

Mihály Eszter – Micsik András: Szerkesztői környezet TEI alapú szövegkiadásokhoz

(2023. április 13., A terem)

Az OSZK Digitális Bölcsészeti Központ által fejlesztett új platform, a dHUpla (Digital Humanities Platform – dhupla.hu) elsősorban digitális szövegkiadások publikálására jött létre. Ennek háttérében a szövegkorpuszok előállításához egy teljes szerkesztőségi rendszer kialakítására volt szükség, amely egyrészt felhasználóbarát módon teszi lehetővé a szerkesztést, másrészt funkcióival támogatja a szövegek jelölőnyelvi kódolását. Ezt a szerkesztői felületet a SZTAKI és az OSZK közösen fejleszti egy XML editor bővítményeként.

A szövegkódolás alapja a TEI (Text Encoding Initiative), a framework e nemzetközi ajánlás bonyolult konstrukcióinak bevitelét többek között kontextus-menüvel, beszúrható mintákkal, Schematron-validációval segíti. A fejlesztésben központi szerepet játszanak a szöveg feldolgozását támogató további eszközök is: névterekkel, adatbázisokkal való összekapcsolás, MI-alapú névelemfelismerés, valamint különböző automatizált műveletek, úgy mint PDF- és képkonverzió, vagy adatvizualizáció támogatása. A manuálisan és a gép által végezhető részfolyamatok minden esetben kiegészítik egymást, megteremtve ezzel egy minőségi, a digitális közeg adta lehetőségeket kiaknázó szövegkiadási módszert.