

AIO

Editing a cura di Domenico Lombardini, traduzione dei contributi 2 e 10 del testo originale a cura di Annalisa Lanzi; traduzione dei contributi 1, 7, 8, 9, 12, 13, 14, 15 e 16 del testo originale a cura di Sara Grizzo; traduzione del contributo 11 del testo originale a cura di Lavinia Tomasinelli e Sara Galluccio.

Machine Translation

Come usarla in modo professionale

a cura di

Domenico Lombardini

Diego Cresceri

Francesco Mana

Contributi di

Aljoscha Burchardt

Diego Cresceri

Christian Eisold

Florian Faes

Sara Grizzo

Silvia Hansen-Schirra

Kim Harris

Domenico Lombardini

Arle Lommel

Francesco Mana

Jean Nitzke

Jörg Porsiel

Moritz Schaeffer

Michael Schneider

Uta Seewald-Heeg

Ilona Wallberg

Marion Wittkowsky



Traduzione parziale dell'opera originale

Machine Translation: What Language Professionals Need to Know
edited by Jörg Porsiel
BDÜ Fachverlag, Berlin 2017



Aracne editrice

www.aracneeditrice.it
info@aracneeditrice.it

Copyright © MMXIX
Gioacchino Onorati editore S.r.l. – unipersonale

www.gioacchinoonoratieditore.it
info@gioacchinoonoratieditore.it

via Vittorio Veneto, 20
00020 Canterano (RM)
(06) 45551463

ISBN 978-88-255-2874-9

*I diritti di traduzione, di memorizzazione elettronica,
di riproduzione e di adattamento anche parziale,
con qualsiasi mezzo, sono riservati per tutti i Paesi.*

*Non sono assolutamente consentite le fotocopie
senza il permesso scritto dell'Editore.*

I edizione: novembre 2019

Indice

- 7 Prefazione
D. Cresceri
- 11 Machine Translation e traduzione brevettuale
D. Lombardini, F. Mana
- 19 Traduzione automatica. A che punto siamo?
J. Porsiel, A. Burchardt
- 25 Il settore dei servizi linguistici e la seconda era delle macchine
F. Faes
- 35 Lingua controllata e traduzione (automatica) nella comunicazione tecnica
M. Wittkowsky
- 45 Integrazione della traduzione automatica nei CAT tool
U. Seewald-Heeg
- 57 Il ruolo della terminologia nella traduzione automatica
C. Eisold
- 71 Perfezionamento della traduzione automatica. Divario tra gli approcci della ricerca e le esigenze del settore
K. Harris, A. Burchardt
- 83 La gestione della qualità nella traduzione
A. Lommel, A. Burchardt
- 103 DIN EN ISO 18587. Una norma sul processo di post-editing
I. Wallberg
- 109 La formazione dei post-editor
U. Seewald-Heeg

- 115 Post–editing. Termini e definizioni
 S. Hansen–Schirra, M. Schaeffer, J. Nitzke
- 129 Il post–editing come specializzazione. Una testimonianza
 S. Grizzo
- 135 Post–editing. Come trovare un equilibrio tra costi e guadagno
 M. Schneider
- 143 Autori

Prefazione

D. CRESCERI*

Intelligenza artificiale, deep learning, big data, robot intelligenti con sembianze sempre più umane e in grado di simulare perfettamente movimenti, a volte emozioni, della nostra specie: l'elenco di tecnologie che hanno rivoluzionato ogni aspetto della nostra vita potrebbe andare avanti all'infinito.

I servizi linguistici non sono certo immuni.

Dopo “soli” quindici anni in questo meraviglioso settore, se potessi avere una macchina del tempo per tornare al mio primo incarico di traduzione troverei un mondo difficilmente riconoscibile, dove l'unica tecnologia a mia disposizione erano CAT tool che non assomigliano neppure lontanamente a quelli di oggi. Il cloud computing era poco più che un'idea, e la collaborazione tra traduttori era limitata a gruppi di mailing invasivi, difficili da seguire e onestamente alquanto fastidiosi.

Oggi abbiamo a disposizione una vasta gamma di tecnologie, dai CAT basati su cloud che integrano funzionalità di intelligenza artificiale, ai sistemi di vendor management automatizzato, passando per la machine translation basata su machine learning e i sistemi di collaborazione avanzati.

Non sono però solo i nostalgici del passato a chiedersi quali sono i rischi di questa ondata di innovazione, e se, a conti fatti, i vantaggi portati da questo progresso incessante saranno in grado di superare i relativi svantaggi.

Senza dubbio stiamo assistendo a un progresso tecnologico particolarmente rapido. Tuttavia, non è certo la prima volta che una ventata di innovazione genera contemporaneamente grandi speranze e altrettanto grandi paure.

Basti pensare al movimento luddista, che prese il nome da un (probabilmente immaginario) tale Ned Ludd, che nel 1779 avrebbe distrutto un telaio in segno di protesta e sabotaggio della rivoluzione industriale. Dal XVIII secolo ad oggi si sono susseguiti numerosi cicli di innovazione, e con il rapido avanzare dell'intelligenza artificiale, molti si sono interrogati

* Diego Cresceri è fondatore e CEO di Creastive Words, società di localizzazione con sede a Genova. Dal 2010 insegna machine translation post-editing in diverse università, corsi postlaurea e aziendali.

ed espressi sulle possibili conseguenze. L'essere umano diventerà schiavo delle tecnologie create da lui stesso? I robot intelligenti prenderanno il sopravvento? Ad affrontare l'argomento imprenditori, scienziati e filosofi.

Oggi come in passato, una delle paure più grandi è sicuramente legata alle possibili ripercussioni della tecnologia sul livello occupazionale.

Eppure, analizzando attentamente il contesto economico, è evidente quanto ad ogni grossa rivoluzione sia sicuramente corrisposto un altrettanto forte shift occupazionale, ma con benefici evidenti a livello globale.

L'avvento del personal computer è un caso esemplare: quando fu introdotto su larga scala rese disoccupati centinaia di migliaia di dattilografi, segretarie e numerosi altri profili, creando però oltre 19 milioni di nuovi posti di lavoro, solo negli Stati Uniti e solo tra il 1980 e il 2015, con un incremento netto dei posti di lavoro di 15,7 milioni.

Lo stesso risultato, pur con cifre diverse, si è registrato all'inizio del secolo scorso con l'invenzione dell'automobile.

Prevedere l'impatto dell'attuale ondata tecnologica non è umanamente possibile. McKinsey stima che entro il 2030 il 14% della forza lavoro mondiale dovrà cambiare occupazione per evitare l'obsolescenza. Altri studi hanno restituito risultati simili, ma la verità è forse che nessuno è in grado di misurare davvero la velocità del cambiamento che stiamo vivendo.

Lo diciamo ancora una volta: il settore linguistico non è esente da tutto questo! Se è vero che i lavori più a rischio sono quelli altamente routinari e automatizzabili, è comprensibile che chi si occupa di traduzioni possa considerare la tecnologia e l'intelligenza artificiale come una grande minaccia.

La domanda che bisogna porsi è quindi: ci si può preparare per non diventare obsoleti e non trovarsi ad affrontare il destino di quei 3,5 milioni di persone a cui il personal computer ha sottratto il posto di lavoro?

La traduzione automatica, nata negli anni Trenta del secolo scorso, ha ottenuto, pur con varie battute di arresto, risultati importanti. Ma è a inizio 2017, con l'introduzione della Machine Translation neurale da parte di Google, che abbiamo osservato un cambiamento epocale, con grandi progressi in termini di qualità e scorrevolezza del testo tradotto. Da allora, tra un proclama iperbolico e l'altro, sono numerose le aziende che hanno più o meno messo da parte i vecchi modelli rule-based o statistici e hanno investito su questa nuova tecnologia. Se la Machine Translation neurale riuscirà o meno a mantenere le promesse non è ancora dato saperlo. Quello che è certo è che ha raggiunto livelli di qualità mai visti, e non ci sono motivi di pensare che non continui sulla strada del miglioramento.

Nonostante l'indiscutibile progresso, la Machine Translation è ahimè ancora vittima di una visione distorta e parziale, in Italia e non solo.

Per numerosi clienti, o buyer, la traduzione automatica è il termine di paragone per un livello di qualità ben al di sotto delle aspettative. « Sembra una traduzione fatta da Google Translate » è la frase preferita dai clienti quando, a ragione o (spesso) meno, non sono soddisfatti del lavoro ricevuto.

Un'altra schiera (in continua crescita) di clienti, al contrario, considera la traduzione automatica come la panacea di tutti i mali, per ottenere ottima qualità in tempi brevissimi. « La traduzione me la faccio io con Google Translate così faccio prima e risparmio » è più o meno l'espressione usata da questo tipo di clienti.

Ovviamente, nessuno dei due siede dalla parte della ragione. Il primo ignora del tutto i miglioramenti della Machine Translation a cui abbiamo assistito negli ultimi anni, mentre il secondo non considera che, per la maggior parte dei contenuti, la Machine Translation non è (ancora, quanto meno) sufficiente senza un intervento di post-editing da parte di linguisti specificamente formati, e ignora le conseguenze di una strategia di localizzazione incentrata esclusivamente sul risparmio a tutti i costi.

Per quanto riguarda le agenzie di traduzione, non di rado la traduzione automatica viene utilizzata esclusivamente come un modo rapido e relativamente facile per rispondere alle pressioni di prezzo imposte dai buyer. Ne consegue che, troppo spesso, le agenzie mancano di un approccio strutturato e serio all'implementazione della Machine Translation nel proprio flusso di lavoro, figuriamoci la consapevolezza che sia necessaria una formazione specifica al proprio team (interno ed esterno) per affrontare questo (relativamente) nuovo servizio.

Infine, la maggior parte dei traduttori professionisti, in particolare chi ha una esperienza di traduzione "tradizionale" medio-lunga, considera la Machine Translation come l'ultimo ritrovato utilizzato dai loro clienti per spremere economicamente l'ultimo anello della catena di fornitura, al pari dell'introduzione degli strumenti commerciali di traduzione assistita negli anni Novanta.

Esiste per fortuna qualche eccezione.

Alcuni clienti hanno abbracciato la Machine Translation in modo corretto e consapevole, come l'unico mezzo possibile per far fronte alla necessità di tradurre volumi sempre crescenti di contenuto, e adottando un approccio strutturato e pianificato.

Numerose agenzie di traduzione, dalle più piccole alle più grandi, stanno inoltre approcciando la Machine Translation con un atteggiamento proattivo anziché reattivo, instaurando un rapporto più collaborativo con il proprio team (interno o esterno). Ne conseguono una maggiore preparazione e trasparenza e, last but not least, la consapevolezza che sia fondamentale formarsi in modo specifico su traduzione automatica e post-editing. Infine, un numero crescente di traduttori sta realizzando quanto non abbia senso continuare a nascondere

la testa sotto la sabbia fingendo che la tecnologia non esista e quanto sia opportuno, se non abbracciarla, quanto meno conoscerla a fondo.

A riprova di questa aumentata apertura alla traduzione automatica, nel 2019 sono aumentati esponenzialmente i corsi sul post-editing offerti dalle principali associazioni di settore, nonché la richiesta da parte delle agenzie di traduzione di formazione specifica su Machine Translation e post-editing.

Mentre clienti, aziende e professionisti si stanno adattando, con ritmi diversi, al cambiamento epocale promesso (e, a mio modesto parere, avviato con successo) dalla Machine Translation, il mondo dell'istruzione arranca per restare al passo con i tempi. I programmi formativi della maggior parte dei corsi in traduzione, triennali e pure magistrali, dedicano poco spazio alla tecnologia applicata alla lingua, sia essa la traduzione assistita (che, ricordiamo, è stata introdotta negli anni Novanta!) o la traduzione automatica.

Con la tecnologia che migliora a un ritmo incessante e sempre crescente, la naturale conseguenza della lentezza di adattamento da parte del mondo accademico è un gap sempre maggiore tra le competenze sviluppate durante il percorso di studi e le effettive esigenze del mercato. Problema non solo italiano questo.

Questo gap di conoscenza, però, non è l'unico problema quando si parla di tecnologia e istruzione.

Ancora troppo spesso il mondo del lavoro è portato a giudicare i nostri laureati in base alla loro conoscenza della materia (che, appunto, è spesso comunque scarsa) piuttosto che in base alle loro capacità di problem-solving e di pensiero creativo, competenze fondamentali in un ambiente di lavoro con una presenza massiccia dell'intelligenza artificiale.

Il sistema formativo va adattato senza ritardo per preparare i giovani a una nuova era in cui gli esseri umani dovranno lavorare molto più intensamente insieme alle macchine intelligenti.

Non si può però delegare totalmente questa responsabilità al mondo accademico: anche il mondo del lavoro può e deve fare la propria parte in questo percorso.

Parafrasando le parole di Lyndon Johnson, presidente degli Stati Uniti tra il 1963 e il 1969, l'ignoranza e inerzia sono i nostri peggiori nemici, mentre l'automazione può essere un perfetto alleato ma solo nel momento in cui ci prepariamo con una pianificazione adeguata.

Consapevoli che poche pagine non possano essere sufficienti per colmare i gap di conoscenza che chi si affaccia oggi nel mondo del lavoro, né tanto meno a facilitare lo sviluppo di competenze di problem-solving e pensiero creativo, speriamo con questo volume di dare il nostro piccolo contributo verso una maggiore presa di consapevolezza su cosa sia la Machine Translation e su come la si possa implementare a vari livelli con risultati positivi per tutti gli attori coinvolti. Buona lettura!

Machine Translation e traduzione brevettuale

D. LOMBARDINI, F. MANA*

Spesso molti indulgono in facili infatuazioni per le nuove tecnologie. Siamo immersi in una società in cui il feticismo tecnologico pervade gli stili di consumo, e ogni giorno vediamo sul mercato nuovi beni e servizi di talvolta molto dubbia utilità reale. Nondimeno, al primo apparire delle soluzioni di *Machine Translation* (MT) soltanto i più reazionari tra i linguisti e in generale tra coloro impegnati nel mondo dei servizi linguistici non ne hanno colto o intuito le potenziali opportunità. La nostra esperienza, molto focalizzata sulle traduzioni brevettuali, rappresenta sì un caso di studio specifico ma è anche un esempio d'interesse generale, sia per il traduttore *freelance* che volesse comprendere meglio le potenzialità della MT, sia per l'LSP (*Language Service Provider*) che ancora recalcitra dinanzi a uno strumento che sta diventando sempre più un ausilio indispensabile per sopravvivere in un settore ad alto contenuto tecnologico, quale è quello dell'industria della traduzione.

L'aumento esponenziale di contenuti di cui si richiede una traduzione, ancorché approssimativa (si pensi alle traduzioni delle descrizioni dei prodotti nei portali di e-commerce), e il relativo aumento del peso dell'economia rivolta all'export hanno fatto sì che il mercato della traduzione vedesse, malgrado un periodo di crisi economica, un costante e robusto incremento, anche in Italia. La MT, in questo caso, si è resa indispensabile e ha essa stessa alimentato una domanda di servizi di traduzione che altrimenti non avrebbe potuto svilupparsi, perché i costi di traduzioni integralmente umane sarebbero stati letteralmente proibitivi. In sintesi, si può dire che questa tecnologia ha fatto da volano alla domanda e che la domanda ha, a sua volta, suscitato il miglioramento della tecnologia, gettando le basi per un circolo virtuoso. È un tipico esempio di un'offerta (in questo caso una nuova tecnologia) che alimenta una domanda (in questo caso di servizi linguistici). I brevetti, tuttavia, sono un caso a sé stante e non sovrapponibile a nessun altro. Il brevetto riassume in sé tutto ciò che di più di ostico e deli-

* Domenico Lombardini, laureato in Scienze biologiche, è fondatore e proprietario di ASTW (www.a-stw.com). Francesco Mana, laureato in traduzione e interpretariato all'Università degli Studi di Genova, lavora come traduttore tecnico.

cato possiamo trovare in un testo da tradurre. Il brevetto è, infatti, un testo avente valore legale (con tutte le ricadute che può avere tale eccezione) e che è anche il compendio, la *summa* di una tecnica nella sua espressione più avanzata e *cutting-edge*. Ne consegue che un testo del genere dovrà essere tradotto con particolare attenzione, sia dal punto di vista terminologico, sia in termini di fedeltà al testo d'origine. La traduzione di un brevetto non può tollerare nessun tipo di errore, sebbene vi sia una gerarchia di errori di crescente gravità, che va dalla non appropriatezza terminologica alle omissioni o aggiunte rispetto al testo originale.

I primi strumenti MT che abbiamo provato per velocizzare la traduzione brevettuale sono stati quelli *rule-based*, ossia i software di traduzione automatica basati su regole. La macchinosità, la lentezza e la necessità di specificare dettagliatamente ogni lemma inserito nel dizionario di tali programmi ne scoraggiava l'utilizzo, sebbene i risultati fossero talvolta molto buoni, una volta raggiunto una dimensione di dizionario appropriata. Un altro difetto fondamentale di tali sistemi era la non perfetta integrazione con gli strumenti di traduzione assistita (*Computer-Aided Translation tool*, CAT tool): spesso, per procedere con il post-editing del testo, era infatti necessario preparare una memoria di traduzione derivante dall'allineamento tra testo originale e testo tradotto dalla macchina, memoria che sarebbe stata poi importata nel CAT tool d'interesse e utilizzata per rivedere e correggere (post-editing) il testo tradotto dalla macchina avendo a fronte il testo *source*. Come si può intuire, tale procedimento consumava un bel po' di tempo, e tale tempo andava a sottrarsi al guadagno di tempo che si sperava ottenere con la traduzione automatica. Da ultimo, ma non per importanza, l'output di una macchina *rule-based* risulta spesso innaturale, in quanto tali sistemi non imitano in maniera adeguata il linguaggio naturale di un traduttore "umano". Si rendeva quindi necessario un sistema MT perfettamente integrabile con l'ambiente di traduzione assistita utilizzato, con tempi di risposta adeguatamente veloci, e che potesse eventualmente "imparare" nuove risorse terminologiche e, possibilmente, preferenze e stili traduttivi, senza un eccessivo intervento "umano" che ne scoraggiasse l'utilizzo, e il cui output fosse più vicino al linguaggio naturale.

L'avvento delle macchine statistiche e basate su "reti neurali" ha in gran parte risolto tali problemi e oggi ve ne sono sul mercato di diversi tipi. È possibile ad esempio usare macchine "vuote", il cui motore viene creato *ex-novo* utilizzando risorse terminologiche proprie, la maggior parte delle volte memorie di traduzione. In tal caso, è necessario disporre di memorie di traduzione di adeguate dimensioni per avere un motore sufficientemente buono. Altrimenti, è possibile usare sistemi pronti, "chiavi in mano", sulla stregua di Google Translate, anch'essi perfettamente integrabili con l'ambiente di traduzione assistita. Sebbene sarebbe preferibile, almeno in

linea teorica, avere un motore “allenato” con dati propri e *domain-specific*, la nostra esperienza ci dice che anche avvalendosi di motori apparentemente “generici” come Google Translate la qualità dell’output è spesso molto buona, sia in termini di resa della lingua, sia in termini di appropriatezza terminologica. Ovviamente, nella scelta della soluzione MT da utilizzare è necessario tenere conto anche del prezzo. A titolo di esempio, tuttavia, i sistemi MT come Google Translate e quello di SDL hanno prezzi veramente abbordabili per qualsiasi tasca. Ovviamente, data l’importanza dei testi in questione, il post-editing dell’output della macchina di traduzione non potrà che essere un *full post-editing*, ossia una revisione integrale del testo tradotto dalla macchina tale per cui la qualità della traduzione dovrà essere, da tutti i punti di vista, equivalente a quella di un testo tradotto *ex-novo* a partire dal solo testo *source* (e quindi senza l’intervento a monte della MT). Questo potrebbe gettare qualche dubbio sull’uso della MT applicata alla traduzione brevettuale, ma la nostra esperienza suggerisce, invece, che tale approccio, se usato in modo consapevole, può aumentare decisamente la produttività dei linguisti e offrire testi con una qualità del tutto corrispondente alle aspettative del cliente. Ciò è almeno in parte attribuibile alle peculiarità dei testi brevettuali, nei quali sono ravvisabili un’elevata tecnicità lessicale, uno stile molto asciutto e impersonale e un’elevata ridondanza intra- e intertestuale. Pertanto, il testo brevettuale presenta tutte le caratteristiche per essere trattato in modo vantaggioso tramite l’uso combinato di traduzione automatica e traduzione assistita (assieme a glossari specifici per il mantenimento della coerenza terminologica), nell’ottica dell’aumento di produttività del linguista e del mantenimento della qualità finale della traduzione. Ovviamente, ai fini della qualità di una buona traduzione brevettuale, rivestono una notevole importanza la formazione e la bravura del traduttore, che soltanto la pratica traduttiva e la disposizione personale possono garantire. Fermo restando quanto sopra, la nostra esperienza suggerisce altresì che l’uso combinato di traduzione automatica e traduzione assistita permette di accorciare la curva di apprendimento dei traduttori brevettuali neofiti e di migliorarne anche la qualità di vita lavorativa. Abbiamo infatti dimostrato empiricamente, tramite un semplice esperimento, che, per uno stesso testo, l’uso della MT nella traduzione brevettuale impone al traduttore di battere a tastiera un numero di volte apprezzabilmente minore rispetto a una traduzione tradizionale. È noto, da questo punto di vista, quanto i problemi osteoarticolari rappresentino una croce per molti traduttori, e poter battere a tastiera un numero di volte minore, imponendo una minore rigidità alla postura, rappresenta un vantaggio oggettivo per il linguista professionista.

1. Panoramica sul mercato dei brevetti in Europa

Al momento della stesura di questo articolo, il mercato dei brevetti, e di conseguenza della traduzione brevettuale, in Europa è quanto mai florido e in costante crescita negli ultimi cinque anni (a parte una lieve flessione nel 2016): solo nel 2018 sono state presentate 147.317 domande di brevetto allo *European Patent Office*. E l'Italia, con 4.399 domande e una crescita dello 0,9% rispetto al 2017, si attesta al sesto posto per numero di richieste tra gli stati appartenenti all'EPO.

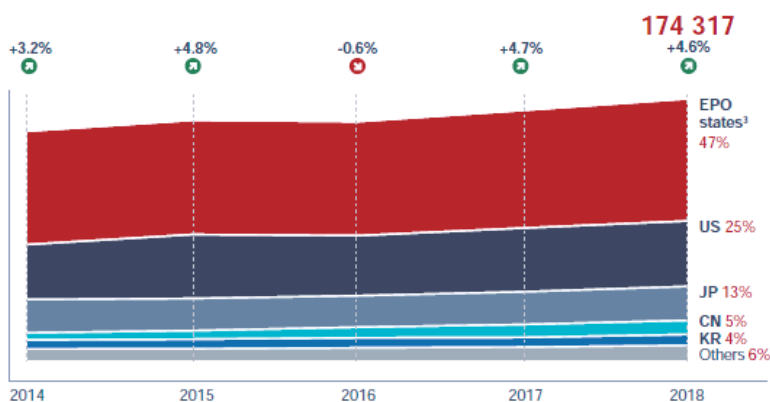


Figura 1. Tendenza delle domande di brevetto presentate negli ultimi cinque anni e suddivisione per aria geografica.

I settori della tecnica maggiormente attivi sono, come prevedibile, quelli che coinvolgono le nuove tecnologie, e quindi le biotecnologie, la comunicazione digitale, l'informatica, l'energia e i trasporti, e la maggiore presenza sul mercato viene registrata dalle grandi multinazionali (che rappresentano il 71% del mercato), con Siemens, Huawei e Samsung a guidare la classifica delle principali richiedenti nel 2018. Sulla base di un ampio campione rappresentativo delle domande di brevetto presentate all'EPO, emerge un dato significativo: nel 2018 una domanda su cinque è stata presentata da PMI o singoli inventori.

In questo scenario, e allo stato attuale della legislazione, per tutelare la proprietà intellettuale di un'invenzione nei diversi Stati industrializzati, una richiedente deve coinvolgere tutti gli Stati in cui desidera ottenere un brevetto; è facile quindi intuire come i costi di traduzione siano determinanti nella procedura di deposito.

Capita spesso che i costi da sostenere per la traduzione di un brevetto superino i costi per il deposito dello stesso; questo fattore, insieme all'assoluta

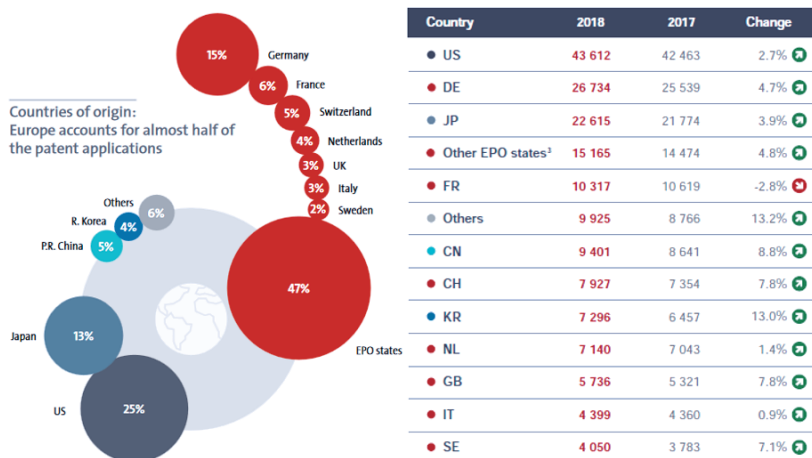


Figura 2. Dati relativi alla provenienza dei brevetti.

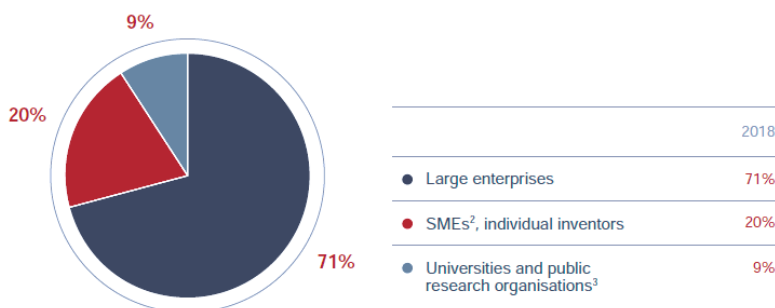


Figura 3. Dati rilasciati dall'EPO nel report annuale 2018.

necessità che la versione tradotta di una domanda di brevetto sia impeccabile per evitare il rischio di renderla nulla, rende difficile l'incontro tra domanda e offerta. Date le premesse, l'adozione della traduzione automatica può rivelarsi una soluzione *win-win*: l'LSP o il traduttore vedrà aumentare sensibilmente la propria produttività — la domanda di brevetto rientra in una tipologia testuale che si presta particolarmente alla traduzione automatica — sarà in grado di gestire volumi maggiori e di garantire tempi di consegna più brevi, senza tuttavia compromettere la qualità del proprio lavoro, e a sua volta il committente (nella maggior parte dei casi un intermediario) otterrà una tariffa più vantaggiosa e potrà beneficiare di una gestione del lavoro più fluida grazie soprattutto alla riduzione dei tempi che intercorrono tra l'assegnazione dell'incarico e la consegna del lavoro ultimato.

2. Uso della traduzione automatica e applicabilità alla traduzione brevettuale

Abbiamo visto che cos'è un brevetto e stabilito che l'approccio del post-editing della traduzione automatica è una soluzione vantaggiosa per tutte le parti coinvolte; ma quali sono i motivi che rendono il testo di una domanda di brevetto così adatto all'MPTE?

Per arrivare a questa risposta, è necessario fare qualche passo indietro e analizzare brevemente alcuni aspetti.

Autore e destinatario. Il brevetto viene redatto da un tecnico (l'inventore) per un tecnico o gruppo di tecnici (commissione dell'ufficio brevetti, esperti del ramo che vorranno implementare l'invenzione).

Linguaggio. Il linguaggio è quindi caratterizzato da una forte settorialità e da un elevato pragmatismo. Lo scopo meramente descrittivo e divulgativo della domanda di brevetto prevede il minor grado di interpretazione possibile e per questo motivo sono tollerate ridondanze terminologiche e di concetto purché venga raggiunto l'obiettivo (di conseguenza, il traduttore o post-editor non dovrà "perdere tempo" in elucubrazioni per trovare il giusto sinonimo per evitare di ripetere la stessa parola all'interno della stessa proposizione).

Struttura. Tutte le domande di brevetto seguono una struttura ben definita (Titolo, Descrizione della tecnica precedente, Scopo dell'invenzione, Sommario dell'invenzione, Descrizione dei disegni, Descrizione delle forme di realizzazione, Rivendicazioni, Tavole dei disegni); allo stesso modo la sintassi utilizzata è molto lineare (anche se al traduttore brevettuale capiterà spesso di trovarsi di fronte proposizioni da centinaia di parole).

Requisito di letterarietà della traduzione. Perché la domanda di brevetto abbia valore legale nello Stato in cui essa viene presentata, viene richiesta la massima aderenza al testo di partenza, in altre parole deve esserci una corrispondenza 1:1 tra testo originale e testo tradotto. In questo ambito traduttivo, un errore di omissione è considerato grave tanto quanto una traduzione errata di un termine, poiché potrebbe compromettere la validità della tutela: quello che c'è scritto da una parte deve essere scritto anche dall'altra, errori compresi!

Il pragmatismo del linguaggio e la rigidità della sintassi si adattano bene al tipo di elaborazione messo in opera dai motori di traduzione automatica, inoltre considerato l'obiettivo puramente informativo e il carattere stilistico

di questa tipologia che non ha nessuna “pretesa di bellezza” fanno sì che l’intervento del post–editor sia ridotto al minimo indispensabile (per dirla in parole povere, tutte le brutture della *Machine Translation* sono accettabili purché si ottenga un risultato corretto dal punto di vista linguistico e dei contenuti).

Vantaggi e limiti della traduzione automatica e benefici per il post–editor. I vantaggi che derivano dall’utilizzo della traduzione automatica sono molteplici e non si limitano all’aumento della produttività (che in ogni caso, per un traduttore esperto, si attesta tra le 500 e le 1500 parole tradotte all’ora, in base alla complessità degli argomenti trattati).

Uno dei vantaggi più apprezzabili in questo ambito è legato al requisito di massima aderenza al testo di partenza di cui abbiamo fatto cenno prima: soprattutto se si utilizzano motori di traduzione RBMT (*Rule–Based Machine Translation*) o SMT (*Statistical Machine Translation*), il post–editor sarà ragionevolmente sicuro che nell’output proposto dal motore di traduzione non verrà tralasciata alcuna informazione, sebbene, a seconda della coppia linguistica, le macchine statistiche possano talvolta incorrere in errori quali piccoli inserimenti od omissioni. La stessa cosa non vale del tutto se si utilizza un motore di traduzione neurale (NMT) che, nonostante sia in grado, grazie all’intelligenza artificiale e alle reti neurali, di elaborare un linguaggio molto simile al linguaggio di un essere umano, talvolta omette o rielabora arbitrariamente le informazioni presenti nel testo di partenza: l’output appare scorrevole e incredibilmente naturale, ma non è detto che sia sempre fedele all’originale; diventa quindi cruciale che il post–editor sia consapevole di quale tipo di traduzione automatica sia stata utilizzata, così da prendere i dovuti accorgimenti. Se di fronte a un output generato da RBMT o SMT, un post–editor esperto è in grado di valutare la correttezza della traduzione anche senza controllare continuamente il testo di partenza, con un output generato da NMT è sempre bene verificare con attenzione la corrispondenza tra traduzione e originale. Va detto però che i miglioramenti della NMT, come quella proposta da DeepL o da Google (che sin dal 2011 ha siglato una partnership proprio con l’EPO), sono evidenti e costanti.

Un altro dei vantaggi che caratterizzano questo approccio alla traduzione riguarda la sensibile riduzione di alcune tipologie di errore tipiche della *Human Translation*, primo fra tutti l’errore di battitura: è evidente che riducendo drasticamente il numero di battute necessarie per la produzione di un testo, si riducono di conseguenza anche i refusi.

La traduzione automatica consente inoltre di garantire un elevato grado di coerenza terminologica: la macchina tenderà a proporre per uno stesso termine sempre la stessa traduzione, e anche laddove la traduzione fosse

errata, la macchina tenderà a sbagliare nello stesso modo (dando quindi al post-editor la possibilità di correggere il tiro con una semplice sostituzione globale).

Ovviamente, quando si parla di *Machine Translation* e brevetti non sono tutte rose e fiori, oltre ai limiti propri della traduzione automatica, ve ne sono alcuni legati esclusivamente alla traduzione brevettuale. Nel caso dei testi brevettuali, accade spesso che il carattere inventivo dell'oggetto del brevetto riguardi anche la terminologia a esso associata, ed è evidente che il motore di traduzione non sarà in grado di coniare di propria sponte un termine nuovo e l'intervento del post-editor con la sua immaginazione diventerà quanto mai necessario.

Come abbiamo già visto in precedenza, normalmente l'autore di un brevetto non è un professionista della scrittura e non è raro trovare all'interno dei documenti veri e propri strafalcioni linguistici, che la macchina non è ovviamente in grado di individuare come tali, producendo quindi un output errato.

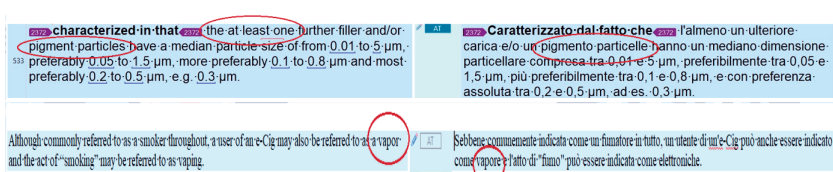


Figura 4. Dati rilasciati dall'EPO nel report annuale 2018.

E infine, le domande di brevetto rientrano in una tipologia testuale altamente specialistica. Fino a che punto il post-editor potrà fidarsi dell'output della macchina? Non c'è una risposta a questa domanda, è proprio qui che entra in gioco l'imprescindibile giudizio dell'essere umano che garantirà la bontà del lavoro.

Riferimenti sitografici

[http://documents.epo.org/projects/babylon/eponet.nsf/o/68C67CCD2D7FA5D3C12583B4004CCA6A/\\$File/epo_annual_report_2018_infographic_en.pdf](http://documents.epo.org/projects/babylon/eponet.nsf/o/68C67CCD2D7FA5D3C12583B4004CCA6A/$File/epo_annual_report_2018_infographic_en.pdf).

<https://www.epo.org/about-us/annual-reports-statistics/annual-report/2018.html>.

Traduzione automatica

A che punto siamo?

J. PORSIEL, A. BURCHARDT*

Andiamo subito al punto: sistemi di traduzione automatica diversi — siano motori basati su regole o sintagmi, ibridi o neurali — producono risultati diversi. In base alla coppia linguistica e/o al settore di specializzazione ogni approccio propone la “sua” soluzione, che può essere più o meno utilizzabile. A fare la differenza sono prima di tutto la qualità e la tipologia del testo di partenza. È impossibile fare un confronto diretto tra i vari approcci, pertanto dichiarazioni come « Il sistema X è meglio del sistema Y » vanno sempre prese con la dovuta cautela.

La traduzione automatica ha vissuto molti alti e bassi fin dai primi sviluppi negli anni Trenta. I suoi detrattori ricordano spesso come da quasi ottanta anni gli scienziati decantino i passi in avanti della tecnologia affermando: « Tra cinque anni non avremo più bisogno di alcun traduttore ». Ma dopo una battuta d’arresto, da un paio d’anni la traduzione automatica è di nuovo sulla bocca di tutti. È merito, tra l’altro, della capacità di calcolo e della potenza dei processori, che non rappresentano più un limite né tecnico né finanziario, e dei dati bilingue, ormai disponibili nella quantità necessaria per la creazione dei sistemi e talvolta anche nella qualità richiesta.

Eppure le opinioni sull’argomento non potrebbero essere più discordanti. Per alcuni la traduzione automatica non produce risultati utili, pertanto è irrilevante per il lavoro del traduttore; e di conseguenza la ricerca in questo campo è fine a sé stessa. Questa tesi è supportata dalla qualità — in parte ancora scadente — delle traduzioni e dal fatto che molti degli utenti potenziali non dispongono della tecnologia necessaria né di personale qualificato per utilizzarla con profitto. A contraddire questi fatti ci sono i circa

* Dr. Aljoscha Burchardt è ricercatore senior e responsabile di laboratorio del Language Technology Lab Centro di ricerca tedesco sull’intelligenza artificiale (Deutsches Forschungszentrum für Künstliche Intelligenz, DFKI GmbH). Jörg Porsiel è laureato in Lingue straniere con specializzazione in giurisprudenza presso l’Istituto di Traduzione e Interpretariato (IÜD) dell’Università Ruprecht Karl di Heidelberg. Si occupa di di lingua controllata per la Volkswagen AG a Wolfsburg ed è responsabile del dipartimento aziendale di traduzione automatica.

100 miliardi di parole tradotte automaticamente ogni giorno da Google¹. E le campagne di marketing che promettono una qualità sempre più simile a quella delle traduzioni umane². Esiste infine una terza opinione, non meno diffusa, per cui in generale non vale la pena investire nella lingua e nella traduzione, né dedicare loro una particolare attenzione. Forse i fautori di questo gruppo sono quelli che avrebbero più bisogno di ricevere chiarimenti sull'argomento.

1. Un breve excursus storico

La storia della traduzione automatica o, in inglese, *Machine Translation* (MT)³ inizia negli anni Trenta del XX secolo, ovvero oltre ottanta anni fa. La tecnologia suscitò per la prima volta un grande interesse, anche da parte del mondo scientifico, verso la fine della seconda guerra mondiale, quando gli inglesi e gli americani tentavano di tradurre quasi simultaneamente i messaggi radio tedeschi. Durante la guerra fredda le attività di ricerca si intensificarono ulteriormente, culminando nell'esperimento congiunto della Georgetown University e della IBM nel 1954, i cui risultati alimentarono previsioni euforiche. Gli scienziati e l'opinione pubblica erano convinti che mancasse pochissimo a una rivoluzione epocale per cui le difficoltà di comunicazione o di traduzione sarebbero diventate un lontano ricordo: la cosiddetta traduzione automatica basata su regole (*Rule-Based Machine Translation*, RBMT) sembrava la soluzione a tutti i problemi linguistici. Forse questo è il motivo per cui oggi abbiamo a che fare con *science fiction*, più che con *science facts*.

Ma la doccia fredda arrivò già nel 1966 con il rapporto ALPAC, che concludeva «there is no immediate or predictable prospect of useful machine translation»⁴. Di conseguenza per decenni le attività di ricerca e soprattutto i finanziamenti nel campo della traduzione automatica hanno subito una battuta d'arresto. Negli anni Ottanta è stata la volta di nuovi approcci, tra cui la traduzione automatica statistica (*Statistical Machine Translation*, SMT). Ma dato che il cambiamento radicale non arrivò nemmeno con questi motori, che pur migliorando alcune cose, ne peggioravano altre, fu logico tentare la strada di un approccio ibrido. Si presumeva, infatti, che combinando RBMT

1. <https://blog.google/products/translate/ten-years-of-google-translate/>.

2. <https://research.googleblog.com/2016/09/a-neural-network-for-machine.html>.

3. Poiché la versione originale di questo libro è bilingue, ci riserviamo il diritto di utilizzare le abbreviazioni comuni in inglese nei testi in italiano, laddove ciò sia ragionevole.

4. J.R. PIERCE, J.B. CARROLL *et al.*, *Language and Machines: Computers in Translation and Linguistics*, ALPAC report, National Academy of Sciences, National Research Council, Washington, DC 1966: <https://www.nap.edu/read/9547/chapter/1>.