Digitization of Health Records in Rural Villages

Rajan Vaish[1], Sascha T. Ishikawa[1], Jing Liu[1], Stephanie C. Berkey[2], Philip Strong[3], James Davis[1]

[1]Computer Science Department
University of California
Santa Cruz, CA, USA
{rvaish, stishika, jingliu, davis}@soe.ucsc.edu

[2]College of Health Professionals
Western Governors University
Santa Ana, CA, USA
stephanie.c.berkey@gmail.com

[3]Health Records for Everyone
Aptos, CA, USA
strongp1@pamf.org

***Abstract*— We present a study that reviews current available methods for obtaining electronic health records (EHRs) to facilitate the provision of health services to patients from rural villages in developing countries. The study compares processes of digitizing health records by means of manual transcription, both by hiring a professional transcriptionist and by using online crowdsourcing platforms. Finally, a cost-benefit analysis is conducted to compare the studied transcription methods to an alternate technology-based solution that was developed to support in-the-field direct data entry.**

Keywords*— Crowdsourcing, human computation, health care, mobile clinics, electronic health records, digitization, transcription, data entry

I. Introduction

The introduction of digital health records has many practical and beneficial uses, even for small rural health clinics in developing countries. Clinics often generate thousands of pages of health records each week. These hard copy records are difficult to share or analyze.

Computer terminals and digital access to electronic records in the clinic is an attractive solution. Organizations such as eHealthNigeria, Inveneo, Baobab Health, and Partners in Health have taken initiatives to deploy on-site solutions. There is also a rich ecosystem of software frameworks and toolkits aimed at speeding up the software development cycle, e.g. ODK, EpiCollect, OpenMRS.

However, rural areas in developing countries typically lack the general infrastructure needed to support technology-heavy solutions [1]. As a result, the additional supporting technology that is required can be expensive and impractical. The lack of resources and/or education to handle digital technology causes its slower acceptance in the region. Furthermore, the software innovation to support technology-heavy solutions is both costly and time-consuming to develop. Thus some researchers have concluded that there is still a need for paper copies [2][3][4].

An alternate approach is to use automated digitization solutions after they are collected, rather than to instrument the clinic with a direct-to-digital technology based system. However, there are a number of obstacles in the way of digitizing health records. Hand written content can be illegible, can float beyond margins of form, and can contain medical jargons or abbreviations. All these factors render automated digitization solutions ineffective for most real medical records given the current state of the art. As a result, the digitization of health records is inevitably reliant upon human data entry.

In this work we investigate the question of who should be doing this data entry. We evaluate several options using cost, speed, and quality of digitized results.

One option is to hire a local nurse to use a computer co-located with the doctor at the rural clinic. To investigate this method, we built a complete electronic health record system and deployed it in a clinic in rural Ethiopia to investigate this method. We found that our system was extremely costly to develop and introduced workflow bottlenecks into the clinic which prevented complete records from being collected.

Another option is to let a trained medical transcriptionist digitize the forms in bulk after they have been recorded on paper. We had a USA educated licensed nurse transcribe a stack of medical records to investigate this option. We found this system to be accurate, of medium cost, and slow when the number of records becomes large.

A last option is to use a generic form digitization service based on microwork. This option promises high quality results at low cost. We evaluate this option by submitting a set of forms to CompanyA, CompanyB and CompanyC, three different commercial microwork based services that are capable of handling form digitization. We find that digital records can potentially be obtained quickly and accurately, but that these services performed unevenly, leaving questions of overall system robustness.

The main contribution of this paper is the evaluation and discussion of alternative system designs for obtaining digital health records, a task that is repeated at thousands of clinics worldwide.

II. Related Work

Automated form digitization has been studied in depth. At the robust end of the spectrum, optical mark recognition (OMR) simplifies forms to a set of bubbles which are either marked or unmarked. Dell et al. [5] successfully recorded vaccine statistics through bubble forms in rural health centers of Mozambique by introducing an android phone application (leveraging in-built camera). Automated handwriting recognition is considerably more difficult than bubble recognition, and this is an active research field with many surveys [6][7][8][9].

A study by Chrons, et al. [10] found that optical character recognition (OCR) methods produced 85% accuracy, while augmenting such methods with the innate human abilities for visual pattern recognition produced 99% accuracy. Several similar studies [11][12] have shown that the gap between what the current state of the art is capable of achieving, and what actually needs to be done can be filled by bringing humans into the loop.

The term "crowdsourcing" [13] was coined precisely to describe a decentralized approach to labor that employs workers from around the globe on a massively parallel level. Captricity is an online service that uses crowdsourcing approach to data entry. At Captricity, users upload and define templates of digital documents, which are shredded into fragments and sent to Amazon's Mechanical Turk for transcription [2]. The resulting workflow was reported to significantly reduce both the time and cost of data entry compared to traditional services.

Digital records can also be captured at the same time the original paper forms are recorded. Ratan, et. al. [4] introduce a handheld electronic slate prototype, where anything written on the paper is simultaneously digitized and validated, producing both paper records and digital copies. The device costs approx. $100 and battery runs for 5-6 hours. The report cites improvements in recording times, accuracy, and record completeness.

It is also possible to implement electronically assisted data entry systems in places where technology literacy is limited. Parikh et al. [1] implemented CAM, which leverages the ubiquity of camera-equipped mobile phones to scan barcodes that identify each field in the paper form. The mobile application is able to recognize the code and prompt users to enter the corresponding values with their phone's keypad.

Along the lines of assisted technology for data collection, DeRenzi et al. [14] implemented an electronic version of diagnostic procedures, dubbed "e-IMCI," which operates on a personal digital assistant (PDA) and assists health workers with procedural tasks while collecting health information. It was shown to successfully reduce mistakes such as skipped steps, miscalculations, and branching-logic errors.

III. CONTEXT OF OUR WORK

A. Rural Clinic

Health Records for Everyone (HR4E) [15] is a 501(c)(3) not-for-profit charitable and educational organization whose goal is to support community health assessment projects in developing countries. Furthermore, HR4E is dedicated to promoting the widespread adoption of EHRs, which are to be stored in a standards-based format that promotes data interchange, analysis and aggregation.

HR4E sponsored a pilot study during a medical mission that deployed mobile health clinics in rural Ethiopian villages. According to a report [16] by the nonprofit organization RAD-AID, in 2007, the physician-population ratio in Ethiopia was 1:38,000 compared to the 1:1,700 in India and 1:400 in the United States. Furthermore, out of 932 Ethiopian physicians trained at home and abroad between 1987 and 2006 only 20% are practicing in the Ethiopian public health sector, most in urban areas. The purpose of the medical mission was to conduct health assessments and create health records for children in local community schools. The health services provided include education on basic hygiene and disease prevention and the treatment of acute disorders.

During their daily operations, patients were directed to form a queue at the entrance to the clinic. Health-related information was collected from each patient throughout various stages of the assessment process, where at each station a clinician completed the corresponding section on a paper form.

B. Existing Paper Form

The health evaluation form, filled out by nurses and physicians, documents patient information including demographics and vital signs, health history, results from a physical examination, and an assessment plan. The form was designed to meet the needs of healthcare workers in a busy clinic rather than facilitate the digitization process, and thus has a mixture of data types and non-precise field localization. We separated this form into several conceptual types during our analysis.

1) Distinct Data Fields: Distinct data fields hold short text such as patient names, numerical measurements such as vital signs, and multiple-choice selections. Because they are consistent within the same form, they are much easier to digitize and workers were able to do so, in general, with high accuracy. Figure 1 shows some examples of fields of this type. Notice that the markings extend into unexpected regions, but that we generally expect a transcriptionist to be able to understand what was intended.

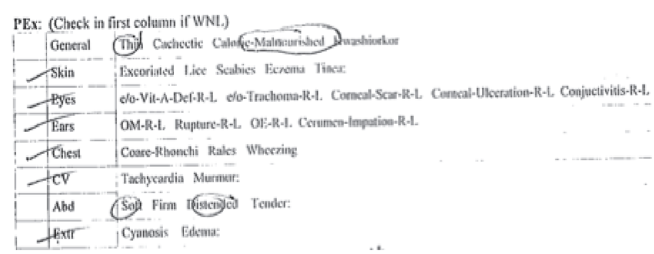

Fig. 1. Sample of the "Physical Examination" section of the health evaluation form. This section consists of multiple-choice-style fields. A check mark is required on the left side if a row is empty. The latter exemplifies the use of medical acronyms that may not be intuitive to the casual reader.

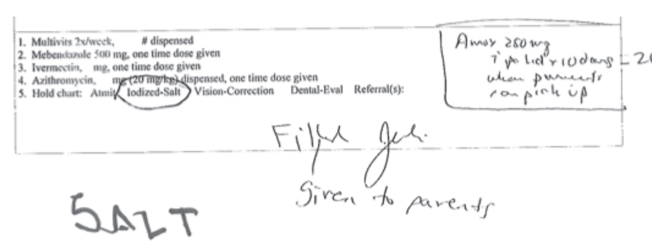

Fig. 2. Sample from the "Assessment/Plan" section of the health evaluation form. This part contains some multiple-choice fields; however, it is also used to compute dosages and provide additional hand-written comments that may be difficult to read.

2) Unstructured text: Unstructured text is found in some areas of the form. These include notes that have been made in designated areas enclosed by a box, as well as scribbles along the margins. Medical jargons, dosage computations and illegible handwriting make the process of transcription difficult (especially for non-medical professionals). Figure 2 provides an example of a portion of the form with unstructured text notes.

IV. ON-SITE DIGITAL RECORDS

As discussed in the introduction of this paper, on-site digital recording was our first approach to capture the medical records electronically at the time of patient visits. The clinic we worked with expressed a preference for this method, and sponsored the development of an electronic health record system that would suit their needs.

We initially expected this development to be straightforward, leveraging existing form building tools, already tailored for medical clinics. In practice, clinics serve diverse patient populations with varying needs, which warrant different assessments and treatments. Any sufficiently general forms to cover many clinics will be too complex for efficient use in a crowded busy context. Therefore, the corresponding digital forms must be tailored to suit the specifications dictated by each clinic.

The digital form was developed as part of a joint project involving HR4E, a computer science class, and a paid engineer. The class met during the spring of 2011 and again in 2012, for a total of 10 weeks. The students in the classes evaluated usefulness of existing tools, such as ODK and OpenMRS, and built applications, yet all of them lacked some key functionality desired by this clinic. As a result, the final software was completely custom designed by the paid graduate student/engineer [17].

The software was field-tested in the medical clinic during October 2011. During this time, an assistant at each station entered data directly into a laptop, while the doctor wrote on the original paper form. This system was used for 1000+ records over a period of two weeks. However, during busy days, the software failed to keep up with the influx of patients and hence, it was discontinued requiring further usability refinements. An attempt to redesign the software during the second iteration of class failed.

Although many organizations recognize the cost of deploying computer hardware, the cost of custom software is frequently ignored. In this instance, greater than US$20,000 of software engineering time was used (assuming a modest annual engineering wage of US$60,000), producing a complete first version and an incomplete second version of software. Each version took four months of engineer's half-time job. Since every clinic has a unique set of forms, and indeed the clinic we worked with revised their form between the time we started and ended the project, we believe this is an ongoing operational cost, not a one-time upfront cost that can be amortized across many clinics.

In summary, building an electronic health record system proved costly and interfered with existing clinic workflow.
Of course the authors believe that it is possible to design software that works; indeed many clinics worldwide have high quality electronic record systems. But we also believe that our experience is likely to represent the experience of a typical clinic. It is going to cost more and take longer to achieve a system that works than is initially expected.

V. TRANSCRIPTION BY TRAINED NURSE

Many paper records held by large medical organizations and corporations are simply transcribed after the fact. Although labor intensive, we asked a registered nurse to transcribe hardcopies of the forms. Instead of cumbersome spreadsheet, the data was entered in a digital version of the health evaluation form.

The mobile form, shown in Figure 3, was created using Open Data Kit (ODK) [20], a suite of open-source tools that can be used to create XML forms for data collection on mobile devices.

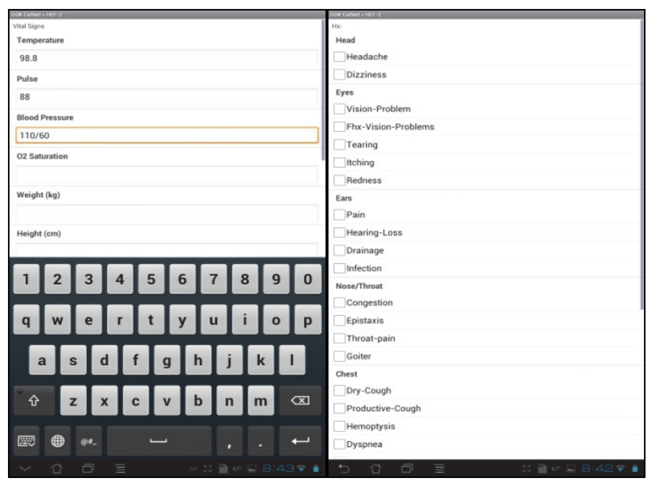

Fig. 3. Screenshots of our electronic form implemented using Open Data Kit. Documents were transcribed with Asus TF101 tablet with keyboard.

Note that this form was substantially more basic than the one deployed in the clinic, taking only a day to develop. The transcriptions were deemed of near perfect quality by a doctor who was part of the campaign in Ethiopia. The time required to transcribe each document ranged from 1-5 minutes, depending on the content. A set of 100 documents took approximately 4 hours to transcribe. According to the U.S. Bureau of Labor Statistics [18] the median pay in 2011 for a medical transcriptionist was $16.37/hr. Given this information, we estimate that a set of 1000 forms, an amount comparable to the number of forms produced during a health mission, would take 40 hours and approximately US$650 to transcribe.

To summarize, our professional-level translation was very accurate with costs that may be appropriate for some clinics. However, as the number of forms increases (1000+) the latency to receive transcribed results becomes too slow for some needs. The clinic we worked with was a mobile deployment, active only a few times per year for two weeks each session. Thus the total number of records per year is in the 1,000-10,000 range. For this clinic, professional transcription appears to be cheaper than custom software.

VI. TRANSCRIPTION BY CROWD

Crowdsourced microwork has allowed some jobs to be completed faster, cheaper, and better than through traditional outsourcing companies. Transcription in particular has been investigated by Chen et al. [2] and Gupta et al. [19]. In this project, we tried to leverage its benefits and conducted experiments to compare speed, cost, and accuracy, on different platforms. Beginning with a pilot study of 5 forms, we conducted a thorough study on a set of 100 forms over three platforms, which we'll state as CompanyA, CompanyB and CompanyC.

The three commercial services were selected because of their differentiating approaches. Both CompanyA and CompanyB use template-based micro task assignments that designate fields for individual transcription. Each field is distributed to multiple workers who have access only to the individual field they are transcribing. CompanyA takes care of template creation, while users are required to create one at CompanyB. CompanyC is a more generic provider for microwork. At the time of writing, CompanyC's transcription service operated by assigning two workers to transcribe entire digital copies of forms. If a reasonable consensus was reached between the two workers, the transcriptions were saved onto a spreadsheet and available for download. In addition to comparing the services, we investigated whether image quality affected the transcription results.

A. Obtaining Digital Images

The set of 100 forms were photographed using both an office scanner and a mobile phone camera. The office scanner automatically loaded and scanned the stack of documents producing high quality black and white, 595 x 842 pixel PDF files (Figure 5, right). While a 5-megapixel phone camera produced color images at a resolution of 1920 x 2560 (Figure 5, left). Although the resolution was greater than the black and white scanned copies, the images from the mobile phone were subject to uneven lighting conditions, which produced shadows, poor contrast, and occasional blurriness. Figure 4 shows an ad hoc "scanning device" built by using a milk crate and placed near window for adequate light. The manual process to photograph all 100 forms took less than 20 minutes, and caused occasional rotation/shifting of forms by a small amount. However, no efforts were made to correct for color or alignment before they were submitted for digital transcription. This was done in an attempt to mimic the clinic environment that has a limited capacity to post-process photos.

Fig. 4. Setup that was used to set up the camera station, consisting of a camera-enabled mobile phone zip-tied to a milk crate.

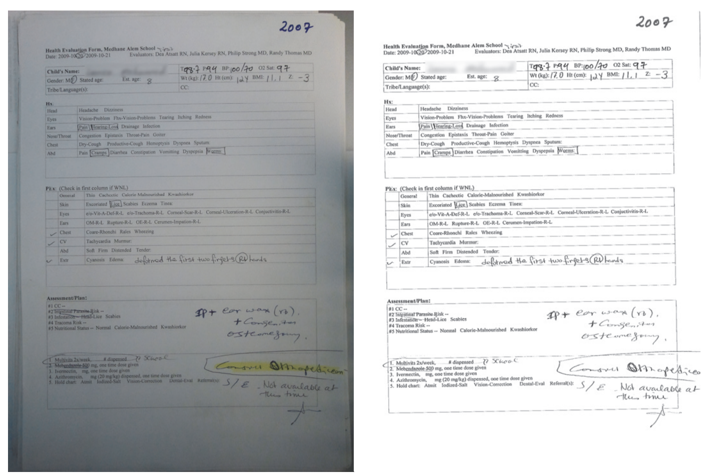

Fig. 5. Samples of digitized versions of the paper forms using a mobile phone camera (left) and an office scanner (right).

B. Evaluating Patient Names

Since different portions of the form had different characteristics, we discuss patient names, distinct data, and free-form text separately.

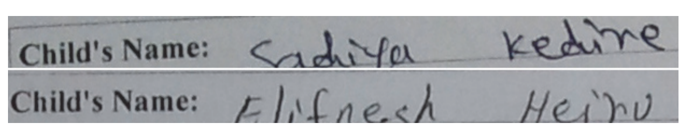

Fig . 6. Examples of the name field from the health evaluation form.

In order to conduct a quantitative assessment of the transcription services, we evaluated the accuracy of transcribing patient names written in definitive boxes. Among the local population, spelling for names is phonetic and inconsistent. Consequently, it was assumed that with the lack of specific cultural context, workers would be more likely to resort to recognizing individual characters. This enables us to quantitatively measure the performance of workers' abilities as they pertain to the transcription of unfamiliar text in a variety of handwritings, which parallels the job of transcribing medical forms. Figure 6 shows some examples of the name field.

We used the Levenshtein distance to measure the number of edits required to change a block of text into the corresponding text from the ground truth. Given the edit distance as a metric, the transcription accuracy was calculated from the ratio of correct characters to the total number of characters entered.
The results are given in Table 1. Notice that two of the services performed at better than 95% accuracy, surprising given the ambiguity in the source data.

TABLE I. EVALUATION OF PATIENT NAME TRANSCRIPTIONS

	A	B	C
Avg. Accuracy[1]	0.88	0.97	0.98
Avg. Distance[2]	1.68	0.38	0.33

[1] The average accuracy was calculated using the ratio of correct characters to the total number of characters entered.[2] The average distance represents the number of edits required to change the transcribed names into the correct ground truth names.

C. *Evaluating Distinct Data Fields*

The form contained many fields with precise boolean, numeric, or multiple choice values. These included items such as gender, age, language, temperature etc. A total of 29 fields were included in the aggregate statistics below.

The digitized fields were manually compared to the ground truth obtained by a US educated nurse. Since a single character change often results in meaningless data, any discrepancy between the ground truth and the transcribed value resulted in the field being labeled "Incorrect." In addition to comparing transcription across services, we compared different approaches to imaging the paper documents. Table 2 shows results for images both from a mobile phone camera, as well as from a document scanner. Notice that there was no significant difference between the imaging modes. Notice also that all of the services performed well on this kind of transcription task.

TABLE II. AVERAGE ACCURACY OF DISTINCT DATA FIELDS

	A	B	C
Scanned	0.91	0.96	0.97
Camera Phone	N/A	0.96	0.98

We made no effort to register form position before submitting to the various services. This had little effect on CompanyC, because workers could view the form in its entirety during transcription. However, CompanyB uses templates that only display portions of the form to its workers during transcription, potentially cutting off or adding extra data when the image is misaligned.

D. *Evaluating Free-Form Text*

The form had multiple sections of free-form text. We evaluated the free-form text from the "Plan" section of each form by comparing it to the ground truth. Results were evaluated by a registered nurse, who manually graded the fields using the four categories: *completely correct*, *incomplete but correct*, *mostly correct* and *incorrect*; where *mostly correct* accepted an entry with one or two errors.

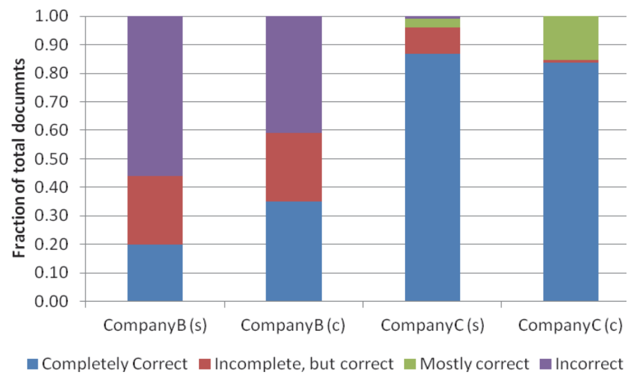

Fig. 7. Evaluation of free-form text using CompanyB and CompanyC by a registered nurse; (s) symbolizes scanned, while (c) symbolizes camera phone.

Figures 7 provide results from CompanyB and CompanyC. Note, first, that using a camera phone to capture images did not noticeably degrade results. Second, note that overall, transcriptions of free-form text made with CompanyC had much greater accuracy than from CompanyB. This is because workers from CompanyC were given access to entire forms to transcribe, whereas CompanyB segments each form into independent fields that are distributed to its workers. The CompanyB model is, in theory, more robust because many redundant entries can be obtained for each individual field and aggregated into a final transcription. However, the free-form text of our medical records was often written in arbitrary locations, which proved too problematic for a template-based transcription service. CompanyA also uses a template model, which we believe lowered the quality of their results sufficiently that we are not reporting them.

In summary, free-form text was clearly more difficult than distinct fields, but at least one of our crowd transcription services performed very well, obtaining perfect accuracy on 85% of records, and partial accuracy on 99%.

E. *Cost and Time*

Based on our experiments, we can estimate the cost, in terms of time and money, to transcribe paper forms to obtain electronic health records. Figure 8 graphically illustrates the results. The crowdsourced transcription services, CompanyB and CompanyC, cost US$0.20 and US$0.50 per page. CompanyA provided us with a free evaluation, so we do not have data on their cost. For distinct data fields, both services provided comparable transcriptions that were 97% and 98% correct. However, there was a marked difference in the transcription quality of the free-form text. Using the higher quality results, and higher price, of CompanyC to obtain digital records, the total costs would amount to US$50 per 100 forms. For comparison, we found that the same task performed by a medical transcriptionist would cost approximately US$65.48 and provide near perfect results. This latter price is not including overhead and operating costs for the transcription company, thus it seems clear that crowdsourced transcription can provide some cost savings, at the price of accuracy in difficult parts of a form.

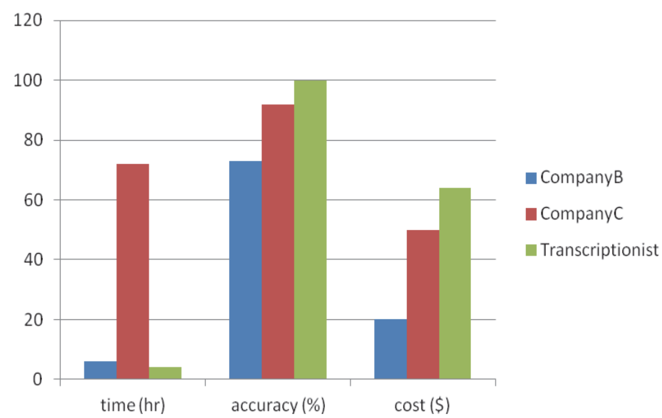

Fig. 8. Summary of a cost-benefit analysis comparing form digitization approaches using crowdsourcing versus manual transcription. Note that for a small number of documents, it requires less time (but more money) to hire a transcriptionist. However, for a large number of forms, the crowdsourcing approach is more efficient, both in terms of cost and time.

Our experiments show that a medical professional is able to transcribe 100 records in approximately 4 hours, the required time for crowdsourced services is significantly longer. The set of 100 forms took ~70hr and ~8hr for CompanyC and CompanyB, respectively. However, on large datasets crowdsourced solutions have greater parallelism and results 100 times faster than dedicated data entry clerks have been reported [2].

V. CONCLUSION

We have investigated several methods for producing electronic health records in rural clinics: on-site computerized forms, professional transcription, and crowdsourced transcription.

On-site computerized forms were by far the preferred solution by our partners running the clinic. Although we did deploy a live form in their clinic, it was more costly and time consuming to develop than expected, and created a bottleneck in clinic workflow when deployed. This resulted in loss of data as some portions of the clinic simply quit using the digital system. Our code is now pending revision to match both design changes as well as form changes instituted by the clinic.

Professional transcription was originally dismissed as too costly. However, our results seem to indicate that the particular clinic we are working with could obtain digital records for only a few thousand US dollars per year. This is far less than the amount of money being spent on developing and maintaining a fully computerized system.

Crowdsourced transcription was the least costly of the methodologies investigated, but also the least accurate. Accuracy was over 95% on some portions of the form, making it suitable for some tasks, such as producing aggregate statistics to report to sponsoring agencies. Much of the loss of accuracy was due to challenging form design. Simple changes such as converting “circle the answer” to “checkboxes,” and drawing lines to clearly delineate field boundaries on the paper form would substantially increase the transcription accuracy. We are currently planning another field trial in order to investigate how each of these strategies converges (or does not) to an acceptable solution after another iteration of design.

REFERENCES

[1] Parikh, T.S., Lazowska, E.D. Designing an architecture for delivering mobile information services to the rural developing world. In *Proc. of the 15th international conference on World Wide Web*, 2006. P.791-800.

[2] Chen, K., Kannan, A., Yano, Y., Hellerstein, J.M., Parikh, T.S. Shreddr: pipelined paper digitization for low-resource organizations. In *Proc. of ACM DEV*, 2012.

[3] Singh, G., Findlater, L., Toyama, K., Helmer, S., Gandhi, R., Balakrishnan, R. Numeric paper forms for NGOs. In *Proc. of ICTD*, 2009.

[4] Ratan, A.L., Chakraborty, S., Chitnis, P.V., Toyama, K., Ooi, K.S., Phiong, M., Koenig, M. Managing Microfinance with Paper, Pen, and Digital Slate. In *Proc. of ICTD,* 2010.

[5] Dell, N., Breit, N., Chaluco, T., Crawford, J., Borriello, G. Digitizing paper forms with mobile imaging technologies. In *Proc. of ACM DEV*, 2012.

[6] Casey, R.G., Lecolinet, E. A survey of methods and strategies in character segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 18,7 (July 1996), 690-706.

[7] Trier, Ø.D., Jain, A.K., Taxt, T. Feature extraction methods for character recognition-A survey. *Pattern Recognition*. 29, 4 (April 1996), 641-662.

[8] Plamondon, R., Srihari, S.N. Online and off-line handwriting recognition: a comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 22,1 (January 2000), 63-84.

[9] Pal, U., Chaudhuri, B.B. Indian script character recognition: a survey. *Pattern Recognition.* 37,9 (September 2004).

[10] Chrons, O., Sundell, S. Digitalkoot: Making Old Archives Accessible Using Crowdsourcing. In *AAAI Workshops*, 2011.

[11] von Ahn, L., Maurer, B., McMillen, C., Abraham, D., Blum, M. reCAPTCHA: Human-Based Character Recognition via Web Security Measures. *Science*, 321 (5895), 1465-1468.

[12] Narula, P., Gutheim, P., Rolnitzky, D., Kulkarni, A., Hartmann, B. MobileWorks: A Mobile Crowdsourcing Platform for Workers at the Bottom of the Pyramid. In *AAAI Workshops*, 2011.

[13] Howe, J. Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business. Crown Publishing Group New York, NY, USA 2008.

[14] DeRenzi, B., Lesh, N., Parikh, T., Sims, C., Maokola, W., Chemba, M., Hamisi, Y., Shellenberg, D., Mitchell, M., Borriello, G. e-IMCI: Improving Pediatric Health Care in Low-Income Countries. In *CHI 2008 Proceedings*.

[15] HR4E: http://healthrecordsforeveryone.com

[16] RAD-AID. http://www.rad-aid.org/pdf/Ethiopia-Country-Report2.pdf

[17] Gainer, A., Roth, M., Strong, P., Davis, J. A Standards-Based Open Source Application to Gather Health Assessment Data in Developing Countries. In *Proc. of IEEE GHTC*, 2012.

[18] U.S. BLS: http://www.bls.gov/oes/current/oes319094.htm

[19] Gupta, A., Thies, W., Cutrell, E., Balakrishnan, R. mClerk: Enabling Mobile Crowdsourcing in Developing Regions. In *CHI 2012 Proceedings.* Austin, TX, May, 2012.

[20] Hartung, C., Anokwa, Y., Brunette, W., Lerer, A., Tseng, C., and Borriello, G. Open data kit: Building information services for developing regions. In *Proc. of ICTD*, 2010.