

An Integrated Environment for Scientific Data Entry and Management on mobile systems

M. Ancona, G. Dodero, V. Gianuzzi, S. Locati, A. Romagnoli

*DISI, Università di Genova
Via Dodecaneso, 35
16146 Genova, Italy*

Contact Person:

Gabriella Dodero

Email: dodero@disi.unige.it

Fax: + 39 010 353 6699

Abstract

In this paper, an integrated approach to the problem of entering and manipulating scientific data on mobile devices connected in a wireless network is presented, and its ability to support mobile users operating on the field in indoor and outdoor environments is briefly discussed. This approach is based on experiences collected by the authors in the development of three mobile applications, for cultural heritage and healthcare, with mobile computers connected via a wireless network.

Keywords: Mobile Computing, Multi Modal Input, Man-Machine Interfacing, Interfacing Components.

1. Introduction

The use of mobile devices in Wireless LAN mobile computing opens new applications but also brings new problems due to their limitations: this paper shall address user interfacing issues, especially detailing input facilities. The need of being operational without assuming a user sitting at a desk has important consequences: writing and drawing on a mobile device is not easy, and this may limit the success of mobile applications. Thus, for small devices like PDAs and Handheld PCs, differences with desktop human computer interaction are due to limited display and memory capabilities, and lack of mouse and keyboard.

In fact, desktops exhibit a never-ending need for larger screens, where multiple overlapping windows, icons collection, scroll bars can be comfortably seen and selected by clicking. The display size of mobile devices is typically 6", hence leading to insufficient space for multiple overlapping windows and other wizards; even the thumb, or elevator, in a scroll bar can be very difficult to manipulate. The recognition of these difficulties has encouraged manufacturers to equip mobile devices with other input tools, such as pointers and microphones, which are not so common on desktops. Hand-held PCs and PDAs may however show a small virtual keyboard on the screen (thus further reducing the available display area for output), but its use may be frustrating and scarcely productive when standing up.

In this paper we present our experiences in three projects, that is, two outdoor applications in cultural heritage, and an indoor healthcare environment: in these applications, we need to mix text and images not only as outputs but also for input. In the latter case, such input shall be entered by means of different input technologies, in accordance both with the working environment and with personal preferences. These

needs do not arise with traditional software for desktop computers. Thus, as the working paradigm is shifting from popular desktop to mobile computer layouts, processing of structured, multi medial documents becomes more and more important [Ste00].

Based on our experiences, we propose an architectural model called *Integrated Environment for Scientific Data Entry and Management*, **IESDEM** for short. It is based on the conceptual separation of input (and output) operations from the application layer by means of an intermediate data representation layer, based on the document information model XML. This layer allows to introduce descriptive information, i.e. meta data, on the structure and content of the documents, useful at various levels: as common data exchange protocol between applications residing on the mobile device and on the fixed server, as internal representation of documents and related components, as content description for pages visualised using different layouts, in accordance with current screen size and properties.

1.1 Classification of mobile devices

Mobile wireless computers, also called PDAs, Personal Digital Assistants, can be divided in three categories, depending on physical features and destination of use. *Handheld* computers have typically a half VGA screen, and up to 64-Mb memory, with optional disk memory. Since they often include a mini keyboard, they can be easily used as “portable office” and provide complex applications. *Palm sized* computers have a smaller, touch sensitive screen (a quarter of VGA) and less memory (up to 32 MB). A pen or a speech recognizer performs input. Their use is mainly as personal organizer, Internet Navigator and portable terminal for data entry, and information supplier in industrial applications. Finally, *smartphones* are cellular phones providing a few Mb of memory and a little screen (an exception is the Ericsson R380, with a 320x200-pixel screen). They are used as personal agenda and Internet Navigator.

Our applications mainly concern palmtops and handheld computers without keyboards, and in the following, we will refer to them generically as mobile devices or PDAs, even if this term applies to a wider set of computers. Operating systems are Palm Os, WindowsCE / PocketPC and Epoc. The first one is the most diffused for palmtop applications, especially designed to be operational also with a small memory. The other two systems are not so widespread yet, but they are more powerful, allowing for example to develop applications in different programming languages and to access an Oracle database directly from the PDA. The mobile devices we have been working with are HP 820, Fujitsu PenCentra 130, Compaq Aero 2000 and Casio E-105 Cassiopeia. The new Compaq Ipaq device is the next candidate for prototype developments.

2. Three mobile computing projects: RAMSES, WardInHand and PAST

As said above, the authors in the recent past have designed the architecture of three different mobile computing systems: RAMSES, WardInHand and PAST. Their architecture is basically the same: a collection of pen-based PDAs, connected among them and to a fixed host-server using a radio-based network. Due to the kind of applications we chose to use unlicensed radio frequency bands, which are well open to multi-vendor solutions, compatible with IEEE802.11 standard. Two available technologies, Frequency Hopping Spread Spectrum (FHSS) and Direct Sequence Spread Spectrum (DSSS), have been considered, and both have been experienced. Other communication technologies, like digital (TDMA/CDMA or GSM) telephone networks, or digital packed data (CDPD), seemed not convenient at the moment, mainly for operation costs and security implications.

The fixed host-server supports interfacing to mobile clients, records documents, provides access to/from legacy systems, and interfaces internet users, as shown in Figure 1.

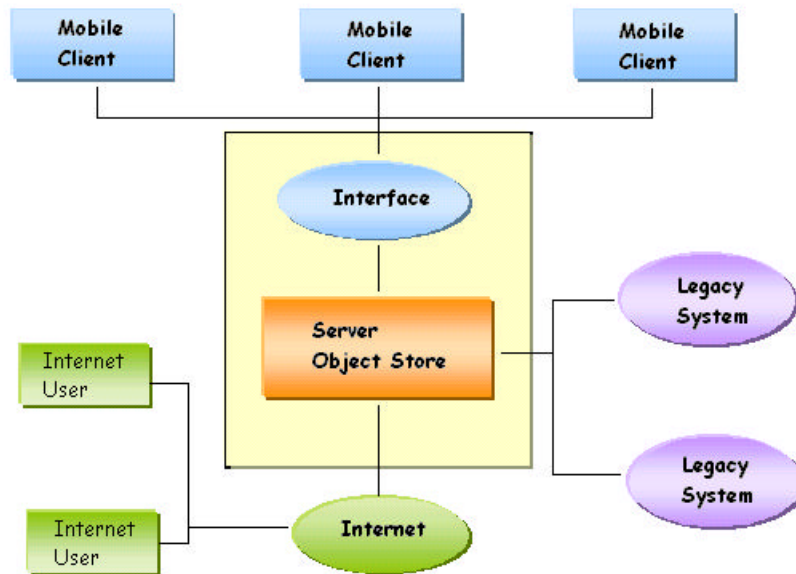


Figure 1: System architecture

Let us consider more in depth the user requirements for these projects.

2.1 RAMSES: Remote Archaeological Mobile Support Enhanced System

It is a 5-years research project (started in 1997) where the authors cooperate with archaeologists of I.I.A.S (Istituto Italiano di Archeologia Sperimentale). It is granted by the Italian National Research Council, Cultural Heritage Project, and it aims at making archaeological on-site research easier [Anc99].

Since excavation sites are usually non-connected to computer centers, if during an excavation campaign a database or a colleague has to be consulted for additional information, someone has to physically travel to or from the site. To collect relevant daily work information, archaeologists usually bring a (paper) diary in the field, where they take daily notes and sketch collected finds. These notes are later on copied to standard forms, and stored in databases.

The architecture of Figure 1 is straightforwardly implemented in the excavation field, where the server, hosted in some nearby building, allows communication with remote scientists by an Internet connection (by telephone cables or satellite radio links), acts as object store, and avoids the need of rewriting and reorganizing the notes and sketches.

The mobile input interface is composed by three tools: one for entering free text, called WordTree, one for drawing sketches, and an application specific tool (called Archeo) that integrates the functions of drawing with that of data entry in order to compose objects to be stored in the electronic archive (A.D.E) on the server, where the complete classification of finds is kept.

2.2 WardInHand: Mobile Workflow support and informAtion distribution in hospitals via voice-opeRateD, wIreless-Networed HANdheld PCs.

It is an EU funded project (1999 IST 10479) started at the beginning of 2000, with three pilot Hospital Units and four IT partners, including the Department of the authors. The first prototype is under development and shall be operational in March 2001.

Using WardInHand technology, doctors and nurses will be equipped with individual PDAs linked via wireless networks to a server, where all clinical records and information on patients is collected. WardInHand allows to completely avoid traditional paper-based management of clinical records, and to have a bedside access to the Electronic Healthcare Patient Records.

The project exploits small and cheap PDAs, with human-computer interfaces leveraging upon graphics, pen-based and voice-based interaction. The text entry tool developed for RAMSES will be included, and available speech recognition tools are under test. The main problem to be solved is to allow a natural user interaction with a smooth transition between input tools, that is, mainly from text to voice input and vice versa, in order to match environmental constraints (noisy areas) or individual preferences (hands-free input) [Anc00].

2.3 PAST: exPeriencing Archaeology across Space and Time

It is an EU funded Project (1999 IST 20805) just started (third quarter 2000). The PAST Consortium consists of three Pilot Users, each one with an archaeological site and an archaeological museum, an archaeology University department, and four IT partners (including the Department of the authors) [Anc00a].

The goal of the project is to facilitate archaeological site visits: a tourist in an archaeological site will carry a PDA, connected via wireless network to a PAST Server. This server shall collect a few data from the tourist, in order to profile him/her and to organize a personalized plan for the visit. The current position of the visitor is known at any time, in order to guide him/her across the site, pointing out specific items and delivering to the PDA context-specific, relevant information. For spatial positioning different technologies are under consideration: GPS, triangulation by means of network access points references, comparison between images taken by the PDA's webcam and known profiles.

3. The Integrated Environment for Scientific Data Entry and Management

One of the most relevant requirements of PDA users is to select, among multiple input methods for a given operation, the best method with respect to external conditions and personal preference at any moment. Unfortunately, almost all PDA applications define their own input systems and their own facilities for storing and managing it; they typically do not offer facilities for data sharing or interchanging, such as those found on desktops.

In the prototype design and development for WardInHand, we followed a new approach for multi-modal data entry, which we found successful and which we plan to use in future projects like PAST. This approach is called *Integrated Environment for Scientific Data Entry and Management*, **IESDEM** for short. The need for such an environment arises from text entry on keyboard-less touch-screens using

different methodologies, including speech recognition; but it is extensible to support other data entry mechanisms. For example, in case of acquisition of images or drawings, IESDEM supports object-based hierarchical drawing; if localization information are required, it may collect them from different devices and so on.

Key features are:

- Input modules are stored in the mobile devices at the lowest (Operating System) layer, in order to be available to any application, without need of rewriting them for reuse. These input modules might perform simple filters or "standard" conversions, like for example spell checking for text; geographical coordinate transformations for GPS; or contour extraction for images.
- Separation of the data input layer from the data representation layer and the application layer.
- Use of open tools both for visualization (GUI) and for data transfer to/from the server.

The resulting architectural model is shown in Figure 4.

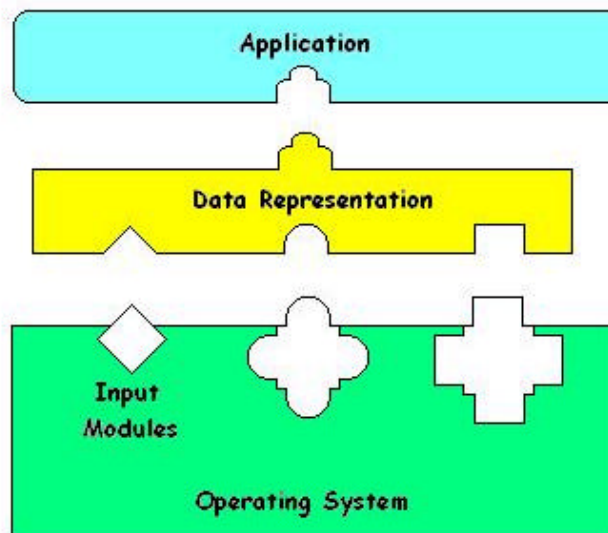


Figure 4: Relationships among the three layers

3.1 Input Data Layer

This layer consists on a set of components, where each component is the smallest independent and self-managing part of a large system spanning over several environments. Usually, components are distributed objects: for example, Corba components or ActiveX controls. Components support rapid application development and a high degree of customizability and produce robust applications that are relatively inexpensive to develop and maintain. Components are characterized by their interface since they have to be used in a variety of systems and applications, by considering their internal structure as a black box.

IESDEM inserts input components in the underlying operating system (Windows, PocketPC or Epos) in order to be available to all applications, where actual input technique selection is performed. In our case WTx has been implemented as Pocket PC (Windows CE) component that can write data into other

applications: the receiving application must conform to the data format produced by it, and must provide a compatible support where data is stored. A speech recognition component may be interfaced much the same way, without need for application modifications.

This interchangeability is quite similar to the use of plug-ins to augment browser facilities. The novelty in this approach lies in providing this kind of features at the operating system layer, rather than within an application as in a browser; and to apply these new features for inputs (i.e. to create multimedia data) rather than for outputs only, as it usually happens for browsers.

3.2 Data Representation Layer

The unified data representation layer allows the management of meta data documents: let us show the use of this decomposition with some examples:

- In RAMSES, graphics and text describing a find are merged in the same document, together with position data of the find inside the grid, and with additional information (meta-data) about its interpretation, like time, stratigraphic unit, links to related objects and so on.
- In PAST, information presented to visitors can be thought of as electronic notes stucked on interesting places: the mobile application sends to the server information on current positioning; the server identifies documents relevant for that position, and forwards them to the visitor.
- In WardInHand, the PDA shall display (portions of) the electronic patient record, where data derived from a plethora of different medical equipment has to be integrated and stored, together with doctors' notes, vital sign measures taken by nurses and so on.

Standards are an important means to achieve common representation schemes and systems interoperability, hence they can play a pivotal role in exploiting meta data. Presently, XML appears the best candidate for this role, for its self-describing nature, and for its platform and vendor neutrality. Many standard DTDs and RDFs have already been proposed for each specific application, and, given the wide acceptance of such a methodology by many firms, a final success is foreseen. Whilst designed initially for display of documentation distributed via Web, XML has also been widely adopted in very different application fields, such as: to interchange information between different applications and computer architectures, to distribute processing load from Web server to Web client, and to improve data base search performed by intelligent Web agents on behalf of individual users.

In the European healthcare environment, the use of XML for messaging has been investigated as part the XML/EDI Project. The project has already delivered a draft set of XML document type definitions (DTDs) related to Request, Provide and Notification Messages (defined by CEN/TC 251 [TC251] group) and a draft set of best practice guidelines for the development and implementation of XML/EDI.

In the US, the HL7 (High Level 7) group is working to develop standards for the electronic interchange of clinical, financial and administrative information among independent health care oriented computer systems. The group is currently investigating the possibility of changing its messaging structure to an XML format (HL7 Version 3) to increase interoperability. HL7 [HL7] has developed the Patient Record Architecture, an XML-based clinical document architecture that provides an exchange model for documents of varying levels of complexity.

Similar efforts are just starting in the cultural heritage field. In both our applications, RAMSES and PAST, a data element is composed by a collection of simple data, like information that defines the bound of the

area covered, temporal information, spatial descriptions, such as the geometric appearance and the position of the data. The same data composition can be found in many applications, such as the FieldNote system described in the paper of J. Pascoe [Pas98], or in any pervasive computing applications.

These data may be recorded in different formats, so that it is difficult to combine them, or use the same representation layer. Also in this case, however, this limitation can be overcome using XML together with a dedicated text language to describe geographic features, that is GML (Geographical Markup Language), proposed by the Open GIS Consortium as "common platform for applications that relate to geographical metadata", and which is, or will quickly be, supported by many GIS vendors [GML]. An object represented in GML has a geometric textual representation and is associated with a Spatial Reference System describing the location where the document is defined.

The wide use of XML and GML will simplify the update and the extension of a document, as opposed to what happens in most legacy systems, where the flat representation of the document makes it difficult the integration with new different information.

4. Data Entry

In this Section we provide additional details on data entry tools we have experienced and on their implementation. On a keyboardless device there are two standard ways of entering a text: a virtual keyboard and an handwriting recognition tool. In both cases a part of the screen is devoted to display the keyboard or as graphical input area; applications automatically receive inputs from these areas.

In some cases an application may be designed in order to receive other inputs from menus or comboboxes, and keeps screen layout under control. This is however possible only if text has to be selected among existing options, not for free texts, and requires some programming effort in the application.

Space considerations do not allow to describe a similar approach, that is treatment of graphic input, be it a hand-made drawing or a digital snapshot; and spatial coordinate input, such as those collected by a GPS, a compass or some other device.

4.1 Text input with WTx

The virtual QWERTY keyboard approach is the most traditional text input method for a PDA, implemented by most software developers including Microsoft [WinCE]. However, a QWERTY keyboard has been designed for ten fingers almost parallel typing, while sitting with the keyboard placed on a table: the typing speed that can be attained on a PDA is no more than 3 keys per second, and often less [Mas99]. Another input tool is the T9 program of Tegic Communication [Kus98], based on the 9-keys telephone keyboard. It allows the efficient generation of any desired text by assigning multiple letters to each key. Comparing a sequence of keystrokes to words in a database solves ambiguities.

WordTree [Anc99a] starts from a different point of view: as suggested by its name, it is a sophisticated form of smart keyboard allowing users to select a word from a dictionary organized into a trie of words dynamically displayed on the screen. In designing WTx, the aim was not to allow *writing under dictation*, but to allow *writing while thinking*, meaning that the tool has an optimal compromise between easiness, comfort and speed.

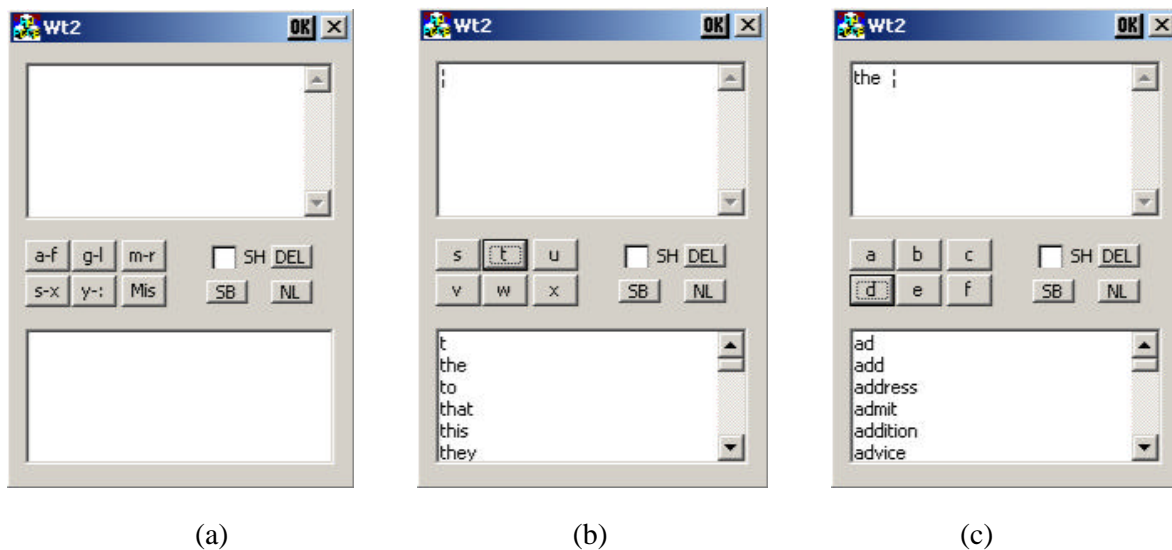


Figure 2: WTx Graphical Interface

The WTx GUI is divided into a typing area and a selection area. In Figure 2a, the typing area shows two groups of buttons: control buttons on the right (e.g. delete a character, insert a blank, etc), and a "multi-level virtual keyboard" on the left. Here, at the beginning, every button in the virtual keyboard is labeled with a group of characters. Tapping on it (e.g. on the "s-x" button) will expand the group over all the buttons of the virtual keyboard. When tapping on a button with a single label, the selected character is appended to the word under construction. This will branch the set of known words (dictionary), and the selection area will list words having the prefix typed so far as their root. This list is ordered by frequency of occurrence. As an example, after tapping on the button «t», the resulting display is shown in Figure 2b. The user can directly select the interested word («the» in this case) with a tap; then, to select the word «address», the user needs three taps, the group «a-f» and then two letters («a» and «d») because after the first one, the desired word was not available (Figure 2c).

WTx has been tested on a group of users, who needed some training period, and later on considered it worthwhile because pen movements are minimized, and words are ordered by frequency, thus resulting in very fast typing speed.

4.2. Speech recognition tools on a PDA

As opposed to the wide variety of tools which can be found on desktops for both voice commands and for dictation, only a few companies so far have developed software for PDAs. They include *Fonix*, whose FAAST (Fonix Accelerated Application Solutions Technology) software development framework contains Fonix Automatic Speech Recognition and Text-To-Speech technology for devices operating on Microsoft Windows CE. The Automatic Speech Recognition Engine is a compact neural net recognition system, which is user-independent and noise-tolerant.

ARTrecognition provides handheld software for both WindowsCE and PocketPC devices. It supports any language or accent thanks to ART's signal processing and pattern recognition technologies. An example of speech enabled wireless software is Pocket commander, a voice recognition application for Windows CE

based computers that allows to launch applications, open documents, send e-mails, run macros, browse Contacts database and more by simply speaking into the PDA's microphone.

In our experience, the performance of speech recognition tools on mobile devices in hostile environments, such as a Hospital ward, is not yet comparable with similar tools working on desktops; thus, it is necessary to wait for improvements for a really significant use, in both WardInHand and PAST. The IESDEM technology shall however be ready to integrate similar features when available.

4.3 Implementing a Server In-Process

Under Windows CE, the technology that allows IESDEM to be supported is called Server In-Process or just SIP. A SIP is a COM (Component Object Model) object with server functionalities, that is, it should define and make available some given functionalities to any process requesting them. It must be implemented as a DLL (Dynamic Link Library) which can be linked to the requesting process. The server contained in the SIP is then invoked by the shell, acting as a client of the input method implemented as the service.

In implementing IESDEM, we have several implementations of such an input method, that is the predefined ones like the virtual keyboard, as well as WTx or some speech recognition software, all integrated by one SIP. The underlying technology resides at the API level, that is, below the application interface which is supported by the MFC.

5. Conclusions

The impact that a user interface has on the success or failure on mobile computing is determinant. We identified a stringent need for facilities for entering text on PDA, but the whole interfacing of these devices to collect input data is lacking of unifying concepts. Looking at our experience in mobile, wireless applications, we pointed out that the ideal user input interface is fundamentally different from that used for desktops or workstations, which relies on mouse and keyboard. Pen and voice input support a more natural user interface, but in order to be really friendly, some considerations must be taken into account.

For text input we developed a tool based on direct selection of words from a trie structured dictionary: a walk on the trie made by few icons clicking is sufficient to select the required word. This method avoids typos and spelling errors and when it is used while standing up, it is superior, in speed, to all other methods, including virtual keyboards and hand-writing recognizers.

We have experienced a significant uncertainty inherent in input methodologies, like handwriting or speech recognition: in these cases the user must be given visual and audio cues to be confirmed, and command actions must be made reversible or recoverable.

In order to integrate and reuse the input tools that we developed for the various applications, we propose a three layers approach. The input layer may make use of the components technology, in order to make the input processing components available to all applications. The middle layer may easily integrate emerging standards in the various application fields, like XML-EDI or GML, based on meta-data representations in XML. Data sharing among applications is thus made straightforward to implement.

References

- [Anc99] M.Ancona, G.Dodero, V.Gianuzzi, "RAMSES: a mobile computing system for field archaeology", *Proc. HUC99 (Int. Symp. on Handheld and Ubiquitous Computing)*, Lecture Notes in Computer Science, 1707, Springer-Verlag, pp.222-233, 1999.
- [Anc99a] M.Ancona, D.Comes, "WordTree: a pen based editor of short texts", *Proc. IGS99*, June 1999, Singapore, pp. 263-268.
- [Anc00] M.Ancona, G.Dodero, V. Gianuzzi, F.Minuto, M.Guida, "Mobile computing in a hospital: the WARD-IN-HAND project", *Proc. ACM Symp. On Applied Computing 2000*, Como (Italy), March 2000.
- [Anc00a] M.Ancona, G.Dodero, V.Gianuzzi, O.Bocchini, A.Vezzoso, A.Traverso, E.Antonacci "Exploiting wireless networks for virtual archaeology: the PAST project", *Proc. VAST Conference*, Arezzo, November 2000.
- [GML] <http://gis.about.com/science/gis>
- [Hje00] J. Hjelm, M. Nilsson, "Position-dependent services using metadata profile matching" *Proc. INET 2000*
- [HL7] Health Level-7 Standards home page <http://www.hl7.org>
- [Kus98] C. Kushler, "AAC Using A Reduced Keyboard", *Proc. CSUN 98*, Los Angeles (CA), March 1998.
- [Mas99] T.Masui, "POBox: an efficient text input method for handheld and ubiquitous computing", *Proc. HUC99 (Int. Symp. on Handheld and Ubiquitous Computing)*, Lecture Notes in Computer Science, 1707, Springer-Verlag, pp.289-300, 1999.
- [Pas98] J. Pascoe, "Adding generic Contextual Capabilities to Wearable Computers", *Proc. 2nd Int. Symp. On Wearable Computers*, Oct.19-20, pp.92-131, 1998.
- [Ste00] C.Stephanidis, "From User Interfaces for All to an Information Society for All: recent achievements and future challenges", *Proc. 6th ERCIM Workshop "User Interfaces For All"*, Florence, Oct. 2000.
- [TC251] European Committee for Standardisation, Technical Committee for Health Informatics <http://www.centc251.org/>
- [WinCE] <http://www.microsoft.com/windowsce>