UNIVERZITA MATEJA BELA V BANSKEJ BYSTRICI
FILOZOFICKÁ FAKULTA

AUTOMATICKÁ TRANSKRIPCIA RUKOPISNÝCH HISTORICKÝCH TEXTOV NA PRÍKLADE VYBRANÝCH KANONICKÝCH VIZITÁCIÍ

Diplomová práca

8ac9b304-af0a-4c4a-b866-e7f852e8709d

Študijný program : História

Študijný odbor: Historické vedy

Pracovisko: Katedra histórie Filozofickej fakulty Univerzity Mateja Bela

Vedúci diplomovej práce: PhDr. Oto Tomeček, PhD.

Banská Bystrica 2022 **Bc. Martin Katreniak**

ČESTNÉ VYHLÁSENIE

Čestne vyhlasujem, že som svoju diplomovú prácu s názvom Automatická transkripcia rukopisných historických textov na príklade vybraných kanonických vizitácií vypracoval samostatne, pod odborným vedením vedúceho práce, s použitím uvedenej literatúry.

Banská Bystrica 2022 .....

Martin Katreniak

POĎAKOVANIE

Chcel by som sa poďakovať najmä vedúcemu mojej diplomovej práce PhDr. Otovi Tomečkovi, PhD. za poskytnutie potrebných materiálov, za cenné rady a pripomienky pri tvorbe tejto práce ako aj za jeho profesionálny prístup.

ABSTRAKT

KATRENIAK, Martin: *Automatická transkripcia rukopisných historických textov na príklade vybraných kanonických vizitácií.* [Diplomová práca]. Univerzita Mateja Bela v Banskej Bystrici. Filozofická fakulta. Katedra histórie. Školiteľ: PhDr. Oto Tomeček, PhD. Stupeň odbornej kvalifikácie: Magister. Banská Bystrica : Filozofická fakulta UMB, 2022. 79 s.

Diplomová práca sa zaoberá možnosťami metódy automatickej transkripcie rukopisných historických textov. Hlavný dôraz sa kladie na preskúmanie možností a postupov automatickej transkripcie dokumentov, skladajúcich sa z viacerých rukopisov. Práca sa tiež sústreďuje na možnosti využívania tzv. „malých modelov", vytvorených na čo najmenšej tréningovej vzorke. Diplomová práca predstavuje rozličné postupy automatickej transkripcie, pričom ponúka ich podrobný rozbor, potenciálne využitie a analýzu ich chybovosti. Výskum je prevádzaný využitím platformy Transkribus, ktorá slúži ako program na automatické rozpoznávanie historických rukopisov. V rámci jednotlivých pokusov práca prináša niekoľko podstatných zistení. Pri vytváraní modelu automatickej transkripcie boli dosiahnuté veľmi dobré výsledky aj pri využití podstatne menšej vzorky textu ako odporúčajú tvorcovia Transkribusu. Spoločné trénovanie viacerých rukopisov sa ukázalo v mnohých smeroch efektívnejšie oproti ich samostatnému trénovaniu. Využitie base modelu sa preukázalo ako jednoznačne výhodné pri vytváraní nových modelov založených na jednom, ale aj viacerých rukopisoch. Práca je rozdelená do štyroch kapitol a obsahuje štrnásť obrazových príloh.

Kľúčové slová: Kanonická vizitácia. Automatická transkripcia. Transliterácia. Transkribus. Špania Dolina. Staré Hory. Motyčky.

ABSTRACT

KATRENIAK, Martin: *Automatic transcription of manuscript historical texts on the example of selected canonical vistitations*. [Master thesis]. Matej Bel University, Banská Bystrica. Faculty of Arts. Department of History. Supervisor: PhDr. Oto Tomeček, PhD. Professional qualification degree: Master. Banská Bystrica : Faculty of Arts UMB, 2022. 79 p.

The Master thesis deals with the possibilities of the method of automatic transcription of manuscript historical texts. The main focus is on the investigation of the possibilities and procedures of automatic transcription of documents consisting of several manuscripts. The thesis also concentrates on the possibilities of using so-called "small models", created on the smallest possible training sample. This thesis presents different automatic transcription procedures, offering a detailed analysis of them, their potential applications and an analysis of their error rates. The research is carried out using the Transkribus platform, which serves as a program for automatic recognition of historical manuscripts. Within the individual experiments, the thesis presents several significant findings. Very good results have been obtained in the development of the automatic transcription models, even when using a much smaller sample of text than recommended by the creators of Transkribus. Joint training of multiple manuscripts proved to be more effective in many ways compared to their individual training. The use of base model proved to be clearly advantageous when creating new models based on both single and multiple manuscripts. The thesis is divided into four chapters and contains fourteen pictorial attachments.

Keywords: Canonical visitation. Automatic transcription. Transliteration. Transkribus. Špania Dolina. Staré Hory. Motyčky.

PREDHOVOR

K napísaniu tejto práce ma viedlo niekoľko dôvodov. Zrejme tým najvýznamnejším bol môj dlhodobý osobný záujem o tému digitalizácie historických dokumentov a sprístupňovania archívneho dedičstva. Myslím si, že postupná systematická digitalizácia slovenského archívneho dedičstva by mala byť v aktuálnej dobe jedným z popredných cieľov v rámci historického výskumu a archívnictva. Postupnou digitalizáciou môžeme sprístupniť historické pramene, ktoré boli doposiaľ minimálne využité, resp. prehliadané a otvoriť tak brány úplne novým výskumom. Zdigitalizované pramene ponúkajú omnoho efektívnejší a jednoduchší prístup, ktorý môže byť pre mnohých bádateľov lákavejší. Význam digitalizácie a sprístupňovania archívneho dedičstva sa výrazne preukázal najmä v posledných rokoch, pod vplyvom globálnej pandémie, z dôvodu ktorej dochádzalo k dlhodobému zatváraniu archívov a knižníc. Prístup k archívnym dokumentom sa tak stal extrémne komplikovaný, až nemožný. Výsledkom bolo výrazné ochromenie práce bádateľov, ktorých výskum bol závislý od priamej práce s archívnymi dokumentmi. Ukázalo sa, aké cenné môžu byť práve digitálne zbierky a archívy, ktoré by boli dostupné nepretržite.

Druhým dôvodom pre spracovanie tejto témy bol môj záujem o paleografiu, ktorý vzrástol najmä vďaka absolvovanému predmetu *Základy paleografie*. V tomto čase sa navyše Katedra histórie Filozofickej fakulty UMB stala riešiteľom projektu s názvom SKRIPTOR, ktorého zameraním je inovatívne sprístupnenie písomného dedičstva Slovenska prostredníctvom systému automatickej transkripcie historických rukopisov. Tento projekt teda spájal na jednej strane digitalizáciu a sprístupňovanie historických prameňov a na strane druhej paleografiu reprezentovanú transkripciou. V rámci tohto projektu navyše vznikol priestor na vypracovanie záverečnej práce so zameraním na automatickú transkripciu historických rukopisných textov. Keď sa teda naskytla možnosť spracovať túto tému v podobe záverečnej práce, rozhodol som sa ju využiť. Jedná sa o tému, ktorá spája dva výskumné smery, o ktoré sa zaujímam. Pozitívum pre mňa predstavoval aj fakt, že sa jedná o prakticky zameranú prácu, ktorá pozostávala z priamej práce s historickými prameňmi.

OBSAH

Úvod 8
1. Kanonická vizitácia ako historický prameň 13
1.1 Vznik a priebeh kanonických vizitácií 14
1.2 Štruktúra kanonických vizitácií 17
1.3 Využitie kanonických vizitácií ako historického prameňa 19
1.4 História kanonických vizitácií 24
2 Vybrané kanonické vizitácie 28
2.1 Cirkevný vývoj a situácia na území vybraných farností 28
2.2 Vnútorná štruktúra vybraných kanonických vizitácií z roku 1783 33
3. Metóda automatickej transkripcie a možnosti jej využitia v historickom výskume 39
3.1 Postup pri využití metódy automatickej transkripcie 39
3.1.1 Využitie metódy automatickej transkripcie v historickom výskume 41
3.2 Nahrávanie dokumentov, segmentácia a vytváranie Ground Truth vzorky 44
4. Automatická transkripcia vybraných kanonických vizitácií 49
4.1 Vytváranie referenčných modelov 50
4.2 Vytváranie spoločného modelu na základe viacerých rukopisov 54
4.3 Využitie base modelu pri vytváraní modelov 59
4.4 Porovnanie chybovosti a efektivity využitých postupov 62
Záver 66
Zoznam bibliografických odkazov 69
Zoznam obrazových príloh 72
Obrazová príloha 73

Úvod

Digitalizácia a sprístupňovanie historických prameňov je téma, ktorá v posledných rokoch značne naberá na popularite. Systematické sprístupňovanie archívneho dedičstva môže postupne otvárať brány úplne novým výskumom na poli histórie a mať tak celkovo pozitívny vplyv na rozvoj slovenského historického výskumu. Mnohé archívne fondy zostávajú dlhodobo nespracované, čo výrazným spôsobom komplikuje prácu s nimi. Bádateľská činnosť s týmito dokumentami sa tak stáva časovo veľmi náročná a značne neefektívna. Vzniká tak situácia, kedy sú archívy častokrát plné veľkého množstva obsiahlych historických rukopisných dokumentov, ktoré zostávajú prakticky nevyužité. Práve odstránenie týchto problémov je jedným z hlavných prínosov digitalizácie. Zdigitalizované archívne zbierky a elektronické archívy predstavujú podstatný krok k celkovému sprístupneniu historických prameňov, či už odbornému publiku, alebo širokej verejnosti. Prednosťou takýchto zbierok je najmä celková dostupnosť a výrazná efektivita pri práci. Úspešne tak môžu eliminovať niektoré z aktuálnych problémov pri skúmaní archívnych historických dokumentov.

V našej práci sa zameriame na možnosti využitia špecifického prístupu k digitalizácii a sprístupňovaniu historických rukopisných prameňov, ktorým je automatická transkripcia. Automatická transkripcia predstavuje metódu automatického textového rozpoznávania historických rukopisov. Najväčším prínosom tejto metódy v rámci digitalizácie je sprístupňovanie historických rukopisných dokumentov, z ktorých mnohé môžu byť pre bádateľov nečitateľné alebo nezrozumiteľné a teda aj práca s nimi značne problematická. V skratke táto metóda umožňuje efektívne transformovať ťažko čitateľný historický rukopisný text do modernej digitálnej podoby. Práca s takýmto dokumentom je následne dostupnejšia, jednoduchšia a časovo efektívnejšia. Výskum spojený s automatickou transkripciou budeme v rámci tejto práce vykonávať za pomoci platformy Transkribus. Možnosti tejto metódy budeme skúmať na príklade dokumentov kanonických vizitácií. Tie sú pre nás obzvlášť zaujímavé z dôvodu, že typ rukopisu sa v rámci kníh kanonických vizitácií často mení, čo môže predstavovať pre metódu automatickej transkripcie istú komplikáciu.

Problematike a výskumu kanonických vizitácií ako prameňa už bola v rámci slovenskej historiografie venovaná značná pozornosť. Zrejme najvýraznejšie sa s týmto výskumom spája meno Z. Lopatkovej, ktorá je autorkou celej rady vedeckých štúdií

venujúcich sa vývoju a využitiu kanonických vizitácií ako prameňa. Uviesť tu môžeme štúdiu *Kanonické vizitácie v slovenských dejinách,*[1] v ktorej sa venuje vývoju tohto prameňa v rámci dejín vzťahujúcich sa na územie dnešného Slovenska. Ďalšou jej významnou prácou v tomto smere je vedecká štúdia *Kanonická vizitácia – k jej interpretácii, limitom a metodologickým otázkam v historickom výskume*.[2] Jedná sa o veľmi prínosnú štúdiu zameranú na využívanie kanonických vizitácií ako historického prameňa. Práca poukazuje na príklade Smolenického dekanátu na možnosti toho, ako sa dá kanonická vizitácia ako historický prameň interpretovať, aké sú jej celkové limity v historickom výskume a približuje aj základné metodologické otázky a problémy, ktoré môžu vznikať pri práci s týmto druhom prameňa. Problematike kanonických vizitácií ako prameňa sa venovala aj M. Bizoňová v práci *Kanonické vizitácie na Spiši v 17. a prvej polovici 18. storočia ako historický prameň,*[3] a tiež A. Ostertagová v práci *Kanonické vizitácie ako historický prameň,*[4] kde veľmi prehľadne prezentuje rozličné smery výskumu, ktoré môžu čerpať a využívať kanonické vizitácie ako historický prameň. Kanonickým vizitáciám po stránke ich vývoja ako aj využitia sa venovala aj M. Kyseľová vo svojej často citovanej práci *Kanonické vizitácie na Spiši*.[5] V neposlednom rade je pri výskume kanonických vizitácií ako prameňa potrebné spomenúť aj obsiahly zborník *Kanonické vizitácie po Tridentskom koncile,*[6] ktorý obsahuje celú radu vedeckých štúdií venujúcich sa rozličným aspektom tohto prameňa.

Čo sa týka problematiky automatickej transkripcie rukopisných historických textov, jedná sa v rámci slovenského územia o doteraz pomerne málo preskúmanú tému. Výraznejšie sa tejto téme venuje profesor D. Katuščák, ktorý je jedným z popredných predstaviteľov tohto výskumu. Celkovo problematiku aplikácie informačných a komunikačných technológií v spoločenských a humanitných vedách a disciplínach rozoberá v štúdii *Digital humanities a automatická transkripcia rukopisných textov.*[7] V tejto

[1] LOPATKOVÁ, Z.: Kanonické vizitácie v slovenských dejinách. In: *Studia historica Tyrnaviensia. VII*. Trnava : Filozofická fakulta Trnavskej univerzity, 2008, s. 96 – 106.
[2] LOPATKOVÁ, Z.: Kanonická vizitácia – k jej interpretácií, limitom a metodologickým otázkam v historickom výskume. In: *Prameň – jeho funkcia, význam, interpretácia a limity v historickom výskume*. Nitra : Univerzita Konštantína Filozofa v Nitre, 2010, s. 69 – 77.
[3] BIZOŇOVÁ, M.: Kanonické vizitácie na Spiši v 17. a prvej polovici 18. storočia ako historický prameň. In: *6. študentská vedecká konferencia : zborník príspevkov*. Prešov, 2010, s. 459 – 469.
[4] OSTERTAGOVÁ, A.: Kanonické vizitácie ako historický prameň. In: *Kanonické vizitácie po tridentskom koncile*. Prešov : Vydavateľstvo Michala Vaška, 2007, s. 5 – 10.
[5] KYSEĽOVÁ, M.: Kanonické vizitácie na Spiši. In: *Slovenská archivistika*, roč. 18, 1983, č. 2, s. 110 – 128.
[6] HIŠEM, C. – FEDORČÁK, P. (eds.): *Kanonické vizitácie po Tridentskom koncile*. 1. vyd. Prešov : Vydavateľstva Michala Vaška, 2007, 231 s. ISBN 978-80-7165-639-5
[7] KATUŠČÁK, D.: Digital humanities a automatická transkripcia rukopisných textov. In: *Informačné technológie a knižnice*, 2020, č. 1, s. 6 – 16.

štúdii predstavuje aj samotnú platformu a nástroj automatickej transkripcie rukopisných textov – Transkribus. Svoje poznatky, získané pri experimentálnej transkripcii rukopisných listov Andreja Kmeťa zhrnul v štúdii *Automated Transcription of Handwritten Text: READ and TRANSKRIBUS (An Experiment with Transcribing Letters of Andrej Kmeť)*.[8] Problematike automatickej transkripcie sa venoval aj v štúdii *Pokrok v sprístupňovaní švabachu a rukopisného písomného dedičstva s použitím umelej inteligencie*.[9] V tejto práci tiež približuje aktuálny stav výskumu automatickej transkripcie a dlhodobejšie ciele spojené s ním. V neposlednom rade ešte spomenieme vedeckú štúdiu I. Nagya *Možnosti aplikácie metódy digitálnej transkripcie historických rukopisných textov pri sprístupňovaní archívnych fondov*.[10] V tejto práci autor na príklade vlastného výskumu katalógu korešpondencie Kováryovcov predstavuje jednotlivé fázy prípravy modelu a následnej automatickej transkripcie daných dokumentov s využitím platformy Transkribus.

V rámci našej práce najskôr predstavíme kanonickú vizitáciu ako historický prameň, pričom sa zameriame najmä na možnosti jej využitia. Tiež priblížime rozličné údaje, ktoré sa dajú z tohto dokumentu získať a ich využitie v historickom výskume.

Naším hlavným cieľom bude najskôr predstaviť samotnú metódu automatickej transkripcie a postup, ktorý je potrebné dodržiavať pri jej využívaní. Následne zameriame pozornosť na preskúmanie možností automatickej transkripcie dokumentov založených na viacerých rukopisoch. Pokúsime sa zistiť, akým spôsobom je ideálne pristupovať k automatickej transkripcii takéhoto typu dokumentov. Ďalej sa zameriame na to, akým spôsobom ovplyvňuje pridávanie viacerých rukopisov celkovú presnosť modelu a ako sa rozličné rukopisy vzájomne ovplyvňujú pri vytváraní modelu. Po vykonaní automatickej transkripcie niekoľkých vybraných kanonických vizitácií vyhodnotíme chybovosť transkripcie, určíme kde vznikali chyby, akého charakteru boli a zhodnotíme s čím mala platforma Transkribus pri automatickej transkripcii problém. V neposlednom rade sa pokúsime zhodnotiť, ktorý z využitých postupov bol najefektívnejší pri automatickej

[8] KATUŠČÁK, D.: *Automated Transcription of Handwritten Text: READ and TRANSKRIBUS (An Experiment with Transcribing Letters of Andrej Kmeť)*, 2019. Dostupné online: https://www.researchgate.net/publication/336676993_Automated_Transcription_of_Handwritten_Text_READ_and_TRANSKRIBUS_An_Experiment_with_Transcribing_Letters_of_Andrej_Kmet

[9] KATUŠČÁK, D.: *Pokrok v sprístupňovaní švabachu a rukopisného písomného dedičstva s použitím umelej inteligencie*, 2021. Dostupné online: https://www.researchgate.net/publication/355586057_Pokrok_v_spristupnovani_svabachu_a_rukopisneho_pisomneho_dedicstva_s_pouzitim_umelej_inteligencie

[10] NAGY, I.: Možnosti aplikácie metódy digitálnej transkripcie historických rukopisných textov pri sprístupňovaní archívnych fondov. In: *Slovenská archivistika*, roč. 51, 2021, č. 2, s. 53 – 67.

transkripcii dokumentov skladajúcich sa z viacerých rukopisov a aké je praktické využitie týchto postupov.

Čiastkovým cieľom v rámci našej práce bude tiež preskúmanie možností vytvárania tzv. malých modelov, založených na čo najmenších cvičných vzorkách. Pokúsime sa zistiť, aké sú minimálne podmienky na vytvorenie funkčného a čo najpresnejšieho malého modelu, skladajúceho sa z jedného, ale aj viacerých rukopisov. K naplneniu týchto cieľov budeme v rámci nášho výskumu využívať ako skúmané dokumenty vybrané kanonické vizitácie.

Pri tvorbe tejto práce sme využili niekoľko základných metód. *Priamu* a *nepriamu* metódu rovnako ako aj *induktívnu* a *deduktívnu* metódu využívame vo všetkých etapách tejto práce. *Progresívnu metódu* sme využívali najmä pri výskume cirkevného vývoja na území, z ktorého pochádzajú naše skúmané vizitácie. Pri práci a analýze jednotlivých rukopisov kanonických vizitácií sme využili *komparatívnu metódu*, ktorá nám poslúžila aj pri rozbore štruktúry našich skúmaných vizitácií, ktorých časti a informácie obsiahnuté v nich sme porovnávali medzi sebou. Podstatnou metódou, ktorú sme pri našom výskume využívali bola *metóda paleografickej transliterácie*. Aplikovali sme ju pri odbornom prepise časti skúmaných kanonických vizitácií a vytváraní Ground Truth vzorky. V nadväznosti na ňu sme použili *metódu automatickej transkripcie*, ktorú sme aplikovali na naše skúmané rukopisné dokumenty a pomocou ktorej sme sa dopracovali k výsledkom nášho výskumu.

Práca sa skladá zo štyroch hlavných kapitol, ktoré sú rozdelené na príslušné podkapitoly. V rámci prvej kapitoly predstavíme kanonické vizitácie ako historický prameň. Táto kapitola bude slúžiť na priblíženie tohto typu prameňa po rôznych stránkach. Najskôr predstavíme spôsob, akým kanonické vizitácie vznikajú a ich základnú štruktúru, ktorá je väčšinou dodržaná. Následne sa zameriame na rôzne smery historického výskumu, v rámci ktorých si našiel tento prameň využitie. Na záver prvej kapitoly stručne predstavíme históriu a dejinný vývoj kanonických vizitácií.

Druhá kapitola predstavuje prechod od kanonických vizitácií vo všeobecnosti, ku špecifickým kanonickým vizitáciám, ktoré v rámci nášho výskumu využívame. Táto kapitola je rozdelená na dve podkapitoly. V rámci prvej podkapitoly sa v stručnosti venujeme cirkevnému vývoju na území, z ktorého pochádzajú využité kanonické vizitácie. Jedná sa o tri farnosti v Starohorskej oblasti, ktorými sú Špania Dolina, Staré Hory a Motyčky. V tejto podkapitole objasníme počiatky vzniku týchto farností, cirkevnú štruktúru, konfesionálnu situáciu a jej vývoj v rámci tohto územia približne do konca 18. storočia, kedy vznikli dané vizitácie. V rámci druhej podkapitoly priblížime špecifickú vnútornú štruktúru využívaných kanonických vizitácií. Predstavíme hlavné body, z ktorých

sa tieto vizitácie skladali, čo bolo ich obsahom a aké informácie sa z nich dali čerpať. V rámci jednotlivých bodov tiež uvedieme niekoľko príkladov priamo z vizitácií.

Tretia kapitola približuje metódu automatickej transkripcie a možnosti jej využitia v historickom výskume. Slúži na predstavenie postupu, ktorý je potrebné dodržiavať pri využívaní tejto metódy. V rámci tejto kapitoly tiež predstavíme základnú terminológiu a pravidlá, ktoré budeme počas nášho výskumu dodržiavať. Táto kapitola je rozdelená do dvoch podkapitol. V rámci prvej podkapitoly sa venujeme metóde automatickej transkripcie vo všeobecnosti. Predstavíme samotnú platformu Transkribus, pomocou ktorej vykonávame náš výskum a priblížime na akom princípe funguje. Následne v skratke vysvetlíme základné fázy automatickej transkripcie. V druhej časti tejto podkapitoly priblížime využitie tejto metódy v rámci historického výskumu a potenciál, ktorý vidíme v jej širšej implementácii. Druhá podkapitola už detailne opisuje náš postup v počiatočných fázach procesu automatickej transkripcie. Najskôr opisujeme ako sme pristúpili k nahraniu dokumentov do Transkribusu. Následne predstavujeme akým spôsobom sme vykonali proces segmentácie našich dokumentov a na aké komplikácie sme počas neho narazili. V záverečnej časti tejto podkapitoly priblížime špecifické pravidlá a postupy, ktoré sme dodržiavali počas transliterácie našich dokumentov, teda vytvárania tzv. Ground Truth vzorky.

Posledná, štvrtá kapitola predstavuje už konkrétne postupy a výsledky, ktoré sme rôznymi postupmi automatickej transkripcie získali. Rozdelená je do štyroch podkapitol. V rámci prvých troch podkapitol prezentujeme rozličné postupy automatickej transkripcie, ktoré sme vykonali a výsledky, ku ktorým sme sa v rámci nášho výskumu dopracovali. V prvej podkapitole sa venujeme vytváraniu referenčných modelov založených na základe jedného rukopisu. Tiež tu skúmame minimálne požiadavky na vytvorenie funkčného modelu. V druhej podkapitole sa venujeme vytváraniu modelu, skladajúceho sa z viacerých rukopisov. Sledujeme akým spôsobom ovplyvňuje pridávanie viacerých rukopisov presnosť modelu a akým spôsobom sa navzájom jednotlivé rukopisy v rámci modelu ovplyvňujú. V tretej podkapitole skúmame možnosti využitia tzv. base modelu pri vytváraní nových malých modelov. Sledujeme aký vplyv má aplikovanie base modelu na nový model, založený na základe jedného, ale aj viacerých rukopisov. V poslednej štvrtej podkapitole porovnávame efektivitu a potenciál jednotlivých postupov, ktoré sme v rámci tejto kapitoly preskúmali. Pokúsime sa zhodnotiť, aký postup je najideálnejší pri automatickej transkripcii dokumentov, skladajúcich sa z viacerých rukopisov. Zhodnotíme tiež chybovosť jednotlivých postupov a ich využitie.

1. Kanonická vizitácia ako historický prameň

Kanonické vizitácie patria medzi najdôležitejšie a obsahom najbohatšie dokumenty cirkevnej proveniencie, charakterizujúce pomery v jednotlivých farnostiach. Z. Lopatková upozorňuje, že pojem kanonická vizitácia môžeme chápať dvoma rôznymi spôsobmi. Prvým z nich sú pravidelne opakujúce sa prehliadky rôznych cirkevných ustanovizní, organizované cirkevnou vrchnosťou. V tomto prípade sa teda jedná o kanonickú vizitáciu chápanú ako istú udalosť. Druhým významom, ktorý sa dá chápať pod pomenovaním kanonická vizitácia je už samotný výstup, ktorý vzniká ako výsledok spomínanej prehliadky. Jedná sa o vizitačnú zápisnicu, ktorá obsahuje zistený stav na základe vizitácie.[11] Z. Lopatková vo svojej práci *Kanonická vizitácia – k jej interpretácii, limitom a metodologickým otázkam v historickom výskume* tiež uviedla istú učebnicovú definíciu, ktorá dobre vystihuje podstatu kanonickej vizitácie. Podľa nej sa jedná o pravidelné prehliadky cirkevných ustanovizní diecézy, ktorých cieľom je zistenie náboženského, hospodárskeho a sociálneho stavu danej ustanovizne. Jedná sa teda o inventár zariadenia kostola, majetkové a finančné pomery, príjmy a výdavky, ale aj súkromný život zamestnancov a plnenie ich záväzkov.[12]

Samotný pojem kanonická vizitácie môže obsahovať ešte aj niekoľko ďalších významov. Môže sa ním označovať napríklad celá séria vizitácií, ktorá sa následne pomenúva buď po organizátorovi daných vizitácií, samotnom vizitátorovi, ktorý ich prevádzal, alebo po územnom celku.[13]

Ďalšia historička venujúca sa problematike kanonických vizitácií M. Bizoňová definovala kanonické vizitácie ako pravidelné kontroly kapitúl, fár, kláštorov a iných cirkevných ustanovizní prevádzaných cirkevnou vrchnosťou. Ich účelom bolo zistiť skutočný stav cirkevných ustanovizní z hľadiska hospodárskeho, právneho a náboženského. Vyhotovovali sa inventáre, zisťoval sa fyzický stav budov, príjmy a výdavky a tiež život a činnosť zamestnancov inštitúcie. V neposlednom rade sa sledovala aj vzdelanostná a mravná úroveň obyvateľstva. Takto zadefinovala Bizoňová kanonickú vizitáciu vo svojej

[11] LOPATKOVÁ, Z.: Cirkevné dejiny raného novoveku z pohľadu kanonických vizitácií. In: *Pontes ad fontes : církevní dějiny ve světle pomocných věd historických a příbuzných oborů*. Hradec Králové : Filozofická fakulta Univerzity Hradec Králové, 2011, s. 176.
[12] LOPATKOVÁ, Kanonická vizitácia ..., s. 70.
[13] BOŠANSKÝ, M.: *Kánonické vizitácie Nitrianskej diecézy v rokoch 1674 – 1831*. Dizertačná práca. Katedra cirkevných dejín, RKCMBF UK, Bratislava, 2017. s. 18.

práci *Kanonické vizitácie na Spiši v 17. a prvej polovici 18. storočia ako historický prameň*.[14]

Historička F. Čechová definovala kanonické vizitácie ako občasné návštevy vykonávané cirkevnou vrchnosťou za účelom vytvoriť si spoľahlivý úsudok o stave duchovných a hmotných záležitostí obcí, nápravy rôznych vád a kontroly cirkevnej disciplíny.[15]

Na záver ešte uvedieme pomerne často citovanú definíciu V. Bucka, podľa ktorej je kanonická vizitácia osobná prehliadka cirkevnej disciplíny prostredníctvom príslušného cirkevného predstaveného, alebo jeho splnomocnenca. Cieľom je zistenie stavu ustanovizne a oblastí, ktoré treba napraviť. Jedná sa teda o prehliadku cirkevnej obce, kostola a ostatného cirkevného majetku a preskúmanie života duchovných a veriacich po stránke nábožensko-mravnej.[16]

1.1 Vznik a priebeh kanonických vizitácií

Po definovaní samotného pojmu kanonickej vizitácie je vhodné venovať sa aj tomu, ako a prečo vlastne kanonické vizitácie vznikali a tiež čo bolo ich cieľom. Jednotlivé kanonické vizitácie samozrejme vznikali rôznymi spôsobmi a z rôznych dôvodov. V tejto časti sa teda pokúsime zadefinovať skôr akýsi všeobecný postup toho, ako vizitácie vznikali a čo bolo ich hlavným cieľom. Jedná sa o prvky, ktoré sa dajú nájsť svojim spôsobom pri vzniku väčšiny kanonických vizitácií, aj keď sa detaily a špecifiká rôznych kanonických vizitácií môžu líšiť.

V Uhorsku sa inštitút kanonických vizitácií začal v rímskokatolíckych ako i v protestantských zboroch zavádzať zhruba v polovici 16. storočia.[17] Kanonické vizitácie bývali obvykle nariadené najvyšším predstaviteľom cirkvi v diecéze alebo arcidiecéze. Nariadiť ju mohol samostatne, alebo za pomoci panovníka, ktorý nariadil vizitáciu mandátom.[18] Osôb, ktoré vykonávali vizitáciu mohlo byť niekoľko. V prvom rade to býval biskup, no bežná bola aj alternatíva, kedy ho zastupoval poverený kanonik, alebo vikár.

[14] BIZOŇOVÁ, Kanonické vizitácie na Spiši v 17. a prvej polovici 18. storočia..., s. 459.
[15] ČECHOVÁ, F.: Význam a záväznosť kanonických vizitácií v dejinách nitrianskej diecézy v 20. storočí. In: *Kanonické vizitácie po tridentskom koncile*. Prešov, 2007, s. 167.
[16] LOPATKOVÁ, Kanonické vizitácie v slovenských dejinách ..., s. 97.
[17] BORZA, P.: Kanonické vizitácie mukačevského biskupa Michala Manuela Olšavského. In: *Kanonické vizitácie po tridentskom koncile.* Prešov, 2007, s. 126.
[18] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 176.

Prednosť pred biskupom mali ešte kardináli, pápežskí legáti a jeho vlastný arcibiskup.[19] Obvykle bol vizitátor doplnený o sprievod, ktorý mával rôzne veľkosti. Mohli v ňom byť cirkevní hodnostári, miestny farár, prísažný, richtár, notár, či sudca.[20]

Kanonické vizitácie podávali cirkvi dobrý obraz o tom, ako prebieha život v diecéze po stránke duchovnej aj hospodárskej. Kontroloval sa aktuálny stav farnosti mesta, či dediny. Nejednalo sa teda len o príležitostné návštevy, ale išlo o výslovne kontrolné návštevy farností.[21] Kanonické vizitácie slúžili v prvom rade pre vnútornú potrebu cirkvi. Vďaka kontrole života kňazov a farníkov sa starali o istú nápravu cirkvi zvnútra, ktorá bola obzvlášť významná v období protireformácie. Ich prvotným cieľom bolo zachytiť skutočný stav, či už miest, osôb, alebo vecí v čase konania vizitácie.[22] Druhým cieľom bolo odhalenie nedostatkov a ich náprava. Opakovanie týchto vizitácií umožňovalo sledovať aj to, ako sa postupuje na odstraňovaní odhalených nedostatkov a prípadne trestať členov spoločenstva, ktorí porušovali stanovené predpisy.[23] Jednalo sa teda aj o isté sondy, ktoré sledovali to, ako sa farnosti vysporiadali so zistenými problémami.

Ešte predtým ako samotná kanonická vizitácia začala, predchádzalo jej zverejnenie a rozposlanie istých inštrukcií do všetkých vizitovaných farností. Tieto inštrukcie sa nazývali „puncta interrogatoria“ a začali sa bežne objavovať od prvej polovice 18. storočia. Rozosielané boli najčastejšie dištriktuálnym dekanom a prostredníctvom nich všetkým farárom.[24] Obsahovali zoznam tém a otázok, na ktoré si farári mohli dopredu pripraviť odpovede, čím výrazne urýchlili priebeh vizitácie. Samotní vizitátori si ich už počas vizitácie len overovali. Aj keď sú hlavné body v rámci otázok väčšinou jednotné, autori brali ohľad aj na odlišnosti jednotlivých farností, čo dokazuje aj mierna zmena „punct“ v niektorých prípadoch.[25] B. Fabová, ktorá sa venuje téme puncta interrogatoria hlbšie vo svojej práci *Kanonické vizitácie v Abovsko-turnianskej župe v 18. storočí,* poukazuje okrem ich primárneho významu, ktorým je ich pôsobenie ako podklad pre vznik vizitačných protokolov, aj na sekundárnu úlohu, ktorá nebola pôvodným zámerom jej autorov. Týmto sekundárnym významom je uľahčenie orientácie a práce s vizitačnými protokolmi a tiež

[19] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 175.
[20] OSTERTAGOVÁ, A.: *Historicko-poznávacia hodnota kánonických vizitácií.* Dizertačná práca. Katedra slovenských dejín, FF UK, Bratislava, 2008. s.22.
[21] JURKO, J.: Kanonická vizitácia košického biskupa Andreja Szaboa v Bardejove v roku 1813. In: *Kanonické vizitácie po tridentskom koncile.* Prešov, 2007, s. 219.
[22] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 5.
[23] FIALOVÁ, I.: Kanonické vizitácie ako zdroj poznania života farností Šaštínskeho arcidiakonátu v prvej polovici 17. storočia. In: *Kanonické vizitácie po tridentskom koncile*. Prešov, 2007, s. 85.
[24] FABOVÁ, B.: Kanonické vizitácie v Abovsko-turnianskej župe v 18. storočí. In: *Človek a spoločnosť*, roč. 14, 2011, č. 2, s. 21.
[25] FABOVÁ, Kanonické vizitácie v Abovsko-turnianskej župe ..., s. 22.

poskytnutie prehľadu o údajoch obsiahnutých v zápisniciach. Puncta Interrogatoria bývajú zachované vo viacerých formách. Buď predstavujú začiatok kníh vizitačných protokolov, alebo tvoria súčasť textu. V tomto prípade môže byť puncta na ľavej strane a k nej patriace odpovede na pravej strane. Druhou alternatívou je členenie textu nadpismi zodpovedajúcimi hlavám punct. V tomto prípade je ich rekonštrukcia ťažšia a vyžaduje si porovnávanie väčšieho počtu protokolov.[26]

Po rozoslaní otázok „puncta interrogatoria" sa jednotlivé fary začali pripravovať na nadchádzajúcu vizitáciu. Pripraviť si bolo potrebné nielen odpovede na dané otázky, ale aj samotnú faru na návštevu vizitátora, ktorý cestoval väčšinou so sprievodom rôznych veľkostí. Jednou z kľúčových problematík v tomto smere bola otázka samotnej veľkosti sprievodu, ktorý sprevádzal vizitátora počas vizitácie. Stávalo sa totiž, že vizitátori prišli do menších fár s príliš veľkým sprievodom a zapríčinili tak vizitovaným inštitúciám príliš veľké výdavky, s ktorými sa malé inštitúcie len ťažko vyrovnávali. Vizitácia sa obvykle vykonávala za pomoci aspoň dvoch duchovných z katedrálnej, alebo kolegiátnej kapituly.[27]

Aj samotné trvanie vizitácie sa v rôznych prípadoch líšilo. Základným pravidlom ale bývalo, že vizitátor sa mal zdržať na fare čo najkratšie, aby nespôsoboval svojim dlhším pobytom väčšie výdavky a nespomaľoval prácu. Samotný vizitátor bol tiež zaneprázdnený čo znamenalo, že v mnohých prípadoch sa na fare nezdržal ani jeden celý deň. Dá sa konštatovať, že priemerná vizitácia mohla trvať približne deň.

Počas vizitácie mohol vizitátor verejne kázať, spovedať a trestať previnenia. Kanonickej vizitácii podliehali osoby, katolícke inštitúcie, posvätné miesta a veci, ktoré sa nachádzali v obvode diecézy.[28] Preskúmať sa teda malo všetko, čo sa týkalo viery, mravov, vykonávania cirkevných obradov, zachovávania sviatkov a cirkevnej disciplíny.[29] Na samotnej vizitácii sa okrem vizitačného sprievodu mohol zúčastňovať aj pomerne veľký počet ľudí. Bežne sa tam nachádzali obyvatelia obce, richtár, učiteľ, pastor a poprípade aj zemepán. Miesto konania vizitácie sa tiež líšilo, obvykle prebiehali na fare, prípadne v kostole, no ak to zemepán umožnil mohla vizitácia prebiehať aj u neho.[30] Väčší počet zúčastnených na vizitácii bol prospešný aj v tom smere, že umožnil vizitátorovi vypočuť si všetky strany a dosiahnuť tak ucelený pohľad na život vo farnosti. Po vykonaní vizitácie

[26] FABOVÁ, Kanonické vizitácie v Abovsko-turnianskej župe ..., s. 22.
[27] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 175.
[28] LOPATKOVÁ, Kanonické vizitácie v slovenských dejinách ..., s. 97.
[29] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 175.
[30] SZEGHY, A.: Reálie južného Zemplína vo vizitáciách seniora evanjelickej reformovanej cirkvi Istvána Csulyáka Miskolciho. In: *Kanonické vizitácie po tridentskom koncile*. Prešov, 2007, s. 73.

vznikla písomná správa, nazývaná zápisnica kanonickej vizitácie, ktorá obsahovala správu o výsledkoch návštevy.[31] Jej obsah býval veľmi bohatý a poskytoval kompletný pohľad na život vo farnosti. *Vizitačná zápisnica* je mnohokrát označovaná aj samotným slovným spojením *kanonická vizitácia*.

1.2 Štruktúra kanonických vizitácií

Štruktúra rôznych kanonických vizitácií sa líši na základe špecifickostí vizitovanej fary. Je ale možné vytvoriť určitú všeobecnú štruktúru, resp. prototyp štruktúry kanonickej vizitácie skladajúci sa z tých najdôležitejších bodov, ktorým sa vo väčšej, alebo menšej miere venovali všetky kanonické vizitácie. Sféry, ktoré musel vizitátor dokumentovať a prešetrovať by sa dali rozdeliť do troch hlavných tematických okruhov, ktoré sa dajú sledovať vo väčšine kanonických vizitácií, a ktoré v istej miere pretrvávajú v centre pozornosti počas celého oficiálneho konania kanonických vizitácií.

Prvým sledovaným okruhom bol samotný kostol. Vizitátor si obvykle všímal stav zachovania rôznych častí, materiál, patrocínium a hodnotil celkový stav v akom sa kostol nachádzal.[32] Pri prehliadke kostola sa vizitátor zaujímal aj o stav kostolných pokladníc a rôznych skutočností s nimi spojenými.[33] V neposlednom rade sa sledovalo aj vybavenie kostola. To znamenalo zachytenie informácií o oltári, relikviách, sv. olejoch, krstnej vode, ale aj informácií týkajúcich sa vedenia matriky alebo správneho slávenia sviatkov.

Druhým všeobecným okruhom, ktorý vizitátor sledoval bol samotný farár a jeho osoba. Zameriaval sa na rôzne aspekty jeho pôsobenia a správania po stránke morálnej aj hospodárskej. Ako hlavný činiteľ v danej farnosti mal farár dôležitú úlohu. Staral sa o duchovný život a blaho obyvateľov.[34] Informácie o povahe a vlastnostiach farára preto prichádzali často práve od farníkov. Opísané zvyklo byť v skratke jeho predchádzajúce pôsobenie a vzdelanie. Následne sa pozornosť obracia na jeho vlastnosti a povahu. Tu sa môžu vyskytnúť občasné sťažnosti najmä v prípade alkoholizmu, alebo zanedbávania jeho povinností. Objaviť sa mohli aj sťažnosti na nesprávne vedenie bohoslužieb, alebo ich dĺžku.[35] Okrem morálnej stránky sa sledovali aj ďalšie skutočnosti okolo osoby farára. Objavujú sa informácie napríklad o tom či má breviár, či vlastní nariadené knihy, dodržiava

[31] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 175.
[32] SZEGHY, Reálie južného Zemplína ..., s. 75.
[33] FIALOVÁ, Kanonické vizitácie ako zdroj poznania ..., s. 87.
[34] FIALOVÁ, Kanonické vizitácie ako zdroj poznania ..., s. 88.
[35] SZEGHY, Reálie južného Zemplína ..., s. 76.

predpísaný rítus sviatostí a pod.[36] V neposlednom rade sa venovala pozornosť aj výške jeho príjmov. Zaznamenávali sa dávky, ktoré kňaz obdržal v naturáliách a tiež v rámci peňažného platu. Mohlo sa zaznamenať aj to, akým spôsobom sú tieto dávky zbierané, napríklad či sa skladala celá obec, alebo bola dávka stanovená pre jednotlivé kategórie obyvateľstva osobitne.[37] Osobitnú kategóriu mohli tvoriť aj poplatky za cirkevné úkony, akými boli napríklad sobáš, krst, pohreb a pod. Pozornosť sa venovala nielen farárovi, ale aj ostatným členom fary, obvykle to bol učiteľ a organista. Opäť sa zvyklo opísať ich predchádzajúce pôsobenie a vzdelanie. Pri učiteľoch sa zaznamenával rozsah ich povinností a tiež príjmy. Informácie o tom, čo a kedy učia mohli byť tiež spomenuté. V prípade existencie školy sa väčšinou opísala jej budova a počty žiakov.[38]

Poslednú tematickú skupinu tvorili informácie o farníkoch, resp. obyvateľoch danej obce. Informácie sa vizitátor dozvedá najčastejšie od farára alebo učiteľa. Jadrom tejto časti bývajú informácie o plnení si náboženských povinností obyvateľov a tiež informácie o vierovyznaní.[39] Zapisovalo sa či farníci navštevujú kostol a ako pristupujú k rôznym slávnostiam, poprípade spovediam. Vizitátor mohol zaznamenať aj starosť o platby a odovzdávanie desiatkov.[40] Mohli sa tu objaviť aj rôzne sťažnosti predovšetkým na ľahostajnosť alebo ignorovanie bohoslužieb. Poniektoré vizitačné zápisnice mohli v rámci tejto časti obsahovať aj informácie o počte a etnicite obyvateľov farnosti.[41]

Tieto tri hlavné okruhy sa dajú považovať za istý prototyp štruktúry kanonickej vizitácie. Ako už bolo spomenuté, kanonické vizitácie sa od seba líšili na základe rôznych detailov a špecifickostí, ktoré boli vlastné vizitovaným farám. Prakticky pri všetkých kanonických vizitáciách je ale možné v istej podobe sledovať tieto tri opísané okruhy záujmu, ktorými boli kostol, farár a farníci. Keď sa pozrieme na koniec vizitačných protokolov môžeme tam nájsť aj istý záver, ktorý obsahoval príkazy a nariadenia, vyhotovené vizitátorom pre farára. Tieto nariadenia sa označujú pojmom *commissa*.[42]

Vývoju štruktúry kanonických vizitácií v priebehu storočí ako aj limitom a metodologickým otázkam pri výskume kanonických vizitácií sa venovala Z. Lopatková v už spomínanej práci *Kanonická vizitácia – k jej interpretácii...* na príklade kanonických

[36] KURUC, Š.: Vizitácia spišského prepošta Jána Peltza z roku 1731. In: *Kanonické vizitácie po Tridentskom koncile*. Prešov, 2007, s. 99.
[37] FIALOVÁ, Kanonické vizitácie ako zdroj poznania ..., s. 88 – 89.
[38] LOPATKOVÁ, Kanonická vizitácia ..., s. 71.
[39] FIALOVÁ, Kanonické vizitácie ako zdroj poznania ..., s. 90.
[40] SZEGHY, Reálie južného Zemplína ..., s. 78.
[41] FIALOVÁ, Kanonické vizitácie ako zdroj poznania ..., s. 90.
[42] FABOVÁ, Kanonické vizitácie v Abovsko-turnianskej župe ..., s. 21.

vizitácií Smolenického dekanátu, ktoré sú dobre zachované už od druhej polovice 16. storočia až po začiatok 20. storočia. Môžeme ich teda využiť na predstavenie všeobecného vývoja štruktúry kanonických vizitácií.

Štruktúra týchto kanonických vizitácií bola v 16. storočí značne stručná. Ich hlavným cieľom bolo informovať o postupe reformačných síl v Uhorsku a informácie o živote na fare teda neboli v centre záujmov.[43] Štruktúra sa výrazne mení a začína formovať svoju stálejšiu podobu až v 17. storočí. Zmeny sa týkajú nielen obsahovej, ale aj formálnej stránky. Práve tu sa už dajú rozlíšiť hlavné tematické kategórie, ktorým sme sa venovali v predchádzajúcej časti.[44] Objavujú sa už aj nové údaje týkajúce sa hospodárskeho a majetkového stavu.

Vizitácie z 18. storočia považuje Lopatková za najbohatšie a najúplnejšie po stránke obsahovej. Zaujímavé je, že obsahujú väčší počet informácií nielen keď ich porovnávame s vizitáciami zo skorších storočí, ale aj keď ich porovnáme s vizitáciami z neskorších storočí, nakoľko tie už „skĺzli do mechanicky reprodukujúceho sa obsahu, skladajúceho sa zo strojopisných údajov o farnostiach“.[45] V rámci 18. storočia majú svoje významné miesto najmä vizitácie z 2. polovice 18. storočia, presnejšie sa jedná o roky 1780 až 1788. Označujú sa ako baťániovské kanonické vizitácie,[46] ktoré prinášali nové jedinečné informácie. Lopatková dokumentuje aj dôležitú zmenu po stránke formálnej, pričom spomínané baťániovské vizitácie boli vyhotovené takým spôsobom, že každá farnosť mala vlastný samostatný vizitačný protokol vo veľkom fóliovom formáte, zviazaný v koženej väzbe, oproti vizitáciám z predchádzajúcich storočí, ktoré bývali zviazané v jednej knihe, ktorá zahŕňala územie celého arcidiakonátu, alebo dekanátu. Tiež poznamenáva, že vizitačný protokol sa už skladal z niekoľkých častí. Tieto časti predstavovala samotná kanonická vizitácia, porovnanie označované ako Combinatorium, tabuľka označená ako Tabella, nariadenia vizitátora a napokon prepísané dokumenty, týkajúce sa údajov uvedených vo vizitácií.[47]

1.3 Využitie kanonických vizitácií ako historického prameňa

[43] LOPATKOVÁ, Kanonická vizitácia ..., s. 71.
[44] LOPATKOVÁ, Kanonická vizitácia ..., s. 71.
[45] LOPATKOVÁ, Kanonická vizitácia ..., s. 72.
[46] Jozef Baťán bol uhorským rímskokatolíckym kňazom, biskupom a kardinálom. Pôsobil ako ostrihomský arcibiskup a uhorský prímas v rokoch 1776 – 1799.
[47] LOPATKOVÁ, Kanonická vizitácia ..., s. 72.

Na základe rozboru všeobecnej štruktúry kanonických vizitácií vidíme skutočne veľké množstvo informácií, ktoré tento historický prameň ponúka. Kanonické vizitácie predstavujú výnimočný a v mnohých smeroch aj špecifický prameň, ktorého využitie je podľa slov Z. Lopatkovej prakticky nevyčerpateľné. Ich jedinečná hodnota spočíva aj v tom, že predstavujú jediný historický prameň, ktorý po obsahovej stránke prináša také veľké množstvo informácií o živote farností.[48] V mnohých smeroch teda predstavujú pre niektoré výskumné smery základný a najbohatší historický prameň. Mnohokrát sú jediným dobovým písomným materiálom viažucim sa k dejinám jednotlivých farností. Využitie si tak kanonické vizitácie nachádzajú čoraz viac nielen v striktne historických štúdiách, ale aj v rôznych výskumoch, ktoré stoja na pomedzí multidisciplinárneho výskumu medzi históriou a príbuznými vedami.[49] Predstavujú významný prameň aj pre množstvo smerov, ktoré sú značne špecificky zamerané a kanonické vizitácie tak pre ne môžu predstavovať jediný historický prameň s dostatočnou informačnou hodnotou. Rôznym vedným odborom, ktoré môžu čerpať informácie z kanonických vizitácií sa venovala v detaile predovšetkým A. Ostertagová vo svojej štúdii *Kanonické vizitácie ako historický prameň*. V rámci tejto podkapitoly sa pokúsime predstaviť aspoň niektoré smery výskumu, pre ktoré majú kanonické vizitácie ako prameň kľúčovú úlohu.

Bádatelia pracujúci s kanonickými vizitáciami nachádzajú ich hlavné využitie najmä pri výskume cirkevných dejín. Predstavujú spoľahlivú pramennú základňu najmä pre výskum reformácie a následnej rekatolizácie, pričom ponúkajú informácie o ich priebehu na lokálnej úrovni.[50] Zaujímavý výskum v tomto smere vykonala M. Bizoňová, ktorá použila kanonické vizitácie na výskum šírenia reformácie a zmenu konfesionálnej štruktúry obyvateľstva na území Spiša v 17. a 18. storočí.[51] V rámci cirkevných dejín môžu kanonické vizitácie predstavovať aj unikátnu možnosť nahliadnuť do života jednotlivých zborov.

Podobne dôležité sú kanonické vizitácie aj pre regionálne dejiny, pretože môžu obsahovať unikátne informácie o náboženských aj politických pomeroch z danej oblasti. Obzvlášť významné miesto zastávajú práve pri menších obciach, pre ktoré neexistuje väčší počet zachovaných prameňov, v ktorých by sa nachádzali také detailné údaje o danej lokalite v danom časovom období. Podstatné je aj to, že kanonické vizitácie boli pravidelne opakované, čo ponúka ďalší rozmer výskumu skladajúci sa zo sledovania vývoja danej

[48] LOPATKOVÁ, Kanonická vizitácia ..., s. 76.
[49] HOFERKA, M.: Kánonická vizitácia farností Záhorského dekanátu Bratislavského prepošstva z roku 1673. In: *Studia Historica Tyrnaviensia XVIII*. Trnava, 2015, s. 101.
[50] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 9 – 10.
[51] BIZOŇOVÁ, Kanonické vizitácie na Spiši v 17. a prvej polovici 18. storočia ..., s. 459 – 469.

oblasti.[52] Kanonické vizitácie môžu ponúknuť pre výskum dejín vybraného územia aj cenné informácie o cirkevnej organizácii na danom území, čím môžu napomôcť k jej rekonštrukcii.[53]

Informácie z kanonických vizitácií čerpajú aj bádatelia skúmajúci v oblasti dejín umenia. Vizitácie opisujú ako vonkajšiu, tak aj vnútornú podobu kostola a tiež jeho aktuálny stav. Ponúkajú cenné informácie aj o aktuálnom stave sakristie, fary, kaplniek, zvoníc a informujú o rôznych pamiatkach v exteriéri ako sú kríže, kalvárie, božie muky a podobne. Tiež umožňujú výskum architektonického vývoja týchto objektov, nakoľko obsahujú informácie aj o rekonštrukciách, deštrukciách a použitých materiáloch.[54] Rovnako významné a podstatné poznatky pre tento odbor prináša aj pohľad na výbavu kostolov. Opisy oltárov, krížov, sôch alebo obrazov predstavujú dôležitú pramennú základňu pre výskum v rámci dejín výtvarného umenia. Detailnejšie vizitácie môžu obsahovať aj cenné informácie týkajúce sa autorov, hodnoty alebo dátumu vzniku rôznych diel.[55]

Kanonické vizitácie ponúkajú tiež mimoriadne významné informácie pre výskum dejín školstva. Nachádzajú sa v nich informácie o tom, či má daná obec školu a tiež informácie o jej stave a prípadne aj vybavení. V rámci detailnejších vizitačných protokolov môžeme nájsť aj informácie o spôsobe vyučovania ako sú napríklad vyučovacie predmety, počty žiakov, učebné pomôcky alebo aj jazyk výučby.[56] Významné sú najmä kanonické vizitácie z obdobia, kedy boli ľudové školy pripojené k farám. Učiteľ a žiaci sa tým pádom aktívne zúčastňovali aj na cirkevnom živote. Okrem toho mali učitelia väčšinou aj určité povinnosti voči cirkvi, pričom boli častokrát zároveň aj organistami alebo zvonármi.[57]

Jedinečné poznatky prinášajú kanonické vizitácie aj pre výskum hospodárskych dejín. Ponúkajú informácie o výške a spôsobe vyberania desiatku, zoznamoch dlžníkov, veriteľov a súm. Okrem toho prinášajú informácie o majetku fary, kostola, farára a rôznych príjmoch z rozličných zdrojov.[58]

Spomedzi špecifickejších výskumných smerov predstavujú kanonické vizitácie kľúčový prameň pre výskum zvonolejárstva a dejín zlatníctva. A. Ostertagová spomína, že pre tieto smery neexistuje lepší prameň ako kanonické vizitácie. Môžu totiž obsahovať

[52] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 9.
[53] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 184.
[54] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 7.
[55] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 8.
[56] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 9.
[57] NEMEŠOVÁ, A.: Vývoj ľudového školstva na Spiši vo svetle kanonických vizitácií. In: *Kanonické vizitácie po Tridentskom koncile*. Prešov, 2007, s. 105.
[58] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 9.

informácie o tom, kedy a kde boli zvony uliate a poprípade aj to kto ich ulial. Z kanonických vizitácií sa niekedy dajú zistiť aj osudy zvonov, to sa ale týka už neskorších kanonických vizitácií od 19. storočia.[59]

Podobne špecifickým smerom výskumu je muzikológia a dejiny hudby. Pre tieto smery výskumu ponúkajú kanonické vizitácie cenné informácie o hudobných nástrojoch, no najmä organoch. Jedná sa o informácie ako napríklad kde a kým bol organ zhotovený, o počte registrov alebo o ladení píšťal. Okrem toho môžu ponúkať informácie o využívanej hudobnej literatúre a spevníkoch.[60]

Ďalším zo špecifických smerov výskumu, pre ktoré sú kanonické vizitácie ako prameň dôležité je výskum knihovedy. Vizitačné zápisnice môžu v niektorých prípadoch obsahovať informácie o knihách nachádzajúcich sa na farách. Obvykle sú ale dosť stručné, pričom obsahujú len názov diela a poprípade meno autora.[61]

Kanonické vizitácie sú využívané mnohými výskumnými smermi aj ako doplňujúci prameň. Je tak napríklad v prípade jazykovedy, ktorá môže čerpať z kanonických vizitácií rôzne mená a priezviská, ale tiež môže využívať aj obsiahnuté topografické názvy potokov, chotárov, lesov alebo lúk.[62] V prípade, že zápisnice evidujú aj počty obyvateľov, poprípade ich zloženie či už po stránke náboženskej alebo národnostnej, môžu slúžiť aj pre výskum demografie a štatistiky. V neposlednom rade je potrebné spomenúť aj dejiny zdravotníctva, ktoré môžu z kanonických vizitácií získavať cenné informácie pri výskume rôznych morových alebo cholerových epidémií.[63]

Ako sme poukázali v tejto podkapitole, kanonické vizitácie predstavujú kľúčový prameň pre veľké množstvo výskumných smerov a ich historická hodnota je skutočne vysoká. Práca s kanonickými vizitáciami pomáha lepšie pochopiť myslenie v danej dobe, pričom ponúka pohľad do takmer každej sféry ľudského života. Predstavujú vzácny doklad života ľudí so všetkými jeho aspektmi v konkrétnej historickej dobe a lokalite.[64]

Pri práci s kanonickými vizitáciami je ale potrebné brať ohľad aj na to, aká je skutočná relevantnosť údajov v nich obsiahnutá. Kanonické vizitácie boli vyhotovované príslušníkmi cirkvi, ktorí mohli mať tendenciu niektoré údaje skresľovať.[65] Stávať sa tak mohlo najmä v období reformácie, kedy nemusia byť všetky údaje v kanonických

[59] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 8.
[60] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 8.
[61] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 9.
[62] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 10.
[63] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 10.
[64] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 187.
[65] FABOVÁ, Kanonické vizitácie v Abovsko-turnianskej župe ..., s. 21.

vizitáciách úplné.[66] Na druhú stranu prechádzali informácie obsiahnuté v kanonických vizitáciách dvojitou kontrolou a v prvom rade slúžili pre vnútornú potrebu cirkvi, čo znamená, že nejaké rozsiahle manipulovanie a skresľovanie údajov v kanonických vizitáciach nemalo prakticky význam. Kanonické vizitácie sú všeobecne brané ako dôveryhodný prameň najmä preto, že predstavovali úradný dokument, pričom ich vierohodnosť zvyšoval aj samotný proces ich vzniku, ktorý bol ukotvený v platných právnych normách. Aj napriek tomu, že sa kanonické vizitácie dajú považovať za hodnoverný historický prameň, je potrebné k nim pristupovať kriticky a posudzovať ich v kontexte s dobou, v ktorej vznikali.[67]

Spomenieme teda niekoľko metód, ktoré sa bežne využívajú pri práci s kanonickými vizitáciami ako historickým prameňom. Prvou je priama metóda a využitie kritiky. Ako už bolo povedané, kanonická vizitácia je považovaná za hodnoverný prameň. No aj napriek tomu sa môže stať, že sa v nej nachádzajú chyby, ktoré mohli vznikať pri zapisovaní. Podobne sa mohli chyby objaviť aj pri odpovediach na otázky, pričom mohli byť mylné domnienky uvedené ako fakty.[68] Okrem tohto je bežné aj využívanie induktívnej a deduktívnej metódy no aj rôznych medziodborových metód, ktoré hraničia s historiografiou. Taká je napríklad biografická metóda, využívaná najmä pri životopisoch farárov, uvedených vo vizitácii. Treba ale poznamenať, že detailnejšie životopisné informácie o farároch sú skôr ojedinelé. Osobný profil farárov sa uvádzal najmä v období reformácie, kedy budúcnosť farnosti z hľadiska náboženstva záležala práve na ich rozhodnutiach.[69] V súvislosti s toponomastikou sa využíva filologická metóda. Viaže sa na rôzne miestne a chotárne názvy, ktoré ponúkajú svedectvo o charaktere krajiny.[70] Pri identifikácii a lokalizácii miestnych názvov sa používa geografická metóda. Geografické okolnosti a prostredie malo výrazný vplyv na fungovanie a chod farností. Pri identifikácii názvov je ale potrebné využívať väčší počet literatúry zaoberajúci sa touto problematikou, nakoľko pri tvorbe vizitačného protokolu mohlo dôjsť k skomoleniu názvov.[71] Tiež je pri využití geografickej metódy prakticky nevyhnutné pracovať s mapami.

Pri priamej práci s kanonickými vizitáciami existuje aj niekoľko metód, ktoré je potrebné využiť pri takmer akomkoľvek výskume, ktorý čerpá a pracuje s kanonickými

[66] LOPATKOVÁ, Kanonická vizitácia ..., s. 76.
[67] LOPATKOVÁ, Kanonická vizitácia ..., s. 76.
[68] LOPATKOVÁ, Kanonická vizitácia ..., s. 73.
[69] LOPATKOVÁ, Kanonická vizitácia ..., s. 74.
[70] LOPATKOVÁ, Kanonická vizitácia ..., s. 75.
[71] LOPATKOVÁ, Kanonická vizitácia ..., s. 75.

vizitáciami ako historickým prameňom. Prvou z nich je využitie latinčiny. Nakoľko je jazykom kanonických vizitácií latinčina je prakticky nevyhnutné spraviť presný a precízny preklad textu vizitácie, aby s ňou bolo možné náležite pracovať.[72] Dôvodom je najmä to, že slovník kanonických vizitácií je značne špecifický. Latinské termíny v kanonických vizitáciách majú často odlišný význam od termínov v klasickej latinčine a preto je potrebné využívať najmä slovník stredovekej latinčiny.

Druhá metóda, ktorá je nesmierne dôležitá pri práci k kanonickými vizitáciami je paleografický prepis textu. Paleografický prepis si nájde svoje využitie pri práci s prakticky akýmikoľvek dobovými rukopisnými textami. Pri niektorých môže slúžiť na zjednodušenie práce, pri iných môže byť priam nevyhnutný k samotnému výskumu. Každopádne predstavuje významný krok k zefektívneniu práce s dobovými rukopismi.

Paleografický prepis tvorí významnú súčasť priamej práce s dobovými rukopismi. Samotný prepis rukopisov, ktoré môžu mať stovky až tisíce strán ale predstavuje obzvlášť zdĺhavý proces, ktorý môže mnohokrát odradiť bádateľov od detailnejšieho venovania sa týmto dokumentom. Aj z tohto dôvodu zostáva veľké množstvo informácií zahrnutých v rôznych kanonických vizitáciách nevyužité. Je to práve problematika paleografického prepisu a transkripcie rukopisných historických prameňov, ktorá stojí v jadre tejto diplomovej práce venujúcej sa automatickému prepisu rukopisných historických textov. Automatický prepis rukopisných textov môže odstrániť jednu z najväčších prekážok pri práci s veľkým počtom rukopisov, výrazne zefektívniť výskum a sprístupniť tak nesmierne množstvo informácií.

1.4 História kanonických vizitácií

Aj keď sa kanonické vizitácie oficiálne objavujú v slovenských dejinách až zhruba od 14. storočia, samotné vizitovanie má omnoho starší pôvod. Zrejme úplne prvé doklady o konaní sa vizitácií pochádzajú z Byzantskej ríše a to počas pôsobenia carihradského patriarchu Gregora v rokoch 329 – 390.[73] Výraznejšie doklady o vizitovaní priamo z územia Európy sa objavujú už na začiatku 6. storočia, no je vysoko pravdepodobné, že sa kanonické vizitácie v Európe konali aj predtým. Informujú o tom niektoré synody, ktoré sa v tomto časovom období konali, ako napríklad Tarragonská synoda (516), synoda v Brage (572),

[72] OSTERTAGOVÁ, Kanonické vizitácie ako historický prameň ..., s. 6.
[73] KYSEĽOVÁ, M.: Kanonické vizitácie na Spiši ..., s. 113.

alebo synoda v Tolede (633). Kanonické vizitácie sa už v našom zmysle slova objavujú aj vo Franskej ríši za vlády Karola Veľkého (768 – 814). Nachádzajú sa v rôznych snemových ustanoveniach, v ktorých sa hovorí o rozšírení právomocí vizitátora aj na svetské veci. Ohľadom kanonických vizitácií nás neskôr informuje aj koncil v Trevíri (895) a Štvrtý carihradský koncil (869 – 870).[74]

Kanonické vizitácie sa na území Slovenska stávajú bežné až v období novoveku. Existujú samozrejme aj výnimky a je takmer isté, že sa kanonické vizitácie konali aj v období stredoveku, no chýbajú presnejšie informácie o tom, kedy a akou formou prebiehali.[75] Čo sa týka zachovaných kanonických vizitácií, tak z obdobia stredoveku máme zachovanú len kanonickú vizitáciu Ostrihomskej arcidiecézy z roku 1397.[76] Ďalšie vizitácie v Ostrihomskom arcibiskupstve sú až z rokov 1560 – 1564 a boli vykonané z iniciatívy Mikuláša Oláha.[77] Aj keď nie sú skoršie kanonické vizitácie zachované, existujú zmienky, ktoré ich existenciu potvrdzujú. Dobrým príkladom je územie Spiša. Počiatky kanonických vizitácií sa tu dajú datovať už do 13. storočia. Je to na základe dohody medzi spišským prepoštom a Bratstvom farárov XXIV kráľovských miest. Dohoda hovorí o ročnom vizitovaní miest spišským prepoštom.[78] Zmienky o skorších vizitáciách pochádzajú aj z niekoľkých synod, na ktorých sa riešili okolnosti spájajúce sa s vizitovaním. Jedná sa o Budínsku synodu z roku 1279, alebo Nitriansku synodu z roku 1494.[79]

Kľúčovou udalosťou, ktorá výrazne zasiahla do cirkevných dejín bola reformácia. Tá značne poznamenala aj samotný vývoj kanonických vizitácií. Bola to práve reakcia na reformáciu, ktorá posunula vývoj kanonických vizitácií ďalej. Touto reakciou bola tzv. rekatolizácia a s ňou spojený Tridenský koncil, ktorý predstavuje najvýznamnejší bod v historickom vývoji kanonických vizitácií. Kanonické vizitácie do 16. storočia bývali častokrát zanedbávané alebo zjednodušované. Toto sa výrazne mení po Tridentskom koncile. Dekréty na ňom prijaté sa stávajú smerodajné z hľadiska vykonávania kanonických vizitácií pre nasledujúce storočia. Vizitáciám diecéz sa venovalo 24. zasadnutie konané 11. 11. 1563. Tu bola zdôraznená vizitačná povinnosť biskupov, pričom k vizitovaniu diecéz malo dochádzať každý rok a vo väčších diecézach aspoň raz za dva roky.[80] Vizitácie sa mali

[74] BOŠANSKÝ, *Kánonické vizitácie Nitrianskej diecézy* ..., s. 26
[75] LOPATKOVÁ, Kanonická vizitácia ..., s. 70.
[76] LOPATKOVÁ, Kanonická vizitácia ..., s. 70.
[77] BIZOŇOVÁ, Kanonické vizitácie na Spiši v 17. a prvej polovici 18. storočia ..., s. 461.
[78] BIZOŇOVÁ, Kanonické vizitácie na Spiši v 17. a prvej polovici 18. storočia ..., s. 461.
[79] BOŠANSKÝ, *Kánonické vizitácie Nitrianskej diecézy* ..., s. 28.
[80] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 177.

svojim spôsobom stať určitou kontrolnou silou voči reformácií a protestantskej náuke, nakoľko slúžili ako dôsledný nástroj na sledovanie rozširovania sa reformačných učení.

Na našom území svojím spôsobom predbehol dobu Mikuláš Oláh, ktorý nariadil celkovú kanonickú vizitáciu Ostrihomského arcibiskupstva ešte predtým, ako vstúpili dekréty Tridentského koncilu do platnosti. Táto vizitácia prebiehala od roku 1560 do roku 1564.[81] Po smrti Mikuláša Oláha ale dochádza k opätovnému úpadku kanonických vizitácií a prestávajú sa vykonávať podľa nariadení schválených na Tridentskom koncile. S týmto problémom sa zaoberali najmä dve synody na začiatku 17. storočia, ktoré prebehli v Trnave. Ich cieľom bolo prebrať nariadenia Tridentského koncilu, správnym spôsobom ich aplikovať na našom území a prispôsobiť ich na miestne pomery. Prvá z týchto synod prebehla v roku 1611 a predsedal jej ostrihomský arcibiskup František Forgáč.[82] Vizitovať sa malo skutočne všade a prostredníctvom kňazov sa mali vizitovať aj opustené lokality. Ani po tejto synode ale nedošlo k dostatočným zmenám a kanonické vizitácie sa v celoplošnom rozmere naďalej nekonali. Aj z tohto dôvodu sa uskutočnila ďalšia synoda v Trnave a to v roku 1629. Tá prebiehala už za pôsobenia ostrihomského arcibiskupa Petra Pázmaňa, ktorý je známy svojimi rekatolizačnými snahami. Aj na tejto synode boli jednou z hlavných tém opäť kanonické vizitácie a slúžila aj ako doplnenie predpisov a nariadení, ktoré boli prijaté na synode v roku 1611. V centre kanonickej vizitácie malo byť odhalenie nedostatkov kňazov a ich odstránenie a náprava.[83] Kanonické vizitácie teda pod vplyvom Petra Pázmaňa naberali ďalší významný rozmer ako nástroj proti reformácii, ktorý mal slúžiť už aj na nápravu cirkevných pomerov. Zameral sa aj na učiteľov, pričom sa malo preskúmať aj to, do akých škôl posielajú veriaci svoje deti a to, či učitelia vyučujú deti v duchu kresťanskej viery.[84] Samotný P. Pázmaň vykonal kanonické vizitácie Ostrihomského arcibiskupstva v rokoch 1624 až 1634. Spomínané dve Trnavské synody a predpisy Tridentského koncilu definovali vývoj kanonických vizitácií na našom území a udali im smer, ktorým sa po zvyšok novoveku uberali.

Nasledujúci vývoj kanonických vizitácií bol oproti udalostiam zo 16. a 17. storočia pomerne kľudnejší. V prvej polovici 18. storočia predstavuje významnejšiu udalosť vydanie Resolutio Carolina Karolom III. v roku 1731. Jednalo sa o nariadenie, ktoré uprednostňovalo katolícku cirkev a čiastočne jej podriaďovalo aj evanjelickú cirkev. Ihneď po vydaní tohto

[81] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 177.
[82] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 178.
[83] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 178.
[84] LOPATKOVÁ, Kanonické vizitácie v slovenských dejinách ..., s. 100.

nariadenia bola vykonaná kompletná vizitácia Ostrihomského arcibiskupstva. V rámci tejto vizitácie sa stretávame už aj so spomínanými otázkami „puncta interrogatoria“.[85]

Počas druhej polovice 18. storočia zasiahol výraznejšie do cirkevných pomerov a spolu s tým aj do vývoja kanonických vizitácií Jozef II. Okrem vydania Tolerančného patentu boli práve za jeho vlády vykonané tzv. baťániovské kanonické vizitácie (1781 – 1783), ktoré sú považované za najlepšie spracované kanonické vizitácie. Pomenované sú po ostrihomskom arcibiskupovi Jozefovi Baťánim, na základe ktorého nariadení boli vykonané.[86] Podľa veľkosti fary mali 24 až 30 bodov pričom boli rozsiahle a podrobné.[87]

Aj vďaka baťániovským vizitáciám tak kanonické vizitácie v 18. storočí dosiahli po svojej obsahovej stránke najvyššiu úroveň, pričom predstavujú nesmierne významné historické pramene. Ako už bolo spomenuté, kanonické vizitácie v nasledujúcich storočiach skĺzli skôr do opakujúcich sa obsahov zložených zo strojopisných údajov, čím utrpela aj ich obsahová stránka. Na druhú stranu sa v 19. storočí dá hovoriť o istom zvýšení kritického prístupu ku kanonickým vizitáciám, čo rovnako napomáha aj k zvýšeniu dôveryhodnosti prameňov z tohto obdobia.

[85] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 179.
[86] LOPATKOVÁ, Cirkevné dejiny raného novoveku ..., s. 179, 180.
[87] LOPATKOVÁ, Kanonické vizitácie v slovenských dejinách ..., s. 102.

2 Vybrané kanonické vizitácie

2.1 Cirkevný vývoj a situácia na území vybraných farností

Kanonické vizitácie, s ktorými v práci narábame pochádzajú z roku 1783 a týkajú sa farností Motyčky, Staré Hory a Špania Dolina.[88] Farnosť Motyčky mala v roku 1783 filiálky v osadách Bukovec, Môce, Bully, Kunst, Mistríky, Mišúty, Donovaly, Hanesy, Sliačany, Jergaly, Dúvodné a Štubne. Farnosť na Starých Horách mala filiálky v osadách Polkanová, Richtárová, Dolná Turecká, Horná Turecká, Dolný Jelenec, Horný Jelenec, Valentová, Bachlačka, Rybie a Prašnica. Farnosť na Španej Doline mala v tomto roku filiálku len v osade Piesky. V čase vzniku daných osád bola ich zemepánom Banská komora v Banskej Bystrici, ktorá predstavovala pre tieto sídla rozhodujúci orgán. Aj napriek tomu, že tieto sídla ležali na historickom banskobystrickom donačnom území, nepovažovali Banskú Bystricu za svoju vrchnosť a pôdu, na ktorej stáli vlastnila komora. Dozvedáme sa to aj z priamych výpovedí predstaviteľov dotknutých sídiel v dotazníkoch vypracovaných pri príležitosti tereziánskej urbárskej regulácie, v ktorých predstavitelia týchto osád prehlásili, že pracujú pre kráľovskú cisársku komoru a neuznávajú mesto Banská Bystrica za svoju vrchnosť.[89] Na základe výkazu banskobystrickej banskej komory z roku 1788 vieme, že pod jej správu spadalo až 96% osadníkov obývajúcich osady Starohorskej oblasti.[90]

Spomenuté filiálky predstavovali pomerne menšie sídla. Ich veľkosť v období, kedy sa tu konali naše vizitácie nám dokáže priblížiť Jozefská evidencia z roku 1787. Tá eviduje, že tieto sídla mali medzi 4 až 21 domami, pričom najmenšie boli Mišúty so 4 evidovanými domami a najväčšie boli Donovaly s 21 evidovanými domami.[91] Na základe vizitácií z roku 1783 hovorilo obyvateľstvo spomenutých farností slovenskou rečou.[92] To isté nám potvrdzujú aj informácie z vizitácie konanej v roku 1804 Gabrielom Zerdahelym a vizitácie z rokov 1822 a 1829 vykonané biskupom Jozefom Belanským.[93]

---

[88] Tieto vizitácie sa nachádzajú v Archíve Rímskokatolíckeho biskupstva a diecézy v Banskej Bystrici so sídlom v Badíne (ďalej DA BB), fond kanonické vizitácie, CV č. 10.
[89] TOMEČEK, O.: *Drevorubači a uhliari v lesoch Banskej Bystrice*. 1. vyd. Banská Bystrica : Fakulta humanitných vied UMB, 2010. s. 154.
[90] TOMEČEK, *Drevorubači a uhliari v lesoch Banskej Bystrice* ..., s. 158.
[91] TOMEČEK, *Drevorubači a uhliari v lesoch Banskej Bystrice* ..., s. 88 – 90.
[92] Archív DA BB, Visitatio Canonica Ecclesiae Parochialis Moticsko , 1783, CV č. 10; Archív DA BB, Visitatio Canonica Ecclesiae Parochialis Vallis Dominorum , 1783, CV č. 10; Archív DA BB, Visitatio Canonica Ecclesiae Parochialis Vetero Montanae , 1783, CV č. 10.
[93] TOMEČEK, *Drevorubači a uhliari v lesoch Banskej Bystrice* ..., s. 116.

Vznik farnosti na Starých Horách uvádzajú cirkevné schematizmy v 15. storočí a vznik farnosti v Španej Doline už dokonca v 13. storočí, pričom Kostol v Španej Doline býva považovaný za najstarší kostol v blízkom okolí Banskej Bystrice. Informácie o existencii týchto farností v danej dobe sú ale nepodložené. Tejto problematike sa venovala Z. Mičková v práci *Banícka osada Špania Dolina od neskorého stredoveku k ranému novoveku,* kde poukazuje na protiargumenty existencie kostola v Španej Doline minimálne do prvej polovice 15. storočia. Aj v rámci 15. storočia sa skôr jednalo len o kaplnku ako o kostol.[94] Ako dôkaz neexistencie farností v Španej Doline a Starých Horách je možné uviesť mapku farností, ktorá bola vyhotovená podľa zoznamov pápežských decimátorov z rokov 1332 – 1337, na ktorej nenachádzame ani jednu z týchto farností.[95] Podobne sa ani jedna z farností nenachádza v testamente významného banskobystrického ťažiara Michala Königsberga z roku 1503. Farnosť v Španej Doline nie je uvedená ani v kanonických vizitáciách Blažeja Zathaya z roku 1561 a podobne absentuje aj vo vizitácii Šimona Bánovského z roku 1634. Dá sa predpokladať, že cirkevná organizácia na tomto území sa dotvárala až v období novoveku, pričom existenciu farnosti na Španej Doline je možné potvrdiť s určitosťou až v polovici 17. storočia.[96]

Charakteristickou črtou pre cirkevný vývoj tejto oblasti bol dlhodobý zápas medzi evanjelickou a kresťanskou cirkvou. Slobodné kráľovské mestá a banské mestá disponovali už v stredoveku patronátnym právom a nebolo pre ne problémom povolávať či prepúšťať kazateľov v závislosti od toho, aká ideová orientácia si našla stúpencov medzi členmi mestských rád.[97] Šírenie reformácie v stredoslovenských banských mestách sa dá z veľkej miery pripísať aj prítomnosti početnejšieho nemeckého etnického elementu, ktorý udržiaval s materskými oblasťami intenzívne kontakty.[98] Protestantizmus sa uchytil veľmi úspešne aj medzi baníkmi v okolí Španej Doliny, ktorí sa usilovali o samostatnejší náboženský život. Údajne sa už v roku 1526 kázalo v celej Starohorskej doline v Lutherovom duchu. V tomto období ale ešte nepredpokladáme v Španej Doline väčší sakrálny objekt, no určite sa tu už nachádzala kaplnka.[99] Napriek samostatnejšiemu cirkevnému životu, ktorý sa rozvíjal v Španej Doline existuje aj možnosť, že prvotným cirkevným centrom tejto oblasti mohli

[94] MIČKOVÁ, Z.: *Banícka osada Špania Dolina od neskorého stredoveku k ranému novoveku*. 1. vyd. Banská Bystrica : Právnická fakulta UMB, 2013. s. 90.
[95] MIČKOVÁ, *Banícka osada Špania Dolina* ..., s. 88.
[96] TOMEČEK, *Drevorubači a uhliari v lesoch Banskej Bystrice* ..., s. 160.
[97] KOWALSKÁ, E.: Ozbrojená moc v procese rekatolizácie. In: *Vojenská história*, 9, 2005, č. 2, s. 46.
[98] TOMEČEK, O.: Náboženské zápasy v zázemí mesta Banská Bystrica po nástupe reformácie. In: *Acta Facultatis Philosophicae Universitatis Ostraviensis. Historica*, 13, 2006, s. 57.
[99] MIČKOVÁ, *Banícka osada Špania Dolina* ..., s. 96 – 97.

byť Staré Hory, kde už stál neskorogotický kostol.[100] Významným bodom vo vývoji miestnej cirkevnej situácie bolo vypracovanie vyznania siedmych stredoslovenských banských miest v roku 1559. Toto vyznanie sa dá chápať aj ako priama odozva na snahy M. Oláha, ktorý sa pokúšal o rekatolizáciu banských miest. Dá sa predpokladať, že po vzore Banskej Bystrice, sa k tomuto vyznaniu pridali aj okolité osady.[101] Takto už v druhej polovici 16. storočia boli osadníci v okolí Španej Doliny verní protestantizmu a v roku 1593 bola pod dohľadom ich kazateľa Jána Nostitia špaňodolinská kaplnka prestavaná na kostol.[102] Špaňodolinská farnosť patrila v čase, keď bola spravovaná evanjelikmi pravdepodobne do Zvolenského seniorátu.[103]

Šírenie luteránskeho učenia sa spomalilo až v prvej polovici 17. storočia pod vplyvom rekatolizačných snáh Františka Forgáča a následne Petra Pázmaňa. Do pôsobenia miestnych kazateľov čoraz výraznejšie zasahovali aj vyššie politické záujmy. Tlak bol vyvíjaný najmä na komorských úradníkov v Banskej Bystrici, ktorých vplyv bol kľúčový pre okolité osady a sídla. Banskobystrická banská komora sa tak stala jedným z efektívnych nástrojov rekatolizácie tohto územia. Na miesta komorských úradníkov boli postupne dosádzaní výlučne len katolíci. Tlak sa presunul prostredníctvom komorských úradníkov aj na samotných baníkov evanjelického vyznania. Dochádzalo tak k efektívnej rekatolizácii radového robotníctva banskobystrickej banskej komory.[104] Významný rekatolizačný vplyv priniesli aj jezuiti. Prví dvaja jezuiti prišli do Banskej Bystrice 1. mája 1648 z rozhodnutia ostrihomského arcibiskupa Juraja Lipaia a usadili sa v dome kráľovskej komory.[105] Už v roku 1652 sa jezuiti usadili aj v Španej Doline, kde založili svoj misijný dom. Prvým jezuitským misionárom bol Leopold Rott a neskôr tu začali pôsobiť dvaja pátri, jeden nemecký a druhý slovenský.[106] Farnosť zotrvávala aj napriek priamemu pôsobeniu jezuitov evanjelická, čo nám potvrdzuje vizitačný zápis zvolenského archidiakona Štefana Bartóka, v ktorom sa spomína vyše 2000 evanjelikov a len 60 katolíkov v rámci tejto farnosti.[107]

Tvrdá forma rekatolizácie sa v Španej Doline začala prejavovať v 70. rokoch 17. storočia pod zostreným kurzom Leopolda I. voči evanjelikom. Evanjelickí duchovní z celej krajiny boli zvolávaní na Bratislavské súdy. Predvolaný bol aj špaňodolinský farár Adam

[100] MIČKOVÁ, *Banícka osada Špania Dolina* ..., s. 98.
[101] TOMEČEK, Náboženské zápasy v zázemí mesta Banská Bystrica ..., s. 58.
[102] MIČKOVÁ, *Banícka osada Špania Dolina* ..., s. 101.
[103] TOMEČEK, *Drevorubači a uhliari v lesoch Banskej Bystrice* ..., s. 168.; problematike sa tiež venuje MIČKOVÁ, *Banícka osada Špania Dolina* ..., s. 103.
[104] TOMEČEK, Náboženské zápasy v zázemí mesta Banská Bystrica ..., s. 59.
[105] MIČKOVÁ, *Banícka osada Špania Dolina* ..., s. 105.
[106] MIČKOVÁ, *Banícka osada Špania Dolina* ..., s. 105 – 106.
[107] TOMEČEK, *Drevorubači a uhliari v lesoch Banskej Bystrice* ..., s. 119.

Brestovinus, ktorý sa ale nedostavil. Veliteľ cisárskeho vojska si následne predvolal predstavených zo Špania Doliny, od ktorých žiadal kľúč od kostola a vyhnanie farára spolu s diakonom, rektorom a kantorom. Neuposlúchnutie tejto výzvy znamenalo vypuknutie otvorenej vzbury proti rekatolizácii v Špania Doline.[108] Cisárske vojská prišli do Špania Doliny dňa 13. októbra 1673 a vzbúrenci boli potlačení. Porážka vzbury znamenala začiatok nástupu rekatolizácie v celej starohorskej oblasti. Farár Brestovinus bol nútený odísť do vyhnanstva a farnosť bola udelená do správy jezuitom, čím sa stala katolíckou. Krátko nato rozšírili jezuiti svoje pôsobenie aj do Starých Hôr, ktoré boli v tomto období ešte len filiálkou farnosti v Špania Doline. Významnou osobnosťou v týchto časoch bol jezuita Alexander Metzger. Počas jeho približne 30 ročného pôsobenia sa podarilo rekatolizovať takmer celé obyvateľstvo Špania Doliny, Starých Hôr a priľahlých uhliarskych osád.[109]

Situácia sa zmenila opäť v prospech evanjelikov počas stavovského povstania Imricha Tökölihо, kedy došlo k vyhnaniu jezuitov. V júni 1682 tak kostol v Špania Doline opäť pripadol evanjelikom. Táto situácia ale nebola dlhodobá. Na základe matričných zápisov vieme, že jezuiti začali farnosť opäť spravovať najneskôr v roku 1694. Samotná farnosť v Špania Doline ale ešte čelila istým zmenám. Fara sa totiž v roku 1705 znovu dostala do rúk evanjelikom, tentoraz pod vplyvom Rákociho povstania. Komisia posudzujúca náboženské pomery rozhodla o odobratí kostola a fary katolíkom. Zápis v špaňodolinskej matrike potvrdzuje, že k odobratiu došlo koncom roku 1705.[110] V Špania Doline tak opäť vznikla komplikovaná situácia, ktorá bola umocnená častým prechádzaním z jednej strany na druhú. Odrážalo sa to aj na početných sťažnostiach domácich obyvateľov, ktoré predkladala komisii jedna aj druhá strana. Bežné boli najmä konflikty o kostolné vybavenie a rozdeľovanie financií.[111]

Jezuiti počas tohto obdobia zotrvali na Špania Doline a po vojenských neúspechoch Rákociho sa pomery opäť menili. Kostol bol evanjelickej cirkvi odobratý v roku 1709 a definitívne sa dostal do rúk katolíkov. Na začiatku 18. storočia sa tak katolícka konfesia na Špania Doline už úspešne upevnila, pričom podľa kanonických vizitácií Františka Ladislava Medňanského z roku 1713 sa v Špania Doline evidovalo 1505 katolíkov a len 3 evanjelici. Na Starých Horách to bolo 249 katolíkov a 44 evanjelikov. Tiež vieme, že v roku 1713 spravovali farnosť Špania Dolina dvaja pátri, jeden nemeckej a druhý slovenskej národnosti,

[108] TOMEČEK, Náboženské zápasy v zázemí mesta Banská Bystrica ..., s. 61.
[109] MIČKOVÁ, *Banícka osada Špania Dolina* ..., s. 107 – 108.
[110] TOMEČEK, *Drevorubači a uhliari v lesoch Banskej Bystrice* ..., s. 122.
[111] MIČKOVÁ, *Banícka osada Špania Dolina* ..., s. 112.

ktorých príjem činil 300 toliarov poskytovaných od baníckeho bratstva.[112] Táto situácia bola pomerne unikátna pre širšie okolie Banskej Bystrice. Dôvodom tohto vysokého počtu obyvateľov katolíckej konfesie v starohorskej oblasti je zrejme už spomínaný tlak, ktorý vyvíjala banskobystrická banská komora na svojich zamestnancov. Naopak vplyv komory nebol tak silný na samotné mesto, nakoľko cirkevné záležitosti v Banskej Bystrici riadila mestská rada a samotné obyvateľstvo nebolo závislé na práci pre banskú komoru.[113]

Činnosť jezuitov v tejto oblasti neutíchala ani v nasledovnom období a konfesionálne pomery sa po zvyšok 18. storočia vyvíjali už len v prospech katolíckej konfesie. V rokoch 1722 – 1723 došlo pod vplyvom jezuitov k veľkej prestavbe kostola v Špranej Doline, ktorý už nadobudol jednotný ráz a interiér dnešnú barokovú podobu.[114] V roku 1754 jezuiti začali s novou misiou na Motyčkách, kde vybudovali v roku 1755 kostol. Dochádza tak k úprave cirkevnej správy tohto územia, pričom sa horná časť Starohorskej doliny, od Motyčiek až po donovalské sedlo, priradila pod správu jezuitov na Motyčkách. To, že katolícka konfesia na tomto území našla silné zázemie potvrdzujú aj ďalšie kanonické vizitácie z druhej polovice 18. storočia. Kanonická vizitácia Jozefa Battyániho z roku 1754 nám už hovorí, že všetko obyvateľstvo farnosti v Špranej Doline a jej filiálok je považované za katolíkov.[115] Pozitívny rozvoj katolíckej konfesie na tomto území nespomalilo ani zrušenie jezuitského rádu, pričom jezuiti pôsobili v Špranej Doline až do roku 1777, teda ešte štyri roky po oficiálnom zrušení rádu.[116] Z tohto roku máme informáciu, že v obvode Staré Hory žilo 1381 obyvateľov, ktorí boli všetci uvedení ako katolíci.[117] V obvode Motyčky bola situácia totožná, pričom tu žilo 1016 obyvateľov. Aj na základe vizitácií z roku 1783, s ktorými narábame v tejto práci je obyvateľstvo vo farnostiach Staré Hory a Motyčky evidované ako čisto katolícke.[118] Konfesionálnu situáciu na začiatku 19. storočia je možné zistiť najmä z cirkevných schematizmov banskobystrického biskupstva. Tie vypovedajú, že územie bolo v tomto čase považované za úplne katolícku oblasť s minimálnym počtom obyvateľov evanjelického vyznania. Na základe týchto informácií sa tak dá táto oblasť po konfesionálnej stránke označiť ako definitívne katolícka už od polovice 18. storočia.

[112] CHLADNÁ, Z.: Vizitácia Špranej Doliny a Starých Hôr v roku 1713. In: *Bystrický permon,* 6, 2006, č. 2, s. 5.
[113] TOMEČEK, Náboženské zápasy v zázemí mesta Banská Bystrica ..., s. 60.
[114] MIČKOVÁ, *Banícka osada Špania Dolina ...,* s. 112.
[115] TOMEČEK, *Drevorubači a uhliari v lesoch Banskej Bystrice ...,* s. 123.
[116] MIČKOVÁ, *Banícka osada Špania Dolina ...,* s. 113.
[117] TOMEČEK, *Drevorubači a uhliari v lesoch Banskej Bystrice ...,* s. 124.
[118] Archív DA BB, Visitatio Canonica Ecclesiae Parochialis Moticsko , 1783, CV č. 10; Archív DA BB, Visitatio Canonica Ecclesiae Parochialis Vetero Montanae , 1783, CV č. 10.

Z pohľadu cirkevnej štruktúry bol pre toto územie obzvlášť významný rok 1776, kedy došlo k zriadeniu Banskobystrického biskupstva, ktorého prvým biskupom sa stal František Berchtoldt (1776 – 1793). Územie, z ktorého pochádzajú naše kanonické vizitácie spadalo pod novovzniknutý vicearcidiakonátny dištrikt Hornozvolenský. Samostatné farnosti sa nachádzali okrem Španej Doliny už aj na Starých Horách a Motyčkách. K výraznejším zmenám v cirkevnej správe na tomto území došlo ešte na konci 18. a začiatku 19. storočia, kedy sa cirkevná správa Banskobystrického biskupstva dotvárala. Vyčlenením z Hornozvolenského dištriktu vznikol Banskobystrický a neskôr aj Brezniansky a Ľubietovský dištrikt. Z celkového počtu 6 dištriktov, ktoré mala Banskobystrická diecéza pri svojom vzniku sa vytvorilo po vnútornej reorganizácii až 12 menších dištriktov. Na vyššej úrovni ešte došlo k vytvoreniu 4 arcidiakonátov banskobystrickým biskupom Antonom Makayom. Vznikol Katedrálny, Turčiansky, Tekovský a Nitriansky arcidiakonát.[119]

2.2 Vnútorná štruktúra vybraných kanonických vizitácií z roku 1783

Základná štruktúra využitých kanonických vizitácií sa skladá zo 14 hlavných bodov, uvádzacej časti a záveru, v ktorom sa k vizitácii vyjadruje samotný banskobystrický biskup.[120] Vstupné uvedenie vizitácie obsahuje názov farnosti, jej filiálok a následné zaradenie do cirkevnej štruktúry. V tomto prípade sa jedná o farnosti Banskobystrickej diecézy, patriacej do Hornozvolenského dištriktu. Ďalej uvádza samotného banskobystrického biskupa, ktorý dal tieto kanonické vizitácie vyhotoviť. V našom prípade sa jedná o Františka Berchtoldta. Záver vstupného uvedenia obsahuje dátum nariadenia patričnej kanonickej vizitácie.

Prvý bod vizitácií má názov *De Ecclesia* (O kostole) a venuje sa rozličným aspektom týkajúcim sa samotného kostola. Obsahuje opis stavu kostola a podáva informácie o jeho zasvätení a konsekrovaní. Nájdeme tu aj informácie o samotnej budove a jej rozmeroch. Napríklad v prípade kostola v Španej Doline sa dozvedáme, že bol 14 a trištvrte siah dlhý, 5 siah široký a 4 a pol siahy vysoký.[121] Ďalšiu dôležitú časť tohto bodu predstavujú údaje o oltároch, ktoré informujú najmä o ich typoch a počte. Ako príklad môžeme uviesť kostol

[119] TOMEČEK, *Drevorubači a uhliari v lesoch Banskej Bystrice ...*, s. 163 – 165.
[120] Vo vizitácii Starých Hôr, ktorá má len 10 hlavných bodov sa nenachádzajú body: *De Capellanis; De Regularibus; De Xenodochiis et aliis piis Fundationibus* a *De Haereticis*.
[121] Originálna jednotka, ktorá je vo vizitáciách použitá je tzv. *Orgia*. Jedná sa o dobovú jednotku dĺžky, ktorá predstavuje siahu, teda 1,9 m.

v Špánej Doline, kde sa spomínajú štyri pevné a päť prenosných oltárov. V Motyčkách sa nenachádzali žiadne pevné oltáre, len dva prenosné a v Starých Horách sa hovorí o štyroch pevných a štyroch prenosných oltároch. Samostatná pozornosť sa venuje hlavnému oltáru a jeho výzdobe a nachádzajú sa tu aj informácie o organe a zvonoch. V osobitnom bode sa vizitácia venuje posvätným olejom a spôsobom ich očistenia a posvätenia. Ďalej ponúka informácie o krstiteľnici a uchovávaní krstnej vody. Zo všeobecnejších informácií ponúka táto časť aj informácie o rozostavení a orientovaní lavíc, uchovávaní eucharistie, alebo umiestnení svätenej vody. Samostatne sa spomínajú spovednice. Dozvedáme sa, že v kostole v Špánej Doline boli dve spovednice, ktoré sú opísané ako pevné a nepoškodené. V Motyčkách sa spomína len jedna spovednica a v Starých Horách až štyri. V tomto bode môžeme nájsť aj informácie o kryptách, v našom prípade sa nachádzala len v starohorskom kostole. Záverečná časť je venovaná sochám a vyobrazeniam v danej farnosti. Ako príklad môžeme uviesť farnosť Špania Dolina, kde sa spomína kríž na cintoríne, medená socha sv. Trojice a kamenná socha sv. Jána Nepomuckého.

Druhý bod vizitácií nesie názov *De Supellectili Ecclesiae* (O vybavení kostola) a približuje vybavenie kostola. Jedná sa o obsiahly zoznam, ktorý obsahuje všetok relevantný kostolný inventár a základnú charakteristiku daných predmetov. Ako prvé tu nachádzame informácie o tzv. humeráloch. Vo vizitácii farnosti Špania Dolina sú tieto odevy rozdelené na poškodené a neporušené. Táto časť nám ponúka aj informácie o rúchach a prestieradlách. Pozornosť sa následne presúva na rôzne kalichy, čase, misky, podnosy, kadidelnice alebo cibórie. Osobitne sa tento bod venuje liturgickým rúcham, pričom sa opisuje ich vzhľad, farba a počty. Nachádzame tu aj informácie o svietnikoch, krčahoch alebo džbánikoch na rôzne účely. Záverečná časť tohto bodu hovorí o tom, či má farnosť pečať. V našom prípade mali všetky tri farnosti vlastnú pečať a ako príklad môžeme uviesť farnosť v Starých Horách, ktorá mala na pečati nápis *Sigillum Ecclesiae Parochialis Vetero Montanae, Dioecesis Neosoliensis.*

Tretí bod vizitácií sa nazýva *De proventu, et Dote Ecclesiae* (O príjmoch a daroch farnosti). Tento bod obsahuje podrobnejší rozpis všetkých príjmov farnosti z rôznych zdrojov a väčšinou je v podobe tabuľky. Hlavnú časť tohto bodu predstavuje rozpis ročných príjmov farnosti. Môžeme tu nájsť napríklad príjmy za zvonenie zvonov, prípravy pohrebov alebo financie, ktoré získala farnosť z darov poskytovaných veriacimi kostola. Tiež sa tu uvádza, či má farnosť vo vlastníctve nejaké polia, lúky alebo záhrady. Samostatnú časť tvoria dlhy farnosti. Zistiť sa tak dá od koho si farnosť požičala peniaze, resp. komu dlhujú. Takéto pôžičky sledujeme len v rámci farnosti Staré Hory, ktorá tu mala zaevidovaných

1000 florénov od mesta Banská Bystrica, 600 od mestečka Slovenská Ľupča, 200 od osoby *Mathias Hritz* a 70 od osoby *Catharina Mozor*. V závere sa dozvedáme aj o tom, ako a kým sú financie farnosti spravované. Tiež sa tu spomínajú rôzne inštitúcie, ktoré majú vplyv na financie farnosti. V tomto bode sú zahrnuté aj príjmy v naturáliách. Väčšinou sa jedná o olej, vosk, alebo víno.

Štvrtý bod vizitácií má názov *De Parocho eiusque obligatione* (O farárovi a jeho povinnostiach). Cieľom tohto bodu je predstaviť osobu farára a poskytnúť základné informácie o tom, akým spôsobom sa stará o farnosť. Prvá časť obsahuje opis osoby farára po rôznych stránkach. Prezentuje jeho meno, vek a jazyky, ktoré ovláda. Nájdeme tu aj informácie o meste, kde nadobudol svoje vzdelanie, ako dlho už je farárom a poprípade aj prehľad jeho predchádzajúcich pôsobísk. Ako príklad môžeme uviesť osobu farára v Starých Horách. Bol to istý *František Xaver Hesső*. Mal 52 rokov, hovoril dobre po slovensky, nemecky a čiastočne ovládal aj maďarčinu. Svoje štúdium dokončil vo Viedni, pričom farárom bol už 20 rokov. Podobné informácie vieme nájsť vo všetkých troch vizitáciách. Po predstavení osoby farára nasledujú informácie o tom, ako farár celebruje a káže. Ponúka sa tu aj prehľad toho, ako farár vykonáva svätú službu v nedeľu, alebo ako prebiehajú slávnosti rôznych sviatkov. Následne sa táto časť venuje tomu, ako farár prevádza katechézu a šírenie kresťanstva. Obzvlášť zaujímavým môže byť v rámci tejto časti bod č. 8, ktorý sa venuje knihám farnosti. Sú tu menované všetky knihy farnosti a hodnotí sa či farnosť používa knihy zodpovedajúce nariadeniam diecézy a katechizmu tridentského koncilu.

Piaty bod vizitácie sa nazýva *De Proventibus Parochi, et Domo Parochialis* (O príjmoch farára a farárskom dome). Na úvod sa táto časť venuje domu, v ktorom býval farár. Následne obsahuje detailný rozpis rôznych príjmov farára. Tieto príjmy sú rozdelené do rozličných kategórií. Väčšinou sa medzi kategóriami objavujú rôzne pevné príjmy, ktoré získava farár ročne, ale aj premenlivé príjmy, ktoré vychádzajú najmä z cirkevných úkonov, ako sú napríklad krsty, svadby alebo pohreby. Zaznamenané sú tu aj rôzne majetky, z ktorých by mohol farár nadobudnúť nejaké príjmy. Pevné príjmy farár získaval väčšinou priamo od mesta, alebo relevantnej inštitúcie. Ako príklad môžeme uviesť farára z Motyčiek, ktorého najväčším príjmom bolo 400 florénov, ktoré dostával ročne od banskobystrickej komory. Nachádzame tu aj informácie o tom, či farár dostáva časť svojho príjmu v naturáliách, alebo z vyzbieraných darov.

Šiesty bod vizitácií sa nazýval *De Capellanis* (O kaplánoch). Predstavuje kaplánov, ktorí vo farnosti pôsobili. Obsahuje detailnejší opis ich osoby a povinností. Kaplán sa

spomína len vo vizitácii Špan ej Doliny. Jednalo sa o 29 ročného kaplána menom *Anacletus Balázsy.* Narodený bol v Banskej Bystrici a ovládal slovenský a nemecký jazyk. Pozornosť sa následne obracia na rozličné aspekty kaplánovho pôsobenia. Sú tu informácie o jeho plate a dlhoch, ale aj informácie o jeho práci, o tom či vedie omše, alebo píše kázne. Ďalej sa spomína aj to, ako sa správa a či má dobré zvyky.

Siedmy bod nesie názov *De regularibus* (O rádoch). Vizitácie Motyčiek a Špan ej Doliny, ktoré tento bod obsahujú hovoria len o tom, že sa v daných farnostiach nenachádzalo žiadne rehoľné sídlo.[122] V prípade Motyčiek ale vieme, že pred zrušením jezuitského rádu sa tu nachádzala jezuitská misionárska stanica.

Ôsmy bod sa nazýva *De dominis Terrestribus* (O zemepánoch). Táto časť obsahuje informácie o možných zemepánoch a patrónoch danej farnosti. Po informačnej stránke sa jedná o skromnejší bod, pričom prevažne oboznamuje len s inštitúciou, ktorá je zemepánom a patrónom farnosti. V prípade farnosti Motyčiek sa tu ako zemepán a zároveň patrón farnosti uvádza Cisársko-kráľovská komora v Banskej Štiavnici (*Caesareo Regia Camera Schemniczien sis)*. Vo farnosti na Starých Horách je uvedená Cisársko-kráľovská výsosť (*Sacratissima Caesareo Regia Maiestas)*, teda panovník, ktorý je zemepánom aj patrónom farnosti. Vo vizitácii Špan ej Doliny sa podobne uvádza Cisársko-kráľovská a apoštolská výsosť (*Sacratissima Caesarea Regia et Apostolica Maiestas)* ako zemepán aj patrón farnosti.

Deviaty bod vizitácií sa nazýva *De oeconomo, seu Inspectore Ecclesiae* (O hospodároch a inšpektoroch farnosti). Tento bod sa venuje tomu, či má farnosť nejakých hospodárov, alebo inšpektorov. Okrem samotnej osoby hospodára, alebo inšpektora sa zaoberá aj tým, ako je zaobchádzané s príjmami farnosti a ako prebieha ich rozdeľovanie. Ani v jednej zo sledovaných vizitácií sa nespomína žiadny priamy hospodár alebo inšpektor. Jediné meno sa nachádza vo vizitácii Motyčiek. Tu je uvedený istý *Andreas Longauer*, ktorý ako kostolník zastupoval niektoré z týchto povinností.

Desiaty z hlavných bodov sa nazýva *De Parochianis* (O farníkoch). Nachádzame tu základné informácie o obyvateľstve, ktoré patrilo do danej farnosti. Na úvod sa venuje počtom obyvateľov patriacich do farnosti. Počty sa uvádzajú v tzv. „dušiach" (*animae*). Obvykle sú obyvatelia rozdelení na dospelých a deti. Táto časť informuje aj o jazyku, ktorým obyvatelia v rámci farnosti hovorili. Vo farnosti Motyčky sa registrovalo celkovo 1053 obyvateľov, z toho 769 dospelých. Všetci hovorili po slovensky. Vo farnosti na Starých

[122] *Domicilium Regularium hic loci non datur.*

Horách bolo celkovo 473 obyvateľov, z toho 341 dospelých a aj tu hovorili všetci po slovensky. Vo vizitácii farnosti Španej Doliny sa spomína 1989 obyvateľov, z čoho bolo 1554 dospelých. Všetci hovorili po slovensky, no spomínala sa menšia skupina, ktorá hovorila aj po nemecky. Tento bod sa následne zameriava na to, ako farníci napĺňajú svoje cirkevné povinnosti. Venuje sa tomu ako sú obyvatelia oddaní viere, ako často navštevujú kostol a či dôkladne rešpektujú farára. Nachádzame tu aj informácie o dodržiavaní a napĺňaní veľkonočných zvykov. Samostatná časť hovorí o tom, či sa vo farnosti nachádzajú nejakí „previnilí obyvatelia" ako napríklad závažní hriešnici, úžerníci, cudzoložníci alebo porušovatelia pôstu. Záverečná časť informuje o konfesionálnej situácii medzi farníkmi. Spomína sa, či boli objavení nejakí odpadlíci od viery a ponúka aj informáciu o tom, či existujú v rámci farnosti nejaké zosobášené páry rozličných vyznaní.

Jedenásty bod sa nazýva *De Servis Ecclesiae* (O osobách slúžiacich vo farnosti). Táto časť sa venuje rozličným osobám, ktoré sa podieľajú na chode farnosti. Je rozdelená do niekoľkých menších podkapitol. V rámci prvej podkapitoly sa venuje najmä učiteľovi a škole. Približuje základné informácie o osobe a príjmoch učiteľa, no obsahuje aj informácie o tom, kedy a ako vyučoval. Pri škole väčšinou nachádzame aj informáciu o počte žiakov. Ako príklad uvedieme farnosť v Španej Doline, kde pôsobili dvaja učitelia. Prvým bol *Josephus Derer,* ktorý mal 30 rokov a druhým bol *Josephus Schmidt,* ktorý mal 26 rokov a zároveň bol organistom. Učili čítanie, písanie a počítanie. Celkový počet žiakov bol 106, pričom žiaci tu boli rozdelení do štyroch skupín. Druhá podkapitola v rámci tohto bodu sa nazýva *De Obstertrice* (O pôrodných babách). V stručnosti predstavuje pôrodné baby, ktoré pôsobili vo farnosti. Napríklad v Španej Doline sa nachádzali dve pôrodné baby: *Catharina Mrva,* ktorá mala 68 rokov a *Anna Zsilla,* ktorá mala 49 rokov. V Starých Horách sa nachádzala len jedna pôrodná baba a v Motyčkách až štyri. Poslednou podkapitolou tohto bodu je *De Vespillone* (O hrobárovi), ktorá predstavuje hrobára a náplň jeho práce. Nájdeme tu aj informácie o hroboch a o tom, či sú dôsledne rozostavené a dostatočné. V Španej Doline boli dvaja hrobári. Prvým bol *Franciscus Vaitzik* a druhým bol *Joannes Szobor.* V Motyčkách bol hrobárom *Matthias Longauer* a v Starých Horách *Joannes Zakaliczki*.

Dvanásty bod vizitácií sa nazýva *De Xenodochiis, et aliis piis Fundationibus* (O chudobinci) Aj keď vizitácie farností v Španej Doline a Motyčkách tento bod obsahujú, ani v jednej z nich sa chudobinec nespomína. Tento bod sa venuje aj chudobným osobám, resp. žobrákom. V Španej Doline sa udávajú 4 muži a 8 žien a v Motyčkách 4 muži a 2 ženy. Samostatná časť hovorí aj o tom, či sa vo farnosti nachádzajú príslušníci cigánskeho etnika

(*Zingari*). Vizitácia Špranej Doliny a rovnako tak aj Motyčiek hovorí, že sa vo farnosti žiadne takéto osoby nenachádzali.

Trinásty bod vizitácií má názov *De Haereticis* (O heretikoch). Vizitácie Motyčiek ani Špranej Doliny nespomínajú žiadnych heretikov. V rámci vizitácie Starých Hôr sa táto časť vôbec nenachádza.

Posledný bod sledovaných vizitácií sa nazýva *De Filialibus* (O filiálkach). Tento bod predstavuje všetky filiálky danej farnosti, pričom ponúka základné informácie o každej z nich. Treba ale poznamenať, že množstvo informácií, ktoré sa k jednotlivým filiálkam vzťahuje je značne premenlivé. Na úvod sa v tomto bode udáva celkový počet filiálok, ktoré patria pod danú farnosť. Jednou z prvých informácií je ich vzdialenosť od centra farnosti. V niektorých prípadoch môžeme nájsť aj informácie o samotnej ceste, ktorá vedie do jednotlivých filiálok. Napríklad v prípade farnosti Špania Dolina sa uvádza, že cesta do filiálky Piesky je nerovná horská cesta, ktorá môže byť v zime ťažko dostupná. Ďalej táto časť obsahuje informácie o kaplnkách alebo kostoloch, ktoré sa mohli vo filiálkach nachádzať. Môžu sa tu objaviť aj informácie o školách a učiteľoch, ako napríklad v prípade vizitácie Motyčiek. Vo filiálke Donovaly je spomenutý učiteľ menom *Michal Longauer*. Tento bod ďalej ponúka aj informácie o počte obyvateľov v jednotlivých filiálkach. Najdetailnejší pohľad na filiálky ponúka vizitácia farnosti v Starých Horách. Všetkým jej filiálkam sa venuje osobitne a uvádza najväčšie množstvo informácií. Nachádzajú sa tu informácie o vzdialenosti filiálky a ceste, existencii kostolov a kaplniek a pri každej filiálke sa uvádza aj počet veriacich, pričom sú rozdelení na dospelých a deti. Rovnako je tu popísaný aj jazyk, akým obyvatelia hovoria a venuje sa aj cirkevným zvykom a povinnostiam veriacich vo filiálkach. Spomína sa tu aj to, ako často navštevujú kostol a ako aktívne sa podieľajú na cirkevnom živote.

3. Metóda automatickej transkripcie a možnosti jej využitia v historickom výskume

3.1 Postup pri využití metódy automatickej transkripcie

Automatickú transkripciu vybraných rukopisných textov budeme zhotovovať využitím platformy Transkribus, ktorá predstavuje základný kameň nášho výskumu. Platforma Transkribus je program na automatické textové rozpoznávanie historických rukopisov, založený na metódach umelej inteligencie a strojového učenia. Transkribus bol vytvorený Univerzitou v Innsbrucku za spolupráce s rozličnými európskymi výskumnými skupinami ako súčasť výskumného projektu Horizon 2020 (READ).[123] V súčasnosti sa projekt naďalej rozvíja a ponúka sa na komerčné, vedecké a vzdelávacie využitie v rámci združenia READ-COOP SCE, ktoré má v aktuálnej dobe už 102 členov z 26 krajín.[124]

Proces automatickej transkripcie sa skladá z niekoľkých hlavných krokov. Predstavíme tu v skratke len tie najhlavnejšie, ktoré je potrebné vykonať v rámci akéhokoľvek postupu automatickej transkripcie s využitím platformy Transkribus, pričom detailnejšie možnosti samotného postupu sa môžu líšiť od prípadu k prípadu.

Prvým krokom je digitalizácia a následné nahranie dokumentu, s ktorým mienime pracovať do platformy Transkribus. Rukopisný dokument je potrebné najskôr zdigitalizovať, teda previesť ho do digitálneho grafického súboru, pričom Transkribus je kompatibilný so všetkými najpoužívanejšími formátmi ako JPEG, PNG alebo PDF. Keď je dokument úspešne nahraný do Transkribusu je možné pristúpiť k ďalšiemu kroku automatickej transkripcie, ktorým je segmentácia.

Segmentácia predstavuje rozčlenenie rukopisného dokumentu na textové časti. Prostredníctvom rôznych nástrojov vytvoríme detailnú štruktúru textu v danom dokumente. Tento krok slúži na to, aby program Transkribus dokázal rozoznať kde má v rámci dokumentu hľadať a čítať text. Rukopisný text teda rozdelíme do textových polí podľa potreby a následne v rámci týchto polí vyznačíme riadky a poradie, v ktorom budú čítané. Segmentáciu je možné aj automatizovať. V tomto prípade sa platforma Transkribus pokúsi samostatne nájsť text a automaticky ho vyznačiť a včleniť do jednotlivých polí a riadkov.

[123] https://readcoop.eu/transkribus/. Prevzaté 11.4.2022.
[124] NAGY, I.: Možnosti aplikácie metódy digitálnej transkripcie historických rukopisných textov ..., s. 54.; tiež: https://readcoop.eu/members. Prevzaté 11.4.2022.

Samozrejme automatická segmentácia vo veľkej miere závisí nielen od kvality digitalizácie, ale aj od samotnej kvality a zachovania dokumentu. V prípade dobre nafoteného dokumentu s jasným riadkovaním je automatická segmentácia dostatočne spoľahlivá a na používateľa zostáva už len prípadná kontrola a korekcia. Ak je ale dokument alebo jeho digitálna kópia v horšom stave, je potrebné vykonať segmentáciu manuálne. Takáto situácia môže vzniknúť najmä pri preexponovaných, alebo príliš tmavých fotografiách, kedy platforma Transkribus nedokáže dostatočne rozoznať text od pozadia. Pri manuálnej segmentácii používateľ využíva nástroje platformy Transkribus, aby manuálne vyznačil textové polia a následne jednotlivé riadky.

Po vykonaní či už automatickej alebo manuálnej segmentácie rukopisného dokumentu je možné prejsť na ďalšiu časť automatickej transkripcie, ktorou je vytváranie modelu. Alternatívne je možné využiť aj jeden z už existujúcich modelov, ak je písmo skúmaného dokumentu s daným modelom dostatočne podobné. V prípade, že používateľ chce model, ktorý sa naučí čítať špecificky písmo v danom dokumente, je potrebné vytvoriť samostatný model. Vďaka segmentácii už platforma Transkribus vie zistiť kde sa text, ktorý má čítať nachádza. Cieľom vytvorenia a trénovania modelu je naučiť Transkribus ako čítať daný rukopis a jednotlivé znaky v ňom. Jednoducho povedané, musíme program naučiť aký alfanumerický znak má Transkribus priradiť k jednotlivým znakom nachádzajúcim sa v originálnom texte. To akým spôsobom sa platforma Transkribus učí čítať rukopisný text zadefinoval I. Nagy nasledovne. „Transkribus je založený na strojovom učení, pričom obrazovej predlohe znaku na origináli priraďuje konkrétny alfanumerický znak. Transkribus tento prevod porovná s „Ground Truth“, skutočnou hodnotou, ktorú by mal tejto predlohe priradiť. Ak zistí nesprávnu hodnotu, vyradí ju z množiny možných riešení. Opakovaním tohto procesu sa množina možných riešení, ktoré Transkribus priraďuje obrazovej predlohe znaku postupne zmenšuje až na hodnotu rovnú skutočnej hodnote znaku. Stroj sa takto naučí ako má konkrétny znak čítať. Následne mu už bude pri každom jeho výskyte priraďovať rovnakú hodnotu.“[125]

Z toho teda vyplýva, že po segmentácii dokumentu je potrebné pripraviť tzv. Ground Truth vzorku textu, ktorá vznikne odborným prepisom istej časti dokumentu, s ktorým pracujeme.[126] Na základe tejto vzorky naučíme Transkribus ako má čítať jednotlivé znaky v texte. Tvorba tejto vzorky predstavuje zrejme tú najdôležitejšiu časť automatickej

[125] NAGY, Možnosti aplikácie metódy digitálnej transkripcie historických rukopisných textov ..., s. 56.
[126] Ground Truth predstavuje tzv. skutočnú hodnotu. Jedná sa vzorku písma, ktorá vznikne odbornou transliteráciou časti dokumentu.

transkripcie, nakoľko od nej závisí jej celková presnosť. Je teda potrebné vytvoriť paleografickú transliteráciu časti nášho dokumentu. Pri tejto transliterácii je nutné dbať na presnosť a precíznosť daného prepisu, pričom samotný prepis musí byť exaktný. To znamená, že prepis musí obsahovať aj všetky chyby, ktoré sa v texte nachádzajú, bez akýchkoľvek korekcií. Aj z tohto dôvodu je pri vytváraní Ground Truth vzorky potrebná minimálne základná znalosť paleografických pravidiel, jazyka v ktorom je dokument napísaný a tiež znalosť typu prameňa s ktorým narábame. Takýmto spôsobom priradíme ku každému riadku, ktorý sme v rámci segmentácie vyznačili presné alfanumerické znaky.

Po vytvorení Ground Truth vzorky môžeme pristúpiť k trénovaniu nami vytvoreného modelu. Cieľom trénovania modelu je zníženie jeho chybovosti, aby sme jeho následným využitím dosiahli čo najpresnejšie výsledky pri automatickej transkripcii zvyšku dokumentu. Pri trénovaní modelu využijeme spomínanú Ground Truth vzorku, ktorú sme vytvorili. Túto vzorku rozdelíme na dve časti. Prvou časťou je cvičný súbor (Training set) a druhou overovací súbor (Validation set). Trénovanie modelu prebieha v tzv. cykloch. V rámci každého cyklu sa Transkribus pokúsi prečítať strany v cvičnom súbore a zapamätá si, kde urobil chyby. Tieto chyby už v nasledujúcom cykle nezopakuje a takýmto spôsobom sa presnosť modelu zdokonaľuje. Znalosti, ktoré Transkribus nadobúda čítaním cvičného súboru následne aplikuje na overovací súbor. Pri overovacom súbore vzniká rozdiel v tom, že Transkribus si nepamätá ako tento súbor prečítal predtým a pristupuje k nemu vždy ako k novému textu. Tu sa dozvedáme reálnu presnosť modelu, ktorý trénujeme.

Keď máme k dispozícii model, či už niektorý z existujúcich, alebo svoj vlastný môžeme pristúpiť k automatickej transkripcii zvyšku rukopisného dokumentu. Pred automatickou transkripciou je potrebné ešte vykonať segmentáciu strán, ktoré chceme automaticky transkribovať. Následne už zostáva len zvoliť model a môžeme spustiť proces automatickej transkripcie. Výstupom by mal byť prepis zvyšku dokumentu s približnou chybovosťou závislou od využitého modelu. Takýto text je následne pripravený na ďalší výskum a prácu s ním.

3.1.1 Využitie metódy automatickej transkripcie v historickom výskume

Širšia implementácia metódy automatickej transkripcie v rámci historického výskumu môže mať veľmi pozitívny vplyv na rozvoj slovenského historického výskumu

hneď v niekoľkých smeroch. Výrazné zefektívnenie môže predstavovať nielen v rámci väčších výskumných úloh, zameraných na sprístupňovanie prameňov, ale aj pre výskum jednotlivcov, ktorí pri svojej práci narábajú s väčším počtom historických rukopisných prameňov.

Najvýraznejší dopad metódy automatickej transkripcie sa dá očakávať v rámci digitalizácie a sprístupňovania historických rukopisov. Digitalizácia predstavuje proces konverzie analógových materiálov do digitálnej formy.[127] Proces digitalizácie je už dlhodobým fenoménom aj v rámci výskumu histórie, pričom jej význam v dnešnej dobe už nemožno podceňovať. Širšia digitalizácia historického písomného dedičstva môže otvoriť brány úplne novým výskumom využívajúcim rukopisné dokumenty, ktoré boli doteraz prakticky nedotknuté a zlepšiť tak stav slovenského historického výskumu ako celku. Doteraz leží v archívoch veľké množstvo rozsiahlych historických rukopisných dokumentov, ktoré predstavujú veľké, no minimálne využité a prakticky neprístupné archívne bohatstvo. Jedným z hlavných prínosov digitalizácie v praxi je najmä dostupnosť. Digitálne zbierky si môže časovo efektívne prezerať prakticky ktokoľvek a kedykoľvek. Potenciál digitalizácie zdôraznil D. Katuščák, ktorý uviedol, že digitalizácia umožňuje premenu pamäťových a fondových inštitúcií na moderné organizácie, ktoré dokážu reagovať na potreby a požiadavky spoločnosti. Digitalizácia predstavuje prostriedok, pomocou ktorého inštitucionalizovaná kultúra dokáže získať lepšiu kontrolu nad kultúrnym dedičstvom.[128]

Ako ale poznamenal I. Nagy digitalizácia slovenských archívnych fondov má v dnešnej dobe ešte stále veľké rezervy, pričom funkčný elektronický archív je záležitosťou vzdialenej budúcnosti. Dôvodom sú mnohé nespracované archívne fondy bez opisov, inventárov, katalógov a registrov, čo sťažuje orientáciu a vyhľadávanie v nich.[129] Práca s takýmito dokumentami sa stáva časovo náročnou a výrazne neefektívnou bádateľskou činnosťou, ktorá mnohokrát odrádza historikov a tvorí prekážku v potencionálnom historickom výskume. Práve tu sa dá hľadať primárny prínos rozvoja a širšej implementácie automatickej transkripcie.

Digitalizácia historických rukopisov kráča prakticky ruka v ruke s ochranou a najmä sprístupňovaním tohto archívneho dedičstva. Digitalizácia písomného dedičstva je len

[127] KATUŠČÁK, D.: *Digitalizacia. Úvod do digitalizácie písomného dedičstva. Distanční studijní text*. Opava. 2019. s. 7. Dostupné na internete: https://is.slu.cz/el/fpf/zima2020/UBKVKKNP03/Digitalizace_a_dlouhodoba_ochrana_fondu_1.pdf?lang=en
[128] KATUŠČÁK, *Digitalizacia. Úvod do digitalizácie písomného dedičstva* ..., s. 10.
[129] NAGY, Možnosti aplikácie metódy digitálnej transkripcie historických rukopisných textov ..., s. 53.

prvým krokom, ktorý zaručí uchovanie a ochranu dokumentov. Prirodzeným druhým krokom je následné sprístupňovanie tohto písomného dedičstva pre rôzne účely, či už vedecké, výskumné alebo vzdelávacie. Práve automatická transkripcia môže umožniť veľmi efektívne sprístupnenie rozsiahleho množstva historických rukopisných dokumentov nielen odbornej komunite, ale aj širokej verejnosti. Zdôrazniť sa tu dá veľký potenciál najmä v oblasti publikovania pramenných edícií. Samostatne treba zmieniť aj spomínaný elektronický, resp. digitálny archív, ktorý by dokázal ponúknuť rozsiahle slovenské archívne dedičstvo v digitálnej podobe na jednom mieste. Dôkaz o tom, že niečo takéto je možné netreba hľadať ďaleko. Vynikajúcim príkladom v tomto smere je maďarský portál kultúrneho dedičstva Hungaricana, ktorý disponuje skutočne rozsiahlymi a kompletne zdigitalizovanými archívnymi databázami. Jedná sa pritom o rozsiahly projekt, ktorý spája archívy po celom Maďarsku a ponúka historické dokumenty z vyše 100 rozličných inštitúcií.[130] Táto databáza obsahuje celkovo stotisíce zdigitalizovaných historických listín a dokumentov. Osobitne spomenieme niektoré významnejšie databázy, napríklad databáza stredovekých dokumentov, databáza Urbára Márie Terézie z roku 1767, ktorá obsahuje vyše 300 000 strán, databáza kráľovských kníh z rokov medzi 1526 až 1918 a v neposlednom rade zbierka Národného archívu Maďarska: *Urbaria et Conscriptiones*, ktorá obsahuje vyše 300 000 zdigitalizovaných strán obsahujúcich informácie o pozemkových majetkoch a zemepánoch. Tu vidíme potenciál, ktorý môže priniesť digitalizácia a sprístupňovanie písomného dedičstva pre rozvoj historického výskumu a rovnako tak aj potenciál jedného z možných využití platformy Transkribus.

Pri detailnejšom pohľade na možnosti automatickej transkripcie sa ako najvýraznejšia výhoda javí možnosť tzv. plnotextového vyhľadávania, ktoré by výrazným spôsobom zefektívnilo prácu najmä s rozsiahlymi rukopisnými dokumentami. Ak teda bádateľ potrebuje extrahovať presné údaje z rukopisných dokumentov, nemusel by už prácne prechádzať celé zbierky, alebo dokumenty od začiatku do konca. Po využití metódy automatickej transkripcie, by vedel veľmi efektívne vyhľadať potrebné údaje v danom dokumente a jednoducho nájsť všetky ich výskyty. Navyše ak sa jedná len o osobnú pomôcku pri výskume, je možné prehliadnuť aj rozličné chyby, ktoré pri automatickom prepise vznikli, nakoľko dôraz by sa kládol na nájdenie relevantných údajov a nie na potencionálnu presnosť prepisu. Dôraz na presnosť prepisu by sa zase kládol v prípade, že by sa metóda automatickej transkripcie využívala pri tvorbe pramenných edícií, ktoré majú

[130] Portál maďarského kultúrneho dedičstva Hungaricana: https://hungaricana.hu/en/about/

byť následne publikované. Vidíme teda pomerne všestranné využitie, ktoré môže mať automatická transkripcia v rámci historického výskumu. Či už ako osobná pomôcka pri výskume, ale aj ako nástroj pre efektívne publikovanie prameňov.

Na záver je potrebné poznamenať, že rozvoj platformy Transkribus a automatickej transkripcie na území Slovenska je ešte len v začiatkoch. Aktuálne prebieha počiatočná fáza adaptácie tejto platformy na historické rukopisy zo slovenského prostredia. Táto činnosť prebieha prevažne prostredníctvom Katedry histórie FF UMB a odborných pracovníkov Štátnej vedeckej knižnice v Banskej Bystrici. Ako už bolo poznamenané jednu z najdôležitejších častí automatickej transkripcie tvorí transliterácia využívaných dokumentov, ktorá musí byť presná a odborná. Aj z tohto dôvodu stojí rozvoj automatickej transkripcie historických rukopisných dokumentov prevažne na poli historického výskumu. Automatickú transkripciu je ale jednoznačne možné považovať za interdisciplinárnu metódu, ktorá čerpá nielen z odboru histórie, ale aj paleografie, knižničnej vedy a archívnictva. Ako naznačujú prvé výsledky v tomto smere výskumu, automatická transkripcia a Transkribus určite má svoje miesto v rámci rozvoja slovenského historického výskumu, pričom s pozitívnym vplyvom širšej implementácie metód automatickej transkripcie na vývoj historického výskumu sa dá len ťažko polemizovať.

3.2 Nahrávanie dokumentov, segmentácia a vytváranie Ground Truth vzorky

Predtým ako sa dostaneme k jednotlivým postupom vytvárania modelov, predstavíme niekoľko krokov, ktoré sú spoločné pre všetky tieto postupy. Jedná sa o nahratie dokumentov, ich segmentáciu a vytváranie Ground Truth vzorky.

Prvým krokom bolo nahratie využitých rukopisných dokumentov do platformy Transkribus. Dokumenty, s ktorými pracujeme sú tri zápisnice kanonických vizitácií, nafotené v dvojstranách, ktoré tvoria celkovo 22 obrazov. Nakoľko sa jedná o menší počet obrazov, rozhodli sme sa ich pomocou grafického editoru rozdeliť na jednotlivé strany, ktoré sme nahrali do Transkribusu. Celkovo sme tak získali 38 jednotlivých strán. Kanonické vizitácie farností Špania Dolina a Staré Hory obsahovali po 12 strán a kanonická vizitácia farnosti Motyčky 14 strán. Toto rozdelenie strán nám umožnilo najmä lepšie a detailnejšie narábanie s dokumentmi pri ich rozdeľovaní na cvičný a overovací súbor.

Ako už bolo spomenuté, našim cieľom je zistiť ako sa dokáže malý model adaptovať na odlišné rukopisy. Všetky tri kanonické vizitácie, ktoré využívame sú napísané v mierne odlišných rukopisoch. Kvôli prehľadnosti ich budeme označovať aj pomocou písmen. Kanonickú vizitáciu farnosti Špania Dolina budeme označovať ako rukopis „A“, kanonickú vizitáciu farnosti Staré Hory ako rukopis „B“ a kanonickú vizitáciu farnosti Motyčky ako rukopis „C“.

Po nahratí dokumentov do platformy Transkribus sme prešli k druhému kroku, ktorým bola segmentácia. Pri každej z vizitácii sme využili možnosť automatickej segmentácie. Ako prvej sme sa venovali vizitácii Špánej Doliny, čiže rukopisu A. Automatická segmentácia v tomto prípade našla všetky riadky a väčšinu z nich aj správne vyznačila, pričom nemala problém ani so zmenami naklonenia riadkov. Stretli sme sa ale aj s väčšími komplikáciami, ktoré po automatickej segmentácii vznikli. Prvým a zrejme tým najvýraznejším problémom automatickej segmentácie tohto dokumentu bolo pomerne časté nesprávne vyznačovanie konca riadkov. V praxi to vyzeralo tak, že posledné písmeno v riadku, alebo jeho časť bola odseknutá, čo vyžadovalo manuálnu úpravu vyznačenia riadku. Táto chyba sa objavovala po segmentácii tohto dokumentu skutočne často, pričom bolo potrebné manuálne upravovať približne každý tretí riadok. Efektívnosť automatickej segmentácie tak bola v prípade tohto dokumentu značne znížená. Zrejme ako hlavný dôvod nesprávneho vyznačovania konca riadkov sa dá určiť „ohraničenie“ strán. Okolo textu sa totiž nachádzali čiary, ktoré vyznačovali hranice, do ktorých sa vpísaval text. Všimli sme si, že ak riadok končil príliš blízko tejto čiary, mala automatická segmentácia problém so správnym vyznačením takýchto riadkov (pozri obr. 1).

Druhý problém, ktorý vznikal pri automatickej segmentácii sa týkal číslovania. Kanonické vizitácie sú rozdelené na jednotlivé časti, pričom každá z nich je očíslovaná. Toto číslovanie bolo automatickou segmentáciou vyznačené nepresne vo veľkej väčšine prípadov. Dôvodom môže byť kombinácia väčších čísel v spojení s menším textom, čo spôsobovalo spomenutú nepresnosť segmentácie. Z tohto dôvodu sme museli číslovanie segmentovať manuálne v rámci celého dokumentu (obr. 2).

Posledným výraznejším problémom, na ktorý sme narazili pri automatickej segmentácii boli tabuľky. Jednalo sa o jednoduché tabuľky priamo v texte bez výraznejšieho vyznačenia ich riadkov a stĺpcov. Automatická segmentácia bola pri takýchto tabuľkách značne nepresná. Chyby tu vznikali vo viacerých aspektoch. Riadky týchto tabuliek boli mnohokrát „vybodkované“. To spôsobovalo Traskribusu problém zistiť kde sa nachádza, resp. končí text a mnohokrát to viedlo k nesprávnemu vyznačeniu riadkov. Problémy

vznikali aj pri segmentácii čísel v týchto tabuľkách, ktoré bývali často nepresne vyznačené a odseknuté. Komplikáciu tvoril aj oddeľovací znak alebo pomlčky a čiarky, ktorými boli vyznačené prázdne miesta v tabuľkách (obr. 3). Podobne ako pri číslovaní bolo aj pri týchto tabuľkách najlepšou voľbou vykonať ich segmentáciu manuálne. V rámci tohto dokumentu sme sa stretli aj s občasným miernym presvitaním textu z druhej strany, s čím ale automatická segmentácia nemala žiadny problém. Podobne ani občasné fľaky na papieri nemali viditeľne negatívny vplyv na efektivitu automatickej segmentácie.

Ako druhú sme vykonali automatickú segmentáciu kanonickej vizitácie Starých Hôr, čiže rukopisu B. Tá predstavovala v tomto smere špecifický dokument. Dôvodom bolo, že na stranách tohto dokumentu bol veľmi výrazne prepitý atrament z opačnej strany (obr. 4). Automatická segmentácia ale prekvapivo nemala takmer žiadne problémy s týmto výrazným presvitaním opačných strán. Riadky boli vyznačené presne a zdá sa, že Transkribus nemal problém medzi rozoznávaním textu nachádzajúcom sa na strane a textom, ktorý presvital z opačnej strany. Vyskytlo sa ale aj pár ojedinelých prípadov, kedy automatická segmentácia vyznačila ako text aj malú časť presvitajúceho textu. Stávalo sa tak najmä v prípade, keď nebol okolo presvitajúceho textu žiadny skutočný text.

Aj v rámci tohto dokumentu sme pri automatickej segmentácii narazili na podobné problémy ako pri predchádzajúcom dokumente. Problém s nesprávnym zaznačením konca riadkov sa objavoval pomerne často aj v tomto dokumente. Podobne ako pri vizitácii Špane Doliny, aj tu sa stretávame s výraznými okrajovými čiarami, ktoré spôsobovali nepresné vyznačenie konca riadkov. V mnohých prípadoch, kedy sa riadok končil v bezprostrednej blízkosti tejto čiary bol jeho koniec nesprávne vyznačený a posledné písmeno v riadku bolo odseknuté. Zaujímavé v tomto smere bolo, že v riadkoch, ktoré pisár natiahol výraznejšie za túto čiaru sa s týmto problémom nestretávame. Tieto riadky sú automatickou segmentáciou väčšinou vyznačené správne až do ich konca, pričom to že prechádzajú cez túto čiaru sa nejaví ako komplikácia. To nám opäť naznačuje, že dôvodom nesprávneho segmentovania koncov riadkov môže byť práve okrajová čiara, nakoľko riadky, ktoré ju presahujú boli segmentované väčšinou správne a riadky, ktoré končili v jej blízkosti mali koniec posledného slova odseknutý.

Aj v prípade tohto dokumentu sa stretávame s nesprávnym segmentovaním číslovania. Celkovo sa dá povedať, že v rámci tohto dokumentu bola automatická segmentácia pri vyznačovaní spomínaného číslovania presnejšia. Napriek tomu bolo ale stále efektívnejšou voľbou vykonať segmentáciu tohto číslovania manuálne. Mierne lepšie

výsledky pri segmentácii číslovania v rámci tohto dokumentu sa dajú zrejme pripísať tomu, že číslo a písmená pri ňom sú v približne rovnakej veľkosti.

Tabuľky sa objavujú aj v tejto vizitácii a opäť predstavovali pre automatickú segmentáciu problém. V tomto dokumente mali tabuľky skutočne voľnú podobu, nakoľko sa tu nenachádzali žiadne vyznačené bunky. Rovnako ako pri predchádzajúcej vizitácii aj v tomto prípade bolo jednoduchšie a efektívnejšie vykonať segmentáciu týchto tabuliek a zoznamov manuálne.

Ako poslednú sme vykonali automatickú segmentáciu vizitácie farnosti Motyčky, čiže rukopisu C. Pri tomto dokumente sme dosiahli automatickou segmentáciou najlepšie výsledky. Riadky boli vyznačené prevažne presne na ich začiatku aj konci. V tejto vizitácii sa bočné čiary na ohraničenie strany síce nachádzajú, no sú výrazne svetlejšie oproti textu. Opäť sa nám teda potvrdzuje teória, že dôvodom nepresnosti segmentácie koncov riadkov v predchádzajúcich vizitáciách bola práve bočná čiara, nakoľko konce aj začiatky riadkov boli v tomto dokumente po automatickej segmentácii väčšinou vyznačené presne. Problém s nedotiahnutím riadku sa nám ale objavil aj tu. Dôvodom bolo zahnutie papiera spôsobené väzbou knihy. Tento text sa teda po odfotení a zdigitalizovaní dokumentu javil mierne zahnutý, s čím mala automatická segmentácia miestami problém. Opäť sa ale jednalo skôr o ojedinelý problém.

Aj pri tejto vizitácii sme sa stretli s problémom pri segmentácii číslovania, pričom bolo opäť potrebné vyznačiť toto číslovanie manuálne. Tabuľky a zoznamy boli v tejto vizitácii podobné ako vo vizitácii Starých Hôr. Napísané boli voľne a nemali jednoznačne vyznačené bunky. Stretávame sa tu teda s podobnými problémami a to najmä s nepresným zaznamenaním čísiel a rozličných znakov v rámci tabuliek. Samotný text v tabuľkách bol ale väčšinou vyznačený presne. Nebolo teda potrebné ich segmentovať celé manuálne, ale stačila len korekcia istých častí.

Po segmentácii sme pristúpili k vytvoreniu Ground Truth vzorky, ktorú sme vytvárali prepisom istej časti našich dokumentov. Predstavíme teda postup a pravidlá, ktoré sme pri tomto prepise všeobecne naprieč všetkými našimi dokumentmi dodržiavali. Tvorcovia Transkribusu odporúčajú dva hlavné spôsoby prepisu. Prvým je „paleografická transkripcia“ a druhým „mierna normalizácia“.[131] Spôsob prepisu je potrebné zvoliť hneď na začiatku a následne je nevyhnutné dodržiavať daný postup pre celý text. Zmena v priebehu prepisu by mala negatívny dopad na presnosť modelu. Pri paleografickej transkripcii vykonávame

[131] How to Transcribe Documents with Transkribus – Introduction. Dostupné online: https://readcoop.eu/transkribus/howto/how-to-transcribe-documents-with-transkribus-introduction/

prepis najpresnejšie a najautentickejšie ako je to možné. Jedná sa najmä o grafémy, ktoré sa už aktívne nepoužívajú. Pri paleografickej transkripcii by sme takéto grafémy označili ich presným relevantným Unicode znakom. Na druhú stranu pri miernej normalizácii by sme využili namiesto týchto dobových grafém len ich základný znak. Ako príklad môžeme uviesť dvojité „i“ alebo „ij“, ktoré sa v dobových latinských textoch označuje často znakom „ÿ“. Pri paleografickej transkripcii by sme ho prepísali práve v tejto originálnej podobe. Pri miernej normalizácii by sme takýto znak prepísali jednoducho ako „y“, nakoľko sa jedná o základný znak. Tvorcovia Transkribusu odporúčajú užívateľom, ktorý sa zaujímajú nielen o správny prepis, ale aj o historický vzhľad a vývoj grafém využívať práve paleografickú transkripciu. Naopak ak ide užívateľom len o obyčajný prepis dokumentu, je mierna normalizácia vhodnejšia. My sme sa rozhodli pri našom prepise využiť postup miernej normalizácie. Dôvodom bolo najmä to, že sme pracovali so značne obmedzeným počtom strán. Z tohto dôvodu nám prišlo skôr kontraproduktívne učiť Transkribus čítať znaky, ktoré sa v texte objavujú len ojedinele. Jedinú výnimku predstavovalo už spomínané písmeno „ÿ“, ktoré sme prepisovali v originálnej podobe. Normalizáciu sme zvolili napríklad pri písmene „s“, ktoré sa v dobových latinských rukopisoch písalo dvoma spôsobmi. Prvým je klasické písané „s“, využívané aj dnes. Druhým spôsobom, akým sa toto písmeno písalo bolo tzv. „dlhé s“. Jednalo sa o variant minuskulného písmena „s“ (ſ). Pre zjednodušenie sme obe tieto grafémy prepisovali rovnakým spôsobom ako obyčajné „s“. V texte sme sa stretli aj s ligatúrou „et“, ktorú sme prepisovali v jej originálnej podobe, teda ako znak „ampersand“ (&). Rozlišovali sme ale medzi napísaným „et“ a ligatúrou „&“.

Pri prepise sme odstránili niekoľko znakov z tréningového procesu. Jednalo sa o špeciálne znaky, ktoré sa v texte objavovali veľmi ojedinele a ich podoba bola značne nekonzistentná. Nakoľko sme pracovali s malým počtom strán a slov, usúdili sme, že pridávať tieto znaky do tréningového procesu by mohlo mať pri tejto malej vzorke skôr negatívny vplyv. Boli to najmä znaky označujúce menu ako napríklad florény či denáre. Neprepisovali sme ani pomlčky a čiarky, ktoré slúžili na vypĺňanie miesta v tabuľkách. Za využitia týchto pravidiel a postupov sme teda vytvárali naše Ground Truth vzorky, ktoré sme použili pri tvorbe a trénovaní modelov.

4. Automatická transkripcia vybraných kanonických vizitácií

Cieľom našich pokusov je preskúmanie možností automatickej transkripcie pri využití kanonických vizitácií ako skúmaných dokumentov**.** Z pohľadu toho, ako sú napísané tvoria kanonické vizitácie pomerne špecifický prameň. Samotné knihy vizitačných protokolov sú pomerne obsiahle dokumenty, ktoré môžu obsahovať stovky strán. Problémom je, že sa skladajú z množstva rozličných vizitačných protokolov, ktoré sú väčšinou písané odlišnými pisárskymi rukami. Vzniká tu potenciál vytvorenia všeobecného modelu, ktorý by dokázal čítať celú knihu kanonických vizitácií. Takýto model by mohol byť prospešný najmä pre osobné využitie a výskum, kedy bádateľ nepotrebuje najvyššiu možnú presnosť prepisu a chce sa len dostať k informáciám obsiahnutých v texte.

Komplikácia vzniká v tom, že aj keď sú knihy vizitačných protokolov obsiahlymi dokumentami, časté zmeny rukopisov môžu predstavovať problém pri vytváraní modelu, ktorý by ich dokázal spoľahlivo prečítať. Tento fakt nám neumožňuje vytvoriť tzv. veľký model, ktorý by dokázal čítať celý dokument s vysokou presnosťou po natrénovaní jedného rukopisu. Jedným z cieľov tejto práce bude preskúmať možnosti vytvorenia tzv. malého modelu, založeného na základe menšej vzorky viacerých rukopisov. Prostredníctvom niekoľkých postupov sa zameriame na preskúmanie limitov pri vytváraní modelu založeného na niekoľkých mierne odlišných rukopisoch. Tiež zistíme, ktorý z využitých postupov je pri vytváraní modelu najefektívnejší a najvhodnejší na automatickú transkripciu už spomenutého typu dokumentov. Keďže vytvárame malý model, založený na menšom počte strán a viacerých rukopisoch, tak aj faktory, na ktoré sa pri jeho tvorbe zameriame budú mierne odlišné od veľkých modelov založených na jednom rukopise. Pri väčších modeloch je väčšinou v centre pozornosti dosiahnutie čo najnižšej chybovosti, pričom takéto modely môžu dosahovať hodnotu chybovosti zhruba 1 – 3%. V našom prípade sa na chybovosť modelu zameriame len čiastočne. Aj keď sa pokúsime dosiahnuť čo najvyššiu presnosť našich modelov, hlavným zameraním nebude len presnosť, ale aj reálne využitie modelu na rôzne typy podobných rukopisov a samotný potenciál automatickej transkripcie za využitia týchto modelov ako osobnej pomôcky pri výskume. To znamená, že sa pokúsime nájsť rovnováhu medzi čo najvyššou presnosťou modelu a jeho schopnosťou čítať viac rukopisov.

4.1 Vytváranie referenčných modelov

Cieľom nášho prvého pokusu bolo vytvorenie istej referenčnej hodnoty, ktorá bude slúžiť ako porovnávacie kritérium v ďalších pokusoch. Venovali sme sa samostatne každému z rukopisov, aby sme zistili možnosti základnej automatickej transkripcie vzhľadom na naše rukopisné dokumenty. Položíme tu teda základy nášho malého modelu a detailnejšie sa v tomto pokuse pozrieme na to, aké sú minimálne požiadavky na vytvorenie využiteľného modelu. V rámci tohto pokusu sme vytvorili tri nové modely, pričom pri každom z nich sme použili len jeden rukopis. Sledovali sme aký presný model dokážeme vytvoriť využitím jednotlivých rukopisov a obmedzeného počtu strán, ktoré máme k dispozícii. Informácie, ktoré týmto spôsobom získame budú kľúčové pre ďalší výskum a vytváranie modelu skladajúceho sa z viacerých rukopisov. Hodnotiť úspešnosť a využitie našich modelov budeme na základe ich chybovosti. Keďže v tomto prípade je hlavný dôraz kladený na zrozumiteľnosť výsledného transkribovaného textu, môžeme si dovoliť čiastočne prehliadnuť istý typ chýb, ktoré nemajú výrazný vplyv na celkovú zrozumiteľnosť textu. Ako príklad môžeme uviesť nesprávne rozoznanie majuskuly od minuskuly, alebo chyby v interpunkcii či medzerách slov. Presnosť modelu budeme sledovať na základe chybovosti prezentovanej hodnotou CER (character error rate). Táto hodnota predstavuje celkové percento nesprávne rozoznaných znakov v texte a je sledovaná v rámci cvičného aj overovacieho súboru. Kľúčová pre nás bude chybovosť práve v rámci overovacieho súboru, nakoľko tá predstavuje reálny pohľad na to, ako dokáže daný model rozpoznávať text, s ktorým sa ešte nestretol. Chybovosť cvičného súboru je teda menej relevantná pre zistenie skutočného využitia modelu. Na záver ešte dodáme, že všetky modely, ktoré budeme vytvárať v rámci tejto práce budú trénované v 50 cykloch/epochách.

Ako prvej sme sa venovali kanonickej vizitácii Špane Doliny, ktorá predstavovala rukopis A. Aby sme zistili možnosti a minimálne požiadavky na vytvorenie funkčného modelu, začali sme s trénovaním nového modelu na minimálnom počte strán. Prvý model sme teda založili len na piatich stranách, ktoré tvorili cvičný súbor. Tieto strany obsahovali celkovo 1214 slov. Model, ktorý sme týmto spôsobom vytvorili mal hodnotu chybovosti v overovacom súbore 13.21% CER, čo znamená, že až vyše 13% zo všetkých znakov Transkribus nedokázal správne rozoznať.[132] Jedná sa o značne vyššiu chybovosť, ktorá sa pohybuje na hranici, kedy je možné spochybňovať zmysluplnosť takéhoto prepisu. Pri

[132] Chybovosť v cvičnom súbore bola 0.21%

pohľade na typ chýb sme zistili, že chyby sa skutočne nachádzali vo väčšine slov a v mnohých prípadoch sa jednalo o chyby, ktoré boli príliš závažné a spôsobovali, že veľké časti textu boli nezrozumiteľné. Korekcia takéhoto veľkého počtu chýb by bola časovo veľmi náročná. Navyše spochybniť môžeme aj využiteľnosť takéhoto modelu ako osobnej pomôcky, nakoľko veľký počet nesprávne prepísaných písmen môže spôsobiť, že samotné vyhľadávanie v texte by nemuselo byť spoľahlivé.

Rozhodli sme sa teda do cvičného súboru pridať väčší počet strán a vytvoriť nový model. Do cvičného súboru sme teda pridali ďalšie dve strany, čím sa zväčšil už na 7 strán, ktoré celkovo obsahovali 1628 slov. Výsledkom bol model, ktorého chybovosť v overovacom súbore predstavovala hodnotu 9.76% CER.[133] Vidíme, že pridaním len dvoch strán, na ktorých sa nachádzalo 414 slov sa nám podarilo zlepšiť chybovosť modelu o 3.45%. Týmto modelom sme sa tiež dostali pod hranicu 10%, ktorá sa všeobecne považuje za hraničnú chybovosť, kedy je možné hovoriť o zmysluplnosti automatickej transkripcie.[134] Značná časť textu je v tomto prípade už prepísaná správne. Pri pohľade na typ chýb vidíme tiež isté zlepšenie. Naďalej sledujeme chyby, kedy sa v jednom slove nachádza viac nesprávnych písmen, no častejšie sa už jedná len o jedno písmeno, ktoré je nesprávne. Veľkú časť tvorili aj chyby spôsobené nesprávnym rozoznaním veľkých a malých písmen na začiatku slov a podobne aj v interpunkčných znamienkach, ktoré zatiaľ vo väčšine prípadov chýbali. Ako problematické sa javili aj čísla, ktoré predstavovali častý problém, nakoľko vyše polovica čísel v texte bola nesprávne prepísaná. Celkovo už ale môžeme hovoriť o využiteľnejšom modeli, nakoľko korekcia takéhoto textu by teoreticky nezabrala viac času ako manuálna transkripcia. Zredukovanie počtu chýb, ktoré sme v tomto modeli dosiahli by už teoreticky mohlo umožniť aj lepšie plnotextové vyhľadávanie, no ako spoľahlivé by sme ho ešte nenazvali. Naďalej sa pohybujeme na hranici funkčnosti a využiteľnosti takéhoto modelu.

Rozhodli sme sa teda cvičný súbor rozšíriť ešte raz, nakoľko pridanie spomenutých dvoch strán prinieslo značné zlepšenie. Tentokrát sme do cvičného súboru zahrnuli desať strán, ktoré celkovo obsahovali 2168 slov. Výsledkom tohto tréningu bol model s chybovosťou 8.75% CER v overovacom súbore.[135] Zlepšenie chybovosti, ktoré sme dosiahli pridaním ďalších 540 slov do cvičného súboru bolo 1.01%. Opäť teda vidíme zlepšenie, aj keď v tomto prípade už menšie. Nepresnosti v rámci tohto modelu sa opäť javia

[133] Chybovosť v cvičnom súbore bola 0.34%
[134] NAGY, Možnosti aplikácie metódy digitálnej transkripcie historických rukopisných textov ..., s. 59.
[135] Chybovosť v cvičnom súbore bola 0.50%

ako menej závažné oproti predchádzajúcemu. Chyby, ktoré robili text úplne nezrozumiteľný sa už pri tomto modeli prevažne nevyskytovali. Vo väčšine slov, v ktorých sa nachádzala chyba dochádzalo už len k výmene jedného písmena. Bežným problémom naďalej zostávalo rozoznávanie veľkých a malých písmen, rovnako ako aj interpunkcia. Celkovo sa už ale text dal považovať za zrozumiteľný a chyby, ktoré vznikli vyžadovali menšiu korekciu. Plnotextové vyhľadávanie by už pri takejto transkripcii bolo určite možné, aj keď jeho úplnú spoľahlivosť naďalej spochybňujeme. Celkovo tento model vieme označiť za pomerne funkčný a to aj napriek vyššej chybovosti. Vychádzame z toho, že veľká časť chýb, ktorá pri transkripcii vznikla nemala závažný dopad na zrozumiteľnosť textu. Rozhodli sme sa, že pridávať viac strán do tohto modelu už nebudeme, nakoľko pracujeme s obmedzeným počtom strán a dalo sa predpokladať, že potenciálne zlepšenie chybovosti by už bolo minimálne. Tento model bude teda predstavovať referenčnú hodnotu pre rukopis A, a tiež základný model, na ktorom budeme stavať náš ďalší výskum. Odkazovať na tento model budeme ako na „Referenčný model A“ (obr. 5).

Druhý z našich referenčných modelov sme vytvorili na základe rukopisu B, ktorý predstavuje kanonická vizitácia farnosti Staré Hory. Keďže sme pri vytváraní predchádzajúceho modelu, založeného na rukopise A dosiahli prijateľné výsledky využitím desiatich strán, rozhodli sme sa aj v tomto prípade využiť podobný počet strán. Do cvičného súboru sme teda vložili deväť strán z kanonickej vizitácie Starých Hôr. Celkovo sa jednalo o 2991 slov. Výsledkom trénovania bol model, ktorého chybovosť v overovacom súbore dosiahla hodnotu 2.42% CER.[136] Jedná sa teda o model so skutočne nízkou mierou chybovosti, čo sa dá považovať za vynikajúci výsledok. Treba poznamenať, že pri trénovaní tohto modelu sme využili o 823 slov viac ako pri našom najpresnejšom modeli z rukopisu A, ktorý dosiahol spomínanú hodnotu chybovosti 8.75%. Jedná sa teda o prekvapivo presný výsledok a plne využiteľný model, ktorý sa nám podarilo vytvoriť na základe rukopisu B a využitia len 9 strán, resp. 2991 slov (obr. 6). Presnosť tohto modelu sa nám potvrdila aj pri pohľade na overovací súbor. Väčšina textu v ňom bola prepísaná bezchybne. Chyby, ktoré sa v overovacom súbore nachádzali mali minimálny dopad na zrozumiteľnosť textu. Takmer všetky chyby, ktoré sa v overovacom súbore vyskytli, sa týkali interpunkcie (väčšinou chýbajúca bodka, alebo čiarka) a čísiel, s ktorými mal aj tento model problémy (dôvodom je zrejme nedostatočná vzorka čísel v cvičnom súbore). Chyby sa miestami vyskytli aj v nesprávnom prepísaní niektorých veľkých a malých písmen na začiatku slov. Stávalo sa

[136] Chybovosť v cvičnom súbore bola 0.46%

tak najmä pri písmenách „l“ a „s“. Dôvodom týchto chýb mohla byť aj naša nepresnosť pri vytváraní Ground Truth, pričom práve pri týchto písmenách sme mali miestami pri prepise problém. Celkovo sa ale jedná o definitívne použiteľný model. Jeho presnosť sa nachádza v rozmedzí, ktoré sa považuje za vynikajúci výsledok a zrozumiteľnosť textu je prakticky nepoškodená, len s občasnou výnimkou nesprávneho písmena. Môžeme predpokladať, že plnotextové vyhľadávanie v takomto texte bude veľmi presné (obr. 7). Získavame tak pomerne prekvapujúci výsledok, ktorý nám naznačuje, že aj pri využití menšieho počtu slov a znakov je možné vytvoriť pomerne presný a jednoznačne funkčný model, nakoľko ako už bolo spomenuté k tomuto výsledku nám stačilo využiť necelých 3000 slov. Určite treba zohľadniť aj fakt, že využitý rukopis vizitácie farnosti Starých Hôr bol prehľadný a skutočne konzistentný naprieč celým dokumentom, čo určite výrazne prispelo k presnosti modelu. Aby sme si overili presnosť tohto modelu, vykonali sme dodatočnú automatickú transkripciu doteraz nevyužitých strán kanonickej vizitácie Starých Hôr. Prepis ktorý sme dostali zodpovedal presnosti modelu. Aj keď týmto spôsobom nevieme získať percentuálnu chybovosť, dokážeme odhadnúť, že bola primeraná presnosti modelu. Jediné výraznejšie viditeľné chyby boli v interpunkcii, najčastejšie zámena bodky za čiarku. Nesprávne prepísané písmeno bolo skutočne ojedinelé, pričom sa objavilo približne raz za riadok. Potvrdila sa nám tak presnosť tohto modelu a tiež fakt, že ak je rukopisný dokument prehľadný a konzistentný, dá sa vytvoriť spoľahlivý model aj so značne menšou cvičnou vzorkou, ktorá v našom prípade predstavovala len 2991 slov. Odkazovať na tento model budeme ako na „Referenčný model B“.

Posledný referenčný model sme vytvorili na základe kanonickej vizitácie farnosti Motyčky, teda rukopisu C. Do cvičného súboru sme v tomto prípade vložili 12 strán. Celkovo tak mal cvičný súbor 3232 slov. Výsledkom tréningového procesu bol model s hodnotou chybovosti 3.94% CER v overovacom súbore.[137] Aj v tomto prípade sa teda dá hovoriť o modeli s nízkou mierou chybovosti (obr. 8). Vidíme tu ale mierne vyššiu chybovosť ako pri modeli vytvorenom na základe vizitácie Starých Hôr a to aj napriek tomu, že sme v rámci trénovania tohto modelu využili viac slov. Toto sa dá zrejme vysvetliť menšou konzistenciou rukopisu pri písaní niektorých znakov. Jednalo sa o písmená „s“, „l“ a najmä „t“, ktoré autor tohto rukopisu písal viacerými spôsobmi. Môžeme predpokladať, že pri využití väčšieho cvičného súboru a teda rozšírení Ground Truth vzorky, by sa tieto problémy zminimalizovali. Celkovo ale môžeme aj tento model považovať za použiteľný,

[137] Chybovosť v cvičnom súbore bola 0.50%

pričom v priemere sme zaznamenali menej ako dve chyby na riadok. Aj v tomto prípade sa jednalo o chyby, ktoré nemali závažnejší vplyv na zrozumiteľnosť textu a vyžadovali len jednoduchú korekciu. Medzi chybami sledujeme pomerne často chyby v interpunkcii a vo veľkých a malých písmenách na začiatku slov. Zo závažnejších chýb je to občasné zamenenie už spomínaných písmen. Napríklad *Cæsareo* Transkribus prepísal ako *Cælareo, Terrestris* ako *serrestris,* alebo *Presbyter* ako *Prelbyter*. No ako už bolo spomenuté, určite sa jedná o funkčný a využiteľný model s prijateľnou mierou chybovosti do 5% CER a chybami, ktoré majú vo veľkej väčšine prípadov minimálny dopad na zrozumiteľnosť textu. Vo zvyšku práce budeme na tento model odkazovať ako na „Referenčný model C".

Týmto výskumom sa nám teda podarilo zistiť, že v prípade konzistentného rukopisu je možné vytvoriť model s nízkou mierou chybovosti už pri využití len zhruba 3000 slov ako cvičného materiálu pre model. Tieto tri modely, ktoré sme vytvorili na základe jednotlivých rukopisov budú ďalej slúžiť ako referenčná hodnota pre náš výskum, zameraný na vytvorenie jednotného modelu skladajúceho sa zo všetkých troch rukopisov.

4.2 Vytváranie spoločného modelu na základe viacerých rukopisov

Cieľom nášho druhého pokusu bolo vytvoriť jednotný model, ktorý by dokázal s čo najväčšou presnosťou čítať všetky tri využité rukopisy. Zameriame sa na to, ako ovplyvní presnosť modelu fakt, že ho trénujeme na viacerých typoch rukopisu a tiež na to, ako sa tieto rozličné rukopisy navzájom ovplyvňujú v procese trénovania nového modelu. Tiež chceme zistiť, či je v prípade dokumentu s viacerými rukopismi efektívnejšie venovať sa im v rámci automatickej transkripcie samostatne, alebo spoločne. Výsledkom tohto pokusu by mal byť čo najpresnejší použiteľný model, ktorý by dokázal s dostatočnou presnosťou vykonať automatickú transkripciu všetkých troch rukopisov. Postup, ktorý sme pre tento pokus zvolili bude pozostávať z využitia našich referenčných modelov a vytvorenia dvoch dodatočných modelov. Prvý z nich bude založený na spojení dvoch rukopisov a to vizitácii Špania Dolina a Staré Hory. Druhý bude založený na spojení všetkých troch rukopisov. Tieto modely budeme vytvárať zakaždým úplne od začiatku. Pri ich vytváraní využijeme tie isté strany, ako sme využili pri vytváraní našich referenčných modelov v cvičnom aj overovacom súbore. To nám umožní porovnávať presnosť a chybovosť našich novovytvorených kombinovaných modelov, založených na viacerých rukopisoch a to nielen medzi sebou, ale aj priamo s referenčnými modelmi vytvorenými na základe jednotlivých rukopisov. To, že

pri trénovaní modelov použijeme tie isté overovacie strany nám umožní ich vzájomne porovnať. Zistíme tak, akým spôsobom ovplyvnilo začlenenie viacerých rukopisov celkovú presnosť našich modelov a ako bola ovplyvnená presnosť prepisu jedného rukopisu pridaním dodatočných rukopisov.

Prvý model, ktorý sme v rámci tohto pokusu vytvorili sa skladal z vizitácii farností Špania Dolina a Staré Hory. Do cvičného súboru sme teda vložili celkovo 19 strán, 10 z vizitácie Špania Dolina a 9 z vizitácie Staré Hory, ktoré obsahovali celkovo 5159 slov. Do overovacieho súboru sme vložili dve strany, jednu z vizitácie Špania Dolina a jednu z vizitácie Staré Hory. Po tréningovom procese sme dostali výsledný model, ktorého chybovosť bola 4.46% CER (obr. 9). [138] Keďže sme pri vytváraní tohto modelu využili tie isté strany ako pri našich referenčných modeloch, dokážeme ich presnosť vzájomne porovnať. Chybovosť Referenčného modelu A bola 8.75% a chybovosť Referenčného modelu B predstavovala hodnotu 2.42%. Priemerná chybovosť medzi týmito dvomi modelmi je 5.59%. Jedná sa teda o vyššiu mieru chybovosti, ako ktorú sme dosiahli keď sme tieto rukopisy trénovali spolu. Vytvorenie jednotného modelu pre tieto dva rukopisy teda dosiahlo o 1.13% lepšiu presnosť ako bol priemer referenčných modelov. Vidíme, že rôzne rukopisy sa počas trénovania modelu ovplyvňujú. Spoločným trénovaním dvoch rukopisov sme dosiahli lepší výsledok ako bol priemer ich samostatných modelov. To nám naznačuje, že pri práci s viacerými rukopismi je samostatné venovanie sa každému z rukopisov menej efektívne ako ich spoločné trénovanie. Teda aj keď sa jedná o mierne odlišné rukopisy, ich vzájomná podobnosť istým spôsobom napomáha Transkribusu pri učení sa ich príslušných znakov. Samozrejme toto platí len v prípade, že rukopisy sú aspoň čiastočne podobné. Na tento model skladajúci sa z dvoch rukopisov budeme odkazovať ako na „Kombinovaný model AB“. Pozrieme sa teda na to, akým spôsobom ovplyvnilo spoločné trénovanie dvoch rukopisov ich príslušnú chybovosť. Vykonáme to porovnaním overovacích strán novovzniknutého kombinovaného modelu s overovacími stranami ich samostatných referenčných modelov. Ako prvú porovnáme overovaciu stranu z rukopisu A. Hneď na prvý pohľad vidíme zlepšenie oproti referenčnému modelu. Celkovo sme v rámci Referenčného modelu A videli na overovacej strane chybu v 65 slovách. V našom Kombinovanom modeli AB sme na tej istej strane našli chybu v 50 slovách. Samozrejme tento údaj úplne nevypovedá o skutočnom zlepšení presnosti prepisu, nakoľko nehovorí o celkovom počte chýb alebo ich type. Relevantnejšie je v tomto smere zamerať sa na periodickosť chýb. Tu

[138] Chybovosť v cvičnom súbore bola 1.09%

vidíme jednoznačné zredukovanie chýb oproti referenčnému modelu (obr. 10). Aj keď typ chýb zostal väčšinou rovnaký ako tomu bolo pri referenčnom modeli (veľké a malé písmená na začiatku slov a interpunkcia zostávala naďalej problémom), počet a periodickosť závažných chýb bola jednoznačne nižšia. Spoločné trénovanie rukopisu A a rukopisu B malo teda jednoznačne pozitívny dopad na čítanie a presnosť prepisu rukopisu A.

Pozornosť teraz zameriame na druhý rukopis v rámci tohto modelu, ktorý predstavuje kanonická vizitácia Starých Hôr. Referenčný model B, založený na rukopise vizitácie Starých Hôr dosiahol hodnotu chybovosti 2.42%, čo hovorí, že už samotný referenčný model bol značne presný a tak tu veľa miesta na potenciálne zlepšenie nezostávalo. Na prvý pohľad je počet chýb na overovacej strane rukopisu B podobný overovacej strane referenčného modelu. Sledujeme, že niektoré slová, ktoré boli v rámci referenčného modelu nesprávne prepísané, boli v tomto prípade už prepísané správne. Tiež sa ale objavili chyby v slovách, ktoré sa pri referenčnom modeli nenachádzali. V prípade rukopisu B tak môžeme konštatovať, že pridanie ďalšieho rukopisu nemalo na presnosť jeho rozoznávania väčší dopad. Dôvodom môže byť fakt, že samotný Referenčný model B dosiahol takú nízku chybovosť, že pridaním ďalšieho rukopisu už prakticky nič nezískal. Chyby sa objavovali prevažne v tých istých slovách a problém s rozoznávaním veľkých a malých písmen na začiatku slov a tiež interpunkcie naďalej pretrvával. Isté zlepšenie ale sledujeme v prepisovaní čísel, ktorých presnosť bola oproti referenčnému modelu jednoznačne vyššia.

Celkovo môžeme zhodnotiť, že trénovanie jedného modelu skladajúceho sa z dvoch rukopisov malo pozitívny dopad na celkovú presnosť oproti samostatnému trénovaniu. Výraznejšie zlepšenie sledujeme najmä pri rukopise A, ktorého chybovosť sa viditeľne znížila. Prepis rukopisu B, ktorého chybovosť bola nízka už pri našom referenčnom modeli zostal približne rovnako presný. Tento model tak môžeme jednoznačne považovať za funkčný model, s nízkou mierou chybovosti. V rámci overovacieho súboru, ktorý sa skladal z jednej strany rukopisu A, a jednej strany rukopisu B, sme dosiahli hodnotu chybovosti 4.46% CER. Model dokáže s dostatočne vysokou presnosťou čítať obidva rukopisy a môžeme konštatovať, že ich spoločné trénovanie malo pozitívny dopad na celkovú presnosť tohto modelu.

S týmito poznatkami sa teda môžeme posunúť k druhej časti tohto pokusu a skompletizovaniu nášho cieľa. Vytvoríme druhý kombinovaný model, ktorý sa bude skladať zo všetkých troch rukopisov, ktoré máme k dispozícii. Model opäť vytvoríme od začiatku, s využitím tých istých strán, ako boli využité v referenčných modeloch. Do

cvičného súboru sme vložili celkovo 31 strán, z toho 10 z vizitácie Špania Dolina, 9 z vizitácie Staré Hory a 12 z vizitácie Motyčky. Celkový počet slov tak dosiahol hodnotu 8391. V overovacom súbore sa nachádzali tri strany, jedna z každého rukopisu. Výsledkom bol model, ktorého chybovosť predstavovala hodnotu len 3.70% CER v overovacom súbore.[139] Tento model budeme označovať ako „Kombinovaný model ABC“ (obr. 11). Opäť sledujeme zlepšenie presnosti oproti našim trom referenčným modelom, ktorých priemerná chybovosť predstavovala hodnotu 5.04%. Náš Kombinovaný model ABC, v ktorom sme trénovali všetky tri rukopisy spoločne, dosiahol o 1.34% lepší výsledok ako keď sme trénovali rukopisy samostatne. Zlepšenie sledujeme aj oproti predchádzajúcemu modelu, skladajúcemu sa z dvoch rukopisov. V tomto smere sme pridaním vzorky tretieho rukopisu dosiahli celkové zníženie chybovosti modelu o 0.76%. Opäť sa nám teda potvrdzuje tvrdenie, že v prípade práce s viacerými podobnými rukopismi je výhodnejšie trénovať ich spoločne v rámci jedného modelu, ako trénovať ich samostatne. Dôvodom je, že ak sú rukopisy dostatočne podobné, dokážu počas trénovania ovplyvňovať svoju vzájomnú presnosť.

Opäť sa pozrieme na to, akým spôsobom bola ovplyvnená chybovosť jednotlivých rukopisov v rámci tohto modelu pridaním rukopisu C. Pri pohľade na rukopis A vidíme len minimálne zmeny oproti jeho stavu v rámci Kombinovaného modelu AB. Sledujeme, že niektoré chyby zmizli, no tiež sa tu objavilo niekoľko nových chýb. Môžeme ale konštatovať, že celkový počet chýb sa znížil najmä v slovách, ktoré obsahovali viac ako jednu chybu. Celkovo sa javí, že pridanie tretieho rukopisu neprinieslo väčšie zmeny presnosti v rámci rukopisu A. Dôvodom môže byť fakt, že zatiaľ čo rukopis A a B sú pomerne podobné, rukopis C je mierne odlišný a tak nevzniká výraznejšie prelínanie, ktoré by zlepšilo presnosť čítania rukopisu A.

Podobný výsledok vidíme aj v prípade rukopisu B. Pridanie rukopisu C malo minimálny dopad na presnosť rukopisu B. Všimli sme si, že niektoré chyby sa už neobjavili, no pribudlo aj niekoľko nových chýb, medzi ktorými môžeme spomenúť chyby v interpunkcii. Schopnosť modelu čítať rukopis B sa tak pridaním tretieho rukopisu výrazne nezlepšila, ale ani nezhoršila. Náš výsledný model dokáže naďalej čítať rukopis B s veľmi vysokou presnosťou, bez chýb, ktoré by ovplyvnili zrozumiteľnosť textu.

To ako bola ovplyvnená presnosť čítania rukopisu C, ktorý sme pridali do tohto modelu zistíme porovnaním overovacej strany tohto rukopisu s overovacou stranou

[139] Chybovosť v cvičnom súbore bola 1.71%

Referenčného modelu C, ktorý bol vytrénovaný čisto na rukopise C. V tomto smere vidíme jednoznačné zlepšenie v presnosti nášho modelu. Spoločné trénovanie tohto rukopisu spolu s rukopisom A a B, malo pozitívny dopad na presnosť jeho čítania a prepisu. Celkový počet chýb sa znížil a chyby, ktoré sme spomínali v rámci referenčného rukopisu C, týkajúce sa nesprávneho rozoznávania písmen „s“, „l“ a „t“ boli menej pravidelné. Môžeme tak konštatovať, že spoločné trénovanie troch rukopisov malo jednoznačne pozitívny vplyv na presnosť čítania rukopisu C.

Výsledkom nášho výskumu je malý model, ktorý sme natrénovali na základe troch rukopisov a dosiahol chybovosť len 3.70% CER. Jedná sa teda o model s nízkou mierou chybovosti. Vytrénovaný bol na základe 8391 slov, pochádzajúcich z troch rozličných rukopisov. Na otázku, či je možné vytvoriť funkčný model z menšej vzorky textu, ktorý by bol využiteľný na čítanie niekoľkých podobných rukopisov teda môžeme určite odpovedať kladne. Aj keď tvorcovia platformy Transkribus prevažne uvádzajú, že na vytváranie nových modelov odporúčajú minimálne 15 000 slov, podarilo sa nám vytvoriť pomerne uspokojivý model s približne polovicou týchto znakov, ktoré navyše pochádzajú z troch mierne odlišných rukopisov. Ako sa nám potvrdilo, tento model dokáže s dostatočnou presnosťou čítať všetky obsiahnuté rukopisy. Tiež môžeme skonštatovať, že spoločné trénovanie viacerých rukopisov v rámci jedného modelu malo pozitívny vplyv na celkovú presnosť modelu. Keďže bol typ písma v jednotlivých rukopisoch podobný, dokázali do istej miery vzájomne zlepšiť presnosť prepisu. Najmenej získal zo spoločného trénovania rukopis B, ktorého samostatný prepis bol už sám o sebe značne presný a tak výraznejšie neprofitoval z pridania ďalších rukopisov. Pri rukopisoch A a C ale vidíme jednoznačne pozitívny vplyv na ich presnosť po pridaní ďalších rukopisov do modelu. Tento postup tak môžeme jednoznačne označiť ako efektívnejší oproti samostatnému trénovaniu jednotlivých rukopisov, nakoľko výsledky, ktoré sme ním dosiahli boli viditeľne lepšie. Je tiež potrebné zdôrazniť, že väčšina chýb, ktorá pri využívaní tohto modelu vznikla nemala vážny dopad na zrozumiteľnosť textu. Mnohé z chýb boli späté s interpunkciou, pričom dochádzalo najmä k zámene bodiek a čiarok. Ďalšou bežnou chybou bolo zamieňanie veľkých a malých písmen na začiatku slov a komplikácie vznikali aj pri prepisoch čísel. Ako ale bolo spomenuté, tieto chyby neovplyvňujú samotnú zrozumiteľnosť textu. Môžeme teda konštatovať, že model, ktorý sme týmto postupom vytvorili sa určite dá považovať za funkčný a vhodný na ďalšie postupné rozširovanie a prácu s nim. V prípade postupu spoločného trénovania viacerých rukopisov a vytvorenia funkčného malého modelu sme dosiahli pozitívny výsledok.

4.3 Využitie base modelu pri vytváraní modelov

V rámci tohto testu využijeme alternatívnu metódu vytvárania modelov, ktorou je využitie tzv. „base modelu“, alebo základného modelu. Jedná sa o postup, kde namiesto toho, aby sme vytvárali model úplne od začiatku, využijeme nejaký existujúci model ako základ pre tento model a pridáme k nemu cvičný a overovací súbor našich dokumentov. Takýmto spôsobom prakticky prispôsobíme a adaptujeme už existujúci model na náš špecifický dokument/rukopis. Nakoľko base model už obsahuje isté natrénované informácie, znamená to, že by tým pádom mala stačiť aj menšia vzorka nášho vlastného rukopisu, aby sme docielili presnejší model. Samozrejme ako base model nemôžeme využiť akýkoľvek model. Je nevyhnutné aby bol rukopis, na ktorom bol daný model vytrénovaný aspoň približne podobný rukopisu dokumentov, s ktorými pracujeme. V tomto smere treba brať ohľad najmä na jazyk dokumentu, časové obdobie, kedy daný text vznikal a aj na samotný rukopis. Tiež treba počítať s tým, že platforma Transkribus ponúka len obmedzený počet verejne dostupných modelov. Ak teda bádateľ pracuje so špecifickejším rukopisom, respektíve jazykom, môže sa stať, že nenájde voľne dostupný model, ktorý by vedel využiť. Každopádne ale treba spomenúť, že najmä latinčina a nemčina sú v rámci verejne dostupných modelov dostatočne zastúpené.

Base model sa dá využiť v dvoch hlavných smeroch. Prvým je už spomínané využitie base modelu ako istého základu, ktorý prispôsobíme pre svoje vlastné dokumenty. Druhé využitie spočíva v rozširovaní našich už existujúcich modelov o ďalšie Ground Truth strany. Pri práci s rozsiahlejšími dokumentami je možné vytvoriť model na menšom počte strán, následne využiť tento model na automatickú transkripciu ďalšej časti strán, ktoré po náležitej korektúre týmto spôsobom pripojíme k nášmu už existujúcemu modelu, čím zvýšime jeho presnosť. To, že je využívanie base modelu pri rozširovaní už existujúceho modelu efektívnejšie ako vytváranie úplne nového modelu preukázal I. Nagy v rámci svojho experimentu.[140] V našom pokuse sa zameriame na prvú zo spomenutých možností, ktorou je prispôsobenie base modelu rukopisom, s ktorými pracujeme. Tento postup je odporúčaný práve pri vytváraní malých modelov, nakoľko pri práci s rozsahovo menšími dokumentami máme k dispozícii menej strán na vytvorenie modelu, čo môže mať negatívny dopad na jeho presnosť. Z tohto dôvodu sa javí využitie base modelu ako ideálna možnosť, nakoľko nám

[140] NAGY, Možnosti aplikácie metódy digitálnej transkripcie historických rukopisných textov ..., s. 62.

ponúkne už pomerne presný základ, ktorý len upravíme na náš vlastný rukopis. Teoreticky nám to tak umožní dosiahnuť čo najpresnejší výsledok s využitím čo najmenšieho počtu strán. V tomto smere teda vykonáme dva testy. V prvom zistíme akým spôsobom ovplyvní využitie base modelu presnosť nášho Kombinovaného modelu ABC, skladajúceho sa zo všetkých troch rukopisov. V druhom sa pozrieme ako ovplyvní využitie base modelu presnosť modelu skladajúceho sa len z jedného rukopisu. Na toto využijeme náš Referenčný model A, ktorý dosiahol najvyššiu chybovosť spomedzi referenčných modelov, a to 8.75% CER.

V obidvoch prípadoch využijeme rovnaký base model s názvom *NeoLatin_Ravenstein_1643-1772*. Tento model nám umožňuje nahliadnuť do cvičných a overovacích strán, na základe ktorých bol trénovaný. Zvolili sme ho práve z dôvodu, že dokumenty, na ktorých bol trénovaný pochádzajú približne z rovnakého obdobia ako naše dokumenty a samotný rukopis je podobný rukopisom, s ktorými pracujeme. Model je založený na prepise výročných listov *Litterae Annuae Parochiae Ravensteijn SJ ab Anno 1643 ad Annum 1772,* ktoré sa momentálne nachádzajú v Katolíckom dokumentačnom centre v Leuvene v Belgicku. Autormi tohto modelu sú T. Gribnau, P. Boer a L. Nellissen.[141] Model obsahuje 64 435 slov v 6 864 riadkoch (115 strán v cvičnom súbore a 8 strán v overovacom súbore). Chybovosť modelu predstavuje hodnotu 3.58% CER v overovacom súbore a 4.51% CER v cvičnom súbore.

Ako prvému sme sa venovali nášmu Kombinovanému modelu ABC, ktorý sa skladá zo všetkých troch rukopisov. Do cvičného a overovacieho súboru sme pridali tie isté strany ako pri jeho vytváraní. Jediným rozdielom bolo, že v tomto prípade sme použili spomínaný *NeoLatin Ravenstein 1643 – 1772* model ako base model. Následne sme spustili tréningový proces. Výstupom tohto procesu bol model, ktorého chybovosť v rámci overovacieho súboru dosiahla hodnotu len 3.22% CER (obr. 12). [142] Jedná sa o zlepšenie oproti nášmu pôvodnému modelu, ktorý bol vytváraný bez base modelu a dosiahol chybovosť 3.70% CER. Využitím base modelu sme teda dosiahli zlepšenie o 0.48%. Aj keď sa toto zlepšenie javí ako pomerne malé, je potrebné zdôrazniť, že náš model sa pohybuje na úrovni veľmi nízkej chybovosti, kedy je už pomerne náročné dosiahnuť akékoľvek výraznejšie zlepšenie presnosti. Práve z tohto dôvodu je zlepšenie o približne pol percenta o to významnejšie, nakoľko sme ho dosiahli jednoducho využitím base modelu, bez pridania dodatočných strán a prakticky bez

[141] Jedná sa o súčasť ich publikácie: GRIBNAU, T. et al.: *Martiaal en theatraal. De jezuïeten in Ravenstein (1643-1772).* Nijmgen : Uitgeverij Valkhof Pers, 2019. 312 s. ISBN 978 90 5625 514 5.
[142] Chybovosť v cvičnom súbore bola 1.71%

akejkoľvek ďalšej práce. Z tohto výsledku sa teda javí, že využitie base modelu pri vytváraní malých modelov je jednoznačne efektívne a výhodné. Samozrejme to platí len za predpokladu, že existuje model, ktorý je svojim rukopisom aspoň približne podobný rukopisom s ktorými pracujeme.

Robiť detailnejší rozbor chýb v tomto prípade už nemá väčší význam, nakoľko typ a charakter chýb zostal väčšinou rovnaký ako bol aj v rámci Kombinovaného modelu ABC. Zmena prichádza len v mierne nižšej periodicite chýb. Presnosť tohto modelu už skutočne dosahuje vysokú hodnotu a to najmä v rukopise B a C. Napríklad v rámci overovacej strany z rukopisu B už nachádzame len 11 chybne prepísaných písmen v rámci celej strany, ktorá obsahuje 1822 znakov. Podobne dobré výsledky sledujeme aj v rámci overovacej strany z rukopisu C, kde sa nachádza 15 chybne prepísaných písmen z 1560 znakov (obr. 13, 14). Zvyšné chyby sú rovnaké ako pri predchádzajúcich modeloch, teda nesprávna interpunkcia, alebo nesprávne veľké a malé písmená na začiatku niektorých slov. Výsledok tohto pokusu tak môžeme považovať za jednoznačne kladný.

Zatiaľ čo v predchádzajúcom pokuse sme sa zamerali na vplyv využitia base modelu na model skladajúci sa z viacerých rukopisov, v rámci tohto dodatočného pokusu budeme sledovať vplyv využitia base modelu na model založený na jednom rukopise. Ako base model využijeme opäť rovnaký voľne dostupný model *NeoLatin Ravenstein 1643 – 1772*. V rámci tohto pokusu využijeme strany z vizitácie Špranej Doliny, teda rukopisu A. Ten dosiahol v rámci našich referenčných modelov najvyššiu mieru chybovosti a tým pádom tu vidíme aj najväčší potenciál na zlepšenie aplikovaním base modelu. Vytvoríme teda nový model s využitím tých istých strán, ako sme využili pri vytváraní Referenčného modelu A. Teraz ale aplikujeme aj spomínaný base model a následne vzájomne porovnáme overovacie strany týchto dvoch modelov. Zistíme tak, aký vplyv malo aplikovanie base modelu.

Vytvorili sme teda nový model, do ktorého sme vložili rovnaké cvičné a overovacie strany, ako sme použili pri vytváraní Referenčného modelu A. Tentokrát sme ale aplikovali aj spomínaný base model *NeoLatin Ravenstein 1643 – 1772*. Výstupom bol model, ktorého chybovosť dosiahla hodnotu 5.50% CER v overovacom súbore.[143] Keď túto hodnotu porovnáme s chybovosťou nášho Referenčného modelu A, v ktorom sme nepoužili base model vidíme zlepšenie až o 3.25% CER. Aj v tomto prípade malo využitie base modelu viditeľne pozitívny dopad na výslednú presnosť modelu. Len využitím verejne dostupného

[143] Chybovosť v cvičnom súbore bola 0.40%

base modelu sa nám podarilo zlepšiť pôvodnú chybovosť 8.75% na 5.50% a to bez použitia dodatočných strán v rámci tréningového procesu.

Opäť sa nám teda potvrdzuje, že využitie base modelu v prípade vytvárania malého modelu na základe menšej vzorky textu sa javí ako veľmi vhodný postup. V oboch prípadoch sme dosiahli značné zlepšenie presnosti našich modelov a to prakticky bez dodatočnej práce, čo robí tento postup výrazne efektívnym.

Na záver ešte uvedieme jedno zistenie, ktoré sme dosiahli pri experimentovaní s využívaním base modelu. Okrem vyššie spomenutých testov sme vykonali ešte niekoľko dodatočných testov, kedy sme použili v pozícii base modelu niektoré z našich referenčných modelov. Využili sme napríklad náš Referenčný model B pri vytváraní dvoch nových modelov, založených na rukopise A a B. Vo všetkých prípadoch sme ale dosiahli minimálne až žiadne zlepšenie v presnosti nového modelu. Aj z tohto dôvodu sme sa rozhodli tento pokus neskúmať hlbšie. Na základe našich zistení môžeme usúdiť, že naše referenčné modely nie sú vhodné na využitie v pozícii base modelu pri trénovaní odlišných rukopisov. Aj keď je typ písma podobný, problém pravdepodobne vznikol vo veľkosti týchto modelov. Neponúkali zrejme dostatočný základ pre efektívne adaptovanie na trénovanie nového rukopisu. Z tohto dôvodu usudzujeme, že pri výbere base modelu je potrebné sledovať aj samotnú veľkosť modelov, nakoľko sa javí, že väčšie modely sú vhodnejšie na využitie v podobe base modelu.

4.4 Porovnanie chybovosti a efektivity využitých postupov

Všetky tri postupy, ktoré sme v rámci nášho výskumu preskúmali, sa v rozličných mierach javia ako použiteľné v praxi pri vytváraní modelov založených na viacerých rukopisoch. V našom prvom pokuse sme sa venovali osobitne každému z rukopisov. Podarilo sa nám vytvoriť niekoľko pomerne presných modelov, ktoré dosahovali hodnotu chybovosti medzi 8.75% až 2.42%. Prostredníctvom tohto pokusu sa nám podarilo zistiť, že je možné vytvoriť model s nízkou mierou chybovosti, aj keď využijeme len menšiu Ground Truth vzorku. Napríklad pri rukopise kanonickej vizitácie Starých Hôr sa nám podarilo vytvoriť model s chybovosťou 2.42% a to využitím len deviatich strán v cvičnom súbore. Môžeme teda zhodnotiť, že v prípade konzistentného rukopisu je skutočne možné vytvárať presné „malé modely“ s využitím približne 3000 slov, ako tomu bolo v prípade našich modelov. V tomto smere sa nám teda podarilo preskúmať akési minimálne limity na

vytvorenie použiteľného modelu. Všetky tri modely, ktoré sme vytvorili ako naše referenčné modely pre náš ďalší výskum môžeme považovať za použiteľné. Hlavný význam takýchto modelov vidíme najmä v podobe osobnej pomôcky pri výskume a práci so staršími rukopisnými dokumentami. Ich presnosť umožní vyhľadávať relevantné informácie v texte s využitím plnotextového vyhľadávania. Typy chýb, ktoré sa v prepisoch vyhotovených týmito modelmi nachádzali navyše nemali vo väčšine prípadov výrazný dopad na zrozumiteľnosť textu. Bádateľ, ktorý má skúsenosť s prácou s daným typom dokumentu a ovláda aspoň základ jazyka, v ktorom je dokument napísaný, by teda nemal mať problém pochopiť podstatu obsahu výsledného textu po automatickej transkripcii. Hlavným cieľom nášho výskumu ale bolo zistiť, ako pristupovať k automatickej transkripcii dokumentov skladajúcich sa z viacerých rukopisov. Po vykonaní našich pokusov sa javí venovanie sa každému z rukopisov samostatne ako jednoznačne najmenej efektívna alternatíva. Pri porovnaní tohto postupu s nasledujúcimi dvoma sa dá výrazne spochybniť jeho zmysluplnosť.

V rámci nášho druhého pokusu sme preskúmali možnosti vytvorenia jedného modelu, založeného na viacerých rukopisoch a dopad, ktorý malo pridávanie viacerých rukopisov na ich vzájomnú chybovosť. Spoločné trénovanie viacerých rukopisov v rámci jedného modelu sa javí ako jednoznačne lepšia alternatíva oproti predošlému postupu. Presnosť modelov, ktoré sme týmto spôsobom vytvorili bola lepšia, ako keď sme sa im venovali samostatne. Dôvodom bolo, že aj keď boli rukopisy mierne odlišné, niektoré ich aspekty boli dostatočne podobné, aby sa počas tréningového procesu ovplyvňovali a zlepšovali tak navzájom celkovú presnosť. Dôležitým predpokladom efektívnosti tohto postupu je teda značná podobnosť jednotlivých rukopisov, ktoré v modeli spoločne trénujeme. Rukopisy, s ktorými pracujeme túto podmienku spĺňali a odrazilo sa to aj na presnosti modelov. Pridávaním ďalších rukopisov sa nám podarilo nielen zlepšiť celkovú presnosť nášho modelu, ale do istej miery aj presnosť v rámci jednotlivých rukopisov. Výsledný model, ktorý sme týmto postupom vytvorili dosiahol hodnotu chybovosti len 3.70%. Model dosahoval lepšie výsledky ako naše referenčné modely a dokázal s dostatočnou presnosťou čítať všetky tri rukopisy, na ktorých sme ho trénovali. Model, ktorý sme týmto spôsobom vytvorili je naďalej možné považovať za malý model, keďže sme v rámci tréningového procesu využili len 8391 slov. Jedná sa teda o veľmi dobrý výsledok, najmä keď vezmeme do úvahy, že tento model dosiahol chybovosť len 3.70% a dokázal čítať tri rukopisy. V prípade rozsiahlejších dokumentov, v ktorých dochádza k občasných zmenám pisárskej ruky sa tak javí tento postup ako efektívny a určite použiteľný pri práci

s dobovými rukopisnými dokumentami. Navyše tu vidíme aj jednoznačný potenciál na postupné rozširovanie takéhoto modelu. Tento model by mohol predstavovať vhodný základ, ktorý by bolo dodatočnou prácou možné transformovať na veľký model, ktorý by dokázal čítať celý rad rukopisov z daného obdobia.

V rámci nášho tretieho postupu sme sa zamerali na možnosti využitia base modelu pri trénovaní jedného rukopisu, ako aj kombinácie viacerých rukopisov. Využitím base modelu sa nám podarilo vytvoriť model s chybovosťou len 3.22%. Tento model bol rovnako ako predchádzajúci model založený na tých istých troch rukopisoch, pričom bez využitia base modelu bola chybovosť modelu vyššia a to už spomínaných 3.70%. V rámci druhého pokusu sme aplikovali base model na trénovanie modelu založeného len na jednom rukopise. V tomto prípade sme získali model s chybovosťou 5.50%, oproti originálnemu modelu, ktorého chybovosť bola až 8.75%. V obidvoch prípadoch sa nám teda potvrdilo, že aplikovanie base modelu do tréningového procesu malo viditeľne pozitívny dopad na presnosť trénovaného modelu. Tento postup sa javí ako najefektívnejší a celkovo sme s nim dosiahli najlepšie výsledky. Využitie base modelu totiž nevyžadovalo žiadnu nadbytočnú prácu a v obidvoch prípadoch prinieslo jednoznačne pozitívny vplyv na presnosť našich trénovaných modelov.

Tento postup ale má aj svoje negatíva. Zrejme tým najvýraznejším je obmedzená dostupnosť verejných modelov. Ako už bolo spomenuté, pri tomto postupe využívame už existujúci model, ktorý adaptujeme na svoje dokumenty. V tomto smere sme teda závislí od toho, či nájdeme vhodný model, ktorý je dostatočne podobný naším dokumentom a spĺňa všetky podmienky. Problém môže vzniknúť najmä ak pracujeme s neobvyklým jazykom dokumentu, alebo špecifickým rukopisom. V takýchto prípadoch nemusíme nájsť vhodný model a tak je samotná efektivita tohto postupu obmedzená.

Druhým negatívom môže byť fakt, že nie všetky verejne dostupné modely umožňujú nahliadnutie do cvičného a overovacieho súboru, na ktorom boli trénované. V takomto prípade nevieme úplne presne zistiť o aký rukopis sa jedná a musíme sa spoliehať len na opis poskytnutý autorom. Tiež nám to neumožňuje získať informáciu na základe koľkých strán a akého druhu dokumentov bol náš nový model trénovaný. Toto samozrejme predstavuje negatívum len v prípade, že sú pre nás tieto informácie relevantné. Jedinú informáciu, ktorú v prípade takýchto modelov získavame je okrem ich percentuálnej presnosti ešte celkový počet slov, na základe ktorých bol daný model trénovaný.

Aj napriek spomenutým negatívam sa na základe nášho výskumu javí tento postup ako najefektívnejší pri práci s viacerými rukopismi. V prípade, že nenájdeme vhodný model,

ktorý by sme mohli využiť ako base model, podobne vhodný sa javí aj postup spoločného trénovania viacerých rukopisov v rámci jedného kombinovaného modelu. Pri pohľade na percentuálnu chybovosť sme dosiahli pri oboch týchto postupoch pozitívne výsledky a to 3.70% pri spoločnom trénovaní troch rukopisov bez využitia base modelu a 3.22% s využitím base modelu.

Záver

V diplomovej práci sme sa zaoberali preskúmaním možností metódy automatickej transkripcie historických rukopisných textov na príklade vybraných kanonických vizitácií. V rámci tejto práce sme sa snažili zistiť, aké sú možnosti automatickej transkripcie historických rukopisných textov, skladajúcich sa z viacerých rukopisov. Tiež sme sa zamerali na odhalenie minimálnych požiadaviek na vytvorenie funkčného malého modelu, skladajúceho sa z jedného, ale aj viacerých rukopisov.

Pri automatickej segmentácii nášho textu sme zistili, že problémy vznikali najmä v dvoch smeroch. Prvým bolo nesprávne vyznačenie koncov riadkov, ktoré bolo spôsobené čiarami vyznačujúcimi okraje textu. Druhým bolo nesprávne segmentovanie tabuliek, ktoré nemali vyznačené bunky a nachádzali sa voľne v texte. Tu bolo dôvodom nepresnosti väčšinou to, že Transkribus mal problém vyznačiť, kde presne sa nachádzal text a čísla. Dôvodom boli najmä bodky, čiarky a rozličné znaky, ktoré slúžili na vyznačenie prázdneho miesta. Segmentovať tieto časti manuálne sa osvedčilo ako najlepšia alternatíva.

V rámci nášho prvého pokusu sme sa zamerali na zistenie limitov vytvárania malých modelov, založených na jednom rukopise. V tomto smere odporúčajú tvorcovia Transkribusu využiť pri tvorbe nového modelu okolo 15 000 slov. Postupným pridávaním strán do nášho modelu sme zistili, že na vytvorenie funkčného modelu sme potrebovali vložiť do cvičného súboru len približne 3000 slov. Overili sme teda, že efektívny model sa dá vytvoriť aj na relatívne menšej vzorke znakov. Presnosť samotného modelu sa výrazne odvíjala aj od konzistentnosti rukopisu, pričom sa nám podarilo vytvoriť modely s chybovosťou medzi 8.75% až 2.42%.

V druhom pokuse sme zisťovali možnosti automatickej transkripcie dokumentov, skladajúcich sa z rôznych rukopisov. Vytvorili sme dva modely, jeden založený na dvoch a druhý na troch rukopisoch. Snažili sme sa zistiť, ako vplývajú rôzne rukopisy na celkovú presnosť trénovaného modelu a ako sa navzájom počas tohto procesu ovplyvňujú. Podarilo sa nám zistiť, že spoločné trénovanie viacerých rukopisov v rámci jedného modelu malo pozitívny vplyv na celkovú presnosť tohto modelu. Zistili sme to porovnaním priemernej chybovosti našich referenčných modelov s modelmi, ktoré sme vytvorili kombináciou viacerých rukopisov (zlepšenie z 5.59% na 4.46% a zlepšenie z 5.04% na 3.70%). Počas tréningu modelu, v ktorom sme použili dostatočne podobné rukopisy sa nám potvrdilo, že sa navzájom ovplyvňujú a zlepšujú tak celkovú presnosť modelu. Okrem celkovej presnosti

sme týmto postupom dosiahli lepšie výsledky aj v rámci presnosti jednotlivých rukopisov, oproti hodnote, ktorú dosiahli počas samostatného tréningu. Chyby, ktoré sme v texte transkribovanom pomocou týchto modelov našli, nemali výrazný vplyv na zrozumiteľnosť textu. Veľká časť chýb spočívala najmä v nesprávnom rozoznaní veľkých a malých písmen na začiatku slov a nesprávnej transkripcii čísiel. V konečnom dôsledku boli obidva modely, ktoré sme vytvorili spoločným trénovaním viacerých rukopisov funkčné a efektívne, pričom dokázali s dostatočnou presnosťou čítať všetky obsiahnuté rukopisy.

V rámci tretieho pokusu sme preskúmali možnosti využitia base modelu pri trénovaní nových modelov skladajúcich sa z jedného, ale aj viacerých rukopisov. Využitie base modelu malo pozitívny vplyv na celkovú presnosť modelu v obidvoch prípadoch. Pri aplikovaní base modelu na jeden rukopis sme dosiahli zlepšenie chybovosti z 8.75% na 5.55%. Pri aplikovaní na model skladajúci sa z viacerých rukopisov sme dosiahli zlepšenie z 3.70% na 3.22%. Využitie base modelu pri vytváraní nových modelov sa nám teda ukázalo ako prínosné a efektívne. Objavili sme ale aj limitácie využitia base modelu, medzi ktorými môžeme spomenúť napríklad fakt, že rukopis tohto modelu musí byť dostatočne podobný rukopisu využívaných dokumentov. Potenciálny base model musí byť tiež dostatočne rozsiahly, aby bolo jeho využitie efektívne.

Pri automatickej transkripcii dokumentov skladajúcich sa z viacerých podobných rukopisov sa nám teda potvrdilo, že spoločné trénovanie týchto rukopisov je efektívnejším postupom oproti samostatnému trénovaniu jednotlivých rukopisov. Spoločným trénovaním sme dosiahli nielen lepšiu celkovú presnosť, ale zlepšila sa aj presnosť v rámci jednotlivých rukopisov. Tiež sa potvrdilo, že v prípade ak máme k dispozícii vhodný model, ktorý sa dá využiť v pozícii base modelu, je jeho začlenenie do postupu automatickej transkripcie jednoznačne výhodné. Využitím base modelu sme totiž dosiahli výrazné zlepšenie presnosti našich modelov s minimálnym úsilím navyše. Na základe nášho výskumu sa javí tento postup ako najefektívnejší. Modely, ktoré sme ním vytvorili sú vhodné na prácu s rukopisným textom v podobe osobnej pomôcky, ktorá umožní lepší prehľad a zefektívni prácu s daným textom. Takto vytvorené modely majú tiež potenciál ďalšieho rozširovania a transformovania na tzv. veľké modely, ktoré dokážu čítať a transkribovať celé rady rozličných rukopisov.

Prínos tejto práce predstavuje najmä preskúmanie rôznych postupov a prehĺbenie všeobecných znalostí spojených s metódou automatickej transkripcie rukopisných historických textov a programom Transkribus. Nakoľko je využívanie tejto metódy u nás zatiaľ len v počiatkoch, dajú sa akékoľvek nové poznatky a prístupy považovať za prínosné.

Hlavný prínos vidíme najmä v predstavení rôznych postupov a možností automatickej transkripcie, špecificky vo vzťahu ku transkripcii dokumentov skladajúcich sa z viacerých rukopisov, ktoré sú v našom prípade reprezentované kanonickými vizitáciami. Za rovnako podstatné považujeme aj celkové preskúmanie a vyhodnotenie chybovosti a typu chýb, ktoré pri tvorbe našich modelov vznikali, ako aj celkové zhrnutie efektívnosti a praktického využitia tzv. malých modelov, skladajúcich sa z jedného alebo viacerých rukopisov. V neposlednom rade by táto práca mala predstavovať istý prínos aj v spojení s už spomínaným projektom SKRIPTOR, ktorý sa zameriava na implementáciu najnovších poznatkov z výskumu automatického textového rozpoznávania historických dokumentov na Slovensku.

Predpoklady ďalšieho možného výskumu v rámci tejto témy vidíme najmä v zameraní sa na väčšiu vzorku rukopisov. Obzvlášť podnetný môže byť v tomto smere výskum zameraný na možnosti rozširovania malých modelov na veľké. Potenciálne pokračovanie tohto výskumu vidíme aj v preskúmaní možností vytvárania spoločných modelov, založených na rozličných typoch písma, ako napríklad latinka a grécka abeceda (alfabeta). V neposlednom rade vidíme ako jeden z hlavných smerov tohto výskumu snahu o vytváranie skutočne rozsiahlych modelov, ktoré by dokázali čítať a transkribovať historické rukopisy v rámci istého obdobia a jazyka samostatne, bez akéhokoľvek dodatočného trénovania.

Zoznam bibliografických odkazov

Použité pramene:

Archív Rímskokatolíckeho biskupstva a diecézy v Banskej Bystrici so sídlom v Badíne (ďalej DA BB), fond kanonické vizitácie, CV č. 10.

1. Archív DA BB, Visitatio Canonica Ecclesiae Parochialis Moticsko, 1783, CV č. 10
2. Archív DA BB, Visitatio Canonica Ecclesiae Parochialis Vallis Dominorum, 1783, CV č. 10
3. Archív DA BB, Visitatio Canonica Ecclesiae Parochialis Vetero Montanae, 1783, CV č. 10

Použitá literatúra:

1. BIZOŇOVÁ, M.: Kanonické vizitácie na Spiši v 17. a prvej polovici 18. storočia ako historický prameň. In: *6. študentská vedecká konferencia : zborník príspevkov*. Prešov, 2010, s. 459 – 469.
2. BORZA, P.: Kanonické vizitácie mukačevského biskupa Michala Manuela Olšavského. In: *Kanonické vizitácie po tridentskom koncile*. Prešov, 2007, s. 111 – 121.
3. BOŠANSKÝ, M.: *Kánonické vizitácie Nitrianskej diecézy v rokoch 1674 – 1831*. Dizertačná práca. Katedra cirkevných dejín, RKCMBF UK, Bratislava, 2017, 199s.
4. ČECHOVÁ, F.: Význam a záväznosť kanonických vizitácií v dejinách nitrianskej diecézy v 20. storočí. In: *Kanonické vizitácie po tridentskom koncile*. Prešov, 2007, s. 149 – 167.
5. FABOVÁ, B.: Kanonické vizitácie v Abovsko-turnianskej župe v 18. storočí. In: *Človek a spoločnosť*, roč. 14, 2011, č. 2, s. 21 – 29.
6. FIALOVÁ, I.: Kanonické vizitácie ako zdroj poznania života farností Šaštínskeho arcidiakonátu v prvej polovici 17. storočia. In: *Kanonické vizitácie po tridentskom koncile*. Prešov, 2007, s. 85 – 93.
7. GRIBNAU, T. et al.: *Martiaal en theatraal. De jezuïeten in Ravenstein (1643-1772)*. Nijmgen : Uitgeverij Valkhof Pers, 2019. 312 s. ISBN 978 90 5625 514 5.

8. HIŠEM, C. – FEDORČÁK, P. (eds.): *Kanonické vizitácie po Tridentskom koncile*. 1. vyd. Prešov : Vydavateľstvo Michala Vaška, 2007, 231 s. ISBN 978-80-7165-639-5
9. HOFERKA, M.: Kánonická vizitácia farností Záhorského dekanátu Bratislavského prepošstva z roku 1673. In: *Studia Historica Tyrnaviensia XVIII, Cirkevné dejiny – súčasť národných dejín*. Trnava, 2015, s. 101 – 123.
10. CHLADNÁ, Z.: Vizitácia Špranej Doliny a Starých Hôr v roku 1713. In: *Bystrický permon*, IV, 2006, č. 2, s. 5.
11. JURKO, J.: Kanonická vizitácia košického biskupa Andreja Szaboa v Bardejove v roku 1813. In: *Kanonické vizitácie po tridentskom koncile*. Prešov, 2007, s. 219 – 230.
12. KATUŠČÁK, D.: *Digitalizacia. Úvod do digitalizácie písomného dedičstva*: *Distanční študijný text*. Opava, 2019, 79 s. Dostupné na internete: https://is.slu.cz/el/fpf/zima2020/UBKVKKNP03/Digitalizace_a_dlouhodoba_ochrana_fondu_1.pdf?lang=en
13. KOWALSKÁ, E.: Ozbrojená moc v procese rekatolizácie. In: *Vojenská história*, 9, 2005, č. 2, s. 45 – 54.
14. KURUC, Š.: Vizitácia spišského prepošta Jána Peltza z roku 1731. In: *Kanonické vizitácie po Tridentskom koncile*. Prešov, 2007, s. 94 – 102.
15. KYSEĽOVÁ, M.: Kanonické vizitácie na Spiši. In: *Slovenská archivistika*, roč. 18, 1983, č. 2, s. 110 – 128.
16. LOPATKOVÁ, Z.: Cirkevné dejiny raného novoveku z pohľadu kanonických vizitácií. In: *Pontes ad fontes : církevní dějiny ve světle pomocných věd historických a příbuzných oborů*. Hradec Králové, 2011, s. 174 – 187.
17. LOPATKOVÁ, Z.: Kanonická vizitácia – k jej interpretácií, limitom a metodologickým otázkam v historickom výskume. In: *Prameň – jeho funkcia, význam, interpretácia a limity v historickom výskume*. Nitra, 2010, s. 69 – 77.
18. LOPATKOVÁ, Z.: Kanonické vizitácie v slovenských dejinách. In: *Studia historica Tyrnaviensia. VII : Pramene k slovenským dejinám*. Trnava, 2008, s. 96 – 106.
19. MAREK, M.: *Glossarium mediae et infimae latinitatis slovaciae. Vol. I (A-K)*. Trnava : Filozofická fakulta Trnavskej univerzity v Trnave, 2012. 474 s.
20. MAREK, M.: *Glossarium mediae et infimae latinitatis slovaciae. Vol. II (L-Z)*. Trnava : Filozofická fakulta Trnavskej univerzity v Trnave, 2012. 440 s.

21. MIČKOVÁ, Z.: *Banícka osada Špania Dolina od neskorého stredoveku k ranému novoveku*. 1. vyd. Banská Bystrica : Právnická fakulta UMB, 2013. 129 s. ISBN 9788055705248
22. NAGY, I.: Možnosti aplikácie metódy digitálnej transkripcie historických rukopisných textov pri sprístupňovaní archívnych fondov. In: *Slovenská archivistika,* roč. 51, 2021, č. 2, s. 53 – 67.
23. NEMEŠOVÁ, A.: Vývoj ľudového školstva na Spiši vo svetle kanonických vizitácií. In: *Kanonické vizitácie po Tridentskom koncile*. Prešov, 2007, s. 103 – 110.
24. NOVÁK, J. et al.: *Paleografická čítanka. Pomôcka na čítanie rukopisov z 12.-19. storočia*. Martin : Matica slovenská, 2001. 175 s. ISBN 80-7090-599-9
25. OSTERTAGOVÁ, A.: *Historicko-poznávacia hodnota kánopnických vizitácií*. Dizertačná práca. Katedra slovenských dejín, Filozofická fakulta Univerzity Komenského, Bratislava, 2009.
26. OSTERTAGOVÁ, A.: Kanonické vizitácie ako historický prameň. In: *Kanonické vizitácie po tridentskom koncile*. Prešov, 2007, s. 5 – 10.
27. SZEGHY, A.: Reálie južného Zemplína vo vizitáciách seniora evanjelickej reformovanej cirkvi Istvána Csulyáka Miskolciho. In: *Kanonické vizitácie po tridentskom koncile*. Prešov, 2007, s. 68 – 84.
28. TOMEČEK, O.: *Drevorubači a uhliari v lesoch Banskej Bystrice*. 1. vyd. Banská Bystrica : Fakulta humanitných vied UMB, 2010. 216 s. ISBN 978-80-557-0094-6
29. TOMEČEK, O.: Náboženské zápasy v zázemí mesta Banská Bystrica po nástupe reformácie. In: *Acta Facultatis Philosophicae Universitatis Ostraviensis. Historica*, 13, Ostrava, 2006, s. 57 – 67.

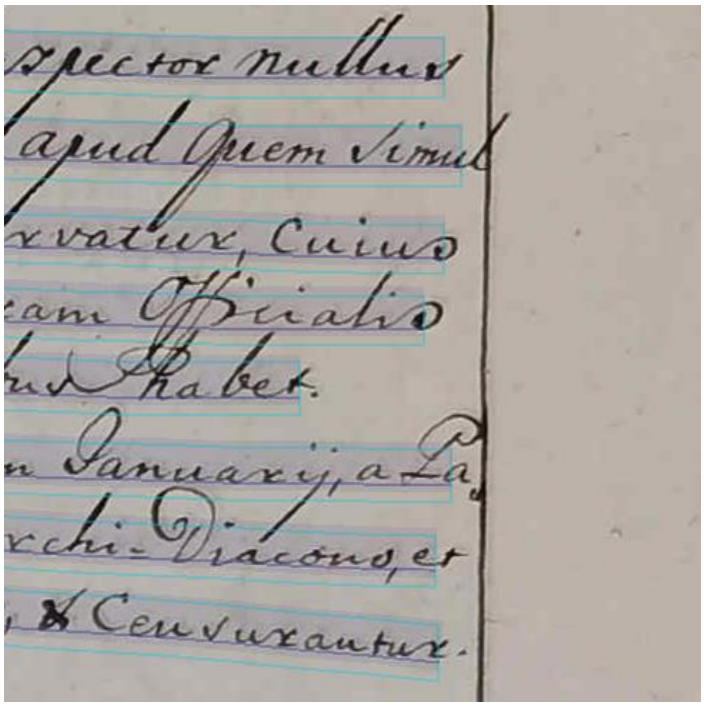
Online zdroje:

1. How to Transcribe Documents with Transkribus – Introduction: https://readcoop.eu/transkribus/howto/how-to-transcribe-documents-with-transkribus-introduction/
2. Hungaricana – Portál maďarského kultúrneho dedičstva: https://hungaricana.hu/en/
3. Platforma Transkribus: https://readcoop.eu/
4. Transkribus Beginners‘ Webinar [English]: https://www.youtube.com/watch?v=Zjm97vkx1PQ

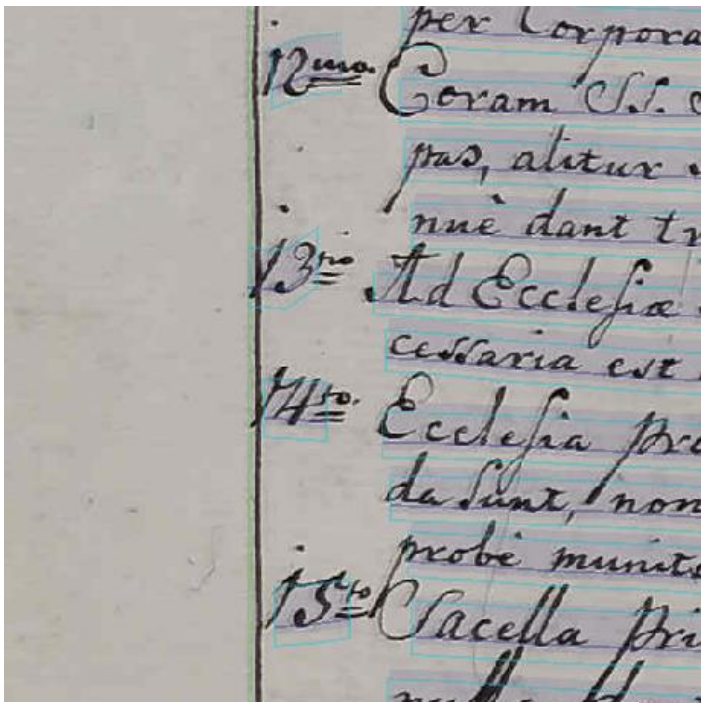
Zoznam obrazových príloh

1. Ukážka nepresnej automatickej segmentácie konca riadkov. Zdroj: Transkribus
2. Ukážka nepresnej automatickej segmentácie číslovania. Zdroj: Transkribus
3. Ukážka nepresnej automatickej segmentácie tabuľky. Zdroj: Transkribus
4. Príklad presvitajúceho textu v rámci kanonickej vizitácie farnosti Staré Hory (rukopis B) . Zdroj: Transkribus
5. Referenčný model A. Zdroj: Transkribus
6. Referenčný model B. Zdroj: Transkribus
7. Ukážka chýb v automaticky transkribovanom texte. Zdroj: Transkribus
8. Referenčný model C. Zdroj: Transkribus
9. Kombinovaný model AB. Zdroj: Transkribus
10. Porovnanie zmien v chybovosti medzi Referenčným modelom „A“ a Kombinovaným modelom „AB“ . Zdroj: Transkribus
11. Kombinovaný model ABC. Zdroj: Transkribus
12. Model skladajúci sa z troch rukopisov, vytvorený využitím base modelu. Zdroj: Transkribus
13. Ukážka chybovosti na rukopise B. Zdroj: Transkribus
14. Ukážka chybovosti na rukopise C. Zdroj: Transkribus

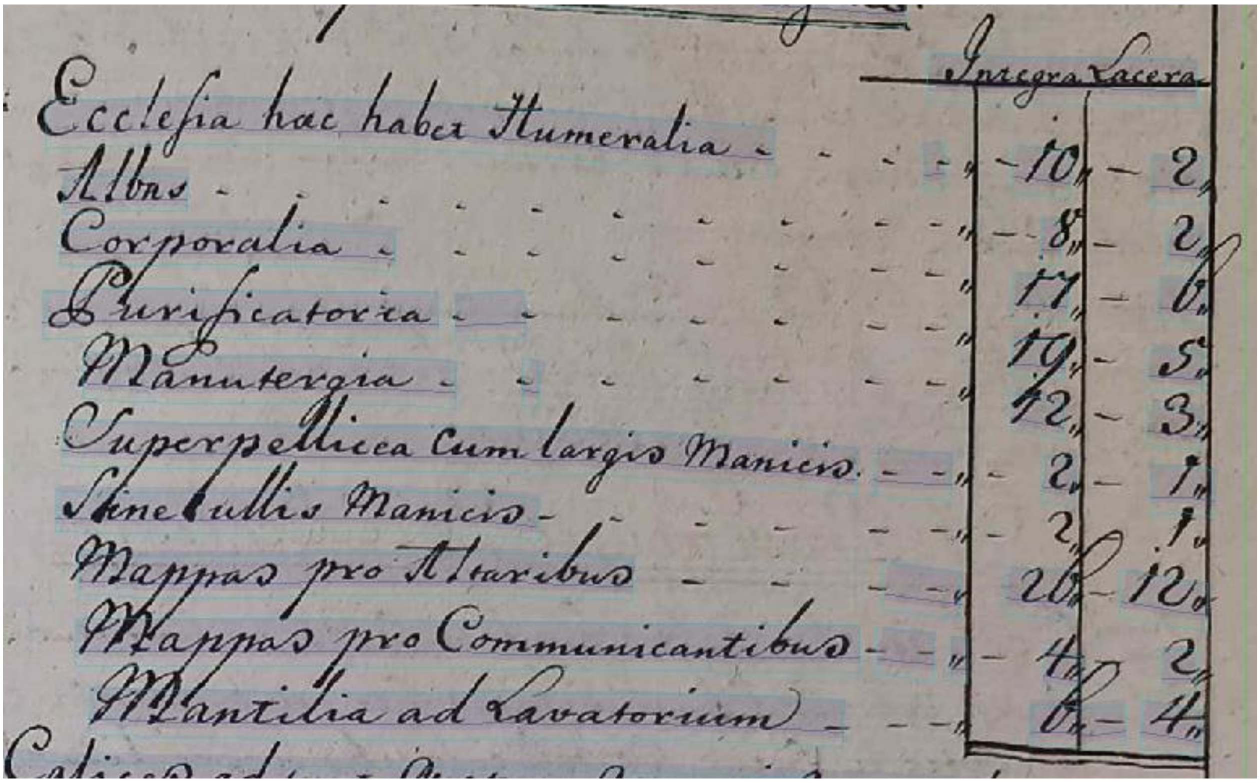
Obrazová príloha



Obrázok 1: Ukážka nepresnej automatickej segmentácie konca riadkov

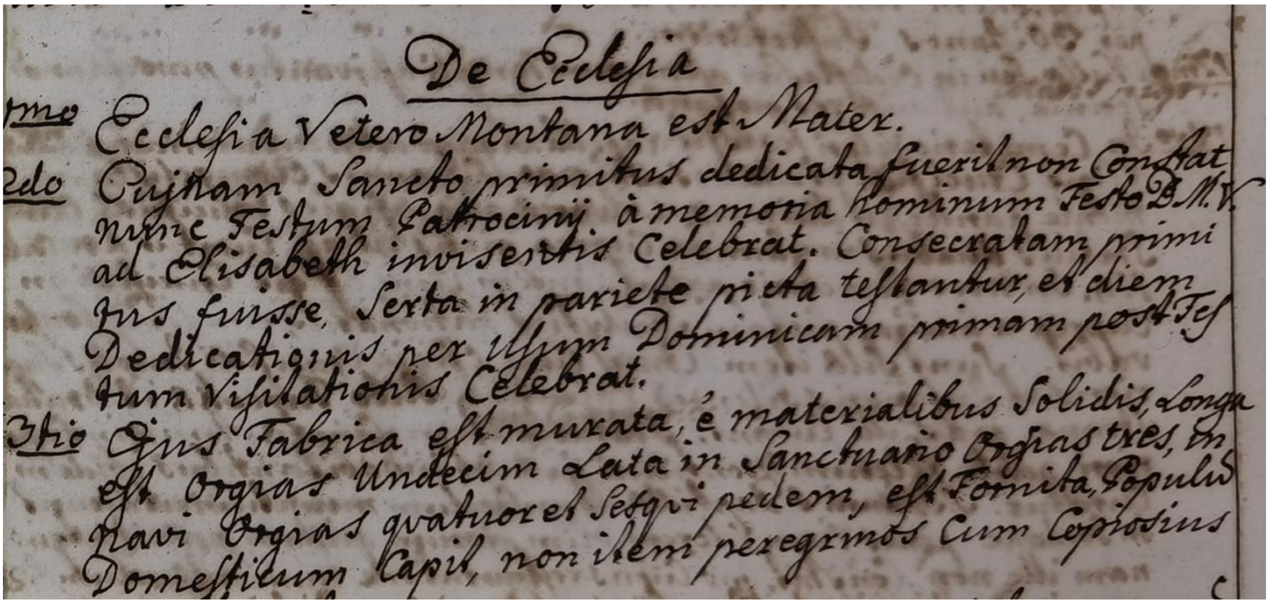


Obrázok 2: Ukážka nepresnej automatickej segmentácie číslovania



	Integra	Lacera
Ecclesia hac habet Humeralia	10	2
Albas	8	2
Corporalia	17	6
Purificatoria	19	5
Manutergia	12	3
Superpellicea cum largis Manicis	2	1
Sine tullis Manicis	2	1
Mappas pro Altaribus	26	12
Mappas pro Communicantibus	4	2
Mantilia ad Lavatorium	6	4

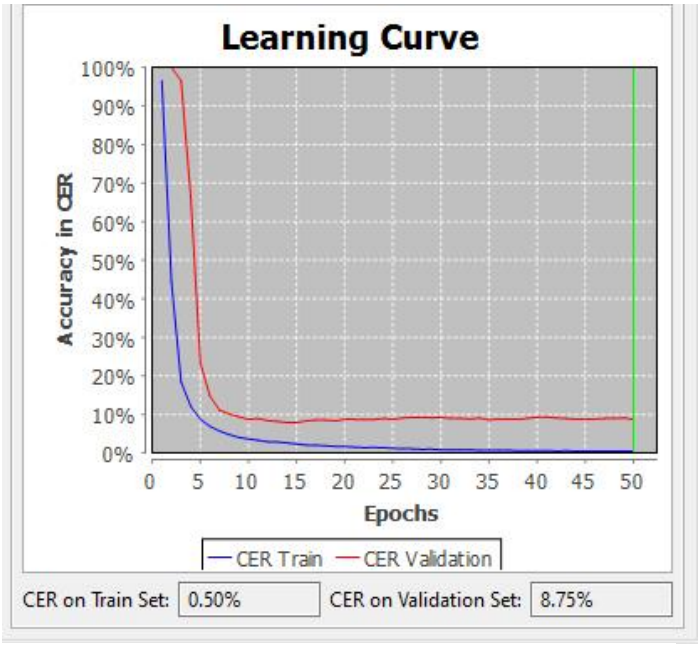
Obrázok 3: Ukážka nepresnej automatickej segmentácie tabuľky



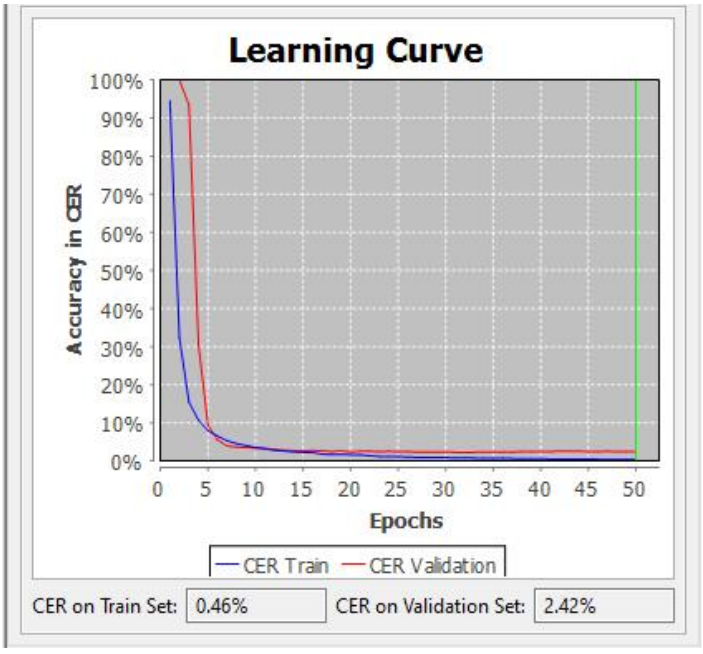
De Ecclesia

1mo Ecclesia Vetero Montana est Mater.
2do Cujnam Sancto primitus dedicata fuerit non Constat
nunc Festum Patrocinij à memoria hominum Festo B. M. V.
ad Elisabeth invisentis Celebrat. Consecratam primi
tus fuisse, Serta in pariete picta testantur, et diem
Dedicationis per illum Dominicam primam post Fes
tum Visitationis Celebrat.
3tio Ejus Fabrica est murata, é materialibus Solidis, Longa
est Orgias Undecim Lata in Sanctuario Orgias tres, in
navi Orgias quatuor et Sesqui pedem, est Fornita, Populum
Domesticum Capit, non item peregrinos Cum Copiosius

Obrázok 4: Príklad presvitajúceho textu v rámci kanonickej vizitácie farnosti Staré Hory (rukopis B)



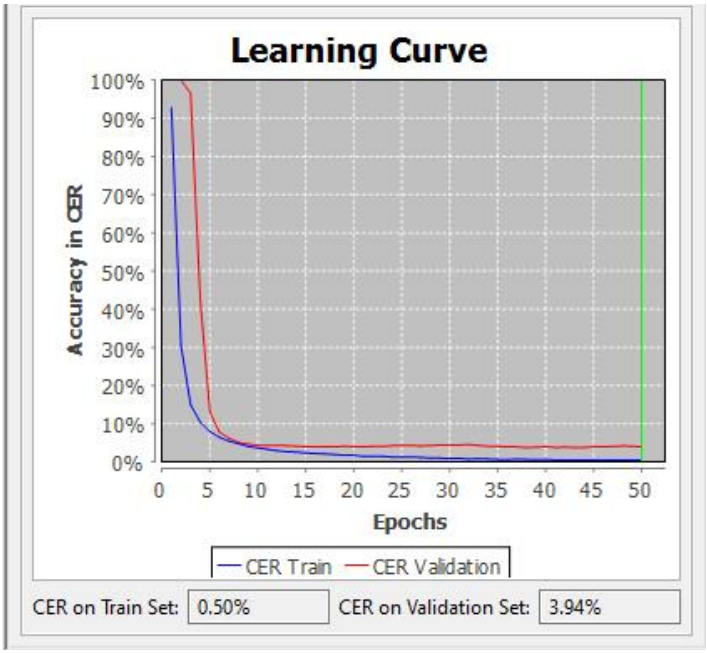
Obrázok 5: Referenčný model A



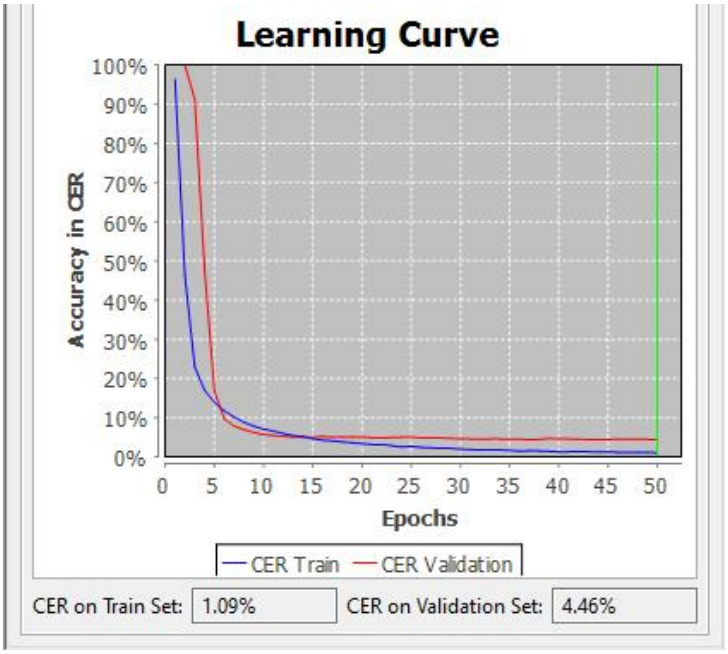
Obrázok 6: Referenčný model B

Reperiuntur in ea animæ 26. ~~Videlic et~~ Videlicet Confessionis
Capaces 14. Incapaces Confessionis 12. Idioma Loqvuntur
slavicum.
5to
Populus est ~~Devotus~~ Devotus, in Ecclesia ~~freqvens,~~ freqvens Parochi reverens
bene instructus Rudimentis Fidei ~~a~~ & Romano Catholicus.
6to
Diversæ Religionis Conjugati nulli dantur.
Qvarta Filialis est Tureczka Inferior. Habitationes
Laboratorum Montanorum.
Distat a Matre solida media hora Via incomoda Lapidosa
~~eet~~ et Contra montem.
3tio
non existit hic ulla Ecclesia aut Sacellum.
4to
Reperiuntur in ea animæ. 141. ~~Videlicet:~~ videlicet: Confessionis
capaces 96. Incapaces Confessionis 45. Idioma Loqvuntur
~~Plavicum.~~ slavicum.

Obrázok 7: Ukážka chýb v automaticky transkribovanom texte



Obrázok 8: Referenčný model C

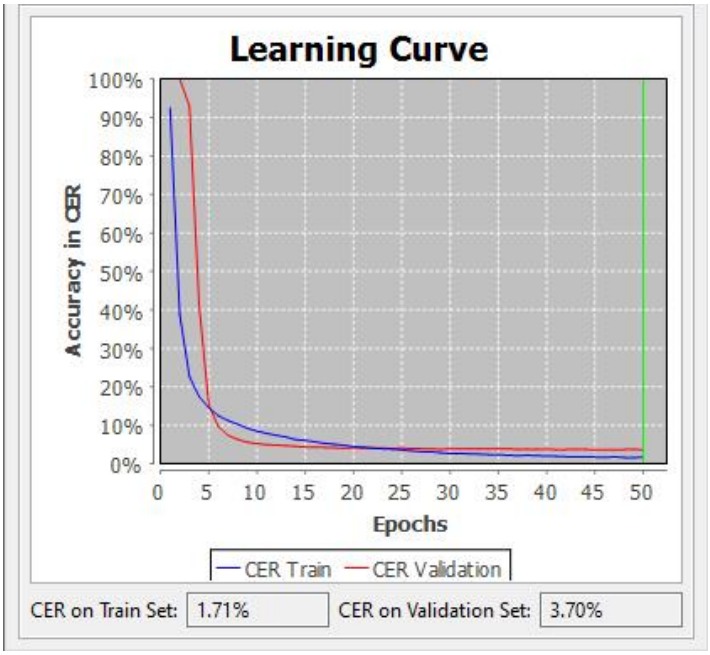


Obrázok 9: Kombinovaný model AB

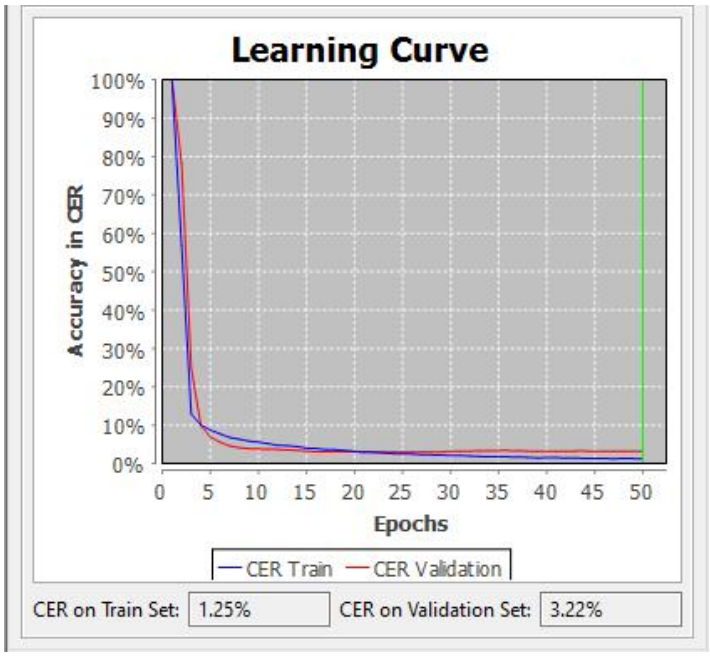
~~Aqvaductum,~~ Aqvæductum, Stabulum, et Currule. In hac
domo Parochus debitam ~~Nunditiem~~ Munditiem servat,
~~iesqve Iuas ordimate~~ Resqve suas ordinate habet.
Domum Parochialem, et ~~dccessoria Patronatis~~ accessoria Patronatus
dum opus est reparat, ~~iis demqne~~ iisdemque Conservan
dis ~~iutentus~~ intentus est.

Aqvæductum, Stabulum, et Currule. In hac
domo Parochus debitam Munditiem ~~servat~~ servat,
Resqve suas ordinate habet.
Domum Parochialem, et accessoria ~~Patronatis~~ Patronatus
dum ~~Opus~~ opus est reparat, ~~iisdenmque conservan~~ iisdemque Conservan
dis ~~iitentus~~ intentus est.

Obrázok 10: Porovnanie zmien v chybovosti medzi Referenčným modelom „A“ a Kombinovaným modelom „AB“



Obrázok 11: Kombinovaný model ABC



Obrázok 12: Model skladajúci sa z troch rukopisov, vytvorený využitím base modelu

non existit hic ulla Ecclesia, aut ~~sacellum.~~ Sacellum.
4to
Reperiuntur in ea animæ ~~2 6.~~ 26. Videlicet Confessionis
Capaces 14. Incapaces Confessionis 12. Idioma Loqvuntur
slavicum.
5to
Populus est ~~Devotus~~ Devotus, in Ecclesia freqvens Parochi reverens
bene instructus Rudimentis ~~Fidei,~~ Fidei & Romano Catholicus.
6to
Diversæ Religionis Conjugati nulli dantur.
Qvarta Filialis est Tureczka Inferior. Habitationes
Laboratorum Montanorum.
Distat a Matre solida media hora Via incomoda Lapidosa
et ~~Contra,~~ Contra montem.
3tio
non existit hic ulla Ecclesia aut ~~sacellum.~~ Sacellum.

Obrázok 13: Ukážka chybovosti na rukopise B

De ~~Oeconomo,~~ Oeconomo seu Inspectore
Ecclesiæ.
1mo
In defectu œconomi, seu ~~Insfectoris~~ Inspectoris Ecclesiæ, Ædituus est
~~Antreas~~ Andreas Longauer qvi vices Inspectoris ~~suplet,~~ suplet. Ignobilis
subsistit comode, constitutus a Patribus societatis Paro
chiam hanc administrantibus.
2do
In incassandis Ecclesiæ proventibus est accuratus.
3tio
~~Progationes~~ Erogationes pro Ecclesiæ ~~necessitate,~~ neccessitate fiunt ordinarie præsen
te Ædituo ~~Templi~~ Templi, utpote alteram a Cassa Ecclesiæ ~~clavem~~ clavim
habente, expensas connotante Parocho.
4to
Ecclesiæ Cassa, seu Cista asservatur apud Parochum.

Obrázok 14: Ukážka chybovosti na rukopise C