Quo Vadis, Crowdsourcing and Online Collaborative Translation?

Miguel A. Jiménez-Crespo
Rutgers University
jimenez.miguel@rutgers.edu

ABSTRACT: Online collaborative translation experienced a meteoric rise in the first decade of the 20th century thanks to the affordances provided by the Web 2.0. Two distinct models emerged, solicited and unsolicited models. In the first one, a company or institutions request the help of volunteer or participants with translation tasks. In unsolicited models, fans, activists or different collectives self-organize to start a translation initiative. These practices quickly attracted the attention of the Language Industry after large corporations implemented crowdsourcing models (Google, Facebook or Twitter). Translation providers and tech companies explored collaborative initiatives in a context of exciting possibilities for growth, while it quickly became a serious cause for concern for professional translators and professional associations such as the International Federation of Translators (FIT) or the American Translators’ Association (ATA). As a rapidly growing phenomenon, Translation Studies scholars were quickly drawn to research this emerging set of phenomena. Initially, the main issues that appeared on Translation Studies literature were related to motivation, epistemological/conceptual research, ethics, translator visibility, or the description of existing initiatives. The second decade of the 20th century saw the consolidation of these activities through technological developments and innovative workflows and the expansion to non-profit ventures, while new technology-driven models based on collaborative micro-task approaches emerged, such as “paid crowdsourcing”. By 2020, the number of providers offering translations crowdsourcing has been dramatically reduced. Many start-ups have been absorbed or have disappeared, while non-for-profit models of translation collaboration, such as educational, NGO or activist initiatives
continue to grow. This paper offers a critical analysis of the evolution of translation collaboration on the web and potential future directions, as well as a review of existing research trends within Translation Studies. The paper ends with an exploration of potential future research trends and directions in this ever-changing area driven by technological innovation.

**KEYWORDS:** translation crowdsourcing, collaborative translation, volunteer translation, fansubbing, motivation to translate

1. **Introduction**

Since the advent of the digital era, translation practices have been in constant evolution thanks to technological innovation and the digital transformation brought by the WWW (Jiménez-Crespo & Ramírez-Polo, 2021). Translation memory, machine translation (MT), lights-out project management, AI or
specialized workflows have revolutionized the ways in which translations are commissioned, produced, managed, distributed, used or repurposed. These developments have set the stage for a number of technology-driven phenomena, such as translation collaboration in online contexts. This set of practices has grown exponentially during the last decade, increasing the visibility of translation in society at large. It has become a consolidated, vibrant area of research within Translation Studies (TS), fostering intra- and interdisciplinary connections. After fifteen years of research on this area, the main objective of this paper is to present a critical review of current research trends and directions, as well as potential future trends and directions.

For the purpose of this paper, a detailed critical bibliographical review of recent studies was conducted using key terms such as “crowdsourcing”, “fansubbing”, “online collaborative translation” or “translation collaboration” in the Bibliography of Translation and Interpreting (BITRA) and in Google Scholar. The thematic analysis of this bibliographical review yielded several broad areas of interest that will be touched upon in the following sections:

1. Epistemological and terminological issues
2. Ethics
3. Motivation
4. Case studies on different contexts and settings
5. Research methodologies

In addition to these main topics, other areas of interest are the relation between translation quality and non-professional models, economic issues such as the impact of collaboration on the profession (Jiménez-Crespo, 2021) or the uberization of translation (Fırat, 2021), as well as the introduction of distributed and extended cognition paradigms.

2. First Things First: Two Distinctive Models of Collaboration

Online translation collaboration experienced a meteoric rise thanks to the affordances provided by the Web 2.0 and beyond, fueled by the ever-growing presence of internet connectivity, personal computers and mobile technologies. It blossomed in both commercial and non-profit settings with
two distinctive models: (1) translation initiated and controlled by companies, institutions or NGOs, and (2) those initiated by volunteer or fan groups. This distinction is referred to as “solicited” and “non-solicited” models by O’Hagan (2011), while Jiménez-Crespo (2017) refers to these two models as “crowdsourcing” and “online collaborative translations”. The first term, translation “crowdsourcing”, is widely used in TS literature, while the latter is often subject to terminological debates that will be discussed later in this section.

“Crowdsourcing” is characterized by the use of dedicated technology-driven workflows that subdivide texts into smaller discrete units, usually sentences, that are subsequently translated and evaluated by large pools of participants without monetary compensation (Jiménez-Crespo, 2017; Morera-Mesa et al., 2013). Meanwhile, non-solicited models are initiated by volunteers, activists or fans. Examples of non-solicited translation are wikis or fansubbing, that is, “fan-produced, translated, subtitled version of a Japanese anime program” (Díaz-Cintas & Muñoz Sánchez, 2006, p. 37). These models also use technological workflows and solutions to manage collaboration to varying degrees, depending on the initiative. Nowadays, it can be confidently claimed that the main difference between these two models rests primarily on the initiator of the process and how it is controlled (Jiménez-Crespo, 2017). In solicited or vertical models, an established company, institution or non-profit organizes the translation process and distributes a call for participation (open or closed) among a pool of potentially motivated or interested volunteers. Meanwhile, in non-solicited or horizontal models, the collective of users themselves organize the process and the call for participation. Here, the locus of control remains within the collective of participants, even when, over time, some initiatives tend to move to some form of hierarchical organization with different levels of access, control and gateways for participation (see for example Orrego-Carmona, 2012; Wang, 2017). Historically, non-solicited translation collaboration by fans and volunteers existed prior to the emergence of the internet and the WWW. Nevertheless, these types of initiatives quickly adapted and grew exponentially thanks to the communicative affordances provided by these technological developments. The early internet forums and subsequent online collaborative platforms helped multiply the possibilities for participation (Fan, 2020; Jiménez-Crespo, 2019).
The chronological review of the emergence of crowdsourcing is traced back by Jiménez-Crespo (2017) to initiatives such as Google in Your Language in 2001 or Yakushite.net (Shimobata et al., 2001). This last project intended to harvest the collective intelligence to assist MT tasks. The combination of MT and volunteer participation was a groundbreaking model that continues to be popular with companies such as Unbabel. Several large-scale initiatives by popular social networking sites in the mid 2000’s led professionals and scholars to demand an examination of the potential implications of these practices. Technological giants such as Facebook, Twitter or Skype could implement existing crowdsourcing models (Brabham, 2013) to harness the collective intelligence of their users for translation purposes. Obviously, the industry as a whole, and social networking sites with captive audiences that could reach billions, were glad to explore innovative new avenues to increase efficiency and cost-effectiveness while minimizing the impact of time constraints. It was initiatives such as Facebook Translate in 2007 or LinkedIn (Kelly et al., 2011) that led translation associations (Federation Internationale de Traducteurs (FIT) or the American Translators’ Association (ATA)) to ring the alarm on the potential impact on the socioeconomic structure of translation as a profession. FIT, for example, released a statement warning about the dangers associated with these volunteer practices (FIT, 2015). Institutions such as the European Union produced extensive publications analyzing this phenomenon (European Commission, 2011). Translation quality was often the supposed concern, but the potential shift from paid to unpaid models, and the possibility of economic loses through market share shifts, could arguably be the primary focus. The social and economic recognition of the profession has always been an arduous path, and crowdsourcing seemed to go against this struggle. All types of volunteer or below-market-rates initiatives were perceived by professionals as a direct threat to their working conditions. Online translation collaboration, both in commercial and non-profit settings, were thus quickly perceived as disruptors for translation as a profession (Sakamoto, 2018). Translation Studies reacted with publications exploring issues that could affect negatively the status and social standing of translators, such as ethics, professionalism, recognition, translation quality or visibility of translation, to name a few (Flanagan, 2016; McDonough Dolmaya, 2012; Zwischenberger, 2021).

The evolution of online collaborative translation continues to date, with
models exploring MT post-editing and “paid crowdsourcing” (García, 2015; Jiménez-Crespo, 2021; Jiménez-Crespo & Ramírez-Polo, 2021), the third stage in the evolution of collaborative translation practices. Paid crowdsourcing emerged around 2008 and it is based on the premise that participants collaborate online to produce a single translation but their compensation rates are below professional ones. Collaboration is still conducted through a micro-task approach, that is, splitting a single translation task among a pool of participants in real time using specialized technology-driven workflows. Its origins can be traced back to early attempts to expand free crowdsourcing models and the difficulties associated with extending volunteer models to a range of content types and genres. Obviously, there are limits such as motivation or the lack of a pool of interested participants. In fact, industry publications soon indicated that free crowdsourcing would work well only for “certain specific purposes and in very narrowly defined contexts” (Kelly et al., 2011, p. 92). Companies attempted to solve the issue of motivation by offering low payments to encourage participants to engage with the “non-appealing bits” (García, 2015). Paid crowdsourcing represents a legitimate attempt at cutting costs while using technological innovations to increase efficiency, speed, and of course, profits. It “unabashedly aims at serving not translators, but clients” (García, 2015, p. 24) and its technological nature means that “the use of a ‘platform’ is the key for maximising its effectiveness” (Sakamoto, 2018, p. 88). In this model, quality outcomes were classified based on the critical nature of the content (legal or medical as opposed to social media updates), lifespan or shelf life of texts (e.g., an email that is read only once as opposed to a webpage that is read millions of times by end users) (Jiménez-Crespo, 2018).

The third decade of the 20th century thus started with the unabated growth of non-solicited fan or “prosumer” (Toffler, 1980) models, while crowdsourcing workflows are now implemented in all kinds of settings, from content with low shelf-life such as tweets, to content where professional quality is expected (i.e., medical or legal texts). From a conceptual and epistemological point of view, this expansion has opened up debates about the search for a metaconcept that can include all types of web-mediated collaborative translation practices. In the early days, it was initially defined by distinct features, such as the volunteer (e.g., Pym, 2011), collaborative (Cordingley & Frigau Manning, 2016) or non-professional (McDonough Dolmaya & Sánchez Ramos, 2019; O’Hagan, 2011) nature of the participation, as well as
the technological foundation (Jiménez-Crespo, 2019). Nowadays, the search for a metaconcept has proven partly unsuccessful due to the dynamic and fluid nature of this phenomenon as explored in the next section.

3. Epistemological Issue and the Search for a Metaconcept

The dynamic and shifting nature of collaborative translation has led to a number of studies that focus on epistemological and terminological issues (Hebenstreit, 2019; Jiménez-Crespo, 2017; McDonough Dolmaya & Sánchez Ramos, 2019; O’Hagan, 2011; Pym, 2011; Zwischenberger, 2021). This includes workshops dedicated to untangling this terminological fuzziness in TS surrounding online collaborative translation (such as the workshop Translation on and over the Web organized by the University of Vienna, Nov. 2021). Translation collaboration has always been quite common, and it is “evident in all types of translation scenarios and across the whole process of translation, from authors, publishers, to translation agencies and to translators” (O’Brien, 2011, p. 17). In this context, it is productive to separate collaboration into two broad notions (O’Brien, 2011). In a narrow sense, it refers to the actual collaboration between two or more translation agents to produce a single translation, the final product being the result of more than one participant. In the second notion translators collaborate with others, such as authors and translators or localization engineers and localizers. The type of online collaborative translation is understood as the first one, the joint efforts of a crowd or collective to produce a single translation.

Over the years, the main perspective to define collaborative translation has been shifting, from a focus on “voluntarism” or “web mediation”, to “non-professionals” or “collaboration”. The first terminological debates separated translation collaboration using professional vs. non-professional settings, as well as paid vs unpaid volunteers. This terminological debate is still ongoing. Pym (2011), for example, initially proposed the term “volunteer translation” as a metaconcept. The focus was on the fact that participants were not remunerated for their work. Nowadays, crowdsourcing also appears in the “paid crowdsourcing models” (Garcia, 2015; Jiménez-Crespo, 2021, 2018; Sakamoto, 2018) and, therefore, the emphasis on voluntarism is not inclusive enough to encompass all existing practices. The most widely
used categorization was proposed by O’Hagan (2011) and it is based on the locus of control of the initiative, that is, solicited vs. non-solicited models. Jiménez-Crespo (2017), after a critical review of proposed concepts, suggests “crowdsourcing” and “online collaborative translations”, even when he claims that the latter could be used as a metaconcept for all types of web-mediated translation collaboration. More recently, Hebenstreit (2019, p. 150) proposes the notion of “social media-driven translations”, using three classifiers to substantiate this term: users, instruments and process. A recent article by Gambier and Kasper (2021, p. 14) proposes “participatory or collective translation (with implied crowdsourcing)”, two classifiers that were already discussed in previous works such as Jiménez-Crespo (2017) and Cordingley and Frigau Manning (2016) as inadequate to encompass the wide range of practices. Finally, Zwischenberger (2021) proposes the notion of “online collaborative translations” as the metaconcept, a term used by Jiménez-Crespo (2017) for non-solicited models. Nevertheless, the issue here is that horizontally managed, non-solicited collaboration is left then without a concept. As a proposal, current debates suggest that the notion of “crowdsourcing” (including “paid-crowdsourcing”) is widely established in TS literature. Similarly, horizontal models in which the locus of control is within the collective could be referred to as “non-solicited” collaboration (including fan and activist translation) to subdivide these two models. Finally, the metaconcept “online collaborative translations” proposed by Zwischenberger (2021) and Jiménez-Crespo (2017) and could be used as a metaconcept to help solve the terminological confusion.

As a final point, the use of “online” or “web-mediated” should also be discussed. This review highlights the fact that translation collaboration is primarily mediated by the WWW, either through technological workflows or through forums or different types of cloud collaboration (cloud-based translation memory (TM) tools, cloud subtitling tools such as Amara, cloud drive repositories such as Google Drive or others). Often the use of “online” indicates primarily the use of technological solutions to collaborate, while in a limited number of cases it refers to the use of the WWW as either a cloud repository to collaborate on shared documents or as a communicative platform (emails, instant messaging, forums, etc.). It is likely that given the encroachment of the WWW on modern lives, the use of “online” might be redundant at this point, and the use of terms such as “crowdsourcing” vs. “non-solicited” collaborative
translations might be a potential solution to this terminological confusion. Similarly, in terms of a metaconcept, “collaborative translation”, with or without the “online” or “web-mediated” adjective, might represent in itself a productive term given the inseparable technological nature of modern societies.

4. Ethics of Crowdsourcing and Non-solicited Collaborative Translation Models

Ethical issues continue to be one of the most significant issues analyzed by scholars in this area (e.g., Basalamah, 2020; Borodo, 2020; Zwischenberger, 2021). Initially, ethical concerns were raised in relation to whether deontological or professional ethics would apply equally to volunteer and professional settings. In this regard, Chesterman (2001, p. 146) questioned whether debates on ethics would “include amateurs as well as “professionals”. He proposed to distinguish “between someone “who is a translator” and someone “who does translations (sometimes)”. Pym (2012) also discusses in this same regard the potential distinction between “ethics of translation” and “translator ethics”. Here, Pym wonders whether any study into ethics should recognize that the translator is not always professional: “translator ethics has now more to consider than the professional translator” (2012, p. 84), and consequently “this opens up new terrain for ethical inquiry” (Pym, 2012, p. 4). This is precisely the new area of enquiry that has been developed over the years. Debates on ethics primarily focus on “crowdsourcing” or “solicited” models, those in which a company or institution profits from the labor of volunteer participants. This concern also exists in activist or NGO settings, as was discussed in the paper by Piróth and Baker (2020) on Translators without Borders. They argue that NGO volunteer translation might entail potentially unethical consequences if the translation assets produced using volunteer work (translation memories or terminology banks) are used to develop for-profit assets for companies, such as training MT systems for large corporations.

In general, most scholars argue that even when participants might enjoy or receive intrinsic satisfaction from their participation, crowdsourcing represents a legal, but unethical activity (O’Hagan, 2016). Zwischenberger (2021), for example, argues that from a “consequentialist ethics centered on
the actual and possible (long-term) consequences of actions”, crowdsourcing represents an unethical exploitative activity. In addition, she adds that attention should be paid to consequences “that are not immediately apparent” (Zwischenberger, 2021, p. 11). The ethical implications of for-profit crowdsourcing models might therefore entail consequences far beyond the immediate actors or time of application. For example, she claims that these practices might lead to a potential downward pressure on professional rates or a negative impact on translator status. This is so even when collaborative translators might “see themselves as benefitting in various ways through their work”. Her claims are based on Wertheimer's (1999) notion of exploitation and how it might impact third parties down the road and consequentialist approaches to ethics, “which unlike a deontological ethics does not prescribe and evaluate single actions but focuses on the moral acceptability of (long-term) consequences of actions” (Zwischenberger, 2021, p. 11). The broad adoption of crowdsourcing, according to Zwischenberger (2021), might also lead to an examination of the need for a specific university-based education and training in translation or translation studies in general.

A common methodology to research ethical issues has been the study of narratives and discourses of commercial and non-profit initiatives in order to persuade or motivate participants to translate pro bono (Kang & Hong, 2020; Piróth & Baker, 2020). This approach is justified as the work of participants is often described as a positive contribution to moral and social obligations. Nevertheless, it is pointed out that corporations often provide monetary donations in the form of grants and they end up benefitting from the altruistic work of participants in critical situations (e.g., Piróth & Baker, 2020). Similarly, Kang and Hong (2020) study the discourses on social responsibility used by Coursera, a for-profit educational platform. They discuss how Massive Open Online Course (MOOCs) are associated with greater social benefit and goals, but nevertheless, the use of these discursive devices to motivate participants might mask their for-profit motives. This case illustrates the evolution and a more fine-grained approach to ethical issues in for-profit scenarios as society at large, as well as speakers of languages of lesser diffusion, might benefit from volunteer “digital labor”.

The ethical implications of non-solicited models such as fansubbing, in relation to potential conflicts, have not attracted much attention in current literature. Some issues such as the breach of copyright laws by fans have been
studied in the past (e.g., Lee, 2011). Scholars such as O’Hagan have argued that crowdsourcing represents a legal, unethical activity while fansubbing is an illegal, ethical activity. It is also of interest here the study by Lee (2011) using ethnographic or interventionist methods on the analysis of copyright interpretation by fan communities in several countries. The study concludes that fansubbers perceive their activity as ethical. However, they tend to be more protective of how others might use their subtitles once they are released. In fact, fansubbers felt the use of their subtitles without authorization was problematic and unethical to a large degree.

The last case of recent interest is the participation of volunteers in emergency crises and situations. This area was initially studied in crowdsourcing research with cases such as the Haiti earthquake and the use of the Haitian diaspora to translate emergency SMS messages using the Ushahidi platform (Munro, 2010). Recent studies have introduced the notion of “linguistic first aid”, or in other words:

a situation in which a person with some proficiency in a foreign language can provide immediate help to another person or a group of people who are not able to communicate in a language to help them overcome linguistic barriers (Probirskaja, 2020, p. 340).

In this case, professionals and non-professionals alike participate as a personal choice derived from the appreciation and love to their neighbors and communities rather than from a duty related to their profession. Parra Escartín and Moniz (2020) discuss the ethical implications of using crowdsourcing in cascading crises (earthquakes, hurricanes, pandemics, etc.). They discuss how the issue of remuneration and the wider potential impact on the profession down the road might not be of interest in these scenarios, but procedural and workflow issues suddenly emerge as the locus of ethical choices. They frame their discussion around what they call “procedural accountability”. This is especially significant in cases in which the lack of time might override other considerations in order to quickly save lives. Issues such as selecting who has access to the crowdsourcing platform used, how the data is distributed for translation, how sensitive or confidential data is handled, the support mechanisms available for the crowd for potentially traumatic experiences (Parra Escartín & Moniz, 2020). In addition, specific issues to
address for those developing and implementing crowdsourcing relate to data anonymization of both participants and victims, as well as protection of the crowd from post-traumatic stress disorder.

As a whole, these recent studies provide a richer and more nuanced approach to ethical issues in different volunteer scenarios that go beyond the most commonly studied cases of crowdsourcing implemented by large corporations in previous research (often social networking sites or tech companies). These studies open up the field to the wide range of scenarios in which a single approach to ethical issues might not suffice, even in for-profit settings such as the case of educational or edutainment initiatives (Coursera or TED Talks) vs. social networking sites. Similarly, they illustrate that discourses and narratives of non-profit initiatives might also entail ethical dilemmas if the “digital labor” or work of love of participants might end up falling in the hands of for-profit companies, often for training MT or AI systems.

5. Motivation to Participate

Motivation to participate is still one of the main research questions that attracts a large number of studies (e.g., Alonayq, 2021a, 2021b; Camara, 2015; Dombek, 2013; Mesipuu, 2012; Moreno García, 2020). Theoretically, most previous studies differentiate between intrinsic and extrinsic motivations using the framework provided by studies on Free and Open Software (FOSS) (Frey, 1997). Intrinsic motives are those related not to financial compensation or reward but to a feeling of obligation to a specific community rather than to personal enjoyment, such as self-improvement or enjoyment of the task. Examples of this motivation type are to gain intellectual stimulation or make information available to other language speakers. Extrinsic motivations are related to direct or indirect rewards, such as personal benefit (i.e., gaining more clients or reputation, getting presents or the potential to attract customers). In the review by Jiménez-Crespo (2017, pp. 220-223) of motivations in previous studies, the main motivations across the board were intrinsic ones such as:

1. Making information in other languages accessible to others
2. Helping the organization with their mission or a belief in the organization’s principles
3. Achieving intellectual stimulation and intellectual reasons

Another set of intrinsic and extrinsic motivations appears in most studies as secondary or less prevalent such as:

4. The desire to practice a second language
5. Professional motivations related to the need to gain translation experience or increase one's reputation
6. The satisfaction of completing something for the good of the community

More recent studies have questioned the exclusive emphasis on intrinsic vs. extrinsic motivations, as well as the difference between initial motivations to participate vs motivations to continue participating. In the Arabic context, Alonayq (2021a, 2021b) uses a socio-narrative approach to explore common narratives, rather than surveys or interviews. The researcher compiled a corpus of narratives by organizations such as Kalima, the Arab Organization for Translation, Taghreedat, and the Translation Challenge. It was found that in Arabic-speaking volunteer contexts, specific distinctive issues, such as religion, gender, and diglossia, appear as motivating factors that were not identified in previous studies focused on Western contexts. He advocates for the study of common narratives to identify potential motivating factors in distinctive geographic areas. Similarly, a recent study into Chinese fansubbers in Spain offered a new perspective by separating the motivating factors to start subtitling from those that keep fansubbers active in the long run (Moreno Garcia, 2020). The motivations to start participating identified were participants' passion for the content and activity, as well as their desire to attain experience. Conversely, the main motivations to continue participating were finding friends and having an online volunteering platform. The study also identified a relation between the possibility of producing content themselves, the prosumer (Toffler, 1980) and/or “produser” (Bruns, 2009) paradigm, with increased motivation to participate.

These recent studies offer two distinct pathways for future research. A common one is the study of different geographical areas and type of initiatives in order to identify potential differentiating motivations, and a different one related to differences in motivation from initial stages to later ones. Differences in motivation between professionals vs. non-professionals would
also provide interesting insights, such as those identified in humanitarian initiatives where professionals see their motivation as a desire to help their neighbors rather than a duty derived from their profession (Probirskaja, 2020).

6. Settings and Constellations in Translation Collaboration

Possibly one of the most exciting contributions of recent research in TS is the widening up of research in this area in terms of descriptive and empirical research that encompass a growing number of settings and geographical areas. These recent studies are helping provide a richer and more nuanced treatment of recent trends in research into collaborative translation. The settings are many and varied. In some cases, studies have delved into distinctive translation modalities and settings. Studies have explored the rich and varied constellation of fan translation in audio visual translation (AVT) (Fan, 2020), or issues of interest such as the use of crowdsourcing and MT in subtitling (Svobodová, 2018). One of the main areas of fan translation is videogame localization, including romhacking. This term refers to the process of modifying the ROM (Read-Only Memory) data which is not meant to be changed by the user of a video game, to alter various aspects of the game, including the game's language (O'Hagan & Mangiron, 2013, p. 10). Recent studies have primarily studied it in Arabic contexts (Al-Batineh & Alawneh, 2021), as well as Western contexts and other geographical areas (Capellini, 2021), including comparisons between professional and non-professional versions in areas such as Turkey (Sarigül & Ross, 2020). The interest in the study of activist translations also continues to grow, with studies into the role of collaborative translation in different movements such as the Arabic Spring (Baker, 2016) or Taiwan's Sunflower Student Movement in 2014 (Chang, 2020). Similarly, the role of collaborative translation in specific settings such as religious translation has been the focus of research, such as its role in Christian free cultures (Sen, 2021). Other areas of interest are those related to edutainment or educational initiatives such as the study of quality and revision in TED Talks (Malaczkov, 2020) or the study on narratives of volunteer participation on Coursera (Kang & Hong, 2020). Similarly, citizen science and the role of collaborative translation practices have also drawn attention, such as the Zooniverse initiative (Desjardins, 2021; Heinisch, 2021). Last but
not least, specific translation types such as literary translation also continue to attract the interest of scholars (Yang, 2020). Among these descriptive studies, the organization of collaboration is a topic of interest from various perspectives, such as the study of the Chinese platform Yeeyan of Yu (2019) or the seminal study by Orrego-Carmona (2012) on the organization of the Argentinian fansub team Argenteam. The study by Yang (2020), for example, showed that communication can help mitigate organizational and quality risks in online collaborative translation.

In addition, the potential of crowdsourcing to mediate in cases of languages of lesser diffusion or in developing countries has been the focus of a growing number of studies. txtEagle is an example of an early initiative (Eagle, 2009). This project was created in Africa in order to crowdsource micro tasks through SMSs. Participants translated segments into different long tail languages and they received cellphone credit as compensation. The same system has been expanded to Asia with other initiatives such as Mobile Works or mClerk. This last initiative, for example, involved text embedded in images distributed through SMS messages that users could either translate or transcribe. Recently, research in this area has explored using crowdsourcing to prepare an MT system for less commonly spoken languages in Africa such as Bambara (Tapo, 2020). There is a long way ahead, but voluntarism to fill the gap in languages of lesser diffusion or those with a lack of professional translators remains a much-needed area of research.

7. Research Methodologies

In terms of research methodologies, the majority of studies into online collaborative translation can be located within the so-called “sociological turn” in TS (Wolf & Fukari, 2007). This turn has inspired empirical participant-oriented and context-oriented research in the discipline (Saldanha & O’Brien, 2014), focusing on the interaction between agents, texts and the contexts of production and reception. Many studies have used ethnographic methods, including “netnographic” approaches in which the informant immerses herself in online communities (Kozinets, 2010). Examples of this last approach is the study by Dombek (2013) on motivation to participate in Polish Facebook Translate or the studies in the Chinese platform Yeeyan by Yu (2019).
This has often been combined in mixed methods approaches, combining online surveys with “netnography”. Online surveys have continued to be part of the study of collaborative translation, especially in issues such as motivation or perception of activities both for professionals and volunteer participants alike. Similarly, the use of documentary research methods to study common narratives and discourses has also been a recurrently used methodology, such as the study of professional blogs to analyze professionals’ attitudes towards crowdsourcing by Flanagan (2016), study of participants’ online blurbs by Olohan (2014) to study motivations, documents by professional associations by Zwischenberger (2021), documents in Arabic volunteer initiatives by Alonayq (2021a, 2021b) or the documents released by the educational MOOC initiative Coursera (Kang & Hong, 2020).

Product and process-based methodologies have also been used to a lesser degree, such as the corpus-based study on the naturalness of the user interface of the Spanish-Spain version of Facebook as compared to locally produced sites (Jiménez-Crespo, 2013). This study was subsequently used in a process-based study of user reception in which 80 Spanish-speaking subjects were asked to select the most natural constructions for User Interface items (Jiménez-Crespo, 2016). Both studies confirmed that the iterative model implemented by Facebook and other social networking sites identify the most frequent constructions, and by extension the most naturally sounding translations, in non-translated original social networks. Similarly, other initiatives have built corpora of educational materials translated through crowdsourcing (Sosoni et al., 2018), or the parallel corpus TED Multilingual Discourse Bank (TED-MDB) in the English-Turkish combination (Zeyrek et al., 2020). This last study has allowed the much-needed study of general features of crowdsourced translations, such as explicitation or implicitation (Zeyrek et al., 2020). As a whole, more corpus, user-reception and process-based studies are needed in order to obtain a clearer picture of the many under-researched areas in this field dominated by the sociological trends in Translation Studies.

8. Where Do We Go from Here? The 2020’s and Collaborative Translation

The study of online collaborative translation will continue to be a vibrant area
with dedicated panels in the main TS conferences and special issues journals. Issues related to AVT and fansubbing will continue to attract the attention of a large number of scholars in different geographical areas. Studies will continue to offer interdisciplinary and multidisciplinary intersections that will enrich our understanding of these phenomena. Similarly, intradisciplinary connections with overlapping TS areas will continue to grow, such as non-professional translation, MT or AVT. It can be predicted with confidence that the main research trends described in this paper (ethics, motivation or descriptive studies of different settings, language combinations or geographical regions) will continue to be highly productive areas of investigation. One topic that will probably produce studies of interest is the intersection of online collaborative translation modes and quality. Since the early days, one of the main areas of interest was the potential lack of quality in texts produced by non-professional or volunteer initiatives. In this regard, the development of technological workflows and different modes of gatekeeping, aggregation of participation or collaborative evaluation has moved the debate beyond simple comparisons of quality outcomes with professional translations. Nowadays, studies have introduced a more nuanced approach to the study of translation quality that includes economic factors (Jiménez-Crespo, 2018), such as the uberization of translation in the digital economy (Fırat, 2021), fitness for purpose models or the recurrent debate in crises that pits quality against saving lives (Hunt et al., 2019; Parra Escartín & Moniz, 2020), to name a few. Recent studies in this area have examined initiatives such as quality management in Translators without Borders (Krimat, 2021). Similarly, the use of online collaborative translation in translation education is expected to be part of the ongoing debate given its widespread presence in online contexts. As an example, recent studies investigate the use of crowdsourcing in localization (Sánchez Ramos, 2021) or AVT courses (Beseghi, 2021). Finally, the vibrant area of Cognitive Translation Studies needs to make inroads into the study of collaborative translation in technological settings such as indicated by Risku et al. (2016) and Jiménez-Crespo (2017). The introduction of distributed and extended cognition paradigms in TS according to which a collective system can be studied as a unit (Risku & Windhager, 2013) can be a productive research avenue. These models have widened the cognitive object of research from some individual, isolated mental operations and systems to socio-cognitive issues, including the social and artefact-mediated processes
that were an essential part of human cognition. This has been studied previously in TS in relation to TM by Christensen (2011) that argued that the shared use in real time of translation memories can be considered as a distributed cognition unit. Similarly, a recent study by Pleijel (2021) advocates for the introduction of group-level cognition to study translations cognitively processed by translation teams.

Throughout history, translation has often been a collaborative act. The communicative affordances provided by the WWW has exponentially multiplied its possibilities and have become a breeding ground for new practices that simply could not have existed before. As Désilets (2007) wrote in the early days of online collaborative translation, “massive online collaboration introduce[s] exciting new opportunities that simply were not on our minds before” (n.p.). Translation Studies has responded to the challenges posed by these new practices, with a new established research area that combines what is known as the “technological turn” with theoretical and methodological foundations on the “sociological turn” (Jiménez-Crespo, 2020). Online collaborative translation is here to stay, and advances in AI and MT will continue to impact the developments of new initiatives and models of participation, as well as research in this area.

References


Dombek, Magdalena. (2013). *A study into the motivations of Internet users contributing
to translation crowdsourcing: The case of Polish Facebook user-translators [Doctoral dissertation]. Dublin City University.


Quo Vadis, Crowdsourcing and Online Collaborative Translation?

Creative Industries Journal, 3(3), 237-252. https://doi.org/10.1386/cij.3.3.237_1


Professional Profile

Miguel Ángel Jiménez-Crespo holds a PhD in Translation and Interpreting Studies from the University of Granada, Spain. He is Full Professor in the Department of Spanish and Portuguese, Rutgers University. He authored *Crowdsourcing and Online Collaborative Translations: Expanding the Limits of Translation Studies* (John Benjamins, 2017) and *Translation and Web Localization* (Routledge, 2013). He has been the co-editor of the John Benjamins journal *JIAL: The Journal of Internationalization and Localization*. His research focuses on the intersection of translation theory, translation technologies, digital contexts, translation training and corpus-based translation studies.