

《智能信息处理》课程考试

## 基于商品音乐领域本体的信息检索模型研究

王树声

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

2021 年 12 月 16 日

# 基于商品音乐领域本体的信息检索模型研究

王树声

(大连海事大学 理学院, 辽宁省大连市 中国 116026)

**摘 要** 商品音乐是我们通常所说的“流行音乐”，它是商业性的音乐消遣娱乐，深受现代广大年轻人的喜爱。由于互联网的快速发展和网络资源量急速飞升，当人们使用传统信息检索方式搜索音乐的时候，往往不能得到满意的结果。同时音乐领域内复杂程度高、专业性强，传统的信息检索方式无法清晰辨析包含的深层语义。而语义信息检索是信息检索领域的新技术，它以传统信息检索技术为基础，结合领域本体、数据挖掘和自然语言来实现高效搜索。本体不仅可以基于语义层面上搭建检索过程中所使用到的模型结构，而且还支持推理。因此，利用本体的优点来实现基于语义层面上的查询，将有助于这些系统性能的提升。本文通过对商品音乐领域相关资料的收集与研究，搭建了商品音乐领域本体；针对不同的查询请求类型，给出了对应的语义查询算法；还提出了一种新的方法用于计算概念间相似度大小，综合考虑了传统方法中所提及到的影响因素以及本体层次结构的特点。

**关键词** 商品音乐；语义检索；本体；语义相似度

## Research on information retrieval model based on commercial music domain ontology

Wang Shusheng

(Department of Mathematics, Dalian Maritime University, Dalian, 116026)

**Abstract** Commercial music is what we usually call "pop music". It is a commercial music entertainment, which is deeply loved by the majority of modern young people. Due to the rapid development of the Internet and the rapid rise of network resources, when people use the traditional information retrieval method to search music, they often can not get satisfactory results. At the same time, the music field is highly complex and professional, and the traditional information retrieval methods can not clearly distinguish the deep semantics contained. Semantic information retrieval is a new technology in the field of information retrieval. It is based on traditional information retrieval technology and combines domain ontology, data mining and natural language to achieve efficient search. Ontology can not only build the model structure used in the retrieval process based on the semantic level, but also support reasoning. Therefore, using the advantages of ontology to realize query based on semantic level will help to improve the performance of these systems. Through the collection and research of relevant data in the field of commercial music, this paper constructs the ontology of commercial music field; According to different query request types, the corresponding semantic query algorithms are given; A new method is also proposed to calculate the similarity between concepts, which comprehensively considers the influencing factors mentioned in the traditional methods and the characteristics of ontology hierarchy.

**Key words** Commercial music, Semantic retrieval, Ontology, Semantic similarity

## 1 引言

21世纪现阶段人类已经进入信息化社会,且正在向知识化社会转变。大量的计算机应用需要学科专业知识的支持,音乐知识系统也不例外(音乐知识查询和音乐智能教学)。即使在一般的自然语言理解中,音乐学知识也是必需的,因为音乐是一般文学作品和新闻报道等文章中经常涉及的内容。

“音乐是凭借声波振动而存在、在时间中展现、通过人类的听觉器官而引起各种情绪反应和情感体验的艺术门类<sup>[1]</sup>。”首先,音乐学科具有社会科学的一般特点,同时,又由于音乐离不开声音,声音是有物质基础的,从而音乐学科与物理、数学甚至医学都有天然而广泛的联系。总之,在计算机领域,音乐学科知识的表示是十分必要而且具有相当难度的。

根据最新的统计结果可以看出<sup>[2]</sup>,我国的网络普及率随着经济的高速的发展正在飞速上涨。但是目前网络上比如百度、网易云等大部分的信息查询系统采用的技术都说基于关键词匹配的,查询结果不总是理想的。而随着商品音乐行业中计算机与互联网技术的大规模使用,各类信息的高度重合与复杂繁乱,用户搜索音乐也经常不尽人意,所以需要提升查询音乐信息资源的高效性和准确性。

为了解决上述问题,引入了语义检索技术。而语义网为语义信息检索技术的实现提供可能性<sup>[3]</sup>,语义网不仅在很大程度上克服了互联网信息交互中语义异构问题,还对领域知识的表示提出了标准化规范。本体作为其中的核心内容,它为整个语义网提供了知识表达的领域模型,本体机制是实现语义理解和交流的关键,是解决语义层次上网络信息共享和交换的基础。在信息检索中加入本体相关技术,将简单的语法匹配提升至深层语义匹配<sup>[4]</sup>,使最后的查询结果信息能更好地符合用户的需求。

## 2 研究背景

### 2.1 语义网的相关概念以及层次结构

#### 2.1.1 语义网

语义网的概念是由万维网联盟的蒂姆·伯纳斯-李(Tim Berners-Lee)在1998年提出的一个概念,实际上是基于很多现有技术的,也依赖于后来和

text-and-markup与知识表现的综合。其渊源甚至可以追溯到20世纪60年代末期的Collins、Quillian、Loftus等人的研究,还有之后70年代初Simon、Schank、Minsky等人陆续提出的一些理论上的成果。其中Simon在进行自然语言理解的应用研究时提出了语义网络(Semantic Network)的概念。当时人们甚至发明了以逻辑为基础的程序设计语言Prolog。

语义网就是能够根据语义进行判断的智能网络,实现人与电脑之间的无障碍沟通。它的目的在于将网络上的文本信息转化为能够被机器识别和处理的语义信息,主要包括文本信息的数据部分以及事物之间的关系。和万维网相比,语义网更容易理解<sup>[5]</sup>,如表1所示。

表1 万维网与语义网的对比

特性	万维网	语义网
基本组件	非结构化的内容	形式化的陈述
主要用户	人类用户	应用程序
链接	标识位置	标识位置和含义
主要词汇	格式编排指令	语义和逻辑
逻辑	非形式化/非标准	描述逻辑

语义网的目标并非取代万维网,而是通过易于使用的、标准的语义来拓展万维网,以结构化的文本数据为中心,通过推理功能的引入,有助于更好地理解文本数据中所包含的语义信息,使得机器对文本实现更好地理解,语义网共包含7层的层次结构<sup>[5]</sup>,具体的层次结构如图1所示。

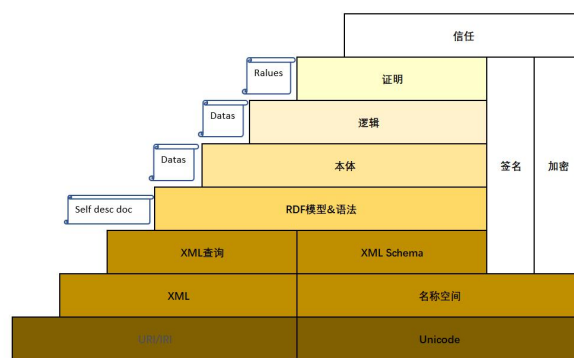


图1 语义网层次结构图

(1) 基于Unicode和URI/IRI的基础层,语义网中通过URI来对网络上的资源进行标识,并且每个资源的标识唯一且不重复。Unicode用于对资源进行编码,世界范围内的语言均支持该项标准,是语义网支持多元化元素的根本。

(2) XML+NS+XML Schema 语法层,用于描述资源的内容与结构。在结构方面XML是文档构建的一个重要支体。从内容方面URI是发现与鉴别实体的标

识, 构成了名称空间(NS)。XML文档通过XML Schema在语言范围内开启管控。

(3)RDF模型&语法 数据层,按照已有的规则对各类资源以及其对应的类型进行描述。其中,RDF支持简单的语义描述,RDF Schema则是对RDF的优化与完善,提供了定义类和属性的方法,并且实现了类和属性的双重约束,提供了对约束的检测方法。

(4)Ontology vocabulary 本体层,根据抽象化的表述方法准确地表达出特定领域内的各类信息及其内部联系。

(5)Logic 逻辑层,为了实现推理的功能提供了强大的逻辑语言。

(6)Proof 证明层,对语言的专业性进行分析并检验程序,根据实际场景开展工作。

(7)Trust 信任层,为了提供安全而又稳妥的信息传递,信任层形成了信任机制用以实现此目的。

### 2.1.2 XML

W3C的可扩展标记语言(*e Xtensible Markup Language, XML*)于1998年2月被发布,XML语言主要用于信息交换,实用性较高,被广泛应用于很多领域。XML语言与市面上所提供的众多解析器相互兼容。此外,XML语言还有效地解决了在共享的过程中所出现的各类语法问题。XML语言虽然有很多优点,但是XML语言去无法解决共享过程中的数据层次问题。

### 2.1.3 RDF

由W3C发布资源描述框架(*Resource Description Framework, RDF*),主要用途是为语义传达内容上搭建相互沟通互相联系的桥梁,语义信息在语义网传输的重要元素。RDF的内容主要包括资源(*Resource*)、属性(*Property*)和陈述(*Statement*)。

(1)资源:资源指的是通过商业音乐领域内的文本信息所抽取出的概念信息,它可以是音乐、音乐风格、人物团体、专辑等。通过资源标识符(*URI*)进行描述,例如网页、网页中的一个元素等都可以作为URI使用。

(2)属性:属性指的是资源间的各类关系,比如“音乐歌手”、“音乐格式”等。和资源一样,属性也需要URI来进行标记。

(3)陈述:由资源和属性构成,使用三元组(*Triple*)的形式来规划化元数据。

### 2.1.4 OWL

OWL:从专业的理论知识看,网络本体语言(*Web Ontology Language, OWL*)是由W3C所提出的

一种对本体进行描述的语言规范,以DAML和OIL语言为基础,是对RDFS的扩展,具备更丰富的语义表达。同时,在网络资源的描述中,它不仅具备更强的表达能力和逻辑推理能力,而且其语言特性更便于计算机的理解和处理,比XML、RDF、RDFS等描述语言的功能更加强大和完善。OWL语言可以分为

OWL Lite、OWL DL和OWL Full三类。

(1)OWL Lite:表达能力最弱,语法结构简单。

(2)OWL DL:其约束能力和表达能力介于OWL Lite和OWL Full之间,具有逻辑推理功能,表达能力丰富。

(3)OWL Full:具有最丰富的表达能力,是RDFS语言的扩展。由于OWL DL和OWL Lite中的约束已经取消,用户对领域知识的表达达到了最大,但无法保证推理能力和可计算性。

## 2.2 本体基本理论

### 2.2.1 本体

最初,本体(*Ontology*)的概念最先出现于哲学之中,后来,本体被引入计算机领域,用来表达人们对于领域的共同理解。目前,关于本体的定义说法很多,目前获得较多认同的是1998年R.Studer等在对前人定义进行深入研究基础上给出的解释:

“本体对概念体系的明确的形式化的可共享的规范说明”。本体的定义具有4个特征:概念化、形式化、规范化和共享。

目前本体的主要功能是提高信息检索的准确度,通过语义检索得到更为准确的查询结果,而不仅仅是收集到带有模糊关键字的网页信息。本体基于语义级别的检索不仅可以找到信息中所描述的语义内容,而且还可以通过推理系统发现查询信息中所包含的隐含信息,弥补了查询结果信息不足的问题。

### 2.2.2 本体的分类

根据形式化程度、应用领域,概念粒度的相对大小,以及对于领域依赖的程度可以得到多个关于本体的分类。

基于本体形式化程度可以分为高度非形式化的本体、结构化非形式化的本体、半形式化的本体和严格形式化的本体。按照领域依赖程度,本体可以分为顶层、领域、任务和应用本体4类;按照主题可分为知识本体、通用本体、领域本体、术语本

体和任务本体。目前在国内外已经出现了很多成型的知识本体,典型的有SUMO、WordNet、SENSUS、CYC和知网<sup>[6]</sup>。

### 2.1.3 本体的组成

在网络资源应用中,本体用于对各个专业领域的具体术语进行解析的作用,通过对知识的分类和具体关系来搭建概念模型。根据大量学者的调研,现今Perez的利用度较高,具体元素如下:概念(Concepts)、关系(Relations)、函数(Functions)、公理(Axioms)以及实例(Instances)<sup>[7]</sup>:

- (1)概念(Concepts):指的是将客观世界中具有相同特点的资源进行抽象化表示,抽象出资源间所具有的共同特征,具备分类层次关系,也被称作“类”。
- (2)关系(Relations):关系表示的是概念之间的关联程度。通过有效地定义和规范,不仅可以划分出不相关的资源,而且还可以体现出不同资源的性质。
- (3)函数(Functions):对多个个体、整体、概念的互相联系进行表示。
- (4)公理(Axioms):主要在限定事物本身的原则性、公正性、合理性上起裁判结果的作用。
- (5)实例(Instances):在类或概念的具体事物工作中开展,也就是现实世界中客观存在的个体。

## 3 本体的构建技术以及在商品音乐领域的应用

### 3.1 本体的构建技术

#### 3.1.1 本体的构建规则与方法

通过对本体技术的分析中可以得到选定领域的不同、项目类型的不同均会影响到本体构建,因此本体的构建是没有绝对的标准,需要将所研究的领域和自身的需求相结合进行分析。Gruber所提出的构建准则使用最为广泛,设计人员可以根据特定领域的具体现状,得到最适合的方法进行构建过程。在维护本体的稳定原则前提下,Gruber所提出的准则在不同角度下适当的提高了本体构建的效果。通过对本体构建的逐步探索,关于本体构建方法的研究成果不断涌现。由于研究者所研究领域和本体构建的不同条件,同时又没有规范化、程序化、科学化的系统方案构建本体,因此他们采用不同的方法来构建本体,每种方法都有自己的特点,可以根据具体的情况进行选择。其中比较具有影响力

的本体构建方式骨架法、METHONTOLOGY法、TOVE法、七步法、IDEF5法、循环获取法、五步循环法等。

在构建本体的过程中,本文采用的是七步法思想。七步法是由Noy和Mc Guinness所提出的一种本体构建方法。七步法的特点在于,它一共分为七步,每一步都对应着具体的工作内容,依次递进。该方法是由Protege本体建模工具配合实现的。其具体步骤包括:

- (1)界定研究领域和范围,明确领域相关知识;
- (2)研究借鉴现有本体的可行性;
- (3)总结研究领域内的重要术语或概念;
- (4)分析研究领域内的概念,定义对应的层次体系;
- (5)根据领域知识,创建类的数据属性和对象属性;
- (6)定义属性的定义域以及值域;
- (7)创建实例。

七步法的优点在于步骤规划相对合理,根据不同专业领域的知识与具体实践相结合,具体的构建流程如图2所示:

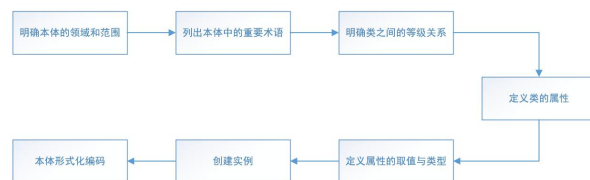


图2 七步法的构建流程

#### 3.1.2 本体的构建工具

在许多领域中都涉及到了本体,不同的研究领域对本体的构建过程有不同的要求和目标,因此,许多组织已经在各自的行业中开拓出需要的构建工具,让本体的开发使用更为广泛,例如: Apollo、OILED、Onto Edit、Protege等均是闻名世界的应用,这里主要针对Protege进行介绍与分析。Protege在实际的操作中可以把类和属性互相结合,其图形化界面操作更为简洁,用户直接通过概念来展开本体模型的构建而不需要建立对应的本体描述语言。Protege不仅可以对类、属性、类的继承关系等各类本体要素进行描述,还可以建立各类约束规则,例如:属性的定义域、值域、默认值等。

### 3.2 商品音乐领域本体以及推理规则的构建

#### 3.2.1 定义概念层次结构

本文以商品音乐作为顶层概念,采用自上向下的方法,对商品音乐领域进行了相应的研究,通过查阅音乐相关书籍与文献<sup>[9]</sup>,并且对qq音乐、网易

云等各类音乐软件进行分析后,结合音乐的相关理论知识,分成若干个二级类。如下图4所示。



图4 商品音乐本体层次分类图

通过查阅相关的文献对二级类进行进一步地细化,例如,人物团体划分为歌手、作词、作曲、编曲、混音、录音、母带,而歌手类又可以进一步细分为网络歌手、选秀歌手、创作型歌手、实力型歌手。本文主要考虑对用户影响最大的格式、MV、人物团体、音乐风格和心情主题几个二级类进行类分类。具体商品音乐领域本体的类分类层次图如图5-图9所示。



图5 格式类分类层次

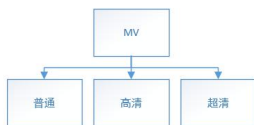


图6 MV类分类层次

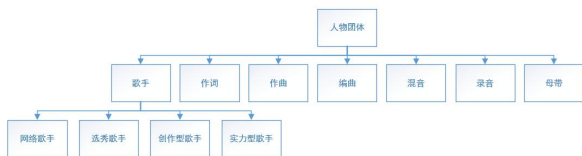


图7 人物团体类分类层次

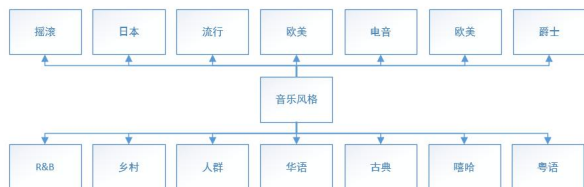


图8 音乐风格类分类层次



图9 心情主题类分类层次



图10 Protégé中的类层次关系面板

### 3.2.2 定义类的属性

类的层次结构仅仅是一个总体框架,不能对所研究领域中和信息进行充分的描述。本体中的属性切入事物内部对概念以及概念间的关系进行描述,因此在本体构建的过程中属性的定义及其约束的定义是非常重要的。属性具有内外之分,内部属性指的是类的一些独有特征,外部属性指的是两者之间的交互作用,表现为类之间的关联关系。由于此本体的概念比较复杂,图11仅列出类商品音乐领域本体中具有代表性的概念之间的逻辑关系图。其中歌手、作词、作曲又属于二级类人物团体

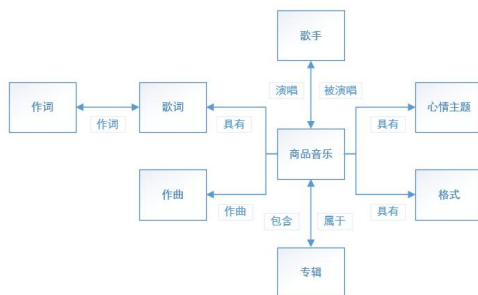


图11 商品音乐本体主要概念逻辑关系图





$$weight_1 = \begin{cases} 1, same \\ 1/2, kind\ of \\ 2/3, instance\ of \end{cases}$$

## (2) 节点深度

节点深度指的是从根节点出发到达指定节点间所经过的最短路径中所包含的边的数量。由于本体是一个从根开始逐层细化所有的概念信息得到的树形结构，叶子节点包含的信息量是最多最具体的。由同一节点衍生出的任意两个子节点，在最短路径相同的前提下，节点间的相似程度与其深度差成反比。节点深度权重具有以下关系：

$$weight_2 = (1/2^{dep(c)} + 1/2^{dep(c)-1} + \dots + 1/2)$$

其中， $weight_2$ 代表概念的节点深度权重值， $dep(c)$ 代表本体层中概念 $c$ 的深度，概念 $c$ 为概念 $c_1$ 和概念 $c_2$ 的最近公共父节点。

## (3) 节点密度

在本体的树形结构中，从上到下概念越来越复杂，抽象程度越来越低，除了考虑上述两个影响因子外还应考虑节点密度对于语义相似度的影响。在一棵本体树中，若对某一概念节点进行很大程度的细化的话就代表其拥有很多个子概念节点，有向边的密度给权重的值带来的影响为：

$$weight_3 = \frac{outDegree(c_1) + outDegree(c)}{2outDegree(root)}$$

其中， $outDegree()$ 表示某概念在本体中有向无环图的出度。 $outDegree(c)$ 表示概念 $c$ 的出度，概念 $c$ 为概念 $c_1$ 和概念 $c_2$ 的最近公共父节点。 $outDegree(root)$ 表示本体树根节点的出度。

通过上述分析，得到了一个综合关系路径权重，如下公式所示：

$$weight_{all} = \alpha * weight_1 + \beta * weight_2 + \lambda * weight_3$$

其中， $\alpha + \beta + \lambda = 1$ 。

语义距离指的是通过计算任意节点间在本体树中相互连通的边集合来判定其相似程度，是非常重要的一个影响因素。随着距离的增加或缩短，节点间的相似程度就会随之降低或者上升。若 $c_1, c_2$ 为本体树中的两个节点概念，二者的语义距离记作 $dist(c_1, c_2)$ ，由于语义距离与权重互为倒数的关系，如下式所示：

$$dist(c_1, c_2) = \frac{1}{weight_{all}(c_1, c_2)}$$

定义：概念 $c_1$ 与 $c_2$ 的语义相似度为 $sim_{dist}(c_1, c_2)$ ，则 $dist(c_1, c_2)$ 与 $sim_{dist}(c_1, c_2)$ 存在着如下对应关系：若 $dist(c_1, c_2)$ 越大， $sim_{dist}(c_1, c_2)$ 越小；反之则越大。所以，可调节语义相似度和语义距离的影响参数为 $\alpha$ ，如下式：

$$sim_{dist}(c_1, c_2) = \frac{\alpha}{dist(c_1, c_2) + \alpha}$$

其中 $\alpha$ 为调节因子。

(1)当 $sim_{dist}(c_1, c_2)=0$ 时，代表 $c_1, c_2$ 是两个完全不同的概念，即 $c_1, c_2$ 之间的语义距离趋于无穷大；

(2)当 $sim_{dist}(c_1, c_2)=1$ 时，代表 $c_1, c_2$ 是两个完全相同的概念，即 $c_1, c_2$ 之间的语义距离趋于0；

(3)除上述两种情况外，其余的语义相似度取值均介于0到1之间。此外，概念 $c_1, c_2$ 深度相同时，他们的最近公共父节点(LCA)必定会被两概念间的路径距离 $dist(c_1, c_2)$ 所经过。如下式：

$$dist(c_1, c_2) = dist(c_1, LCA) + dist(LCA, c_2)$$

## 4.2 基于信息量的语义相似度计算

节点间的信息含量的多少也会影响到二者相似程度的高低，二者成正比关系<sup>[11]</sup>。通过本体所构建的树形结构可以得到，本体结构中任意一个子节点都是其父节点的详细化，涵盖了其父节点的全部内容，因此，父节点的信息含量对其子节点间的相似度取值也是有一定影响的。对各个概念信息的量化由下列公式表示：

$$sim_{ic}(c_1, c_2) = \frac{2IC(c)}{IC(c_1) + IC(c_2)}$$

$$IC(c) = -\log \frac{Num(c)}{Num}$$

式中各项解析： $Num$ 代表本体分类树中所有概念的总数， $Num(c)$ 代表概念在本体树中出现的频率，概念 $c$ 为概念 $c_1$ 和概念 $c_2$ 的最近公共父节点。

## 4.3 基于属性的语义相似度计算

属性的不同，对概念的描述就会有所不同，因此概念所具有的属性集合可以作为该概念区别于其它概念的标志，节点间的语义相似程度与其公共属性的多少成正比。用如下公式表示：



$$sim_{dist}(c_1, c_2) = \frac{count(c_1) \cap count(c_2)}{count(c_1) \cup count(c_2)}$$

其中,  $count(c_1)$  代表概念  $c_1$  所拥有的属性个数,  $count(c_2)$  代表概念  $c_2$  所拥有的属性个数。

#### 4.4 语义相似度计算的改进方法

本文通过对传统的计算方法进行分析, 发现传统的方法忽略了对类的层次结构分析。任意两节点相连接的路径可能有多条, 每条路径上所经过的边的关系、节点所处位置以及节点附近的稠密程度都会影响到节点间的相似程度。因此, 本文综合考虑了上述的各个影响因素, 通过对每个影响因素所占比重的分析设置其对应的权重取值, 提出了一种新的方法, 用于计算概念间的相似度大小。如下所示:

$$sim(c_1, c_2) = \alpha * sim_{dist}(c_1, c_2) + \beta * sim_{ic}(c_1, c_2) + \lambda * sim_{attr}(c_1, c_2)$$

其中,  $\alpha$ 、 $\beta$ 、 $\lambda$  分别表示路径距离、信息含量、属性的调节系数,  $\alpha + \beta + \lambda = 1$ 。

### 5 结 语

本文对基于语义网和本体技术的商品音乐领域的逻辑概念进行分析, 在此基础上设计了语义本体模型, 实现了基于语义的检索方法, 为构建大众参与的个性化产品协同创意设计平台打下了基础。随着音乐领域信息的不断变化和增加, 会导致现有的知识结构的变化和新的知识结构的产生。因此为了保证检索的准确度, 需要修改现有知识结构并且将新的知识结构加入到已有的结构当中去, 对本体结构进行实时迭代更新以及如何通过创意设计知识库构建、产品设计要素特征提取与分类、知识迭代与融合及设计交互与智能整合, 形成知识与设计的协同演进机制、实现产品个性化定制与创意设计的高效对接是今后研究重点。

### 参 考 文 献

[1] 缪天瑞. 音乐百科词典[M]. 北京: 人民音乐出版社, 1998

[2] Jordanous A. Enhancing information retrieval and resource discovery from data using the Semantic Web[C]. International Symposium on Emerging Trends and Technologies in Libraries and Information Services. IEEE, 2015: 105-110

[3] Shilpa S. Laddha, Pradip M. Jawandhiya. Semantic tourism information retrieval interface[C]. International Conference on Advances in Computing, 2017: 694-697

[4] Chin-I Lee, Tse-Chih Hsia, Hsiang-Chih Hsu, et al. Ontology-based tourism recommendation system[C]. International Conference on Industrial Engineering and Application, 2017: 376-379

[5] John Hebel, Matthew Fisher, Ryan Blace, Andrew Perez-Lopez. Semantic Web Programming [M]. Indianapolis, Indiana: Wiley Publishing, 2009

[6] Dong Zhendong, Dong Qiang. 知网 —— How Net Knowledge Database[EB]. <http://www.keenage.com>.

[7] Wikimedia Foundation. Ontology (information science) [EB/OL]. <https://en.wikipedia.org/wiki/Ontology>, 2019-04-27

[8] Jena2-A Semantic Web Framework for Java [EB/OL]. <http://jena.sourceforge.net/index.html>, 2019-05-22.

[9] LAN Chunqiu, Li Ying. Research and construction of music domain ontology [J]. Computer CD software and application, 2014, 17 (03): 76 - 78

(兰春秋, 李樱. 音乐领域本体的研究与构建[J]. 计算机光盘软件与应用, 2014, 17(03): 76 - 78)

[10] Chandramouli A, Gauch S, Eno J. A cooperative approach to web crawler URL ordering[M]. Human-Computer Systems Interaction: Backgrounds and Applications 2. Springer Berlin Heidelberg, 2012: 343-357

[11] G. Zhu, C. A. Iglesias. Computing semantic similarity of concepts in knowledge graphs[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28: 1-14