

Správa z konferencie
Digital humanities – nástroje sprístupňovania historického dedičstva

DOI: <https://doi.org/10.24040/ahn.2022.25.02.103-109>

Jedným z hlavných poslanií historikov a archivárov je ochrana historického dedičstva a jeho následné sprístupnenie odbornej i laickej verejnosti, ktorá má o históriu záujem. To sa týka aj písomného dedičstva, ktorého sprístupňovanie rôznymi inovatívnymi spôsobmi bolo hlavným bodom konferencie *Digital humanities – nástroje sprístupňovania historického dedičstva*. Konferencia prebiehala v dňoch 12. a 13. októbra 2022 v Banskej Bystrici, v historických priestoroch Župného domu, kde dnes sídli Štátna vedecká knižnica, ktorá bola spolu s Filozofickou fakultou Univerzity Mateja Bela (FF UMB) jej organizátorom.

Toto podujatie a zborník abstraktov¹, ktorý dostali účastníci konferencie k dispozícii, boli výstupom z projektu APVV -19-0456 SKRIPTOR – Inovatívne sprístupnenie písomného dedičstva Slovenska prostredníctvom systému automatickej transkripcie historických rukopisov. Hlavným cieľom projektu SKRIPTOR je tvorba modelov automatickej transkripcie, ktoré rezonovali vo viacerých príspevkoch konferencie. Na to slúži inovatívny nástroj na transkripciu historických rukopisných dokumentov s názvom *Transkribus*, ktorý funguje s cieľom digitalizácie, prepisu a editovania historických dokumentov.

Referáty hostí sa členili do dvoch základných blokov. Prvý niesol názov *Digital humanities* a ponúkal komplexný pohľad na digitalizáciu a modernizáciu elektronických služieb v rámci pamäťových inštitúcií, ktoré mali na konferencii zastúpenie (štátne archívy, Ústredná knižnica SAV, Slovenský bankský archív a pod.). V tomto bloku bol taktiež reflektovaný stav digitalizácie prameňov (J. Šedivý) a možnosti vyučovania *Digital humanities* v rámci štúdia historických vied (P. Labanc). Druhým blokom bol blok *Transkribus*, v ktorom vystupujúci v príspevkoch predstavili výsledky svojej výskumnej činnosti a práce s platformou *Transkribus* na konkrétnych písomných prameňoch.

Prvý deň konferencie začal úvodným slovom prof. PhDr. Dušan Kaťaščák, PhD., v ktorom zhrnul účel konferencie a výskumné ciele projektu v kooperácii s platformou *Transkribus*. Taktiež sa vo svojom vystúpení

¹ MALINIÁK, Pavol – NAGY, Imrich (eds.): *Digital humanities : nástroje sprístupňovania historického dedičstva : zborník abstraktov*. Banská Bystrica : Štátna vedecká knižnica v Banskej Bystrici, 2022, 73 s. Dostupné na: <https://repo.umb.sk/handle/123456789/39> [cit. 20. 10. 2022]

venoval doterajšiemu úsiliu o digitalizáciu písomných prameňov na Slovensku a v okolitých štátoch.

Ďalej nasledoval prvý referát z bloku *Digital humanities* s názvom *Digitalizácia a modernizácia elektronických služieb štátnych archívov v európskom kontexte*, ktorý predniesla Mgr. Monika Péková a Dan Ryška z Ministerstva vnútra SR. Vykreslili v ňom význam a aktuálny stav digitalizácie v štátnych archívoch. Prezentovali Národnú koncepciu rozvoja štátnych archívov s výhľadom do roku 2028 a jej následné ciele. Z nich možno spomenúť vytvorenie podmienok na komplexnú digitalizáciu novoprichádzajúceho archívneho obsahu, dobudovanie funkcionálov a kapacít Elektronického archívu Slovenska, modernizáciu bádateľní o príručné skenery a počítače a pod. Zaoberali sa aj úmyslom Ministerstva vnútra SR zriadiť desať digitalizačných pracovísk a vybaviť ich potrebnou technikou a infraštruktúrou. Tieto plány odôvodnili snahou o skvalitnenie služieb občanovi, šetrením finančných prostriedkov a ochranou životného prostredia.

Svojím vystúpením ďalej pokračovala doc. Mgr. Ivona Kollárová, PhD., ktorá pôsobí v Ústrednej knižnici SAV. Jej príspevok niesol názov *Digitalizácia historických zbierok v Ústrednej knižnici SAV v. v. i. a reflektoval základné míľniky vývoja digitalizácie v tejto inštitúcii. Patrí medzi ne začiatok spolupráce s Centrom vedecko-technických informácií SR či partnerstvo so spoločnosťou Piql Slovakia, ktorá sa zaoberá inovatívnou bezmigracnou technológiou digitálneho filmu. V referáte bola spomenutá podpora Európskej únie ohľadom projektu *eArchiving of Engineering and Science Library*. Cieľom tohto projektu je vývoj systémov digitálneho uchovávanía dát.*

Ing. Mgr. Juraj Michelík zo špecializovaného pracoviska Slovenský banský archív, spadajúceho pod Slovenský národný archív, predstavil *Fond Hlavného komorskogrófskeho úradu v Banskej Štiavnici a jeho digitalizáciu*. V referáte vyzdvihol význam samotného fondu pre historický výskum a prezentoval stav jeho digitalizácie po desaťročnom fungovaní špecializovaného digitalizačného pracoviska s veľkoformátovým skenerom.

Tento blok konferencie svojim vystúpením obohatil aj prof. PhDr. Juraj Šedivý, MAS, PhD., ktorý pôsobí na Katedre archívnictva a muzeológie FF Univerzity Komenského v Bratislave. Jeho referát niesol názov *Digitalizácia prameňov k dejinám Bratislavy* a predstavil v ňom jednotlivé fázy digitalizácie prameňov na Slovensku, ktorú z hľadiska záberu prameňov rozlíšil na 1. inštitucionálnu, 2. rezortnú, 3. tematickú, 4. komplexnú. Upozornil na problematiku rezortných databáz a chýbajúceho komplexného metaportálu. Tu kritizoval portál *Slovakiana*, pri ktorom zdôraznil zlú kategorizáciu, ktorá neumožňuje vyhľadávanie konkrétneho prameňa. Ďalej porovnával jednotlivé pamäťové portály z hľadiska

vyhľadávania prameňov k dejinám Bratislavy. Vyzdvihol portál *Pam-Map.sk*, a to z dôvodu množstva privátnych prameňov. Cieľom jeho referátu bolo porovnať jednotlivé portály a podrobiť ich kritike koncového užívateľa.



Obr. 1. Prof. PhDr. Dušan Katuščák, PhD. a doc. Mgr. Imrich Nagy, PhD. otvárajú konferenciu Digital humanities. Autor fotografie Dušan Jarina.

Ďalším vystupujúcim bol doc. Mgr. Peter Labanc, PhD., ktorý otvoril zaujímavú tému *Možnosti, výhody, prínosy, prekážky a nástrahy vyučovania Digital humanities v historických vedách*. Vo svojom vystúpení vytýčil niekoľko postrehov z príprav vysokoškolského vyučovacieho predmetu, ktorého náplňou sú práve Digital humanities a ich inovácie. Ide o dvojsemestrálny predmet, ktorý je začlenený do odporúčaného študijného programu na Katedre histórie a Katedre klasickej archeológie Filozofickej fakulty Trnavskej univerzity v Trnave so začatím výučby v akademickom roku 2022/2023. Zdôraznil finančnú náročnosť projektu, a to najmä z hľadiska zabezpečenia IT nástrojov a infraštruktúry, ale aj jeho obrovskú pridanú hodnotu a potrebu vzdelávania budúcich historikov a archeológov i v tejto oblasti.

Posledným z referátov na tému Digital humanities bol príspevok zahraničného host'a. Bol ním Zoltán Szatucsek z Maďarského národného archívu v Budapešti, ktorý sa venoval problematike digitalizácie prameňov. Vo svojom príspevku vyzdvihol potrebu automatizovaných prostriedkov

transkripcie a spoluprácu s dobrovoľníkmi počas procesu transkripcie rukopisov. Venoval sa aj platforme *Transkribus*, jej prínosom, ale aj obmedzeniam a otvoril možnosť vývoja vlastného systému, ktorý by slúžil na účely digitálnej transkripcie. Z obmedzení upozornil najmä na najmä ťažkú prácu s tabuľkami alebo nepoužívanie jazykového modelu.

Prvý deň konferencie zakončilo vystúpenie Mgr. Michaely Mikuškovej a Mgr. Lucie Nižníkovej z Univerzitetnej knižnice Univerzity Mateja Bela v Banskej Bystrici na tému *Využitie softvéru Transkribus na automatickú transliteráciu štyroch typov fontu štvorjazyčnej historickej tlače*. Zároveň sa týmto referátom začala druhá časť konferencie venovaná už konkrétnejšie práci s platformou *Transkribus*. Vystupujúce vo svojej výskumnej činnosti pracovali s dielom J. A. Komenského *Orbis Pictus*, ktoré je napísané v štyroch jazykoch, a tým ponúka zaujímavú možnosť na overenie schopností *Transkribusu*. Autorky uprednostnili metódu transliterácie dokumentu s cieľom zachovať pôvodné grafémy. Následne oboznámili prítomných s tréňovaním jednotlivých modelov v *Transkribuse* a ich výsledkami.

Druhý deň podujatia začal referátom prof. PhDr. Dušan Katuščák, PhD., ktorý sa venoval téme *Transkripcia historických dokumentov v kontexte Digital humanities*. Profesor Katuščák pôsobí v Štátnej vedeckej knižnici a na Sliezskej univerzite v Opave. Autor sa venoval téme transkripcie a niektorým konkrétnym výsledkom transkripcie dokumentov prostredníctvom projektu SKRIPTOR. Ďalej spomenul metodologický kontext transkripcie v Digital humanities a globálne trendy v oblasti dokumentácie v pamäťových a fondových inštitúciách, príp. systémoch. Taktiež na základe svojich skúseností s platformou *Transkribus* opísal jednotlivé kroky a postupy transkripcie dokumentov a s tým spojené náklady. Venoval sa konkrétnym dokumentom, s ktorými pracoval, výsledkom ich transkripcie (časopis *Jitřenka*, publikácia *Jánošík*) a modelom transkripcie. Zároveň načrtnol úvahu o inštitucionalizovaní transkripcie historických dokumentov na Slovensku.

Doc. PhDr. Pavol Maliniak, PhD., ktorý je pracovníkom Katedry histórie FF UMB predniesol príspevok s názvom *Automatická transliterácia rukopisnej kazateľskej tvorby Izáka Abrahamidesa (prvé modely, predpoklady ďalšieho postupu)*. Čo sa týka transliterovaného dokumentu, tak ide o zbierku kázni, ktorá nebola doteraz bádateľsky sprístupnená v edícii, čím ponúka širokú perspektívu na prácu v *Transkribuse*. Zároveň ide o text písaný rovnakým rukopisom v krátkom časovom úseku, čo zvyrazňuje potenciál dokumentu na spracovanie v tejto platforme. Jazykom, v ktorom je písaný, je slovakizovaná čeština (vyskytujú sa aj germanizmy a hungarizmy). Vyskytujú sa v ňom pasáže v latinčine a gréčtine, príp. hebrejčine, ale autor sa rozhodol grécke a hebrejské pasáže do tréňovania

modelov na transliteráciu nezahrnúť z dôvodu ich ojedinelosti. Následne vystupujúci oboznámil prítomných s fázami transliterácie a výsledkami práce s jednotlivými modelmi v systéme. Výsledkom bolo zistenie, že efektívnosť modelu sa zvyšuje s počtom slov, ktoré obsahuje cvičná vzorka. Potvrdilo sa tým odporúčanie vývojárov *Transkribusu* pracovať aspoň s 15 000 slovami vzhľadom na paleografickú náročnosť dokumentu.

S ďalším príspevkom vystúpila Bc. Klára Kováčová, študentka Ústavu bohemistiky a knihovníctva na Filozoficko-prírodovedeckej fakulte Slezskej univerzity v Opave. Autorka sa zaoberala transkripciou v platforme *Transkribus* na príklade vzácnej kuchárskej knihy z roku 1667. Hlavným cieľom tohto referátu bolo poukázať na možnosti uplatnenia nástrojov umelej inteligencie v digitalizácii, rozpoznávania textu pomocou umelej inteligencie, prepisu a vyhľadávania historických dokumentov v akomkoľvek jazyku v rámci transkripcie. Samotný dokument, ktorý autorka použila, má vyše 800 strán a je písaný kurentom. Na tomto príklade bola overená možnosť práce s tzv. zariadením ScanTent, veľmi nápomocným pri vytváraní digitálnych kópií dokumentov. Samotný výskum bol súčasťou bakalárskej práce autorky.

V bloku *Transkribus* pokračoval Mgr. Martin Katreniak z Novohradského múzea a galérie v Lučenci a Mgr. Patrik Kunec, PhD., pracovník Katedry histórie FF UMB. Spolu predniesli referát na tému *Automatická transkripcia historických prameňov obsahujúcich viac rukopisov na príklade kanonických vizitácií*. Zdôraznili uplatniteľnosť automatickej transkripcie pri prepisoch špecifických prameňov k cirkevným dejinám, keďže poskytujú historikom zaujímavé informácie nielen o podobe náboženského života, ale aj ohľadom iných oblastí. Na automatickú transkripciu v platforme *Transkribus* autori vybrali vizitačné protokoly z rokov 1756 a 1783, ktoré sa skladali z viacerých rukopisov. Hlavným cieľom výskumu bolo zistiť možnosti aplikovania metódy automatickej transkripcie na daný dokument a jeho špecifiká. Ďalej sa autori zaoberali chybovosťou transkripcie na základe jednotlivých modelov a následne sa pokúsili dopracovať k jednotnému modelu, ktorý by vykazoval uspokojivé výsledky. Taktiež bola preskúmaná možnosť použitia už vytvoreného tzv. základného modelu (base model). Výsledky výskumu priniesli zistenie, že pri transkripcii dokumentov pozostávajúcich z viacerých rukopisov sa javí ako efektívnejšia alternatíva oproti samostatnému trénovaniu skôr spoločné trénovanie. Výhodným sa zdá aj využívanie samotného base modelu. Autori avizovali pokračovanie vo svojej výskumnej činnosti týmito postupmi.

Nasledovalo vystúpenie PhDr. Ota Tomečka, PhD. z Katedry histórie FF UMB, ktorý sa venoval téme *Vytváranie modelu na automatickú transkripciu novolatinského rukopisu reambulačného protokolu Banskej Bystrice*

v prostredí platformy *Transkribus*. Autor považoval za veľký benefit automatickej transkripcie dokumentov ich eventuálne sprístupnenie odbornej a širšej laickej verejnosti v zrozumiteľnej, čitateľnej podobe. Pri automatickej transkripcii dokumentu s názvom *Metales* bola použitá platforma *Transkribus*. Následne sa vystupujúci venoval jednotlivým fázam práce s dokumentom v *Transkribuse* aj s použitím base modelu a predstavil jej výsledky. Hlavným výsledkom bolo vytvorenie modelu, ktorý je možné použiť na automatickú transkripciu daného dokumentu.

Mgr. Mária Bôbová, PhD., pôsobiaca v Štátnej vedeckej knižnici v Banskej Bystrici, predstavila *Nové možnosti digitálneho prostredia vo výskume dejín knižnej kultúry*. V referáte sa zaoberala procesom tvorby modelu rukopisu, využívajúc platformu *Transkribus*. Objektom autorkinho výskumu bol historický rukopisný katalóg cirkevnej knižnice Kňazského seminára sv. Karola Boromejského v Banskej Bystrici.

Mgr. Alica Kurhajcová, PhD. z Katedry histórie FF UMB predniesla referát venujúci sa automatickému rozpoznávaniu rukopisu J. M. Hurbana v platforme *Transkribus*. Zdôraznila potenciál tohto nástroja pri zverejňovaní a editovaní korešpondencie slovenských osobností 19. storočia. Autorka sa konkrétne venovala listom Jozefa Miloslava Hurbana, ktoré boli napísané najmä v dobovej slovenčine, češtine, ale občas boli obohatené o latinské a maďarské výrazy alebo nárečové slová z okolia Myjavy. Do latinčiny písaných textov vstupuje ojedinele nemčina písaná kurentom, prípadne ruština a srbčina zapísaná azbukou. Vo vystúpení autorka predstavila hlavné etapy práce a jednotlivé postupy, ktoré použila pri tvorbe konkrétneho modelu na automatickú transkripciu spomínaného dokumentu. Ďalej sa zaoberala mierou chybovosti pri trénovaní modelu, jej následným znižovaním a stavom chybovosti v záverečnej fáze procesu, kde overila funkčnosť modelu na doposiaľ nespracovaných stranách.

Konferenciu zakončil svojim vystúpením jeden z hlavných organizátorov podujatia doc. Mgr. Imrich Nagy, PhD. z Katedry histórie FF UMB, ktorý sa venoval práci s Csákósovým katalógom korešpondencie Koháryovcov. V referáte zdôraznil problematiku manuálnej segmentácie jednotlivých strán, ktorú pokladal za najnáročnejšiu fázu automatickej transkripcie, a to najmä z časového hľadiska. Ďalej oboznámil poslucháčov s procesom trénovania jednotlivých modelov a úrovňou ich chybovosti. Tá sa menila v závislosti od počtu ručne transkribovaných strán zaradených do cvičného súboru. Podľa autora už pri chybovosti modelu okolo 3% môžeme hovoriť o praktickej využiteľnosti *Transkribusu* pri danom dokumente. Najviac chýb sa vyskytovalo v oblasti interpunkcie a diakritiky, čo ale nemusí ovplyvňovať zrozumiteľnosť textu nad únosnú mieru. Autor sa venoval aj nákladom na automatický prepis s konkrétnymi odhadmi pri transkripcii spomínaného katalógu. Tie odhadol

v rozmedzí 0,18 € – 0,33 € za spracovanú snímku digitalizátu obsahujúceho dvojstranu.

Súčasťou konferencie bol spoločenský večer, na ktorom si účastníci neformálnymi rozhovormi vymieňali skúsenosti z ich vedecko-výskumnej činnosti a nadväzovali nové pracovné kontakty. Na záver podujatia sa uskutočnila exkurzia do Digitalizačného centra Slovenskej národnej galérie v Banskej Bystrici. *Konferencia Digital humanities – nástroje sprístupňovania historického dedičstva* zanechala pozitívny dojem a v rámci diskusie nastolila viacero tém ohľadom ďalšieho smerovania výskumu, sprístupňovania a využívania digitalizovaného písomného historického dedičstva.

Matej Smida