

VALÓS TÉRBEN – AZ ONLINE TÉRÉRT

Networkshop 31: országos konferencia

2022. április 20–22.
Debreceni Egyetem

Szerkesztette: Tick József, Kokas Károly, Holl András

HUNGARNET Egyesület
Budapest, 2022



A kötet megjelenését támogatta az
Energiaügyi Minisztérium

Szerkesztette: Tick József, Kokas Károly, Holl András

Tipográfia és tördelés: Vas Viktória

Workshop

2022. április 20–22. Debreceni Egyetem, konferencia előadásainak közleményei

ISBN 978-615-82243-0-7

DOI: [10.31915/NWS.2022](https://doi.org/10.31915/NWS.2022)

Kiadja a HUNGARNET Egyesület
az MTA Könyvtár és Információs Központ közreműködésével
Budapest
2022

Borítókép: [freepik.com](https://www.freepik.com)

TARTALOMJEGYZÉK

Előszó	5
Lencsés Ákos: A nyílt tudomány pénzügyi vonatkozásai	7
Farkas Katalin: Centenáriumi média-adattár és virtuális kiállítás létrehozásának tanulságai az SZTE Klebelsberg Könyvtárban	13
Bódog András: A nyílt archívumi információs rendszer (OAIS) szabványának honosítása.....	20
Perlaki Attila: Oktatást segítő gamifikációs alkalmazások, mint szakdolgozati témák	27
Csapó Noémi – Dani Erzsébet: APPropó fejlődés – A Bács-Kiskun Megyei Katona József Könyvtár mobilapplikációja.....	32
Simon András: Integrált könyvtári rendszerek tranzakciós rekordjainak vizsgálata, a könyvtári állomány digitalizálásának tervezésekor.....	41
Németh Márton: Az OSZK Webarchívum nemzetközi kapcsolatai.....	58
Antal Péter: A mesterséges intelligencia kihívásai a XXI. század társadalmára	70
Hajdu Csaba – Szilágyi Zoltán: Modern robotikai technológiai ismeretek oktatása „Teljes spektrumú” oktatási módszerrel	77
T. Nagy László – Boda István Károly – Tóth Erzsébet: E-tananyagfejlesztés virtuális 3D környezetben.....	84
Palencsárné Kasza Marianna: Digitális átállás – Minőség – lehetőségek az EQAVET terén.....	92
Nagy Gyula: Nemzetközi kitekintés a felsőoktatási könyvtárak világára: a EUGLOH könyvtári workshopja	99
Babocsay Gergely: Az európai természettudományi gyűjtemények digitális integrációja: határ a csillagos ég.....	108
Somorjai Noémi: Egyenlőtlenségek a tudományos kutatás területén. Az amatőr kutatók szerepe	114
Molnár Dániel – Dani Erzsébet: Robotok a könyvtárban: Hogyan válhat a robotika a könyvtári mindennapok részévé?	122
Horváthné Felföldi Helga: Digitalizáció a szakképzésben. A Szakmajegyzékben szereplő szakmák digitáliskompetencia jártassági szintjeinek felülvizsgálata	130
Kalcsó Gyula: Ne csak útra csomagoljunk! Miért fontos a csomagolás a digitális megőrzésben?	138
Karsa Zoltán István – Szeberényi Imre: A CIRCLE felhő elmúlt évtizede	146
Bobák Barbara – Kasza Péter: Az MI lehetőségei a kora újkori filológiában: Johannes Michael Brutus <i>Rerum Ungaricarum</i> libri kéziratának digitális kiadása (esettanulmány)	154
Egyed-Gergely Júlia – Vajda Róza, Gárdos Judit – Horváth Anna – Meiszterics Enikő – Micsik András – Martin Dániel – Marx Attila – Pataki Balázs – Siket Melinda: Szociológia, kutatási adatok, mesterséges intelligencia: lehetőségek és tapasztalatok	161
Szemes Botond – Bajzát Tímea – Fellegi Zsófia – Kundráth Péter – Horváth Péter – Indig Balázs – Dióssy Anna – Hegedüs Fanni – Pantyelejev Natali – Sziráki Sarolta – Vida Bence – Kalmár Balázs – Palkó Gábor: Az ELTE Drámakorpuszának létrehozása és lehetőségei.....	170



Sebestyén Ádám: Az ELTEdata szemantikus adatbázis legújabb fejlesztései.....	179
Szlamka Erzsébet: Új trendek a tanulási eredmények tanúsításában	185
Tóth Máté – Héjja Balázs: Webshop indítása közkönyvtári környezetben.....	192
Etlinger Mihály – Hernády Judit: A kiadás hagyatéka / a hagyatéka kiadása: A Régi Magyar Költők Tárának hálózati kiadásáról.....	199
Varga Emese – Makkai T. Csilla: „Ki a fenének kell collstok?” A digitális szöveg rejtett mértékegységei	204
Dobás Kata – Fazekas Júlia: ITIdata – Egy irodalmi adatbázis fejlesztése Wikibase alapon és ennek hasznosítása Kosztolányi Dezső forrásjegyzékénél	211
Sörény Edina: Kézai Simon Program – digitális családi fotóarchívum.....	219
Fülöp Tiffany – Molnár Tamás – Hoczopán Szabolcs: Open Monograph Press e-könyvplatform a Szegedi Tudományegyetemen	227
Palkó Gábor: Mesterséges intelligencia, digitális bölcsészet, kulturális örökség: trendek és eredmények.....	235
Pergéné Szabó Enikő – Bátfai Mária Erika: A tudományos publikálás támogatása a Debreceni Egyetemi és Nemzeti Könyvtárban	241
Csirmazné Rezi Éva: Nemzetközi kiadványazonosítók és kötelezpéldányok kezelése az OSZK OKP (Országos Könyvtári Platform) rendszerében	250
Alföldi István – Dióssy Anna Laura: Digitálisan született kutatási anyagok megőrzése: a relációs adatbázis mint born-digital objektum	262
Fekete Norbert: HTR-modellépítés és kézírásfelismerés nagyméretű, többszerzős szövegtörzseten. A Transkribus alkalmazása az Arany János hivatali iratokon.....	271
Horváth Péter – Kundráth Péter – Palkó Gábor: ELTE Népdalkorpusz – magyar népdalok gépileg annotált adatbázisa	276
Nagy György: IKT eszközök alkalmazása az alsó tagozatos környezetismeret órákon.....	284
Köpösdí Zsuzsa – Molnár Tamás: Multimédiás, interaktív és adaptív tananyagok létrehozásának lehetőségei H5P keretrendszerrel	289
Jankó Tamás: Munka 4.0 – Ipar 4.0 – Szakképzés 4.0 – : A digitális kompetencia jövőbeni fejlesztési útjai	296
Békésiné Bognár Noémi Erika – Nagy Andor: Megújuló könyvtári statisztika: az egységes adatstruktúra és a korszerű megjelenítés kialakításának útján	304
Bolya Máttyás: Kézírtos dallamlejegyzések feldolgozása MI-vel támogatott digitális környezetben	310
Maróthy Szilvia – Seláf Levente – Vigyikán Villó: Régi magyar verskorpusz összeállítása stilometriai és számítógépes metrikai kutatásokhoz	324
Szűcs Kata Ágnes: Kézírtos források transzformációinak lehetőségei a közgyűjteményekben.....	330
Fellegi Zsófia: A digitális filológia infrastruktúrái. A DigiPhil megújulásáról.	338
Mihály Eszter: Mi az a dHUpla? A Digitális Bölcsészeti Platform bemutatása.....	345
Nemeskey Dávid Márk – Palkó Gábor: Szemantikus névelém-azonosítás magyar nyelvű szövegeken (a HuWikifier bemutatása)	359

HTR-modellépítés és kézírásfelismerés nagyméretű, többszerzős szövegkorpuszon A Transkribus alkalmazása az Arany János hivatali iratokon¹

Fekete Norbert
Miskolci Egyetem-DigiPhil
feketenorbert1986@gmail.com

Creation of HTR models and handwritten text recognition on a large, multi-author text corpus

Using Transkribus on documents from the office of János Arany

In our study, we show how Transkribus has been used on a multi-author text corpus to create a HTR model that can process documents from the office of Arany efficiently. The test results so far are encouraging, and allow the features of model creation to be applied to other archives. To do this, it is worth creating HTR models based on one or two dominant handwritings first, and then other handwritings can be added to increase efficiency.

keywords: handwritten text recognition (HTR), Transkribus, János Arany

kulcsszavak: kézírásfelismerés (HTR), Transkribus, Arany János

Tanulmányunkban azt vizsgáljuk, hogy a Transkribus kézírásfelismerő (HTR) program hogyan használható egy több ember kézírását tartalmazó, eltérő dokumentumtípusokból álló szövegkorpuszon. A vizsgálathoz szükséges korpuszt az Arany János akadémiai főtőkárságának ideje alatt (1865–1877) keletkezett hivatali iratok szolgáltatták. Az Arany műveiből készülő kritikai kiadás új sorozatának előkészítő munkálatai során eddig feltérképezett korpusz kb. 9200 dokumentumot, azaz mintegy 30000 képfájlt jelent. Ennek a nagyméretű irategyüttesnek a kézírásról szöveggé történő átírása megoldhatónak tűnik a Transkribus segítségével. Ehhez első lépésként egy olyan HTR-modellt szükséges létrehozni, amely képes kellő hatékonysággal megküzdeni a korpuszt létrehozó számos kézírással, ugyanis a vizsgált anyag nemcsak az Arany által írt dokumentumokat, hanem más, az MTA-val kapcsolatba kerülő szerzők iratait is tartalmazza. A kutatócsoport tagjai az eddigi digitális filológiai kutatások során Arany kézírásával kapcsolatban két modellt dolgoztak ki.² Alkalmazkodva a korpusz sajátosságaihoz egy általánosabb harmadik modell létrehozására törekedtünk, amelynek eredményeit a következőkben szeretnénk bemutatni.

¹ A tanulmány a Digitális Örökség Nemzeti Laboratórium támogatásával valósult meg.

² Bobák Barbara, Gábori Kovács József, „Kézírásfelismerés Arany János levelein”, In: *Networkshop 2019*, szerkesztette Tick József, Kokas Károly, Holl András, 38-44, Budapest: Hungarnet Egyesület, 2019.



A hivatali iratok és jelentőségük

Arany 1859-től az MTA rendes tagja volt, majd 1865-ben főtitkárrá választották, amely pozíciót egészen 1877-ig töltötte be. Arany tizenkét éven keresztül gyakorlatilag az Akadémia harmadik, az elnöki és az alelnöki pozíciót követően legfontosabb tisztségét birtokolta. Az általa végzett hivatali munka jelentős, mivel főtitkársága alatt kialakította és hosszú évtizedekre meghatározta az MTA, mint a hazai tudományosság egyik legfontosabb intézményének működési rendjét. Munkája során számos napi, a tudományszervezéshez kapcsolódó adminisztratív teendőt látott el: bejegyezte az üléseken tartott előadásokat, szerkesztette a társaság folyóiratait, intézte a kiadványok ügyeit, felbontotta és megválaszolta a postát, elkészítette az MTA éves beszámolóit, vezette az ülések jegyzőkönyveit, illetve intézte a pályázatok ügyeit. Akadémiai munkáját a kezdetektől egy öt fős titkárság segítette.³ Az Arany-érában keletkezett iratok vizsgálata és digitális kritikai kiadásának létrehozása nemcsak a költő főtitkári pályájának mélyebb megértéséhez vezet el, hanem rámutat az MTA történetére, illetve az 1860-as és 1870-es évek tudományos, kulturális és művészeti életére is. Egyrészt a korpusz közzététele azért jelentős, mert számos kutatás tudja majd hasznosítani a jövőben; másrészt a kiadás egy lényeges, megoldandó digitális filológiai probléma, amely a hazai archívumok feldolgozására kínálhat lehetőséget.

A feldolgozás mérföldkövei a Transkribus előtt

A hivatali iratok feldolgozása és közzététele az Arany János összes művei kritikai kiadássorozat 14. kötetében történt meg először. A sajtó alá rendező Gergely Pál elsődleges célja az volt, hogy a költő hivatali működését leíró legfontosabb iratokat és dokumentumtípusokat bemutassa. Az általa feltárt mintegy 2500 dokumentumból mindössze 717-et tett közzé, melyek között természetesen nem csak Arany kezétől származó iratokat találunk.⁴ A kritikai kiadás új sorozatának munkálataihoz kapcsolódva 2014-ben folytatódott az iratokat feldolgozó munka. Ennek keretében feltérképezték és digitalizálták a Magyar Tudományos Akadémia Könyvtár és Információs Központ Kézirattár Régi Akadémiai Levéltárában (MTA KIK, Kt, RAL) található gyűjteményt. Ez összesen kb. 9200 dokumentumot, vagyis kb. 30000 fényképfelvételt jelent. Ebből a korpuszból kb. 2000 dokumentum szövegének átírását végezte el Gábori Kovács József, Szlanyiszló Lilla, Antal Alexandra és Fekete Norbert. A kutatók a szövegeket kézzel, wordben, illetve a Hegedüs Béla által létrehozott LyX kritikai kiadás készítésére alkalmassá tett programban végezték. A kiadás koncepciója többször változott: a jelenlegi cél, hogy a főtitkár kezén átment teljes anyag kiadása megvalósuljon, mely lehetővé teszi a teljes főtitkári munka feltárását.⁵ A RAL feltárt állománya sem tekinthető azonban teljesnek, melyet a digitális kiadás előkészületeihez tartozó munkálatok során magunk is tapasztaltunk: Arany autográf kézírásával készült számos olyan további dokumentum található, melyek a költő hivatali tevékenységéhez köthetők. Ezek alapján látható, hogy a korpusz a digitalizált állomány feldolgozását követően tovább fog bővülni.

3 Gábori Kovács József, „Arany János, a hivatalnok”, In: *Önarckép álarckokban: Kiállításkatalógus*, szerkesztette Vaderna Gábor, Kalla Zsuzsa, Kaszap-Asztalos Emese, Sidó Anna, 350–366, Budapest: Petőfi Irodalmi Múzeum, 2018, 357–361.

4 Arany János, *Hivatali iratok 2.: Akadémiai évek (1859–77)*, szerkesztette, a jegyzeteket készítette Gergely Pál, Budapest: Akadémiai Könyvkiadó, 1964. (Arany János összes művei, 14.)

5 Gábori Kovács József, „Arany János hivatali levelei az Akadémián: A kritikai (újra)kiadás szükségessége és lehetőségei”, In: „... és palota épül a puszta beszédéből”: Akadémiai tudományos ülésszakok a 200 éves Arany Jánosról, szerkesztette Gábori Kovács József, Major Ágnes, 212–221, Budapest: reciti Könyvkiadó, 2017, 217–218; Bobák Barbara, Gábori Kovács József, „Kézírásfelismerés...”, 5.

A feldolgozás lehetőségei és eredményei Transkribusban

A hivatali iratok feldolgozásához az Innsbrucki Egyetem által 2016-ban kifejlesztett Transkribus programot használtuk fel, amelynek nagy előnye, hogy nyílt forráskódú, felhasználóbarát felülettel bír, és egyszerre rendelkezik OCR és HTR modullal, így a nyomtatott és a kéziratos források feldolgozására is alkalmas. Ennek köszönhetően nagy mértékben csökkenthető vele a digitális szövegkiadások elkészítésének ideje, hiszen lerövidíti a szövegek átírására fordított időt. Hátránya azonban, hogy nem ingyenes.⁶ Az eredményes kézírásfelismeréshez szükséges egy tanító modell összeállítása, amelynek létrehozásához egy legalább 15000 szóból álló korpusznak kell rendelkezésre állnia. A DigiPhil csoport és az ELKH Bölcsészettudományi Kutatóközpont Irodalomtudományi Intézetének közös kutatása során eddig két Arany-modell készült, melyeket olyan szövegekből hoztak létre, amelyek csak Arany kézírását tartalmazták. A két modell között lényeges eltérés figyelhető meg: egyrészt bővült a modellekhez kiválogatott és átírt szövegek mennyisége, illetve a második modell hatékonysága jelentősen javult az elsőhöz képest. A Digitális Örökség Nemzeti Laboratórium, az ELTE, az ELKH BTK Irodalomtudományi Intézet és a Miskolci Egyetem kutatóiból álló kutatócsoport a hivatali iratok feldolgozásához egy új, harmadik modell létrehozását tűzte ki célul, amely már nemcsak Aranytól, hanem más szerzőktől is tartalmaz szövegeket. Távlati cél az, hogy az új modell képes legyen bármilyen 19. századi magyar nyelvű kézírás kiolvasására a lehető legkisebb hibaarányal.

A harmadik modell kialakításához kapcsolódóan több megfontolást is figyelembe kellett venni. Először át kellett tekintenünk a korpusz egy részét: erre azért volt szükség, mivel nem állt rendelkezésünkre olyan átfogó összeírás, amely tartalmazná a hivatali iratok összességét, ezért ezt párhuzamosan kellett elvégeznünk. A korpusz részleges áttekintése lehetővé tette, hogy megalapozzuk a harmadik modellhez tartozó tanítóanyag létrehozását. A források közül döntően magyar nyelvű, lehetőleg teljes dokumentumokat tartalmazó képeket válogattunk ki. Fontos szempont volt, hogy a képeken ne csak pár sornyi szöveg legyen, és az egyszerűség kedvéért kerültük a táblázatos források kiválasztását. A korpusz egyrészt az 1865-ben keletkezett és az első két modellhez használt képeket, valamint 1868-as és 1869-es dokumentumokat tartalmazza. A harmadik modell létrehozásához összesen 974 képfájl használtunk fel, melynek megoszlását az 1. táblázat mutatja.

Évkör	Korpusz	Mennyiség
1865	1. modellhez	100
1865	2. modellhez	200
1865	Ring Adorján	101
1868–1869	3. modellhez	573

1. táblázat A modellépítéshez használt Arany korpuszok

Az 1865-ös anyagból látható, hogy az egyik alkorpusz Ring Adorjának, Arany egyik írnokának a szövegeit tartalmazza, aki pozíciójából kifolyólag nagy számú hivatali szöveget hozott létre. Ebből következően a hozzá kötődő korpusz jelentősen tudja javítani egy többszerzős modell létrehozását.

Az új Arany-modell rendkívül heterogén, ugyanis kb. 30 kéztől tartalmaz különböző iratokat. A korpusz 75%-át teszik ki Arany és a Főtitkári Hivatal dolgozóinak iratai, míg 15%-át adja az MTA többi meghatározó személyisége (elnök, osztályelnökök stb.), akik pozíciójukból kifolyólag gyakran leveleztek a főtitkárral. Emellett mintegy 10%-a származik azoktól a levélíróktól, akik csupán néhány levélváltás erejéig szerepelnek a korpuszban.

⁶ Transkribus <https://transkribus.eu/Transkribus/> (hozzáférés: 2022. 06. 07.)

A modellkészítés során az anyagot a második modell segítségével ismertettük fel, majd ezt kézzel javítottuk. A harmadik modellhez kapcsolódó átírások javítását összesen tizenegy fő végezte, míg a második modellhez készült szövegeket korábban Gábori Kovács József ellenőrizte.⁷ A harmadik modellhez kapcsolódó textusok javítására fordított idő függött az írás terjedelmétől, az íráskép olvashatóságától. Arany kézírását a Transkribus a második modell alapján jó hibahatárral ismerte fel, így az ő szövegeinek javítása más szerzőkével szemben problémamentes volt. A legrövidebb javításra fordított idő kb. 5 perc, míg a leghosszabb akár 60–70 percet is igénybe vehetett. Átlagosan 15–20 percet fordítottunk egy átírat átnézésére.

A harmadik modell eredményei jócskán felülmúlták az előzőekét, amely leginkább a Character Error Rate (CER on Validation set) értéken mérhető le. Ez alapján megállapítható, hogy a program milyen hibaszázalékkal képes elvégezni a kézírásfelismerést, vagyis képenként a felismert karakterek hány százalékát rontja el. A korábbi eredményekhez képest (lásd 2. táblázat) a CER Validation értéke csökkent, miközben a korpuszt alkotó szószám jelentősen növekedett. A CER Validation értéke az első modell esetében 10,42% volt, míg a másodikonál 8,08%-ra csökkent, viszont a harmadik esetében 9,3%-ra növekedett. Ez utóbbi azonban jó aránynak tekinthető, főleg egy ennyire heterogén korpusz esetében.

	CER Train	CER Validation	Szavak száma
1. modell	5,73%	10,42%	17636
2. modell	2,71%	8,08%	29498
3. modell	5,30 %	9,30%	124070

2. táblázat Az Arany-modellek CER mutatói

A harmadik modell eredményeit a Bobák Barbara és Gábori Kovács József által alkalmazott vizsgálat módszerét figyelembe véve is felmértük.⁸ Ennek alapja az ún. token, vagyis a szóköztől szóközig terjedő karaktersorban ejtett hibák száma, melyhez kapcsolódva a szerzőpáros felmérte a karakterekben ejtett hibákat is, vagyis azt, hogy hány karaktert tévesztett a program az átírás elkészítésekor. A harmadik modell eredményeit a 3. táblázatban foglaltuk össze, melyben szerepeltettük Bobák és Gábori Kovács (BB-GKJ) adatait is. A teszthez két darab, egy-egy képre kiterjedő dokumentumot jelöltünk ki, melyeken csak egy kéztől származó kézírás szerepelt. Az egyik kép Arany Jánosé (AJ) és két levél piszkozatát tartalmazza, míg a másik képen található egyetlen levél Szabenyi Ferencről (SzF), a pesti Királyi Magyar Egyetemi Nyomda gondnokától származik.⁹ Mind a két választott képen lefuttattuk mindhárom modellt, melynek eredményeképpen megállapítható, hogy az első két modell nagyobb sikerrel ismerte fel a karaktereket és kevesebb tokenben hibázott az Aranytól származó szöveg esetében. Ez többek között annak is köszönhető, hogy ezek Arany szövegeit felhasználva készültek. A választott Arany-leveleken végzett tesztelés közel hasonló eredményeket hozott mint a Bobák és Gábori Kovács által végzett teszteké. A harmadik, heterogén modell azonban sikeresen ismerte fel mind a két képen található szöveget, és már csak minimális hibák mutatkoztak a tokenek és a karakterek szintjén egyaránt. A tesztek más példák esetében eltérő számokat mutathatnak, de az arányok jól szemléltetik, hogy a nagyobb korpuszon alapuló harmadik modell közel azonos hatékonysággal képes felismerni

7 A harmadik modellhez kapcsolódó szövegek átírásában közreműködtek: Bajzát Tímea Borbála, Bobák Barbara, Fazekas Júlia, Fekete Norbert, Fellegi Zsófia, Hernády Judit, Horváth Anna, Jakab Éva, Szolnoki Anna, Vámos Violetta, Vétek Bence.

8 Gábori Kovács, Bobák, „Kézírásfelismerés...” 6.

9 Arany János Radnótfáy Sámuelnek, Pest, 1868. jan. 8., MTA KIK, Kt, RAL, 10/868; Arany János Karvassy Ágostonnak, Pest, 1868. jan., MTA KIK, Kt, RAL, 10/868; Szabenyi Ferenc a m. kir. vallási és közoktatási ministeriumnak, Buda, 1869. máj. 18., MTA KIK, Kt, RAL, 1869/1246.

az Aranyétól eltérő kézírásokat is, így a modell gyakorlatilag alapjául szolgálhat a 19. századi kézírásokat jó hatásfokkal felismerni képes modellhez.

	Token BB-GKJ	Karakter BB-GKJ	Token AJ	Karakter AJ	Token SzF	Karakter SzF
1. modell	128/200	276/954	69/133	121/814	109/127	436/793
	64%	29%	52%	15%	86%	55%
2. modell	48/200	116/954	34/133	51/814	79/127	102/793
	24%	12%	26%	6%	62%	13%
3. modell	-	-	7/133	9/814	7/127	9/793
	-	-	5%	1%	6%	1%

3. táblázat Az Arany-modellek teszteredményei

Összegzés

Tanulmányunkban azt mutattuk be, hogy miként sikerült egy egyre bővülő és több szerzői kéztől származó szövegtörzset betanítani a Transkribus és létrehozni egy olyan HTR-modellt, amely hatékonyan tudja kezelni a hivatali iratokat. Az eddigi teszteredmények bizakodásra adnak okot, és valószínűleg lehetővé teszik, hogy a modellalkotás sajátosságait más hasonló archívumok esetében is alkalmazni lehessen. Ehhez először is érdemes olyan modelleket létrehozni, amelyek egy-két meghatározó kézen alapulnak, majd újabb kezek hozzáadásával a hatékonyság tovább növelhető.

Bibliográfia

Arany János Radnótfáy Sámuelnek, Pest, 1868. jan. 8., MTA KIK, Kt, RAL, 10/868.

Arany János Karvassy Ágostonnak, Pest, 1868. jan., MTA KIK, Kt, RAL, 10/868.

Arany János, Hivatali iratok 2.: Akadémiai évek (1859–77). szerkesztette, a jegyzeteket készítette Gergely Pál, Budapest: Akadémiai Könyvkiadó, 1964. (Arany János összes művei, 14.)

Bobák Barbara, Gábori Kovács József, „Kézírásfelismerés Arany János levelein”, In: *Networkshop 2019*, szerkesztette Tick József, Kokas Károly, Holl András, 38-44, Budapest: Hungarnet Egyesület, 2019. <https://doi.org/10.31915/NWS.2019.5>

Gábori Kovács József, „Arany János, a hivatalnok”, In: *Önarckép álarckokban: Kiállításkatalógus*, szerkesztette Vaderna Gábor, Kalla Zsuzsa, Kaszap-Asztalos Emese, Sidó Anna, 350–366, Budapest: Petőfi Irodalmi Múzeum, 2018

Gábori Kovács József, „Arany János hivatali levelei az Akadémián: A kritikai (újra)kiadás szükségessége és lehetőségei, In „... és palota épül a pusztából”: Akadémiai tudományos ülésszakok a 200 éves Arany Jánosról, szerkesztette Gábori Kovács József, Major Ágnes, 212–221, Budapest: reciti Könyvkiadó, 2017.

Szebenyi Ferenc a m. kir. vallási és közoktatási ministeriumnak, Buda, 1869. máj. 18., MTA KIK, KT, RAL, 1869/1246.

Transkribus <https://transkribus.eu/Transkribus/> (hozzáférés: 2022. 06. 07.)