

民航不安全事件语义词典构建及应用研究*

马慧¹, 赵捧未¹, 王洪俊², 孙辛博²

(1. 西安电子科技大学经济与管理学院, 西安 710126; 2. 北京拓尔思信息技术股份有限公司, 北京 100101)

摘要: 针对民航不安全事件领域缺乏统一知识模型, 从而导致信息上报质量不高的问题, 对不安全事件知识表达及相关应用进行了研究。采用本体知识表示方法, 根据本体的组成元素分析了不安全事件的知识来源, 进而基于本体建模“七步法”, 结合循环获取不断完善的思想构建了民航不安全事件语义词典, 并研究了不安全事件日志的语义标引, 进行了逻辑语义结构分析和完整度分析, 有利于提高民航不安全事件信息的上报质量。

关键词: 民航不安全事件; 语义词典; 本体

分类号: G254; TP391

DOI: 10.3772/j.issn.1673-2286.2015.09.005

1 引言

随着经济全球化的发展, 高效、便捷、经济的航空运输方式受到人们青睐, 同时飞机的安全问题也日益引起大众的关注。民用航空不安全事件, 是指与民用航空器运行相关的不安全事件, 包括民用航空器飞行事故、民用航空地面事故、民用航空器飞行事故征候及虽未构成飞行事故、事故征候, 但影响或可能影响飞行安全的事件^[1]。民航不安全事件产生的信息是安全管理活动所依赖的资源, 高质量的事件上报信息在安全管理系统中具有超前和预防意义, 是安全管理的生命线。因此要不断提高民航安全水平, 就必须充分依赖和利用不安全事件信息。

近年来, 我国民航安全管理水平的提高使得安全管理取得实质性的成效, 但是仍然存在一些问题和不足: (1) 民航不安全事件所涉及的知识繁多, 包括天气、飞机部件、事件类型、处理方式等, 而民航领域目前缺乏统一的知识模型, 造成个体的描述不同, 信息上报的数量和质量有待提高, 信息上报不全面, 存在一定的漏报、瞒报现象; (2) 该领域全信息的知识共享和交流比较困难; (3) 信息的利用程度比较低, 缺乏对不安全事件信息科学、客观地处理、分析。因此, 如何对民航不安全事件信息实现及时、高质量的上报并分析成为目前民航领域亟待解决的问题。

本文基于我国民航某单位飞行事件检索和智能挖掘系统的建设需求, 需要根据该单位已有的“不安全事件情况上报数据库”智能构建一个行业语义词典, 采用增量增加的方式逐步完善行业词典, 从而为其他智能分析技术所应用。本文采用本体建模“七步法”构建民航不安全事件语义词典, 并通过该词典实现对不安全事件日志的相关应用研究。

2 相关知识概述

2.1 知识表示

知识表示是知识的形式化和符号化的过程, 是指把知识客体中的因子与知识关联起来, 便于人们理解知识^[2]。随着知识交叉和复用的增加, 传统的知识表示形式, 如逐条表示的方法, 已经不能够满足人们对于信息的充分需求。当在网络上查找某一领域的相关知识时, 如果文章不包含用户给出的关键字, 即便文章属于该领域也无法被检索到。因此需要给出一种更加有效的知识表示方式。

本体 (ontology) 一词由17世纪的德国经院学者郭克兰纽 (Goclenius, 1547-1628) 首先使用。近几十年里, 这个词被应用到计算机界, 并在人工智能、计算机

语言以及数据库理论中扮演着越来越重要的作用。本体的定义较多,这里给出1998年Studer等人提出的定义:“本体是共享概念模型的明确的形式化规范说明”^[3]。采用本体的知识表示方式可以获取、描述民航不安全事件领域的知识,提供对本领域知识的共同理解,确定对该领域的共同认可的概念,清楚描述出概念之间的相关关系,从而起到计算机与人或者计算机之间没有交流障碍的桥梁作用^[4]。

在民航领域,张桂平^[5]在分析了航空术语的基本特点的基础上根据HowNet和KDML描述语言构建了面向航空领域的术语语义知识库,对民航信息处理有重要的作用;王红^[6]以民航突发事件应急预案、应急案例、应急资源和应急规则等应急知识为基础,构建了民航应急管理领域词典,为提高民航突发事件应急决策的智能性提供了一种新的思路;陈铭^[7]结合JASC航空器部件分类规则使用protégé工具建立航空事故成因本体的层次结构,用于描述事故部件之间的关系,以事故、事故原因、整改意见三层结构来描述一个事故及解决方法,对于适航审定和事故原因分析都有辅助作用。

本文针对民航知识的领域特性和知识结构的复杂性,选择了基于本体的知识表示方式来构建航空领域的语义词典。语义词典是处理语言信息的基础,应用于文献检索、自然语言查询、机器翻译等多领域,利用不同的语义概念和属性标记,结构化地存储和表达了文字之间上下位、近反义等多种语义关系。语义词典定义比较松散,具有语义功能的词典都可以称为语义词典。本文所构建的语义词典为民航不安全事件领域词典,通过该语义词典,我们可以将非结构化的文本数据格式化为易读写、易查询的结构化数据,也可以结合文本挖掘技术对不安全事件进行语义分析,方便了行业内各个组织对领域知识的共享、集成和扩展。

在阅读资料的过程中,笔者发现经常有语义词典和本体的概念混用的情况,笔者认为本体有着严格的界定,是一个领域的概念化描述,具有规范性、权威性,用于语义分析中的消歧和推理机制;而语义词典则定义松散,它们对事物的描述一般不太严格规范,不具有较高的权威性。语义词典描述的范围也比领域本体宽泛,包含了汉语中词级的语法和语义知识,用于小句分析中的功能成分分析,如英文语义词典WordNet和中英文常识库HowNet都是语义词典。然而本体和语义词典又是不可分开的,本体可以从语义词典中得到。复旦大学的陈壮坚^[8]认为具有严格关系且有领域权威性的,且支持推

理查询的语义词典就是本体。

2.2 知识来源

Perez等人^[9]提出本体由概念(Concepts)或类(Class)、关系(Relations)、函数(Functions)、公理(Axioms)、实例(Instances)五个基本元素(建模元语)构成。通过分析各元素之间的关系以及搜寻文献资料,可得到该语义词典的知识来源。

(1)从《BOEING英汉对照民用航空词典》、《民用航空不安全事件的处置程序》、新版《民用航空器事故征候》等相关民航领域的词典、文件中获取民航不安全事件及其他相关的概念知识。

(2)从相关领域的综合性叙词表中获得概念的分类标准和层次结构,即获取领域的知识关系,如可以从《航天科学技术叙词表》、《美国航空航天局叙词表》等获得民航领域的相关概念之间的关系。

(3)为了明确详细地说明某个需要表达的概念时,需要声明若干公理并对定义的概念术语进行约束和限制。领域专家启发式的先验经验知识中蕴含着概念的公理函数等本体知识,这些知识不容易编码,需要与专家反复交流获得。

(4)从企业相关信息管理系统中获取该词典所需的属性、实例等领域知识,本文所得的民航不安全事件日志均来自民航总局提供的“不安全事件情况上报数据库”。

(5)民航领域相关的报告、文件、论文等文献资源、民用航空不安全事件样例等技术资料,以及丰富的网络资源都可以作为民航不安全事件本体的知识来源。

3 构建民航不安全事件语义词典

近几年来所涌现的比较有代表性的本体工程构建方法主要有IDEF-5法、骨架法、TOVE法、Methontology法、循环获取法、七步法以及国内学者提出的螺旋原型法等^[10]。本文参考斯坦福大学医学院开发的七步法,并结合循环获取法的逐步修改不断完善的思想,把本体构建的过程融入于语义词典的构建过程中,提出民航不安全事件的语义词典构建步骤,如图1所示。

(1) 确定专业领域和范畴

构建领域本体或词典,首先要明确本体的专业领域和范畴。民航不安全事件的专业领域即民航飞机在滑行起飞、爬升、巡航、下降、进近、着陆阶段飞机各部

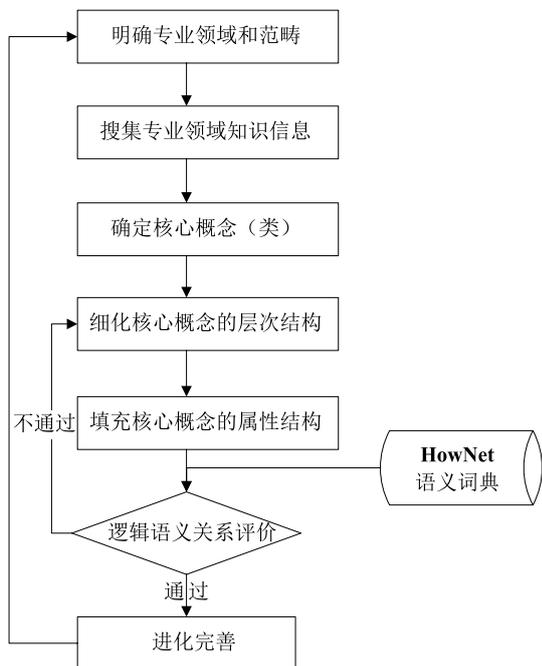


图1 民航不安全事件语义词典构建流程图

件可能产生的不安全事件及相应的状态、影响、处理等领域; 事件的处理对象涉及机组、机务、场务等人员。构建该语义词典的目的是提供对本领域知识的共同理解, 确定对该领域的共同认可的概念, 描述出概念及属性之间的相关关系, 方便行业内各个组织对领域知识的共享。

(2) 搜集专业领域知识

确定了专业领域和范畴后, 需要搜集相关不安全事件信息, 详细了解该领域知识。

通过查阅上文知识来源中所述专业资料、分析现有数据库日志描述, 以及与民航总局领域专家反复沟通交流, 得到本领域核心的概念集合以及各概念层次结构关系表示。

(3) 确定核心概念(类)

经过上一步骤的资料搜集和对已有的数据信息进行分析研究, 运用头脑风暴方法以及专家咨询方法提取出民航不安全事件领域的核心概念如下:

①“情境”: 是民航不安全事件产生的背景信息, 包括航班信息、发生时间、发生地点、飞行阶段、天气情况等, 描述了事件的环境背景差异, 是分析事件产生的重要因素之一。

②“飞机部件”: 是事件发生的载体, 不安全事件的产生往往伴随飞机部件出现异常, 异常的位置通常与所遭受的事件类型具有密切关系。因此, 飞机部件是进

行不安全事件分析和处理的重要概念。

③“事件起因”: 导致不安全事件发生的原因, 反映不同的事件类型, 例如爆胎、鸟击、雷击、故障失效等, 是事件发生的直接表现形式, 与天气环境、人为因素等都有一定的关系。

④“事件处理”: 对不安全事件发生的反应与处理方法, 包括处理决策、检测手段、处理方式等内容, 是民航不安全事件知识的最直接的应用目的。

⑤“人员机构”: 事件中所涉及的反应主体, 一般是事件处理的主体, 主要包括机组、场务、机务等机构, 是民航不安全事件知识中的人为因素。

⑥“设施”: 事件的产生有可能造成周边设施的损毁, 例如飞机冲出跑道造成防护栏损坏。

⑦“状态”: 描述当发生不安全事件时, 经检测后发现的飞机部件及设施所处的状态, 包括异常状态和正常状态, 可根据此分类进而对检查出有异常状态的事件进行挖掘分析。

⑧“事件影响”: 是本次民航发生不安全事件经处理后所得到的最终结果, 包含对本航班的影响以及对其他航班的影响, 例如航班延误、航班取消等。

以上核心概念作为概念模型的顶级概念, 基本能够涵盖民航不安全事件领域涉及的所有知识。

(4) 细化核心概念的层次结构

当语义词典的核心概念确定之后, 需要由上及下地扩展整个语义词典的概念模型。对每个核心概念进行细化的结构分析, 进一步划分出更小的概念类, 将整个事件解析为层次概念模型。

该过程需要细化每个概念, 例如对于“飞机部件”概念, 通过查阅《航天科学技术叙词表》、《民航百科全书》等资料并结合咨询领域专家, 充分获取飞机层次结构知识; 根据《民用航空不安全事件的处置程序》、《民用航空器事故征候》获取“不安全事件”概念的层次结构知识等。

经过对核心概念模型的扩展, 得到民航不安全事件语义词典的概念层次模型, 如图2所示。

(5) 填充核心概念的属性结构

有了概念层次只是有了本体的大体框架, 还需要定义概念的属性结构以填充词典框架。属性是概念中的实体性质, 是对象的性质与对象间的关系的统称, 其中描述对象自身性质的属性为数值属性, 描述对象间关系的属性为对象属性^[1]。

经过对大量已有的数据库日志进行分析挖掘后,

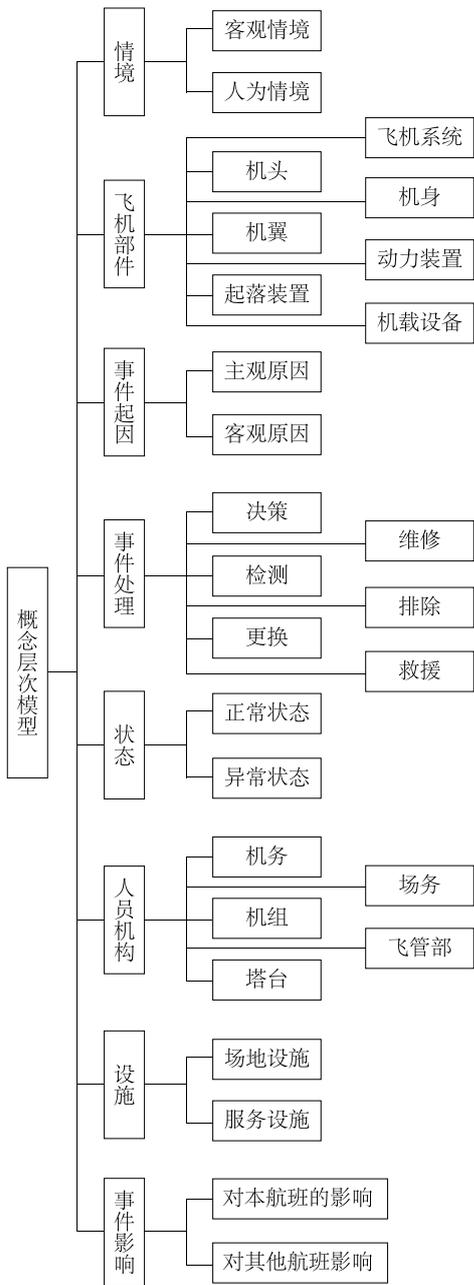


图2 民航不安全事件层次概念模型

得出该领域词典的概念属性模型，如图3所示。

(6) 逻辑语义关系判断

在本文构建领域词典的过程中，由于文献资料的组织方式不同，以及加入多人手工方法添加规则描述，可能会导致词典中概念之间语义关系存在错误。结合外部知识库HowNet，通过查询对比本文领域词典中的概念及属性之间的语义关系是否符合HowNet中定义语义关系，包括上下位关系、同义关系、反义关系、部件—整体关系、事件—角色关系、属性—宿主关系等。例如“窗”与

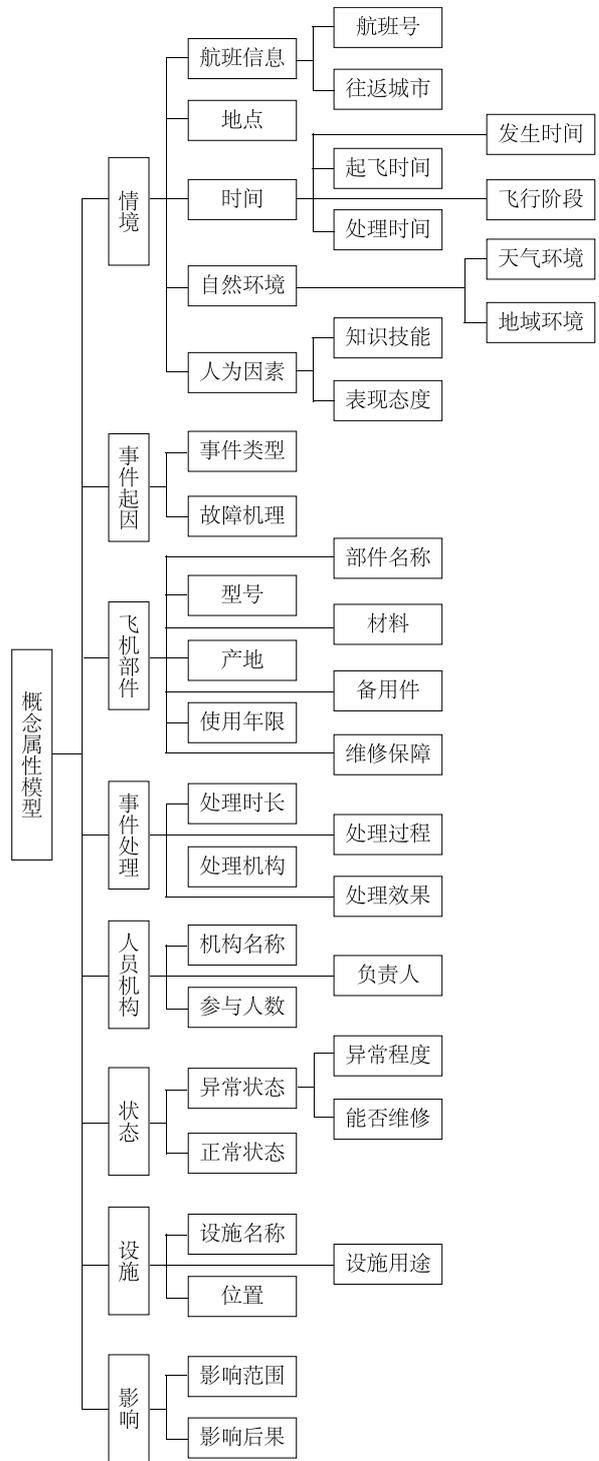


图3 民航不安全事件概念属性模型

“窗户”为同义关系，“机翼”与“左大翼”为上下位关系，“发动机”与“发动机涵道”为整体—部件关系。通过HowNet匹配检测判断语义关系是否通过^[12]，若未通过则对其进行修正并补充，使词典更加正确、完整、有效。

(7) 进化完善

由于领域知识的无穷性, 和领域知识之间的交叉性, 该语义词典需要在使用中不断的进行增量迭代和进化, 构建的整个过程需要螺旋上升循环获取^[13]。通过本阶段先建立起基本的民航不安全事件领域语义词典, 后续过程会继续扩展进化, 从而保证领域本体的完整性、权威性。

通过以上步骤构建完成的民航不安全事件语义词典的结构如图4所示。

该结构描述图包含了构成民航不安全事件的核心概念、核心概念的部分属性和层次结构、以及这些概念之间的关系, 其中实线表示概念与概念之间的相关关系, 虚线表示概念与属性之间的相关关系。

4 民航不安全事件语义词典的相关应用

4.1 日志语义标引

标引是对资源进行特征提取和标示引导的过程, 它不仅能够合理组织和管理文档日志, 也是实现检索的前提基础。传统标引通过词频统计提取关键词, 再根据关键词对文档日志进行索引^[14]。语义标引则是根据该领域语义词典对数据库中的文本数据标引出概念类、概念属性相关特征词的过程。目前国内外对语义标引的研究均呈现不断增长的趋势, 文献[15]通过文献

计量法对国内外语义标注相关文献进行统计分析, 包括研究内容分布、时间分布以及研究力量分布等多个方面, 总结对比了国内外语义标注的研究进展; 文献[16]提出了基于领域本体的文献语义标引模式, 为语义检索打下基础; 文献[17]以基于潜在语义标引方法进行文本聚类实验, 通过多组实验分析跨语言文本聚类中的基于潜在语义标引性能和发展方向。文献[18]提出了一种基于领域本体的空间向量模型的语义标引方法来建立本体概念与词汇的映射关系, 并结合石油产品领域对该方法进行实例说明。

本文使用语义词典进行语义标引的基本思路如下: 首先, 从民航不安全事件日志中抽取代表该领域的特征词汇, 收集形成特征词汇集合; 然后, 对一篇新的日志进行领域分词, 通过相似度匹配^[19]将特征词汇集合中不同概念的词汇与该日志分出的词汇关联起来; 最后, 将这些匹配到的特征词汇与语义词典中的概念再进行匹配并完善语义词典。这样, 就形成了事件日志与语义词典中具体概念之间的映射关系。

基于语义词典的语义标引示意图如图5所示。



图5 语义标引示意图

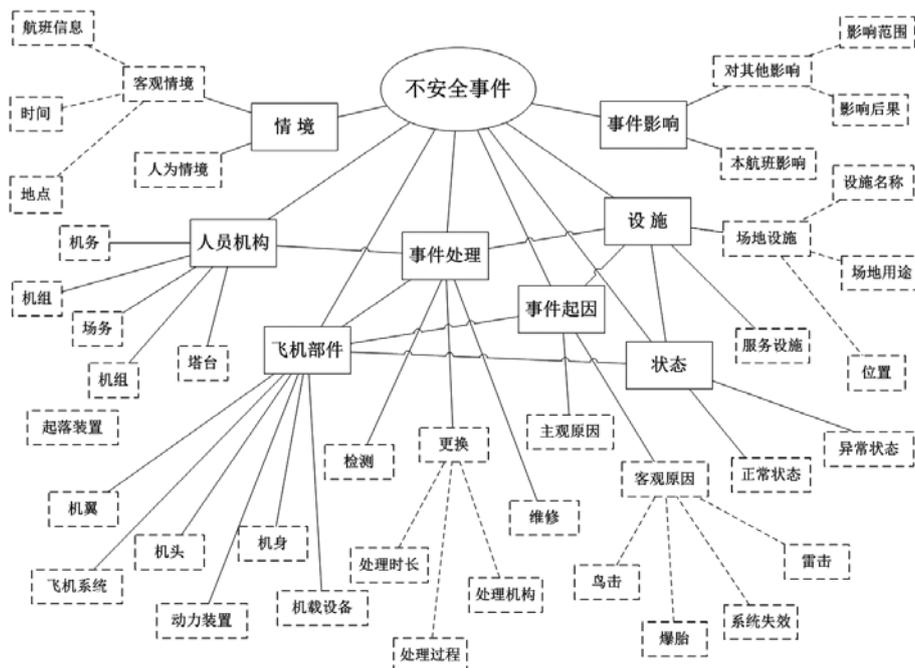


图4 语义词典结构描述图

由于民航不安全事件语义词典对领域内的概念、属性及之间的关系有着详细精确的描述,通过此方法对日志进行标引后可以把日志的隐含语义信息表示出来,同时也可对提取的特征词汇进行较为准确的概念划分,例如将事件日志中的“左发右涵道”提取出来,并给该特征词划为“飞机部件->动力装置->发动机部件->涵道”的层次概念属性。

通过语义标引提取特征词汇,能够在全面把握中心思想的基础上,实现对上报的中文文本内容进行虚词过滤和特征词汇识别,提取出若干个代表主要语义内容的词汇或短语,相关结果可用于精化阅读、语义结构分析和完整性分析等。

4.2 逻辑语义结构分析

由于不安全事件日志多由人工编写,描述无统一格式,记录内容的语法、结构也会随着记录人员的不同而有很大差异,所以这些日志属于非结构化文本。这种日志中事件要素遗漏的现象很常见,导致整体信息质量不高,影响了信息的处理和分析。为了提高信息上报质量,需要制定信息上报标准或对上报信息格式化,提供给信息填报人员参考的方法。

逻辑语义是一种语义关系,诸如施事、动作、受事、空间、材料、方式、新质等。在特定的交际环境下,某一语言片断的各个基本单元之间必然存在着某种逻辑关系,这些逻辑关系称为逻辑语义,逻辑语义的集合称为逻辑语义结构^[20]。例如,在描述为“场务检查机翼”的日志中,“场务”为施事,“检查”为动作,“机翼”为受事,施事、动作、受事组成了该描述的逻辑语义结构。

根据语义词典和HowNet知识库等,将已标引的特征词汇按照该词的概念属性进行逻辑语义结构分析,不同的语义结构包含的逻辑语义也不相同。根据需求将不规则的描述结构化、格式化,构建出具体描述的语义结构模型,以方便用户对该具体描述进行语义梳理和内容分析。

4.3 完整性分析

不安全事件日志的描述随着记录人员的不同有很大差异,因此事件要素遗漏现象十分常见,影响对信息内涵项、内涵项对应的所有可能信息的挖掘整理。通过上文逻辑语义结构分析,可实现对不安全事件日志描述的完整性分析。

(1) 对整体事件要素的完整性分析

使用文本挖掘技术,结合语义词典,根据不同事件在语义词典中所必须包含的核心概念,对日志中的事件要素进行检测和分析,根据该类型事件中应有的事件要素来判断描述文本中的事件要素是否完整,若发现缺失的事件要素,可对日志描述提出更改意见,从而提供信息质量自动审核和评估能力。

例如,一段事件起因类型为“雷击”的描述文本中,没有提取到处理方法和事件要素,那么我们就判断这段描述是不完整的,缺少处理方法的相关描述,可返回用户进行修改。

(2) 对具体语义结构的完整性分析

通过上文的语义结构分析,可以实现对一篇日志的结构化描述,得到特定的语义结构模型。在对具体某项时间的语义分析中,若文本本身不完整,则会缺失必须的语义要素,根据语义词典中各个核心概念的关系,对具体语义描述进行完整性分析,发现并提示缺失要素。

例如,在对一段描述文本进行分析的时候,已提取出检测方法、飞机部件、状态信息的关键词,我们可以初步判定描述中有对飞机进行检测的语义,但是对比语义关系中的关联规则,发现缺少了“人员机构”的相应特征词,那么这段描述将被判断为存在要素缺失,缺失关于“人员机构”的相应描述。

通过完整性分析功能,可以通过机器分析,在大量文本中挖掘出疑似信息不完整的描述信息,通过给出的缺失要素,用户可以更方便快捷地审核和修改原有文本,而实现提升信息质量的目的。

4.4 实例分析

以 JAVA、JavaScript、HTML为编程语言,以Oracle数据库和TRS全文数据库系统(TRS Database Server)作为数据存储系统,以SSH(Struts, Spring, Hibernate)作为系统开发框架,通过Eclipse开发平台,依靠已经构建的航空不安全事件本体模型,利用文本分词、关键词标引、文本分类、语义相似度计算等中文文本挖掘技术,可设计开发出一款B/S结构的“航空不安全事件报告质量管理体系”。

从民航某单位“不安全事件情况上报数据库”提供的数据中抽取一篇事件日志做实例分析,日志描述为“7月31日,某机型飞机执行哈尔滨-济南航班,济南过

站检查,发现右发遭鸟击,经检查飞机未受损,正常执行后续航班。”首先对该描述文本分词,然后通过文本相似度匹配,与语义词典中已包含的具体领域特征词汇集合“济南”、“过站检查”、“右发”、“遭鸟击”、“未受损”、“正常执行后续航班”匹配成功,例如“右发”是语义词典中核心概念“飞机部件”下子概念“动力装置”所包含的概念词汇,由此形成了事件日志与语义词典中具体概念之间的映射关系,完成了该日志的语义标引。

根据特征词汇在语义词典中的所属概念以及概念之间存在的语义关系,结合关键词的语序,对已标引出的特征关键词进行逻辑语义结构分析,如图6所示。

根据实际需求,需要分析出该日志所涉及的时间、地点、事件类型、飞机部件、人员机构、处理方法、检测状态及事件影响,根据语义分析,可将特征词汇分为

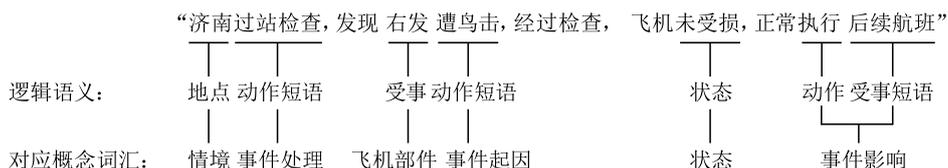


图6 逻辑语义分析及概念对应

```
<terminated> AnalysisDriver [Java Application] C:\Program Files(x86)\Java\jdk1.6
```

```
analysis use time:25
```

```
情境描述:
```

```
时间: value = null \异常缺失
```

```
地点: value = 济南
```

```
category = 情境
```

```
categoryPath = 情境 \ 地点
```

```
事件描述:
```

```
事件类型: value = 遭鸟击
```

```
category = 事件起因
```

```
categoryPath = 事件起因 \ 客观原因
```

```
飞机部件: value = 右发
```

```
category = 飞机部件
```

```
categoryPath = 飞机部件 \ 动力装置 \ 发动机
```

```
处理过程描述
```

```
人员机构: value = null \异常缺失
```

```
处理方法: value = 过站检查; 检测
```

```
category = 事件处理
```

```
categoryPath = 事件处理 \ 检测; 事件处理 \ 检测
```

```
检测状态: value = 未受损
```

```
category = 状态
```

```
categoryPath = 状态 \ 正常状态
```

```
影响描述:
```

```
事件影响: value = 正常执行后续航班
```

```
category = 事件影响
```

```
categoryPath = 事件影响 \ 对本航班影响
```

图7 逻辑语义结构分析输出结果

“情境描述”、“事件描述”、“处理过程描述”、“影响描述”4块结构。该日志描述的输出结果如图7所示。

通过以上分析结果可发现处理过程中“情景描述一时间”异常缺失,“人员机构”异常缺失,因此这段不安全事件信息的描述是不完整的,经自动审核后给出提醒,方便信息录入工作者补充缺失信息,从而完善提高了信息上报质量。

5 结语

基于民航某单位对不安全事件分析的实际需求,本文对民航不安全事件知识进行了语义词典的构建及相关应用研究。采用构建本体的方法来构建语义词典,利用该语义词典可以初步实现对不安全事件日志的语义标引、语义结构分析,并对日志的完整度进行分析,

从而有效促进民航不安全事件信息的积累和共享应用,进而提高信息上报质量。本研究所构建的语义词典具有较高的准确性和规范性,可为整个行业共享、复用提供借鉴。但是语义词典的构建是一个较为复杂的过程,尤其是在构建初期阶段和维护阶段,为了保证词典的完整性及准确率,需要花费大量的时间和人力,因此本词典的构建尚处于实验阶段,在今后还需要进一步完善。

参考文献

- [1] 中国民用航空总局航空安全办公室. 民用航空不安全事件的处置程序[J]. 中国民用航空总局公报,2004 (05).
- [2] 张文秀,朱庆华. 领域本体的构建方法研究[J]. 图书与情报,2011, 155(1):16-19, 40.
- [3] Erdmann M, Studer R. Ontologies as Conceptual Models for XML Documents[C]// Proceedings of the 12th Workshop on Knowledge Acquisition, Modeling and Management(KAW '99),Banff, Canada, October 1999.

- [4] Breitman K K, Filho A H, Haeusler E H. Using ontologies to formalize services specifications in multi-agent systems, NASA/TM-2006-214296[R]. Washington DC: National Aeronautics and Space Administration, 2006.
- [5] 张桂平,刁丽娜,王裴岩. 基于HowNet的航空术语语义知识库的构建[J]. 中文信息学报, 2014, 28(5):92-101.
- [6] 王红,杨璇,王静,等. 基于本体的民航应急决策知识表达与推理方法研究[J]. 计算机工程与科学,2011, 33(4):129-133.
- [7] 陈铭. 航空事故案例库设计及检索方法研究[D]. 南京:南京航空航天大学,2009.
- [8] 陈壮坚. 面向本体的专业领域语义词典及其实现[D]. 上海:复旦大学,2006.
- [9] Gómez-Pérez A, Benjamins R. Overview of knowledge sharing and reuse components: ontologies and problem-solving methods[C]// Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI-99). Stockholm: the IJCAI-99 Workshop, 1999: 1-15.
- [10] 李景. 领域本体的构建方法与应用研究[D]. 北京:中国农业科学院,2009.
- [11] 丁晨春,傅柱. 基于航天叙词表的领域本体半自动化构建研究[J]. 情报理论与实践, 2011, 34(11):113-116.
- [12] 陆文豪. 基于关系数据库的专业领域语义词典构建研究[D]. 上海:复旦大学,2009.
- [13] 周扬,李青. 飞机故障知识的本体建模及语义检索[J]. 计算机工程与应用, 2011, 47(16):12-15.
- [14] 刘紫玉. 多专业领域本体的构建及语义检索研究[D]. 北京:北京交通大学,2009.
- [15] 邱均平,牟楠,楼雯,等. 国内外语义标注研究进展分析[J]. 现代图书情报技术,2014, 37(05):12-16.
- [16] 王欣,孔庆杰,徐宝祥. 基于领域本体的文献语义标引模式的构建研究[J]. 情报科学,2013, 31(9):88-92.
- [17] 邓三鸿,万接喜,王昊. 基于特征翻译和潜在语义标引的跨语言文本聚类实验分析[J].现代图书情报技术,2014(01):28-35.
- [18] 时念云,杨晨. 基于领域本体的语义标注方法研究[J].计算机工程与设计,2007, 24(28):5985-5987.
- [19] 张会昌. 基于领域词典的中文文本相似度匹配[D]. 济南:山东大学,2014.
- [20] 董振东. 逻辑语义及其在机译中的应用[DB/OL].[2015-08-14]. http://www.keenage.com/html/c_index.html.
- [21] 王红,杨璇,王静,等. 基于本体的民航应急决策知识表达与推理方法研究[J]. 计算机工程与科学,2011,33(4):129-133.
- [22] 王功辉,黄奇,秦超,等. 本体构建中的语义分析方法研究[J]. 图书情报工作,2013, 57(7) :106-111.
- [23] 马中杰. 基于领域本体的语义检索系统研究[D]. 合肥:安徽大学,2011.
- [24] 于娟. 基于文本的领域本体学习方法及其应用研究[D]. 大连:大连理工大学,2010.
- [25] 刘义俊. Web2.0环境中的领域本体构建方法研究[D]. 南京:南京航空航天大学,2013.

作者简介

马慧, 女, 1991年生, 在读研究生, 研究方向: 知识组织、信息资源管理, E-mail: emmahui0805@126.com。
赵捧未, 男, 1958年生, 博士, 教授/博导, 研究方向: 信息存储与检索、信息资源管理。
王洪俊, 男, 1975年生, 博士, 研究方向: 文本挖掘与信息检索。

Research on Construction and Application of Civil Aviation Unsafe Events Semantic Dictionary

MA Hui¹, ZHAO PengWei¹, WANG HongJun², SUN XinBo²

(1. School of Economics and Management, Xidian University, Xi'an 710126, China;

2. Beijing TRS Information Technology CO., LTD, Beijing 100101, China)

Abstract: In view of the poor quality of information reporting which resulted from the lack of unified knowledge model in the field of civil aviation security events, the ontology knowledge representation method is adopted to study the representation and its related applications of security events, in which the knowledge source of security events is analyzed based on the elements of ontology. Furthermore, on basis of "seven steps" of Ontology Modeling, Civil Aviation Unsafe Events Semantic Dictionary is compiled with continuous effort and improvement. Meanwhile, semantic indexing of unsafe event log and its logical semantic structure and integrity are investigated as well, which are in favor of improving the quality of information reporting in the civil aviation security events.

Keywords: Civil Aviation Unsafe Events; Semantic Dictionary; Ontology

(收稿日期: 2015-08-26; 编辑: 王立学)