

LITHME WG1-WG7 Joint workshop – Bridging the gap between technology and professionals

*(Eötvös Loránd Tudományegyetem Bölcsészettudományi Kar,
Nyelvi Közvetítés Intézete, Fordító- és Tolmacsképző Tanszék,
Budapest, 2023. augusztus 28–30.)*

Van Waarden Franciska

ELTE BTK Nyelvi Közvetítés Intézete, Fordító- és Tolmacsképző Tanszék
ELTE BTK Nyelvtudományi Doktori Iskola, Fordítástudományi Doktori Program
E-mail: van.waarden.franciska@btk.elte.hu
<https://orcid.org/0009-0000-2155-3884>

Az *Európai Tudományos és Technológiai Együttműködés* (European Cooperation in Science and Technology, COST) 2020–2024. közt meghirdetett LITHME programja (Language in the Human-Machine Era) a nyelv, az ember és a technika kapcsolódási pontjait kutató szakembereket fogja össze. Az érdeklődők nyolc különböző munkacsoportba csatlakozhatnak és ezáltal lehetőségük nyílik a határon átnyúló közös interdiszciplináris tudományos munkára. 2023 augusztusában az 1-es számú (*Nyelvtechnológia*) és a 7-es számú (*Nyelvi munka, nyelvi szakemberek*) munkacsoportjai tartottak közös workshopot, amelynek az ELTE Fordító- és Tolmacsképző Tanszéke adott otthont. A projekt egyik fő célja, hogy közelebb hozza egymáshoz a nyelvészeket és a technológusokat. A workshop keretein belül mindkét táborhoz tartozó résztvevőknek lehetőségük nyílt bemutatni új kutatási eredményeiket.

A háromnapos rendezvényt Barta Péter, az ELTE BTK Nyelvi Közvetítés Intézetének leköszönő intézetigazgatója nyitotta meg. A LITHME tagjait Rui Sousa Silva, az 1-es Munkacsoport vezetője és Seresi Márta, a 7-es Munkacsoport vezetője köszöntötte.

Az első szekcióban (*Gépek és az őket használó emberek*) elsőként Ilan Kernerman előadását hallhattuk, aki az ember részvételének fontosságát hangsúlyozta az újfajta mesterséges intelligencia és a nagyméretű nyelvi modellek kialakításában. Ezt követően Eszenyi Réka és Per van der Wijst mutatták be közös kérdőíven alapuló kutatásukat, amelyben azt vizsgálták, hogy a neurális gépi fordításnál

szükség van-e utószekeresztésre, és ha igen, abban az esetben pozitív vagy negatív változásról beszélünk-e, és mennyire befolyásolja a kreativitás, valamint a monotonitás arányát a fordítási munka során. A szekciót Anna Setkowicz-Ryska előadása zárta, amelyben lengyelországi perspektívából ismertette a gépi fordítói kompetencia szükségességét a jogászok esetében. A plenáris előadást Cristina España-Bonet, a Németországi Mesterséges Intelligenciai Kutatóközpont (DFKI) munkatársa tartotta, szemléletes példákon keresztül bemutatva, milyen emberi előítéletek, elfogultságok jelennek meg a többnyelvű korpuszokban, gépi fordításban.

A nap második felében a munkacsoportok vezetői zártkörű megbeszélést tartottak a készülő konferenciakötettel kapcsolatban, míg a LITHME többi tagja részt vehetett egy networking eseményen, ahol az előadóknak lehetőségük nyílt személyes kapcsolatokat teremteni.

A rendezvény második napját (*Gépek és minőség*) Maarit Koponen, a 7-es munkacsoport vezetője nyitotta meg. Yang Győző, a Nyelvtudományi Kutatóintézet munkatársa plenáris előadásában ismertette az új magyar BERT, GPT-2, mBART nyelvmodelleket, amelyek többnyelvű kódolásra és dekódolásra lesznek majd képesek, továbbá a ChatGPT fordításra való alkalmazhatóságát járta körbe. Ezután Carlos Manuel Hidalgo Ternero mutatta be a gApp és a Paidiom rendszereket, amelyeknek elsődleges célja, hogy az idiómák és egyéb bevett szókapcsolatok gépi fordítását ún. transzlemmatizálás segítségével zökkenőmentesebbé tegye. Ezt követően Robin Edina és Seidl-Pécs Olívia mutatta be kutatási projektjének első eredményeit, amelynek keretében az Európai Bizottság által fordított szövegeket, lektorált változataikkal vetettek össze a DGT hibatipológiája alapján. Felermino Ali válaszolt az olyan nyelvek gépi fordításának nehézségeit, ahol a nagymértékű nyelvi variációk miatt a helyesírás egységesítése is komoly problémát okoz – mindezt egy makuva-portugál nyelvű párhuzamos korpusz példáin keresztül ismertette. A szekció utolsó előadásában Maria Onyshchuk mutatta be a TCCCC (Tactical Combat Casualty Care Course, azaz Taktikai harci sérültellátási tanfolyam) projektjét, amely egy ukrán-lengyel-angol nyelvű terminológiai adatbázisban vannak összegyűjtve a harci sérültek ellátásához szükséges kifejezések, amelyeken újabb korpusznyelvészeti vizsgálatok váltak elérhetővé az AntConc eszköz segítségével.

A workshop második napjának szakmai programját kerekasztal-beszélgetés zárta, amely a technológiai fordulat témáját tárgyalta Seresi Márta, az ELTE Fordító- és Tolmácsképző Tanszékének egyetemi adjunktusa moderálásában. A kerekasztal további résztvevői Carlos Manuel Hidalgo Ternero, a Malagai Egyetem tanára, Maarit Koponen, a Kelet-finnországi Egyetem tanára, Susanna Nocchi, a dublini Műszaki Egyetem tanára, valamint Robin Edina, az ELTE Fordító- és Tolmácsképző Tanszékének habilitált egyetemi docense voltak. A résztvevők megállapították, hogy a technológiai fordulat drámai változást hozott a nyelvészetbe, megváltoztatva az addigi kutatási módszereket, és új kompetenciák alkalmazását igényelte. Seresi Márta, Robin Edina és Susanna Nocchi szerint mindez, de főként a gépi fordítás robbanásszerű fejlődése a fordítás oktatására is nagy hatással van,

hiszen a fordításoktatóknak nagyobb figyelmet kell fektetni arra, hogy a gépi fordítást tiltják, türik, vagy pedig támogatják az oktatásban. Maarit Koponen szerint egy súlyos probléma, hogy például az audiovizuális fordításban a profi, képzett fordítók helyét átvehetik a gyengébb nyelvi és fordítói kompetenciákkal rendelkező egyének, akik erőteljesebben támaszkodnak a gépi fordításra. Carlos Manuel Hidalgo Ternero szerint nagyobb teret kellene adni a nyelvészeknek a gépi fordítás fejlesztésében, nem kizárólag informatikai háttérrel rendelkező szakemberek feladata a jövő gépesített nyelvi rendszereinek kialakítása. A keddi nap zárásaként az előadók fakultatív budapesti városnéző sétán, és közös networking vacsorán vehettek részt.

A workshop utolsó napján (*Gépek az emberi tanulási folyamatban*) elsőként Olesea Bodean-Vozian, a Moldovai Állami Egyetem tanára ismertette az intézményükben folyó tolmácsképzést, valamint, hogy a koronavírus-járvány okozta korlátozások milyen technológiai változásokat hoztak az oktatási módszerekbe. Ezután Susanna Nocchi ismertette a #LanguagesPathways projektet, amely a virtuális valóság segítségével tenné szélesebb körben elérhetővé az olasz nyelv tanításának lehetőségét online, interaktív módon. Elif Sari, a törökországi Trabzon Műszaki Egyetem tanára mutatta be kutatásának eredményeit, amelynek keretein belül az automatizált visszajelzések pszichológiai hatásait vizsgálta az angol nyelvet tanulók körében. A szekciót Yolanda Morató előadása zárta, amelyben a kreatív szövegtípusok fordításában alkalmazható speech-to-text beszédfelismerő rendszerek használhatóságát járta körül.

A háromnapos rendezvényt kerekasztal-beszélgetés zárta, amely azt a kérdést igyekezett megválaszolni, hogy tíz év múlva mi várható a nyelv és technológia viszonylatában. A beszélgetés résztvevői Rui Sousa Silva, a Portói Egyetem tanára, Eszenyi Réka, az ELTE Fordító- és Tolmácsképző Tanszékének egyetemi adjunktusa, Yang Győző, a Nyelvtudományi Kutatóintézet tudományos munkatársa, Ilan Kernerman, a Lexicala by K Dictionaries alapítója és ügyvezető igazgatója, Antonio Pareja az Alcalá Egyetem adjunktusa, valamint Henrique Lopes Cardoso, a Portói Egyetem kutatója. A résztvevők véleménye megoszlott abban a kérdésben, hogy tíz év múlva a nyelv vagy a technológiáé lesz-e a vezető szerep, abban azonban egyetértettek, hogy az új technológiák szabályozására nagyobb hangsúlyt kell helyezni. Továbbá úgy vélik, a gépek és rendszerek az elkövetkezendő tíz évben még nem, de húsz-huszonöt éven belül már képesek lehetnek arra, hogy értsék és reagáljanak az emberi nyelvre.

A workshop végén Maarit Koponen és Rui Sousa Silva, a 7-es, illetve az 1-es Munkacsoport vezetői fejezték ki reményüket a további interdiszciplináris együttműködésre a COST (European Cooperation in Science and Technology) keretein belül.