TECHNICAL AND LEGAL POSSIBILITIES TO USE ARTIFICIAL INTELLIGENCE TO PROVIDE TRANSLATION AND INTERPRETATION IN CRIMINAL PROCEEDINGS

IT Law Lab Task Solution

Arina Stivrina
Liva Rudzite
Marko Jaeger

Master’s Exam Part I
Lecturers:
Martin Ebers, Liisi Jürgen,
Marta Cantero Gamito,
Karmen Turk, Liisi Adamson
Consulting Entity:
Kristi Hallas
Office of the Prosecutor General

Tartu

2020
# TABLE OF CONTENTS

INTRODUCTION ................................................................................................................................. 3

1. STATE-OF-THE-ART OF MACHINE TRANSLATION APPROACHES ...................................... 4
   1.1. Rule-based and corpus-based (statistical) MT approaches ....................................................... 4
   1.2. Neural MT approach .................................................................................................................. 5
   1.3. Speech-to-speech MT ............................................................................................................... 8
   1.4. Human parity ................................................................................................................................. 9
   1.5. Post-editing with MT .................................................................................................................. 10
   1.6. Open source MT systems ............................................................................................................ 11

2. LEGAL BASIS UPON THE EUROPEAN UNION AND INTERNATIONAL LAW FOR
   AUTOMATED OR MACHINE TRANSLATION AND INTERPRETATION ................................. 12
   2.2. Possibility to adequately explain the circumstances of the crime ........................................ 16
   2.3. Specific provisions of Directive 2012/29/EU regarding translation ........................................ 18
   2.4. The right to a fair trial ................................................................................................................. 20

3. CURRENT LEGAL REGULATION IN ESTONIA AND POSSIBILITY OF USING
   AUTOMATED AND MACHINE-ASSISTED TRANSLATION AND INTERPRETATION .......... 22
   3.1. General provisions on translation and interpretation ................................................................. 22
   3.2. Special provisions for the translator or interpreter .................................................................... 24
   3.3. Challenges with providing translation and interpretation in Estonia ....................................... 25
   3.4. Possibility of using automated and machine assisted translation and interpretation considering current Estonian regulation .................................................................................................................. 27
   3.5. Amending the law ....................................................................................................................... 28

SUMMARY ......................................................................................................................................... 30

OVERVIEW OF THE CONTRIBUTION ......................................................................................... 32

REFERENCE LIST ........................................................................................................................... 33
INTRODUCTION

The objective of the IT Law Lab Task is to analyse the different approaches of machine translation and interpretation, the state of current technology and the compatibility of those technologies with the European Union, international and Estonian law of criminal proceedings. To achieve this objective relevant sources of literature, legal texts in Estonian and foreign languages, as well as the case-law of the national, regional and international courts in the field of criminal proceedings have been studied and analysed. Methods used in work comprise: 1) data gathering, processing; 2) interpreting, presenting the results for each chapter; 3) chronologically analysing the historical pattern nationally, systematically; 4) comparing the related fields regionally, internationally to inductively, deductively and qualitatively conduct conclusions.

This resulting analysis is relevant at this time and age as the world globalises and a variety of languages need to be translated or interpreted; hence, hindering the possibility to find enough proficient people to provide the necessary services, especially in a small state like Estonia. The cost of human translation is rising; thus, helpful tools to alleviate that cost could be beneficial. Also, human translations are slow in comparison with the capabilities of computers and using those technologies could make criminal proceedings more effective and less time-consuming.

The main problems analysed in work are: 1) if the state of technology has developed far enough to allow the usage of dependable machine operated solutions; a) which solutions are possible regarding technology; b) would it constitute a saving of funds; 2) international and local law has to be considered as it sets boundaries on the usage of technology, especially in a highly regulated field like criminal proceedings considering an immense effect of the criminal proceedings on the people and society if the state wants to function as a real democracy.

The work is split into three main chapters. The first gives an overview and analyses the technical possibilities and approaches and tries to find an answer to the question if current technical capabilities are suitable and beneficial for use in complex translations. The second chapter gives an overview of the European Union and international law regarding criminal proceedings. These have to be considered for the Estonian national law to comply when trying to adopt new technologies into their own system of criminal proceedings. The third chapter provides an analysis of Estonian national law on the subject, points out the current challenges of providing translation and interpretation and tries to find an answer to the question if technologies could be used under the barriers set by the national law or if some amendments would have to be done.

Keywords: machine, translation, interpretation, criminal, law.
1. STATE-OF-THE-ART OF MACHINE TRANSLATION APPROACHES

The term machine translation (MT) refers to computerized systems responsible for the production of translations\(^1\) with or without human assistance. At a very basic level, machine translation algorithm identifies a word in a source language, finds the corresponding word in a target language and replaces one with another. However, such a direct approach usually does not produce good results, as natural languages are complex. Words may have multiple meanings in different contexts; language and grammar structures may differ significantly.

To face these challenges, many MT approaches have been developed. In terms of technology, MT is classified into rule-based (including transfer-based, interlingual, dictionary-based), and corpus based\(^2\) (including statistical and example-based or phrase-based MT\(^3\)-\(^4\)). Other approaches include neural MT\(^5\) and hybrid MT\(^6\). To determine the state-of-the-art in MT, the authors will further review three main MT paradigms\(^7\) or approaches such as rule-based, corpus-based and neural MT.

1.1. Rule-based and corpus-based (statistical) MT approaches

Rule-based MT (RBMT), also called knowledge-based MT or classical approach, is used to describe MT systems, where linguistic information about the source and target languages is retrieved from dictionaries and MT relies on a set of rules and exhaustive descriptions within three phases: analysis, transfer and synthesis components. RBMT systems were the first commercial MT systems\(^8\). Doug A. categorized the following problems in the classical approach to MT: 1) analysis problem, referring to difficulties to infer the content, deriving from ambiguity

---

\(^4\) Example-based or phrase-based MT utilizes similarities of phrase examples available in dictionaries.
and imperfect grammatic structure of the source text; 2) transfer problem, which refers to rules of transfer between (dramatically different) language structures; 3) synthesis problem, referring to the correct choice of grammatic structure; 4) the problem of description, which refers to amount of rules and descriptions analysis-transfer-synthesis components must contain\(^9\), leading to great resources and costs consumption, deeming RBMT systems expensive.

The corpus-based or statistical MT (SMT) approach relies on calculation of probabilities for possible representations of the given phrase or word in a source language for the target language\(^10\). A phrase (in phrase-based approach), word with the highest probability is chosen. The probability for a given word to appear in a sentence before or after other words also might be analysed. SMT usually requires large parallel translation texts\(^11\) for calculation of probabilities. Prior to neural MT, phrase-based MT (PBMT) approach was considered state-of-the-art MT\(^12\).

Figure 1. Advantages and disadvantages of RBMT and SMT systems\(^13\)

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBMT</td>
<td></td>
</tr>
<tr>
<td>Based on linguistic theories</td>
<td>Requires linguistic rules and dictionaries</td>
</tr>
<tr>
<td>Adequate for languages with limited resources</td>
<td>Human Language inconsistency (i.e. exception to the rules)</td>
</tr>
<tr>
<td>Does not require many computational resources</td>
<td>Disambiguation problems</td>
</tr>
<tr>
<td>Easy to perform error analysis</td>
<td>Local translations, Language dependent</td>
</tr>
<tr>
<td></td>
<td>Expensive to maintain and extend</td>
</tr>
<tr>
<td>SMT</td>
<td></td>
</tr>
<tr>
<td>No linguistic knowledge required</td>
<td>Requires parallel text</td>
</tr>
<tr>
<td>Reduces the human cost</td>
<td>Requires high computational resources</td>
</tr>
<tr>
<td>Easy to build</td>
<td>Difficult to perform error analysis</td>
</tr>
<tr>
<td>Easy to maintain (if data available)</td>
<td>Problems with pairs of languages with different morphology/order</td>
</tr>
<tr>
<td>Trained with human translations (extract knowledge from corpus)</td>
<td>No linguistic background</td>
</tr>
</tbody>
</table>

1.2. Neural MT approach

A neural MT (NMT) approach involves artificial intelligence-powered neural networks. Artificial intelligence (AI) refers to systems that display intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals\(^14\). Neural network (NN) consists of many simple; connected processors called neurons, which get

---


\(^10\) Doug A. See reference 9, p.139.


\(^14\) Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe, Brussels, 25.4.2018 COM (2018) 237 final.
activated either through input from environment or through weighed connections from previously active neurons. The “learning” part of NN is about finding weights that make the NN exhibit desired behaviour and, if a problem required long causal chains of computational stages, where each stage may transform the activation of the network, “deep learning” is about correctly assigning weights across many such stages. The main advantage of deep learning is that, by developing the right architecture, the system automatically learns features from data without the need for explicitly designing them. Most NMT architectures are based on encoder-decoder approach, where an input in the source language is decoded into a sequence of numbers that represent the sentence meaning.

Figure 2. Standard architecture of NMT systems (2017)

The new paradigm of NMT has yielded outstanding results and improved state-of-the-art results for several language pairs, outperforming PBMT. Since 2015, NMT is considered the state-of-the-art in the field of MT. By 2016, most of the best MT systems used NN.

Several risks and drawbacks of NMT should be discussed. First, the translation quality of NMT depends on training data, and therefore, if NMT system is trained on biased or outdated data, it’s output will be skewed accordingly. While the obvious solution is to train NMT system

---

16 Ibid.
on high-quality, up-to-date corpus, the availability of high-quality parallel corpus is a showstopper for NMT development in many language pairs.

The biased data problem, however, might not really have a viable solution, as biased data origin is ourselves, our human biases which make their way into training data. Second, NMT systems perform poorly when translating inputs which are drastically different from those encountered during training. Thus, NMT system which is trained on high-level European Union legislative corpus will perform poorly on casual friends’ conversation. Third, the current state-of-the-art sentence-level MT makes it difficult for NMT system to derive the document-level or paragraph-level, or, in extreme examples, even the next sentence context. The fourth drawback is the inability to correctly assign the contextual meaning, as illustrated by figure 4, Latvian translation of the displayed phrase asserts “I’m a huge metal ventilator”. However, the English-Russian translation of the phrase did not present the same context error.

Figure 3. “Google translate” NMT context error

The translation quality of NMT systems depends strongly on the training data size; however, sufficient amounts of parallel data are not available for many language pairs. While a lot of research is focused on languages with large native speaker bases, such as English, Spanish, Chinese, French, Arabic, Spanish, Russian, Japanese and corresponding language pairs, smaller languages, such as Estonian, Latvian and Lithuanian, and language pairs are disadvantaged in terms of available linguistic resources and technological approaches, enabling a technological gap between the two groups of languages. All three Baltic languages, however, benefitted from

---

inclusion in the set of translation tasks of the Conference on Machine Translation (WMT). Latvian was featured at WMT 2017\textsuperscript{28}, Estonian at WMT 2018\textsuperscript{29} and Lithuanian at WMT 2019\textsuperscript{30}.

Therefore, state-of-the-art of NMT may differ significantly between language pairs with a large amount of parallel data and those with the insufficient amount.

### 1.3. Speech-to-speech MT

In order to fully deploy MT in criminal proceedings, speech-to-speech (“STS”) MT must reach human parity. STS MT is an existing technology in production stage\textsuperscript{31} and, in principle, requires adding “speech-to-text” automatic voice recognition before the encoder and “text-to-speech” layer after the decoder, additional layers somewhat slowing down the MT speed. However, the direct sequence-to-sequence MT system “Translatotron”, which can directly translate speech from one language to speech in another language, without intermediate text representation, taking source spectrograms in one language and generating spectrograms of the translated content in another language, is in development by Google\textsuperscript{32}. As demonstrated by Google, it achieved high translation quality on two Spanish-to-English datasets, although performance is not as good as a baseline cascade of speech-to-text and text-to-speech models\textsuperscript{33}.

Figure 4. “Translatotron” STS MT model architecture\textsuperscript{34}.

![“Translatotron” STS MT model architecture](image_url)


\textsuperscript{33} Ibid.
\textsuperscript{34} Ibid.
1.4. Human parity

Human parity assessment and metrics for human evaluation of MT are subject to controversy. Three main measures to assess human parity can be identified: 1) measurement type, referring to how the measurement is performed (e.g. “Which text is better?”, “How good is this translation from 1-5”, “Assess translation quality in scale from 0 to 100” etc.), 2) evaluators (experts vs bilingual workers), and 3) unit of evaluation, (sentences vs documents). Three main measures to assess human parity can be identified: 1) measurement type, referring to how the measurement is performed (e.g. “Which text is better?”, “How good is this translation from 1-5”, “Assess translation quality in scale from 0 to 100” etc.), 2) evaluators (experts vs bilingual workers), and 3) unit of evaluation, (sentences vs documents).35

Google AI research in 2016 claimed to bridge the gap between human, machine translation, based on assessments shown in Figure 4. Microsoft (MS) AI & Research 2018 study claimed that MS latest NMT system has reached a new state-of-the-art, translation quality being at human parity when compared to professional human translations and significantly exceeding the quality of crowd-sourced non-professional translations. However, both Google (2016) and MS (2018) studies assessed human parity based on evaluation of bilingual crowd worker, not professional translators, as well as used sentences as a unit of evaluation.

Several follow-up reassessments studies have shown evidence that MT systems, including MS in 2018, have not yet reached human parity at the level of documents, using experts as evaluators and performing sufficient power statistical analysis.

---

Findings of the 2019 Conference on Machine Translation (WMT19) indicate that while sentence-level text-to-text MT for some language pairs might be considered to draw conclusions on human parity, the evaluation of isolated segments is not meaningful from the perspective of drawing conclusions on human parity for MT as a whole. The authors also deem the evaluation of isolated segments not useful for purpose of using MT in criminal proceedings. The 2019 declaration by Facebook AI that its MT translations have been declared “superhuman” by the WMT organizers is a public relations statement rather than scientifically approved result.

Thus, it can be concluded that while NMT significantly improved state-of-the-art results of MT, currently, using documents as a unit of evaluation, performing the evaluation with the help of professional translators and performing sufficient power statistical analysis, the state-of-the-art text-to-text MT cannot yet be considered as having reached human parity.

1.5. Post-editing with MT

The conclusion that MT cannot be used in criminal proceedings without human post-editing necessitates the question, can using MT with human post-editing deliver significant time, and thus, cost savings and lead to better quality. To answer this question, the authors will consider text-to-text MT. Results obtained for statistical MT show that post-editing MT for subtitles is on average, 40% faster than translating from scratch. Latest results for neural MT show differences for different language pairs, e.g. the use of NMT enabled professional translators to work 59.74% faster in German–French and 9.26 % in German–Italian. Using NMT with post-editing did not have a negative impact on quality, rather for some language pairs, expert scores were slightly higher.

Therefore, it can be concluded that post-editing MT results could lead to significant cost-savings while not decreasing quality. Thus, text-to-text MT with human-post editing could be considered to decrease translation costs within criminal proceeding system.

41 German-English, English-German, English-Russian.
46 Ibid.
1.6. Open source MT systems

Open source MT systems are often available as open-source implementation toolkits, the most frequently used in research studies, according to Popescu-Bells (2019) being DL4MT: Deep Learning for Machine Translation, Fairseq: Facebook AI Research Sequence-to-Sequence Toolkit, Marian: Fast Neural Machine Translation in C++, Nematus: Open-Source Neural Machine Translation in TensorFlow, Neural Monkey, OpenNMT, Sockeye, Tensor2Tensor (all aforementioned toolkits available at GitHub repository)\(^{47}\).

GitHub search for “neural machine translation” performed by authors returned 959 repository results in more than 10 programming languages\(^{48}\). Top three non-deprecated toolkit:

1. TensorFlow Neural Machine Translation tutorial (permissive Apache 2.0 licence) repository with a tutorial for replicating Google’s NMT system and references to GitHub repositories with Google 2017 NMT model code with encoder and decoder, attention model\(^{49}\).

2. OpenNMT (permissive MIT licence) repository contains sequence-to-sequence models (encoder-decoder) with multiple RNN cells (lstm/gru), attention model, image-to-text and speech-to-text processing\(^{50}\).

3. Facebook research UnsupervisedMT (Creative Commons Attribution-NonCommercial 4.0 International Public License) repository with the original implementation of the unsupervised PBSMT and NMT models, supporting three MT architectures (seq2seq, biLSTM + attention, Transformer)\(^{51}\).

Therefore, it can be concluded that state-of-the-art NMT system model implementations are widely available at GitHub repository as open-source toolkits under permissive, public and non-commercial licenses, which often allow modification and distribution under conditions of publishing the modified code, non-commercial use, and other. Conditional on further exhaustive research of chosen toolkit licensing conditions, \textit{prima facie} it could be possible to use the state-of-the-art models to implement and train open-source NMT system for use in criminal proceedings.

---


\(^{48}\) Available at https://github.com/search?q=neural+machine+translation&s=stars&type=Repositories.


2. LEGAL BASIS UPON THE EUROPEAN UNION AND INTERNATIONAL LAW FOR AUTOMATED OR MACHINE TRANSLATION AND INTERPRETATION


According to the Recital 34 of Directive 2012/29/EU\textsuperscript{52} (hereinafter – Directive 2012/29/EU) possibility to adequately explain the circumstances of the crime, to provide evidence in a manner understandable to the competent authorities, ability to access the rights, to avail treatment in respectful manner are vital to effectively achieve justice.

Therefore, interpretation should be free of charge, made available during questioning of the victim and in order to enable to participate actively in court hearings as well as in accordance with the role of the victim in the relevant criminal justice system. For other aspects of criminal proceedings, the need for interpretation and translation can vary depending on specific issues, the role of the victim in the relevant criminal justice system, involvement in proceedings as well as concerning the possessed rights. As such, interpretation and translation for these other cases need only be provided to the extent necessary for victims to exercise their rights.

What derives from the above mentioned is that Directive 2012/29/EU comprises conditions of translation and interpretation pertaining specifically to victims not addressing analogous aspects of other involved parties in criminal proceedings, for instance, witness, a person against whom a criminal proceeding has been initiated and others. Nonetheless, Recital 66 of the Directive stipulates that it respects fundamental rights and observes the principles recognized by the Charter of Fundamental Rights of the European Union\textsuperscript{53} (hereinafter – the Charter) and the respective rights of the Convention on Human Rights\textsuperscript{54} (hereinafter – ECHR)\textsuperscript{55}. In particular, the Directive seeks to promote the right to dignity, life, physical and mental integrity, liberty and security, respect for private and family life, the right to property, the principle of nondiscrimination, the principle of equality between women and men, the rights of the child, the elderly and persons with disabilities, and the right to a fair trial. Hence, Directive 2012/29/EU should be constituted as \textit{lex specialis} against the fundamental rights embedded in the Charter, and

also towards the intended respective rights of the other involved persons within criminal proceedings. Thus, also, the automated translation should address the relevant aspects.

That could be particularly important, especially regarding the possible application of the completely autonomous machines resulting from artificial intelligence that do not involve a human intervention or human oversight.\(^5\) Namely, the necessity for the machine to distinguish between the different procedural status of a person involved in a criminal proceeding. That could be particularly cumbersome also because Recital 18 of Directive 2012/29/EU states that the protection of individuals should be technologically neutral and not depend on the techniques used to prevent creating a severe risk of circumvention. In addition, Recital 38 in conjunction with Article 11(1) of Directive (EU) 2016/680\(^6\) stipulates that the data subject should have the right not to be subject to a decision evaluating personal aspects based solely on automated processing and which produces adverse legal effects concerning, or significantly affects, a person. Suitable safeguards must be applied, including the provision of specific information to the data subject, the right to obtain human intervention and an explanation of the decision reached after such assessment or to challenge the decision. Whereas, profiling resulting in discrimination against individuals based on sensitive data in relation to fundamental rights should be prohibited.

Hence, it derives that entirely autonomous translation or interpretation machine could hardly fulfil the mentioned criteria, especially the requirement about the technological neutrality and automated processing because in order for a machine to distinguish between the criminal procedural status of the person the machine would have to process, for instance, sensitive data as voice sample, handwritten application, testimony or biometric data\(^7\) that based on Article 9(1) of Regulation (EU) 2016/679 will require the consent of a victim or the data subject since the particular activity falls under the fundamental rights to privacy of the data subject – victim and

---


\(^{9}\) Regulation (EU) 2016/679, Article 9(2)(a).
under the rights to a fair trial as the rights to or not to participate in proceedings, rights to agree to machine translation and automated processing of sensitive data; thus, differs from the ordinary conditions of biometric data processing in law enforcement sector and falls outside the Article 10 of Directive (EU) 2016/680. Another method would be to process pictures to determine the victim that could lead to profiling, result in discrimination depending on the data set by which the machine would have been trained or even to the false results and, consequently, produce adverse legal effects concerning a person if its criminal procedural status would be determined incorrectly by the machine or significantly affect the person who has already suffered from the possible criminal wrongdoings; thus, per se is more vulnerable.

Furthermore, for the consent to be valid it has to be freely given, specific, informed and unambiguous⁶⁰ which means that the data subject should not only be fully aware of the data processing per se but also about the techniques used. That is of particular importance regarding autonomous machines resulting from artificial intelligence in deep or self-learning form leading to singularity or surpassing human intelligence due to the so-called ‘black box’ effect or a lack of transparency of a decision making⁶¹, and; thus, ambiguity as well as the lack of information.

Besides, the data subject also has rights to erasure or rights to be forgotten, namely, rights to withdraw consent and request that its data will no longer be processed⁶² or delinked⁶³. That could be extraordinary cumbersome for entirely autonomous machines that apply the self-learning method, meaning it learns from inserted data. Therefore, once the machine has processed the data, it has already learned, consequently, data erasure or delinking of the data subject will not reverse the already acquired knowledge resulting from previously inserted data. Moreover, fulfilment of the data erasure could be onerous also for the lower artificial intelligence forms regarding automated translation or interpretation due to the nature of biometric data, namely, uniqueness that can lead to identification of the data subject⁶⁴ even if indirectly⁶⁵ since the voice due to physical and behavioural characteristics as a length of the vocal tract, nasal passage, pitch, cadence, accent is as unique for a person as a fingerprint is⁶⁶. That means that pure delinking of

⁶⁰ Regulation (EU) 2016/679, Recital 32.
⁶² Regulation (EU) 2016/679, Article 17(1).
⁶³ CJEU, C-131/12, Google Spain and Google, para 98; CJEU, C-507/17, Google v. CNIL, para 45.
⁶⁴ ECHR 44787/98 P.G. and J.H. v. the United Kingdom, para 59.-60.
⁶⁵ CJEU C-582/14, Patrick Breyer v. Bundesrepublik Deutschland, para. 49.
voice sample and the data subject may still lead to a unique identification. Hence, the only option for not fulfilling the criteria of data protection would be anonymization of data\textsuperscript{67}, but that could not be an option in criminal proceedings due to the necessity to identify the victim.

What is more, even if Directive (EU) 2016/680 does not contain specific provisions regarding data processing of a child, nonetheless, in activities that fall outside the scope of Directive (EU) 2016/680, provision of information society service for a child is only allowed if the child has reached at least 16 years or not younger than 13 years if the national law permits and only if and to the extent that consent is given or authorized by the holder of a parental responsibility over the child\textsuperscript{68}. It means that if automated translation or interpretation service would be provided remotely in the form of information society service due to the wording of Recital 27 of Directive 2012/29/EU (information for the victims may be provided through the press, an official website of the competent authority or a similar communication channel), the age limits, conditions of the consent should be considered especially concerning the age limits of involvement in criminal proceedings by the national law. Nonetheless, Recital 14 of Directive 2012/29/EU does not stipulate the definition of a ‘child’ but refers to the best interests of a child under the Charter and Convention on the Rights of the Child under which child is a human being below the 18 years unless the majority is attained earlier\textsuperscript{69}. Besides according to Recital 42 right to be heard in criminal proceedings should not be precluded solely based on the age, while Article 2(1)(c) stipulates that the term ‘child’ refers to any person below 18 years of age. Thus, in activities that fall under Directive (EU) 2016/680 age limit for a victim–a child is below 18 years.

Recapitulating, currently the most suitable form of automated or machine translation service resulting from artificial intelligence could rather be a machine that involves human oversight in the form of human-in-the-loop (human intervention in every decision cycle of the system), human-on-the-loop (human intervention during the design cycle of the system and monitoring the operation of the system) or human-in-command (oversight of the overall activity of the system including economic, societal, legal and ethical impact as well as the discretion to decide upon usage of the system)\textsuperscript{70} approach and not entirely autonomous machine. The approach would comply with the mentioned guidelines by the European Commission as well with the vision of the Council of Europe regarding the principle ‘under user control’ (users are informed

\textsuperscript{67} Regulation (EU) 2016/679, Recital 26.
\textsuperscript{68} Ibid. Article 8(1).
\textsuperscript{70} Independent High Leven Expert Group on Artificial Intelligence set up by the European Commission. Ethics Guidelines for Trustworthy AI, 8 April 2019.
actors and in control of their choices)\textsuperscript{71} and by the European Parliament regarding human control, verification as the safeguard for the process of automated and algorithmic decision-making\textsuperscript{72}.

In addition, human oversight also resolves the issue of liability. Namely, Article 6(1) of Directive 85/374/EEC\textsuperscript{73} stipulates that product is defective if it does not provide the safety which a person is entitled to expect. Namely, concerning not only the intended purpose, the objective characteristics and properties of the product but also the specific requirements of the users for whom the product is intended\textsuperscript{74}. While Article 7(2) states that the producer shall not be liable if it is probable that the defect which caused the damage did not exist at the time when the product was put into circulation by a producer or that this defect came into being afterwards. Albeit Article 2 of Directive 85/374/EEC stipulates that product comprises electricity but does not state whether a product shall also be attributed to software, nevertheless, Article § 1063 (1) of the Estonian Law of Obligations Act\textsuperscript{75} expressis verbis attribute product definition also to computer software. Henceforth, to prevent uncertainty of liability issues, automated or machine translation system should be explicitly designed or adjusted considering the needs of criminal proceedings as well as the provisions outlined in Directive 2012/29/EU. Besides human oversight should be involved either as verification, control by a human interpreter in the presence or remotely as well as readability to provide translation or interpretation service by a human as allows Article 7(2) of Directive 2012/29/EU. In addition, prior consent from a data subject should be acquired. Thus, depending on the circumstances of the case, liability should be attributed to the respective person.

\textbf{2.2. Possibility to adequately explain the circumstances of the crime}

Recital 34 of Directive 2012/29/EU stipulates a necessity to provide an opportunity to properly explain the circumstances of the crime as well as to provide evidence in a manner understandable to the competent authorities. While Article 3 states the right to understand and to be understood, namely, that communications with victims are given in a simple and accessible language, orally or in writing. Besides Article 5(2) of Directive 2012/29/EU outlines the

\textsuperscript{74} CJEU C-503/13, C-504/13, Boston Scientific Medizintechnik, para 38.
requirement that victims who do not understand or speak the language of the competent authority be enabled to make the complaint with regard to a criminal offence in a language that they understand or by receiving the necessary linguistic assistance. While Article 5(3) states a necessity to provide an opportunity for a victim to receive the translation, free of charge, of the written acknowledgement of its complaint by request, in a language the victim understands.

Hence, Directive 2012/29/EU only specifies the ability for a victim to submit a claim and to receive acknowledgement of it in a language it understands, but for the rest of activities to receive the translation in a language which understands both, the victim and an authority. What is more, Article 7(1) narrows the scope of an obligation to provide translation to interviews, questioning of the victim during criminal proceedings before investigative and judicial authorities, including during police questioning, and interpretation for their active participation in court hearings and any necessary interim hearings. Thus, the right to translation comprises the gathering of information from the victim, in turn interpretation - active participation of the victim in hearings. Besides, the Directive does not specify the term ‘active participation’. Recital 20 leaves discretion to the Member States. For instance, in preliminary case C-603/19 the term ‘active participation’ by the General Prosecution Office of Slovakia is attributed to the minimum standards on the rights of victims, support and protection of victims of crime, ability to participate in criminal proceedings and the right to secure compensation for damage in criminal proceedings. In turn, the Republic of Latvia pertain the term ‘active participation’ to the party of a criminal proceeding that includes not only right to testify, to request for investigation, to participate in the review of the case, express view including the identification, examination, evidence, participate in court hearing and otherwise actively contribute to the evidentiary process. On the contrary, for example, Canada perceives the term not only as passive status but rather as a constant engagement in criminal procedural activities. Thus, ‘active participation’ at least shall mean the ability to execute rights availed by the individual procedural status since Recital 22 states that authorities may also ex officio initiated criminal proceedings by a victim.

---

76 CJEU C-603/19, Úrad špeciáльnej prokuratúry. Request for a preliminary ruling, question No.1.
Furthermore, Article 7(3), 7(4), 7(5) and 7(6) narrows the written or oral translation to the time and place of the trial, information essential to the exercise of the rights in criminal proceedings and shall include at least any decision ending and upon the request by the victim also reasons or a summary of them, except in the case of a jury decision or a decision where the reasons are confidential upon the national law. According to Recital 30 of Directive 2012/29/EU, the term ‘decision’ shall be understood only as a reference to the finding of a guilt or otherwise ending criminal proceedings. While Recital 32 allows to provide also information about release or the escape of the offender where there might be a danger or an identified risk of harm to the victims unless there is an identified risk of harm to the offender.

2.3. Specific provisions of Directive 2012/29/EU regarding translation

Recital 15 of Directive 2012/29/EU in conjunction with Recital 21, 56, 66, Article 2(3) and 22(3) stipulate that in order to identify the needs for a victim, personal characteristics including disability of a victim must be taken into account as may affect the ability to understand or to be understood. To prevent difficulties in understanding or to communicate due to, for instance, hearing or speech impediments equal access to rights stated in Directive 2012/29/EU including access to information in criminal proceedings for victims with disabilities should be provided.

What stems from the mentioned is not only the physical access to the premises but also access to information what is a right rather than a privilege deriving also from Article 9 of the Convention on the Rights of Persons with Disabilities and should be provided to prevent both possible direct discrimination as well as the feasible indirect discrimination for people associated with the disabled. To achieve the aim, information should be adjusted in the form stated in Web Content Accessibility Guidelines 2.0 issued upon W3C Recommendation and approved as an ISO standard: ISO/IEC 40500:2012. ISO/IEC 40500, last reviewed and confirmed in 2019.

---

79 Brooke S. Digital access is a right not a privilege, Raconteur, 8 September 2016. - https://www.raconteur.net/technology/digital-access-is-a-right-and-not-a-privilege (02.01.2020.).
In addition, User Agent Accessibility Guidelines\textsuperscript{84}, Authoring Tool Accessibility Guidelines 2.0\textsuperscript{85}, Directive (EU) 2016/2102\textsuperscript{86} and European accessibility act\textsuperscript{87} sets requirements that must be fulfilled to make available access to web content, web browser, media players, assistive technology as screen readers, alternative keyboards and others\textsuperscript{88} using not only a computer but also mobile applications and other smart devices. Besides, technology should be adjusted in a way to provide also access to alternative content to prevent trigger of post-traumatic stress disorder that may be caused by image or contrast, speech recognition and other features of the information displayed by technology and is of the utmost importance for victims in criminal proceedings who are particularly vulnerable and requires special needs to prevent repeated victimization as derives from Recital 54,55, 58, Article 22(1) of Directive 2012/29/EU.

The mentioned provisions are with the utmost importance also regarding automated or machine translation, interpretation system that also should be adjusted regarding the specific needs of a victim. Besides, high costs, limited financial resources do not justify the violation of human rights.\textsuperscript{89} What is more, Article 17 of Directive 2012/29/EU also requires that for victims that reside in a Member State of the Directive 2012/29/EU other than that where the criminal offence was committed to ensure that the statement, complaint from the victim, is received remotely, if necessary using also video conferencing, telephone conferencing calls upon the Convention on Mutual Assistance.\textsuperscript{90} That means that also automated translation or interpretation system should be adjusted to provide an opportunity to fulfil the stated rights remotely.

To summarize, although Directive 2012/29/EU outlines specific rights and requirements attributed to victims in criminal proceedings, considerable discretion is left for the Member States to determine. However, Recital 66 of Directive 2012/29/EU stipulates respecting the fundamental rights and observes the principles recognized by the Charter as the right to a fair trial, nondiscrimination for the elderly that is of particular importance considering implementation of automated translation technology and to adjust it to for the specific audience. Thus, automated translation technology must address human rights, including the right to a fair trial.

\textsuperscript{84} User Agent Accessibility Guidelines (UAAG) 2.0. UAAG Reference, Explanation, Examples, and Resources for User Agent Accessibility Guidelines 2.0 - https://www.w3.org/TR/UAAG20-Reference/ (02.01.2020.).
\textsuperscript{85}Authoring Tool Accessibility Guidelines (ATAG) 2.0. - https://www.w3.org/TR/ATAG20/ (02.01.2020.).
\textsuperscript{87} European Accessibility act. - https://ec.europa.eu/social/main.jsp?catId=1202 (02.01.2020.).
\textsuperscript{88} Essential Components of Web Accessibility. - https://www.w3.org/WAI/fundamentals/components/ (02.01.2020.).
\textsuperscript{89} ECtHR No. 56027/10, Reshetnyak v. Russia, para.70; ECtHR, No. 20075/03, Shilbergs v. Russia, para 71-79.
\textsuperscript{90} Convention established by the Council in accordance with Article 34 of the Treaty on European Union, on Mutual Assistance in Criminal Matters between the Member States of the European Union - Council Declaration on Article 10(9) - Declaration by the United Kingdom on Article 20. - OJ C 197, 12.7.2000, p. 3–23.
2.4. The right to a fair trial

Article 6(3)(a) of the ECHR stipulates that under rights to a fair trial person who is charged with the criminal offence has rights to be informed promptly, in detail of the nature, cause of the accusation in the understandable language. According to Article 6(3)(e) it also has the right to free assistance of an interpreter if cannot understand or speak the language used in a court. Similar rights also outline Article 14(3) and 14(3)(f) of the International Covenant on Civil and Political Rights (hereinafter - ICCPR). Under rights to interpreter should be understood the transfer of oral content from one language to another, whereas the right to translation comprises written content. In turn, the Human Rights Committee associates the right to the interpreter with the right to defence. The differences exist due to the expressis verbis missing right to translation under ECHR and ICCPR. Since to execute the right to a fair trial a defendant needs to understand oral, written evidence, but it is not absolute if the relevant materials may be linguistically presented; right to high standard interpreter under Article 6 of ECHR also applies to the essential documentary materials also in the pre-trial phase, free of charge, in language which the person adequately understands and that should be proved by the competent authorities and not by the defendant. Namely, the right to the interpreter is said as being more as of a necessity rather the preference. Besides, the right to an interpreter may also extend to any non-conversant defence witnesses. As no ‘high standard interpretation’ definition exists, the fairness of the proceeding including the competence of an interpreter must be safeguarded by the judge to assure that the defendant could not only understand the activities in hearing but also to participate in own defence. No consensus exists as to the right to waive the right to an interpreter.

---

96 ECtHR No. 9783/82, Kamasinski v. Austria, para 72-74; ECtHR No. 6210/73, 6877/75, 7132/75, Luedicke, Belkacem and Koç v. Germany, para 48; ECtHR No. 18114/02, Hermi v. Italy, para 69.
97 ECtHR No. 10964/84, Brozicek v. Italy, para 23.
100 Canada Supreme Court No.23321, R v. Tran, para 61-62.
101 ECtHR No. 3277/96, Cuscani v. the United Kingdom, para 38.
102 ECtHR No. 9783/82, Kamasinski v. Austria, para 72-74.
Furthermore, Article 5(2) of ECHR addresses the right to information in understandable language for the charged or arrested since the initiation of charge and remains until the end of the proceedings also for the appeal. In turn, the Charter does not *expressis verbis* outline the right to receive information in the understandable language; nonetheless, Article 47, 48, 6 guarantees the right to a fair trial, legal advice, representation, defence to be provided as with the ECHR.

Besides, Directive 2010/64/EU addresses rights to oral, written translation, interpretation for the execution of the European Arrest Warrant for the suspended and accused persons; other procedural statuses are not comprised. Article 2 of Directive outlines right to interpretation also remotely during the investigative, judicial phases including hearings, appeals, communication between the person and its lawyer ensuring also service for the persons with disabilities. The national law should outline the criteria to determine the language understanding.

Furthermore, Article 3 guarantees rights to the translation of essential documents including minor offences, key witness statements, other reasonably requested; the right to waive translation rights only upon prior legal advice, voluntarily. The quality of the interpretation should provide the ability to acquire the knowledge of the case and to exercise the right to own defence. Besides, Article 5(2) requests the establishment of registers of the independent, qualified translators and to ensure confidentiality. Moreover, according to Article 6, judges, prosecutors and judicial staff involved in criminal proceedings shall oversight the quality of interpretation. Currently, no need to revise the Directive has been found by the European Commission.

Summarizing, the right to a fair trial depicts additional angles for the right to translation, interpretation, issues with *no consensus*. Hence, the national law should concern all the aspects of implementing automated translation, especially the quality issue but cannot imperil the execution of the EU law including to deprive the automatic translation system of any effectiveness.

104 Ibid, p. 17.
105 CJEU C-175/17, *Belastingdienst/Toeslagen*, para 35.
109 CJEU C-278/16 Sleutjes, para 34.
111 CJEU C-216/14, *Covaci*, para 49, 51.
3. CURRENT LEGAL REGULATION IN ESTONIA AND POSSIBILITY OF USING AUTOMATED AND MACHINE-ASSISTED TRANSLATION AND INTERPRETATION

3.1. General provisions on translation and interpretation

Criminal proceedings in Estonia are regulated with the Code of Criminal Procedure (hereinafter *Code*). As stated in § 1, this Code provides the rules for pre-court and court procedure concerning criminal offences and the rules concerning enforcement of decisions made in criminal matters.\(^{114}\) This means that also, the rules for translation and interpretation are regulated by the Code. It is important to point out that the general provisions concerning translation and interpretation have to adhere to European Union law, international law (like the European Charter of Human Rights), which have been analysed in the previous chapter, and in the case of an arrest or detention also to The Constitution of the Republic of Estonia, the latter stating that everyone who has been deprived of his or her liberty must be informed promptly, in a language and manner which he or she understands, of the reason for the deprivation of liberty and of his or her rights, and be given an opportunity to notify those closest to him or her.\(^{115}\)

Paragraph 10 of the Code regulates the language of criminal proceedings. Subsection 1 states that as a rule, the language of criminal proceedings is Estonian. With the consent of the body conducting criminal proceedings, of participants in proceedings and of the parties to judicial proceedings, the criminal proceedings may be conducted in another language if the body, participants and parties are proficient in that language.\(^{116}\) Although it has to be taken into account that legal scholars have found this problematic or even unconstitutional.\(^{117}\) It is not clear from these provisions if the language of criminal proceedings may vary in different stages of the proceedings or not. If you take into consideration § 144 subsection 1, which states that all procedural documents shall be prepared in Estonian language or at the very least a translation into Estonian shall be appended thereto, it has to be understood that only oral communication in the proceedings may be conducted in another language.\(^{118}\)

Going further § 10 subsection 2 states that suspects, accused, victims, civil defendants and third persons who are not proficient in the Estonian language shall be ensured the assistance of an

---
\(^{117}\) In the Constitution of the Republic of Estonia § 6 states that the States official language is Estonian and § 52 states the same about government agencies.
interpreter or translator. In the case of doubt, proficiency shall be ascertained by the body conducting proceedings. If it is impossible to ascertain proficiency or the individual’s command of Estonian proves to be insufficient, the assistance of an interpreter or translator shall be ensured. This does not only apply to oral proceedings but also written documents. After that, the Code differentiates somewhat between the rights of the suspect and accused and other subjects stated in the aforementioned provision.

If a suspect or accused is not proficient in the Estonian language, the text of the report on detention of the suspect, arrest warrant, European arrest warrant, statement of charges and judgment shall be translated into his or her native language or a language in which he or she is proficient, at least to the extent which is significant from the point of view of understanding the content of the suspicion or charges or for ensuring fairness of the proceedings.\textsuperscript{119} If an interpreter or translator does not participate in a procedural act where the participation of an interpreter or translator is mandatory, the act is null and void\textsuperscript{120}. Thereby it has to be taken into account that if a suspect or accused is not provided with a proper translation or an interpreter, it may be grounds for dismissal of the case. The Supreme Court has said that the interpreter is effectively guaranteed, firstly, if the interpreter takes part in the procedure done in the presence of the accused and makes the processes understandable to the suspect, and secondly by providing the suspect with essential procedural documents translated into his or her native language or language in which he or she is proficient in.\textsuperscript{121} Thereby it is clear that an interpreter has to be present according to the current legal regulation. But in the same case the Court also said that pursuant to the provisions of § 339 (1) 9) of the Code, only a court hearing held in a criminal proceeding without an interpreter in a language which the accused does not speak is an absolute violation of criminal procedural law. However, in the case of pre-court proceedings, the absence of an interpreter can only be regarded as a significant violation of criminal procedural law if it can be identified that it resulted in an unfounded or illegal judgement.\textsuperscript{122} It also has to be noted that an oral interpretation shall be ensured to a suspect and accused immediately, a written translation of the documents shall be ensured to a suspect and accused within a reasonable period of time so that this does not impair the exercise of their rights of defence.\textsuperscript{123}

\textsuperscript{120} Ibid.
\textsuperscript{121} RKKKm 3-1-1-58-10, p 9.
\textsuperscript{122} Ibid, p 12.
Further rules of translation for a victim of a crime are stated in §10 subsection 6\(^1\) of the Code. The victim has the right to request a translation of the text which is essential for understanding the substance of the order on termination of criminal proceedings or the court judgment or for ensuring the fairness of the proceedings.\(^{124}\) But the same subsection also states that if the body conducting proceedings finds that the request for translating other documents is not justified, such body shall formalise the refusal by an order. Thereby there is no absolute right for the victim to demand a translation if the body conducting proceedings finds that this is not needed. Of course, this order can be appealed in court.

Hence, the general rule is that all subjects in criminal proceedings must be able to understand the proceedings, but the rules are stricter for suspects and accused. As seen from the Code it is not important that the translation or interpretation is done into the native language of the subject but into a language that the subject is proficient in. On the other hand, this raises the question of proficiency. There are no regulations in the Code about determining the proficiency in a language. It has to be derived that proficiency is left for the subject and the body conducting proceedings to decide – if he or she understands the proceedings or documents and agrees to use that language.

### 3.2. Special provisions for the translator or interpreter

In addition to the general provisions, there are some special provisions regarding translators and interpreters. An interpreter or translator is a person proficient in language for specific purposes or a person interpreting for a deaf or dumb person. Other subjects of criminal proceedings shall not perform the duties of an interpreter or translator.\(^{125}\) As can be seen from the previous, when considering a technical solution, you would have to keep in mind that the current legislation clearly states that the translator or interpreter has to be a person thereby we have to rule out fully automated solutions immediately. The technical possibilities have been analysed in previous chapters. The next subsection of the Code, even more, stresses this part when stating that an interpreter or translator to whom the oath of interpreters and translators has not been administered shall be warned that he or she may be punished pursuant to criminal procedure for a knowingly false interpretation or translation.\(^{126}\) As only a natural person (for

---

\(^{124}\) *Kriminaalmenetluse seadustik [Code of Criminal Procedure]*. - RT I, 20.12.2019, 8, §10 (6)\(^1\).

\(^{125}\) Ibid, §161 (2).

\(^{126}\) Ibid. §161 (3).
some offences also a legal person) can be subject of criminal proceedings; it follows that a machine on itself could not be seen as a translator or interpreter under current law.

Going further the Code states in order to ensure the correctness of interpretation or translation, an interpreter or translator has the right to pose questions to participants in the proceedings, examine the minutes of procedural acts and make statements concerning the report, and such statements shall be recorded in the minutes.\textsuperscript{127} This means that the interpreter or translator has a very important role in the proceedings. He or she even has a right to pose independent questions, which an autonomous machine at the current level of technology would not be able to do. Moreover, the Code obligates that an interpretation or translation of any aspect of a procedural act rendered by an interpreter or translator shall be precise and complete. If a non-staff interpreter or translator is not sufficiently proficient in language for specific purposes or in the form of expression of a deaf or mute person, he or she is required to refuse to participate in the criminal proceedings. Here it has to be noted that in Estonia, the government agencies conducting criminal proceedings employ translators and interpreters for languages that are needed the most.

The Supreme Court has found that it is clear from those aforementioned provisions that an interpreter has to be involved in the procedural act as soon as it commences and not at a later stage. This is necessary to enable a person who does not speak the language of the proceedings to participate in the act on an equal basis with a person who is proficient in the language.\textsuperscript{128} It becomes clear that translation and interpretation are immensely important in conduction criminal proceedings, and all technological advancements have to take that into account. The body conducting proceedings and the translator or interpreter have to be convinced that proficiency is guaranteed.

\textbf{3.3. Challenges with providing translation and interpretation in Estonia}

The first challenge that is encountered in providing translation or interpretation is the question about suitable people for the task at hand. This is a bigger problem with interpretation as with translation because as stated in the previous subchapter an interpreter has to be involved in the procedural act as soon as it commences and not at a later stage. Documents, however, can be translated at a slightly later time. Given the fact that procedural acts are conducted at all times of

\textsuperscript{128} RKKKm 3-1-1-157-05, p 10.
day and that the places where the procedural acts are not always located near the bigger cities, this provides a problem for immediate interpretation. Not all subjects can be detained at the scene for allowing enough time for providing these services. Although the law does not prohibit video translation. Meaning this is one possibility to consider in those cases. If the interpreter or translator is not sufficiently proficient in language for specific purposes or in the form of expression, he or she is required to refuse to participate in the criminal proceedings. If the interpreter or translator is not employed by the agency conducting the proceedings, it is not always possible to determine their level of proficiency until they arrive at the site and see the case or proceedings directly. This also raises the question about the quality of the service. The body conducting the proceedings has to make sure that the quality of all translations or interpretations is evenly good. If not it has to remove an interpreter or translator if he or she does not perform his or her duties as required or if the quality of the interpretation or translation may impair the exercise of the right of defence of the suspect or accused.\textsuperscript{129} The problem is that the search for suitable translators or interpreters might make the proceedings longer and ineffective. In addition to that, the accused have a right to a fair and public hearing within a reasonable time, which could be hindered if he or she does not understand the proceedings and a proficient translator or interpreter can be found.

The second challenge is the cost of providing these services. There are around 100 translators and interpreters employed by government agencies conducting criminal proceedings\textsuperscript{130}, which in itself is costly. But as the world globalises more and more the languages that need to be translated or interpreted vary so much that it is impossible to provide the necessary people to provide the services to all those that need it. Estonia has to take into account that for exotic languages, they do not have the necessary lecturers in the universities and people would have to study abroad. Therefore, the costs also include fees of outside experts who are needed for translation and study costs. Some translations are done via intermediate languages, meaning an exotic language might be translated into English before translating it into Estonian. This proves rather difficult or even impossible in interpretation as it has to be immediate. Meaning it could be needed to fly an interpreter into the country from another state. This takes time and resources. But also real-time video interpretations are used, which is not prohibited by law. With video interpretations, there can be problems with the quality of the broadcasting service, so using video interpretations in high profile cases cannot be recommended. Sworn

\textsuperscript{130} Estimated number from the contact lists of the agencies available online.
translators or translation businesses can be used as outside service providers. This is mostly possible when translating documents, but proficient outside service providers can also be used as interpreters.

3.4. Possibility of using automated and machine-assisted translation and interpretation considering current Estonian regulation

To tackle the problems related to providing translation and interpretation in criminal proceedings, the help of machine-assisted and automated translation tools have been and are continuously being developed. The current technological possibilities and approaches are analysed in the first chapter of this work.

As already shortly pointed out in previous sub-sections, the current Estonian legislation regarding translators and interpreters is a strictly natural person (human) based. Which means that the technological possibilities analysed in this work are only strictly possible with human oversight. The accountability lies on the person providing the necessary service. Therefore, amendments in the law would be necessary to incorporate more complex automated translation on interpretation tools. Also, the question regarding the liability of an incorrect translation must be addressed as well as under current regulations an incorrect translation or interpretation might make the person accountable punished pursuant to criminal procedure. Furthermore, malfunction of automated or machine translation could infringe not only the Code but also human rights, leading to infringement procedure in the Estonian court system or even the Court of Justice of the European Union. This all means that a back-up solution has to be present if it would be possible to use fully automated translation or interpretation. Another problem is data processing. There have to be special regulations in the law, specifically about processing, access and storage terms as all the data in criminal proceedings have to be protected and handled confidentially.

This does not mean that translation tools may not be used at all. When translating documents, different text-to-text tools are probably already being used by almost every professional translator. This is not different in the field of criminal proceedings. As demonstrated in the previous chapters, even using only text-to-text tools may constitute a significant cost saving in the long term. At the same time, it has to be considered that Estonian is a small language spoken by very few people, so more advanced tools might not be available for the needed purpose. Therefore, Estonia would have to develop separate text corpuses to be used with other languages. For the STS tools, special phonetic corpuses would also have to be compiled.
Even if the tools would allow applying these corpuses to already existing possibilities, it would still prove costly.

It has to be noted at this point that Estonia has a special programme for supporting Estonian language technology since 2006. A new program was approved in 2018 for the years 2018-2027. Through the program, the state supports a field where it is not always profitable for the private sector to take on the risks associated with the development of technology for a language with a small number of speakers. The objective is the development of basic technologies and resources that have not yet been implemented as part of a system and result in a comprehensive software module or systematised language resource, and the introduction of language technology involving the application of technology as part of a system or the creation of a stand-alone application, resulting in the operation of language technology as part of a system or the completion of a language-technology focused application. Some previous projects also include building text and phonetic corpuses and improving the translation quality of machine translation. This means the state, together with businesses and universities are continuing to improve the tools which could also be used in criminal proceedings. Still, you would have to take into account that the proceedings would need specialised wording and not everyday speech.

3.5. Amending the law

The specific amendments to Estonian law depend on what kind of technologies the state would like to implement for use. In the case of translating documents or text-to-text, the law does not need any amendments when human oversight is guaranteed, and the translation is checked and confirmed (signed) by a translator. As shown before some text-to-text tools are already being used. There the accountability would still lie on the human translator who provides necessary oversight and can check for errors. However, if the state would like to implement a fully automated machine translation system of text-to-text translation with no human oversight, the law would have to be amended. As shown in the previous chapter, the national law has to adhere to EU law in these aspects, meaning all amendments to national law have to be allowed by EU directives.

---


132 Website of the National Programme for Estonian Language Technology - https://www.keeletehnoloogia.ee/en (09.01.2020.).
The general regulations in § 10 of the Code would not have to be amended for use with fully automated systems. The problem is with the specific regulations in paragraphs 161 and 162. It would be recommended to draft a special law or regulation in order to use the systems where all the aspects shown in sub-chapter 3.4. should be regulated. An option to regulate some of it in the Code would be adding two subsections to § 161 of the Code as follows:

(2)¹ A fully automated electronic system may be used for translation of documents. The body conducting the proceedings shall guarantee the proficiency and quality of the documents translated by the electronic system. Subsections (3) – (6) of the current paragraph and § 162 do not apply to the use of this electronic system.

(2)² The suspect, accused, counsel, victim or civil defendant may submit a petition of challenge against the use of the system or the outcome of the translation referred to in subsection (2)¹ or the outcome of the translation of the current paragraph.

This, however, does not deal with all the problems listed above.

When dealing with speech-to-speech solutions, no easy amendments can be suggested. As the authors have pointed out the technology is not yet present to guarantee high-quality solutions, which is needed in criminal proceedings. If the technology evolves a new system of law would have to be created in order to allow speech-to-speech interpretation.

To sum it up, it would have to be said that the regulation problems concerning fully automated translation or interpretation are highly complex issues which would need a lot of time to be solved and the co-operation of lawmakers of the entire European Union.
SUMMARY

To attain the objective of the IT Law Lab Task, namely, to analyse the different approaches of machine translation and interpretation, the state of current technology and the compatibility of those technologies with the European Union, international and Estonian law of criminal proceedings, the technical possibilities and approaches were analysed to conclude whether the current technical capabilities are suitable and beneficial for use in sophisticated translations.

The research outlined that Neural MT has yielded outstanding results and improved state-of-the-art results for several language pairs. Neural MT is considered the current state-of-the-art, as most of the best MT systems use neural networks. Quality of neural MT depends on training data size and quality, as well as the match between the training data context and the input context.

Additionally, it is worth to outline that the state-of-the-art of neural MT may differ significantly between the language pairs with a large amount of parallel data and those with the insufficient amount. Overall, it can be concluded that the current state-of-the-art text-to-text MT cannot yet be considered as having reached the human parity. Furthermore, the direct sequence-to-sequence speech-to-speech MT systems are currently in the development phase. Hence, none of the state-of-the-art MT can entirely and autonomously substitute the human interpreter in criminal proceedings.

Nonetheless, human post-editing of MT results could lead to significant cost-savings while not decreasing quality; therefore, MT with human-post editing could be considered to decrease translation costs within the criminal proceeding system. It is worth to outline, that various state-of-the-art neural MT system model implementations are widely available at GitHub repository under open-source licences which often allow modification and distribution of modified code. Therefore, prima facie it would be possible to use available open-source models to implement and train the open-source neural MT system for use in criminal proceedings.

What is more, in order to mitigate the training data size and context problem for Estonian as a small language, input and human translated output anonymised texts from the criminal proceedings could be used to create the future parallel corpus.

In spite of the technical capabilities, an overview of the European Union and international law regarding criminal proceedings outlines that taking into account the necessity to ensure execution of human rights, especially the right to a fair trial, to dignity, integrity, liberty, the principle of nondiscrimination, equality between women and men, the rights of the child, the elderly and persons with disabilities as well as considering that protection of individuals should
be technologically neutral and not depend on the techniques used to prevent creating a severe risk of circumvention or lacking transparency, ambiguity, currently the most suitable form of automated translation or interpretation service resulting from artificial intelligence from the perspective of compliance with the international law could rather be a machine that involves a human oversight than entirely autonomous system.

What is more, the European Union law addresses specific minimal standards that differ amongst the criminal procedural statuses and, simultaneously, have also to comply with the fundamental rights, especially with the right to a fair trial equally for all including the needs for the people with disabilities, child and the persons with the granted procedural status but residing outside the particular country. Addressing all the aspects could be considerably cumbersome.

Moreover, albeit the considerable discretion of the provision of rights to translation and interpretation is left for the states, currently no necessity to revise the EU law has been found as the existing regulation comprises the essential minimal legal and technical aspects of the platform in criminal proceedings and also due to the prohibition to jeopardise the execution of the EU law including deprivation of effectiveness of the automatic translation system by the national law.

It is clear that to ensure a fair process, an interpreter has to be involved in a procedural act as soon as it commences. This is necessary to enable a person who does not speak the language of the proceedings to participate in the act on an equal basis with a person who is proficient in the language. The technology could help in achieving equal basis for subjects who do not speak Estonian.

This leads to different challenges when providing these services. Namely, ensuring the equal quality of the translations or interpretations, the availability of proficient service providers and the rising cost due to globalisation.

The current Estonian legislation regarding translators and interpreters is strictly natural person (human) based. Which means that the technological possibilities analysed in this work are only a strictly possible with a human oversight at this time.

The issues handled in this analysis would surely need further and more in-depth investigation. The technology could help all the parties involved in the proceedings. In the case of Estonia, the smallness of the state and the language has to be considered when implementing technological solutions. Even more, providing fully automated translation or interpretation comes with highly complex legal issues which would need much time to be solved and the co-operation of lawmakers of the entire European Union.
OVERVIEW OF THE CONTRIBUTION

Initial steps regarding the incorporation of machine translation and interpretation in criminal proceedings have been analysed by the group members in the form of project blueprint within the study course ‘Public E-Services’. The project also availed exchange of ideas between the group members for the potentially problematic aspects and the questions to analyse within the framework of the Master’s Exam Part I or IT Law Task as well as the suggestions for the future.

Concrete suggestions for the content of the IT Law Lab Task Solution were also discussed amongst the group members as well as divided tasks, a number of pages and desired outcome.

Arina performed in-depth research of the relevant scientific publications and provided the analysis of extensive body of neural machine translation research for the question included within Chapter I, namely, the state-of-the-art of machine translation approaches. Her views are reflected not only in the chapter, but also she outlined the thesis of the relevant conclusions for the summary part. In addition, Arina also supplemented the reference list with the sources referenced in the mentioned chapter and revised the summary.

Liva profoundly analysed the legal basis upon the European Union and international law for the automated or machine translation and interpretation included in Chapter II. She also drew conclusions regarding the chapter for the summary as well as compiled the overall summary. Additionally, Liva wrote the title page, table of contents, an overview of the contribution, revised and supplemented introduction, compiled the overall work incorporating the created chapters, unifying the formatting, adding page numbering and supplementing with the other necessary parts. Liva also created and supplemented the reference list with sources used in the chapter as well as revised the compliance of the reference list, footnotes with the formatting requirements.

Marko researched the question outlined in Chapter III, namely, current legal regulation in Estonia and possibility of using automated and machine-assisted translation and interpretation. Besides, Marko also contacted the consulting authority addressing the questions from the group. In addition, he wrote an introduction, revised, supplemented an overall formatting, spelling including final spellcheck, outlined the conclusions for the summary, added the referenced sources from Chapter III in the reference list as well as followed to the Guidelines of Writing and Formatting Student Works to overview inclusion of the necessary parts of the structure of the work. Additionally, Marko uploaded the IT Law Lab Task Solution in the moodle system.
REFERENCE LIST

General:
7) Brooke S. Digital access is a right not a privilege, Raconteur, 8 September 2016. - https://www.raconteur.net/technology/digital-access-is-a-right-and-not-a-privilege.


Legal Acts:


10) Convention established by the Council in accordance with Article 34 of the Treaty on European Union, on Mutual Assistance in Criminal Matters between the Member States of the European Union - Council Declaration on Article 10(9) - Declaration by the United Kingdom on Article 20.- OJ C 197, 12.7.2000, p. 3–23.


22) Authoring Tool Accessibility Guidelines (ATAG) 2.0. - https://www.w3.org/TR/ATAG20/.


24) Essential Components of Web Accessibility. - https://www.w3.org/WAI/fundamentals/components/.


Practice:
Expert:

Judicial:

Legal:

Case Law:
1) CJEU C-38/18, Gambino and Hyka.
2) CJEU C-105/03, Pupino.
3) CJEU C-404/07, Katz.
4) CJEU, C-131/12, Google Spain and Google.
5) CJEU, C-507/17, Google v. CNIL.
6) CJEU C-582/14, Patrick Breyer v. Bundesrepublik Deutschland.
7) CJEU C-503/13, C-504/13, Boston Scientific Medizintechnik.
8) CJEU C-603/19, Úrad špeciálnej prokuratúry, Request for a preliminary ruling, question No.1.
9) CJEU, C-303/06, Coleman v. Attridge Law.
10) CJEU C-175/17, Belastingdienst/Toeslagen.
11) CJEU C-278/16 Sleutjes.
12) CJEU C-216/14, Covaci.
13) CJEU C-25/15, Balogh.
14) ECtHR 30562/04, 30566/04, S and Marper v. the United Kingdom.
15) ECtHR 44787/98 P.G. and J.H. v. the United Kingdom.
16) ECtHR No. 56027/10, Reshetnyak v. Russia.
17) ECtHR, No. 20075/03, Shilbergs v. Russia.
18) ECtHR No. 9783/82, Kamasinski v. Austria.
19) ECtHR No. 6210/73, 6877/75, 7132/75, Luedicke, Belkacem and Koç v. Germany.
20) ECtHR No. 18114/02, Hermi v. Italy.
21) ECtHR No. 10964/84, Brozicek v. Italy.
22) ECtHR No. 3277/96, Cuscani v. the United Kingdom.
26) ICTY, Prosecutor v. Zejnil Delalić.
27) Canada Supreme Court No.23321, R v. Tran.
28) RKKKm 3-1-1-58-10.
29) RKKKm 3-1-1-157-05.