

Nástroj DeepL pro překlady akademických textů do angličtiny

Svět a zvláště Evropa jsou i přes proces evropské integrace stále heterogenní, protože jsou multilingvální. To je samozřejmě pro Evropany velký handicap, který vynikne při srovnání s velkými monolingválními populacemi v USA, Číně nebo Rusku. Částečným řešením je výuka cizích jazyků, především angličtiny, ale i přesto naprostá většina Evropanů je schopna komunikovat jen ve svém rodném jazyce. Určitým řešením je vývoj strojových textových překladačů, které posléze vystřídají překladače hlasové. Dosažení cíle, individuálního hlasového překladače pro každého Evropana, bude bezesporu velkým evropským, ale i světovým vývojovým mezníkem. Cesta za úspěchem je ovšem zdoluhavá a velmi obtížná. Zatímco strojové textové překladače již existují a stále se zlepšují, u hlasových překladačů jsme teprve v nukleárním stádiu. Jak stručně vypadá současná situace v textových překladačích?

Zřejmě nejznámější a nejpoužívanější je překladač od Google a potom Bing Translator od Microsoftu. V obou případech se jedná o frázové, statistické strojové překlady založené na technologii srovnávání a obrovských knihovnách dat. I když jsou oba tyto nástroje významnou pomůckou při překladech, mají bohužel do dokonalosti daleko. Ukazuje se, že budoucnost bude patřit překladačům založeným na neuronových sítích.

V roce 2017 se na trhu strojového překladače objevila a postupně stále více trhem otřásá německá firma Linguee se svým překladačem DeepL¹ – efektivním a spolehlivým nástrojem pro akademické překlady. Tento překladač je založen na skvěle vytrénované umělé inteligenci, tvořené neuronovou sítí disponující obrovskou sadou miliard ověřených překladů. Jeho ambicí od počátku bylo co do úrovně překladače překonat překladač Googlu. Významným pokrokem oproti předchozím statistickým strojovým překladačům je pohled překladače DeepL na celou větu a zohlednění jejího kontextu. Nelze ani opomenout také podporu aktivní práce s překladačem ve formě volby z široké nabídky variant překladů jednotlivých slov. V březnu letošního roku se překladač DeepL „naučil“ také česky, takže v současnosti obsahuje 24 jazyků s možností libovolného párování. Zatímco u překladače Googlu kvalita překladače závisí na vybraném páru, např. angličtina – němčina je výrazně lepší než třeba čeština – maďarština, u DeepL to neplatí.

Pro akademické pracovníky je samozřejmě nejdůležitější překlad čeština – angličtina, a proto jsme požádali Čechoameričana a redaktora časopisu *Journal of Separation Science*, prof. Františka Švece, o názor a zkušenost. Jeho český text o 600 slovech přeložil DeepL tak, že bylo potřeba zkorigovat jen tři slova. Český roubík byl přeložen jako „gag“, správně mělo být „rod“ (roubík vznikl právě českou mutací slova „rod“). Kvalitu celkového překladače zhodnotil prof. Švec slovy: „vygenerovaný text byl výrazně lepší než mnohé anglické rukopisy, které dostávám od autorů z celého světa, nevyjímaje i některé rodilé anglické mluvčí.“

DeepL vytváří až neuvěřitelně plynule znějící kvalitní překlady, doporučované k následné dodatečné kontrole a úpravě kvalifikovaným překladačem. Toho však často u akademických textů spolehlivě zastoupí sám zkušený autor odborného textu, znalý vhodné terminologie osvojené soustavným studiem odborné literatury a vlastní publikační činností. Při využívání překladače DeepL je třeba zdůraznit skutečnost, že veškeré texty a opravy vložené do volně přístupné verze DeepL se stávají součástí trénovací množiny neuronové sítě. Je proto třeba dbát určité opatrnosti při práci s nepublikovanými výzkumnými daty. V placené verzi DeepL Pro jsou data ukládána pouze na nezbytně nutnou dobu jejich zpracování a poté jsou ze systému odstraněna. Nicméně i pro tuto verzi je doporučováno neuvádět žádná osobní data a překládané texty vhodně anonymizovat.

Dosavadní testování strojového překladače DeepL na chemických textech páru čeština – angličtina vyvolává nadšení. Na druhé straně hrozí i nedokonalosti. Pro vědecký jazyk je do určité míry charakteristické zavedené stereotypní vyjadřování s omezenou mírou kreativity. Problém může představovat např. možná nejednotná terminologie v různých větech textu nebo vynechávání slov nebo slovních spojení, u kterých nenalezl překladač vhodnou shodu.

Co znamená překladač DeepL a pokrok v překladačelských technologiích pro časopis *Chemické listy* a obecně pro všechna národní vědecká periodika? Evidentně mnoho. Především vyřešení zásadního dilematu: národní chemický časopis publikující převážně česky a slovensky versus co není v dnešním vědeckém světě publikováno v angličtině, tak neexistuje! Logickým východiskem je publikovat duální verze článků, v češtině a v angličtině, právě pomocí nástroje DeepL Pro a budoucích dokonalejších verzích. Pro profesionální korektory je obvykle velmi obtížné upravovat pouze anglický text. Souběžná existence kvalitní české předlohy textu pro anglický překlad tak vytváří optimální prostředí pro využití kvalitního chytrého překladače a následnou úpravu přímo samotnými autory textů. Očekávané dopady překladů do angličtiny jsou pro *Chemické listy* celkem logické a jednoznačné: zpřístupnění poznatků a informací, a to i těch jinak regionálně vázaných právě na české a slovenské vědecké a jazykové prostředí, pro čtenáře z celého světa, navýšení citovanosti a zlepšení jejich bibliometrických ukazatelů. Snad se to promítne i do většího zájmu publikovat v tomto časopise.

Redakce *Chemických listů* začne postupně publikovat duální česko-anglické verze článků a věří, že se to u autorů i čtenářů setká s kladným ohlasem. Zároveň uvítá sdělení vlastních zkušeností, doporučení a ohlasů s praktickým používáním překladače DeepL.

Bohumil Kratochvíl a Zdeněk Bělohlay

LITERATURA

1. <https://www.deepl.com/translator>, staženo 6. 6. 2021.