**Az Országos Széchényi Könyvtár és a Nyelvtudományi Kutatóközpont innovációs együttműködése**

**Az Országos Széchényi Könyvtár (OSZK) és az Eötvös Loránd Kutatási Hálózat Nyelvtudományi Kutatóközpontja (NYTK) 2023. június 21-én aláírt megállapodása egy olyan, a mesterséges intelligencia fejlesztését célzó innovációs együttműködés kezdete, amely nagy nyelvi modellek építését és nyelvfeldolgozó eszközök létrehozását állítja a középpontba.**

Az olvasásnak, mint mindenkor, a jövőben is meghatározó szerepe lesz az ismeretek, a műveltség megszerzésében és a kommunikációban; a könyvtári tevékenység fókuszában az írásban rögzített tudás rendelkezésre bocsátása marad. A megállapodás értelmében az OSZK a Nyelvtudományi Kutatóközpontnak átadja a [Magyar Elektronikus Könyvtárban (MEK)](https://mek.oszk.hu/) és az [Elektronikus Periodika Archívum és Adatbázisban (EPA)](https://epa.oszk.hu/) elérhető szöveges dokumentumokat, a webarchívumok szöveges állományait, valamint a [dHUpla](https://dhupla.hu/) adatbázisban szereplő anyagokat.

A kutatóközpont ezeket a szövegeket felhasználva nagy nyelvi modelleket épít, amelyek finomhangolásával számos nyelvfeldolgozó eszközt lehet fejleszteni. (Ilyenek találhatók többek között a ChatGPT mögött is.) A kutatóközpontnak már van tapasztalata a nyelvi modellek fejlesztése terén; a nevéhez fűződik a Puli-GPT, amely csak magyar nyelvű szöveges anyagokon tanult, és a ChatGPT-hez hasonló funkcionalitással bír: <https://juniper.nytud.hu/demo/puli>.

Az NYTK a nemzeti könyvtár szövegei alapján létrehozott eszközöket az OSZK számára is hozzáférhetővé teszi. A nyelvfeldolgozó eszközök segítségével egyebek mellett a nyelvfelismerés és a webarchiválási tevékenység keretében keletkezett hatalmas mennyiségű tartalom feldolgozása is automatizálhatóvá válik.

Emellett a nemzeti könyvtár és az NYTK kooperációban dolgozza fel és teszi elérhetővé a Magyar Nagylexikon tizenkilenc kötetének anyagát. A terjedelmes dokumentum egyrészt olvasható és kereshető lesz egy webes lekérdezőfelületen a felhasználók számára, másrészt származtatott erőforrások készülnek belőle, amelyeket a mesterségesintelligencia-kutatással foglalkozó szakemberek kiválóan tudnak majd hasznosítani egyéb eszközök fejlesztéséhez.

További információ a sajtó képviselőinek: oszkpress@oszk.hu.