

Szoftveres Cutter-keresés az SZTE Klebelsberg Könyvtárban

Cutter Search Software in the SZTE Klebelsberg Library

Frankó Máté
SZTE Klebelsberg Könyvtár
mate.franko@ek.szte.hu
ORCID: [0009-0009-3782-6571](https://orcid.org/0009-0009-3782-6571)

Zeller Rozália
SZTE Klebelsberg Könyvtár
rozalia.zeller@ek.szte.hu
ORCID: [0000-0003-2501-8760](https://orcid.org/0000-0003-2501-8760)

Absztrakt

A Cutter-számok a könyvtári dokumentumok betűrend szerinti elhelyezéséhez és rendszerezéséhez használt alfanumerikus kódok. Más közgyűjteményekhez hasonlóan az SZTE Klebelsberg Könyvtárban is - a dokumentumban foglalt mű témájára utaló szakjelek mellett - Cutter jelzeteket használunk a szabadpolcos állományunk sorrendezéséhez. Heti rendszerességgel mintegy 300-400 könyv Cutter jelzettel való ellátása a Könyvtári raktározási táblázatok c. segédlet használatával meglehetősen időigényes feladatot jelentett kollégáinknak. Ezért 2022-ben a munkafolyamat fokozott automatizálása mellett döntöttünk. Célunk az volt, hogy létrehozzunk egy olyan elektronikus felületet, amely a szerző, a cím vagy más bibliográfiai adatok megadása után, további emberi beavatkozás nélkül megmutatja, hogy az adott karaktersorozathoz milyen Cutter-szám tartozik.

Első lépésben az eredeti táblázatot kellett a kereső automatizmusok számára használható - koherens és a programnyelvek számára is értelmezhető - formátumra hoznunk, úgy, hogy az eredeti adatszerkezet ne sérüljön. Javítottuk az eredeti táblázat digitalizálása során keletkezett hibákat, átalakítottuk a speciális magyar karaktereket, komplementáltuk a hiányzó vagy hiányos zárótaggal rendelkező cuttereket, egyúttal igyekeztünk megoldást találni a szóközös és szóköz nélküli betűsorok egységesítésére. A táblát MS Excel környezetben, VBA-makrókkal konvertáltuk kódsorokká. A kész felület alatt egy PHP-alkalmazás működik. Előadásunkban a fejlesztés egyes munkafolyamatait, az alkalmazott módszereket, az elkészült rendszer működését és a tervezett fejlesztési irányokat kívántuk bemutatni.

Kulcsszavak: Cutter számok, Cutter táblázat, betűrendi jel, szoftveres keresés, VBA makró

Abstract

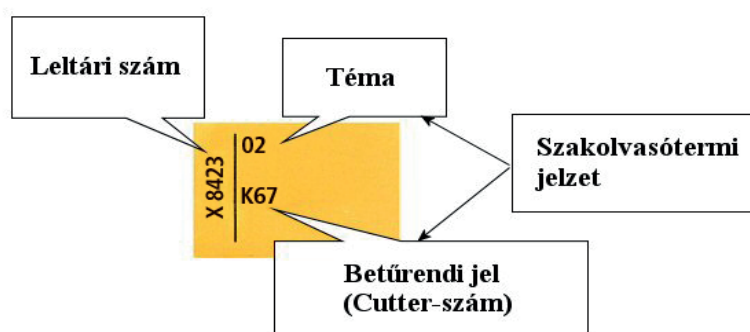
Cutter numbers are alphanumeric codes used to alphabetise and organise library documents. Like in many other libraries, we use the Cutter classification system - in addition to the Universal Decimal Classification system - to organise our open shelf holdings in the SZTE Klebelsberg Library. It was a time-consuming task for our colleagues to assign Cutter numbers to about 300-400 books per week using the Alphabetic Order Table. Therefore, in 2022 we decided to increase the automation of the workflow. Our goal was to create an electronic interface that, after entering the author, title or other bibliographic data, would show the Cutter number associated with a given character set without further human intervention.

As a first step, the Alphabetic Order Table had to be converted into a format that could be used by search automation - coherent and interpretable by programming languages - without damaging the original data structure. We corrected the errors that had occurred during the digitisation of the original table, converted the special Hungarian characters, complemented the truncated cutters, and tried to find a solution to unify the spaced and unspaced strings. The table was converted into lines of code in MS Excel using VBA macros. A PHP application runs under the finished interface. In our presentation we wanted to show the workflows of the development, the methods used, the functioning of the completed system and the planned directions of development.

Keywords: Cutter classification system, Cutter numbers, Alphabetic Order Table, search software, VBA macro

1. Bevezető

A nyomtatott dokumentumok fizikai helyét a legkülönbébb szakjelek, jelzetek, szakcsoportszámok jelölik világszerte a könyvtárakban. Ezek a betűkből és számokból összeállított azonosító jelek a szakemberek számára információkkal szolgálnak a megjelölt publikáció témájára, típusára vagy bibliográfiai adataira vonatkozóan. Használatuk megkönnyíti a dokumentumok tematikus rendszerezését és sorrendezését, illetve biztosítja azok gyors visszakereshetőségét akár százezres vagy milliós példányszámú közgyűjteményi állományok esetén is.



1. ábra. Az SZTE Klebelsberg Könyvtár szabadpolcos állományában használt címke

A Szegedi Tudományegyetem Klebelsberg Kuno Könyvtárának szakolvasótermeiben¹ elhelyezett szabadpolcos állomány azonosítására egy kétosztatú jelölési rendszert használunk. A dokumentumok szakterületi besorolásánál egyszerűsített ETO-számokat tartalmazó szakjelzetekkel, míg a kötetek ábécé szerinti rendezésénél általában a befoglalt mű szerzőjére vagy címére utaló Cutter-számokkal dolgozunk.

2. Cutter-számok és használatuk az SZTE Klebelsberg Könyvtárban

A Cutter-szám egy nyomtatott nagybetűből és egy kétjegyű számból álló alfanumerikus kód. Minden számhoz tartozik egy kezdő és záró érték. A Cutter-számok a két végpont közé eső, illetve azokkal megegyező karaktersorozatokat jelölik. A Cutter-számok hivatalos listáját – a hozzájuk tartozó kezdő és záró tagokkal – a Könyvtári Intézet (KI) által kiadott Könyvtári raktározási táblázatok² című segédkönyv tartalmazza.

1 Az SZTE Klebelsberg Könyvtár olvasói tereinek bemutatása. <http://www.ek.szte.hu/olvasoi-terek/> Hozzáférés: 2023.06.15.

2 Könyvtári raktározási táblázatok. 9. átdolg. javított kiadás. Budapest: Könyvtári Intézet, 2001. <https://www.ki.oszk.hu/dokumentumtar/konyvtari-raktarozasi-tablazatok> Hozzáférés: 2023.06.15

Például a B14-es cutterhez a „Baim” és a „Bakor” karaktorsorok tartoznak. Követve a megadott szabályt minden olyan nevet vagy könyvcímet, amely betűrend szerint a „Baim” után következik, de a „Bakor” szó előtt szerepel az ábécében, a B14-es Cutter-számmal kell jelölni. A B14-es szám alá kerül egyebek közt a „Bajorország” kifejezés és „Baka István” neve is.

A Cutter-számok kiosztását az SZTE Klebelsberg Könyvtárban az egyes tudományterületekért felelős szakreferensek, valamint a Feldolgozó Osztály munkatársai végzik. Korábban a Könyvtári raktározási táblázatok nyomtatott, illetve annak digitalizált és OCR-ezett elektronikus változatából dolgoztak a kollégák, de heti 300-400 dokumentum jelzetelése és átirányítása mellett meglehetősen problémás és időigényes feladat volt az egyes címekhez és szerzőkhöz tartozó Cutter-számok lekeresése.

Néhány évvel ezelőtt a probléma megoldása érdekében készült egy egyszerű keresőprogram, de az sem váltotta be igazán a hozzá fűzött reményeket. Az alkalmazásban kizárólag a kezdő és záró tagokra lehetett keresni, így teljes szavak, nevek keresésére nem volt alkalmas, csak a potenciális lehetőségek körét tudta szűkíteni.

2022-ben a könyvtár érintett vezetői a folyamat fokozott automatizálása mellett döntöttek. A cél az volt, hogy létrehozzunk egy olyan online felületet, amely a szerző, a cím vagy más bibliográfiai adat megadása után automatikusan megmutatja, hogy az adott karaktorsorozathoz milyen Cutter-szám tartozik.

3. Adattisztítás a Cutter-táblázatban

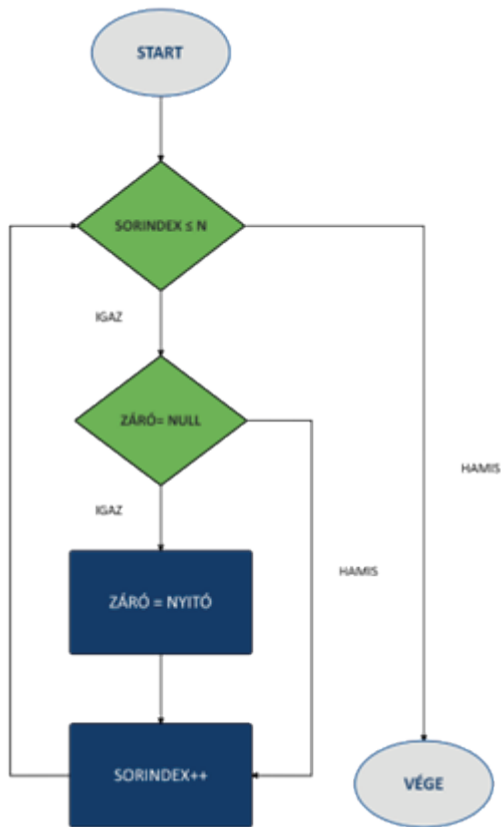
Első lépésben a Cutter-táblázatot kellett koherens és a programnyelvek számára is értelmezhető formára hozni – természetesen úgy, hogy az eredeti adattartalom ne sérüljön. Át kellett írunk a forrásállományban az egyszerűsítő jelöléseket és a számítógépes automatizálás folyamatát megnehezítő speciális karaktereket. A táblázat egyes sorait Microsoft Excelben, VBA makrók segítségével tisztítottuk le.

A Cutter-tábla belső felépítésének, logikai szerkezetének feltérképezése során két olyan, a főszabálytól eltérő típuskivételt találtunk, amelyek a számítógépes feldolgozás során problémát okozhattak volna.

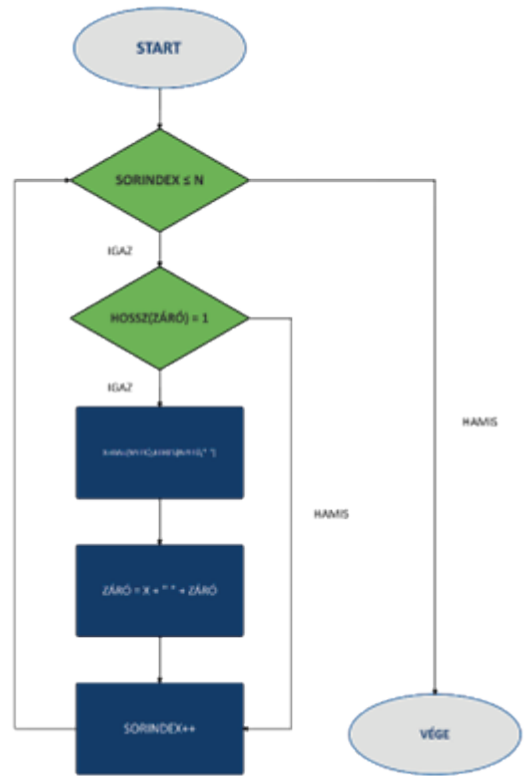
3.1. Hiányzó és hiányos záró tagok

Az első csoportba a hiányzó záró taggal rögzített sorok tartoznak. Ilyen például az A25-ös kódszámhoz tartozó rekord, ahol is az „Ady” nyitó tag mögött nincs záró érték. Könnyen belátható, hogy itt – és az összes hasonló felépítésű sorban – a nyitó és a záró tag megegyezik, vagyis az A25-ös Cutter-szám az „Ady”-tól „Ady”-ig terjedő karaktorsorozatokat jelöli, csak az utóbbit külön nem tüntették fel a segédlet készítői. Ahhoz, hogy a későbbiekben a keresőprogramnak legyen mit-mivel összehasonlítania, utólagosan pótolnunk kellett a hiányzó záró tagokat. Ehhez lefuttattunk egy szkriptet a tábla összes során, utasítva a táblázatkezelőt, hogy ott, ahol a záró érték helyén üres cellát talál, írja be a nyitó tagot.

A második kivételhalmazt a hiányos záró taggal felvitt sorok alkotják. Példaként említhetjük a K77-es Cutter-számot, ahol is a „Kovács E” előtag mögött csak egy „G” betűt tüntettek fel a szerkesztők. Ebben az esetben is egy formai egyszerűsítésről van szó: a „G” betű a „Kovács G” karaktorsorozatot hivatott jelölni, a keresőprogram helyes működéséhez azonban a teljes záró tagnak szerepelnie kell a táblázatban. Ezeknél a soroknál szoftveresen kimásoltuk a szóközt megelőző karaktereket a nyitó tagból és beillesztettük azokat a záró tag elé.



2. ábra. A program működési elve hiányzó záró tag esetén



3. ábra. A program működési elve hiányos záró tag esetén

3.2. Speciális karakterek

Két fontos könyvtári-nyelvtani szabályt kellett még átültetnünk a keresőprogramhoz készített új táblázatba. Az egyik, hogy a könyvtári ábécé nem tesz különbséget a rövid és a hosszú magánhangzók, valamint az „sz” kivételével az egy és a többjegyű betűk között. Ennek megfelelően az „á” betűs szavakat az „a”-hoz, a „cs” betűs szavakat a „c”-hez sorolja. A másik, hogy a szóközök minden más karaktert megelőznek, így például „Abonyi Andor” neve az „Abonyiak” kifejezés elé kerül a betűrendben, függetlenül attól, hogy az „Abonyiak” szót lezáró „k” karakter az „Andor” név „n” betűje előtt van az ábécében.

Figyelembe véve a fenti szabályokat létrehoztunk egy speciális ábécét, ahol minden betűt egy egyjegyű szám vagy az angol ábécé egyik betűje jelöl. Látható, hogy minden betű és szám mögé beillesztettünk egy felkiáltójelet is. Erre a helyes adattípus-deklaráció miatt volt szükség. Például „Baja” város neve az átírás után „1090”-ként jelenne meg a táblázatban. Ha nincsenek felkiáltójelek, a program számként és nem szöveges adatként értelmezné ezt a kifejezést.

Kódolt ábécé									
A a	0!	É é	4!	L l	B!	P p	G!	Ü ü	M!
Á á	0!	F f	5!	Ly ly	B!Q!	Q q	H!	Ú ú	M!
B b	1!	G g	6!	M m	C!	R r	I!	V v	N!
C c	2!	Gy gy	6!Q!	N n	D!	S s	J!	W w	O!
Cs cs	2!J!	H h	7!	Ny ny	D!Q!	Sz sz	J!R!	X x	P!
D d	3!	I i	8!	O o	E!	T t	K!	Y y	Q!
Dz dz	3!R!	Í í	8!	Ó ó	E!	Ty ty	K!Q!	Z z	R!
Dzs dzs	3!R!J!	J j	9!	Ö ö	F!	U u	L!	Zs zs	R!J!
E e	4!	K k	A!	Ő ő	F!	Ú ú	L!	[Szóköz]	!

4. ábra. A Cutter kereső szoftver számára előkészített speciális ABC

4. A Cutter-kereső működése

Az átalakított Cutter-táblázat egy XML fájlként került be a keresőprogramba. Az állomány az eredeti és a kódolt nyitó illetve záró tagokat is tartalmazza. A kódolt változat a kereséshez kell, a felhasználói interfész találati listájában viszont az eredeti alakban jelenik meg az eredmény.

<code><cutter></code>	
<code><number>R49</number></code>	Cutter-szám
<code><firsttag>Régi</firsttag></code>	Nyitó tag
<code><endtag>Regn</endtag></code>	Záró tag
<code><xfirsttag>!!4!6!8!</xfirsttag></code>	Átalakított nyitó tag
<code><xendtag>!!4!6!D!</xendtag></code>	Átalakított záró tag
<code></cutter></code>	

5. ábra. Részlet az XML fájlja alakított Cutter-táblázatból

A kereső első lépésben törli a felesleges karaktereket (például a szövegben hagyott írásjeleket vagy sortöréseket) a felhasználói inputból, majd átalakítja a megadott karaktersort a Cutter-tábla transzformálása során alkalmazott szabályok szerint.

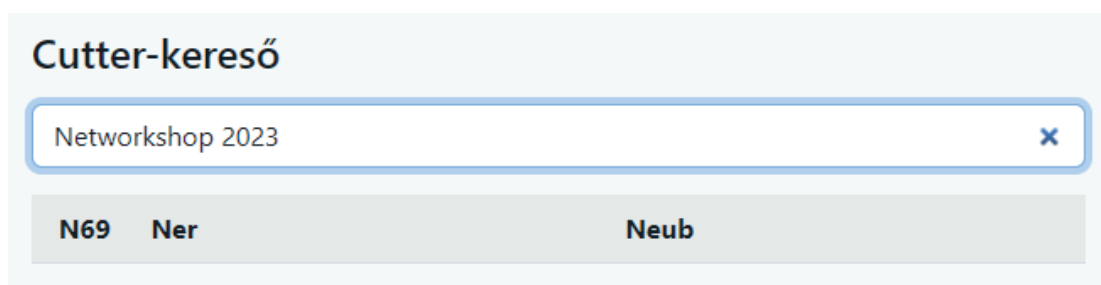
Ezt követően a program végigfut a Cutter-táblázaton, és szükség esetén feltölti plusz karakterekkel a nyitó és záró tagokat. Ha a nyitó és záró tagok rövidebbek, mint a felhasználó által megadott kifejezés, a program hibás eredményt adhat. A nyitó tagok mögé szóközöket (azaz a kódtábla szerint felkiáltójeleket), a záró értékek mögé pedig „z” betűket (azaz „R” betűket) ír a program, így a művelet a betűrendet nem módosítja.

Zárásként megvizsgálja, hogy melyik sorban igaz az, hogy az input nagyobb vagy egyenlő, mint a nyitó tag és kisebb vagy egyenlő, mint a záró. Végül az „IGAZ” értékkel visszatérő sorokat megjeleníti a beviteli szövegmező alatt egy táblázatban.

Fontos megjegyeznünk, hogy bár a program fejlesztése során egyszeres oldali programnyelvet (nevezetesen PHP-t) használtunk, a kezelőpanel valós időben kommunikál a kiszolgálóval. Három karakter megadása után automatikusan lefut a program. Ha a könyvtáros átírja a szövegdobozban megadott kifejezést, az alkalmazás azonnal módosítja a találati listát, ezért nincs szükség „Keresés” gombra.

5. Fejlesztési irányok

A kereső első verziója nyilvánosan is elérhető az SZTE Klebelsberg Könyvtár weboldalán³, így már nem csak a saját kollégáink munkáját tudja megkönnyíteni.



6. ábra. A Cutter-kereső felülete az SZTE Klebelsberg Könyvtár weboldalán

Mindazonáltal van egy apró hiba, ami még javításra szorul a programban. A napi munkavégzés során derült ki, hogy egyes szóközös kifejezések - nevezetesen azok, ahol a szóközők jelöletlenül, a zárótagok végén vannak - az átalakításnál rejtve maradtak, így ezeknél a soroknál két találatot kap a felhasználó. Utólag összesen hét hibás sort azonosítottunk, ezek javítása még folyamatban van.

Ezen felül célszerű lenne elindulni a munkafolyamat teljes automatizálása felé. Hasznos lenne, ha a Cutter-keresés funkció beépülne az integrált könyvtári rendszerek katalogizáló moduljaiba, így a MARC-rekordokból kinyert információk alapján, további emberi beavatkozás nélkül, automatikusan generálhatóvá válna a megfelelő Cutter-szám. Ez a funkció egyes rendszerekben (például a LinLib⁴-ben vagy a TextLib⁵-ben) már elérhető, de a legtöbb IKR-ben még nem vált alapszolgáltatássá.

3 Cutter-kereső az az SZTE Klebelsberg Könyvtár weboldalán. <http://www.cutter.bibl.u-szeged.hu> Hozzáférés: 2023.06.15.

4 LinLib integrált könyvtári rendszer. <http://www.linlib.hu/konyvtarirendszer.htm> Hozzáférés: 2023.07.31.

5 TextLib. A Cutter mező kitöltése. <https://www.textlib.hu/html/cutter.htm> Hozzáférés: 2023.07.31.

The background is a complex digital artwork. It features a grid of squares, each containing a different texture or color. The colors transition from warm oranges and yellows on the left to cool blues and teals on the right. A bright, glowing light source is positioned in the center, creating a lens flare effect that radiates across the grid. The overall effect is one of depth and digital complexity.

ÚJ TECHNOLÓGIÁKKAL, ÚJ TARTALMAKKAL A JÖVŐ DIGITÁLIS TRANSZFORMÁCIÓJA FELÉ

32. Networkshop: országos konferencia

2023. április 12–14.

Pannon Egyetem, Veszprém

ÚJ TECHNOLÓGIÁKKAL, ÚJ TARTALMAKKAL A JÖVŐ DIGITÁLIS TRANSZFORMÁCIÓJA FELÉ

32. Networkshop: országos konferencia

2023. április 12–14.
Pannon Egyetem, Veszprém

Szerkesztette: Tick József, Kokas Károly, Holl András

HUNGARNET Egyesület
Budapest, 2023



Szerkesztette: Tick József, Kokas Károly, Holl András

Tipográfia és tördelés: Vas Viktória

Workshop

2023. április 12–14. Pannon Egyetem, Veszprém konferencia előadásainak közleményei

ISBN 978-615-82243-1-4

DOI: [10.31915/NWS.2023](https://doi.org/10.31915/NWS.2023)

Kiadja a HUNGARNET Egyesület
az MTA Könyvtár és Információs Központ közreműködésével

Budapest

2023

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

TARTALOMJEGYZÉK

Előszó.....	5
Király Sándor, Balla Tamás: Flipped classroom az sqlsuli.hu-ban.....	7
Wirágh András: Abaújszántótól Zombolyáig. Megjegyzések egy új sajtóadatbázishoz	14
Albert Ágota Katalin: Az EGT-tagállamok adatvédelmi felügyeleti hatóságainak szankcionálási gyakorlata az oktatási szektorban a GDPR alkalmazása óta	19
Simon András: Digitális dokumentumok gyűjteménykezelési gyakorlatának támogatása a digitális tartalmak számossága, mérete és féleségeik vizsgálatával	24
Bódog András: Az Annif gépi tárgyszavazó rendszer magyarországi adaptációjának feltételei és lehetőségei	31
Dezső Krisztina: A Pécsi Egyetemtörténeti Gyűjtemény online adatbázisai és digitális gyűjteményei	36
Ungváry Rudolf, Király Péter: Nemzeti könyvtárak és az OSZK MARC21 állományainak összehasonlító elemzése néhány adatmező alapján	42
Szemes-Révész Enikő Evelin: Kapocs a tudáshoz – A könyvtár szerepe a civilek és a tudomány kapcsolatában	50
Tóth Zoltán: Az RO-Crate alapú kutatási objektum csomagolás keretrendszere az ELKH ARP platformban	54
Király Roland, Király Sándor, Palotai Martin Marcell: Neurális hálózatok oktatási alkalmazását támogató keretrendszer Virtual (VR) és Augmented Reality (AR) eszközökkel	60
T. Nagy László: Mesterséges intelligencia, multimédia, tanulástámogatás	69
Horváth Péter: Egy automatikusan generált rímshótár fejlesztése és a magyar kanonikus költészet rímshavainak néhány jellemzője	77
Héjja Balázs, Tóth-Jávorka Brigitta, Tóth Máté: Digitális tartalomfejlesztés közkönyvtári környezetben	85
Koczka Ferenc: Szemelvények egy felsőoktatási rendszer informatikai védelmének tapasztalataiból	91
Bolya Mátyás: A digitális gyűjtésrekonstrukció lehetőségei: az Ethiofolk projekt	99
Dobás Kata, Sidó Zsuzsa, Szabó-Reznek Eszter: A Kolozsvári Állami Magyar Színház jelmezterveinek digitalizációja és felvitele az ITIdata adatbázisba	108
Köpösdí Zsuzsa: H5P-ben létrehozható interaktív és adaptív tananyagok	116
Fülöp Tiffany, Molnár Tamás, Hoczapán Szabolcs: Komplex kutatástámogató szolgáltatási portfólió az SZTE Klebelsberg Könyvtárban	122
Vass Johanna: Az Open Science könyvtári vonatkozásai	129
Antal Péter, Czeglédi László: A digitális oktatás módszertana a gyakorlatban	135
Máray Tamás: A szuperszámítástechnika mint európai stratégiai ágazat	143
Frankó Máté, Zeller Rozália: Szoftveres Cutter-keresés az SZTE Klebelsberg Könyvtárban	151
Zsiborács Judit, Dési Ádám Dániel, Nagy Attila Árpád, Urbán Katalin: Tudományometriai műhely könyvtári környezetben	157



Palkó Gábor, Szekrényes István, Bobák Barbara: A Digitális Örökség Nemzeti Laboratórium webszolgáltatásai automatikus kézírás-felismertetéshez	164
Szűcs Kata Ágnes: Adatvizualizációs lehetőségek a bölcsészettudományban	170
Leitgéb Mária: A BME Építészettörténeti és Műemléki Tanszék repozitóriuma	178
Mihály Eszter, Micsik András: Szerkesztői környezet TEI-alapú szövegkiadásokhoz	186
Dobás Kata, Fellegi Zsófia, Palkó Gábor: A kis gömböc meséje - az ITIdata irodalomtudományos adatbázis fejlesztése 2022–2023-ban	192
Alföldi István, Szemigán Dorottya Henrietta, Palkó Gábor, Fellegi Zsófia: Kutatói e-mail hagyaték archiválása és feldolgozása	199