Krüger, R. & Hackenbuchner, J. (2022). Outline of a didactic framework for combined data literacy and machine translation literacy teaching. *Current Trends in Translation Teaching and Learning E.* 375 – 432.
https://doi.org/10.51287/cttl202211

OUTLINE OF A DIDACTIC FRAMEWORK FOR COMBINED DATA LITERACY AND MACHINE TRANSLATION LITERACY TEACHING

Ralph Krüger, Janiça Hackenbuchner

TH Köln – University of Applied Sciences

Abstract

This paper outlines a didactic framework for combined data literacy and machine translation (MT) literacy teaching for translation and specialised communication students. The framework is being developed in the context of the DataLit^{MT} project, a publicly funded project at the Institute of Translation and Multilingual Communication at TH Köln – University of Applied Sciences, Germany, which develops didactic resources for teaching data literacy in its translation-specific form of MT literacy to students of BA and MA programmes in translation and specialised communication studies. After discussing the high relevance of machine translation literacy and data literacy in professional translation contexts, the paper introduces the DataLit^{MT} project and discusses a framework of professional MT literacy and an MT-oriented data literacy framework, which form the two theoretical pillars of the project. Also, the interface between MT literacy and data literacy will be illustrated by showing how specific (sub)dimensions of data literacy can be mapped to relevant (sub)dimensions of professional MT literacy. Finally, the paper presents some preliminary didactic resources of the DataLit^{MT} project – concerned with social biases in MT, with MT training data preparation and with automatic MT quality evaluation – and discusses how these resources can be used to teach specific (sub)dimensions of data literacy and professional MT literacy to students in the fields of translation/specialised communication studies.

Keywords: data literacy, professional machine translation literacy, neural machine translation, translation didactics, DataLit^{MT}

1. INTRODUCTION

The professional translation industry has been subject to processes of extensive digitalisation and datafication in recent years. The digitalisation of the translation profession is characterised by the continuous evolution of existing translation technologies or the development of new translation technologies geared at supporting or optimising all phases of the professional translation process (cf.

Sandrini, 2017). The datafication of translation, on the other hand, is characterised by the accumulation and process integration of large-scale digital translation resources such as translation memories or multilingual text corpora (ibid.) as well as by processes of gathering metadata (e.g., data on translation speed and quality or on post-editing productivity) and using such metadata for algorithmic decision making (such as translation job assignments) (cf. Garcia, 2017). A prime example of the interplay of digitalisation and datafication in a translation technology context is the data-driven paradigm of neural machine translation (NMT), where an MT system based on current innovations in artificial intelligence (AI) and natural language processing (NLP) is trained with large-scale and high-quality translation corpora and learns to translate previously unseen texts based on this training data. Due to its relatively high output quality, NMT is increasingly being employed in professional translation production networks.[1]

[1] According to the European Language Industry Survey 2022, standard human translation is still the dominant service offered by language service providers. Post-editing of machine translation (MTPE) ranks second and is considered to exhibit the highest growth potential (cf. ELIS Research, 2022, p. 14). Also, among independent language professionals, more than 40% started or increased offering MTPE services in 2021 (ibid., p. 16).

The high relevance and multifaceted impact of NMT in the professional translation process requires translators to develop an adequate degree of *machine translation literacy* (Bowker & Ciro, 2019).[2] MT literacy, which will be discussed in more detail in section 3 below, can be understood as a *digital literacy*, which O'Brien & Ehrensberger-Dow (2020, p. 147) define as "a set of skills and competencies needed to find, interpret, evaluate and handle digital information". In addition to this increasing relevance of MT literacy on the part of professional translators, another digital literacy, i.e., *data literacy*, is gaining importance – both in professional contexts and in wider society (see section 4). In this regard, Ridsdale et al. (2015, p. 8) observe that "[t]the society of the 21st century is arguably a data rich one" and that "[a]ny country that does not have a technology and data savvy citizenry will be left behind both socially and economically" (ibid.).

[2] According to Long & Magerko (2020, p. 2), the term *literacy* traditionally described our human ability to express ourselves and to communicate by means of written language. However, recently the term has increasingly been used to define "skill sets in a variety of disciplines that have the same potential to enable expression, communication, and access to knowledge" (ibid.).

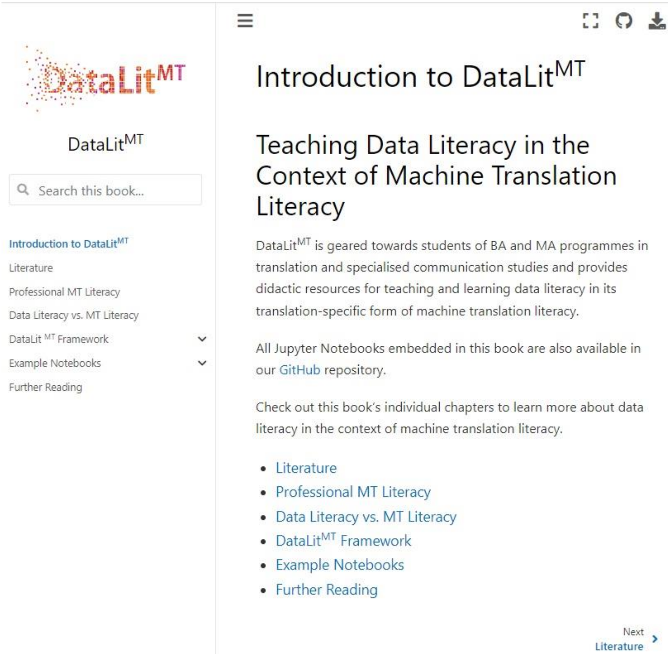
There is an immediate link between data literacy and machine translation literacy, which, in the neural paradigm, is concerned with MT systems the performance of which depends on their training with high-quality translation data (see above). In other words, machine translation literacy in the age of NMT includes various components of data literacy, as will be discussed further in section 4. Translation didactics, which has generally recognised the important role of NMT in the translation profession and the corresponding need to educate students accordingly (see, for example, Massey & Ehrensberger-Dow, 2017; EMT, 2017; Mellinger, 2017), also has to reflect and incorporate this data dimension of modern MT systems and translator training institutions have to find ways to incorporate suitable didactic resources into their curricula.

2. THE DATALIT^{MT} PROJECT

Against this backdrop, we present an outline of a didactic framework for combined data literacy and MT literacy teaching, which is currently being developed as part of the DataLit^{MT} project. The DataLit^{MT} project (DataLit^{MT}, 2022), which runs from June 2021 to February 2023, is publicly funded by the Stifterverband and the Ministry of Culture and Science of North Rhine-Westphalia and is based at the Institute of Translation and Multilingual Communication at TH Köln, Germany. The project forms part of the overall *Data Literacy Initiative* of TH Köln (TH Köln, n. d.) and aims to develop didactic resources for teaching relevant components of data literacy in their translation-specific form of machine translation literacy to students of translation and specialised communication programmes. Initially, the project only intended to develop such didactic resources for students at BA level since, according to Ridsdale et al. (2015, p. 2), "[t]he best place to begin [a data literacy] initiative is the undergraduate curriculum in post-secondary institutions, due in part to their overarching goal of producing globally competitive, critically thinking, well-equipped graduates". After laying the theoretical groundwork for the project, it was decided to extend its didactic scope to cover both BA and MA levels. To this end, learning elements at Basic and Advanced levels will be created which can be used in BA or MA programmes as required. The final project outcomes will be made available as open educational resources under a Creative Commons license in early 2023 and will be provided online in a Jupyter Book format (Jupyter Book Community, 2022), as shown in figure 1 below:

Krüger, R. & Hackenbuchner, J. (2022). Outline of a didactic framework for combined data literacy and machine translation literacy teaching. *Current Trends in Translation Teaching and Learning E.* 375 – 432. **https://doi.org/10.51287/cttl202211**

Figure 1: Jupyter Book interface of the DataLit^{MT} project (draft version)



The DataLit^{MT} project is based on two theoretical pillars – consisting of a Professional Machine Translation Literacy Framework (section 3) and an MT-specific data literacy framework (section 4)[3] –,

[3] Krüger (forthcoming) is specifically concerned with laying the theoretical groundwork of the DataLit^{MT} project and therefore contains a more exhaustive discussion of the Professional Machine Translation Literacy

as well as on a didactic pillar in the form of an MT *cum* data literacy-specific competence matrix, which is currently being developed. Sections 3 and 4 will discuss the two theoretical pillars of the DataLit^{MT} project and section 5 will present some preliminary didactic resources concerned with social biases in MT, MT training data preparation, and automatic MT quality evaluation.

3. THE PROFESSIONAL MACHINE TRANSLATION LITERACY FRAMEWORK

The term *machine translation literacy* was coined by Bowker & Ciro (2019). A concise definition is provided by O'Brien & Ehrensberger-Dow (2020, p. 146), according to whom MT literacy involves "knowing how MT works, how it can be useful in a particular context, and what the implications are of using MT for specific communicative needs". According to Bowker (2021, pp. 26–28), components of MT literacy could include a general understanding of data-driven approaches to MT, an awareness of the need for transparency when working with MT, an awareness of MT-related risks, and strategies for interacting with MT, such

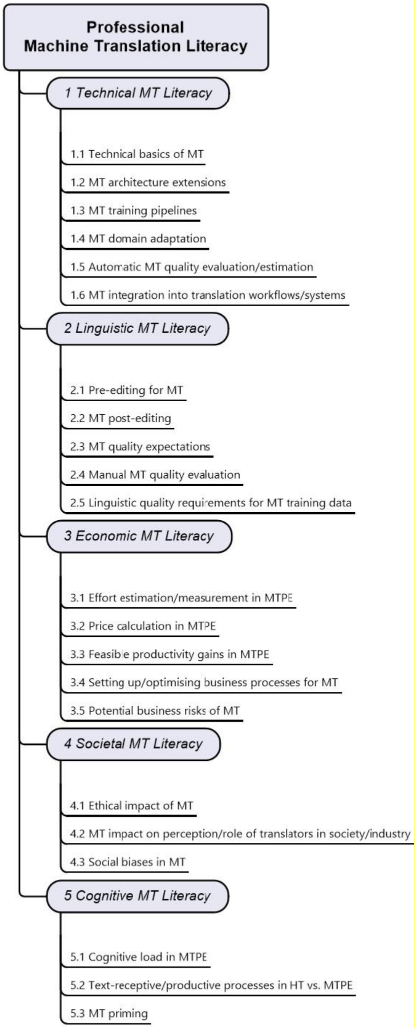
Framework and the MT-specific data literacy framework surveyed in sections 3 and 4 of the present paper.

https://doi.org/10.51287/cttl202211

as pre- or post-editing. Recent work on MT literacy has been concerned with varied audiences having different levels of affinity towards (professional) translation, for example, the scholarly community (Bowker & Ciro, 2019), students (and teachers) of non-translation or non-language specific under-graduate university programmes (Bowker, 2021), students of humanities programmes (Dorst et al., 2022), or undergraduate students in applied languages (Loock & Léchauguette, 2022).

Since the DataLit^{MT} project is geared towards students of translation and specialised communication programmes at BA and MA levels, it is concerned with the expert dimension of MT literacy. This expert dimension could be called *professional* MT literacy, i. e., MT literacy as it is required in professional translation settings. According to Krüger (forthcoming), professional machine translation literacy can be understood as “the full range of MT-related competences professional translators (and other language professionals) may require in order to participate successfully in the various phases of the MT-assisted professional translation process”. This expert dimension of MT literacy is captured by the Professional Machine Translation Literacy Framework depicted in figure 2 below, which forms the first theoretical pillar of the DataLit^{MT} project.

Figure 2: Professional Machine Translation Literacy Framework


Professional Machine Translation Literacy
1 Technical MT Literacy
1.1 Technical basics of MT
1.2 MT architecture extensions
1.3 MT training pipelines
1.4 MT domain adaptation
1.5 Automatic MT quality evaluation/estimation
1.6 MT integration into translation workflows/systems
2 Linguistic MT Literacy
2.1 Pre-editing for MT
2.2 MT post-editing
2.3 MT quality expectations
2.4 Manual MT quality evaluation
2.5 Linguistic quality requirements for MT training data
3 Economic MT Literacy
3.1 Effort estimation/measurement in MTPE
3.2 Price calculation in MTPE
3.3 Feasible productivity gains in MTPE
3.4 Setting up/optimising business processes for MT
3.5 Potential business risks of MT
4 Societal MT Literacy
4.1 Ethical impact of MT
4.2 MT impact on perception/role of translators in society/industry
4.3 Social biases in MT
5 Cognitive MT Literacy
5.1 Cognitive load in MTPE
5.2 Text-receptive/productive processes in HT vs. MTPE
5.3 MT priming

This framework was developed taking into account relevant MT literature from translation studies and translation-oriented NLP and it also reflects the lead author's didactic experience in teaching machine translation to translation students at BA and MA levels at the Institute of Translation and Multilingual Communication at TH Köln, Germany. The framework as presented here is to be understood not as a static structure but as a dynamic construct that can be extended or reduced depending on specific contexts of use or in response to new developments in the field of MT.

The Professional MT Literacy Framework consists of five main MT literacy dimensions, which are differentiated further into specific subdimensions. These are discussed briefly below.

Technical MT literacy is concerned with the technical basics of neural MT (word embeddings, Transformer architecture, etc.), with architectural extensions (such as adaptive, interactive, document-level and multimodal MT), with MT training pipelines (selecting, compiling, organising and cleaning training data, etc.), with methods for MT domain adaptation (including terminology integration), with methods for automatic MT quality evaluation and estimation (BLEU, TER,

https://doi.org/10.51287/cttl202211

etc.) and with integrating MT into translation workflows and systems (APIs and connectors, usability, data security, etc.) (see, e.g., Buj et al., 2020; Koehn, 2020; Koehn & Wiggins, 2021; Rossi & Carré, 2022). To what extent and at which level of detail professional translators actually require such a technical MT literacy will certainly depend on the specific professional context they are embedded in. However, being familiar with the technical intricacies of NMT may generally contribute to translators' intellectual empowerment vis-à-vis this technology (cf. Kenny, 2020, p. 500; Kenny, 2022) and also with regard to other AI technologies which are increasingly influencing our lives at both professional and overall societal levels (cf. the concept of *AI literacy* introduced in Long & Magerko, 2020).

Linguistic MT literacy is concerned with the dimension of overall MT literacy that has traditionally been the most important one for professional translators working in MT-assisted translation scenarios. It covers pre-editing texts for MT (general pre-editing principles, negative translatability indicators, etc.), post-editing (PE) of MT output (PE levels, PE standards and guidelines, types of PE effort, etc.), realistic quality expectations for MT (generic vs. domain-adapted

https://doi.org/10.51287/cttl202211

systems, performance for particular genres and relative to professional human translation, etc.), methods for human MT quality evaluation (adequacy/fluency evaluations, error analyses based on the DQF-MQM[4] framework, etc.) and linguistic quality requirements for MT training data (depending on target audiences, register and genre requirements, etc.) (see, e.g., Nitzke & Hansen-Schirra, 2021; Rossi & Carré, 2022).

Economic MT literacy is particularly relevant in translation project management and is concerned with methods for effort estimation and measurement in MTPE (indicators such as PE time or technical PE effort, etc.), with methods for MTPE price calculation (relevant effort indicators, guidelines for MTPE pricing, etc.), with feasible productivity gains in MTPE (for different language pairs, domains/genres or quality levels), with setting up/optimising business processes for MT integration (translation process standards, MT maturity levels, etc.) and with potential business risks of MT (catastrophic MT errors, liability risks, cyber risks, etc.) (see, e.g., Scansani et al., 2020; Koehn & Wiggins, 2021; ELIS Research, 2022).

[4] DQF = *Dynamic Quality Framework*; MQM = *Multidimensional Quality Metrics* (cf. Lommel, 2018)

Societal MT literacy is concerned with the ethical impact of MT (potential translator disempowerment, trust in MT-assisted translation production networks, etc.), with MT-induced changes in the public/industry-internal perception and role of translators (potential delegitimisation of translators' expert status, etc.), and with potential social biases in MT systems (racial, political, or gender bias, etc.) (see, e.g., Moorkens et al., 2016; do Carmo, 2020; Savoldi et al., 2021).

Cognitive MT literacy, finally, is concerned with factors potentially increasing cognitive load in MTPE (negative translatability indicators, etc.) and strategies for decreasing this cognitive load (offloading cognitive effort for routine translation tasks to MT, etc.), with changes in text-receptive and text-productive processes in human translation vs. MTPE (decreasing relevance of text-productive competences and increasing relevance of text-receptive competences, etc.) and with potential semantic, syntactic and/or stylistic priming effects of MT (machine-translationese, post-editese, etc.) (see, e.g., Daems et al., 2017; Toral, 2019).

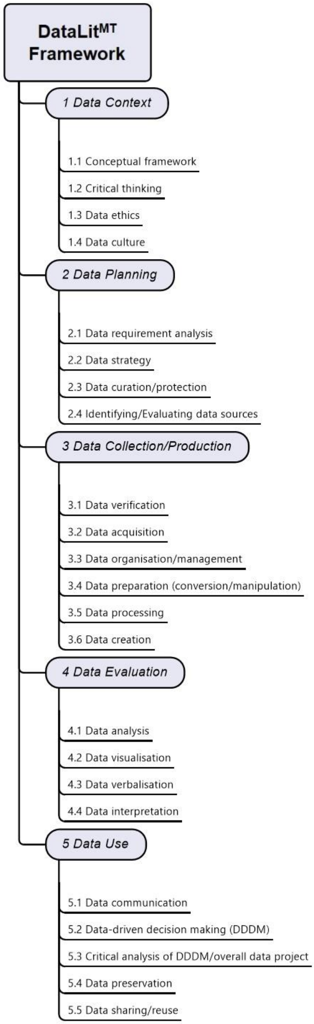
4. THE DATALIT[MT] FRAMEWORK AND POINTS OF CONTACT BETWEEN MT LITERACY AND DATA LITERACY

'Information and data literacy' is one of the key components of the European Union's *Digital Competence Framework 2.0* (1. information and data literacy; 2. communication and collaboration; 3. digital content creation; 4. safety; 5. problem-solving, European Commission, n. d. a), which intends "to provide a common understanding of what digital competence is" and which, at the same time, "provides a basis for framing digital skills policy" (European Commission, n. d. b). Ridsdale et al. (2015, p. 11) define data literacy as "the ability to collect, manage, evaluate, and apply data, in a critical manner". In addition to this rather context-free definition, there are more specific definitions of data literacy, which situate it within specific contexts of use. For example, Schüller (2020, p. 7) situates data literacy within the process of knowledge creation (consisting of the six steps *A) Establish a data culture*, *B) Provide data*, *C) Evaluate data*, *D) Interpret results*, *E) Interpret data*, and *F) Derive actions*) and understands the term as covering both "the creation of data products by the methodologically experienced specialist as well as the competent handling of data by the end user, i.e. the critical and adequate interpretation and application of the data" (ibid.). Misra (2021, p. 8), on the other hand, situates data literacy within the data lifecycle (consisting of *Data Context*, *Data Planning*, *Data Production*, *Data Evaluation* and *Data Use*) and stipulates that "[d]ata literacy should be about the study of data and how it is collected, used and shared by different organisations" (ibid., p. 6). What these different approaches to data literacy have in common is the basic assumption that data literacy is not merely a set of technical skills related to data science, analytics or statistics but that it also involves critically thinking about and handling data in different contexts (e.g., Ridsdale et al., 2015, p. 2; Schüller, 2020, p. 7; Misra, 2021, p. 7).

The DataLit^{MT} Framework (see figure 3 below), which serves as the second theoretical pillar of the eponymous project, combines Ridsdale et al.'s context-free approach to data literacy, Schüller's conceptualisation of data literacy in the context of knowledge creation, and Misra's view of data literacy in the data lifecycle. The five main dimensions of the framework are adopted from Misra's (2021) data lifecycle (with a slight adjustment of dimension 3 concerned with data collection/production), and the individual subdimensions are derived primarily from Ridsdale et al. (2015) and Schüller (2020). These (sub)dimensions – and their interface with the individual (sub)dimensions of the Professional MT Literacy Framework discussed in section 3 – are discussed briefly below.

Krüger, R. & Hackenbuchner, J. (2022). Outline of a didactic framework for combined data literacy and machine translation literacy teaching. *Current Trends in Translation Teaching and Learning E.* 375 – 432. **https://doi.org/10.51287/cttl202211**

Figure 3: The DataLit^{MT} Framework


DataLit^MT Framework
1 Data Context
1.1 Conceptual framework
1.2 Critical thinking
1.3 Data ethics
1.4 Data culture
2 Data Planning
2.1 Data requirement analysis
2.2 Data strategy
2.3 Data curation/protection
2.4 Identifying/Evaluating data sources
3 Data Collection/Production
3.1 Data verification
3.2 Data acquisition
3.3 Data organisation/management
3.4 Data preparation (conversion/manipulation)
3.5 Data processing
3.6 Data creation
4 Data Evaluation
4.1 Data analysis
4.2 Data visualisation
4.3 Data verbalisation
4.4 Data interpretation
5 Data Use
5.1 Data communication
5.2 Data-driven decision making (DDDM)
5.3 Critical analysis of DDDM/overall data project
5.4 Data preservation
5.5 Data sharing/reuse

The *data context* covers "a general awareness of the data and its environment" (Misra, 2021, p. 8) and includes a conceptual framework (general knowledge of data, data use and data applications), the notion of critical thinking (critical awareness of working with data), data ethics (awareness of the ethical dimension of data and the ability to work with data in an ethical manner) and the notion of a data culture (identifying and specifying areas of application where specific tasks could be solved using data, adopting a critical mind-set and showing ethical awareness). Concerning the interface between data literacy and professional MT literacy, the subdimensions of critical thinking and data ethics can readily be linked to the three subdimensions of societal MT literacy listed in the Professional MT Literacy Framework. A critical awareness of the ethical impact of (translation) data may help identifying MT-induced effects of translator disempowerment (e.g., Moorkens et al., 2016) or delegitimisation of professional translators' expert status (e.g., do Carmo, 2020) and it may also be helpful in identifying and remedying social bias patterns such as gender bias (e.g., Savoldi et al., 2021) in MTPE (see section 5.1 below).

Data planning is concerned with "identifying data needs based on real-world scenario[s] and critical inquiry, developing a data strategy or plans [and] considering aspects of data governance" (Misra, 2021, p. 8). The subdimensions of data planning are a data requirement analysis (identifying specific data needs), a data strategy (determines how the data needs identified in the prior analysis can be met), data curation/protection (data retention, storage, accessibility and sharing as well as assessing data security requirements), and identifying/evaluating data sources (assessing the accessibility, relevance, usability and trustworthy-ness of these sources). For example, the identify-ing/evaluating data sources subdimension could be linked to subdimension 1.3 of technical MT literacy concerned with MT training pipelines and to subdimension 2.5 of linguistic MT literacy concerned with linguistic quality requirements for MT training data. The initial step of an MT training pipeline would be the selection of suitable training data, with the suitability of this data being guided by concerns regarding quality requirements for particular language pairs, domains, genres, etc. In order to actually select such data, knowledge of suitable data sources is required, e.g. MT training data repositories such as the commercial *TAUS*

Data Marketplace (TAUS, 2021) or the open-source *OPUS* corpus collection (OPUS, n. d.).

Data collection/production includes "aspects related to data acquisition, design, sourcing, collection, processing, management [and] dissemination" (Misra, 2021, p. 8) as well as the production of new data. The subdimensions of data collection/production are data verification (checking data for correctness, relevance, representativeness and completeness), data acquisition (e.g., downloading data from data sources identified/evaluated previously), data organisation/ management (basic methods and tools for data organisation, metadata creation and use, etc.), data preparation (knowledge of relevant data types and conversion methods, identifying outliers, cleaning data, etc.), data processing (putting previously prepared data to their actual use in the data project) and data creation (creating new data based on previously processed data). Concerning the data literacy and MT literacy interface, the data preparation subdimension could again be linked to subdimension 1.3 of technical MT literacy (MT training pipelines), since central to such a training pipeline are training data preparation steps such as filtering out sentence pairs with identical source and target sides, removing empty or oversized sentences, duplicates, HTML or other markup, performing lowercasing and (subword) tokenisation, etc. (cf. Buj et al., 2020, p. 331 and see section 5.2 below).

Data evaluation includes "the 'working with data' aspects such as interpretation, application, analysis and visualisation of the data" (Misra, 2021, p. 8). The subdimensions of data evaluation are data analysis (applying methods and tools for data analysis, conducting exploratory data analyses, evaluating the results of data analyses, etc.), data visualisation (creating tables or graphical representations of datasets and evaluating their efficiency and accuracy), data verbalisation (verbalising the results of data analysis) and data interpretation (reading and understanding previously visualised/verbalised data, identifying key insights, etc.). Looking at the interface with professional MT literacy, the data analysis subdimension could be linked to subdimension 1.5 of technical MT literacy concerned with automatic MT quality evaluation and estimation and to subdimension 2.4 of linguistic MT literacy concerned with manual MT quality evaluation. In an MT-assisted translation production network, the output of MT engines trained with data collected in the corresponding downstream steps could be analysed automatically and/or manually with regard to its quality (see section 5.3 below), and the results of these data analyses could then be fed into the downstream step of data use.

This *data use* dimension, finally, is concerned with "data communication, critique, engagement, argument and advocacy related to data-driven decision making" (Misra, 2021, p. 8), potentially also including "ethical data sharing and reuse practices" (ibid.). The subdimensions of data use are data communication (communicating previously interpreted data visualisations or verbalisations to stakeholders), data-driven decision-making (DDDM, converting data into actionable information, weighing the impact of data-driven decisions, implementing such decisions, etc.), critical analysis of DDDM/overall data project (evaluating the effectiveness/impact of data-driven decisions and reflecting critically on the effectiveness/impact of the overall data project), data preservation (assessing preservation requirements, methods and tools, etc.) and data sharing/reuse (assessing suitable data sharing methods and platforms and sharing data in a legally correct and ethically adequate manner). In a data use scenario, the subdimension of data-driven decision making could be linked, for example, to

https://doi.org/10.51287/cttl202211

subdimension 3.2 of economic MT literacy concerned with price calculation in MTPE workflows. In this case, the results of the automatic and/or manual quality evaluation of the MT output could guide decisions on how to calculate potential discounts for MTPE services (cf. ELIS Research, 2022, p. 26).

5. PRELIMINARY DIDACTIC RESOURCES OF THE DATALIT^{MT} PROJECT

In the following sections, some preliminary didactic resources/learning elements for teaching specific components of data literacy *cum* MT literacy will be presented.[5] In these learning elements, the individual (sub)dimensions of the Professional MT Literacy Framework discussed in section 3 provide the actual contexts of use in which the various (sub)dimensions of the DataLi^{tMT} Framework discussed in section 4 can be taught. For each learning element presented below, we provide preliminary competence descriptors

[5] The draft versions of all DataLit^{MT} learning elements created so far are available in the following GitHub repository: https://github.com/ITMK/DataLitMT

from the DataLit^MT Competence Matrix currently being developed.[6]

5.1. Teaching Critical Thinking and Data Ethics in the Context of Societal MT Literacy

Competence descriptor for *Critical Thinking*:
Can understand, describe and analyse potential problems, risks and implications of translation data collection/production, evaluation and use practices, and can reflect on the implications of these practices.

Competence descriptor for *Data Ethics*:
Can understand and describe legal and ethical issues associated with collecting/producing, evaluating and using open-source or commercially available translation data.

The critical thinking and data ethics subdimensions of the data context were mapped to subdimension

---

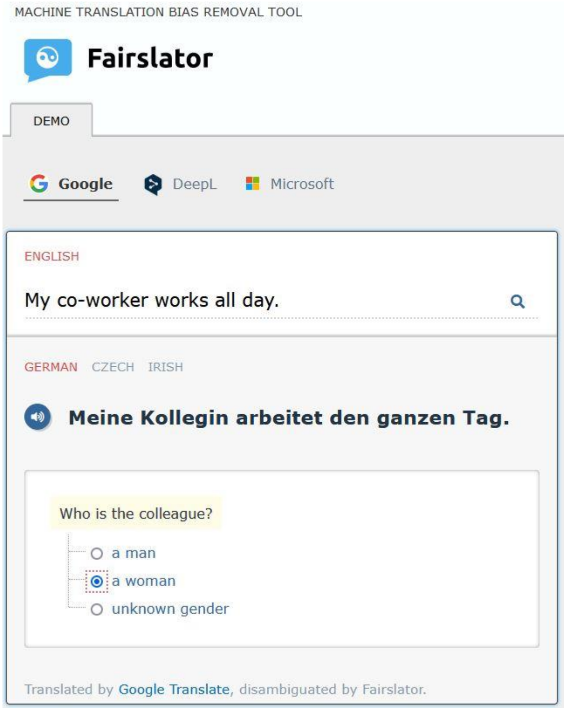
[6] Both the competence descriptors and the didactic resources discussed in the following sections are generic in that they are not specific to the Basic or the Advanced level of the DataLit^MT project yet. It was decided to first develop the competence descriptors and the learning elements at a proficiency-agnostic level in order to test their overall didactic suitability and to then split them into Basic and Advanced level descriptors/resources (which is currently being done).

4.3 of societal MT literacy concerned with social biases in MT in order to create a corresponding learning element. In this learning element, students are presented with prepared source texts (ST) written in a genderless or notional gender language (e.g. English) which include specific language patterns triggering generic MT systems such as DeepL or Google Translate to produce gender-biased target texts (TT) for a grammatical gender language (e.g. German). Examples of such ST language patterns and corresponding gender-biased TTs are the plural generic masculine in German (e.g., *ministers* translated as *Minister* instead of *Ministerinnen und Minister*, *Minister:innen*, etc.), professions prone to gender bias (e.g., *pilots* being translated with the male form *Piloten* and *flight attendants* with the female form *Flugbegleiter-innen*), gender bias-prone adjective-substantive combinations (e.g., *the famous lawyer* being translated with the male form *der berühmte Anwalt* and *the pretty lawyer* with the female form *die hübsche Anwältin*) and gender bias-prone sentence contexts (e.g., *My co-worker works all day* being translated as *Mein Kollege arbeitet den ganzen Tag* and *My co-worker gossips all day* as *Meine Kollegin quatscht den ganzen Tag*) (for all examples, see Hackenbuchner, 2022, pp. 36–41). Students are then asked to machine-translate these source texts using a publicly available MT system such as DeepL or Google Translate and to analyse the output manually for instances of gender bias as discussed above. In order to guide students through this analysis, they are referred to the website of the DeBiasByUs (2021) project. DeBiasByUs (origin-nally named *BiasByUs*) was developed by the co-author of this paper together with other participants of the Artificially Correct Hackathon organised by the Goethe-Institut (Goethe-Institut, 2021; see also Daems & Hackenbuchner, 2022). The DeBiasByUs platform has two goals: 1) to inform its audience about gender bias and its presence in MT, about categories of gender bias and about relevant studies and their results both for general audiences as well as for researchers, and 2) to serve as a platform for gathering data to create a research database of parallel data for multiple languages – specifically grammatical gender languages – containing instances of gender bias. On the DeBiasByUs website, students are asked to paste a source sentence triggering the identified gender bias and the biased target sentence that they encountered in their chosen MT system. Then, they are asked to propose an unbiased alternative according to their preferences (e.g., a binary or non-binary gender translation solution). Also, they are asked to assign an error category and a bias category to the identified instances of gender bias. The DeBiasByUs website is depicted in figure 4 below:

Figure 4: Website of the *DeBiasByUs* project on gender bias in machine translation (proof-of-concept version)

As a further resource for their gender bias analysis of the MT output, students are referred to the website of *Fairslator* (Fairslator, 2022; see also Měchura, 2022), an MT system specifically designed to handle ambiguous MT inputs by asking for user input on ambiguous input strings, thus providing the opportunity to remove MT bias. The Fairslator website is depicted in figure 5 below:

Figure 5: Fairslator website



Students can paste the gender bias-prone ST into the Fairslator interface and choose a preferred unbiased translation solution, which they can then add to the DeBiasByUs database. Fairslator can thus complement students' gender bias analysis guided by the DeBiasByUs website by pointing them directly to bias-prone ST elements and providing them a choice of unbiased TT solutions. In addition to the overall didactic goal of developing students' critical thinking about and awareness of ethical aspects of translation data in the context of societal MT literacy, this may also strengthen students' text-production competence with regard to gender-fair language. Further "material scaffolds" (Sannholm, 2021) provided for this learning element are a detailed documentation on gender bias in MT as well as videos guiding students through the process of gender bias analysis.

5.2. Teaching Data Preparation in the Context of Technical MT Literacy

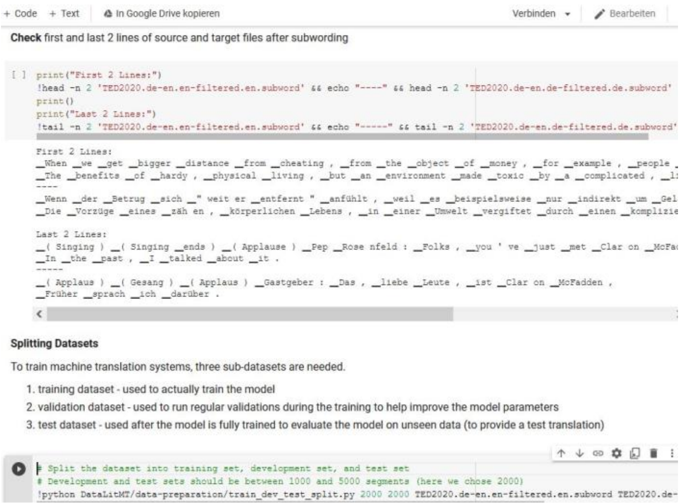
Competence descriptor for *Data Preparation*:
Can understand MT-specific data types and methods for converting and cleaning MT training data and apply such methods for training data preparation in an MT-assisted translation scenario.
The data preparation subdimension of data collection/production was mapped to subdimension 1.3 of technical MT literacy concerned with MT training pipelines in order to create a learning element for MT training data preparation. For this learning element, a Jupyter notebook (Jupyter

https://doi.org/10.51287/cttl202211

Team, 2015) was set up[7], which implements the Python scripts created by Moslem (n. d.) for MT training data preparation. With this notebook and the implemented Python scripts, students will be guided through various data filtering steps such as deleting empty rows, duplicates, source-copied rows and HTML tags, performing standard tokenisation and lowercasing (if required), followed by subword tokenisation and splitting the data into training and development datasets (ibid.). After each step of the data preparation process, students are able to print a section of the training data in order to see how the data was changed through this particular step, as illustrated in figure 6 below:

[7] This Jupyter notebook as well as other notebooks to be developed by the DataLitMT project will be integrated into the Jupyter Book depicted in figure 1.

Figure 6: Jupyter notebook-based Python implementation of MT training data preparation based on Moslem (n. d.) (draft version)



Jupyter notebooks were shown to be a suitable didactic resource for teaching the more technical aspects of machine translation literacy (and data literacy) to students of translation and specialised communication programmes (cf. Krüger, 2021a). These notebooks provide an interactive computing environment, which can be used as a didactic scaffold to guide students through computational processes via detailed documentation and pre-defined code. Students can then run this code and follow the computational processes performed by it without having any prior programming experience. Currently, a video documentation is being developed which guides students through tailoring the data preparation notebook to their individual needs (processing individual texts, performing only selected processing operations, etc.). The didactic benefits of Jupyter notebooks for teaching technical MT literacy are illustrated further in the following section.

5.3. Teaching Data Analysis and Data Visualisation in the Context of Technical MT Literacy

Competence descriptor for *Data Analysis*:
Can understand and apply methods and tools for manually or automatically analysing machine-translated target data produced in an MT-assisted translation scenario.

Competence descriptor for *Data Visualisation*:
Can create tables or graphical representations of MT data analysis results, and can evaluate the effectiveness and accuracy of such tables/representations.

The data analysis and visualisation subdimensions of data evaluation were mapped to subdimension

1.5 of technical MT literacy concerned with automatic MT quality evaluation/estimation in order to create a corresponding learning element. Here, students are provided with the MT output of different MT systems (DeepL, Google Translate or self-trained systems) and with a corresponding human reference translation and are asked to perform a comparative automatic quality evaluation of this output using various automatic MT quality metrics such as the string matching-based metric *chrF* (character n-gram F-score, Popović, 2015) or the word embedding-based metric *BERTScore* (Zhang et al., 2020). In order to compute these scores, students are provided with Jupyter notebooks which include a Python implementation of these metrics. Figure 7 below depicts a notebook section on calculating chrF:

Figure 7: Jupyter notebook section on chrF calculation

2.1.4 chrF

chrF is an acronym for **char**acter n-gram **F**-Score. It was originally proposed in Popovic (2015): chrF: Character n-gram F-Score for Automatic MT Evaluation. The source code of the chrF score function used in this notebook can be found here. chrF is one of the most current MT quality scores based on string-matching. Unlike the scores covered so far, chrF does not operate at the *word* but rather at the *character* level (hence the *chr*). As such, it employs the notion of *character n-grams*, where an 1-gram represents a single character, a 2-gram represents a sequence of two characters, etc. (basically the same idea as with regular n-grams, but here, the atomic units are characters instead of words).

chrF is calculated using the following formula:

$$chrF\beta = (1+\beta^2)\frac{chrP \times chrR}{\beta^2 \times chrP + chrR}$$

Looking at the formula, you may already have guessed that the *F* in *chrF* stands for F-Measure again (which we know already as an autonomous MT quality score and as part of the METEOR score). So, chrF is "the F-score based on character *n*-grams" (Popovic 2015:392). As the previous metrics covered in this notebook (with the exception of METEOR), chrF is completely language-agnostic. In the formula, $chrP$ is the percentage of n-grams in the hypothesis which are also present in the reference (see our general definition of *precision* above) and $chrR$ is the percentage of character n-grams in the reference which are also present in the hypothesis (again, see our general definition of *recall* above). β, finally, is a parameter which assigns β times more importance to recall than to precision (the idea that recall should be assigned more importance than precision should sound familiar from the discussion of METEOR above). When $\beta = 1$, precision and recall are given equal importance.

The best correlations with human quality judgements (remember, this correlation is the currency in which the quality of an MT quality score is measured) were achieved with the *6-gram chrF3 score* (see Popovic 2015:393). This means that the n-grams used in calculating chrF are character 6-grams, in other words, sequences of six characters. 3 is the value for β, which means that recall is given three times more importance than precision when calculating chrF. The *chrf_score()* function implemented in NLTK retains these values in its standard configuration. Run the code below to calculate 6-gram chrF3 for our reference-hypothesis pair.

```
[ ] # Import chrf_score() and word_tokenize() functions
    from nltk.translate import chrf_score
    from nltk import word_tokenize

    # Define reference and hypothesis
    reference_chrF = word_tokenize('This is a simple test sentence')
    hypothesis_chrF = word_tokenize('This is an example sentence')

    #Calculate and print chrF score
    chrF = chrf_score.sentence_chrf(reference_chrF, hypothesis_chrF)
    print(f"chrF: {chrF}")
```

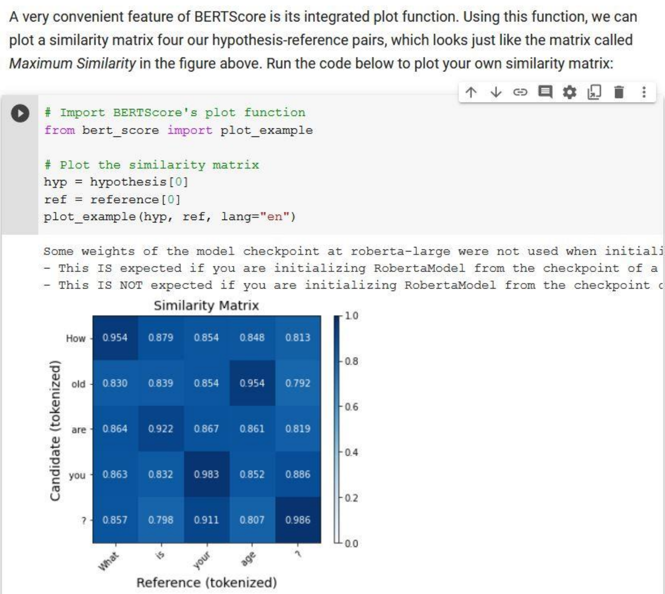
As you can see, for the same hypothesis-reference pair as in our F-Measure calculation in section 2.1.1 above, chrF returns a score which is a little bit lower than the F-Measure score computed above (which was 0.6). If you want to calculate chrF for other character n-gram sequences and for other β values, you have to pass additional arguments to the function. For

The first documentation section of the notebook explains to students how the chrF score works and how it is calculated. The implemented code can then be run by students in order to compute actual chrF scores for specific MT outputs. The docu-

https://doi.org/10.51287/cttl202211

mentation section following this code then gives further information on how to interpret these scores. While the code sections in these notebooks thus help students *analyse* their MT data, the documentation sections may also support them in *verbalising* and *interpreting* this data. Specific MT quality scores, such as BERTScore, also include *visualisation* capabilities, as depicted in figure 8 below:

Figure 8: Jupyter notebook section on BERTScore visualisation



As can be seen in figure 8, the code implemented in this Jupyter notebook creates a similarity matrix for a specific MT output and a human reference translation, which visualises similarity scores for all possible pairings of words in the two strings. A further documentation section could then be provided to help students interpret this visualisation. In addition to such score-internal visualisation functions, there are various open-source Python libraries supporting data visualisation (such as Matplotlib or Seaborn), which can be directly imported into Jupyter notebooks to provide students with powerful data visualisation capabilities.

The Jupyter notebooks discussed in this section on data analysis and visualisation form part of a GitHub repository of Python-based Jupyter notebooks for teaching the technical dimension of machine translation to students of translation studies programmes.[8] Suitable notebooks of the repository, which is presented and discussed in more detail in Krüger (2021b), are currently being integrated into the DataLit^{MT} project.

[8] Link to the GitHub repository: https://github.com/ITMK/MT_Teaching

6. CONCLUDING REMARKS

This paper outlined a didactic framework for combined data literacy and MT literacy teaching, which is currently being designed as part of the DataLit^{MT} project at TH Köln, Germany. After discussing the theoretical pillars of the combined data literacy/MT literacy framework proposed here, some preliminary didactic resources of the DataLit^{MT} project were presented. In addition to resource development, current project work is focused on establishing a complete mapping between relevant data literacy and MT literacy (sub) dimensions, as discussed in a cursory way in the present paper, and with designing a competence matrix consisting of Basic and Advanced level competence descriptors for the individual (sub)dimensions of the DataLit^{MT} Framework.

With its didactic aims, the DataLit^{MT} project is part of a more general trend in contemporary translation didactics to create teaching and learning resources for developing translation students' and other interested parties' MT literacy and related digital literacies. Examples of other such initiatives are the MultiTraiNMT project (cf. Kenny, 2022), which aims at introducing wider audiences to contemporary machine translation technologies, or the

EasyAI project (Fantinuoli, n. d.), which provides an introduction to current AI technologies for translators, interpreters and linguists. Since powerful AI technologies are permeating both the private and the professional spheres of modern societies and are increasingly influencing citizens' lives (often with an invisible hand), these initiatives offer timely contributions to the emergence of empowered, AI-savvy citizenries and workforces who are able to harness the advantages of these technologies in a critical and ethically responsible manner.

ACKNOWLEDGEMENTS

The authors would like to thank the two anonymous reviewers for their valuable comments and suggestions on the manuscript version of this paper.

FUNDING INFORMATION

The DataLit^{MT} project reported on in this paper is funded by the Stifterverband and the Ministry of Culture and Science of North Rhine-Westphalia, Germany, under the funding line Curriculum 4.0 (https://www.stifterverband.org/curriculum-4-0).

REFERENCES

Bowker, L. (2021). Machine translation use outside the language industries: A comparison of five delivery formats for machine translation literacy instruction. In R. Mitkov, V. Sosoni, J. C. Giguère, E. Murgolo & E. Deysel (Eds.), *TRITON 2021: TRanslation and Interpreting Technology Online – Proceedings of the conference* (pp. 25–36). http://triton-conference.org/proceedings/

Bowker, L., & Ciro, J. B. (2019). *Machine translation and global research. Towards improved machine translation literacy in the scholarly community*. Emerald Publishing. https://doi.org/10.1108/9781787567214

Buj, D. M., Ibáñez García, D., Parcheta, Z., & Casacuberta, F. (2020). NICE: Neural integrated custom engines. In A. Martins, H. Moniz, S. Fumega, B. Martins, F. Batista, L. Coheur, C. Parra, I. Trancoso, M. Turchi, A. Bisazza, J. Moorkens, A. Guerberof, M. Nurminen, L. Marg & M. L. Forcada (Eds.), *Proceedings of the 22nd annual conference of the European Association for Machine Translation* (pp.

329–338). European Association for Machine Translation. https://aclanthology.org/2020.eamt-1.35/

Carmo, F. do (2020). 'Time is money' and the value of translation. *Translation Spaces* 9(1), 35–57. https://doi.org/10.1075/ts.00020.car

Daems, J., Vandepitte, S., Hartsuiker, R. J., & Macken, L. (2017). Identifying the machine translation error types with the greatest impact on post-editing effort. *Frontiers in Psychology* 8, 1–15. https://doi.org/10.3389/fpsyg.2017.01282

Daems, J., & Hackenbuchner, J. (2022). DeBiasByUs: Raising awareness and creating a database of MT bias. In H. Moniz, L. Macken, A. Rufener, L. Barrault, M. R. Costa-jussà, C. Declercq, M. Koponen, E. Kemp, S. Pilos, M. L. Forcada, C. Scarton, J. Van den Bogaert, J. Daems, A. Tezcan, B. Vanroy & M. Fonteyne (Eds.), *Proceedings of the 23rd annual conference of the European Association for Machine Translation* (pp. 287–288). European

Association for Machine Translation. https://aclanthology.org/2022.eamt-1.34/

DataLit^{MT} (2022). DataLit^{MT} – Teaching data literacy in the context of machine translation literacy. Project website. https://www.th-koeln.de/informations-und-kommunikationswissenschaften/datalitmt--teaching-data-literacy-in-the-context-of-machine-translation-literacy_85011.php

DeBiasByUs (2021). DeBiasByUs – On a mission towards gender-fair machine translation. https://artificiallycorrec.wixsite.com/biasbyus

Dorst, A. G., Santos Ângelo Salgado Valdez, S., & Bouman, H. M. C. (2022). Machine translation in the multilingual classroom: How, when and why do humanities students at a Dutch university use machine translation? *Translation and Translanguaging in Multilingual Contexts* 8(1), 49–66. https://doi.org/10.1075/ttmc.00080.dor

ELIS Research (2022). *European language industry survey 2022.* https://elis-survey.org

EMT (2017). European Master's in Translation competence framework 2017. *Website of the DG Translation of the European Commission*. https://ec.europa.eu/info/sites/default/files/emt_competence_fwk_2017_en_web.pdf

European Commission (n. d. a). The digital competence framework 2.0. *EU Science Hub*. https://joint-research centre.ec.europa.eu/digcomp/digital-competence-framework-20_en

European Commission (n. d. b). DigComp. *EU Science Hub*. https://joint-research-centre.ec.europa.eu/digcomp_en

Fairslator (2022). Fairslator. Machine translation bias removal tool. https://www.fairslator.com/

Fantinuoli, C. (n. d.). EasyAI. An introductory class to Artificial Intelligence for translators, interpreters and linguists. https://easyai.uni-mainz.de/

Garcia, I. (2017). Translation in the cloud age: Online marketplaces. *Hermes – Journal of Language and Communication in Business* 56, 59–70. https://doi.org/10.7146/hjlcb.v0i56.97202

Goethe-Institut (2021). Artificially correct hackathon 2021. https://www.goethe.de/prj/one/en/aco/art/22548141.html

Hackenbuchner, J. (2022). *Gender bias in machine translation – A comparison of systems trained with biased and unbiased datasets*. MA thesis, TH Köln.

Jupyter Book Community (2022). Jupyter Book. Build beautiful, publication-quality books and documents from computational content. https://jupyterbook.org/en/stable/intro.html

Jupyter Team (2015). Jupyter notebook user documentation. https://jupyter-notebook.readthedocs.io/en/stable/notebook.html

Kenny, D. (2020). Technology and translator training. In M. O'Hagan (Ed.), *The*

Routledge handbook of translation and technology (pp. 498–515). Routledge.

Kenny, D. (2022). *Machine translation for everyone. Empowering users in the age of artificial intelligence*. Language Science Press. https://doi.org/10.5281/zenodo.6653406

Koehn, P. (2020). *Neural machine translation*. Cambridge University Press.

Koehn, P., & Wiggins, D. (2021). Everything that you need to know about customizing a great machine translation engine. *Omniscien Technologies webinar series*. https://omniscien.com/webinars/everything-that-you-need-to-know-about-customizing-a-great-machine-translation-engine/

Krüger, R. (2021a). Using Jupyter notebooks as didactic instruments in translation technology teaching. *The Interpreter and Translator Trainer* (Ahead of Print), 1–21. https://doi.org/10.1080/1750399X.2021.2004009

Krüger, R. (2021b). An online repository of Python resources for teaching machine translation to translation students. *Current Trends in*

Translation Teaching and Learning E 2021, 4–31. https://doi.org/10.51287/cttle20212

Krüger, R. (forthcoming). Integrating professional machine translation literacy and data literacy. *Lebende Sprachen*.

Lommel, A. (2018). Metrics for translation quality assessment: A case for standardising error typologies. In J. Moorkens, S. Castilho, F. Gaspari & S. Doherty (Eds.), *Translation quality assessment. From principles to practice* (pp. 109–127). Springer. https://doi.org/10.1007/978-3-319-91241-7_6

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In R. Bernhaupt, F. Mueller, D. Verweij & J. Andres (Eds.), *CHI '20: Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–16). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376727

Loock, R., & Léchauguette, S. (2022). Machine translation literacy and undergraduate students in applied languages: Report on an exploratory study. *Revista Tradumàtica* 19, 204–225. https://doi.org/10.5565/rev/tradumatica.281

Massey, G., Ehrensberger-Dow, M. (2017). Machine learning: Implications for translator education. *Lebende Sprachen* 62(2), 300–312. https://doi.org/10.1515/les-2017-0021

Měchura, M. (2022). We need to talk about bias in machine translation. The Fairslator whitepaper. *Fairslator*. https://www.fairslator.com/fairslator-whitepaper.pdf

Mellinger, C. D. (2017). Translators and machine translation: Knowledge and skills gaps in translator pedagogy. *The Interpreter and Translator Trainer* 11(4), 280–293. https://doi.org/10.1080/1750399X.2017.1359760

Misra, A. (2021). Advancing data literacy in the post-pandemic world. A primer to catalyse policy dialogue and action. *PARIS21*. https://paris21.org/sites/default/files/inline-files/DataLiteracy_Primer_0.pdf

Moorkens, J., Lewis, D., Reijers, W., Vanmassenhove, E., & Way, A. (2016). Translation resources and translator disempowerment. In L. Devillers, B. Schuller, E. Mover Provost, P. Robinson, J. Mariani & A. Delaborde (Eds.), *Proceedings of the 10th international conference on language resources and evaluation. ETHics in corpus collection, annotation and application workshop* (pp. 49–53). http://www.lrec-conf.org/proceedings/lrec2016/workshops/LREC2016Workshop-ETHICA2_Proceedings.pdf

Moslem, Y. (n. d.). Machine translation (MT) preparation scripts. GitHub repository. https://github.com/ymoslem/MT-Preparation

Nitzke, J., & Hansen-Schirra, S. (2021). *A short guide to post-editing*. Language Science Press. https://doi.org/10.5281/zenodo.5646896

O'Brien, S., & Ehrensberger-Dow, M. (2020). MT literacy—a cognitive view. *Translation, Cognition & Behaviour* 3(2), 145–164.

https://doi.org/10.1075/tcb.00038.obr

OPUS (n. d.). OPUS … the open parallel corpus. https://opus.nlpl.eu/

Popović, M. (2015). chrF: Character n-gram F-score for automatic MT evaluation. In O. Bojar, R. Chatterjee, C. Federmann, B. Haddow, C. Hokamp, M. Huck, V. Logacheva & P. Pechina (Eds.), *Proceedings of the tenth workshop on statistical machine translation* (pp. 392–395). Association for Computational Linguistics. https://aclanthology.org/W15-3049/

Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., & Wuetherick, B. (2015). *Strategies and best practices for data literacy education*. Knowledge synthesis report. Dalhousie University. http://hdl.handle.net/10222/64578

Rossi, C., & Carré, A. (2022). How to choose a suitable neural machine translation solution: Evaluation of MT quality. In D. Kenny (Ed.), *Machine translation for everyone: Empowering users in the age of artificial* *intelligence* (pp. 51–79). Language Science Press. https://doi.org/10.5281/zenodo.6653406

Sandrini, P. (2017). Translation 4.0 – Eine Perspektivenverschiebung. In L. Zybatow, A. Petrova, A. Stauder & M. Ustaszewski (Eds.), *Übersetzen und Dolmetschen: Berufsbilder, Arbeitsfelder, Ausbildung. Ein- und Ausblicke in ein sich wandelndes Berufsbild der Zukunft* (pp. 139–152). Lang.

Sannholm, R. (2021). *Translation, teamwork, and technology – The use of social and material scaffolds in the translation process*. PhD dissertation, University of Stockholm. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1604340&dswid=7197

Savoldi, B., Gaido, M., Bentivogli, L., Negri, M., & Turchi, M. (2021). Gender bias in machine translation. *Transactions of the Association for Computational Linguistics* 9, 845–874. https://doi.org/10.1162/tacl_a_00401

Scansani, R. (2020). How do LSPs compute MT discounts? Presenting a company's pipeline and its use. In A. Martins, H. Moniz, S. Fumega, B. Martins, F. Batista, L. Coheur, C. Parra, I. Trancoso, M. Turchi, A. Bisazza, J. Moorkens, A. Guerberof, M. Nurminen, L. Marg & M. L. Forcada (Eds.), *Proceedings of the 22nd annual conference of the European Association for Machine Translation* (pp. 393–401). European Association for Machine Translation. https://aclanthology.org/2020.eamt-1.42/

Schüller, K. (2020). *Future skills: A framework for data literacy*. Hochschulforum Digitalisierung. https://hochschulforumdigitalisierung.de/de/future-skills-framework-data-literacy

TAUS (2021). Data marketplace. For language data monetization and acquisition. https://datamarketplace.taus.net/

TH Köln (n. d.) DaLI – Data Literacy Initiative. https://www.th-koeln.de/informations-und-kommunikationswissenschaften/dali---data-literacy-initiative_78648.php

Toral, A. (2019). Post-editese: An exacerbated translationese. In M. Forcada, A. Way, B. Haddow & R. Sennrich (Eds.), *Proceedings of Machine Translation Summit XVII: Research track* (pp. 273–281). European Association for Machine Translation. https://aclanthology.org/W19-6627/

Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2020). BERTScore: Evaluating text generation with BERT. *arXiv*. https://arxiv.org/abs/1904.09675