Post-print: Construction of Evaluation Metrics for Baidu Baike Entries from a Multi-dimensional Perspective

Authors: He Zhenyu, Dong Xiangxiang, Zhu Qinghua

Date: 2023-07-26T00:00:00+00:00

Abstract

[Purpose/Significance] This study integrates user usage behavior into the encyclopedia entry evaluation system and conducts a comprehensive evaluation of Baidu Baike entries from a multi-dimensional perspective. [Method/Process] Building upon existing research from scholars both domestically and internationally, and based on 4 dimensions, 12 quantitative indicators were selected as research metrics to conduct a comprehensive evaluation of Baidu Baike entries. [Results/Conclusion] User usage behavior indicators and link indicators play a relatively important role in the comprehensive evaluation of entries; the research enriches the evaluation system for encyclopedia entry studies, but its limitation lies in not considering features such as richness and rigor that are difficult to quantify accurately.

Full Text

Preamble

Volume 63, Issue 12, June 2019
ChinaXiv Cooperative Journal

Construction of Baidu Encyclopedia Entry Evaluation Indicators from a Multi-dimensional Perspective

He Zhenyu, Dong Xiangxiang, Zhu Qinghua
School of Information Management, Nanjing University, Nanjing 210023

Abstract

[Purpose/Significance] This study incorporates user behavior metrics into the encyclopedia entry evaluation framework to conduct a comprehensive assessment of Baidu Encyclopedia entries from a multi-dimensional perspective.

[**Method/Process**] Drawing upon domestic and international research findings, we selected 12 quantitative indicators across four dimensions as our research metrics to comprehensively evaluate Baidu Encyclopedia entries. [**Result/Conclusion**] User behavior indicators and link indicators play a relatively significant role in the comprehensive evaluation of entries. This research enriches the evaluation system for encyclopedia entry studies, though its limitation lies in not considering characteristics such as richness and rigor that are difficult to quantify accurately.

Keywords: Baidu Encyclopedia entry; Grey Relational Analysis; Entropy Weight Method; Comprehensive Evaluation
Classification Number: G250
DOI: 10.13266/j.issn.0252-3116.2019.12.015

With the application of Web 2.0 technologies, online encyclopedias—typical Wiki products—have achieved rapid development, with Baidu Encyclopedia being the most representative product in China. Entries (or articles) constitute the smallest unit of online encyclopedias. As of February 8, 2019, Baidu Encyclopedia contained 15,906,266 entries, with 6.7 million people participating in their compilation. Entries represent the core content of Baidu Encyclopedia, which categorizes them into 11 major categories including arts, science, nature, culture, and geography. The entry structure is highly standardized, comprising the entry title, summary, basic information in table format, table of contents, main text, reference annotations, entry tags, and other elements.

Online encyclopedia content is primarily user-generated and significantly influenced by the subjectivity of editors, making it difficult for users and administrators to effectively measure entry quality. Although Baidu Encyclopedia has established a proprietary review system, entry information quality largely depends on the quality of the vast editor community. Since editors do not always follow rigorous structured processes during editing, and the number of encyclopedia entries has grown explosively, entry quality has become increasingly uneven, negatively impacting users' experience. Comprehensive evaluation of encyclopedia entries has always been an important research problem.

2 Research Status

Current research on online encyclopedia quality evaluation mainly focuses on the following aspects:

(1) Qualitative evaluation studies. These primarily reference traditional encyclopedia standards, conducting qualitative analysis of target entries using expert or individual knowledge for manual review. Many scholars have studied factors affecting entry quality and established objective evaluation criteria. The Wikipedia website [**?**] itself has proposed detailed article quality rating standards, dividing entries into different levels such as featured, A-class, good, B-class, C-class, start, and stub through peer review. B. Stvilia [**?**] et al. explored the impact of different cultural and community backgrounds on infor-

mation quality assessment and the feasibility of using article editing-based metrics for automated quality measurement of Wikipedia in different contexts. S. Lichtenstein [?] et al. established a knowledge production model for Wikipedia and analyzed its information quality. O. Arazy [?] et al. developed a theoretical model explaining how three factors—diversity of member knowledge bases, task conflict, and member roles—interact to determine Wikipedia entry quality, conducting empirical research on 96 entries. J. Liu [?] et al. analyzed the relationship between editor collaboration and entry quality, with empirical research helping to improve Wikipedia quality. E. Yaari [?] et al. selected 64 users to evaluate the quality of five Wikipedia entries, exploring how information consumers assess content quality in collaborative environments. K. Osman [?] et al. analyzed 147 conversations about Wikipedia quality using a grounded theory approach, exploring the role of entry conflict in Wikipedia quality. D. Lewandowski [?] et al. summarized extended discussions on encyclopedia entry quality, developed a heuristic method for evaluating Wikipedia entries, and compared Wikipedia pages in search engine results with manual relevance judgments from experimenters. Ding Jingda [?] examined the connotation of Wikipedia entry information quality from three aspects, summarized domestic and international research status and different evaluation indicators for Wikipedia entry information quality, and constructed a multi-dimensional heuristic evaluation framework containing concepts, relationships, classifications, and methodologies to evaluate Wikipedia entry information quality. Zhao Wenxuan [?] analyzed the characteristics, advantages, working models, and resource comparisons of online encyclopedias in detail, discussed entry information quality, provided quality evaluation criteria for high-quality entries, and elaborated on six factors affecting online encyclopedia entry quality. Huang Linghe [?] studied the dynamic evolution characteristics of online encyclopedia information quality and proposed targeted evaluation strategies, emphasizing process evaluation, considering lifecycle characteristics, and requiring active third-party involvement.

(2) Construction of evaluation indicator systems. This research establishes detailed statistical indicator systems for entry quality based on statistical indicators and quantitative evaluation methods. A. Lih [?] extracted two evaluation indicators, "diversity" and "rigor," from the perspective of editing history information to judge entry quality. F. Chevalier [?] et al. proposed five maturity and quality assessment indicators for Wikipedia articles through visualization methods, significantly reducing evaluation time while maintaining quality accuracy. Jin Yan and Zhou Ting [?] et al. studied collaborative content creation systems in general, established specific evaluation indicator systems using the Analytic Hierarchy Process, constructed a targeted Baidu Encyclopedia evaluation indicator system from Baidu Encyclopedia's actual situation, and conducted verification. Zhang Bo [?] et al. used the Analytic Hierarchy Process, integrated user needs and content characteristics, and established a content quality assessment model with four levels and 14 indicators using Wikipedia as an example. Jin Yan [?] established a quality evaluation indicator system for collaborative content creation systems from a user experience perspective

and proposed quality assurance measures in three aspects: information value, system performance, and user satisfaction.

(3) Automatic evaluation studies. With the development of computer technology, automatic evaluation methods have been introduced to online encyclopedia entry evaluation. By selecting appropriate entry features and combining them with other attributes, machine learning algorithms are used to automatically complete entry rating work. M. Warncke-Wang [**?**] et al., based on the work of B. Stvilia [**?**] et al., ultimately selected five features and also used decision tree algorithms to classify entry quality. J. E. Blumenstock [**?**] proposed a simple method to measure Wikipedia entry quality using only word count and demonstrated through empirical testing that its effect is better than many complex methods. H. Dalip [**?**] et al. explored numerous quality indicators, unified evaluation indicators into the same framework through machine learning methods, and identified indicators most and least relevant to quality evaluation. L. Calzada [**?**] et al. provided a simple information quality model, mainly studying "stable" and "controversial" Wikipedia articles and comparing them with manually evaluated information quality. Qiu Jiangnan [**?**] et al. proposed 15 quantifiable objective indicators, used the decision tree C4.5 classification algorithm to construct an automatic classification model for evaluating Wikipedia page information quality. Li Xinyi [**?**] selected content features and editing history features as core indicators for judging entry quality, used SVM-based classification methods and PageRank algorithms to distinguish and rank high-quality and low-quality entries. Tong Zhaojuan [**?**] et al. selected three webpage feature values, achieved automated evaluation of webpage quality through automated processing and extraction of webpage features, and selection of high-quality classifiers. Xiao Kui and Li Bing [**?**] et al. considered entry attributes and editor behavior, screened 15 entry attributes and four user attributes, and determined Wikipedia entry levels through classification and cosine similarity calculation of entry quality.

(4) Model-based ranking evaluation. Based on specific models, entry quality values are automatically calculated and ranked to complete the evaluation process. M. Hu [**?**] et al. used interaction data between entries and contributors from Wikipedia editing history to propose three quality detection models: BasicModel, PeerReviewModel, and ProbReviewModel, using the NDCG@k indicator for algorithm evaluation and achieving good performance. H. Zeng [**?**], E. Lim [**?**], and P. Dondio [**?**] et al. used different models to calculate entry quality from the perspective of editing behavior and obtained entry quality ranking results. T. Wohner [**?**] et al. proposed an information lifecycle-based method to evaluate Wikipedia entry quality. S. Javanmardi [**?**] et al. modeled the dynamic evolution of Wikipedia entry content quality, evaluated featured and non-featured Wikipedia articles, and conducted case analysis through CalSWIM. Y. Suzuki [**?**] evaluated entry quality from the perspective of entry content survival rate and developed a set of entry ranking methods. X. Li [**?**] et al. developed several models to rank Wikipedia articles through the relationship between trial articles and editors and proposed that using manual evaluation to assist automatic

evaluation is a feasible quality assessment solution.

Some scholars have also used link analysis ideas to calculate entry quality rankings. D. Wilkinson et al. [**?**] used the PageRank algorithm to first grade entries, then analyzed entry quality within the same grade to discover commonalities of high-quality entries. Xiao Kui and Luo Baoshan [**?**] et al., through domain limitation, applied the HITS algorithm idea, used the credibility of editors in specific fields to automatically calculate entry quality values. K. Wu [**?**] et al. conducted network analysis on the revision history of Wikipedia in six different languages, exploring the impact of the network structure of interactions between articles and contributors on UGC quality.

In summary, current research mainly focuses on single-dimensional perspectives such as entry content dimension, network dimension, and editing dimension, while multi-dimensional entry quality evaluation research is still lacking. In addition to the feature indicators used in the above scholars' research, Baidu Encyclopedia entry attributes also include user behavior characteristic indicators: entry user browsing, sharing, and like values, while research on encyclopedia entry evaluation and classification based on such indicators is very rare. Therefore, this study aims to solve two problems: How to incorporate user behavior indicators into the Baidu Encyclopedia entry evaluation system; How to comprehensively evaluate Baidu Encyclopedia entries from a multi-dimensional perspective.

3 Research Methods

To address the two research questions, this study refers to the encyclopedia entry evaluation characteristic indicators proposed by domestic and international scholars, combines them with the accessible content from the 2018 Baidu Encyclopedia user interface, and selects 12 quantitative indicators across four dimensions as research objects.

(1) Content dimension: Content length indicators include summary length, main text length, and image count. Content length indicators represent the richness of entry content. Compared to Blumenstock's use of word count as an indicator, this study more finely distinguishes between summary and main text length, while adding image count to make the data more convincing. Secondary content indicators include reference count, tag count, and infobox count. These three are not directly related to entry quality like entry length and link indicators but reflect entry refinement from a side perspective.

(2) Network link dimension: Outlink and inlink counts. As commonly used indicators in link analysis, outlink and inlink counts can reflect the connectivity of entries in the entire entry network. Entries with high outlink and inlink counts are more likely to attract users to transfer from other entries or to other entries and are important indicators reflecting entry usage probability.

(3) User behavior dimension: Page views, likes, and shares. The quality of

encyclopedia entries is largely associated with user behavior. Entries with high views, likes, and shares inevitably differ in quality from neglected entries.

(4) Editing dimension: Compared to entry creation time, edit count better measures the entry lifecycle, reflecting differences between high-quality entries with high update frequency and low-quality entries with long-term lack of maintenance.

3.1 Indicator Weight Measurement Method

To measure the influence of each indicator on the comprehensive evaluation results, this study uses the entropy weight method to calculate the entropy weight of each indicator. The entropy weight method is an indicator weight assignment method based on the concept of information entropy from Shannon's information theory [**?**], which can evaluate the weight of each parameter without decision-maker preferences. The calculation steps are as follows [**?**]:

(1) Standardize the existing Baidu Encyclopedia entry indicator decision matrix D (with m evaluation indicators and n entry samples) to obtain the standardized matrix R:

$$R = (r_{ij})_{m\times n}$$

where r_{ij} refers to the standardized value of the i-th entry on the j-th indicator.

$$r_{ij} = \frac{D_{ij} - \min\{D_{ij}\}}{\max\{D_{ij}\} - \min\{D_{ij}\}}, \quad 1 \le i \le m, 1 \le j \le n \tag{1}$$

(2) Calculate the entropy value H of all indicators:

$$H_i = -k \sum f_{ij} \ln f_{ij}, \quad i = 1, 2, 3, \dots, m \tag{2}$$

(3) Calculate the entropy weight ω of all indicators:

$$\omega_i = \frac{1 - H_i}{m - \sum_{i=1}^{m} H_i} \tag{3}$$

where $0 \le \omega_i \le 1$, and $\sum_{i=1}^{m} \omega_i = 1$.

The larger the parameter weight calculated by the entropy weight method, the greater the parameter's impact on the evaluation results.

3.2 Entry Quality Evaluation Method

The Grey Relational Analysis (GRA) model was originally proposed by Professor Deng Julong, the founder of grey system theory, in the 1980s. Professor Deng's GRA model is also known as the Deng's Grey Relational Analysis model or Grey Relational Analysis model in application. Currently, the GRA model is widely used in multi-objective optimization. The specific evaluation steps in this study are as follows:

(1) Collect and statistics experimental data according to selected evaluation indicators, construct the decision matrix for standardization (same standardization process as in 3.1).

(2) Calculate the correlation coefficient between each group of entry indicators and the optimal value:

$$\xi_{ij} = \frac{(|z_{ij} - z_{0j}|) + \eta \times \max(|z_{ij} - z_{0j}|)}{|z_{ij} - z_{0j}| + \eta \times \max(|z_{ij} - z_{0j}|)} \tag{4}$$

where η is the resolution coefficient, and $\eta \in (0, 1)$. Generally, the value of η determines the difference between correlation coefficients; the smaller η is, the stronger the resolution. In this study, η is set to 0.5. The correlation coefficient ξ represents the degree of association between the i-th entry and the optimal value on the j-th indicator.

(3) Calculate the correlation degree of each entry based on the weight coefficients obtained by the entropy weight method, and then make a comprehensive evaluation.

3.3 Data Acquisition

The data crawling process used in this study is as follows: (1) **Initial crawling**: By analyzing the link structure of the Baidu Encyclopedia site, links in the form of `http://baike.baidu.com/view/00000010.htm` were found. By traversing digital IDs, a series of initial links could be obtained. Using Python's `urllib2` library, pages could be downloaded locally. (2) **Data cleaning**: First, determine the structural elements of the Baidu Encyclopedia page to be crawled, such as: main title, subtitle, summary, main text, images, internal links, external links, etc. Simultaneously, establish a class named `Page` to describe the above information, where titles are represented by text variables and links by list variables. (3) **Data storage**: Establish a MySQL database, create a table named `entry` in the database, represent the above page structure elements with attributes in the table, and use the `pymysql` library to write the cleaned data into the database. (4) **Secondary crawling**: Through analysis, it was found that the initially crawled pages contained a large number of relative links in the form of `/item/abc`, with link targets being other entry pages. Therefore, using the internal links obtained from the initial page collection as seed URLs, the methods in steps 1-3 were used again for data crawling and cleaning. During secondary crawling, it was necessary to determine whether the URL had already been written into the database; if a corresponding record existed, it would not be crawled repeatedly.

4 Research Results and Analysis

4.1 Indicator Weights

Based on the decision matrix generated from all encyclopedia entry data, the entropy weight method was applied to calculate the entropy weight values of all

indicators as shown in Table 1 and Figure 1 [Figure 1: see original paper]:

Table 1 Encyclopedia Entry Indicator Weight Table | Indicator No. | Indicator Name | Weight | |————|————-|———| | Indicator 1 | Summary word count | 0.022 | | Indicator 2 | Main text word count | 0.041 | | Indicator 3 | Infobox count | 0.011 | | Indicator 4 | Tag count | 0.018 | | Indicator 5 | Outlink count | 0.086 | | Indicator 6 | Inlink count | 0.171 | | Indicator 7 | Reference count | 0.052 | | Indicator 8 | Image count | 0.052 | | Indicator 9 | Page views | 0.160 | | Indicator 10 | Shares | 0.185 | | Indicator 11 | Likes | 0.175 | | Indicator 12 | Edit count | 0.026 |

Figure 1 Encyclopedia Entry Indicator Weights

The results show that the four highest-weight indicators include three user behavior indicators—page views, likes, and shares—and the inlink count weight is comparable to these three. The outlink count, as an "altruistic" indicator, has a weight less than half of the inlink count but is still significantly higher than the remaining indicators.

4.2 Comprehensive Evaluation Results

Based on the indicator weights obtained from the entropy weight method, applied to the grey relational analysis algorithm, the comprehensive evaluation scores for all entries, i.e., entry quality scores, were calculated and ranked accordingly.

Top 10 Entry Quality Rankings:
Taiwan (Region of the People's Republic of China), Liu Shishi, CCTV Documentary Channel, Uchiha Sasuke, Nanjing Road (Shanghai Nanjing Road), Sagittarius (Astrology), Harem Anime, Korean TV Dramas, DOS (DOS Operating System), Wang Leehom

Bottom 10 Entry Quality Rankings:
Related Peak, Game Interface, Lung Kwu Tan Power Plant, Social Isolation, Primary Coast, Compression Algorithm, Compatibility Testing, Tower of Hanoi Problem, Monocrystalline Silicon Solar Cell Module, Carbon Dioxide Tax

The comparison of indicators across dimensions for high-ranking and low-ranking entries is shown in Table 3 :

Table 3 Entry Indicator Comparison | Indicator | CCTV Documentary Channel | Uchiha Sasuke | Lung Kwu Tan Power Plant | |————|————|————|————————| | Summary length | … | … | … | | Main text length | … | … | … | | Infobox count | … | … | … | | Tag count | … | … | … | | Outlink count | … | … | … | | Inlink count | … | … | … | | Reference count | … | … | … | | Image count | … | … | … | | Likes | … | … | … | | Edit count | … | … | … |

As shown in Table 3, high-ranking entries have far higher values for most indicators than low-ranking entries, except for the tag count indicator, where there is

no significant difference among entries. This aligns with the indicator weight results obtained from the entropy weight method, where tag count has the lowest weight among the 12 indicators.

CCTV, as the entry ranked 3rd in comprehensive scoring, compared with the top 5 entries, has no advantage in most indicators except for tag count and edit count, but can rank 3rd simply by having 2,626 inlinks, demonstrating the importance of a privileged position in the entry link network.

4.3 Discussion

Based on the entry quality assessment using the entropy weight method and grey relational analysis, clear quality differences between different entries can be demonstrated. This study not only uses feature indicators from related domestic and international research but also additionally incorporates user behavior indicators. The results show that these indicators have influence comparable to network link indicators and far higher than other indicators in comprehensive entry quality evaluation. Therefore, incorporating user behavior data into encyclopedia entry quality evaluation research is meaningful.

This study, based on the theoretical foundation of domestic and international encyclopedia entry research, extracted 12 quantitative indicators across four dimensions for Baidu Encyclopedia entries and conducted a comprehensive evaluation. Incorporating user behavior dimension indicators into the entry evaluation indicator system is feasible and necessary. The three user behavior indicators all have relatively high entropy weights. The selected indicators are all quantifiable indicators based on statistics. Future research will explore how to combine qualitative features such as text richness, rigor, and editor reputation with quantitative features.

References

[?] WIKIPEDIA. Article assessment[EB/OL]. [2018-10-29]. https://en.wikipedia.org/wiki/Wikipedia:Article_{

[?] STVILIA B, FARAJ A, YI Y J. Issues of cross-contextual information quality evaluation-the case of arabic, english, and korean Wikipediaes[J]. Library & information science research, 2009, 31(4): 232-239.

[?] LICHTENSTEIN S, PARKER C M. Wikipedia model for collective intelligence: a review of information quality[J]. International journal of knowledge and learning, 2009, 5(3/4): 254-272.

[?] ARAZY O, NOVO O, PATTERSON R, et al. Information quality in Wikipedia: the effects of group composition and task conflict[J]. Journal of management information systems, 2011, 27(4): 71-95.

[?] LIU J, RAM S. Who does what: collaboration patterns in the Wikipedia and their impact on article quality[J]. ACM transactions on management information systems, 2011, 2(2): 1-23.

[?] YAARI E, BARUCHSON-ARIBI S, BAR-ILAN J. Information quality assessment of community generated content: a user study of Wikipedia[J]. Journal of information science, 2011, 37(5): 487-498.

[?] OSMAN K. The role of conflict in determining consensus on quality in Wikipedia articles[C]//International symposium on open collaboration. New York: ACM, 2013: 1-6.

[?] LEWANDOWSKI D, SPREE U. Ranking of Wikipedia articles in search engines revisited: fair ranking for reasonable quality?[J]. Journal of the American Society for Information Science & Technology, 2014, 62(1): 117-132.

[?] DING JINGDA. Research on heuristic evaluation framework for Wikipedia entry information quality[J]. Library and Information Knowledge, 2014(2): 11-17.

[?] ZHAO WENXUAN. Analysis of online encyclopedia resource quality[J]. Library Theory and Practice, 2016(4): 38-42.

[?] HUANG LINGHE. Research on dynamic evolution of online encyclopedia information quality[J]. Information Theory and Practice, 2018(2): 127-131.

[?] LIH A. Wikipedia as participatory journalism: reliable sources?[C]//International symposium on online journalism. Austin: ISOJ, 2004: 16-17.

[?] CHEVALIER F, HUOT S, FEKETE J D. Wikipedia viz: conveying article quality for casual Wikipedia readers[C]//Pacific visualization symposium. Taipei: IEEE, 2010: 215-222.

[?] JIN YAN, ZHOU TING, ZHAN LIHUA. Research on quality evaluation system of collaborative content creation system based on analytic hierarchy process: taking Baidu Encyclopedia as an example[J]. Library Theory and Practice, 2015(7): 41-45.

[?] ZHANG BO, QIAO HUAN. Research on content quality assessment model of collaborative knowledge production community: taking Wikipedia as an example[J]. Modern Intelligence, 2015, 35(10): 17-22.

[?] JIN YAN. Quality assurance measures for collaborative content creation system based on user experience: taking Baidu Encyclopedia as an example[J]. Information Theory and Practice, 2016, 39(3): 6-9.

[?] WARNCKE-WANG M, COSLEY D, RIEDL J. Tell me more: an actionable quality model for Wikipedia[C]//International conference on Web intelligence. Washington DC: IEEE, 2006: 81-87.

[?] STVILIA B, TWIDALE M B, SMITH L C, et al. Assessing information quality of a community-based encyclopedia[C]//International conference on information quality. Cambridge: ICIQ, 2005: 442-454.

[?] BLUMENSTOCK J E. Size matters: word count as a measure of quality on Wikipedia[C]//International conference on World Wide Web. New York: ACM,

2008: 1095-1096.

[?] DALIP D H, CRISTO M. Automatic quality assessment of content created collaboratively by web communities: a case study of Wikipedia[C]//ACM/IEEE-CS joint conference on Digital libraries. New York: ACM, 2009: 295-304.

[?] CALZADA G D L, DEKHTYAR A. On measuring the quality of Wikipedia articles[C]//Workshop on Information credibility. New York: ACM, 2010: 11-18.

[?] QIU JIANGNAN, WENG NAN, XU SHENGGUO. Wikipedia page information quality evaluation based on C4.5[J]. Journal of the China Society for Scientific and Technical Information, 2012, 31(12): 1259-1264.

[?] LI XINYI. Research on online encyclopedia entry quality evaluation[D]. Changsha: National University of Defense Technology, 2014.

[?] TONG ZHAOJUAN, XU XIN. Automated evaluation of Baidu Encyclopedia webpage quality[J]. Journal of Information Resources Management, 2015(2): 63-69.

[?] XIAO KUI, LI BING, WU TIANJI. Wikipedia entry quality evaluation method based on user behavior analysis[J]. Journal of Intelligence, 2015(5): 185-189.

[?] HU M, LIM E P, SUN A, et al. Measuring article quality in Wikipedia: models and evaluation[C]//Conference on information and knowledge management. New York: ACM, 2007: 243-252.

[?] ZENG H, AL-HOSSAINI M A, DING L, et al. Computing trust from revision history[C]//International conference on privacy, security and trust: bridge the gap between PST technologies and business services. New York: ACM, 2006: 1-1.

[?] LIM E P, VUONG B, LAU H W, et al. Measuring qualities of articles contributed by online communities[C]//International symposium on online journalism. Austin: ISOJ, 2004: 16-17.

[?] DONDIO P, BARRETT S. Computational trust in web content quality: a comparative evaluation on the Wikipedia project[J]. Informatica, 2007, 31(2): 151-160.

[?] WOHNER T, PETERS R. Assessing the quality of Wikipedia articles with lifecycle based metrics[C]//International symposium on Wikis. Orlando: DBLP, 2009: 1-10.

[?] JAVANMARDI S, LOPES C. Statistical measure of quality in Wikipedia[C]//Proceedings of the first workshop on social media analytics. New York: ACM, 2010: 132-138.

[?] YU S, YOSHIKAWA M. Assessing quality score of Wikipedia article using mutual evaluation of editors and texts[C]//ACM International conference on information & knowledge management. New York: ACM, 2013: 1727-1732.

[?] LI X, TANG J, WANG T, et al. Automatically assessing Wikipedia article quality by exploiting article-editor networks[C]//Advances in Information Retrieval, ECIR. Gewerbestrasse: Springer, 2015: 574-580.

[?] WILKINSON D M, HUBERMAN B A. Cooperation and quality in Wikipedia[C]//International symposium on Wikis. New York: ACM, 2007: 157-164.

[?] XIAO KUI, LUO BAOSHAN, WU TIANJI. A domain-specific network document resource quality evaluation method: taking Wikipedia as an example[J]. Information Theory and Practice, 2016, 39(12): 120-123.

[?] WU K, ZHU Q, ZHAO Y, et al. Mining the factors affecting the quality of Wikipedia articles[C]//International conference of information science and management engineering. Washington DC: IEEE, 2010: 343-346.

[?] LIU SIFENG, CAI HUA, YANG YINGJIE, et al. Research progress on grey relational analysis models[J]. System Engineering Theory and Practice, 2013, 33(8): 2041-2046.

[?] NI JIUPAI, LI PING, WEI CHAOFU, et al. Regional land development and consolidation potential evaluation based on AHP and entropy weight method[J]. Transactions of the Chinese Society of Agricultural Engineering, 2009, 25(5): 202-209.

Author Contributions

DONG Xiangxiang: Paper writing;
HE Zhenyu: Proposed research ideas, designed experimental scheme, data collection and processing, conducted experiments, paper writing;
ZHU Qinghua: Provided paper revision suggestions.

Note: Figure translations are in progress. See original paper for figures.