



HTR, azaz a kézírásfelismerítés

Szűcs Kata
szucs.kata@oszk.hu

2023.05.21.



Tartalom



- *Bevezető gondolatok*
- *Mi az a dHUpIa?*
- *Workflow*
- *Mi az a HTR?*
- *Az első nyilvános magyar nyelvű HTR modell*
- *Próbáljuk ki!*



Digitális Bölcsészeti Központ

<https://dhupla.hu/> – Digital Humanities Platform

A Digitális Bölcsészeti Platform célja digitális szövegkiadások és a szövegfeldolgozáson alapuló kreatív tartalmak közzététele.



dhupla

digital humanities platform /// digitális bölcsészeti platform



szövegkiadások entitástár kreatív a dhupláról

Keresés

Szövegkiadások



Digitális kritikai kiadás

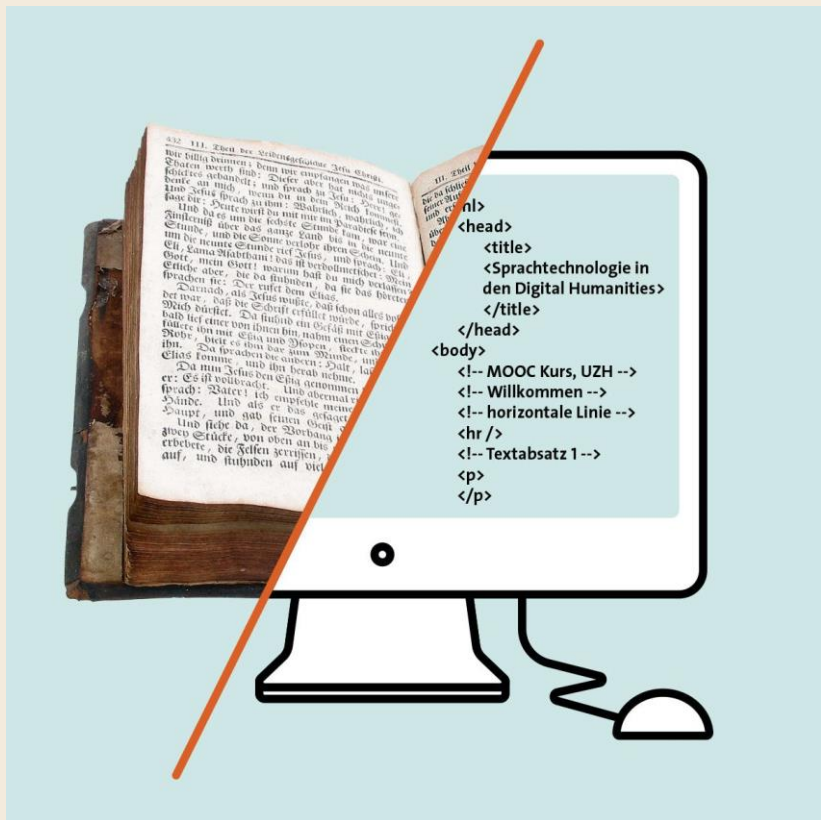
Mórícz Zsigmond levelezése (1892–1913)

2.0 verzió

Az alap kutatás célja Mórícz Zsigmond levelezésének (1892–1913) forrásfeltárása és közzététele digitális kritikai kiadás formájában. A projekt az NKFIH támogatásával valósult meg.

Digitális forráskiadás

Mórícz Zsigmond: Tükör (1914–1917)



<https://dhupla.hu/collection/>

dhupla

digital humanities platform /// digitális bölcsészeti platform



[szövegkiadások](#)

[entitástár](#)

[kreatív](#)

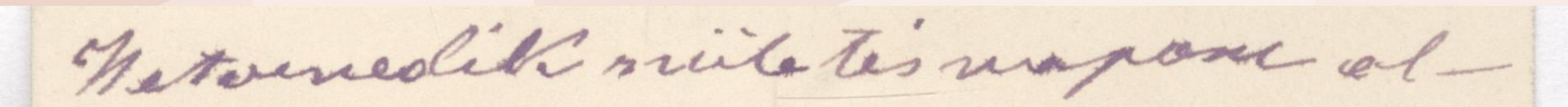
[a dhupláról](#)

Keresés

[Kiss József levelezése](#)

Blaha Lujza – Kiss József (1920)

⚠ Nézetváltáshoz kattints egy sorra vagy egy képre!



Hetvenedik születés napom al-

kalmából jóleső figyelmeért^[1],

hálás köszönetét küldi,

régi jó híve

Blaha Lujza



AZ NKFI ALAPBÓL
MEGVALÓSULÓ
PROJEKT

[Impresszum](#)

[Adatvédelem](#)

Workflow



Git

Workflow



Mi az a **Transkribus**?

A Transkribus történelmi dokumentumok, digitalizált kéziratok filológiai központú kezelésére alkalmas átfogó, nyelvtől független platform, amely bárholnan elérhető.

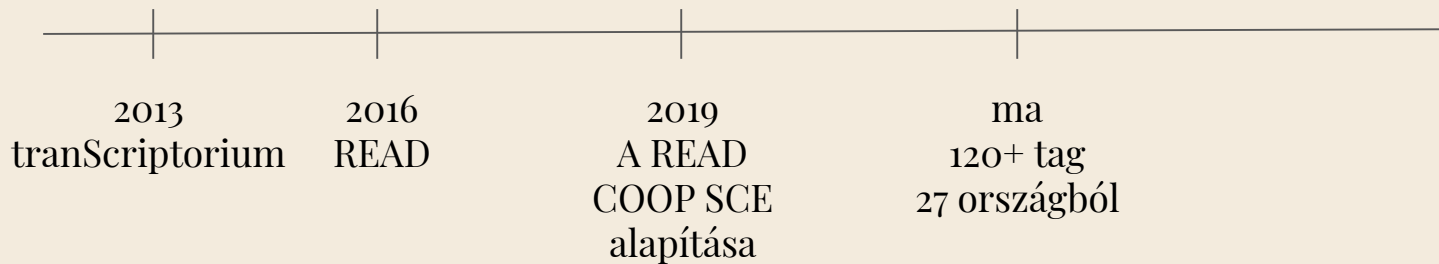


Kik állnak a **Transkribus** mögött?

READ-COOP SCE

Recognition and Enrichment of Archival Documents

Európai szövetkezet a Transkribus platform fenntartására és továbbfejlesztésére.



Mit tesz lehetővé a **Transkribus**?



**Kézzel írott és nyomtatott
dokumentumok automatikus
átírása**



**MI alapú modellek
tanítása**



Együttműködés



**Keresés a
dokumentumokban
hatékony eszközökkel**



**A dokumentumok
szerkezetének és
tartalmának
címkézése/tagelése**



**Dokumentumok
exportálása különböző
formátumokban**

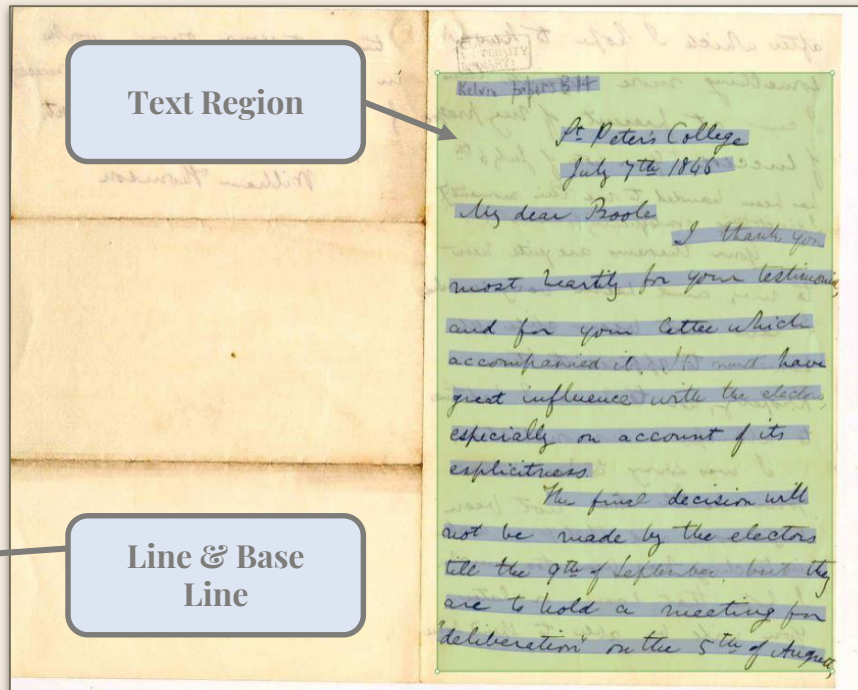
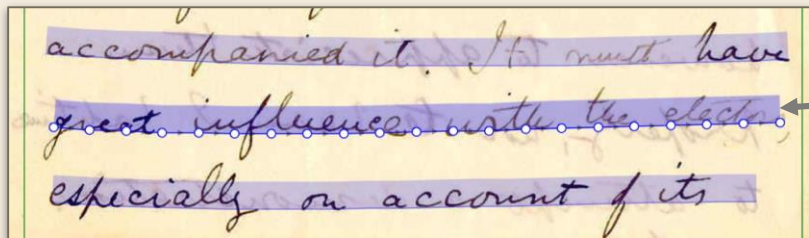
Workflow



Transkribus Workflow

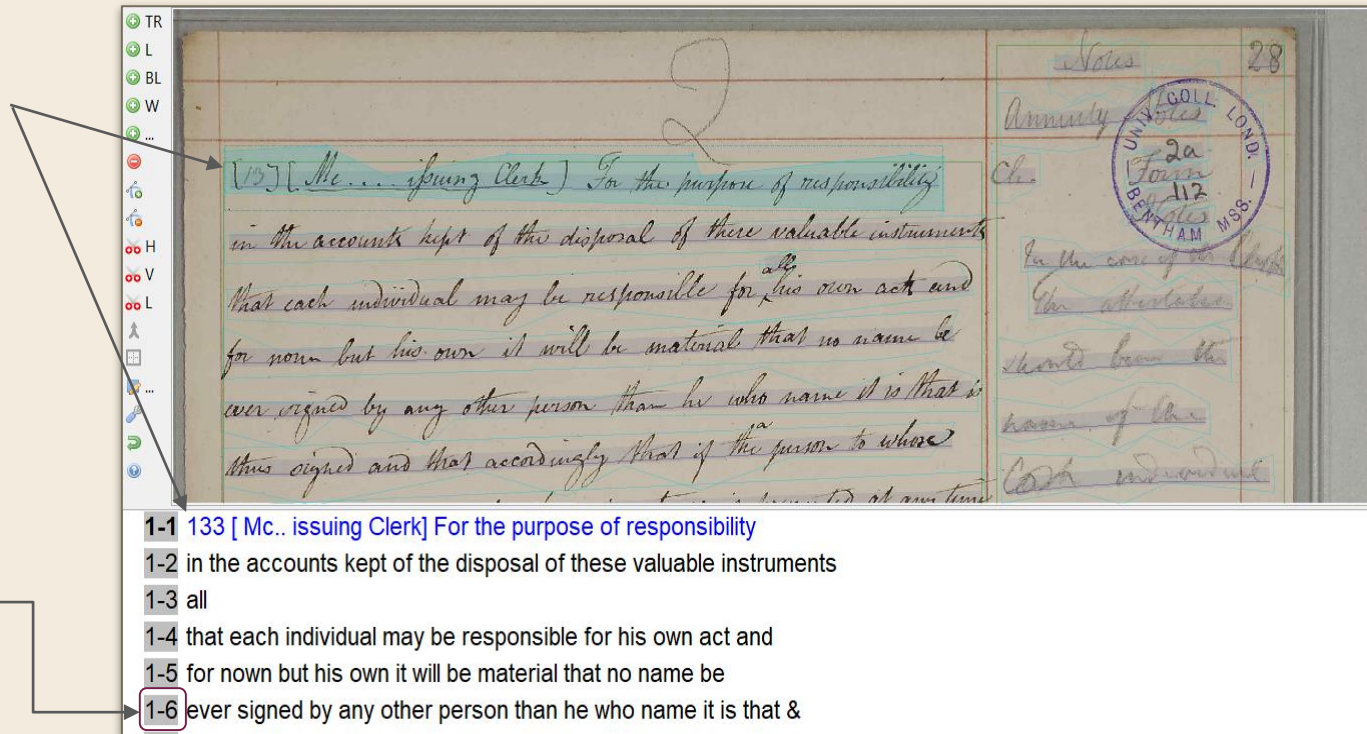
Szegmentálás/Layout kialakítása:

Az átírás létrehozásának
előfeltétele: a dokumentum képi
elrendezésének jelölése



Transkribus Workflow

Szövegsor - képsor
közötti megfelelés: az
egyikre kattintva a
másik is kiemelésre
kerül.



1-1 133 [Mc.. issuing Clerk] For the purpose of responsibility

1-2 in the accounts kept of the disposal of these valuable instruments

1-3 all

1-4 that each individual may be responsible for his own act and

1-5 for nown but his own it will be material that no name be

1-6 ever signed by any other person than he who name it is that &

Első szám: Text
region (szöveg régió)
Második szám: Sor

Mi az a HTR?

Kézírás felismertetés

Can you read this?

It's a legitimate question.
You might or might not
recognize it as an
example of cursive writing.

AR-190319612.jpg

+Page 1

+Block 1

+Paragraph 1

can you read this ?

+Paragraph 2

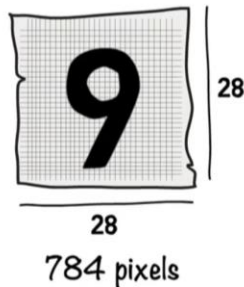
gró a legitimate questi
on . you might or might
not recognize it as an e
xample of cursin writin
g .

Mélytanulás

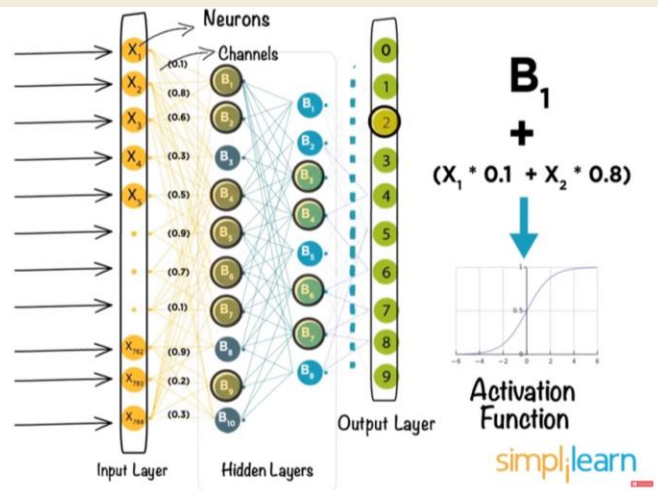
Gépi tanulás

Mesterséges
intelligencia

Handwritten Text Recognition



©Simplilearn. All rights reserved.



Mi a HTR?

**Handwritten Text Recognition (HTR) és az
Optical Character Recognition (OCR)**
ugyanazt a munkát végzik el (átalakítják a
szkennelt képek szövegét gépi kódolású
szöveggé),

DE

Mi a HTR?

OCR	HTR
Nyomtatott/gépelt szöveg	Kézzel írott szöveg + kora újkori nyomtatott könyvek IS
Fókuszban az egyes karakterek	A teljes sort dolgozza fel
Fix eszköz	Egyéni AI modellek tanítása lehetséges
	Jelentősen javult az elmúlt évtizedben a mesterséges intelligenciának és a mély neurális hálózatok kifejlesztésének és alkalmazásának köszönhetően.

Legjobb eredmények, vegyes kézírásmodell 9.19%



Tudom. Szekessy. Ké!
Azt a sok szentelencsélőre, a mielőtt megérkeztem
Szekessy ismétlő önkéntes kérésével megírtam magam.

Kiss József úrnak, a Költőnek.
Budapest.

ter! Évek és évek is írtam, ma
jelen keze három antecessor
ai fogadóból. Készenjáromon
Budapesti Főváros és igazgató
röviden vékonyasági arra látó
nak, hogy a fogadótan halom
ra került. Készenjáromon
Fővárosiemből máhny, feje
hasalozott munkát, a Kri
és esetleg a Kritika után.



Magyarország 1911. évi
Közlöny.

Metameditáció mióta és naponta
Kalmár börtönjében
Kalmár börtönjében

női jó kive
Mafra

Kiss József!
Tudom, hogy a sok szentelencsélőre, a mielőtt megérkeztem
Szekessy ismétlő önkéntes kérésével megírtam magam.

Magyarország 1911. évi
Közlöny.

Kiss József!
Tudom, hogy a sok szentelencsélőre, a mielőtt megérkeztem
Szekessy ismétlő önkéntes kérésével megírtam magam.

Magyarország 1911. évi
Közlöny.

szertartás meddig maradjunk, is hogy
le mag moshon is mindkét oldal
nem fogad jel áránai magadat mert

Kiss József!
Tudom, hogy a sok szentelencsélőre, a mielőtt megérkeztem
Szekessy ismétlő önkéntes kérésével megírtam magam.



Kiss József!
Tudom, hogy a sok szentelencsélőre, a mielőtt megérkeztem
Szekessy ismétlő önkéntes kérésével megírtam magam.

Kiss József!
Tudom, hogy a sok szentelencsélőre, a mielőtt megérkeztem
Szekessy ismétlő önkéntes kérésével megírtam magam.

Kiss József!
Tudom, hogy a sok szentelencsélőre, a mielőtt megérkeztem
Szekessy ismétlő önkéntes kérésével megírtam magam.

Levélíró:

Rosenfeld Bella (1869-1908)

Címzett:

Kiss József (1843-1921)

Kérem tisztelt, kedves szer-
kesztő úr,
írtam egy novellát, de a
„Hét”-től azt ellforálták,
hogy Ön bizonytalan ideig
B. Füreden marad még. Mi-
vel nem szeretném, ha a no-
velles belekerülne a lapba,
a nélkül hogy megjárta
volna az Ön ítéletének fo-
rumát: nagyon kérem, vaje-
z vagy írassa meg nekem, med-
dig marad Füreden? Oda-
küldhetem-e a kézirat?

Version Comparator

☒ Show line numbers

1-1 # kélyen-Mélyen tisztelt, Kedves-kedves szer-
1-2 # kesztő ur-úr,
2-1 # írtam egy novellát, de a
2-2 # a „Héét”-től „Hét”-től azt ellforálták-telefonálták,
2-3 # hogy Ön bizonytalan ideig Ön bizonytalan ideig
2-4 # Sz. B. Füreden sarad meg, hi-marad még. Mi-
2-5 # vel nem szeretném, ha a mo=no-
2-6 # velles belekerülne vella belekerülne a lapba-lapba,
2-7 # a nélkül hogy megjárta-megjárta
2-8 # volna az Ön ítéletének-ítéletének fo-
2-9 # numat-rumát: nagyon kénem, vajez=kérem, írja,
2-10 # vagy Rassa-nég nekem, írassa meg nekem, med-
2-11 # Uig-dig marad Füreden? Pelgg-Füreden? Oda
2-12 # külehetem-e küldhetem-e a kénirast? kézírást?
3-1 # 1



<https://readcoop.eu/transkribus>

Language:

hun

▼

Handwritten

Print

AI model: Hungarian handwriting 19th–20th Century ⓘ

Drag an image here
or

Select a file...

PNG or JPG up to 10 Mb

By uploading an image, you accept our [terms and conditions](#) and our [privacy policy](#)



szucs.kata@oszk.hu

Ha többet szeretnél tudni **Transkribus** és a **READ COOP** témában



Weboldal: <https://readcoop.eu/>

Útmutatók, videók, GYIK és szószeret:
<https://readcoop.eu/transkribus/resources/>

E-mail: info@readcoop.eu

Twitter: [@Transkribus](https://twitter.com/Transkribus)