Bar-Ilan University (RA)
Faculty of Humanities
Library of Information science

אוניברסיטת בר-אילן (ע"ר)
הפקולטה למדעי הרוח
הספריה ללימודי מידע

Tags vs. folders in personal information management/ Noa Gardovitch

Abstract

Traditionally, personal information management systems provided folders for information storage and retrieval. However, with the boosting of Web 2.0, tags had defused into PIM systems. Tag-supporters claim that it has two fundamental advantages over folders: It supports multiple-classification (adding several tags to an information item) and is not reliant upon hierarchical locations.

The aim of the present study was to test users' preference when using systems that support both folders and tags. In order to increase generalizability we conducted two studies in two different working environments. In the Gmail Study we had informed 75 participants of both folder-labeling (moving an email message into a single label) and tag-labeling (adding several relevant labels to a message), waited a month to observe their storage behavior (based on a print screen of their mailbox) and asked them to estimate their retrieval behavior (via a retrieval habit questionnaire).

In the Windows 7 Study we informed 23 participants about tags and asked them to tag all their folders for two weeks and then left them for 5 weeks. Their storage and retrieval habits were tested before the learning session and after 7 weeks, using special category recording software and a retrieval habit questionnaire. A controlled retrieval task and an in-depth interview were conducted in the Windows 7 Study. We also checked the gap between user's attitudes and beliefs towards tags and their real behavior by conducting a survey of 168 participants.

Results of Gmail and Windows 7 studies show a strong preference for folders over tags for both storage and retrieval. Moreover, even in the minority of cases where tags were used, participants typically used only a single tag per information item and even in the small minority of cases when multiple-classification were used when storing information items it had only marginal use for retrieval. Results also show preference for hierarchical over flat storage and for location-based over non location retrieval. The controlled retrieval task showed lower success rate and slower retrieval speed in cases were tags were used (compared to other kinds of retrieval). Results of the attitudes study suggest that though user's attitudes towards tags are positive their actual behavior show preferences for folders.

Possible reasons for participants preferences are the difference between the web 2.0 environment and PIM environment, the cognitive ease that suggests the hierarchical method over tags and the habit. We think that one of the key issues here is familiarity. When looking for content which other users had uploaded to the Web, users can not possibly know where it is located. Therefore searching by tags seems like a much better option. In PIM on the other hand, users are very familiar with their own information organization, after all they had stored it there according to their own subjective needs (Bergman, Beyth-Marom, & Nachmias, 2003, 2008). Moreover, they get more familiar with their organization scheme each time they navigate through it to retrieve their files. Therefore in the large majority of cases they are able to retrieve their own personal information quickly and efficiently (Bergman et al., 2010).

The hierarchical method answers the cognitive needs of the users. The hierarchical method is consistent. This consistency is a key value in human computer interaction (Shneiderman, 1997). Information items are kept in specific folders and remains there till the user decides otherwise. The single classification that comes along with the hierarchical structure seems to ease the cognitive load over the user. Maybe classification is hard but it is done only once for each information item. Also, folders externalize the internal representation of a problem and help its decomposition (Jones et al., 2005; Zhang & Twidale, 2010). Retrieval by navigation eases the cognitive load on the user by use of procedural memory and recognition instead of recall. Furthermore, the location metaphor of folders suits the physical world that we are used to from our early childhood.

The force of the habit can also explain the preferences for folders over tags. The participants are used to use folders for PIM. This is the first possibility they have learned.

Will tags replace folders in PIM? Only time will tell what the future holds for us.

System No. 002363532