JoSTrans
The Journal of Specialised Translation

www.jostrans.org · ISSN: 1740-367X

Moorkens, J. & Lewis, D. (2019). Research Questions and a Proposal for the Future Governance of Translation Data. *The Journal of Specialised Translation, 32*, 2-25.
https://doi.org/10.26034/cm.jostrans.2019.154

Research Questions and a Proposal for the Future Governance of Translation Data

Joss Moorkens, Dublin City University
Dave Lewis, ADAPT Centre, Trinity College Dublin

ABSTRACT

This article seeks to assess the impact of data-driven methods of machine translation (MT), not just on translators, but more broadly on industry and society. We consider translation as a shared resource that is the result of producers and consumers who share the overall objective of providing sustainable access to reliable multilingual information. This view of translation as a shared knowledge resource is based on the reliance of machine learning-based translation techniques on the pooling and leverage of parallel text in the form of existing translations. These techniques build on the well-established practice of leveraging such resources as translation memories. We use the institutional analysis and development (IAD) framework to assist with analysis of the dynamic situations emerging through the use of MT, and consider why changes, such a shared ownership model, would ultimately benefit all stakeholders. We finally suggest some research questions that might help to assess sustainability for translation as a resource, an industry, and an occupation, such as 'what would happen if we left translation to machines?', 'why do translators not act collectively', and 'could principal and end customer buy-in be heightened by quantifying the threat to sustainability of access to translations?'.

KEYWORDS

Translation data, machine translation, digital commons, open data, machine learning, translation technology, artificial intelligence.

1. Introduction

There is growing consensus that artificial intelligence (AI) will have a major impact on work and employment in high-skill tasks previously thought immune from automation. The processes by which technologies impact economic activity involve a complex interplay between economic return, work organisation, and skills management (Valenduc and Vendramin 2016: 45) that make them difficult to analyse and complicate planning of adaptation strategies. A recent report by Manyika *et al.* (2017: 2) highlights that while approximately 50% of current work activities are technically automatable, the actual impact will depend on broader shifts in growth, investment, marketisation of currently unpaid roles, and workforce demographics. However, in advanced economies (such as those in western Europe), the study predicts that up to 33% of workers may need to switch occupational categories as a result of automation in the period up to 2030, with areas involving data processing, data collection, and predictable physical tasks. Further research is therefore urgently required to establish governance structures that allow companies and policy makers to balance the rapid worker adaptation required by AI automation with the broader economic and societal benefits this may bring.

The translation industry is at the forefront of addressing these changes as AI techniques applied to machine translation (MT) become an increasing factor in production of many commercial translations. Several studies have begun to address the impact of MT on the work processes of translators, however the complexities of the broader industrial and societal impact of MT require further study. To assess this wider impact we consider translation as a shared knowledge resource that is the result of producers and consumers who share the overall objective of providing sustainable access to reliable multilingual information. This view of translation as a shared knowledge resource is based on the reliance of AI-based translation techniques on the pooling and leverage of parallel text representing existing translations, building on the well-established practice of leveraging resources such as translation memories (TMs).

In this article, we use the institutional analysis and development (IAD) framework (Hess and Ostrom 2005, Ostrom 2011) to assist with analysis of the dynamic situations emerging through the use of MT. The IAD framework is a tool that assists in designing the institutional structures for the sustainable governance of complex shared resources by helping to build shared understanding of resource dynamics, the diverse interests within the community of users and producers, and the costs and benefits of different governance structures. Hess and Ostrom use the term 'commons' for these shared resources and explain that the term 'information or knowledge commons' is often used for digital information that is "enclosed, commodified, and overpatented" (2007: 4). Bollier notes that current copyright and trademark law originated "in a more static technological and economic context" and reconsideration of ownership of shared resources will grow as "more commerce, academic research, and ordinary social life migrate to Internet platforms" (2007:35)[1]. The IAD framework is centred on the 'action arena' (see Figure 1) where actors make decisions within an action situation constrained by the nature of the resource being managed, the community that use and produce the resource, and the existing rules in operation within the community. This results in specific patterns of interaction between community actors and resources that in turn lead to outcomes that can be assessed under evaluation criteria. We describe the characteristics of these resources in the following section and thereafter look at other categories as defined by the framework.

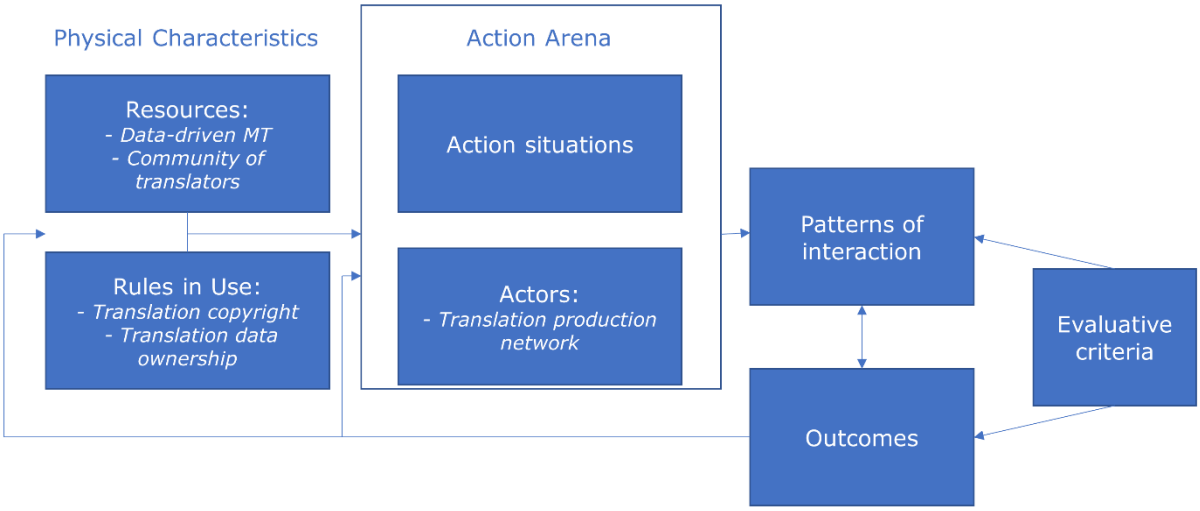

Figure 1. Institutional analysis and development framework (after Hess and Ostrom 2005)

2. Resource Characteristics

To contextualise the analysis of the action arena for decision making for automated translation we break down the resource characteristics into: the core physical attributes of sustainable access to reliable multilingual information in the form of translations; the attributes of the users, providers and policy makers related to providing such translations; and the current rules in use in terms of legal constraints, operational policies and community norms. To constrain the complexity of our analysis, we focus primarily on the context in place in the European Union. While this still represents a large market for translation with 24 official EU supported languages and over 250 different languages in use (Ethnologue 2018), it serves to restrict the analysis of the legal rules in force to a jurisdiction where common rules spanning different language markets are possible, and in some cases already in place.

2.1. Physical Characteristics

New use cases are continually being found for MT output, particularly in scenarios in which the translated content is highly perishable or where erroneous translations are considered to carry little risk (Way 2018: 160). Since the early 2000s, this MT tends to be data-driven (particularly where major languages are concerned), trained and tuned on previous human translations, rather than following any prescriptive rules. The reliance on human translation is becoming more acute as neural machine translation (NMT), trained entirely on bilingual data and with greater data requirements than statistical MT, has gradually become the preferred paradigm (Castilho *et al.* 2017, Forcada 2017). Webcrawling for parallel text is a cost-effective option for data-gathering, particularly for low-resource languages and for domain adaptation (Toral *et al.* 2017). However, neural networks ideally require large amounts of high-quality training data, making human translation data a valuable resource. This may be illustrated by the high

valuation of a company that holds "over 8.7 billion parallel segments of high-quality human TM data which covers over 222 languages" and subsequent stories of a "boom in language data" (Diño 2017a, 2018). The oft-repeated metaphors of data as oil and data as gold, of course, suggest that such data is naturally-occurring, rather than human-created (Dean 2016). The application of machine learning to translation in NMT, where the computer automatically extracts patterns from data rather than being explicitly programmed, mirrors the growth of machine learning in many different areas.

2.2. Community Attributes

The community of users and producers involved in providing sustainable access to reliable translations can be broken down into those involved in commercial translation and those involved in non-commercial translation. Commercially, translations are sought by organisations wishing to convert content they hold in one language into one or more other languages. Such content may be paid-for content or support content. Paid-for content is characterised by the consumer paying to access the content in their chosen language, e.g. in the form of a book, magazine, report, or online content. Support content is usually not paid for directly by its consumer, but rather is provided (and translated) in support of a product being marketed or introduced into different language markets, e.g. product instructions, manuals or software user interfaces. Commercial translation for large institutions is sometimes performed in-house by staff translators, typically when a reliable flow of content requiring translation with a consistent level of quality is required, e.g. in large national or supranational organisations such as the European Commission[2] or the US Department of Defence. However, there is a global trend towards contingent or freelance work, a working status that is particularly common for translators, leaving many disempowered with regard to working conditions and repurposing of translated work[3].

Specialised translators who had invested in TM tools expecting to gain a competitive advantage found instead that expected throughput had increased as word rates decreased. Many translators began to resent new technologies that they considered to be imposed on them (Cadwell *et al*. 2017, Christensen and Schjoldager 2016, Way 2018): first TM with its associated discounts, and more recently MT post-editing, which requires them to accept further discounted rates to fix "fundamental linguistic errors that a trained human translator would rarely generate" (O'Brien 2012: 110). It is rarely made explicit by companies and research groups that specialise in MT that human translation is its necessary basis, with the focus instead on new and better ways to process this trove of pre-existing 'big data' (Kenny 2011).

The gradual limitation of the translator's role has undermined their ability to conform to the ethical code of their profession (Chesterman 2001: 142)

by reducing the translation process to a series of "language-replacement exercises" (Pym 2003: 7). In this way, rather than only eliminating work that is "mechanical and routine" (Kay 1980: 1, LeBlanc 2013), translation fits with the narrative developed by Frischmann and Selinger (2018), in which the encroachment of machine learning in many occupations risks reducing the role of the human to one that is comparable with a simple machine. Furthermore, as the profession has moved from analogue to digital, translators' powerlessness is reflected in continued data dispossession, common for many knowledge workers, and largely unaffected by legal constraints (Huws 2014: 4). This data dispossession, in which digital data is "expropriated from us and put to alien use" (Dean 2016: np), is a wider problem within the digital domain, and may be carried out overtly (when a person is told that their call will be recorded) or covertly, when rights of data ownership are reassigned based on a person's assent to a clause buried within data-use policies (Reijers *et al.* 2016). This has, for example, allowed social network data to be combined, shared, and mined without user knowledge, data that has been shown to "provide psychological information about attitudes and behaviours, including health-related behaviours" (Young 2014: 601).

The importance of translation globally means that any negative perceptions of translation as a career may have a sizeable negative economic impact. A change to the current copyright impasse for translators (as described in Section 3) would make translation appear more attractive as a career option but would also require cooperation from translation and client companies, changes and adherence to standards, and engagement from translators. We consider the ethical, economic, and legal pressures to make these changes, and their possible ramifications in other industries affected by machine learning.

3. **Rules in Use**

Rules define what an actor or community may, must, or must not do and in practice — certainly in translation — there is often a discrepancy between formal legal rules and the less formal rules in use (Cole 2017: 838). Translators in some jurisdictions may have grounds to claim copyright over translated texts as creators of derivative or adapted work, depending on the perceived originality of the translation and subject to the rights of the author of the original work (Troussel and Debussche 2014: 98). Translators retain further copyright over their translation, granting them control and potential royalties when this is used as the basis for a new translation. There are also rights conferred as creators and maintainers of a database, including maintenance work on TMs and terminological or quality annotation on the source or target. Data protection regulations, such as the EU General Data Protection Regulation 2016/679 (GDPR) confer some additional rights within their applicable jurisdiction to control the use of personal data, which may cover translator-identifying metadata or techniques to re-identify translators' attribution from anonymised translations (Lewis *et al*. 2017).

In this section, we consider some metadata options that could allow translator preferences with regard to how their work is repurposed to be respected, and some options for collective action, both bottom-up and top-down, that may result from translator cooperation or may be lobbied for by translator organisations.

3.1. Copyright and Data Ownership for Translators

Early adopters of TM tools when they were first introduced in the early 1990s had the opportunity to increase productivity once past the initial learning curve. However, as they came to be more widely used, the power shifted to (especially larger) clients as discounts based on fuzzy match resemblances began to be applied by translation buyers with very little grounds for negotiation on the part of individual translators (García 2006). By 2012, Marshman found that roughly 50% of 250 language professionals surveyed said that the use of technology impacted their control over remuneration, with the reported impact most notable "among experienced freelance translators" (2012: 8).

The TM file as a by-product of a translation effort is usually, by precedent, handed over to the hiring agency or end client after a translation job, whether or not ownership has been specified in translation project contracts. In practice, some translators keep copies of this data for reuse. In the absence of an enforceable contractual agreement regarding ownership of this parallel linguistic data that Smith (2008: 23) has called the "translation family jewels," their legal status is often ill-defined in national and international laws and is thus unclear (Troussel and Debussche 2014: 128).

A translator may have grounds to claim copyright of a translated text to which they have made an original contribution as a derivative or adapted work, depending on the contract and the perceived degree of originality, and subject to rights of the original author as recognised by the Berne Convention (World Intellectual Property Organisation 1979, Troussel and Debussche 2014)[4]. When the author or translator has worked as part of a crowdsourced effort, their input and therefore degree of ownership is less clear. The creator or maintainer of a database, such as a TM file, is considered to have rights to data ownership, depending on the originality involved with its creation, in countries such as France and Germany, with the option of asserting further *sui generis* rights if the creator has demonstrated a substantial investment in obtaining, verifying, or presenting that database (Troussel and Debussche 2014: 19).

The situation with regard to copyright issues internationally appears fluid. Copyright laws have changed over time in many jurisdictions, and within the EU are further complicated by a number of EU-level directives that are intended as a step to harmonise copyright (Margoni 2016), and to address

new issues raised by unexpected technological advances, permitting mass digitisation of books, for example[5]. Periodic public consultations have taken place, most recently in 2013, which look to address issues with text and data mining, and user-generated content, and have been followed up with the establishment of European Commission working groups. The somewhat fluid state of copyright law has not appeared to affect the reality for ownership of translation data, which (to our knowledge) has never been legally tested.

Although these potentially conflicting claims of copyright for written or translated material are currently ignored, they may at some point create difficulties for enterprises offering MT and, to a lesser extent, collectives sharing MT, as there are so many competing claims on the training data that it becomes impossible to exploit it. For translators, the retasking of TM as parallel text for training MT engines is a particular concern (Moorkens and O'Brien 2017: 121). The leverage of TMs from previous translations is well understood by translators. They understand the role it plays in avoiding unnecessary retranslation of similar segments and the resulting role played by matching scores between available TMs and the source of incoming translation projects in price discounting. The practice of individual translators retaining TMs from previous projects independently of vendors is widespread, as modern desktop translation tools allow them to use these as reminders of previous translations and for term concordancing. These are useful features for individual translators even if the level of useful TM matching leverage with a personal TM is low. These practices seem to indicate a tacit approval by translators of the use of TM leverage. There seems to be an appreciation by some translators that they benefit from the prior work of other translators captured in a shared TM in the same way that other translators will benefit from their work in future (Gough 2011: 206). We can assume there is a degree of collegiality at play among those translators who favour resource sharing, since even if translators producing and consuming translation via TM may not know each other's identities directly, the poor level of TM leverage across domains or client content types means benefitting translators can be assumed to be working in the same broad domain as those who produced the content.

The use of TMs for MT training erodes this traditional acceptance of TM leverage, since translators perceive that the resulting MT system can be used by vendors and clients for translation in very different domains. Furthermore, MT is seen to be useful in classes of translation tasks where little or no translator input is required or feasible (see Way 2018), contributing to the perception that the spread of MT endangers the livelihood of translators.

Although TM data interoperability standards, such as Translation Memory eXchange (TMX) and XML Localization Interchange File Format (XLIFF) enable translator provenance to be recorded, such metadata is typically

stripped from TMs before being returned to clients or used between projects by vendors. In NMT training, data is broken down to the level of words, subwords, or even characters (Lee *et al.* 2017), so that the input of any individual translator is unrecognisable and their contribution to a system trained with very large amounts of data is untraceable. The traditional acceptance of TM leverage means that, outside of a specific translation project, the tracking of the provenance of individual translation to specific translators is not practised, and is not strongly demanded by translators. However, the loss of this provenance data means that there is no way for individual translator contributions to large aggregated language resources to be identified or credited. Further, to do so might create a data security risk as a breach of identifiable data would fall foul of the GDPR, and hence translators are denied the opportunity to specify any preferences on the rights they wish to declare over the use of TMs they return to vendors and clients. However, the provision of a trusted third party to administer the assertion of translation preferences while respecting GDPR concerns would seem feasible with some level of community investment.

4. Action Arena

The action arena proposed here is the principal–agent dyad in a translation production network, a hierarchical relationship between two parties as described by Abdallah (2010) in which a principal (the translation buyer, an intermediary or end client) subcontracts or delegates work to an agent (the translator or post-editor). The resource that the agent brings to this action situation is their translation competences (as identified by EMT Expert Group 2017 and PACTE Group 2017). They carry out their tasks using translational resources that may have been provided by the principal (in which case a discount may be expected), or from publicly available online resources, or from dictionaries and resources that they own. The hierarchy of the relationship is enforced by the agent's working status, with most translators working on a freelance basis, and especially for those who work on a freelance basis for a single employer (Moorkens 2017: 466). Freelance entrepreneurs' relationships in this arena are necessarily transactional, as highlighted by Chul Han (2017), having little or no capacity for relationships that may be enjoyed for the happiness of company alone. The agent can be considered to work within the limits of bounded rationality, their decision-making limited due to their constrained power within the dyad, exemplified by the use of interfaces created with usability considered of secondary importance (Lagoudaki 2008, Moorkens and O'Brien 2017), and often a lack of control or access to language resource data (and the means of processing that data, such as via MT engines).

4.1. Patterns of Interaction

The vendor model for translation means that there is no clear career path for many translators. Moorkens and O'Brien (2017) found that, among their

survey cohort, older translators were more likely to work on a freelance basis or to run their own translation company. As AI makes inroads into the industry, more translators are likely to work with MT and, as mentioned, other roles such as project management may be at least partly automated. Automation has heretofore primarily affected so-called 'bulk translation,' wherein the acceptance of *good enough* or *fit-for-purpose* translation for economic reasons (for example in light post-editing) means a concomitant acceptance of translations that "fall short of excellence, even though they may actually do an excellent job of fulfilling their intended purpose" (Bowker, *forthcoming*). The progress of MT research and pressure on cost, have combined to press MT into action for an increasing number of text types and market sectors that would previously have not been considered appropriate (Moorkens 2017). AI for translation, as with other industries, relies on human training data, yet at present data ownership is almost always held at the principal side of the dyad, as detailed in the Section 3. We consider that these rules may be outdated in terms of payment and copyright.

Several authors have suggested per-hour payment for translation (Abdallah 2010, Jenner 2015), and as MT of varying quality is added to translation workflows, this could be a fairer method of recompense than the more common per-word model, especially considering that MT quality estimation cannot yet accurately predict effort or pricing (Specia and Shah 2018). Hourly payments are not yet the norm, although 50% of 124 respondents to Blagodarna's international survey (2018) report often or always charging for PE work by the hour. The boost in MT quality (and hype; Castilho *et al*. 2018) since the advent of NMT has meant that more LSPs are offering MT workflows (Lommel and DePalma 2016), using data gathered as part of an outdated mutual understanding between translators and translation buyers. The current copyright precedent as detailed in Section 3 is based on leveraging at the segment level within TM tools, however NMT systems repurpose bilingual data at the word and subword or character level. The rules-in-use do not seem likely to attract new entrants to the occupation of translator, despite the current demand for qualified translators and the stipulation that college-educated translators must be employed to achieve the ISO EN-15038 process standard (Jiménez-Crespo 2018), due to the lack of a clear career path, relatively disempowered status of new entrants, and digital dispossession (see Section 2.2). The disparity between MT services and the human-translated corpora might further alienate translators from their work, and add to existing mistrust in MT and in data sharing.

For translation buyers, the situation as it stands has the benefit that, while there are significant one-off costs to creating translations, costs for secondary uses are trivial (Pollock 2018: 97). For occupational translators, however, in relinquishing copyright they also lose any control over future exploitation of the translation data they have generated. When laws of copyright are effectively bypassed in content collection, curation, and

exploitation, this permits resource holders to retain data at a cost (i.e. a fair share of downstream earnings) to disempowered human writers and translators, and also at a cost to end-users of translated content.

The growing application of machine learning to automated translation however serves to widen the opportunities to leverage translators' output at low cost in a wider range of applications, thereby reducing the scope for collective future earnings of translations. Various suggestions put forward for accruing royalties for translation may be difficult to implement in practice: levels of individual input for payment are impossible to identify, as mentioned previously; work practices for translation make it difficult to make payments to the community, as many translators work on part-time basis, often not directly translating but performing an ancillary role.

Despite these difficulties, one option for sustainable control of translation resources may be an application within the translation industry of the proposal of property rights for the digital knowledge commons by Hess and Ostrom (2005). Rushkoff (2019: 63) characterises the commons as "not a winner-takes-all economy, but an all-take-the-winnings economy." This application would necessitate active and developing management within a translation community of "successful collective action and self-governing behaviours; trust and reciprocity" (Hess and Ostrom 2005: 4), a task made more difficult by the geographically-dispersed and digitally-mediated nature of the contemporary translation industry[6]. Hess and Ostrom (2005) identify seven levels of rights that could be attributed to users of the digital commons: (i) the right of access to view, (ii) contribution, (iii) extraction, (iv) the ability to edit or remove one's own contributions, (v) management/participation in regulation of the resource, (vi) the right to determine who has each level of rights, and (vii) the right to sell or lease rights and access. Rules with regard to administration of these rights need to be "flexible and adaptable in order to create effective institutional design and ensure resource sustainability" (Hess and Ostrom 2005: 53). We contend that this proposal would increase the sustainability of the translation industry, and thus would ultimately benefit the principal who wishes to provide continued access to reliable multilingual information.

In the next section, we consider the situation at present using various evaluation criteria, and predict possible outcomes were the current situation with regard to language resource ownership to remain unchanged.

4.2. Outcomes

Ostrom (2011: 16) suggests outcome evaluation using following criteria:

- economic efficiency ("the magnitude of net benefits associated with an allocation of resource"),
- equity through fiscal equivalence ("that those who benefit from a service should bear the burden of financing that service"),

- redistributional equity (redistribution of resources to "benefit particularly needy groups"),
- accountability,
- conformance to values of local actors (how "outcomes fit the values of those involved"), and
- sustainability.

Not all production networks are the same, and thus the balance of economic benefits may change from one dyad to another. What we can say is that the net benefits associated with the allocation of resources tends to accrue to the party in possession of those resources, most usually the principal. The resource-holding principal can choose to reallocate those resources or not, can request work at discount based on leverage from the resources, and can use those resources as a basis for machine learning. Discounts achieved by leveraging these resources may be used to gain the principal a competitive advantage in the marketplace, continuing the ongoing depression of rates paid to translators (Do Carmo 2018). As noted by a reviewer of this article, the pooling of translation resources is not new. However, existing large-scale pooled resources (such as MyMemory and TAUS Data Cloud) are principal-controlled.

While the agent may have some legal rights, as a single freelance agent there is a strong disincentive to asserting those rights against an employer, as to do so may affect the agent's reputation and payment (Smith 2008), a particular risk for translators who work directly with a single agency. The change to a translation commons would mean benefits could accrue to the agent, which could also positively affect retail translation prices as, while the translation industry is growing year on year (DePalma *et al.* 2017), language service providers commonly operate on tight margins. The availability of free online MT (predicated on its inexpensive secondary use of human translations) and the commodification of translation inherent in the vendor model of employment means that translation at all stages in the network may not be highly valued. A community-owned resource would not have the financial wherewithal to compete with multinational corporations on MT research, but the digital commons could grow to an extent whereby the access to high-quality data would be worth the corporation paying for, adding royalties to the community and increasing the value of translation more generally.

The uneven balance between agent and principal in a hierarchical action situation means that at present there is unlikely to be fiscal equivalence between effort put towards that action and the benefits derived by both parties, nor is there likely to be equivalence in an ability to pay for maintenance of resource systems. While the principal retains ownership to translation resources, they also finance and manage secure storage and maintenance of those resources, along with allocation to agents within a production network. Management of resources is not trivial; translations are

commonly created asynchronously, may be appropriate for repurposing only in specific customer projects or scenarios, and are stored in a variety of proprietary and open file and exchange formats. The initial costs of creating a large-scale digital commons would have to be borne by the community or supported via donation, before paid access could begin to support the ongoing costs. In the longer term, community-controlled translation resource management should produce an equalising effect in benefits to both principal and agent.

Translators who do not work directly for end clients rarely have access to the centre of the production network (Abdallah 2010: 20), therefore project managers often prove to be valuable advocates. Some organisations have attempted to save costs to themselves and clients by automating parts of this role, as exemplified by Lionbridge onDemand, SDL Language Cloud[7]. The varied file formats provided for translation by many clients currently require the intervention of a localisation engineer, but the development efforts being put into automation of project management may make it more common in the coming years. Abdallah (2010) has highlighted the importance of trust in relationships within production networks, ideally based on mutuality rather than dependency. If these relationships become purely transactional rather than trust-based, this risks reducing the social quality, and thus the production quality of a translation project (Abdallah 2014: 11). The lack of supportive co-workers has been associated with "lower rates of job satisfaction, life satisfaction, and subjective well-being" (Moorkens 2017: 467). A weakening of communication within the network will also disimprove accountability between groups and individuals working in silos. This is already a problem, as identified by one respondent to the UK Translator Survey (DGT/CIOL/ITI 2017: 24), who believes that there needs to be a "greater sense of support and teamwork" between translators and agencies. Suggestions to use a distributed ledger system may improve accountability, but not support or teamwork. The move to a digital commons for translation, however, would serve to empower translators as owners of resources, and through the quality of additional transactions in the discovery, access, pooling and maintenance of resources. This would not necessarily mean that translator agents are less isolated within the production network, but all parties throughout the network would benefit from the accountability of having a centrally managed resource of attributed and quality-controlled translation data.

As freelance agents, translators who do not work for direct clients are not in a strong position with regard to redistributional equity, as payment comes from the opposite end of the production network, and a weakening of links of trust in these networks is likely to disimprove this position. On the other hand, some established translators are struggling to keep up with demand (Hippe-Heisler 2018), and in that case may set prices that are constant irrespective of the number of intermediaries (from client to multi-language vendor to single language vendor to freelancer, possibly with engineering

and project management costs along the way) in the production network. Translators with experience and social capital are thus more likely to achieve redistributional equity within a principal/agent dyad. Translator agents can build their social capital by conforming to the values of local actors — that is, by completing jobs on time, especially rush jobs, by producing work that meets quality expectations of the principal and end client, and by communicating cordially and responsively, in a way that is meaningful for the receiver. Control of translation resources would further boost this social capital, and improve redistributional equity. As currently, the principal could build trust by giving repeat work leading to a long-term partnership, paying promptly, communicating expectations, and advocating for the agent within the network.

There are several reasons for concern about the sustainability of (especially larger) translation production networks with the current outcomes. The principal/agent dyad as described within translation production networks work for some translators but not others. Amidst predictions of increased employment for translators (Diño 2017b, Bureau of Labor Statistics 2018), some translators are choosing to leave the industry due to job insecurity and pressures on price. Two participants from Moorkens (*forthcoming*) responded that they were moving out of translation, with one writing that the "isolation of self-employed working from home [was] literally killing me, as was RSI [repetitive strain injury] and stress related to tight deadlines and the cut-throat market." One respondent to the UK Translator Survey (DGT/CIOL/ITI 2017: 22) "can't stand freelancing any more so I am looking for work in a different field." Abdallah (2010) found that four of eight interview participants chose to exit the industry rather than tolerate unsatisfactory conditions. Sustainability would appear to be one of the most important reasons for moving to a community-owned digital commons model, circumventing any future legal challenge to the current ambiguous copyright status quo, providing ownership rights and ongoing royalties to the translation community, and offering a resource-anchored hub for improved mutual professional support activities.

How this digital commons would be controlled and created from the outset is less clear than the potential benefits. The most obvious groups to begin the process of building this resource for the translation community[8] are FIT (the International Federation of Translators), who represent 80,000 members of national translator organisations. Requirements for membership, however, differ from one country to another, and not all countries are currently represented. Some start-up organisations (such as Guildtree) aim to define the boundaries of translation as a profession[9] and to lobby for control of resources and redistribution using blockchain-based decentralised data exchange. As noted by Pym *et al.* (2012: 3), the fragmented nature of translation work and employment statuses does not encourage the prospect of professionalisation, and the move to an online translation marketplace has led to instances of online identity theft and false

claims of education and experience that make status signals difficult to verify.

5. Research Questions

In his article on translation and posthumanism, O'Thomas proposes that technology could be used "to create narrow, unimaginative translations that move towards retrogressive practices of standard equivalencies, and it can create virtual networks of translators who operate in a global sphere of mutual cooperation" (2017: 297). This article attempts to provide some justification for the latter prognosis, and to suggest some useful research questions that may lead in that direction. The concept of *homo economicus* assumes that humans act with perfect rationality. For the principal and agent in our translation production dyad, what would perfect rationality look like?

What would happen if we left translation to the machines?
According to press releases, Microsoft NMT has "achieved parity" with human English-Chinese translation (Hassan *et al.* 2018), SDL has "cracked" Russian to English NMT (SDL 2018). As some translation buyers are concerned primarily with cutting the cost of translation, would it be perfectly rational to move entirely to MT? In turning the argument about whether translators should use MT on its head, we must consider the risks inherent in the publication of raw MT. Canfora and Ottman (2018) identify reputational, legal, and communicational risks caused by mistranslation to the client company, along with the risk of harms to the public, such as unusable products, injury, or death. Regular media stories detail unexpected and embarrassing MT errors that could be avoided by following post-editing guidelines that specify the removal of "offensive, inappropriate or culturally unacceptable content" (TAUS 2010: online). More broadly, we have seen how any application of machine learning will reveal bias in its training data or algorithms (Feldman *et al*. 2015, O'Neil 2016, Corbett-Davies *et al*. 2017). The translations would lose the individual variation found in human translation processes (Koehn 2009, Taivalkoski-Shilov 2018), and the addition of new data as provided by human translators in response to natural evolution in language use. What would be the effect, particularly for literary, game, or audiovisual translation, of the constraint to expression that would occur if humans were left out of the loop?

Why don't translators act collectively for their mutual benefit?
Considering that translators' isolation is disempowering, would it not be perfectly rational to work together? Well, they do to an extent. We have previously mentioned translator organisations and FIT, the international umbrella group. There are regular regional translator conferences such as the annual ATA (American Translators Association) and ELIA (European Language Industry Association) Together events. There are sporadic examples of collective action, such as the collective bargaining agreement for Medicaid translators in Washington, USA, and the Italian Literary

Translators' agreement with publishing houses[10]. There are several factors that impede more collective action among translators. As freelance, part-time, and contingent workers, they are geographically dispersed, and Berardi believes that "precariousness and globalisation have jeopardised the social solidarity that is necessary for a long-lasting process of autonomous organisation" of digitally-mediated workers (2015: 329). Many translators do not feel it ideologically or legally appropriate to act collectively. Several participants in the survey by Moorkens (*forthcoming*) believe that the market should set the price, and that "freelancers cannot fix their prices under competition law". Would collective action on the part of translators cause harm to the public and to the industry, or be legally or ethically objectionable?[11]

Are translators exiting the industry? Are students being put off studying translation?

While there are examples of small-scale studies that show translators are leaving the occupation of translation (Abdallah 2010, Pidchamook 2018), there are no large-scale studies that might help to estimate numbers and reasons. Similarly, numbers fluctuate in translation programmes globally, but it would be very difficult to identify a single variable with a causal relationship with this fluctuation. A longitudinal study of student translators in a number of programmes and the rates of attrition could contribute here. Even without this information, the social quality of a translation production network appears to be undervalued. Lagoudaki identified how translators' requirements were (in 2008) of subordinate need to industry demands when developing translation interfaces, and ten years later translators found the same dissatisfactions with their translation tools (Moorkens and O'Brien 2017). Without a direct revenue benefit, the arguments to tailor tools or workflows for translators tend not to filter through the production network.

Empirical research on those exiting the industry would help to answer the question of what would happen if the industry is left to continue as it is. One survey after another shows translator dissatisfaction with payment rates, agency, considering that their role has value and meaning (the most important factor in job satisfaction, according to Krifa/Happiness Research Institute (2017)), and with work at the segment level within current tools (Lagoudaki 2008, Kelly *et al*. 2012, DGT/CIOL/ITI 2017, Moorkens *forthcoming*). The International Labour Organisation's substantive elements of Decent Work include adequate earnings and productive work, social security, stable and secure work, and social dialogue that includes workers' and employers' representation (Ferraro *et al*. 2015). It is arguable whether these elements are part of the translation occupation for many freelancers, and whether their absence makes the occupation less appealing for new entrants. If translators are exiting, yet demand for translators is growing, sustainability must be a concern. It must also be noted that, while many translators struggle with marketing and negotiation skills (DGT/CIOL/ITI 2017), many translators are thriving as freelance entrepreneurs, having

established a niche in translation, and marketing their work successfully via online and offline media. However, if more and more translated work is being repurposed in more ways using machine learning, is it fair or sustainable that the benefits of this leveraging accrue only to the principal side of the dyad, ultimately being stockpiled by large organisations that repurpose it in the form of free or cheap bulk online translation that devalues and commodifies translation, rather than being shared with the community?

Could principal and end customer buy-in be heightened by quantifying the threat to sustainability of access to translations rather than solely seeking to rebalance the translator-principal dyad?
Beyond the translation industry, it is worth considering the societal harms to continuing on the present trajectory. If the industry is unsustainable, will it become more difficult to get trustworthy translations, for which there has been human oversight to minimise risk, and for which there is accountability? In the case of injury or death caused by an incorrect translation, who will accept liability? At present, neither copyright nor liability may be attributable to a non-human (Bridy 2012).

6. Conclusions

As an industry at the forefront of the AI revolution, with a liberalised market and scalable workforce, translation is an important area in which to consider not only the threat (or otherwise) of automation on employment, but also the potential harms to translators and the public of unregulated and inconsistent rules for data ownership and reuse. We propose an application of the digital commons model, as described by Hess and Ostrom (2005), as one suggestion for community data ownership, and consider the outcomes of taking on this model using the IAD framework.

We believe that a move to a community-owned and managed digital commons would ultimately benefit the public and translators by making the industry more sustainable than at present, and argue that, whether or not this proposal is considered, there are several reasons for changing the current copyright and data ownership conditions. These conditions are based on segment-level data sharing and leveraging for a TM environment, whereas bilingual data may now be used for MT and beyond, and are less appropriate to sustain the occupation of translation and to minimise the potential risks and harms to translators and the public.

Acknowledgements

This work was supported by the ADAPT Centre for Digital Content Technology, funded under the SFI Research Centres Programme (Grant 13/RC/2106) and co-funded under the European Regional Development Fund.

References

- **Abdallah, Kristiina** (2010). "Translator's agency in production networks." Tuija Kinnunen and Kaisa Koskinen (eds). *Translator's Agency*. Tampere: Tampere University Press, 11–46.

- **Abdallah, Kristiina** (2014). "Social Quality: Key to Collective Problem Solving in Translation Production Networks." Gunta Ločmele and Andrejs Veisbergs (eds). *Translation, Quality, Costs*. Riga: University of Latvia Press, 5–18.

- **Berardi, Franco 'Bifo'** (2015). *And: Phenomenology of the end*. South Pasadena, CA: Semiotext(e).

- **Blagodarna, Olena** (2018). *Enhancement of Post-Editing Performance: Introducing Machine Translation Post-Editing in Translator Training*. Doctoral dissertation. Universitat Autònoma de Barcelona.

- **Bollier, David** (2007). "The Growth of the Commons Paradigm." Charlotte Hess and Elinor Ostrom (eds). *Understanding Knowledge as a Commons: From Theory to Practice*. Boston: MIT Press, 27–40.

- **Bowker, Lynne** (*forthcoming*). "Fit-for-purpose translation." To appear in Minako O'Hagan (ed). *The Routledge Handbook of Translation and Technology*. Abingdon: Routledge.

- **Bridy, Annemarie** (2012). "Coding Creativity: Copyright and the Artificially Intelligent Author." *Stanford Technology Law Review* 5, 1-28.

- **Bureau of Labor Statistics** (2018). *Occupational outlook handbook, 2017-18 edition: Interpreters and translators*. Washington, DC: U.S. Department of Labor. http://www.bls.gov/ooh/media-and-communication/interpreters-and-translators.htm (consulted 30.05.2018).

- **Cadwell, Patrick, Sharon O'Brien and Carlos S. C. Teixeira** (2018). "Resistance and accommodation: factors for the (non-) adoption of machine translation among professional translators." *Perspectives: Studies in Translation Theory and Practice* 26(3), 301–321.

- **Canfora, Carmen and Angelika Ottmann** (2018). "Of ostriches, pyramids, and Swiss cheese: Risks in safety-critical translations." *Translation Spaces* 7(2), 167–201.

- **Castilho, Sheila *et al*.** (2017). "Is Neural Machine Translation the New State-of-the-Art?" *Prague Bulletin of Mathematical Linguistics* 108, 109–120.

- **Castilho, Sheila *et al*.** (2018). "Evaluating MT for Massive Open Online Courses: A multifaceted comparison between PBSMT and NMT systems." *Machine Translation* 32(3), 255–278.

- **Chesterman, Andrew** (2001). "Proposal for a Hieronymic Oath." *The Translator* 7(2), 139–154.

- **Christensen, Tina Paulsen and Anne Schjoldager** (2016). "Computer-aided translation tools – the uptake and use by Danish translation service providers." *The Journal of Specialised Translation* 25, 89–105.

- **Chul Han, Byun** (2017). *Psychopolitics: Neoliberalism and New Technologies of Power*. Translated by Erik Butler. London: Verso Futures.

- **Cole, Daniel H.** (2017). "Laws, norms, and the Institutional Analysis and Development framework." *Journal of Institutional Economics* 13(4), 829–847.

- **Corbett-Davies, Sam *et al.*** (2017) "Algorithmic decision making and the cost of fairness." Paper presented at the *23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Halifax, Canada, 13-17 August 2017), 797-806.

- **DGT/CIOL/ITI** (2017). *UK Translator Survey: Final report*. Luxembourg: European Commission Directorate-General for Translation.

- **Dean, Jodi** (2016). "Big Data: Accumulation and Enclosure." *Theory & Event* 19(3). https://muse.jhu.edu/article/623988 (consulted 26.09.2018).

- **DePalma, Donald A. *et al*.** (2017). "The Language Services Market: 2017." Boston: Common Sense Advisory.

- **Diño, Gino** (2017a). "Sogou Invests USD 6-7m in Shanghai-Based UTH in Scramble for Translation Data." *Slator*, Aug. 22, 2017. https://slator.com/ma-and-funding/sogou-invests-usd-6-7m-shanghai-based-uth-scramble-translation-data/ (consulted 09.10.2018).

- **Diño, Gino** (2017b). "Interpreters and Translators Top Emerging Career in 2017 According to UC San Diego." *Slator*, Dec. 5, 2017. https://slator.com/industry-news/interpreters-translators-top-emerging-career-2017-according-uc-san-diego/ (consulted 10.12.2017).

- **Diño, Gino** (2018). "Korean Voice Assistant Highlights Tech's Insatiable Hunger for Language Data." *Slator*, Oct. 8, 2018. https://slator.com/technology/korean-voice-assistant-highlights-techs-insatiable-hunger-for-language-data/ (consulted 09.10.2018).

- **Do Carmo, Félix** (2018) "Does 'time is money' apply to translation?" Paper presented at the *KäTU Symposium XVI* (Turku, Finland, 13-14 April 2018).

- **Drugan, Joanna, Ingemar Strandvik and Erkka Vuorinen** (2018). "Translation quality, quality management and agency." Joss Moorkens *et al*. (eds). *Translation Quality Assessment: From principles to practice.* Berlin: Springer, 39–68.

- **Du, Jinhua *et al*.** (2015). "Domain Adaptation for Social Localisation-based SMT: A Case Study Using the Trommons Platform." Paper presented at *WPTP 2015: Post-editing Technology and Practice* (Miami, FL, 3 November 2015), 57–65.

- **EMT Expert Group** (2017) "European Master's in Translation Competence Framework 2017." Brussels: European Master's in Translation (EMT). https://ec.europa.eu/info/sites/info/files/emt_competence_fwk_2017_en_web.pdf (consulted 09.02.2018).

- **Ethnologue** (2018). "Languages of the World: Europe." https://www.ethnologue.com/region/Europe (consulted 12.06.2018).

- **European Trade Union Confederation** (2018). "Council of Europe recognises bargaining rights of self-employed: Now EU should follow." https://www.etuc.org/en/pressrelease/council-europe-recognises-bargaining-rights-self-employed-now-eu-should-follow (consulted 07.01.2019).

- **Feldman, Michael *et al.*** (2015) "Certifying and Removing Disparate Impact." Paper presented at the *21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Sydney, Australia, 16 July 2015), 259-268.

- **Ferraro, Tania, Leonor Pais and Nuno Rebelo Dos Santos** (2015). "Decent work: An aim for all made by all." *International Journal of Social Sciences* 4(3), 30–42.

- **Forcada, Mikel** (2017). "Making sense of neural machine translation." *Translation Spaces* 6(2), 291–309.

- **Frischmann, Brett and Evan Selinger** (2018). *Re-Engineering Humanity*. Cambridge: Cambridge University Press.

- **García, Ignacio** (2006). "Translation memories: A blessing or a curse?" Anthony Pym, Alexander Perekrestenko and Bram Starink (eds). *Translation technology and its teaching*. Tarragona: Universitat Rovira i Virgili, 97–105.

- **Gough, Joanna** (2011). "An empirical study of professional translators' attitudes, use and awareness of Web 2.0 technologies, and implications for the adoption of emerging technologies and trends." *Linguistica Antverpiensia NS - TTS* 10, 195-217.

- **Hassan, Hany *et al.*** (2018). "Achieving Human Parity on Automatic Chinese to English News Translation." Redmond: Microsoft AI & Research paper. https://www.microsoft.com/en-us/research/uploads/prod/2018/03/final-achieving-human.pdf (consulted 20.05.2018).

- **Hess, Charlotte and Elinor Ostrom** (2005). "A Framework for Analyzing the Knowledge Commons." *Libraries' and Librarians' Publications* 21. https://surface.syr.edu/sul/21 (consulted 10.05.2018).

- **Hess, Charlotte and Elinor Ostrom** (2007). "Introduction: An Overview of the Knowledge Commons." Charlotte Hess and Elinor Ostrom (eds). *Understanding Knowledge as a Commons: From Theory to Practice*. Boston: MIT Press, 3–26.

- **Hippe-Heisler, Elisabeth** (2018). "Machine translation in human translation workflows." https://hippe-heisler.blogspot.com/2018/01/machine-translation-in-human.html (consulted 20.05.2018).

- **Huws, Ursula** (2014). *Labor in the global digital economy: The cybertariat comes of age*. New York: Monthly Review Press.

- **Jenner, Judy** (2015). "Should We Charge for Translation Services by the Hour?" *The ATA Chronicle* 31. http://bit.ly/Jenner-Chronicle (consulted 20.05.2018).

- **Jiménez-Crespo, Miguel** (2018). "Crowdsourcing and Translation Quality: Novel Approaches in the Language Industry and Translation Studies." Joss Moorkens *et al*. (eds). *Translation Quality Assessment: From principles to practice.* Berlin: Springer, 69–94.

- **Kay, Martin** (1980). "The proper place of men and machines in language translation." Palo Alto, CA: Xerox Palo Alto Research Center. [Reprinted in: Machine Translation (1997) 12(1-2), 3–23]. http://www.mt-archive.info/70/Kay-1980.pdf (consulted 20.05.2018).

- **Kelly, Nataly, Donald A. DePalma and Vijayalaxmi Hegde** (2012). "Voices from the Freelance Translator Community." Boston: Common Sense Advisory.

- **Kenny, Dorothy** (2011). "The ethics of machine translation." Paper presented at the *New Zealand Society of Translators and Interpreters Annual Conference 2011* (Auckland, New Zealand, 4-5 June 2011).

- **Koehn, Philipp** (2009). *Statistical Machine Translation*. Cambridge: Cambridge University Press.

- **Koskinen, Kaisa and Helle V. Dam** (2016). "Academic boundary work and the translation profession: insiders, outsiders and (assumed) boundaries." *The Journal of Specialised Translation* 25, 254–267.

- **Krifa/Happiness Research Institute** (2017). "Job satisfaction index 2017." Copenhagen: Happiness Research Institute.

- **Lagoudaki, Elina** (2008). *Expanding the Possibilities of Translation Memory Systems: From the Translator's Wishlist to the Developer's Design*. Doctoral dissertation. Imperial College, London. https://spiral.imperial.ac.uk:8443/handle/10044/1/7879 (consulted 10.05.2018).

- **LeBlanc, Mathieu** (2013). "Translators on translation memory (TM): Results of an ethnographic study in three translation services and agencies." *Translation & Interpreting* 5(2), 1–13. http://trans-int.org/index.php/transint/article/view/228/134 (consulted 26.09.2018).

- **Lee, Jason, Kyunghyun Cho and Thomas Hofmann** (2017). "Fully Character-Level Neural Machine Translation without Explicit Segmentation." *Transactions of the Association for Computational Linguistics* 5, 365–378.

- **Lewis, Dave, Joss Moorkens and Kaniz Fatema** (2017). "Integrating the Management of Personal Data Protection and Open Science with Research Ethics." Paper presented at *the First Workshop on Ethics in Natural Language Processing*, 60–65 (Valencia, Spain, 4 April 2017).

- **Lommel, Arle and Donald A. DePalma** (2016). "Europe's Leading Role in Machine Translation: How Europe Is Driving the Shift to MT." Boston: Common Sense Advisory.

- **Manyika, James *et al*.** (2017). "Jobs lost, jobs gained: Workforce transitions in a time of automation." New York: McKinsey & Company.

- **Margoni, Thomas** (2016). "The Harmonisation of EU Copyright Law: The Originality Standard." Mark Perry (ed.). *Global Governance of Intellectual Property in the 21st Century: Reflecting Policy Through Change*. Berlin: Springer, 85–105.

- **Marshman, Elizabeth** (2012). "In the driver's seat: Perceptions of control as indicators of language professionals' satisfaction with technologies in the workplace." Paper presented at Translating and the Computer (London, 29-30 November 2012), 1-12.

- **Moorkens, Joss** (2017) "Under pressure: Translation in times of austerity." *Perspectives: Studies in Translation Theory and Practice* 25(3), 464–477.

- **Moorkens, Joss** (*forthcoming*). "Irish Translator Survey." Translation and Interpreting.

- **Moorkens, Joss and Sharon O'Brien** (2017). "Assessing user interface needs of post-editors of Machine Translation." Dorothy Kenny (ed.). *Human issues in translation technology*. Abingdon: Routledge, 109–130.

- **Moorkens, Joss and Dave Lewis** (*forthcoming*). "Copyright and the Reuse of Translation as Data." To appear in M. O'Hagan (ed). *The Routledge Handbook of Translation and Technology*. Abingdon: Routledge.

- **O'Brien, Sharon** (2012). "Translation as human-computer interaction." *Translation Spaces* 1, 101–122.

- **O'Neil, Cathy** (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Random House.

- **Ostrom, Elinor** (2011). "Background on the Institutional Analysis and Development Framework." *Policy Studies* 39(1), 7–27.

- **O'Thomas, Mark** (2017). "Humanum ex machina: Translation in the post-global, posthuman world." *Target* 29(2), 284–300.

- **PACTE Group** (2017). *Researching Translation Competence*. Amsterdam: John Benjamins.

- **Pidchamook, Wichaya** (2018). "Is 'Good Enough' Really Good Enough? - Ethics and Working Conditions of Part-Time Freelance Subtitlers." Paper presented at Languages & the Media 2018 (Berlin, 3-5 October 2018).

- **Pollock, Rufus** (2018). *The Open Revolution: Rewriting the rules of the information age*. London: A/E/T Press.

- **Pym, Anthony** (2003). "Translational ethics and electronic technologies." Paper presented at the VI Seminário de Tradução Científica e Técnica em Língua Portuguesa A Profissionalização do Tradutor.

- **Pym, Anthony *et al*.** (2012). "The status of the translation profession in the European Union." Luxembourg: European Commission.

- **Reijers, Wessel *et al.*** (2016). "On the need for a global declaration of ethical principles for experimentation with personal data." Paper presented at ETHI-CA 2016 Workshop (Portorož, Slovenia, 24 May 2016).

- **Rushkoff, Douglas** (2019) Team Human. New York: W. W. Norton & Company.

- **Sakamoto, Akiko** (2018). "Disruption in translator-client matching: Paid crowdsourcing platforms vs human project managers." *Revista Tradumatica* 16, 85-94.

- **SDL** (2018). "SDL Cracks Russian to English Neural Machine Translation." Press release, June 19. https://www.sdl.com/about/news-media/press/2018/sdl-cracks-russian-to-english-neural-machine-translation.html (consulted 20.06.2018).

- **Smith, Ross** (2008). "Your own memory." *The Linguist* 47(1), 22–23.

- **Specia, Lucia and Kashif Shah** (2018). "Machine Translation Quality Estimation: Applications and Future Perspectives." Joss Moorkens *et al*. (eds). *Translation Quality Assessment: From principles to practice.* Berlin: Springer, 201–236.

- **Taivalkoski-Shilov, Kristiina** (2018). "Ethical Issues Regarding Machine(-assisted) Translation of Literary Texts." *Perspectives: Studies in Translation Theory and Practice* (online first) https://doi.org/10.1080/0907676X.2018.1520907.

- **TAUS** (2010). "MT Post-editing Guidelines." https://www.taus.net/academy/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines (consulted 10.06.2018).

- **Toral, Antonio *et al.*** (2017). "Crawl and crowd to bring machine translation to under-resourced languages." *Language Resources and Evaluation* 51(4), 1019–1051.

- **Troussel, Jean-Christophe and Julien Gregory Debussche** (2014). *Translation and intellectual property rights*. Report by Bird & Bird for the European Commission DG Translation. Luxembourg: European Commission.

- **Valenduc, Gérard and Patricia Vendramin** (2016). "Work in the digital economy: Sorting out the old from the new." European Trade Union Institute working paper. https://www.etui.org/Publications2/Working-Papers/Work-in-the-digital-economy-sorting-the-old-from-the-new (consulted 10.05.2018).

- **Way, Andy** (2018). "Traditional and emerging use-cases for machine translation." Joss Moorkens *et al.* (eds). *Translation Quality Assessment: From principles to practice.* Berlin: Springer, 159–178.

- **World Intellectual Property Organization** (1979). *Berne Convention for the Protection of Literary and Artistic Works (as amended on September 28, 1979).* Geneva: World Intellectual Property Organization. http://www.wipo.int/wipolex/en/details.jsp?id=12214 (consulted 10.10.2018).

- **Wynn, Michael** (2015). "Organising Freelancers: A Hard Case or a New Opportunity?" Andrew Burke (ed.). *The Handbook of Research on Freelancing and Self-Employment*. Dublin: Senate Hall, 111–120.

- **Young, Sean D.** (2014). "Behavioral insights on big data: using social media for predicting biomedical outcomes." *Trends in Microbiology* 22(11), 601–602.

Biographies

Joss Moorkens is an Assistant Professor at the School of Applied Language and Intercultural Studies at Dublin City University and a researcher affiliated with the ADAPT Centre and the Centre for Translation and Textual Studies. He has authored articles and chapters on translation technology, MT post-editing, translation evaluation, translator precarity, and translation technology standards. He coedited the book *Translation Quality Assessment: From Principles to Practice* with Sheila Castilho, Federico Gaspari, and Stephen Doherty, published by Springer in July 2018.

Email: joss.moorkens@dcu.ie

Professor **Dave Lewis** is an Associate Director of ADAPT Centre with responsibility for its collaborative research with industry and the Centre's role in the EDGE Marie-Curie Postdoctoral Fellowship programme. He is an ADAPT Principal Investigator, leading its research on Data Governance for Digital Content. With over 160 peer-reviewed publications, his research addresses the use of semantic models to manage the data protection and data ethics issues raised by digital content processing.

Email: dave.lewis@adaptcentre.ie

Notes

[1] For details of the history of translation copyright and recent considerations based on technological change, see Moorkens and Lewis (*forthcoming*).
[2] The Directorate General for Translation in the European Commission also employ freelance translators and are actively growing the outsourced portion of their work (Drugan *et al*. 2018: 43).
[3] See Moorkens (2017) for an expanded discussion of this and the difficulty of collective action for freelance translators. See also Wynn (2015) for a discussion of how trade unions would need to change to assist organisation of freelance workers.
[4] The Berne Convention differentiates between a translation, with rights retained by the original author or creator, and an adaptation that adds original elements. For example, in the recent Swedish fansubbing case (see https://www.svd.se/man-doms-for-flera-mil-uppladdad-undertext), the courts ruled that the copyright for directly translated subtitles belongs to the rightsholder for the original film.

[5] Directive 2003/98/EC on the re-use of public sector information, for example, includes translations under its remit and Directive 2012/28/EU on 'orphan works' concerns digitisation of works whose rightsholder is uncontactable.
[6] Such a resource exists for volunteer translation resources: see a description of Trommons in Du *et al*. (2015).
[7] See Sakamoto (2018) for a discussion on this topic.
[8] Estimated at 330,000 (Pym *et al.* 2012: 94).
[9] See Koskinen and Dam (2016) for a discussion about the boundaries of the translation profession.
[10] See Moorkens (2017) for more details.
[11] This is a point of active discussion at the time of writing, as the Committee of Ministers of the Council of Europe decided in late 2018 that there should be a lifting of restrictions on collective bargaining for the self-employed (European Trade Union Confederation 2018).