

《智能信息处理》课程考试

人力资源本体模型的构建

曹宇航

考核	到课[10]	作业[20]	考试[70]	课程成绩[100]
得分				

2021 年 12 月 10 日

人力资源本体模型的构建

曹宇航

(大连海事大学 信息科学技术学院, 大连 116026)

摘 要 本体是概念的集合及概念间关系的集合的集合, 实际上就是对特定领域之中某套概念及其相互之间关系的形式化表达。知识图谱是结构化的语义知识库, 用于以符号形式描述物理世界中的概念及其相互关系。人力资源被视为最宝贵的战略资源, 已成为企业的核心竞争力。本文采用自顶向下的方法构建人力资源知识图谱的本体模型。首先建立人力资源知识图谱的本体, 对实体类和实体关系进行建模, 然后构建起基本的知识图谱。本文利用 OWL 语言构建了人力资源领域知识的相关本体, 并实现了人力资源知识图模型的建立。

关键词 本体; 知识图谱; 人力资源;

Human Resource Ontology Model for Knowledge Graph

Cao Yuhang

(School of Information Science and Technology, Dalian Maritime University, Dalian 116026)

Abstract Ontology is the set of concepts and the set of relations between concepts, in fact, it is the formal expression of a set of concepts and their relations in a specific domain. Knowledge graph is a structured semantic knowledge base, which is used to describe the concepts and their relationships in the physical world in the form of symbols. Human resources are regarded as the most valuable strategic resources, and have become the core competitiveness of enterprises. In this paper, the top-down method is used to construct the ontology model of human resource knowledge map. Firstly, the ontology of human resource knowledge map is established, the entity classes and entity relations are modeled, and then the basic knowledge map is constructed. In this paper, Owl language is used to construct the relevant ontology of human resource domain knowledge, and realize the establishment of human resource knowledge map model.

Key words Ontology; knowledge graph; human resources

1 引言

当今世界, 人是人力资源管理的核心, 如何合理地描述人已成为人力资源业务的基础。虽然现如今已经存在许多种类型的人事信息系统来对人员进行科学管理, 但它们的标准和模型是多种多样的, 人员与周围环境以及其他人员之间存在着大量的关系, 人力资源部门仍然在努力试图用更加高效

的方式掌握人员关系信息。因此, 为了实现人力资源管理的科学化和高效性, 我们可以利用知识图谱技术, 有效整合现有的人力资源管理知识, 构建人力资源管理知识支持系统, 并实现人力资源管理的智能决策。

近年来, 知识组织、管理和表示研究一直是一个重要研究领域, 已经取得了长足的进步, 具有为人力资源提供技术支持的能力。同时, 本体作为知识组织和表示的主要方法, 已经形成了较为成熟的

理论和技术体系,并且广泛应用于知识管理、知识组织、知识集成、知识表示等领域。本体可以规范某一领域内统一的概念和术语,促进领域知识的集成和共享,消除不同背景造成的理解冲突和混淆,为实现基于本体的知识查询和推理奠定基础。基于此,本文拟采用本体技术对人力资源管理领域的文本进行建模。不仅能够更科学地存储人力资源管理领域的知识,提高人力资源管理领域知识共享和利用的效率,也可以为人力资源管理决策提供有力的参考,提高人力资源管理的响应速度和决策效率,促进人力资源管理的智能化和精密化。

2 本体

2.1 本体的定义

本体概念被引入人工智能、知识工程等领域后被赋予了新的含义。不同的专家学者对本体的理解不同,所给出的定义也有所差异。

人工智能领域的学者 Neches(1991)等人对本体进行定义如下:

定义1. 本体是构成相关领域词汇的基本术语和关系,以及利用这些术语和关系构成的规定这些词汇外延的规则的定义^[1]。

Neches 是最早对本体定义进行研究的学者,从内容的角度给出了本体定义,概括出了本体的基本要素,包括领域术语、关系和规则。这为其后各领域学者对本体的定义研究提供了参考借鉴。

美国斯坦福大学 Gruber(1993)所给出的本体定义如下:

定义2. 本体是概念化的规范说明^[2]。

Gruber 给出的本体定义最为经典,但是未能全面概括出本体的本质。随后, Borst 等人(1997)对 Gruber 给出的定义进行补充如下:

定义3. 本体是共享概念模型的形式化规范说明^[3]。

Borst 提出了本体共享的概念,阐明了本体的共享本质,但没有说明概念与概念之间的关系。

德国学者 Studer 等人(1998)又对 Borst 的定义进行了扩展,提出了概念关系之间的“明确”定义,即:

定义4. 本体是共享概念模型的明确的形式化规范说明^[4]。

Studer 给出的本体定义被各领域专家学者高

度认可,其涵盖了本体的基本特征:共享、明确、概念化、形式化,被学术界广泛引用,对于后来的本体研究具有重要意义。

张秀兰教授通过对国内外各领域本体定义的深入研究,总结出的本体定义如下:

定义5. 本体是通过描述、捕获领域知识,确定领域内共同认可的概念和概念间的关系,以用于领域内的不同主体之间交流与知识共享的形式化规范说明^[5]。

2.2 本体描述语言

本体作为概念及其相互关系的形式化表达,需要一种标准化的语言来表达或描述它们。目前,通常用于描述本体的标准化语言主要有两种:一种是基于谓词逻辑的本体描述语言,另一种是基于 Web 语义 Web 的本体描述语言。基于谓词逻辑的本体描述语言在本体的形式化描述方面往往存在局限性,因为特定的概念及其关系难以以谓词逻辑的形式表达。因此,目前广泛使用的本体描述语言是基于 Web 语义 Web 的本体描述语言。

Web 语义网最早由 Berners Lee 提出,他希望通过构建语义网来实现与机器的沟通和对话。目前广泛认可的语义 Web 定义来自 W3C:“语义 Web 提供了一个公共框架,允许在应用程序、企业和社区之间共享和重用数据。” W3C 还提供了三种用于语义分析的技术。其中资源描述框架(resource description framework, RDF)是最常用的符号语义表示模型。

资源描述框架模式(resource description framework Schema, RDFS): RDF 为描述客观世界的事实提供了一个基本框架,但它缺乏模式层的定义方法,如类和属性。RDFS 用于定义术语集、类集和属性集,主要包括以下原语:class(类)、SubClassOf(子类)、Type、property(属性)、SubPropertyOf(子属性)、Domain(定义)、Range(值范围)等。基于这些简单的表达式组件,可以构造最基本的类层次结构和属性系统。

OWL(Web Ontology Language): OWL 可以被视为是 RDFS 的一个扩展,主要在 RDFS 的基础上扩展了表达类和属性约束的能力,允许 OWL 构建更复杂、更完整的本体。这些扩展的本体表达功能包括:复杂类、属性限制、基数限制、属性特征等。

与 RDF 和 RDFS 相比, OWL 可以更清晰地描

述资源及其关系,具有更强大的内容表达能力。它以描述逻辑为主要理论基础,在医疗保健、金融、人力资源管理等领域的知识图构建中具有实际应用价值。基于 OWL 在内容性能方面的强大优势,本文使用 OWL 对人力资源知识图谱进行建模。

2.3 本体的构建准则

研究人员从实践出发提出了许多指导本体构建的原则。然而,目前仍没有构造本体的统一标准,一般采用 1995 年 T. R. Gruber 提出的指导本体构造的五条原则^[6],具体如下:

1) 清晰:领域本体必须能有效地说明所定义术语的含义。定义应该是客观的,与背景独立的;当定义可以用逻辑公理表达时,它应是形式化的,应尽力用逻辑公理表达;定义应该尽可能的完整;所有定义应该用自然语言加以详细说明。

2) 一致:领域本体应该是前后一致的,也就是说,它应该支持与其定义相一致的推理。领域本体所定义的公理以及用自然语言进行说明的文档都应该具有一致性。假如从组公理中推导出来的一个句子与一个非形式化的定义或者实例矛盾,则这个领域本体是不一致的。

3) 可扩展性:领域本体的可扩展性是指其提供一个共享的词汇,这个共享的词汇应该为预期的任务提供概念基础。它应该可以支持在已有的概念基础上定义新的术语,以满足特殊的需求,而无须修改已有的概念定义。

4) 编码偏好程度最小:领域本体与特定的符号即编码无关。也就是说,领域本体的表示形式的选择不应该只考虑表示或实现上的方便,概念的描述不应该依赖于某一种特殊的符号层的表示方法,不能依赖于某种确定的语言,因为实际的系统可能采用不同的知识表示方法。

5) 本体约定最小:本体约定应该最小,只要能够满足特定的知识共享需求即可。也就是说,本体应该对所模拟的事物产生尽可能少的推断,而让共享者自由地按照他们的需要去专门化和实例化这个本体。

上述五条原则给出了构造领域本体的基本思路和框架,但是明显的不足之处就是它们所反映的内容较模糊且难于把握。实际本体构建过程中,以上五原则甚至可能有不一致的情况,所以需要权衡各原则,必要时可能还要参照其他原则,需要灵活

运用本体构建原则才能构建高质量的本体。

3 人力资源本体模型的构建

3.1 人力资源知识图谱的实体建模

针对人力资源数据的特殊性,将人力资源信息系统中的数据作为人力资源的主要数据来源进行分类和归纳,并结合公司的日常需要进行人员检查、岗位调整等业务需求,考虑现有的人力资源数据分为三类:个人、部门和事件。根据要求,每个类别可分为若干子类别。

这些主题中的每一个都可以抽象为 OWL 中的一个类或附加到该类的一个子类。此外,对于这些类中的公共属性,需要定义一个人力资源数据库类对象类,并将一些公共和公共属性抽象到人力资源数据库类中。人力资源知识图谱数据类别及其子类别之间的关系如图 1 所示。

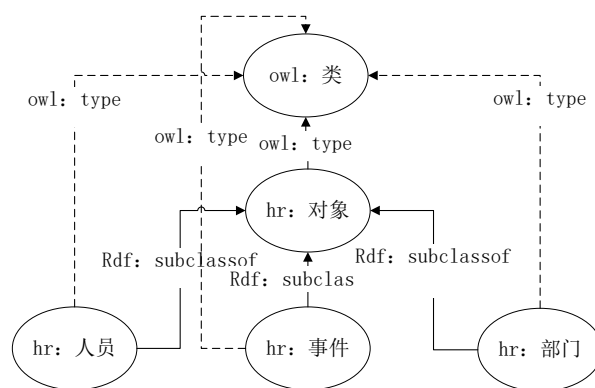


图 1 人力资源知识图谱数据类

在图 1 中, owl 是网络本体语言的标准空间前缀, rdf 和 rdfs 分别是 rdf 资源描述框架和 rdf 模式的标准空间前缀。“hr”是人力资源名称空间的前缀。

下面对图中的数据类别及其子类别进行说明:

1) hr: 对象选择

人力资源知识图的基类。主要用于定义其他类的公共属性。

2) hr: 人员类别

对于人力资源管理业务来说,人是一切人力资源管理业务发展的核心和重点。如图 2 所示,如何界定人员的实体类别将决定人力资源知识图谱从哪个角度协助人力资源部门进行决策。

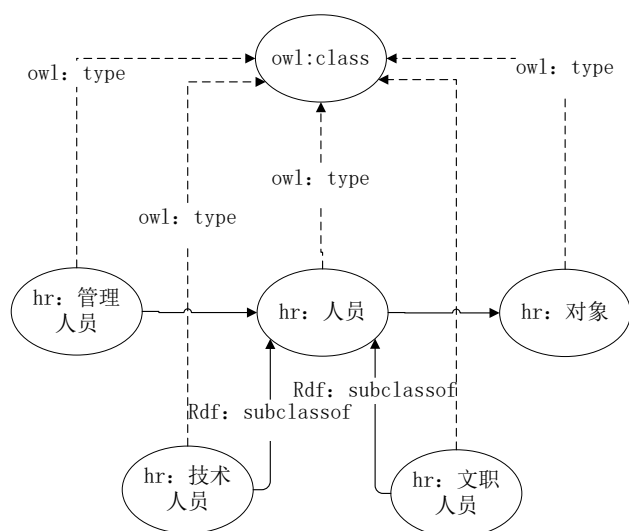


图2 人力资源知识图谱的人员类

以管理人员为例，使用 OWL 语言实现人员类和管理人员子类的管理：

```
<rdf:RDF
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  <owl:Ontology rdf:about="xml:base"/> //OWL Version information
  <owl:Class rdf:ID="Person">
    <rdfs:comment>Person Class</rdfs:comment>
  </owl:Class> // Definition of personnel class
  <owl:DatatypeProperty rdf:ID="Identity Category">
    <rdfs:domain rdf:resource="#Person"/>
    <rdfs:range rdf:resource="&xsd;nonNegativeInteger">
      </owl:Class> // Definition of "identity category" for public attribute of personnel class
    <owl:Class rdf:ID="Administrative Staff">
      <rdfs:comment>Management Person Class</rdfs:comment>
      <owl:intersectionOf rdf:parseType="Collection">
        <rdfs:subClassOf rdf:resource="#Person"/>
        <owl:Restriction>
          <owl:onProperty rdf:resource="#Identity category"/>
          <owl:allValuesFrom rdf:resource="#Management"/>
        </owl:Class> //The definition of "manager" subclass and its relationship with personnel class
```

3) hr: 部门

单位类别。作为人力资源知识图谱的数据源，对不在单位内部工作的人员进行研究是毫无意义的。因此，为了便于各部门人员的管理，有必要建立单位类别，并将人员类别中的每个实体与单位类别中的实体联系起来。一是可以保证人力资源信息的完整性，二是也是人力资源部门进行岗位调整的业务需要。

4) hr: 事件

事件类别。事件类别主要用于记录本单位各部门及下属相关的业绩和事件。通过对事件类别及其各个实体的定义，可以作为人员任用调整、人才选拔和培训的重要依据。

3.2 人力资源知识图谱的关系建模

在上述分析之后，通过 OWL 对人力资源知识图中涉及的实体类和子类进行了描述，并建立了模式图。但是对于知识图谱来说，仅仅实体和实体属性是不够的，关系设计作为另一个重要的描述尚未实现。这是通过关系建模，利用 OWL 来完成人力资源知识图的关系建模，并给出关系模式图。

在人力资源知识图谱中，关系被定义为实体之间的相关连接。这里的实体可以属于同一本体类别，也可以不属于同一本体类别。对于不同类之间的实体关系，经过分析，如图3所示。

给出下列相关定义：

1) 教育经历

人员与单位之间的实体关系。当记录一个人在某个单位有教育或培训经验时，有必要在人力资源知识图中建立这两个实体之间的“经验”关系。

2) 职位

人员与单位之间的实体关系。当记录一个人在某个单位有教育或培训经验时，有必要在人力资源知识图中建立这两个实体之间的“经验”关系。

3) 过往职位

上一岗位人员与单位的实体关系。当记录某人曾在某个单位服务时，有必要在人力资源知识图中建立这两个实体之间的“持有”关系。

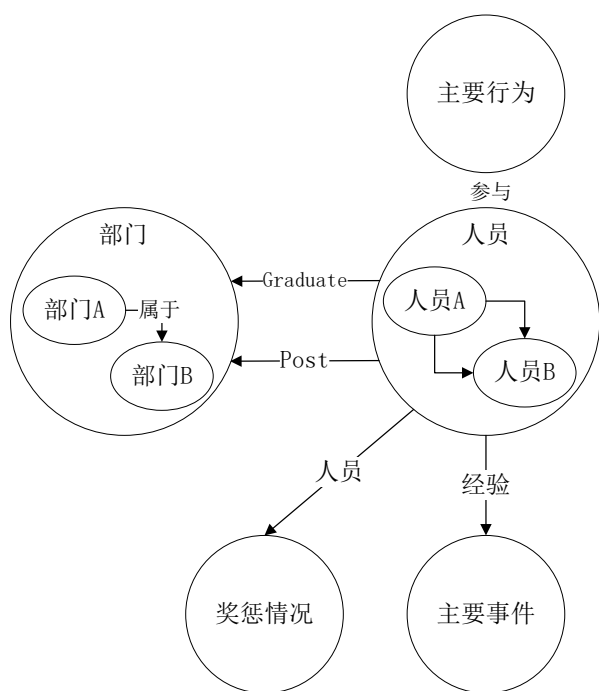


图3 人力资源知识图谱的关系模型

4) 事件关系

它是人员类和事件类之间的实体关系。对于不同的子类别，事件关系也不同，如下所示：

① 参与

参与是人员类别和事件类别下的行动子类别之间的关系。当记录某人参与重大事件和活动时，需要在人力资源知识图中建立这两个实体之间的“参与”关系；

② 经验

人员类别和事件类别下的事件子类别之间的关系。当记录一个人经历了一个重要事件时，需要在人力资源知识图中建立这两个实体的参与“经验”关系；

③ 奖惩情况

人员类别和事件类别中奖惩信息子类别之间的关系。当记录某个人受到某种奖励或惩罚时，有必要在人力资源知识图中建立这两个实体之间的“获取”关系。

4 总结

本文主要构建了人力资源知识图的本体模型。首先介绍了本体建模常用的本体和语义网技术。在分析了常用的三种知识建模语言的基础上，选用OWL构建本体，并给出了人员、单元和事件的详细描述。通过建立三种OWL模型，并以模式图的形式表示三种类型的实体，最后建立了OWL模型和人力资源知识图的实体间关系模式图。

然而，构建的人力资源本体模型还需要进一步改进和完善。虽然本文通过多种渠道收集人力资源管理领域的知识，初步构建了人力资源本体模型，但是一个精确而完整的本体并不是一次性发展起来的，而是一个螺旋式发展的过程。通过反复评价和不断修改，才能最终形成符合开发目标和要求的本体。

参 考 文 献

- [1] Neches R, Fikes R E, Gruber T, et al. Enabling Technology for Knowledge Sharing [J]. AI Magazine, 1991, 12(3): 36-56
- [2] Gruber T. A Translation Approach to Portable Ontology Specifications [J]. Knowledge Acquisition, 1993(
- [3] Borst W N. Construction of Engineering Ontologies for Knowledge Sharing and Reuse [D]. Ph. D thesis, 1997
- [4] Studer B, Benjamins V R, Fensel D. Knowledge Engineering: Principles and Methods [J]. Data and Knowledge Engineering, 1998, 25(1/2): 161-197.
- [5] 汤艳莉, 赖茂生. Ontology 在自然语言检索 中的应用研究 [J]. 现代图书情报技术, 2005(2): 33-36, 5
- [6] T. R.Gruber. Towards Principles for the Design of Ontologies Used for Knowledge Sharing[J]. International Journal of Human Computer Studies, 1995
- [7] 臧根林,王亚强,吴庆蓉,占春丽,李熠.智慧城市 知识 图 谱 模 型 与 本 体 构 建 方 法 [J]. 大 数 据,2020,6(02):96-106
- [8] 金芝, 知识工程中的本体论研究. 世纪之交的知识工程与知识科学. 清华大学出版社 2001: 468-477