Mask, Pseudonymise, Translate: Enhancing MT Quality Under Privacy Constraints*

Rinalds Vīksna[1,2,*,†], Mārcis Pinnis[1,2,†]

[1]Tilde, Vienības gatve 75a, Riga, Latvia
[2]University of Latvia, Raiņa bulv. 19, Riga, Latvia

Abstract

Data-protection regulations such as the General Data Protection Regulation in the European Union may prohibit the transfer of personally identifiable information (PII) to external professional translation vendors or cloud-based machine-translation (MT) providers. Conventional anonymisation removes PII entirely, but in doing so also strips many of the linguistic cues—gender, number, case—that morphologically rich languages rely on, degrading MT quality. We propose a targeted pseudonymisation framework that replaces entities with pseudonyms carefully matched for pseudonym category and morphology. The approach preserves the grammatical information essential for high-quality translation from/to morphologically rich languages and ensures the removal of PII. Additionally, we introduce a new method for evaluating translations of anonymised content—masked translation evaluation—which allows for comparing the quality of translations of texts anonymised and pseudonymised by different methods. Our experiments show that anonymisation using a named entity category offers good context translation performance when using MT systems which support untranslatable entities. Pseudonymisation allows for superior quality for context translation if MT systems do not support untranslatable entities or when translating from morphologically rich languages to English.

Keywords

Anonymisation, pseudonymisation, machine translation

1. Introduction

Driven by the legislative changes (for instance, the General Data Protection Regulation (GDPR)[1] in the EU and the Health Insurance Portability and Accountability Act (HIPAA)[2] in the US), there is an increasing need to strip sensitive and personally identifying information (PII) from documents before disseminating them to the public or using them to train machine learning models. Text anonymisation and pseudonymisation, or simply text sanitisation [1], transforms documents through edit operations such as hiding particular text spans or replacing them with different values in order to protect PII and enable sharing documents publicly or outside an organisation's infrastructure.

When an organisation wants to share its data in a different language from the one in which it is written, the data has to be translated into the target language. Before sending data that contains PII to an external translation service provider or machine translation (MT) service, the data should to be anonymised to avoid sharing sensitive data and not to violate the GDPR. The anonymised text presents less information than the original text to a translator or an MT system, impacting the translation quality. In cases when anonymisation is performed by replacing sensitive terms with masks or labels, the MT engine has to support such placeholders and output appropriate placeholders after translation. However, in cases when anonymisation is performed by replacing sensitive terms with suitable pseudonyms, adaptations are not required on the MT side, as the input and output consist of unmasked (raw) text. To avoid loss of coherence, the pseudonyms must match in gender, inflexion, and grammatical number with the replaced term.

*IVUS 2025: Information Society and University Studies 2025, May 15, Kaunas, Lithuania

*Corresponding author.
†These authors contributed equally.
rinalds.viksna@tilde.lv (R. Vīksna); marcis.pinnis@tilde.lv (M. Pinnis)
0000-0003-2705-508X (R. Vīksna); 0000-0001-6832-5600 (M. Pinnis)

[1]https://eur-lex.europa.eu/eli/reg/2016/679/oj
[2]https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

The pseudonym generation task is commonly solved by generating lists of common first and last names and organisations, and selecting a pseudonym from this list, as in [2]. The lists may be extended to include separate lists for male and female names [3]. Approaches using pre-trained language models make use of the masked language modelling capability of such models [3, 4] to produce pseudonyms, or simply replace the words with similar words in the embedding space [5].

This work makes the following contributions:

- A method for pseudonym generation in morphologically rich languages;
- A novel method for evaluating the impact of anonymisation on machine translation.

2. Background and Related Work

The need for effective and reliable text anonymisation methods has led to the development of various tools and frameworks. For instance, the MAPA [3] project developed a tool for multilingual text anonymisation, providing multiple models used for the legal, administrative, and general domains. The tool includes a module for text pseudonymisation using both rule-based and neural language model-based methods.

A tool named Faker [6] can be used to introduce fake sensitive entities into anonymised text. Commonly used entity types are available in Faker by default, and it supports several languages. However, the generated data is provided in base forms. For morphologically rich languages, pseudonyms must be generated in a morpho-syntactic agreement with the rest of the sentence, matching gender, inflexion, and number.

A rule-based pseudonymisation system for Swedish clinical texts was developed by Dalianis (2019) [2]. The pseudonymisation system is capable of replacing the PIIs (first and last names, phone numbers, locations, dates, ages, and healthcare units) found in the electronic patient records with realistic surrogates. Pseudonyms are either generated by selecting one from a previously prepared dictionary or, for numeric values, some transformation is applied.

The CURLICAT project [4] developed text anonymisation systems for several Eastern European languages. The project used a mixture of approaches to pseudonymisation. The Bulgarian system uses a rule-based approach. The Hungarian system uses a morphology generator to produce inflected pseudonyms. The Romanian approach uses randomly generated pseudonyms and tries to preserve the initial surface form, including characteristics such as suffixes or prefixes. The Slovak approach replaces nouns with compatible nouns in the same inflexion. The Slovak, Croatian, and Polish approach uses the masked language modelling capability of the pre-trained BERT model to find possible replacements for each word. The model produces 40 alternatives, and the final one is selected to match the grammatical form of the original.

A workflow that utilises a human in the loop for post-editing anonymised texts, with the aim of reconciling the competing needs of data privacy and data quality, has been introduced by Chatzitheodorou et al. (2023) [7]. This workflow assumes that an MT system is used to translate both original and pseudonymised text. The translator is presented with the translated pseudonymised text version for post-editing. After post-editing, both non-anonymised machine-translated and pseudonymised post-edited versions are merged, producing the final version of the translated text.

Pissarra et al. (2024) [8] offer six metrics for the evaluation of clinical domain text anonymisation, which are divided into two categories. The first category comprises anonymisation sensitivity metrics, which are based on the Levenshtein metric. Given a list of sensitive entities, each entity is matched against anonymised text to find the longest substring of the entity, which can be found in the text. For the longest substring found in the text, a Levenshtein similarity ratio is calculated, and compared to a preset threshold. Entities with a low Levenshtein similarity ratio are considered successfully anonymised, while entities above the threshold are considered not anonymised. The second category of metrics assesses the impact of anonymisation on the preservation of clinical concepts.

Dalianis (2019) [2] performed document-level human evaluation to assess the quality of pseudonymisation. To avoid forcing the evaluators to read texts with few or no entities, the documents were truncated to at most 20 lines with the highest density of entities. Out of the total 98 documents (40 of which were pseudonymised), two evaluators judged that only 4 and 6 documents were pseudonymised. Out of these, 1 and 2 were original, unedited documents. On average, only 3.5% of documents were correctly judged as pseudonymised.

3. Document Pseudonymisation

If the text to be translated contains personal data or other confidential or identifying information and its open transmission to external machine translation systems is restricted by the provisions of the GDPR (or other laws and/or regulations), then the text must be processed in such a way that the personal data and other sensitive information (e.g. place names, organization names, dates, etc.) are masked before it is sent outside the internal infrastructure of the company or organization. The simplest way to mask such information is to replace it with asterisks or use an appropriate named entity (NE) category as a pseudonym (see Figure 1 for pseudonymisation examples). The task of machine translation systems is to generate the most reliable translation of such anonymised text, trying to translate the context around the masked named entities and, if possible, preserving the masked named entities in the translation.

Given that masking named entities may result in the loss of information necessary for the correct translation of the context, translating anonymised text is objectively expected to result in a lower quality translation than translating non-anonymised text. It is also important that machine translation systems that translate anonymised text can handle the translation of anonymised text. I.e., they must either be able to interpret masks as non-translatable entities or placeholders, or their training data must include examples of anonymised data during training (e.g., [9]).

A more lossless method of anonymising personal data for the translation process is pseudonymisation. In the process of pseudonymisation, named entities are replaced with randomly selected other named entities that belong to the same NE category (see Figure 1). Unlike anonymisation with asterisks or NE categories, in pseudonymisation, NEs are selected so that their inflexions coincide with the inflexions of the replaced NEs. Provided that the selected NEs are well represented in the training data of the machine translation system, the translation quality of the Pseudonymised text should be similar to the translation quality of the original text. Table 1 provides an example sentence that is anonymised and pseudonymised, and further translated with a machine translation system.

Our proposed pseudonymisation method replaces sensitive information with pseudonyms of the same semantic category and in the same grammatical form as original NEs (see Table 1). For many morphologically rich languages, including Latvian and Lithuanian, morphological taggers that can determine the necessary grammatical form are publicly available, while generators to produce the desired form are not available. To solve this problem, we rely on a large monolingual corpus that is processed through a morphological tagger and a named entity recogniser (NER). We then build an extensive dictionary of NEs in various grammatical forms. This dictionary is used to select a pseudonym in the necessary surface form. Therefore, our pseudonymisation method consists of two steps. At first, we construct a pseudonym dictionary, and then we use it as a resource for pseudonymisation to select appropriate candidates.

3.1. Pseudonym Dictionary

First, a large dictionary of entities is constructed using entity lemmas as keys (see Figure 2). Each entity contains a list of grammatical forms and a Wikidata identifier[3] if it is possible to link the entity to Wikidata. In English, usually the list of surface forms is limited to examples in singular and plural forms and nominal and possessive cases. In other, morphologically rich languages, many more forms are available describing case, number, gender, etc. The entities are grouped by the NE category.

We constructed the pseudonym dictionary by applying NER to a large set of monolingual texts extracted from online newspapers and released for the Conference's on Machine Translation (WMT)

[3]https://www.wikidata.org/wiki/Wikidata:Identifiers

Text type	Example
Original text	
Text in source language	Andris Berzins was given a biscuit in Riga by Liga.
Reference translation	Andrim Bērziņam cepumu Rīgā iedeva Līga.
Masked Reference translation	NER cepumu NER iedeva NER.
Machine translation	Andrim Bērziņam cepumu Rīgā uzdāvināja Līga.
Masked Machine translation	NER cepumu NER uzdāvināja NER.
Anonymised text - asterisks	
Text in source language	***** was given a biscuit in ***** by *****.
Machine translation	***** ***** iedeva cepumu ***** .
Masked Machine translation	NER NER iedeva cepumu NER .
Anonymised text - categories	
Text in source language	PER was given a biscuit in LOC by PER.
Machine translation	PER cepumu LOC uzdāvināja PER.
Masked Machine translation	NER cepumu NER uzdāvināja NER.
Pseudonymised text	
Text in source language	Reinis Ozols was given a biscuit in London by Madara.
Machine translation	Reinim Ozolam cepumu Londonā uzdāvināja Madara.
Masked Machine translation	NER cepumu NER uzdāvināja NER.

Table 1
Examples of translations of sentences that are identical in anonymised form (translation of anonymised text is not defined because the anonymised content is ambiguous; it is not possible to prepare a single correct translation for it).

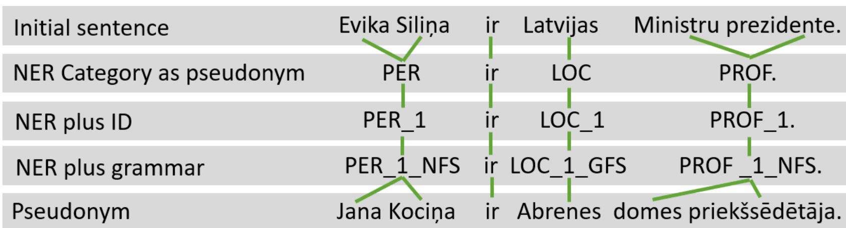

Figure 1: Pseudonymisation methods. The initial sentence, "Evika Siliņa is the Prime Minister of Latvia," is unmasked. Usage of an NE category as a pseudonym is shown in the following row, and in the next two rows, this method is improved by adding a linked entity identifier and grammatical information. The bottom row, "Pseudonym", shows NEs replaced with suitable pseudonyms in the correct inflexion (gender, number, and case).

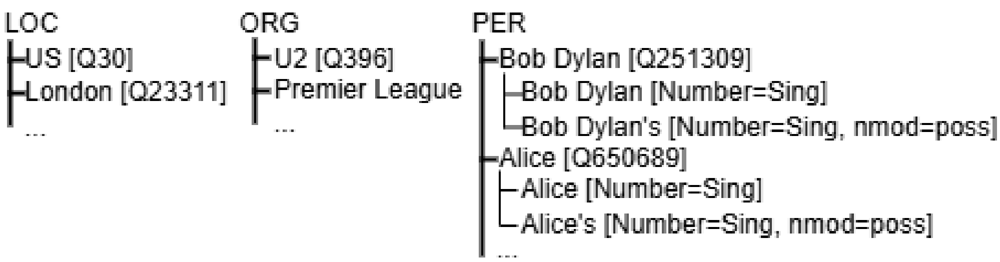

Figure 2: Pseudonym dictionary example.

series of shared tasks[4]. The text is parsed using Stanza [10] to obtain the grammatical information for each sentence and token. From the parsed text, we extract tokens corresponding to the found NEs and save the relevant grammatical information (lemma, part-of-speech, morphological features, headword from dependency parsing) for each NE mention. The entity mentions belonging to an entity are linked

[4]https://data.statmt.org/news-crawl/

using lemma, thus for each entity there is a list of entity mentions defined by their morphological features, e.g., in Figure 2, for the entity "Alice" there are two forms available: singular and singular possessive. We use only the grammatical features of the headword to describe the grammatical form of the whole entity, considering other tokens as contextual elements that do not directly affect the grammatical annotation.

The pseudonym dictionary was limited to the 150 most frequently found entities in each NE category to prevent the selection of obscure words as pseudonyms. The most frequently used entities are usually used in the corpus in various contexts and are found in multiple grammatical forms. This also limits the size of the pseudonym dictionary, allowing for fast pseudonym search. A separate pseudonym dictionary is created for each language used in our experiments.

3.2. Text Pseudonymisation

In order to pseudonymise text, we first must recognise entity mentions using a NER tool. Then, the text is analysed using a syntactic parser (for this, we used Stanza [10]) to determine the grammatical forms of identified entities. If possible, we link entities to a Wikidata identifier. Then, the entities are linked together on document-level using either the Wikidata identifiers or lemmas. This is important as we want to replace entities consistently across the document. Linking produces clusters of entity mentions, which may have multiple grammatical forms (e.g., "*Bob Dylan*", "*Bob Dylan's*").

When candidate NEs for pseudonymisation are identified and linked, we select replacement entities from the pseudonym dictionary in three steps. In this process, we try matching the entity's NE category and all necessary morphological and syntactic properties, and if possible, we prefer replacement candidates that match other attributes, such as token count and the Wikidata category. The three steps are as follows:

1. First, for each entity, we iterate over all entries in the pseudonym dictionary to select a list of candidate replacements having matching morphological and syntactic properties (for at least one entity mention) and the same Wikidata class or superclass. If suitable pseudonym candidate(s) are found, the candidates are ranked by length (counting lemma characters), and the shortest one is selected and added to an exclusion list to avoid selecting the same pseudonym for multiple entities.
2. If no suitable pseudonym candidate(s) are found, then the requirements are relaxed to just matching morphological and syntactic properties.
3. If using the relaxed requirements, no suitable replacement candidates are found, then as a fallback, a randomly selected pseudonym is used from the appropriate NE category.

4. Evaluation of Anonymised Text Machine Translation

The anonymisation process alters the source text, which may have less information than in the original text.As the input data changes, it is to be expected that the translations of the original text and the altered text will differ. When anonymising text with asterisks, asterisks are expected in the translations. When anonymising text with pseudonyms, pseudonym translations are expected in the translations. This means that direct comparison of translations that are obtained by translating a source text that is anonymised using different methods is not possible. To compare the translation quality of texts anonymised using different methods, we instead want to analyse the translation quality of the context (i.e., the text surrounding NEs), ignoring the NEs themselves. Therefore, we designed a new evaluation method for machine translation systems—the masked translation evaluation method—to evaluate the translation quality of texts anonymised using different anonymisation methods.

For the experiments, we use the MultiLeg [11] dataset, which consists of documents from the European Court of Human Rights, originally available in all languages of the European Union. The documents are sentence-level parallel and annotated with NE categories. We use a subset of this dataset covering five languages – English, Estonian, Latvian, Lithuanian and Polish. The original dataset, which contains translations of the original English-language documents into all other languages, is used as a non-anonymised reference. In the result tables (2 and 3), this dataset is labelled "*Original Text*".

All NEs in reference translations are masked using the placeholder "*NER*" (see Figure 3). Masking allows for preserving context, but transforms all NE translations into placeholders irrespective of the anonymisation method used.

Figure 3: Example of masking named entities in a reference translation.

When translating text using a machine translation system, the translation must be aligned with the source text on the token level in order to identify which source tokens got translated into which target tokens. The word (or token) alignment information is necessary to know which token spans in the translation represent, as the masked translation evaluation method requires that all NEs are masked to calculate masked translation quality (i.e., context translation quality). The masking process for pseudonymised text translation is shown in Figure 4. For word alignment, we use simalign [12]. Once the word alignments are obtained, the text spans that represent NEs in the translations are replaced with the placeholder "*NER*", thus obtaining a masked translation. This masked translation will be compared with the masked reference translation to assess the translation quality.

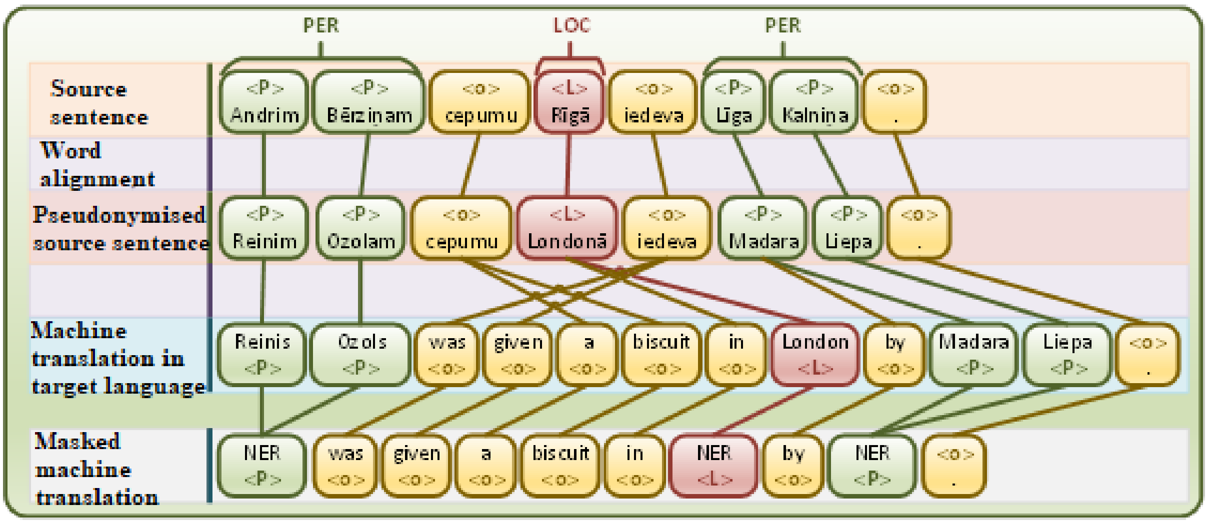

Figure 4: Masking named entities in machine translation of Pseudonymised text.

When reference translations and machine translations of anonymised source texts are masked, we perform automatic evaluation of machine translation quality using the following metrics: the bilingual evaluation understudy (BLEU) [13], character f-measure (ChrF) [14], translation error rate (TER) [15], and COMET [16].

In general, we want the translation of anonymised text to be as similar as possible to the translation of an unanonymised (reference) text. Since we mask NEs in translations and the reference text, we want the translations of the context to be as similar as possible.

For machine translation experiments, we used existing general-purpose neural machine translation (NMT) systems from the Tilde MT platform (production systems). In total, experiments were conducted

Table 2
The results for machine translation systems that translate texts from English.

Translation direction	Text type	NE tag support	BLEU	ChrF	TER	COMET
en -> et	Original text	-	41.23	64.64	48.31	0.8693
	Anonymised(*)	-	32.91	61.71	63.92	0.8336
	Anonymised(C)	-	39.98	62.81	47.73	0.8669
	Pseudonymised	-	41.29	**64.52**	48.29	0.87
	Anonymised(C)	+	**41.81**	64.47	**46.25**	**0.8838**
en -> lt	Original text	-	43.88	66.07	47.25	0.8635
	Anonymised(*)	-	39.33	63.91	52.97	0.8444
	Anonymised(C)	-	44.51	65.16	45.15	0.8671
	Pseudonymised	-	43.93	**65.89**	47.18	0.8626
	Anonymised(C)	+	**44.93**	65.83	**44.32**	**0.8739**
en -> lv	Original text	-	48.72	67.22	48.37	0.8542
	Anonymised(*)	-	41.45	64.75	58.16	0.8372
	Anonymised(C)	-	47.93	66.18	46.70	0.8655
	Pseudonymised	-	48.76	67.11	48.21	0.8526
	Anonymised(C)	+	**49.84**	**67.30**	**45.38**	**0.8727**
en -> pl	Original text	-	44.92	63.07	48.29	0.8494
	Anonymised(*)	-	36.94	60.84	65.89	0.8193
	Anonymised(C)	-	43.84	62.29	47.89	0.8506
	Pseudonymised	-	**44.35**	62.68	48.46	0.8490
	Anonymised(C)	+	44.26	**62.84**	**47.27**	**0.8538**

in eight translation directions – from English into Polish, Estonian, Latvian and Lithuanian and vice versa.

The NMT systems, although supporting non-translatable entities, are not adapted to recognise that a string of asterisks or an NE category should be considered a non-translatable entity. Therefore, for the purposes of the study, new machine translation systems were derived from existing systems, for which additional non-translatable entity recognition rules were defined, which would allow recognising sequences of multiple asterisks and NE tags (PER, LOC, ORG, etc.) as non-translatable entities.

4.1. Evaluation results

The results of the experiments are summarized in two tables. Table 2 summarises the results for machine translation systems that translate texts from English, and Table 3 summarises the results for systems that translate texts into English.

The results of all metrics (BLEU, ChrF, TER, and COMET) show that when translating text from English using anonymisation methods and, in some experiments, also using the developed pseudonymisation methods, it is possible to obtain a better quality translation of the context of the anonymised text than the translation quality of the original text. This may happen when NEs represent rare or unseen words for a machine translation system. When anonymising or replacing the NEs with more common pseudonyms, the machine translation system has an easier task. Considering that English is a language with a fixed word order, anonymising NEs with asterisks or NE categories mostly does not harm translation quality.

Pseudonymised text translation shows good results when translating text from morphologically rich inflectional languages with free word order into English, but relatively lower results than other anonymisation methods when translating from English. This can be explained by the fact that for free word order languages, anonymisation with asterisks or NE categories results in the loss of crucial information on the syntax of the sentence. Since Pseudonymisation preserves morphological and grammatical information, the MT system is presented with mostly lossless information (as long as pseudonymisation was successful). Meanwhile, the fixed word order for English means that the complex

Table 3
The results for machine translation systems that translate texts to English.

Translation direction	Text type	NE tag support	BLEU	ChrF	TER	COMET
et -> en	Original text	-	50.97	67.75	40.14	0.8397
	Anonymised(*)	-	42.26	63.89	52.18	0.8075
	Anonymised(C)	-	47.41	65.65	41.84	0.8303
	Pseudonymised	-	**49.54**	**67.02**	**40.70**	0.8373
	Anonymised(C)	+	49.05	66.75	40.89	**0.8399**
lt -> en	Original text	-	48.45	66.53	40.91	0.8202
	Anonymised(*)	-	41.42	63.02	52.42	0.7910
	Anonymised(C)	-	46.01	64.91	42.31	0.8194
	Pseudonymised	-	**47.32**	**65.86**	**41.30**	0.8184
	Anonymised(C)	+	46.86	65.61	41.73	**0.8222**
lv -> en	Original text	-	53.03	69.44	37.28	0.8314
	Anonymised(*)	-	42.51	65.55	53.80	0.7910
	Anonymised(C)	-	50.01	67.82	38.34	0.8291
	Pseudonymised	-	**51.65**	**68.68**	38.30	0.8201
	Anonymised(C)	+	51.46	**68.69**	**37.58**	**0.8327**
pl -> en	Original text	-	56.85	71.92	34.34	0.8289
	Anonymised(*)	-	41.69	67.20	60.88	0.7689
	Anonymised(C)	-	54.73	70.88	35.17	0.8246
	Pseudonymised	-	**55.55**	**71.24**	**34.88**	0.8279
	Anonymised(C)	+	55.09	71.14	**34.88**	**0.8285**

Table 4
Evaluation results on a subset of 500 manually aligned segments.

Translation direction	Text type	NE tag support	BLEU	ChrF	TER	COMET
en -> lv	Anonymised(*)	-	47.16	66.46	46.49	0.7452
	Anonymised(C)	-	49.58	67.56	44.65	0.7645
	Pseudonymised	-	**52.13**	**68.96**	**43.01**	0.7894
	Anonymised(C)	+	50.88	68.30	43.37	**0.8016**
lv -> en	Anonymised(*)	-	45.33	64.65	41.69	0.4949
	Anonymised(C)	-	47.68	66.40	38.60	0.5663
	Pseudonymised	-	**50.34**	**67.87**	**37.17**	0.5847
	Anonymised(C)	+	**50.34**	67.82	37.33	**0.5955**

process of pseudonymisation may not be necessary to acquire good-quality translations.

In some cases, simalign produces incorrect word alignments for the pseudonymised text translations, which most often occurs for NEs representing dates. E.g., initially the entity "*29 September 2008*" and its translation "*2008. gada 29. septembra*" is masked as a single entity – "*NER*". When translating text anonymised using asterisks or NE categories, the translation remains masked with a single mask; however, pseudonymised texttranslation contains the plaintext pseudonym "*2008. gada 15. septembra*". For this, simalign for some reason outputs incorrect word alignments and the NE is incorrectly masked as "*NER. gada NER. NER*". Therefore, we tasked a translator to revise masking for a subset of random 500 lv-en and en-lv segments to see if the use of simalign may have introduced inaccuracies in the evaluation process. The results of the manually aligned data are shown in Table 4.

The results of this evaluation show that excluding the alignment errors introduced by the automatic aligner, the context translation quality is higher when using pseudonyms according to the BLEU, ChrF, and TER metrics, while the COMET metric shows a higher score when the entities were anonymised using category labels. Since COMET is a trained metric that has not been pre-trained for NE-masked translations, further research would have to be conducted to assess its reliability.

5. Conclusion

This paper addressed the challenge of translating anonymised text with MT systems by analysing various anonymisation and pseudonymisation strategies and introducing a novel evaluation method, masked translation evaluation. This method enables comparison of translation quality around masked named entities without being affected by the entity replacements themselves.

We developed and adapted a pseudonymisation method for machine translation in morphologically rich languages. This method ensures that pseudonyms retain as much information as possible while restricting them to the most frequently encountered entities.

Experimentally, it was found that pseudonymisation ensures the highest quality of contextual translation when translating from morphologically rich languages. This is because named entities contain information essential to the syntax of the sentence, which is lost when NEs are anonymised using simply asterisks or NE categories. At the same time, anonymisation using NE categories allows producing a high-quality contextual translation when translating from English, provided that the MT systems are adapted to translation using placeholders.

We also observed an initially unexpected, but in hindsight very much explainable result – when translating from English, for three out of four language pairs, translation of anonymised text produced better results than translation of the original text. This happened because the original named entities (and named entities in general) represent sparse and rare words with which the MT system struggled. Anonymisation reduced uncertainty and thereby improved translation quality. A crucial prerequisite for this observation was the fact that English has a strict word order, and thereby, replacing content words with placeholders does not result in the loss of syntactic information necessary for correct translation.

Acknowledgments

This research has been supported by the ICT Competence Centre (www.itkc.lv) within the project 2.5 *Daudzvalodīgs dokumentu anonimizācijas un pseidonimizācijas mākslīgā intelekta risinājums* (Multilingual artificial intelligence solution for document anonymization and pseudonymization) of EU Structural funds, ID no 5.1.1.2.i.0/1/22/A/CFLA/008.

Declaration on Generative AI

During the preparation of this work, the authors used Google Translate (https://translate.google.com) to translate text from Latvian into English. All translations were post-edited and carefully revised. The authors also used Grammarly (https://www.grammarly.com) to correct spelling and grammar mistakes. All edits by Grammarly were reviewed by the authors and edited as needed. The authors take full responsibility for the publication's content.

References

[1] A. Papadopoulou, Y. Yu, P. Lison, L. Øvrelid, Neural text sanitization with explicit measures of privacy risk, in: Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Association for Computational Linguistics, Online only, 2022, pp. 217–229. URL: https://aclanthology.org/2022.aacl-main.18.

[2] H. Dalianis, Pseudonymisation of Swedish electronic patient records using a rule-based approach, in: L. Ahrenberg, B. Megyesi (Eds.), Proceedings of the Workshop on NLP and Pseudonymisation, Linköping Electronic Press, Turku, Finland, 2019, pp. 16–23. URL: https://aclanthology.org/W19-6503.

[3] V. Arranz, K. Choukri, M. Cuadros, A. García Pablos, L. Gianola, C. Grouin, M. Herranz, P. Paroubek, P. Zweigenbaum, MAPA project: Ready-to-go open-source datasets and deep learning technology

to remove identifying information from text documents, in: Proceedings of the Workshop on Ethical and Legal Issues in Human Language Technologies and Multilingual De-Identification of Sensitive Data In Language Resources within the 13th Language Resources and Evaluation Conference, European Language Resources Association, Marseille, France, 2022, pp. 64–72. URL: https://aclanthology.org/2022.legal-1.12.

[4] T. Váradi, S. Koeva, R. Garabík, A. Repar, B. Nitoń, V. Štefanec, E. Irimia, Curlicat deliverable 3 anonymization, 2022. URL: https://curlicat-project.eu/assets/downloads/D3_Anonymisation.pdf.

[5] M. Abdalla, M. Abdalla, F. Rudzicz, G. Hirst, Using word embeddings to improve the privacy of clinical notes, Journal of the American Medical Informatics Association 27 (2020) 901–907. URL: https://doi.org/10.1093/jamia/ocaa038. doi:10.1093/jamia/ocaa038. arXiv:https://academic.oup.com/jamia/article-pdf/27/6/901/34152514/ocaa038.pdf.

[6] D. Faraglia, Faker, 2014. URL: https://faker.readthedocs.io/en/master/.

[7] K. Chatzitheodorou, M. Ángeles García Escrivá, C. G. Lacal, Machine translation of anonymized documents with human-in-the-loop, in: M. Nurminen, J. Brenner, M. Koponen, S. Latomaa, M. Mikhailov, F. Schierl, T. Ranasinghe, E. Vanmassenhove, S. A. Vidal, N. Aranberri, M. Nunziatini, C. P. Escartín, M. Forcada, M. Popovic, C. Scarton, H. Moniz (Eds.), Proceedings of the 24th Annual Conference of the European Association for Machine Translation, European Association for Machine Translation, Tampere, Finland, 2023, pp. 429–436. URL: https://aclanthology.org/2023.eamt-1.41/.

[8] D. Pissarra, I. Curioso, J. Alveira, D. Pereira, B. Ribeiro, T. Souper, V. Gomes, A. V. Carreiro, V. Rolla, Unlocking the potential of large language models for clinical text anonymization: A comparative study, 2024. URL: https://arxiv.org/abs/2406.00062. arXiv:2406.00062.

[9] M. Pinnis, R. Krišlauks, D. Deksne, T. Miks, Neural machine translation for morphologically rich languages with improved sub-word units and synthetic data, in: International conference on text, speech, and dialogue, Springer, 2017, pp. 237–245.

[10] P. Qi, Y. Zhang, Y. Zhang, J. Bolton, C. D. Manning, Stanza: A Python natural language processing toolkit for many human languages, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, 2020. URL: https://nlp.stanford.edu/pubs/qi2020stanza.pdf.

[11] R. Vīksna, I. Skadiņa, MultiLeg: Dataset for text sanitisation in less-resourced languages, in: N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, N. Xue (Eds.), Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), ELRA and ICCL, Torino, Italia, 2024, pp. 11776–11782. URL: https://aclanthology.org/2024.lrec-main.1028.

[12] M. Jalili Sabet, P. Dufter, F. Yvon, H. Schütze, SimAlign: High quality word alignments without parallel training data using static and contextualized embeddings, in: T. Cohn, Y. He, Y. Liu (Eds.), Findings of the Association for Computational Linguistics: EMNLP 2020, Association for Computational Linguistics, Online, 2020, pp. 1627–1643. URL: https://aclanthology.org/2020.findings-emnlp.147/. doi:10.18653/v1/2020.findings-emnlp.147.

[13] K. Papineni, S. Roukos, T. Ward, W.-J. Zhu, Bleu: a method for automatic evaluation of machine translation, in: P. Isabelle, E. Charniak, D. Lin (Eds.), Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Philadelphia, Pennsylvania, USA, 2002, pp. 311–318. URL: https://aclanthology.org/P02-1040/. doi:10.3115/1073083.1073135.

[14] M. Popović, chrF++: words helping character n-grams, in: O. Bojar, C. Buck, R. Chatterjee, C. Federmann, Y. Graham, B. Haddow, M. Huck, A. J. Yepes, P. Koehn, J. Kreutzer (Eds.), Proceedings of the Second Conference on Machine Translation, Association for Computational Linguistics, Copenhagen, Denmark, 2017, pp. 612–618. URL: https://aclanthology.org/W17-4770/. doi:10.18653/v1/W17-4770.

[15] M. Snover, B. Dorr, R. Schwartz, L. Micciulla, J. Makhoul, A study of translation edit rate with targeted human annotation, in: Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers, Association for Machine Translation in the

Americas, Cambridge, Massachusetts, USA, 2006, pp. 223–231. URL: https://aclanthology.org/2006.amta-papers.25/.
[16] R. Rei, C. Stewart, A. C. Farinha, A. Lavie, COMET: A neural framework for MT evaluation, in: B. Webber, T. Cohn, Y. He, Y. Liu (Eds.), Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, Online, 2020, pp. 2685–2702. URL: https://aclanthology.org/2020.emnlp-main.213/. doi:10.18653/v1/2020.emnlp-main.213.