A Framework for Supporting Natural Interaction with Printed Matter in Ambient Intelligence Environments

George Margetis, Margherita Antona
Foundation for Research and Technology – Hellas
Heraklion, Greece
e-mail: gmarget@ics.forth.gr, antona@ics.forth.gr

Constantine Stephanidis
Foundation for Research and Technology – Hellas
Heraklion, Greece
Department of Computer Science, University of Crete
Heraklion, Greece
e-mail: cs@ics.forth.gr

***Abstract*—Paper is a widely used material which - through appropriate technological augmentation - has the potential to become a widely accepted means of interaction. Ambient Intelligence bears the promise of smart, adaptive and user-friendly environments, anticipating user needs in an unobtrusive manner. So far, there is no systematic approach to paper augmentation in Ambient Intelligence. Addressing this need, this paper introduces an extensible context-aware interaction framework to enable the integration of printed matter into Ambient Intelligence environments.**

Keywords*—*ambient intelligence; printed matter; interactive paper; natural interaction.

I. INTRODUCTION

Through the centuries, paper prevailed as the major means for information sharing among people. With the invention of the printing press by Gutenberg, a vast burst of information dissemination occurred all over the world, establishing printed matter as an essential part of people's everyday life. Since the early 90's, the idea of digitally augmenting physical paper was intriguing enough to trigger the first research efforts in this direction. Since then, numerous approaches have been proposed, based on paper's affordances, for providing user interaction. Paper-based interaction has the potential to be widely accepted and applied in everyday life, due to a fundamental prop of paper: it is inexpensive and can be found anywhere.

The recent emergence of Ambient Intelligence (AmI) realizes the vision of a technological environment where the emphasis is on greater user-friendliness, provision of more efficient services, user-empowerment, and support for human interactions. In AmI environments, people are surrounded by intelligent intuitive interfaces that are embedded in all kinds of objects, while the environment is capable of recognizing and responding to the presence of different individuals in a seamless, unobtrusive and often invisible way [1]. AmI has profound consequences on the type, content and functionality of the emerging digital products and services, as well as on the way people interact with them, bringing about multiple new requirements.

Aarts and Marzano in [2] discuss the fundamental features that characterize Ambient Intelligence environments, which can be summarized in five concepts: technology embedment, context awareness, personalization, adaptiveness and anticipation. Although several approaches have contributed frameworks that embed technology in everyday environments, anticipate and address everyday life needs in adaptive and personalized ways, and provide natural interaction with the use of physical or smart objects, there are so far no systematic approaches engaging printed matter towards realising such concepts.

This paper discusses a systematic approach to fill this gap by developing an extensible context-aware interaction framework, which will enable the integration of printed matter into AmI environments, thus providing:

- multimodal natural interaction with printed matter
- printed matter augmentation
- a reference model for printed matter context-aware and anticipation mechanisms, based on a proposed ontology.

For the assessment and evaluation of the proposed framework with end users, four Ambient Intelligence applications are presented, constituting indicative real life examples.

The rest of the paper is organized as follows:

Section II discusses related work, highlighting efforts towards printed matter and paper digital augmentation. Section III presents the overall architecture of the proposed framework. Section IV describes a generic approach for printed matter modelling and profiling. Section V discusses the proposed ontology scheme for enabling context awareness in the use of printed matter in AmI Environments. Section VI focuses on two fundamental modules of the proposed framework that are responsible for the augmentation of printed matter and interaction rendering. Section VII presents four example applications that have been designed and developed using the proposed framework. Finally, Section VIII summarises the paper and highlights next steps and future work.

II. RELATED WORK

The idea of digitally augmenting physical paper was firstly introduced in two pioneer systems, namely DigitalDesk [3] and its successor EnhancedDesk [4], which performed physical paper augmentation offering interaction via touch.

Since then, numerous approaches toward physical paper augmentation emerged, setting the frontier for the interactive paper era. Most of these approaches focused on paper's affordances, such as its light weight and the capability of annotating content , but were also based on the fact that

 ISBN: 978-1-61208-421-3

paper constitutes a fundamental means of information dissemination.

In 1999, Mackay and Favard in [5] introduced the term “interactive paper”, signifying the potential role of digitally augmented paper for the forthcoming technologies. One such example is the Anoto system [6], which combines a unique pattern printed on each page with a digital pen to capture strokes made on paper. PapierCraft [7], on the other hand, is a system which uses pen gestures on paper to support active reading and allows users to carry out a number of actions, such as copying and pasting information from one document to another.

Other approaches to digitally augmented paper use touch and gestures as basic interaction technique. For example, the Pacer system [8] provides gestures support and touch based interaction with printed paper through the touch screen of a cameraphone. Pointing and writing in augmented reality environments has also been studied, but the majority of research work is based on proprietary technological artefacts e.g., light pens, pen with pads, and haptic devices [9] [10].

In terms of visualizing physical paper augmentation, a diversity of different approaches have emerged. For example, MagicBook [11] provides augmentation of physical books with 3D graphics and moving avatars through VR glasses, giving to the reader the sense of living pages. Pacer [8] supports printed paper augmentation via smartphones’ camera, acquiring images of the physical paper in real time and displaying them augmented with digital content on a smartphone’s display. Korozi et al. in [12] present two educational mini-games that offer physical interaction on a tabletop setup through printed cards, where a simple webcam monitors the table's surface and identifies the thrown cards, while the digital content is displayed on a nearby screen. In [13], an interactive desk that augments physical papers placed upon its surface with multimedia content and interactive applications is discussed. This system augments physical paper by projecting the digital content either on the paper or laterally to it.

Although a large number of approaches consider physical paper or printed matter as a fundamental means of interaction with technological artefacts in everyday life, there is still a lack of holistic approaches placing digital paper in the context of Ambient Intelligence in terms of technology embedment, context awareness, personalization, adaptiveness and anticipation. In this paper, an integrated solution for the use of printed matter and physical paper in Ambient Intelligence environments is proposed.

III. THE FRAMEWORK

According to Cook et al. [14], any smart environment can be adequately decomposed in four fundamental layers: physical, communication, information and decision. Each layer performs a different role in the environment, facilitating diverse operations and addressing specific requirements.

The discussed framework has been designed in order to facilitate the development of smart systems that use printed matter or physical paper as a main means of interaction in Ambient Intelligence environments.

Figure 1 illustrates the overall architecture of the proposed framework. Beginning bottom-up, the physical layer of the system comprises the hardware accompanied by the necessary software for printed matter recognition and tracking, as well as the supported user interaction techniques. Since the physical layer can consist of heterogeneous and alternative printed matter recognition systems based on different approaches (e.g., computer vision, electronic markers, etc.), the *Printed matter modeller* provides the digital “alter ego” of the physical subject. Furthermore, the *Annotation tool* provides intuitive UIs for printed matter modelling, through which the developers can make available to the system digital information corresponding to the

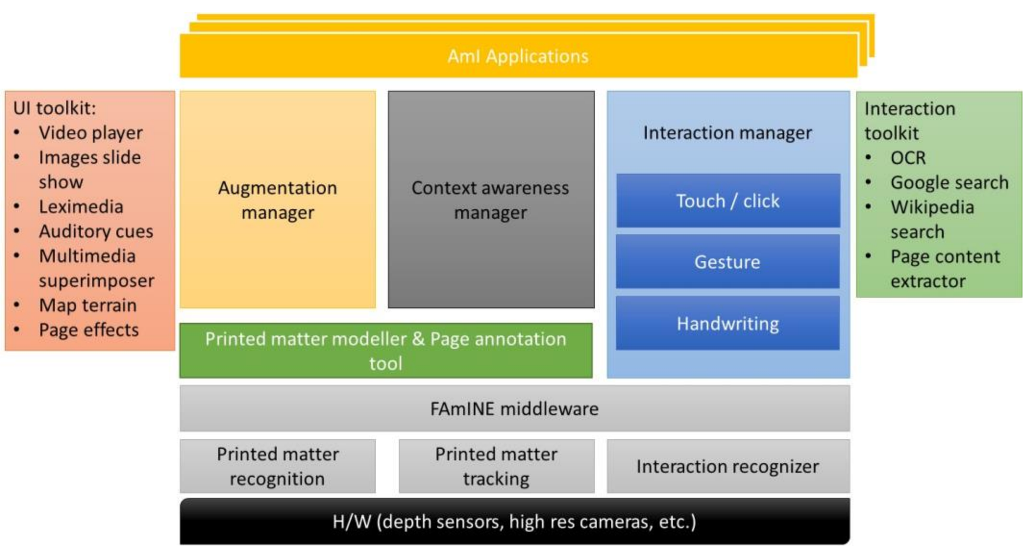

Figure 1. Framework’s architecture

Copyright (c) IARIA, 2015. ISBN: 978-1-61208-421-3

printed matter (e.g., multimedia content, references to internet sources, etc.)

The communication layer is realized by *FAmINE*, a middleware software developed in the context of the FORTH-ICS's AmI Programme and Smart Environments, providing the necessary seamless interoperability of the devices and services that comprise the AmI ecosystem.

The decision layer is implemented by the *Context Awareness Manager*, a fundamental component of the proposed framework, which is responsible for the selection of the appropriate UI components and corresponding content, according to the user's interaction in a specific context of use. The type of interaction is provided by the *Interaction manager*, which undertakes the task of interpreting users' interactions with printed matter. In order for the *Interaction manager* to render the supported types of interaction and for the *Context Manager* to extrapolate the necessary information from the corresponding printed matter, an *Interaction Toolkit* has been implemented including a number of external processes, such as Optical Character Recognition (OCR), information harvesting from various internet sources (e.g., Google search, Wikipedia), and a page content extractor (e.g., extracts text or images from the open pages, etc.)

The information layer consists of the *Augmentation manager*, which is responsible for the rendering of the available UIs provided by the framework for printed matter digital augmentation, adapted to the users' preferences and needs. Moreover, a number of fundamental UI components for printed matter augmentation has been developed and included in the *UI Toolkit*.

Each of the abovementioned components is *printed matter centric*, meaning that they address interaction and augmentation requirements for using printed matter. Furthermore, these components implement the necessary functionality for realizing the fundamental properties of AmI environments. For example, the *Context awareness manager* provides a user / context modelling scheme and ontology-based reasoning enabling personalization, context awareness and anticipation of users' needs. On the other hand, the *Augmentation manager* facilitates adaptivity mechanisms, offering alternative UIs according to the devices where an application is deployed, the profile and preferences of potential users, as well as the type of augmentation supported by an application.

IV. PRINTED MATTER MODEL

In order for the framework to keep structured information about the digital instance of printed matter, the *Printed matter modeller* has been implemented.

This component provides a classification of printed matter in an extended version of the XML description discussed in [24], including the digital representation (e.g. high resolution image) of the printed matter and interactive areas (hotspots) accompanied with their properties, stored in a recognition database. Every single matter in the recognition database is referenced by a unique id and is accompanied by its digital representation path. This digital representation is necessary for printed matter recognition by the framework, but it can also be displayed on any interactive screen near the physical paper or directly on it, using a video projector, enabling therefore the user to interact with hotspots that may be provided.

Every interactive hotspot is declared by a set of coordination points (normalized in order to be independent of the printed matter size), representing the hotspot's bounding path and a set of metadata information regarding the actual content of the hotspot such as the type (e.g.,

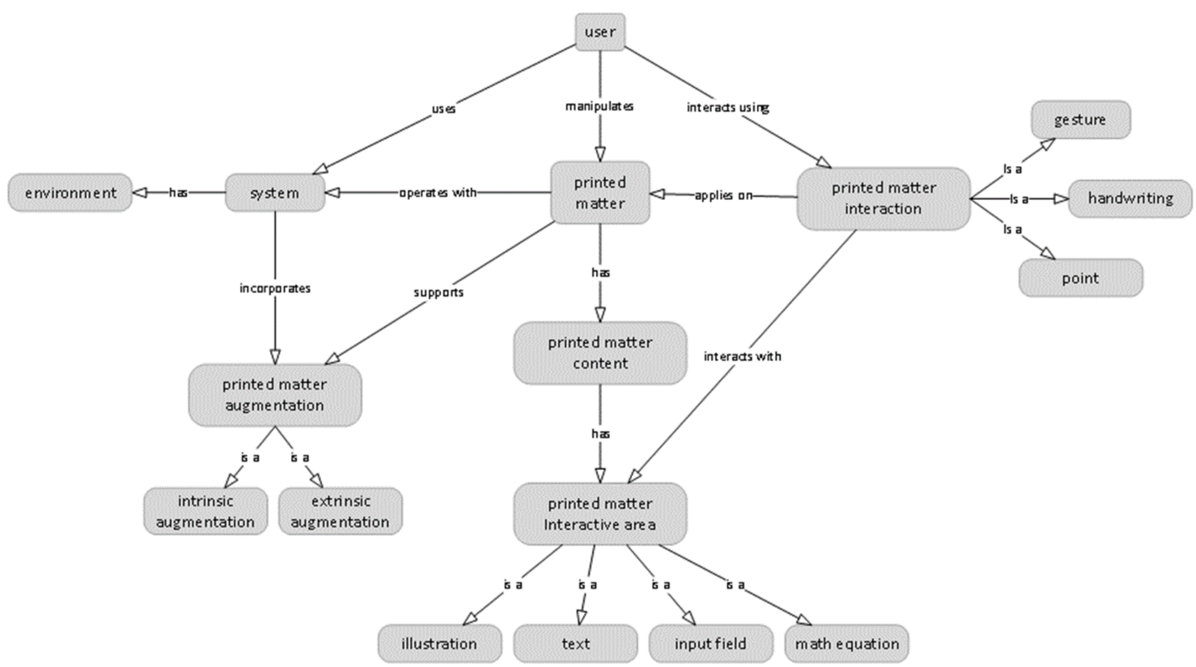

Figure 2. The framework's ontology

image, textual information, input placeholder), a description and a number of keywords.

The aforementioned information can be edited by the developer or the users themselves through the *Page annotation tool.*

The printed matter model constitutes the basic input for the *Context awareness manager* to provide appropriate content to the *Augmentation manager* and to consume interaction events fed by the *Interaction manager.*

V. CONTEXT AWARENESS

As already discussed in the previous section, the proposed framework provides the necessary tools and a software infrastructure for developers to easily integrate paper-based systems in AmI environments. A basic concern in this respect is the support of user modelling and context-awareness. To this end, a new ontology scheme has been designed, illustrated in Figure 2.

For the genericity of the framework's approach, physical paper is considered a subclass of printed matter. It should be noted that white paper doesn't provide any content and therefore it can be incorporated as soon as anything is written or printed on it. Physical paper can be considered as printed matter as soon as something is written or printed on it.

The basic entities of the aforementioned ontology are:

- **user:** The user entity of the proposed ontology captures users' needs in an AmI environment. It can be extended by user attributes or entities, such as users' profile, or their role and potential tasks that they may perform in the context of an environment.
- **printed matter:** It constitutes the generic class that describes the common properties of physical paper or any similar matter that contains handwritten or printed information.
- **printed matter interaction:** It generalizes the potential ways of interaction of the users with printed matter. Children of this class are: (a) gestural interaction (e.g., cycling a paper region, making a pinch gesture), (b) handwriting using a stylus object and (c) pointing / clicking either using fingers or stylus objects.
- **system:** The generic class of AmI stystems, ranging from single smart artefacts to sophisticated platforms (i.e., sophisticated systems comprising multiple sensors, running various services and running on heterogeneous hardware).
- **environment:** The generic class that defines environmental properties. Directly correlated classes are the location of the systems, the time that users' actions are performed and the environmental conditions in terms of temperature, pressure, humidity, lighting and noise.
- **printed matter augmentation:** It is divided in two subclasses (a) intrinsic augmentation that includes all augmentation techniques that apply on the printed matter itself and (b) extrinsic augmentation that regards augmentation techniques, which apply laterally or at a short distance from the printed matter.
- **printed matter content:** The generic class referring to handwritten or printed information on the printed matter. This class can be the generalization of the *document* entity and its subclasses are as defined in [15].
- **printed matter interactive area:** Refers to the types of interactive areas that can be found on printed matter: (a) *illustration* including any type of illustrated picture of figure (e.g., images, graphs), (b) *text* regarding any handwritten or printed textual information, and (c) *input field* referring to any type of printed placeholders that need users' input (e.g., text fields, checkboxes, etc.) and (d) *math equations.*

The proposed ontology can be easily extended using existing ones, for example [16], providing thus an open ontology scheme that can be used for context awareness in AmI environments.

The implementation of the proposed ontology makes use of the Web Ontology Language (OWL) [17]. In conjunction with a reasoning engine (e.g. Apache Jena [18], Jess [19], Microsoft Workflow Foundation Rules Engine [20]), the ontology constitutes the basic component of the *Context manager* module that is responsible for providing context awareness to the applications using the framework.

VI. CONTENT VISUALIZATION AND USER INTERACTION

As already mentioned in Section III, two modules have been implemented for content provision and interaction management, the *Augmentation manager* and the *Interaction manager.*

The *Augmentation manager* handles the visualization of the content in terms of appropriate UIs selection. The key factors that are mainly considered for the visualization of the available content are: (i) whether the provided content will be displayed on the surface of the printed matter or near it, (ii) the visualization output properties (e.g., a projection juxtapose to the printed matter, a nearby display, absence of display), (iii) users' preferences and (iv) the environment (e.g., whether the visualization applies in a noisy or silent environment).

The selection of the appropriate UI is made from a set of basic UIs that have been especially designed in order to address the needs of visualization of appropriate content in the context of printed matter manipulation in AmI environments. The UI components that have been implemented are able to visualize and provide interaction with heterogeneous content, including images, videos, textual information, geospatial information, auditory cues, animations and effects on the physical printed matter or digital representation of it, as well as any combination of these components.

On the other hand, the *Interaction manager* processes the input of several interaction techniques that the framework provides, such as touch, gestures and handwriting on printed matter or on a provided UI. According to the type of interaction, different types of information are provided. For example, if a user points at a hotspot area of a printed matter, then the *Interaction manager* will provide only the corresponding metadata information to the *Context awareness manager* for further processing. On the other hand, if a user makes a rectangular gesture denoting that the designated area should be isolated for further processing

 ISBN: 978-1-61208-421-3

(e.g., annotation by the user), the *Interaction manager* extracts this area in the form of a digital image and also acquires possible referenced information about this area such as the text that it may contain (extracted either from the description of the printed matter model or directly using Online Character Recognition). This information is also provided to the *Context manager* for further elaboration.

Along with the *Interaction manager*, a toolkit is provided containing tools for data extraction from the printed matter and semantic information search from online web services.

VII. APPLICATIONS

For the assessment of the proposed framework four example applications have been developed. These applications were evaluated in terms of usability and user experience according to standard evaluation procedures [17] [22] [23].

The first application (SESIL) [24] is an educational system that incorporates the proposed framework in order to provide stylus-based interaction in different spatial arrangements, such as large interactive surfaces featuring a display with multi-touch capabilities (i.e., for use in a library or at an exposition) using cameras for stylus recognition and tracking.

The system aims at enhancing reading and writing activities on physical books through unobtrusive monitoring of users' gestures and handwriting, as well as the display of information related to the current users' focus of attention. Additionally, it exploits the *Context manager* to decide at run-time the type of additional information and support to be provided in a context-dependent fashion.

The system consists of a desk with a set of three high resolution cameras placed above it to achieve the recognition and tracking of the school books placed on the desk's surface. A nearby large display runs an educational application that provides content-sensitive information to the users, based on their stylus-based interaction with a school book, following the extrinsic type of printed matter augmentation.

The second application (Book of Ellie) [25] is the augmented version of a classic schoolbook for teaching the Greek alphabet to primary school children. The book introduces alphabet letters and their combinations by increasing the difficulty level. For each letter or letter combination, relevant images and text involving the specific letter(s) are provided. The short stories for each letter are structured around dialogues and activities of a typical Greek family, with the protagonist being Ellie, one of the four children. In the augmented version of the book, Ellie has become an animated character, constantly available to assist the young learner by reading phrases from the book, asking questions or providing advice.

In terms of setup, the system consists of a television screen (32'') for visual and audio output, an "Asus Xtion Pro" RGBD camera, and a PC running the software. The RGBD camera is used to recognize and localize book pages and cards, as well as to detect and localize fingertip contacts on the book and table. The physical book and paper cards (e.g., depicting letters, simple objects, or animals) are interactive components of the system.

The third application is an Augmented Reality (AR) study desk [13] [24], which aims at augmenting physical books with digital information. The system consists of a standard definition projector and an ASUS Xtion Pro, both overlooking the surface of a desk. The images acquired by the color camera of the Xtion are used for printed matter recognition and its localization on the desk surface, while the images acquired by the Xtion's depth camera are used for detecting users' finger touch on the printed matter or the desk.

The AR study desk provides context-aware multimedia and interactive applications related to the content of the open book page. Such content is dynamically displayed to enrich the contents of the currently open book page, and is aligned, in real-time, with its 2D orientation upon the desk.

Technically, augmentation is supported by the projector-camera calibration. Given the coordinates of the book or the stylus in the desk coordinate frame, this calibration is used to predict the coordinates of the projector pixel that will illuminate the corresponding region or point of interest.

The last application (Study Buddy) [26] provides an unobtrusive intelligent environment that implements a context aware system targeted to augment the learning process. The system is composed of a smart reading lamp and educational software, called LexiMedia, aiming to provide dictionary information, as well as multimedia information for specific words, thus assisting in language learning.

The smart reading lamp incorporates a small camera and an embedded computer with WiFi connection. The camera of the reading lamp targets to the student's reading area (i.e., the area of the desk where the book is placed). Interaction with the system is initiated when a user indicates a word in the book, by using a black pointer (e.g., pen) and carrying out one of the following gestures: pointing at the word, underlining the word or circling the word.

Whenever the smart reading lamp observes that the reader needs help about a word or a phrase, it scans the area trying to recognize the indicated words, using OCR software. Then, it collects useful information about the recognized words, such as related images and words' definition. Finally, it transmits the aforementioned information to device (e.g., tablet, smart phone, etc.) which runs LexiMedia, placed near the reader .

TABLE I summarizes the interaction modalities and augmentation features that the aforementioned applications provide using the components of the discussed framework.

TABLE I. SUMMARY OF INTERACTION MODALITIES AND DIGITAL AUGMENTATION PROVIDED BY THE APPLICATIONS

Application	Interaction modalities	Augmentation
SESIL	• Handwriting • Stylus gestures • Page flipping • Recognition and localization of books on desk	Provides context-sensitive assistive content to the students on a nearby display.

Application	Interaction modalities	Augmentation
Book of Ellie	• Hand gestures • Page flipping • Printed cards position and orientation recognition • Recognition and localization of books on desk	Provides auditory cues to the students on a nearby device. Provides card-based educational games of varying difficulty according to the students' profile.
Study desk	• Hand gestures • Page flipping • Recognition and localization of books on desk	Provides context-sensitive content on top of or latteraly to the open page.
Study buddy	• Handwriting • Stylus gestures • Page flipping • Recognition and localization of books on desk	Provides thesaurus information and explanatory multimedia content for any english word that can be found on the open page.

VIII. Conclusions

Taking into account the potential of physical paper to become a widely accepted interactive component of an Ambient Intelligence environment, an extensible context-aware interaction framework has been proposed, which enables the integration of printed matter into AmI environments. The proposed framework allows application developers to enrich printed matter with digital information according to their application's envisioned context of use, supports natural multimodal interaction both with physical paper and the digital content resulting from the paper augmentation, and features multimedia output.

Future work will address the issue of multimodal interaction with printed matter in combination with other physical objects coexisting in the environment. Furthermore, the generalization of the proposed framework will be investigated towards incorporating other physical objects. Another aspect to be addressed is the investigation of automated recognition and modelling of printed matter based on its content web semantic analysis. Finally, future efforts will include the evaluation of the framework by software developers of AmI applications.

Aknowledgment

This work was supported by the Foundation for Research and Technology – Hellas, Institute of Computer Science (FORTH – ICS) internal RTD Programme 'Ambient Intelligence and Smart Environments'.

References

[1] IST Advisory Group, Scenarios for Ambient Intelligence in 2010. [Online]. Available from: ftp://ftp.cordis.lu/pub/ist/docs/istagscenarios2010.pdf [retrieved: June, 2015].

[2] E. H. Aarts and S. Marzano, (Eds.). "The new everyday: Views on ambient intelligence", 010 Publishers, 2003.

[3] P. Wellner, "Interacting with paper on the DigitalDesk", Commun. ACM, vol. 36(7), pp. 87–96, 1993.

[4] M. Kobayashi and H. Koike "EnhancedDesk: Integrating paper documents and digital documents", Asia Pacific Computer Human Interaction (APCHI'98), July 1998, pp. 57–62.

[5] W. E. Mackay and A. L. Fayard, "Designing interactive paper: lessons from three augmented reality projects", Proc. International workshop on Augmented reality: placing artificial objects in real scenes, AK Peters, Ltd., 1999, pp. 81-90.

[6] Anoto, Development Guide for Service Enabled by Anoto Functionality, 2002.

[7] C. Liao, F. Guimbreti, K. Hinckley, and J. Hollan, "Papiercraft: A gesture-based command system for interactive paper", ACM Transactions in Computer-Human Interactions, vol. 14 (4), pp. 18 – 55, 2008.

[8] C. Liao, Q. Liu, B. Liew, and L. Wilcox, "Pacer: fine-grained interactive paper via camera-touch hybrid gestures on a cell phone", Proc. SIGCHI Conference on Human Factors in Computing Systems, ACM, 2010, pp. 2441-2450.

[9] A. S. Forsberg, Jr. J. J. LaViola, and R. C. Zeleznik, "ErgoDesk: A Framework for Two- and Three-Dimensional Interaction at the ActiveDesk", Second International Immersive Projection Technology Workshop, May 998, pp. 11-12.

[10] D. Schmalstieg, et al., "The studierstube augmented reality project", Presence: Teleoper. Virtual Environ., vol. 11(1), pp. 33-54, 2002.

[11] M. Billinghurst, H. Kato, and I. Poupyrev, "The MagicBook: Moving Seamlessly between Reality and Virtuality", IEEE Computer Graphics and Applications, vol. 21(3), pp. 6-8, 2001.

[12] M. Korozi, et al., "Ambient educational mini-games", Proc. International Working Conference on Advanced Visual Interfaces, ACM, 2012, pp. 802-803.

[13] G. Margetis, A. Ntelidakis, X. Zabulis, S. Ntoa, P. Koutlemanis, and C. Stephanidis, "Augmenting physical books towards education enhancement", 1st IEEE Workshop on User-Centered Computer Vision (UCCV 2013), IEEE, Jan. 2013, pp. 43-49.

[14] D. J. Cook and S. K. Das, "How smart are our environments? An updated look at the state of the art", Pervasive and Mobile Computing, vol. 3(2), pp. 53–73, 2007.

[15] Bibliographic Ontology Specification. [Online]. Available from: http://bibliontology.com/ [retrieved: June, 2015]

[16] D. Preuveneers, et al., "Towards an extensible context ontology for ambient intelligence", Proc. Second European Symposium Ambient intelligence (EUSAI 2004), Springer Berlin Heidelberg, Nov. 2004, pp. 148-159.

[17] OWL Web Ontology Language Overview. [Online]. Available from: http://www.w3.org/TR/2004/REC-owl-features-20040210/#s1.1 [retrieved: June, 2015]

[18] Jena documentation overview. [Online]. Available from: https://jena.apache.org/documentation/index.html [retrieved: June, 2015]

[19] Jess, the Rule Engine for the Java Platform. [Online]. Available from: http://www.jessrules.com [retrieved: June, 2015]

[20] Introduction to the Windows Workflow Foundation Rules Engine. [Online]. Available from: https://msdn.microsoft.com/en-us/library/aa480193.aspx [retrieved: June, 2015]

[21] J. Nielsen, "Usability engineering", Morgan Kaufmann, pp. 165-206, 1994.

[22] J. Nielsen, "Usability engineering", Morgan Kaufmann, pp. 207-226, 1994.

[23] C. Lewis, “Using the ‘thinking-aloud’ method in cognitive interface design”, IBM Research Report RC 9265, Yorktown Heights, NY, 1982.

[24] G. Margetis, X. Zabulis, P. Koutlemanis, M. Antona, and C. Stephanidis, “Augmented interaction with physical books in an Ambient Intelligence learning environment”, Multimedia Tools and Applications, vol 67(2), pp. 473-495, Jan. 2012.

[25] E. Papadaki, et al., “The Book of Ellie: An Interactive Book for Teaching the Alphabet to Children”, International Conference on Multimedia and Expo Workshops (ICMEW), IEEE, July 2013, pp. 1-6.

[26] G. Margetis, S. Ntoa, M. Bouhli, and C. Stephanidis, “Study-Buddy: Improving the Learning Process through Technology-Augmented Studying Environments”, HCI International 2011–Posters’ Extended Abstracts, Springer Berlin Heidelberg, 2011, pp. 504-508.