Kashihara, A. et al. (Eds.) (2024). Proceedings of the 32nd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Facilitating Thinking like a Historian in Open-Ended Learning Space: A White Box Approach

Aoi MATSUURA*, Yuki HAYASHI & Kazuhisa SETA
Graduate School of Informatics, Osaka Metropolitan University, Japan
*matsuura@ksm.kis.osakafu-u.ac.jp

Abstract: In this study, we adopt the Thinking Like a Historian Approach (TLH), which systematizes the historical inquiry process of historians as a learning model. By expanding Linked Open Data (LOD) using a large language model (LLM) based on an ontology that defines TLH concepts, we construct a white-box-oriented learning resource that ensures comprehensiveness, reproducibility, learning controllability, and maintainability. The historical learning support system developed based on this learning resource has adaptive question generation function and historical information visualization function in an open-ended historical inquiry space.

Keywords: Open-ended learning, thinking like a historian approach, inquiry-based learning, linked open data

1. Introduction

In the field of learning sciences, it is known that providing meaningful questions contribute to deepening learner's understanding. Since the quality of learning depends on the quality of the questions (National Research Council et al., 1999), it is important in the context of self-directed learning to generate meaningful questions themselves and explore with purpose. In the field of history education as well, it is important to learn through inquiry by posing and answering questions. To support such learning, Mandell & Bobbie (2008) proposed the Thinking Like a Historian Approach (TLH), which systematizes how historians explore historical topics. Instead of merely memorizing facts such as the dates, places, and reasons for historical events (e.g., "*The cause of the French Revolution was the citizens' dissatisfaction with the monarchy*"), learners are encouraged to ask questions such as "*Why were the citizens dissatisfied?*" or "*How did the citizens oppose the monarchy?*" and explore these considerations. On the other hand, it has been pointed out that learners do not always generate meaningful questions to deepen their learning while learning new topics (Otero, 2009). Furthermore, the demonstration of independence is essential for such inquiry-based learning (Mandell & Bobbie, 2008). To address these challenges, it is known that computer-based history learning materials that combine content such as statistical data (chronology, maps, etc.) can motivate learners (Greene et al., 2010).

The web-based learning environment, which allows for the exploration of historical topics from various regions and eras around the world, is suitable for exploring knowledge and learning history in line with individually set learning goals (Hill & Hannafin, 1997). Considering the aforementioned background, it is important to address the following challenges in web-based learning: (1) How to promote learners' proactive learning? (2) How to realize adaptive question generation support that leads to learners' meaningful learning? (3) How to present meaningful information that promotes history learning utilizing diverse learning content?

Regarding challenges (1) and (2), we have previously supported web-based learning by developing a system that adaptively generates questions based on ontology and Linked Open Data (LOD) (Jouault et al., 2016), and by realizing a mechanism that integrates semantic

and spatial information of historical topics (Matsuura et al., 2022). As a result, it was suggested that these approaches contribute to deepening learners' understanding and promote them to learn proactively while immersing themselves in the world of history.

In this paper, we propose a historical learning support system equipped with information visualization and question generation functions for historical topics, utilizing ontology and LOD based on the TLH theory. The contribution of this study is to propose new technological mechanisms to address the following research questions: "How can we realize a historical inquiry-based learning environment that dynamically presents multiple data and visual expressions related to various regions and eras around the world in an open-ended space?" and "How can we present questions to learners that capture their interests and lead to meaningful learning?"

2. Method for Expanding LOD to Accommodate Diverse Learner Interests

2.1 Approach for Data Expansion to Support History Learning

To support cross-sectional and longitudinal understanding of historical topics and personal relationships from various regions and periods around the world, which differ in interest for each learner, the system needs to comprehensively manage the knowledge required for such support. To ensure the comprehensiveness of knowledge as a learning resource, this study adopts Linked Open Data (LOD), a machine-readable web resource. LOD where every entity has a unique URI and takes the form of triples (i.e., <subject, predicate, object>), Especially Wikidata has currently over 100 million triples existing. For example, it provides machine-readable information on *when* and *where* the *French Revolution* occurred and that *Louis Antoine* was involved. Wikidata is highly reliable and comprehensive as a general resource because it is collaboratively edited by many people. On the other hand, it does not necessarily include historical-specific relationships crucial for learning history as indicated by the TLH theory (Section 3.1), such as the interrelationships between historical persons or the causal relationships of historical events.

Therefore, this study takes an approach of expanding Wikidata using a Large Language Model (LLM) to increase the comprehensiveness of historical learning resources. Here, since the processing of LLM is a black box, it is necessary to consider issues such as hallucinations and reproducibility. This study tackles these issues by explicitly mapping the knowledge (triples) generated by the LLM to the concepts of TLH. By enabling the system to identify which historical concepts correspond with the dynamically generated questions (Section 3.2), the information referenced by the system and the knowledge utilized become clear (improving reproducibility). This allows learning designers to define their intended controls (ensuring learning controllability). Furthermore, even if incorrect knowledge is generated, the explicit reference to the information allows for corrections (maintainability of knowledge). The following section describes the expansion method of LOD aimed at white boxing the internal structure of learning resources.

2.2 Expansion Method of LOD

As shown in Figure 1, LOD represents knowledge in the form of triples (<subject, predicate, object>). When nodes included in Wikidata and nodes within the triples generated by the LLM refer to the same historical topic, their IDs need to be aligned. To avoid the problem that multiple concepts and entities having the same label, we generate triple sets using LLM, limited to the learner's target historical topics. We then align IDs based on label matching. The types of information and their output formats used to expand LOD are predefined based on the TLH's concepts.

To explain using the 'French Revolution' (Figure 1) as an example: (1) when a learner selects the 'French Revolution' as a learning topic, the system identifies it as a type of 'Revolution' based on the `instance-of` relationship(Figure 1(A)) represented in

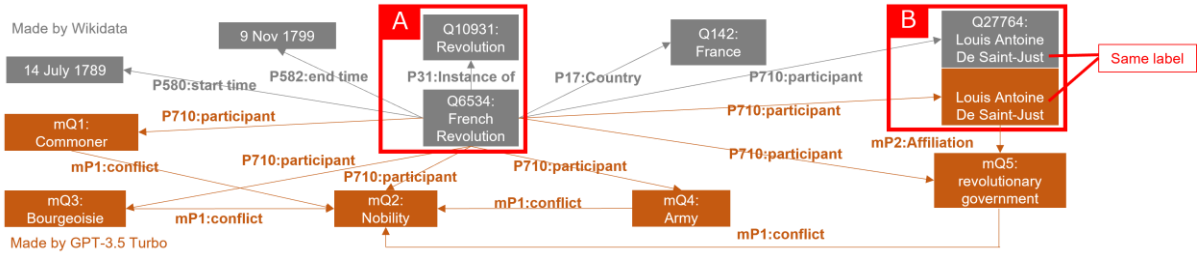

Figure 1. Linked Data Expanded with Wikidata by LLM.

Wikidata. (2) When deepening the consideration of "revolution", TLH theory emphasizes that "when historians research a revolution in depth, they consider deeply into the participants (P710: participant) of the revolution and the relationships between participants." By specifying this kind of historical concepts and relationships between concepts in the ontology, the system can prompt the LLM (GPT-3.5 Turbo) based on this definition to obtain the relevant information. Finally, (3) by matching labels with nodes on Wikidata, the system integrates and expands information useful for historical learning (Figure 1(B)).

One of the notable features of the proposed method is that the system can distinguish between highly reliable human-generated information in Wikidata and the information expanded by LLM, which may contain errors. This allows the system to facilitate interactions based on the distinction and achieve a white-box approach that makes knowledge maintenance easier.

3. Adaptive Learning Support Method for Meaningful Learning

3.1 Methodology for Promoting Meaningful History Learning

The Thinking Like a Historian Approach (TLH), which systematizes how historians explore historical topics as meaningful exploratory learning, has been proposed (Mandell & Bobbie, 2008). This approach emphasizes the importance of learning that involves critically exploring historical topics and forming one's own interpretations, rather than merely memorizing the facts written in textbooks. It is a systematic theory that can transform history learning, which tends to become boring memorization of facts, into creative learning activities.

In TLH, three processes are defined: "Question," "Evidence," and "Interpretation." During the execution of each process, the five perspectives of "Cause and Effect," "Change and Continuity," "Turning Points," "Through Their Eyes," and "Using the Past" are considered ideal for exploratory learning. These three processes detail meaningful learning activities in a systematic manner. For instance, in the "Question" process, about five learning activities are specified for each of the above five perspectives (a total of 22 activities). For example, in "Cause and Effect," the activity involves "Formulating questions that promote consideration from the perspective of who benefited. ; in "Evidence", it involves "Gathering information from multiple perspectives"; and in "Interpretation", it involves "Explaining the connections between historical topics." These activities are systematized along with the questions that initiate such learning activities.

This study proposes a method to guide learners towards meaningful exploratory learning by promoting exploratory activities aligned with the TLH theory.

3.2 Adaptive Question Generation Method

As an adaptive support in an open-ended learning space, Jouault et al. (2016) have constructed a knowledge model that integrates two types of LOD and defined a history domain ontology that enable semantic processing of the integrated knowledge model. Furthermore, by defining a history dependent question ontology associated with these concepts, they proposed a method for generating questions based on the constructed massive knowledge base. This study adopts this mechanism to generate questions based on TLH concept.

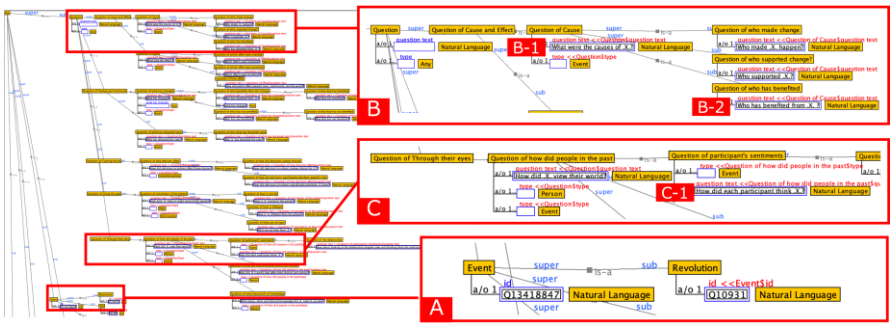

Figure 2. A part of Ontology That Systematized TLH Concepts.

Figure 2 shows the definition of question concepts defined in the developed TLH ontology. Here, the 22 questions mentioned in Section 3.1 are defined in correspondence with the learning activities for each of the processes: Question, Evidence, and Interpretation. For learners who find it difficult to discover questions that deepen their understanding of history, the aim is to promote meaningful inquiry-based learning based on the TLH concept by presenting questions relevant to the learning content.

As a more specific mechanism, the content of the learning materials that the learner engages with is represented as LOD. Therefore, by referring to the TLH ontology, the system can grasp the learner's interests and concerns. The ontology defines what kinds of questions are effective to present in order to deepen understanding when learning about certain historical topics. For example, when a learner is studying the 'French Revolution,' based on the concept definition that "a *Revolution* is a type of *Event*(Figure 2(A)), and understanding an *Event* is deepened by comprehending its *causes*(Figure 2(B-1)) and the *sentiments of the participants*(Figure 2(C-1))," the system can present questions like "*What were the causes of the French Revolution?*(Figure 2(B-1))" and "*Who has benefited from French Revolution?*(Figure 2(B-2))"

4. Historical Learning Support System

Figure 3 shows the interface of our developed history learning support system. When the system is activated, the screen for setting the learning goal and selecting historical topics is displayed. The learner first inputs the learning goal, and then sets the historical topic they want to explore. Historical topics can be narrowed down by name search or by region, period, and category, and the learner can select one from the search results (a list of historical topics) that match the conditions.

By selecting a historical topic, the historical topic exploration screen is displayed (Figure 3 shows an example of selecting the 'French Revolution'). The historical topic exploration screen consists of (A) question list confirmation area, (B) historical topic review area, and (C) interpretation area. The design allows learners to freely navigate among these three areas as they progress with their exploratory learning. The details of each area are described below.

(A) Question list confirmation area (Figure 3(A)): The area corresponding to "Question" process in TLH theory, where learners can check questions presented by the system (e.g., "*What were the causes of the French Revolution?*") related to the historical topics they have selected. For learners who cannot produce meaningful questions themselves, the system provides questions aligned with the TLH theory to ensure the quality of history learning. In addition, sub-questions that further explore each question (Figure 3(A-1)) are presented. For example, related to the question "*What were the causes of the French Revolution?*" a sub-question such as "*Who made change happen?*" is provided. This aims to clarify the direction for learners when they are stuck in their exploration activities and to enable them to investigate from multiple perspectives. Learners can answer each question in the answer area (Figure3 (A-2)), and review them at any time.

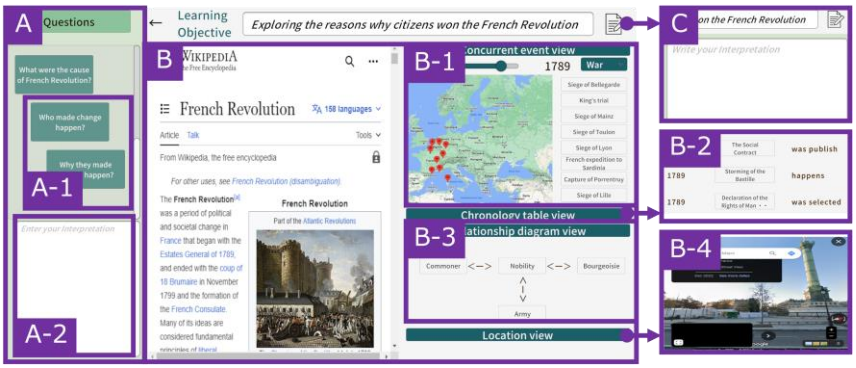

Figure 3. Interface of History Learning Support System.

(B) Historical topic review area (Figure 3(B)): The area corresponds to the "Evidence" process in TLH theory, where learners can review the Wikipedia page of the historical topic and information expressed on various views related to the historical topic. The system implements five different visualization methods based on literature review in history learning: concurrent event view(Figure3 (B-1)), chronological table view(Figure3 (B-2)) (Lee et al., 2013), family tree view (Terao, 2004), relationship diagram view(Figure3 (B-3)), and location view(Figure3 (B-4)) (Lowenthal, 1975). Relevant connections are extracted from the Linked Data discussed in Section 2 and presented according to the visualization method to provide adaptive learning materials corresponding to historical topics from around the world.

(C) Interpretation area (Figure 3(C)): The area corresponds to the "Interpretation" process in TLH theory, where learners can input their own interpretations of the learning goal. The interpretations input by learners are saved in the system, allowing them to review and modify their previous interpretations while reflecting on their exploration activities.

5. Initial Evaluation

5.1 Experimental Setting

An initial experiment was conducted with six undergraduate and graduate students to verify whether the proposed system could lead them to meaningful historical exploration activities.

First, the participants were asked to answer the question (learning goal), "Q1: What impact did the Khmer Empire have on its surroundings?" using only information from Wikipedia via a web browser until they were fully satisfied with their answers. During this activity, they were also asked to answer the question, "Q2: What do you think is important when learning history?" (learning phase (a)). Next, using the proposed system, the participants were asked to revise and expand on their answers to Q1 and Q2 if necessary (learning phase (b)). No time limits were set for either phase (a) or (b). Afterward, they evaluated their own Q1 answers from phases (a) and (b) using the rubric set by TLH (Mandell & Bobbie, 2008, p. 117). Finally, they were asked to answer a questionnaire about their experience using the system. The questionnaire consisted of two items (rated on a five-point scale): "Q3: Was each view information meaningful for learning history?" and "Q4: Was the questions presented by the system meaningful for learning history?"

5.2 Results

The average rubric scores for the Q1 answers in learning phases (a) and (b) were as follows (n=6): 'Level of Analysis': (a) 0.67, (b) 2.00, 'Use of Evidence': (a) 1.17, (b) 1.67, 'Explanation from Perspectives': (a) 1.50, (b) 2.34, 'Explanation of Significance': (a) 0.50, (b) 1.17, 'Explanation of Connections': (a) 0.84, (b) 1.17. For all items, the answers created using the system scored higher. This suggests that participants felt they were experiencing the learning required for practicing TLH theory through history learning using the proposed system. The t-test results showed a significant difference at the 5% level for 'Level of Analysis (p=.025),'

‘Explanation from Perspectives (p=.042),’ and ‘Explanation of Significance (p=.025)’. From the system log, it was confirmed that questions prompting consideration of ‘Use of Evidence’ and ‘Explanation of Connections,’ were presented to all participants when using the system. However, one reason for the relatively lower self-assessment may be the insufficient learning resources available for answering these questions. In addition, answers to Q2 showed a shift from abstract statements like “*Considering the background of historical topics*” in phase (a) to more concrete statements, such as “*It is important to properly understand the cause-and-effect relationships because there must be reasons behind events*” and “*It is important to take a multifaceted view, including the relationships with other countries*”, in phase (b). This indicates that learning with the proposed system enabled the participants to realize the importance of the learning activities required by TLH.

The average scores for the Q3 and Q4 survey results were 3.83 and 4.17, respectively. These results suggest that both information visualization and question generation functions were positively received by the participants and utilized in history learning.

The average time required for the learning tasks in phases (a) and (b) was 62.5 minutes and 39.1 minutes, respectively. Additionally, the average number of characters in the Q1 answers (in Japanese) was 436.5 (phase (a)) and 624.7 (phase (b)). These results suggest that the adaptive questions capturing participants’ interests and the information presented in various views served as stimuli, making participants realize there was room for further consideration in their answers from phase (a), thus promoting further exploratory learning.

6. Conclusion

In this study, we proposed a method for expanding LOD using a white-box approach to address various learning topics in history learning, and developed a history learning support system based on the TLH theory as a learning model. Initial experimental result suggested that the system promotes learners’ independent exploratory learning and the usefulness of support functions. As a future task, we plan to confirm the usefulness of the system through practical use by a larger number of subjects.

References

Greene, J. A., Bolick, C. M. & Robertson, J. (2010). Fostering historical knowledge and thinking skills using hypermedia learning environments: The role of self-regulated learning. *Computers & Education*, 54(1), 230-243.

Hill, J. R. & Hannafin, M. J. (1997). Cognitive strategies and learning from the World Wide Web. *Educational Technology Research and Development*, 45(4), 37-64.

Jouault, C., Seta, K. & Hayashi, Y. (2016). Content-Dependent Question Generation Using LOD for History Learning in Open Learning Space. *New Generation Computing*,34(4), 367-393.

Mandell, N. & Malone, B. (2008). *Thinking like a historian: Rethinking history instruction*. Wisconsin Historical Society.

Matsuura, A., Seta, K. & Hayashi, Y. (2022). Enabling Physical- and Concept-Walk in VR-based Open-ended Historical Leaning Space. *Proc. of 15th Workshop on Technology Enhanced Learning by Posing/Solving Problems/Questions in conjunction with ICCE2022*, 11-15.

National Research Council, et al. (2000). *How people learn: Brain, mind, experience, and school: Expanded edition*. Vol. 1. National Academies Press.

Otero, J. (2009). Question generation and anomaly detection in texts. *Handbook of Metacognition in Education*, 47-59.

Lee, P., Dickinson, A. & Ashby, R. (2013). Children’s ideas about historical explanation. *International review of history education* (pp. 97-115). Routledge.

Terao, T. (2004). Principles of Learning about Historical Figures based on the Theory of Cognitive Constructivism: An Analysis of the Amherst Project in the American High School History Curriculum. *Journal of Japanese Educational Research Association for the Social Studies*, 61, 1-10. (in Japanese)

Lowenthal, D. (1975). Past time, present place: landscape and memory. *Geographical review*,65(1), 1-36.