

Proposal for a Triple Bottom Line for Translation Automation and Sustainability: An Editorial Position Paper

Joss Moorkens, SALIS/ADAPT Centre, Dublin City University

Sheila Castilho, SALIS/ADAPT Centre, Dublin City University

Federico Gaspari, University of Naples "Federico II" and ADAPT Centre, Dublin City University

Antonio Toral, University of Groningen

Maja Popović, ADAPT Centre, Dublin City University

ABSTRACT

This article is both an editorial introduction to the guest-edited special issue of JoSTrans on Translation Automation and Sustainability, and a position paper in which we propose a model for evaluating the sustainable use of automation technology in translation and beyond. As grounding notions, the article reviews definitions of automation and considers the urgency of sustainability. Thereafter we propose an adaptation of Elkington's (1997) triple bottom line, giving equal weight to evaluation based on people, planet, and performance, describing each of these elements in turn. Finally, we introduce the articles from this special issue, in which authors describe various aspects of automation technology in translation with a focus on sustainability.

KEYWORDS

Translation technology, translation automation, artificial intelligence, sustainability, triple bottom line, ethics.

1. Introduction: translation, automation, and sustainability

The advent of Neural Machine Translation (NMT) with overall improved quality and (sometimes deceptive) fluency for an unprecedented number of language combinations and domains has coincided with a huge increase in translation automation. The largest free online MT provider, Google Translate, translated 143 billion words per day in 2016, and had one billion Android app user installs by 2021 (Pitman 2021). However, automation is not 'all or nothing', but rather varies according to the level of human input (Parasuraman *et al.* 2000). Improved quality and hype about its capabilities have pushed NMT, with or without post-editing (PE), into high-stakes use cases for which automation had previously been considered inappropriate (Vieira *et al.* 2021). Predictions that "post-editing will dominate translation production" (Lommel and DePalma 2016: 20) do not seem to have materialised in all segments of the market, with ELIS Research (2023) reporting that only 31% of surveyed European translation organisations offer PE as a product, although this percentage is growing year on year. Beyond full or light PE, there are various other modes of interaction with NMT in translation processes, such as its use as 'just another input' (Cadwell *et al.* 2016), for fuzzy match repair, and for interactive MT. There is no evidence as yet of flexibility for translators to move between these modes, as recommended in contemporary literature on human factors (such as Calhoun 2022).

In a dynamic global industry, NMT is not the only form of automation in

translation workflows, particularly as more data is being gathered from translation projects. Other options for automation include error identification and correction, quality evaluation, terminology consistency checks, project management, job allocation (e.g. Herbert *et al.* 2023), and billing/invoicing functions. The availability of generative tools powered by artificial intelligence (AI) and using large language models (LLMs) broadens these options further, with functionalities still being uncovered in the translation of well-resourced languages (Hendy *et al.* 2023), with contextual awareness that surpasses NMT (Castilho *et al.* 2023), and improved automatic translation evaluation (Kocmi and Federmann 2023). Sánchez-Gijón and Palenzuela-Badiola (2023) found that the capabilities of ChatGPT across preparation, translation, and post-production stages of a translation project are uneven, with some workflow steps carried out successfully and others in a random and inconsistent manner. Commercial Computer-Aided Translation (CAT) tools have begun to introduce these LLM tools, often using an API¹ connection to one or other provider (such as memoQ AGT, which looks to provide contextually appropriate MT based on sharing a number of fuzzy matches). The fact that ChatGPT, GPT-4, and Google Gemini were launched between the submission of articles for this special issue and their publication exemplifies the dynamic nature of technology and the relevance of the topic of automation, which is becoming increasingly crucial across translation, in educational as well as professional settings.

These developments and their repercussions within the translation industry and more broadly throughout society make it clear that narrow methods of evaluation are no longer sufficient to draw conclusions about the utility of technology or to predict the implications of its introduction. Messick (1989: 5) proposed that not only should measures be appropriate, meaningful, and useful, but that we should also consider “the social consequences of their use”, particularly if these measures or scores are used as a basis for action. The vastly exaggerated claims of human parity for Chinese-English MT in the news domain based on direct assessment evaluation had social consequences in terms of media reports that were much less nuanced than the original paper by Hassan *et al.* (2018). Some of us responded to suggest broader evaluation methods (see Toral *et al.* 2018 and Läubli *et al.* 2020), but a difficulty is that these do not present the reader with a simple story. Similarly, Schwartz *et al.* (2020) criticise a narrow focus on performance in evaluation – what they call Red AI – without consideration of efficiency and sustainability.

In this position paper which also serves as an introduction, we build on the insights provided by the papers collected in the special issue and on our individual and collective thinking while editing it, to propose a broader method of evaluating automation technology in connection to sustainability, and its utility within research and private and public organisations, using the framework of Elkington’s (1997) triple bottom line of people, planet, and profit (the latter of which we adjust to ‘performance’). It is important to state from the outset that we are not anti-technology or automation –

far from it, we are all technology researchers, after all – but feel it is important to caution responsibility in its evaluation, reporting, and use, mindful of the broader consequences and high stakes. In the following sections, we present some definitions of automation and its effects on the translation industry, then explain the urgency of sustainability. Thereafter we introduce the triple bottom line and some related discussion from business ethics, and how we might practically go about a holistic evaluation of technology. To this end, taking our cue from well-established models and approaches, we propose six questions that in our view represent a useful starting point to analyse the development and adoption of translation technologies from the point of view of the sustainability of translation and automation.

The idea for this special issue originated from the long-standing common interests and collaboration of the guest editors over multiple research projects and publications in the field of human-informed translation technologies, and in particular NMT. Recent developments in generative AI lend even more meaning and importance to the key motivation for the special issue, which is to investigate the variety of novel implementations of (full or partial) automation in translation and their effects on sustainability, considering both the sustainability of the translation industry and profession as well as ecological sustainability. We discuss the contributions to the special issue prior to our conclusion (and in the video roundtable in this special issue), highlighting how they illuminate the relation between automation and sustainability of translation from multiple complementary perspectives.

2. Defining automation

Automation is the first grounding notion of this position paper. Parasuraman *et al.* (2000: 287) define automation as the “full or partial replacement of a function previously carried out by the human operator”. The corollary is that automation can “vary across a continuum of levels, from the lowest level of fully manual performance to the highest level of full automation” (*ibid.*: 287). Their proposed typology of automation is mirrored by many others in the literature, according to a meta-analysis by Vagia *et al.* (2016). The definition of automation as replacement pits humans against machines, which O’Brien (2023) criticises as an antagonistic dualism that does not afford the potential automation of tasks that were not previously performed by humans.

In translation, few if any translators now receive absolutely no machine assistance, with basic spelling and grammar correction available in most text editing interfaces at a minimum. As we move through the automation typology, users are offered several options (as may happen within a CAT tool) or just one (as may happen when post-editing). As the machine begins to assume more control, humans may have the opportunity to veto or approve the choice (as happens with interactive MT) or may be not involved at all (as in raw MT). However, it should be noted that humans are still

involved in the development, curation of data, and implementations used for raw MT processes, even if in a somewhat less visible, subtle manner (Gaspari 2022).

This sort of scale was applied to translation by Hutchins and Somers (1992) during the era of rule-based MT, who influentially included human-aided MT and machine-aided human translation as transitional areas in a translation automation continuum. Christensen *et al.* (2022: 35) update this model to create a six-point typology of translation automation along the lines of the Society of Automotive Engineers' classification of driving automation, noting that while these roughly align, they also differ in that society has zero tolerance for driving automation errors that will have fatal repercussions, whereas "it seems that society has already grown accustomed to imperfect translations".

In these typologies, the automation level is usually static. Lommel's (2021) proposal for responsive MT suggests a dynamic interaction, with automatic domain adaptation, contextually aware MT systems that are adaptive based on feedback and that automatically match with application and usability requirements. A dichotomy that appears in the automation literature is between adaptable (human-controlled) and adaptive (machine-controlled) dynamic switching between automation levels. Vagia *et al.* (2016) suggest that the former can avoid "some of the pitfalls" of the latter, echoed by Calhoun (2022) who finds improved situation awareness and perceived control on the part of users who can set their desired level of automation.

Way (2013: 2) examines the key variables that influence the feasible or desirable levels of automation in (machine) translation along with some case studies, arguing that the "degree of human involvement required – or warranted – in a particular translation scenario will depend on the purpose, value and shelf-life of the content." More recent work (e.g. Rothwell *et al.* 2023) added risk to this guidance, following Canfora and Ottmann's (2018) typology of translation risk, ranging from the risk of miscommunication to the risk of injury or death. This would seem to still hold true for generative AI tools. Of course, risks to sustainability as detailed in the next section are at a higher level than those originating in the texts themselves.

3. The urgency of sustainability

According to the Brundtland Commission Report (1987: 16) sustainable development "meets the needs of the present without compromising the ability of future generations to meet their own needs." While these early notions of sustainability mostly related to environmental sustainability, this has broadened over time to include stewardship of our social environment, as evidenced by the United Nations Sustainable Development Goals (SDGs; UN General Assembly 2015)².

Two major news stories at the time of writing regard an unprecedented heatwave in many parts of the globe and a prolonged strike by screen writers and actors in the United States. The former follows alarming

research (Thompson *et al.* 2023: 1), predicting that many vulnerable populations will likely be subject to more frequent extreme heat due to climate change, and will be “increasingly exposed because of limited healthcare and energy resources”. Stoddard *et al.* (2021: 654) cite “narrow techno-economic mindsets and ideologies of control” as some of the reasons why three decades’ worth of climate mitigation efforts have been ineffective. The Hollywood screen writers’ strike may appear unrelated, but this was also triggered by narrow techno-economic mindsets and ideologies of control, as writers are forced to argue for project-length contracts and try to prevent their output being used to train AI systems that might eventually displace them.

Both ecological and workplace sustainability are core issues in translation and have been for some time. As Cronin (2019: 516) writes, translation is “inevitably implicated in any discussion of what happens to technology in an age of accelerated climate change”. Our engagement with technology needs to be situated “within the carrying capacity of a planet with finite resources and an ever shortening timeline of climatic viability” (*ibid.*: 516). Yamada’s (2023) more modest aim is to use translation process research to demonstrate that human translation and PE are not so dissimilar, arguing that the devaluation of PE threatens the sustainability of commercial translation. As we shall see in Section 4.2, there is ongoing work in AI and sustainability, with a growing realisation of environmental harms beyond carbon emissions, but efforts to improve efficiency are currently outpaced by the growing size of LLMs as the most effective way to increase output quality.

UN SDG number 8 proposes “sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all”, but a focus on reducing labour costs by imposing automation has affected many industries, including translation. Moorkens (2020) notes that a tendency to atomise some parts of translation work is incompatible with findings on motivation and satisfaction, in that positively motivating factors are often intrinsic, involving meaningful work with recognition and responsibility (Herzberg 1976). According to Docherty, Kira, and Shani (2008: 4), instead of focusing on “short-term, static efficiencies such as productivity and profitability; we must also focus on long-term, dynamic efficiencies such as learning and innovation”. This goes against long-standing business orthodoxies, such as Porter’s (1980) five forces, which advise against allowing suppliers or customers to form an allegiance in order to limit profits by increasing costs or limiting prices. It is with work such as Porter’s in mind that the move to digital translation platforms with controlled methods of communication between translators who work on decomposed portions of text, as parts that may be considered interchangeable, makes business sense, at least in the short term (Agorni and De Bonis 2022).

4. Triple bottom line

Elkington (1997) was not the first to propose that environmental and social

benefits should be of equal importance as economic ones, but his idea of the 'triple bottom line' of people, planet, and profit has stuck. The 'people' in this model are not only workers, but also other stakeholders in an organisation. Following Stakeholder Theory (Freeman 1984), translation stakeholders might include translators, target text end users, institutions and citizens, trainers and educators, members and representatives of professional associations/bodies, shareholders, company owners, project managers, and workers in all areas of translation. We could also include people living in the area where business is conducted or journalists who report on an industry, who Phillips (2003) defines as derivative stakeholders, who have an indirect connection in the organisation. According to the triple bottom line, an organisation should not exploit these stakeholders, and should ideally contribute to their wellbeing, along with all others involved in the production of their output.

The 'planet' part of the model involves not only causing no ecological harm and limiting energy consumption and waste, but also carrying out a full life cycle assessment of products. Brevini (2022) highlights not only the power and water requirements to operate the data centres required for cloud computing behind contemporary AI, but also the rare earth minerals required to build machines and the e-pollution caused by the disposal of outdated hardware in huge dumps, such as those in Kenya and Cambodia. As Cronin (2020: 520) writes, "there is nothing immaterial about the material consequences of virtual technologies". Williams (2011: 355) details the risks to people in the manufacture, disposal, and recycling of technology, such as "exposure to ancillary chemicals used in high-tech processing, in particular making semiconductors", including known carcinogens.

The argument put forward to support the triple bottom line is that environmental sustainability can be more economically profitable for an organisation in the long run (Elkington 1997: 38). Melé (2009) believes that running an organisation in accordance with the common good – for the good of our larger community – should be a successful business approach and does not mean ignoring the needs of shareholders. For Melé (2019: 298), sustainability is "nothing other than the common good for future generations". This fits with Elkington's (1997) view of sustainability as a new form of value that society will demand.

The 'profit' part of the model may seem obvious, but Elkington (1997: 74) expands this to focus on the economic sustainability of an organisation, including not only economic capital, but also human capital ("a measure of the experience, skills, and other knowledge-based assets of the individuals who make up an organization"), intellectual capital, social capital, and natural capital. This view of profit incorporates the real economic benefit enjoyed by the host society and, more broadly, ecosystem. The corresponding concept to the traditional bottom line in business of 'profit' when considering technology is 'performance'. Schwartz *et al.* (2020: 62) write that the "push to improve state-of-the-art performance has focused

the research community's attention on reporting the single best result after running many experiments for model development and hyperparameter tuning" at the expense of improving model efficiency. The tendency is to evaluate the latest (AI-based) technologies based on a narrow view of performance, as we shall see in Section 4.3, ignoring inconvenient elements, such as biases in output, that are paid considerably less attention. Of course, we recognise that we cannot ignore the bottom line of performance or quality, but we suggest that we can, and should, report it alongside the implications for people and the planet.

The triple bottom line model was intended to encourage long-term thinking, a broader view of the purpose of a business or organisation within society, and a move away from simplistic short-term answers, from a single bottom line that looks only at immediate economic profit. As Elkington later wrote (2018: 8) the "stated goal from the outset was system change — pushing toward the transformation of capitalism". Accordingly, our suggestion to use this model for a broader and more truthful evaluation of translation technologies and their consequences moves away from simplistic accounts that tell a deceptively straightforward story towards a more complex and nuanced representation of automation that moves beyond efficiency and short-term gain.

4.1 People

The first consideration for people as users of translation technology is quality, and when quality is decoupled from users, for example by using automatic evaluation metrics that do not correlate well with human judgement, the use of those metrics should be called into question. However, in this article we link quality to performance, as explained in Section 4.3. Reduced translation quality will introduce risk to users, but a key element of the triple bottom line is that each part is crucial and interlinked. As Abdallah (2014) proposes in her three-dimensional model of translation quality, the social quality of translation, asking who does what under what circumstances, is intertwined with the quality of the process and of the final product.

People are also the source of training data for contemporary machine learning technologies (Gaspari 2022). Translators may have contractually agreed (or not) to allow their work to be repurposed for MT system training, but the use of webcrawling for data acquisition is currently standard, without any real legal basis. Some artists have begun to use cloaking tools to prevent their work being used as training data (Shan *et al.* 2023) and website administrators may use the robots.txt file to try to prevent crawling of their online data (if the crawler has been set to respect this instruction). As human data is the gold standard, and machine-generated data begins to appear frequently on the internet (causing deteriorating quality if used for training machine learning systems) human data and "data about human interactions with LLMs will be increasingly valuable" (Shumailov *et al.* 2023, 2). Organisations will want to avoid the "ouroboros effect" (Moorkens

2023a: 18) of systems being trained on their own output, while being cognisant of the rights of people not to have their data used without consent.

Translators and workers within digital platforms often have little choice in whether or not their data is collected and used for MT system training. Their work may even be intended solely to produce data for training or to provide human feedback in order to reduce bias or harms in system output (Ouyang *et al.* 2022). Workers within digital platforms are also often prone to algorithmic management, with little or no human oversight, and no explanation or accountability for decision-making. While this can reduce unpaid work in searching for jobs within a platform, there is the necessity to conform to what the system rewards, with timeliness and polite communication replacing the need to maximise translation quality, and translation norms being displaced by algorithmic norms (Moorkens 2023b).

The increased automation within the translation industry, especially within digital platforms (used regularly by 89% of translator respondents to a survey by Pielmeier and O'Mara (2020)), presents a risk to organisational sustainability as mentioned in Section 3. According to Wheelen *et al.* (2018), companies that focus on business sustainability are rewarded by reduced staff turnover and increased employee effort. They also find that companies that focus on cost, replacing workers with automation where possible, demonstrate the opposite effect. The recommendation is thus to use automation to add value or to diversify the business offering, automating what was not being done before, and making sure to foreground satisfying, motivating work. A sustainable work system is "aimed at the regeneration of the resources it utilizes – human, social, material, and natural resources" (Docherty *et al.* 2008: 4), and the balancing of these and the needs of various stakeholders is difficult and will require constant evaluation and recalibration. These natural resources are where we turn to in the following section.

4.2 Planet

The most obvious concern for environmental sustainability is the energy requirement for automation technologies, sometimes referred to as compute costs. The contribution by Strubell *et al.* (2019) raised awareness of the carbon emissions related to training machine learning models based on the Transformer neural architecture (by far the most common architecture for MT and text generation in general at the time of writing; Vaswani *et al.* 2017), with the graphical processing unit (GPU) emissions for training a large model equated to the output of 1.5 cars over the 20-year lifetime of those cars, without even including the power and cooling requirements for the whole computer, hence their recommendation to "prioritize computationally efficient hardware and algorithms" (*ibid.*: 3646). Luccioni *et al.* (2023) also highlight the additional emissions related to generative AI when compared with task-specific systems, particularly at the training stage.

Wu *et al.* (2021) include the manufacturing and operational costs of equipment in their calculations and report 25% increases in efficiency of machine learning models over a two-year period. There has also been a trend in recent years towards pre-trained models, rather than always training from scratch (see Dođru and Moorkens in this issue), and researchers such as Jooste *et al.* (2022) have examined ways to make machine learning systems more efficient. This promising work should, however, be placed in the context of increasingly large language models. Shterionov and Vanmassenhove (2022) highlight the difference in emissions based on energy sources and location, with training in Ireland utilising more renewables than fossil fuel-based power in the Netherlands. Dodge *et al.* (2022) also add consideration of the time of day for training and the data centre location.

The manufacturing and disposal risks to workers in particular, and more broadly to citizens, are part of a larger sustainability problem across the life cycle of technology as detailed early on by Williams (2011), well before the latest AI-led technology breakthroughs. Problems include the release of toxins in incineration, release of harmful metals and compounds when computer equipment is (inadvertently) mixed with general waste, and burning of some materials within informal recycling processes. Due to high labour costs and environmental restrictions in many countries, these informal processes tend to be pushed to developing regions for reasons of cost effectiveness and, often, more permissive laws. Again, we have no easy answers as to whether a technology is worth developing or not, or – more specifically – under what circumstances it could or should be developed, balancing its benefits and costs. We would, in principle, agree with Williams (2011: 357) who concludes that “[u]nderstanding the interaction of ICTs with economic and social systems presents significant and interdisciplinary methodological challenges”.

Ethical best practice in research has moved on from the edict to ‘do no harm’ to actively do good. Antonopoulos *et al.* (2020) explain how machine learning technology can automate demand response in energy providers, routing energy quickly and effectively, managing costs and resources. This forms part of the wide-ranging proposals for machine learning to improve sustainability by Rolnick *et al.* (2023): ideas include modelling emissions and forecasting demand in electricity provision, reducing the need for standby generators, using sensors to reduce waste and harmful emissions, improving transport efficiency, optimising farms and industries, and (perhaps most ambitiously) modelling and managing emissions at a global level. While some of these ideas are currently fanciful, we would argue that those with concrete impact should be examined and balanced against the immediate negative impacts.

4.3 Performance

The traditional bottom line for technology is performance, which in translation is quality. As detailed in Castilho *et al.* (2018), there are many

ways to monitor and evaluate translation quality, and these may be appropriate for some products, processes, and scenarios but not for others. Quality is “never absolute but depends on both context and situation” (Drugan *et al.* 2018: 42). There is undoubtedly a place for automatic evaluation where human evaluation is too slow or expensive. What is key is not to overgeneralise based on a narrow set of results, for example using sensationalist terms such as the already mentioned ‘human parity’ or ‘superhuman translation’ based on a small-scale evaluation of translation in a single language pair and domain where not all independent variables are given due consideration and controlled properly (Läubli *et al.* 2020). Researchers such as Freitag *et al.* (2020) and Kocmi *et al.* (2021) advise against the use of BLEU (Papineni *et al.* 2002), the most popular automatic evaluation metric in the MT field during the last two decades, on the basis that its overuse has impeded MT development. However, the best available metrics show weaknesses too, such as for named entities or numbers (Amrhein and Sennrich 2022). This is an area of active research, with new automatic evaluation metrics being regularly proposed (e.g. Fernandes *et al.* 2023; Guerreiro *et al.* 2023) and new standards for translation evaluation due to be published by ASTM and ISO³.

Reduced translation quality could not only introduce risk, but could also push the effort in comprehension towards the end user, as already noted by Pym (2012). There may be problems with mistranslation, along with problems of hegemonic, gender, or racial bias for the reader to disentangle. However, we must also be mindful that quality needs differ based on circumstances, as argued by Way (2013). For example, in crisis settings, the “gold-standard expectations enshrined in ISO standards and codes of conduct seem unattainable” (Federici and O’Brien 2019: 11), but one can argue that in critical situations an imperfect translation is better than no translation at all, especially if it meets some urgent and serious needs of people in danger without causing harm. We might usually want excellent quality translation in general, but at times reduced quality translation is better and less risky than none at all, and may in fact be the only viable or possible option for very solid reasons.

If a (translation) technology is to be useful, there will have to be a clear benefit in performance or an economic motivation to its deployment. However, this is where gains in efficiency need to be placed in the context of the human, social, and environmental spheres. These are not to be played off against one another, but to be considered in the round and in how they contribute to overall sustainability. Elkington (1997: 316) stresses the need to change from an extractive approach to “modes which, over time, actively rebuild economic, environmental, human and social capital”. In business, the COVID-19 pandemic demonstrated the lack of resilience in ‘just in time’ supply chains that foregrounded efficiency above all else (Remko 2020). Similarly, a short-term focus on performance and cost without a similar corresponding focus on sustainability is unlikely to lead to real, long-term benefits. By all means, efficiency in comparison with

benchmarks is important, but not at all costs.

5. Six questions for modelling a triple bottom line for translation automation and sustainability

Putting this technology triple bottom line into practice is more complex and involved than carrying out a single static measurement of the quality of a tool or technology. Rather, it should involve continuous evaluation and recalibration of the technology and its deployment. In this section, we propose six questions – two each on people, planet, and performance – that could form a useful basis for analysis before and after the development and introduction of a technology, which we feel are particularly relevant to translation technologies in the context of automation and sustainability in the age of AI, as per the focus of this special issue.

Who are the stakeholders and how does this technology affect them?

The most obvious stakeholders in translation are source text authors, translators, clients and end users. How will this technology change translators' work and if it does, will this be positive or negative? What will be the impact on clients? Will the text or product for the end users change for the better or worse, and might it expose them to risks? What about other stakeholders "to whom the organisation has a moral obligation, an obligation of stakeholder fairness" (Philips 2003: 30) such as the people who create or provide the data, the company that develops or uses the tool and its employees? Will they be treated more or less fairly as a result of this technology? How will society be affected more broadly? There is an argument that the use of AI technologies might have broader repercussions on society, or that overuse of MT might negatively affect a minoritized language (see Aranberri and Iñurrieta in this issue).

What are the consultation, training, and feedback needs when introducing this technology?

Cadwell *et al.* (2018: 317) argue that "translators ought to be included in the change process from the very beginning" when introducing MT. This would seem to be appropriate for technology more broadly, as introduction via discussion and ideally based on consensus is more likely to be accepted than unilateral imposition of a tool, which might lead to tension, if not outright rejection. Cadwell *et al.* (2018: 317) highlight the importance of agency here, not giving the impression that the technology has taken precedence and is "inevitable, no matter how unfitting it might be for the task at hand". Bywood *et al.* (2017), Vieira and Alonso (2018) and others have also stressed the importance of training, but there are many ways to carry out training (see Bell *et al.* 2017) and it needs to be carefully planned. Cadwell *et al.* (2018: 317) also recommend that translators "monitor and improve their own quality and productivity in an engaged way, without

feeling externally monitored”. Ongoing evaluation and recalibration should lead to improved and more effective interaction between people and technology.

What is the environmental impact of developing and using the technology?

How can it be mitigated or ideally offset?

Measuring energy requirements is difficult, with GPU time possibly the most straightforward way to approximate the energy consumption of a machine learning technology. As Shterionov and Vanmassenhove (2022) note, emissions will vary depending on the energy source, e.g. a cloud tool hosted in Iceland using renewable energy will emit less carbon dioxide than one hosted in Malta or Luxembourg⁴. Will the technology save energy elsewhere? During the COVID-19 pandemic, many meetings became video conferences, facilitated by cloud-based tools hosted in data centres. This also meant that participants made fewer journeys. However, the overall costs and benefits are very difficult to measure and compare against each other, also considering the shifts in gains (and losses) across individuals, companies, organisations and institutions.

Is the quality provided by the automation technologies appropriate for the users and purpose?

Is performance efficient and consistent?

Koby and Melby (2013: 178) write that a quality translation “demonstrates required accuracy and fluency for the audience and purpose and complies with all other negotiated specifications, taking into account end-user needs”. There are lots of ways to measure translation quality, both automatically and using human evaluation (see the contributions to Moorkens *et al.* (2018) for many examples). Additionally, if a technology is slow or inconsistent, it will produce erratic output and frustrate users. What is important is that the evaluation method is appropriate for the users and purpose. We also note that in the MT field, evaluation does not normally take the translation brief into account.

Providing answers to these questions that we propose here will give a far more nuanced picture of a technology than a static measure of performance such as a BLEU score. However, we are aware that benchmarking, comparing, and communicating this holistic result will be difficult, introducing new challenges in a world focused on short-term easily defined gains.

6. The contributions to this special issue

Following this position paper, nine articles in this special issue address different aspects of translation automation and sustainability, beginning

with four contributions that focus on PE. Rico Pérez revisits PE guidelines, arguing that the conventional distinction between light and full PE is no longer fit for purpose due to the generally improved quality of state-of-the-art NMT. She proposes a redefined and replicable set of PE guidelines to mitigate factors that may induce tension and to discourage negative attitudes towards PE.

Dai and Liu investigate the impact of source-text readability on PE effort for English-Chinese NMT. Their results show that readability has a significant effect on cognitive effort, and that readability metrics can predict PE effort to a certain degree, although no single formula was able to predict all of the effort indicators, suggesting that a combination of metrics may be useful for effort prediction. Nitzke *et al.* revisit their previously published decision tree (Nitzke, Hansen-Schirra, and Canfora 2019) to aid decisions about a project's suitability for MT or PE. Based on interviews with 19 stakeholders, the authors reduce their model to four main decisions concerning the suitability of the source text, the reliability of the MT output, the purpose of the final text, the quality requirements for the final text, and social sustainability within translation workflows.

Guerberof-Arenas, Valdez and Dorst's study engages master's students of translation at two Dutch universities, who translated and post-edited literary texts in English at the beginning and at the end of taking translation technology modules. The study shows that the students tended to be more creative when translating unaided, although they made significantly fewer errors in PE, especially at the start of the training, than when translating manually. The study provides valuable insights on the complex connections between PE training, creativity, and translation proficiency.

The following two articles relate to subtitling. Guerberof-Arenas, Moorkens and Orrego-Carmona's more than 200 English-speaking participants were surveyed on their narrative engagement, enjoyment, and translation reception of Latin American Spanish to English subtitles, comparing MT, PE, and human translation. The largest disparity in translation reception is between MT and PE, but the study also reveals that achieving publishable PE subtitles requires a substantial number of edits, as indicated by high HTER scores. The authors emphasise the importance of considering factors such as time and remuneration for PE tasks, to ensure acceptable quality and sustainable work processes.

Tamayo and Ros Abaurrea analyse speech recognition software for the intralingual subtitling of news programmes in Basque. Their evaluation identifies room for improvement in recognition, particularly regarding the handling of punctuation, the recognition of proper nouns, and speaker identification. Despite these weaknesses, Basque speech recognition seems promising, considering that speech recognition is still in the early stages of development for a low-resourced language like Basque. Aranberri and Iñurrieta also consider language sustainability for the Basque-speaking community, surveying translators and interpreters, language professionals, and general users to investigate their current and expected use of Basque-

Spanish MT. Respondents are generally positive about the future for MT and Basque, with some concerns expressed about the impact that MT might have on the quality of published Basque texts, particularly administrative texts originating in Spanish.

The final two articles relate to translation workflows. Dođru and Moorkens investigate the impact of data augmentation using TMs for desktop MT fine-tuning in OPUS-CAT, assessing the utility of desktop MT for professional translators by fine-tuning MT engines in three language pairs (English → Turkish, English → Spanish, and English → Catalan) with localization corpora of varying sizes. The results demonstrate promising improvements in translation quality across all three language pairs, underscoring the potential of desktop MT applications to deliver high-quality translations while offering benefits such as privacy, confidentiality, and reduced computation power usage, the latter thanks to the use of pre-trained MT models. Finally, Silva *et al.* revisit the Multidimensional Quality Metric (MQM) evaluation framework, originally oriented towards European languages, and propose an amended error typology to better suit East Asian languages such as Mandarin, Japanese and Korean. They also propose a Quality Estimation (QE) method to predict the MQM scores of MT outputs at scale, showing a fair correlation with the human judgement.

7. Conclusions and avenues for future work

The work of Elkington (1997: 141) and others has been influential in changing mindsets in business, but as he wrote in 1997, “the eco-resource challenge may prove to be the relatively easy part of the sustainability transformation, while the socio-economic challenge looks likely to be more intractable”. A broader notion of sustainability will be important so that people are not left behind. This is the thinking behind the UN SDGs, even if these might compare unevenly (as noted by Buts *et al.* 2023). Sætra (2021: 16) argues that we need to look at AI in context and that “doing so shows that it is intimately tied to severe threats to most of the SDGs”. If a triple-bottom line evaluation of translation technology is to be effective, it needs to be able to identify routes of research and development that are not sustainable and thus should not be pursued.

In revisiting the triple-bottom line, Elkington (2018: 10) sees radical intent as necessary to spur “the regeneration of our economies, societies, and biosphere”. This is inarguably more important now than ever before and can only become more urgent, particularly with the young generations appearing to take a keen interest in these issues, especially with regard to environmental concerns. While sustainability as a topic has rarely been directly and specifically addressed in translation studies research, this special issue, along with pioneering work by Cronin (2017), the special section edited by Buts *et al.* (2023) and other recent work addressing sustainability (e.g. Todorova 2022) shows that researchers in translation studies are interested in and concerned about sustainability. However, as discussed in Section 1, sustainability does not appear to have been the

foremost consideration in the development and deployment of automation technologies in translation. The model proposed in this position paper may not be the single best way forward for evaluation of automation technologies, but it is increasingly clear that it is time for new ideas and that the old methods that focus only on isolated measures of performance need to evolve and be enriched.

References

- **Abdallah, Kristiina.** (2014). "Social Quality. Key to Collective Problem Solving in Translation Production Networks". *Proceedings of the 6th Riga Symposium on Pragmatic Aspects of Translation*, Riga, 5–18.
- **Agorni, Mirella and Giuseppe De Bonis,** eds. (2022). *Collaboration in translation: from training to platforms and publishing*. Napoli: Paolo Loffredo.
- **Amrhein, Chantal and Rico Sennrich.** (2022). "Identifying Weaknesses in Machine Translation Metrics Through Minimum Bayes Risk Decoding: A Case Study for COMET". <https://arxiv.org/abs/2202.05148>.
- **Antonopoulos, Ioannis, Valentin Robu, Benoit Couraud, Desen Kirli, Sonam Norbu, Aristides Kiprakis, David Flynn, Sergio Elizondo-Gonzalez and Steve Wattam** 2020. "Artificial Intelligence and Machine Learning Approaches to Energy Demand-Side Response: A Systematic Review". *Renewable and Sustainable Energy Reviews* 130: 109899.
- **Bell, Bradford S., Scott Tannenbaum, Kevin J. Ford, Raymond A. Noe and Kurt Kraiger** (2017). "100 Years of Training and Development Research: What We Know and Where We Should Go." *Journal of Applied Psychology* 102(3): 305–23. <https://doi.org/10.1037/apl0000142>
- **Brevini, Benedetta** (2022). *Is AI Good for the Planet?* Cambridge, UK: Polity Press.
- **Brundtland, Gro Harlem** (1987). *Report of the World Commission on Environment and Development: Our Common Future*. Geneva: United Nations.
- **Buts, Jan, Hanna Pięta, Laura Ivaska and James Hadley** (2023). "Indirect Translation and Sustainable Development". *Translation Spaces* 12(2): 167–176. <https://doi.org/10.1075/ts.00031.but>
- **Cadwell, Patrick, Sheila Castilho, Sharon O'Brien and Linda Mitchell** (2016). "Human Factors in Machine Translation and Post-Editing among Institutional Translators" *Translation Spaces* 5(2): 222–43. <https://doi.org/10.1075/ts.5.2.04cad>
- **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* 26(3): 301–21. <https://doi.org/10.1080/0907676X.2017.1337210>
- **Calhoun, Gloria** (2022). "Adaptable (Not Adaptive) Automation: Forefront of Human–Automation Teaming" *Human Factors: The Journal of the Human Factors and Ergonomics Society* 64(2): 269–77. <https://doi.org/10.1177/00187208211037457>
- **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, Stephen Doherty, Federico Gaspari and Joss Moorkens** (2018). 'Approaches to Human and Machine Translation Quality Assessment'. Joss Moorkens, Sheila Castilho, Federico Gaspari and Stephen Doherty (eds) *Translation Quality Assessment* Cham: Springer 9–38. https://doi.org/10.1007/978-3-319-91241-7_2
- **Castilho, Sheila, Clodagh Mallon, Rahel Meister and Shengya Yue** (2023). "Do Online Machine Translation Systems Care for Context? What about a GPT Model?" *Proceedings of the 24th Annual Conference of the European Association for Machine Translation*, Tampere, Finland.
- **Cronin, Michael** (2017). *Eco-Translation: Translation and Ecology in the Age of the Anthropocene*. Abingdon: Routledge.
- **Cronin, Michael** (2019). "Translation, Technology and Climate Change" Minako O'Hagan (ed.) (2019) *The Routledge Handbook of Translation and Technology* Abingdon: Routledge, 516–30. <https://doi.org/10.4324/9781315311258-30>
- **Docherty, Peter, Mari Kira and Abraham B. Shani** (2008). "What the World Needs Now Is Sustainable Work Systems" Peter Docherty, Mari Kira, and Abraham B. Shani (eds) (2008) *Creating Sustainable Work Systems: Developing Social Sustainability*, Abingdon: Routledge.
- **Dodge, Jesse, Taylor Prewitt, Remi Tachet Des Combes, Erika Odmark, Roy Schwartz, Emma Strubell, Alexandra Sasha Luccioni, Noah A. Smith, Nicole DeCaro, and Will Buchanan.** (2022). "Measuring the Carbon Intensity of AI in Cloud Instances". In *2022 ACM Conference on Fairness, Accountability, and Transparency*, Seoul Republic of Korea: ACM, 1877–94. <https://dl.acm.org/doi/10.1145/3531146.3533234> (July 15, 2022).
- **Drugan, Joanna, Ingemar Strandvik and Erkka Vuorinen** (2018). "Translation Quality, Quality Management and Agency: Principles and Practice in the European Union Institutions". Joss Moorkens, Sheila Castilho, Federico Gaspari and Stephen Doherty (eds) *Translation Quality Assessment* Cham: Springer, 39–68. https://doi.org/10.1007/978-3-319-91241-7_3
- **ELIS Research** (2023). *European Language Industry Survey 2023: Trends, Expectations and Concerns of the European Language Industry*. ELIA, EMT, EUATC, FIT Europe, GALA, LIND, Women In Localization. <https://elis-survey.org/wp-content/uploads/2023/03/ELIS-2023-report.pdf>.
- **Elkington, John** (1997). *Cannibals with Forks: The Triple Bottom Line of 21st Century Business*. Oxford: Capstone.
- **Elkington, John** (2018). "25 Years Ago I Coined the Phrase "Triple Bottom Line." Here's Why It's Time to Rethink It". *Harvard Business Review*. <https://hbr.org/2018/06/25-years-ago-i-coined-the-phrase-triple-bottom-line-heres-why-im-giving-up-on-it>.
- **Federici, Federico M. and Sharon O'Brien** (2019). "Cascading Crises". Federico M. Federici and Sharon O'Brien (eds) (2019) *Translation in Cascading Crises*, Abingdon: Routledge, 1–22. <https://www.taylorfrancis.com/books/9781000657074/chapters/10.4324/9780429341052-1>
- **Fernandes, Patrick, Daniel Deutsch, Mara Finkelstein, Parker Riley, André F. T. Martins, Graham Neubig, Ankush Garg, Jonathan H. Clark, Markus Freitag and Orhan Firat** (2023). "The Devil Is in the Errors: Leveraging Large Language Models for Fine-Grained Machine Translation Evaluation". <http://arxiv.org/abs/2308.07286>

- **Freeman, R. Edward** (1984). *Strategic Management: A Stakeholder Approach*. Boston: Pitman.
- **Freitag, Markus, David Grangier and Isaac Caswell** (2020). "BLEU Might Be Guilty but References Are Not Innocent". *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online: Association for Computational Linguistics, 61–71. <https://aclanthology.org/2020.emnlp-main.5>.
- **Gaspari, Federico** (2022). "Covert Collaboration among Translators through Data-Driven Technologies and Digital Resources: What You Get Is Not What You See". Mirella Agorni and Giuseppe De Bonis (eds) (2022) *Collaboration in Translation: From Training to Platforms and Publishing*, Naples: Paolo Loffredo Editore, 43–61.
- **Guerreiro, Nuno M., Ricardo Rei, Daan van Stigt, Luisa Coheur, Pierre Colombo and André F.T. Martins** (2023). "xCOMET: Transparent Machine Translation Evaluation through Fine-Grained Error Detection" <http://arxiv.org/abs/2310.10482>
- **Hassan, Hany, Anthony Aue, Chang Chen, Vishal Chowdhary, Jonathan Clark, Christian Federmann, Xuedong Huang, Marcin Junczys-Dowmunt, William Lewis, Mu Li, Shujie Liu, Tie-Yan Liu, Renqian Luo, Arul Menezes, Tao Qin, Frank Seide, Xu Tan, Fei Tian, Lijun Wu, Shuangzhi Wu, Yingce Xia, Dongdong Zhang, Zhirui Zhang and Ming Zhou** (2018). "Achieving Human Parity on Automatic Chinese to English News Translation". <http://arxiv.org/abs/1803.05567>
- **Hendy, Amr, Mohamed Abdelrehim, Amr Sharaf, Vikas Raunak, Mohamed Gabr, Hitokazu Matsushita, Young Jin Kim, Mohamed Afify and Hany Hassan Awadalla** (2023). "How Good Are GPT Models at Machine Translation? A Comprehensive Evaluation". <https://arxiv.org/abs/2302.09210>.
- **Herbert, Sarah, Félix do Carmo, Joanna Gough and Anu Carnegie-Brown** (2023). "From Responsibilities to Responsibility: A Study of the Effects of Translation Workflow Automation" *Journal of Specialised Translation* 40, 9-35.
- **Herzberg, Frederick** (1976). "One More Time: How Do You Motivate Employees?" Michael M. Gruneberg (ed.) (1976) *Job Satisfaction — A Reader*, ed.. London: Palgrave Macmillan 17–32. http://link.springer.com/10.1007/978-1-349-02701-9_2
- **Hutchins, William John and Harold L. Somers** (1992). *An Introduction to Machine Translation*. London: Academic Press.
- **Jooste, Wandri, Rejwanul Haque and Andy Way** (2022). "Knowledge Distillation: A Method for Making Neural Machine Translation More Efficient" *Information* 13(2): 88.
- **Kocmi, Tom, Christian Federmann, Roman Grundkiewicz, Marcin Junczys-Dowmunt, Hitokazu Matsushita and Arul Menezes** (2021). "To Ship or Not to Ship: An Extensive Evaluation of Automatic Metrics for Machine Translation". *Proceedings of the Sixth Conference on Machine Translation*, Online: Association for Computational Linguistics, 478–94. <https://aclanthology.org/2021.wmt-1.57>
- **Kocmi, Tom and Christian Federmann** (2023). "Large Language Models Are State-of-the-Art Evaluators of Translation Quality". <https://arxiv.org/abs/2302.14520>.
- **Läubli, Samuel, Sheila Castilho, Graham Neubig, Rico Sennrich, Qinlan Shen, Antonio Toral** (2020). "A Set of Recommendations for Assessing Human–Machine Parity in Language Translation". *Journal of Artificial Intelligence Research* 67. <https://jair.org/index.php/jair/article/view/11371>

- **Lommel, Arle** (2021). *Responsive Machine Translation: The Next Frontier for MT*. Boston: CSA Research. <https://csa-research.com/Blogs-Events/Blog/responsive-machine-translation>
- **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. http://cracker-project.eu/wp-content/uploads/Europes_Leading_Role_in_MT.pdf.
- **Luccioni, Alexandra Sasha, Yacine Jernite and Emma Strubell** (2023). "Power Hungry Processing: Watts Driving the Cost of AI Deployment?" <http://arxiv.org/abs/2311.16863>
- **Melé, Domènec** (2009). *Business Ethics in Action: Seeking Human Excellence in Organizations*. New York: Palgrave Macmillan.
- **Messick, Samuel** (1989). "Meaning and Values in Test Validation: The Science and Ethics of Assessment". *Educational Researcher* 18(2): 5–11.
- **Moorkens, Joss** (2020). "'A Tiny Cog in a Large Machine': Digital Taylorism in the Translation Industry". *Translation Spaces* 9(1): 12–34. <https://doi.org/10.1075/ts.00019.moo>
- **Moorkens, Joss** 2023a. "Generative Tools Using Large Language Models and Translation". *Marie Curie Alumni Association Newsletter* 35: 16–18. <https://www.mariecuriealumni.eu/newsletters/35th-mcaa-newsletter/special-issue-generative-tools-using-large-language-models-and>
- **Moorkens, Joss** 2023b. "'I Am Not a Number': On Quantification and Algorithmic Norms in Translation" eds. Michael Tieber and Stefan Baumgarten. *Perspectives*. <https://doi.org/10.1080/0907676X.2023.2278536>
- **Nitzke, Jean, Silvia Hansen-Schirra and Carmen Canfora** (2019). "Risk Management and Post-Editing Competence". *The Journal of Specialised Translation* 31: 239–59.
- **O'Brien, Sharon** (2023). 'Human-Centered Augmented Translation: Against Antagonistic Dualisms'. *Perspectives*. <https://doi.org/10.1080/0907676X.2023.2247423>
- **Ouyang, Long, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike and Ryan Lowe** (2022). "Training Language Models to Follow Instructions with Human Feedback". <http://arxiv.org/abs/2203.02155>
- **Papineni, Kishore, Salim Roukos, Todd Ward, and Wei-Jing Zhu** (2002). "Bleu: A Method for Automatic Evaluation of Machine Translation". Pierre Isabelle, Eugene Charniak, and Dekang Lin (eds) *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Philadelphia: Association for Computational Linguistics, 311–18. <https://aclanthology.org/P02-1040>
- **Parasuraman, R., T.B. Sheridan and C.D. Wickens** (2000). "A Model for Types and Levels of Human Interaction with Automation". *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 30(3): 286–97.
- **Paulsen Christensen, Tina, Kristine Bundgaard, Anne Schjoldager and Helle**

- Dam Jensen** (2022). "What Motor Vehicles and Translation Machines Have in Common - a First Step towards a Translation Automation Taxonomy". *Perspectives* 30(1): 19–38.
- **Phillips, Robert** (2003). *Stakeholder Theory and Organizational Ethics*. 1st ed. San Francisco: Berrett-Koehler.
 - **Pielmeier, Hélène and Paul O'Mara** (2020). *The State of the Linguist Supply Chain*. Boston: Common Sense Advisory.
 - **Pitman, Jeff** (2021). "Google Translate: One Billion Installs, One Billion Stories". *The Keyword*. <https://blog.google/products/translate/one-billion-installs/>.
 - **Porter, Michael E.** (1980). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. New York: Free press.
 - **Pym, Anthony** (2012). *On Translator Ethics: Principles for Mediation between Cultures*. Amsterdam: John Benjamins. <http://www.jbe-platform.com/content/books/9789027272997>
 - **Remko, Van Hoek** (2020). "Research Opportunities for a More Resilient Post-COVID-19 Supply Chain – Closing the Gap between Research Findings and Industry Practice". *International Journal of Operations & Production Management* 40(4): 341–55.
 - **Rolnick, David, Priya L. Donti, Lynn H. Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, Andrew Slavin Ross, Nikola Milojevic-Dupont, Natasha Jaques, Anna Waldman-Brown, Alexandra Luccioni, Tegan Maharaj, Evan D. Sherwin, S. Karthik Mukkavilli, Konrad P. Kording, Carla Gomes, Andrew Y. Ng, Demis Hassabis, John C. Platt, Felix Creutzig, Jennifer Chayes and Yoshua Bengio** (2023). "Tackling Climate Change with Machine Learning". *ACM Computing Surveys* 55(2): 1–96.
 - **Rothwell, Andrew, Joss Moorkens, María Fernández-Parra, Joanna Drugan and Frank Austermuehl** (2023). *Translation Tools and Technologies: A Practical Guide for Students and Translators*. Abingdon: Routledge.
 - **Sætra, Henrik Skaug** (2021). "AI in Context and the Sustainable Development Goals: Factoring in the Unsustainability of the Sociotechnical System". *Sustainability* 13(4): 1738.
 - **Sánchez-Gijón, Pilar and Leire Palenzuela-Badiola** (2023). "Analysis and Evaluation of ChatGPT-Induced HCI Shifts in the Digitalised Translation Process". Constantin Orăsan, Ruslan Mitkov, Gloria Corpas Pastor and Johanna Monti (eds) (2023) *Proceedings of Hit-IT: The International Conference on Human-Informed Translation and Interpreting Technology*. Naples: Incoma, 227–67.
 - **Schwartz, Roy, Jesse Dodge, Noah A. Smith and Oren Etzioni** (2020). "Green AI". *Communications of the ACM* 63(12): 54–63. <https://dl.acm.org/doi/10.1145/3381831>
 - **Shan, Shawn, Jenna Cryan, Emily Wenger, Haitao Zheng, Rana Hanocka and Ben Y. Zhao** (2023). "Glaze: Protecting Artists from Style Mimicry by Text-to-Image Models". <https://arxiv.org/abs/2302.04222>
 - **Shterionov, Dimitar and Eva Vanmassenhove** (2022). "The Ecological Footprint of Neural Machine Translation Systems". <http://arxiv.org/abs/2202.02170>
 - **Shumailov, Ilia, Zakhar Shumaylov, Yiren Zhao, Yarin Gal, Nicolas Papernot**

- and Ross Anderson** (2023). "The Curse of Recursion: Training on Generated Data Makes Models Forget". <https://arxiv.org/abs/2305.17493>
- **Stoddard, Isak, Kevin Anderson, Stuart Capstick, Joanna Depledge, Keri Facer, Frédéric Hache, Claire Hoolohan, Martin Hultman, Niclas Hällström, Sivan Kartha, Sonja Klinsky, Magdalena Kuchler, Eva Lövbrand, Naghmeh Nasiritousi, Peter Newell, Glen Peters, Youba Sokona, Clive L. Spash, Andy Stirling, Matthew Stilwell and Mariama Williams** (2021). "Three Decades of Climate Mitigation: Why Haven't We Bent the Global Emissions Curve?" *Annual Review of Environment and Resources* 46(1): 653–89.
 - **Strubell, Emma, Ananya Ganesh and Andrew McCallum** (2019). "Energy and Policy Considerations for Deep Learning in NLP". <http://arxiv.org/abs/1906.02243>
 - **Thompson, Vikki, Dann Mitchell, Gabriele C. Hegerl, Matthew Collins, Nicholas J. Leach and Julia M. Slingo** (2023). "The Most At-Risk Regions in the World for High-Impact Heatwaves". *Nature Communications* 14(1): 2152.
 - **Todorova, Marija** (2022). "The Role of Translation in Environmental Protection: An Inclusive Approach". *The Translator* 28(4): 415–28.
 - **Toral, Antonio, Sheila Castilho, Ke Hu and Andy Way** (2018). "Attaining the Unattainable? Reassessing Claims of Human Parity in Neural Machine Translation". *Proceedings of the Third Conference on Machine Translation: Research Papers*, Brussels, Belgium: Association for Computational Linguistics, 113–23. <https://aclanthology.org/W18-6312>
 - **UN General Assembly** (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*. United Nations. <https://www.refworld.org/docid/57b6e3e44.html>.
 - **Vagia, Marialena, Aksel A. Transeth and Sigurd A. Fjerdings** (2016). "A Literature Review on the Levels of Automation during the Years. What Are the Different Taxonomies That Have Been Proposed?" *Applied Ergonomics* 53: 190–202.
 - **Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser and Illia Polosukhin** (2017). "Attention Is All You Need". <http://arxiv.org/abs/1706.03762>
 - **Vieira, Lucas Nunes, Minako O'Hagan and Carol O'Sullivan** (2021). "Understanding the Societal Impacts of Machine Translation: A Critical Review of the Literature on Medical and Legal Use Cases". *Information, Communication & Society* 24(11): 1515–32.
 - **Way, Andy** (2013). "Traditional and Emerging Use-Cases for Machine Translation". *Proceedings of Translating and the Computer 35*, (London, 28-29 November). <https://aclanthology.org/2013.tc-1.12>.
 - **Wheelen, Thomas L., J. David Hunger, Alan N. Hoffman and Charles E. Bamford** (2018). *Concepts in Strategic Management and Business Policy: Globalization, Innovation, and Sustainability*. Fifteenth edition. Harlow: Pearson.
 - **Williams, Eric** (2011). "Environmental Effects of Information and Communications Technologies". *Nature* 479(7373): 354–58.
 - **Wu, Carole-Jean, Ramya Raghavendra, Udit Gupta, Bilge Acun, Newsha Ardalani, Kiwan Maeng, Gloria Chang, Fiona Aga Behram, James Huang,**

Charles Bai, Michael Gschwind, Anurag Gupta, Myle Ott, Anastasia Melnikov, Salvatore Candido, David Brooks, Geeta Chauhan, Benjamin Lee, Hsien-Hsin S. Lee, Bugra Akyildiz, Maximilian Balandat, Joe Spisak, Ravi Jain, Mike Rabbat and Kim Hazelwood (2021). "Sustainable AI: Environmental Implications, Challenges and Opportunities". <https://arxiv.org/abs/2111.00364>

- **Yamada, Masaru** (2023). "Post-Editing and a Sustainable Future for Translators". Isabel Lacruz (ed.) (2023) *American Translators Association Scholarly Monograph Series*, Amsterdam: John Benjamins, 39–56. <https://benjamins.com/catalog/ata.xx.03yam>

Biographies

Joss Moorkens is an Associate Professor at the School of Applied Language and Intercultural Studies in Dublin City University (DCU), Challenge Lead at the ADAPT Centre, and member of DCU's Institute of Ethics, and Centre for Translation and Textual Studies. He is General Co-Editor of Translation Spaces with Dorothy Kenny and coauthor of the textbooks *Translation Tools and Technologies* (Routledge 2023) *Translation Automation* (Routledge 2024).

ORCID: 0000-0003-0766-0071



E-mail: joss.moorkens@dcu.ie

Sheila Castilho is an Assistant Professor in SALIS at Dublin City University. She worked as an Irish Research Council Research Fellow at the Adapt Centre on the DELA Project, which involved testing sentence-level metrics for document-level machine translation evaluation and establishing best practices. Sheila has actively contributed to various EU projects. Her research output includes over 40 publications, covering topics on translation technology, post-editing of MT, user evaluation of MT, and translators' perception of MT.

ORCID: 0000-0002-8416-6555



E-mail: sheila.castilho@dcu.ie

Federico Gaspari teaches English language and translation at the Department of Political Science of the University of Naples “Federico II” (Italy) and collaborates with the ADAPT Centre of Dublin City University (DCU, Ireland) on EU-funded international research projects. His main research interests include language and translation technologies, applied English linguistics, corpus linguistics and corpus-based translation studies, and he has published widely in these areas.

ORCID: 0000-0003-3808-8418



E-mail: federico.gaspari@unina.it

Antonio Toral is an Associate Professor in Language Technology at the University of Groningen, where he coordinates the Computational Linguistics research group. His research interests include the application of machine translation (MT) to literary texts, MT for under-resourced languages and the computational analysis of translations produced by machines and humans. Prior to starting a faculty position, he was a postdoctoral researcher and research fellow at Dublin City University, and before that a PhD student at the Universitat d'Alacant and at the Istituto di Linguistica Computazionale. He coordinated the Abu-MaTran project, which was flagged as a success story by the European Commission and won the best paper award at MT Summit 2019 for his work on post-editing.

ORCID: 0000-0003-2357-2960



Email: a.toral.ruiz@rug.nl

Maja Popović is a Research Fellow at ADAPT Centre at the School of Computing in Dublin City University (DCU). Her main interests are machine translation, multilingual natural language processing, as well as human and automatic evaluation methods for NLP. She has over 80 scientific publications in book chapters, journals, conferences and workshops. She is associate editor of the LREV (Language Resources and Evaluation) journal.

ORCID: 0000-0001-8234-8745



Email: maja.popovic@adaptcentre.ie

Notes

¹ Application Programming Interface.

² See Sætra (2021) for an analysis of AI using Sustainability Development Goals.

³ These are expected to appear at <https://www.astm.org/workitem-wk46396> and <https://www.iso.org/standard/80701.html> after publication.

⁴ See https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Renewable_energy_statistics