



知识图谱

用户指南

文档版本 01

发布日期 2019-09-12

华为技术有限公司



版权所有 © 华为技术有限公司 2020。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编： 518129

网址： <https://www.huawei.com>

客户服务邮箱： support@huawei.com

客户服务电话： 4008302118

目录

1 创建知识图谱简介	1
2 开通 KG 服务	3
3 管理本体	5
3.1 本体简介	5
3.2 创建本体	6
3.2.1 控制台创建本体	6
3.2.2 OBS 导入本体	8
3.3 预览本体	8
3.4 编辑本体	9
3.5 导出本体	9
3.6 删除本体	10
4 自定义信息抽取模型	11
4.1 信息抽取模型简介	11
4.2 准备训练数据	12
4.3 创建模型	13
4.4 管理版本	18
4.4.1 创建新版本	18
4.4.2 发布版本	19
4.4.3 修改版本	19
4.4.4 删除版本	20
5 步骤一：选择本体	21
6 步骤二：配置数据源	23
7 步骤三：配置信息抽取	26
7.1 配置信息抽取简介	26
7.2 普通抽取	27
7.2.1 信息抽取函数	27
7.2.2 交互界面配置	28
7.2.3 格式化配置	31
7.3 算法抽取	32
8 步骤四：配置知识映射	34

9 步骤五：配置知识融合..... 38

10 融合验证..... 42

11 管理图谱..... 44

11.1 图谱管理简介..... 44

11.2 查看图谱详情..... 44

11.3 发布图谱版本..... 46

11.4 预览图谱..... 46

11.5 全量更新图谱..... 47

11.6 增量更新图谱..... 47

11.7 复制配置快速创建图谱..... 48

1 创建知识图谱简介

知识图谱服务（Knowledge Graph，简称KG）为自然语言处理服务（Natural Language Processing Service，简称NLP）的子服务，是面向各类企业的一站式知识计算平台，提供图谱创建、本体管理、图谱管理等功能，帮助用户从0到1快速构建知识图谱，管理知识图谱流水线任务。

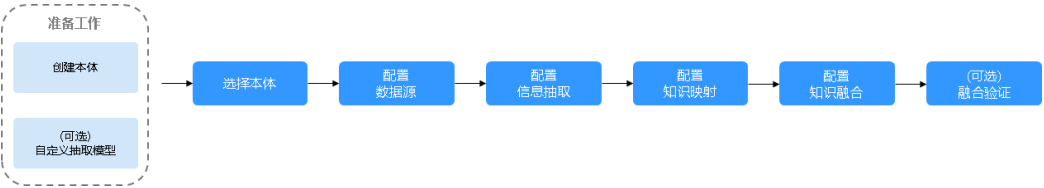
知识图谱在语义搜索与推荐、智能问答系统、大数据分析与决策等场景下均有广泛应用，具体请参见[产品介绍>适用场景](#)。

本文档介绍如何在知识图谱管理控制台快速构建知识图谱，并管理本体和知识图谱任务。

创建图谱流程

在KG服务上构建知识图谱的基本流程可归纳为5个步骤：创建本体、配置基础数据、选择本体、配置信息抽取、配置知识映射、配置知识融合。具体的创建流程如[图1-1](#)和[表1-1](#)所示。

图 1-1 构建知识图谱



知识图谱的基本概念和创建图谱的流程概念请参见[产品介绍>基本知识](#)。

表 1-1 创建流程表

流程	说明	详细指导
创建本体	本体是某个领域中抽象概念的集合，能够描述某个范围内一切事物的共有特征以及事物间的关系。 在创建知识图谱之前，您需要创建图谱的本体，用于组织图谱信息结构。	创建本体

流程	说明	详细指导
(可选) 自定义抽取模型	如果您在创建知识图谱时，选择使用自定义模型进行信息抽取，您需要在创建知识图谱之前，自定义抽取模型。	创建模型 创建新版本 发布版本
选择本体	您可以选择已创建的本体，或者从OBS中选择本体，用于组织待创建图谱的信息结构。	步骤一：选择本体
配置数据源	数据是知识图谱的基础。 基于已存储在OBS中的基础数据，您可以从OBS导入用于构建知识图谱的基础数据。	步骤二：配置数据源
配置信息抽取	信息抽取是从基础数据中抽取待创建图谱的实体、属性信息以及实体间的相互关系。 通过输入实体类型、抽取函数及抽取前后的数据字段，完成信息抽取的配置。	步骤三：配置信息抽取
配置知识映射	知识映射是建立从基础数据抽取出的结构化信息与知识图谱本体的映射关系。 通过设置映射前后的相关字段，完成知识映射的信息配置。	步骤四：配置知识映射
配置知识融合	知识融合融合来自多个数据源的关于同一个实体或概念的描述信息。 通过设置知识融合相关属性及相似函数参数，完成知识融合的信息配置。	步骤五：配置知识融合
融合验证	若在创建图谱时，配置了知识融合，则需要进行融合验证，即验证当前知识融合的配置产生的结果是否符合预期。	融合验证

2 开通 KG 服务

目前KG服务处于公测阶段，首次使用KG服务需要开通公测权限。由于KG服务依赖其他服务，所以需要在开始使用KG服务前获得OBS服务，KG与相关服务的关系请参见[与其它服务的关系](#)。

开通KG服务，必须先申请公测权限，然后再申请获取相关服务的授权。

前提条件

- 1. 已[注册](#)华为云账号，并完成实名认证。具体申请信息请见[账号注册](#)和[实名认证](#)。
- 2. KG服务当前为公测阶段，是免费的。但是其相关服务OBS为商用服务，根据您使用的功能或资源按需收费，因此在使用KG服务时，为保证KG服务所有功能正常工作，建议您确保当前账号未欠费。

申请公测权限

根据国家法律规定，所有用户必须完成实名认证后才能使用云服务。

- 1. 登录[KG服务管理控制台](#)。
 - 如果您的账号已获得公测权限和相关服务授权，系统将自动跳转至KG服务管理控制台。
 - 如果您的账号未开通服务公测，系统将自动跳转到“申请KG服务权限”页面。根据界面提示，表示当前账号无公测权限，单击“申请公测”跳转到“申请公测”页面。
- 2. 在“申请公测”页面，根据界面提示填写所需信息，申请信息说明请参见[表2-1](#)，填写完成后，勾选“同意公测试用服务协议”，然后单击“申请公测”完成申请操作。

表 2-1 申请公测参数

参数名	是否必填	参数说明
应用场景	是	当前项目/业务所在行业，例如web、移动与社交应用程序、数字媒体与市场营销等。
业务描述	是	当前项目/业务的实际情况，便于平台为您提供精准的服务。总字符长度不小于15。

参数名	是否必填	参数说明
联系人	是	申请人的称呼。
联系电话	是	申请人的手机号。便于平台为您提供技术支持。
电子邮箱	否	申请人的邮箱地址。

图 2-1 申请公测权限

申请的公测产品/服务

自然语言处理服务-知识图谱

申请信息

* 应用场景:

其他

* 业务描述:

您可以简单描述下您目前项目及业务的实际情况，请输入至少15个字。

0/80

* 联系人:

郭**

* 联系电话:

+86(中国大陆)

158****8420

电子邮箱:

☒ 同意《公测试用服务协议》

申请公测

3. 申请公测操作完成后，系统将自动跳转至“我的公测”页面。当“审批状态”显示为“审批通过”时，表示您已经获得了KG服务的公测权限。

3 管理本体

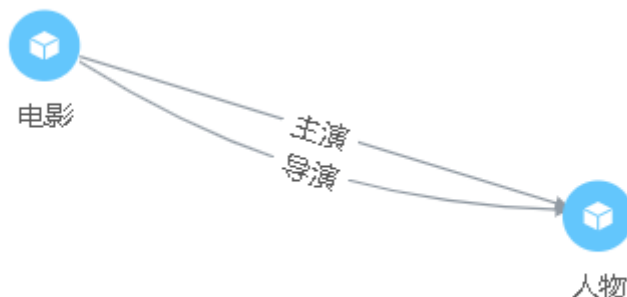
3.1 本体简介

在创建知识图谱之前，您需要创建图谱的本体，用于组织知识图谱信息结构。

基本概念

- 实体
实体是有可区别性且内于其自身而独立存在的某种事物。例如本体如图3-1所示，某一具体的电影（或人物或关系），如电影“霸王别姬”，可称作实体。
- 概念
概念指是抽象的、普遍的想法，是充当指明实体、事件或关系的范畴或类的实体。例如本体如图3-1所示，“电影”、“人物”可称作概念。
- 概念间关系
概念间关系是概念之间的相关联系。例如本体如图3-1所示，“电影”、“人物”两者之间的关系“主演”、“导演”均可称作概念间关系。关系名称不能重复。
- 本体
本体是某个领域中抽象概念的集合，能够描述某个范围内一切事物的共有特征以及事物间的关系。例如图3-1可称作一个本体。

图 3-1 本体



操作指引

- **创建本体**：在创建知识图谱之前，创建用于组织图谱信息结构的本体。
- **预览本体**：创建本体后，预览描述本体信息的结构图。
- **编辑本体**：创建本体后，编辑本体信息，更新本体。
- **导出本体**：把创建的本体导出至OBS文件夹中。
- **删除本体**：删除已创建的本体，以适应业务变化。

创建本体方式

- 方式一：**控制台创建本体**
- 方式二：**OBS导入本体**

3.2 创建本体

3.2.1 控制台创建本体

进入创建本体页面

1. 登录**KG服务管理控制台**，在左侧菜单栏中单击“本体管理”，进入“本体管理”页面。
2. 单击左上角“创建本体”，进入“创建本体”页面。
3. 在“创建本体”页面，您可以手动创建概念和概念间的关系，并编辑概念，如图3-2所示。

说明







在创建本体时，所有概念都默认有一个属性为“name”。

图 3-2 创建本体



创建本体操作

在“创建本体”页面手动创建概念和概念间的关系。

- 填写本体名称
在“本体名称”文本框中填写待创建的本体名称，在“本体描述”文本框中填写待创建的本体描述。
- 编辑概念
 - a. 单击“概念编辑模式”，进入概念编辑模式。首次默认进入“概念编辑”模式。
 - b. 拖拽编辑框左上角的概念节点悬放在编辑框的目标位置。
 - c. 单击编辑框内的概念节点。
右侧弹出概念编辑框。
 - d. 在概念编辑框中的“名称”文本框里填写概念的名称。
概念名称只能是字母、数字、下划线组成的合法字符串。
 - e. 在默认属性“name”下方，单击“新增属性”可添加属性文本框，在文本框中填写概念对应的属性。
属性只能是名称只能是字母、数字、下划线、中划线或者中文字符组成的合法字符串。
 - f. 单击属性文本框右侧, 可删减属性。
- 删除概念
右键单击概念节点, 单击“Delete”即可删除概念节点。
- 编辑概念间关系
 - a. 单击“关系编辑模式”，进入关系编辑模式。
 - b. 添加概念之间的关系。
 - 创建单概念的关系曲线
双击概念节点.
 - 创建概念间的关系
单击概念节点, 拖拽出关系虚线至目标概念节点，单击目标概念节点。
 - c. 编辑关系名称。
关系名称不能重复。
 - i. 单击关系曲线。
右侧弹出关系编辑框。
 - ii. 在关系编辑框中的“名称”文本框里填写关系的名称。
名称只能是字母、数字、下划线组成的合法字符串。
- 删除关系
右键单击关系曲线，单击“Delete”即可删除关系曲线。

本体的所有概念节点和关系曲线均编辑完成后，单击“保存”，即可成功创建一个新的本体。

创建完本体后，您可以在“本体管理”页面查看已创建的本体，并通过“操作”栏的操作[预览本体](#)、[编辑本体](#)、[导出本体](#)、[删除本体](#)。

3.2.2 OBS 导入本体

前提条件

将待导入的本体文件（json格式）上传至OBS中，详情请参见[OBS添加桶](#)和[OBS上传文件或文件夹](#)。

从 OBS 导入

1. 登录[KG服务管理控制台](#)，在左侧菜单栏中单击“本体管理”，进入“本体管理”页面。
2. 在“本体管理”页面，单击“OBS导入”，弹出“导入本体”对话框。
3. 在导入本体对话框中，填写如下信息：
 - “本体名称”：填写本体的名称。
 - “OBS桶”：选择提前准备的本体文件存放的OBS桶。
 - “存储路径”：选择提前准备的本体文件存放在OBS桶中的路径。
4. 确认信息后，单击“确定”，跳转至“本体管理”页面。
您可以查看从OBS导入的本体，并通过“操作”栏的操作预览本体、编辑本体、导出本体、删除本体。

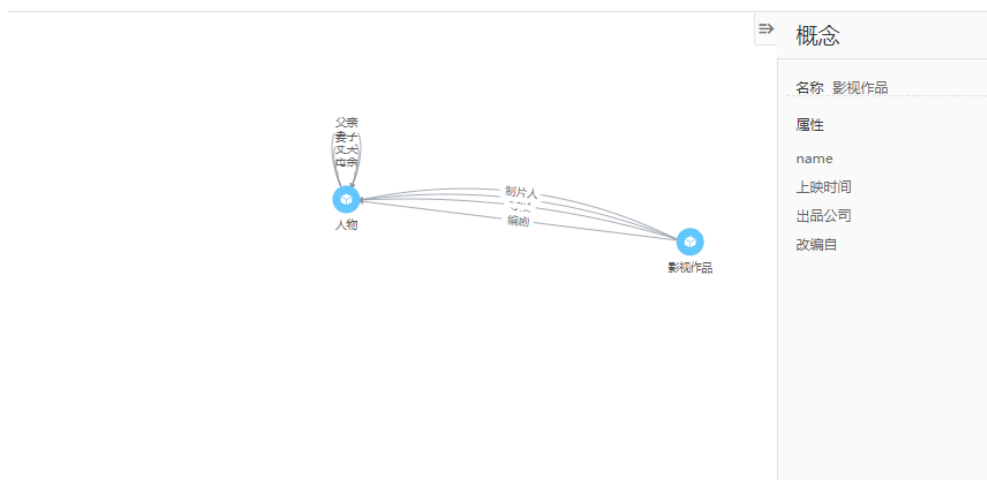
3.3 预览本体

在KG服务控制台创建本体后，您可以预览描述本体信息的结构图。

1. 登录[KG服务管理控制台](#)，在左侧菜单栏中单击“本体管理”，进入本体管理页面。
2. 在本体列表中，单击操作栏的“预览”，即可预览本体信息的结构图。

图 3-3 预览本体

本体预览



3. 单击编辑框的概念节点和关系曲线。
右侧弹出信息框，可查看概念和关系的名称和属性。






3.4 编辑本体

在KG控制台创建本体后，您可以通过编辑本体信息，更新本体。

1. 登录**KG服务管理控制台**，在左侧菜单栏中单击“本体管理”，进入本体管理页面。
2. 在本体列表中，单击操作栏的“编辑”，进入编辑页面。

您可以在本体编辑页面，对本体信息进行编辑，可进行的具体操作如表3-1所示。

表 3-1 编辑本体操作说明

操作名称	操作步骤
修改本体描述	在“本体描述”文本框中填写待创建的本体描述。
增加概念节点	拖拽编辑框左上角的概念节点  悬放在编辑框的目标位置。
删除概念节点	右键单击概念节点  , 单击“Delete”即可删除概念节点。
编辑概念名称和属性	<ol style="list-style-type: none">1. 单击编辑框内的概念节点。 右侧弹出概念编辑框。2. 在概念编辑框中的“名称”文本框里编辑概念的名称。 概念名称只能是字母、数字、下划线组成的合法字符串。3. 在默认属性“name”下方，单击“新增属性”可添加属性文本框，在文本框中编辑概念对应的属性。 属性只能是名称只能是字母、数字、下划线、中划线或者中文字符组成的合法字符串。
增加关系曲线	<ul style="list-style-type: none">● 创建单概念的关系曲线。 双击概念节点。● 创建概念间的关系。 单击概念节点, 拖拽出关系虚线至目标概念节点，单击目标概念节点。
删除关系	右键单击关系曲线，单击“Delete”即可删除关系曲线。
编辑关系名称	<ol style="list-style-type: none">1. 单击关系曲线。 右侧弹出关系编辑框。2. 在关系编辑框中的“名称”文本框里填写关系的名称。 名称只能是字母、数字、下划线组成的合法字符串。

3. 修改完成后，单击“保存”，完成本体编辑。

3.5 导出本体

针对在KG服务控制台创建的本体，您可以把本体导出至OBS文件夹中。

1. 登录**KG服务管理控制台**，在左侧菜单栏中单击“本体管理”，进入本体管理页面。

2. 在本体列表中，单击操作栏的“导出”，弹出对话框。
3. 在对话框中选择OBS存储本体的路径：
 - 在“OBS桶”和“文件选择”文本框中选择导出本体的存储OBS桶和文件夹。
 - 在“导出文件名”文本框中填写存储本体的文件夹名称。
4. 单击“导出本体”，完成导出本体操作。

3.6 删除本体

针对在KG服务控制台创建的本体，您可以删除本体以适应业务变化。

1. 登录**KG服务管理控制台**，在左侧菜单栏中单击“本体管理”，进入本体管理页面。
2. 在本体列表中，单击操作栏的“删除”。
3. 单击“确定”，确认删除本体。

4 自定义信息抽取模型

4.1 信息抽取模型简介

KG服务提供自定义信息抽取模型功能，如果您希望在信息抽取时使用自定义的信息抽取模型，您可以在KG模型管理页面创建抽取模型、创建并发布模型版本，用于创建知识图谱过程中的信息抽取。

在KG控制台上自定义信息抽取模型是一种基于schema约束的中文信息抽取模型，即从自然语言文本中抽取三元组的模型，因此仅适用于抽取基础数据格式为txt文本的自然语言短句。

自定义模型流程

自定义信息抽取模型的流程如表4-1所示。

表 4-1 自定义信息抽取模型流程

流程	说明	操作指引
准备训练数据	提前准备用于训练模型的数据。	准备训练数据
创建模型	基于您的训练数据（即已标注数据），您可以在KG服务中创建信息抽取模型。	创建模型
创建新版本	创建模型后，您需要创建模型版本，方便后续版本发布和模型调优。	创建新版本
发布版本	基于您的业务要求，您可以发布已满足要求的模型版本，用于信息抽取。 您还可以对版本进行修改和删除操作，不断优化模型。	发布版本 修改版本 删除版本

4.2 准备训练数据

在创建抽取模型时，需要您提前准备用于训练模型的数据上传至OBS目录，数据格式为txt文本的自然语言短句。KG服务当前支持的数据类型请参见[训练数据类型介绍](#)。

准备数据流程如下：

1. [准备待标注的数据](#)
2. [定义三元组类型（schema）](#)
3. [标注数据](#)
4. [上传至OBS](#)

准备待标注的数据

准备待标注的训练数据，并保存至本地。数据格式为短文本txt格式，文件内容示例如下：

张三的生日是1990年1月1日，身高175cm，出生于北京。

李四，著名导演，毕业于电影学院，代表作有《电影1》、《电影2》。

... ..

定义三元组类型（schema）

为了训练自定义的信息抽取模型，需要定义该抽取模型能够抽取的三元组类型，并在训练数据中进行标注。

三元组是知识图谱构建的基本元组，三元组是一个主谓宾结构，由主语（subject）、谓语（predicate）、宾语（object）组成。自然语言中的知识可以建模为三元组的格式，如“中国的首都是北京”这句话中，就有“中国-首都-北京”这样的三元组。

为了将三元组分类，我们定义三元组的类型由主语类型（subject type）、谓语（predicate）类型、宾语类型（object type），如上述三元组“中国-首都-北京”的类型是“国家-首都-城市”。通常我们将这些三元组类型组成的集合称为模型的schema。

标注数据

为了训练自定义的信息抽取模型，需要在训练数据中标注三元组类型。

您可以选择使用[BRAT](#)、[MODELARTS](#)或其他标注手段进行标注。

上传至 OBS

将用于训练模型的数据上传至OBS目录：

- 创建OBS桶和文件夹的操作指导请参见[创建桶](#)和[新建文件夹](#)。确保您使用的OBS目录与KG服务在同一区域。
- 上传训练数据请参考[上传文件](#)。

4.3 创建模型

在创建知识图谱之前，您可以在KG服务上通过选择模型框架和上传训练数据来创建模型，用于后续创建图谱过程中的信息抽取。

训练模型框架介绍

KG服务提供不同类型的关系抽取模型，以应对用户不同的条件与需求。KG服务当前仅有一个训练模型框架供您自定义模型：MRC-BM。

MRC-BM是一个非常简单的、基于BERT模型、端到端的关系抽取模型，它对于标注数据量的要求非常低，平均每个关系仅需要200组左右标注数据即可训练出一个较好（实测F1值在0.7左右）的模型，平均每个关系标注数据在2000左右模型训练效果达到一个峰值（实测F1值在0.8左右），数据量越多，质量越好，效果会越好。

训练数据类型介绍

在创建抽取模型时，需要您提前准备训练数据上传至OBS目录。KG服务当前支持的数据类型包括“DATASET”、“BRAT”、“MODELARTS”。

- **DATASET**

DATASET数据需要用户指定的文件夹中包含两个指定格式与名称的文件：schema.json和train_data.json。前者用来指定需要抽取的关系三元组类型，后者用来存放所有的标注数据。

schema.json一行一句json字符串，表示一种关系三元组类型：predicate表示谓词，subject表示主语类型，object表示宾语类型。例如：

```
{ "predicate": "出生地", "subject": "人物", "object": "地点" }
```

train_data.json一行一句json字符串，表示一条标注数据：其中text是文本，spo_list是spo三元组的list，每个spo三元组包含7个元素，predicate为三元组谓词，object_type为宾语类型，subject_type为主语类型，object_position为宾语在文中出现的位置，subject_position为主语在文中出现的位置，object为宾语词，subject为主语词。例如：

```
{ "text": "比如钱钟书和杨绛、鲁迅和许广平都只有一个子女", "spo_list": [ { "predicate": "妻子", "object_type": "人物", "subject_type": "人物", "object": "杨绛", "subject": "钱钟书" }, { "predicate": "丈夫", "object_type": "人物", "subject_type": "人物", "object": "钱钟书", "subject": "杨绛" } ] }
```

- **BRAT**

BRAT是一个开源的文本标注工具，主要用于对文本的结构化标注，您可以前往[官网](#)下载。KG服务的预置模型框架支持直接利用BRAT关系三元组标注的结果来训练模型。注意您的标注任务仅限于BRAT的关系三元组标注，而且不要使用BRAT标注系统中的高级内容（如定义关系的自反性、传递性等）。

一个典型的BRAT文本标注结果文件夹包含如下文件：annotation.conf，file.txt，file.ann。其中annotation.conf是BRAT标注配置伟岸，file.txt是待标注文本文件，可以有多个，每个文本文件都有一个同名的file.ann文件作为标注结果文件。

Annotation.conf：配置文件需要满足BRAT标注系统的要求，一个典型的配置文件如下：

```
[entities]
人物
Date
图书作品
[relations]
出生日期 Arg1: 人物, Arg2: Date
```

```
作者 Arg1: 图书作品, Arg2: 人物  
[events]  
[attributes]
```

File.ann: 用户上传文本文件同名的标注结果文件, 上面标注后结果文件一般如下:

```
历史人物 54151 54153 蔡京  
T3153 Text 54181 54183 北宋  
人物 54151 54153 蔡京  
T3155 Date 54154 54164 1047年2月14日  
朝代 Arg1:T3152 Arg2:T3153  
出生日期 Arg1:T3154 Arg2:T3155
```

● MODELARTS

ModelArts是华为云面向开发者的一站式AI开发平台, 提供了一系列文本标注工具, 您可以在ModelArts上面进行标注, 详细操作指导请参见[ModelArts用户指南](#) > [文本三元组数据标注](#), 按照要求将标注结果导出至OBS。在[发布数据集](#)时, “版本格式”选择“Default”。

ModelArts发布的标注数据集是.manifest格式的文件, 其中包含多行, 每行是一个JSON格式的标注样例。

例如典型的标注样例如下:

```
{"usage": "train", "source": "content://张三, 字树人, 某某省某某市人, 某市钢铁集团公司第十一子公司  
退休职工, 生于1941年", "id": "32f2f201e522c83742e7a23ab9f89084.000000001", "annotation":  
[{"name": "人物", "property": {"@modelarts:start_index": 0, "@modelarts:end_index": 3}, "id": "E1",  
"type": "modelarts/text_entity", "creation-time": "2019-11-27 16:35:16", "annotated-by": "human/  
account/user"}, {"name": "地点", "property": {"@modelarts:start_index": 8, "@modelarts:end_index":  
14}, "id": "E2", "type": "modelarts/text_entity", "creation-time": "2019-11-27 16:35:16", "annotated-  
by": "human/account/user"}, {"name": "出生地", "property": {"@modelarts:from": "E1",  
"@modelarts:to": "E2"}, "id": "R1", "type": "modelarts/text_triplet", "creation-time": "2019-11-27  
16:35:16", "annotated-by": "human/account/user"}]}
```

其中usage字段表示其用途, source字段为待标注文本信息, annotation字段为标注结果, 里面包含两类结果: text_entity和text_triplet, 由type字段区分。

text_entity为实体标注结果, name为实体类型标签, id为实体ID, property为标注的文本信息 (@modelarts:start_index为头, @modelarts:end_index为尾), creation-time为标注时间, annotated-by为标注账号; text_triplet为三元组标注结果, name为三元组类型标签, id为三元组ID, property为标注的三元组信息 (@modelarts:from为标注的头实体id, @modelarts:to为标注的尾实体id), creation-time为标注时间, annotated-by为标注账号。

前提条件

准备训练数据并上传至OBS目录, 详细步骤请参见[准备训练数据](#), 数据类型和要求请见[训练数据类型介绍](#)。

操作步骤

1. 登录[KG服务管理控制台](#), 在左侧菜单栏中单击“模型管理”, 进入“模型管理”页面。
 2. 在模型列表左上方, 单击“创建模型”。
- 弹出“创建模型”对话框, 如[图4-1](#)所示。

图 4-1 创建模型

创建模型

* 模型名称

test1122

* 选择训练模型

SimpleBertModel

* 选择数据类型

DATASET

* 选择OBS桶

* 选择文件夹

.. 返回上一级

文件

brat_data_10000/

brat_data_100000/

确定

取消

3. 按[表4-2](#)填写相关信息。

表 4-2 创建模型参数说明

参数	说明
模型名称	模型的名称。暂不支持修改。
选择训练模型	选择预置的训练模型。KG服务预置模型仅有“MRC-BM”，详情请见 训练模型框架介绍 。
选择数据类型	选择训练模型的数据类型，即待抽取的信息数据类型。KG服务支持的数据类型包括“DATASET”、“BRAT”、“MODELARTS”，详情请见 训练数据类型介绍 。
选择OBS桶	选择训练数据存储的OBS桶。
选择文件夹	选择训练数据存储在OBS桶中的文件夹。
描述	模型的描述。

4. 单击“确定”，完成模型创建。
- 跳至“模型管理”页面，新创建的模型显示在模型列表中。

查看模型

- 查看模型列表
- 完成模型创建后，登录[KG服务管理控制台](#)，您可以在左侧菜单栏中单击“模型管理”，查看模型列表，包括模型的“状态”、“更新时间”、“OBS路径”、“描述”、“操作”，其中模型状态包括“可用”、“训练中”、“不可用”，如图4-2所示。
- 您也可以在列表右上方选择状态过滤查看模型，或在搜索文本框中输入模型名称，搜索查看模型。

图 4-2 模型列表

模型名称	状态	更新时间	OBS路径	描述	操作
test19	不可用	2019-11-20 18:39:03	/kqp-test-mzdnss/1.0.2/data/dataset_data_100...	test_voy	查看 删除
test	可用	2019-11-20 16:42:22	/kqp-test-mzdnss/1.0.2/data/dataset_data_100...	-	查看 删除

- 查看模型详情
- 登录[KG服务管理控制台](#)，在左侧菜单栏中单击“模型管理”，在模型列表中单击操作栏的“查看”，进入模型详情页面。
- 模型详情页面包含信息如表4-3所示。

图 4-3 模型详情

已发布版本

抽取模型

SimpleBertModel

模型评估

F1: 0.71 | 准确率: 0.79 | 召回率: 0.65

训练数据

更新时间

2019-11-20 16:42:22

状态

已发布

(b53547fd-31bf-44a2-8a55-8bb6e63345bc)

查看 schema

版本管理

创建新版本

您还可以训练 2 个版本。

版本	描述	模型评估 ①	更新时间	状态	操作
<div> <div>▼</div> <div>b53547fd-31bf-44a2-...</div> </div>		<div> <div>①</div> <div>0.71</div> <div>②</div> <div>0.79</div> <div>③</div> <div>0.65</div> </div>	2019-11-20 16:42:22	<div>已发布</div>	<div>停用</div>
<div> <div>▼</div> <div>fad2a090-fa87-4bb5-...</div> </div>	test19	<div> <div>①</div> <div>0.50</div> <div>②</div> <div>0.79</div> <div>③</div> <div>0.37</div> </div>	2019-11-20 16:42:22	<div>停用</div>	<div>修改</div> <div>发布</div> <div>删除</div>
<div> <div>▼</div> <div>ffe8027b-e3a2-4ab2-...</div> </div>	-	<div> <div>①</div> <div>0.37</div> <div>②</div> <div>0.82</div> <div>③</div> <div>0.24</div> </div>	2019-11-19 19:15:47	<div>停用</div>	<div>修改</div> <div>发布</div> <div>删除</div>

表 4-3 模型详情

详情分类	参数	说明
已发布版本	抽取模型	已发布版本的模型所选择的预置模型。
	模型评估	已发布版本的模型评估参数。 <ul style="list-style-type: none">准确率：在被所有预测为正的样本中实际为正样本的概率。召回率：在实际为正的样本中被预测为正样本的概率。F1：综合考虑准确率和召回率的影响，由两者计算而来，越接近1代表模型越好。




详情分类	参数	说明
	训练数据	已发布版本的模型，其训练数据的OBS存储目录。
	更新时间	已发布的模型版本的更新时间。
	状态	已发布的模型版本状态，包括“训练中”、“训练完成”、“训练失败”、“已发布”、“停用”。
	查看schema	KG服务把信息抽取任务定义为schema约束下的关系三元组抽取任务，schema表明从文本信息中抽取的知识类型。 <ul style="list-style-type: none">• subject_type：主语• predicate：谓语• object_type：宾语 单击可查看最新发布的模型版本可抽取的知识种类，如图4-4所示。
版本管理	版本	版本的名称。
	描述	版本的描述。
	模型评估	版本的模型评估参数。 <p> P：准确率（Precision）是在被所有预测为正的样本中实际为正样本的概率。</p> <p> R：召回率（Recall）是在实际为正的样本中被预测为正样本的概率。</p> <p> F1：F1值综合考虑准确率和召回率的影响，由两者计算而来，越接近1代表模型越好。</p>
	更新时间	版本的更新时间。
	状态	版本的状态，包括“训练中”、“训练失败”、“训练完成”、“已发布”、“停用”。
	操作	针对不同版本状态，可进行对应的操作。 <ul style="list-style-type: none">• 修改：修改当前版本的模型参数。• 发布：发布当前版本。• 删除：删除当前版本。• 停用：停用当前版本，使模型不能用于信息抽取。

图 4-4 查看 schema

查看schema

subject_type	predicate	object_type
人物	祖籍	地点
人物	父亲	人物
企业	总部地点	地点
人物	出生地	地点
生物	目	目
行政区	面积	Number
机构	简称	Text
影视作品	上映时间	Date
人物	妻子	人物
歌曲	所属专辑	音乐专辑

10 ▾

总条数: 50

< 1 2 3 4 5 >

取消

4.4 管理版本

4.4.1 创建新版本

创建抽取模型后，您需要创建模型的版本，才能进行后续的发布操作，使模型在创建知识图谱时用于信息抽取。

每个模型最多可创建5个模型版本。

前提条件

在KG管理控制台创建抽取模型，详情请见[创建模型](#)。

操作步骤

1.

登录[KG服务管理控制台](#)，在左侧菜单栏中单击“模型管理”，进入模型管理页面。
1.

在模型列表中，单击模型名称，进入模型详情页。
2.

在“版本管理”区域单击“创建新版本”。
弹出“创建新版本”对话框。
3.


配置模型参数的参数值，输入新版本的描述内容。
可将鼠标移至参数右侧, 查看各参数的含义和推荐值。

图 4-5 创建新版本

创建新版本

★ 模型参数

参数	参数值
batch_size ?	<input type="text" value="32"/>
epochs ?	<input type="text" value="10"/>
max_len ?	<input type="text" value="192"/>
learning_rate ?	<input type="text" value="0.00005"/>

描述

0/255

确定

取消

4. 信息填写完毕后，单击“确定”，成功创建模型的新版本。

4.4.2 发布版本

创建模型版本后，您需要通过发布操作，发布模型版本后，才能使对应的版本模型在创建知识图谱时用于信息抽取。

前提条件

已[创建模型](#)，并针对已创建的模型[创建版本](#)。

操作步骤

1. 登录[KG服务管理控制台](#)，在左侧菜单栏中单击“模型管理”，进入模型管理页面。
2. 在模型列表中，单击模型名称，进入模型详情页。
3. 在版本管理列表中的“操作”栏，单击“发布”，待状态栏处于“已发布”状态，对应模型版本即可用于创建图谱过程中的信息抽取。

4.4.3 修改版本

创建抽取模型后，针对处于“训练完成”和“停用”状态的模型版本，您可以根据自身业务需要，对模型版本进行修改优化。版本状态请参见[查看模型](#)。

操作步骤如下：

1. 登录[KG服务管理控制台](#)，在左侧菜单栏中单击“模型管理”，进入模型管理页面。
2. 在模型列表中，单击“模型名称”，进入模型详情页。
3. 在版本管理列表中的“操作”栏，单击“修改”，弹出“修改版本”对话框。
4. 在“修改版本”对话框中，填写“模型参数”的参数值和“描述”，如[图4-6](#)所示，鼠标移至[?]可查看参数含义和推荐值。

图 4-6 修改版本

修改版本

★ 模型参数

参数	参数值
batch_size [?]	<input type="text" value="32"/>
epochs [?]	<input type="text" value="10"/>
max_len [?]	<input type="text" value="192"/>
learning_rate [?]	<input type="text" value="0.00005"/>

描述

4/255

确定

取消

5. 单击“确定”，完成模型版本的修改。

4.4.4 删除版本

根据自身业务需要，您可以删除模型版本。处于“训练完成”和“停用”状态的模型版本才能进行删除操作，版本状态请参见[查看模型](#)。

操作步骤如下：

1. 登录[KG服务管理控制台](#)，在左侧菜单栏中单击“模型管理”，进入模型管理页面。
2. 在模型列表中，单击模型名称，进入模型详情页。
3. 在版本管理列表中的“操作”栏，单击“删除”。

5 步骤一：选择本体

在创建知识图谱时，首先您需要把已创建的本体导入至待创建的图谱中，用于组织待创建图谱的信息结构。

前提条件

已创建本体，具体创建方式见[创建本体](#)。

操作步骤

1. 登录[KG服务管理控制台](#)，默认进入“我的图谱”页面，单击 $+$ ，进入“新建图谱”页面。
2. 在“新建图谱”页面，填写相关信息。
 - “图谱名称”：在“图谱名称”文本框中填写待创建图谱的名称，名称只能是字母、数字、下划线组成的合法字符串。
 - “图谱规格（边数）”：选择待创建的图谱规格，当前可选“一万”、“一百万”。
3. 在“新建图谱”页面下方，单击“本体选择”，右侧弹出本体选择列表。
4. 在本体列表中勾选合适的本体，单击“保存”，完成本体选择。
可单击操作列的“预览”，预览本体。

图 5-1 本体列表选择本体

本体选择

	名称	更新时间	实体个数	关系个数	操作
<input type="radio"/>	car	2020-01-16 22:24:30	4	6	预览
<input type="radio"/>	gene_world	2020-01-06 16:56:45	2	1	预览
<input type="radio"/>	gene0106	2020-01-06 16:51:50	2	2	预览
<input type="radio"/>	gene	2020-01-06 16:51:19	2	2	预览
<input type="radio"/>	ceshi002	2020-01-03 09:27:30	4	6	预览
<input type="radio"/>	MC_MOXING	2020-01-02 23:43:19	2	2	预览
<input type="radio"/>	ceshi001	2020-01-02 23:24:44	3	3	预览
<input type="radio"/>	asdf	2020-01-02 23:14:52	2	1	预览
<input type="radio"/>	testaaa	2019-12-24 16:09:29	2	2	预览
<input type="radio"/>	xindian	2019-12-20 10:32:53	8	9	预览

10

总条数: 35

< 1 2 3 4 >

保存

取消

选择本体后，进入下一步操作，配置数据源，具体操作步骤请见[步骤二：配置数据源](#)。

6

步骤二：配置数据源

数据是知识图谱的基础。在创建知识图谱时，基于已存储在OBS中的基础数据，您选择本体之后，需要从OBS导入用于构建知识图谱的基础数据。

数据格式要求

- 多行紧凑json文件

该类型为结构化数据输入格式。文件中，每一行为一个完整的紧凑（单行）JSON字符串。文件名需以.json结尾，如“input.json”。

```
{"entity_type": "Person", "中文名": "康时", "国籍": "中国", "职业": "演员", "出生日期": "1985年3月3日", "url": "/film.kg.huawei.com/康时/165882308"}
{"entity_type": "Person", "中文名": "戴奇", "国籍": "韩国", "职业": "演员, 歌手", "出生日期": "1987年6月22日", "url": "/film.kg.huawei.com/戴奇"}
```

[最佳实践>非结构化数据创建图谱](#)中给出数据样例下载链接，您可根据自身需要下载使用。

- 多行单句文本文件

该类型为非结构化输入格式。文件中，每一行为一个自然语言短句，文件名需以.txt结尾，如“input.txt”。

项羽，名籍

项羽是楚国武将

文昭甄皇后史称甄夫人

文昭甄皇后中山无极（今河北省无极县）人，上蔡令甄逸之女



文昭甄皇后魏文帝曹丕的妻子，魏明帝曹叡的生母

.....

前提条件

- 已创建用于存储基础数据的OBS桶及文件夹，详情请见[添加桶](#)。
- 待创建图谱的基础数据已[上传至OBS](#)，上传至OBS的基础数据格式请见[数据格式要求](#)。
- 已创建本体，具体创建方式见[创建本体](#)。
- 已进入“新建图谱”页面选择本体，详情请参见[步骤一：选择本体](#)。


操作步骤

1. 在“新建图谱”页面，单击“数据源配置”，右侧弹出“数据源配置”对话框。
2. 在“数据源配置”对话框，填写相关信息。
 - “选择数据源文件”：单击，弹出“选择数据源文件”对话框，选择数据源存放在OBS的路径：L
 - i. “OBS桶”：选择数据源文件存放的OBS桶。
 - ii. “存储路径”：选择数据源文件存放在OBS桶的文件路径。选择完成后，单击“确定”。
 - “实体类型字段”：在“实体类型字段”文本框中填写基础数据中描述实体类型的函数字段。
3. 填写完信息后，单击“保存”。
 - 如果您待创建的图谱只需要数据一个数据源，那么您已完成数据源的配置。
 - 如果您待创建的图谱需要输入多个数据源，请进入步骤4。
4. 单击本体选择右侧的，添加新的数据源，然后回到步骤1，配置新的数据源。

配置数据源后，进入创建图谱的下一步操作，信息抽取，具体操作步骤请见[步骤三：配置信息抽取](#)。

示例


以创建一个有关于电影的知识图谱为例，展示如何在创建图谱时配置数据源。

1. 在“新建图谱”页面，单击“数据源配置”，右侧弹出“数据源配置”对话框。
2. 在“数据源配置”对话框，单击“选择数据源文件”右侧的：单击，弹出“选择数据源文件”对话框。
3. 在“选择数据源文件”对话框中，选择数据源文件存放的“OBS桶”和“存储路径”，单击“确定”。

“数据源配置”对话框下方可预览基础数据，如[图6-1](#)所示。

图 6-1 数据源配置

数据源配置

* 选择数据源文件 

* 实体类型字段②

```
{
  "entity_type": "Person", "中文名": "严焱", "国籍": "中国", "职业": "演员、导演、制片人、编剧、教师、歌手",
  "entity_type": "Person", "中文名": "章岩", "国籍": "中华人民共和国", "职业": "演员", "出生日期": "1988年",
  "entity_type": "Film", "中文名": "红颜知己", "票房": "34.15亿", "上映时间": "1986年1月22日", "导演": "滕",
  "entity_type": "Film", "中文名": "西部牛仔", "票房": "21.50亿", "上映时间": "2000年9月10日", "导演": "滕",
  "entity_type": "Person", "中文名": "陶园", "国籍": "中国([link]@中国://film.kg.huawei.com/中国)", "职业": "演员",
  "entity_type": "Person", "中文名": "平庆", "国籍": "中国", "职业": "演员", "出生日期": "1981年3月6日",
  "entity_type": "Film", "中文名": "满目疮痍", "票房": "29.35亿", "上映时间": "1986年7月3日", "导演": "戴",
  "entity_type": "Film", "中文名": "不止英雄", "票房": "19.50亿", "上映时间": "1987年6月6日", "导演": "滕",
  "entity_type": "Person", "中文名": "孔婉", "国籍": "中国", "职业": "演员、学生", "出生日期": "1997年10月",
  "entity_type": "Person", "中文名": "祁露", "国籍": "中国", "职业": "演员、歌手", "出生日期": "1967年7月",
  "entity_type": "Person", "中文名": "郎珍", "国籍": "中国", "职业": "演员、模特", "出生日期": "1991年9月",
  "entity_type": "Person", "中文名": "葛璐", "国籍": "中国", "职业": "演员、导演", "出生日期": "1974年4月",
  "entity_type": "Person", "中文名": "伍蕊", "国籍": "中华人民共和国", "职业": "演员、歌手", "出生日期": "1989年4月",
  "entity_type": "Person", "中文名": "乐翰", "国籍": "中国", "职业": "演员、歌手", "出生日期": "1989年4月",
  "entity_type": "Film", "中文名": "穿越太空", "票房": "34.18亿", "上映时间": "1954年12月25日", "导演": "滕",
  "entity_type": "Person", "中文名": "穆竹", "国籍": "中国", "职业": "演员、导演、出品人", "出生日期": "1989年4月",
  "entity_type": "Film", "中文名": "霞光万道", "票房": "9.14亿", "上映时间": "1992年5月8日", "导演": "贝利",
  "entity_type": "Film", "中文名": "保卫和平", "票房": "39.04亿", "上映时间": "1958年1月16日", "导演": "滕",
  "entity_type": "Person", "中文名": "计广", "国籍": "中国", "职业": "演员、歌手", "出生日期": "1980年12月",
  "entity_type": "Person", "中文名": "卫希", "国籍": "中国", "职业": "演员、歌手", "出生日期": "1967年1月",
}
```

- 在“实体类型字段”文本框中填写数据中标识该实体属于何种类型的字段“entity_type”。
- 单击右下角“保存”，完成数据源配置。

7 步骤三：配置信息抽取

7.1 配置信息抽取简介

在创建知识图谱时，当您选择本体后，您需要配置信息抽取，输入实体类型、抽取函数及抽取前后的数据字段，才能进行配置信息映射、配置知识融合等操作。

基本概念

信息抽取是从基础数据中抽取待创建图谱的实体、属性信息以及实体间的相互关系。

配置方式

信息抽取分为普通抽取和预置算法抽取，其适用范围和抽取方式如表7-1所示。

表 7-1 配置方式说明

配置方式	适用范围	具体方式	操作指引
普通信息抽取	基础数据格式为json格式	交互界面配置： 通过输入实体类型、抽取函数及抽取前后的数据字段，完成信息抽取配置。 所支持的信息抽取函数请见 信息抽取函数 。	交互界面配置
		格式化编辑： 将交互界面中的配置转换为json格式配置，通过编辑json配置信息，完成信息抽取配置。 所支持的信息抽取函数请见 信息抽取函数 。	格式化配置
算法信息抽取	基础数据格式为txt文本的自然语言短句	通过选择合适的算法，完成信息抽取配置。 可选择已有的预置算法，也可选择您自定义的模型。	算法抽取

说明

在通过普通抽取方式进行信息抽取时，交互界面配置与格式化编辑配置可以相互切换。

7.2 普通抽取

7.2.1 信息抽取函数

普通抽取适用于json格式的基础数据，选择普通抽取方式信息抽取时，可分为交互界面配置和格式化配置。

在通过普通抽取方式进行信息抽取时，即用交互界面配置或用格式化编辑进行信息抽取时，需要配置引用字段的抽取函数。

KG服务支持的信息抽取函数如表7-2所示。其中，“字段”表示基础数据中的字段名，在抽取函数中引用字段时，使用格式为“\${字段}”，例如抽取基础数据中属性“name”的字段，在抽取函数中引用字段的格式为“trim(\${name})”。

表 7-2 信息抽取函数说明表

抽取函数	函数说明	抽取函数示例	抽取前数据示例	抽取后数据示例
trim（字段）	裁剪字段前后的空白字符	trim(\${name})	"name": "mike "	"mike"
substring (字段, int pos, int len)	从共有len个字符的字段值中，获取从该字段值第pos个字符（从0开始）开始的子串字段值	substring (\$ {id}, 1, 6)	"id": "a000111"	"000111"
split（string pattern）字段）	将字段值以pattern为分隔符分割，得到一个列表。pattern可以是一个正则表达式，所有与pattern匹配的子串都作为分隔符。	split(\${roles}, ',')	"roles": "role1,role2"	["role1", "role2"]
concat（字段1, 字段2, ...）	将字段1、字段2...依次拼接成一个长字符串。此处字段值也可以是一个字符串常量。	concat(\${name}, '--', \${id})	"name": "mike", "id": "a000111"	"mike--a000111"

抽取函数	函数说明	抽取函数示例	抽取前数据示例	抽取后数据示例
regexp_replace (字段, string pattern, string replacement)	将字段值中与正则表达式pattern匹配的部分替换为replacement	regexp_replace(\${info}, 'hello', 'world')	"info": "hello world"	"worldworld"
upper (字段)	将字段中英文字母全部转为大写	upper(\${name})	"name": "mike"	"MIKE"
lower (字段)	将字段中英文字母全部转为小写	lower(\${name})	"name": "JOHN"	"john"
initcap (字段)	将字段中英文首字母转为大写	initcap(\${city})	"city": "hangzhou"	"Hangzhou"
coalesce (字段1, 字段2, ...)	返回字段1, 字段2, ...中第一个不为空的字段	coalesce(\${name}, \${名字}, \${中文名} ...)	"name": " ", "名字": " ", "中文名": "小明"	"小明"
parse_array (字段)	尝试将字段值解析为一个列表	parse_array(\${list})	"list": ["item1", "item2"]	['item1', 'item2']
array_get (字段, int index)	获取列表类型的字段中的第index个值 (index从0开始)	array_get(parse_array(\${list}), 0)	"list": ["item1", "item2"]	"item1"

7.2.2 交互界面配置

前提条件

- 已完成本体选择和配置数据源，具体操作步骤见[步骤一：选择本体](#)和[步骤二：配置数据源](#)。
- 熟悉KG服务支持的[信息抽取函数](#)。


操作步骤

- 在“新建图谱”页面，单击“信息抽取”，右侧弹出“信息抽取”对话框。
- 在“信息抽取”对话框中，“抽取方式”选择“普通抽取”，“编辑方式”选择“交互界面”。

说明

在“交互界面”页签填写信息抽取前后的相关字段时，可单击“格式化编辑”，切换至“格式化编辑”页签，查看json代码。

- 选择“保留源数据字段”开关。

开启 ，KG服务保留基础数据的数据字段作为抽取信息后的“数据字段”。

4. 在“信息抽取”对话框中填写信息抽取项：按表7-3填写“实体类型”，“数据字段”和“字段函数”。

表 7-3 信息抽取项属性说明

属性	说明
实体类型	基础数据中描述实体类型的字段。
数据字段	抽取信息后的属性字段。
字段函数	抽取信息时所使用的引用字段的信息抽取函数，具体请见 信息抽取函数 。

- 增加信息抽取项：单击“增加信息抽取项”，即可增加一项抽取信息类型。

- 删除信息抽取项：勾选待删除的信息抽取项，单击上方“删除”，或单击待删除的信息抽取项操作栏的“删除”，即可删除信息抽取项。
5. 确认所填写的信息后，单击“保存”，完成信息抽取。

说明

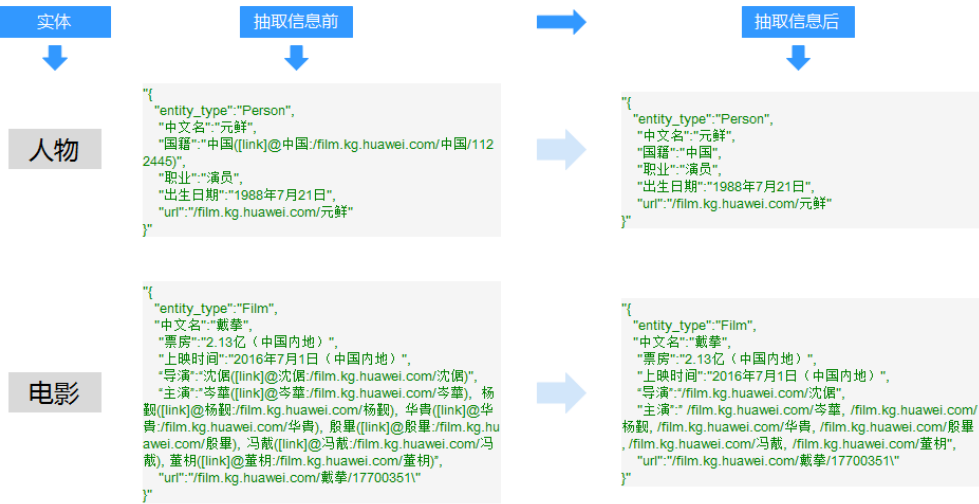
若创建多个数据源，请完成所有数据源的信息抽取配置。配置数据源请参见[步骤二：配置数据源](#)。

配置信息抽取后，进入下一步操作，配置知识映射，具体操作步骤见[步骤四：配置知识映射](#)。

示例

以创建一个有关于电影的知识图谱为例，抽取信息前后的实体信息如图7-1所示，展示如何在创建图谱时配置信息抽取。

图 7-1 信息抽取



1. 按步骤一：选择本体为待创建的图谱选择本体。
2. 按配置数据源>示例为待创建的图谱配置数据源。


- 3. 在“新建图谱”页面，单击“信息抽取”，右侧弹出“信息抽取”对话框。
- 4. 在“信息抽取”对话框中，“抽取方式”选择“普通抽取”，“编辑方式”选择“交互界面”。
- 5. 关闭“保留源数据字段”开关.
- 6. 在“信息抽取”对话框中填写信息抽取项，按表7-4填写。

表 7-4 信息抽取项示例

实体类型	数据字段	字段函数
Person	中文名	regexp_replace(\${中文名}, '\\([link\\]@.*?:/film.*?)\\', '')
	国籍	regexp_replace(\${国籍}, '\\([link\\]@.*?:/film.*?)\\', '')
	职业	\${职业}
	出生日期	\${出生日期}
	url	\${url}
Film	中文名	\${中文名}
	票房	\${票房}
	上映时间	\${上映时间}
	导演	regexp_extract_all(\${导演}, '\\([link\\]@.*?:/film.*?)\\', 1)
	主演	regexp_extract_all(\${主演}, '\\([link\\]@.*?:/film.*?)\\', 1)
	url	\${url}

示例中字段抽取函数主要有三类，示例如下：

- \${职业}
直接引用“职业”字段的值，不做任何操作。
- regexp_replace(\${国籍}, '\\([link\\]@.*?:/film.*?)\\', '')
将“国籍”字段中符合正则模式的“\\([link\\]@.*?:/film.*?)\\”替换成空字符串，即删除符合这个pattern的字符串。例如从字段“中国([link]@中国:/film.kg.huawei.com/中国/1122445)”抽取信息“中国”。
- regexp_extract_all(\${导演}, '\\([link\\]@.*?:/film.*?)\\', 1)
抽取导演字段中，与模式“\\([link\\]@.*?:/film.*?)\\”匹配的数据值，返回的类型是一个列表。例如“导演:庄伟建([link]@庄伟建:/film.kg.huawei.com/庄伟建)，苏万聪([link]@苏万聪:/film.kg.huawei.com/苏万聪)”，抽取信息后为“导演:/film.kg.huawei.com/庄伟建， /film.kg.huawei.com/苏万聪”。

- 7. 确认所填写的信息后，单击“保存”，完成信息抽取。


7.2.3 格式化配置

前提条件

- 已完成本体选择和配置数据源，具体操作步骤见[步骤一：选择本体](#)和[步骤二：配置数据源](#)。
- 熟悉KG服务支持的[信息抽取函数](#)。

操作步骤

1. 在“新建图谱”页面，单击“信息抽取”，右侧弹出“信息抽取”对话框。
2. 在“信息抽取”对话框中，“抽取方式”选择“普通抽取”，“编辑方式”选择“格式化编辑”。
3. 选择“保留源数据字段”开关。

开启“”，KG服务保留基础数据的数据字段作为抽取信息后的“数据字段”。

4. 在“信息抽取”对话框下方编辑json格式代码。

说明

在“格式化编辑”页签编辑json代码配置知识映射时，可单击“交互界面”，切换至“交互界面”页签，查看知识映射前后的相关字段。

5. 单击“保存”，完成信息抽取。

说明

若创建多个数据源，请完成所有数据源的信息抽取配置。配置数据源请参见[步骤二：配置数据源](#)。

示例

以一个电影实体为例，抽取信息前后的实体信息如[图7-1](#)所示，代码示例如下所示：

```
{
  "ie_type": "function",
  "function_ie_configs": {
    "ie_configs": {
      "Person": [
        {
          "key": "中文名",
          "value_function": "regexp_replace(${中文名}, '\\(\\[link\\]@.*?:/film.*?\\)', '')"
        },
        {
          "key": "国籍",
          "value_function": "regexp_replace(${国籍}, '\\(\\[link\\]@.*?:/film.*?\\)', '')"
        },
        {
          "key": "职业",
          "value_function": "${职业}"
        },
        {
          "key": "出生日期",
          "value_function": "${出生日期}"
        },
        {
          "key": "url",
          "value_function": "${url}"
        }
      ]
    }
  }
}
```

```
"Film": [
  {
    "key": "中文名",
    "value_function": "${中文名}"
  },
  {
    "key": "票房",
    "value_function": "${票房}"
  },
  {
    "key": "上映时间",
    "value_function": "${上映时间}"
  },
  {
    "key": "导演",
    "value_function": "regexp_extract_all(${导演}, '\\(\\(\\[link\\]\\)@.*?:(/film.*?)\\)', 1)"
  },
  {
    "key": "主演",
    "value_function": "regexp_extract_all(${主演}, '\\(\\(\\[link\\]\\)@.*?:(/film.*?)\\)', 1)"
  },
  {
    "key": "url",
    "value_function": "${url}"
  }
]
},
"default_config": false
}
```

7.3 算法抽取

前提条件

- 已完成本体选择和配置数据源，具体操作步骤见[步骤一：选择本体](#)和[步骤二：配置数据源](#)。
- 如果选择使用自己自定义模型，您需要在创建知识图谱之前，自定义用于信息抽取的模型，具体操作请见[自定义信息抽取模型](#)。

操作步骤

1. 在“新建图谱”页面，单击“信息抽取”，右侧弹出“信息抽取”对话框。
2. 在“信息抽取”对话框中，“抽取方式”选择“模型抽取”。
3. 根据自身业务需要，您可以选择“预置模型”或者“用户自定义模型”，二者仅需选择一项。
“预置模型”：当前仅提供“RESBM”和“RelationExtraction”两种预置模型。
“用户自定义模型”：您需提前在控制台自定义模型，详情请参见[自定义信息抽取模型](#)。
4. 可在下方查看从文本信息中抽取的知识类型，如[图7-2](#)所示，知识类型包括主语“object_type”、谓语“predicate”、宾语“subject_type”，即可查看到可抽取的主语、谓语、宾语组合。

图 7-2 抽取模型

信息抽取

抽取方式

普通抽取

模型抽取

预置模型

请选择预置模型

用户自定义模型

kg_person_film

subject_type	predicate	object_type
人物	母亲	人物
影视作品	上映时间	Date
人物	祖籍	地点
影视作品	出品公司	企业
人物	丈夫	人物
人物	出生日期	Date
人物	毕业院校	学校
影视作品	改编自	作品
人物	身高	Number
人物	国籍	国家

10

总条数: 18

<

1

2

>

保存

取消

5. 单击“保存”，完成信息抽取。

说明

若创建多个数据源，请完成所有数据源的信息抽取配置。配置数据源请参见[步骤二：配置数据源](#)。

8

步骤四：配置知识映射

在创建知识图谱时，当您配置信息抽取后，您需要配置信息映射，设置抽取映射前后的相关字段。

基本概念

知识映射是建立从基础数据抽取出的结构化信息与知识图谱本体的映射关系。

前提条件

已按序完成本体选择、基础数据配置、信息抽取配置操作，具体操作步骤见[步骤一：选择本体](#)、[步骤二：配置数据源](#)、[步骤三：配置信息抽取](#)。

配置方式



- 交互界面配置：通过填写基础数据映射前后的相关字段，完成知识映射配置。具体步骤请见[交互界面配置步骤](#)和[交互界面配置示例](#)。
- 格式化编辑：通过编辑json格式代码，完成知识映射配置。具体步骤请见[格式化编辑步骤](#)。

交互界面配置步骤

1. 在“新建图谱”页面，单击“知识映射”，右侧弹出“知识映射”对话框。
2. 在“实体唯一标识字段”文本框中填写识别每个数据实体的字段。
3. “编辑方式”选择“交互界面”。
4. 在知识映射页面右上方，单击“交互界面”，填写知识映射的相关信息。

说明

在“交互界面”页签填写知识映射前后的相关字段时，可单击“格式化编辑”，切换至“格式化编辑”页签，查看json代码。

- 配置实体的类型映射：在对应实体下方的“类型映射”区域填写映射前的实体类型字段。
- 配置实体的属性映射：在对应实体下方的“属性映射”区域填写映射到实体各个属性的数据字段。单击可添加多个数据字段。
- 配置实体的关系映射：在对应实体下方的“关系映射”区域填写映射到各个实体关系的“数据字段”和“尾实体类型”。单击可添加多个数据字段。

5. 单击“保存”，完成知识映射。

📖 说明

若创建多个数据源，请完成所有数据源的知识映射配置。配置数据源请参见[步骤二：配置数据源](#)。

配置知识映射后，进入下一步操作，配置知识融合，具体操作步骤请见[步骤五：配置知识融合](#)。

交互界面配置示例

以创建一个有关于电影的知识图谱为例，展示如何在创建图谱时配置知识映射。

- 1. 按[步骤一：选择本体](#)为待创建的图谱选择本体。
- 2. 按[配置数据源>示例](#)为待创建的图谱配置数据源。
- 3. 按[配置信息抽取>示例](#)为待创建的图谱配置信息抽取。
- 4. 在“新建图谱”页面，单击“知识映射”，右侧弹出“知识映射”对话框。
- 5. 在“知识映射”对话框中，在“实体唯一标识字段”文本框中填写识别每个数据实体的字段“url”。
- 6. 在“知识映射”对话框中，选择编辑方式为“交互界面”。
- 7. 在“知识映射”对话框中，按照抽取后的信息与本体的映射关系，填写相关信息。
 - 知识与实体电影的映射关系，如[图8-1](#)所示。

图 8-1 实体电影的知识映射关系



- 在实体“电影”下方的“类型映射”区域填写映射前的实体类型字段“Film”。
- 填写映射到实体“电影”的数据字段，如[表8-1](#)所示。

表 8-1 实体电影的知识映射字段

属性/关系名称	数据字段
name	中文名
票房	票房
上映时间	上映时间
导演	导演
主演	主演

- 知识与实体人物的映射关系，如图8-2所示。

图 8-2 实体人物的映射关系



- 在实体“人物”下方的“类型映射”区域填写映射前的实体类型字段“Person”。
- 填写映射到实体“人物”的数据字段，如表8-2所示。

表 8-2 实体人物的知识映射字段

属性	数据字段
name	中文名
国籍	国籍
职业	职业
出生日期	出生日期

8. 单击“保存”，完成配置知识映射。

格式化编辑步骤

1. 在“新建图谱”页面，单击“知识映射”，右侧弹出“知识映射”对话框。

2. 在“实体唯一标识字段”文本框中填写识别每个数据实体的字段。
3. “编辑方式”选择“格式编辑”。
1. 在“格式化编辑”页签的编辑框中，输入json格式代码，以映射关系如图8-2为例，代码示例如下所示：

```
{
  {
    "label": "人物",
    "source_keys": [
      {
        "value": "中文名"
      }
    ],
    "target_key": "name",
    "type": "string"
  },
  {
    "label": "人物",
    "source_keys": [
      {
        "value": "国籍"
      }
    ],
    "target_key": "国籍",
    "type": "string"
  },
  {
    "label": "人物",
    "source_keys": [
      {
        "value": "职业"
      }
    ],
    "target_key": "职业",
    "type": "string"
  },
  {
    "label": "人物",
    "source_keys": [
      {
        "value": "出生日期"
      }
    ],
    "target_key": "出生日期",
    "type": "string"
  }
]
```

说明

在“格式化编辑”页签编辑json代码配置知识映射时，可单击“交互界面”，切换至“交互界面”页签，查看知识映射前后的相关字段。

2. 单击“保存”，完成配置知识映射。

9 步骤五：配置知识融合

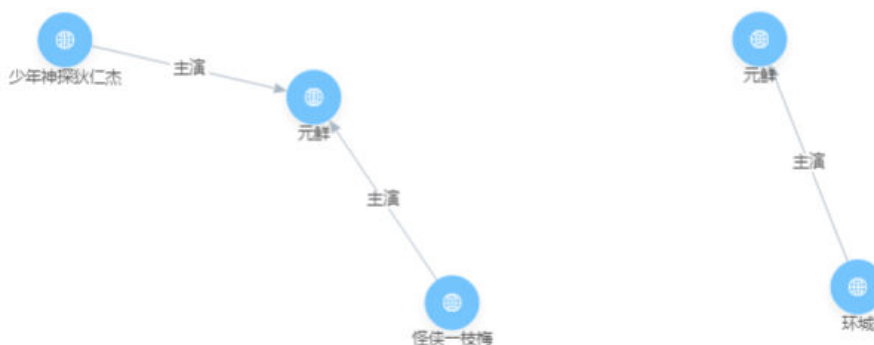
在创建知识图谱时，当您配置知识映射后，您需要配置知识融合，设置知识融合判断属性及相似度函数参数，完成新知识图谱的创建。

基本概念

知识融合是指融合来自多个数据源的关于同一个实体或概念的描述信息，对来自不同数据源的知识在统一规范下进行异构数据整合、消歧。

例如如图9-1所示的两条数据，这两条数据中的“元鲜”实际上是同一个人，因此需要对这两条数据进行融合。

图 9-1 知识融合示例



配置方式

- 交互界面配置：通过填写知识融合判断属性及相似度函数参数，完成知识融合配置。具体步骤请见[交互界面配置步骤](#)和[交互界面配置示例](#)。
- 格式化编辑：通过编辑json格式代码，完成知识融合配置。具体步骤请见[格式化编辑步骤](#)。

前提条件

已按序完成本体选择、基础数据配置、信息抽取配置、知识映射配置操作，具体操作步骤见[步骤一：选择本体](#)、[步骤二：配置数据源](#)、[步骤三：配置信息抽取](#)和[步骤四：配置知识映射](#)。

交互界面配置步骤

1. 在“新建图谱”页面，单击“知识融合”，右侧弹出“知识融合”对话框。
2. 在“知识融合”对话框，
 - 若不需要进行知识融合，则关闭“是否融合”开关，进入[步骤6](#)。
 - 若需要进行知识融合，则打开“是否融合”开关，进入[步骤3](#)。
3. 在“知识融合”对话框的“编辑方式”中，选择“交互界面”。

说明

在“交互界面”页签填写知识融合判断属性及相似度函数参数，可单击“格式化编辑”，切换至“格式化编辑”页签，查看json代码。

4. 配置每个实体的知识融合参数。

“融合标识符”：单击 ▾ 勾选标识知识融合的本体属性，即融合标识符。可添加多个融合标识符。

“属性配置”：填写判断融合的属性配置信息。

- 选择判断融合的属性：在“新实体属性”栏和“待融合实体属性”栏单击 ▾，分别选择判断融合的属性，融合后新数据使用新实体属性的对应值。
- 填写相似度函数：在“相似度函数”栏单击 ▾，选择判断属性融合的相似度函数，当前支持两种相似度函数，一种是“EditSim”（编辑距离），一种是“JaccardSim”（Jaccard距离）。
- 填写相似度阈值：在“相似度阈值”栏填写相似度阈值，也可单击 + 和 - 调整阈值大小。

单击“添加实体类型”可添加多个判断融合的属性配置信息，最多添加9个。单击操作栏“删除”可删除属性配置。

单击“添加分组”，增加一组判断融合的属性配置，按以上提示填写判断融合的属性配置信息。

说明

融合条件：当两条数据在融合标识符相似的前提下，存在某一组所有判断融合的属性相似度大于相似度阈值时，融合这两条数据。

5. 单击“保存”，完成配置知识融合。
6. 在“新建图谱”页面，单击“运行”，完成图谱的创建。

说明

若创建多个数据源，请完成所有数据源的知识融合配置。配置数据源请参见[步骤二：配置数据源](#)。

完成知识融合配置后，即成功创建一个新的知识图谱。

交互界面配置示例

以创建一个有关于电影的知识图谱为例，对人物的属性“name”相似的数据进行知识融合判断，当属性“name”的相似度大于1.00，且属性“职业”的相似度大于1.00

时，数据进行融合。对电影的属性“name”相似的数据进行知识融合判断，当属性“name”的相似度大于1.00，且属性“上映时间”的相似度大于1.00时，数据进行融合。相似度函数均为编辑距离函数“EditSim”。

1. 按**步骤一：选择本体**为待创建的图谱选择本体。
2. 按**配置数据源>示例**为待创建的图谱配置数据源。
3. 按**配置信息抽取>示例**为待创建的图谱配置信息抽取。
4. 按**知识映射>示例**为待创建的图谱配置知识映射。
5. 在“知识融合”对话框，打开“是否融合”开关。
6. 在实体“电影”对应的“融合标识符”栏单击 $+$ ，在弹出的文本框中单击 ∇ ，选择标识知识融合的本体属性“name”。
7. 单击“添加分组”，填写判断融合的属性配置信息，如**图9-2**所示。

图 9-2 知识融合

电影

融合标识符 $\textcircled{+}$ name \times $+$

属性配置

添加分组

新实体属性	待融合实体属性	相似度函数	相似度阈值	操作
name	name	EditSim	<div><div>-</div><div>1.00</div><div>+</div></div>	<div>删除</div>
上映时间	上映时间	EditSim	<div><div>-</div><div>1.00</div><div>+</div></div>	<div>删除</div>

+

 增加

您还可以添加 8 行

人物

融合标识符 $\textcircled{+}$ name \times $+$

属性配置

添加分组

新实体属性	待融合实体属性	相似度函数	相似度阈值	操作
name	name	EditSim	<div><div>-</div><div>1.00</div><div>+</div></div>	<div>删除</div>
职业	职业	EditSim	<div><div>-</div><div>1.00</div><div>+</div></div>	<div>删除</div>

+

 增加

您还可以添加 8 行

8. 在实体“人物”对应的“融合标识符”栏单击 $+$ ，在弹出的文本框中单击 ∇ ，选择标识知识融合的本体属性“name”。
9. 单击“添加分组”，填写判断融合的属性配置信息，如**图9-2**所示。
10. 单击“保存”，回到“新建图谱”页面。
11. 在“新建图谱”页面，单击“运行”，完成图谱的创作。

格式化编辑步骤

1. 在“新建图谱”页面，单击“知识融合”，右侧弹出“知识融合”对话框。
2. 在“知识融合”对话框，

若不需要进行知识融合，则关闭“是否融合”开关，进入**步骤6**。

若需要进行知识融合，则打开“是否融合”开关，进入**步骤3**。
3. 在“知识融合”对话框的“编辑方式”中，选择“格式化编辑”。
4. 在“格式化编辑”页签的编辑框中，输入json格式代码。

说明

在“格式化编辑”页签编辑json代码配置知识融合时，可单击“交互界面”，切换至“交互界面”页签，查看知识融合判断属性及相似度函数参数。

示例

以一个电影实体为例，对属性“name”相似的数据进行知识融合判断，当属性“production_region”的编辑距离大于0.04，且属性“name_en”的Jaccard距离大于0.05时，数据进行融合。代码示例如下所示：

```
{
  "do_conflate": true,
  "block_keys": {
    "Person": [],
    "Film": [
      "name"
    ]
  },
  "blocker": "ngram",
  "similarity_functions": {
    "Person": [],
    "Film": [
      {
        "source_key": "production_region",
        "target_key": "production_region",
        "similarity_function": "EditSim",
        "threshold": 0.04
      },
      {
        "source_key": "name_en",
        "target_key": "name_en",
        "similarity_function": "JaccardSim",
        "threshold": 0.05
      }
    ]
  }
}
```

5. 单击“保存”，回到“新建图谱”页面。
6. 在“新建图谱”页面，单击“运行”，完成图谱的创建。

创建完新图谱后，您可以在“我的图谱”页面查看图谱的运行状态，图谱的状态包括“创建中”、“初始化中”、“可用”、“运行中”。

10 融合验证

若在创建图谱时，在[步骤五：配置知识融合](#)配置了知识融合，即在[步骤2](#)打开知识融合开关，并配置知识融合的信息，需要进行融合验证，即验证当前知识融合的配置产生的结果是否符合预期。

前提条件

已创建完知识图谱，即完成图谱创建，并在[步骤五：配置知识融合](#)[步骤2](#)打开知识融合开关配置知识融合的信息。

操作步骤

1. 登录[KG服务管理控制台](#)，单击右侧导航栏“我的图谱”，在“我的图谱”页面单击图谱名称，进入图谱详情页。
2. 单击“图谱版本”切换至“图谱版本”页签，单击“操作”列的“验证”，进入“版本验证”页面。
图谱版本状态处于“已上线”时，仅可以执行“操作”列的“查看”操作。
图谱版本状态处于“未上线”时，可以执行“操作”列的“验证”、“发布”、“查看”、“删除”操作。
3. 在“版本验证”页面，有随机采样的部分融合样例，若您认可融合的结果，单击 ☒，若不认可，单击 ☐。
右侧可查看融合验证的结果，包含“准确率”、“准确数”、“不准确数”、“不确定数”。

图 10-1 融合验证



11 管理图谱

11.1 图谱管理简介

针对已经创建的知识图谱，您可以预览不同结构形式的知识图谱，还可以基于已有图谱，通过全量和增量操作更新图谱，通过复制配置快速创建图谱。

图谱管理相关操作及说明如下：

- **预览图谱**：针对已创建的知识图谱，您可以通过搜索实体，预览不同结构形式的知识图谱。
- **全量更新图谱**：针对已经创建的知识图谱，您可以通过使用新的数据替换原有的基础数据，全量更新图谱。
- **增量更新图谱**：针对已经创建的知识图谱，您可以在原有图谱的数据基础上增加新的数据，增量更新图谱。
- **复制配置快速创建图谱**：针对已经创建的知识图谱，您可以通过复制配置，在原有图谱配置的基础上更改配置，快速创建新的图谱。

11.2 查看图谱详情

针对已创建的知识图谱，您可以查看图谱详情。

查看图谱详情

1. 登录**KG服务管理控制台**，单击“我的图谱”，进入“我的图谱”页面。
2. 在“我的图谱”页面，单击待查看的图谱名称，进入图谱详情页面，如**图11-1**所示。

图 11-1 图谱详情



表 11-1 图谱信息说明

名称	说明
“图谱预览”	单击可预览图谱，详情请参见 预览图谱 。
“流水线预览”	单击可预览该图谱各个创建流程的配置信息，包括本体选择、数据源配置、信息抽取、知识映射、知识融合。
“图谱统计数据”	包括该图谱的实体个数和关系个数。
“版本统计”	包括图谱版本的上线个数和未上线个数。
“更新日志”	更新日志列表，包括“运行编号”、“事件类型”、“本体”、“状态”、“操作时间”、“操作”。
“图谱版本”	<p>图谱版本列表，包括“版本名称”、“状态”、“实体个数”、“关系个数”、“更新时间”、“操作”。</p> <p>其中，操作包括“验证”、“发布”、“查看”、“删除”。</p> <ul style="list-style-type: none"> ● 验证：融合验证，详情请见步骤五：配置知识融合。 ● 发布：发布图谱版本。版本处于“已上线”状态时不能执行发布操作。详情请见发布图谱版本。 ● 查看：查看图谱创建流程的配置信息。 ● 删除：删除当前图谱版本。版本处于“已上线”状态时不能执行删除操作。

11.3 发布图谱版本

针对新创建的图谱、全量更新的图谱、增量更新的图谱、通过复制快速创建的图谱，您可以发布图谱版本。只有发布图谱版本后，才能查询图谱中的实体数据。

操作步骤如下所示：


1.

登录[KG服务管理控制台](#)，单击“我的图谱”，进入“我的图谱”页面。

2.

在“我的图谱”页面，单击待发布图谱的名称，进入图谱详情页面。

3.

单击“图谱版本”，切换至“图谱版本”页签。
您可以单击版本“描述”列的，编辑图谱版本名称。

4.

选择图谱版本，单击“操作”列的“发布”，弹出“发布确认”对话框。

5.

单击“发布”，发布图谱版本。
发布的版本“状态”变成“发布中”，等待“状态”变为“已上线”，表示图谱的版本发布成功。

11.4 预览图谱

针对已创建的知识图谱，您可以预览指定实体的不同结构形式知识图谱。

前提条件

- 已创建知识图谱，创建图谱流程请参见[创建图谱流程](#)。
- 已发布图谱版本，详情请参见[发布图谱版本](#)。

操作步骤如下：

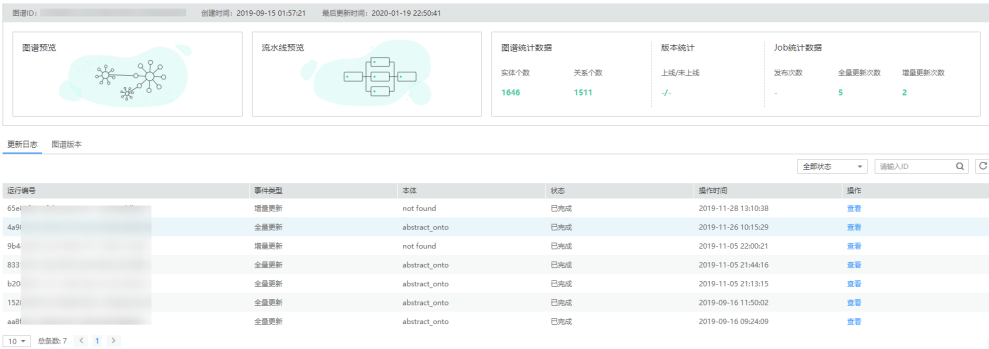
1.

登录[KG服务管理控制台](#)，单击“知识图谱”，进入“我的图谱”页面。

2.

在“我的图谱”页面，单击已创建图谱的名称，进入图谱详情页面，包含信息如[图11-2](#)所示。

图 11-2 图谱详情



3.

单击“图谱预览”，选择匹配方式：
 - 选择“完全匹配”，输入“实体名称”，按回车进入图谱预览页面，预览该实体的知识图谱。
 - 选择“模糊匹配”，输入实体名称部分字段，按回车进入图谱预览页面，预览实体名称含该字段的所有实体的知识图谱。

举例：输入实体名称“张艺谋”，按回车进入图谱预览页面，预览实体“张艺谋”的知识图谱。

4. 右上角功能区域可放大/缩小图谱、导出图谱图片、切换图谱布局。

11.5 全量更新图谱

针对已经创建的知识图谱，您可以全量更新图谱，即使用新的数据源更新知识图谱。

前提条件

在全量更新知识图谱之前，您需要创建用于组织新图谱信息结构的本体，方式如下：

- 方式一：[控制台创建本体](#)
- 方式二：[OBS导入本体](#)

进入全量更新页面

1. 登录[KG服务管理控制台](#)，单击“知识图谱”，进入“我的图谱”页面。
2. 在“我的图谱”页面，单击“全量更新”，进入“全量更新”页面，依次完成[表 11-2](#)中的配置操作。

说明

“全量更新”页面入口与“创建知识图谱”页面入口不同，但是全量更新中的配置操作均与创建知识图谱操作相同。

表 11-2 全量更新图谱操作

操作顺序	操作名称	操作指引
1	选择本体	步骤一：选择本体
2	更新基础数据	步骤二：配置数据源
3	更新信息抽取	步骤三：配置信息抽取
4	更新知识映射	步骤四：配置知识映射
5	更新知识融合	步骤五：配置知识融合
6	融合验证	融合验证

11.6 增量更新图谱

针对已经创建的知识图谱，您可以增量更新图谱，即在原有图谱的数据基础上增加新的数据更新知识图谱。

进入增量更新页面

1.

登录[KG服务管理控制台](#)，单击“知识图谱”，进入“我的图谱”页面。
2.

在“我的图谱”页面，单击“增量更新”，进入“增量更新”页面，依次完成[表11-3](#)中的配置操作。

说明

“增量更新”页面入口与“创建知识图谱”页面入口不同，但是增量更新中的配置操作均与创建知识图谱操作相同。

表 11-3 增量更新图谱操作

操作顺序	操作名称	操作指引
1	选择本体	无需配置
2	增加基础数据	步骤二：配置数据源
3	更新信息抽取	步骤三：配置信息抽取
4	更新知识映射	步骤四：配置知识映射
5	更新知识融合	步骤五：配置知识融合
6	融合验证	融合验证

11.7 复制配置快速创建图谱

当您希望创建与已有图谱类似的新图谱时，您可以复制原有知识图谱的相关配置，快速创建新的图谱。

复制配置操作指在已有图谱配置的基础上更改配置信息，重新创建一个知识图谱。

前提条件

在复制配置快速创建新图谱之前，您需要提前创建新的本体，创建方式如下：

- 方式一：在知识图谱中创建本体，具体步骤见[创建本体](#)。
- 方式二：将待导入的本体（json格式）上传至OBS中，详情请参见[OBS添加桶](#)和[OBS上传文件或文件夹](#)。

进入复制配置页面

1.

登录[KG服务管理控制台](#)，单击“知识图谱”，进入“我的图谱”页面。
2.


在“我的图谱”页面，单击图谱卡片中的，进入“复制图谱”页面，依次完成[表11-4](#)中的配置操作。

图 11-3 复制图谱



说明

复制配置页面入口与创建知识图谱页面入口不同，但是复制配置中的配置操作均与创建知识图谱操作相同。

表 11-4 复制配置图谱操作

操作顺序	操作名称	操作指引
1	选择本体	步骤一：选择本体
2	配置基础数据	步骤二：配置数据源
3	信息抽取	步骤三：配置信息抽取
4	知识映射	步骤四：配置知识映射
5	知识融合	步骤五：配置知识融合
6	融合验证	融合验证