

维基百科生产者特征分析及其启示

王丹丹^{1,2}

(1.中国科学技术信息研究所,北京 100038 ;2.河南科技大学 管理学院,河南 洛阳 471003)

摘要: 维基百科是 wiki 技术和协作模式成功的典范,证明了开放共享、协同写作这一新的内容创建形式和协作模式的可行性和成功之处。笔者从社区的角度对维基百科的成长动力进行分析,主要是从社区得以运行的源头,以此为切入点揭示英文维基百科取得巨大成功的原因,造成中文维基百科发展滞缓的因素,及对其他 wiki 项目建设管理所带来的启示。

关键词: 维基百科;社区;开放内容

中图分类号: G350 **文献标识码:** A **文章编号:** 1007-7634(2012)01-121-04

Study on the Wikipedia' Producers and Getting Enlightenment from It

WANG Dan-dan^{1,2}

(1. Information Science and Technology Institute of China, Beijing 100038, China;

2. Management School, Henan University of Science and Technology, Luoyang 471003, China)

Abstract: Wikipedia is a successful model of wiki technology and cooperation modes which prove the open sharing, collaborative writing are feasibility of successes. Take it as a community, the author analysis the growth dynamic of Wikipedia, reveal the factors which has made a huge success of English Wikipedia, and the reasons which slow the development of Chinese Wikipedia. The conclusion can be used for other wiki platforms.

Key Words: wikipedia; community; open content

维基百科是 wiki 技术和协作模式成功的典范,证明了开放共享、协同写作这一新的内容创建形式和协作模式的可行性和成功之处。自创立以来,他就呈现出指数型成长的态势^[1],尤其是英文维基百科,吸引了研究者和实践者的眼球。但中文维基百科的发展却与汉语的使用人数不相匹配。由于维基百科的成长和发展在很大程度上是由其内容的生产者决定的^[2],其本身也符合虚拟社区的特征^[3]。

1 研究方法

1.1 档案分析法

Wiki 在直观的意义上是交互的、协同写作的社

区知识构建工具,在更深层次上它包含了元数据和源信息^[4],笔者依靠的主要方法就是基于这些源信息进行档案分析。具体分三步。

第一步,在中、英文维基百科中随机选取若干条目,访问其历史页面,再次随机选取若干用户。

第二步,对选取的用户做进一步的分析,一是访问该用户的用户页面,通过阅读了解部分信息;二是通过维基百科特殊页面类别中,“用户贡献”工具页面对所选择的用户的贡献情况做调研。

第三步,在通过上述方法对所选择的若干用户做完整个分析之后,试图总结出规律性的东西。

1.2 内容分析法

内容分析法是一种对文献内容做客观系统的定

收稿日期:2011-09-04

作者简介:王丹丹(1980-),女,河南洛阳人,讲师,博士后,主要从事网络化信息服务研究。

量分析的专门方法,其目的是弄清或测验文献中本质性的事实和趋势,揭示文献所含有的隐性情报内容,对事物发展作情报预测^[5]。本研究中所运用的内容分析法主要体现在对维基百科对话页互动内容进行定量统计分析。

2 基于档案分析法的生产者特征对比

Wikipedia由非营利组织维基媒体基金会负责,目前有超过270种语言版本。英文维基百科始于2001年1月,直到2001年10月才真正出现第一个条目。中文则始于2002年10月。截至2009年12月14日,英文维基百科已拥有3,125,162个条目,是中文维基百科条目数的10.9倍。英文维基百科拥有注册用户11,207,531个,是中文维基百科注册用户数的14.9倍,其中英文维基百科注册用户中管理员有1706名,是中文维基百科管理员数量的21.6倍;英文维基百科编辑总次数为353,112,521,是中文维基百科编辑总次数的29.6倍^[6-7],如表1。

表1 中、英文维基百科基本统计数据

项目类别	注册用户	管理员	内容页面	编辑次数
英文版	11,207,531 (14.9)	1706(21.6)	3,125,162 (10.9)	353,112,521(29.6)
中文版	754,478	79	286,772	11,945,984

遵循上述三个步骤,按照用户总数量万分之一的标准,在英文维基百科中随机抽取用于分析的用户样本1120名,在中文维基百科中随机抽取用于分析的用户样本75名,从基本信息,用户动机和编辑贡献情况三个层面展开进行归纳分析。

2.1 生产者基本特征比较

从性别分布来看,中文维基百科生产者中男性占56%,女性占24%;英文维基百科男性占39%,女性占13%。无论是中文还是英文维基百科,男性生产者所占的比例都远高于女性生产者。

从年龄分布来看,中文维基百科生产者中20岁以下的占20%,大于20岁小于30岁的占45%,大于30岁的占23%。英文中20岁以下的占4%,20岁到30岁之间占15.2%,30岁到40岁之间的占5.4%,40岁到50岁之间的占4%,大于50岁的占3%。无论是中文还是英文维基百科,生产者均以年轻人居多。中文与英文相比,主要生产者群体更为年轻,而且主要是在校学生,其中出国留学人士占据了相当大的比例。英文维基百科中,除了在校大学生外,具有工作经历的用户也占据相当大的比例。

从生产者参与不同版本维基百科的编撰情况来看,中文维基百科中79%的生产者除了参与中文维基百科的编撰外,还参与了其他多个语言版本的维基百科编辑活动。而英文维基百科中60%的生产者仅参与了英文维基百科的编撰。

从生产者个人受教育的程度看,中文维基百科中本科以上学历的占据了48%。英文中本科学历的占2%,本科学历的占27%,硕士以上学历的占12.5%。对比发现无论是中文还是英文维基百科生产者受教育程度都比较高,尤其是英文维基百科中硕士以上学历的生产者占据了一定的比例。

2.2 生产者动机比较

这里主要是分析生产者的参与动机与离开动机。就参与动机而言,大致可以归为五类:①对新事物充满好奇,故而进行探索、尝试或者学习;②认同维基百科的价值观,对维基百科感兴趣,愿意与他人共享知识,或者为维基百科的发展贡献力量;③闲着没事,打发时光;④以维基百科为平台,展示自己,结交朋友;⑤享受写作。分析发现,无论是中文还是英文维基百科,用户对其定位、知识生产与共享模式和价值观的认同是参与贡献的主要动机。另外,相当一部分用户通过维基百科的用户页展示自己,在用户页提供到个人博客或个人网站的链接,以维基百科为平台结交朋友。这一类型的动机在教育水平高的用户中更为普遍,很多高学历用户特别注重自己用户页的编撰,一方面展示自己的研究成果,另一方面希望以此结交更多兴趣相同的研究者。此外,英文维基百科的一些生产者指出,相对于其他语言版本的维基百科而言,英文维基百科大规模的用户群体和相对较高的质量水平,是吸引他们参与贡献的决定性因素,因为这使他们感觉参与编辑的过程更为有趣。

中文维基百科中随机抽取用于分析的75名用户中,有16位用户在其用户页对离开的动机进行了说明,对其内容进行分析,笔者认为生产者不愿意继续贡献的原因也大致可以归为五类:①时间问题;②能力问题;③健康问题;④维基百科自身的问题;⑤懒惰。其中9名用户(占57%)指出,淡出维基百科,贡献不积极的主要原因在于时间和精力有限,其他任务的优先性和重要性超过了维基百科;4名用户(占25%)指出中文维基百科自身的一些因素,除此之外,用户自身的惰性、健康问题和能力问题也是导致用户离开的因素。英文维基百科随机抽取的这

1120名用户中,只有17名用户在用户讨论页中表达了其不再贡献的原因。考虑到样本量太小,进一步随机选取了100名英文维基百科用户,这100名用户中,又有43名用户对离开的原因进行了反馈。结果如图1。

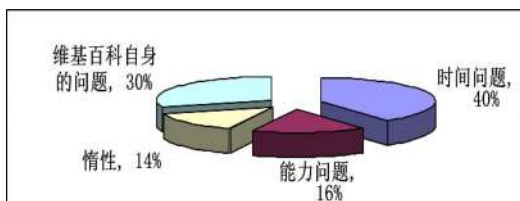


图1 英文维基百科生产者离开的动机分布

将维基百科中不同荣誉类型(为了鼓励参与者,“维基参与荣誉”依照服务的时间、参与编辑的次数、新条目的数目、特殊贡献的多寡等编辑贡献情况将授予用户的荣誉分为三个等级,新人荣誉、进阶荣誉和高级荣誉^[8])中注册用户与管理员用户的数量进行对比,将维基天使^[9](熟悉维基百科的运作与维护,并乐意与其他人交流的维基百科人)中注册用户与管理员用户的数量进行对比。结果表明,在维基百科中,管理员参与了大量的编辑活动,往往也是较高荣誉的获得者,更愿意承担起与新用户沟通的义务。

2.3 生产者编辑贡献情况比较

从生产者身份类型来看,中文维基百科生产者中普通注册用户占77%,管理员用户占23%。英文维基百科生产者中,普通注册用户占71%,管理员用户占29%。

从编辑次数来看,中文维基百科管理员用户平均每年的编辑次数是5800次,平均每月编辑483次;普通注册用户平均每年的编辑次数是2607次,平均每月编辑217次。英文维基百科每个管理员用户平均每年的编辑次数是12609次,平均每月编辑1051次;每个普通注册用户平均每年的编辑次数是4447次,平均每月编辑371次。无论是英文维基百科管理员用户,还是普通注册用户都比与中文维基百科编辑贡献更为活跃。

对管理员和普通注册用户不同命名空间^[10-11]的贡献情况进行对比分析,发现无论是中文还是英文维基百科,无论是管理员用户还是普通注册用户,条目页的贡献活动都是其主要活动内容。但是,与普通注册用户相比,管理员更多的参与了对话页面(包括条目对话页和用户对话页)的沟通交流活动:中文管理员用户对话页活动占12.22%(条目对话页

2.8%,用户对话页9.42%),普通注册用户对话页活动占7.72%(条目对话页2.41%,用户对话页5.31%);英文管理员用户对话页活动16.05%(条目对话页6.68%,用户对话页9.37%),普通注册用户对话页活动占15.04%(条目对话页6.03%,用户对话页9.01%)。无论是普通注册用户还是管理员,用户对话页沟通活动都超过了条目对话页的沟通活动。将中、英文维基百科作对比,发现英文维基百科中的交流活动所占的比例大于中文维基百科。

3 基于内容分析法的生产者特征对比

对话页是维基百科中最常见的交流形式,它以类似论坛的形式组织,没有预先确定的主题框架,当开始一条新的讨论主题时,才发展页面的结构,每一条讨论之后都显示作者和确切的时间。很多对话页的规模都超过了实际的内容本身。例如“不适合维基百科的文章”的对话页^[12]的长度要远大于“不适合维基百科的文章”本身。除了条目,用户页和所有其他的维基百科页面也都有各自的对话页。对话页互动内容的分析有助于从微观层次把握社区中互动的深刻涵义以及互动的作用和功能。

在中文维基百科中随机选取23个条目(其中特色条目11个),对其对话页的全部交互内容进行分析;在英文维基百科中选取11个特色条目,对其最近的三个对话页中的交互内容进行分析。结果表明,无论是中文还是英文维基百科,条目对话页中所有的讨论都与条目的改善建议有关,主要是围绕条目内容的准确性、权威性和可证实性、凝聚性、完整性、文字流畅性、一致性和相关性等问题进行交流协商。从讨论模式来看,协调的请求和建议改变都是条目对话页的讨论中最流行的模式;从对话页的交互风格来看,条目对话页中很少有明确的赞扬或者肯定,贡献者很少通过使用ID针对某一个人,交流很客观。特色条目对话页的交互次数和参与人数都高于普通条目。从交流的方式看肯定、鼓励、承认或者感谢在特色条目对话页出现的频率明显高于普通条目。

从特色条目对话页的交互情况看,中文维基百科参与交互的人数和交互的次数,都远小于英文维基百科:中文维基百科的平均交互次数是28次,参与交互的人数是13人;英文维基百科的平均交互次数是781次,参与交互的人数是251人。通过对特色条目对话页中注册与匿名用户交互的内容进行比

较发现:①匿名用户在进行编辑前通常是提建议,相对于注册用户而言,他们的行为更可能会引起注意和评论;②与注册用户相比,匿名用户发表的消极的言论多于积极的言论,所提出的建议被接受的少,被拒绝的多;③积极地言论相对于消极的言论更可能得到回应,但是总体而言,有相当一部分发起的讨论话题被忽略或者没有得到响应。

在每个条目对话页中选取交互次数最多的用户对其用户对话页进行分析,发现:用户对话页是维基人被社区内其他维基人认识的重要渠道。创建对话页可以使用户自己在社群内的知名度提高,方便其他用户与其联络,了解其长处进而开展协作;通过用户对话页,其他社群成员可以针对该用户编辑内容和行为中存在的问题直接与其进行交互,既指出存在的问题,又有针对性的对其进行指导、教育和帮助,如此以来既有利于问题及时得到解决,也有利于用户自身积累经验快速成长发展;通过用户对话页,管理员或者有经验的维基人,可以对新手表示欢迎,让其感受到来自社群的关怀与温暖,帮助其快速的熟悉和适应维基的环境,这对于吸纳新用户、发展潜力维基人具有重要的作用;此外,还可以通过用户对话页进行各种情感交流,比如给予用户鼓励、对用户的工作给予肯定、赞扬,当然也可以表达不满、发泄情绪,用户对话页交互效果的好坏直接影响到用户进一步参与贡献的动机。

4 结 语

维基百科这种以 wiki 技术为依托,基于社群合作的内容构建模式,代表着一种新型的社区——开放内容社区,这里的开放体现在两个层面,一是内容开放,二是内容创建过程开放。成员和内容是决定开放内容社区能否呈现良性发展态势的两大因素。这里的成员是指内容生产者,他们是社区中知识创造和信息交流的主体,社区内词条内容的创建与完善完全来自于社区成员的贡献。内容具体而言涉及内容的增长和内容质量的稳定两个层面。

通过分析,得出以下几点推论:①男性相对于女性而言更容易采纳和掌握 wiki 技术;②用户的个人爱好是社区词条创建的主要来源,知识质量的高低取决于贡献者对某一词条知识贮备的程度;③交互的质量、密度以及社群成员愿意付出的完成这些交互的时间,都会对其贡献行为产生影响,群组中的成员需要感觉到他们对群组做什么产生了影响,否则

将丧失参与的动机;④维基之所以能够融合大家的智慧,最主要来自于方针的执行与共识的形成。目前,中文维基百科的生产以翻译为主,原创性内容生产所占比例不高,生产者的学历层次低于英文;无论管理员还是普通注册用户的编辑贡献都远不如英文活跃,在条目层面的交互人数和频次也都远低于英文,方针的执行与共识的形成与其条目的成长不相匹配,都在一定程度上减缓了中文维基百科的发展。

参考文献

- 1 Bongwon Suh, Gregorio Convertino, Ed H. Chi et al. The Singularity is Not Near: Slowing Growth of Wikipedia[EB/OL]. <http://www.wikisym.org/2009/09/08/wikisym-2009-paper-the-singularity-is-not-near-slowng-growth-of-wikipedia/>, 2009-12-04.
- 2 Jan Marco Leimeister, Pascal Sidiras, Helmut Kremer. Success Factors of Virtual Communities from the Perspective of Members and Operators: An Empirical Study.[EB/OL].http://users.tkk.fi/molaine3/ml_mastersthesis_310506_public.pdf, 2009-12-04.
- 3 Giorgos Cheliotis. From Open Source to Open Content: Organization, Licensing and Decision Processes in Open Cultural Production[J]. Decision Support Systems,2009,(47):229244.
- 4 B. Stvilia, M. Twidale et al. Information Quality Work Organization in Wikipedia[EB/OL].http://mailer.fsu.edu/~bstvilia/papers/stvilia_wikipedia_infoWork_p.pdf, 2009-07-21.
- 5 邱均平. 文献计量内容分析法[M]. 北京:国家图书馆出版社, 2008:11.
- 6 Statistics[EB/OL].<http://en.wikipedia.org/wiki/Special:Statistics>, 2009-12-04.
- 7 统计 [EB/OL]. <http://zh.wikipedia.org/zh-cn/Special:%E7%B6%9F%E5%AD%A4%E6%81%A6%E4%BF%A1%E6%81%AF>, 2009-12-04.
- 8 维基荣誉[EB/OL].<http://zh.wikipedia.org/zh-cn/Wikipedia:%E7%B6%AD%E5%9F%BA%E6%A6%E8%AD%BD>, 2009-12-04.
- 9 维基天使[EB/OL]. <http://zh.wikipedia.org/zh-cn/Wikipedia:%E5%B0%8F%E5%A4%A9%E4%BD%BF%E5%B7%A5%E4%BD%9C%E5%B0%8F%E7%B5%84>, 2009-12-04.
- 10 命名空间[EB/OL].<http://zh.wikipedia.org/zh-cn/%E5%90%8D%E5%AD%97%E7%A9%BA%E9%97%B4>, 2009-12-04.
- 11 Namespace[EB/OL].<http://en.wikipedia.org/wiki/Namespac>, 2009-12-04.
- 12 Wikipedia_talk:What_Wikipedia_is_not [EB/OL].http://en.wikipedia.org/wiki/Wikipedia_talk:What_Wikipedia_is_not, 2009-03-04.

(责任编辑:刘凤琴)