

ISSN 1673-9418 CODEN JKYTA8  
Journal of Frontiers of Computer Science and Technology  
1673-9418/2000/00(00)-0000-00  
doi: 10.3778/j.issn.1673-9418.1506033

E-mail: fctst@vip.163.com  
<http://www.ceaj.org>  
Tel: +86-10-89056056

## 针对环境污染突发事件领域的事件本体模式\*

刘 炜<sup>+</sup>, 丁 宁, 杨竣辉, 王 旭, 张雨嘉, 刘宗田  
上海大学 计算机工程与科学学院, 上海市 200444

### Event Ontology Pattern for Domain of Environmental Pollution Emergency<sup>\*</sup>

LIU Wei<sup>+</sup>, DING Ning, YANG Junhui, WANG Xu, ZHANG Yujia, LIU Zongtian

School of Computer Engineering and Science, Shanghai University, Shanghai 200444, China

+ Corresponding author: E-mail: liuw@shu.edu.cn

**LIU Wei, DING Ning, YANG Junhui, et al. Event Ontology Pattern for Domain of Environmental Pollution Emergency. Journal of Frontiers of Computer Science and Technology, 2000, 0(0): 1-000.**

**Abstract:** Event ontology design pattern (EODP) is a solution to facilitate event ontology development, and to help reducing arbitrariness design. This paper firstly proposes a general event model structure for representation of general event information on Web, without making assumptions about the domain-specific vocabularies used. And then, this paper presents the concept of EODP and the development method of EODP, by which an environment pollution event ontology pattern is proposed based on shared vocabulary. At last, this paper discusses the formalization of the environment pollution event ontology pattern by using the description logic, which verifies the effectiveness and concept-consistency while modeling domain ontology by using the event ontology pattern.

**Key words:** Environmental Pollution Emergency; Event Ontology Pattern; Description Logic

**摘 要:** 事件本体设计模式是加快事件本体构建、避免设计歧义以及在建模过程中重复出现相同错误的一种解决方案。提出了一种通用的事件模型结构,该模型结构可以在没有特定领域词汇的情况下表示 Web 文本中最一般性的事件信息。在此基础上,提出了事件本体模式概念和事件本体模式的提取方法,并利用该方法建立基于共享词汇的环境污染突发事件本体模式。最后,利用描述逻辑对环境污染突发事件本体模式进行形式化,验证了使用事件本体模式实现领域事件本体建模的有效性和概念一致性。

**关键词:** 环境污染突发事件; 事件本体模式; 描述逻辑

**文献标志码:** A **中图分类号:** TP391

\*The National Natural Science Foundation of China under Grant Nos. 61305053, 61273328 (国家自然科学基金).

Received 2015-06, Accepted 2015-09.

## 1 引言

环境突发事件是目前国内常发的一类突发事件,它包括各种污染事件(如雾霾、饮用水污染、核材料泄漏、工厂偷排污染物等)以及污染事件引发的一系列社会事件(如化工项目引发的群体事件、抢购饮用水和口罩等),对社会稳定影响很大。这些事件通过新闻报道或者通过网络用户在论坛、微博、微信等互联网应用中发布,通常会引发网络用户激烈的讨论,而这些事件的在互联网中的讨论和演化反过来也会影响到现实世界。环境突发事件应急预案系统用于从互联网上搜集和处理与污染事件相关联的各种事件信息,并帮助环境管理者和突发事件响应部门监测、分析突发事件并及时采取有效措施是十分必要的。传统的环境突发事件预警系统普遍采用框架方法<sup>[1]</sup>或网络模型<sup>[2]</sup>描述突发事件信息,采用框架方法描述突发事件具有较好的完整性,但是无法表示突发事件的动态过程,而且无法对知识进行形式化,不能支持严密的推理。采用网络模型描述突发事件,虽然可以描述事件发生过程的状态迁移,但是对突发事件的内部结构无法进行描述,而且对于事件之间的语义关系(尤其是非分类关系)也缺乏相应的表示方法。利用本体表示突发事件信息并实现推理预警是近几年出现的一种新方法<sup>[3][4]</sup>,本体方法能够对环境突发事件进行概念分类,基于描述逻辑的本体形式化基础有助于实现突发事件在语义级别上统一表达,从而达到知识的共享和互操作。因此,建立相关本体在环境突发事件领域正变得越来越重要,是污染突发事件报警系统的不可缺少的组成部分。然而,以概念为中心的传统本体在构建突发事件领域本体知识时存在着诸多缺陷:(1)概念离散的问题,突发事件概念和突发事件的参与者、地点和时间不是作为一个有机整体存在,因此在构建本体过程中存在大量的概念分类工作;(2)传统本体往往容易忽视突发事件的动态性,比如很难描述突发事件随时间变化所呈现出的不同状态;(3)传统本体中概念的关系单一,不能对事件之间的语义关系进行充分的描述,造成推理能力的不足。

以事件为中心的本体建模能够捕获特定领域的动态特征。事件能够很好地体现人物、地点、动作和对象之间的复杂关系。事件关系为事件及相关

概念的语义描述和推理提供了更精确的表示。本文首先提出一种针对 Web 文本分析与处理的事件模型结构来表示文本通用的事件信息。该结构充分考虑文本事件描述中所包含的动作、地点、参与者、时间、事件过程中的状态以及文本语言的表现形式,该结构试图对不同领域的事件及事件类进行描述,因此没有针对具体领域的事件术语做假设。此外,该事件结构具有最小的语义承诺从而保证了所建立的事件或事件类之间具有最大的互操作性。在此基础上,考虑利用现有的本体语言或术语词典(包括标准的 OWL 2 和人物、地点本体)来实现不同事件要素的描述,力求达到对事件要素类进行复用。例如,重用 FOAF<sup>1</sup>中的 Agent 类来描述事件的施动者,使用时间本体(Time Ontology<sup>2</sup>)来描述事件中的时间要素。

针对环境突发事件的事件本体构建过程通常是十分复杂的。特别是当本体的规模和复杂度不断增加的时候,构建本体的难度也会大幅上升。本体构建者往往都是特定领域的专家,而非逻辑专家。他们更关注的是特定领域的概念,而不是本体的表示方法。此外,在构建环境突发事件本体时,本体构建者常常会建立和遵循一套通用的建模方案来控制本体建模的复杂度。而在构建具体的事件本体时(如空气污染本体和水污染本体),不同事件本体之间存在着大量相似的事件类、事件要素以及事件关系。这些相似的部分可以通过本体模式进行抽象。事件本体设计模式的作用是从形式上捕获重现的事件本体模型,并实现已存在的事件本体词汇的重用,来简化事件本体构建过程,并避免一些常见的错误。本文提出一种基于共享词汇的环境污染事件本体模式,并设计了从多个领域事件本体中抽取事件本体模式的算法。最后,利用描述逻辑对本体模式进行形式化,通过实例验证了事件本体模式实现领域事件本体建模的有效性和概念一致性。

## 2 事件本体相关概念

近年来,研究人员针对基于事件的知识表示提出了多种模型结构<sup>[5-8]</sup>,这些模型结构由于不同的侧重点、不同的关注领域以及形式化层次而各不相

<sup>1</sup> FOAF Vocabulary Specification. <http://xmlns.com/foaf/spec/>

<sup>2</sup> Time Ontology in OWL. <http://www.w3.org/TR/owl-time>

同。但是大多数的事件模型结构都是类似于 5W 模型(Who, What, When, Where, How)。我们在研究事件结构的过程中发现在对事件进行表示时, 非常重要的一点是对事件发生过程中内部状态的变化进行描述。同时在构建事件知识库的时候, 由于多数是通过人工方式构建, 因此往往忽视了事件知识在文本中的语言表现, 而事件的语言表现能对计算机进行文本事件的自动侦测和识别起到重要的作用。

## 2.1 事件规范说明

**定义 1** (事件) 本文将事件定义为在某个特定的时间和地点发生的, 由若干角色参与, 表现出若干动作特征, 并伴随着对象内部状态变化的一件事情。其形式化表示为:

$$Event ::= \langle A, O, T, P, S, L \rangle$$

从知识表示的角度, 我们将事件定义为包含了动作 (A)、对象 (O)、时间 (T)、地点 (P)、内部状态 (S) 和语言表现 (L) 等六个要素的知识单元。其中 A 表示事件所包含的动作或动作序列的集合, 在文本中, 动作通常是作为识别一个事件的触发词; O 表示一个事件中的对象集合, 包括事件中的所有的参与者和涉及到的对象, 我们将事件对象分为主体和客体; T 表示事件发生的时间段, 事件时间可以是绝对时间也可以是相对时间, 两类时间都可以通过计算转换成形如  $[t_1, t_2]$  的序偶表示, 以此描述事件的开始、发展和结束时间, 当开始时间和结束时间一样时, 表示事件发生在瞬间。P 表示事件发生的地点, 可以表示物理地点 (如中国香港、302 国道), 也可以表示逻辑位置 (如淘宝网、微信朋友圈); S 表示事件发生过程中对象的状态集合, 由事件发生的前置条件、后置结果集合组成。前置条件指该事件发生前满足事件发生的触发条件的对象状态集合; 后置结果是事件发生后对象结果状态的集合。L 表示事件的语言表现, 主要包括事件核心词表现、事件核心词搭配, 核心词表现为事件在句子中常用的标志性词汇, 通常也是计算机识别事件的触发词, 核心词搭配是指核心词与其他词汇的固有搭配。在事件的六个要素中, 前五个要素是事件的内在要素。

**实例 1.** 这些天, 上海遭遇雾霾围城, PM 2.5 值突破了 300。

**事件名:** 上海遭遇雾霾

**动作 (触发词):** 遭遇雾霾、突破

**对象:** PM2.5 值

**时间:** 这些天

**地点:** 上海

**状态:** 前置状态={PM2.5 值正常、空气质量正常}, 后置状态={PM2.5 大于 300、空气质量恶劣}

**语言表现:** 雾霾事件={雾霾}, 事件搭配={“遭遇”+事件名}

**定义 2** (事件类) 指具有共同特征的事件的集合, 它由事件的外延和内涵构成。事件类的外延指的是由属于该事件类的事件实例组成的集合。事件类的内涵指的是属于该事件类的所有事件实例在某个要素上具有的共同特性。

**定义 3** (事件关系) 指存在于事件类或事件之间的语义关系, 事件系分为两类: 分类关系和非分类关系。其中分类关系指的是事件类之间的包含关系或父子关系, 非分类关系指的是事件或事件类之间内在的语义关系, 包括组成关系 (*isComposedOf*)、跟随关系 (*follow*)、因果关系 (*causal*) 和并发关系 (*concurrency*)。

**包含关系 (subsumption 或 is\_a):** 指事件类存在种属关系。对于事件类  $EC_1 = \{E_1, C_{1A}, C_{1O}, C_{1T}, C_{1P}, C_{1S}, C_{1L}\}$  和事件类  $EC_2 = \{E_2, C_{2A}, C_{2O}, C_{2T}, C_{2P}, C_{2S}, C_{2L}\}$ , 存在  $EC_1 \subset EC_2$  当且仅当

$C_{1A} \subset C_{2A} \wedge C_{1O} \subset C_{2O} \wedge C_{1T} \subset C_{2T} \wedge C_{1P} \subset C_{2P} \wedge C_{1S} \subset C_{2S}$ 。包含关系只用于事件类之间, 比如污染事件类和水污染事件类就是包含关系。事件类包含关系形式化为  $R_{is\_a}$ 。

**组成关系 (isComposedOf):** 当事件类  $EC_1$  中的一个事件实例由事件类  $EC_2$  中的某个事件实例组成时, 称事件类  $EC_1$  由事件类  $EC_2$  组成, 或称事件类  $EC_2$  是事件类  $EC_1$  的组成事件类。“大气污染”事件类由“污染物排放”、“检测污染源”等事件类组成。组成关系形式化为  $R_{comp}$ 。

**因果关系 (Causal):** 事件类  $EC_1$  的实例事件发生以一定的概率导致了事件类  $EC_2$  的实例事件发生, 发生的概率大于给定的阈值, 则称两事件类之间存在因果关系, 称  $EC_1$  是  $EC_2$  的因,  $EC_2$  是  $EC_1$  的果。因果关系形式化为  $R_{cause}$ 。因果不但反映了事件之间的相互影响, 还在时间上反映了事件发生的先后关系。如:

6 月 1 日, 阿富汗坎大哈市内的一座清真寺发生爆炸事件, 造成包括喀布尔市警察总监在内的至少 40 人死亡, 另有 60 人受伤。

其中“爆炸”和“死亡”、“受伤”是因果关系。

**跟随关系 (follow):** 在一定长度的时间段内,





模式命名为 EODP(Event Ontology Design Pattern)。事件本体设计模式能够改进事件本体模型的可重用行和可扩展性,提高本体开发效率,同时也能够使不同应用领域的异构的、基于事件的语义数据的集成工作简化和标准化。

事件本体模式的粒度因不同的抽象程度而不同。一般来说,抽象程度越高,可重用的范围越大,本体模式的粒度也越大,但是可重用的元素就越少。相反,抽象程度越低,可重用的范围越小,本体模式的粒度也就越小,但是可重用元素越多。因此,建立事件本体模式需要在抽象程度和可重用元素的数量之间找到平衡点。我们的观点是,在不同的应用领域建立相应的事件本体模式,例如在环境污染突发事件领域、交通事故、恐怖袭击领域等。事件本体模式通常是通过对具体的领域本体进行抽象得到的。也就是说,可以将不同事件本体中公有的事件类、事件要素以及事件关系结构提取出来,抽象出一个通用的事件本体模型。当需要创建一个新的具体的事件本体时,相应的事件类或事件关系可以按照本体设计模式进行重用。即参照建模的场景,对模式中的事件类(包括事件要素和事件关系)进行继承、扩展或覆盖。

事件本体模式的建立通常按照以下几个步骤:

(1) 抽象出事件本体的公共部分。针对要建立事件本体模式的领域,对该领域中不同类型的具体突发事件案例进行本体建模(例如建立具体的空气污染、水域污染和工厂排污等本体),并通过不断的分析和验证,抽取出不同子领域本体模型中可复用的抽象事件类和事件要素;

(2) 建立事件本体模式。利用步骤(1)中抽取出的事件类和事件要素建立抽象事件本体模型,即事件本体模式。

(3) 验证和修正本体模式。通过复用步骤(2)中的事件本体设计模式,为其他子领域建立事件本体模型,验证事件本体设计模式与新建事件本体在事件类、事件要素和事件关系上的一致性,并对本体模式进行修正。

下面阐述环境污染突发事件领域的事件本体模式构建过程。

环境污染问题具有多样性,包括空气污染、水污染、土壤污染,核污染或电磁污染等。因此,根据不同的污染问题构建不同的本体是十分困难的,并且需要花费大量的时间。通过分析不同类型的污

染事件以及由其引发的社会事件的基本特征不难发现,在污染突发事件的事件本体模型中总是存在许多相似的事件类和事件要素,以及组成结构相似的事件关系。例如,空气污染事件本体和水污染事件本体都包含事件类“污染(Pollution)”,而事件类“污染(Pollution)”在要素层面总是包含一个参与者“污染方(polluters)”和一个对象“污染物(pollutant)”。因此,可以将不同类型的环境突发事件本体中公有的事件类、事件要素以及事件关系结构提取出来,抽象出一个通用的事件本体模型。

### 3.1 事件本体构建

本节首先分别为不同应用场景的环境污染进行本体建模。由于篇幅所限,本文仅讨论两种具体本体:“工厂废气排放污染事件本体”和“交通工具化学品泄漏水域污染事件本体”。同时,将尝试发现这两个本体之间诸多相似的事件类和事件要素。通过对来自互联网超过 60 篇有关空气污染和水污染的文章进行分析和事件标注。然后在相关环境专家的指导下,创建了两个本体,如图 2 和图 3。

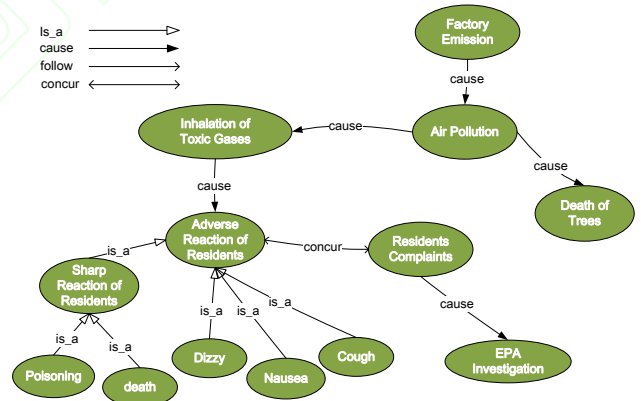


Fig.2 Event ontology of air pollution incident caused by factory gas emissions

图 2 工厂排放废气事件本体

图 2 包含了空气污染突发事件中的几个关键的事件类以及它们之间的事件关系(图中省略了事件要素)。该本体描述了“工厂排污”(Factory Emission)引发“空气污染”(Air Pollution);“空气污染”(Air Pollution)又导致了“居民吸入有毒气体”(Inhalation of Toxic Gases)和“树木死亡”(Death of Trees);“居民吸入有毒气体”(Inhalation of Toxic Gases)又引发了“健康危害”(Health Hazard)以及“居民投诉”(Residents Complaints);“健康危害”(Health Hazard)包括“头晕”(Dizzy)、“恶心”(Nausea)、“咳嗽”(Cough)等“居民不良反应”

(Adverse Reaction of Residents), 还包括“中毒”(Poisoning) 和“死亡”(Death) 两种“严重的不良反应”(Sharp Adverse Reaction of Residents); “居民投诉”(Residents Complaints) 则导致了“环保部

门调查”(EPA Investigation)。

表 1 罗列了空气污染事件本体中关键的事件类, 以及与其单向关联的事件类和事件关系类型。

Table 1 Event classes and relations in event ontology of air pollution incident caused by factory gas emissions  
表 1 工厂废气排放污染事件本体事件类及事件关系

主要的事件类名	关联事件类 (含关联关系类型)	抽象事件类
Factory Emission	<<causal>>_Air Pollution	Environment Emergency
Air Pollution	<<causal>>_Inhalation of Toxic Gases; <<causal>>_Death of Trees	Pollution
Inhalation of Toxic Gases	<<causal>>_Adverse Reaction of Residents	Exposure to Toxic Substances
Death of Trees	<<null>>	Ecological Damage
Adverse Reaction of Residents	<<is_a>>_Sharp Reaction of Residents; <<is_a>>_Dizzy; <<is_a>>_Nausea; <<is_a>>_Cough; <<concur>>_Residents Complaints	Adverse Reaction of Residents
Sharp Reaction of Residents	<<is_a>>_Poisoning; <<is_a>>_Death	Sharp Adverse Reaction of Residents
Residents Complaints	<<causal>>_EPA Investigation	Residents Complaints
EPA Investigation	<<null>>	Official Investigation

图 3 描述的是交通工具化学品泄漏引发水域污染的事件本体。该本体包含的几个主要的事件类及其事件关系: “交通工具化学品泄漏”(Vehicle Chemical Leakage) 造成了“河流污染”(River Pollution); “河流污染”导致“居民饮用水源受污”(Residents Drinking Water Pollution)以及“动物死亡”(Animal Death); 随之而来的是“居民饮用了受污染的水”(Residents Drink Polluted Water)而导致“健康危害”(Health Hazard of Residents)。当居民遭受了健康危害, 会抱怨政府部门或通过互联网发布其遭遇, 即“居民投诉”(Residents Complaints)。并且, 一些居民还可能会“抢购饮用水”(Panic Buying Drinking Water)。同时, “健康危害”又可以细分为“居民不良反应”(Adverse Reaction of Residents), 如“腹泻”(Diarrhea)、“恶心”(Nausea)、“呕吐”(Emesis)、“头晕”(Dizzy) 等, 和“严重的健康损害”(Severely Heath Injury of Residents), 如“中毒”(Poisoning) 和“死亡”(Death)。“居民的投诉”引起“记者秘密调查”(Reporter Secretly Investigation); 进而导致“环保部门调查”(EPA

Investigation)。随后, 环保部门采取措施“紧急对饮用水源进行清理”(Emergency Purify Drinking Water Source)。

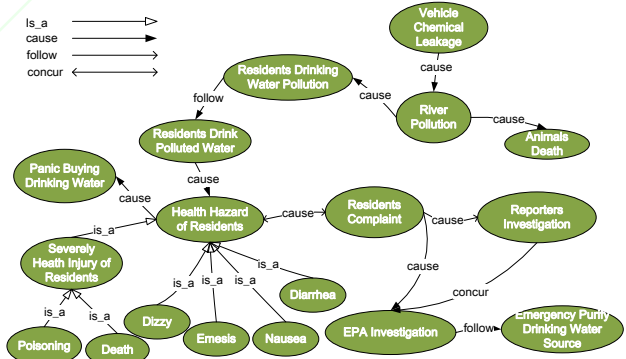


Fig.3 Event ontology model of water pollution incident caused by vehicle chemical leakage

图 3 交通工具化学品泄漏水域污染事件本体  
表 2 列出了交通工具化学品泄漏引发水域污染本体中主要的事件类, 以及与其单向关联的事件类和事件关系类型。

Table 2 Event classes and relations in event ontology of water pollution incident caused by vehicle chemical leakage  
表 2 交通工具化学品泄漏水域污染事件本体主要事件类及事件关系

主要的事件类名	关联事件类	抽象事件类
Vehicle Chemical Leakage	<< causal >>_River Pollution	Environment Emergency



<i>River Pollution</i>	<<causal>>_Residents Drinking Water Pollution; <<causal>>_Animal Death	<i>Pollution</i>
<i>Residents Drinking Water Pollution</i>	<<follow>>_Residents drink polluted water	<i>Pollution</i>
<i>Residents drink polluted water</i>	<<causal>>_Adverse Reaction of Residents	<i>Exposure to Toxic Substances</i>
<i>Animal Death</i>	<<null>>	<i>Ecological Damage</i>
<i>Panic Buying Drinking Water</i>	<<null>>	<i>Residents Resist Pollution</i>
<i>Adverse Reaction of Residents</i>	<<is_a>>_Dizzy; <<is_a>>_Nausea; <<is_a>>_Emesis; <<is_a>>_Diarrhea; <<concur>>_Residents Complaints;	<i>Adverse Reaction of Residents</i>
<i>Severely Heath Injury of Residents</i>	<<is_a>>_Poisoning; <<is_a>>_Death	<i>Sharp Adverse Reaction of Residents</i>
<i>Residents Complaints</i>	<<causal>>_EPA Investigation <<follow>>_Reporter Secretly Investigation	<i>Complaints</i>
<i>Reporter Secretly Investigation</i>	<<causal>>_EPA Investigation	<i>Private Investigation</i>
<i>EPA Investigation</i>	<<causal>>_Emergency Purify Water Source	<i>Official Investigation</i>
<i>Emergency Purify Water Source</i>	<<null>>	<i>Official Response Action</i>

3.2 构建事件本体模式

构建一个领域内的本体模式首先需要对具体领域事件本体中的事件类进行分析，然后对这些事件类进行抽象（见表 2 与表 3 中的抽象事件类列）。例如，事件类“动物死亡”（*Animal Death*）可以抽象为一个父事件类“生态破坏”（*Ecological Damage*）。如果不同本体存在共同的抽象类，那么相关的事件类、事件要素、事件关系则可以被用来构建本体模式。本文提出一种构建事件本体模式的算法，见表 3 算法。该算法描述了从多个特定的事件本体中抽取出抽象的事件类、事件要素（对象 *objects* 和施动者 *actors*）以及事件关系。在本体语言描述的事件类中，时间要素和地点要素总是被定义为一种非正式的表达式，或者引用外部本体词汇。例如，

事件类“河流污染”（*River Pollution*）的地点要素可以定义为一个由外部词汇而来的静态概念“水域地址”（*WaterAddress*），也可以用自然语言来进行描述，如“临近河流的区域”。事件的状态和动作总是定义为带约束的逻辑断言。例如，环境突发事件的前置状态被定义为 *discharged(pollutants, place)*。可见，时间、地点要素和事件类状态很难被机器自动识别和精确比较。所以，当创建一个事件本体模式时，通常需要对模式中事件类的时间、地点和状态要素进行人工提炼。

利用表 3 算法可以从空气污染事件本体和水污染事件本体中提取抽象事件类和通用的事件要素，如“污染者”（*polluters*），“污染物”（*pollutants*），“居民”（*residents*），“调查者”（*investigator*）等。

Table 3 Algorithm of event classes and elements extraction for event ontology pattern  
表 3 针对事件本体模式的事件类和要素提取算法

输入：事件本体集合 <i>eventOntology</i> 。
步骤 1：遍历 <i>eventOntology</i> 中所有的事件本体，找出存在于 <i>eventOntology</i> 中 所有事件本体的同名父事件类，将这些父事件类保存在 <i>abstractEventClass</i> 集合；
步骤 2：从 <i>abstractEventClass</i> 集合选择一个事件类 <i>abstractEventClassA</i> ，然后在所有的事件本体中，各自找出 <i>abstractEventClassA</i> 的子事件类保存在集合 <i>eventClassA</i> ，在 <i>eventClassA</i> 集合中查找这样的事件类，该事件类具有作为起始端相同的关联关系 <i>relationA</i> 而且关联的终端事件类具有同名的父事件类 <i>abstractEventClassB</i> 。则建立一个抽象关系 <i>abstractRelation</i> ，该关系与 <i>relationA</i> 类型相同，抽象关系 <i>abstractRelation</i> 的端点事件为 <i>parentEventClassA</i> ，终点事件类为 <i>parentEventClassB</i> 。将 <i>abstractRelation</i> 加入到 <i>abstractEventClassA</i> 的关系集合中。
步骤 3：遍历 <i>eventClassA</i> 中的事件类，找出存在于 <i>eventClassA</i> 中所有事件类的同名对象和同名参与者，保存在对象集合 <i>objectA</i> 和 <i>actorA</i> 。将对象集合 <i>objectA</i> 追加到 <i>abstractEventClass</i> 的对象集合， <i>actorA</i> 追加到 <i>abstractEventClass</i> 的参与者集合中。
步骤 4：重复步骤 2，直到 <i>abstractEventClass</i> 中所有的事件类都被处理完毕。
输出： <i>abstractEventClass</i> 集合。

利用算法 1 对表 1 和表 2 事件本体中的事件类进行抽象和提取，在此基础上结合人工提取抽象的事件时间、地点和状态，形成由抽象事件类构成的

环境突发事件本体模式，见图 4。通过对图 4 中的事件本体模式进行扩展、复用等，可以快速构建诸如城市雾霾、核泄漏污染等突发事件的本体模型。

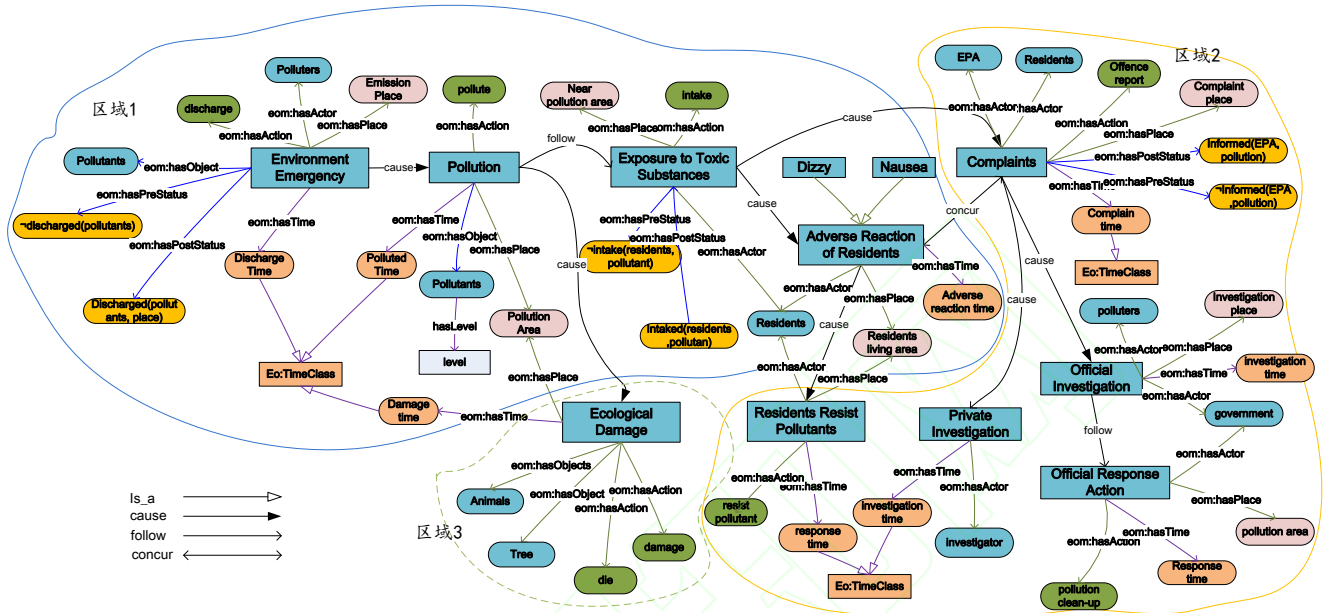


Fig.4 Event Ontology Pattern for Environmental Pollution Emergency

图4 环境污染突发事件本体模式

#### 4 事件本体模式的形式化

事件本体模式将具有共性的事件和事件类进行抽象,在特定领域建立通用的事件本体模型。在图4中的环境污染突发事件本体模式中,将所有相关的事件类参照各自所描述的对象分成了三个主题:污染突发事件主题(pollution incident subject, 区域1),社会效应事件主题(social event category, 区域2),自然效应事件主题(natural event category, 区域3)。每个部分包含了语义上相关的事件类。不同的事件类之间通过事件关系进行关联,从而形成了一个事件类网络。因此,可以在特定的领域使用通用的本体模式对不同事件本体进行构建。本章将分别阐述在本体模式中的事件类和事件要素的语义,并使用描述逻辑<sup>[13-15]</sup>对事件类和要素进行形式化。我们采用描述逻辑中的 $DLP_3$ 对事件及要素概念进行表示,表示方法支持可判定推理,可保证对本体模式中事件类和事件要素进行继承和实例化时概念的一致性。

污染突发事件主题包含了基本的事件类(*Environment Emergency*),并通过因果关系与社会效应事件主题(如投诉事件, *Complaint*)和自然效应事件主题(如树木损毁事件, *Tree Damaged*)两个部分进行关联。这部分是整个本体模式的起点和

主线,包含了两个重要的事件类:环境突发事件类(*Environment Emergency*)和污染事件类(*Pollution*),并且两个事件类之间构成了因果关系。这样的结构将“工厂排放废气(*Factory Emission*)”、“交通事故化学品泄漏(*Vehicle Chemical Leakage*)”等事件都归结成了这样一种统一的模型:一个诱因性质的突发事件与一个结果性质的污染事件。而不同污染事件的本体的差异特性可以反映到“污染”事件类的各个要素集合上去。具体地说,环境突发事件类包含了“排放”动作(*Emission*),动作的施动方“污染者”(*Pollutor*),和动作的客体对象“污染物”(*Pollutants*),以及时间和地点要素。这些事件要素就是区分不同污染突发事件或事件类的关键,同时也是形成统一模式的依据。

环境突发事件类(*Environment Emergency*)和污染事件类(*Pollution*)之间的因果关系可以形式化如下:

$$Environment\ Emergency \rightarrow Pollution \quad (1)$$

如果将事件类扩展到基本的事件要素,可以用如下规则进行表示:

$$Pollutors(x) \wedge Discharge(x, z) \wedge Pollutants(z) \quad (2)$$

$$\rightarrow Discharged(z, p) \wedge Place(p)$$

$$Discharged(z, p) \wedge Pollutants(z) \wedge Place(p) \quad (3)$$

$$\rightarrow Pollute(z, w) \wedge Area(w)$$



用描述逻辑  $DLP_3$  表示如下：

$$Pollutants \hat{=} Discharge^- . Pollutors \quad (4)$$

$$\hat{=} \exists Discharged . Place$$

$$Pollutants \sqsupseteq Discharged . Place \quad \exists Pollute . Area \quad (5)$$

其中，概念 *Pollutors* 表示排污者，概念 *Pollutants* 表示污染物，角色 *Discharge* 表示排放，概念 *Place* 代表事件的地点，角色 *Discharged* 代表污染物被排放在某地，概念 *Area* 表示地方区域，它与 *Place* 一样也是地点要素，很多时候其对应的实例是同一个。但有时 *Area* 表示的地点范围可能更大，因为污染影响的区域 *Area* 可能在范围上超过了原先排放污染物的地点 *Place*。

当创建一个具体的环境突发事件本体时，该本体中的事件类可从本体模式中对事件类继承获得。参考面向对象方法思想，继承是一种包含关系 (*is\_a*)。例如，环境突发事件 (*Environment Emergency*) 包含交通工具化学品泄露 (*Vehicle Chemical Leakage*)，因此其可以表示如下。

$$Vehicle \text{ Chemical Leakage} \hat{=} Environment \text{ Emergency} \quad (6)$$

如果扩展到具体本体中的事件要素，依然可以通过继承本体模式中对应的要素进行创建。这些事件要素被定义为概念和角色，能够采用一般化概念包含规则 (general concept inclusion, GCI) 进行继承。在图 3 所描述的交通工具化学品泄漏水域污染事件本体中，排污者是交通工具 (*Vehicle*)，污染物是化学品 (*Loaded Chemicals*)，而河流污染 (*River Pollution*) 是污染的一种。所以，可以表示如下。

$$River \text{ Pollution} \hat{=} Pollution \quad (7)$$

$$Vehicle \hat{=} Pollutors \quad (8)$$

$$LoadedChemical \hat{=} Pollutants \quad (9)$$

上述规则中概念之间的继承关系实现了事件要素之间的对应与关联。通过这种方法，将可以将本体模式的污染突发事件部分应用于该事件本体。

参照上述公理将事件类扩展到基本的事件元素，可以将该部分内容表示为如下规则：

$$Vehicle(x) \wedge Discharge(x, z) \wedge LoadedChemical(z) \quad (10)$$

$$\rightarrow Discharged(z, p) \wedge River(p)$$

$$Discharged(z, p) \wedge LoadedChemical(z) \wedge River(p) \quad (11)$$

$$\rightarrow Pollute(z, w) \wedge River(w)$$

公理 10 和公理 11 转换成描述逻辑表示如下：

$$LoadedChemical \hat{=} Discharge^- . Vehicle \quad (12)$$

$$\hat{=} \exists Discharged . River$$

$$LoadedChemical \hat{=} Discharged . River \quad (13)$$

$$\hat{=} \exists Pollute . River$$

污染突发事件部分的其他事件类和事件要素也可以通过同样的方法进行形式化。

图 4 中本体模式的社会效应部分包含了一系列因污染事件导致的人类活动，包括构造了“接触有毒物质” (*Exposure to Toxic Substances*)、“居民不良反应” (*Adverse Reaction of Residents*)、“抱怨” (*Complaints*)、“非官方调查” (*Private Investigation*)、“官方调查” (*Official Investigation*)、“官方回应行动” (*Official Investigation*) 多个事件类。不同事件类之间按照非分类关系进行连接：“接触有毒物质”事件类与“抱怨”和“居民消极反应”两个事件类之间都是因果关系，同时“抱怨”和“居民消极反应”两个事件类之间是并发关系。“抱怨”事件类与“非官方调查”、“官方调查”事件类之间分别是跟随关系和因果关系。“官方调查”与“官方回应行动”两个事件类之间是跟随关系。这几个事件类组成的关系网络，对于实现以统一的解决方案构造各类环境污染突发事件本体中的诸多社会效应事件类有着重要的借鉴意义。

$$Exposure \text{ to Toxic Substances} \rightarrow Complaints \quad (14)$$

$$Complaints \rightarrow Official \text{ Investigation} \quad (15)$$

$$Complaints \rightarrow Private \text{ Investigation} \quad (16)$$

$$Official \text{ Investigation} > Official \text{ Response Action} \quad (17)$$

$$Exposure \text{ to Toxic Substances} \rightarrow \quad (18)$$

$$Adverse \text{ Reaction of Residents}$$

$$Adverse \text{ Reaction of Residents} \rightarrow \quad (19)$$

$$Residents \text{ Resist Pollutants}$$

上述公理是基于事件元素的规则描述。事件类“接触有毒物质” (*Exposure to Toxic Substances*) 指的是居民在污染物排放地点接触到了其中的有毒物质。事件类“投诉” (*Complaints*) 描述的是居民因污染事件向环保部门投诉，并且环保部门了解到了该事件。所以，公理 14 可以扩展为以下程序：

$$Pollutants(z) \wedge Pollute(z, w) \wedge Area(w) \wedge Residents(r) \quad (20)$$

$$\wedge LocateIn(r, w) \rightarrow Intaked(r, z)$$

$$Intaked(r, z) \wedge Pollutants(z) \wedge Residents(r) \wedge \quad (21)$$

$$OffenceReport(r, x) \wedge EPA(x) \rightarrow Informed(x, z)$$

公理 20 和公理 21 采用描述逻辑可以表示如下：

$$\begin{aligned} \exists \text{Area} \circ \text{LocatedIn}^- . \text{Residents} \hat{=} \\ \text{Area} \circ \text{Pollute}^- . \text{Pollutants} \end{aligned} \quad (22)$$

$$\text{LocatedIn} \circ \text{Pollute}^- \hat{=} \text{Intaked} \quad (23)$$

$$\begin{aligned} \exists \text{Residents} \circ \text{OffenceReport} . \text{EPA} \hat{=} \\ \text{Residents} \circ \text{Intaked} . \text{Pollutants} \\ \text{OffenceReport}^- \circ \text{Intaked} \hat{=} \text{Informed} \end{aligned} \quad (24)$$

其中, 概念 *Pollutants* 代表环境突发事件中被排放的污染物。概念 *Area* 指的是事件发生的区域。角色 *Pollute* 描述的是某物污染某地的动作要素, 关联了概念 *Pollutants* 和概念 *Area* 的实例。角色 *LocatedIn* 表示的是某人居住在特定的区域, 该动作要素关联了参与者概念 *Residents* 和地点概念 *Area*。角色 *Intaked* 指的是某人或动物吸入某物质。角色 *OffenceReport* 代表特定人群 (如居民 *Residents*) 对某对象的投诉并被公开报道。概念 *EPA* 指的是环保部门。

公理 15 描述的是居民“投诉” (*Complaints*) 导致了“官方调查” (*Official Investigation*), 两个事件类之间属于因果关系。通过将事件类扩展到事件要素的各个状态上, 可以生成如下规则:

$$\begin{aligned} \text{EPA}(x) \wedge \text{Informed}(x, z) \wedge \text{Pollutants}(z) \rightarrow \\ \text{GovernmentDepartment}(x) \wedge \\ \text{Investigate}(x, p) \wedge \text{Pollutors}(p) \end{aligned} \quad (26)$$

公理 26 采用描述逻辑语法规则可以表示为:

$$\begin{aligned} \text{EPA} \hat{=} \text{Informed} . \text{Pollutants} \hat{=} \\ \exists \text{GovernmentDepartment} \hat{=} \text{Investigate} . \text{Pollutors} \end{aligned} \quad (27)$$

其中, 角色 *Informed* 表示某人被告知了某事。角色 *Investigate* 表示环保部门对污染者进行调查。公理 16 的形式化方法与公理 15 相似, 这里不再赘述。

公理 17 描述的是官方即政府有关部门在对污染事件调查之后作出反应并采取行动 (如清除污染物)。公理 17 的事件类可以扩展为如下规则:

$$\begin{aligned} \text{GovernmentDepartment}(x) \wedge \text{Investigate}(x, p) \wedge \\ \text{Pollutors}(p) \rightarrow \text{EmergencyResponseDepartment}(y) \\ \wedge \text{Clean-up}(y, q) \wedge \text{Pollutants}(q) \end{aligned} \quad (28)$$

转换成描述逻辑可表示如下:

$$\begin{aligned} \text{GovernmentDepartment} \hat{=} \text{Investigate} . \text{Pollutors} \hat{=} \\ \hat{=} \exists \text{EmergencyResponseDepartment} \\ \hat{=} \text{Clean-up} . \text{Pollutants} \end{aligned} \quad (29)$$

由于篇幅所限, 本文不再对本体模式中社会效

应部分的其他事件类的信息进行具体说明。

环境污染突发事件本体模式中的自然效应部分描述的是一系列由污染引发的自然事件。本文在该本体模式中定义了一个通用的自然事件类 *Ecological Damage*, 表示动物或植物的死亡或者受到损害。这里以“动物因污染死亡”为例, 将其形式化后的规则表示如下:

$$\begin{aligned} \text{Pollute}(z, w) \wedge \text{Area}(w) \wedge \text{Pollutants}(z) \rightarrow \\ \text{Animal}(x) \wedge \text{LiveIn}(x, w) \wedge \text{Died}(x) \end{aligned} \quad (30)$$

采用描述逻辑  $\text{DLP}_3$  进行表示, 可以得到如下公理:

$$\text{Area} \hat{=} \text{Pollute}^- . \text{Pollutants} \hat{=} \text{LiveIn}^- . \text{Animal} \quad (31)$$

$$\text{Animal} \hat{=} \exists \text{Died} \quad (32)$$

其中, *Pollutants*, *Animal*, *Area* 和 *Died* 都是概念。角色 *Pollute* 描述的是某物污染某地的动作要素, 关联了概念 *Pollutants* 和概念 *Area* 的实例。角色 *LiveIn* 表示某些动物在某地生活或栖息。

上述公理和形式化方法可以应用于创建新的具体事件本体。例如, 事件本体“鱼类因河流污染而死亡”可以参照本体模式的这部分内容进行形式化。

$$\begin{aligned} \text{Pollute}(z, w) \wedge \text{River}(w) \wedge \text{ToxicSubstance}(z) \rightarrow \\ \text{Fish}(x) \wedge \text{LiveIn}(x, w) \wedge \text{Died}(x) \end{aligned} \quad (33)$$

采用描述逻辑  $\text{DLP}_3$  进行表示, 得到下公理:

$$\text{River} \hat{=} \text{Pollute}^- . \text{ToxicSubstance} \hat{=} \text{LiveIn}^- . \text{Fish} \quad (34)$$

$$\text{Fish} \hat{=} \exists \text{Died} \quad (35)$$

$$\text{River} \hat{=} \text{Area} \quad (36)$$

$$\text{ToxicSubstance} \hat{=} \text{Pollutants} \quad (37)$$

其中, *River*、*ToxicSubstance* 和 *Fish* 都是概念。

此外, 利用上述所定义的本体模式中相应的事件类和事件要素概念可以对文本中的事件进行实例化和形式化表示。例如, “位于台北市地铁市府站旁的中油直营加油站, 日前遭台北市环保局稽查发现, 地下水每公升含‘甲基第三丁基醚’ *MTBE* 达 3.48 毫克, 超过管制标准。今早当地兴雅里民齐聚加油站抗议, 要求加油站搬迁”。

这里仅抽取一个事件的逻辑程序公理以说明。

$$\text{Area} \hat{=} \text{Pollute}^- . \text{Pollutants} \hat{=} \text{LiveIn}^- . \text{Animal} \quad (38)$$

$$\text{People} \hat{=} \text{Animal} \quad (39)$$

将文本示例中的事件要素实例抽取出来, *a* 为地点要素“台北市地铁市府站旁的中油直营加油

站”；*mtbe* 为污染物实例“甲基第三丁基醚”；*p* 为市民实例。存在以下断言： $Area(a), Pollutants(mtbe), People(p), LiveIn(p,a)$ 。令解释  $I = (\Delta^I, \cdot^I)$ ，其中非空概念集合  $\Delta^I$  和函数  $\cdot^I$  构成的二元组， $\Delta^I$  为  $I$  的论域，则可以得到：

$$Area^I \cap (Pollute^I.Pollutants^I) \subseteq (LiveIn^I.Animal^I) \quad (40)$$

$$People^I \subseteq Animal^I, a^I \in Area^I, \quad (41)$$

$$mtbe^I \in Pollutants^I, p^I \in People^I$$

本节通过将事件中的不同要素类抽象为描述逻辑中的概念和角色，并将事件关系中的要素联系通过形式化公理进行表示，形成能够有效概括此类事件的通用模式。结合一些传统本体中的概念层次，可以将概念一致的类和相应的实例应用到该模式当中，见公理 (1-32)。最后通过事件本体中的概念构建和文本中事件实例的描述验证了事件本体模式在构建新的本体（包括概念和实例）过程中的有效性和概念一致性。

## 5 结束语

本体模式作为解决本体复用等问题的解决方案被提出<sup>[16]</sup>，获得了越来越多的研究人员关注。目前，本体模式已经被应用到了医药信息<sup>[11,17]</sup>、地理信息系统<sup>[18-20]</sup>、生物科学<sup>[21]</sup>等各个领域。但是针对基于事件的本体模式研究，国内外研究仍处与探索和起步阶段。由于人类认知的差异性，针对不同应用领域进行事件本体建模具有很高的难度。本文提出了一种具有一般性的事件模型结构和针对 Web 事件知识进行表示和推理的事件本体模型。该事件结构将不同类型的离散的事件要素有机的结合在一起，建立起事件和事件要素的语义联系，形成一个统一的事件知识表示单元；同时利用事件状态这一要素实现了对事件发生时动态过程的表示；五种事件关系可以较完整地描述事件或事件类之间的语义关系。此外，提出了一种应用于环境污染突发事件领域的事件本体模式，用以实现在该领域进行类似于面向对象设计模式的可重用的本体设计方案。该模式能够提高本体建模的效率，降低构建事件本体的难度。基于  $DLP_3$  的模式形式化方法，简化了事件之间复杂的语义关系，使模式中的事件类和要素等概念支持可判定推理。然而，本文所涉及的研究工作仍然存在局限性，如缺乏标准的事件本体语言，以及对事件间的非分类关系进行可判定推理等。这

些工作都需要在今后做进一步的研究。同时，如何在构建事件本体模式时对可重用性和有效性之间进行合理的权衡也是未来需要研究的内容。

## References:

- [1] Liao zhenliang, Liu yanhui, Xu zuxin. Environmental pollution emergency response plan system based on case reasoning[J]. Environmental Pollution and Control, 2009, 31:86-89
- [2] Shao Quan, Wengwen Guo, He Changhong, et al. Model hierarchical network representation method in emergencies model base [J], Journal of Tsinghua University: Natural Science Edition, 2009, 5:625-628
- [3] S. H. Jihan and Segev A., Humanitarian Assistance Ontology for Emergency Disaster Response[J], IEEE Intelligent Systems, 2014, 29(3):6-13
- [4] Wang Tao, Yan-Zhang, Lu Yanxia. Research on public emergencies meta-event model based on ontology[J]. Journal of Dalian University of Technology, 2012, 52: 458-463
- [5] Hage, Willem Robert, et al. Abstracting and reasoning over ship trajectories and web data with the Simple Event Model (SEM)[J]. Multimedia Tools & Applications, 2012, 57(1):175-197
- [6] Raimond Y, Abdallah S, Sandler M, et al. The music ontology[C]// Proceedings of the 8th International Conference on Music Information Retrieval. ACM, 2007: vol.[S.I.], 417-422.
- [7] Scherp, Ansgar, et al. F-a model of events based on the foundational ontology dolce+ DnS ultralight[C]//Proceedings of the 5th international conference on Knowledge capture, ACM, 2009: 137-144.
- [8] Shaw, Ryan, R. Troncy, and L. Hardman. LOD: linking open descriptions of events[J]. Lecture Notes in Computer Science, Springer, Heidelberg, 2009, 5926:153-167.
- [9] Di L, Zhao P. Geospatial semantic web, interoperability[J]. Encyclopedia of GIS, 2008, pp.70-77.
- [10] Daga, E., Blomqvist, E., Gangemi, A., et al. NeOn D2.5.2 Pattern based ontology design: methodology and software support. NeOn project. <http://www.neon-project.org>, 2010
- [11] Mortensen J M, Horridge M, Musen M A, et al. Applications of ontology design patterns in biomedical ontologies[C]//AMIA Annual Symposium Proceedings, Chicago, 2012, pp. 643-652,
- [12] W. Liu, et al., Extending OWL for Modeling Event-oriented Ontology[C]//Proceeding of 4th International Conference on Complex, Intelligent and Software Intensive Systems, IEEE Computer Society, 2010: 581-586
- [13] Straccia U. A fuzzy description logic for the semantic web[J]. Capturing Intelligence, 2006, 1:73-90.
- [14] Gasse F, Sattler U, Haarslev V. Rewriting rules into SROIQ axioms\\Poster at 21st International Workshop on Description Logics, 2008.
- [15] Krotzsch M, Rudolph S, Hitzler P. ELP: Tractable rules for OWL 2[C]// Lecture Notes in Computer Science 5318: Proceedings of the 7th International Semantic Web Conference (ISWC-08), Springer, 2008: 649-664
- [16] Gangemi A. Ontology Design Patterns for Semantic Web Content[C]// Lecture Notes in Computer Scienc 3729: Proceedings of the 4th International Semantic Web Conference, 2005: 262-276.
- [17] Mart íez-Costa, C., Karlsson, D., Schulz, S.: Ontology Patterns for Clinical Information Modelling[C]//Proceeding



- of 5th Workshop on Ontology and Semantic Web Patterns, 2014: 61-72
- [18] Yingjie Hu, Krzysztof Janowicz, David Carral, et al. A Geo-Ontology design pattern for semantic trajectories[C] \\\ Lecture Notes in Computer Science 8116: Proceedings of 11th International Conference On Spatial Information Theory, Springer, 2013: 438-456.
- [19] Carral, D., Scheider, S., Janowicz, K., et al. An ontology design pattern for cartographic map scaling[C] \\\ Lecture Notes in Computer Science 7882: Proceedings 10th ESWC2013, Montpellier, France, May 2013. Springer, 2013: 76-93.
- [20] Carral, D., Janowicz, K., Hitzler, P.: A logical geo-ontology design pattern for quantifying over types[C] \\\ Proceeding of SIGSPATIAL'12, 2012: 239-248.
- [21] Aranguren M E, Antezana E, Kuiper M, et al. Ontology design patterns for bio-ontologies: a case study on the cell cycle ontology[J]. BMC Bioinformatics, 2008, 9(1): 279-282.

### 附中文参考文献:

- [1] 廖振良, 刘宴辉, 徐祖信. 基于案例推理的突发性环境污染事件应急预案系统[J]. 环境污染与防治, 2009, 31: 86-89.
- [2] 邵荃, 翁文国, 何长虹, 等. 突发事件模型库中模型的层次网络表示方法[J]. 清华大学学报: 自然科学版, 2009, 5: 625-628.
- [4] 王涛, 王延章, 鲁艳霞. 突发公共事件基于本体的元事件模型研究[J]. 大连理工大学学报, 2012, 52: 458-463.



LIU Wei was born in 1978. He received the Ph.D. degree from Shanghai University in 2005. Now he is an associate professor at Shanghai University. His research interests include knowledge representation and reasoning, semantic network and ontology technologies, etc.

刘炜(1978-), 男, 江西赣州人, 2005年于上海大学获得博士学位, 现为上海大学计算机工程与科学学院副研究员, CCF 会员, 主要研究领域为知识表示与推理, 语义网与本体技术等。发表学术论文 40 余篇, 主持国家自然科学基金项目 1 项、上海市自然科学基金项目 1 项, 以主要成员参与国家自然科学基金项目 4 项。



DING Ning was born in 1986. He is a master candidate at Shanghai University. His research interests include knowledge representation and machine learning, etc.

丁宁(1986-), 男, 上海人, 上海大学计算机工程与科学学院硕士研究生, 主要研究领域为知识表示, 机器学习等。



YANG Junhui was born in 1981. He is a Ph.D. candidate at Shanghai University. His research interests include knowledge representation, natural language processing and web data mining, etc.

杨竣辉(1981-), 男, 江西赣州人, 上海大学计算机工程与科学学院博士研究生, 主要研究领域为知识表示、自然语言处理和 Web 数据挖掘等。发表学术论文 10 余篇, 主持省级课题 2 项, 以主要成员参与国家自然科学基金 2 项。



WANG Xu was born in 1989. He is a master candidate at Shanghai University. His research interests include natural language processing and machine learning, etc.

王旭(1989-), 男, 江苏睢宁人, 上海大学计算机工程与科学学院硕士研究生, 主要研究领域为自然语言处理与机器学习。



ZHANG Yujia was born in 1992. She is a master candidate at Shanghai University. Her research interests include knowledge representation and machine learning, etc.

张雨嘉(1992-), 女, 浙江杭州人, 上海大学计算机工程与科学学院硕士研究生, 主要研究领域为知识表示, 机器学习等。



LIU Zongtian was born in 1946. He received the M.S. degree from Beijing University of Aeronautics and Astronautics in 1982. Now he is a professor and Ph.D. supervisor at Shanghai University. His research interests include artificial intelligence and software engineering, etc.

刘宗田(1946-), 男, 1982 年于北京航空航天大学计算机科学与工程系获得硕士学位, 现为上海大学计算机工程与科学学院教授、博士生导师, 主要研究领域为人工智能, 软件工程等。