**Gépi kézírásfelismerő modell készült az Országos Széchényi Könyvtárban**

**Megjelent az első publikus magyar nyelvű számítógépes kézírásfelismerő modell. Az ún. HTR- (Handwritten Text Recognition-) modellt az Országos Széchényi Könyvtár Digitális Bölcsészeti Központja (OSZK DBK) hozta létre, és mindenki számára elérhetővé tette a Transkribus nevű szoftver komponenseként. Alapját Kiss József, a 19–20. század fordulóján élt magyar költő, a *Nyugat* előfutáraként számontartott *A Hét* című hetilap szerkesztőjének szakmai és személyes levelezése képezi.**

A kézírásfelismerést mesterséges intelligencián alapuló algoritmus végzi. Ezt először be kelltanítani, hogy a különböző kézírásokat felismerje, majd a megadott minták alapján felépít egy modellt, amellyel képessé válik arra, hogy egy általa ismeretlen, még nem látott kézírásképet is értelmezzen. Minél többféle anyagon tanítjuk, annál jobban teljesít különböző szövegeken. A jelenlegi modell először Kiss József írásán, majd a levelezőpartnerek vegyes kézírásán tanult.

Az eddig felhasznált kéziratok a Petőfi Irodalmi Múzeumban (PIM) találhatók, a tanulásra felhasznált szavak száma körülbelül 75 ezer. A dokumentumok között szerepelnek borítékok, képeslapok, hagyományos és fejléces levelek, névjegykártyák is. A levélírók Kiss József és családja, illetve a századforduló írói, újságírói és művészei, mint Ady Endre, Móricz Zsigmond vagy Tömörkény István. Ez összesen 300 darab változó hosszúságú és minőségű levelet jelent, amelyeket a DBK folyamatosan közzétesz az olvasók számára a [dhupla.hu/collection/kiss-jozsef-levelezes](https://dhupla.hu/collection/kiss-jozsef-levelezes) oldalon. A levélváltások további kéziratai jelenleg is feldolgozás alatt állnak az OSZK-ban és a PIM-ben.

A modell jelenleg 9,19 hibaszázalékkal működik, ami azt jelenti, hogy a projektben közel 90%-os pontossággal meg tudja állapítani, milyen karakterekből áll a szöveg.





A különböző magyar nyelvű projekteknek, amelyek automatikus kézírásfelismertetést használnak, a jövőben össze kell fogniuk, hogy a saját szövegkorpuszokon betanított modelleket egymásba építve egyre általánosabb érvényű eszköz jöjjön létre a magyar kéziratos források digitalizációjához. Közös érdekünk, hogy e közgyűjteményekben rejtőző kulturális kincsek a digitális térben szövegként hozzáférhetővé, olvashatóvá, kereshetővé, számítógépes eszközökkel is feldolgozhatóvá, kutathatóvá váljanak, ahogyan a dhupla.hu oldalon közzétett szövegkiadások, illetve a hozzájuk tartozó kreatív tartalmak esetében is láthatjuk. A most nyilvánosságra hozott első magyar kézírásfelismerő modell ennek a folyamatnak fontos mérföldköve.

Az új magyar kézírásfelismerő modell megtalálható a Transkribus asztali alkalmazásban és a webes felületen is ([https://readcoop.eu/model/hungarian/)](https://readcoop.eu/model/hungarian/).

További információ a sajtó képviselői számára: press@oszk.hu.