UTILIZATION OF MACHINE TRANSLATION IN IP PRACTICE,
HOW WILL MACHINE TRANSLATION AFFECT IP PRACTICE?

Shigeru Inaba

Japanese patent attorney

UNH Law-JPAA Joint IP Seminar

Friday, September 7, 2018

OUTLINE

- Background

- Introduction of MT (machine translation) services

- Security issue when using free online MT services

- Evaluation of Quality of NMT (neural machine translation)

- Drawbacks of NMT

- Practical use of MT in IP practice

- How MT will affect IP practice

TECHNOLOGY HAS IMPROVED PRODUCTIVITY IN IP FIRM IN THE LAST 30 YEARS

- Handwriting/typing → Word processor
- Filing by mail → Online filing
- Manual search → Online search
- Telex/Fax → E-mail
- Manual docket control → Software

IP firms have become more efficient due to the use of technology and many of tasks that used to be done by humans are being replaced by machines.

WHAT'S NEXT?

- Human translation ➡️ Machine translation?

- Drafting and prosecution ➡️ AI??

Debate hosted by CIPA on the proposition that "This house believes it is inevitable that, within 25 years, a patent will be filed and granted without human intervention".

http://www.cipa.org.uk/policy-and-news/latest-news/cipa-artificial-intelligence-debate-videos/

DEVELOPMENT OF MACHINE TRANSLATION

- Rule-Based Machine Translation (RBMT) came about in 1970s

- Statistical Machine Translation (SMT) came about in the late 80s and early 90s

- Phrase-Based SMT (PBSMT) came about around 2003
  - ✓Still practically available

- Neural Machine Translation (NMT) came about around 2014
  - ✓NMT has dramatically improved the quality of machine translation.

MACHINE TRANSLATION -PROS AND CONS-

Pros

· **Cheap**

✓Substantially no cost if free online MT tool is used

· **Fast**

✓Turnaround time: a few to several seconds for translating an entire patent specification

Cons

· Poor Quality

· Potential Costs Associated with Poor Quality

✓**With the advent of NMT, quality has been improved and is constantly improving!**

NEURAL MACHINE TRANSLATION
-A POWERFUL TOOL-

"Neural machine translation is an emerging technology. It is based on **huge neural network models that "learn" from previously translated sentences.** The specificity of neural machine translation (compared to previous "phrase based" statistical methods) is that **it produces more natural word order, with particular improvements seen in so-called distant language pairs, like Japanese-English or Chinese-English......** (an excerpt from WIPO website)."

A LARGE SET OF HIGH-QUALITY TRAINING DATA AVAILABLE

- The more foreign patent filing, the more training data.

- "The high level of accuracy of the Chinese-English translation is the result of the training of the neural machine translation tool, which **compared 60 million sentences from Chinese patent documents** provided to WIPO's PATENTSCOPE database by the SIPO **with their translations as filed at the USPTO**(an excerpt from WIPO website)."

- "In co-operation with the national patent offices in its member states and with other key partner offices, **the EPO has provided Google with millions of official, human-translated patent documents.** These were used to train the translation engine to handle technical subject-matter and the specific style and format used for patent documents. Every year, **millions of documents** are added to the EPO's databases. These new documents too are fed into the system, continually improving the engine(an excerpt from EPO website)."

MUST-HAVE PRODUCTIVITY TOOL

- "Patent translation not only supports innovation by providing access to technical information, but it also represents a massive potential leap in productivity and efficiency. **Back in 2013, one study we carried out led us to estimate that it would take 16,000 man-years to translate the Chinese patent documentation available at the time into English**. Patent Translate provides all that documentation instantly – and the additional documents published since then – in English, French and German"

    https://blog.epo.org/patent-information-2/new-benchmark-patent-translation/

- It should be noted that the number of Chinese applications has been increasing and more than **10,000,000** applications (invention + utility model) have been filed since 2013.

GOVERNMENTAL ORGANIZATIONS SUPPORT DEVELOPMENT OF MT

- Recent Development of MT for patent documents has been led mainly by governmental organizations not by private sectors.

- The IP 5 offices (USPTO, EPO, JPO, KIPO, and SIPO) as well as WIPO promote the use of machine translation to remove language barriers from patent documentation.

- MT enables their examiners to access global patent information for prior art searches and such.

- The governmental organizations such as EPO and JPO will continue supporting the development of NMT technology for their patent searching platforms such as Espacenet and Jplatpat as well as examiners.

MT SERVICES PROVIDED BY WIPO

- Patentscope with MT tools including WIPO translate, Google translate, Being translate and Baidu translate.
  https://patentscope.wipo.int/search/ja/search.jsf

- WIPO translate ("domain-aware technique")
  https://patentscope.wipo.int/translate/translate.jsf?interfaceLanguage=ja
  "WIPO Translate is trained exclusively with huge amounts of patent texts and includes a "domain-aware technique" that translates according to the specificity of the invention. The tool internally integrates 32 technical domains taken from the International Patent Classification. ....The technology takes into consideration the specific domain when translating a particular sentence thereby yielding more accurate translations(an excerpt from WIPO website)."

MT SERVICES PROVIDED BY EPO
-PATENT TRANSLATE OPTIMIZED FOR PATENTS-

- Espacenet with Patent Translate
  https://worldwide.espacenet.com/

- Patent Translate is the EPO's machine translation service .

- Patent Translate uses the same translation engine approach as the general Google Translate service and has been specially trained for patent documents.

EPO EXAMINATION GUIDELINES STIPULATE THE USE OF MT

"In order to overcome the language barrier constituted by a document in an unfamiliar non-official language, it might be appropriate for the examiner to rely on a machine translation of said document (see T 991/01), which should be sent to the applicant (see B-X, 9.1.3).... .. (G-IV 4.1)"

MT SERVICES PROVIDED BY JPO
-PATENT SEARCHING PLATFORM (JPLATPAT) WITH MT TOOLS-

- Automatic machine translation of searched Japanese publications to English text (Jplatpat English site)

- Searching non-Japanese patent documents using Japanese key words as well as searching Japanese patent documents using English key words (Jplatpat Japanese site/English site)

- Automatic machine translation of searched Chinese and Korean publications to Japanese text

  http://www.ckgs.jpo.go.jp/gazette_number

  http://www.ckgs.jpo.go.jp/full_text

- Machine translation of file wrapper documents at OPD (One Portal Dossier) site

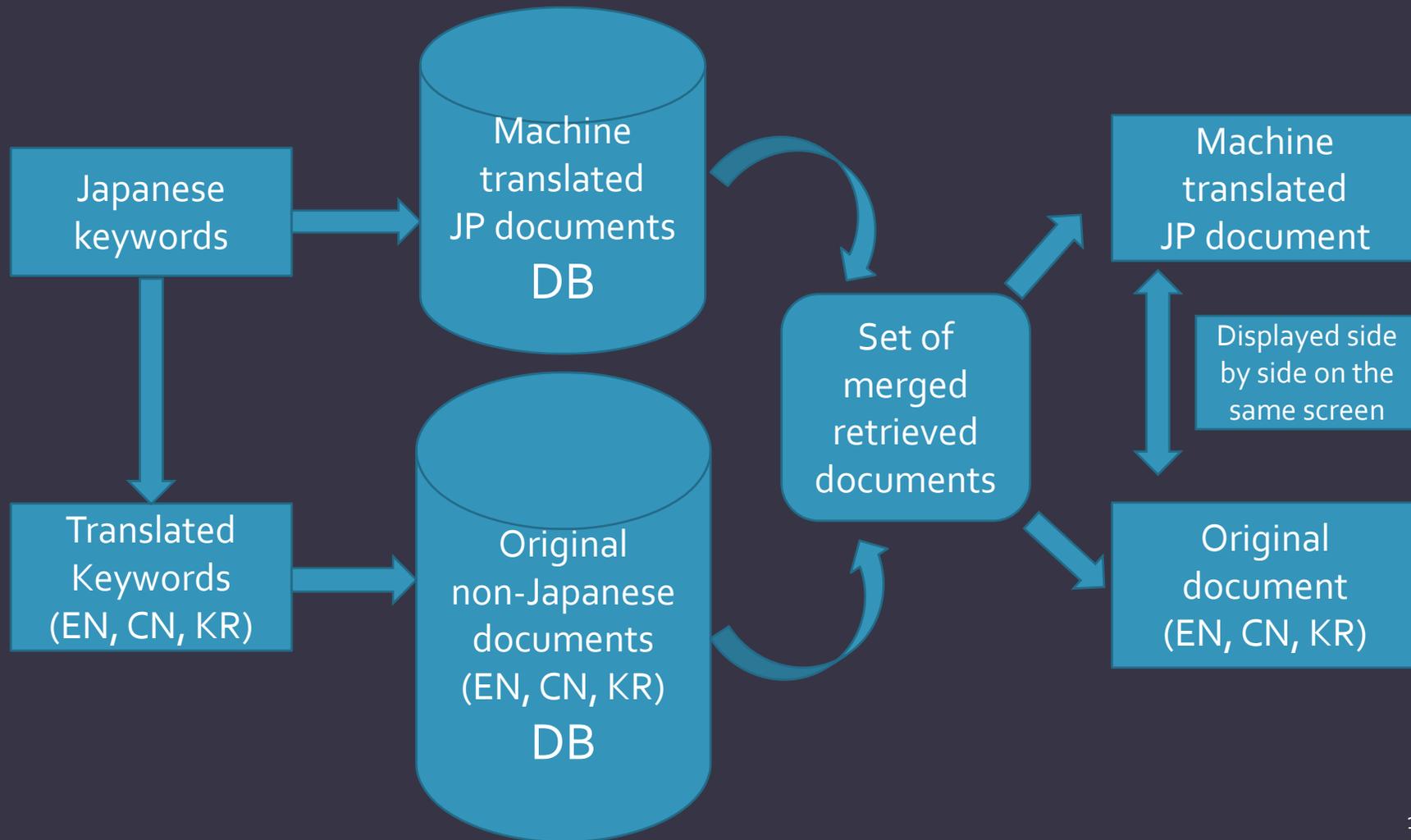 https://www10.j-platpat.inpit.go.jp/pop/all/popd/POPD_GM101_Top.action

USE OF MT BY JPO EXAMINERS
-高度検索システム検証プロトタイプ-
WWW.TOKUGIKON.JP/GIKONSHI/280/280TOKUSYU03.PDF

- JPO Examiners utilize a MT engine developed by NICT(National Institute of Information and Communications Technology).

- Majority of patent documents disclosed in recent years are available only in Chinese or Korean. To ensure high-quality patent examination, it is important to search non-Japanese documents including Chinese and Korean documents as well as English documents.

- The JPO employs a hybrid non-Japanese documents searching system which uses both a first combination of Japanese keywords and machine-translated JP documents and a second combination of machine-translated keywords (EN, CN and KR) and the original non-Japanese (EN, CN and KR) documents.

PATENT SEARCHING SYSTEM IN JPO FOR NON-JAPANESE DOCUMENTS

Japanese keywords

Machine translated JP documents DB

Translated Keywords (EN, CN, KR)

Original non-Japanese documents (EN, CN, KR) DB

Set of merged retrieved documents

Machine translated JP document

Displayed side by side on the same screen

Original document (EN, CN, KR)

A VARIETY OF ONLINE MT SERVICES -FREE SERVICES VS PAID SERVICES-

- Free services
  - Patent searching platforms with MT tools such as
    - ✓ Espacenet with Patent Translate in collaboration with Google
    - ✓ Patentscope with MT tools
    - ✓ Jplatpat with MT tools
  - Browser's built-in translate functionality such as
    - ✓ Chrome's built-in Translate functionality
  - WIPO translate ("domain-aware technique")
  - Google Translate
  - Microsoft Translator
  - Bing Translator
  - Baidu Translate
  - みんなの自動翻訳＠TexTra®by NICT

- Paid services (Commercial on-premises or cloud-based services)

WHO IS A FRONT RUNNER?
-PATENT SEARCHING PLATFORMS WITH MT TOOLS-

➢ **Espacenet with Patent Translate**
✓ Use of Google NMT engine optimized for patent documents

➢ **Jplatpat**
✓ Familiar to the Japanese
✓ NMT system for Jplatpat will be upgraded next year in collaboration with NICT (We will see….).

➢ **Patentscope**
✓ Selection of MT engine (Google Translate, WIPO translate..)
✓ Frequently requires image verification such as "please count the number of fire trucks" in the middle of searching (Frustrating!)

CHROME'S BUILT-IN TRANSLATE FUNCTIONALITY -CONVENIENT TOOL-

- Can be used for any patent and trademark searching platforms with or without MT tools
  https://patents.google.com/
  http://www.wipo.int/branddb/en/

- Can be used for any HTML documents (HTML形式の電子出願用ファイル, e.g.) stored in a computer connected to an Internet

- Use of Google NMT engine

- No limitation on language pairs

HOW SAFE IS
FREE ONLINE MT SERVICES?

- "While there are always quality concerns with automatic translation tools, the bigger issue businesses should be aware of is security. **Everything you type into Google Translate runs through the tech giant's gigantic data mill**.... **Google Translate is no place for confidential business information or private details that should be kept secure**."

http://www.translateplus.com/blog/machine-translation-quality-isnt-concern-using-translation-technology/

- "**When content is entered into a free web-based machine translation tool, it is being delivered outside of your firewall and into the possession of another company.** This probably isn't a big deal when I'm trying to translate my Aunt Sophia's family meatball recipe from Italian, ...But what about sensitive or regulated information? ...."

http://www.sdlgov.com/risks-of-free-machine-translation/

GOOGLE TERMS OF SERVICE
-RISK OF BREACH OF CONFIDENTIALITY-

- When you upload, submit, store, send or receive content to or through our Services, you give Google (and those we work with) **a worldwide license to use, host, store, reproduce, modify, create derivative works (such as those resulting from translations, adaptations or other changes we make so that your content works better with our Services), communicate, publish, publicly perform, publicly display and distribute such content.** The rights you grant in this license are for the limited purpose of operating, promoting, and improving our Services, and to develop new ones.

✓ "This raises the question: Is your "confidential" information still confidential when you use Google Translate? Based on what's described in the company's terms and services, it doesn't seem like it is."

http://daily.unitedlanguagegroup.com/stories/editorials/update-google-translate-bad-business

ASK YOURSELF BEFORE USING FREE ONLINE MT SERVICES

- Free online MT is to be used for translation of publicly available documents such published patent documents including International Publication and file wrapper documents including office actions, amendments, and response filed.

- You had better **NOT** use free online MT services for translating confidential texts such as unpublished patent documents.

✓ Free Online MT services may be used by carefully editing an unpublished document to remove confidential portions such as novel portions.

✓ Free Online MT services may be used if it is not necessary to keep contents of application secret until the publication of application.

EVALUATION OF QUALITY OF NMT

- Evaluate quality of NMT output obtained by using Chrome's built-in Translate functionality

- Pick up a rather short patent publication relating to a simple invention ("a pillow" for non-Japanese publications).

- It is difficult to conduct an objective evaluation. Quality of MT may depend on various factors including language pairs, MT engines, technical areas, and description of source texts, and the result of evaluation may depend on evaluation factors (accuracy, fluency, term consistency and etc.).

- Please note that the evaluation in the following chart is a rough evaluation made based mainly on fluency of target texts.

- Practically, whether or not it is worth using NMT is an issue.

ROUGH EVALUATION MADE MAINLY ON FLUENCY OF TARGET TEXT

Source	Target	A	B	C	D	E	F
English text(US20180071133)	Japanese	4	4	5	4	4	4 - 5
English text(WO2015196135)	Japanese	4	4	5	4	4	4 - 5
Korean text (WO2018026215)	Japanese	5	5	6	5	4	5
Korean text (WO2016200110)	Japanese	5 - 6	5 - 6	6	5 - 6	5	6
Chinese text (CN106913778)	Japanese	4	4 - 5	- - -	4 - 5	4	5
Chinese text (WO2018098794)	Japanese	4	3 - 4	6	3 - 4	3	5 - 6
Japanese text (JPA2017-131808)	English	5	4 - 5	5 - 6	5	4	4 - 5
Japanese text (JPA2014-069016)	English	4 - 5	4 - 5	5	4 - 5	4	4 - 5

Part of document	
A	Background of the invention
B	Summary of the invention
C	Brief Description of drawings
D	Detail description (embodiments)
E	Independent claim
F	Dependent claims

Evaluation	
6	As good as human translation
5	Very good
4	Good
3	Acceptable(enough for gisting)
2	Rather bad
1	Bad

EXAMPLE (KOREAN TO JAPANESE) -AS GOOD AS HUMAN TRANSLATION-

発明の効果

本発明は、睡眠時に枕に内蔵されたスピーカーを介して発生される雨音、波の音、風の音などの音響やクラシック、ポップス、外国語などの音源が寝姿勢に関係なく、睡眠者に効果的に伝達される効果がある。そして、本発明は、枕にスピーカーが内蔵されていても、睡眠者が異物感を感じず、快適な睡眠姿勢をとり続けるようにすること、睡眠を誘導したり、学習が行われるようにすることができる効果がある。また、枕のタイマー機能を設定し、高速睡眠を誘導することができる効果がある。さらに、電磁波認証用が可能だけでなく、枕に異物が付着して吸収されても、スピーカーの損傷を防止することができる効果がある。

図面の簡単な説明

図1は、本発明の一実施形態に係るスピーカを備えた枕の構造を示した斜視図であり、

図2は、本発明の一実施形態に係るスピーカを備えた枕に設置されたスピーカーと音響ガイド部の構造を示す平面図であり、

図3は図1のスピーカーを備えた枕を上側から見た構造を示した平面図であり、

図4は図1のスピーカーを備えた枕を正面から見た構造を示した正面図であり、

図5は、本発明の一実施形態に係るスピーカを備えた枕をユーザーが使用する様子を示す使用状態図である。

EXAMPLE(CHINESE TO JAPANESE) -AS GOOD AS HUMAN TRANSLATION-

- 本発明は、スマート枕の分野に関し、より詳細には、圧電ケーブルの円筒状巻線構造を有するスマート枕に関する。
- 要約すると、本発明の技術的解決策は、上記目的を実現するのに十分に効果的であり得る、および本発明の構造的及び機能的原理は、完全に本から逸脱することなく、所望の効果及び目的を達成するために、実施例で確認されています本発明の原理および本質に基づいて、本発明の実施形態に対して様々な変更または修正を行うことができる。したがって、本発明は、特許出願に記載された範囲内の全ての置換内容を含み、本発明の特許出願の範囲内でなされた同等の変更は、本出願の特許範囲内に入るものとする。
- 本発明は、伝統的な漢方薬の技術分野に関し、特に血圧を下げる伝統的な漢方薬の枕に関する。
- 本発明の目的は、従来技術の上記の欠点を克服し、血液脂肪を減らし血管を清潔にし、圧力を降下させて調節することができる漢方薬枕を提供することである。

EXAMPLE(ENGLISH TO JAPANESE)
-COMMON MISTRANSLATION BY GOOGLE-

図面の簡単な説明

本発明の好ましい実施形態を、添付の図面を参照して、単なる例として以下に説明する。

イチジク。図1は、本いびき破壊システムの好ましい実施形態を示す概略図を示す。

イチジク。図2は、最大膨張位置における本いびき破壊システムの拡張ユニットの実施形態を示す概略図を示す。

イチジク。図3は、最大収縮位置での本いびき破壊システムの拡張ユニットの実施形態を示す概略図を示す。

✓ "イチジク" which means "figs fruit" in Japanese abruptly occurred in the section of "brief description of the drawings". No word "Figs (abbreviation of Figures)" in the original English text.

EXAMPLE(JAPANESE TO ENGLISH) -AUTOMATIC SUPPLEMENT OF ARTICLES IN CLAIM!!!-

A housing,

An electrode assembly provided in an internal space of the housing and provided with at least one set of a positive electrode and a negative electrode,

A water intake formed in the housing and taking in external water into the internal space,

A discharge port formed in the housing for discharging the water taken into the internal space to the outside,

筐体と、

筐体の内部空間に設けられ、少なくとも一組の正極・負極を備えた電極アセンブリと、

筐体に形成され、内部空間に外部の水を取り込む吸水口と、

筐体に形成され、内部空間に取り込まれた水を外部に吐出する吐出口と

IS NMT RELIABLE?

- "Recent studies show that NMT generally produces **fluent but inadequate translations** (Tu et al. 2016b; Tu et al. 2016a; He et al. 2016; Tu et al. 2017)." Neural Machine Translation Advised by Statistical Machine Translation, Xing Wang, et al.

- "Translations generated by NMT systems **often lack of adequacy**. It has been widely observed that NMT tends to **repeatedly translate some source words while mistakenly ignoring other words.** When generating target words, the decoder often repeatedly selects some parts of the source sentence while ignoring other parts, which leads to **over-translation and under-translation** (Tu et al. 2016b). Neural Machine Translation with Reconstruction, Zhaopeng Tu, et al

- "A significant weakness in conventional NMT systems is **their inability to correctly translate very rare words**", Addressing the Rare Word Problem in Neural Machine Translation, Minh-Thang Luong, et al.

CRUCIAL DRAWBACKS OF NMT

- Translation is fluent but not faithful to the original text.
- ✓ Over-translation and under-translation

Source text

Target text

Under-translation
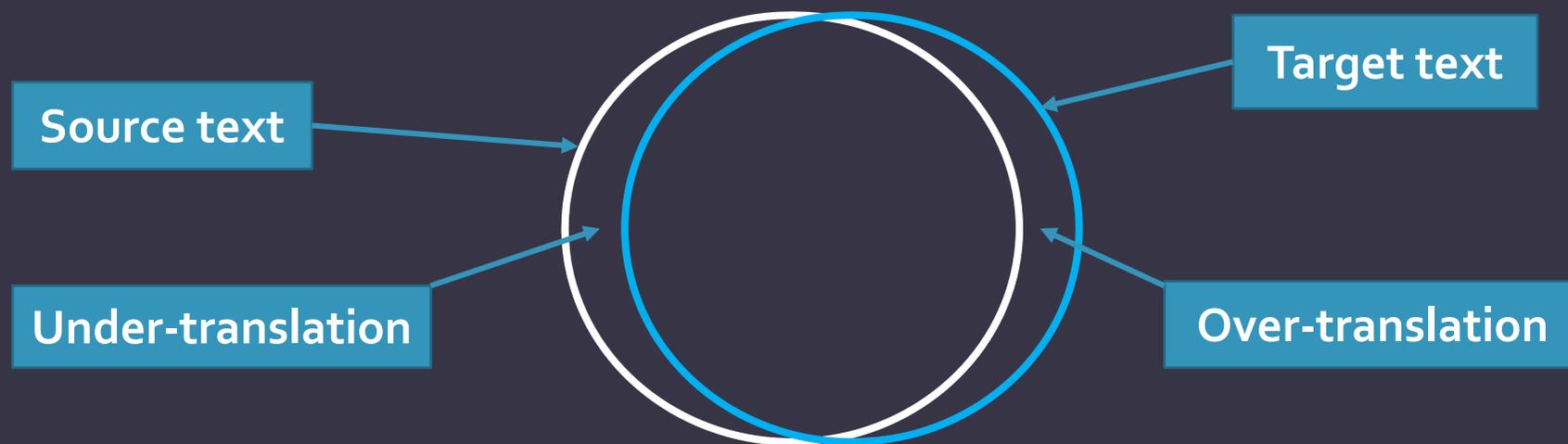
Over-translation

- ✓ Inconsistencies between target terms for the same source term

CRUCIAL DRAWBACKS OF NMT -CONTINUED-

- The drawbacks such as over-translation, under-translation, and inconsistencies conflict with the requirements for translating patent documents to be filed with patent offices.

- Human intervention (post-editing) is necessary.

- Due to fluency, it is sometimes difficult to detect discrepancies and mistranslations. Post-editors must examine NMT output very carefully.

OVER-TRANSLATION
-CHINESE TO JAPANESE/ENGLISH-

- Source sentence

  本发明涉及中药技术领域 (CN)

- Target sentences

  本発明は、伝統的な漢方薬の技術分野に関する（JP）

  The invention relates to the technical field of traditional Chinese medicine(EN)

- ✓No word corresponding to "伝統的" as well as "traditional" in the original Chinese text.

- ✓Very likely to overlook discrepancies if a target sentence is fluent.

DISCREPANCIES, INCONSISTENCIES -CHINESE TO JAPANESE-

1.一种降血压中药枕，包括枕芯和枕套，所述枕芯内装有降血压中药混合物，其特征在于，所述混合物包括下述质量份的原料：蚕沙15-25份、降压草15-25份、玉米须10-18份、茶叶20-25份、决明子10-20份、荞麦壳15-25份、金银花10-20份、夏枯草10-20份、蒲公英10-20份、绿豆衣5-10份、罗布麻叶5-10份、桑叶5-10份、甘草5-10份、薰衣草3-5份、白菊3-5份。

2.根据权利要求1所述的降血压中药枕，其特征在于，所述枕芯和枕套之间还设有若干块磁石。

枕コアと枕ケースとを備え、枕コアに血圧降下漢方剤が充填された血圧降下漢方薬枕であって、前記混合物は、蚕砂15-25部と血圧草15-25部、トウモロコシ10-18部、茶20-25部、アカシア10-20部、。ソバ15-25部、ハニーサックル10-20部、プルネラ10-20部、タンポポ10-20部緑豆衣類5-10部、アポカニム葉5-10部、桑の葉5-10部、甘草5-10部、ラベンダー3-5部、白菊3-5部である。

前記ピローコアと前記ピローケースとの間に複数の磁石がさらに配置されていることを特徴とする請求項1に記載の血圧降下剤。

INCONSISTENCIES, RARE WORDS UNDER-TRANSLATION -ENGLISH TO JAPANESE-

- By turning on the inflator **12**, the _bladder_ **16** is inflated, expanding the _expander unit_ **11**, causing the _snorer_'s head that is located on the pillow to gently move up or in any other direction depending on the position of the head **during a predefined period of time**. After the predefined period of the inflator **12** being on and the _expander unit_ **11** reaching a predefined maximum expansion, the inflator **12** is turned off by the control system **22**, allowing the _bladder_ **16** to depressurize through the valve **13**, causing the _expander unit_ **11** to reduce in cross sectional area, resulting in the _snorer_'s head to move down or in any other direction depending on the position od the head.

- インフレータ**12**を作動させることによって、膀胱**16**は膨張され、_拡張ユニット_**11**を膨張させ、枕上に配置された_snorer_の頭部を、頭の位置に応じて上向きに、または任意の他の方向に、時間。インフレータ**12**の所定の期間がオンであり、_膨張器ユニット_**11**が所定の最大膨張に達した後、インフレータ**12**は制御システム**22**によってオフにされ、ブラダ**16**がバルブ**13**を介して減圧され、_膨張器ユニット_**11** 断面積を減少させて、頭部の位置に応じて_嗅覚者_の頭部を下方に、または他の方向に移動させる。

RARE WORD?
-KOREAN TO JAPANESE-

- 床に安着されて、睡眠者の頭を支持することにより、上記の睡眠が睡眠姿勢をとることができようにし、複数の他工法（110）が貫通形成された枕（100）と、前記枕（100）の内部の下面に埋め込まれた形で設置されて音源や音響が発生するようにするスピーカー（200）と、前記スピーカー（200）から発生される音源や音響が伝播される方向をガイドして、睡眠者の耳側に伝達されるように一側が上記のスピーカー（200）と近接した位置に配置され他側が上記枕（100）の表面と近接した位置に配置され、上記の枕（100）に形成された他工法（110）に埋め込まれた形で設置されるガイド管（310）を備えた音ガイド部（300）を、含み、前記スピーカー（200）の上部には、上記のスピーカー（200）で発生される前記音源や音響が、前記ガイド管（310）に誘導されるように、上記のスピーカー（200）を外部と遮蔽させる別の遮蔽部材（320）が備えられ、上記ガイド管（310）は、・・・・

- The target sentence seems to be fluent but a claim becomes incomprehensible due to one mistranslation of 타공(110).

- [Korean to English] A pillow (100) having a plurality of perforations (110) penetrating through which…

DISCREPANCIES, OVER-TRANSLATION, AND UNDER-TRANSLATION -JAPANESE TO ENGLISH-

2. The method according to claim 1,

The high water absorption fiber is in the form of a plate, the water absorbent resin is granular, and the granular water absorbent resin is sealed in the water permeable bag in a state of being in contact with the surface of the plate-Wherein the cooling device is a cooling device.

3. The method according to claim 2,

The cooling device according to claim 1 or 2, wherein the second water absorption rate is equal to or lower than the first water absorption rate.

【請求項２】前記高吸水繊維は板状であり、前記吸水樹脂は粒状であり、前記粒状の吸水樹脂は、前記板状の高吸水繊維の表面に接触した状態で前記前記透水性袋体内に封入される、請求項１に記載の冷却具。

【請求項３】前記第２の吸水率は前記第１の吸水率と同等か前記第１の吸水率よりも小さい、請求項１、２いずれかに記載の冷却具。

UNDER-TRANSLATION -JAPANESE TO ENGLISH OBTAINED BY *PATENT TRANSLATE-*

An electrode assembly provided in an internal space of the casing and including at least one pair of a positive electrode and a negative electrode; a water intake formed in the casing for taking in external water into the internal space; , A discharge port for discharging the water taken into the internal space to the outside, a filter provided in the water intake port, rotary blades provided so as to rotate in the internal space of the housing, A motor for rotating the blades, and a current supply means to the electrode assembly, wherein in the internal space, the water intake port and the filter, the discharge port and the rotary blade are located on opposite sides of the electrode assembly

筐体と、
筐体の内部空間に設けられ、少なくとも一組の正極・負極を備えた電極アセンブリと、
筐体に形成され、内部空間に外部の水を取り込む吸水口と、
筐体に形成され、内部空間に取り込まれた水を外部に吐出する吐出口と、
吸水口に設けられたフィルターと、
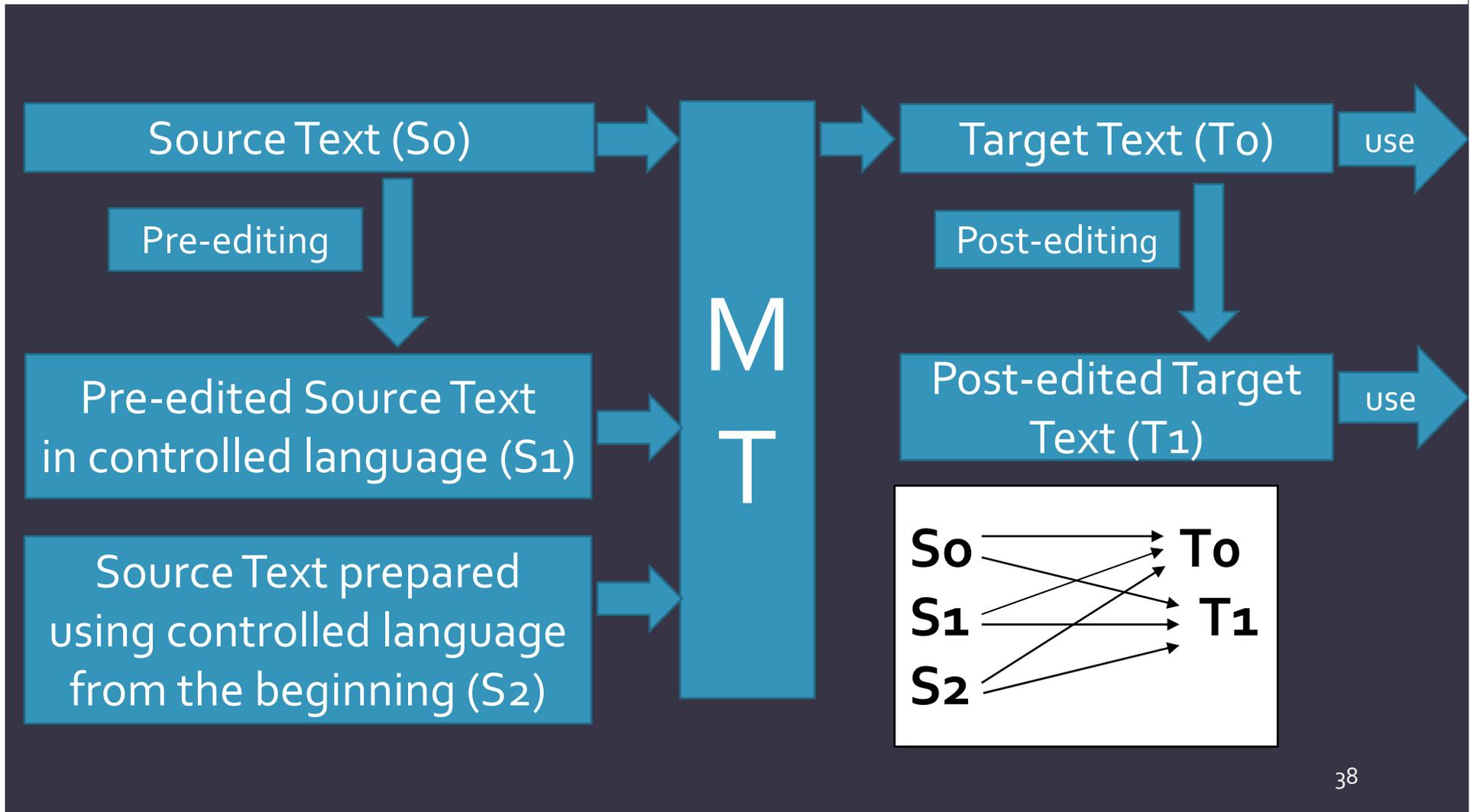筐体の内部空間で回転するように設けられた回転羽根と、
筐体に設けられ、回転羽根を回転させるモータと、
電極アセンブリへの通電手段と、
を備え、
前記内部空間において、前記吸水口及び前記フィルターと、前記吐出口及び前記回転羽根
は、前記電極アセンブリを挟んで反対側に位置しており、

PRACTICAL USE OF MT
-OVERALL SCHEME FOR USE OF MT-

Source Text (So)

Pre-editing

Pre-edited Source Text in controlled language (S1)

Source Text prepared using controlled language from the beginning (S2)

MT

Target Text (To) — use

Post-editing

Post-edited Target Text (T1) — use

So → To
S1 → T1
S2

PRE-EDITING & CONTROLLED INPUT
-MACHINE TRANSLATION-FRIENDLY LANGUAGE-

- Pre-edited or controlled input (source text) boosts output (target text) quality.

- Pre-editing is the process of adjusting a source text before machine translation in order to improve output of MT and reduce the amount of work required in the post-editing.

- Pre-editing may include simplifying the structures by using the short sentences and avoiding complex syntactic structures, avoiding ambiguities and term inconsistencies and such.

- Pre-editing requires language skills of source language.

- 特許ライティングマニュアル「産業日本語」provides examples of controlled language.
  https://www.tech-jpn.jp/tokkyo-writing-manual/

POST-EDITING

- Post-editing is the process of reviewing and correcting the MT output.

- Post-editing may be defined as "proofreading in conjunction with MT."

Human translation (3 days, e.g.)	→	Proof-reading (a few hours, e.g.)
Machine translation (a few seconds, e.g.)	→	Post-editing (several hours or 1 day, e.g.)

- Pre-editing may not totally eliminate the need for post-editing.

- Post-editing requires language skills of both source and target languages.

PRACTICAL USE OF MACHINE TRANSLATION IN IP PRACTICE

- Translation of searched documents /documents cited in office action

- Translation of file wrapper documents

- A (rough) translation of patent documents to be filed with Patent offices

TRANSLATION OF SEARCHED DOCUMENTS/ DOCUMENTS CITED IN OFFICE ACTION

- MT output makes gisting (getting a basic idea) of patent documents such as searched documents and documents cited in O.A. possible without any human intervention.

- Translations of patent documents such as searched documents and documents cited in O.A. can be sent to foreign associates or USPTO as IDS documents (Disclaimer may be necessary).

- MT can also be used for translating description of goods/services for trademark publication/application.

TRANSLATION OF FILE WRAPPER DOCUMENTS

- MT output may make gisting of office action and argument possible.

- Translating of office action and argument sometimes require interpretation of context and reading between the lines but machine cannot understand context.

- Pre-editing and/or post-editing may be necessary but human translation may be more efficient and accurate.

- Drafting office action using controlled language may improve the quality of translation for office actions.

A ROUGH TRANSLATION OF PATENT DOCUMENTS TO BE FILED WITH PATENT OFFICES

Pre-editing *optional*	→	A rough translation by MT	→	Human translation or post-editing

- MT can be used as "a productivity tool" for translating patent documents for filing purpose.

- Use of a rough translation (下訳) obtained by MT may improve the translation productivity.

- If the quality of a target text is very high (more than a rough translation), hybrid use of machine translation and post-editing may work.

44

HOW WILL MACHINE TRANSLATION AFFECT IP PRACTICE?

- Neural Machine Translation (NMT) has dramatically improved the quality of machine translations.

- NMT is a powerful productivity tool.

- It is inevitable for IP firms to embrace an emerging MT technology to improve productivity.

- On the other hand, for some Asian IP firms, the use of MT is not merely a matter of productivity as patent document translation work is one of the major sources of income for the firms.

Will MT totally replace human translation in respect of patent translation ?

WILL MT TOTALLY REPLACE HUMAN TRANSLATION IN RESPECT OF PATENT TRANSLATION ?

The answer may be "No."

- MT is not perfect and at least NMT has crucial drawbacks such as under-translation, over-translation, and term inconsistencies for translating patent documents for filing purposes.

- Translation of patent documents for filing with patent offices require extremely high quality. Even if MT produces an almost-ready translation, mistranslation of even one word in the claim could become a serious problem.

- Post-editing by human at least is necessary even if the drawbacks will be eliminated.

IS IT WORTH USING NMT FOR TRANSLATING PATENT DOCUMENTS FOR FILING PURPOSE?

Other things (output quality) being equal, which is more efficient "machine translation with human intervention" or "human translation"?

- At the present quality of NMT, the answer may depend on cases.

- Pre-editing and post-editing take time. If the output quality is very low, post-editing may be extremely time-consuming work.

- If MT produces an almost-ready translation, post-editing may be easier.

- In any case, it should be noted that MT produces a lengthy target text in a few to several seconds.

- Computer-assisted pre-editing/post-editing tools may be developed and introduced in the future.

WHICH IS MORE EFFICIENT?
## - USE OF MT WITH POST-EDITING MAY DOMINATE IN THE FUTURE -

| Human translation, T1 | → | Proof-reading, T2 |

Replaced?

| Pre-editing, T3 | → | MT, T4 ≒ 0 | → | Post-editing, T5 |
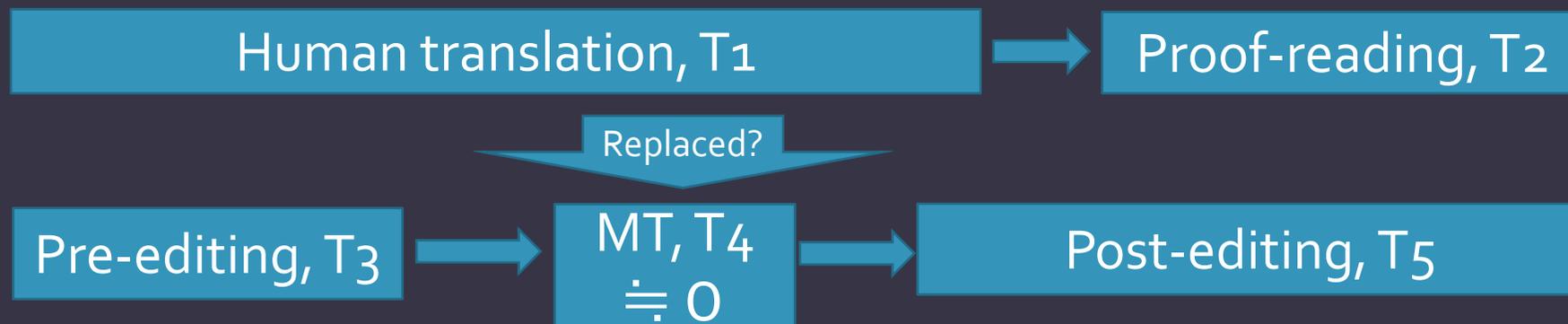
T refers to "time required, namely workload, for the task".
T3 is zero if pre-editing is not conducted.

➤ If "T1 + T2 >> T3 + T5", it make sense to use MT.

➤ If "T1 + T2 ≦ T3 + T5", it does not make any sense to use MT.
  手直しするよりも最初から自分で訳した方が早いケース！

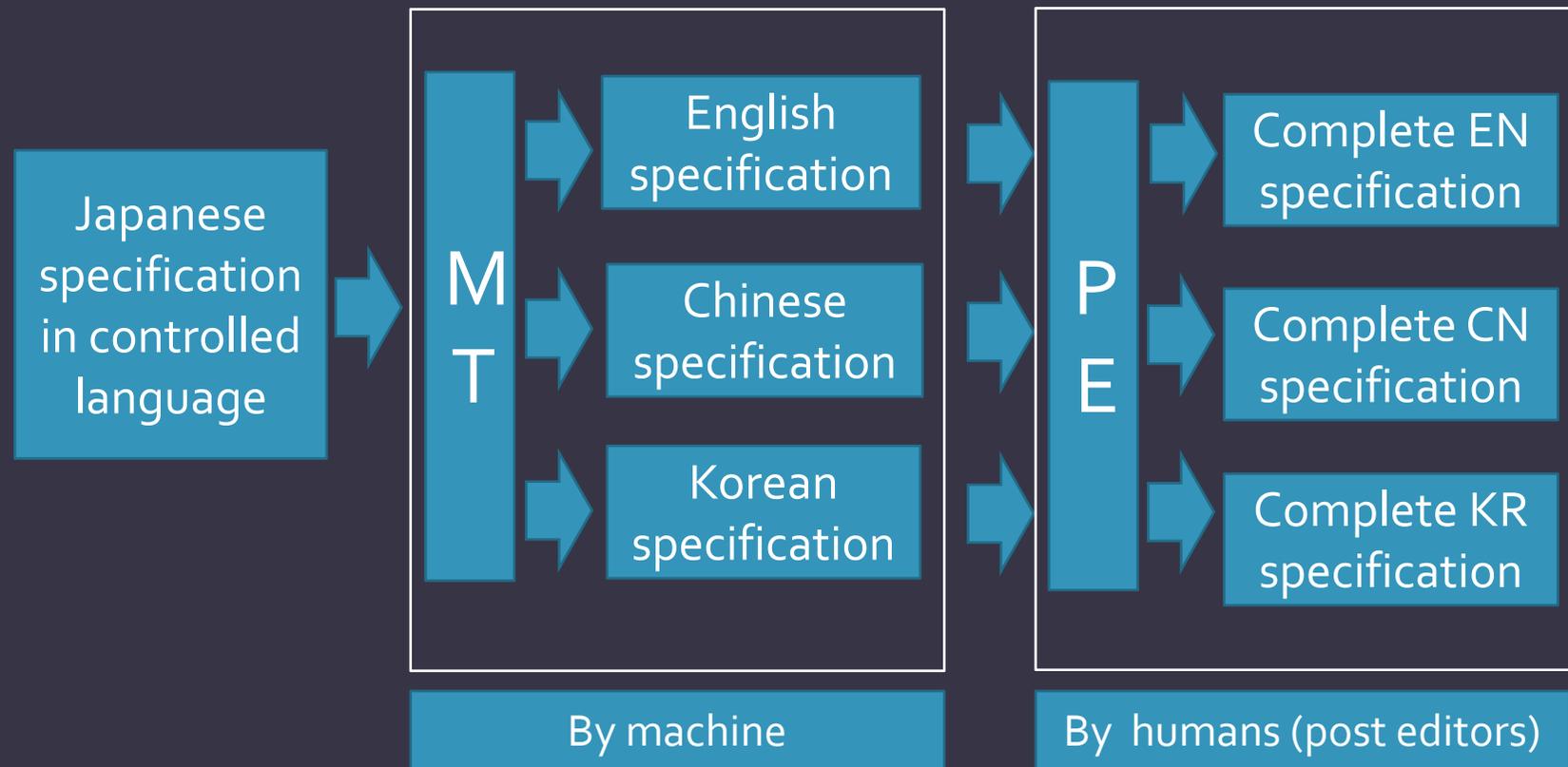# FUTURE PRACTICAL USE OF MT FOR PATENT TRANSLATION
-TRANSLATION WORK (AT LEAST, POST-EDITING) WILL REMAIN-

Human translation → Proof-reading

Replaced ↓

Pre-editing *Optional* → Machine translation → Post-editing ↑
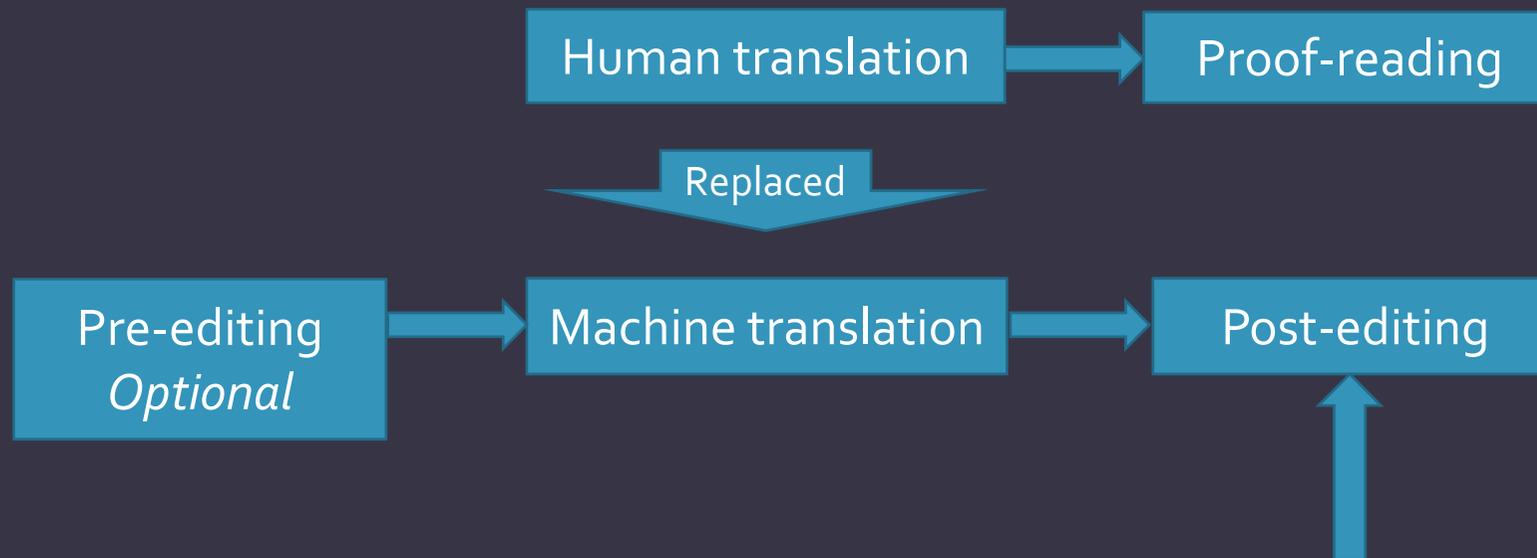
Post-editing skills will become vital in patent translation. Post-editing skills include skills for efficiently detecting and correcting discrepancies, knowledge and understanding of patent specification and claims as well as language skills of both source and target texts.

POSSIBLE FUTURE SCHEME FOR PREPARING DOCUMENTS FOR FOREIGN FILING

Japanese specification in controlled language

MT

→ English specification

→ Chinese specification

→ Korean specification

By machine

PE

→ Complete EN specification

→ Complete CN specification

→ Complete KR specification

By humans (post editors)

PERSPECTIVE

- It is inevitable for IP firms to embrace MT technology to improve productivity.

- MT will make global patent searches more efficient.

- Use of MT with human intervention may be mainstream for translating patent documents in the future.

- Translation work (at least, post-editing) will remain and post-editing skills will become vital in patent translation.

- Overall translation cost may be lowered because the use of MT will improve translation productivity.

- Hopefully, lower translation costs may encourage SMEs to consider filing foreign patent applications.

THANK YOU