

Personalized Reading Resources Organization

Roman BURGER*

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 3, 842 16 Bratislava, Slovakia
roman.arnold.burger@gmail.com*

With vast amount of accessible information and resources through the web, one may start to carefully choose what to read and what to ignore. Anyhow we try, chances are that our workspace will eventually get overwhelmed with quantum of resources stored for later reading. The very common cause is that resources processing speed of the user is lower than resources retrieval speed.

Keeping track of resources organization and structure then becomes tedious task, possibly distracting us from reading of individual resources. Users need to manually manage their workspaces which produces significant overhead while trying to work with stored resources. It is not uncommon for users to actually abandon the management due to convenience. Especially for user's convenience we seek for means of automatic organization management of resources.

Related works [1, 2] focus mostly on content analysis of documents and thus producing merely semantic relations between said documents. While semantic relations are very important, they usually cannot fully capture project relations. By project relations we mean relations defining collections of all resources used for particular projects. In this representation project could be anything from semestral work to as simple as collection of blogs about interesting topics.

In this project, we will propose a method for automatic organization for emerging resources. To actually capture project relations we will take into account two sources of information. The first one will be documents metadata, the second will be documents context obtained via implicit user feedback. Metadata is chosen because of their convenience and precision. Part of metadata could be actually keywords extracted from content, but we will focus on less traditional metadata such as creation time or last edit time. These metadata have higher chance discovering project relations, because they are describing usage patterns of the user. Context of documents is used to aid determining project relations of resources, because resources in the same project should have the same or very similar context.

Our proposed method (Figure 1) first extracts metadata from added resources. After metadata extraction, clustering algorithm will be deployed to generate initial

* Supervisor: Mária Bielíková, Institute of Informatics and Software Engineering

collections of resources. User will be given tools to personalize collections to their needs and by logging these actions we will get implicit feedback on precision of generated resources structure. The next time user adds resource to system, we can use context of existing collections to cluster new resources.

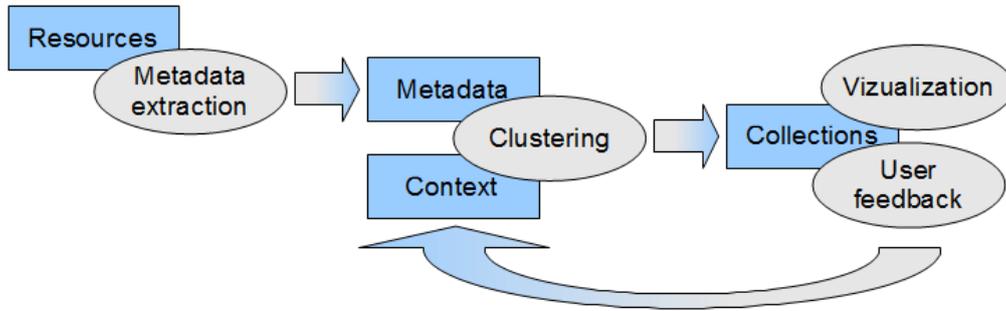


Figure 1. Proposed process of generating automated personalized collections.

On top of resources organization, recommendation of reading order could be employed. With automatic personalized organization and processing order recommendation, user would need little to none effort maintaining workspace. User would only need to pop out first item from recommended resources list while automatic organization and visualization should take care of giving the user the right collection.

Final solution will be evaluated in user experiment. We will try to evaluate overall usefulness and focus mostly on organization method and visualization. Validation of the organization method will be done through user experiment where we will measure amount of user correction in generated structure. Hypothesis says that user intervention in organization should decrease over time.

Acknowledgement. This work was partially supported by the Slovak Research and Development Agency under the contract No. APVV-0208-10.

References

- [1] Corrêa, R. F., Ludermir, T. B.: Semantic mapping and K-means applied to hybrid SOM-based document organization system construction. In: *Proc. of the 2008 ACM symposium on Applied computing*, ACM, (2008), pp. 1112-1116.
- [2] Prinz, W., Zaman, B.: Proactive support for the organization of shared workspaces using activity patterns and content analysis. In: *Proc. of the 2005 international ACM SIGGROUP conference on Supporting group work*, ACM, (2005), pp. 246-255.