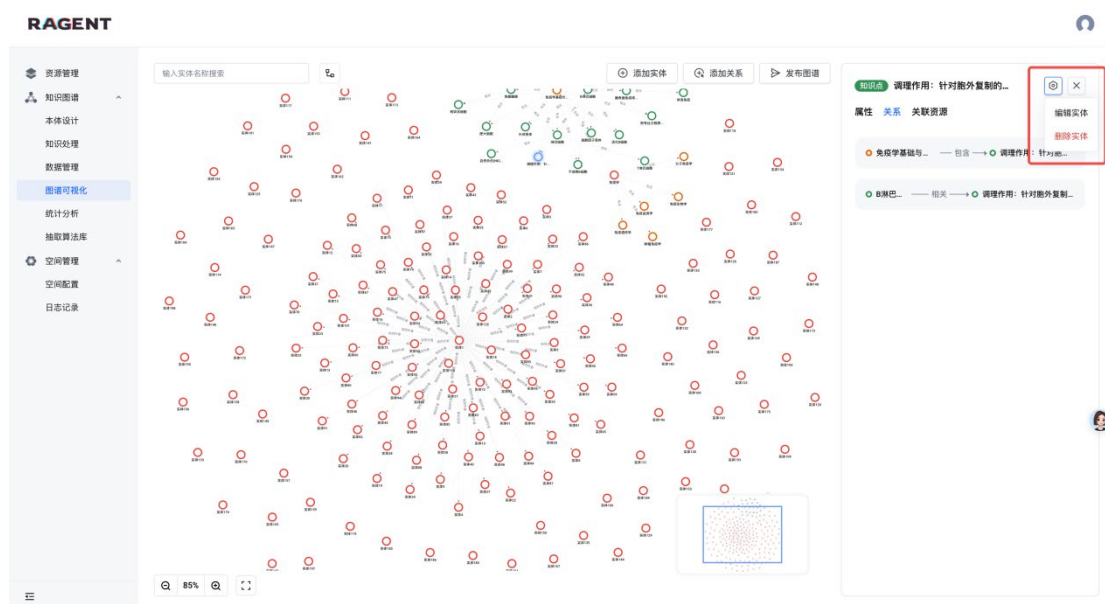
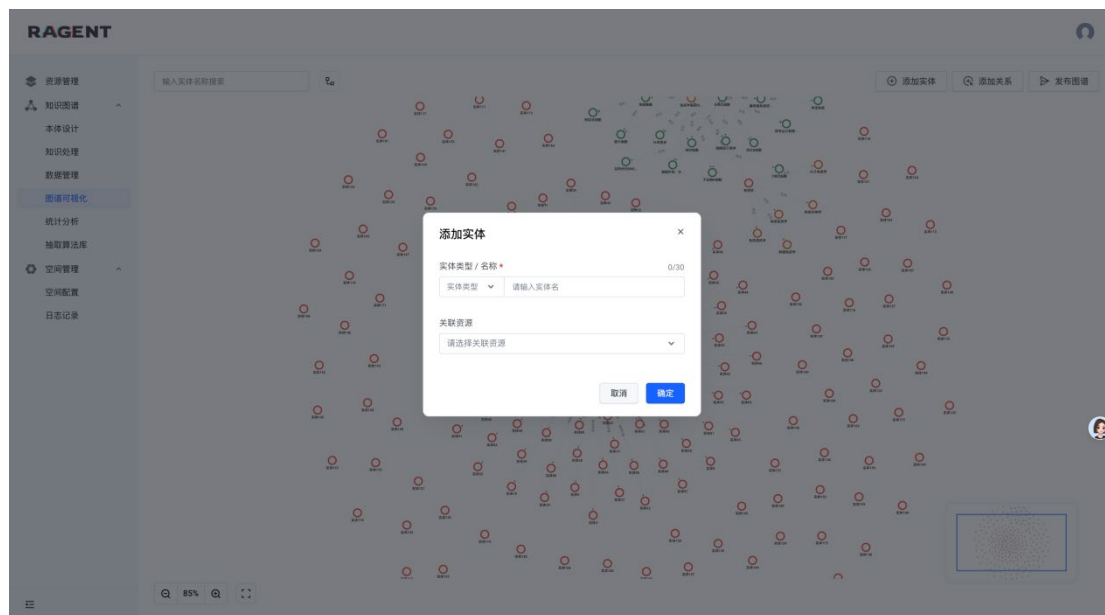
本体中不包含的关系无法创建，需要先进行本体设计。



2.7 图谱可视化

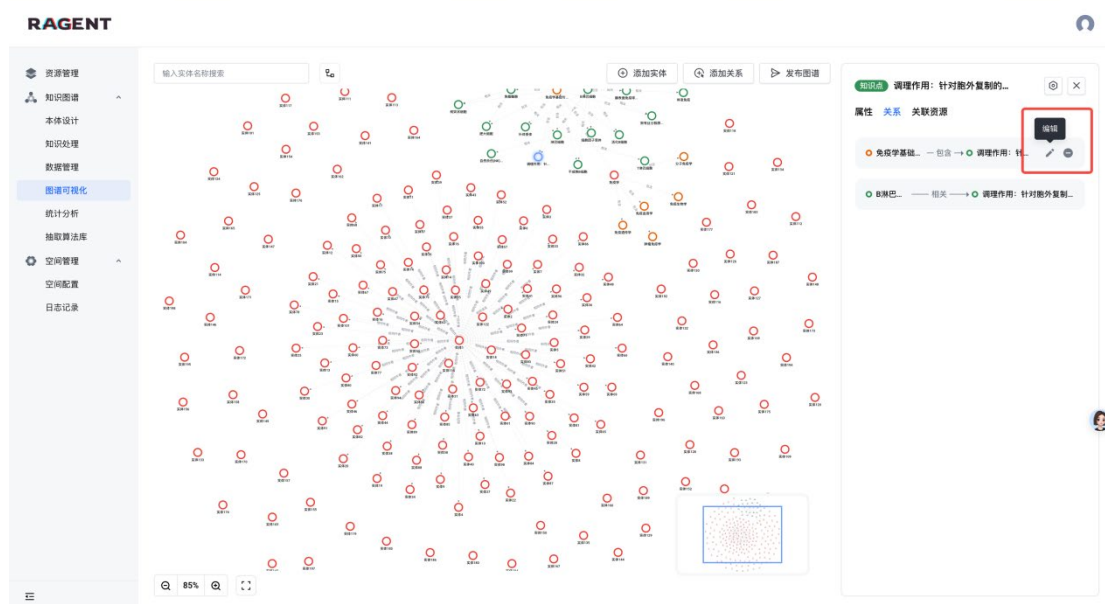
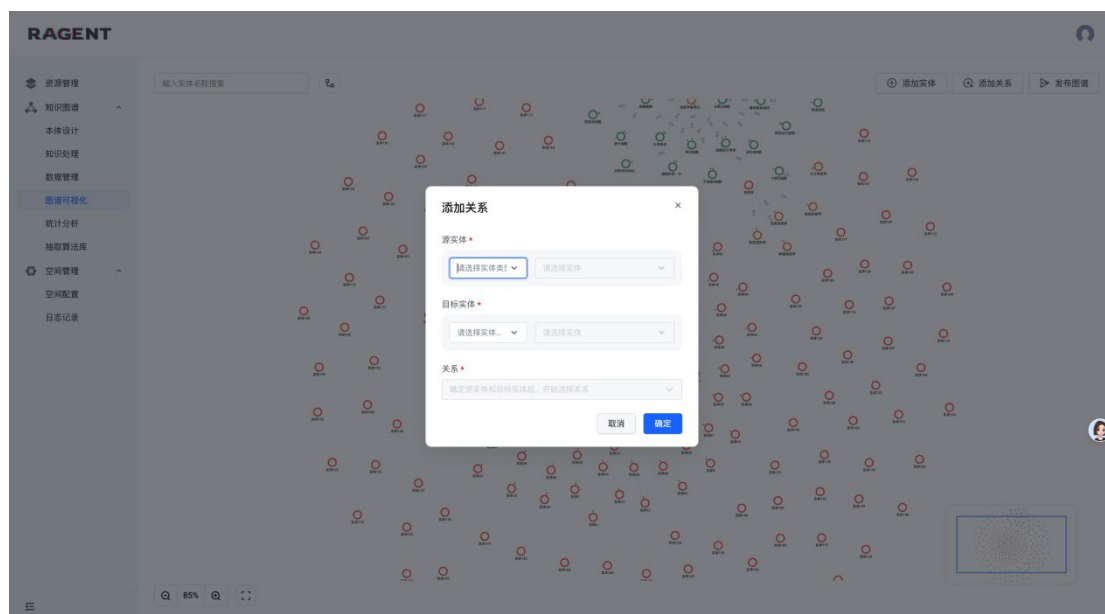
2.7.1 添加/编辑/删除实体

在图谱可视化页面中也支持手动添加实体，便于在浏览图谱过程中随时补充数据。



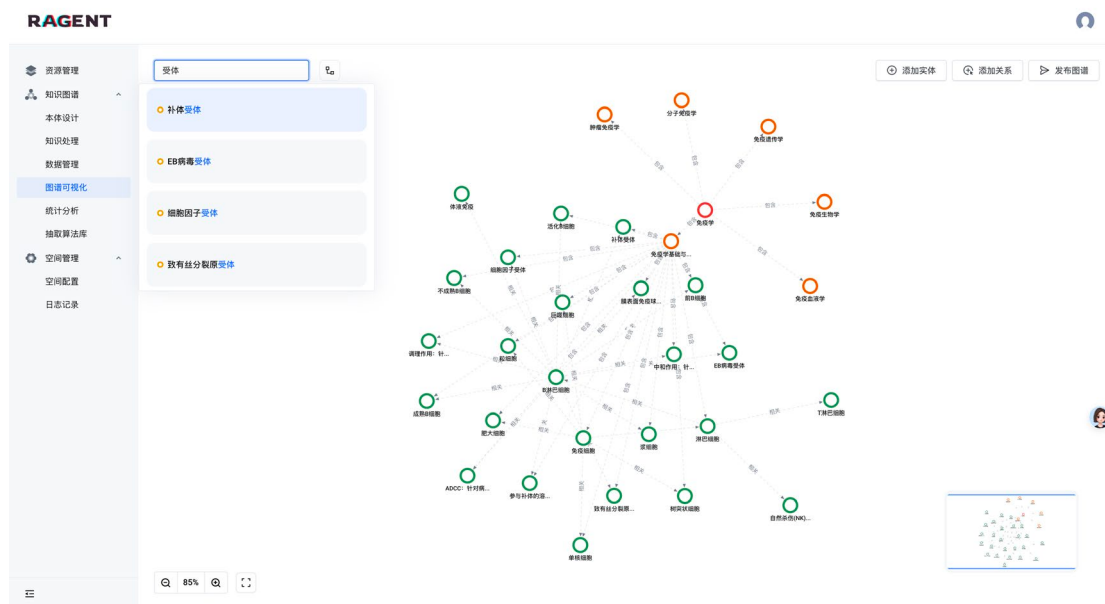
2.7.2 添加/编辑/删除关系

在图谱可视化页面中也支持手动添加关系，便于在浏览图谱过程中随时补充数据。



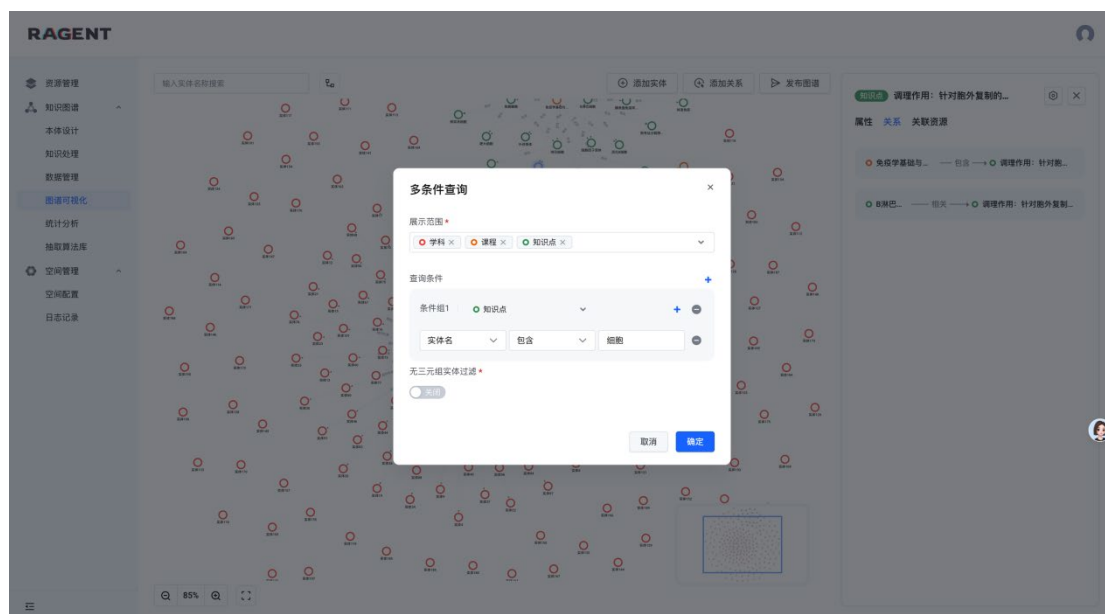
2.7.3 图谱查看

点击知识图谱-图谱可视化功能，进行图谱查看，支持搜索实体、查看实体详情等功能。



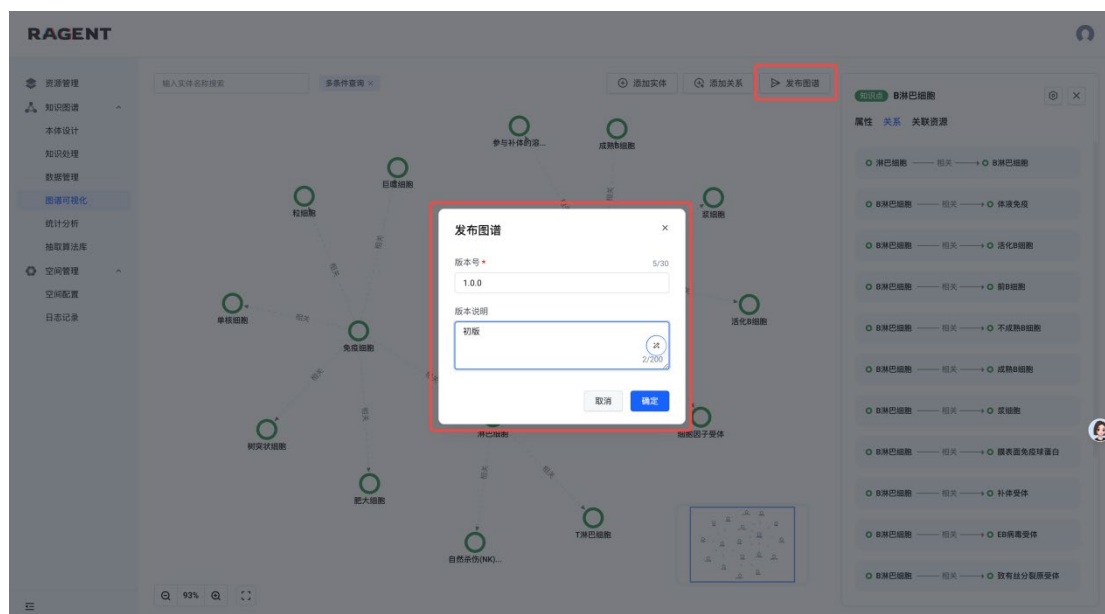
2.7.4 多条件查询

支持设置多条件查询规则，将大图谱过滤筛选成子图，方便用户能专注查看某个场景下的图谱数据。

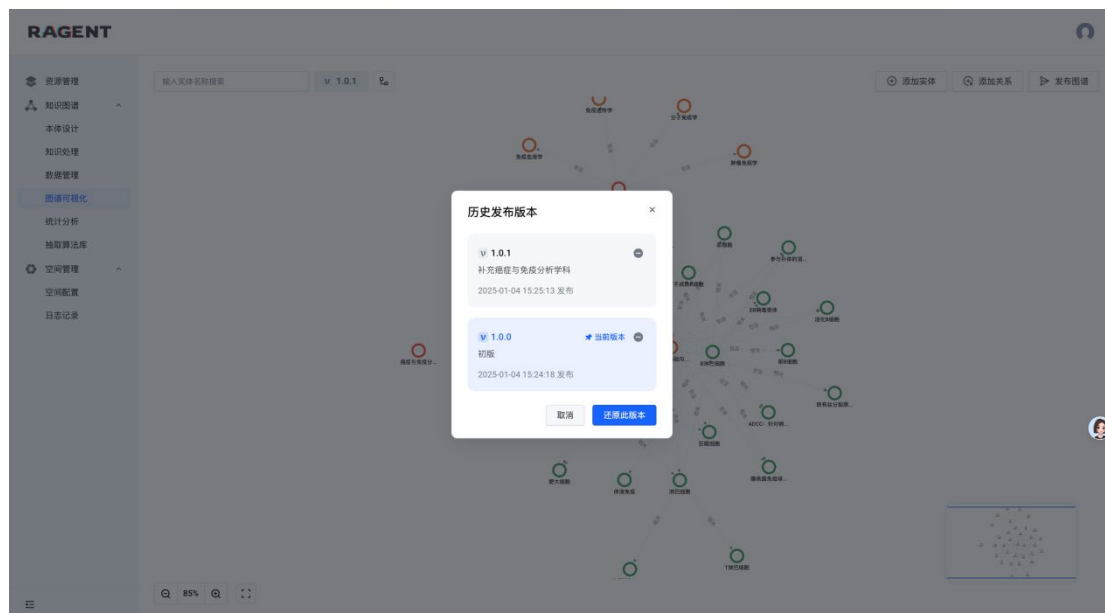


2.7.5 发布图谱

支持将已有图谱数据进行发布，发布后将会留存一个独立的图谱版本，在外部系统获取图谱数据时，将按照当前版本提供数据。如果有数据变化建议都发布版本来留存，方便后面出现异常时进行图谱版本回滚。

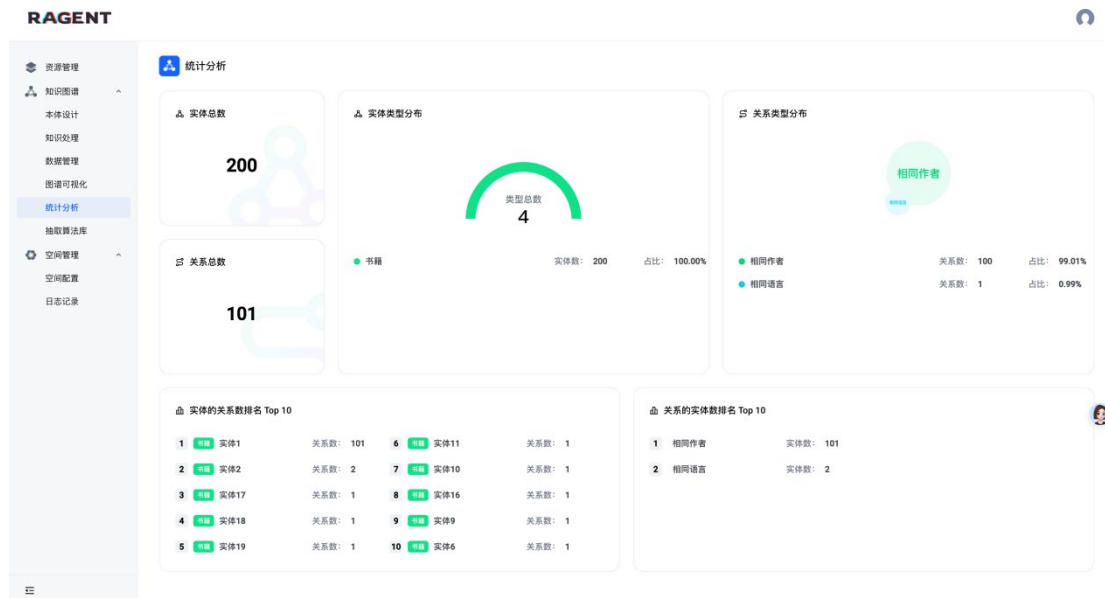


支持将图谱还原成历史版本，方便图谱构建出现异常时进行图谱版本回滚。



2.8 统计分析

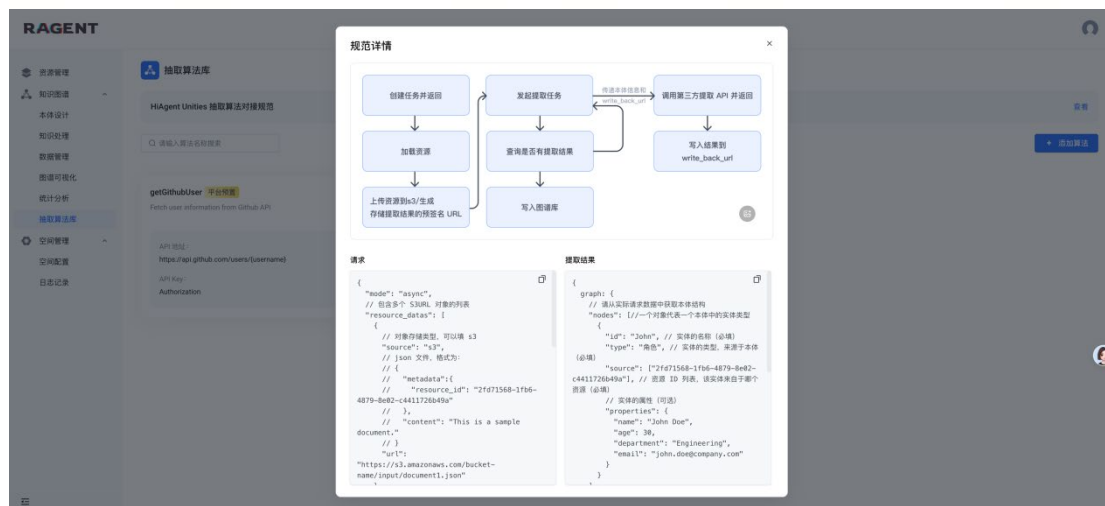
每一个知识空间均可以构建一个知识图谱，每个知识图谱都能通过统计分析功能，查看实体数量、关系数量、实体类型分布、关系类型分布、关系数排名等信息统计，方便用户评估图谱的数据规模。



2.9 抽取算法库

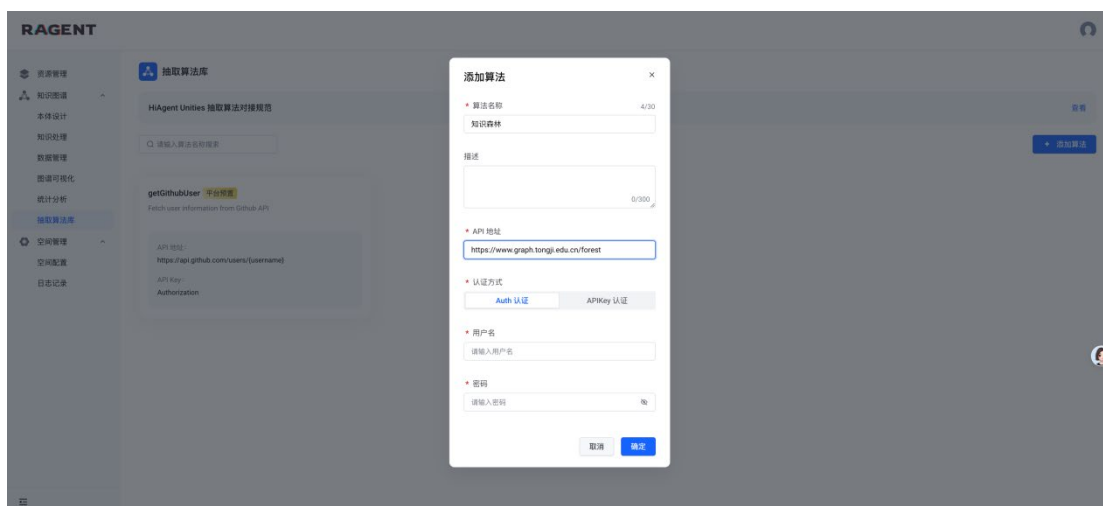
2.9.1 算法规范

支持查看抽取算法接入规范，平台调用外部抽取算法时，将上传资源至指定 S3 并在请求中提供预签名的 URL、本体信息；外部抽取算法服务需要去 S3 获取具体资源，完成抽取后按照提取结果规范写回 url；平台将定时查询是否写回提取结果，获取到提取结果后抽取任务状态将变为已完成。



2.9.2 添加算法

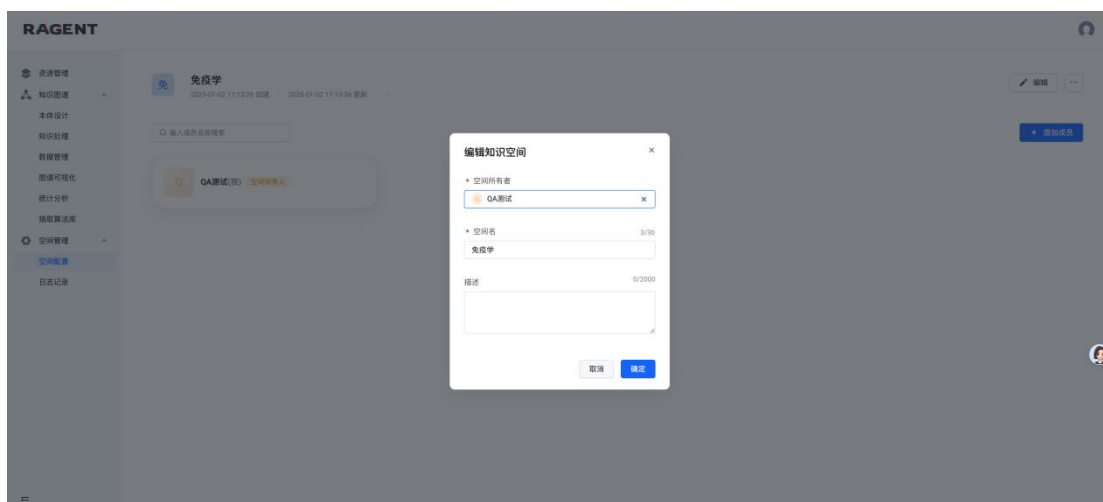
按照算法接入规范提供算法服务 API 地址，即可将外部抽取算法接入平台。为了保障抽取算法安全，必须要添加 API 认证。



2.10 空间配置

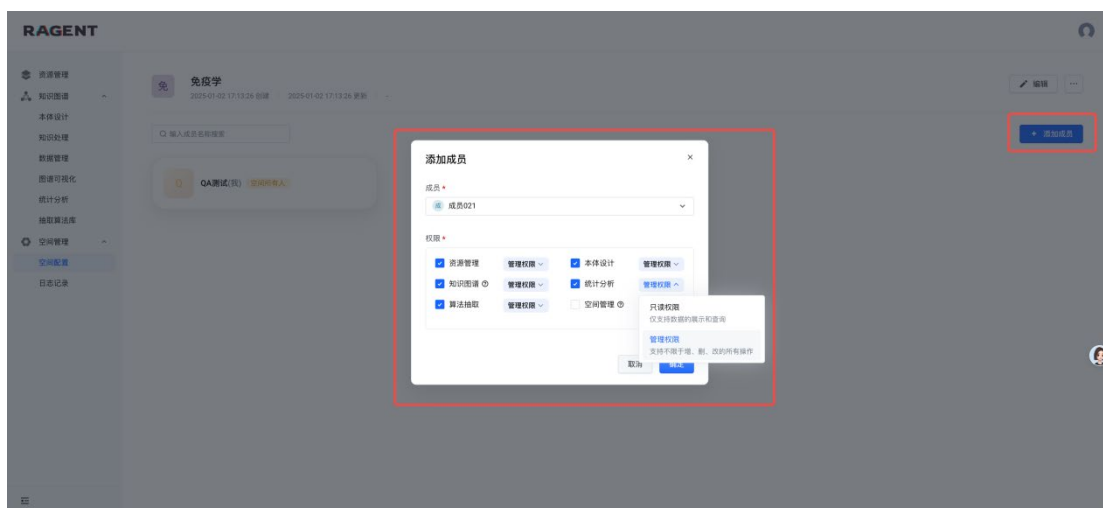
2.10.1 编辑知识空间

支持编辑知识空间的基本信息，如更换空间管理员、修改空间名等。

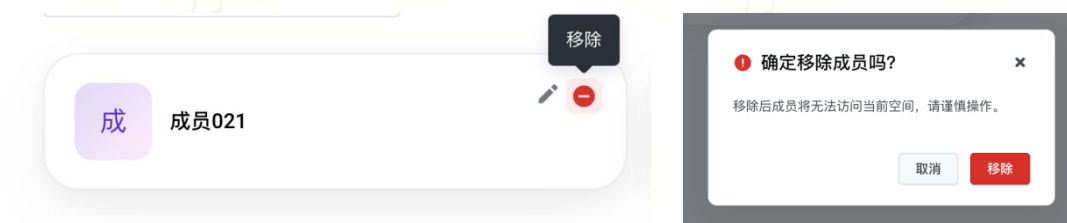


2.10.2 管理空间成员

(1) 添加/编辑空间成员：支持添加/编辑成员并赋予合适的空间功能权限。



(2) 移除成员：支持将某位成员移出知识空间，移除后该用户将无法访问当前空间。



2.11 日志记录

支持查看空间内所有操作记录。

