PalimPost: Information Convergence Using Sticky Notes

Li Bian
MIT Media Laboratory
75 Amherst St.
Cambridge, MA
libian@media.mit.edu

Roy Shilkrot
MIT Media Laboratory
75 Amherst St.
Cambridge, MA
roys@media.mit.edu

ABSTRACT

In today's world, the digital information retrieval experience is inherently a sparse device-centric activity. Users rely on the ability of the currently used device to supply the requested information, in some disconnection from past activities on other devices. There is a growing need to develop new methods of connecting cross-context information retrieval sessions. We present PalimPost, a converged system for storing, searching, and sharing digital and physical world information using sticky notes and mobile devices. PalimPost extracts contextual cues from a user's physical environment and activities, and connects them to the user's digital world research. Subsequently, the system presents systematically categorized information that is relevant to the moment of interaction in a just-in-time manner. PalimPost uses physical sticky notes with embedded QR codes, as well as virtual sticky notes on mobile devices. The system incorporates Automatic Speech Recognition (ASR), Optical Character Recognition (OCR), and Natural Language Processing (NLP) techniques for understanding and categorizing the content.

Categories and Subject Descriptors

H.5.2 User Interfaces,
H.3 INFORMATION STORAGE AND RETRIEVAL - Content Analysis and Indexing, Information Storage, Information Search and Retrieval

General Terms

Algorithms, Design, Human Factors.

Keywords

Internet of things, embedded information technology, augmented reality, cloud computing, tangible user interfaces.

1. INTRODUCTION

In recent years advances in desktop and mobile technologies have enabled a myriad of techniques for synchronization of information gathered at remote locations. However, even in simple everyday tasks, such as buying a shirt, there is still a great difference between online browsing and in store browsing. PalimPost suggests a new method of connecting such discrete activities into coherent and continuous experiences. The system allows the users to stick to existing habits, such as using sticky-notes and scanning QR codes, so that the learning curve of using the system is minimized.

PalimPost is a system for information storage and retrieval that works both in desktop and mobile environments. At the desktop environment users already have an established set of tools to perform and store research, such as Evernote [1] and OneNote [2]. These tools allow users to save files of almost all types of media, such as text, audio, image and video; they also allow synchronization of information across devices and platforms. However, the interaction is still device-centric, does not present a physical aspect, and keeps the acquired data strictly in its original form.

PalimPost makes use of sticky notes to give digital information physical embodiment and to allow tangible arrangement and sharing of the information. The online exploration and thoughts that users may have in their research process can be saved into a physical sticky note. This is achieved by means of saving browsing sessions, capturing the screen or recording audio and video with the embedded camera, and uploading the information to a cloud-based storage. So far, sticky-notes have been used to save a very concise piece of information, as a mental reminder of much bigger and complex information. Our system allows the sticky note to have a much larger storage capability, in order to help users retain more information while keeping the tangibility of a paper note.

By using centralized cloud storage and algorithmic analysis, PalimPost allows information to be integrated seamlessly among devices, with minimal user instruction. The system uses ASR and OCR technology to extract keywords from spoken sentences or images. The keywords are used not only to tag the information for later use, but also to look for links between incoming information and existing stored information. To search for these links, we employ techniques in NLP, based on *bags-of-words*, dictionary definitions and statistical classifiers of the extracted information. The links create a more seamless experience for the users, as they are almost completely relieved of the process of categorizing.

PalimPost re-creates the experience of research for information across multiple locations and devices, while giving a physical embodiment to the stored digital data. In Section 3 we will describe the many aspects of the system, and how they are integrated.

2. RELATED WORK

2.1 Sticky Notes

Much research has been done in the past based upon people's behavior patterns using sticky notes or paper in general. Recently, effort has been devoted into bridging the gap between the physical and digital worlds, where paper is used to form seamlessly connected experiences. Specifically, Quickies [3] is a notable project that combines normal sticky notes with Artificial Intelligence, Natural Language Processing, and Ink Recognition technologies to help people who are not tech savvy to store information digitally by writing on paper sticky notes. PaperLink [4] augments paper documents with electronic features, using a highlighter pen that comes with a camera, and computer vision and pattern recognition techniques. Insight Lab [5] creates an

immersive teamwork environment by linking paper, whiteboard printouts, and multimedia data using barcode technologies. WebStickers [6] is a system that allows associating web bookmarks with barcodes that are attached to sticky notes. However this work was done before the age of smart-phones and web cameras embedded in personal computers, and therefore it was confined to desktop usage coupled with a barcode scanner. The proposed system is of a more ubiquitous nature, and suggests usage scenarios relating to mobile, desktop and the connection between them.

2.2 Mobile and Desktop Convergence for Information Storage

Many solutions exist for synchronization of simple information between the desktop and mobile, such as Catch [7], MobileNoter [8] and SHIFD [9]. These applications have minimal support for saving audio-visual information, and concentrate on short textual notes, much like physical sticky notes. However more competent solutions exist, such as Evernote [1], Microsoft's OneNote [2] and Google's Docs [10], that allow many types of data to be shared between mobile and desktop devices. These applications, although massively adopted by users, only present a data-storage driven solution, while our proposed system concerns data convergence and saving the context rather than only raw information.

3. PALIMPOST

3.1 What is PalimPost?

PalimPost is a system for digital-physical and cross-device information convergence using sticky notes and mobile devices. The mobile application extracts contextual cues from a user's physical environment and activities, and connects them to the user's digital world research. An automatic topic categorization process aids the users in organizing the relevant information into the research domains in a just-in-time manner. Whether a user is writing down a shopping list on a sticky note after surfing the internet at home or checking out hundreds of products at hand in a physical store, PalimPost integrates information from different research sessions to form a seamlessly connected experience for the user. Figure 1 depicts a high-level explanation of how the system is used.

PalimPost is inspired by people's existing behavior and habits of using sticky notes and mobile devices. When a user is sitting in front of a computer at home, she still retains the habit of writing down important information in a short message on a sticky note and attaching the note right next to her computer screen. The content in the message can often include a person's contact information such as name and phone number, a shopping list with product specifications as a result of browsing e-commerce websites, or the address of a physical location. All this information, coupled with the user's Internet activity such as webpages opened, represents the state of interaction between the user and information carrier. To save this state of interaction at the specific moment, a user can simply scan the physical sticky note with QR code printed on it. Upon scanning, the message on the sticky note becomes a tag of the digital information, containing the state of work.

When the user is on the go, he no longer has his physical sticky note available with him. However, the information stored through scanning the sticky note earlier can now be retrieved from the cloud. A mobile phone application allows the user to access all

the previously saved information, including the tag and the digital information. Moreover, the application allows the user to add
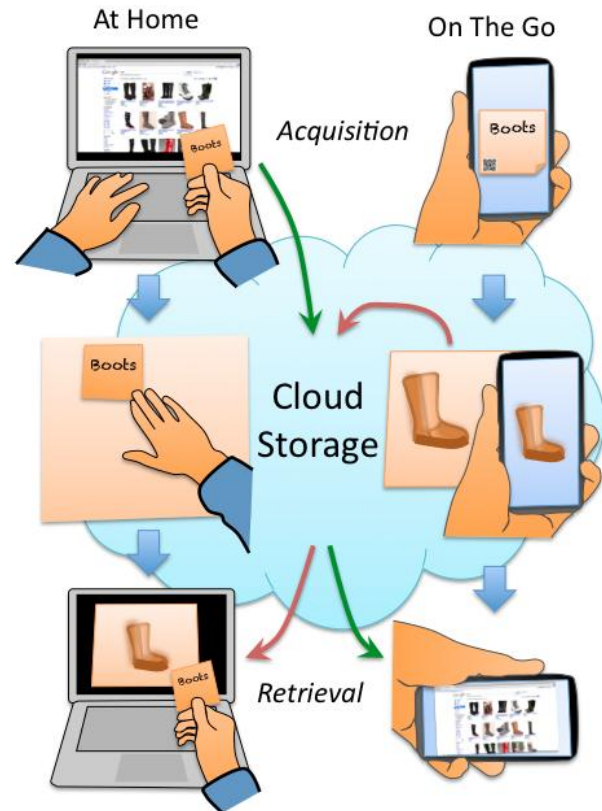

Figure 1. The usage flow of the PalimPost system. Both at home and on the go the data is stored in the cloud. The data is linked, and so any device can retrieve information acquired in all devices.

new information to the stored sessions, such as information retrieved by scanning a product barcode in the store, a photo taken of a store product, the logo or name of manufacturer on the product, a store assistant's recommendation, or the user's own momentary thoughts regarding a product. The mobile phone application contains virtual sticky notes, with which the user can continue taking notes, through ASR, and tagging the new information he encounters on the go. Section 3.3 details how PalimPost enables this experience.

The PalimPost system automatically groups and categorizes information obtained at home and on the go by statistically analyzing the content from both physical and virtual sticky notes. This enables a highly clean interface to the database of existing notes, in which the user would like to store new acquired information. Section 3.4 details the techniques of finding links in the PalimPost database on the fly.

3.2 Usage Scenarios

In this section, we present a few application scenarios to further explain how a user can use the PalimPost system in various contexts. All these scenarios involve a user conducting different activities at different locations and time. Yet, with PalimPost, the user can access information obtained through these activities

anywhere and anytime while retaining her usual habit of using sticky notes and mobile devices.

Buying New Shoes. With the flourishing of e-commerce websites in the past decade, online shopping experience has improved tremendously. A customer no longer needs to compare prices among stores on the street. All the brands and prices are neatly

Figure 2. User speaks to record a measurement at the store.

arranged in tables on webpages, including affordable shipping options. However, only at a physical store can a customer touch and feel the actual products. In fact, for any wearable products such as shoes, a customer often risks wasting time buying products that don't end up fitting nicely due to the discrepancy between expectation from online research and findings of the real products in physical stores.

Assuming a user has searched at home on Amazon, Ebay and various websites for the shoes she likes, she now wants to try the shoes on to see how they fit and look on her. She writes down on a physical PalimPost sticky note "buy shoes," scan the note with her computer, upon which a virtual PalimPost note with the same content is created and the user's browsing sessions on the websites are saved and linked to the sticky note.

Later, the user goes by a shoe store. She accesses the "buy shoes" sticky note on her mobile phone and looks up all the brands and models she liked from online research at home. Browsing through the physical store's inventory, she finds a few models from her list and tries them on. She takes pictures of the pairs that actually fit and adds the pictures to the "buy shoes" sticky note on her mobile phone. A store assistant tells the user that this year's popular trend is brown boots; the user says to the phone "brown" to take notes. The user also notices the subtle difference of the color of one pair of shoes from the picture and under physical lighting condition. She clicks on the picture of this pair and says to her phone "darker brown" to add information about this pair of shoes.

When the user returns home, she scans the "buy shoes" sticky note in front of her computer; the browsed webpages from home online research and all the pictures of her wearing the shoes at the store, along with the added notes, show up on the computer screen. Now the user can make a better decision, having explored both online and store offers.

Buying a Webcam. A user needs to buy a webcam for his laptop to video chat with friends. He goes online trying to find the best option. Finally, he sees a post from a friend's Facebook wall that says "Logitech webcams rock!" He goes on to Logitech's company website, decides that he likes the model HD Pro

Webcam C910. He then grabs a sticky note from his desk and writes down "Logitech HD Pro C910". However, an Amazon review says that the cable of this model is too short. So the user scans his sticky note in front of the computer to store all his present research information, and heads out to a nearby BestBuy. When the user reaches BestBuy, he takes out his phone and opens up the PalimPost app, which contains the model of the specific webcam he liked, and other information he found while researching. The user asks the store assistant to open up the box to take a look, and decides that the cable is not too short for his laptop. He buys the webcam at the store and goes home satisfied. Figure 2 illustrates how the mobile application may be used in a shopping environment.

Dinner Preparation. Cooking food is considered by many as a form of art, as it requires instincts, feelings and inspiration to create an exciting dish. Millions of recipes exist online already, let alone in cookbooks, on TV shows and any other form of media, yet chefs and aspiring cooks keep inventing more and more. When cooking original food, there is a stage of research, a stage of inspiration or improvisation, and a stage of decision.

A sous-chef for an upcoming restaurant is given the task to create a new main course dinner dish that includes lobster meat in it. Initially she goes online to search for existing dishes using lobster meat, and saves her findings on a PalimPost sticky note marked "Lobster Dish". She also records a short voice message of herself saying, "Lobster could go very nicely with asparagus or garlic." Later, she goes to the market. She finds some asparagus and records a new message about the smell and texture of it. The system links her transcribed voice message with the existing PalimPost sticky note regarding asparagus. When she returns, the chef scans the physical PalimPost note to retrieve all her online exploration and the sensory experience in the market. She now has a better idea of how to cook a lobster with asparagus.

In the next section, we will describe in detail how the PalimPost

(a) A physical note   (b) A virtual note

Figure 3. PalimPost notes, physical and virtual.

system works.

3.3 Technical Description of the System

PalimPost uses physical sticky notes with QR code printed on them, as well as virtual sticky notes on mobile devices. The system incorporates ASR, OCR, and NLP techniques for understanding and categorizing sticky note contents. It transfers information through the cloud, searches and retrieves relevant results based upon contextual cues harvested from the content. Figure 3 shows an example of virtual and physical notes used by our system.

Physical Sticky Notes. In a desktop environment, the users can use a physical sticky note to store the current state of work. They can use a pen to write a short message on the note, and in the moment of scanning the digital information is attached to the note.
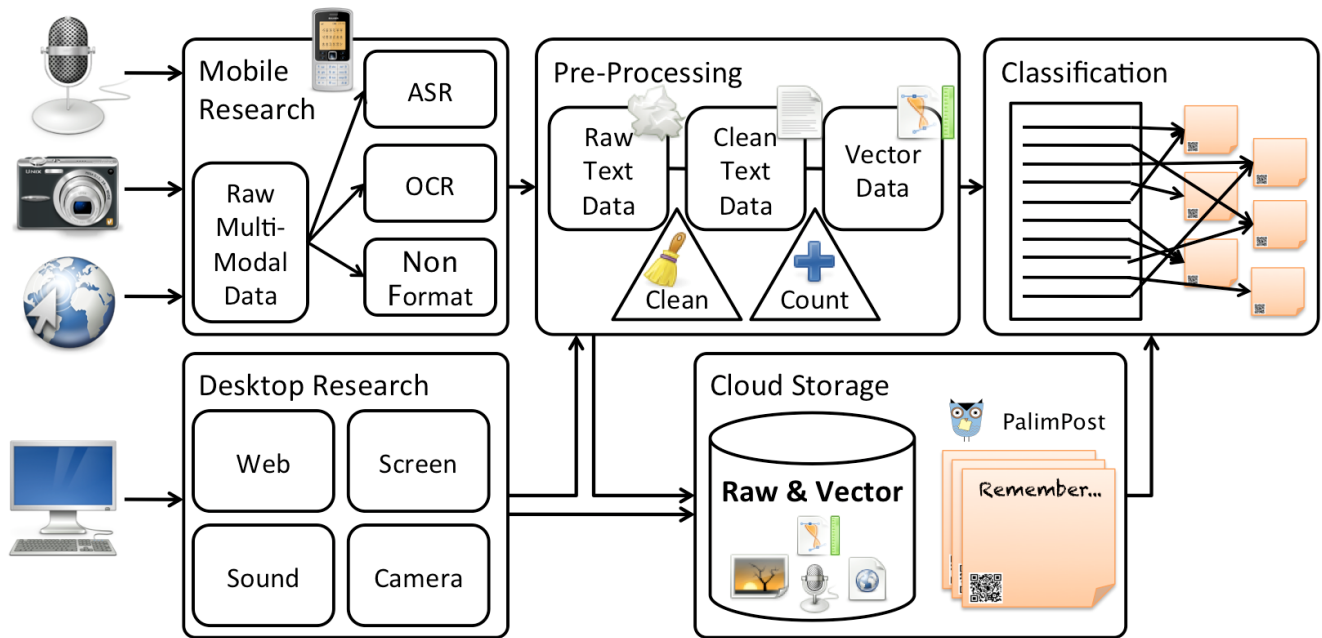
Figure 4. The flow of information and data processing in the PalimPost system. Raw data is acquired either on the desktop or mobile, and then converted into raw textual format. Then, the data is cleaned and converted to vectors, and finally categorized into existing PalimPost notes.

Scanning is done using the camera attached to the computer, which is constantly activated as waiting for a QR code to be scanned. When the code is scanned, in order to capture a full sensory view of the computer, the application takes a screenshot capture, records the URLs of the open tabs of the Internet browser, and records a sound message and the image seen through the camera. Each of these aspects carry a wealth of contextual information on the intention and status of the user in the moment of recording, so when they are replayed, the users can easily be reminded of what they were doing. All the information is sent to a web-service over an HTTP session, which stores the information associated with the unique ID that was given to the note.

When re-scanning a PalimPost sticky note, the desktop application queries the server for existing data. If such data exists, it is fetched and displayed back on the screen. The URLs are re-opened in a browser window, the screenshot is shown and the sound is played back.

Virtual Notes. When a user is out of the house, for example in the store, he can use his phone's camera to scan QR codes on products, or take pictures, sound and video clips. The recommendations that other people give to the user at the store, whether sales representative or fellow customers, can also be recorded and provide a more lively version of product reviews. An application on the smart-phone creates a virtual PalimPost note on the screen, where the recorded information accumulates. The user's speech is transcribed and appears as text on the virtual note, and textual cues are also searched in the images, while the data from any scanned QR codes is added. Naturally, any acquired information is sent to the server.

On the server side, aside from saving the information, the server tries to link the content of the virtual note to relevant information that the user has been researching at home. Since the information retained in the desktop session is rich in textual data, we use a statistical approach to match it with incoming information, which is discussed in the next section. Once a strong connection between incoming and existing data is determined, a link is created

between the virtual note and the physical note at home, and all information is bundled.

Sharing PalimPost Notes. The users can share their recorded experiences, either physically or virtually via their social network. They can post the information on Facebook or Twitter, by clicking on the share option on the mobile interface. On the other hand, they may share the actual physical sticky note with a friend to reveal the associated digital information within it on the friend's own computer or mobile device.

3.4 Finding Links in PalimPost Data

We approach the problem of finding links in the data stored on PalimPosts as a Text Categorization (TC) problem. In our system we strive to *normalize* the multi-modal data, e.g. images, sound, etc., into chunks of textual data. We do that using standard methods of ASR, OCR, and extraction from formatted data, e.g. HTML. All the information is kept in the original natural form, for re-playing purposes, alongside the extracted textual information. This means that every PalimPost note includes a *bag-of-words* to represent it, and therefore we chose to use *bag-of-words* based techniques for TC. See Figure 4 for a description of the classification method.

Working with Unclean Data. As the quality of ASR and OCR algorithms is not perfect, in many cases the results contain words with no meaning and words recognized as punctuation marks. We work with a dictionary [11] to prune only meaningful words, and completely remove punctuation marks. As we are not interested in understanding the meaning of the sentences in the text, but rather their global domain, we are free to break the sentences apart. Prior to using the textual data we also remove *function words*, i.e. articles, prepositions, conjunctions, etc., which have no domain-specific information in them. This approach, on the other hand, confines us to using only *unigrams*, as word adjacency has no contextual meaning.

In our system, some of the text inputs may be very short. Such is the case of transcribed text from audio, where the recorded

message is usually short and concise. When such a short message is detected, via a threshold on the number of words that remain after cleaning, we expand it using dictionary definitions. Using word definitions as an aid for statistical semantic inference is widely practiced [12,13]. After expansion, we clean the text again.

Choosing Terms. A canonical problem in TC is *term selection*. It originates from the need to represent a document as a vector of numbers, using the *bag-of-words*. Most TC techniques use *word-counting* methods, and usually, for efficiency concerns, do not count the occurrence of every possible word in the document. Designers of TC systems choose a method to select which words to count in the documents, to reduce the dimensionality of the problem, but still keep a high classification rate.

Various techniques for *term selection* exist. However, for simplicity purpose, we chose to use a standard method of term selection using the standard highest *inverse document frequency* (*idf*) measure [14,15]. We compute the *idf* measure for each term, being the number of documents containing the term, and keep the 150 words with highest score as our *terms*, also known as *features*. For a *term frequency* measure we use the normalized *term frequency – inverse document frequency* (*tfidf*) measure. We then calculate the normalized *tfidf* measure of these *features* in each *bag-of-words* and get a vector representation for it.

Classification. Although a large number of types of classifiers exist, we chose to take a standard approach, and use a *Naïve Bayes* (NB) Classifier [14,16,17] in a *ranking classification* sense. Training the classifier: First we calculate the representative vector using the method described earlier, for each piece of information in the PalimPost note, and label it with the ID of the note. We then train the NB classifier with the matrix of labeled vectors.

When an incoming *bag-of-words* query arrives in the server, we first transform it to a vector, and then use the NB classifier to get a probability for it to belong to each of the existing PalimPost notes. Then we decide on a link between the query and an existing note by putting the threshold on the probability at 0.5. Although many methods for finding optimal thresholds exist [14,18], we chose to go with a rather naïve one, as we allow for user interaction. In case of any ambiguity, i.e. the query fits for more than one note, the server responds to the mobile application to allow the user to choose from a narrow list of options.

Linking. A link will be determined after classification, and possibly a quick user interaction. The information on the virtual PalimPost note is joined with the information already in storage. This means their IDs are merged under a new ID, and the ID of the physical note will refer to the new one.

Rearranging of the information requires re-training of the classifier, to let new coming information be categorized correctly. However we decided not to re-train the classifier after addition of each new piece of information. Re-training only occurs when a new PalimPost note is created, whether by a merge or generation of a new note.

4. IMPLEMENTATION

Software Supporting Physical Notes. We implemented the desktop application in C++ environment on MacOSX operating system, although the code is easily portable. The framework consists of:

1. An application that operates the camera, built over *OpenCV* 2 C++ APIs, and the QR-scanning library *ZXing*.

2. An application to record the sound off the computer's microphone using *Java Swing/AWT* for UI and *Java Sound API* for audio recording.

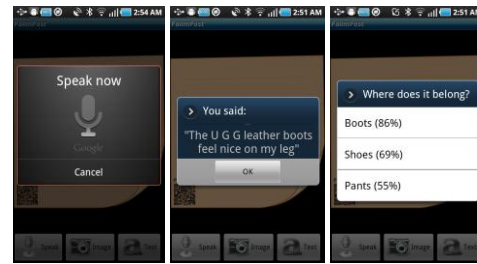3. A script for capturing the Internet browser's state built using

Figure 5. Recording and transcribing menus, followed by the user action to associate the information with an existing note.

MacOSX *AppleScript*.

4. A script for taking a screenshot image using MacOSX built-in function *screencapture*.

Software Supporting Virtual Notes. We implemented the mobile application using Android 2.2 APIs. For Speech Recognition we used the internal Android engine.

Server Implementation. We implemented the PalimPost server in Python environment over a standard Linux distribution. The different aspects of the server were implemented using the following:

1. For Character Recognition we used *Tesseract 3* engine with bindings for *Python*. For image transcoding we used *ImageMagick++* bindings for *Python*.

2. For information storage and retrieval we used a simple file-based architecture with *Python Pickle* for serialization to disk of non-binary data. Binary files, e.g. images, sound clips, etc., were saved as-is as files.

3. The HTTP server was built using *Tornado* standalone *Python* server.

4. For sound transcoding and compression for storage we used *sox* and *ffmpeg* respectively, and glue scripts in *bash*.

5. For Natural Language Processing we used *NLTK*.

5. CONCLUSION AND FUTURE WORK

We presented PalimPost, a converged system for storing, searching, and sharing digital and physical world information using sticky notes and mobile devices. Inspired by people's existing behavior and habits of using sticky notes and mobile devices, PalimPost extracts contextual cues from a user's physical environment and activities, connects them to the user's digital world research, and subsequently presents to the user systematically categorized information that is relevant to the moment of interaction in a just-in-time manner.

PalimPost allows the user to not only input information with a physical sticky note but also retrieve information with the same note. In this way, the physical sticky note becomes a physical embodiment, or tag, of dynamic information with changing content. The amount of content varies according to the user's digital and physical world activities. The system incorporates Speech Recognition, Optical Character Recognition, and Natural

Language Processing techniques into its desktop and mobile platforms for understanding and categorizing information contents. It transfers information through the cloud, searches and retrieves relevant results based upon contextual cues.

In the next steps, we would like to conduct a series of user studies to test our application. We believe that the user studies will give us insights for validating or improving the user interactions we have designed. Moreover, we will further experiment with technologies alternative to those that we have been using. For example, since Speech Recognition is playing a big role in the PalimPost system for users to take notes on the go, we will also try out Nuance's Dragon Naturally Speaking software to see if there is any improvement in the recognition result given our current performance with Android's Speech Recognition.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] "Evernote." http://www.evernote.com/

[2] "OneNote." Microsoft, http://office.microsoft.com/en-us/onenote/

[3] P. Mistry and P. Maes, "Quickies: intelligent sticky notes," *4th International Conference on Intelligent Environments (IE 08)*, 2008, p. P10-P10.

[4] T. Arai, D. Aust, and S.E. Hudson, "PaperLink: a technique for hyperlinking from real paper to electronic content," *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1997, pp. 327-334.

[5] B.M. Lange, M.A. Jones, and J.L. Meyers, "Insight lab: an immersive team environment linking paper, displays, and data," *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press/Addison-Wesley Publishing Co., 1998, p. 550–557.

[6] P. Ljungstrand, J. Redstr\\"om, and L.E. Holmquist, "WebStickers: using physical tokens to access, manage and share bookmarks to the Web," *Proceedings of DARE 2000 on Designing augmented reality environments*, 2000, pp. 23-31.

[7] "Catch." Android, http://www.appbrain.com/app/catch-notes/com.threebanana.notes

[8] "MobileNoter." http://www.mobilenoter.com/

[9] "SHIFD." http://shifd.com/welcome

[10] "Google Docs Mobile."

[11] C. Fellbaum, *WordNet: An electronic lexical database*, The MIT press, 1998.

[12] M. Lesk, "Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone," *Proceedings of the 5th annual international conference on Systems documentation*, 1986, pp. 24-26.

[13] T. Pedersen, S. Patwardhan, and J. Michelizzi, "WordNet:: Similarity: measuring the relatedness of concepts," *Demonstration Papers at HLT-NAACL 2004 on XX*, 2004, pp. 38-41.

[14] F. Sebastiani, "Machine learning in automated text categorization," *ACM computing surveys (CSUR)*, vol. 34, 2002, pp. 1-47.

[15] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval* 1," *Information processing & management*, vol. 24, 1988, pp. 513-523.

[16] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features," *Machine Learning: ECML-98*, 1998, pp. 137-142.

[17] T. Joachims and C.-M.U.P.P.A.D.O.F.C. SCIENCE., "A probabilistic analysis of the Rocchio algorithm with TFIDF for text categorization," 1996.

[18] Y. Yang, "An evaluation of statistical approaches to text categorization," *Information retrieval*, vol. 1, 1999, pp. 69-90.