Research on Big Data-Based Knowledge Sharing Methods: Postprint

Authors: Gu Xinjian, Ma Buqing, Dai Feng

Date: 2023-10-08T00:00:00+00:00

Abstract

[Purpose/Significance] Knowledge evaluation and the assessment of knowledge workers' competency levels represent two major challenges confronting Chinese enterprises in their pursuit of innovation. [Method/Process] This study proposes a big data-based knowledge sharing methodology, which aims to acquire knowledge big data and knowledge evaluation big data through the construction of relevant systems, standards, and platforms, thereby achieving full traceability of the entire knowledge management process and transparency in both enterprise employees' knowledge sharing and their knowledge levels. [Results/Conclusion] On this basis, enterprises provide fair and impartial incentives that encourage employees to willingly share knowledge and actively participate in knowledge evaluation, ultimately establishing a transparent and fair enterprise environment that fosters an orderly knowledge network and a positive atmosphere where employees enthusiastically engage in knowledge sharing.

Full Text

Research on Knowledge Sharing Methods Based on Big Data

Gu Xinjian, Ma Buqing, Dai Feng
Institute of Mechanical Engineering, Zhejiang University, Hangzhou 310027

Abstract

[**Purpose/Significance**] Knowledge evaluation and the assessment of knowledge workers' expertise levels represent two major challenges confronting Chinese enterprises in their innovation efforts. [**Method/Process**] This paper proposes knowledge sharing methods based on big data, attempting to obtain large-scale knowledge data and knowledge evaluation data through the construction of relevant systems, standards, and platforms. This approach enables traceability throughout the entire knowledge management process and transparency in enterprise employees' knowledge sharing activities and expertise levels. [**Result/Conclusion**] On this foundation, enterprises can provide fair and equitable incentives, encouraging employees to willingly share knowledge and actively participate in knowledge evaluation. Ultimately, this establishes a transparent and fair enterprise environment, forming an orderly knowledge network and a positive atmosphere where employees actively share knowledge.

Keywords: big data, knowledge sharing, knowledge evaluation, knowledge network, knowledge management

1 Significance of Knowledge Sharing Based on Big Data

This paper advances the following perspective: China's current economy must undergo transformation and upgrading to escape the middle-income trap and achieve sustainable development, transitioning from a manufacturing giant to a manufacturing powerhouse. Such transformation is exceedingly difficult—few nations or regions worldwide have successfully upgraded from developing to developed status. The key to transformation lies in innovation, and integrity at both enterprise and individual levels constitutes a critical factor affecting innovation and knowledge sharing. The internet and big data can track credit histories, helping to build an integrity-based society. Knowledge sharing based on big data facilitates collaborative innovation and supports China's economic transformation. As shown in [Figure 1: see original paper], this viewpoint receives support from several sources.

1.1 The World Bank's Perspective: Developing Countries' Poverty Stems Primarily from Attribute Knowledge Gaps

The World Bank distinguishes between technical knowledge and attribute knowledge. Technical knowledge refers to technological know-how, while attribute knowledge concerns product quality, borrower creditworthiness, or employee diligence. The uneven distribution of technical knowledge is termed the "knowledge gap," while imbalances in attribute knowledge are called "information problems." As illustrated in [Figure 2: see original paper] [1], the World Bank's 1998/99 World Development Report argues that opportunistic behavior is difficult to succeed in both traditional societies with low personnel mobility and in knowledge economies where computer networks can trace credit histories. However, intermediate societies may not share this characteristic [1].

1.2 The Academic Perspective: The Third Hand—Network-Based Collaborative Big Data [2]

Adam Smith asserted in 1776 that in pursuing private goals, individuals are guided by an "invisible hand" that achieves optimal social resource allocation and enhances social welfare. This first invisible hand represents market forces, which can fail because a key assumption—transparent information—rarely holds in reality. Markets are vast, and countless customers and operators always possess opaque, asymmetric information about products, leading to blind production and fraudulent behavior by selfish operators.

The second invisible hand comprises social and moral norms, based on the theory that "people have a natural altruistic tendency." However, at present, individuals' selfish tendencies often overwhelm altruistic ones, rendering moral norms relatively powerless.

Network-based collaborative big data can create a transparent society, making various transactions increasingly visible and transforming the traditionally flexible atmosphere governed by moral norms into a rigid environment under sunlight, where social supervision ruthlessly eliminates bad actors. Cooperative competition and win-win strategies become optimal choices, and enterprise goals must align with social objectives. Consequently, network-based collaborative big data is becoming the third hand, following market laws and moral norms. See [Figure 3: see original paper].

1.3 The Chinese Government's Perspective: Big Data as a New Driver of Economic Transformation

(1) On September 30, 2013, the Political Bureau of the CPC Central Committee held its ninth collective study session on implementing innovation-driven development strategies. General Secretary Xi Jinping emphasized that implementing innovation-driven development strategies determines the future and destiny of the Chinese nation. China must keenly grasp global sci-tech innovation trends and seize opportunities from the new round of sci-tech revolution and industrial transformation, making innovation-driven development a major future strategy [3]. New-generation information technology is the main driving force of this new sci-tech revolution and will support knowledge sharing and collaborative innovation based on big data. See [Figure 4: see original paper].

(2) On May 8, 2015, the State Council released *Made in China 2025*, with the main theme of promoting manufacturing innovation and quality-efficiency improvement. It emphasizes accelerating the integration of new-generation information technology with manufacturing, advancing intelligent manufacturing, strengthening industrial foundation capabilities, improving comprehensive integration levels, perfecting multi-level talent systems, and promoting industrial transformation to achieve the historic leap from manufacturing giant to manufacturing powerhouse [4].

(3) On August 31, 2015, the State Council issued the *Action Outline for Promoting Big Data Development* (hereinafter "the Outline") as Document No. 50 [2015]. The Outline defines big data as data collections characterized by large volume, diverse types, fast access speeds, and high application value. It is rapidly evolving into a field that collects, stores, and correlatively analyzes massive, dispersed, and multi-format data to discover new knowledge and create new value. The Outline identifies big data as a new driver of economic transformation, stating that data flows will guide technology, material, capital, and talent flows, profoundly affecting social division of labor and collaboration patterns and promoting intensive and innovative production organization methods. Big data changes traditional production methods and economic operation mechanisms, significantly improving economic efficiency. It continuously stimulates business model innovation and spawns new formats, becoming a crucial driver for business innovation and value enhancement in emerging fields like the internet. The Outline's Mass Innovation Big Data Project includes knowledge service big data applications, utilizing big data and cloud computing technologies to integrate knowledge across fields, build hierarchical, comprehensive, and accurate knowledge resource libraries, establish national knowledge service platforms and resource centers, and form a national knowledge service system covering major economic fields.

2 Different Knowledge Sharing Approaches

Different enterprises employ vastly different knowledge sharing methods, as shown in [Figure 5: see original paper].

(1) Tacit Knowledge Inheritance Enterprises: These rely on oral or hands-on knowledge transfer—the "master-apprentice" model that has dominated human knowledge inheritance since antiquity. The problem: inefficient, limited-scope knowledge transfer cannot meet today's complex product innovation demands. While some tacit knowledge "can only be understood, not expressed," requiring such approaches, the efficiency remains inadequate.

(2) Mandatory Knowledge Contribution Enterprises: These enforce commands requiring employees to publish a certain number of knowledge items or suggestions annually, or to summarize experiences after completing work. While somewhat effective for knowledge accumulation, employees often consider their own interests and are reluctant to upload valuable knowledge—their "bread and butter"—fearing "teaching the apprentice starves the master." Consequently, uploaded knowledge often becomes "junk knowledge."

(3) Simple Incentive Knowledge Sharing Enterprises: These provide simple statistics-based incentives, such as converting published knowledge items into points for economic and spiritual rewards. Employee responses include publishing low-value knowledge, publishing minimally, or not publishing at all, as the incentives fail to reflect true knowledge value. For instance, a Chinese research institute offered 400 RMB per knowledge item, yet enthusiasm remained low.

(4) Transparent and Fair Enterprises: These use information technology to make employees' knowledge sharing degrees and processes transparent, providing the fairest possible incentives accordingly. This creates an internal knowledge market. The difficulty lies in evaluating knowledge and sharing degrees accurately. The internet and big data provide opportunities to establish transparent, fair enterprises. Siemens' ShareNet connects global experts, allowing them to share and expand knowledge; each published knowledge item receives comments from the entire community, and contributors receive ShareNet "shares" or reward points for valuable sharing [6-7]. Similarly, China Telecom Shanghai Research Institute's Innovation Dream Factory based on Web 2.0 supports employees in publishing ideas, scoring them, conducting reviews, classifying, systematically organizing, and mining information, supporting the complete process from idea generation to productization [8].

Transparent and fair enterprises encourage active knowledge contribution and collaborative innovation, gradually forming knowledge sharing habits and culture, enabling enterprises to evolve into happy enterprises.

(5) Happy Enterprises: Employees feel genuinely happy, consider the enterprise as home, and selflessly contribute knowledge. Happy enterprises rely on culture and values to motivate knowledge contribution—the ideal state. Japan's Kazuo Inamori culture represents this type, influencing some Chinese private enterprises like Ningbo Zhongxing Precision. Shenzhen Huawei is also building such an enterprise. However, becoming a happy enterprise is extremely difficult under current market and institutional environments, requiring long-term, unremitting effort. Most importantly, enterprise leaders must first be selfless with strong personal charisma. Partnership systems, Inamori's "Amoeba Management" philosophy, plus internet and big data-driven enterprise transparency can promote happy enterprise growth, creating "spiritual, destiny, goal, and interest communities" that inspire employees to unleash their potential.

3 Knowledge Sharing Methods Based on Big Data

Knowledge is rapidly generated every moment in enterprises, distributed across different media and stored in various formats. After long-term accumulation, enterprises accumulate considerable knowledge data—exhibiting big data's 3V characteristics: volume, velocity, and variety—termed "knowledge big data." Simultaneously, employee knowledge evaluations are rapidly generated, distributed across different media and stored in various formats. After long-term accumulation, these evaluations also form considerable data volume, called "knowledge evaluation big data."

Big data technology helps solve two major challenges in current knowledge sharing: (1) Knowledge explosion creates disorderly knowledge repositories, reducing utilization efficiency. Employee behavior big data can automatically evaluate knowledge value and relationships, while mass employee collaborative evaluations form evaluation big data that filters out useless, duplicate, and outdated knowledge, ordering the knowledge big data into an orderly knowledge network. (2) Employees lack willingness to share and evaluate knowledge. Based on the quantity and quality (mass evaluation results) of shared knowledge, enterprises can evaluate and assess employees' knowledge sharing degrees and expertise levels, providing corresponding incentives.

The basic process of knowledge sharing methods based on big data is shown in [Figure 6: see original paper], forming a closed loop. Knowledge big data includes: (1) internal knowledge (employee experiences and ideas, work records and summaries, failure cases) and (2) external knowledge (design manuals, patents, standards, journal literature, online articles). Knowledge evaluation big data includes employee knowledge usage behavior data such as comments, ratings, downloads, reading, and citations.

3.1 Mass Knowledge Sharing

Through appropriate systems, standards, and platforms, mass knowledge sharing is supported, making individual knowledge organizational and tacit knowledge explicit.

(1) System Construction: Includes: (i) Enterprise internal intellectual property system—internalizing external IP systems; employees own copyright for self-created knowledge published in the enterprise repository; whoever publishes first owns the copyright; if knowledge is unpublished and others publish it, copyright belongs to the publisher; if the knowledge generates benefits in future product/technology development, the original publisher receives corresponding compensation. (ii) Knowledge disclosure system—determining knowledge disclosure scope and usage rules. (iii) Employee knowledge sharing system—mandating annual knowledge sharing quantities for different levels and positions, linked to annual assessments and promotions. (iv) Position and process knowledge construction system—requiring responsible persons to organize and improve essential position and process knowledge to quality standards, as this knowledge serves as infrastructure for new employee training and veteran employee development.

(2) Standard Construction: Includes: (i) Knowledge document template standards; (ii) Knowledge description specification standards; (iii) Knowledge citation standards; (iv) Knowledge classification standards; (v) Knowledge integration standards.

(3) Platform Construction: Includes enterprise-level knowledge sharing subsystems at individual, team, department, enterprise, and group levels. Employees can publish knowledge in different subsystems based on confidentiality levels and needs. These subsystems are integrated into one platform that not only supports employee knowledge publication but also integrates knowledge from different enterprise information systems such as PDM/PLM, ERP, and maintenance management systems.

3.2 Mass Knowledge Evaluation

Through appropriate systems, standards, and platforms, mass knowledge evaluation is supported to promote knowledge ordering.

(1) System Construction: Includes: (i) Employee knowledge evaluation system—mandating evaluation tasks for different levels and positions, particularly regarding knowledge classification, association, network construction, and identification of core and key knowledge; (ii) Knowledge collaborative evaluation system; (iii) Patent collaborative evaluation and patent map collaborative establishment system.

(2) Standard Construction: Includes: (i) Knowledge evaluation standards; (ii) Knowledge ontology standards; (iii) Knowledge network standards; (iv) Knowledge semantic model standards.

(3) Platform Construction: Includes: (i) Knowledge evaluation subsystem based on user behavior; (ii) Knowledge ontology evaluation subsystem based on user behavior; (iii) Knowledge ontology management subsystem; (iv) Knowledge lifecycle tracking subsystem.

Based on mass knowledge evaluation data, knowledge value can be assessed. Knowledge value evaluation considers both professionalism factors and popularity factors, expressed as:

$$V_k = f(P, G)$$

where V_k represents the evaluation value of specific knowledge k, P represents professionalism factors (e.g., user ratings), and G represents popularity factors (e.g., reading, downloading, commenting, tagging, recommending, collecting, forwarding).

The professionalism factor is calculated based on user ratings. For knowledge k, the professionalism factor can be expressed as:

[Formula representation]

where X_{ik} represents user i's rating of knowledge k (with $X_{ik} \in \{1, 2, 3, 4, 5\}$), and $\bar{X}_k$ represents the average rating of knowledge k by all users.

This calculation doesn't consider user rating weights and habitual factors. Introducing user rating weights, converting user i's rating of knowledge k to standard scores and normalizing, the professionalism factor calculation can be optimized as:

[Formula representation]

where $\bar{X}_i$ represents the average rating by user i across all knowledge, σ_i^X represents the standard deviation of user i's ratings, and W_i is user i's rating weight.

The popularity factor is calculated based on user operations such as reading, downloading, commenting, tagging, recommending, collecting, and forwarding. For knowledge k, the popularity factor can be expressed as:

[Formula representation]

where Y_{ki} represents the number of operations by user i on knowledge k. Here, knowledge comments are temporarily grouped with other operations due to lack of effective semantic understanding means. $\bar{Y}_i$ represents the average number of operations by user i across all knowledge, and σ_i^Y represents the standard deviation.

When calculating the popularity factor, time factors must be considered. Since more recent daily evaluation operations better reflect knowledge popularity, the formula can be optimized by introducing time decay:

[Formula representation]

where *date* represents the number of days elapsed since the user's specific evaluation operation, and β is the time decay factor.

Based on these formulas, knowledge value evaluation results are calculated as:

[Formula representation]

where δ_p and δ_g are weight coefficients for each factor, with $\delta_p + \delta_g = 1$.

User ratings reflect knowledge professionalism, while user operations within 30 days reflect knowledge popularity. These calculation methods comprehensively consider both professionalism and popularity factors, ensuring that knowledge with higher professional levels and greater user popularity receives higher value evaluations.

3.3 Knowledge Sharing Evaluation

Through appropriate systems, standards, and platforms, employees' knowledge sharing levels and expertise levels are evaluated to help enterprises better understand knowledge workers' performance. This not only helps fully exploit and develop employee potential but also facilitates proactive knowledge delivery to employees.

(1) System Construction: Includes: (i) Knowledge sharing evaluation norms; (ii) Employee knowledge sharing evaluation credit norms.

(2) Standard Construction: Includes: (i) Employee knowledge sharing evaluation standards; (ii) Employee expertise level evaluation standards; (iii) Employee knowledge sharing evaluation weight standards.

(3) Platform Construction: Includes: (i) Employee knowledge sharing evaluation subsystem; (ii) Employee expertise level evaluation subsystem; (iii) Team knowledge sharing evaluation subsystem.

Employee knowledge sharing level and expertise level evaluation results consider: employee professional skill level factors, knowledge sharing participation factors, and knowledge sharing contribution factors, expressed as:

$$W_i = f(T, J, C)$$

where T represents employee professional skill level factors (e.g., professional titles, position levels), J represents knowledge sharing participation factors (e.g., user operations on knowledge), and C represents knowledge sharing contribution factors (e.g., value of shared knowledge). These three main factors are weighted.

Enterprises can set grading standards for employee professional skill factor T based on actual position ratings and titles, giving experts higher T values and newcomers lower values, with $T_i \in [0, 1]$.

Employee knowledge sharing participation factor J reflects the degree of participation in knowledge sharing evaluation, calculated based on employee operations (reading, downloading, commenting, tagging, recommending, collecting, forwarding, rating):

[Formula representation]

where J_i represents the number of operations by employee i within 30 days, and $\bar{J}$ represents the average number of operations by all employees within 30 days.

Employee knowledge sharing contribution factor C reflects the contribution degree of knowledge sharing behavior, calculated as:

[Formula representation]

where $U_{\text{rank}i}$ represents the ranking of the average rating of knowledge shared by employee i among all employees, and n is the total number of employees in the system.

Based on these formulas, employee knowledge scoring weight is calculated as:

[Formula representation]

where δ_t, δ_j, and δ_c are weight coefficients for each factor, with $\delta_t + \delta_j + \delta_c = 1$.

These calculation methods comprehensively consider employee professional skill level, knowledge sharing participation, and contribution factors, ensuring that employees with higher professional levels, better sharing performance, and more active participation receive higher evaluations [9].

3.4 Knowledge Sharing Incentives

Through appropriate systems, standards, and platforms, employee knowledge sharing behaviors are evaluated to encourage sustained sharing of high-quality knowledge and active participation in knowledge evaluation.

(1) System Construction: Includes: (i) Knowledge sharing assessment system; (ii) Knowledge sharing reward system; (iii) Knowledge sharing title management system.

(2) Standard Construction: Includes: (i) Employee knowledge sharing assessment standards; (ii) Employee knowledge sharing incentive standards.

(3) Platform Construction: Includes: (i) Employee knowledge sharing assessment subsystem; (ii) Employee knowledge sharing incentive subsystem. The knowledge sharing incentive platform is not independent but part of the knowledge sharing platform, integrated with human resource management, production management, and product data management systems.

3.5 Mass Knowledge Application

Through appropriate systems, standards, and platforms, knowledge application is supported.

(1) System Construction: Includes: (i) Proactive knowledge delivery norms—pushing needed knowledge to users based on their research fields and levels; (ii) Knowledge application norms—requiring employees to search and apply relevant knowledge from the repository during project pre-research, feasibility analysis, and solution selection before extending outside the repository, and publishing newly discovered knowledge in the repository.

(2) Standard Construction: Includes: (i) Proactive knowledge delivery standards; (ii) Knowledge search standards; (iii) Technology evolution map standards; (iv) Technology roadmap standards.

(3) Platform Construction: Includes proactive knowledge delivery, knowledge search, knowledge download and reading, technology evolution map construction based on knowledge networks, and technology roadmap construction based on knowledge networks.

The new round of industrial revolution represents an intersection point for China between new-generation information technology and enterprise transformation. Seizing this opportunity may enable China's manufacturing sector to break through and escape the "middle-income trap." The key to transformation is innovation. Big data-based knowledge management can promote knowledge sharing and collaborative innovation, truly creating a mass innovation scenario.

Big data-based knowledge management features: (1) Big data helps achieve transparency and traceability throughout the knowledge management process;

(2) Big data-based knowledge management covers larger scopes with more knowledge accumulation and full participation; (3) Knowledge big data and knowledge evaluation big data facilitate employee knowledge sharing and expertise level evaluation, promoting knowledge sharing; (4) Knowledge usage big data helps automate knowledge ordering and improve utilization efficiency; (5) Big data supports collaborative IP protection, promoting collaborative innovation; (6) Big data supports establishing complete enterprise knowledge systems, position-level knowledge networks, process-level knowledge networks, and strategic-level knowledge networks; (7) Big data helps enterprises comprehensively evaluate and fully utilize employees; (8) Big data helps enterprises find the most suitable partners.

The future enterprise knowledge management should be transparent, fair, and characterized by conscious, active participation. Knowledge sharing should become everyone's habit. Our vision: people share knowledge without reservation and collaborate innovatively; Lei Feng-like contributors no longer suffer losses; when veteran employees leave, their knowledge remains; when new employees arrive, systematic knowledge is available for learning; enterprise knowledge networks become smarter with use; enterprise and employee knowledge sharing and collaborative innovation performance become transparent and traceable; employees grow through knowledge sharing; enterprises strengthen through knowledge accumulation.

References

[1] World Bank. *1998/99 World Development Report: Knowledge for Development* [M]. Beijing: China Financial & Economic Publishing House, 1999.
[2] Yang Peifang. Social Collaboration: The Third Force in the Information Age [N]. *Guangming Daily*, 1999-10-22(6).
[3] Xinhua News Agency. Political Bureau of CPC Central Committee Holds Ninth Collective Study Session, Xi Jinping Presides [EB/OL]. [2015-10-01]. http://www.gov.cn/ldhd/2013-10/01/content_{2499370}.htm.
[4] State Council. *Made in China 2025* [EB/OL]. [2015-05-08]. http://news.china.com/domestic/945/20150519/
[5] State Council. *Action Outline for Promoting Big Data Development* [EB/OL]. [2015-08-31]. http://politics.people.com.cn/n/2015/0905/c1001-27545655.html.
[6] Yu Feiliu. Research on Application of Knowledge Management 2.0 in Siemens [D]. Shanghai: East China Normal University, 2011.
[7] Liu Jun. What Can Siemens and Xerox Teach Manufacturing Enterprises? [EB/OL]. [2005-11-21]. http://www.amteam.org.
[8] Liu Xianghua, Xia Jinghua. Wisdom Exchange: Innovation Dream Factory Supports Innovation 2.0 Implementation [EB/OL]. [2015-09-16]. http://group.vsharing.com/Article.aspx?aid=909575.
[9] Dai Feng. Research on Knowledge Network Theory and Integrated Application for Complex Product Development Processes [D]. Hangzhou: Zhejiang University, 2015.

Author Contributions:
Gu Xinjian: Wrote the paper, proposed the main ideas;
Ma Buqing: Participated in writing and data collection;
Dai Feng: Participated in writing and algorithm research.

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