Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 169 (2020) 96–99

Procedia Computer Science

www.elsevier.com/locate/procedia

Postproceedings of the 10th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2019 (Tenth Annual Meeting of the BICA Society)

A Hybrid System for building a Personal Knowledge Base

Roman E. Prilipsky[a], Margarita A. Zaeva[a,*]

[a]*National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoe shosse, 31, Moscow, 115409, Russian Federation*

Abstract

With the rise of the information age and the ever-increasing amount of available digital information the practice of personal knowledge management is becoming more and more important for everyday life. There are various methods of personal knowledge management. However, existing practical implementations of those methods tend to be somewhat narrow in scope, performing specific sets of tasks. This paper considers an approach that combines multiple ways of managing personal knowledge into a hybrid system in an attempt to compensate for shortcomings of its components.

Keywords: information management; digital curation; software; learning; organization

1. Introduction

OAvailability of information is no longer a problem in the modern world indeed, the opposite problem, information overload, has been noted to be an obstacle in the process of work or studying[1].

Vast amounts of digital information sources of different types, such as books, articles, recordings, web pages and more contain useful data that can be extracted, structured and contextualized in order to apply it more efficiently[2].

However, extracting and storing information from unorganized sources can be a task in and of itself - a task that increases in difficulty the bigger the list of sources becomes.

This poses the problem of efficient information acquisition and extraction, as well as organization and retrieval. These processes can be broadly grouped into a single discipline of personal knowledge management (PKM). This topic is somewhat underrepresented both in scientifical publications and in practical applications, with much greater attention given to the practice of business knowledge management - a related, but separate field, with different methods applicable.

* Margarita A. Zaeva. Tel.: +7-903-719-9188.
E-mail address: MAZayeva@mephi.ru

10.1016/j.procs.2020.02.120

Out of all potential ways of personal knowledge management, software-assisted methods present the most potential. Various software tools can be used to streamline the process of learning, offloading many parts of it onto external systems. However, existing software solutions for personal information or knowledge management tend to focus on singular applications, such as note-taking, file tagging or concept mapping, which makes them well-suited only for specific tasks related to knowledge management.

This article proposes a combined approach to knowledge acquisition, curation, organization and retrieval that applies several techniques, such as concept mapping, tagging, flash cards and hyperlinking in order to create a flexible and extendable system for PKM.

2. Personal Knowledge Base

The term ”Personal Knowledge Base” is defined by S. Davies as ”an electronic tool through which an individual can express, capture, and later retrieve the personal knowledge he or she has acquired”[3].

Such a tool in its original definition is mostly concerned with managing the user’s personal knowledge - that is, instead of the sources of information it would contain fragments of knowledge the user derived from those sources. However, there is also obvious value in tracking sources of assimilated knowledge within the knowledge base. A tool with such capabilities would also act as a personal digital archive, extending PKM with personal information management (PIM). The functionality of this tool could potentially allow it to be integrated into all phases of personal information lifecycle - aquisition, organization and retrieval [4].

A personal knowledge base contains knowledge elements - basic components of the base that the user creates and interacts with [3]. These elements can represent text notes, diagram nodes, fragments of external documents or links to them, and more. Integrating external documents into a database can be performed and managed in a number of ways [5].

Versatility of a specific personal knowledge base is expressed through formats of knowledge elements it supports, ways these elements can be linked to each other, and possible representations of contained data. A system that can process and store multiple element types, by design, has more flexibility, which is offset by complexity of realization. That complexity is one of major reasons for the current lack of modern full-fledged PKM systems on the market.

With that being said, the complexity of system realization can be reduced by unifying the structure of different knowledge elements, allowing the system to organize, display and transform those elements in a standartized way.

To discern possible approaches to system simplification, let us first list common knowledge management methods that can be found in available PKM-related software tools.

3. Knowledge management methods

Knowledge management methods applicable to PKM can be split across phases of personal knowledge lifecycle that match phases of the previously mentioned personal information lifecycle. The absolute majority of existing PKM solutions focus on a single method from the following list.

3.1. Aquisition

- **Note-taking.** The practice of creating and organizing textual documents used for knowledge capture. This method fits into both aquisition and organization phases. This is one of the most common methods of personal knowledge management.
- **Incremental reading.** The process of reading documents in portions and extracting key points into separate elements of the base. This method of acquisition is typically used to create sets of flashcards for use with the spaced repetition method.

3.2. Organization

- **Tree.** A structural framework that stores elements in a hierarchy, in which every element has one parent. A classic organizational concept applied in computer filesystems, easily comprehensible and navigable.

- **Tagging.** Provides a mechanism of user-defined metadata assigned to elements of the base. Useful for search and freeform categorizing, as, unlike with the tree model, there is no limit on the number of tags that can be assigned to a single element.
- **Mind and concept maps.** Types of diagrams used to visualize, structure, and classify information and knowledge. The graphical representation of diagrams' contents lend themselves well to spatial organization purposes.
- **Hypertext.** A text format with references to other elements of the base. The element referred to is not nessesarily a text document. A commonly seen subtype of hypertext knowledge organization systems are Wiki websites.
- **Flashcards.** Storage format used by spaced repetition method, flashcards provide information in a format that facilitates the process of active recall. They can be considered a specialized type of a note. This storage format is frequently used in language studies for word memorization.

3.3. Retreival

- **Search.** Aside from title and full-text search that can be performed on elements of the base, the search method can leverage the tagging mechanism or map structures to perform complex base queries.
- **Graphical representation.** Stored diagrams and the structure of the base itself can be displayed in a visual and easily comprehensible form. Manipulating this representation to alter connections between elements of the base may also be easier for the user.
- **Spaced repetition.** This technique is utilized for knowledge memorization over a prolonged period of time. It works best with sets of flashcards, with each card containing a separate small knowledge fragment.

4. Combinations

The aforementioned methods of knowledge management, if applied in a software solution, can be combined in multiple ways. These combinations can allow for realization of several knowledge manipulation techniques that cannot be replicated in single-method tools.

Tags applied to incremental reading materials can be used to organize extracted knowledge elements in the base. Tags can be applied to files, text notes and flashcards - in fact, tags can be applied to any element of the base no matter the format. Inheriting all base elements from an abstract taggable element type simplifies the representation and storage.

Internal data of the base can be represented in all primary PKB structural frameworks - trees, graphs, spatial and categorical [6]. Metadata of the tagging system can be interpreted as categories, clusters in a graph or sections of a spatial layout. Graph representations can be displayed as mind maps. Hypertext links also represent a variation of a graph framework and can be displayed similiarly.

Displaying elements of the base in the graph framework allows for intuitive manupulation of connections between displayed elements.

Tags themselves can be organized into a hierarchical system under the tree framework. This allows the tagging system to express tree-like hierarchical connections in addition to freeform categorical organization it provides by default.

Reinterpretation of the contents of the base in different formats increases its versatility, as all structural frameworks have various advantages and disadvantages. Users may work with the base through the interface they find most comfortable.

The process of knowledge extraction can be extended to all types of elements stored in the base. Elements extracted from source material during incremental reading don't have to be specifically flashcards - text notes and diagrams can also be stored in the base.

Conclusions

The described variant of a personal knowledge base extends several personal knowledge management methods previously separately applied in practice. The intended implementation of the variant is a single software application

that integrates all of the described functionality. Indeed, the key point of this PKB is integration, as combinations of listed methods allow for unique interactions to emerge.

This knowledge base can serve as a potential replacement to multiple types of currently existing software tools, such as note-taking applications, mind map editors and tag-based file managers.

Interestingly, the resulting structure somewhat resembles T. Nelson's ZigZag structure [7], described as potentially more general than any of the PKB structural frameworks [6].

It is in our opinion that such a system represents a new kind of solution for the majority of PKM-related tasks that provides several advantages previously not seen in any of the available tools.

Acknowledgements

This work was supported by Competitiveness Growth Program of the Federal Autonomous Educational Institution of Higher Professional Education National Research Nuclear University MEPhI (Moscow Engineering Physics Institute).

References

[1] Rachfall, Thomas & Williamson, Elizabeth & Temple, Bryan. (2014) The information overload phenomenon: the influence of bad - and (ir)relevant information. *International Journal of Research in Engineering and Technology. 10.15623/ijret.2014.0329005.* **88**: 27–32.

[2] Beagrie, N. (2005) Plenty of room at the bottom? Personal digital libraries and collections. *D-Lib Magazine.* **11** (6) http://www.dlib.org/dlib/june05/beagrie/06beagrie.html

[3] Davies, S. (2011) Still building the memex, *Communications of the ACM.* **54** (2): 80–88 doi:10.1145/1897816_1897840

[4] Williams, Peter, Leighton John, Jeremy, Rowland, Ian. (2009) The personal curation of digital objects: A lifecycle approach. *Aslib Proceedings. 10.1108/00012530910973767.* **61** : 340–363.

[5] Buneman, P., Chapman, A., and Cheney, J. (2006) Provenance management in curated databases. *In Proceedings of the 2006 ACM SIGMOD International Conference on Management of Data.* ACM, NY, 539–550.

[6] Davies, Stephen; Velez-Morales, Javier; and King, Roger. (2005) Building the Memex Sixty Years Later: Trends and Directions in Personal Knowledge Bases. *CU-CS-997-05"Computer Science Technical Reports.* **931** https://scholar.colorado.edu/csci_techreports/931

[7] Nelson, T. (2006). A Cosmology for a Different Computer Universe: Data Model, Mechanisms, Virtual Machine and Visualization Infrastructure. *Journal of Digital Information.* **5** (1). https://journals.tdl.org/jodi/index.php/jodi/article/view/131/129