Nectar: Support for Semi-Structured On-Field Data Collection

Dhyanesh Narayanan
Department of Computer Science
Stanford University
Stanford CA 94305 USA
dhyanesh@.cs.stanford.edu

Ron B. Yeh
Department of Computer Science
Stanford University
Stanford CA 94305 USA
ronyeh@.cs.stanford.edu

Scott R. Klemmer
Department of Computer Science
Stanford University
Stanford CA 94305 USA
srk@.cs.stanford.edu

Abstract
This paper presents Nectar, an infrastructure to support field workers for effective collection of semi-structured data. The utility of Nectar is described in the specific context of collecting patient data from remote rural areas of developing countries such as India. Nectar is architected to leverage and complement the ButterflyNet infrastructure [7].

Keywords
Mobile Capture/Access, Augmented field-work support, Technology-enhanced Health Care

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous, J.3. Life and Medical Sciences: Medical information systems

Introduction
Indian hospitals organize "Village of the Month" programs in which a specific village is chosen for survey to collect indicative health parameters from villagers. The collected data has to be consolidated and acted upon, for gauging village health conditions and taking suitable measures. In addition, the data has to be maintained over time for future retrieval and analysis.

Need for Semi-Structured On-Field Data Collection Support

During health surveys, medical professionals typically use pen and paper as the primary tools to gather patient data. This imposes significant cognitive overhead in asking a set of questions to each villager and making note of the responses. The medical professional also has to ensure that the villager (who is mostly illiterate) understands the questions clearly. In addition, since the villager would likely get annoyed if asked to repeat some piece of information, or may not repeat it with the same level of accuracy as the first time s/he said it, the health official has to be quick in jotting down responses. To achieve this, the official may resort to using some (cryptic) short forms, which when looked at later may not be easily understood.

Some of the notes made by medical professionals may be lost before they are brought back to the hospital for entry into the Patient Information Systems (PIS). It is more often the case that notes get damaged because they are accidentally dropped on muddy ground, get drenched in rain etc (typical of village settings).

If multiple forms of data need to be obtained, such as pictures and/or audio, the medical professional has to switch between devices, for each form of data for each patient. In addition, a mapping has to be maintained between the responses (paper notes), pictures, and audio that belong to a specific patient. Despite this, it may be difficult to later correlate the different pieces of data about a specific patient into a meaningful whole.

Thus there is a significant pain-point in terms of gathering patient data in a simple, fluid, & dependable manner. This creates a need for offering support to field workers engaged in semi-structured data collection activities such as rural patient data acquisition.

Nectar

Technology augmented support for field workers engaged in semi-structured data collection activities can be an effective means of improving productivity and accuracy. Nectar is an effort in this direction and is architected to specifically target the pain-points faced by medical professionals during rural patient data collection. Nectar also leverages and extends the ButterflyNet infrastructure [7].

Nectar entails the use of a Pocket PC to gather data during medical surveys in villages. Doctors use a simple interface on a handheld Pocket PC, to quickly fill out key details for each rural patient they screen. The data is stored locally on the Pocket PC and synchronized with the PIS in batch mode. The PIS data can later be mined for aggregated information on village health conditions and suitable health measures can be taken.

The motivation for Nectar stems from an earlier effort called Rural Data Capture System (RDC), undertaken by the first author. RDC was a minimal functional prototype built using requirements gathered from doctors at the IIT Madras Hospital (India). Nectar is the result of rethinking RDC design for complete leverage and full-fidelity integration with ButterflyNet [7].

Nectar Interface Requirements

To design a suitable interface for Nectar that effectively supports medical professionals on-field, we need to obtain requirements that specify the different pieces of patient data to be captured during a health survey. An indicative set of rural patient data that would be

desirable, as gathered from doctors at the IIT Madras Hospital (when building the RDC prototype), are presented below:

First Name, Last Name, Age, Sex, Blood Group, Current Medications if any, Food Habits, Bowel Habits, Bladder Habits, Other Habits (Smoking etc), Self-rating of Overall Health, Extent of Social Support, Occupation, Patient's Picture (if available)

Nectar Interface Design

A closer look at the requirements highlights that much of the interaction with Nectar is in terms of *data* rather than *control* – there are possibly just two control actions: one to commit the data to the Pocket PC local store and the other to synchronize the data with the Hospital. It is hence a better design choice to build a *data-centric* UI rather than a *task-centric* UI for Nectar. The following discussion is driven by this design choice.

To aid in convenient and fluid data entry, the patient data to be gathered is grouped into *semantic clusters* that localize to subsets of related data about a patient. This places minimal cognitive overhead on the medical professional when filling out details relevant to one semantic cluster, whilst enabling a simple direct jump to the next cluster when finished with the current cluster. Revisiting the list of indicative rural patient data, the following semantic clusters can be identified:

- [Cluster 1: Personal Information] {First Name, Last Name, Age, Sex, Blood Group, Occupation}
- [Cluster 2: Medical Information] {Self-rating of Overall Health, Current Medications if any, Extent of Social Support}
- [Cluster 3: Habit Information] {Food, Bowel, Bladder, Other Habits (Smoking etc)}
- [Cluster 4: Media Information] {Patient's Picture}

Clusters of information that have to be gathered, map elegantly to the tab UI widget. A set of semantically related information can be grouped as a single tab and a meaningful name can be associated with the tab. In addition, a tabbed interface is very appropriate in context of Nectar because it roughly quadruples the use of the limited real estate available on the Pocket PC (assuming one tab for each cluster).

It would also be desirable to design the Nectar interface in such a way that it requires minimal learning on the part of the medical professional. Accordingly, the Nectar interface is built as consisting of UI components that health officials are already familiar with. Typically, any contemporary health official is conversant with using word processors, Email clients, and filling out electronic forms online. Thus they have a basic UI skill set such as making radio button selections, using checkboxes, picking items from a dropdown list and navigating through simple forms. The Nectar Interface leverages these existing skills, nevertheless, requiring a minimal amount of training for the application of these skills to a Pocket PC interface.

Nectar Design Space

It is interesting to study the design space that we have carved out to focus our attention on, while building Nectar. The positioning of Nectar along six different design axes of interest is shown in Figure 1. This figure serves to illustrate the target design space and implicit

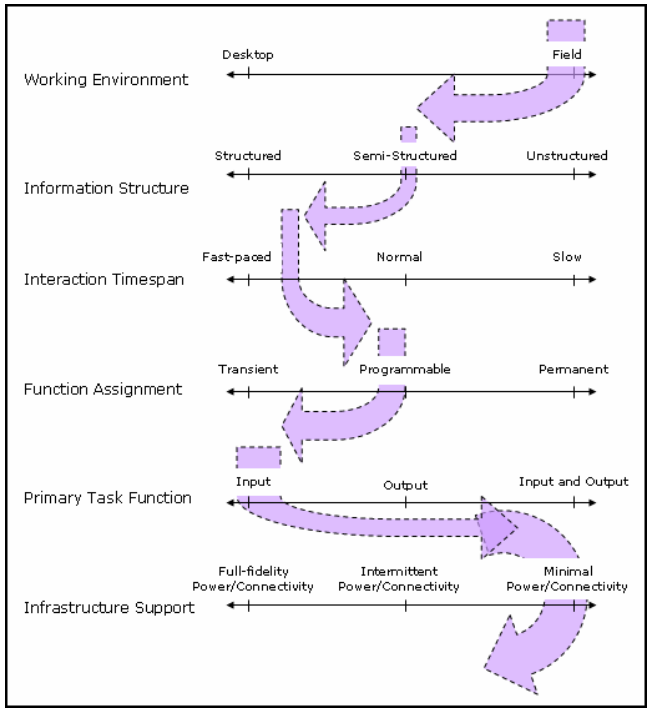

Figure 1: Nectar Design Space

trade-offs in Nectar system design, and is motivated by a related presentation in [3]

Nectar Prototype

A prototype of Nectar has been developed to target the Pocket PC runtime using C#/.NET. Screenshots of the Nectar interface running on a Pocket PC 2002 Emulator are presented in Figures 2-5. Note in Figure 5 that we have support for including an auxiliary image, in addition to the patient's image. This may be used for

Figure 2: Personal Information Tab

capturing a snapshot of a specific part of the patient's body that may be affected by a disease. We also leverage the built-in support available on most Pocket PC's for recording audio, to capture a short audio recording. This illustrates the integrated support for different types of media that Nectar affords, in addition to enabling effective collection of semi-structured data on-field.

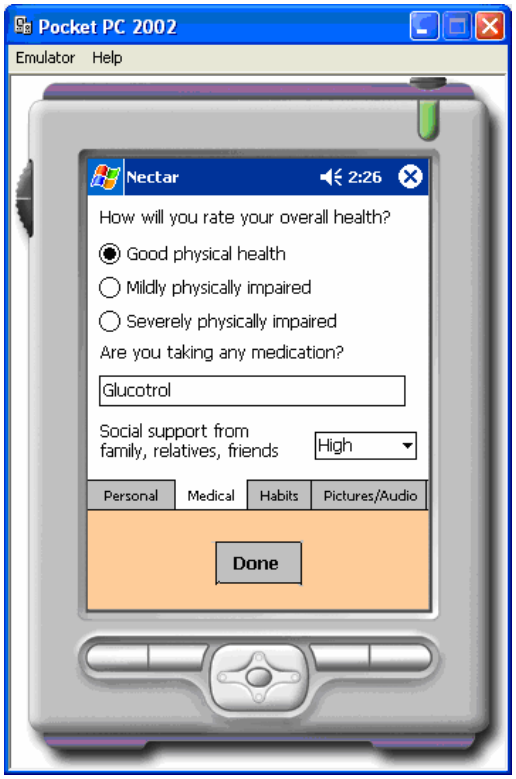

Figure 3: Medical Information Tab

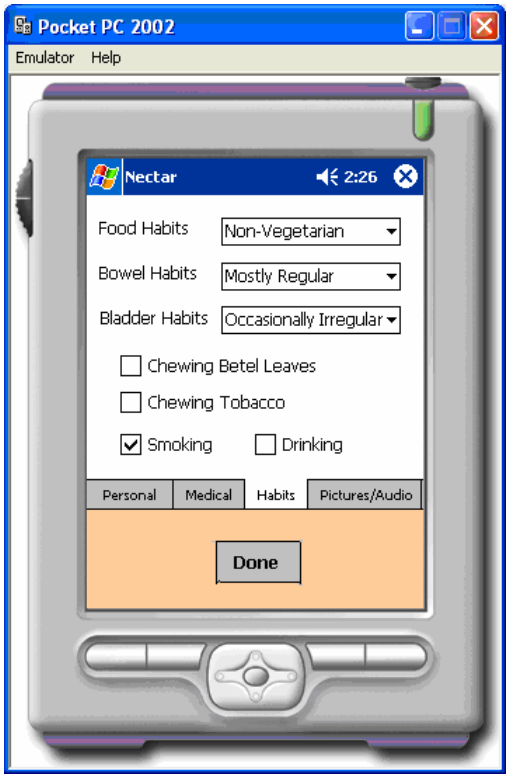

Figure 4: Habits Information Tab

Nectar is also architected to leverage and extend the ButterflyNet infrastructure [7]. ButterflyNet is a mobile capture and access system for use by field scientists engaged in biology research. A screenshot of the ButterflyNet spreadsheet, augmented for display of Nectar data is shown in Figure 6.

Related Work

The TIER project [1] aims to address the challenges in bringing the IT revolution to the developing regions. [2] presents software for use on PDAs and intended for collection of conservation biology data. [4] is likely to provide overall directions for the addition of automatic UI generation capabilities to Nectar, an interesting direction of future development. [5] could also inform further development of Nectar if one were to model a health interview like an oral history interview.

Figure 5: Media Information Tab

[6] presents CAM, a mobile document processing system that employs a camera-equipped mobile phone.

Conclusion

The key contribution that we posit in this paper is identifying and articulating the design space appropriate for semi-structured on-field data collection, and illustrating the same by building a prototype in the domain of technology-enhanced health care. It is our

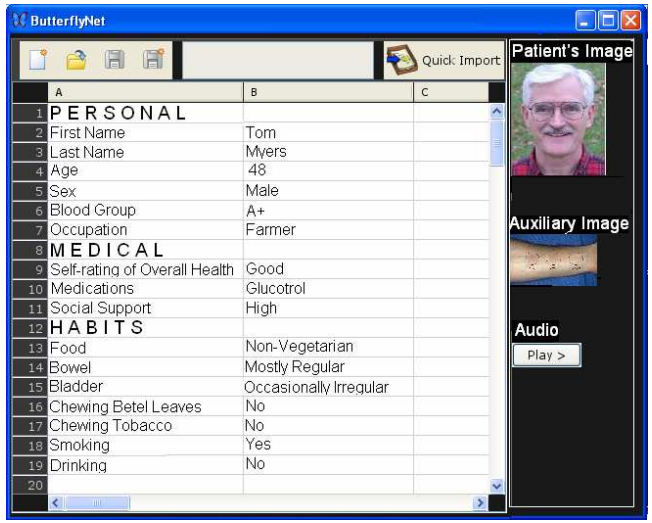

Figure 6: Integration with ButterflyNet

belief that this effort lends itself to interesting directions for further exploration.

References

[1] Brewer A. E., TIER: Technology & Infrastructure for Emerging Regions, http://tier.cs.berkeley.edu

[2] CyberTracker, http://www.cybertracker.co.za/

[3] Fitzmaurice W. G. et al, Bricks: Laying the foundations for graspable user interfaces, *CHI 95*

[4] Gajos K., Weld S. D., SUPPLE: Automatically Generating User Interfaces, *IUI 2004*

[5] Klemmer R.S.,et al, Books with Voices: Paper Transcripts as a Tangible Interface to Oral Histories

[6] Parikh S. T., CAM, *IEEE Symposium on Visual Lang. & Human-Centric Computing*

[7] Yeh B. R. et al, ButterflyNet: A Mobile Capture and Access System for Field Biology Research, *CHI 2006*