Moniz, Helena and Carla Parra Escartín (eds) (2023). *Towards Responsible Machine Translation: Ethical and Legal Considerations in Machine Translation*. Cham: Springer, pp. 233, €149.99 (hardback) / €39.99 (MyCopy Softcover). ISBN 978-3-031-14688-6.

This new volume in the *Machine Translation: Technologies and Applications* series is a very welcome addition to the ever-growing literature on crucial topics in Machine Translation, including Fair MT (Kenny et al. 2020) and MT literacy (Bowker and Buitrago Ciro 2019). With a focus on the ethical and legal considerations of developing and using machine translation systems, the various chapters in *Towards Responsible Machine Translation: Ethical and Legal Considerations in Machine Translation* edited by Helena Moniz and Carla Parra Escartín offer essential points of discussion for anyone working with MT, including developers, professionals in the language industry, researchers, translation educators and MT users outside of the field. Due to the growing prominence of MT in society and the need to raise citizens' awareness of MT (i.e. to enhance overall MT literacy levels), those teaching topics within MT and those developing the systems must lead from the front if we want responsible MT to be adopted as well as informed and critical citizens as end-users of technology in the future.

This edited volume consists of a foreword from the series editor Andy Way, who argues that this volume "fills an important gap in the field" (v) and deems it "essential reading for us all" (vi), with which I agree. Following this is an introduction by the volume's editors and ten chapters, which are split into three parts concerned with different dimensions of responsible Machine Translation (MT): Part I focuses on ethical, philosophical and legal aspects (4 chapters); Part II focuses on the end-user perspective (3 chapters); and Part III focuses on the societal impact of MT (3 chapters). Included at the beginning is also a useful list of abbreviations for ease of reference.

In the introduction, the editors Helena Moniz and Carla Parra Escartín address the topic of "responsible artificial intelligence", the key factors involved in achieving responsible MT as well as highlighting the complexity of doing so. They envisage this book volume as a useful starting point for analysing how responsible MT could be visualised in the (near) future.

The first two chapters in Part 1 consider the ethics of Machine Translation. In chapter 2, Wessel Reijers and Quinn Dupont aim to establish a connection between the ethics of MT and the field of philosophy of technology. They question what machines 'do' to translation (17) and conclude that machines do not translate words, but rather transcribe "symbolic notations" (18). To this end, they address two philosophical questions linked to logocentrism (the extent to which machines translate

https://doi.org/10.26034/cm.jostrans.2024.4731

the 'presence' of lived experience) and to phonocentrism (the extent to which machines transcribe the spoken word). Their critical introduction to the ethics of MT explores the notions of "sacrifice," "linguistic hospitality," and various virtues in the translation profession, and reflects on the extent to which MT can accommodate and maintain these characteristics. In chapter 3, Alexandros Nousias highlights concerns regarding parameters for Natural Language Processing (NLP) systems since they have ethical implications on end-users. Ethical considerations are crucial in the initial design, implementation and use phases of the technology. Therefore, alongside ethics, the author argues for a deep understanding of language as the basis for shaping the future of NLP, and in particular for avoiding issues such as demographic misrepresentation, modelling overgeneralisation and data exclusion. Developing ethical principles for "what we design, how we design it and what is the impact of such design" (46) allows for optimised end-user experiences.

The next two chapters tackle pressing issues related to the data used for corpus-based MT (CBMT) systems. In chapter 4, Mikel L. Forcada discusses the "complicated matter" (50) of authorship, licenses and copyright of data used to train current CBMT systems. This chapter builds nicely on previous work by Moorkens and Lewis (2019) on copyright and the reuse of translation for this purpose. Forcada asks whether "authors and translations should get additional compensation for [this] unintended use of their works" (50), since monolingual and parallel texts were "not created with MT in mind" (50). He notes, however, that "recoverability is virtually impossible in NMT output" (63), due to how the training data are "broken down to the level of words, subwords or even characters" (63), making the input of an individual translator unidentifiable and their overall contribution undetectable. The chapter offers an overview of how corpora as training data for MT systems are created, the problems associated with the licensing and publication of corpora, and it brings the reader on a deep dive into the world of copyright. What is clear from this chapter is that copyright and data ownership will remain crucial topics in the quest for a sustainable translation industry. In chapter 5, Miguel L. Lacruz Mantecón continues the discussion by looking at how the introduction and use of new technologies subsequently raise issues regarding ownership of rights over machine or "mechanical" (74) translations and regarding the use of translation data to ultimately improve MT algorithms. The chapter begins with a clear and succinct discussion of "translation as an intellectual work" (71-72) and how intellectual property laws from various countries regulate or approach the work of translation. As a rule, the translator holds the rights to their work. However, when MT and AI enter the frame, things look slightly different, and in turn raise questions about ownership and reuse of translations. The author concludes that copyright might be the solution to the unauthorised use of translations. To this end, he suggests an interesting proposal called "equitable compensation" (90), which is a payment distributed fairly to translators who create data for the development of MT systems.

The three chapters in Part II consider responsible MT from the end-user perspective. In chapter 6, Celia Rico and María del Mar Sánchez Ramos identify three key ethical dilemmas facing any translator working with machine translation post-editing (MTPE) and address these from the perspective of a translation ecosystem, adapted from Krüger's (2016a, 2016b) model. MTPE as an object of study is becoming more relevant than ever with the latest developments in neural MT. The need to explore human agency and ethical aspects in relation to MTPE thus becomes even more imperative. The authors analyse the post-editor's status, the post-editor's commitment to quality, digital ethics and the post-editor's responsibility. In doing so, they approach the post-editor as a figure situated within this ecosystem, with the translator at the centre and the artefacts (e.g. technology) and people (e.g. cooperation partners) around them. Questions arise regarding the actual position of the post-editor in the translation ecosystem, the ability of the post-editor to ensure quality in a field where current PE levels (e.g. full vs. light) will shortly become obsolete and where new metrics will be used, and regarding the levels of MT literacy skills that will be necessary among post-editors to conduct PE work in a digitally ethical manner.

In chapter 7, Ana Guerberof-Arenas and Joss Moorkens introduce two important concepts for modelling MT use cases: "level of automation" and "risk" – the lower the risk, the less human revision required. While first providing use cases of this continuum on various text types, they follow with results from two experimental environments to give voice to users of translation in three different modalities: human translation (HT), post-edited MT output (MTPE) and raw MT (MT). In the first one, in a technical environment, eye-tracking, semi-structured interviews and Retrospective Think Aloud protocols were used to gauge user feedback. The factors measured to analyse usability highlighted that the majority were not aware they were working in MT modality, although they also reported experiencing difficulties understanding the terms used. In the second one, user experience of creativity in translation was explored by comparing the same modalities. Online surveys recorded user feedback and the findings indicated that MT had a strong (negative) effect on the reader experience. Consequently, the authors call for the prioritisation of translation reception studies to ensure no long-lasting effects on society, which could occur if there is an overreliance on MT solutions. Finally, they call for stakeholders to step up and uphold their ethical responsibilities if the objective of MT is to offer more fair and optimal communication in our societies.

In chapter 8, Federico M. Federici, Christophe Declercq, Jorge Díaz Cintas and Rocío Baños Piñero highlight the positive role played by MT in crisis settings, and simultaneously call for the consideration of essential ethical questions posed by technological innovations (e.g. "communication in crisis settings pivots around notions of trust, credibility, and social equality"

[152]). They raise awareness around various challenges that hinder the efficient and appropriate application of technologies in such settings (e.g. quality check of data sets, the frequent use of English as a pivot language, risks of raw MT, crucial lack of translated audiovisual material). They also voice concerns around deploying translation and interpreting technologies uncritically in multilingual, crisis settings — while it might be a necessity, it still poses ethical challenges.

The three chapters in Part III focus on the importance of implementing responsible MT given the impact the tools currently have and will come to have on society. In chapter 9, Federico Bianchi, Tommaso Fornaciari, Dirk Hovy and Debora Nozza discuss two "Language Invariant Properties" (161): age and gender identity. In the authors' opinion, these properties should not change during the translation process, as, together with content, they contribute equally to the success of an utterance being communicated. It is argued that language is "more than just information content" (180) and carries important social aspects too. Yet, the current commercial MT systems tested in this chapter (Bing, DeepL and Google) fail in this regard. The findings show that messages translated using these systems reveal significant disparities in the "perceived demographics" (162) in the translations, which made the authors of the original texts "appear older" and "more male than they are" (162).

In chapter 10, Dimitar Shterionov and Eva Vanmassenhove focus on the ecological footprint of neural MT systems. With significant advancements in AI, including MT, the authors point to the need for more training data, computing power and memory to run these systems which, for instance, use graphics processing units instead of central processing units. This results in higher electrical power consumption and higher carbon dioxide emissions, both of which entail considerable environmental concerns. To counter this, the authors present *quantization* as one of several techniques for optimising models towards "greener MT" (209). They also call to action for researchers to seriously consider how more ecological solutions can be developed for MT, thus contributing to a more responsible and sustainable future with MT systems.

In chapter 11, Isabel Trancoso, Francisco Teixeira, Catarina Botelho and Alberto Abad highlight ethical concerns regarding spoken language technologies. What they present might be both insightful and disconcerting to the average reader: speech, a "biometric signal" (215), can determine several pieces of information about a user, including some more obvious ones such as age range and accent, and some not so evident ones such as personality traits and political opinions. Privacy and security issues become ever more pressing in relation to this technology, especially since there is a potential for misuse of speech, e.g. voice cloning, taken from applications run on cloud-based platforms. The authors thus call for "ethical speech processing" (230) and for speech to be "legally regarded as Personally

Identifiable Information" to help ensure that current user data protection (e.g. GDPR) guarantees user safety against ethical and legal issues posed by the rise in speech-to-speech MT systems.

The strengths of this book are many: the flow of the texts, moving from the more theoretical questions and essential reflections to discuss the legal aspects of the technology and ending with user-studies that highlight current issues facing societies today and in the (near) future; the fact that each chapter can be read as a standalone article and that state-of-the-art research is referenced throughout, allowing the reader to delve deeper into the subject matter independently; the inclusion of experts in the field who offer their individual take on the topics; the references to various freely-available corpora that could be excellent resources for students, researchers and professional translators alike; and the fact that the reader is provided with a unique insight into the most relevant and contemporary issues concerning the legal and ethical considerations in MT design and use in a very accessible manner.

Towards Responsible Machine Translation is a valuable contribution to the ongoing efforts to raise awareness for a more ethical, sustainable and greener approach to MT. The volume covers all the important topics including research on MT and its development, the challenges in training translators and post-editors, and the perspective of the end-users of technology. I thoroughly recommend this book and would say it is a must-read for translator educators, translation students, translation professionals and MT developers. Furthermore, it is hoped that the volume ideally reaches a wider audience of MT users outside of Translation Studies.

References

- **Bowker, Lynne and Jairo Buitrago Ciro** (2019). *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Bingley: Emerald Publishing Limited.

- **Kenny, Dorothy, Joss Moorkens and Félix do Carmo** (2020). "Fair MT: Towards ethical, sustainable Machine Translation." *Translation Spaces* 9(1), 1–11. DOI: 10.1075/ts.00018.int.

- **Krüger, Ralph** (2016a). "Situated LSP translation from a cognitive translational perspective." *Lebende Sprachen* 61(2), 297–332. DOI: 10.1515/les-2016-0014.

- **Krüger, Ralph** (2016b). "Contextualising computer-assisted translation tools and modelling their usability." *trans-kom* 9(1), 114-148. http://www.trans-kom.eu/bd09nr01/trans-kom_09_01_08_Krueger_CAT.20160705.pdf

- **Moorkens, Joss and Dave Lewis** (2019). "Copyright and the re-use of translation as data." Minako O'Hagan (ed.) (2019). *The Routledge Handbook of Translation and Technology.* London/New York: Routledge, 469-481.

Marian Flanagan
University of Copenhagen
E-mail: marianflanagan@hum.ku.dk