

Ignacio Garcia*

Translating in the Cloud Age: Online Marketplaces

Abstract

This article explores the influence of cloud computing on translation, including professional translation. Cloud computing reshapes the management of labour in ways that unsettle the traditional relations between managers and workers. It gives managers live control over how a project progresses, and a minute description of how each worker performs. These real-time metrics in turn facilitate rapid, semi-automated performance assessment that focuses on the desired parameters and saves sifting through résumés. Now, even larger translation projects conducted under more conventional project management protocols are being pushed to online marketplaces in the cloud and one of the reasons behind the decision may be so that the generated metadata can be analysed to profile the most efficient procedures and identify the best performing translators.

While the advantages of those online marketplaces to vendors and their clients are apparent, the model's success requires cloud entrepreneurs to sell it to potential workers as well. The marketing therefore emphasises democratised opportunity, whereby the suitably skilled (amateur or professional) can exploit their linguistic and topic knowledge by working where and when they wish, on more (and more varied) projects. Personal profiling is promoted as an added benefit that helps ensure each individual gets the most suitable offers. From a labour studies perspective however, this paper argues the cloud environment may reduce translators' bargaining power.

Keywords

cloud computing, online marketplaces, translation crowdsourcing, paid crowdsourcing, post-editing, labour studies

1. Introduction

Unbabel, an online translation marketplace established in 2013, illustrates how technology is shaping translation. The sole requirement for joining is to have knowledge of two languages and an email address. On login-in, simple source segments are tabulated on screen beside their machine-translated feeds, which the user must suitably fix (i.e. post-edit). Those who pass this point become new-fledged 'editors' whose first tasks will be pro-bono, although Unbabel promises to begin paying 'soon'. No other training is required.

In a thorough study, Codagnone et al. (2016) review emerging trends on digital labour. They distinguish between what they call *manual* (sometimes *mobile*) labour marketplaces, platforms that connect buyers and sellers for the provision of some *manual* service, usually through the smartphone, and *online* labour marketplaces, created mostly for the exchange of information tasks. Transportation network Uber is the best-known example of the first type; Unbabel is a good example of the second. Here we refer to all those platforms on which translation services are exchanged as online marketplaces.

When Unbabel first appeared and to entice translators to join, *Politics*, *Gossip* and *Sex & Relationships* were the topic choices presented. Temperamentally unsuited for painstaking erudite translations, it can well accommodate, however, the kind of social media stream that its own editors are likely helping to generate, and are thus thoroughly familiar with. That familiarity in many

* Ignacio Garcia
I.Garcia@westernsydney.edu.au

cases will extend to other areas such as current affairs, politics and opinion, which still may find their way to professional translators – who are equally welcome to participate in these projects too, of course. It is all fun: a game with a purpose (GWAP, von Ahn 2006) to help machines solve a problem – translation – that machines cannot solve (yet) on their own. Users, or ‘editors’ in this instance, can play as long or as little as they want, and on the topics of their choice. And, once they overcome the first hurdles, they can make some money on the side.

Unbabel’s business model is a prime example of the disruptive effects cloud computing may bring to translation. While cloud computing has specific meanings in computer science, we use the term here in its lay sense. Suffice to say, for the purpose of following our argument, that we access cloud computing every time the software and the data that appear on our screen is stored not on our local computer but in external massive data centres accessed through our web browser. The cloud is where some enterprises keep business information so that updates can reach the smart phones of employees 24/7. Whenever we use Facebook or Twitter, we are in the cloud, as we are whenever we upload our photos, videos or files to Instagram, YouTube, or Dropbox.

Cloud computing provides the infrastructure required for the blossoming of online marketplaces such as Unbabel. This paper represents a first attempt to study how they may affect translation, including professional translation. Based on recent literature analysing the effects of technological change on employment, Section 2 offers an overview on how these online marketplaces are reshaping the management of labour. Section 3 looks at the emergent marketplaces of this type specifically geared towards translation. Section 4 considers the range of translation projects that can be carried out on them. Section 5 discusses their influence on payment rates. Finally, section 6 addresses their influence on the working conditions of translators employed within and outside them.

2. Online marketplaces and paid crowdsourcing

Early in the nineteenth century, Taylor devised a system of splitting complex tasks into small, simple consecutive ones that could be timed and then executed in an assembly line. A century later, in the digital world, complex information tasks are also being split the same way, in a virtual assembly line. Industrial work would typically be done by employees at an appropriately equipped facility; with *information* work, the Internet provides a virtual facility that allows independent contractors to perform piece work at a venue of their choosing, whilst remaining under perfect, big-data encoded supervision.

Online marketplaces essentially implement a virtual assembly line that seeks to respond to demand as it occurs. To achieve this, labour is outsourced as needed from a large digital (rather than *geographical*) pool of contractors, under what has become known as the crowdsourcing model after the term coined by Howe (2006). Such work can be paid or unpaid. Unpaid crowdsourcing can help produce some remarkable accomplishments. Individuals will engage with it for amusement, or to gain some kind of recognition or experience. Wikipedia and Linux are often cited as successful mega-examples. It exemplifies the peer-to-peer (P2P) friendly side of what is being referred to as the sharing economy, a misnomer according to many scholars some of whom prefer to use the collaborative consumption moniker to refer to it (Benkler 2006).

Paid crowdsourcing uses the same online marketplaces, but in a peer-to-business (P2B) mode. They offer users the opportunity of activating underused time and skills to gain, for many, a significant – if not the sole – source of income as we will see below. Through online marketplaces, requesters – often called commissioners, buyers or clients – can find a massive pool of casual, inexpensive labour that can be tapped on demand. Providers – often called suppliers, workers, or occasionally in their own vernacular *turkers* (Iperiotis 2010, Berg 2016) – function as independent contractors.

Scholars classify these online marketplaces between those involved only in micro-tasking and those allowing for the carrying out of complete tasks. Of the first type, Amazon Mechanical Turk

(AMT) is perhaps the canonical example. AMT (Artificial Artificial Intelligence is its by-line) best illustrates the concept of human computation (Law/von Ahn 2011): human input is there to solve the problems that machine scripts presently cannot, and to simultaneously assist in fine-tuning those scripts so they can function more autonomously. Most work is performed by the computers, which simply export and then re-import that requiring human involvement. In AMT providers log into the system to first search for suitable HITs (Human Intelligence Tasks). Example of HITs may be “Extract purchased items from a shopping receipt” or “Input specific values displayed in the image”. At the time of writing (October 2016), AMT has half a million registered providers. Other micro-tasking platforms that boast an even higher number are data mining companies Crowdflower (five million) and Clickworker (seven hundred thousand).

The largest online marketplaces allowing for the carrying out of complete tasks are Upwork and Freelancer. Web and mobile developers, graphic designers, and writers (and within this category, a subcategory of translators) figure in both platforms as the first advertised skills, with hundreds other listed in their directories. At the time of writing (October 2016), Upwork, formerly Elance-oDesk, boasts in its ‘About Us’ five million registered requesters, twelve million freelance providers, three million jobs posted annually and one billion dollars plus worth of work done annually. Freelancer has over twenty million registered users.

While most online marketplaces are generic, some focus on specific areas such as design (i.e., 99designs) and software testing (Testbirds). Codagnone et al. (2016: 23) reviewed 39 such marketplaces categorised as manual or online, P2P or P2B, for micro-tasking or not, generic or specific. Only one of them, LingJob, is listed for translation. There are more, and will be covered in the next section.

Trust between requesters and providers is gained via sophisticated reputation systems which provide feedback on each other in the form of ratings or reviews (Thierer et al. 2016). Cognitive tasks, however, may involve on the part of the requester not only an assessment of the performance of the provider, but also of the usefulness of the work performed. The literature shows how quality assurance (QA) has often been a major concern, in particular when dealing with micro-tasks (Kaganer et al. 2013, Thierer et al. 2016). Beyond the standard reputation system first introduced by the AMT, several strategies have been developed to help: allowing requesters to seed *test* tasks into the workload and reject the work of providers who fail, first applied by Crowd-Flower; or requiring experienced workers to review tasks completed by less experienced ones, as pioneered by Clickworker (Kaganer et al. 2013).

Recent studies indicate that many workers engage with online marketplaces not as a preferred choice but for lack of alternative opportunities. Berg’s (2016) study of AMT and Crowdflower found that 33% of the sample at the time of the survey (November-December 2015) was unemployed. Some consider this work as their primary source of income - 37% in Berg’s study (9), and 63% in a study conducted by Elance-oDesk (now Upwork)(2014: 31).

In a labour environment designed around saving costs, it comes as no surprise that fair payment is a major concern. According to Marvit (2014), in the US, where 7.25 dollars is the minimum hourly wage, those working on these online marketplaces can consider themselves fortunate if they make 6 dollars per hour; most on average earn no more than 2-3, although 63 per cent of them hold a college degree, against 23 per cent among the general population. In theory, workers enjoy the flexibility to choose when and for how long they work; to make an income similar to that of the average wage, however, they must be ready to accept most of the work directed to them and ensure an untainted reputation. The P2B economy does not have time for dabblers and prefers tested and committed workers, with studies claiming, for example, for AMT, 10% of *turkers* performing 75% of completed tasks (Codagnone et al. 2016: 42). The legal safeguards that apply to off-line work (i.e. minimum wage, workers’ compensation) do not apply online (Felstiner 2011, Schmidt 2013).

There are no hard data on the volume of work conducted on these online marketplaces. A report released by PricewaterhouseCoopers in 2014 calculated the value of the whole sharing economy,

thus also including manual as well as online marketplaces, at 15 billion US dollars in 2015, forecasting 335 billion for 2025 (Vaughan/Hawthornthwaite 2013). According to economists Katz/Krueger (2016: 1) permanent full-time work as the benchmark for employment is losing ground, with alternative arrangements, defined as “temporary help agency workers, on-call workers, contract company workers, and independent contractors or free-lancers” having raised from 10.1 percent in 2005 to 15.8 percent in 2015. They also estimate the number of workers who identify customers through an online marketplace at 0.5 percent (Katz/Krueger 2016: 3). This estimate is raised by a McKinsey Global Institute report (2015: 33) to about 1% of the US labour force. The comprehensive Codagnone et al. study further rises this figure: “In countries such as the UK and the US those working regularly for digital labour markets (every week) can be estimated at least between 1% and 2% of the labour force (staying on the conservative side)” (2016: 52).

This then is the paid crowdsourcing model which is also being applied to translation, as we will now consider below.

3. Translation within online marketplaces

A favourable environment for crowdsourcing to thrive occurs when tasks can be parsed into manageable chunks. Translation is often mentioned in the general literature (e.g. Frei 2009, Kaganer et al. 2013) because it allows precisely for this approach. A project can be divided by document, section, even by sentence, with each separate component performed simultaneously by different translators.

In the translation studies sphere, two journals – *Linguistica Antverpiensia* (O’Hagan 2011) and *The Translator* (Susam-Sarajeva/Pérez-González 2012) – have devoted special issues to crowdsourcing, albeit under other titles such as community translation 2.0 (*Linguistica Antverpiensia*) or non-professional (*The Translator*) or collaborative or volunteer translation – all these tags holding content that while never identical often overlap. Most analysis within translation studies of this phenomenon, on those two publications and elsewhere, have focussed on unpaid translation crowdsourcing. We will thus not consider it here, insofar as it can be disentwined from the paid variety. Effectively, the unpaid and paid forms often coexist on the same platforms, although some only target unpaid volunteers and in others membership is meant to be restricted to paid professionals.

In the commercial sector, we can identify two broad types of online translation marketplaces. The first type simply connects translators with clients, applying to translation the P2P, eBay model. Of this type, ProZ, established in 1999 and with 784,356 registered users at time of writing (October 2016), is the undisputed leader. In somewhat distant second place comes TranslatorsCafe, with 248,863 registered users. Both ProZ and TranslatorsCafe have been studied in detail by, respectively, Risku/Dickinson (2009) and McDonough Dolmaya (2007). The cited studies mainly consider them simply as translators’ networks, namely places where translators collaborate by responding to terminology queries, sharing feedback on clients, or providing and receiving training. Yet, it should be noted that their primary role is that of labour bourses where translators bid for jobs posted by clients. What interests us here, however, is the second type, the P2B online marketplaces that follow the models reviewed in the previous section.

We referred in Section 2 to both generic and focus-specific marketplaces. Translation services do appear in the generic micro-tasking ones. A ‘translation’ search on AMT in October 2016 produced 33 HITs – most with researcher Chris Callison-Burch as a requester. He has used it extensively, but for academic rather than commercial purposes (Callison-Burch 2009; Zaidan/Callison-Burch 2011). Clickworker introduced a “New Self-Service Marketplace tool ‘Translations’” in 2016 (Clickworker 2016).

Beyond micro-tasking, the main marketplaces dealing with complete tasks Upwork and Freelancer have also, and with some prominence, Translation as a sub-skill within Writing. A ‘translation’ search on Upwork produced 3,207 ‘Translation Jobs’ and, on Freelancer 279, ‘results’.

As we saw before, apart from generic marketplaces there are also others that focus on specific fields. Amongst the 39 specific marketplaces listed on the Codagnone et al. table referred to above (2016: 23) only one, LingJob, is set ‘for writing and translation jobs’, listed under a ‘micro-tasking’ label and with 3,000 ‘contractors’. Garcia (2015) identified a dozen other translation specific ventures (which did not include Unbabel).

The volume of work these translation specific online marketplaces carry out is difficult to ascertain. Not having found a better criterion to identify the best performing ones, we have just chosen investment funding received as an indicator of potential growth. Table 1 lists, out of the fifteen identified (twelve by Garcia (2015) plus LingJob and Unbabel) the five which have received over one million US dollars in funding, as per TechCrunch’s CrunchBase. TechCrunch is an often-quoted publisher of technology industry news; the information in other columns has been gathered from their respective websites.

Marketplace	Year funded	Paid to translator (base)	Charged to client (base)	Translator pool size	Funding received (CrunchBase)
Gengo	2008	0.03	0.06	21,000	25.6 million
One Hour Translation	2008	0.054	0.087	17,000	10 million
TextMaster	2011	0.014	0.066	n/a	8.36 million
Transfluent	2011	n/a	0.12	50,000	3.29 million
Unbabel	2013	8 x hour	0.03	40,000	3 million

Table 1. Best funded online translation marketplaces (with figures in US dollars, current at October 2016)

A closer look at the consequences of the rise of these online marketplaces is warranted.

4. From micro-tasks to complex projects

As seen in section 2, online marketplaces – as with any other mass online enterprise including translation marketplaces, and whether paid or unpaid – require requesters, workers (those who would do the translation in our case), and a platform to connect both and manage the projects. Some follow the micro-tasking AMT model (i.e., Unbabel), and its adoption is on occasion even acknowledged as a badge of pride (see Toto 2010 for Gengo), while others work with whole projects only (i.e., One Hour Translation – OHT). In this second case, these translation marketplaces act as conventional agencies, but dispense with most of the administrative overheads by simply setting fixed prices per word for clients and for translators.

In a crowdsourcing project the source text is normally presented within some kind of computer-assisted translation (CAT) tool, accompanied or not by relevant memories and glossaries. It might also implement machine translation (MT), either to post-edit no-match segments in the CAT tool, or as the sole help, as in Unbabel. When it comes to using post-edited MT, most platforms are happy to offer clients the choice; only One Hour Translation (OHT) explicitly forbids the practice. Accordingly, the source text can be in the form of an entire document or even a massive complex project shared between many translators, as in any other translation/localisation project, but can also take the form of a micro-task, i.e. as a sentence, which, although the default unit when using a CAT tool, never has been done in standard commercial translation.

Micro-tasks in particular can be well suited for unpaid work. The prototypical example of a successful micro-tasked project is the translation of Facebook’s user interface by its own users. While the micro task level is the default system for unpaid work (see also subtitling Coursera’s videos with Transifex), some paid work is also done that way, such as through Unbabel.

Most paid (and some unpaid) crowdsourcing is handled at the document level. In a typical paid setting, such as OHT, a document will be made available to a pre-screened group of translators; the first one to access it locks it away exclusively as his or her job for the duration of the permitted time to complete the translation.

As we saw in Section 2, QA is of special concern on online marketplaces since jobs are allocated to a pool of users, not to a particular user previously vetted through a bidding process. QA can be even more important with translation tasks as in some cases the client will not be in a position to assess the output. Tests are conducted on entry, before allowing the translator to get paid for tasks. Often this testing will, simply, as explained at the beginning in the Unbabel case, take the form of unpaid tasks checked by established translators. Afterwards, QA relies on customer feedback or, when charged separately, on revision by the best performing translators, i.e. those boasting most words translated for the most demanding clients and with the least number of complaints.

Nowadays, even larger projects of the kind once performed under conventional project management protocols can be pushed to online marketplaces. What is more, conventional translation agencies, Translated.net, Elanex and Straker, for example, are starting to follow basically the same business model: no bidding required for approved translators; once a job is picked by one it will be locked out of the reach of the others. What conventional agencies do which marketplaces do not is to allow them to add their desired rates upon registering for work. Even major players Lionbridge (with onDemand) and SDL (with Language Cloud) include some sort of paid crowdsourcing among their services. Apart from increasing choice for clients, there are additional benefits to the provider, such as, by forcing everyone involved to work always online, *live* control of the progress of the project, and useful metadata to profile translators (and clients and platforms) in regard to costs and quality.

This will be explored further below.

5. Online marketplaces and payment trends for translators

Most professional translators – 75.9 per cent according to the 5th ATA T&I Services Survey conducted in 2014 (ATA 2016: 8), 74 per cent in Pym et al. (2012-2013: 3) - work as independent contractors. The standard form of payment to self-employed translators is a monetary rate paid per source word. If operating as a typical sole trader, the translator sets his or her prices, subject to some negotiation with the requester (be it a direct client or agency). Price setting and expectations will depend on language pair (interlinked with supply and demand), volume of work, turnaround times, complexity of the subject matter, the desired quality and inherent consequences of errors, and other variables such as experience, specialisation and so forth. However, as we will explore below, things can operate rather differently in the online marketplaces.

There is a perception that, since the start of the global financial crisis (and the emergence of online translation marketplaces), pay rates for translation have fallen (see, for example, Muzii 2013). The best hard numbers we have on translators' earnings are those of the latest ATA's Translation and Interpreting Compensation surveys. The results of its Fourth survey, conducted in 2006, showed personal gross compensation from T&I language services and language related activities at 60,423 US dollars (Six 2008: 13). In its Fifth, in 2014, that figure was 52,323, although, on an upbeat note, its executive summary also reads: "Nearly half of the respondents reported that their T&I language services compensation increased in 2014 compared to 2013" (ATA 2016: 13).

In discussing pricing and remuneration, we firstly require some kind of benchmark as a departure point for comparisons. ProZ gives some data on the rates charged (or, at least, requested) by freelancers, based on an aggregate view of the figures they have entered in their profiles – which, being used to filter job notifications, can be considered fairly representative. Thus, for October 2016, for English to Spanish (13,875 translators having provided their data) we find that the standard average rate amongst ProZ members is 11 cents, and the minimum, 8 cents; the aver-

age rate per hour, 34.86 dollars and the minimum 26.21 (Proz 2016). TranslatorsCafe (2016) has a similar feature and – again, for English to Spanish – yields similar figures, while both sets are also consistent with those of a survey held in 2011 by the Chartered Institute of Linguists (CIoL) in the UK (Gardam et al. 2012).

From a labour perspective it is clearly of interest to convert payment per word into hourly wages, so accordingly we need to quantify average translator productivity. Wide individual variation is to be expected. An informal survey conducted by Allen (2004) reported averages of 275-300 words. Calculating averages is not easy: do we ask translators, and assume their self-reporting to be accurate? Do we ask agencies, or end clients as Allen also did? Data can easily be gained through short trials, but can such bursts of productivity be maintained throughout an entire day or week? Moreover, Allen found that those who reported speeds in words-per-hour gave a higher average figure than those who estimated their output over a day, week or month, or in pages. For certification purposes, bodies such as the American Translators Association (ATA) have since the eighties standardised on a 250-word passage as being suitable for candidates to complete in one hour. We will adopt this figure as a reasonable conservative estimate for sustained average output.

While it may be difficult to agree on a figure that represents good or fair remuneration for translation, the yardstick of hourly minimum wage figures makes it somewhat easier to delineate what would be marginal or precarious. Thus, in order to make the US hourly minimum wage of 7.25 dollars, a translator working at 250 words per hour would need to receive 3 cents per word. Does this mean that TextMaster's base offer of 1.4 cents per word – equating to less than half the US subsistence target – is exploitative and should be illegal? Or should notions of fairness also consider the conditions and opportunities afforded to workers in the countries where they effectively reside and work? In Mexico, the minimum hourly wage of US 0.62 equates to some 44 words at TextMaster rates.¹ In that case the chance of earning 3.5 dollars per hour may even justify TextMaster's (2016) claim of having set its 1.4 cent minimum precisely “to guarantee stable pay for the [translator] and to avoid competition for the lowest offer”. The same cloud technology that enables online marketplaces also fosters global outsourcing, pitting translators in developed nations against peers in countries with costs of living that are significantly lower overall.

Next is the matter of achieving full rather than partial or sporadic employment. Working at bargain rates does not necessarily mean one can expect eight billable hours per day, every day. Priority will be given to translators who receive good feedback from customers. As noted in section 2, for the self-employed, the only sure way to earn consistent money is to monitor job offers continually and be prepared to work long hours, and work hard. With payment by the word and rates trending down, the only response to maintain one's income level is to do more, and faster.

Given these pressures, translators might ultimately feel it is preferable to secure an hourly rate than to continue performing what is, effectively, piece work. Again, in the cloud environment, that conclusion requires wariness: Straker (2016) for example extols the virtues of hourly pricing, but it is clients who reap the principal benefit from what the firm calls a natural selection environment:

Not all translators are created equal – some produce better quality translations at faster speeds, and we use natural selection to pick the best. Only 2 out of every 10 translators we test meet our per hour benchmarks. [. . .] There are only 2 ways to influence translator cost – pay translators less (the crowd-sourcing approach) or make them more efficient. We can have a vastly greater impact on cost by linking the time it takes to complete a translation and improving productivity than focusing on driving down word rates.

Interestingly, where those per-hour speed benchmarks are concerned, Straker asserts that “[t]he industry average is only 360 and in over 20 language pairs Straker translators work at 687 words per hour”. Both quoted figures are notably higher than the 250 we have used above, and while Straker avers it does not focus on driving down word rates, it is difficult not to see this as an au-

¹ Minimum wages figures taken from OECD data quoted from https://en.wikipedia.org/wiki/Minimum_wage

automatic collateral effect of having to work at two or three times the industry average speed in order to secure work at a flat hourly rate. It is tempting to ponder whether translators who are that fast might be better served by working at the kind of per-word rates surveyed over ProZ, TranslatorsCafe or the CIOlL. However, speculation is difficult in the absence of details concerning actual hourly pay, if and how pay grades are linked to language pair, specialisation, performance, and – harking to our earlier comment on bursts versus sustained output – translator churn through burnout/attrition, and so forth.

When all work is carried out in the cloud, whether in online marketplaces or through conventional agencies, an enticing array of monitoring possibilities becomes available, even beyond tracking who is working fast or not, and stimulating translator efficiency.

We will delve into them in the next section.

6. Caught in the cloud?

In the mid-nineties, CAT tools revolutionised the translation industry. Two decades later, cloud computing is revolutionising it again. At present, translators can work within essentially three main modes:

- (a) Off-grid: what we might still consider as traditional or pre-industrial translation practices based around translator autonomy, involving mainly word processing and a local environment and resources. Present examples might include transcreation; official paper document translations requiring certification; or high-security applications on segregated computer systems.
- (b) CAT-enabled: translators exploit their local translation memory tools and resources, and use the Internet for client access, peer networking, and research; however, this mode can be problematic for harmonising ongoing or collaborative projects due to tool compatibility issues, and can permit use of unauthorised resources such as non-vetted glossaries or free online MT.
- (c) Cloud: by tightly circumscribing the working environment (CAT tool, memories, glossaries, MT), this mode gives clients and agencies all the advantages (and more) of working with CAT-enabled translators, with few of the disadvantages.

A notable feature of CAT-enabled translation, which has been the paradigm since the 1990s, is that it ranges worker autonomy against industry standardisation and control. It is therefore easy to imagine the emerging cloud paradigm soon becoming the dominant one, since it resolves an inherently unstable dynamic in favour of those who ultimately pay for the required service – namely clients and agencies. To use a Marxian term, it means extracting more surplus value out of translators.

The increasing value chain along the three modes is apparent. Off-grid translators simply deliver a target text. CAT-enabled translators deliver a target text *plus* reusable language data in the form of bilingual language databases that permit subsequent leveraging and discounting by clients for exact and *fuzzy* matches. Cloud translators deliver the target text *plus* the language data *plus* metadata – of great enterprise value because it allows managers to profile translators by productivity (speed and time), and by quality (from edits made at revision stage, or from feedback by clients or end-users).

Admittedly, given time and effort, off-grid translations could be mined for language data (by aligning legacy files), and CAT translations can yield limited metadata via translator identification and time stamps on translation units in the memory. However, this approach entails costs that might not be entirely recouped, and neither is as rich nor as easy as that of having translators work in the cloud, where the metadata is acquired *live* and, once the scripts are in place, interpreted and

actioned immediately. If clients want to know who the fastest translators in their pools are, and some do as we have seen, they must ensure all translation work is done in a browser.

A pre-condition for cloud translation to succeed is to have translators working online at all times. That should not be such a hurdle, since translators already rely on the web for term lookups and information searches. Similarly, powerful web-based CAT tools may ultimately seem for translators a more pragmatic choice compared to desktop software with its licences and updates.

Most brands now offer a web interface, and all can provide web access if absolutely required. When Memsource, one of the best known, writes in its blog that over 50 percent of translations from English to Spanish (and that applies to many other language pairs) had been done on its tool “combining translation memory with a machine translation engine”, it is relying on the relevant metadata. That same metadata also allows Memsource to perform “translation memory analysis” and “post-editing analysis” to calculate the edit distance between the target text and the memory or machine-seeded segments. This information may be of limited interest to translators, but the client can use it to apply discounts and calculate productivity; so the same blog can announce that post-editing “boost[s] throughput from 2,000 to 4,000-8,000 words per day” (Memsource 2016). Features that provide such productivity and edit distance calculations are becoming standard for established tools (XTM, memoQ) and for emerging ones (MateCat, SmartCat). Productivity tells the client how much time each translator takes in translating a given number of words; edit distance, how well the machine translation engine being used performs.

Web-based CAT entails a balance of pros and cons that will not equally suit all potential parties. Clients certainly have control over which tools translators use, and on which settings, thus harmonising post-processing tasks; and can identify which translators work fastest and best. However, they must develop or purchase the necessary expertise to gather, store and exploit the metadata. For translators, the tools can be comparatively inexpensive, if not free, and require no installation or updating; the downside, apart from potential system lag times, is relinquishment of control, and being monitored and mined by clients for metadata (if that is a concern).

Once accustomed to an online work interface, and comfortable with (or unmindful of) what happens with their metadata, translators are primed in turn for online marketplaces. As with web-based CAT, the online marketplaces have their own mix of pros and cons. Owners and clients can enjoy automation of most administrative tasks and elimination of tendering: no more wasting time on quotes or purchase orders; lower costs, so that they can translate more words for the same money; and access to the type of metadata that, for example, allows Straker to screen out eighty percent of aspirants to its pool. However, there is potential for high churn rate among translators, and quality is highly process-dependent. The attractions to professional translators are not so clear – they may resort to it for some extra cash in their down-time, which is how Gengo advertises its platform to them; it is amateurs who may see the most benefit in crowdsourcing, perhaps as a rung on the ladder to professionalization. The downside is low rates (e.g. 1.4 cent per word), loss of independence, and living an on-call existence.

Prima facie, harnessing the power of metadata cannot but contribute to making the translation industry more productive. Which is precisely the aim of the Translation Automation User Society (TAUS), its declared mission “to enable better translation through innovation and automation”. In fulfilment of this mission, TAUS released its Quality Dashboard in June 2015. Provided that projects are conducted fully in the cloud, metadata can be entered and aggregated so that managers can track and benchmark the productivity of projects by language pair, technology (using post-editing or crowdsourcing, or a combination of both), sector, and domain. It will also allow translator profiling by speed, quality, and translation type (TAUS 2015).

TAUS also proposes an *efficiency score* – calculating not just words per hour but also edits per hour – as a more sophisticated performance parameter, and one by which competent translators can distinguish themselves. But can the approach offer long-term benefits to translators? Or could it risk unfairly ignoring important qualities that are not so easily measured? Unlike an ordinary productivity score, efficiency calculations require all translation tasks to be performed in

the cloud. At the lowest end of the spectrum, translators earning an income at 1.4 cent per word through cloud labour presumably have no other choice; if the efficiency score were to be widely adopted, even elite performers could be forced toward the cloud, as the only real place to obtain such a ranking.

The cloud propitiates measurement, and the metadata it generates can be aggregated and compared, and evidence-based decisions made. Cloud technology offers perfect transparency for clients, who can access all the data on all translators while the translators can access only their own data. Thus, while tools such as MateCat may allow an individual translator to access to her own productivity data, a project manager has access to the data of all translators in his pool: he can compare the performance of each with that of her peers. In the TAUS Quality Dashboard, translators have free access to aggregated data on productivity and language pair, but only managers can access itemised data that tells them how each project and each translator is performing. In general, managers know more about the value of individual translators than do the translators themselves. The difference here is that managers know vast amounts, and that many decisions are going to be made by automated processes.

7. Conclusion

With some exceptions, translators have thus far been reluctant to adopt web-based CAT tools. If the Yahoo support group number of members is a reliable indicator, freelancers still prefer to buy a paid version of Wordfast rather than using Wordfast Anywhere – the freely available web-based version that comes with all advanced features, including access to Wordfast’s Very Large TM and predictive typing for sub-segmental reuse: the Wordfast Classic list counts with 6931 members against 594 in the Wordfast Anywhere (October 2016). This might be partly attributed to practical issues such as internet connectivity and download speeds, which are steadily improving of course.

It might also be attributed to a not entirely unjustified wariness. The potential of cloud computing for social good is enormous, but so is its potential for surveillance and exploitation.

Translation buyers are becoming aware of the advantages lying within the cloud. Crowdsourcing, argue those on the management side, allows translators to work when they want and on the projects they choose. Profiling just helps optimise translators’ time by allocating them the right tasks, while increased translator productivity will compensate for discounts applied and falling rates. As translators may see it, up to now there were standard rates to pay for translation, usually by the word, in some cases by the hour and, through the tendering process, there was some room for the individual translator to negotiate better ones. How good a translator performed would be known by the client only after having worked with him or her on several projects. In the online marketplace model, however, the elimination of tendering limits the translator’s room for negotiation while, through the metadata, the client gains a clearer picture of each translator’s worth.

Mainstream and emerging players alike will attempt what Straker already says it does, to pick just the best and fastest and to pay them at standard hourly rates, albeit couched in more inclusive and upbeat language. Unbabel, the gamified marketplace with which we started the article, has already done so, in fact, even more aggressively: “Only 15% of applicants pass our quality evaluation tests” (Unbabel 2016). As quoted above, the pass percentage for Straker was 20.

Ultimately, the cloud scenario is looming as a major fork in the course of translation as a – hitherto – expert human activity. This may mean that, except for the translation of sensitive material the quality of which must be ensured *before* publication, translators may become interchangeable and expendable, existing simply as data sets. Can a place be ensured for those who wish to assert their worth as something more? Whatever the case, the modern penchant for big data, and its practical expression found in cloud technology, already pose major questions for educators and students about preparing for and assessing the future of translation as a profession and career. Meanwhile, we may envision more amateur aspirants being attracted to cloud translation as a pathway

to online work, hoping to achieve top-ranked status and continue translating, albeit not always for fun nor even entirely for profit.

References

- Allen, Jeff 2004: Translation speed versus content management. In *Multilingual* 62, 50-53.
- American Translation Association (ATA) 2016 The ATA Translation and Interpreting Services Survey, Fifth Edition. American Translators Association.
- Benkler, Yochai 2006: *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven and London: Yale University Press.
- Berg, Janine 2016: Income Security in the On-Demand Economy: Findings and Policy Lessons from a Survey of Crowdworkers. In *Comparative Labor Law & Policy Journal* 37(3), 1-26.
- Callison-Burch, Chris 2009: Fast, cheap, and creative: evaluating translation quality using Amazon's Mechanical Turk. In *Proceedings of EMNLP '09*, 286-295.
- Clickworker 2016: About us [online]. <https://www.clickworker.com/about-us/> (accessed 1 October 2016).
- Codagnone, Cristiano/Abadie, Fabienne/Biagi, Federico 2016: *The Future of Work in the 'Sharing Economy'. Market Efficiency and Equitable Opportunities or Unfair Precarisation?* Institute for Prospective Technological Studies, JRC Science for Policy Report EUR 27913 EN, doi:10.2791/431485.
- Elance-oDesk 2014: Annual Impact Report [online]. <https://blog-static.odesk.com/content/Elance-oDeskAnnualImpactReport2014.pdf> (accessed 10 October 2016).
- Felstiner, Alek 2011: Working the Crowd: Employment and Labor Law in the Crowdsourcing Industry. In *Berkeley Journal of Employment and Labor Law* 32 (1), 143-204.
- Frei, Brent 2009: Paid Crowdsourcing. Current State & Progress toward Mainstream Business Use. White Paper [online]. <http://www.3rd-idea.com/wp-content/uploads/2009/09/Paid-Crowdsourcing-Sept-2009-PreRelease-Version-1-Smartsheet.pdf> (accessed 10 October 2016).
- Garcia, Ignacio 2015: Cloud marketplaces: Procurement of translators in the age of social media. In *The Journal of Specialised Translation* 23, 18-38.
- Gardam, John/Jorg, Udo/Pocock, Christine 2012: Your rates revealed. In *The Linguist* 51 (2), 12-15.
- Howe, Jeff 2006: The rise of crowdsourcing. In *Wired*, June 14.
- Ipeirotis, Panos 2010: Demographics of Mechanical Turk. In *Working Paper CeDER-10-01*.
- Kaganer, Evgeny/Carmel, Erran/Hirschheim, Rudy/Olsen, Timothy 2013: Managing the Human Cloud. In *MIT Sloan Management Review* 54 (2), 189-200.
- Katz, Lawrence F./Krueger, Alan B. 2016: The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015 [online]. https://krueger.princeton.edu/sites/default/files/akrueger/files/katz_krueger_cws_-_march_29_20165.pdf (accessed 10 October 2016).
- Law, Edith/von Ahn, Luis 2011: *Human Computation. Synthesis Lectures on Artificial Intelligence and Machine Learning*. San Rafael CA: Morgan & Claypool.
- McKinsey Global Institute 2015: *A Labor Market that Works: Connecting Talent with Opportunity in the Digital Age. Highlights*. Washington: McKinsey & Company.
- Marvit, Moshe Z. 2014: How Crowdworkers Became the Ghosts in the Digital Machine. In *The Nation*, Feb 5.
- McDonough Dolmaya, Julie 2007: How Do Language Professionals Organize Themselves? An Overview of Translation Networks. In *Meta: Translators' Journal*, 52 (4), 793-815.
- Memsources 2016: Post-editing in Memsources [online]. <http://blog.memsources.com/post-editing-in-memsources/> (accessed 10 October 2016).
- Muzii, Luigi 2013: Is Quality Under Pressure? Or Is Translation? In *TMT Conference* [online]. <http://www.slideshare.net/muzii/pressure-text> (accessed 1 October 2016).
- O'Hagan, Minako 2011: (ed.) Community Translation 2.0. *Linguistica Antverpiensia New Series – Themes in Translation Studies* 10.
- Proz 2016: Average rates reported by language pair [online]. <http://search.proz.com/employers/rates> (accessed 10 October 2016).
- Pym, Anthony/Grin, François/Sfreddo, Claudio/Chan, Andy L. J. 2012-2013: *The status of the translation profession in the European Union*. Luxembourg: European Commission.
- Risku, Hanna/Dickinson, Angela 2009: Translators as Networkers: The Role of Virtual Communities. In *Hermes – Journal of Language and Communication Studies* 42, 49-70.

- Schmidt, Florian Alexander 2013: The Good, The Bad and the Ugly: Why Crowdsourcing Needs Ethics. In *Cloud and Green Computing (CGC), Third International Conference*, 531-535.
- Six, Shawn 2008: Summary of ATA's Latest Translation and Interpreting Compensation Survey. In *The ATA Chronicle*, February, 12-15.
- Susam-Sarajeva, Sebnem/Pérez-González, Luis 2012: (eds.) Non-professionals Translating and Interpreting. Participatory and Engaged Perspectives. In *The Translator* 18 (2).
- Straker 2016: Per-hour Translation Infographics [online]. <https://www.strakertranslations.com/translation-pricing/#per-hour> (accessed 10 October 2016).
- TAUS 2015: Taus releases the Quality Dashboard [online]. <https://www.taus.net/think-tank/news/press-release/taus-releases-the-quality-dashboard> (accessed 10 October 2016).
- TextMaster 2016: General Questions [online]. <https://www.textmaster.com/client-faq> (accessed 10 October 2016).
- Thierer, Adam/Koopman, Christopher/Hobson, Anne/Kuiper, Chris 2016: How the Internet, the Sharing Economy, and Reputational Feedback Mechanisms Solve the “Lemons Problem”. In *University of Miami Law Review* 70, 830-878
- Toto, Serkan 2010: MyGengo Is Mechanical Turk For Translations. In *Techcrunch*, January 11.
- TranslatorsCafe 2016: Translation Rates Statistics [online]. <http://www.translatorscafe.com/cafe/en-US/CommunityRates.asp> (accessed 10 October 2016).
- Unbabel 2016: Meet the Unbabel community [online]. <https://unbabel.com/customer-service/> (accessed 10 October 2016).
- von Ahn, Luis 2006: Games with a Purpose. In *IEEE Computer Magazine*, June, 96-98.
- Zaidan, Omar F./Callison-Burch, Chris 2011: Crowdsourcing Translation: Professional Quality from Non-Professionals. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language technologies*, 1220-1229.