

Relationship-based Business Process Crowdsourcing?

Jacki O'Neill & David Martin

Xerox Research Centre Europe, 6 Chemin de Maupertuis, 38240, Meylan, France

{Jacki.Oneill, David.Martin}@xrce.xerox.com

Abstract. New technologies do not always benefit the worker, especially when harnessed by organisations seeking ever cheaper labour. Crowdsourcing is a technology-enabled way of working which offers the potential to bring work to far flung communities. However, it is something of a double-edged sword and there are many socio-technical and ethical challenges. In the micro-task market crowdsourcing platforms tend to be designed largely for the advantage of the organisation requesting work, rather than the worker. This paper contributes to research calling to redress this balance [2, 6]. It describes the findings of an ethnographic study of an outsourced business process – healthcare form digitization – as performed by workers in-office (India) and @Home (USA). It reveals the complexities of the relationships between worker and organisation and argues that designing some aspects of these relationships into crowdsourcing platforms and applications is as beneficial for the organisation as it is for the worker.

Keywords: Crowdsourcing, ethnography, business process outsourcing, relationship-based crowdsourcing

1 Introduction

New technologies offer new possibilities for working and there are on-going government and private initiatives in most countries throughout the world which aim to harness the power of information and communication technologies (ICTs) for socio-economic development. For example, government initiatives in rural India have led to the use of ICTs to create jobs in rural areas from enabling small producers to take part in large order production¹ to rural Business Process Outsourcing² (BPO). However, new applications of ICTs are not always primarily motivated to benefit the worker since work reconfiguration and technology design is often motivated or harnessed by organisations seeking ever cheaper labour pools. Common approaches include a) deskilling, typically dividing complex tasks into smaller parts and using ICTs to assist lower skilled workers to complete these parts; b) offshoring; distributing the work to areas where labour is cheap, or a combination of both a) and b). A poignant example

¹ E.g. <http://ropeinternational.com/aboutus.html>

² E.g. <http://desicrew.in/>

of this is the outsourcing movement, which typically undertakes to complete non-essential or non-central functions for various customer entities at a lower cost. The primary push for outsourcing from the customers perspective is cost reduction through, for example, labour arbitrage i.e. by taking a business function outside of the employer the new workers do not have a right to the same terms and conditions as in-house employees. For the outsourcer, profits are achieved not only through economies of scale, but also through deskilling and offshoring. This is not inherently negative, since for example BPO has had a positive impact on the Indian economy³, providing much needed employment and propelling infrastructure development. However, it is not a completely rosy picture either, with an frequently undervalued workforce doing low skilled work.

In some domains such as customer care it is questionable whether deskilling is an appropriate strategy, since undoubtedly more professional agents would be better able to support customer needs. However, there is also a large body of work that cannot yet be automated but which in theory could be completed by anyone with a reasonable level of literacy, e.g. data entry of handwritten forms. In this paper we report on a study that was undertaken to see what it would mean to crowdsource this work. Crowdsourcing is a new and growing mode of organizing activities. It has the potential to provide work in developing countries because the workers and work requester (i.e. the enterprise putting out a task to be completed) do not have to be collocated [1]. Thus it is not necessary for the enterprise to set up offices in a particular country, with the attendant cost-benefit analysis, internal and external political choices and so on that that involves. Workers can come from anywhere as long as they have the right skills and a reasonable technical infrastructure. Thus whilst outsourcing has not taken off to the same extent in Africa as in India⁴, due to a poorer technology infrastructure and the (comparatively) high costs and difficulties of setting up companies (labour costs, corruption, lack of government incentives, etc.) crowdsourcing could, in theory bring some of this work to Africa⁵.

Crowdsourcing is defined as the act of taking a task traditionally performed by an employee or contractor, and outsourcing it to an undefined, generally large group of anonymous people, in the form of an open call⁶. It is frequently used for work which cannot be trivially automated, e.g. requiring semantic knowledge. The term has been widely applied but we focus here on the crowdsourcing of microtasks which are completed for pay and thus fall clearly into the category of work and indeed typically into the category of piecework [2]. Typically a crowdsourcing vendor mediates between the employer and the worker, generally providing a platform through which

³ For example, the Information Technology-Information Technology Enabled Services (IT-ITES) sector (US\$100 billion industry) alone has increased its contribution to India's GDP from 1.2% in FY1998 to 7.5% in FY2012 and plays a strong role in generating Employment in India.

⁴ Even though South Africa and Ghana might be considered hotspots of outsourcing compared to the rest of Africa, they are nowhere near Indian levels.

⁵ Currently the technical infrastructure in much of Africa is poor, however, initiatives such as those to tap into WACS (the West African Cable System) promise to change this

⁶ <http://en.wikipedia.org/wiki/Crowdsourcing>

crowdsourcing can be undertaken. Amazon Mechanical Turk⁷ (AMT) is probably the best known Crowdsourcing micro-task platform. It enables individuals or organizations to post small, usually low skill, tasks (digitization, translation, search, image labeling, etc.) in large volumes to be taken up by individuals for execution. The workers post back their work for evaluation and get paid on acceptance. AMT has thousands of micro-tasks, which can be executed in seconds or minutes, with payments typically in the order of few cents.

However, like outsourcing, crowdsourcing can be considered to be something of a double-edged sword and there are many socio-technical and ethical challenges. Crowdsourcing offers the potential for providing well needed income for people in developing countries, however employment models tend towards if not quite the idea of getting something for nothing, then certainly something for very little. The appeal to the enterprise is the ability to get work done quickly, without the financial burden of a contracted workforce. Much of the research focuses on (minimum) incentive schemes, how to deal with bad work and spammers and so on. Whilst not denying these are important issues, crowdsourcing tools tend to be somewhat biased towards the needs of the enterprise often to the detriment of the worker [3]. However, as with outsourcing we do not believe that crowdsourcing is necessarily a negative work model, it can just appear that way because, with some notable exceptions [e.g. 1, 4] the tools, technologies and work models employed thus far have been designed to fit a particular organisational perspective. Ethnographic research within domains such as HCI and CSCW is well positioned to provide a critique of this perspective, since it typically reveals the human expertise necessary to carry out even apparently low skilled work [5, 6, 7].

If crowdsourcing is to fulfill its promise of providing employment in poorer communities, companies need to find ways to introduce crowdsourcing into their business processes, thus providing more consistent ongoing work opportunities. This is likely to require considerable process redesign and technology innovation. In this paper we describe an ethnographic study of outsourced form digitisation, which was undertaken as a first step in understanding the possibility/requirements for work and technology (re)design if this work were to be put out to the crowd. Rather than taking the purely organisational perspective of cheaper work done quicker, we want to investigate models where crowdsourcing could be beneficial for both parties - worker and organisation. To this end we consider factors such as the skills and knowledge involved in even this low skilled work and how the nature of the relationship between worker and employer impacts on performance (of both parties!). This study reveals some of the challenges that need to be overcome if such work is to be crowdsourced, given its ongoing nature, strict turnaround times and high quality requirements. In addition it also demonstrates how a more relationship-based approach to crowdsourcing is likely to be as beneficial to the organization requesting the work as it is to the worker. The concept of relationship-based crowdsourcing has been introduced by [3] as part of a call to take the workers perspective in crowdsourcing. Whilst we might hope that all enterprises would aim to treat their workers fairly, in practice this is more likely when

⁷ <http://www.mturk.com>

they can see a clear benefit in their day-to-day activities. If crowdsourcing is to be a viable work option to support socio-economic development, it needs to rid itself of its exploitative image. This paper is a first step in showing how, for this type of BPO work, doing so will likely be as beneficial to the organization as to the worker.

2 Crowdsourcing Literature

Much of the crowdsourcing research thus far has focused on AMT and the bias towards the organizational perspective is clear. However recently more research is re-addressing this balance and we describe work from both perspectives here.

Handling bad data or poor quality work, whether by scammers or genuine workers is a major issue on AMT. Approaches include task design, especially getting the work done iteratively (e.g. find, fix, verify for correcting documents [8]) and/or rating workers through reputation schemes. Whilst it is possible to simply weed out bad work/workers, a more promising approach is to distinguish between scammers and genuine workers and to enable on the job learning so that genuine workers can improve. [10] demonstrated how including initial training sets and gold standard data (i.e. with a known correct output) throughout the task can be used to both judge worker quality (and weed out spammers) and provide feedback enabling genuine workers to improve. Whilst AMT provides means to rate workers' reputations, there is no equivalent means in the tool itself of rating work requesters [3, 11] despite the problem of unscrupulous requesters being well known. To compound this inequality, workers' reputations are harmed if work requesters reject their work. Furthermore AMT deliberately hides the relationship between work requester and worker, often to the detriment of the worker [12]. To begin to address this inequality, Irani and Silberman developed a plugin⁸ to AMT which enables crowd workers to rate work requesters [13]. This plugin provides a very useful resource for crowd workers enabling them to identify good and bad requesters on the basis of their collective experience, but it is telling that Amazon have not taken up the clear need for such functionality by embedding it in their tool.

Putting tasks out to the crowd typically involves work redesign and since AMT is a microtask platform, the work needs to be broken down into suitable chunks. An innovative example of this comes from IBM for the OCR correction of scanned documents [14]. They present all the characters which the OCR engine identifies as being the same but with a low confidence rating together rather than having the keyer move from form to form. This improves productivity, removes security concerns (the worker only sees a table of letters with no semantic information) and since it makes the task very low skill, they say it makes it easy to recruit workers on-demand, although this short paper does not document the latter. This research is the closest to our application domain, but not all work is suited to such ultra-micro-tasking, for example the handwritten forms commonplace in our application domain would be hard to split up in this way. Since we are employing people, with all their myriad skills, we believe it

⁸ <http://turkopticon.differenceengines.com/>

makes sense to capitalize on their abilities, rather than to decompose the task into its smallest, lowest skill parts. As Ipeirotis⁹ described in his blog, whilst ultra-microtasking can help unskilled new workers, as workers become experienced this embedded workflow can get in the way of them using their knowledge and expertise to produce higher quality work more efficiently. People may also work better, even on simple tasks, when not in complete isolation. [15] found that collaboration amongst crowd members, e.g. asking for help with unknown words during translation, improved quality and promoted learning.

Pay is key concern in crowdsourcing. From the workers side research shows that whilst incentives are important they are not the only factor in determining which work gets done. [11] found that higher pay increases completion rate, time spent on the task and quality but that qualified workers are less affected by pay. In contrast [17] found that pay increased throughput but not quality and [18] found that wage per job was not the sole motivator – rather workers focused on their ability to reach salient targets, e.g. whether there was enough work in this category for them to earn some target amount. Unfortunately much of the research into pay is based on experimental studies rather than real tasks and thus it is hard to really understand how pay impacts performance and indeed take up of jobs in real crowdsourcing work. Certainly from the worker's side, as with employees in any workforce, there is concern about getting fair pay for the work done [3], especially given that very low wages are rife [3, 11, 12]. When we take into account that not only does completion time vary but some tasks never get completed on AMT [19], we can see why some requesters have rejected AMT in favour of other models of working such as ODesk¹⁰ – an online market for contract labour. [19] used ODesk and chose to pay an hourly rate because their job included hard tasks that needed to be completed to high quality and these are the sort of tasks which might likely fail to be completed satisfactorily using AMT. [19] examines fairness in AMT in more detail and discusses why fairness until now has been something of a side issue in crowdsourcing discourses, the design of the platform and the design of crowdsourcing tasks.

Whilst at the moment it might be a choice between AMT and forms of contracted labour, it should be possible to design a different sort of crowdsourcing system which balances the advantages of AMT (rapidity, diversity and access to non-experts) with concerns about ethics [11], completion time and quality¹¹. [3] proposes that a more relationship-oriented approach, between requesters and workers, would be beneficial for both – giving requesters higher quality work and workers more fair conditions. As we will describe in the following sections this is particularly important for the sort of work we are hoping to crowdsource. The fieldwork revealed a number of features which would seem to make it more suited to relationship-based crowdsourcing.

Despite its problems, various companies and researchers are already examining crowdsourcing as a means to socio-economic development. Samasource¹² is an exam-

⁹ <http://www.behind-the-enemy-lines.com/2012/02/need-for-standardization-in.html>

¹⁰ <https://www.odesk.com/>

¹¹ Indeed, researchers have built mobile platforms such as in [4] for more 'ethical' work, but because of the domain constraints our focus is on computer-based crowdsourcing.

¹² <http://samasource.org/>

ple of an alternative ethical crowdsourcing company which trains people living in poverty to complete specific tasks for a living wage. Such companies support relationships between work requester and worker – through training, trusted workers and so on. [1] point to a number of issues to be addressed if crowdsourcing is to be extended to low income workers in developing countries including interface design (see [21] for some design solutions), greater regulation of requesters; training; peer-to-peer skills transfer; and predictability of income – which is currently hard to achieve. For our application domain, we would be likely to target the educated underemployed rather than the really low income sectors described in [1, 21] because the nature of the task requires a high level of literacy in English.

2 Method and Study Sites

To understand what it would take to introduce crowdsourcing into a healthcare form digitization workflow, we embarked on an ethnographic study of that work as it is undertaken now in an outsourced environment. Ethnomethodological ethnographies have long been used to understand the nature of the workplace and to conceptualise innovative design solutions (see for example, [22]). Ethnography has proved useful for design because it reveals the underlying skills, knowledge and practices involved in even routine, mundane, low skilled processes. By understanding the work as it is undertaken now and given a particular technology scenario, in this case Crowdsourcing, we can begin to map out what is required to undertake that work in this new setting, as has previously done with for example mobile payments [23]. We believe the ethnographic approach also provides a useful way in to thinking about ethical design, because of its focus on workers skills and the value that they bring to the process.

Approximately five working weeks were spent by one ethnographer (1st author) in the outsourcer's offices in Bangalore and Kochi, India between March and August 2011 undertaking an ethnographic study of the form digitization process for healthcare forms. The entire Indian workflow for three different clients was studied primarily through observation of ongoing work, supplemented by in situ interviewing. The researcher sat with and observed agents doing data entry work at all skill levels, shadowed team leads, supervisors and quality control. In addition a half-day visit was made to the workflow control team, Production Control, who demonstrated to the researcher how they managed their workflows. A further two weeks were spent by a second ethnographer (2nd author) studying home-based workers and support staff in Utah, USA in 2012. The researcher visited agents and supervisors at home and spent some days interviewing and observing the US-based Production Control team. The data collected is qualitative: field notes, audio recordings, photographs and relevant artefacts such as rule books and procedures were collected. It was analysed from an ethnomethodological perspective [24] and used to create in-depth descriptions of the observed work for each client. Ethnomethodology is a non-theoretical analytic orientation uses qualitative data to explicate the endogenous social means and methods by which participants in a setting carry out, organize and reason about their activities.

Presented here are summary descriptions of particular common elements of the work of relevance to transforming those processes for crowdsourcing.

Healthcare forms arrive at dedicated mail rooms in the USA. They are scanned and if possible OCR'd. The scanned forms are routed to the relevant onshore and/or off-shore workforces (depending on client agreements). The following steps are undertaken for all clients: 1) Data entry: keying forms not suitable for OCR by hand or checking OCR errors; 2) Verification: a second agent enters the data and is prompted if there are discrepancies; 3) Review: experienced agents check specific fields. There are a number of other activities that are carried out to ensure the smooth running of this workflow, including 1) quality audits and six sigma projects; 2) supervision and management, including shift organisation, reporting and floor-walking; 3) monitoring and managing the flow of forms through the process by production control.

3 Findings

As the work is already outsourced one might think that the challenges of distribution have been solved and to an extent this is true in the sequential steps of the workflow through which each claim passes. However, our findings show that even low skill data entry for non-OCR forms poses a number of challenges for crowdsourcing. We group these into four themes which emerged from the data: workplace ecology, skills and knowledge for data entry, making the workflow work and collaboration. For each theme we discuss the features in each setting (in office and @home) and then describe the implications for crowdsourcing. In the discussion we focus on the points which lead us to propose relationship-based crowdsourcing as an appropriate solution. Before we turn to the themes however we say a few words on the nature of this sort of BPO work and the workplace in which it is undertaken.

This work might be characterized as 'white collar factory work' and this idea deserves some unpacking. In the earlier 20th Century clerical work was considered a higher status work, offices were the places of work where the quality of environment was better than a factory, dress was more formal, and pay was generally higher. In the latter 20th Century, however, in a trend that continues, manufacturing principles were applied to administrative business processes. This involved an increasing decomposition, separation and standardization of administrative tasks within business workflows. This activity particularly impacted on the most basic tasks – such as data entry and digitization – that apparently needed little professional knowledge and skills. These tasks were hived off to be completed as low-pay, low-status, high volume piece work, especially when they became outsourced. Outsourcers, in wishing to keep their costs down, typically spend little on facilities, with the technical infrastructure being most important.

This is borne out by our experiences; the offices of the workers, despite being situated in fancily named 'tech parks' are typically basically fitted out. The work is mainly piece rate with only certain supervisor duties or management paid by the hour. Consistent high performance can earn employees a bit more than minimum wage, while poor performance over a period of months will earn employees the sack. How-

ever, compared to similar call centre environments, the workers are at least in control of their own pace of work. Indeed, some employees attain a mix of skills and knowledge that allows them to consistently perform very well, while not placing a stressful burden upon themselves. They attain high speeds, high concentration and have particular types of sensory-motor skills (e.g. being able to notice at-a-glance that something looks either correct or problematic) while also being able to attain the required knowledge. These workers are prized and often promoted. This understanding begins to give the lie to the idea that this is wholly routine work that just about any motivated person could do well. Particular skills are involved that not everybody possesses. Retention of good staff is the challenge and goal in this type of industry, because they are not always so easy to find, and recruitment and training costs are relatively high. It is important that the employer has confidence in their workforce delivering consistently. For the employee, of key importance is the ease with which they can carry out the work, the stress levels they deal with, the stability of work volume and the predictability (and control of) their wage.

3.1 Workplace Ecology

The workplace ecology for @home workers in the US and in-house workers in India is unsurprisingly quite different. Each provides various challenges and benefits for the configuration and management of the work; technical, legal, organizational and social. In India the outsourcer's offices in both locations span several floors of office blocks in technical parks. Employees work on data entry for single clients and are grouped together by client in particular seating areas or access-controlled units. The essence of the work site is a highly controlled environment, whereby security and surveillance prevails. Security is ensured technically (through systems) and through the ability to monitor staff activities through the shared physical space, which affords visual and auditory monitoring. The actual environment is standard office space with some open plan and divided areas. By contrast for people working at home in the US, their office space is naturally in their home – with each being different. Their basic requirement is to have a separate office/study space. This might be a dedicated room or a space within a room that is dedicated for work. It must be private (for the period they are working), i.e. if others are in the house with them during their working hours they must not be sharing the work space. Being home offices, the design and arrangement is up to the worker, apart from the technologies supplied by the company. They have access to the rest of their house and the home comforts that provides (for example, seasoned workers can do much of the job while listening to the radio or TV). The home environment affords greater freedom than the in-office environment. Although agents work can be monitored through the system, e.g. the supervisor can view the agents desktop at any time, there is not the co-present supervision at play in the office.

Data Security in-Office. Since health care forms contain personal information including social security numbers, names and addresses, data security is governed by

US laws, in particular HIPPA compliance. This strictly controls who may access the data and protects against unauthorized distribution and use. Data security is currently enforced through physical, contractual, social and technical means. Physically, access to the workspace is restricted, with passes required to enter the technical park and offices and security guards at various entrance points throughout. Employees also have to leave their bags and phones in lockers during the shift and pens and paper are not allowed in the office. Contractually, employees agree to the various company regulations and non-compliance has consequences including dismissal. Socially security is enforced through supervision with team leads walking the floor and answering queries, whilst keeping an eye on the workers. Workers are trained in HIPPA compliance. In addition there are a variety of technical solutions, with the systems and work-flows designed to ensure maximum security. For example, using thin clients which do not store information and have no USB ports. Plus data is stored in the US and is pulled in in batches for processing. Each batch only resides locally during processing. Agents cannot access client systems and can only see the current form.

Data Security @Home. The home workers' environment is less secure by nature, because there is no co-present supervision. However, their environment must meet certain criteria (access to broadband, a separate space to work and a conducive family environment). The workers are supplied with a secure terminal which they route through their internet connection in order to connect to the company's network and work in secure thin client mode. All staff members working from home have completed all necessary training, including HIPPA compliance, and are generally experienced and trusted members of staff. Thus, data security is currently enforced primarily through contractual and technical means. Physical and social means still play a part, but cannot be so tightly controlled and monitored, since although they have to have a separate working area, there is no way of ensuring no one else is present. Ultimately there is a reliance on trust. Supervisors cannot walk the floor and have an overview of what everyone is doing at any one time and cannot police the use of cameras and so on, but they can view an operator's desktop at any time. Working at home is an earned privilege. In order to be selected, employees need to be experienced and high performing, and informal assessments are made as to whether they are good candidates for the remote and autonomous environment. If there are any concerns about their work they can be brought back into the office environment.

Implications for Crowdsourcing. The main implication for Crowdsourcing comes from the distribution of the workers from controlled office or home environments into uncontrolled home environments or unsecured Internet cafes and their lack of a contractual relationship with the company. The control the outsourcer can exercise over the people doing the work is necessarily reduced. In effect then, security can no longer be enforced by contractual, physical and social means and solutions to data security will have to be wholly technical. One solution is to take the IBM approach [14] and split the data such that semantic information cannot be derived from the task by the worker. However, this is not appropriate for all tasks, especially for handwritten

forms which cannot undergo OCR. In this case, an alternative, and conceptually simple, solution is to separate out the parts of the form with sensitive data from those without. If a hybrid workforce were engaged i.e. a mixture of contracted and crowdsourced labour, the non-sensitive parts of the form could be directed to the crowdsourced workforce, whilst the sensitive parts could be done in-house.

3.2 Skills and Knowledge for Data Entry

In-Office. Agents have a basic level of education (graduate or undergraduate), good English language skills and typing speed. The work is known as ‘key what you see’ and is considered low skill; nonetheless the learning curve of a new entrant is around seven weeks. This is because in reality data entry is not simply ‘key what you see’ rather agents must interpret what they see according to an extensive rule set. To illustrate, the most straightforward data entry task is a standard form for claiming medical insurance (a HCFA). The name field alone has around 13 rules for how the name should be entered (for example decided which is the first name surname, middle name and so on) and it is just one of 33 fields. Each client has a variety of different forms and other documents: agents need to learn how the rules apply for all the different form types. Furthermore, task complexity is situational. It differs between and within form types. To illustrate, although data entry for correspondence is only four fields compared to 33 for HCFAs, correspondence can take considerably more time as the information may be anywhere or nowhere on the form. Identifying that information is not present often takes longer than finding and entering information because of the need to double or triple check. Agents are paid per keystroke or per form (depending on job type) with quality taken into account, thus the speed with which they can work is of immediate concern to them.

Complexity also differs across individual claims: 1) Handwriting or poorly printed forms can be difficult to read. 2) If forms do not fit the criteria for data entry because of poor scanning or because they do not comply with the rules (e.g. two patients named on one form) they need to be rejected and this takes more time because it requires double or triple checking. 3) Non-standard means non-standard; a piece of correspondence may have a cover sheet with all the required information on it or the information may not be found in the document at all. Within form complexity can only be determined on a document by document basis and would not be easy to predict in advance.

@Home. The at home workers are trained in the office and the best workers (fastest, highest quality, most reliable) are permitted to work from home if they so wish. Interestingly whilst the skills and knowledge requirements apply equally to the agents working at home, the difference in the nature of the two workforces (Indian and US) was obvious during our observations. The Indian workforce was a more transient one, with many agents leaving and needing to be replaced by newcomers. Whilst some agents had been working there for a couple of years, most were considerably newer, with all teams having a number of new starters on board. In comparison, the at home

workforce consisted of the best agents (as only they are allowed to work from home) and throughput was minimal, with agents typically having been employed with the company for a number of years (some more than 15). Thus the at home workforce, as a whole, was highly experienced and they worked with apparent ease interpreting data and spotting and correcting errors with little hesitation. Indeed it was common for agents to listen to the TV whilst working as they had become so skilled at data entry.

It is important to note we are not implying some fundamental difference in the workforces skills, indeed the more skilled Indian agents worked just as adeptly although they did not get the opportunity to listen to anything. Instead the difference is one of stability of workforce and the benefits which are accrued from that, i.e. the @Home workforce had on average more knowledgeable and skillful workers due to their greater experience. Working from home confers a number of advantages on the employees who choose it. The primary one being flexibility of working hours – workers choose at what time to work their hours each day, enabling them to fit their work around their other commitments (the vast majority of the workforce are stay at home mums). Other advantages also come from reduced supervision, a more relaxed and comfortable environment and regular and predictable income. For many the work is attractive because it provides what might be considered to be an acceptable compromise; it's not the most interesting job and many workers have the ability to do higher skilled work, but it's 'local, it fits into their lives and so is the preferred option.

Implications for Crowdsourcing. Whilst it is clear that an office environment is not essential for this work, it is also clear that both agents and employer benefit from having a well-trained experienced workforce, particularly given the strict quality requirements in the SLA between the client and the outsourcer. For crowdsourcing, we need to think of how to manage the advantage that experience bestows, plus how the learning curve might be accommodated. Approaches could include 1) reducing the required learning; e.g. splitting the form in to sections, reducing the number of rules agents must learn, contextual presentation of rules and embedding training within the dataset [10], and 2) encouraging workers to become skilled in particular areas; whilst specialization does not have to be requester specific and indeed one of the founding concepts of crowdsourcing is that it would not be, for the work requester, there is likely to be benefit from being the 'first choice' for the best workers. Being the favored work requester for the best workers, is likely to be a major drive towards relationship-based crowdsourcing (see Discussion).

The fieldwork also raises the question of how to determine incentives to ensure that all work is completed, given its situational complexity. Pay scales are currently determined by form type, but social and organisational pressures encourage agents to complete the batches they pull, whatever their complexity i.e. they are held accountable by supervisors. If crowdsourced, what's to stop agents from rejecting difficult work? The outsourcer is unlikely to move to hourly wages for crowdsourcing, given that they already use piece rate for their contracted employees. We therefore must assume such rejections will occur and identify and handle them automatically whether through dynamic incentive and reputation schemes or hybrid models of in-house and crowd workers.

It is also clear that there could be potential benefits from being a crowdsourced worker (over an in office worker) as, given a fair rate of pay, the benefits which stem from working conditions at home would carry over. At the moment, homeworking is only available within the US, largely because of data security concerns, but if a wholly technical solution to data security could be found then this work could be crowdsourced and thus done from home anywhere in the world with a good internet connection. We are not underestimating the restrictiveness of this latter condition, but as connections improve the work can be done further afield. For such work to be possible in internet cafes, it would need to be adequately compensated to pay for the use of the infrastructure and the facilities [25].

3.3 Making the Workflow Work

In-Office. The Service Level Agreement (SLA) is the contract between the out-sourcer and client, covering data security, Turnaround Time (TAT) and quality. Whilst pay is a strong motivating factor for agents, it is not enough on its own to ensure the SLA is met. Team and quality leads put in extra work to make the agents accountable for their performance. To illustrate, for one client the data entry of medical records, must be completed by 8am. A key concern of the 6am shift is to complete all remaining records before the deadline. The team leaders communicate the pressures of the queue to the agents ('agents key fast' 'everyone on medical records'). In addition once the deadline has passed the team leader calls up three or four agents at a time to discuss (rather publically) their performance. For example, "Krishna, you only did 16 medical records in an hour. What was that about?" "Ok the target of 50 is not attainable but I would expect you to be doing at least 30. You need to improve." Thus the agents are called upon to account for their performance and team leads work to make the targets achievable for *their* team. Rather than telling someone only doing 16 medical records they should be doing 50, which might seem unachievable and be demotivating, they set a more realistic target. Agents are similarly made accountable when they reject forms, as team leads monitor rejections and question agents. If a batch is rejected a few times they assign it to someone whom they insist must complete it. In this way a balance is achieved between the agents' desire to do the easy work quickly and the requirement to get all the work completed in a timely manner.

@Home. Home work is governed by the same criteria as the in-office work. However, more flexibility is built into scheduling and in general supervision is more light touch. The volumes of work were fairly dependable, however scheduling was still decided on a daily basis according to the volumes of work. The managers would do a daily calculation based on volume and staffing levels and determine a number of hours per employee, which they could complete according to their preferences (and constraints) within the day, e.g. they could work at whatever times they wanted, splitting shifts as they wished, as long as they completed their hours. Generally, the highest volumes came in on Monday and Friday with the lowest volume on Wednesday. This meant that to some extent they could organize their home life according to these

trends. Supervisors did intervene if they were concerned about worker productivity on any given day (by email, phone, pager or IMS), or there were particular pressures or requirements during a given day. The vast majority of home workers are high performers so while assessments and training and motivation is dispensed according to employee needs it is not a particularly regular or notable feature of the working day.

Production Control. On the ‘shop floor’ there is a concern for making sure the volume of jobs is adequately managed as we noted in the sections above, for example by making sure someone takes up a previously rejected batch, by pushing workers to up their speed, or getting homeworkers to change their shift hours. Production control, on the other hand, is the function specifically responsible for ensuring that the workflow flows smoothly and that jobs and batches do not get held up such that they miss their deadlines. To do this they have various tools at their disposal. A workflow management tool moves batches between several dedicated applications (on dedicated servers), including scanning and uploading, OCR and a tool for handling all the (human and machine) steps concerned with data processing. Given the shift patterns and productivity across locations, one of the tasks of production control is to see whether work should be reassigned. They assess how the processing work is unfolding across different locations, technical difficulties, skill coverage and so forth. They also check that batches are complete as they move between applications and that files or cases (or indeed batches) are not lost or held up. Hold ups can derive from technical problems, e.g. OCR application crashes or human factors, e.g. has someone on the shop floor has checked and reassigned documents in the rejected queue. The production line is constantly rolling and it constant monitoring is required to ensure nothing is missed.

Finally, for the various data entry steps (keying, verification and review) the workflow needs to be dynamically managed such that each claim passes through each step in a timely manner. Agents are assigned prioritized task permission lists, i.e. each employee has a list of tasks that they are qualified to undertake and these are prioritized. For example, an agent with OCR correction as priority 1, will be assigned an OCR correction task as their next task each time one becomes available. If no OCR tasks are available, the task rated as priority 2 *for that agent* will be assigned. Given an the unfolding work and a given set of employees, production control decide whether to manually reorder these lists and re-prioritise and re-assign tasks to ensure their timely completion. Thus the work of production control involves both technical and organizational interventions and is crucial to ensuring that the workflow unfolds in a timely and complete manner.

Implications for Crowdsourcing. The current model of work is a push model with a) the work being assigned dynamically (by the workflow tool and production control) according to agents’ skill set and queue and b) various social (and financial) pressures. In contrast crowdsourcing involves a pull model: work is self-selected by the agents who will not have the same accountability. The importance of meeting deadlines and managing the pace of the workflow for the outsourcer are clear. Any

crowdsourcing system needs to be designed to ensure that the work is completed in a timely manner to good quality, for which coordination of the workflow will be key. A major issue is that, without doubt, including crowdsourcing in the workflow will increase its complexity. There will be more places for faults to occur and documents get lost or held up. What will happen once a batch has been sent to a crowdworker? How will you know if a crash has occurred at their local PC or whether they have stopped working on the batch? How and when can you pull a batch from them and inject it back into another part of the organization/workflow? These are potentially complex issues both technically and organizationally, and there are clearly implications in terms of how relationships function and are managed with the crowd. Certainly as much of the workflow management as possible would need to be automated to deal with this increased complexity. However, it seems likely that even with the best automated tools a greater burden will be placed on production control.

In addition, it is interesting to note that even in this highly controlled piece rate environment, financial incentive is not considered enough to ensure that targets are met and all sorts of social factors come into play in determining how agents perform. It would be foolish to think that crowdsourced workers will have any less complex motivations, as indeed the research on incentive schemes has hinted at [3, 11, 17, 18]. Whilst pay is undoubtedly highly important, other factors should not be neglected (such as notions of respect, trust, fairness).

3.4 Collaborative Working

Collaboration has largely been designed out of the workflow. Claims progress through workflow steps from agent to agent and country to country automatically as each prior step is completed.

In-Office. However, in the office the work is collaborative at the claim level. That is, the routine troubles encountered in data entry are solved with colleagues or floor-walkers and it is not uncommon to see a group of two or three people around a screen discussing an issue. Typical issues include deciphering handwriting or determining which rule applies to this circumstance. All sorts of phenomena may be found on the forms (arrows showing a name has been entered the wrong way round, handwritten clarifications, etc.) not all of which are described in the rule set. Team leads walk the floor answering such queries, however they are not always immediately available and since time is money for the agents they may turn to their colleagues for help. Where the issue is deciphering handwriting, agents will typically ask their neighbour first, only turning to the team lead if this fails. For questions of rules the team lead, or other 'expert' users are the first port of call. Such troubles rarely take long to solve, but it is undoubted that this collaboration improves both speed and quality. In addition, newcomers, and agents who have fallen below a quality threshold, have their inputs checked by a supervisor before saving them in the system. This supports their learning whilst maintaining the high quality scores of the whole team.

@Home. A clear difference between at home work and in-office is that collaboration is minimized at home – there are few opportunities for at home workers to collaborate amongst themselves, share knowledge and experience. Communication is between the employees and the supervisor either in the form of 1-1 or global communication between the supervisor and the team. It is largely electronic, they use IM and email, supplemented by mobile phones and pagers, in some lines of business. The phone generally used as a means of stepping up, i.e. due to the failure of an electronic channel to achieve the desired aim due to lack of response or complexity of the communication.

Implications for Crowdsourcing. The @Home scenario might lead us to think that collaboration is not important for the work and therefore not of importance for crowdsourcing. However, it should be remembered that the homeworkers are trained in-office – so undertake their learning in the company of peers – and are the highest skilled workers. It is more likely that the by and large the crowdworkers will more closely resemble new in-office workers than the @Home workers. We might therefore want to think about enabling collaborative set ups amongst crowd members [15] enabling them to help one another, or to have sub-crowds with particular skills, such as handwriting deciphering, to which data fragments might be sent.

4 Discussion

There are a number of ethical issues which have been raised around crowdsourcing in the relatively low skill, micro-task domain, including low wages and the bias towards the work requesters' needs. However, crowdsourcing itself is not inherently unfair. Rather these issues are an artefact of AMT (and other platform) design and that much of the research so far has been driven by work requesters concerns, where the requirement to drive down costs can too easily override all other interests especially ones which are hard to operationalize such as fairness [20]. We believe that the particular constraints revealed in our domain through the ethnographic study, provide a starting point for understanding how addressing these concerns not only benefits the worker but also the organisation who is requesting the work. The idea of relationship-based crowdsourcing, described in the literature, came out of research examining the crowd workers perspective when undertaking traditional AMT tasks. When thinking about crowdsourcing this type of on-going BPO work, it can be seen to make sense from the organisational perspective as well. The fieldwork has revealed that even for this low skill piecework, it is the subtle and complex relationship between the organisation and worker that plays a major role in getting the work done such that it can meet the strict quality and turn-around time requirements.

The fieldwork highlighted a range of considerations and challenges for crowdsourcing. It is not the aim of this paper to address them all¹³, although we hope

¹³ A discussion of some of them, in particular form decomposition (to address data security and other issues) and the extreme distribution of the workflow can be found in [26].

that raising them will provide a resource for the designers of crowdsourcing platforms and plug-ins. Instead, we focus the discussion on the findings which demonstrate the subtleties of the employer-employee relationship in the work as done now and those that lead us as a consequence to propose relationship-based crowdsourcing as a model for this work. A major argument for relationship-based crowdsourcing is that both agents and employer benefit from a well-trained, experienced workforce (for both speed and quality). Even work, which on first glance, is low skill and indeed is even known as ‘key what you see’ actually has a considerable learning curve. Whilst we can certainly think of techniques to reduce this learning curve (specialisation, embedded contextual rules, training data and feedback embedded in batches), the same is likely to hold true for crowdsourcing. That is, if the organisation can attract a (relatively) well skilled crowd who spend time becoming experts in their tasks it is likely to benefit both. One way to attract such a crowd, is to become (one of) the best work requesters. Perhaps equally important is that our fieldwork reveals that even in this low skill piecework environment, pay whilst important is not enough on its own to get the work done in time to the right standards. Rather a variety of social factors come into play. There is no evidence to suggest that pay will be the only driving factor in crowdsourcing either. Indeed the mixed findings of the incentive schemes studies [3, 11, 17, 18] and the very call for relationship-based crowdsourcing [3] suggests that it would be foolish to think that it would. So if pay alone is not an adequate driver, what is? We propose that ideas of trust and fairness in the relationship between the requester and worker are likely to play an equally important part. Nonetheless accountability of workers will be reduced and any crowdsourcing system needs to include tools and technologies which enable the smooth running of the workflow with minimum human intervention, otherwise the role of production control will become overly burdensome and problematic issues will be missed. Finally, whilst data security might seem like an obviously technical question, and indeed many of the solutions in place now are technical, it too is social. As seen in the homeworking situation, acceptable security is actually reached in the relationship between the worker and the organisation. It is contractual, but it goes beyond that, stemming from the homeworkers being valued and trusted employees. However, when it comes to crowdsourcing, even with relationship-based crowdsourcing, the relationship will not be of the type or strength to enable trust on this scale, and thus data security solutions will have to be technical.

What then do we mean by relationship-based crowdsourcing in this situation? As suggested in the literature [3, 20], the idea is to create a fair and balanced crowdsourcing relationship between the organisation requesting the work and the worker doing work. Subtle notions of trust and fairness come into play. It naturally includes factors such as decent pay for the work done, paying on time, fair and transparent accepting or rejecting of work, embedded training and feedback, ways of rating work requester as well as the worker and so on. As part of the drive for higher quality on both sides of the relationship, it may also include supporting relationships between crowdworkers, enabling them to learn off one another [1] and even collaborate.

It is certainly not about trying to recreate employee-employer relations, just without the same contract and benefits. We do believe that crowdsourcing as a way of working and getting work done can confer benefits on both workers and requesters. Our fieldwork showed that working from home, at least in the US, offers certain comforts and whilst some of these will be dependent on socio-economic status, others, such as not having to migrate or travel to work, to manage your work around your other commitments and to fit work into 'spare' hours will be more universal. There are certainly many people worldwide who would see benefits from working from home. Even working from internet cafes will have some benefits (e.g. less direct surveillance) but additional economic costs. Although not suggesting that crowdsourcing is a panacea to underemployment and socio-economic development, we do believe that it offers real opportunities to bring work to various underserved populations. As crowdsourcing opportunities grow, workers should be able to pick and choose the work which is of interest to them, capitalizing on the ability to work for many different organisations depending on what work is available. Additionally the work described here has advantages over other more typically crowdsourced work, since it is a large-scale business process it can offer more predictable income to workers. For the organisations requesting work to be done, even with fair wages, they will gain cost savings from massively reduced infrastructure costs, plus the ability to easily ramp up and ramp down the workforce and if taking the relationship-based crowdsourcing approach, greater chance of accessing the best workers in the crowd, for that type of work (as with the homeworkers).

However, the current crowdsourcing model is too biased towards the organisations requirements and the worst of crowdsourcing – piecework with unfairly low remuneration - does not fit with the benefits which accrue from experience: better performance, quality and loyalty from the best workers. If crowdsourcing for this type of BPO work is to be sustainable over time, then a different approach needs to be taken. Whilst relationship-based crowdsourcing may be a somewhat optional model for some types of tasks (e.g. one offs such as image labeling) for this work, because of all the factors discussed (TAT and quality and on-going nature of work; skills and knowledge and even security) it would seem to be the most logical choice for both organisation and worker. The question remains of how to organize and implement it? Just as the organizations' perspective has been designed into current platforms such as AMT, the platforms and interfaces to support relationship-based crowdsourcing will need to embody new more equitable labour market models, which is a major research challenge for the future.

5 References

1. Thies, W., Ratan, A & Davis, J. Paid crowdsourcing as a vehicle for global development. CHI Workshop on Crowdsourcing and Human Computation, (2011).
2. Felsteiner, Alek. Working the Crowd. Employment and Labor Law in the Crowdsourcing Industry. *Berkeley Journal of Employment & Labor Law*. Vol 32:1
3. Silberman, M.S., Irani, L., Ross, J. Ethics and tactics of professional crowdwork. XRDS • Winter 2010 • Vol.17 • No.2. 39-43

4. Eagle, N. txeagle: Mobile Crowdsourcing N. Aykin (Ed.): Internationalization, Design, LNCS 5623, pp. 447–456, (2009)
5. Sachs, P. Transforming Work: Collaboration, Learning and Design. Communications of the ACM, Vol. 38, No 9, pp. 36-44. (1995).
6. Suchman, L. Making Work Visible. Communications of the ACM. 38:9. 56-ff (1995)
7. Hartswood, M. Procter, R., Rouncefield, M., Slack, R. Making a Case in Medical Work: Implications for the Electronic Medical Record. CSCW. Vol 12:3. 241-266 (2003)
8. Kochhar, S., Mazzochi, S., Paritosh, P. The Anatomy of a Large-Scale Human Computation Engine. HCOMP'10. ACM. New York. 10-17
9. Bernstein, M., Little, G., Millar, R., Hartmann, B. Ackerman, M. Karger, D., Crowell, D. Panovich, K. Soylent: A word processor with a crowd inside. UIST '10 ACM. 313-322 (2010)
10. Le, J., Edmonds, A., Hester, V., Biewald, L. Ensuring quality in crowdsourced search relevance evaluation: The effects of training question distribution. SIGIR'10. 17-20 (2010)
11. Adda, G., Sagot, B., Fort, K., Mariani, J. Crowdsourcing for Language Resource Development. 5th Language and Technology Conference. Poland (2011)
12. Bederson, B. B. and Quinn, A. J. Web workers unite! addressing challenges of online laborers. Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems (CHI EA '11). ACM, New York, 97-106. (2011)
13. Irani, L. and Silberman (M.S.) Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. To appear in CHI 2013
14. Karnin, E. D., Walach, E. and Drory, T. 2010. Crowdsourcing in the document processing practice. ICWE'10. Springer-Verlag, Berlin, Heidelberg, 408-411 (2010)
15. Rzeszutarski, J Worker Collaboration in Crowdsourcing Markets. ACM SIGCHI 2011 Workshop on Crowdsourcing and Human Computation (2011)
16. Kazai, G. An Exploration of the Influence that Task Parameters have on the Performance of Crowds. CrowdConf 2010, San Francisco, USA. (2010)
17. Mason, W., Watts, D.J. Financial Incentives and the “Performance of Crowds”. HCOMP '09. ACM, Paris, France (2009)
18. Horton, J. J., Chilton, L. B. The labor economics of paid crowdsourcing. EC '10. ACM, New York, NY, USA, 209-218 (2010)
19. Huang, E. Zhang, H. Parkes, D.C., Gajos, K.Z., Chen, Y. 2010. Toward automatic task design: a progress report. HCOMP '10. ACM, New York, NY, USA, 77-85 (2010)
20. Silberman, M.S. What's fair? Rational action and its residuals in an electronic market. Unpublished manuscript, 2010. <http://www.scribd.com/doc/86592724/Whats-Fair>
21. Khanna, S., Ratan, A., Davis, J., Thies, W. Evaluating and Improving the Usability of Mechanical Turk for Low-Income Workers in India. ACM DEV'10 (2010)
22. Bentley, R., Hughes, J.A, Randall, D., Rodden, T., Sawyer, P., Shapiro, D., Sommerville, I. Ethnographically-informed systems design for air traffic control CSCW'92. 123-129. (1992)
23. Kumar, D., Martin, D., and O'Neill, J. The times they are a-changin': mobile payments in india. CHI'11. 1413-1422 (2011)
24. Garfinkel, H., (1967) Studies in ethnomethodology. Englewood Cliffs, N.J.: Prentice-Hall
25. Wyche, S.P., Schoenebeck, S.Y., Forte, A. “Facebook is a Luxury”: An Exploratory Study of Social Media Use in Rural Kenya. In Press (CSCW 2013).
26. O'Neill, J., Roy, S., Grasso, A. and Martin, D. Form Digitization in BPO: From outsourcing to crowdsourcing? To appear in CHI 2013 (2013)