Enterprise Knowledge Extraction Theory and Practice Research Postprint

Authors: Wu Qinghai

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

Abstract

[Purpose/Significance] As enterprises increasingly demand higher quality of organizational knowledge content, researching knowledge extraction methods to promote the mining, processing, and refinement of organizational knowledge, and to enhance the replication, promotion, and application of organizational knowledge, holds significant importance. [Method/Process] By studying academic literature and industry practices related to knowledge extraction, this paper analyzes and clarifies the concept of knowledge extraction, proposes the PREFS® process method and STAR® content model for knowledge extraction, and applies them in enterprise practice for empirical validation. [Results/Conclusion] Knowledge extraction work is essential for enterprise knowledge management, and the PREFS® process method and STAR® content model for knowledge extraction have practical value. Through knowledge extraction, not only can organizational best practices be transformed into knowledge products that inherit organizational wisdom, but it can also serve as a lever to promote the implementation or expansion of organizational knowledge management strategies.

Full Text

The Study on Theory and Practice of Enterprise Knowledge Extraction

Wu Qinghai
SunXZ Network Technology (Beijing) Co., Ltd., Beijing 100083

Abstract

[Purpose/Significance] With enterprises placing increasingly high demands on the quality of organizational knowledge content, research on knowledge extraction methods to promote organizational knowledge mining, processing, and

refinement, and to enhance the level of organizational knowledge replication, promotion, and application, holds significant importance. [**Method/Process**] Through research on relevant academic literature and industry practices in knowledge extraction, this paper clarifies the concept of knowledge extraction, proposes the PREFS® process method and STAR® content model, and applies them in enterprise practice for empirical validation. [**Result/Conclusion**] Knowledge extraction work is essential for enterprise knowledge management, and both the PREFS® process method and STAR® content model possess practical value. Through knowledge extraction, not only can organizational best practices be transformed into knowledge products that preserve organizational wisdom, but they can also serve as a lever to drive the implementation or expansion of organizational knowledge management strategies.

Keywords: knowledge extraction; experience extraction; PREFS® process method; STAR® content model; knowledge management

---

In recent years, discussions about organizational knowledge extraction and experience extraction have begun to increase in the business community, driven primarily by two fields: education and training, and knowledge management. The former's focus on experience extraction stems from training courses and related teaching case development, while the latter has long embraced concepts and traditions of knowledge processing, knowledge refinement, and knowledge harvesting. With emerging demands for stronger business orientation and greater attention to content quality, the concept of knowledge extraction has gained prominence. This paper reviews relevant typical models and methods, proposes the PREFS® process method and STAR® content model for knowledge extraction, and validates and refines them through practical application in enterprises.

1.1 Academic Research Review

Taking knowledge extraction as the core research object, we selected the China National Knowledge Infrastructure (CNKI) China Academic Journals Network Publishing Database as our data source, retrieving literature with titles or keywords containing "knowledge extraction," "experience extraction," "knowledge processing," "knowledge refinement," or "knowledge harvesting." The search formula was: TI="knowledge extraction"+"experience extraction"+"knowledge processing"+"knowledge refinement"+"knowledge harvesting" OR KY="knowledge extraction"+"experience extraction"+"knowledge processing"+"knowledge refinement"+"knowledge harvesting." After removing duplicate and irrelevant documents, 30 valid articles were obtained, and an annual distribution map of knowledge extraction literature was drawn (see [Figure 1: see original paper]). The literature spans from 1979 to 2015. As shown in Figure 1, research on knowledge extraction remains limited. Literature analysis reveals that relevant studies either relate to modern teaching and learning [**?**] or to skill training [**?**], representing research outcomes from the education and

training community. The knowledge management community rarely mentions knowledge extraction; although some early work touched upon it [**?**], it lacked substantive content on knowledge extraction. Recently, researchers have begun studying knowledge refinement and knowledge construction [**?**], which is undoubtedly a positive sign.

1.2 Industry Practice Status

In enterprise practice, numerous activities require knowledge extraction, including project summaries, case development, experience refinement, thematic research, work summaries, report writing, and patent analysis. These activities involve mining and refining tacit knowledge accumulated during processes, organizing and processing explicit knowledge, and ultimately forming high-quality knowledge products that can be promoted, replicated, and applied. The resulting knowledge products, presented through text, audio, images, and other media, hold significant value, prompting many researchers and practitioners to specialize in developing and teaching these methods. Representative domestic offerings include Li Wende's "Organizational Experience Refinement" course, Song Hongbo's "Organizational Experience Extraction" course, Liu Baigong's "Organizational Experience Extraction and Efficient Course Development" course, Wang Xingquan's "Extracting Organizational Best Practices" course, and Zhang Lizhi's "Five-Step Knowledge Extraction" course. These instructors primarily come from training backgrounds, and their content typically approaches the topic from training management and course development perspectives.

Currently, few approaches emerge from the knowledge management field itself. However, since knowledge management inherently includes knowledge processing, knowledge refinement, and knowledge harvesting, many knowledge management practitioners unconsciously apply relevant methods in practice and have gradually begun reflecting on and summarizing knowledge extraction approaches. For instance, Qiu Zhaoliang authored the monograph *Fupan+: Transforming Experience into Capability* [**?**] and developed corresponding courses. The author also shared on the theme of "Enterprise Knowledge Extraction Methods and Practical Operations" in WeChat friend circles, sparking enthusiastic discussion and active reflection among practitioners.

1.3 Conceptual Analysis of Knowledge Extraction

Extraction is originally a chemical concept referring to a method of using a solvent to extract a solute from another solvent based on differences in solubility [**?**]. For example, mixing iodine water with carbon tetrachloride or benzene, shaking, and then distilling to obtain iodine crystals. Extraction is one of the methods used in organic chemistry laboratories to purify and refine compounds. Knowledge extraction, as the name suggests, is the process of extracting and refining essential knowledge from large amounts of data, information, text, and experience through appropriate methods and tools (see [Figure 2: see original paper]). In the IT domain, knowledge extraction refers to extracting new knowl-

edge content that can be read and understood by machines from structured or unstructured data, similar to information extraction in natural language processing.

In current practice, the conceptual scope of knowledge extraction has gradually expanded to include the integration, processing, and refinement of both tacit and explicit knowledge. Mining and refining experience represents the process of making tacit knowledge explicit, while organizing and processing documents represents the process of standardizing explicit knowledge. Taking experience as an example, experience refers to the understanding people gain about the phenomena and external relationships of objective things through contact with them. Where does experience come from? It emerges from problem identification and resolution, and from pioneering exploration. Experience often represents the essence behind phenomena, frequently appearing in the form of "reasons"—the rationale for success—and reflects certain regular patterns of things. Experience typically exists as tacit knowledge in people's minds. If knowledge extraction can make experience explicit and form case-based 沉淀, it becomes easier to inherit, reference, and create greater value.

2.1 Typical Methods and Models

Several knowledge extraction-related methods and models have emerged in the industry, notably Lenovo's Fupan, the U.S. Army's After Action Review (AAR), Huawei's knowledge harvesting, Ikujiro Nonaka's SECI model, and Shanghai Taoyu Consulting's ERMS model. The following provides brief analyses of the first three methods, which emphasize practical operation.

(1) Lenovo Fupan. "Fupan" originates from the Go game terminology and has gradually evolved into one of the important methodologies in Lenovo's corporate culture [**?**]. Lenovo considers Fupan the best learning approach. After completing work, it is necessary to review and rehearse it, continuously examining and correcting goals, analyzing gains and losses in the process, and deepening understanding and summarizing patterns. This work method, 延续 from Liu Chuanzhi's era to the present, has become an iconic method within Lenovo. Lenovo's Fupan consists of four steps: Review goals; Evaluate results; Analyze reasons; Summarize patterns. Using the Fupan template (see [Figure 3: see original paper]), practitioners simply need to fill in the blanks according to template requirements at each step to apply the method.

(2) U.S. Army AAR. AAR is an important step in knowledge management's "learn before, learn during, learn after" framework. AAR is a post-task review method originally adopted by the U.S. Army, defined as: a professional discussion of an event, focusing on performance standards, enabling participants to discover for themselves what happened, why it happened, and how to maintain strengths and improve weaknesses [**?**]. When conducting AAR, the U.S. Army typically asks four important questions [**?**]: What were the original mission and objectives? What actually happened? Why did it happen? What will we

do next time? Through this question-answering process, experience is extracted and, depending on circumstances, transformed into U.S. Army management systems or norms, elevating it to organizational combat effectiveness.

(3) Huawei Knowledge Harvesting. In the domestic knowledge management field, Huawei's knowledge harvesting practices have gradually attracted attention in recent years. However, few articles introduce Huawei's knowledge harvesting, with occasional introductions heard only at public sharing events by Huawei knowledge management professionals. Specifically, Huawei's knowledge harvesting primarily targets internal projects, including two components: experience harvesting (Retrospect meetings) and document harvesting (valuable document organization). The former emphasizes tacit knowledge, while the latter emphasizes explicit knowledge, with both forming important parts of Huawei's knowledge harvesting.

Huawei's knowledge harvesting follows fixed routines and processes, mainly consisting of four steps: Project selection: First, identify organizational capability gaps and select prioritized knowledge harvesting projects based on these gaps, meaning prioritizing project knowledge harvesting from a strategic height; Single project knowledge harvesting: After determining specific projects, conduct preparation by clarifying objectives and scope, identifying knowledge harvesting organizational personnel, participants, and other resources. Then, through Retrospect meetings, document harvesting meetings, and other methods, systematically harvest tacit and explicit knowledge from projects, organize the final harvested content, and conduct review and storage; Organizational asset refresh: After determining the refresh scope of organizational knowledge assets, assemble teams to batch upload and refresh project-harvested experiences and documents into Huawei's knowledge repository, formally transforming them into organizational knowledge assets; Knowledge transfer: Finally, identify potential target scopes and audiences for newly uploaded knowledge assets, and proactively organize knowledge transfer activities by knowledge management personnel to ensure these new knowledge assets reach appropriate audiences.

2.2 The PREFS® Process Method for Knowledge Extraction

Based on the above models and practical experience, the author proposes the PREFS® method and related models for knowledge extraction. The PREFS® method emphasizes the knowledge extraction process (see [Figure 4: see original paper]). The entire process includes five steps:

(1) Plan. Knowledge extraction should not be conducted blindly but requires purposeful planning and topic selection based on organizational strategy, business, and project requirements. Resources and personnel within the organization must then be matched accordingly. For knowledge extraction within enterprises, prioritization can follow the 5-15-80 principle: strategic-level knowledge extraction may account for 5%, tactical key-level for 15%, and other general-level for 80%. For specific selected knowledge extraction projects, after determining tar-

get scope, preparation can begin. This stage also requires collecting as detailed materials as possible about the extraction target, including but not limited to background materials, situation introductions, process documents, and project summaries.

(2) Retrospect. This stage can employ various methods such as narration, interviews, reviews, field research, and collective discussions to reconstruct the original scenario as thoroughly as possible and discover underlying reasons and patterns. This stage is crucial and ideally should be led by a senior person with extensive knowledge extraction experience who can adopt a third-party objective perspective to conduct Fupan of the parties and situations involved, while deeply mining experiences and lessons. During review, both individual interviews with different parties and collective discussions and dialogues are needed to stimulate ideas not previously articulated by participants. This phase can draw on Lenovo's Fupan, the U.S. Army's AAR, Huawei's knowledge harvesting, and other methods. Knowledge extraction leaders require sophisticated facilitation skills and consulting capabilities to elicit resonance and associations from participants through deep dialogue, thereby extracting more "dry goods" (substantive content).

(3) Extract. This stage requires digesting the basic materials collected earlier, analyzing first-hand materials obtained during the retrospect phase, and then extracting useful content through refinement and processing. This stage tests writing ability and abstract thinking most severely, requiring rapid formation of the framework and thinking for the final knowledge deliverable product, including drafting documents and video outlines and scripts. This phase can reference the McKinsey Pyramid Principle to hone thinking, writing, and problem-solving logic, or draw on peer-summarized methods such as Zhang Lizhi's five-step method of "finding commonalities, identifying differences, digging into essence, elevating perspectives, and forming models."

(4) Form. This is the stage where knowledge extraction deliverables are finalized and produced. Through the above steps, the final knowledge product is formed and presented through text in reports, manuals, treasure books, cases, and other formats, or through multimedia interactive methods such as audio, video, and virtual reality. In the new internet economy era, as people's acceptance of multimedia continues to increase, knowledge extraction workers also need to keep pace with the times. How to use modern methods such as video and virtual reality to make learning enjoyable, enabling more people to willingly accept information and engage in deep learning and understanding, remains in the exploratory stage.

(5) Spiral. When the above knowledge products are delivered, knowledge extraction does not end. The first version may not be perfect and requires the craftsman spirit of continuous optimization of knowledge extraction deliverables through iterative improvement to achieve spiral ascent. Simultaneously, knowledge extraction practitioners need to adopt a marketer's mindset, proactively considering target audiences and application scopes for knowledge products,

designing, packaging, promoting, and marketing them to maximize their value.

2.3 The STAR® Content Model for Knowledge Extraction

While the PREFS® method for knowledge extraction emphasizes process, when refining and processing to form final knowledge deliverables, many people focus on whether good reference models exist. When determining the content framework structure for knowledge products, the author frequently uses the STAR® content model. Due to its strong versatility, it is described here in detail.

The STAR® content model organizes content according to the following framework structure:

(1) Situation. Description of historical environment and current context. Specific elements may include background introduction (e.g., social background, industry background, organizational background, story background, current situation), drivers (problems, pain points, status quo, history), etc. Heuristic questions include: What was the situation at the time (time, place, people, background)? What caused this situation? Who was involved? What were the surrounding circumstances?

(2) Target. Clear, focused objectives. Specific elements may include relevant vision, mission, goals, plans, tasks, etc. Heuristic questions include: What was the task at the time? What was the purpose or objective? What were the challenges and difficulties in achieving the task? What was the mood when facing the task?

(3) Approach. The path, method, and plan used to achieve objectives. This requires vivid case demonstrations and specific intuitive examples, including relevant practices (examples, stories, mechanisms, IT, incentives, organization, tools, forms, templates), team introductions, leadership messages, etc. Heuristic questions include: What was the reaction to the situation? What was actually done or said? What were the main action steps and details? Why were these actions taken? What experiences and lessons were learned?

(4) Results. The final outcomes and their significance and value. Specific elements may include relevant value, effects, feedback, data, gains, reflections, changes, etc. Heuristic questions include: What were the direct results of the actions taken? What impact was created? Were they effective and appropriate?

When writing content according to the STAR® framework, appropriate flexibility is allowed. For example, an introduction may be added at the beginning and a summary at the end. Keywords may also be added as tags for knowledge extraction deliverables, or to indicate their classification, applicable scope, and target audience. After Results, a "Future Vision" section may be added to describe prospects and ideas for future development. An "Expert Commentary" section may also be added, inviting senior experts to provide objective evaluations that inspire readers to think and engage in extended reading. The author once organized multiple companies to crowdsource the writing of the

book series *Knowledge + Practice Secrets* [**?**], in which each enterprise's knowledge management case adopted the STAR® content model, achieving certain practical application effects.

3. Application Case Analysis of Knowledge Extraction

Regarding practical application cases of knowledge extraction, the author focuses on the practice of knowledge extraction for benchmarking project cases from the Benchmarking Management Office of COFCO Group, an organization the author has served.

3.1 Project Background As an enterprise-level organizational transformation, benchmarking management is one of the most important management methods and tools for COFCO Group. COFCO Group began implementing benchmarking management in 2012, experiencing stages of initiation, deepening, and expansion, and has now entered a phase of comprehensive systematic advancement. COFCO Group's benchmarking management began with benchmarking evaluation and assessment. After completing benchmarking management training, launching benchmarking projects, determining annual budget targets based on benchmarking, building benchmarking databases, and establishing benchmarking workshops, the approach now emphasizes competitive benchmarking and implementation, integrating benchmarking work into daily operations. Through benchmarking management, organizations can establish systems that use a complete thinking method to drive organizations to re-examine and improve various tasks, pursue excellence, and make organizations wiser. In 2015, over 200 benchmarking optimization projects were launched, with each project cycle lasting approximately eight months. However, the process documents and key milestone deliverables of these projects were scattered among various project managers, and much of the experience was not well summarized. Once project teams were disbanded, experience became difficult to transform into organizational-level knowledge assets. However, there was no specialized experience in how to conduct knowledge extraction for these benchmarking optimization projects. Previously, the Benchmarking Management Office had produced benchmarking management project case collections, but the results were not effective and did not achieve effective dissemination.

To address these issues, the author and the Benchmarking Management Office's promotion leader jointly explored solutions. The leader once stated that "one characteristic of world-class enterprises is that 'best practices in any business area can be rapidly replicated within the organization.' " For the best practices generated by COFCO's various business units through benchmarking projects, how to conduct in-depth exploration through knowledge extraction and achieve effective dissemination and application within the organization was undoubtedly the key to solving the problem. Therefore, after discussion, several excellent benchmarking improvement projects were selected from completed projects for knowledge extraction to form benchmarking cases presented through text

manuscripts and micro-videos.

3.2 Knowledge Extraction Process Based on preliminary work, from December 2015 to March 2016, the author led a team deep into the front lines, using the PREFS® process method and STAR® content model for knowledge extraction to support the Benchmarking Management Office in extracting knowledge from relevant benchmarking projects.

Considering the strategic demonstration significance, project benefit indicators, and future replicability and expandability of benchmarking projects, and after screening with the Benchmarking Management Office, three projects were ultimately selected for knowledge extraction: COFCO Huangha's "Improving Soybean Meal Moisture and Protein Control Capability," COFCO Trading's "Optimizing Corn Business Division Accounts Receivable Cycle," and COFCO Trading's "Strategic Tribe Implementation to Advance Benchmarking Management."

In terms of team building, the author established a three-person professional knowledge extraction service team with a project leader, manuscript writing consultant, and video production personnel as the core. Simultaneously, the Benchmarking Management Office was required to coordinate project managers, project members, project sponsors, senior managers, and coaching advisors from specific benchmarking projects to participate jointly in the knowledge extraction joint project team.

The project team found that completing knowledge extraction for a benchmarking project required a minimum of two weeks. The knowledge extraction deliverables (i.e., final knowledge products) mainly consisted of two components: a text manuscript (averaging approximately 10,000 words) and a micro-video (averaging approximately six minutes).

In actual operation, the time spent often exceeded expectations. This characteristic was particularly evident when benchmarking projects involved many stakeholders. This is easily understandable, primarily because scheduling time with project stakeholders was often difficult to concentrate, inevitably extending the timeline, thus requiring mental preparation. Excluding these factors, the average person-days for the professional knowledge extraction service team alone were as follows: Field research and interviews: 3.0 people × 3 working days; Text manuscript organization and writing: 1.5 people × 7 working days; Micro-video filming and production: 1.5 people × 10 working days.

After completing knowledge extraction, the final knowledge products were promoted and disseminated in various settings (e.g., management annual meetings, professional seminars, senior management briefings) and received excellent feedback. For example, COFCO Huangha's "Improving Soybean Meal Moisture and Protein Control Capability" project challenged its own limits even when existing production control levels were already high. Through systematic thinking and design, six quick-win improvements and seven improvement proposals were

implemented to achieve precise control of soybean meal moisture and protein ratios according to actual operating conditions, thereby exceeding original targets and generating annual economic benefits of 4.9 million yuan. Moreover, this case's experience applies to other agricultural product processing involving precise control of protein and moisture content, such as soybean meal, feed, rice, DDGS, and barley malt, thus having tremendous replication effects for similar factories.

Through case mining of benchmarking management projects, the author discovered that knowledge extraction work is highly necessary, demonstrated by its effectiveness, systematicity, and replicability. Knowledge extraction effectively identifies key project factors, analyzes important project nodes, and systematically consolidates experience and knowledge to form knowledge products, laying a foundation for broader application in homogeneous and similar projects.

3.3 Future Outlook for Projects After completing the pilot knowledge extraction project for benchmarking management projects, participants gained new understanding and insights compared to previous case summary methods, both in terms of case depth and breadth.

In terms of case depth, using video formats to synthesize multi-perspective viewpoints makes case presentation more vivid and diverse. Simultaneously, employing the PREFS® process method and STAR® content model makes experience refinement results more professional and comprehensive.

In terms of case breadth, understanding of its value should not be limited to extracting best practices but should emphasize disseminating best practices, matching most suitable scenarios, and contributing to business success. Moreover, the focus should not be solely on benchmarking optimization projects themselves but should increasingly emphasize connecting people, connecting business, and connecting experts.

Furthermore, regarding the medium- and long-term vision for benchmarking management, in addition to continuing knowledge extraction for benchmarking projects, cases can serve as the main lever and breakthrough point. Combined with key business scenarios such as projects and courses, a comprehensive knowledge management platform can be constructed (see [Figure 5: see original paper]). As a foundational project, knowledge management will become the solid base beneath the iceberg, gradually transforming individual knowledge into organizational capabilities and advancing from organizational capabilities to organizational wisdom.

4. Conclusion

Through practical application of the PREFS® process method, STAR® content model, and other approaches, the author has realized that knowledge management can thereby gradually penetrate the business level, extracting valuable

knowledge "dry goods," representing direct contribution and value to the business. New eras require new approaches, and using micro-videos and other methods to present knowledge extraction results will become an inevitable choice for more and more organizations. Simultaneously, the professional value of personnel's knowledge extraction capabilities will become prominent. How to cultivate professional talent and service capabilities in knowledge extraction will become one of the future hot topics. Good content requires strengthened operation—even good wine fears a dark alley—and greater effort must be spent on promoting and applying high-quality knowledge. From this perspective, knowledge extraction urgently needs breakthroughs in both theory and practice. If we can create a "one-stop" service encompassing knowledge extraction-related writings, training, consulting, and implementation, a new blue ocean of professional services will span both education and training and knowledge management, coming directly toward us.

References

[?] Yang Zhi. Research on Generalized Tacit Knowledge Extraction and Learning [?]. Hunan Social Sciences, 2015(1): 210-213.
[?] Zhou Yan. Application of Knowledge Extraction in Skill Training [?]. Technology and Enterprise, 2014(24): 27.
[?] Zhuang Qiming. Research on Knowledge Extraction and Construction in Knowledge Management [?]. New Taipei: Tamkang University, 1995.
[?] Jiang Yongchang. Knowledge Refinement and Application Based on Web2_0 Comprehensive Integration in Knowledge Construction [?]. Library and Information Service, 2014, 58(21): 116-123.
[?] Qiu Zhaoliang. *Fupan+: Transforming Experience into Capability* [?]. Beijing: Mechanical Industry Press, 2015.
[?] Baidu Baike. Extraction [?]. [2016-04-10]. http://baike.baidu.com/.
[?] Liu Chuanzhi. Fupan is the Best Learning Method [?]. Human Resources Development of China, 2013(24): 11.
[?] Xu Jigang. Two Steps for Successful Knowledge Management Implementation [?]. China Information World (E-Manufacturing), 2014(7): 55-57.
[?] Wu Qinghai, Wang Meng, Xia Jinghua. *Knowledge + Practice Secrets* [?]. Beijing: World Knowledge Press, 2015.

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