

# 学习资源的语义众包标注系统设计\*

丁国柱<sup>1,2</sup>, 余胜泉<sup>1,2</sup>, 潘升<sup>1,2</sup>

(1.北京师范大学 教育学部 教育技术学院, 北京 100875; 2. “移动学习”教育部—中国移动联合实验室, 北京 100875)

**摘要:** 语义标注是Web3.0环境下学习资源组织方式的基础和前提。目前学习资源的语义标注多是通过本体对文档的实体进行识别和标注, 一般揭示的是显性的、有确定规则的知识, 对人文性学科中蕴含情境性、文化性特征可多元解释的隐性知识揭示不足, 针对目前这个问题, 该文提出基于众包的学习资源语义标记的思路, 充分利用人的集体智慧, 解决海量学习资源的复杂多元语义标注问题。研究以学习元平台为支撑, 构建基于众包的语义标注系统, 通过社会信任计算和关联词汇判别过滤机制进行质量控制, 实现学习资源中隐性知识的标注, 通过实验证明基于众包能够有效解决人文性学科中的隐性知识标注。

**关键词:** 语义标注; 众包; 学科资源; 学习元

**中图分类号:** G434 **文献标识码:** A

## 一、引言

语义标注是指为文档添加规范化知识表示的过程<sup>[1]</sup>, 是实现文档机器可读的重要途径, 也是学习资源语义检索、资源库转向知识服务等应用的关键环节。虽然目前已有不少机器自动处理的办法来实现语义标注, 但一些学科特别是人文社会类学科的实例具有动态变化的特点, 一些知识实例存在的隐性关联是机器难以自动处理的, 如语文学科中的人文性知识是动态发展的, 而且具有一定的主观性, 单纯靠机器来实现语义标注比较困难, 特别是隐性知识存在情境性, 受文化、民族传统甚至宗教的影响, 机器标注的正确率低; 而依赖少数专家的人工标注则存在巨大的人力和时间成本, 需要其他方式来解决这个问题。众包利用互联网将工作分配出去<sup>[2]</sup>, 利用群体的智慧完成工程巨大的任务。利用众包来解决机器难以完成的部分学习资源的语义标注是一个值得研究的课题。

## 二、研究现状

语义标注的本质是实体识别和实体标注, 即基于本体对文档中的实体进行提取, 并对实体进行标注, 从而使得计算机可以理解文档的内容。目前语义标注的方法主要有自动化标注、半自动标注、

手工标等类型<sup>[3]</sup>, 自动标注有基于机器学习的方式(Machine Learning-based)、基于模式的方式(Pattern-based)和综合方式(Multi Strategy, 结合机器学习和模式方式); 半自动标注结合了自动标注和手工标注两种方式, 有助于提高标注的准确性。标注工具方面, 手工工具有SHOE Knowledge Annotation; 半自动工具有SMORE、MnM、Melita、Semantic Word等, 自动化工具有AeroDAML等<sup>[4]</sup>。总的来看, 语义标注包括如下几个方面: (1)实体标注, 实体标注是指将某个实体指向为本体中的某个概念(或称为类), 即建立实例和概念的关联, 其形式化表示为(I, rdf:type, C), 其中I表示实例, C表示概念; (2)实体关联, 实体关联则是基于本体中的对象属性建立2个实例的关系, 形式化表示为(I1, R, I2), 其中I1、I2为2个不同的实例, R为对象属性; (3)实体属性标注, 即标注一个实体的XML数据类型信息, 如人的出生日期、身高等, 形式化表示为(I1, R, d), 其中I1为实体, R为数据属性, d则是XML数据类型的相关数据; (4)新增概念, 当一个实例要标注的概念系统未记录时, 这个概念就是新增的概念, 可以纳入本体中; (5)新增概念关系, 即增加本体中的对象属性; (6)新增概念属性, 即增加本体中的数据属性。

\* 本文受北京市未来教育高精尖创新中心项目“智慧学伴平台设计与开发”(项目编号:BJAICFFE2016DR-001)资助。

对于学习资源而言,合理的语义标注可以实现资源聚合、语义检索、资源推荐等应用,然而当前语义标注方法多是针对显性知识进行标注,对隐性知识的标注能力不强,如“廉颇老矣”简单四字,却涉及了历史、情感和隐喻,机器往往难以标注全面,因此现实中很大一部分学习资源的标注还是要依赖于人工。

基于众包对信息进行标注可以追溯到Web2.0时代,如图片分享网站Flickr的Folksonomy,用户群体将照片贴上标签并且以此来组织图片资源;众包的标志性成果则是路易斯·冯·安开发的reCAPTCHA,完成纸质书的数字化<sup>[5]</sup>;利用同样的原理Sophia等开发faceCAPTCHA,完成性别标注<sup>[6]</sup>;研究者设计了一款基于众包的游戏Foldit,仅用了三周时间,玩家就解决了一个艾滋病方面的难题<sup>[7]</sup>,而这个难题已经困扰了科学家15年;也有学者利用众包来构建情感词汇<sup>[8]</sup>,构建词汇联想网络<sup>[9]</sup>等。此外亚马逊、爱立信、华为等500强企业也以各种方式利用众包,让群体智慧为自己的企业解决部分难题。从已有的研究来看,众包至少有3个特点:

#### (1)问题转化:

将机器难以处理的任务转化成人类可以完成的任务;(2)化整为零:将一个巨大的任务分成数量较大的小任务,这些小任务分配下来后用户可以较轻松完成;(3)聚合群体智慧:对于每一个任务而言,互联网中个体的简单行为(如创建、投票、浏览)可以看成是一个微小N维向量,一定数量用户的行为产生的向量之和则可以产生足够的影响力;即单独个体的行为与智能简单有限,但相当数量的个体行为相互作用可以形成巨大的智慧。

从众包的特点来,众包可以充分利用人本计算(Human-based Computation),将人视为计算资源,即基于网络的分众性和协同性,利用人的智慧来完成计算机难以做好的事情。因此利用众包来完成学

习资源语义标注,可以解决学习资源隐性知识的语义标注问题。

### 三、基于众包的学习资源语义标注模型

建构主义认为知识的获取过程来自于他人(包括教师和学习伙伴)的帮助,他人知识的显性化可以促进个体的认知。学习资源的标注可以看成是个体知识显性化的过程。当标注活动成为群体行为时,可产生知识共享与协同知识建构,与此同时也解决了学习资源的语义标注问题。具体实现步骤如下:(1)任务分配:根据学习与用户特征将任务分配下去;(2)用户参与机制:学习管理系统的目标是为了促进用户的学习,因此应该设计合理的机制,使得用户在完成学习任务的过程中自然地完成任务;(3)群体智慧的聚合:利用聚合算法实现用户参与结果计算,使得大量用户相对无序的工作变成智慧的产物。基于上述步骤,本研究构建了基于众包的语义标注模型,如图1所示。基于众包的学习资源语义标注主要包括任务分配器、语义标注工具、众包质量控制机制三个部分。

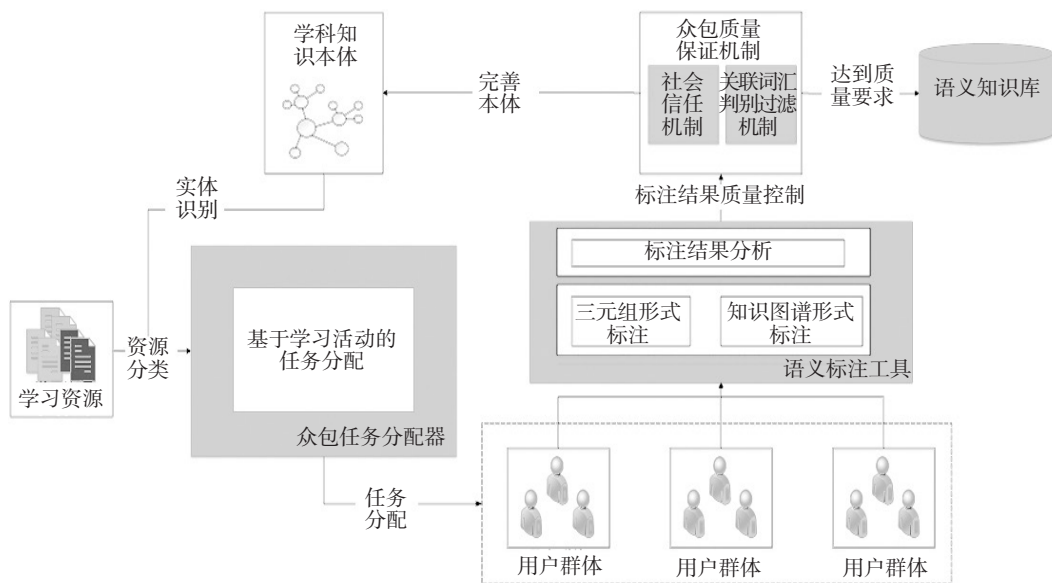


图1 基于众包的语义标注模型

#### (一)众包任务分配器

“学习活动”通过支持教学方法、教学策略、教学活动的重用,在教学层面上实现了教育过程与活动的高水平共享<sup>[10][11]</sup>。网络学习环境下,教学目标的达成可以由一系列的学习活动组成,通过恰当的教学策略与方法,设计和实施有效的学习活动,促进学生与教师之间、学生与学生之间的交流,使得学生积极地投入到网络学习中,实现知识的习得

与内化。因此“活动”是学习设计的重要载体，如课堂/小组讨论、问题解决、角色扮演等，可以将学习资源的语义标注任务转化为学习活动，既促进学习者的学习，也实现学习资源的语义标注。语文学习中大量的知识需要学生进行构建和关联，通过学习活动让学生建立相关的知识关联，一方面可以促进学生的知识内化，教师也可以发现学生的认知缺陷，进而有针对性地指导学生；另一方面也可以利用学生的智慧完善语义标注。如图2所示，教师发布语义标注的学习活动，学生参与学习学动，进而构建更完整的知识体系，而教师也可以根据学生完成情况进行个性化指导，基于学习活动的任务分配可以起到一举多得的效果。

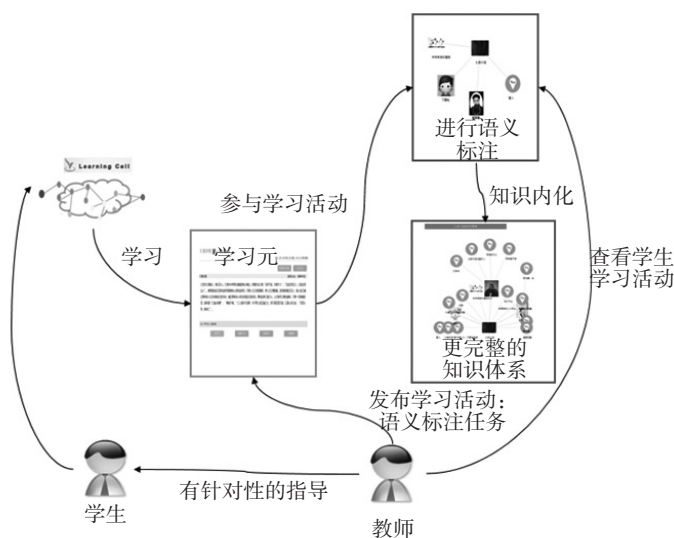


图2 基于学习活动的任务分配

## (二)用户参与机制设计

当一个语义标注任务分配下来后，相关群体需要一个适合的语义标注工具来完成任务。考虑到参与人员大部分是普通的教师，用户对语义这个概念的了解程度比较有限，可以利用一些辅助工具来降低语义标注的难度。语义标注包括的实体标注、实体关联、实体属性标注、新增概念、新增概念关系、新增概念属性等内容，其本质来看都是描述事物A、事物B和2者关系R的三元组，标记为<A, R, B>，故最简单的办法可能是让用户直接填充这个三元组，然后再通过算法识别是哪一个类型的语义标注。因此一个完整的语义标注工具包括了用户前端标注和标注后端分析。

用户前端标注，本文使用了两种用户标注方法，一是直接填充三元组法：通过三个文本输入框，让用户直接填充知识，如右图3所示；二是构建个人知识图谱法：通过图像界面，显示知识点与

知识点之间的关系，通过构建个人知识图谱完善语义标注。如图4所示。

编号	知识点	知识关系	知识点	关联人	投票
1	周恩来	听说	中华不振	G6_王忱哲	已进化
2	周恩来	同时代伟人	毛泽东	G6_王忱哲	已进化
3	周恩来	人物评价	政治家	李威	已进化
4	为中华之崛起而读书	相关背景	中华不振	潘升	已进化
5	周恩来	相关文章	十里长街送总理	李威	已进化
6	帝国主义	侵略	中国	G6_王忱哲	已进化
7	周恩来	名人名言	为中华之崛起而读书	李威	已进化
8	周恩来	妻子	邓颖超	李威	已进化
9	周恩来	读书	东关模范学校	G6_王忱哲	已进化
10	周恩来	相关事件	十里长街送总理	潘升	已进化

共28条数据, 3页 &lt; 23 &gt; 尾页

图3 利用直接填充三元组法实现语义标注



图4 利用知识图谱实现语义标注

用户完成语义标注后，具体标注类型还是未知的，系统可以通过算法进行分类。具体算法如下：

(1)一个语义标注<A, R, B>，若B为概念，R为是、是一个、为、是一类、类型是、类型等类别词汇，同时A不属于一个概念时，则<A, R, B>为实体标注。如标注<刘禹锡，是，唐朝诗人>，唐朝诗人在系统中已作为一个关键概念存在，而刘禹锡不是概念，则这个语义标注为实体标注，将刘禹锡标注为一个实体，并作为唐朝诗人的一个实例。

(2)一个语义标注<A, R, B>，若A和B都不是概念，同时A和B是分别或者同时是某个概念的实例，则<A, R, B>为实体关联。如标注<李白，创作，静夜思>，李白和静夜思都不是概念，而分别属于唐朝诗人和唐朝诗词的实例，那么这个标注为实体关联。

(3)一个语义标注<A, R, B>，若A为某个概念的实例，B既不是实例也不是一个概念，则<A,



R, B>为实体属性标注。如<杜甫, 字, 子美>, 杜甫为唐朝诗人的实例, 子美在系统中不是实例也不是一个概念, 那么这个标注可以视为属性标注, 即将“字”作为杜甫的一个属性。

(4)一个语义标注<A, R, B>, 若A为某个概念的实例或者A不是实例也不是一个概念, R为是、是一个、为、是一类、类型是、类型等类别词汇, B在系统中没有对应的概念, 则<A, R, B>为新增概念。如标注<杜甫, 是, 现实主义诗人>, 已知杜甫是唐朝诗人的实例, 而现实主义诗人在系统中没有作为一个概念存在, 现实主义诗人就作为一个新增的概念加入系统。

(5)一个语义标注<A, R, B>, 若A是一个概念, B在系统中没有对应的概念, 同时B有其他关系连接, 则<A, R, B>为新增概念关系。如标注<司马迁, 创作, 史记>, 系统中<司马迁, 是, 史学家>的记录, 同时也有<史记, 被誉为, “史家之绝唱, 无韵之离骚”>, 此时可以把“创作”这个作为史学家的一个概念关系。

(6)一个语义标注<A, R, B>, 若A是一个概念, B在系统中没有对应的概念, 同时B没有其他关系连接, 则<A, R, B>为新增概念属性。如标注<卫青, 字, 仲卿>, 卫青是西汉大将军, 而仲卿则未有其他节点与之关联, 则可以将“字”视为大将军的一个属性。

### (三)群体智慧的聚合

互联网环境下, 参与众包的用户量巨大, 需要有效的汇聚群体智慧的机制。本文参考王蓉博士的群体认知聚合模型算法<sup>[12]</sup>, 从两个方面进行语义标注聚合计算: 一是个体标注的叠加效果计算, 如多个用户都标注了相同的语义信息; 二是群体对语义标注聚合结果实施的行为表现出来的认可程度, 如浏览、投票次数等, 对每个标注 $W_i$ 聚合算法如下:

$$W_{it} = dW_{it-1} + R + yV + zL \quad (1)$$

其中:  $W_{it}$ 为 $t$ 时刻标注 $W$ 的权重;  $W_{it-1}$ 为 $t-1$ 时刻标注 $W$ 的权重;  $d$ 为时间衰减因子;  $R$ 为当前标注的可信度;  $y$ 为投票权重系数;  $V$ 为对 $W$ 的投票人数;  $z$ 为浏览权重系数;  $L$ 为浏览过 $W$ 的人数。

根据算法, 用户信任评价、投票、浏览参数的设定决定了群体智慧的价值取向和群体智慧进化方向, 当一个学习资源的标注的聚合程度超过一定阈值时, 可以认为围绕这个学习资源形成了一个群体认知。其中 $R$ 表示的标注可信度是能否有效促进学习资源语义众包的关键, 本文从社会信任评估和机器判断2个方面进行可信度计算, 公式如下:

$$F(x) = f(x) + g(x) \quad (2)$$

基于社会信任评估机制。利用开放知识社区中的信任评估模型<sup>[13]</sup>, 系统为每个用户建立信任评估值 $UT$ , 对于对每一条语义标注信息, 用户可以投赞成或者反对票。算法如下:

$$f(x) = \sum_{n=1}^n K * UT_n \quad (3)$$

其中,  $f(x)$ 为具体一条语义标注信息可信度,  $UT_n$ 为具体投票用户的可信度,  $K$ 表示用户的意见, 投赞成票时为1, 反对票为-1。当 $f(x)$ 值超过系统预设的阈值时, 可以认为这条语义标注信息可信, 可以存入语义数据库。

基于Word2vec进行关联词匹配判断。语义标注本质上是把学习资源中的关键知识通过三元组进行表示, 每一个三元组都可以看成是描述主语和宾语之间的关联, 理论上主语和宾语之间存在一定的共现概率。当一个新的三元组被录入时, 可以通过分析和主语一起相关词汇, 按照关联度可以帮助判断这个语义标注的准确性。Word2vec是Google在2013年发布基于深度学习的自然语言处理开源工具<sup>[14]</sup>, Word2vec可以将词汇表征为向量, 核心思想是将语料库作为训练集, 把对文本内容的处理简化为 $K$ 维向量空间中的向量运算, 因此可以利用word2vec来计算词与词之间的相似度, 进而实现语义标注的质量控制。算法流程如下表所示。

基于深度学习的本体进化质量控制算法表

输入: 众包生成的三元组<S, O, P>  
输出: 实例三元组<S, O, P>的可信度 $x$

关键步骤:  
Step1 对每一个众包生成的三元组<S, O, P>, 利用Word2vec得到S相关词汇集 $W: \{[w_1, d_1], [w_2, d_2], \dots, [w_n, d_n]\}$   
Step2 匹配P是否在相关词汇集 $W$ 中, 如果存在, 则输出对应词汇 $w$ 的相似度值 $g(x)=d$ , 否则 $g(x)=0$

## 四、基于学习元平台的语义众包实验

为验证系统对特定学科语文语义标注效果, 以学习元平台(Learning Cell System, 简称LCS)<sup>[15]</sup>作为支撑, 基于人教版初中语文课本中的《孔子语录十则》, 要求学生完善儒家相关知识, 验证结合学科知识本体和众包机制对学习资源进行语义标注的质量, 运行时间为1个月。

### (一)实验对象

本次实验选取某初级中学七年级学生43人作为实验对象, 实验前已对学生进行了语义标注的相关操作培训。

## (二)实验过程

### 1.任务描述

《孔子语录十则》的教学目标之一是“积累成语、格言、警句等，了解儒家相关诸子百家信息，并能用儒家的思想指导自己的言行”。这节课含有较多的人文性知识——儒家思想是中国传统文化的核心之一，从孔子和儒家展开，可以呈现波澜壮阔的中国传统文化。学生通过了解儒家相关的知识，进行语义标注，实现知识共享与构建。

### 2.实验过程

(1)以人教版初中七年级语文课程《孔子语录十则》为基础，学生学习相关学习资源，包括中国传统文化、儒家经典欣赏、百家争鸣和儒家思想形成等；

(2)教师基于教学目标，针对每个学习资源发布语义标注任务；

(3)学生在学习元平台上阅读相关课程内容；

(4)学生完成语义标注活动。

### 3.实验结果分析

(1)语义标注结果分析：学生使用三元组标注和构建个人知识图谱法，一共生成了134条三元组；系统对语义标注的情况进行质量控制，包括用户投票和利用Word2vec进行机器判断，其中一共有6833条用户投票记录，其中赞成票6703条，赞成比例为98%，反对票130条，反对比例为2%；Word2vec方面针对134条三元组，一共生成1356条候选词汇，其中有236条记录关联词汇被确认相关；最终有126条记录存储进了语义数据库。

(2)隐性知识标注分析，134条标注中有48条人文性知识，集中在儒家相关名言上，如<儒家，名言，“三人行，必有我师”><“三人行，必有我师”，表示，要善于向其他学习>，不单揭示围绕《论语十则》的隐性知识，还表现出学生对知识的理解。

## 六、结论与展望

学习资源的语义标注是构建学科知识库的重要环节，本研究提出结合学科知识本体和众包解决学习资源语义标注问题的方法，本方法可以利用人类计算(Human Computation)解决学科知识中的蕴涵的大规模隐性知识的语义标注问题，为深层次的学科知识库构建提供了一种新思路。本研究不足之处是体智慧汇聚还需探讨更有效的算法，以确保群体智慧可以有效汇聚，同时在调动用户参与众包上的激励机制还可改进，如何设计高效有用的用户激励机制是众包成功的关键。因此，本文的后续工作是针

对学习社区群体，研究更有效的学习者智慧汇集算法，设计有吸引力同时相对低耗的激励机制，让更多的用户参与语义标注，为实现学科知识本体的进化打下基础。

### 参考文献：

- [1] 荆涛,左万利等.中文网页语义标注:由句子到RDF表示[J].计算机研究与发展,2008,(7):1221-1231.
- [2] 杰夫·豪(JEFF HOWE).众包:群体力量驱动商业未来[M].北京:中信出版社,2011.
- [3] 邱均平等.国内外语义标注研究进展分析[J].情报理论与实践,2014,(5):12-16.
- [4] 王传清,毕强.数字图书馆自动化语义标注工具系统模型研究[J].现代图书情报技术,2014,(6):17-24.
- [5] Von Ahn,Luis,Benjamin Maurer,Colin McMillen,David Abraham,and Manuel Blum.recaptcha:Human-based character recognition via web security measures[J].Science,2008,(321):1465-1468.
- [6] Kim,Jonghak,Sangtae Kim,Joonhyuk Yang,Jung-hee Ryu,and KwangYun Wahn.Facecaptcha: a captcha that identifies the gender of face images unrecognized by existing gender classifiers[J].Multimedia Tools and Applications,2013,(72): 1-23.
- [7] 胡昭阳.众包科学:网络传播语境下的公众参与创新[D].合肥:中国科学技术大学,2015.
- [8] 邵阳.基于众包的语料标注系统设计与实现[D].大连:大连理工大学,2013.
- [9] 丁宇,车万翔,刘挺,张梅山.基于众包的词汇联想网络的获取和分析[J].中文信息学报,2013,(3):100-106.
- [10] 杨现民,余胜泉.开放环境下学习资源内容进化的智能控制研究[J].电化教育研究,2013,(9):83-88.
- [11] 余胜泉,杨现民.辨析“积件”“学习对象”与“学习活动”——教育资源共享的新方向[J].中国电化教育,2007,(12):60-65.
- [12] 王蓉,刘妍妍,陈禹,方美琪.社会性群体协作学习环境中的认知聚合研究[J].信息系统学报,2010,(1):65-75.
- [13] 杨现民.泛在学习环境下的学习资源进化研究[D].北京:北京师范大学,2012.
- [14] Tomas Mikolov.Word2vec project[EB/OL].https://code.google.com/p/word2v.2015-12-06.
- [15] 余胜泉,杨现民,程罡.泛在学习环境中的学习资源设计与共享——“学习元”的理念与结构[J].开放教育研究,2009,(1):47-53.

### 作者简介：

丁国柱：在读博士，研究方向为知识本体技术、计算机教育应用(laoding1982@qq.com)。

余胜泉：教授，博士生导师，研究方向为移动教育与泛在学习、信息技术与课程整合、网络学习平台关键技术、区域性教育信息化等(yusq@bnu.edu.cn)。

潘升：硕士，研究方向为研究方向为知识本体技术、计算机教育应用(ps199052@163.com)。

(下转第127页)

Education,2008,30(9):1183-1200.

[17] Oon PT, Subramaniam R. On the Declining Interest in Physics among Students-From the perspective of teachers[J]. International Journal of Science Education,2011,33(5):727-746.

[18] Robinson C, Sebba J. Personalising learning through the use of technology[J]. Computers & Education,2010,54(3):767-775.

[19] Haussler P, Hoffman L, Langeheine R, Rost J, Sievers K. A typology of students' interest in physics and the distribution of gender and age within each type[J]. International Journal of Science Education,1998,20(2):223-238.

[20] Ardito C, Costabile MF, De Marsico M, Lanzilotti R, Levialdi S, Roselli T, et al. An approach to usability evaluation of e-learning applications[J]. Universal access in the information society,2006,4(3):270-283.

作者简介:

翟小铭: 在读博士, 研究方向为物理教育、信息技术与课程整合、教育评价与测量等(xiaomingzh@mail.bnu.edu.cn)。

孙伟: 中学高级教师, 研究方向为物理教育(sunweiw@zbyz.net)。

郭玉英: 教授, 博士生导师, 研究方向为物理教育、科学教育等(yyguo@bnu.edu)。

Meilan Zhang: 副教授, 博士生导师, 研究方向为教育信息技术等(mzhang2@utep.edu)。

### Smart Classroom: An Evaluation of Its Implementations and Impacts

—Based on the Longitude Data of Physics Learning in a High School

Zhai Xiaoming<sup>1, 2</sup>, Sun Wei<sup>3</sup>, Guo Yuying<sup>1</sup>, Zhang Meilan<sup>4</sup>

(1.Department of Physics, Beijing Normal University, Beijing 100875; 2.College of Education, University of Washington, Seattle Washington 98195; 3.Shandong Zibo NO.1 Middle School, Zibo Shandong 255200; 4. Department of Teacher Education, University of Texas, El Paso Texas 79968)

**Abstract:** The paper selected 454 samples from a high school with smart-classroom system, conducted a followed-up study for a year, and surveyed the use of Smart-classroom, the extent to which it changed the traditional physics learning, etc., and examined the in-class and after-class impacts of using frequency and using time on physics interest and physics achievement for students with different learning ability levels and of different genders, based on panel data. Results indicate that the overall using time and frequency are high with fluctuant in specific use, and the student-driven and complexity use are relatively rare; the types of use are simplex, and mainly focuses on real-time interactive and accessing use; only augment traditional physics learning without transformation; however, still has significant impacts on students' interest and achievements, though fluctuates among different levels and gender of students.

**Keywords:** Smart-classroom; Smart Education; Intellectualization; Learning Environment; In-depth Integration; Fixed Effects Model

收稿日期: 2016年5月5日

责任编辑: 宋灵青

(上接第95页)

### The Annotation of Learning Resources Based on Crowdsourcing

Ding Guozhu<sup>1,2</sup>, Yu Shengquan<sup>1,2</sup>, Pan Sheng<sup>1,2</sup>

(1.School of Educational Technology, Faculty of Education, Beijing Normal University, Beijing 100875; 2.The Joint Laboratory for Mobile Learning, Ministry of Education-China Mobile Communications Corporation, Beijing 100875)

**Abstract:** Semantic annotation is the foundation and precondition of learning resources organization which based on Web3.0. But there are little paper Focus on the annotation of tacit knowledge. In this paper, we put forward a new way to solve this problem by crowdsourcing which can make full use of human computation to work out the magnanimity learning resources' tacit knowledge. further on ,we created a crowdsourcing based semantic tagging system based on LCS, via vote by the user and the machine for quality control, realized tacit knowledge during the learning resources by the above work and to prove based on crowdsourcing can effectively improve the quality of the annotation of learning resources by experiment.

**Keywords:** Semantic Annotation; Crowdsourcing; Subject Learning Resources; Learning Cell

收稿日期: 2016年6月13日

责任编辑: 赵兴龙