

基于“众包”的软件开发模式^①

李勇军, 缙西梅

(中原工学院 软件学院, 郑州 450007)

摘要: 随着 Internet 技术的快速发展, “众包”成为一种灵活有效的解决问题方式. 虽已在很多领域中得到广泛的应用, 但在软件开发领域的应用仍存在很多问题, 如缺乏统一标准及开发方法, 没有成熟的应用平台等. 针对此问题, 本文提出一个基于“众包”的软件开发平台及“众包”软件开发方法, 一定程度上为利用“众包”进行软件开发提供了一种思路和借鉴.

关键词: 众包; 软件开发; 开发平台; 开发方法

Software Development Model Based on Crowdsourcing

LI Yong-Jun, GOU Xi-Mei

(Software College, Zhongyuan University of Technology, Zhengzhou 450007, China)

Abstract: With the rapid development of Internet technology, Crowdsourcing becomes a flexible and effective solution. Although it has been widely used in many fields, the application in the field of software development still exists many problems, such as lack of unified standard and development methods, no mature application platform, etc. Aiming at these problems, this paper puts forward a software development platform and a software development method, which can provide a train of thought and reference for the use of crowdsourcing software development to a certain extent.

Keywords: web information extraction; knowledge presentation; data intensive web pages; ontology-based keyword library

1 引言

目前, 企业为开发出成功的产品、降低生产成本、提高效率、在市场上占有一席之地, 他们必须寻找出新的创新途径. 同时随着经济全球化和互联网技术的日新月异, 来自行业内的竞争压力让企业不得不去开拓视野, 以谋划更独特的创意, 以及进行更高效的生产. 因此, 作为一种基于互联网的新兴合作模式——众包(Crowdsourcing)正在悄然流行. Howe Jeff^[1]于 2006 年 6 月在美国《连线》杂志的一篇文章中首次提出了众包的概念, 用来描述一种新的商业模式, 即企业利用互联网将工作分配出去, 发现创意或解决技术问题. 企业和组织可以通过互联网利用志愿者群体的创意和能力, 这些志愿者具备完成任务的技能, 愿意利用自己的业余时间工作, 满足于对其服务收取小额报酬, 或者暂时并无报酬, 甚至仅仅满足于能够获

得更多报酬的前景, 尤其对于软件业和服务业, 众包提供了一种组织劳动力的全新方式. 因此, 能将“众包”商业模式的精髓与软件开发模式创新结合起来, 有助于软件开发企业以低成本满足客户多样化的需求运作模式, 充分发挥客户价值, 增强企业竞争力, 并且也能为学术界提供关于软件开发模式创新研究的新领域.

“众包”的实质是管理的问题, 它对我们的冲击力是思维层面的, 或者说是媒体层面的. 众包给我们提供的无非是一个在线协作的平台. 在这个平台上, 对用户了解不透, 也是“众包”模式难以实行的一大阻力. 没有区配社区 CRM(用户管理系统), 没有合理的信用机制, 又使“众包”在软件开发领域中难以得到应用. 再加上我国目前的软件开发没有达到一定的规范, 各个软件公司各自为政, 致使大部分软件项目不能进

^①收稿时间:2013-10-21;收到修改稿时间:2013-12-09

行跨企业、跨区域进行开发. 目前仍没有一个基于“众包”的软件开发平台, 可供开发人员、相关软件企业进行业务“众包”服务, 如何解决这些问题, 并将“众包”应用到软件开发领域变得越来越重要.

2 文献综述

众包作为一个新理念, 在相关文献中只给出一种思想性、描述性的概念, 而没有明确定义. 一些学者根据研究需要从不同视角给出众包的特征描绘. Howe^[1]在维基百科上为众包定义如下, “指一个公司或机构把过去由员工执行的工作任务, 以自由自愿的形式, 外包给特定的(通常是大型的)大众网络的做法”. 文献^[2]指出, 所谓“众包”其实就是组织把特定的工作任务, 通过指定的网络公示给外部受众, 并向按需要完成组织任务的毛遂自荐者支付约定报酬的大众承包模式. 文献^[3]指出, 众包是将软件开发领域中开放源代码的方法应用到其它领域. 价值中国网的 CEO 林永青(2009)^[4]则认为, 众包模式是网络社会的社会生产, “这不只是一场互联网的变革, 而是整个商业模式的变革; 这也不只是一场商业模式的变革, 而是整个社会的变革”. 文献^[5]中认为: 众包是利用共享网络外包给公众个人.

众包和普通意义上的外包不同点在于, 外包的任务是派给确定的个体, 而众包则是外派给不确定的群体. 此外, 外包强调的是高度专业化, 而众包则与之相反, 更注重自由和创意, 跨专业的创新往往蕴含着巨大的潜力. 对于众包与外包的关系, 文献^[6]从实施时间、实施条件、体现关系、发包对象选择、发包对象数量、产品生产者及付费情况等几个方面进行了比较. 而宝洁公司负责科技创新的副总裁 Larry Huston 评价道, “外包是指我们雇佣人员提供服务, 劳资双方的关系到此为止, 其实和雇佣关系没什么两样. 但是现在的做法是从外部吸引人才的参与, 使他们参与到这广阔的创新与合作过程. 这是两种完全不同的概念”. Huston 一语道破众包和外包的区别——众包的核心包含着与用户共创价值的理念.

众包的实施模式可以多种多样, 目前已经有很多成功的应用, 如 Amazon 的 Mechanical Turk 平台^[7]是基于经济奖励制度的方式让大众解决实际难题, 企业用户针对的是那些需要以数美分起价外包简单计算任务的公司, 而个人用户将通过完成某项工作获得小额

的报酬. Foodpickle^[8]则通过将众包与微博结合产生了关于食品的即时问答平台. TopCoder^[9]采用竞争性群体软件开发模式, 仅用 5 个月时间完成美国在线 AOL 需要 1 年时间开发的通信后端系统, 并且最终产品的功能完全达到客户的要求, 并且程度中每一千行代码平均只有 0.98 个漏洞, 远低于业内每千行 6 个的标准. 美国的 Zaarly^[10]嫁接了 LBS(位置服务)和众包模式只用一个月便吸引到了 100 万个“任务”. 人人猎头^[11]利用众包玩招聘, 采用企业悬赏招聘方式, 在众多网络平台发布悬赏职位, 鼓励用户推荐或者自荐来应聘, 是移动互联网时代众包招聘的创新者. 易到用车^[12]创新出中国互联网架构下“共享加盟+服务平台”的全新商业模式, 是中国第一家专业提供专车约租车服务的电子商务网站. 在机器翻译^[13]、微博信息的可信计算^[14]、语音研究^[15]、数据库研究领域^[16]、众包测试^[17]均有众包的应用.

3 “众包”软件开发平台系统架构

要建立“众包”软件开发平台, 首先要做好平台规划. 平台规划的好坏与系统的生命周期有直接的关系, 与系统将来的扩展性和业务发展变化的适应能力有很大关系. 本软件平台主要能够包括庞大的用户群存储、客户群管理 CRM、中控系统及支付系统. 系统逻辑架构如图 1 所示: 系统分为三个体系, 五个层次. 三个体系分别为: 标准规范体系、运维管理体系、安全保障体系. 五个层次从下至上依次为: IT 基础设施层、数据层、应用支撑层、业务层及门户层.

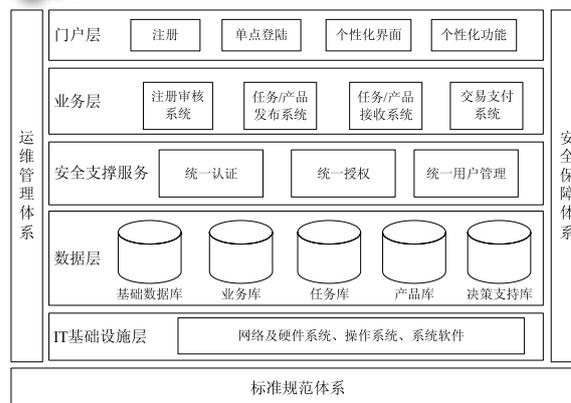


图 1 “众包”软件开发平台的系统架构

3.1 三个体系

为使自由工作者在“众包”软件开发模式中公平、安全、形成规模且任务需求者整体软件开发费用降低,我们为此平台构建了标准规范体系、运维管理体系及安全保障体系等 3 个体系. 标准规范体系主要是一个行业内公认的、统一的软件开发标准,包括技术标准和管理规范化,即软件工程的规范化;运维管理体系主要是保障系统正常运行,尤其是服务器的 Web 引擎和数据库引擎应尽可能地去预防各种错误,对于突发情况,尽可能地去修复;安全保障体系主要是对各种安全问题进行综合管理.

3.2 五个层次

为使基于“众包”的软件开发平台畅通运行,我们给出了 5 个层次: IT 基础设施层、数据层、安全支撑服务层、业务层及门户层. IT 基础设施层主要是提供用户需要的硬件系统、基础网络设施、操作系统及系统软件;数据层主要负责用户相关信息的存储、任务需求者需求的存储、自由工作者提供的任务存储以及自由工作者自发提交的产品存储等;安全支撑服务主要是保证任务需求者及自由工作者相关信息的统一管理;业务层主要是为保证自由工作者及任务需求者二者能畅通的交流创意,完成任务的部署、接受、验证等;门户层主要便于任务需求者及自由工作者使用此平台.

在信息审核、供需信息交流、安全支付等方面构建了如下子系统:

1)注册审核系统

该系统功能是审核用户身份.为保障平台用户权益,平台采用会员制,对会员注册的基本信息进行审核.只有通过审核验证之后,才能在平台上发布自己的任务需求,而未注册此平台的游客只能浏览其权限允许范围内的一些信息.

2)任务/产品发布系统

该系统功能主要分为两个方面,一个是平台用户(任务需求者)在该系统中发布需要解决的任务、相关的需求、愿意支付的酬金(或无报酬形式),任务需求者也可采用招标形式来获得满意的解决方案.平台的其他用户(自由工作者)可注册后进行该系统的需求库,找到愿意解决的任务或投标;另一个是用户在产品库中发布其开发出的创新产品,比如某种应用程序以及价格,其他用户可进入产品库,选择满意的产品.

3)任务/产品接收系统

该系统的功能主要是验证平台用户提交的任务及产品,如何符合标准规范,则可存入到任务库或产品库中.

4)交易支付系统

该系统主要针对付费产品或有报酬的任务.通过这个电子支付系统,拥有任务的用户通过与平台机构签订合同,并缴纳保证金,只要客户一经确认中标任务,中标者就可通过电子支付系统获得自己的报酬,从而才能保护自由开发者的积极性,降低风险.平台可使用现有的支付系统,比如支付宝、财付通,也可考虑开发“众包”软件开发平台的专属支付系统.

4 “众包”软件开发方法

在第 3 部分提到的“众包”软件开发平台之上,任务需求者和自由工作者可进行任务的发布、产品的发布、任务的分派及验证等.在传统的软件开发过程中,如何引入“众包”模式来进行开发?图 2 给出一种基于“众包”软件开发方法,其过程如下:

- 1) 公司 A 准备开发及使用一款软件,公司 B 是一个将自己的软件整体或部分的工作外包给大众个体的公司.我们假定公司 A 将此软件的整体或部分外包给公司 B,如图 2 中(1)所示;
- 2) 公司 B 将软件分成一个或多个部分,并且为每一部分设定相应的规格说明书及恰当的测试用例,另外如不是免费的,还应给出每一部分的价格,并将其利用网络派发给不同的开发人员,如图 2 中(2)所示;
- 3) 自由工作人员看到软件需求及测试用例后,如果觉得能够完成,可提供他们已有的经验及过去所做的相关工作,并提出他们的详细规格说明书及测试用例;
- 4) 公司 B 依据自由工作人员提供的资料来决定最后由谁来完成此任务,如图 2 中(3)所示,自由工作人员 a、b、c 三个开发人员成功接受了相应的任务;
- 5) 自由工作人员设计并实现所接受的任务,并自测此任务能够满足图 2(3)中所提供的规格说明书及测试用例;
- 6) 公司 B 接收任务,验证自由工作人员完成的程序满足了已设定好的规格说明书及测试用例,若满

足, 则按初期约定报酬给予支付, 如图 2(4)所示;
 7) 公司 B 将各个自由工作人员开发的软件进行集成, 并进行整体测试. 最后将集成后的软件交付给公司 A, 如图 2 所示.

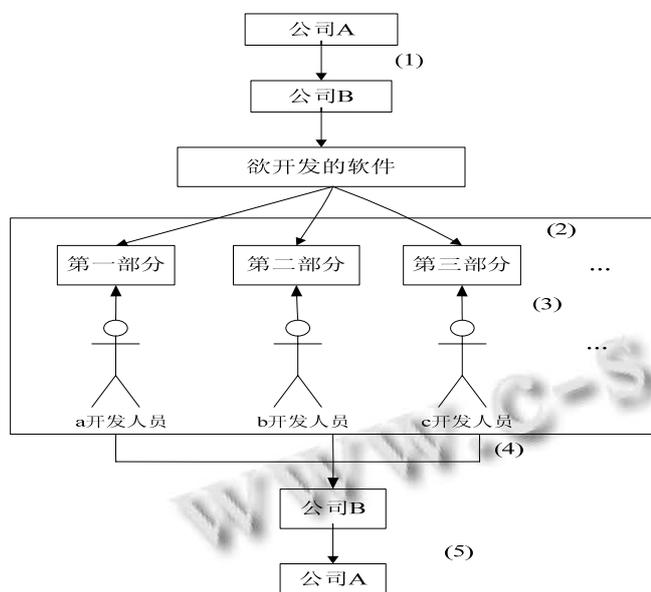


图 2 “众包”软件开发方法

5 结束语

本文所提出的“众包”软件开发平台能够充分发挥互联网的开放性, 激发大众的创作热情, 满足不同层次客户的个性化需求. 但在其发展过程中, 也会产生一系列的问题, 如诚信问题、如何依据自由工作者背景和任务特征进行个性化的任务分配、如何设定正确的检测点来保证用户提交任务的质量、知识产权纠纷等, 这些仍有待进一步的研究.

参考文献

- 1 Howe J. The rise of crowdsourcing. *Wired*, 2006, 14(6):176-183.
- 2 肖岚,高长春.众包——改变企业的创新模式. *上海经济研究*, 2010,(3):35-41.
- 3 仲秋燕,王彦杰,裘江南.众包社区用户持续参与行为实证研究. *大连理工大学学报(社会科学版)*,2011(3):1-6.
- 4 林永青.众包:网络社会的社会生产. *IT 经理世界*, 2009,(10): 82-83.
- 5 Doan A. , Ramakrishman R, Halevy AY. *CrowdSourcing*

- Systems on the World-Wide Web. Communications of the ACM*, 2011, 54(4): 86-96.
- 6 张利斌,钟复平,涂慧.众包问题研究综述. *科技进步与对策*, 2012,29(6):154-160.
- 7 Yan TX, Kumar V, Ganesan D. CrowdSearch: Exploiting crowds for accurate real-time image search on mobile phone. *Proc. of the International Conference on Mobile Systems, Applications and Services*. San Francisco, USA. 2010. 77-90.
- 8 Alonso O, Rose DE, Stewart B. Crowdisourcing for relevance evaluation. *Journal of SIGIR Forum(SIGIR)*, 2008, 42920: 9-15.
- 9 Archak N. Money, glory and cheap talk: analyzing strategic behavior of contestants in simultaneous crowdsourcing contests on TopCoder. com. *Proc. of the 19th International Conference on World Wide Web (WWW'10)*. 2010.21-30.
- 10 王晶.Zaarly:LBS 上的威客一族. *IT 经理世界 · CEOCIO*, 2011,(13):56-57.
- 11 百度百科. <http://baike.baidu.com/view/4720431.htm>.
- 12 百度百科. <http://baike.baidu.com/view/5530065.htm>.
- 13 Callison-Burch C. Fast, cheap, and creative: Evaluating translation quality using Amazon’s mechanical turk. *Proc. of the Conference on Empirical Methods in Natural Language*. Singapore. 2009. 286-295.
- 14 Castillo C, Mendoza M, Poblete B. Information credibility on Twotter. *Proc. of the WWW*. Hyderabad, India. 2011. 675-684.
- 15 Bigham JP, Jayant C, Ji H, et al. VizWiz: Nearly real-time answers to visual questions. *Proc. of the UIST*. New York City, USA. 2010. 333-342.
- 16 Franklin MJ, Kossmann D, Kraska T, et al. CrowdDB: Answering queries with crowdsourcing. *Proc. of SIGMOD*. Athens, Greece. 2011. 61-72.
- 17 张恒.众包是一种哲学:从小米手机众包模式说起. <http://mobile.51cto.com/news-300701.htm>.