Collaborative Annotation Sharing in Physical and Digital Worlds

Jan Grbac, Matjaž Kljun, Klen Čopič Pucihar, Leo Gombač

FAMNIT, University of Primorska, Koper, Slovenia
jan.grbac@student.upr.si,
{matjaz.kljun,klen.copic,}@famnit.upr.si,
leo.gombac@student.upr.si

Abstract. Despite the existence of a plethora of annotating software for digital documents, many users still prefer reading and annotating them physically on paper. While others have proposed the idea of merging these two worlds, none of them fits all the design requirements identified in this paper (working in real-time, use readily available hardware, augment physical annotations with digital content, support annotation sharing and collaborative learning). In this paper we present the implemented prototype and a focus group study aimed at under-standing studying habits and how the system would fit in these. The focus group revealed that paper material is often discarded or archived and annotations lost, web resources are not saved and fade with time, and that the prototype proposed fits in their studying habits and does not introduce any privacy concerns — be it ones related to the prototype's camera (used in public setting) or ones related to annotations sharing.

Keywords: document annotation · change detection · computer vision · physical-digital interaction ·

1   Introduction

Digital revolution has changed a large part our lives, including the way in which we read annotate and interact with documents. During the early ascent of personal computers many have predicted the end of the paper. Quite contrary, we use more paper than ever before [1]. Moreover, we create more information than any other generation before us — be it video, audio or textual. It is the vast abundance of content that is so easily created in the digital form that has contributed to more printing. Nevertheless, many still prefer reading on paper for various reasons: reading on paper eases the learning process [2], using a pen is a ubiquitous process compared to all annotating software [3, 4] and the sheer physicality of the paper provides a sense of possession, control and manipulation. Reading and creating annotations is thus a process that the physical world supports very well and affordances of the physical world are hard to surpass [3, 4]. Dog-ears, immediately visible bookmarks, high density of information available at the same time, ease of visual scanning, persistence of infor-

mation are just a few from the list. However, digital world brings other advantages to the table such as the ease of editing, sharing, duplicating, versioning, archiving, collaborating and searching especially when a large number of documents is in question.

The gap between digital and physical worlds has not diminished. We still struggle to keep both worlds in sync, which is notable also in the domain of learning and acquiring new knowledge, and which happens through annotations as one of the basic practices in these processes. Despite various attempts to bring physical and digital annotation together as described in next section, we still do not have a viable solution for merging them. In this paper we propose a system that would perform real-time capture of physically created annotations and mirror them to digital documents using a laptop and a smartphone camera. The proposed system aims at expanding physically created annotations with digital world capabilities such as real-time translations, broadening content with URL links, digital archiving (preserve annotations even when paper gets discarded) and most importantly sharing annotations with others. By filtering, ranking, moving and modifying the system strives to support users in integrating annotations of others into their mental model map that fits their personal learning style. In addition to solution design, we presented an early prototype of such a system to users in focus groups who provided an insight into how such a system would be accepted (e.g. acceptance of camera use in a public space such as library) and how sharing would fit into their learning process (e.g. how is intimacy, privacy and sense of ownership of one's annotations perceived).

2 Previous Work

There have been many attempts to bring physical world closer to the digital. The earlier attempts were to synchronise the files between an electronic and digital file cabinets. For example, the Self-organising file cabinet enhanced physical file cabinet and allowed users to import physical documents in a digital form and then annotate, organize, update and find information in the digital and consequently in the physical world [5]. Another such example is Protofoil that let users find paper documents in an electronic file cabinet [6].

Several attempts were made to track physical documents in the working environment with the help of digital technology. DigitalDesk used a camera and a projector mounted on the top of an office desk and captured user's interaction with physical documents and projected related digital information providing tangible manipulation of digital content [7]. Kimura used also other sensors that tracked user's movements and interactions in the physical and digital domains and showed these as activity montages on a wall-sized display [8]. Similarly, Magic Touch tracked user's movements on the desk with wearable computing [9]. Tracking of physical documents has been done also by printing special codes on the margins of the papers such in PaperSpace [10]. This computer vision based system allowed users to locate paper copies of printed digital documents and retrieve their digital versions based on these codes. In addition, a set of instructions (annotate, open, link, email, information) were printed in the margins of a physical document as well, which could be activated by system-recognised gestures on them (e.g. by selecting an email icon on the margin of a physi-

cal document resulted in attached digital counterpart to an email). Systems designed and tested by Wendy et. al looked at ways of digitizing handwritten annotations using graphics tablet placed below the physical document and a PDA as an interaction lens for attaching digital annotations such as electronic documents [11]. Other systems went even further and tried to bring the physicality of paper documents in the digital domain through augmented reality (AR) such in BubbleFish [12] that projected digital documents in the physical environment or Pacer [13] that allowed highlighting in physical documents through phone's screen. Similar "direct manipulation" AR approach was implemented in several recent prototypes such as in [4, 14–17]. Using different approaches and different levels of immersion, the above presented research successfully blended the two worlds together — however, it lacked particular focus on study processes and requires system setup that are not easily accessible or easily transferable, or create annotations in digital domain, which are visible in physical space only through phone's screen.

It has been noted in the academic literature that reading occurs most often in conjunction to writing than not [4, 18]. Reading accompanied with writing (drawing, underlining, highlighting) as a support process helps users to form a conceptual understanding of the text while these secondary tasks are requiring no or very little cognitive attention. In such contexts the benefits of paper outweigh that of digital documents. A few systems have been proposed in this line of research that tried to bring physical annotations back to digital documents.   , Paper Augmented Digital Documents (PADD) [21], and S-notebook [22]. However, these systems require specially developed input devices (so called digital pens with a high precision micro camera integrated), and micro patterned paper that supports transference of physical annotations (writing and drawing) to digital documents. In some cases, they are restricted to only a set of recognisable patterns and only on predefined areas on forms. There have even been attempts to bring the physical affordances to the digital world such as in [23], OneNote and similar annotation software. However, we are not discussing these since they try to eliminate paper, which, as explained above provides tangibility, persistence and other affordances that can better support studying processes.

The limitation of the a above presented studies is that they require either specially equipped paper and input devices. In addition to this, none of them fulfils all of requirements we highlight in the design requirement section, particularly the collaboration aspect of visualising other users' annotations that allows grouping and grading these based on one's personal style of learning. As the designed system focuses on collaborative learning this paper presents also an exploratory study into (i) annotation sharing and possible intimacy/privacy implications users might perceive in doing it, and (ii) intimacy/privacy concerns of using camera phone when studying in public spaces such as libraries or other people's dwellings.

3   System Design

For our initial prototype design, we set up several requirements gathered from the literature. Studying commonly shifts from public/shared (library) to private/intimate environments (room) and requires a portable system with lowest possible amount of energy needed to move and set it up at a new location [24, 25]. Beside portability,

social acceptance of the designed system need to be taken into account [26]. For example, one of the reasons why Google Glass has not been successful is the fact that its benefits failed to overshadow its social acceptance. Adding the camera and see-through display to traditional eyeglasses was too intrusive and obvious, and attracted unwanted attention to people wearing them (imagine someone with such glasses coming into our intimate space such as home and filming it). It is also very important for such a system to be effective and require limited to no additional effort when adding new annotations — introducing any additional burden to the study process is likely to discourage users, particularly as studying happens over longer periods of time [27]. Annotations also need to be digitalized individually and linked to a particular digital content being annotated in order to enable indirect search of created annotations, its underlying text, and annotation interactivity (e.g. gaining explanations of words being highlighted or pointed at). Besides, such a system also needs a possibility to (i) expand annotation with additional digital content (e.g. images, video, translations/definitions<) [28], and (ii) retrieve this content when needed. The final requirement relates to supporting collaborative learning introducing the need to: (i) view and select annotations curated by others, (ii) rate annotations, and (iii) organize them according to one's personal style of learning [29, 30].

3.1   Solution Design and Prototype Implementation

We present a solution using a laptop computer and mobile device on a stand as such devices are readily available within student population. Laptop is the processing and rendering unit, whilst the mobile phone is used for video capture of a printed document being annotated (see **Figure 1**). Such setup has sufficient processing power (particularly important during annotation synchronization and rendering phase), battery capacity (the system is required to run for an extended amount of study time), screen estate (important for annotation sharing, grouping and organization) and video capturing capabilities (an important aspect of annotation digitalization).

   We provide the user with two views: the virtual view (**Figure 1** left) and the augmented view (**Figure 1** right). In the virtual view, the laptop screen shows the page of the document the user is currently reading. The document is overlaid with personal and shared annotations. The second view is called augmented view and provides a virtual mirror view of the book. In this case laptop screen renders a live video feed of a webcam captured through the prism glass. Depending on the configuration of the optics being used we alter the video in such a way that it shows the captured scene in correct orientation and provides users with virtual reflection (similar to one used by the Osmo[1] system) of their hands and the document segment. As long as the position of the camera in relation to the paper is known, one can augment the document with shared annotations. In addition to this, as long as it is possible to track the pen, it is possible to make the captured surface fully interactive.

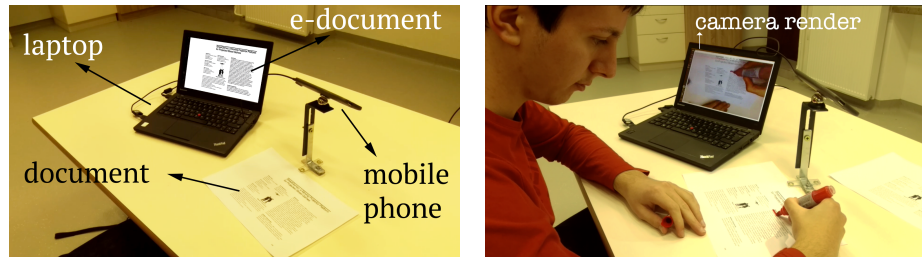
[1] Osmo: https://www.**playosmo**.com

Fig. 1: Left: virtual view in which highlighted text on the paper results in highlighted text on the screen. Right: Augmented view shows a real-time video of user annotating with augmented content on a screen

From software implementation perspective the solution is divided into 3 parts: (1) annotation synchronization; (2) annotations actions in digital domain; (3) annotation sharing, grouping and organization that suits one's personal learning style. Each segment is discussed individually within the following subsections.

Fig. 2: Prototype showing four options (Dictionary, Wiki, Define, Image Search) that can be executed on digitalized highlight "Moving Projector Platform"

Annotation synchronization:

The solution utilizes video stream captured by the phone's camera in order to retrieve the e-version of the printed page. After a page is identified, the system looks for differences between the digital and physical version of the document and transcribes user's annotations to the digital version utilizing colour based tracking. This is done in a 3 stage approach. In first stage a colour filter is applied for blob detection, which is then used in second stage for detecting location and size of these blobs. The final stage is transformation of annotation locations to the document coordinate system. In order to do this, we need to know the position and orientation of the camera phone in relation to the page, which can be done using various camera pose tracking techniques [31, 32]. The prototype (see Figure 2) is implemented using openCV library and is able to digitize annotations coloured in red, green or pink. Using the position and size of the blob, the prototype appends annotation to the text segment being overlaid.

Annotations actions in digital domain:

There are two basic annotation actions that the system supports: digital search actions based on the text below the annotation (e.g. word translations, wiki search, ...) and augmentation of knowledge with content from the web (e.g. appending URL links to curated annotations). Saved digital links can be revisited by mouse clicks on the digital document or through pointing gestures on a physical document. The prototype supports this feature and allows users to trigger 4 different actions (Dictionary, Wiki, Define, Image Search) using a URL driven API's. By clicking the right button or shortcut on the keyboard, a web search query URL is generated and opened in a web browser. If users find a valuable web resource, they can save its URL to the annotation, which becomes augmented with additional digital information.
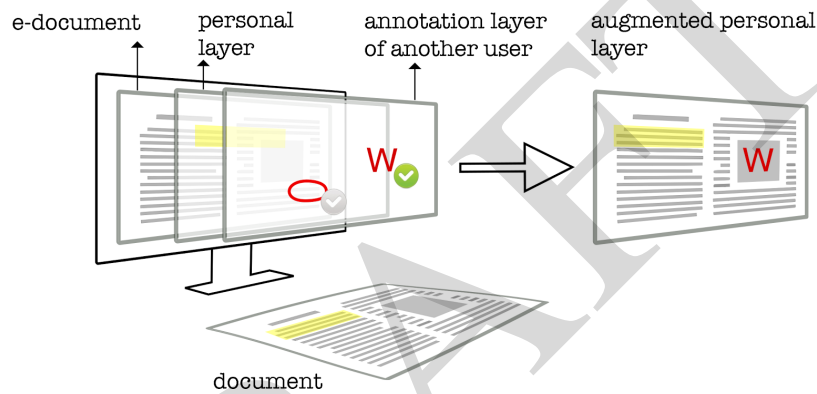

Fig. 3: Design concept for sharing and organizing annotations. Different layers belong to different users. By ticking annotations at various layers of other users, the user adds them to their personal layer. Once on a personal layer, these digital annotations can be moved and modified in order to fit the learning style of the user. The expanded digital content is also transferred to the personal layer.

Annotation sharing, grouping and organization:

Besides viewing one's personal annotations such as highlights, notes and sketches, the page can also be overlaid by annotations curated by other users. The system allows users to choose which annotations they want to keep, where they want to place them on the document and rate their importance. This is important in order to customize the learning environment according to one's personal learning style as demonstrated on Figure 3. This segment of the prototype has not yet been implemented, hence we provide here a concept sketch of the system (see Figure 3).

4   Method

For acquiring the insight into how users would perceive intimacy and privacy of (i) sharing their personal annotations, (ii) people around when using a system and (iii) social acceptance of setting up such a system (in both a public space such as a library

or private space such as schoolmate's bedroom) we conducted a preliminary study before fully developing the prototype. The method chosen was a focus group with students as studying a variety of subjects occupies a large part of their lives. The focus group allowed us and participants to develop a rich group discussion around different opinions. To assure the diversity of studying practices we posted the announcement of the study at all departments at our university. 15 participants answered the announcement (4 females, 11 males) with an average age of 21 from 5 different departments: bioinformatics, applied mathematics, applied economics and finance, computer science, and biology. To further diversify the groups, we held two focus groups with mixed students from different departments as well as of different gender.

The focus group questions have been concentrated around four topics: (i) participants' usual studying settings (e.g. room, kitchen, library) and surroundings (people, studying material, devices), (ii) studying resources used (digital, physical), (iii) annotation sharing, intimacy and privacy concerns, and (iv) participants' opinion about the prototype. For the latest, participants have been shown the system as seen on **Figure 1**. The purpose of the demonstration was to support participants in visualising the style of studying with such a system and further the discussion of the focus group sessions. Both sessions have been filmed, videos transcribed and coded by two researchers. The main findings are presented in following section.

5 Discussion

Discussion section is presented around four topics of the focus group session. The empirical findings provide implications and uncover possible limitations of the proposed prototype.

5.1 Studying environment

Except for one participant who studies whenever opportunity arises (e.g. while commuting, between lectures on a bench, etc.), the majority mainly use their room desks and library study rooms as studying environments. In both settings users have access to all technology needed for the prototype to work: laptop, mobile phone, power plug and internet connection. While room desk provides an intimate and private studying environment where *"things can be left as they are" (p10),* library environment requires users to prepare and clean their studying setup. For the later, the prototype needs to be portable and easy to setup.

5.2 Studying resources

As our system supports annotations synchronisation between physical and digital document we wanted to find out what materials students use when studying. In line with other studies, participants expressed a preference for studying and annotating on paper [1, 4]. However, the kind of studying material available depended on a course subject — hence it was not always possible or rational to print out the materials prepared for a lecture. The mentioned reasons were: printing large quantities of material

is expensive, lecture slides have little or no text to be worth printing, lecturers deliver content on a blackboard, and some subjects (e.g. mathematics) do not require learning by reading. In such cases, taking notes in exercise books is a preferred way of creating the material that participants later study from. This is a limitation for our prototype since exercise books of each participant in the course are unique, therefore, even though our system could capture created annotations (for the user to extend them with digital content) such annotations could not be searched and shared with others in a meaningful way. To achieve such annotation sharing, both, the annotations and the exercise book need to be shared.

Regardless of study material format (printed documents or exercise books) participants also use other resources (books, web) to supplement existing material and to clarify the created content. Interestingly, participants never store links to supplementary resources, but transcribe the relevant content to the main study material. If they need to revisit this resource, they browse for it again (in a book or on the web). Such practice does not externalise these mental links, which cannot be shared with other users and fade with passing time [33 p159].

During the discussion, participants identified the loss of links as problematic, but also highlighted that an even bigger problem is the fact that paper material (either exercise books or annotated printed material) is either archived and never looked at again (it is difficult to access and search through such material) or discarded due to lack of space and required effort. They also mentioned that on several occasions, having such content would come in handy, but it was not possible to use it. On the other hand, the majority of participants kept their digital content as it does not take up physical space and was easily retrieved by searching. If used, our prototype mitigates the loss of physical study material with digitalised annotations and enhances such annotations with indirect digital search capacity through such material.

5.3   Annotation sharing, intimacy privacy concerns

When asked about sharing their annotations, focus group members did not highlight any intimacy and privacy concerns. Some questioned if such annotations would be meaningful to other participants, but not all agreed and highlighted that they regularly share annotations with classmates who photocopy them. However, the later group stressed out that shared annotations are not used in current format, but recycled and integrated into one's own annotations. This is not surprising as annotations represent personal mental model maps, hence are most useful for the person who created them [34]. The implication for the prototype is that the system needs to focus on enabling user's highest possible flexibility when integrating shared annotations into one's personal layer.

5.4   The prototype

Participants generally liked the prototype and found its use socially acceptable in private and public domain. No concerns were raised when participants were asked about acceptability of the system if used by their neighbours in the library or colleagues studying in their rooms (as long as the system did not make additional noise).

The use of camera does not seem to cause any concern amongst participants. This was expected as the camera is pointed towards the table and only captures tabletop surface in-front of the user, which is very unlikely to raise privacy issues. If used in silent mode, mobile phone and laptop use in public and others' people private environments, is nowadays acceptable and even supported through the provision of internet, power access and laptop renting.

Participants found the system as too cumbersome to move for daily use at lectures. However, most agreed that during exam periods, they do not see mobility as problematic because they stay in the same place for extended period of time or study in private setting where clearing one's desk after use is not required. When asked about the extended set of features they would like to see, participants highlighted that they would like to be able to create links to a particular segment of the webpage. This idea was expended to videos where participants expressed the need to create link to a particular segment of a video.

6   Conclusion

Real-time digitalisation of physical annotations in order to archive, share, search, and expand them can bring added value to the process of acquiring new knowledge while digitally preserving it for the future. The implemented prototype demonstrates that such a system is viable on hardware that is readily available within the student population. In addition to this, the presented focus group sessions also highlighted that such hardware configuration is acceptable in private and public domains. The sessions also revealed that finding supplementary digital information resulted in a failure to link it to study material on paper and losing it in the long run (e.g. writing down URLs as annotations is not always a suitable solution), and that even paper material is often discarded, lost or archived in a way which makes it difficult to use again. In addition, the focus group also highlighted that the prototype proposed fits in their studying habits and does not introduce any privacy concerns — be it ones related to the prototype's camera (used in public or others' people private setting) or ones related to annotations sharing. At last, sharing annotations as supported by our prototype was seen a valuable feature complementing and expanding sharing that is already happening in physical world (students are photocopying notes from one another) where users recycle their colleagues' annotations and make them fit their own studying process and mental models.

We are currently building a full prototype, which will be studied both in the lab and in the wild. The former will measure usefulness, usability, and scalability (e.g. how many users can use it together) of the prototype in a predefined task that will include reading a selected text, free annotating the text and viewing (selecting, rating) annotations of other users (researchers). After this study, we plan to use the prototype in a long-term study run as part of university course which is based on reading research papers.

References

1. Sellen, A.J., Harper, R.H.R.: The Myth of the Paperless Office. MIT press (2001).
2. Mangen, A., Walgermo, B.R., Brønnick, K.: Reading linear texts on paper versus computer screen: Effects on reading comprehension. Int. J. Educ. Res. 58, 61–68 (2013).
3. Marshall, C.C., Bly, S.: Turning the page on navigation. Proc. 5th ACM/IEEE-CS Jt. Conf. Digit. Libr. (JCDL '05). (2005).
4. Mueller, P. a, Oppenheimer, D.M.: The Pen Is Mightier Than the Keyboard: Advantages of Longhand Over Laptop Note Taking. Psychol. Sci. (2014).
5. Lawrie, D., Rus, D.: A self-organized file cabinet. Proceedings of the eighth international conference on Information and knowledge management - CIKM '99. pp. 499–506. ACM Press, New York, New York, USA (1999).
6. Rao, R., Card, S.K., Johnson, W., Klotz, L., Trigg, R.H.: Protofoil: storing and finding the information worker's paper documents in an electronic file cabinet. Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '94. pp. 180–185. ACM Press, New York, New York, USA (1994).
7. Newman, W., Wellner, P.: A desk supporting computer-based interaction with paper documents. Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '92. pp. 587–592. ACM Press, New York, New York, USA (1992).
8. Voida, S., Mynatt, E.D., MacIntyre, B., Corso, G.M.: Integrating virtual and physical context to support knowledge workers. IEEE Pervasive Comput. 1, 73–79 (2002).
9. Pederson, T.: Magic Touch: A Simple Object Location Tracking System Enabling the Development of Physical-Virtual Artefacts in Office Environments. Pers. Ubiquitous Comput. 5, 54–57 (2001).
10. Smith, J., Long, J., Lung, T., Anwar, M.M., Subramanian, S.: PaperSpace: a system for managing digital and paper documents. Proceedings of the SIGCHI conference extended abstracts on Human factors in computing systems - CHI EA '06. p. 1343. ACM Press, New York, New York, USA (2006).
11. Mackay, W.E.: The missing link: integrating paper and electronic documents. Proc. 15th French-speaking Conf. human-computer Interact. 15eme Conf. Francoph. sur l'Interaction Homme-Machine. 1–8 (2003).
12. Fallman, D.: The BubbleFish: Digital Documents Available On Hand. INTERACT '01: Proceedings of IFIP TC13 Conference on Human-Computer Interaction. , Waseda University Conference Center, Shinjuku, Tokyo, Japan (2001).
13. Liao, C., Liu, Q., Liew, B., Wilcox, L.: Pacer: fine-grained interactive paper via camera-touch hybrid gestures on a cell phone. … SIGCHI Conf. Hum. …. 2441–2450 (2010).
14. Hardy, J.: Experiences: a year in the life of an interactive desk. Proceedings of the Designing Interactive Systems Conference on - DIS '12. p. 679. ACM Press, New York, New York, USA (2012).
15. Hincapié-Ramos, J., Roscher, S.: tPad: designing transparent-display mobile interactions. Proc. DIS '14. 161–170 (2014).
16. Hincapié-Ramos, J.D., Roscher, S., Büschel, W., Kister, U., Dachselt, R., Irani, P.: cAR: Contact Augmented Reality with Transparent-Display Mobile Devices.

Proceedings of The International Symposium on Pervasive Displays. pp. 80:80–80:85 (2014).

17. Kunze, K., Tanaka, K., Iwamura, M., Kise, K.: Annotate me - Supporting active reading using real-time document image retrieval on mobile devices. UbiComp 2013 Adjunct - Adjunct Publication of the 2013 ACM Conference on Ubiquitous Computing. pp. 231–234 (2013).

18. Adler, A., Gujar, A., Harrison, B.L., O'Hara, K., Sellen, A.: A diary study of work-related reading: design implications for digital reading devices. Proceedings of the SIGCHI conference on Human factors in computing systems - CHI. pp. 241–248 (1998).

19. West, D., Quigley, A., Kay, J.: MEMENTO: A digital-physical scrapbook for memory sharing. Personal and Ubiquitous Computing. pp. 313–328 (2007).

20. Johnson, W., Jellinek, H., Klotz, L., Rao, R., Card, S.K.: Bridging the paper and electronic worlds: the paper user interface. Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems. pp. 507–512 (1993).

21. Guimbretière, F.: Paper augmented digital documents. Proc. 16th Annu. ACM Symp. User interface Softw. Technol. UIST 2004. pp, 51–60 (2003).

22. Pietrzak, T., Malacria, S., Lecolinet, É.: S-Notebook. Proc. Int. Work. Conf. Adv. Vis. Interfaces - AVI '12. (2012).

23. Pearson, J., Buchanan, G., Thimbleby, H.: The reading desk: applying physical interactions to digital documents. In: Jones, M.J. and Palanque, P. (eds.) Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11. p. 3199 (2011).

24. Applegate, R.: The Library Is for Studying. J. Acad. Librariansh. 35, 341–346 (2009).

25. Harrop, D., Turpin, B.: A Study Exploring Learners' Informal Learning Space Behaviors, Attitudes, and Preferences. New Rev. Acad. Librariansh. 19, 58–77 (2013).

26. Roesner, F., Kohno, T., Molnar, D.: Security and privacy for augmented reality systems. Commun. ACM. 57, 88–96 (2014).

27. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: Toward a unified view. MIS Q. 27, 425–478 (2003).

28. Sakar, A., Ercetin, G.: Effectiveness of hypermedia annotations for foreign language reading. J. Comput. Assist. Learn. 21, 28–38 (2005).

29. Marshall, C.C., Brush, a. J.B.: Exploring the relationship between personal and public annotations. Proc. 2004 Jt. ACM/IEEE Conf. Digit. Libr. 2004. 349–357 (2004).

30. Marshall, C.C., Brush, A.J.B.: From personal to shared annotations. CHI '02 extended abstracts on Human factors in computing systems - CHI '02. p. 812. ACM Press, New York, New York, USA (2002).

31. Koller, D., Klinker, G., Rose, E., Breen, D., Whitaker, R., Tuceryan, M.: Real-time vision-based camera tracking for augmented reality applications. Proc. ACM Symp. Virtual Real. Softw. Technol. - VRST '97. 87–94 (1997).

32. Bradski, G., Kaehler, A.: Learning OpenCV: Computer Vision with the OpenCV Library. (2008).

33. Jones, W.: Keeping Found Things Found: The Study and Practice of Personal Information Management. Morgan Kaufman, Burlington, MA (2007).

34. Payne, S.J.: Users' Mental Models: The Very Ideas. HCI Models, Theories, and