

UNIVERZITA KARLOVA
FAKULTA HUMANITNÍCH STUDIÍ

Orální historie – soudobé dějiny



Rozhovory 2.0

Využití AI pro orální historiky

Diplomová práce

Ing. Marek Haubert

Vedoucí práce: Mgr. Jiří Hlaváček, Ph.D.

Praha 2024

Prohlášení

Prohlašuji, že jsem práci vypracoval/a samostatně. Všechny použité prameny a literatura byly řádně citovány. Práce nebyla využita k získání jiného nebo stejného titulu.

V Praze dne 27. dubna 2024

.....

Marek Haubert

Poděkování

Na tomto místě bych chtěl vyjádřit svůj dík Magdalence, Marvinovi, rodině a přátelům, kteří mi pomáhali a podporovali mě při studiu a přípravách této práce. Speciální poděkování patří mému vedoucímu práce, Mgr. Jiřímu Hlaváčkovi, Ph.D., za jeho cenné rady, odborné vedení a trpělivost. Dále děkuji svým spolužákům a kolegům za jejich povzbuzení, inspirativní diskuse a neocenitelnou podporu.

Abstrakt

Tato diplomová práce se zabývá propojením orální historie s moderními informačními technologiemi (IT), a především umělou inteligencí (AI). Cílem je prozkoumat, jak mohou tyto technologie obohatit praxi orálních historiků a zefektivnit zpracování orálně-historických rozhovorů. Práce na praktických ukázkách demonstruje možnosti integrace AI a IT služeb ve všech fázích orálně-historického výzkumu, tedy od přípravy na rozhovor přes jeho realizaci a následnou transkripci, analýzu, interpretaci až po jeho zabezpečení, archivaci a publikaci veřejnosti. V diplomové práci je kladen důraz na praktické ukázky použití jednotlivých technologií a na rešerši dostupných služeb a nástrojů, které mohou usnadnit nahrávání rozhovorů, jejich přepis, analýzu sentimentu či vytváření metadat.

Klíčová slova

umělá inteligence, informační technologie, orální historie, analýza sentimentu, metadata

Abstract

This diploma thesis explores the integration of oral history with modern information technologies (IT), especially Artificial Intelligence (AI), aiming to investigate how these technologies can enrich the practice of oral historians and make the processing of oral historical interviews more efficient. It demonstrates, through practical examples, the possibilities of integrating AI and IT services at all stages of oral historical research, from interview preparation to realization, subsequent transcription, analysis, interpretation, and up to its security, archiving, and public publication. The thesis emphasizes practical demonstrations of technology use and research on available services and tools that can facilitate recording interviews, their transcription, sentiment analysis, or metadata creation.

Keywords

artificial intelligence, information technology, oral history, sentiment analysis, metadata

Obsah

1 Úvod.....	1
2 Metodologie výzkumu.....	3
3 Teoretický rámec.....	4
3.1 Orální historie a moderní technologie.....	4
3.2 Digitalizace.....	5
3.3 Umělá inteligence.....	9
3.3.1 Sní androidi o elektrických ovečkách?.....	9
3.3.2 Hrozby a omezení AI.....	11
3.3.3 Budoucnost AI v orální historii.....	14
3.3.4 Zpracování přirozeného jazyka.....	18
4 Empirická část: moderní technologie a možnosti jejich využití v jednotlivých fázích orálně-historického výzkumu.....	22
4.1 Příprava rozhovoru.....	22
4.1.1 Rešerše.....	22
4.1.2 Simulace rozhovoru.....	23
4.2 Terénní výzkum.....	26
4.3 Zpracování dat.....	29
4.3.1 Vylepšení kvality audio záznamu.....	29
4.3.2 Přepis rozhovorů.....	30
4.3.3 Metadata.....	39
4.4 Analýza dat.....	40
4.4.1 Analýza rozhovoru.....	40
4.4.2 Analýza orálně-historických dat v jiných vědních disciplínách.....	41
4.4.3 Analýza sentimentu.....	42
4.5 Archivace dat.....	49
4.6 Případová studie: 100 studentských revolucí.....	53
5 Závěr.....	65

1 Úvod

V této diplomové práci se zaměřím na propojení orální historie s moderními informačními technologiemi a obzvláště s některými nástroji umělé inteligence, které v posledních měsících a letech zažívají obrovský rozmach a podle mnohých predikcí bude tento trend v následujících letech pokračovat.



ARTIFICIAL INTELLIGENCE (AI) MARKET SIZE, 2022 TO 2032 (USD BILLION)



Obrázek 1: Predikce vývoje velikosti trhu umělé inteligence od roku 2022 do roku 2032. Dostupné z: [<https://www.precedenceresearch.com/artificial-intelligence-market>]

Tyto moderní technologie si již našly, nebo právě nacházejí, cestu do téměř všech vědních oborů včetně orální historie, kde díky svým schopnostem zpracovávat a analyzovat velké objemy nestructurovaných dat a rozpoznávat vzory nabízí nové možnosti pro historický výzkum a interpretaci. Orální historie jako disciplína zabývající se shromažďováním, zaznamenáváním a interpretací vzpomínek jednotlivců se ocitá v bodě, kdy může z těchto technologických inovací významně profitovat.

Hlavním cílem této práce je prozkoumat a zhodnotit, jak mohou moderní informační technologie a nástroje umělé inteligence (dále jen „AI“) obohatit praxi orálních historiků a historiček (dále jen „orální historici“) a přispět k efektivnějšímu a kvalitnějšímu zpracování rozhovorů s historickou tematikou. Tohoto cíle hodlám dosáhnout řešerší a praktickým testováním nejrůznějších služeb a nástrojů, které mohou být pro účely orálně-historického výzkumu užitečné, a to ve všech jeho

fázích. AI nástroje lze použít například pro generování nápadů ve fázi příprav výzkumného projektu, nebo při přípravě před samotným rozhovorem. Mohou také pomoci při formulaci oslovovacího emailu nebo dopisu nátorům. Moderní technologie používáme i při samotném nahrávání rozhovoru pro získání kvalitní a přesné nahrávky. Poté nám může umělá inteligence automaticky vytvořit přepis rozhovoru a také asistovat při jeho následné analýze a interpretaci. Informační technologie využijeme i v závěru výzkumu pro archivaci dat, případně pro jejich sdílení s dalšími výzkumníky nebo pro publikaci výsledků výzkumu široké veřejnosti. Po celou dobu výzkumu hrají důležitou roli také moderní technologie na ochranu dat, aby nedošlo k jejich zničení, odcizení, či neoprávněné manipulaci.

Moje motivace pro tuto práci vychází ze zájmu o kombinaci technických a humanitních oborů. V pozici programátora se zájmem o společenské a humanitní vědy vnímám, že kombinace těchto oborů není zcela ojedinělá, přesto jde obecně spíše o výjimky než o pravidlo. Mezi těmito obory panují vzájemné předsudky, humanitně vzdělaní lidé často v technických oblastech tápou, zatímco technici a programátoři mnohdy nechápou nebo nedoceňují význam humanitních oborů. Absence základní technické gramotnosti pak může vést ke zbytečnému plýtvání časem na činnosti, které by mohly být snadno a efektivně automatizovány.¹ Výzkumník by tak mohl čas věnovaný rutinním činnostem věnovat samotnému výzkumu. Přestože AI nástroje jako např. ChatGPT, Google Gemini, Mistral či Claude pronikly svojí popularitou i mezi širokou veřejnost (ostatně heslo ChatGPT se stalo nejvyhledávanějším heslem anglické Wikipedie roku 2023)², jedná se pouze o pomyslnou špičku ledovce možností, které mohou historické vědě moderní technologie a AI nabídnout. Možnosti AI jsou mnohem rozsáhlejší a mohou přinést zásadní změny ve způsobu, jakým se v humanitních vědách pracuje. Tato diplomová práce je proto zaměřena na prozkoumání těchto možností a demonstraci, jakým způsobem mohou být nové technologie pro obor orální historie prospěšné.

Práce je strukturována tak aby nejprve představila teoretické pozadí oboru orální historie, moderních informačních technologií a umělé inteligence. Následně se v praktické části zaměří na jednotlivé fáze orálně-historického výzkumu a pokusí se identifikovat činnosti, které by mohly být podpořeny, usnadněny či zefektivněny počítačovými technologiemi. To vše bude podepřeno praktickými a užitečnými ukázkami a návody. Nebude chybět také řešerše konkrétních volně dostupných služeb či nástrojů, které lze pro daný účel použít, a prozkoumání jejich výhod či

1 HLAVÁČEK, Jiří. *Digitální historie: teoretická a metodologická východiska pro výzkum soudobých dějin 2.0*. In: *Soudobé dějiny*. 2024, 31(2), s. 17 (rukopis).

2 WIKIMEDIA FOUNDATION. *Announcing Wikipedia's most popular articles of 2023*. Wikimedia Foundation [online]. 2023 [cit. 2024-01-06]. Dostupné z: <https://wikimediafoundation.org/news/2023/12/05/announcing-wikipedias-most-popular-articles-of-2023/>

omezení. Závěr práce pak zhodnotí dopady vyzkoušených technologií na práci orálních historiků a zhodnotí možné směry budoucího vývoje spolupráce těchto dvou oborů.

2 Metodologie výzkumu

Umělá inteligence je s námi již několik desítek let. Přesto o ní běžný uživatel informačních technologií měl až do nedávna povědomí především zásluhou sci-fi románů. Nástroje jako e-mail, překladač nebo sociální sítě od velkých softwarových společností (Google, Meta, Microsoft) sice umělou inteligenci na pozadí používají již více než deset let³, ale do širšího povědomí a k praktickému využití se AI dostala v podstatě až od zveřejnění služby ChatGPT a to poměrně nedávno na konci roku 2022.⁴ Tištěných pramenů zabývajících se tímto tématem je k dispozici naprosté minimum a práce se v tomto ohledu opírá především o online dostupné vědecké práce a články. Mimo téma umělé inteligence se však práce zabývá také například tématy digitalizace historických dat, kde vychází především z článků amerického orálního historika Douga Boyda, českého orálního historika a současně vedoucího této práce Jiřího Hlaváčka či českého technologického publicisty Petra Koubského.

V empirické části práce pracuji metodou case study. Vybrané softwarové nástroje a služby umělé inteligence jsou prakticky testovány na orálně-historických nahrávkách rozhovorů, které byly pořízeny v rámci projektu *100 studentských revolucí*⁵, který pod vedením prof. Miroslava Vaňka realizoval v letech 2017–2018 tým tazatelů a tazatelek z Centra orální historie při Ústavu pro soudobé dějiny AV ČR, v. v. i.

Práce je primárně určena pro humanitně zaměřené čtenáře a pro použití mnohých z nástrojů není nutná pokročilá znalost IT. Ve většině případů lze nástroje ovládat pomocí přehledného webového uživatelského rozhraní, do něhož zpravidla stačí pouze nahrát soubor, nastavit základní parametry a po zpracování souboru si pouze stáhnout výsledek. Práce s chatboty a AI asistenty je také velmi snadná. Pouze v několika případech je v práci uveden kód, který je napsán v programovacím jazyce Python⁶. Uvedenou ukázkou kódu ale vždy postačí pouze zkopírovat či přepsat do programu, který dokáže s jazykem Python pracovat a následně zde tento kód spustit. Není tedy nutné mít jakékoliv

3 Na vylepšení služby Google Translate pomocí neuronových jazykových modelů má významnou zásluhu také český vědec Tomáš Mikolov.

4 Viz například MARR, Bernard. *A Short History Of ChatGPT: How We Got To Where We Are Today* [online]. 2023 [cit. 2024-03-02]. Dostupné z: <https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/>

5 Viz ÚSTAV PRO SOUDOBÉ DĚJINY AV ČR. Studenti. *100 studentských revolucí* [online]. 2019. Dostupné z: <http://www.studenti89.usd.cas.cz/studenti/>

6 Python je rychlý, efektivní, multiplatformní, relativně jednoduchý programovací jazyk a je jedním z nejpoužívanějších programovacích jazyků na světě. Není-li již na počítači předinstalovaný, lze jej nainstalovat pomocí návodu na webových stránkách: <https://www.python.org/downloads/>

zkušenosti s programováním. Jedním z takových jednoduchých nástrojů pro exekuci jazyka Python je například Jupyter Notebook⁷.

3 Teoretický rámec

3.1 Orální historie a moderní technologie

Výzkumná metoda orální historie se zaměřuje na zaznamenávání a analýzu orálních svědectví a vzpomínek. Tento přístup nabízí unikátní pohled na minulost skrze osobní zkušenosti a příběhy jednotlivců, často osvětlující aspekty historie, které nejsou zaznamenány v tradičních písemných pramenech. Orální historie má kořeny v dávných ústně tradovaných vyprávěních, ale jako akademická disciplína se začala rozvíjet teprve ve 20. století. Pro účely mé práce je příznačné, že této proměně významně přispěly právě technologické inovace, umožňující záznam a uchování mluveného slova. V 60. a 70. letech 20. století došlo k výraznému rozvoji orální historie v souvislosti s rostoucím zájmem o sociální dějiny a takzvanou historií „zdola“, která se snaží zachytit vzpomínky obyčejných, nepriviligovaných lidí.⁸

Orální historie využívá rozhovoru jako hlavního nástroje pro shromažďování dat. To vyžaduje pečlivou přípravu, stanovení cílů a dodržování etických zásad, včetně zajištění souhlasu účastníků a zachování důvěrnosti. Výzkum vyžaduje velmi citlivý přístup k narátorům. Důležitým aspektem je také interpretace a analýza shromážděných rozhovorů, při níž se výzkumníci musí potýkat především se subjektivitou a selektivností paměti.

Orální historie přináší perspektivy, které mohou doplnit a někdy i vyvrátit tradiční historické narativy. Oproti mnohým jiným metodám historického výzkumu pronikne hlouběji do sociálních, kulturních a osobních aspektů historie a umožňuje vnést do historického výzkumu i hlasy, které by jinak zůstaly nevyslyšeny. Tato metoda je zvláště významná pro pochopení nedávné historie a pro zachycení zkušeností takřkajíc z první ruky.⁹

Orální historie čelí několika výzvám, včetně otázek objektivity a spolehlivosti paměti. Její kritici poukazují především na to, že vzpomínky mohou být zkreslené a subjektivní. Proto je důležité orální zdroje pečlivě konfrontovat s jinými dostupnými historickými materiály. Klíčovou roli hrají etické otázky, jako je zacházení s citlivými informacemi a ochrana soukromí respondentů.

7 Jupyter Notebook lze snadno nainstalovat pomocí návodu na webových stránkách: <https://jupyter.org/install>

8 VANĚK, Miroslav a MÜCKE, Pavel. *Třetí strana trojúhelníku: teorie a praxe orální historie*. 2., přepracované a doplněné vydání. Orální historie a soudobé dějiny. Praha: Univerzita Karlova v Praze, nakladatelství Karolinum, 2015. ISBN 978-80-246-2931-5, s. 14–19.

9 Tamtéž, s. 27–30

V posledních desetiletích výrazně přispěly k rozvoji orální historie moderní technologie, jako jsou digitální nahrávací zařízení a softwarové nástroje pro přepis a analýzu orálních nahrávek, které umožňují snadnější nahrávání, uchování a analýzu orálně-historických rozhovorů.¹⁰ Důležité jsou také nástroje pro digitální archivaci a databázové systémy pro správu a vyhledávání v rozsáhlých sbírkách orálně-historických nahrávek rozhovorů. Právě moderní technologie postavené na AI nástrojích mohou značně zjednodušit a zefektivnit tyto procesy.

Orální historikové potřebují své rozhovory plánovat, nahrávat, analyzovat a dále zpracovávat, což zahrnuje proces výběru a oslovení respondentů, formulaci otázek, přípravy rozhovorů, přípravy technických prostředků pro nahrávání a podobně. Kromě shromažďování a analýzy dat je pro orální historiky důležité umět svá zjištění prezentovat a předat vědecké poznatky širší veřejnosti. Během celého procesu musí zároveň brát zřetel na etické aspekty výzkumu. Musí zabezpečit ochranu soukromí a důvěrnosti informací, získat informovaný souhlas od účastníků a s respektem zacházet s jejich osobními příběhy. Moderní technologie a média hrají v tomto procesu důležitou roli.

3.2 Digitalizace

Digitalizace představuje pro orální historiky zásadní posun ve způsobu nahrávání, archivace a analýzy orálně-historických rozhovorů. Zároveň se jedná o nezbytnou podmínku pro jakékoliv následné počítačové zpracování dat. Digitalizace umožňuje převod tradičních analogových nahrávek do digitální podoby, což přináší řadu výhod, od lepší přístupnosti po snazší manipulaci s daty. Digitalizace značně usnadňuje i archivaci rozhovorů a jejich sdílení, jelikož digitální data mohou být snadno ukládána, katalogizována a vyhledávána v elektronických databázích, což umožňuje historikům rychlejší přístup k potřebným informacím a jejich efektivnější sdílení v rámci akademické komunity i mimo ni. V digitálních orálně-historických sbírkách mohou být ovšem kromě samotných rozhovorů ukládána i další související multimediální data, jako jsou například fotografie, přepisy rozhovorů, karta narátora, protokol o rozhovoru, informovaný souhlas a další.¹¹

Digitalizace nemusí představovat pouze převod záznamu do digitální podoby, ale také lze v jejím průběhu pomocí moderních technologií původní záznam (audio i video) vylepšit (odstranit šum a jiné rušivé zvuky na pozadí), což může zlepšit srozumitelnost nahraného rozhovoru.¹²

10 Tamtéž, s. 21–27

11 Viz například HLAVÁČEK, Jiří. *Metodické postupy archivace orálně-historických pramenů v digitálním věku*. Archivní časopis. 2019, 69(4), 389–401. ISSN 0004-0393, s. 395.

12 Návod, jak prakticky zlepšit kvalitu nahrávky, se nachází v pozdějších kapitolách, viz kapitola: 4.3.1 Vylepšení kvality audio záznamu.

Digitalizovaný záznam lze pak dále počítačově zpracovávat a analyzovat. Lze vytvořit automatický přepis rozhovoru.

Digitalizovaná data jsou méně náchylná k poškození než analogové záznamy a jejich další kopírování je na rozdíl od analogových záznamů dále bezztrátové. I po mnohonásobném překopírování je zachována původní kvalita. Lze tak vytvářet další a další kopie, které lze uložit na několika na sobě nezávislých místech, a tím zajistit větší ochranu před jejich ztrátou. Digitalizovaná data lze snadno začleňovat do digitálních prezentací, online výstav, vzdělávacích materiálů a dalších multimediálních aplikací, což rozšiřuje možnosti prezentace orální historie a jejího šíření mezi širokou veřejností.

Kromě zmíněných výhod ale vyžadují digitalizovaná data zcela jiný formát ochrany před manipulací, odcizením nebo zneužitím, než analogové záznamy. Digitální archivy připojené k internetové síti¹³ vyžadují nejen ochranu před fyzickým poškozením, ale také ochranu softwarovou. Neblahým důsledkem digitalizace může být také opomíjení kontextu, v němž se hledaná informace nachází. Pomocí fulltextového vyhledávání lze vyhledávat v rozhovorech konkrétní citace, avšak bez znalosti kontextu celého rozhovoru může být takto selektivní výběr zavádějící.

Zdaleka ne všechny druhy dokumentů jsou vhodné ať už k omezenému nebo plnému zveřejnění. V případě práce s dokumenty s omezeným přístupem, tajnými a citlivými dokumenty, jejichž necenzurované či neanonymizované zveřejnění by mohlo výzkumníka či narátora vystavit nechtěné pozornosti, posměchu, nebezpečí nebo hrozbě právního postihu, je naopak nutné, takové záznamy co nejlépe ochránit a zamezit přístupu neoprávněným osobám. Před samotným archivováním dokumentů je nutné nejprve definovat bezpečnostní pravidla, tedy kdo bude mít k materiálům přístup a jakým způsobem ho lze získat. Dále pak jakým způsobem se bude v rámci konkrétní digitální sbírky nakládat s tajnými a citlivými dokumenty¹⁴ a jaké technologie budou používány pro uchování dat.

Pevné disky, flash paměti a další digitální nosiče mají svoji životnost a poruchovost. Rovněž i analogové záznamy mají omezenou životnost a je otázkou času, zda za několik desítek let budou přehrávače schopné přehrát analogové záznamy nadále existovat i v jiné formě než pouze jako muzejní exponáty. V ideálním případě by data měla být zároveň zálohována na několika na sobě nezávislých úložištích. Zrcadlení dat na větší množství disků snižuje riziko ztráty dat, ale pokud budou přesto veškeré kopie umístěny v téže budově, budou například v případě požáru v ohrožení

13 Rešerše softwarových řešení pro digitální archivy se nachází v pozdějších kapitolách, viz kapitola: 4.5 Archivace dat.

14 Viz například HLAVÁČEK, Jiří. *Metodické postupy archivace orálně-historických pramenů v digitálním věku*. Archivní časopis. 2019, 69(4), 389–401. ISSN 0004-0393, s. 394.

všechny kopie zároveň. V případě velmi důležitých a cenných dokumentů stojí za uvážení uchování kopií ve zcela jiném státě. Historie nám ukázala, že změny politického režimu často vedou k zajišťování nebo likvidaci archivů, zejména v případě nedemokratických systémů.

Cílem archivace záznamů je jejich uchování po dobu minimálně několika desítek let, dokud nenastane vhodná doba pro přenos na lepší, spolehlivější a výkonnější nosiče, které nám budoucnost přinese. V minulosti se tomu tak již několikrát stalo. Z prvních fonografů se audio záznamy postupně přenášely na gramofonové desky a následně na magnetofonové pásky. Přenos záznamu z jednoho nosiče na druhý ale nikdy nebyl zcela dokonalý a byl kvalitativně ztrátový. V posledních dekádách dochází k převodu analogových záznamů na digitální, jejichž další kopírování a přenášení lze již provádět zcela bezztrátově. Digitalizovat lze jak textové záznamy, tak i veškeré audiovizuální záznamy.

Digitalizace archivu orální historie vyžaduje mnoho času a úsilí. Digitalizací musí projít dokumenty rozličných formátů např. audio kazety, VHS, kompaktní disky i písemné dokumenty, jako jsou přepisy rozhovorů, životopisy narátorů a další související dokumenty. Po samotné digitalizaci musí být každý dokument detailně popsán, opatřen metadaty, která ho definují a umožňují snadnější orientaci a vyhledávání a na závěr musí být určeno, kdo a za jakých podmínek bude mít ke kterým dokumentům přístup. Splňuje-li stanovené podmínky, může mít k digitalizovanému online archivu okamžitý přístup jakýkoliv výzkumník z celého světa.¹⁵

S digitalizací vědeckých dat úzce souvisí projekt *Open Science*, který si klade za cíl podporovat a zveřejňovat vědecký výzkum a jeho výsledky tak, aby byly přístupné a užitečné pro širokou veřejnost¹⁶. Digitalizace dat v orální historii tak slouží nejen ke zlepšení kvality výzkumu, ale právě kvůli projektům jako je *Open Science* se stává pro splnění požadavků tohoto projektu přímo nezbytností. Digitalizace umožňuje, aby byly orálně-historické rozhovory, jsou-li vhodné pro zveřejnění a neobsahují-li citlivé informace, snadno přístupné online. Digitalizovaná data lze zveřejňovat v online archivech, kde jsou dostupná jak vědcům, tak i veřejnosti. Iniciativa dále zdůrazňuje transparentnost a reprodukovatelnost vědeckého výzkumu. Digitalizace dat v orální historii, obdobně jako v jiných vědních disciplínách, umožňuje podrobné dokumentování postupů a výsledků výzkumů, což napomáhá transparentnosti a umožňuje jiným výzkumníkům ověřit a reprodukovat výsledky, snadno sdílet data a kombinovat je s daty z jiných oborů, což otevírá další možnosti pro nové výzkumné projekty a spolupráce.

15 CARPENTER, Layne. LAUPUS HEALTH SCIENCES LIBRARY. *Oral History Digitization Project*. East Carolina University [online]. 2019 [cit. 2024-01-19]. Dostupné z: <https://hsl.ecu.edu/2019/08/15/oral-history-digitization-project/>

16 UNESCO. *An introduction to the UNESCO Recommendation on Open Science*. Open Science [online]. 2022 [cit. 2024-02-04]. Dostupné z: <https://doi.org/10.54677/XOIR1696>, s. 2.

Kromě transparentnosti a spolupráce jsou dalšími hlavními pilíři iniciativy *Open Science* rovnost příležitostí, zodpovědnost, flexibilita a udržitelnost.¹⁷ To znamená, že se projekt snaží zamezit exkluzivnímu přístupu k výsledkům vědeckého výzkumu. Staví se proti tomu, aby byly výsledky výzkumů (programy, zdrojové kódy, metodické postupy) publikovány v placených časopisech a byly přístupné pouze výzkumníkům z rozvinutých zemí, kteří si přístup k nim mohou finančně dovolit.¹⁸ Zveřejňování informací pomáhá šířit historické povědomí mezi širokou veřejností. Zveřejněné výsledky vědeckého bádání mohou být využitelné i pro vzdělávání na školách. Digitální archivy také poskytují stabilní a dlouhodobé řešení pro uchování historických záznamů, čímž zajišťují udržitelnost znalostí i pro budoucí generace, což je další z hlavních pilířů, na kterých stojí iniciativa *Open Science*.

Pro zvýšení opětovné využitelnosti samotných výzkumných dat se používá označení FAIR data. Jedná se o akronym z anglických výrazů pro principy, které by taková data měla dodržovat, tedy aby byla dohledatelná (findable), přístupná (accessible), interoperabilní (interoperable) a opětovně využitelná (reusable).¹⁹ FAIR data by měla být snadno naležitelná i pro stroje. Některé z charakteristik splňuje i systém URL²⁰, který je jedním ze základních pilířů internetu a slouží k přesné specifikaci umístění zdrojů. Zdroj na URL adrese ale nemusí být dohledatelný, jelikož dokument může být z dané adresy jeho vlastníkem kdykoliv smazán. A jelikož může být i libovolně upravován a měnit se v čase, nemůže být tento zdroj ani opětovně využitelný a pro účely historického výzkumu se na takové pomíjivé zdroje nelze spolehnout.²¹ Pro FAIR data se tedy namísto přiřazení URL adresy přiřazuje jedinečný a perzistentní identifikátor (např. DOI²²), který kromě samotného dokumentu umožňuje i jeho podrobný popis metadaty²³. Dále je pak doporučeno, aby k datům, nebo alespoň k jejich metadatům, byl umožněn bezplatný a otevřený přístup. Pro data i metadata by měly být používány standardizované výrazy, měla by být správně popsána a opatřena jasnou a dostupnou licencí. S daty, která splňují veškeré tyto principy, může snadno a rychle pracovat člověk i stroj a efektivně je využívat pro další výzkum.

17 Tamtéž, s. 9.

18 Tamtéž, s. 5.

19 UNIVERZITA KARLOVA. *FAIR data*. Centrum pro podporu open science [online]. [cit. 2024-03-02]. Dostupné z: <https://openscience.cuni.cz/OSCI-64.html>

20 Anglická zkratka pro Uniform Resource Locator (jednotný lokátor zdroje).

21 HLAVÁČEK, Jiří. *Digitální historie: teoretická a metodologická východiska pro výzkum soudobých dějin 2.0*, s. 8.

22 DOI (digital object identifier) je systém perzistentních identifikátorů objektů v prostředí digitálních sítí. Viz například https://cs.wikipedia.org/wiki/Digital_object_identifier.

23 Metadata jsou data, která poskytují informace o jiných datech. V rámci orální historie mohou metadata k nahrávce rozhovoru zahrnovat například jména tazatelů a narátorů, název výzkumného projektu, hlavní témata rozhovoru, místo a datum pořízení nahrávky apod.

3.3 Umělá inteligence

3.3.1 Sní androidi o elektrických ovečkách?

Umělá inteligence je téma, které prochází světovou literaturou už od dávných dob. Objevuje se již v legendě o pražském Golemovi, románu *Frankenstein* od *Mary Shelleyové*, dramatu *R.U.R.* od *Karla Čapka*²⁴ ve sbírce povídek *Já, robot* od *Issaca Asimova*. Od roku 1968 děsí diváky vražedný *HAL 9000* ve filmové opeře *Stanleyho Kubricka 2001: Vesmírná odysea*. Název této podkapitoly pak pochází ze slavného stejnojmenného sci-fi románu amerického spisovatele *Philipa K. Dicka* z roku 1968, na jehož námět později vzniklo pod názvem *Blade Runner*²⁵ ještě slavnější filmové zpracování a který stál u počátku uměleckého subžánru s názvem kyberpunk. Tento žánr, pro který je umělá inteligence hlavním tématem, ale plně etabloval až román *Williamy Gibsona Neuromancer* z roku 1984. V následujících letech pak zájem o tento žánr nadále rostl a vrcholu popularity dosahoval koncem milénia, kdy kinosály po celém světě okupoval film *A.I. Umělá inteligence*, filmová série *Matrix*, nebo japonská série *Ghost in the Shell*.

Počátek umělé inteligence jako vědního oboru se ale datuje již do 50. let 20. století. V roce 1956 se v Dartmouthu konala konference, kde se sešli v podstatě všichni průkopníci oboru a která je považována za oficiální zrození AI jako akademické disciplíny.²⁶ V následujících letech byl ale kvůli přehnaným očekáváním a následným zklamáním z pomalého pokroku omezen přísun financí pro další výzkum. K výraznému oživení oboru AI došlo až na počátku 21. století. Zásadním průlomem byl vývoj hlubokého učení, které umožňuje strojům učit se z obrovského množství dat pomocí umělých neuronových sítí. Právě tato technologie umožnila vznik většiny AI nástrojů²⁷, které jsou zmíněny v této práci a které nalézají uplatnění v oboru orální historie. Velkým milníkem posledních let se v oblasti AI stalo také vítězství počítačového programu AlphaGo nad světovým šampionem v čínské deskové hře Go.

Za hlavního otce umělé inteligence je nicméně považován britský matematik Alan Turing, jenž během druhé světové války stál za vývojem stroje, který sloužil k rozluštění německého šifrovacího stroje Enigma. Turing se konference v Dartmouthu v roce 1956 nezúčastnil, jelikož tragicky zemřel již o dva roky dříve. V roce 1950 ale Turing navrhl test²⁸, jehož cílem je posoudit, zda stroj

24 Viz například ŠTRÁFELDOVÁ, Milena. *Karel Čapek v R.U.R. předpověděl umělou inteligenci*. Radio Prague International [online]. 2019 [cit. 2024-03-02]. Dostupné z: <https://cesky.radio.cz/karel-capek-v-rur-predpovedel-umelou-inteligenci-8113136>.

25 Americký sci-fi film režiséra Ridleyho Scotta z roku 1982.

26 KOUBSKÝ, Petr. *Umělá inteligence vznikla v polovině dvacátého století, když se ještě věřilo na ústřední mozek lidstva*. Deník N [online]. [cit. 2024-03-09]. Dostupné z: <https://denikn.cz/63598/umela-inteligence-vznikla-v-polovine-dvacateho-stoleti-kdyz-se-jeste-verilo-na-ustredni-mozek-lidstva/>

27 Jedná se o například o rozpoznávání mluveného slova, analýzu videa a fotografií, automatický překlad apod.

28 Turing test původně nazval „The imitation game“. Dnes se častěji označuje pouze jako „Turingův test“.

vykazuje inteligentní chování srovnatelné s člověkem, a to na základě porovnání odpovědí, které člověk a stroj poskytnou na zadanou otázku.

Hlavní ocenění v kontroverzní soutěži *Loebnerova cena*, mělo být uděleno počítačovému programu, který projde Turingovým testem, zahrnujícím porozumění zvukovému a vizuálnímu obsahu a psanému textu. Toto ocenění však nebylo nikomu uděleno. Pořádání soutěže bylo nicméně ukončeno v roce 2019, nedlouho po úmrtí jejího zakladatele Hughua Loebnera²⁹, ale zároveň poměrně krátce před zveřejněním velkých jazykových modelů (ChatGPT, Gemini, Claude). Podle názoru některých vědců by velké jazykové modely již byly schopny hlavní cenu v této soutěži získat, zatímco druhá skupina vědců tvrdí, že obdobně jako technologie se vyvíjí i lidské poznání o nich. Porotci soutěže by měli předpokládat, že soutěže by se nyní účastnily i velké jazykové modely, které by, znajíce slabiny těchto systémů, byli schopni vhodně formulovanými otázkami snadno odhalit. V zásadě stačí pouze vyvést jazykový model z jeho komfortní zóny a zeptat se na téma, které nebylo součástí jeho trénovacích dat.³⁰

Nicméně již v roce 1980 předložil americký filosof John Searle takzvaný „*argument čínské pokoje*“, jehož pomocí se snažil dokázat, že samotný princip Turingova testu (schopnost smysluplně odpovídat na otázky) neimplikuje schopnost myslet a porozumět.³¹ Zjednodušeně řečeno je v tomto myšlenkovém experimentu člověk, který neumí čínsky, uzavřen v místnosti, která je plná čínsky psaných textů. Člověk obdrží otázku napsanou čínskými znaky a je schopen podle návodu v knihovně najít odpovídající odpověď, tu přepsat na papír a ten poté odevzdat zpět jako odpověď. Člověk, který položil otázku, by se mohl domnívat, že člověk uvnitř pokoje čínštinu ovládá, přestože nerozuměl ani otázce, kterou obdržel, ani odpovědi, kterou opsal. Postupoval pouze podle návodu.

Známkou lidské inteligence je schopnost abstrahovat z běžné znalosti a uplatnit tuto znalost na problém, který jsme ještě nikdy předtím neřešili.³² Přestože některé úzce specializované aplikace umělé inteligence takto abstrahovat a používat znalosti dokážou, velké jazykové modely toho ze své podstaty prozatím nejsou schopny, jelikož se jejich chování velmi podobá chování člověka v čínském pokoji. Na jednu stranu ohromují schopností plnit různé úkoly, ale principiálně se jedná

29 Skloňování cizího jména viz CHYTIL, Lukáš. *Pravopisné a morfologické adaptace přejatých antroponym zakončených na [a], [á], [o], [ó], [u], [ú]*. Olomouc, 2011. Magisterská diplomová práce. Univerzita Palackého v Olomouci, Filozofická fakulta, Katedra bohemistiky, s. 65.

30 BIEVER, Celeste. *ChatGPT broke the Turing test — the race is on for new ways to assess AI*. Nature [online]. Springer Science and Business Media, 2023, 619, 686-689 [cit. 2024-03-09]. Dostupné z: doi:<https://doi.org/10.1038/d41586-023-02361-7>, s. 687.

31 KOUBSKÝ, Petr. *Umělá inteligence vznikla v polovině dvacátého století, když se ještě věřilo na ústřední mozek lidstva*.

32 BIEVER, Celeste. *ChatGPT broke the Turing test — the race is on for new ways to assess AI*, s. 688.

pouze o generátor slov, který na základě statistiky hledá další nejpravděpodobnější slovo ve větě na základě obrovského množství trénovacích dat. To vše je doplněno pravidly, která dopisují lidští trenéři (anotátoři), aby výsledné odpovědi ještě více vyladili.³³ V odpovědích jazykových modelů se tak v podstatě jen zrcadlí odpovědi jiných lidí, které byly zaneseny do trénovacích dat, nebo dodány lidskými anotátory.

3.3.2 Hrozby a omezení AI

Pomineme obecné hrozby, kterými by umělá inteligence mohla představovat nebezpečí pro člověka a případně pro celé lidstvo, jako je například dosažení technologické singularity³⁴, nebo zneužití umělé inteligence pro vojenské účely. Zaměříme se pouze na hrozby, které jsou relevantní při použití umělé inteligence v rámci orálně-historického výzkumu.

Velkým tématem je v rámci orální historie především zajištění bezpečnosti citlivých dat, kterému bude ale později věnována samostatná kapitola.³⁵ Jedním z hlavních problémů je etický aspekt. AI může původní obsahy rozhovorů zkusit nebo je interpretovat způsobem, který není v souladu s autentičností a kontextem vyprávění. Další hrozbou může být ztráta lidského elementu v analýze, kde mohou být přehlédnuty drobné nuance, podtexty, sarkasmy. Jazykové modely mohou být omezené ve své schopnosti rozpoznat specifika konkrétní kultury nebo jazyka, což může vést k nedorozuměním a chybným interpretacím.

Jak již bylo zmíněno, nutnou prerekvizitou při použití umělé inteligence pro analýzu orálně-historických rozhovorů je digitalizace archivů. Ta může být složitá, nákladná a vyžaduje jiná zajištění specifických druhů ochrany vůči ztrátě, zneužití a neoprávněné manipulaci dat, než jaká si žádá ochrana analogových záznamů a papírových dokumentů. Jelikož orální historie často obsahuje velmi citlivé osobní informace, je třeba přidělovat otázkám soukromí a ochrany dat maximální prioritu. V roce 2014 například otrásl komunitou orální historie bezprecedentní „*Bostonský případ*“³⁶, který vyvolal řadu etických a právních otázek.³⁷ Orálně-historický projekt univerzity Boston College se zaměřoval na rozhovory převážně s bývalými členy irské vojenské organizace IRA a jejich zapojení do akcí během konfliktu v Severním Irsku. Přestože byli narátoři projektu ujištěni, že nahrávky rozhovorů zůstanou důvěrné a nebudou zveřejněny až do jejich smrti, vydaly

33 Tamtéž s. 686.

34 Jednou z hypotéz technologické singularity je stvoření sebezdokonalující se umělé inteligence, která by mohla představovat hrozbu pro celé lidstvo.

35 Viz kapitola: 4.5 Archivace dat.

36 Nezaměňovat se stejnojmenným americkým filmem z roku 1968, který byl inspirován skutečným případem řádění sériového vraha, který získal přezdívku Bostonský škrtič.

37 ORAL HISTORY ASSOCIATION. *Oral History Association response to developments in Boston College case*. Oral History Association [online]. 2014 [cit. 2024-03-09]. Dostupné z: <https://oralhistory.org/2014/05/05/oral-history-association-response-to-developments-in-boston-college-case/>

americké soudy na základě žádosti britských úřadů rozhodnutí, že některé nahrávky musí být britským úřadům vydány. Po předání těchto nahrávek pak bylo několik narátorů zatčeno a postaveno před soud v souvislosti s vraždou Jean McConvillové z roku 1972.³⁸ Tento případ nesouvisí přímo s moderními technologiemi či AI, ale ilustruje, jak citlivé informace jsou v orálně-historických sbírkách uchovávány a jaké důsledky by mohl mít jejich únik a zveřejnění.

Kvůli pokroku v oblasti umělé inteligence se budou muset uživatelé a výzkumníci v rámci orální historie potýkat s neustále se zvyšujícími riziky spojenými s falzifikací dat. Například technologie deepfake³⁹, která umožňuje vytváření velmi přesvědčivých úprav fotografií, videí i zvukových nahrávek, představuje pro autenticitu historických záznamů nezanedbatelnou hrozbu. Přestože tyto nástroje mohou být i v rámci orální historie v určitých situacích velmi užitečné, mohou být i zneužity k vytvoření falešných audio i video záznamů, na první pohled nerozeznatelných od těch skutečných. Mohou být také upraveny již existující archivní záznamy. V archivních dokumentech můžeme například nalézt nahrávky Václava Havla, s využitím technologií pro imitaci hlasu, syntézu řeči a úpravy videa mu můžeme připsat výroky, které



Obrázek 2: Deepfake fotografie papeže Františka v bílé nafukovací bundě měla na sociálních sítích více než 20 milionů shlédnutí. [<https://www.nytimes.com/2023/04/08/technology/ai-photos-pope-francis.html>]

nikdy nepronesl, umožnit mu setkání s libovolnou osobou a na obrazovce ho nechat vykonávat prakticky jakoukoliv činnost. Obrana proti šíření takto upravené nahrávky, která je téměř nerozeznatelná od skutečnosti, je velice složitá a důsledky mohou být devastující. S tím obecně souvisí rostoucí riziko šíření dezinformací a propagandy. V době, kdy je snadné sdílet informace na sociálních médiích a online platformách, může být sice AI využita k rychlému šíření nejrůznějších falešných informací, ale na druhé straně je třeba poznamenat, že AI může být zároveň využita i jako

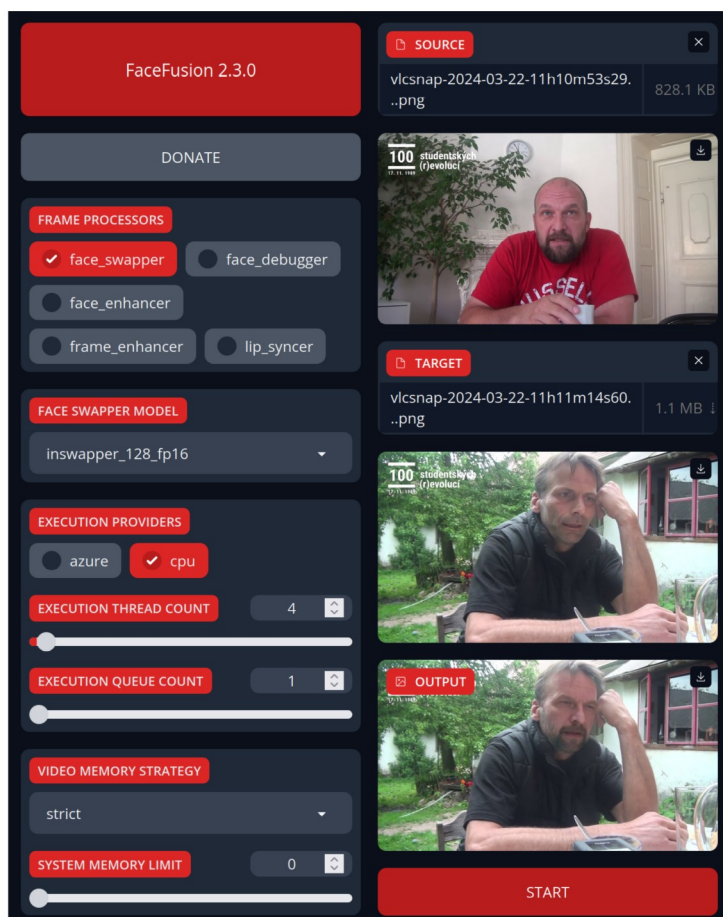
38 MORIARTY, Gerry. *Jean McConville's murder, the Boston tapes, Gerry Adams and the Ivor Bell trial*. *The Irish Times* [online]. 2019 [cit. 2024-03-09]. Dostupné z: <https://www.irishtimes.com/news/ireland/irish-news/jean-mcconville-s-murder-the-boston-tapes-gerry-adams-and-the-ivor-bell-trial-1.4053933>

39 Deepfake je technologie pro realistickou úpravu videa využívající umělou inteligenci, pomocí které je možné například upravovat tvář osoby na videu, nebo změnit obsah toho, co osoba ve videu říká při zachování původního hlasu.

obrana proti těmto hrozbám a sloužit k odhalování podvodů a falzifikací. Falšování historických pramenů není ani v naší historii žádnou novinkou, kterou by umožnil až nástup moderních technologií. Například retušované fotografie, ze kterých zmizeli popravení komunističtí poslanci Rudolf Slánský nebo Vladimír Clementis pocházejí již z 50. let 20. století.⁴⁰

Mnohé z AI nástrojů, které jsou zmíněny v této práci a lze je využívat pro potřeby orální historie, jsou založeny na technologiích zpracování přirozeného jazyka⁴¹ a jazykových modelech. Proto je nutné zohlednit omezení, která z toho plynou. Jazykový model je obvykle trénován na rozsáhlé datové sadě obsahující texty z internetu, včetně knih, článků a webových stránek. Z toho vyplývá, že do výsledného modelu se dostane obrovské množství různorodých textů včetně těch, které obsahují rasové, xenofobní či sexistické předsudky. Jazykový model je zpravidla ještě doplněn o opatření, které by mělo snížit riziko, že výsledky produkované modelem budou těmito daty kontaminovány, ale riziko zkreslení výsledků nelze zcela vyloučit. Akademičtí pracovníci si všimli, že mnoho galerií, knihoven, muzeí a archivů má ve svých sbírkách stále zakomponována historická zkreslení, hluboce zakořeněná skrze staletí kolonialismu.⁴² Některé nekontrolované použití AI pro automatickou indexaci takových sbírek by mohlo ještě více ukotvit předsudky vůči jiným kulturám.

Jazykový model *GPT-3.5* od společnosti *OpenAI*, který byl na počátku roku 2023 nejrozšířenějším jazykovým modelem na světě, byl trénován pomocí tréninkové sady obsahující texty z knih,



Obrázek 3: Ukázka vytváření deepfake videa záměnou obličejů narátorů pomocí programu FaceFusion

40 Viz například PITTHARDOVÁ, Veronika. *Muž, který zmizel. Clementise pověsili a vygumovali. Zbyla jen čepice*. IDnes.cz [online]. 2024 [cit. 2024-03-09]. Dostupné z: https://www.idnes.cz/zpravy/domaci/slavne-fotografie-retus-cenzura-unor-1948-vlado-clementis-gottwald.A240221_063521_domaci_pitt

41 Viz kapitola: 3.3.4 Zpracování přirozeného jazyka.

42 JAILLANT, Lise a Arran REES. *Applying AI to digital archives: trust, collaboration and shared professional ethics*. *Digital Scholarship in the Humanities* [online]. 2023, 2023-06-01, 38(2), 571-585 [cit. 2024-03-17]. ISSN 2055-7671. Dostupné z: doi:10.1093/lc/fqac073, s. 580.

webových stránek, Wikipedie a dalších online zdrojů, přičemž celkový počet slov dosahuje více než 300 miliard.⁴³ Jeho následující verze (*GPT-4*) již byla uzavřená a přesné informace o rozsahu datové sady nebyly společností *OpenAI* zveřejněny. Je ale pravděpodobné, že vzhledem k technologickému pokroku a vysokým finančním investicím jde oproti předchozí verzi o mnohonásobné množství tréninkových dat.

3.3.3 Budoucnost AI v orální historii

Orální historik Michael Frisch již před sedmnácti lety předpověděl možné budoucí využití moderních technologií pro archivní materiály orální historie, podle něhož by každý mohl vytvářet vlastní *dokumentární filmy* z video či audio nahrávek.⁴⁴ Tato možnost je již nyní možná. AI služeb, které dokáží převést text na digitálně vytvořené video, existuje celá řada. Této službě tak postačí pouze poskytnout například přepis rozhovoru a hraný či animovaný film, který by vizualizoval témata probíraná v rozhovoru, je vygenerován během okamžiku. Služeb pro generování videa z textového zadání již existuje několik, ale patrně nejimpozantnější výsledky doposud vytvořil model *Sora* od společnosti *OpenAI*. *Sora* prozatím ještě není otevřená služba a nelze ji testovat, ale ukázka výsledku promptu⁴⁵ na vytvoření minutového klipu ženy procházející se neony osvětlenými ulicemi Tokia⁴⁶ vypadá impozantně a od reálného videa ji lze jen těžko odlišit.

43 HUGHES, Alex. BBC Science Focus. *ChatGPT: Everything you need to know about OpenAI's GPT-4 upgrade* [online]. 2023 [cit. 2024-01-06]. Dostupné z: <https://www.sciencefocus.com/future-technology/gpt-3>

44 FRISCH, Michael. Oral history and the digital revolution: Toward a post-documentary sensibility. In: PERKS, Robert a THOMPSON, Alistair (eds.). *The Oral History Reader*. 2nd ed. London, UK: Routledge, 2006, s. 102–114.

45 Prompt označuje instrukci, kterou uživatel zadává programu k vykonání.

46 Dostupné z: <https://openai.com/sora>



Obrázek 4: „A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. She walks confidently and casually. The street is damp and reflective, creating a mirror effect of the colorful lights. Many pedestrians walk about.“, prompt, Sora, OpenAI, <https://openai.com/sora>



Obrázek 5: „Beaten students, Velvet revolution 1989 National street Czechoslovakia Prague“, prompt, Leonardo.Ai, Leonardo interactive, 6.1.2024, <https://leonardo.ai/>

Přestože na vytváření realistických videí si budeme muset ještě chvíli počkat, jiné oblasti generativní AI, která se využívá k tvorbě textů, obrázků a fotografií, jsou k dispozici již teď v poměrně dobré kvalitě. Pro generování obrázků si služby jako DALL-E, Midjourney, Leonardo.Ai nebo Adobe Firefly získaly mezi uživateli oblibu. Jak je patrné z přiložených ilustračních příkladů, nejsou tyto nástroje ještě zcela dokonalé a z obrázků zachycujících historické události jsou nedokonalosti v dobových detailech na první pohled poměrně patrné.



Obrázek 6: „Sametová revoluce v Československu 1989“, prompt, DALL-E, verze 3, OpenAI, 6.1.2024, <https://labs.openai.com/>



Obrázek 7: „Velvet revolution 1989 Václav Havel Czechoslovakia Prague“, prompt, Leonardo.Ai, Leonardo interactive, 6.1.2024, <https://leonardo.ai/>

Při pokusu o vytvoření scén z událostí listopadu 1989 v Československu bylo nezbytné vyzkoušet více AI nástrojů a provést řadu úprav, aby se postupně odstranily alespoň některé anachronismy, které byly do fotografií zakomponovány, jako jsou nevhodné oděvy, nesprávné vyobrazení uniforem a výzbroje zasahujících příslušníků Státní bezpečnosti nebo státních vlajek vlajících nad hlavami protestujících studentů. Přestože konečné výsledky stále nejsou zcela bezchybné, jsou již poměrně zdařilé a schopnost těchto nástrojů zobrazovat realistická a přesná vizuální díla se neustále zlepšuje. Další zlepšování lze v budoucnu očekávat i od dalších AI služeb, které mohou být užitečné pro potřeby orální historie a které jsou prozkoumávány v následujících kapitolách. U automatické transkripce rozhovorů lze očekávat zlepšení přesnosti přepisu, stejně jako kvality překladu u strojových překladačů. Zlepšení lze také očekávat u analytických nástrojů určených pro analýzu a interpretaci textů, sentimentu, emocí či nonverbální komunikace. Dalším zajímavou invencí může být automatizace archivních systémů, které by mohly inteligentně organizovat, kategorizovat a efektivně spravovat velké objemy orálně-historických dat, což by přispělo k jednoduššímu, rychlejšímu a efektivnějšímu procesu archivace a snadnějšímu zpřístupnění archivních záznamů širší veřejnosti.

Jednou z fascinujících možností může být využití AI k virtuální rekonstrukci historických postav, což by umožnilo provádět s nimi interaktivní rozhovory. To by mohlo otevřít nové cesty pro výuku a výzkum. Představme si například mobilní aplikaci, nebo rovnou software pro virtuální realitu, který by dokázal ze všech dostupných historických pramenů (rozhovorů, videozáznamů, osobních dopisů, deníků, zápisků a dalších ego-dokumentů) vytvořit realistickou imitaci významné (i nevýznamné) postavy třeba i z dávné minulosti.



Obrázek 8: „Náhled mobilní aplikace s virtuální rekonstrukcí historických postav, Tomáš Garrigue Masaryk“, prompt, DALL-E, verze 3, OpenAI, 6.1.2024, <https://labs.openai.com/>

Ještě donedávna se jednalo pouze o námět vědeckofantastických románů. Nyní je to vize velmi blízké budoucnosti.

Nejen začínající orální historici by se tak například mohli na těchto virtuálních avatarech trénovat ve schopnosti vedení rozhovoru a v přípravě otázek. Taková technologie by jistě mohla najít uplatnění i ve výuce dějin na školách i pro další edukační účely.

3.3.4 Zpracování přirozeného jazyka

Obor umělá inteligence, navzdory rozšířenému povědomí, nezahrnuje pouze ChatGPT a strojový překlad jazyků, ale patří sem i jiné technologie a metodiky, včetně strojového učení, robotiky, počítačového vidění⁴⁷, expertních systémů, ovládání autonomních vozidel, lékařské diagnostiky a mnohých dalších.⁴⁸ Uplatnění tak nachází v nejrůznějších oborech od organizování výroby, přes finančníctví až ke zdravotnictví nebo dopravě. Pro potřeby orální historie (obecně humanitních věd) je klíčová především jedna subdisciplína AI s názvem „počítačové zpracování přirozeného jazyka“ (dále jen „NLP“⁴⁹).

Obor NLP se zaměřuje na interakci mezi počítači a lidským jazykem. NLP není překvapivě žádnou novinkou a je vyvíjen již desítky let. Z důvodu značných nároků na výpočetní výkon ale až v posledních letech začalo být možné využívat jej i jinde než pro vědecké účely na akademické půdě.

NLP představuje most mezi lingvistikou a informatikou a jeho cílem je umožnit počítačům rozpoznat, porozumět, interpretovat a generovat lidský jazyk tak, jak je používán ve skutečné komunikaci. To zahrnuje řadu úkolů, jako je rozpoznávání řeči, analýza sentimentu, automatické generování textu, strojový překlad a rozpoznávání a zodpovídání otázek⁵⁰. Tuto technologii využívají jak jazykové překladače, nástroje pro přepis rozhovoru tak i zmíněný ChatGPT. V rámci orální historie lze NLP využívat například k automatickému přepisu nahrávek, identifikaci klíčových témat a osob v rozhovorech, analýzu sentimentu a emocí z orálně-historických rozhovorů, a jinak zpracovávat a analyzovat velké objemy dat.

47 Počítačové vidění (computer vision) je obor zabývající se zpracováním a analýzou digitálního obrazu. Viz například *Počítačové vidění*. In: Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2024-02-27]. Dostupné z: https://cs.wikipedia.org/wiki/Počítačové_vidění

48 COPELAND, B. J. *Artificial intelligence*. Encyclopedia Britannica [online]. 2024 [cit. 2024-02-04]. Dostupné z: <https://www.britannica.com/technology/artificial-intelligence>.

49 NLP je zkratka z anglického „Natural language processing“. Nezaměňovat s psychologickou metodou „neurolingvistické programování“, která používá v češtině i angličtině totožnou zkratku.

50 JENTSCH, Patrick a Stephan PORADA. *From Text to Data: Digitization, Text Analysis and Corpus Linguistics*. In: Digital Methods in the Humanities: Challenges, Ideas, Perspectives [online]. 2021, s. 89–128 [cit. 2024-03-10], s. 116.

Aby byl NLP schopen pracovat, využívá množství různorodých algoritmů, které mu umožňují rozpoznat lidskou řeč a mluvené slovo převést na text. Další algoritmy jsou pak nutné k provedení syntaktické a sémantické analýzy a analýzy sentimentu textu, tedy k rozložení vět na jednotlivé gramatické segmenty, určení významů slov a frází v daném kontextu a nalezení a třídění informací v textu.⁵¹

Pro práci s NLP se používají nejrůznější softwarové knihovny. Mezi nejpopulárnější a nejpoužívanější patří SpaCy a NLTK⁵² (dále např. PyTorch, TextBlok, Polyglot). Většina výše zmíněných knihoven má mnoho společných vlastností. Jedná se o „open-source“⁵³ knihovny, které jsou napsány v programovacím jazyce Python. Vykonání kódu probíhá přímo na lokálním stroji. Data se nikam neodesílají, neposkytují se žádné třetí straně a je tedy bezpečné takto pracovat i s citlivými údaji. Díky nízkým nárokům na výkon lze rychle zpracovat i poměrně velké objemy dat na běžném osobním počítači nebo laptopu. Práce s knihovnami je relativně snadná a dostupná i pro osoby bez programátorských dovedností.

Každá z knihoven poskytuje čtyři základní funkce: tokenizaci (rozdělení na jednotlivá slova nebo věty), lemmatizaci (převedení slova do základního tvaru⁵⁴), větný rozbor a rozpoznávání pojmenovaných entit (například jména osob, názvy organizací, data).⁵⁵ Mezi nadstavbové funkce, obsažené pouze v některých ze zmíněných knihoven, patří například překlad do jiného jazyka nebo analýza sentimentu z textu. Pro ukázkou práce s NLP zde poslouží softwarová knihovna SpaCy⁵⁶, která je moderní, velmi rychlá, přesná a disponuje velkým množstvím funkcí.

SpaCy disponuje podporou více než 50 jazyků. Mezi podporované jazyky patří i čeština, nicméně prozatím bohužel pouze pro základní funkce. Proto zde uvádím ukázkou zpracování anglicky psaného textu. Obdobně jako projekt *100 studentských revolucí* se i tento text věnuje událostem

51 Tamtéž, s. 117.

52 Viz například SOFFOS AI. *Top 8 Python Libraries For Natural Language Processing (NLP) in 2023*. Medium [online]. 2023 [cit. 2024-03-03]. Dostupné z: <https://medium.com/@soffosdotai/top-8-python-libraries-for-natural-language-processing-nlp-in-2023-5963bfa53296>., nebo KOZACZKO, Dominik. *9 Best Python Natural Language Processing (NLP) Libraries*. Sunscrapers [online]. 2023 [cit. 2024-03-10]. Dostupné z: <https://sunscrapers.com/blog/9-best-python-natural-language-processing-nlp/>

53 Termín open-source označuje software, jehož zdrojový kód je veřejný a dovoluje-li to jeho licence, je i volně šiřitelný.

54 Například pro počítání výskytu slova v textu je nezbytné převést všechna slova na základní tvar. Například u slov jako pracoval, pracují, nebo pracuj by se každé ze slov počítalo zvlášť. Lemmatizovaným tvarem je pro slovesa v češtině infinitiv v přítomném čase (pracovat).

55 JENTSCH, Patrick a Stephan PORADA. *From Text to Data: Digitization, Text Analysis and Corpus Linguistics*, s. 117.

56 Podrobnější informace jsou dostupné na oficiálních webových stránkách: <https://spacy.io/>

17. listopadu 1989. Jedná se přepis rozhovoru⁵⁷ s *Monikou MacDonagh-Pajerovou*, který v roce 2017 poskytla pro *Radio Prague International*.

Nejprve je nezbytné nainstalovat do počítače pomocí příkazové řádky⁵⁸ potřebné nástroje.

1. `pip install -U pip setuptools wheel`
2. `pip install -U spacy`
3. `pip install -U textblob`
4. `python -m spacy download en_core_web_sm`

Po jejich instalaci je možné spustit následující Python skript.⁵⁹ Pro lepší přehlednost je barevně zvýrazněna syntaxe kódu. Komentáře jsou uvozeny symbolem „#“, jsou zvýrazněny šedou barvou a slouží pouze pro popis kódu. Klíčová slova, rezervovaná programovacím jazykem Python, jsou vyznačena modře. Zeleně jsou pak zvýrazněny názvy modulů, proměnných a funkcí. Hodnoty, jako jsou textové řetězce a čísla, jsou vyznačeny červenou barvou. Pokud je před exekucí skriptu nutné některou část kódu upravit (změnit název souboru, umístění složky, datového formátu souboru), je tato část kódu žlutě podbarvena.

```
1. import spacy
2. from textblob import TextBlob
3. from collections import Counter, defaultdict
4.
5. # Načtení jazykového (anglického) modelu
6. nlp = spacy.load("en_core_web_sm")
7.
8. # Načtení textového souboru s přepisem rozhovoru
9. with open('/monika-macdonagh-pajerova-how-the-spirit-of-the-velvet-revolution-turned-to-cynicism.txt', 'r', encoding='utf-8') as file:
10.     text = file.read()
11.
12. # Zpracování textu
13. doc = nlp(text)
14.
15. def categorize_value(value):
16.     if value > 0.33:
17.         return "pozitivní"
18.     elif value < -0.33:
19.         return "negativní"
20.     else:
21.         return "neutrální"
22.
23. category_dict = defaultdict(set)
24. sentiments = []
```

57 Rozhovor je dostupný z: <https://english.radio.cz/monika-macdonagh-pajerova-how-spirit-velvet-revolution-turned-cynicism-8176460>.

58 Česky psaný návod, jak pracovat s příkazovou řádkou, je dostupný například zde: <https://nauce.python.cz/lessons/beginners/cmdline/>

59 Výrazem *Python skript* je zde myšlen jednoduchý algoritmus, který je napsán v programovacím jazyku Python a který vykonává určitý úkol.

```

25.
26. # Zpracování každé věty
27. for sentence in doc.sents:
28.     # Identifikace pojmenovaných entit
29.     for entity in sentence.ents:
30.         category_dict[entity.label_].add(entity.text)
31.
32.     # Určení sentimentu věty
33.     sentiments.append(categorize_value(TextBlob(sentence.text).sentiment.polarity))
34.
35. # Vypsání pojmenovaných entit podle kategorií
36. for category, names in category_dict.items():
37.     print(f"{category}: {' '.join(sorted(names))}")
38.
39. # Vypsání statistiky sentimentu
40. print(dict(Counter(sentiments)))

```

Vykonáním kódu získáme dva textové výsledky. První (řádek č. 37.) vypíše seznam pojmenovaných entit, kterými jsou například:

- vlastní jména (*Alexander Dubcek, Dubcek, Gorbachev, Hilgenberg, Jacques Rupnik, Miloš Jakeš, Jan Opletal, Jesus Christ, Jiri Dienstbier, Josef Vavrousek, Klaus, Marek Benda, Marian Calfa, Mikhail Gorbachev, Misha Glenny, Timothy Garton Ash, Vaclav Havel, Vaclav Klaus, Viktor Orban, Walter Komarek*),
- názvy geografických míst (*Czechoslovakia, Dresden, East Germany, Halle, Hungary, Leipzig, Liberec, Petrova, Poland, Prague, Russia, Western Germany*),
- data (*1939 to 1989, 1968, 1989, 1992, 2 or 3 years before, 50 years, 68, 85, 86, November 1989, a year, several years, spring 1990, spring 68, that spring, 1990, the 17th of November, the 50 years, the 90s, the wild 90s, today, two previous years, two, three years*)
- a mnoho dalších názvů organizací, událostí apod.

Druhým výstupem (řádek č. 40) je statistika sentimentu, podle které rozhovor obsahoval celkem 149 vět, z čehož bylo identifikováno 130 vět neutrálních, 14 vět pozitivních⁶⁰ a 5 negativních⁶¹.

Ačkoliv je tento návod tedy prozatím užitečný pouze pro anglicky psané texty, plná podpora pro češtinu může být doplněna v dohledné době. Anglický model pak stačí pouze zaměnit za český a zbytek kódu bude beze změny fungovat naprosto totožně. Jak postupovat při analýze česky psaného textu pomocí NLP je uvedeno v závěru práce v rámci případové studie⁶² na projektu *100 studentských revolucí*.

60 Například věty: They are great people, absolutely great people. | And there was this sort of happy time.

61 Například věty: It was all a secret police. | But it was so difficult, it was almost impossible.

62 Viz kapitola: 4.6 Případová studie: 100 studentských revolucí.

4 Empirická část: moderní technologie a možnosti jejich využití v jednotlivých fázích orálně-historického výzkumu

4.1 Příprava rozhovoru

4.1.1 Rešerše

Ve fázi přípravy na orálně-historický rozhovor lze pro usnadnění práce použít nástroje, které sice nejsou určeny výhradně pro potřeby orální historie, ale svojí univerzalitou jsou i zde užitečnou pomůckou. Využít lze například takzvané chatboty⁶³, jako jsou například již zmíněné a veřejně dobře známé služby ChatGPT od společnosti OpenAI, Google Gemini, Claude 3 nebo Mistral, které si získaly mezi veřejností obrovskou popularitu. Na začátku roku 2024 používá službu ChatGPT již více než 180 milionů aktivních uživatelů.⁶⁴ Tyto nástroje mohou být významnou oporou při přípravě na orálně-historické rozhovory, poskytují prostor pro konzultaci či brainstorming⁶⁵.

Novinář a analytik IT médií Petr Koubský uvádí ve své přednášce následující příklad, který, dle mého názoru, velmi dobře popisuje, jakým způsobem je možné chatboty efektivně používat:

„Prosímte, mám za úkol napsat takovýto článek. Hoď mi asi do sedmi bodů nějakou osnovu, co by v něm všechno mělo být. On mi hodí osnovu do sedmi bodů. Já si je přečtu. Čtyři z těch věcí jsou triviality, na které bych okamžitě přišel sám, jinak tam nemám co dělat na té židli (pozn. myšleno novinářské židli). Dvě jsou úplné blbosti, které nesmím použít, jinak budu za naprostého troubu. Ta sedmá věc, to je ta, o kterou jde. To je netriviální nápad, na který bych možná přišel a možná nepřišel, to nevím, ale rozhodně ne hned. Trvalo by mi to. A kvůli tomu sedmému bodu to dělám. Ten mi něco rozsvítí v hlavě, možná ho ani nepoužiju, ale někam mě posune. Přiměje mě nějakým způsobem přemýšlet.“⁶⁶

S AI lze diskutovat zaměření projektu, výzkumné cíle, témata rozhovorů i navrhované otázky. Avšak vždy je třeba mít na paměti, že na AI generované odpovědi se nikdy nelze zcela spolehnout, jelikož mohou produkovat nepřesné nebo zavádějící informace, a tak je potřeba k nim přistupovat kriticky. Pro tuto vlastnost se v médiích rozšířil nepřesný termín halucinovat⁶⁷ a logicky vyplývá ze samotného principu, jak velké jazykové modely fungují, jelikož model pouze hledá statisticky

63 Chatbot je počítačový program určený ke komunikaci mezi uživatelem a počítačem. Namísto programovacího jazyka nebo grafického uživatelského prostředí se ale pro komunikaci používá textová konzole a přirozený jazyk.

64 Viz například <https://www.demandsage.com/chatgpt-statistics/>

65 Technika pro generování nápadů. Standardně je provozována ve skupině lidí, ale namísto lidí lze použít i chatbota.

66 LLionTV. Petr Koubský: Umělá inteligence: Příslib i hrozba (MFF-PMF 18.1.2024) [01:10:51-01:11:37]. In: Youtube [online]. 2024 [cit. 2024-02-27]. Dostupné z: <https://www.youtube.com/watch?v=LwJYjkctdyU>

67 Viz například <https://www.unite.ai/cs/co-jsou-llm-halucinace-vyvolava-etické-obavy-prevence/>

nejpravděpodobnější slovo, které bude v daném kontextu následovat. Nemá-li o daném tématu dostatek informací, dokáže s přesvědčivostí informace vymýšlet a svá tvrzení podpořit odkazováním na vymyšlené autority a jejich vymyšlené publikace. Nicméně, s kritickým přístupem a vhodnou metodikou práce mohou být chatboty cennými pomocníky.

V orální historii jsou s narátorem zpravidla vedeny dva samostatné rozhovory⁶⁸, přičemž prvním z nich je zpravidla životopisné vyprávění. Při přípravě na druhý rozhovor může být takový chatbot velmi užitečný. Dokáže pomoci identifikovat klíčová témata, ke kterým by bylo vhodné se vrátit, nebo odhalit nekonzistence, které naznačují prostor pro další zkoumání. Může také asistovat v návrhu otázek na následující rozhovor, upravit jejich pořadí, aby na sebe tematicky navazovaly. Přestože chatboty nemohou zkušenosti a schopnosti zkušených orálních historiků plně nahradit, mohou přinést užitek v situacích, kdy je někdo takřikajíc slepý k tomu, co mu leží přímo před očima, nebo když takzvaně pro stromy nevidí les. Mohou i nabídnout podporu při identifikaci témat, která si zaslouží další pozornost a také upozornit na potenciálně emocionálně náročná témata a případně varovat před možností vyvolání vzpomínek na traumatické události, varovat před pokládáním otázek, které by mohly narušit rozhovor a eventuálně i vztah mezi tazatelem a narátorem.

4.1.2 Simulace rozhovoru

Chatboty nabízejí zajímavou možnost simulovat rozhovory pro trénink nebo přípravu, umožňující tazatelům získat představu o tom, jak by skutečný rozhovor s danou osobou mohl probíhat. Výzkumníci mohou chatbotům poskytnout životopisné údaje a přepisy předchozích rozhovorů, aby chatbot mohl lépe reagovat v duchu dané osoby. Zde je ale opět vhodné upozornit na bezpečnostní rizika. Chatbotu nikdy nesmíme svěřit žádné citlivé informace, aby nedošlo k jejich zneužití. Pokud poskytovaná data obsahují citlivé informace, existují dvě řešení. Prvním je velmi pečlivá revize všech dat a případné odstranění citlivých údajů (jména, hesla, osobní údaje, citlivá témata) předtím, než jsou odeslány AI ke zpracování. V případě, že rozhovor obsahuje obzvláště citlivý obsah, doporučuje se jeho sdílení zcela vyhnout, aby se předešlo jakémukoli riziku úniku nebo zneužití těchto údajů. Druhou možností je využití open-source softwarového řešení, které je možné spustit na vlastním hardwaru, funguje bez připojení k internetu a data se nikam neodesílají. Takovou open-source alternativou je například AI chatbot od společnosti NVIDIA⁶⁹ nebo projekt GPT4All⁷⁰.

68 Viz VANĚK, Miroslav a MŮCKE, Pavel. *Třetí strana trojúhelníku: teorie a praxe orální historie*, s. 166.

69 Viz například CLAYTON, Jesse. *Custom Chatbot to NVIDIA RTX AI PCs*. NVIDIA Blog [online]. 2024 [cit. 2024-03-10]. Dostupné z: <https://blogs.nvidia.com/blog/chat-with-rtx-available-now/>

70 GPT4All umožňuje lokální provozování služby ChatGPT, která ale i tak bez internetového připojení nefunguje a stále odesílá data do sítě. Kromě toho je ovšem možné používat mnoho alternativních jazykových modelů (Mistral, Hermes, Wizard, Orca), které fungují pouze lokálně a offline. Dostupné z: <https://gpt4all.io/>

Přípravu před rozhovorem budu ilustrovat na rozhovorech s panem Radkem Kotlabou, který v rámci projektu *100 studentských revolucí* odpovídal na tři otázky, které byly v projektu položeny každému z narátorů.

- *Co se vám vybaví, když se řekne listopad 1989?*
- *Jak vnímáte současnou mladou generaci?*
- *Co považujete za své životní mezníky?*

Použitím nástrojů umělé inteligence se můžeme pokusit předem simulovat možné odpovědi pana Kotlaby na položené otázky, což nám umožní předem identifikovat a řešit potenciální problematická témata. Tento proces nám dává jedinečnou možnost předem ověřit, zda se probíraná témata nemohou ukázat jako kontroverzní, citlivá, nebo dokonce vyvolávat u vypravěče nepříjemné či traumatické reakce. Souběžně s tím nám AI může pomoci ověřit, že otázky nejsou povrchní či triviální anebo zda naopak podporují směřování rozhovoru k hlubšímu dialogu.

Pro přípravu před rozhovorem bude využit chatbot ChatGPT-4, který je instruován, aby podle dostupných dat co nejpřesněji imitoval pana Radka Kotlabu. Chatbot potvrzuje, že o panu Kotlabovi zatím nemá zcela žádné informace a není tak zatížen žádnými nepřesnostmi, předsudky, negativními informacemi a podobně. Chatbot obdrží krátký životopis narátora, který je dostupný přímo z oficiálních zdrojů projektu *100 studentských revolucí*⁷¹, avšak bez přístupu k přepisům rozhovorů, aby se zabránilo možnému zkreslení. Následně mu je poskytnut ještě detailnější životopis a profil narátora, který byl sestaven v rámci projektu *Paměť národa*, pro který pan Kotlaba také poskytl rozhovory.⁷² Více dat umožní modelu se lépe ztotožnit s postavou a poskytnout autentičtější odpovědi.

Nyní, když je AI model připravený a informovaný o kontextu a postavě pana Kotlaby, můžeme mu položit předem připravené otázky. Na otázku týkající se listopadu 1989 poskytl model z pozice pana Kotlaby následující odpověď:

„Když se řekne listopad 1989, vybaví se mi především pocit naděje a změny. Byla to doba, kdy se společně s mými spolužáky a učiteli na Pedagogické fakultě v Ústí nad Labem náhle otevřela možnost aktivně se zapojit do dění a vyjádřit nesouhlas s režimem, který nás celá léta omezoval. Pocit sounáležitosti, když jsme se shromažďovali na demonstracích, a vzrušení z možnosti přispět k demokratickým změnám, které se zdály být na dosah. Tyto události výrazně ovlivnily můj další životní a profesní směr, a

71 Stručný životopis narátora z projektu *100 studentských revolucí*. Dostupné z: <http://www.studenti89.usd.cas.cz/studenti/kotlaba-radek/>

72 Podrobný životopis a rozhovor z projektu *Paměť národa*. Dostupné z: <https://www.pametnaroda.cz/cs/kotlaba-radek-20170515-0>

přinesly mi uvědomění síly občanské angažovanosti a důležitosti stát se aktivní součástí společnosti.“

Když porovnáme simulovanou odpověď se skutečnou odpovědí, kterou na tuto otázku odpověděl pan Kotlaba, zjistíme určité rozdíly, což bylo očekávané vzhledem k obecné povaze otázky. Pan Kotlaba se soustředil spíše na specifické, ikonické momenty, jako jsou demonstrace, klíče, Národní třída a datum 17. listopadu, zatímco simulovaná odpověď byla zaměřena obecněji na pocit naděje a změny a dopad těchto událostí na život narátora. Přesto se v obou odpovědích objevují společné prvky, jako jsou zmínky o škole, kolejích a atmosféře okupační stávky, což ukazuje na částečnou shodu mezi predikcí modelu a skutečnou reakcí pana Kotlaby.

Odpověď na druhou otázku zde bude přeskočena, jelikož v průběhu skutečného rozhovoru byla formulována lehce odlišně, se zaměřením specificky na budoucnost dcery narátora, namísto obecného vnímání současné mladé generace. Na poslední otázku týkající se životních mezníků pak chatbot reagoval následujícím způsobem:

„Nejvýznamnějšími mezníky v mém životě jsou bezpochyby mé aktivní zapojení do listopadové revoluce v roce 1989, což byl pro mě moment hlubokého osobního i společenského obratu. Dalším důležitým mezníkem bylo zahájení mé pedagogické kariéry, která mi umožnila sdílet své znalosti a hodnoty s mladými lidmi. Neméně důležitá je pro mě i moje role v divadle, která mi poskytuje kreativní vyústění a umožňuje mi spojovat se s komunitou na velmi osobní úrovni. Tyto momenty odrážejí mé hlavní životní zájmy a vášně a formují mé chápání světa a mého místa v něm.“⁷³

Obě odpovědi na dotaz týkající se životních mezníků ukazují na důležité okamžiky, avšak z odlišných úhlů pohledu. Společným bodem je zmínka o začátku pedagogické dráhy, která je v obou případech určena jako jeden z klíčových životních mezníků. V reálné reakci pana Kotlaby se však větší pozornost věnuje rodinnému životu a víře, zatímco chatbotem zprostředkovaná simulace se více soustředí na listopadovou revoluci a zapojení do divadla.

Simulovaný rozhovor nabízí poměrně užitečný náhled na možný průběh skutečného rozhovoru a naznačuje odpovědi, které bychom mohli očekávat, což může poskytovat poměrně dobrou přípravu na rozhovor. Důležitým prvkem je v tomto případě detailnost a rozsah informací poskytnutých AI modelu, které umožňují přesnější a věrohodnější simulaci odpovědí. Přesto je samozřejmě třeba mít na paměti, že lidské chování a reakce nikdy nelze zcela přesně předpovídat. V tomto kontextu je simulace rozhovoru užitečným nástrojem, avšak nikdy nemůže nahradit hloubku reálného rozhovoru. Z praktického hlediska je také zřejmé, že příprava takové simulace zabere nějaký čas a

73 ChatGPT, verze 4, OpenAI, 6.1.2024, <https://chat.openai.com/>

nelze, a rozhodně ani není nutné, dělat ji před každým rozhovorem. Ovšem před obzvláště citlivým, důležitým, nebo kontroverzním rozhovorem by to stálo za úvahu.

Před provedením skutečného rozhovoru je možné ještě využít chatbot k posouzení vhodnosti témat a otázek, které plánujeme položit. Opět jako ve všech předchozích případech platí, že AI model je nespolehlivý, dělá chyby a jeho výstupy je třeba pečlivě kriticky posuzovat. Chatbot by měl sloužit primárně jen jako nástroj pro generování nápadů a získávání inspirace. Odpověď chatbotu na zhodnocení výběru tematických okruhů pak vypadá následovně:

„Listopad 1989 – Tato otázka je sice historicky významná, ale neměla by být kontroverzní, pokud bude pan Kotlaba ochoten mluvit o svých zkušenostech a pocitech z té doby. Je důležité umožnit vyprávějíci sdílet jeho osobní perspektivu bez nutnosti hodnotit události, které se staly.

Současná mladá generace – Tato otázka může vyvolat širokou škálu názorů. Je vhodné ji položit tak, aby pan Kotlaba mohl vyjádřit svůj názor na pozitivní i možná negativní aspekty současné generace bez toho, aby se cítil nutný zaujímat konkrétní postoj.

Životní mezníky – Tato otázka je velmi osobní a měla by být formulována s respektem k osobnímu životu a zkušenostem pana Kotlaby. Je důležité nechat prostor pro to, aby mohl vybrat a sdílet ty momenty, které považuje sám za klíčové, aniž by se cítil tláčen do rozdělení něčeho, co by si raději nechal pro sebe.

Pokud máte obavy, že by se pan Kotlaba mohl některým otázkám vyhnout, je dobré mít připravené doplňkové otázky, které mohou téma rozšířit nebo nabídnout alternativní úhel pohledu. Je také užitečné předem vysvětlit účel rozhovoru a zajistit, aby se vyprávějíci cítil pohodlně se sdílením svých myšlenek a zkušeností.“⁷⁴

4.2 Terénní výzkum

Během samotného nahrávání rozhovoru příliš AI nástrojů využít nelze. Výzkumník by se měl plně soustředit na probíhající rozhovor. Pro kvalitní nahrávku rozhovoru je nejlepší volbou profesionální nahrávací zařízení (popř. kamera v případě videonahrávky). Často však postačí i jednoduchý diktafon a v posledních letech i chytrý telefon, který je vždy po ruce a zároveň díky široké škále dostupných aplikací poskytuje i řadu funkcí, prospěšných pro potřeby natáčení orálně-historických rozhovorů. Jednotlivé aplikace se liší především v ceně. Některé jsou zcela zdarma, některé placené a některé poskytují základní funkce zdarma a nadstandardní za poplatek. Některé mohou obsahovat reklamy, jiné ne. Dále se odlišují v tom, pro kterou ze dvou dominantních mobilních platform jsou určeny, zda pro Android či iOS nebo jsou nabízeny pro obě platformy zároveň. Dále je třeba zvážit i jejich unikátní vlastnosti, které mohou být právě pro potřeby orální historie velmi prospěšné.

74 ChatGPT, verze 4, OpenAI, 6.1.2024, <https://chat.openai.com/>

Následující rešerše neslouží jako kompletní průvodce všemi dostupnými aplikacemi pro nahrávání rozhovorů, jelikož těch existují stovky a tisíce. Seznam mobilních aplikací pro rešerši byl sestaven na základě seznamu pěti aplikací⁷⁵ pro nahrávání hlasu vhodných pro orálně-historické rozhovory.⁷⁶

Aplikace *Rev* je zdarma a je určená pro obě mobilní platformy. Umožňuje stříh nahraného audio záznamu přímo v aplikaci a sdílení souborů prostřednictvím mnoha různých způsobů. Touto aplikací nelze nahrávat telefonní hovory, což je relativně omezující. Orálně-historické rozhovory vedené po telefonu byly až do nedávné doby poměrně ojedinělé. Avšak s propuknutím koronavirové pandemie na začátku roku 2020 se situace změnila a obavy z šíření nákazy narušily i formu nahrávání orálně-historických rozhovorů. Rozhovor vedený prostřednictvím telefonního hovoru se tak stal poměrně běžnou záležitostí.⁷⁷ Během nahrávání audio záznamu prostřednictvím aplikace *Rev* lze zároveň používat i jiné mobilní aplikace. Pro potřeby orálních historiků poskytuje zajímavou funkci, a sice možnost prostřednictvím aplikace přímo objednat provedení přepisu rozhovoru, u kterého je garantována 90% přesnost v případě automatického strojového přepisu a až 99% přesnost v případě manuálního přepisu provedeného člověkem. Služba je to jistě zajímavá, ale pro českého výzkumníka neúčinná, jelikož manuální přepis je dostupný pouze pro angličtinu, a s cenou od \$ 1,5 za každou minutu nahrávky i poměrně drahá.⁷⁸

Android aplikace *Smart Recorder* je zdarma, snadná pro ovládání, navržena pro nahrávání dlouhých rozhovorů ve vysoké kvalitě. Během nahrávání lze také používat jiné aplikace a nelze nahrávat telefonní hovory. Zajímavou funkcí je možnost automatického sestříhání tichých pasáží v nahrávce.⁷⁹ To může být užitečné například pro nahrávání pracovních porad, ale pro orální historiky, pro které je i ticho podnětem k interpretaci, není tato funkce příliš vhodná.

Další aplikací je *TapeACall*, která na rozdíl od dvou výše zmíněných umí nahrávat telefonní hovory a je to i její primární účel. Aplikace je dostupná pro obě mobilní platformy, její stažení je zdarma, ale pro používání je vyžadována platba pravidelného měsíčního poplatku. Po ukončení hovoru je také možné prohlédnout automaticky vygenerovaný přepis rozhovoru. Kromě toho žádné mimořádné funkce nenabízí.⁸⁰

75 Jedná se o aplikace *Rev*, *Smart Recorder*, *TapeACall*, *Voice Memos* a *Voice Recorder*.

76 MORTON, Sunny Jane. *5 Voice-Recorder Apps for Oral History Interviews*. Family Tree Magazine [online]. 2021 [cit. 2024-01-13]. Dostupné z: <https://familytreemagazine.com/resources/apps/voice-recorder-apps/>

77 Viz HLAVÁČEK, Jiří. *Covido, ergo zoom! Metodologické, technické a etické aspekty distančně vedených rozhovorů během pandemie Covid-19*, s. 60.

78 Viz REV. *Record & Transcribe Interviews with Rev's Mobile App*. Rev [online]. [cit. 2024-01-13]. Dostupné z: <https://www.rev.com/apps-and-tools/voice-recorder>

79 MORTON, Sunny Jane. *5 Voice-Recorder Apps for Oral History Interviews*. Family Tree Magazine [online]. 2021 [cit. 2024-01-13]. Dostupné z: <https://familytreemagazine.com/resources/apps/voice-recorder-apps/>

80 Viz TAPEACALL. *TapeACall* [online]. [cit. 2024-01-13]. Dostupné z: <https://www.tapeacall.com/>

Aplikace určená pouze pro systém iOS s názvem *Voice Memos* je zdarma, pochází přímo z produkce společnosti Apple, nenahrává telefonní hovory, ale má kromě nahrávání vysoce kvalitního audio záznamu široké možnosti editace nahraných záznamů. Ty lze přímo v aplikaci stříhat, rozdělovat, vkládat do nich části jiných záznamů a podobně.⁸¹ Z důvodu zachování autenticity nahraného zvukového záznamu opět nejsou tyto funkce pro potřeby orální historie příliš vhodné.

Android aplikace *Voice Recorder* je zdarma, obsahuje reklamy, ale za jednorázový poplatek je možné reklamy z aplikace odstranit. Nahrávky lze snadno sdílet prostřednictvím mnoha různých aplikací. Neumožňuje nahrávat telefonní hovory.⁸²

Pokud je prioritou pořídit co možná nejkvalitnější nahrávku s čistým a srozumitelným zvukem, lze zde ještě zmínit aplikaci *Dolby On* od společnosti *Dolby*, která se zabývá výzkumem audiotechniky a kvalita audionahrávek je tak jejich prioritou. Softwarové nástroje této aplikace umožňují v nahrávce automaticky odstraňovat šum na pozadí, vyrovnávat hlasitost, zjemňovat zvukové přechody a podobně.⁸³ Pro čitelnost rozhovoru se rozhodně jedná o užitečné nástroje, ačkoliv i zvuky v pozadí a změny hlasitosti v řeči narátora sebou nesou informace, které mohou být v některých případech podstatné a jejich odstraněním či úpravou jsou tyto informace z nahrávky nenávratně odstraněny.

Tabulka 1: Přehled mobilních aplikací pro nahrávání rozhovorů

Aplikace	Platforma	Cena	Nahrávání tel. hovorů	Funguje na pozadí	Funkce navíc
Rev	Android/iOS	Zdarma	Ne	Ano	Přepis rozhovoru
Smart Recorder	Android/iOS	Zdarma	Ne	Ano	Sestřihání ticha
TypeACall	Android/iOS	Měsíční poplatek	Ano	Ano	Přepis rozhovoru
Voice Memos	iOS	Zdarma	Ne	Ano	
Voice Recorder	Android	Zdarma/ reklamy	Ne	Ne	
Dolby On	Android/iOS	Zdarma	Ano	Ano	Kvalitní zvuk

Mobilních aplikací pro nahrávání audio záznamů existují další stovky až tisíce. Žádná však není primárně určena pro účely orální historie, ani neposkytuje specifické funkce, které by pro tento obor

81 Viz APP STORE. *Voice Memos*. App Store [online]. [cit. 2024-01-13]. Dostupné z: <https://apps.apple.com/us/app/voice-memos/id1069512134>

82 MORTON, Sunny Jane. *5 Voice-Recorder Apps for Oral History Interviews*. Family Tree Magazine [online]. 2021 [cit. 2024-01-13]. Dostupné z: <https://familytreemagazine.com/resources/apps/voice-recorder-apps/>

83 Viz GOOGLE PLAY. *Dolby On: Record Audio & Music*. Apps on Google Play [online]. [cit. 2024-01-13]. Dostupné z: <https://play.google.com/store/apps/details?id=com.dolby.dolby>

mohly být během nahrávání rozhovoru výrazně prospěšné. Výjimkou je pouze automatický přepis rozhovoru, který je ale problematický v tom, že je ve většině případů placený a poměrně drahý. Zároveň je ale také omezený v rozsahu podporovaných jazyků a oproti manuálnímu přepisu velmi nepřesný a nespolehlivý. Proto, pokud nám pro specifické potřeby projektu nepřijde k užítku nějaká konkrétní doplňková funkce některé aplikace, bych při výběru aplikace kladl důraz výhradně na spolehlivost a na kvalitu nahrávky.

Například aplikace jako *Google Recorder* od společnosti *Google*, *Samsung Voice Recorder* od *Samsungu* nebo již dříve zmíněný *Voice Memos* od společnosti *Apple* jsou aplikace od spolehlivých vývojářů, které často bývají v mobilním telefonu již předinstalovány a pro pouhé nahrání audiozáznamu jsou plně dostačující. Zároveň jsou jednoduché na používání, spolehlivé, otestované, zdarma, bez reklam, a lze se spolehnout i na podporu ze strany jejich tvůrců v budoucnosti.

4.3 Zpracování dat

4.3.1 Vylepšení kvality audio záznamu

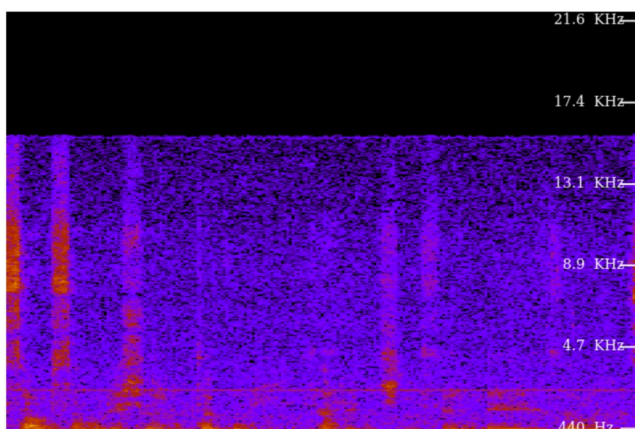
Nahrávání rozhovoru je pouze prvním krokem v procesu orálně-historického výzkumu. Před samotným přepisem rozhovoru, ať už prováděným manuálně nebo s pomocí umělé inteligence, je vhodné zvážit možnost dodatečných úprav audio nahrávky. Často se stává, že nahrávka obsahuje rušivé zvuky, jako jsou hlasy v pozadí, ruch ulice, nebo zvuky větru, které mohou ztížit její přepis a snížit kvalitu záznamu.

Pro vyčištění nahrávky lze využít řadu nástrojů. Například funkce *Enhance Speech*⁸⁴, která je součástí služby *Adobe Podcast AI*, efektivně odstraní nežádoucí zvuky na pozadí a zlepší srozumitelnost nahrávky. Použití aplikace je velmi jednoduché. Na webové stránce se stačí pouze zaregistrovat a pokud nahrávka splňuje určené limity, lze ji zde zdarma nahrát a obratem stáhnout novou verzi souboru bez rušivých zvuků v pozadí. Alternativou může být také populární, multiplatformní, open-source software *Audacity*, který lze zdarma a bezpečně provozovat na lokálním počítači a který nabízí mnoho možností pro nahrávání, analýzu a editaci zvukových záznamů.⁸⁵

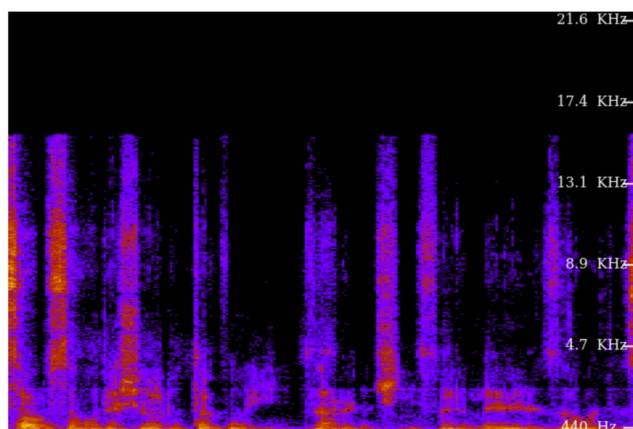
84 Podrobnější informace dostupné z: <https://podcast.adobe.com/enhance/>

85 Podrobnější informace dostupné z: <https://www.audacityteam.org/>

Výsledek této úpravy lze vizualizovat například pomocí nástroje *Spectrum Analyzer*⁸⁶, s jehož pomocí je možné porovnat patrné změny v audio spektru nahrávky před a po úpravě. Veškerá místa obsahující hlasový záznam byla zachována, zatímco veškerý šum okolí byl úspěšně odstraněn.



Obrázek 9: Původní zvukové spektrum rozhovoru



Obrázek 10: Upravené zvukové spektrum rozhovoru

Přestože pro některé výzkumné účely může být důležité zachovat v přepisu vady narátorova projevu nebo pauzy, které mohou být důležité pro následnou interpretaci, v mnoha situacích mohou být tyto prvky překážkou. Zvláště pokud se jedná o výrazné řečové vady nebo časté paralingvistické jevy, jako jsou mlaskání, povzdechy, nebo výplňová slova, mohou přepis značně komplikovat. V takovém případě je možné tyto prvky odstranit například pomocí nástrojů jako je *Cleanvoice AI*⁸⁷, který nabízí široké možnosti úprav audio záznamu. Tento nástroj umožňuje odstranit například výplňková slova, koktání, nepřírozené pauzy, zvuky úst, hlasité dýchání nebo rušivé pozadí, včetně hrající hudby. Kromě toho je schopen vyrovnat hlasitost mezi různými mluvčími a přizpůsobit nahrávku tak, aby byla co nejsrozumitelnější.

V případě video nahrávek nabízí *Cleanvoice AI* funkci, která tyto nežádoucí zvuky pouze ztlumí, ale neodstraňuje je ze záznamu, aby nedošlo k nesouladu mezi audio a video stopou. Tímto způsobem lze zajistit, že i přes odstranění rušivých prvků zůstává nahrávka synchronizovaná a původní atmosféra rozhovoru zůstává zachována.

4.3.2 Přepis rozhovorů

Automatizovaný přepis mluveného slova je jednou z klíčových aplikací AI v orální historii. Pomocí NLP technologií umožňujících rozpoznávání řeči mohou být video nahrávky i zvukové záznamy rychle a relativně přesně převedeny do textové podoby. Prozatím ovšem neexistuje žádný široce

86 Dostupné z: <https://academo.org/demos/spectrum-analyzer/>

87 Dostupné z: <https://cleanvoice.ai/>

používaný profesionální nástroj, který by byl určený primárně pro potřeby přepisu orálně-historických rozhovorů a který by plně vyhovoval specifickým potřebám oboru orální historie. S určitými kompromisy lze pro tento účel využít jiný z široké palety nástrojů pro přepis mluveného slova, které se běžně užívají pro více komerční účely. Lze například využít software pro tvorbu titulků pro podcasty, pro videa určená ke sdílení na sociálních sítích, nebo pro automatické zápisy pracovních schůzí velkých firem. Některé z těchto služeb poskytují kromě samotného přepisu i další funkce, které mohou být pro potřeby orálních historiků užitečné. Lze například na základě hlasové analýzy automaticky rozpoznat počet účastníků rozhovoru a na konci rozhovoru či pracovní schůzky stačí pouze přiřadit konkrétní jména k identifikovaným hlasům a automaticky tak označit, zda právě mluví tazatel, narátor či další osoba, která ať už úmyslně nebo omylem vstoupila do rozhovoru. Automaticky jsou také vygenerovány časové značky určující přesný čas, kdy někdo na nahrávce mluví. Po ukončení rozhovoru může být také automaticky vygenerován přehled rozhovoru, který mnohdy splňuje veškeré požadavky pro protokol o rozhovoru⁸⁸ používaný v orální historii. Tento automaticky generovaný přehled tak může protokol o rozhovoru plnohodnotně nahradit, nebo alespoň posloužit jako podklad pro jeho snadnější vytvoření, jelikož obsahuje základní fakta o rozhovoru jako je místo a datum nahrávání, délka trvání, seznam účastníků a s využitím umělé inteligence často již obsahuje také krátké shrnutí probíraných témat, seznam klíčových slov anebo například seznam všech pojmenovaných entit⁸⁹ (osobních jmen, názvů organizací, geografických jmen, čísel, dat apod.), které v rozhovoru zazněly.

Ohledně samotného automatického přepisu je nutné předem konstatovat, že technické možnosti nejsou ještě natolik vyspělé, aby byl automaticky generovaný přepis rozhovoru pro potřeby orální historie dostatečně přesný. Výsledky přepisu anglicky mluveného textu, na který jsou zpravidla velké jazykové modely primárně a nejlépe trénovány, je již na velmi dobré úrovni. Výsledky přepisu českého mluveného slova jsou z mnoha důvodů⁹⁰ o poznání horší. Nejedná se ovšem pouze o české specifikum. Obecně platí, že čím menší je počet rodilých mluvčích daného jazyka a tím méně je dostupných dat nebo překladů knih pro trénink jazykových modelů, tím nižší kvalitu přepisu lze očekávat. Situace se ještě zhoršuje u různorodých méně běžných dialektů již podporovaných jazyků.

88 Viz VANĚK, Miroslav a MŮCKE, Pavel. *Třetí strana trojúhelníku: teorie a praxe orální historie*, s. 173.

89 Viz kapitola: 3.3.4 Zpracování přirozeného jazyka

90 Čeština má na internetu oproti angličtině výrazně menší zastoupení. Velký vliv má také komplikovanější česká gramatika, výrazný rozdíl mezi spisovnou a obecnou češtinou apod.

Problémy s nářečím se ovšem vyskytují i při rozpoznávání angličtiny. Podle jedné studie⁹¹ vykazovalo pět nejmodernějších systémů využívajících automatické rozpoznávání řeči (Google, Amazon, Apple, IBM a Microsoft) téměř dvojnásobnou chybovost v rozpoznávání slov pronesených afroamerickými mluvčími ve srovnání s bělochy. V této studii byly k analýze rozdílů v chybovosti slov použity sociolingvistické rozhovory s Afroameričany. Problém opět pramení z nerovného zastoupení minoritních skupin ve vstupních datech, které byly použity pro trénink jazykových AI modelů, což je příčinou rasového zkreslení výsledků. Tento problém je možné řešit použitím jiného vhodnějšího jazykového korpusu⁹² v jazykovém modelu, který by více odpovídal sociální skupině narátorů a mohl tak poskytovat lepší výsledky. Pro anglicky vedené sociolingvistické rozhovory s Afroameričany byl například vytvořen jazykový korpus⁹³ s názvem CORAAL⁹⁴. Obdobně lze volit různé jiné jazykové korpusy na základě specifik konkrétního projektu. Pro český jazyk je největším volně dostupným jazykovým korpusem Český národní korpus⁹⁵.

Pro zlepšení kvality přepisu je možné vylepšit či přizpůsobit samotný používaný jazykový model, aby více vyhovoval specifickým potřebám daného projektu. Skupina německých výzkumníků například zkoumala vylepšení rozpoznávání řeči pro německé rozhovory z oblasti orální historie. Autoři studie se zaměřili na adaptaci systému rozpoznávání řeči na specifika těchto rozhovorů, které často zahrnují nespisovnou řeč, dialekty a specifickou terminologii. S využitím přibližně jednoho tisíce hodin tréninkových dat získaných z předešlých orálně-historických výzkumů a aplikací umělého zlepšení kvality audio záznamů dosáhli u automatického přepisu rozhovorů velmi významného snížení chybovosti rozpoznávání slov.⁹⁶

Kvalita automatického přepisu se u jednotlivých služeb výrazně liší a záleží také na metodě, jakou je výsledný přepis generován. Informace nejsou obsaženy výhradně v samotném narativu, ale také v přestávkách a mezerách v něm. Gesta, výrazy obličeje, ticho a další neverbální interakce nesou stejně mnoho informací jako mluvené slovo.⁹⁷ Z toho důvodu se pro účely orální historie zásadně

91 PESSANHA, Francisca a SALAH, Almila Akdag. 2021. *A Computational Look at Oral History Archives*. Journal of Computer and Cultural Heritage. [online]. 15(1), Článek 6 (únor 2022), 16 stran. DOI: 10.1145/3477605. Dostupné z: <https://doi.org/10.1145/3477605>, s. 5.

92 Jazykový korpus je soubor různých textů určitého jazyka a prostředí, ze kterých se AI jazykový model učí. Chceme-li analyzovat psané texty spisovné češtiny, použijeme jiný jazykový korpus, než pro analýzu veršovaných básní, nebo pro přepisy nahrávek mluveného brněnského hantecu.

93 PESSANHA, Francisca a SALAH, Almila Akdag. 2021. *A Computational Look at Oral History Archives*, s. 6.

94 Jedná se o anglickou zkratku pro *Corpus of Regional African American Language*.

95 Dostupné z: <https://www.korpus.cz/>.

96 GREF, M., SCHMIDT, C. a KOEHLER, J., 2018. *Improving Robust Speech Recognition for German Oral History Interviews Using Multi-Condition Training*. In: *Speech Communication; 13th ITG-Symposium*. Oldenburg, Germany. s. 1–5. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/8578034>, s. 4.

97 PESSANHA, Francisca a SALAH, Almila Akdag. 2021. *A Computational Look at Oral History Archives*, s. 6.

vypracovává doslovný přepis včetně přerázků a vatových slov. Přepis obsahuje i doplňkové informace⁹⁸, které nejsou v audio nahrávce patrné.⁹⁹ Některé služby poskytují syrový neredigovaný přepis včetně všech chyb, přerázků, nespisovných¹⁰⁰ a slangových výrazů či nářečí¹⁰¹, který je pro potřeby orálně-historické analýzy a interpretace ideální.

Jiné nástroje se naopak pokoušejí záznam redigovat, nespisovné výrazy nahrazovat a na základě slovníku hledat co možná nejbližší spisovný tvar, čímž se z rozhovoru ztrácí mnoho informací, nezbytných pro úplné pochopení vyprávění. Tento výstup může být nicméně také užitečný například pro prezentaci širokému publiku, kde je namísto autentičnosti prioritou čitelnost. Takový druh přepisů je například použit na webových stránkách projektu *100 studentských revolucí*¹⁰². Nutno však podotknout, že během převodu na spisovný jazyk tyto nástroje velmi často chybují, a přestože je výsledný text, na rozdíl od jeho syrového originálu, čitelnější a spisovnější, často se v něm výrazně změnil původní významy slov a pro potřeby analýzy a interpretaci je tak přepis téměř nepoužitelný.

Manuálně pořízený přepis rozhovoru je sice přesnější, ale vyžaduje až desetinásobek reálné doby trvání nahrávky.¹⁰³ Neredigovaný automaticky vytvořený přepis rozhovoru je sice nedokonalý a bude patrně vždy nutné si rozhovor znovu přehrát a manuálně opravit problémová místa, ale skutečnost, že již máme vytvořený alespoň základ přepisu, máme označené účastníky rozhovoru, doplněné časové značky, může čas potřebný pro manuální přepis výrazně snížit. Orálně-historické rozhovory obsahují relativně často popisy traumatických událostí. V takových případech, kdy i manuální přepis je pro přepisujícího emocionálně vyčerpávající, může být automatizovaný přepis užitečný.

Pro praktickou ukázkou jsem vybral nahrávku rozhovoru s Pavlem Žáčkem, který byl součástí projektu *100 studentských revolucí*. Konkrétně jsem se zaměřil na pasáž, ve které pan Žáček odpovídá na otázku: „*Co se vám vybaví, když se řekne listopad 1989?*“. Manuálně jsem vytvořil přesný přepis části tohoto rozhovoru, který bude sloužit jako referenční text pro porovnání kvality výstupů automatických přepisů.

98 Například: [narátorovi zazvonil telefon], [smích], [pláč] apod.

99 HLAVÁČEK, Jiří. *Metodické postupy archivace orálně-historických pramenů v digitálním věku*. Archivní časopis. 2019, 69(4), 389–401. ISSN 0004-0393, s. 396.

100 Například: bejt, řek, moh, těma dvěma chlapama, vtevřít vokno. Viz například https://cs.wikipedia.org/wiki/Obecná_čeština

101 Například brněnský hantec: *Celé Štatl se chodil na křivó věžku kókat, někdo se temu pecnil, někdo ne a některé jenom mávl haluzó, jako že se mu to ekluje*. Viz například <http://www.hantec.cz/hantec/povidky/krivavez.html>

102 Viz ÚSTAV PRO SOUDOBE DĚJINY AV ČR. Studenti. *100 studentských revolucí* [online]. 2019. Dostupné z: <http://www.studenti89.usd.cas.cz/studenti/>

103 PESSANHA, Francisca a SALAH, Almila Akdag. 2021. *A Computational Look at Oral History Archives*, s. 3.

No ohromná svoboda. Něco, co jsme předtím nemohli. Prostě expanze do světa, expanze do společnosti, expanze i vůči státu. Ale myslím, že obrovská míra svobody, která od té doby ustupuje a dnes prostě to, co člověk tehdy cítil. Tak jsme teda v naprosto jako zvláštní situaci a myslím, že všichni, kdo to srovnávají, kdo kdo nejsou unaveni, tak tak je to zase pro ně jakési motivační, jo protože mě bylo 20 v roce 1989 prostě na prahu, mně bylo 20 v květnu 89 mi bylo 20. Takže vlastně byl člověk na prahu nějaké dospělosti, něco málo chápal, ale moc toho moc toho samozřejmě nebylo. Spíše se divil.

Tento segment rozhovoru je následně poskytnut několika AI službám, které se specializují na přepis mluveného slova. Cílem je zjistit, do jaké míry jsou schopny tyto služby generovat přesný a věrný přepis v porovnání s manuálně vytvořeným přepisem.

S výjimkou služby Whisper jsou všechny ostatní testované softwarové nástroje, určené pro automatický přepis rozhovorů, velmi jednoduché pro použití. Tyto služby vždy vyžadují registraci uživatele a většinou poskytují omezený počet bezplatných minut pro testování předtím, než je vyžadována platba za prémiové služby. Navzdory jejich primárnímu zaměření na tvorbu titulků pro videa, přepisování podcastů nebo přepisy pracovních schůzek, jsou tyto nástroje obecně užitečné i pro potřeby orální historie. Co je nejdůležitější, všechny testované služby mají mezi podporovanými jazyky i češtinu. Každá z těchto služeb, s výjimkou služby Whisper, disponuje jednoduchým webovým rozhraním, kam postačí nahrát rozhovor, nastavit několik vstupních parametrů (jazyk rozhovoru, počet účastníků) a po krátkém zpracování stáhnout výsledný přepis. Přepis je téměř vždy možné stáhnout dle potřeby v několika různých datových formátech: tabulkovém (.tsv), pro počítačové zpracování (.json), titulky pro video (.srt, .vtt) a prostý text (.txt).

Kromě základních funkcí přepisu nabízí většina těchto služeb i generování nejrůznějších textových analýz, souhrnů rozhovorů a tematických analýz, stejně jako detekci pojmenované entity v textu (osoby, místa, data). Každý z těchto nástrojů má však i své další jedinečné funkce a omezení, které jsou uvedeny u každé služby zvlášť.

Speechmatics

No ohromná svoboda.

Něco, co jsme předtím nemohli.

Prostě expanze do světa, expanze do společnosti, expanze i vůči státu.

A myslím, že obrovská míra svobody, která od té doby ustupuje, je dnes prostě to, co člověk tehdy cítil tak naprosto jako zvláštní situaci.

A myslím, že všichni, kdo to srovnávají, kdo kdo nejsou unaveni, tak tak je to pro ně jakési motivační, protože mně bylo dvacet, že v roce 1989 prostě na Prahu nebylo dvacet, v květnu 89 bylo dvacet, takže vlastně byl člověk na prahu nějaké dospělosti.

Něco málo chápal, ale moc toho moc toho samozřejmě nebylo, spíše se divil.¹⁰⁴

104 Dostupné z: <https://www.speechmatics.com/>

Sonix

No ohromná svoboda. Něco, co jsme předtím nemohli. Prostě expanze do světa, expanze do společnosti, expanze i vůči státu. A myslím, že obrovská míra svobody, která od té doby ustupuje, je dnes prostě to, co člověk tehdy cítil tak naprosto jako zvláštní situaci. A myslím, že všichni, kdo to srovnávají, kdo kdo nejsou unaveni, tak tak je to pro ně jakési motivační, protože mně bylo dvacet, že v roce 1989 prostě na Prahu nebylo dvacet, v květnu 89 bylo dvacet, takže vlastně byl člověk na prahu nějaké dospělosti. Něco málo chápal, ale moc toho moc toho samozřejmě nebylo, spíše se divil.¹⁰⁵

Služba Sonix nabízí kromě přepisu i možnost identifikace hlasů pro rozlišení jednotlivých účastníků rozhovoru a poskytuje také analýzu sentimentu. Kromě toho umožňuje generování titulků a překlady do mnoha různých jazyků.

Trint

Ohromná svoboda. Něco, co jsme předtím nemohli. Prostě expanze do světa, expanze do společnosti, expanze vůči státu. No ale myslím, že obrovská míra svobody, která od té doby ustupuje. Dnes prostě to, co člověk tehdy cítil. Tak jsme teda v naprosto zvláštní situaci a všichni to srovnávají. Kdo nejsou unaveni, tak tak je to pro ně jakési motivační já, protože mně bylo 20. V roce 1989 prostě na Prahu přišlo 20. Nebo 20. V květnu jsme se najednou po roce, takže vlastně byl člověk na prahu nějaké dospělosti, něco málo chápal, ale moc toho moc toho samozřejmě nebylo, spíše se divil.¹⁰⁶

Trint vyniká výjimečným uživatelským prostředím pro snadnou manuální revizi přepisů, umožňuje jednoduchou navigaci v textu a přidávání poznámek, což výrazně urychluje proces manuálních oprav přepisu.

Google Cloud

No, ohromná svoboda, eh, něco, co jsme předtím nemohli, prostě expanze do světa, expanze, do společnosti, expanze, i vůči státu, ale myslím si obrovská míra svobody, která od té doby ustupuje a a dnes prostě to, co člověk tehdy cítil, tak jsme teda v naprosto jako zvláštní situaci, já myslím, že všichni, kdo to srovnávají, kdo kdo nejsou unavení, tak, tak, je to pro mě jakési motivační, jo, protože mě bylo 20, v roce 1989, prostě na Prahu, tě bylo 20 rovnejch, mě bylo 20, v květnu 89 mě bylo 20, no, takže vlastně byl člověk na Prahu nějaké dospělosti, eh, něco málo chápal, ale moc toho, moc toho samozřejmě nebylo, spíše se divil.¹⁰⁷

Condens

No ohromná svoboda. Něco, co jsme předtím nemohli prostě expanze do světa expanze do společnosti expanze i vůči státu. Ale myslím, že obrovská míra svobody, která od té doby ustupuje a dnes prostě to, co člověk tehdy cítil, tak jsme teda naprosto jako

105 Dostupné z: <https://my.sonix.ai/>

106 Dostupné z: <https://app.trint.com/>

107 Dostupné z: <https://console.cloud.google.com/speech/transcriptions/>

zvláštní situaci a myslím, že všichni, kdo to srovnávají, kdo kdo nejsou unaveni, tak tak je to zase pro ně jakési motivační. Jo, protože mě bylo 20 v roce 19, 89. Prostě na prahu přibylo 20 rovnej nebylo 20, v květnu 89 bylo 20. No, takže vlastně byl člověk na prahu v nějaké dospělosti? Něco málo chápal, ale moc toho moc toho samozřejmě nebylo. Spíše se divil.¹⁰⁸

Plain Scribe

00:00:00 *No, ohromná svoboda, něco, co jsme předtím nemohli, prostě expanze do světa,*
00:00:08 *expanze do společnosti, expanze i uči státu, ale myslím si, že obrovská měna svobody,*
00:00:16 *která od té doby ustupuje a dnes prostě to, co člověk tehdy cítil,*
00:00:21 *tak jsme traf naprosto zvláštní situaci a myslím, že všichni, kdo to srovnávají,*
00:00:29 *kdo nejsou únavení, tak je to základně motivační.*
00:00:34 *Protože mě bylo 20, že byl v roce 1989, prostě byl na Prahu.*
00:00:40 *Mě bylo 20, v květnu 1989 bylo 20.*
00:00:44 *Takže vlastně byl člověk na Prahu v nějaké dospělosti, něco málo chápal,*
00:00:50 *ale moc toho samozřejmě nebylo, spíše se divil.¹⁰⁹*

Česky.AI

No, ohromná svoboda, něco, co jsme předtím nemohli, prostě expanze do světa, expanze do společnosti, expanze i uči státu, ale myslím, že obrovská měna svobody, která od té doby ustupuje a dnes prostě to, co člověk tehdy cítil, tak jsme traf naprosto zvláštní situaci a myslím, že všichni, kdo to srovnávají, kdo nejsou únavení, tak je to základně motivační.
Protože mě bylo 20, v roce 1989, prostě byl na Prahu. Mě bylo 20, v květnu 1981 bylo 20. Takže vlastně byl člověk na Prahu v nějaké dospělosti, něco málo chápal, ale moc toho samozřejmě nebylo, spíše se divil.¹¹⁰

Služba Česky.AI je specificky orientována na český jazyk a oproti ostatním službám nabízí i česky popsané uživatelské rozhraní. Dále také nabízí funkci, která automaticky opravuje gramatické chyby. Ačkoliv taková oprava nemusí být vhodná pro orálně-historickou analýzu, kde jsou přesné a neupravené záznamy mluveného slova klíčové, pro účely publikace přepisu širšímu publiku je text upravený tímto způsobem mnohem přístupnější a srozumitelnější.

Beey.io

No, ohromná svoboda, něco, co jsme předtím nemohli, prostě expanze do světa, expanze, do společnosti, expanze i vůči státu, ale myslím, že obrovská míra svobody, která od té doby ustupuje a dnes prostě to, co člověk tehdy cítil, jsme teda v naprosto jako zvláštní situaci. Já myslím, že všichni, kdo to srovnávají, kdo nejsou unaveni, tak je to pro ně jakési motivační, protože mně bylo 20, v roce 19 89, prostě na Prahu přibylo

108 Dostupné z: <https://condens.io/>

109 Dostupné z: <https://www.plainscribe.com/>

110 Dostupné z: <https://cesky.ai/>

20, drobnej nebylo 20, v květnu 81 bylo 20, no, takže vlastně byl člověk na Prahu v nějaké dospělosti, něco málo chápal, ale moc toho samozřejmě nebylo, spíše se divil.¹¹¹

Good Tape

No, ohromná svoboda, něco, co jsme předtím nemohli, prostě expanze do světa, expanze do společnosti, expanze i vůči státu. Ale myslím si, obrovská míra svobody, která od té doby ustupuje a dnes prostě to, co člověk tehdy cítil, tak jsme teda v naprosto zvláštní situaci a myslím, že všichni, kdo to srovnávají, kdo nejsou unavení, tak je to jakési motivační. Protože mě bylo 20 v roce 1989 na Prahu. Mě bylo 20 v květnu, v roce 1981 mě bylo 20. Takže vlastně byl člověk na Prahu v nějaké dospělosti, něco málo chápal, ale moc toho samozřejmě nebylo. Spíše se divil.¹¹²

Whisper

No, ohromná svoboda, něco, co jsme předtím nemohli, prostě expanze do světa, expanze do společnosti, expanze i vůči státu, ale myslím si obrovská míra svobody, která od té doby ustupuje a dnes prostě to, co člověk tehdy cítil, tak jsme teda v naprosto zvláštní situaci a myslím, že všichni, kdo to srovnávají, kdo nejsou unavení, tak je to jakési motivační. Protože mě bylo 20 v roce 1989, prostě byl na Prahu. Mě bylo 20 v Ketnu, v 1981 mě bylo 20, takže vlastně byl člověk na Prahu v nějaké dospělosti, něco málo chápal, ale moc toho samozřejmě nebylo, spíše se divil.¹¹³

Služba Whisper, kterou vyvíjí nezisková organizace OpenAI, stojící i za službou ChatGPT, se od ostatních testovaných služeb výrazně odlišuje. Whisper na rozdíl od ostatních služeb nedisponuje jednoduchým webovým rozhraním pro ovládání a vyžaduje tak mírně pokročilejší znalost IT.¹¹⁴ Nabízí však možnost stažení jazykového modelu do vlastního počítače. Automatické přepisy je tak možné generovat zcela zdarma v libovolném množství a především bezpečně. Po stažení a nainstalování balíčku již pracuje zcela bez přístupu k internetu a přepisované nahrávky nejsou, na rozdíl od ostatních testovaných služeb, odesílány na cizí server a potenciálně zneužitelné. Použití je velmi jednoduché, v příkazové řádce postačí spustit následující dva příkazy.¹¹⁵

1. `pip install -U openai-whisper`
2. `whisper rozhovor.wav --language Czech --model large`

První příkaz nainstaluje do počítače knihovnu Whisper a druhý příkaz provede přepis. Přepisovat lze audio nahrávky různých formátů (.wav, .mp3, .mp4, a další). V prvním parametru (`--language`) se volí jazyk nahrávky. Druhý parametr (`--model`) určuje jazykový model. Na výběr je 5 modelů (tiny, base, small, medium, large), které se od

111 Dostupné z: <https://www.beey.io/>

112 Dostupné z: <https://goodtape.io/>

113 Dostupné z: <https://openai.com/research/whisper>

114 Viz OPENAI. Introducing Whisper. *OpenAI* [online]. 2022 [cit. 2024-03-16]. Dostupné z: <https://openai.com/research/whisper>

115 Podrobnější informace jsou dostupné z: <https://github.com/openai/whisper>

sebe liší velikostí, počtem parametrů. Čím vyšší model, tím kvalitnější přepis, ale zároveň větší nároky na čas a na výkon hardware. Nutno podotknout, že přepis audio nahrávek je hardwarově velmi náročná operace. Nejméně kvalitní přepis (model tiny) je velmi rychlý a je možné ho používat i na běžném laptopu, ale výsledný přepis je opravdu velmi špatný. Na druhou stranu nejkvalitnější přepis (model large) nabízí výborné výsledky, ale vyžaduje přibližně 10 GB VRAM na grafické kartě. Takovým výkonem disponují pouze výkonné herní počítače. Na laptopu lze omezení obejít připsáním dodatečného parametru (`--device=cpu`) na konec příkazu, který pro přepis využije namísto grafické karty výkon procesoru. Výrazně se tím ale prodlouží doba exekuce příkazu. Hodinový rozhovor se tak při nejvyšší kvalitě (model large) bude na běžném laptopu přepisovat přibližně 1-2 hodiny.

Pro zhodnocení kvality automatických přepisů porovnáním s referenčním manuálním přepisem byla použita služba *Similarity Checker*¹¹⁶, která umožňuje objektivně posoudit shodu mezi oběma verzemi. Bylo ovšem nutné vzít v úvahu, že určité rozdíly v přepisu nejsou vždy chybou AI (například rozdíly ve formátování čísel či interpunkci). Tedy aspekty, jako je zápis čísla dvacet (slovy nebo číslicemi), nebo rozdělení vět, byly ignorovány a neovlivňovaly celkové hodnocení. Při analýze bylo zjištěno, že přesnost automatických přepisů se pohybovala kolem 90 %. Přestože výsledek ukazuje na poměrně vysokou úroveň přesnosti pro přepis mluveného slova, pro účely orální historie to stále není zcela dostačující úroveň.

Tabulka 2: Přesnost přepisů pomocí různých AI nástrojů¹¹⁷

Služba	~ přesnost (v %)
<i>Speechmatics</i>	93,5
<i>Sonix</i>	93,5
<i>Trint</i>	88,6
<i>Google Cloud</i>	88,1
<i>Condens</i>	92,6
<i>Plain Scribe</i>	88,2
<i>Česky.AI</i>	88,0
<i>Beey.io</i>	93,5
<i>Good Tape</i>	92,7
<i>Whisper</i>	91,0

Automatický přepis rozhovorů lze tedy rozhodně s určitými omezeními používat a je jen na zvážení výzkumníka, zda automatický přepis použije, nebo ho využije pouze jako základ, který bude dále opravovat do správné podoby, nebo rozhovor přepíše (či nechá přepsat) zcela manuálně.

116 Dostupné z: <https://toolsaday.com/text-analysis/similarity-checker>

117 Porovnání přesnosti transkriptu provedené pomocí nástroje Similarity Checker. Dostupné z: <https://toolsaday.com/text-analysis/similarity-checker>

4.3.3 Metadata

Lze si ovšem představit situaci, v níž možnosti umělé inteligence oproti schopnostem jednotlivce dominují zcela jednoznačně a tou je zpracování velkého množství dat. Projekty orální historie mají téměř ve všech případech velmi omezené prostředky a z minulosti je známá celá řada neukončených projektů, jimž v průběhu výzkumu došly finanční zdroje a projekty skončily pouhým sběrem nahrávek s neidentifikovatelnými a nezpracovanými rozhovory.¹¹⁸ V archivu se tak postupně hromadí ohromné množství neidentifikovaných audio záznamů. Nejsou k nim dostupná žádná metadata. Nevíme, do kterého projektu náleží konkrétní nahrávky, co přesně obsahují a stojí-li vůbec za probádání. Ve výsledku se jedná o stovky až tisíce hodin audio záznamů. Lze si představit situaci, že se k tomuto archivu po několika letech či desetiletích dostane výzkumník s ambicí tyto záznamy prozkoumat, zanalyzovat a identifikovat, například zda některé z těchto záznamů neobsahují vzácné unikátní nahrávky. První možností řešení takového úkolu je takzvaně pomocí hrubé síly, tedy manuálně vyslechnout veškeré audio záznamy, vypracovat přepisy, rozčlenit do projektů, identifikovat narátory, což pro jednotlivce může představovat tisíce hodin práce. Při použití dříve zmíněných AI nástrojů lze provést základní analýzu všech rozhovorů, vypracování kondenzovaných rozhovorů a vygenerování bohaté sady metadat (identifikaci tazatelů a narátorů a destilaci klíčových slov a témat)¹¹⁹ v řádu několika minut či nižších jednotek hodin. Veškerá tato metadata mohou poté sloužit pro podrobný popis archivovaného dokumentu a následně pro snadnější a rychlejší vyhledávání v rozsáhlých archivech orální historie. V tomto konkrétním případě užití jsou přepisy rozhovorů s několika chybami a s absencí popisu paralingvistických prostředků více než dostačující.

Prozatím neexistuje mezinárodní metadatový standard pro orální historii. Jednotlivé instituce zabývající se orální historií si udržují vlastní metadatová schémata. V českém prostředí například ÚSD¹²⁰ dělí metadata pro orálně-historické rozhovory do třech úrovní (narátor, rozhovor, technické údaje). První úroveň metadat slouží k popisu narátora (jméno, příjmení, titul, datum narození, projekt). Ve druhé úrovni následují metadata popisující samotný rozhovor (název souboru, formát rozhovoru, transkript, identifikace tazatele, datum, délka, jazyk a místo rozhovoru) a třetí úroveň je věnována technickým údajům nahrávky (originál záznamu, typ, formát, velikost, licence).¹²¹ Ze zahraničí lze například doporučit orálně-historický projekt Historic Pittsburgh, jehož metadatové schéma je veřejně dostupné a přehledně popsáno.¹²²

118 VANĚK, Miroslav a MÜCKE, Pavel. *Třetí strana trojúhelníku: teorie a praxe orální historie*, s. 153.

119 Příklady viz kapitola: 3.3.4 Zpracování přirozeného jazyka

120 Ústav pro soudobé dějiny AV ČR, v.v.i.

121 HLAVÁČEK, Jiří. *Metodické postupy archivace orálně-historických pramenů v digitálním věku*, s. 397-398.

122 Orálně-historický projekt Historic Pittsburgh funguje na knihovním systému University of Pittsburgh. Metadatové schéma je dostupné z: https://github.com/uls-mad/islandora_metadata/wiki/Oral-History-Collections-Metadata

4.4 Analýza dat

4.4.1 Analýza rozhovoru

Pro účely analýzy obsahu rozhovoru lze v první řadě využít AI ke shrnutí obsahu rozhovoru, čehož lze dosáhnout použitím široce používaných velkých jazykových modelů. To ovšem přináší zde již diskutovaná úskalí, omezení a rizika. Tyto služby jsou většinou placené, soubory se odesílají ke zpracování na cizí servery a služby lze většinou ovládat výhradně prostřednictvím webového rozhraní. Dávkové zpracování většího množství souborů je tak navíc zdlouhavé a komplikované. Řešením může být v tomto případě například program GPT4All, který zde byl již také zmíněn a který umožňuje používat vybrané velké jazykové modely na lokálním počítači zdarma a bezpečně.

Software stačí nainstalovat do počítače jedním příkazem v příkazové řádce.

```
1. pip install gpt4all
```

Ovládat chatovací nástroj lze pomocí přehledného grafického rozhraní, nebo pomocí jednoduchého Python skriptu, který udělá shrnutí pro každý transkript rozhovoru, nacházejícího se v uvedené složce. Jedním příkazem tak lze zpracovat velké množství rozhovorů najednou.

```
1. from gpt4all import GPT4All
2. import glob
3.
4. model = GPT4All("wizardlm-13b-v1.2.Q4_0.gguf")
5. source_directory = "/Prepisy/"
6. file_pattern = "*.txt"
7. prompt_template = "Udělej shrnutí následujícího rozhovoru: \"{texts}\""
8.
9. # Procházení všech textových souborů ve složce
10. for file_path in glob.glob(source_directory + file_pattern):
11.     with open(file_path, 'r', encoding='utf-8') as input_file:
12.         texts = input_file.read()
13.         prompt = prompt_template.format(texts=texts)
14.         # Vykonání promptu
15.         output = model.generate(prompt)
16.         print(output)
```

Mluvená řeč není pouze o slovech, ale obsahuje kromě verbálních informací i velké množství těch, které se nedají vyjádřit slovy, které nám dávají nahlédnout nejen do toho, co si mluvčí myslí, ale i do jeho/jejích emocí. Už zde bylo zmíněno, že automatické přepisy rozhovorů postrádají mnoho paralingvistických prostředků. Paralingvistika ale kromě vatových slov, mimoslovních zvukových projevů (smích, pláč, kašlán, funění, sténání apod.) a ticha mezi promluvami, které se v přepisu

orálně-historického rozhovoru zpravidla poznamenávají v hranatých závorkách, zkoumá také například kvalitu hlasu a akustické parametry mluveného slova, jako je intonace, rytmus, tón a důraz.¹²³ Paralingvistické složky řeči jsou užitečné pro rozpoznávání řeči, ale i pro interpretaci záměrů narátorů a analýzu rozhovoru. Nástroje umělé inteligence umožňují provádět automatickou počítačovou paralingvistickou analýzu rozhovorů. V současné době se ale zpracováním archivů orální historie pomocí nástrojů paralingvistiky zabývá pouze velmi málo vědeckých prací.¹²⁴

Dle mého názoru je to způsobeno jednak tím, že vývoj technologických prostředků není ještě na dostatečně vysoké úrovni, aby výsledky jejich analýzy byly pro vědecké účely dostatečně spolehlivé a užitečné a zároveň se zde naráží na etické problémy s tím spojené. Přestože nonverbální gesta jsou pro orální historii důležitá a video nahrávky se běžně detailně analyzují, na návrh doplnit kamerový záznam rozhovoru například o detektor lži a/nebo detektor nervových synapsí mozku pro lepší monitoring nonverbálních signálů by patrně většina narátorů nepřistoupila. Následující přehled technologických prostředků využívajících umělou inteligenci tak slouží více pro demonstraci a zamyšlení nad možnostmi, které moderní AI nástroje pro analýzu rozhovoru nabízejí, než že by bylo možné zmíněné technologie pro analýzu běžného rozhovoru v praxi použít.

4.4.2 Analýza orálně-historických dat v jiných vědních disciplínách

Orální historie má společné zájmy a praktiky s mnoha dalšími blízkými disciplínami, jako je například etnografie, antropologie, sociologie, literární a kulturní studia a psychologie. Vědci těchto blízkých vědních oborů rovněž vytvářejí protokoly a audio nahrávky rozhovorů, které se v mnohém podobají archivům orální historie. Pro orálně-historické nahrávky tak lze najít uplatnění i v jiných oborech, což bylo samozřejmě možné i dříve bez použití umělé inteligence. S jejím využitím se ale v tomto směru otevírá mnoho nových originálních možností.¹²⁵

Velmi zajímavou možností je například hledání korelací mezi neverbálními signály a problémy s duševním zdravím, jako je deprese nebo posttraumatická stresová porucha (PTSD). Mnoho orálně-historických sbírek obsahuje vyprávění přeživších o traumatických událostech. Výzkum orální historie často klade důraz na traumata a trauma se zde stalo často probíraným tématem. Nové nástroje umělé inteligence mají potenciál analyzovat verbální i neverbální komunikaci pro potřeby orálních historiků a datových vědců, ale také psychologů a psychiatrů.¹²⁶ Pro příklad lze uvést

123 Viz například <https://cs.wikipedia.org/wiki/Paralingvistika>

124 PESSANHA, Francisca a SALAH, Almila Akdag. 2021. A Computational Look at Oral History Archives, s. 7.

125 Tamtéž, s. 2

126 Tamtéž, s. 6

výzkum¹²⁷, který byl proveden na univerzitě Stony Brook v New Yorku. V tomto výzkumu byly použity stovky orálně-historických rozhovorů uskutečněných s policisty, lékaři a zdravotníky, kteří pomáhali zachraňovat přeživší po útoku na Světové obchodní centrum 11. září 2001 v New Yorku. Výzkum zkoumal vztah mezi jazykovými vzorci a symptomy PTSD (posttraumatické stresové poruchy). Psychologové v této studii využili umělé inteligence k analýze jazyka užívaného v rozhovorech. Výsledky studie naznačují, že určité jazykové charakteristiky, jako jsou projevy úzkosti či deprese, používání krátkých slov a krátkých vět, vyprávění v první osobě jednotného čísla a další charakteristiky mohou u pacientů předpovídat propuknutí PTSD v blízké budoucnosti. Na základě těchto údajů byla následně také s využitím AI vytvořena trajektorie symptomů, s jejíž pomocí bylo možné již předem předpovídat, zda se u některého z narátorů či narátorek symptomy PTSD zhoršují a bylo možné zahájit případnou medikaci ještě předtím, než se tato duševní porucha plně projevila.¹²⁸

4.4.3 Analýza sentimentu

Umělá inteligence rovněž nabízí pokročilé možnosti v oblasti analýzy sentimentu. Tato technologie dokáže identifikovat a analyzovat emoce, postoje a názory vyjádřené v mluveném slově. Na rozdíl od dále zmíněných analytických nástrojů je eticky nejméně problematická. Není nezbytně nutné analyzovat původní audio či video záznam rozhovoru. Lze například použít pouze výsledný psaný přepis rozhovoru, ve kterém je již většina nevědomých nonverbálních projevů odfiltrována. Tento nástroj je již běžnou součástí široce používaných AI velkých jazykových modelů (ChatGPT, Gemini). Zároveň aktivně slouží například pro detekci spamu v emailu, nebo pro automatickou kategorizaci emailů pro zákaznické linky, kde podle obsahu zprávy vyhodnotí její obecný sentiment a rozdělí zprávy podle spokojenosti se službou na spokojené, zklamané či rozčilené a tím může stanovit priority pro jejich řešení. Analýzu sentimentu lze samozřejmě provádět i na audio nahrávkách. Využívání takového softwaru například na již zmíněné zákaznické lince pro účely určení dominantní emoce volajících zákazníků se jeví eticky naprosto v pořádku a v praxi již běžně používané. Analyzovat sentiment z hlasového projevu narátora v rámci orální historie může být již eticky diskutabilní.

V kapitole věnované automatickému přepisování rozhovorů bylo zmíněno, že například služba Sonix poskytuje kromě samotného přepisu i širokou paletu dalších funkcí, využívajících možností

127 SON, Y. a kol., 2020. World Trade Center responders in their own words: predicting PTSD symptom trajectories with AI-based language analyses of interviews. *Psychological Medicine*. [online]. 53, 918–926. Dostupné z: <https://doi.org/10.1017/S0033291721002294>

128 Tamtéž, s. 924 (7)

umělé inteligence. Jednou z jich je i analýza sentimentu, která umožňuje slova z rozhovoru nejen přepsat, ale zároveň interpretovat a analyzovat jejich emoční zabarvení.

Stejně jako u přepisu rozhovorů lze tuto analýzu předvést například na odpovědi pana Žáčka na otázku týkající se jeho osobního vnímání událostí listopadu 1989. Analýza sentimentu může pomoci identifikovat, jaké emoce v něm tyto vzpomínky vyvolávají, jak pan Žáček vnímal tento historický moment, ať už s nadějí, nostalgií, možná i se smíšenými pocity nebo s pocitem zklamání z toho, jak se situace od té doby vyvíjela.

[00:00:00-00:01:50] Sentiment: Smíšený

- *Mluvčí začíná pozitivně, když vzpomíná na nově nabytou svobodu a rozšíření do světa a společnosti, což naznačuje období optimismu a pozitivních změn. Fráze jako "obrovská míra svobody" a "otevřely se dveře do světa" podporují tento pozitivní sentiment.*
- *Nicméně, sentiment se mění na negativní, když mluvčí přemýšlí o poklesu této svobody v průběhu času a uvědomuje si, že společenské zlepšení nebylo tak významné, jak se očekávalo. Fráze jako "ustupuje", "moc toho samozřejmě nebylo" a "nečekal" přispívají k negativnímu sentimentu.*

Kontextová analýza: Počáteční pozitivní sentiment vychází z mluvčího zkušenosti s transformačním obdobím v historii, spojeným s osobním růstem a společenskými změnami. Negativní sentiment vzniká z mluvčího zklamání nad nakonec malým pokrokem a uvědoměním trvajících problémů z minulosti, což podtrhuje rozdíl mezi minulými nadějemi a současnou realitou.¹²⁹

Analýza sentimentu může být zvláště užitečná v situacích, kdy se historik snaží interpretovat složitější nebo jemnější nuance v emotivních projevech narátorů, které mohou být závislé na konkrétních okamžicích jejich vyprávění.

V rámci orální historie má historik obvykle k dispozici originální audiozáznam rozhovoru, často doprovázený i videozáznamem. Analýza emocí a sentimentu se tak nemusí omezovat pouze na analýzu textových přepisů rozhovorů, ale může využívat i originálních audio a video nahrávek, které poskytují bohatý zdroj emocionálních a nonverbálních projevů. Audio nahrávky mohou odhalit změny v tónu hlasu, rychlosti řeči nebo pauzy, které mohou být indikátorem sentimentu a emocí. Video nahrávky pak umožňují analyzovat gesta, mimiku a další tělesné projevy, které mohou být stejně významné pro pochopení emocí a postoje narátorů.

Sonix a jemu podobné AI nástroje se specializují na analýzu sentimentu založenou výhradně na textových prepisech. Pro hlubší zkoumání emocí z hlasového projevu je možné využít služby jako *Vokaturi*¹³⁰, která umožňuje rozšířenou analýzu sentimentu přímo z audio záznamů. Následující

129 Analýza sentimentu z přepisu rozhovoru s panem Žáčkem provedená pomocí nástroje Sonix. Dostupné z: <https://my.sonix.ai/>

130 Hlasový analyzátor *Vokaturi*. Dostupné z: <https://vokaturi.com/>

Python skript¹³¹ načte audio nahrávku a pro každou emoci (neutrální, štěstí, smutek, zlost, strach) vypočítá procentní zastoupení. Předtím je potřeba stáhnout *OpenVokaturi* knihovnu¹³² a po stažení ji uložit do stejné složky, kde budeme mít uložen následující skript.

```
1. import sys
2. sys.path.append("OpenVokaturi-4-0/api")
3. import scipy.io.wavfile
4. import Vokaturi
5. import platform
6. import struct
7.
8. # Načtení knihovny Vokaturi pro různé operační systémy
9. platform_system = platform.system()
10. arch = "64" if struct.calcsize("P") * 8 == 64 else "32"
11. if platform_system == "Darwin":
12.     library_path = f"OpenVokaturi-4-0/lib/open/macos/OpenVokaturi-4-0-mac.dylib"
13. elif platform_system == "Windows":
14.     library_path = f"OpenVokaturi-4-0/lib/open/win/OpenVokaturi-4-0-win{arch}.dll"
15. elif platform_system == "Linux":
16.     library_path = "OpenVokaturi-4-0/lib/open/linux/OpenVokaturi-4-0-linux.so"
17. else:
18.     raise OSError("Nepodporovaný operační systém")
19.
20. Vokaturi.load(library_path)
21.
22. # Vložení zvukového souboru k analýze
23. file_name = 'Zacek_1.wav'
24. sample_rate, samples = scipy.io.wavfile.read(file_name)
25.
26. # Příprava audio bufferu
27. buffer_length = len(samples)
28. c_buffer = Vokaturi.float64array(buffer_length)
29.
30. if samples.ndim == 1: # Mono
31.     c_buffer[:] = samples / 32768.0
32. else: # Stereo
33.     c_buffer[:] = (samples[:, 0] + samples[:, 1]) / 2 / 32768.0
34.
35. # Analýza emocí
36. voice = Vokaturi.Voice(sample_rate, buffer_length, 0)
37. quality = Vokaturi.Quality()
38. emotionProbabilities = Vokaturi.EmotionProbabilities()
39.
40. voice.fill_float64array(buffer_length, c_buffer)
41. voice.extract(quality, emotionProbabilities)
42.
43. if quality.valid:
44.     print(f"Neutrální: {emotionProbabilities.neutrality:.3f}")
45.     print(f"Štěstí: {emotionProbabilities.happiness:.3f}")
46.     print(f"Smutek: {emotionProbabilities.sadness:.3f}")
```

131 Kód byl sestaven na základě ukázky kódu dostupné z: <https://vokaturi.com/using-vokaturi/example-code-for-python-batch>

132 Knihovna OpenVokaturi je ke stažení na webových stránkách: <https://vokaturi.com/downloads/download-the-sdk>

```
47. print(f"Hněv: {emotionProbabilities.anger:.3f}")
48. print(f"Strach: {emotionProbabilities.fear:.3f}")
49.
50. voice.destroy()
```

Skript na výstupu vypíše procentuální zastoupení emocí ve zvukovém souboru.

- Neutrální: 0,956
- Štěstí: 0,000
- Smutek: 0,044
- Hněv: 0,000
- Strach: 0,000

Základní analýza sentimentu provedená na stejném segmentu rozhovoru s panem Žáčkem poskytuje výsledky srovnatelné s textovou analýzou sentimentu. Hlasový projev se nijak výrazně neodlišuje od obsahu rozhovoru. V hlasovém projevu nedominuje žádná specifická emoce, z 95,6 % má projev neutrální emoční zabarvení. Smutek tvoří 0,044 % emocionálního obsahu. Emoce jako štěstí, hněv a strach jsou v hlasovém projevu zastoupeny zcela minimálně.

Výrazy obličeje

Počítačové rozpoznávání obličejových výrazů je již poměrně vyspělým odvětvím používaným pro interakci mezi člověkem a počítačem, pro porozumění lidského chování a rozpoznávání duševního stavu člověka. Přestože se může zdát, že se jedná o triviální záležitost, jelikož člověk dokáže identifikovat výrazy obličeje velmi dobře, opak je pravdou. I jednoduchý úsměv může obsahovat nuance, které běžný člověk nedokáže rozlišit. Podle Paula Ekmana existují základní emocionální výrazy obličeje (šťěstí, strach, hněv, překvapení, smutek, zhnusení), které jsou až na výjimky rozpoznatelné napříč kulturami.¹³³ Ovšem pouze do určité míry. Kromě kultury se některé sociální signály liší i na základě pohlaví, identity či věku. Obzvláště patrné je to u zpracování řeči starších osob nebo dětí. Tyto skupiny mají velmi odlišné akustické vlastnosti hlasu, které se výrazně liší od ostatních věkových skupin.¹³⁴

Automatizované analytické nástroje tak mohou na základě výrazu obličeje, pozice hlavy a směru pohledu poměrně dobře detekovat emoční rozpoložení člověka. A to nejen na úrovni základních emocí, ale také jejich kombinace (šťastně překvapený). Potenciál analýzy obličeje pro orálně-historické archivy je zřejmý, jelikož mnoho archivů obsahuje videozáznamy zachycující obličeje

133 EKMAN, Paul. *Basic emotions*. In: DALGLEISH, Tim a POWER, Mick J. (eds.). *Handbook of Cognition and Emotion*. John Wiley & Sons, 1999, s. 46.

134 PESSANHA, Francisca a SALAH, Almila Akdag. 2021. *A Computational Look at Oral History Archives*, s. 11.

narátorů.¹³⁵ Koronavirová pandemie, kvůli které se distančně vedený rozhovor stal běžnější praxí, paradoxně dopomohla tomu, že máme ještě více orálně-historických rozhovorů ve formě videozáznamu. Při běžném vedení rozhovorů se zpravidla vždy nahrává zvukový záznam, ale videozáznam je nahráván pouze v některých případech. Při distančně vedeném rozhovoru je ale kvůli zajištění vizuálního kontaktu zcela běžné komunikovat prostřednictvím videohovoru.¹³⁶

Pro analýzu obrazu lze opět vybírat z velkého množství nejrůznějších nástrojů. Jedním z nich je například služba *DeepFace*¹³⁷, která poskytuje sadu zajímavých funkcí. „*Face Verification*“ například umožňuje porovnat obrázky obličejů a ověřit, zda se jedná o totožného člověka (počítačové ověření fotografie na občanském průkazu). Funkce „*Face Recognition*“ ověřuje, zda se poskytnutý obrázek nalézá v konkrétní obrázkové galerii (například pro ověření, zda se na obrázek nevztahují autorská práva). Dále nabízí službu pro vyhodnocování obrazu z web kamery v reálném čase. Pro účely orální historie je nejužitečnější funkce „*Facial Attribute Analysis*“, která slouží pro analýzu různých obličejových atributů, jako jsou odhadovaný věk, rasa, pohlaví a zejména emoční rozpoložení osoby. K tomu účelu používá řadu AI modelů, které dokážou rozlišit různé emoční stavy jako radost, smutek, překvapení, hněv a další. Použití této služby lze demonstrovat například na rozhovoru s Klárou Jirovcovou Pospíšilovou, který opět pro projekt *100 studentských revolucí* vedl prof. Miroslav Vaněk.

Nejprve je potřeba pomocí příkazového řádku nainstalovat knihovnu *DeepFace* do počítače.

```
1. pip install deepface
```

Následné použití služby je už velmi jednoduché. K provedení analýzy postačí opět Python skript, který má pouhé dva řádky.

```
1. from deepface import DeepFace  
2. DeepFace.analyze(img_path='jirovcova_pospisilova_1.png')
```

135 Tamtéž, s. 8.

136 HLAVÁČEK, Jiří. *Covido, ergo zoom! Metodologické, technické a etické aspekty distančně vedených rozhovorů během pandemie Covid-19*, s. 65.

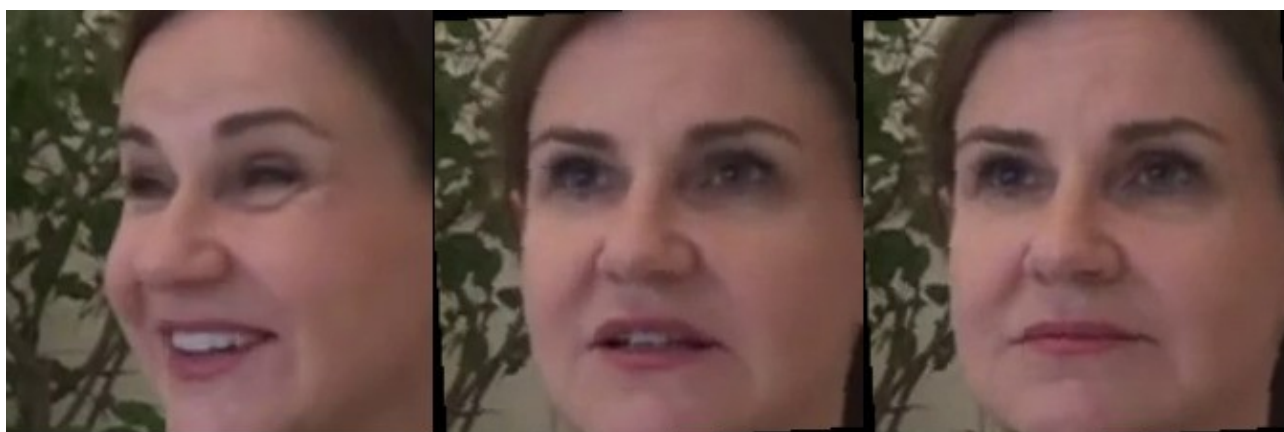
137 Podrobnější informace dostupné z: <https://viso.ai/computer-vision/deepface/>

Po provedení analýzy emocí na fotografii vyjmuté z videa rozhovoru bylo přiřazeno nejvyšší procento emocím štěstí (96,64 %), což naznačuje, že se dotyčná osoba pravděpodobně usmívá. Region určuje souřadnice detekovaného obličeje v rámci obrazu nebo video rámečku. Odhadovaný věk osoby je 47 let. Je zde také procentní vyjádření předpovědi pohlaví, které ukazují 99,19% jistotu, že se jedná o ženu a 0,81% pravděpodobnost, že se jedná o muže. Dominantním (pravděpodobnějším) pohlavím je tedy rozhodně žena. Následují hodnoty pravděpodobnosti pro různé rasové kategorie: asijská, indická, černá, bílá, středovýchodní a latinoamerická. Nejvyšší pravděpodobnost má bílá rasa (51,32 %), následovaná latinoamerickou (22,90 %) a středovýchodní (21,73 %). Dominantní rasou je tedy bílá. Celkově tato analýza naznačuje, že analyzovaná osoba je přibližně 47letá bílá žena, jejíž aktuálně převládající emocí je štěstí.

Pro orální historii nicméně nejsou údaje o rasové příslušnosti nebo odhadovaném věku narátora příliš relevantní. Klíčovým aspektem je zde emoční analýza a identifikace převažujících emocí. Při implementaci by bylo možné si službu přizpůsobit tak, aby v reálném čase analyzovala emoce narátora během přehrávání videozáznamu. Výsledky analýzy emocí mohou odhalit, jak narátor reaguje na konkrétní otázky nebo téma, a mohou být použity k doplnění textového přepisu o další vrstvu informací. Tímto způsobem může emoční analýza pomoci odhalit skryté nuance a kontext rozhovoru, který by mohl být při pouhém poslechu nebo čtení přepisu přehlédnut.

```
[ 1 item
  0 : {
    emotion : {
      angry : 0.2509000877708092
      disgust : 4.441981541531818e-05
      fear : 0.006315641050783866
      happy : 96.63987795831228
      sad : 0.2368383758558779
      surprise : 8.352116224406402e-05
      neutral : 2.8659375693276443
    }
    dominant_emotion : happy
    region : {
      x : 473
      y : 150
      w : 190
      h : 190
    }
    age : 47
    gender : {
      Woman : 99.19222593307495
      Man : 0.8077717386186123
    }
    dominant_gender : Woman
    race : {
      asian : 1.5917524695396423
      indian : 2.048054151237011
      black : 0.41932323947548866
      white : 51.31601095199585
      middle eastern : 21.72670066356659
      latino hispanic : 22.89816588163376
    }
    dominant_race : white
  }
]
```

Obrázek 11: Výstup analýzy sentimentu z fotografie pomocí služby DeepFace



Obrázek 12: Vybrané záběry z rozhovoru s Klárou Jirovcovou Pospíšilovou pro účely analýzy emocí

Tabulka 3: Výsledky analýzy sentimentu pomocí služby DeepFace

	Obrázek č. 1	Obrázek č. 2	Obrázek č. 3
Štěstí	96,63987795831228	0,68055395968258	87,97074556350708
Strach	0,00631564105078	35,96886694431305	0,00555760198039
Hněv	0,25090008777080	26,39312744140625	0,45052529312670
Překvapení	0,00008352116224	3,07525638490915	0,03243505198042
Smutek	0,23683837585587	19,89728063344955	1,48170161992311
Zhnusení	0,00004441981541	13,50337117910385	0,00053241942623
Neutrální	2,86593756932764	0,48154592514038	10,05850732326507
Dominantní emoce	šťěstí	strach	šťěstí

Řeč těla

Důležitou součástí neverbální komunikace jsou i další tělesné projevy, které dokážou o náladě a vztahu k okolí mnohé prozradit. Gestikulace rukou může nahradit nebo doplnit mluvený projev. Pro úplné pochopení rozhovoru je interpretace tělesných signálů a gest velmi důležitá. Některá gesta, například třepotání rukou, mohou být spojena s úzkostí nebo depresí, dotýkání se obličeje může zase signalizovat zmatení nebo soustředění. Štěstí a vzrušení se projevují aktivnějšími pohyby, zatímco smutek je spojen s příkrčeným tělem a sklíčenou hlavou. Postoje těla, směr pohledu a gesta odhalují postoj narátora a naznačují projevy souhlasu či nesouhlasu.¹³⁸ Široce používaným nástrojem pro tyto účely je například software *OpenPose*¹³⁹. Ten automaticky vytvoří trojrozměrný model lidské kostry, který poté modeluje na člověka ve videonahrávce a umožňuje měřit a

138 PESSANHA, Francisca a SALAH, Almila Akdag. 2021. *A Computational Look at Oral History Archives*, s. 8.

139 Podrobnější informace například zde: <https://viso.ai/deep-learning/openpose/>

analyzovat postoj, pohyby těla, gesta (počet, razanci), rychlost pohybu a tak podobně.¹⁴⁰ Užití této techniky v orální historii je ovšem omezené a problematické. Opět vyvstává otázka, zda je etické analyzovat nevědomé tělesné projevy narátorů. Druhý problém je ryze praktický. Jelikož se orálně-historické rozhovory odehrávají většinou vsedě u stolu a na video nahrávce je vidět pouze horní část těla, musí se tato analýza omezit pouze na gestikulaci rukou.

Detekce dýchání a srdečního tepu

Dýchání a srdeční tep se s různými emocemi mění, i tyto projevy lze detekovat a analyzovat. Vědecké studie ukázaly souvislost mezi emocemi jako radost, hněv, strach a smutek a zpětnou vazbou na ně v podobě změny dýchání a změny frekvence srdečního tepu. Například vyprávění bolestných událostí může způsobit zřetelnou změnu v dýchání. Vzpomínka na traumatickou událost může vyvolat její „znovuprožívání“.¹⁴¹ Povzdechy vyjadřují úlevu, touhu nebo nudu a pomáhají získat kontrolu nad dýcháním. Smích také ovlivňuje dýchání a vede k velkým výdechům, které pomáhají snižovat vyvolaný stres, což může vysvětlovat použití humoru v některých životopisných příbězích o traumatických událostech. I zde je ovšem zapotřebí brát v úvahu, že vyjadřování emocí není univerzální a v rámci různých kultur se může odlišovat. Vliv mohou mít i věk, fyzická kondice či pohlaví narátorů.¹⁴²

Opět se jedná o eticky problematickou analytickou metodu a její použití v praxi by bylo komplikované. Přesné měření by totiž během rozhovoru vyžadovalo napojení narátora na detekční přístroje, což si lze jen těžko v praxi představit. Alternativou je analýza archivních záznamů opět s využitím nástrojů poháněných umělou inteligencí. V tomto případě je možné vybírat z nástrojů oboru takzvaného počítačového vidění (*computer vision*), které dokáže frekvenci dýchání a srdečního tepu odvodit z video záznamu rozhovoru na základě hýbání hrudníku a pohybu tepen a žil na obličeji, krku a rukou. Ovšem, stejně jako relevance a přesnost měření, je i praktické využití v orální historii v tomto případě diskutabilní.

4.5 Archivace dat

V oblasti orální historie je ochrana citlivých dat narátorů velmi důležitá, i zde může být pro zlepšení zabezpečení těchto dat využita umělá inteligence. AI lze například použít pro identifikaci a

140 PESSANHA, Francisca a SALAH, Almila Akdag. 2021. *A Computational Look at Oral History Archives*, s. 8-9.

141 Tamtéž, s. 9

142 Tamtéž, s. 10

anonymizaci osobních údajů v rozhovorech, což zvyšuje ochranu soukromí jednotlivců.¹⁴³ Například s využitím NLP nástrojů lze identifikovat a zamaskovat veškeré výskyty osobních jmen, adres a dalších citlivých informací v prepisech rozhovorů. Do procesů archivace a katalogizace mohou být integrovány nástroje, které by mohly poskytovat varování nebo upozornění na potenciálně nespolehlivý obsah. Dalším klíčovým aspektem je zabezpečení datových úložišť. Lze implementovat bezpečnostní a monitorovací systémy s využitím AI, které detekují a reagují na pokusy o neoprávněný přístup do systému nebo jinou podezřelou aktivitu. Systémy založené na AI mohou vyhodnocovat velké objemy dat a identifikovat potenciální slabiny v zabezpečení. To umožňuje podniknout preventivní kroky k zabezpečení datových úložišť a zamezit možným únikům informací. AI může také pomoci s automatizovaným šifrováním dat, aby byly citlivé informace chráněny i v případě, že dojde k jejich úniku během bezpečnostního incidentu.¹⁴⁴ Zároveň je ale nezbytné zajistit, aby použití AI pro zabezpečení dat nevytvářelo další rizika.

V kontextu orální historie je nezbytné se zaměřit na zabezpečení dat, a to jak při jejich vytváření, tak i při jejich ukládání a sdílení. Začínáme u samotné nahrávky, která je po ukončení rozhovoru uložena na nahrávacím zařízení nebo případně na mobilním telefonu. Zde je klíčové zabezpečit audio soubor proti neautorizované manipulaci, přehrání nebo smazání. Toho lze dosáhnout šifrováním dat či opatřením datového souboru digitálním podpisem nebo permanentním identifikátorem¹⁴⁵, který zajistí jeho integritu.¹⁴⁶ Pro mobilní zařízení existují různé nástroje, jako je například aplikace *Folder Lock*, která poskytuje možnost uzamknout a šifrovat soubory uložené na mobilním zařízení. Pro šifrování souborů uložených na desktopových zařízeních lze také vybírat z mnoha možností v závislosti na používaném operačním systému. OS Windows disponuje pro šifrování disku již integrovaným nástrojem BitLocker. Apple pro systém macOS nabízí obdobnou službu FileVault a Linuxové distribuce zase nástroj LUKS (Linux Unified Key Setup). Kromě toho existují mnohá další univerzální řešení. Lze doporučit například šifrovací software Cryptomator. Jedná se o open-source řešení, které je vyvíjeno od roku 2017, je distribuováno zdarma a dostupné pro všechny hlavní desktopové (Windows, macOS, Linux) i mobilní platformy.¹⁴⁷

143 JAILLANT, Lise a Arran REES. *Applying AI to digital archives: trust, collaboration and shared professional ethics*, s. 571.

144 Tamtéž, s. 571.

145 Viz kapitola: 3.2 Digitalizace, kde byl persistentní identifikátor zmíněn v souvislosti s FAIR daty. Persistentní identifikátor (např. DOI) slouží k identifikaci digitálních objektů na internetu. Samotný digitální objekt je zde doplněn o množství metadat, která ho popisují. Nechybí ani

146 Viz například UNIVERZITA KARLOVA. *Perzistentní identifikátory*. Centrum pro podporu open science [online]. aktualizováno 17. 10. 2022 [cit. 2024-03-03]. Dostupné z: <https://openscience.cuni.cz/OSCI-193.html>

147 CRYPTOMATOR. *Cryptomator - Free Cloud Encryption for Dropbox & Co* [online]. [cit. 2024-04-14]. Dostupné z: <https://cryptomator.org/>

Dalším krokem je zabezpečení dat při jejich přenosu. Pro přenos souborů a jejich sdílení existuje také mnoho možností, jako jsou e-maily, chatovací aplikace či cloudové služby typu *Dropbox*¹⁴⁸ nebo *Google Drive*. Každý z těchto nástrojů má svá specifika a s tím spojené bezpečnostní hrozby. Je tedy důležité zvážit rizika spojená s používáním konkrétního nástroje a zajistit opatření pro minimalizaci potenciálních hrozeb, jako je například použití šifrování během přenosu a ověřování pravosti přijímaných dat. Na šifrovanou a zabezpečenou komunikaci se zaměřuje například mobilní aplikace *Signal*¹⁴⁹, která umožňuje šifrovat jak textovou komunikaci, tak i odesílané soubory. Pro zabezpečení dat a zajištění kybernetické bezpečnosti lze v akademické sféře (výzkumné instituce, univerzity) využívat služeb sdružení CESNET.¹⁵⁰

Poslední fází je trvalé uložení souborů v digitálním archivu. Software digitálního archivu by měl být vybaven robustním bezpečnostním systémem, který chrání data před krádeží, neautorizovanými změnami a smazáním. Důležitá je pravidelná aktualizace softwaru, udržování aktivní antivirové ochrany a pečlivá správa přístupových práv. Správce systému nese hlavní zodpovědnost za bezpečnost, ale i každý uživatel má svůj díl odpovědnosti a zodpovídá za své přihlašovací údaje a ručí za to, že je neposkytne neautorizované osobě. Vzhledem k citlivosti orálně-historických dat je zabezpečení jednou z velmi důležitých oblastí a je nutné ji neustále sledovat, udržovat a zlepšovat.¹⁵¹

Bez ohledu na to, zda byl rozhovor s narátorem nahrán prostřednictvím mobilního telefonu nebo digitálního nahrávacího zařízení, je poté v případě řádného orálně-historického výzkumu zpravidla nahrán do archivního systému, který slouží k archivaci, správě a analýze orálně-historických dat. Do archivního systému lze společně s nahrávkou vložit metadata, která ji blíže popisují (jména narátorů, tazatelů, probíraná témata, klíčová slova, místo a datum pořízení nahrávky apod.).¹⁵² Některé moderní archivní systémy ovšem již využívají AI nástroje a umožňují automatické generování těchto dat. I těchto systémů existuje na trhu velké množství a každý z nich disponuje jinými funkcemi, odlišnými možnostmi použití, primárním zaměřením, cenou, softwarovou podporou. Pro příklad jsem si vybral dva velmi rozdílné a velmi rozšířené systémy.

148 Podrobnější informace dostupné z: <https://www.dropbox.com/>

149 Viz například SNOW, John. *Signal is secure, as proven by hackers*. Kaspersky official blog [online]. [cit. 2024-03-20]. Dostupné z: <https://www.kaspersky.com/blog/signal-hacked-but-still-secure/45273/>

150 Podrobnější informace dostupné z: <https://www.cesnet.cz/sluzby/bezpecnost-4>

151 ORAL HISTORY ASSOCIATION. *Archiving Oral History*. Oral History Association [online]. 2019 [cit. 2024-01-19]. Dostupné z: <https://oralhistory.org/archives-principles-and-best-practices-overview/>

152 Tamtéž

Systém *OHMS (Oral History Metadata Synchronizer)*¹⁵³ vznikl ve středisku *Louie B. Nunn Center for Oral History* na *University of Kentucky Libraries*. První verze vznikla již v roce 2008 a systém byl navržený pro zlepšení přístupu k orálně historickým záznamům, snadnému vyhledávání a správě metadat a je distribuován zcela zdarma.¹⁵⁴ Za vývojem OHMS stojí poměrně známý orální historik a odborník na archivaci a digitální technologie Douglas A. Boyd¹⁵⁵. Nyní je tento systém již značně zastaralý a v září 2023 proběhla integrace tohoto systému s modernějším systémem *Aviary Platform*¹⁵⁶, který systém OHMS zcela nahradil. Před ukončením vývoje byl však systém OHMS stále ještě využíván 700 aktivními uživateli ze 60 zemí celého světa. Lze tedy předpokládat, že na tomto a jemu podobných systémech dodnes závisí mnoho archivů po celém světě. Nový systém *Aviary* je moderní a využívá některých výše zmíněných AI nástrojů (automatický přepis rozhovorů, generování metadat). Na rozdíl od OHMS se nezaměřuje výhradně na správu archivů orální historie, ale jde o obecný nástroj pro správu jakéhokoliv audiovizuálního obsahu, který nemusí ve všech ohledech splňovat specifické nároky orální historie.¹⁵⁷ Přestože *Aviary* platforma je placenou službou, na modul OHMS se vztahuje výjimka a základní plán je stále zcela zdarma. Je ovšem limitován na pouhých 10 zdrojů (rozhovorů). Po naplnění limitu je nutné přejít na placený plán¹⁵⁸.

Alternativním řešením je například archivní systém *Aureka.AI*. Jedná se o moderní archivní systém pocházející z prostředí *Freie Universität* v Berlíně a lze ho použít pro potřeby archivů, knihoven i muzeí. Umožňuje pro audio a video nahrávky vytvořit automatický přepis rozhovoru, vygenerování metadat pomocí AI, překlady do jiných jazyků, kontroluje pravopis vložených textů a umožňuje integraci s mnoha dalšími systémy.¹⁵⁹ Jak je patrné na ilustračním záznamu obrazovky aplikace *Aureka*, aplikace poskytuje automatický přepis rozhovoru a dokáže i odlišit, kolik osob celkem v nahrávce mluví a automaticky přiřadí mluvený text ke konkrétnímu mluvčímu. Přepis lze dále manuálně upravovat, pokud by v něm byly chyby, a lze do něj také manuálně vkládat další poznámky. Jedná se opět o placené řešení, ale ceník služeb není zveřejněn a konečná cena je vypočítána až podle požadovaného řešení.

153 Viz LOUIE B. NUNN CENTER FOR ORAL HISTORY. UNIVERSITY OF KENTUCKY LIBRARIES. *OHMS Oral History Metadata Synchronizer* [online]. Dostupné z URL: <https://www.oralhistoryonline.org/>

154 BOYD, Doug, 2013. *OHMS: Enhancing Access to Oral History for Free*. *The Oral History Review*. [online]. 40(1), 95-106. DOI: 10.1093/ohr/oht031. Dostupné z: <https://www.tandfonline.com/doi/full/10.1093/ohr/oht031>, s. 3.

155 Viz například ASHENFELDER, Mike. *Doug Boyd and the Power of Digital Oral History in the 21st Century*. *The Signal* [online]. 2013 [cit. 2024-03-17]. Dostupné z: <https://blogs.loc.gov/thesignal/2013/01/doug-boyd-and-the-power-of-digital-oral-history-in-the-21st-century/>

156 Viz AVIARY. *Aviary Platform* [online]. Dostupné z URL: <https://www.aviaryplatform.com/>

157 BOYD, Douglas. *Announcement: Integrating OHMS and Aviary*. *Digital Omnium* [online]. 2023 [cit. 2024-01-06]. Dostupné z: <https://digitalomnium.com/announcement-integrating-ohms-and-aviary/>

158 Podrobnější informace včetně ceníku služeb jsou dostupné zde: <https://www.aviaryplatform.com/pricing#ohms-aviary-subscription>

159 AUREKA. *AI Cataloguing Assistant*. *Aureka - The AI Research Assistant* [online]. 2023 [cit. 2024-01-13]. Dostupné z: <https://www.aureka.ai/solution-for-archives-libraries-museums>

2022-12-15_karin_baume

Search with aureka.ai

TRANSCRIPT ANNOTATIONS METADATA TRANSLATIONS EXPORT

4 Pages 1 Article 1 Article 1 Exports 4 Exports

Low transcription accuracy

Annotation for 00:00:10

Topic

Annotation

Sprecherin 1 : Ich bin 1955 in Schwerin geboren, als erstes Kind einer riesen Großfamilie, die nach dem Zweiten Weltkrieg aus Osteuropa kam, Litauen und Masuren, und für die der Aufbau des Sozialismus ein avantgardistisches Projekt war. Avantgardistisch die nicht gekannt, den Begriff den sag ich jetzt, aber die haben alles auf den Prüfstand gestellt. Mein Vater als Bauernjunge war in russischer Gefangenschaft, kam als Kommunist wieder und hat die Enteignung der Großgrundbesitzer mitorganisiert und es war in meiner Familie ein ganz großes Privileg, dass meine Kita im Schloss war, wo die Großgrundbesitzer verjagt wurden. Ja, der Sozialismus ist sozusagen meine Herkunft. Natürlich war ich contra, wie alle Kinder, zu dem was sie bekommen. Ich bin die erste in meiner Familie die studiert hat und ich hab Soziologie studiert, und ich habe in der DDR, ich dann schon das erste Mal schwanger war – etwas typisches in der DDR – Stadtsoziologie der Bundesrepublik bearbeitet.

Sprecherin 1 : Das hatte damit zu tun, dass ich schon schwanger war und Bibliotheksarbeit gemacht habe, und weil die Westliteratur in der DDR sehr viel gelesen wurde im Osten also bei Architekten, Stadtplanern, Soziologen, die gabs ja auch nicht für jeden, aber für jeden, der dazu studiert hat. Und ich habe meine eigenen Gegenstandsauffassung der Stadtsoziologie entwickelt, später auch in Weimar bei Architekten und Stadtplanern gearbeitet und gelehrt, so dass ich einen ganz guten Blick hatte, als die Mauer fiel, weil das mein Berufsfeld war eine kritische Auseinandersetzung mit der westdeutschen Stadtsoziologie. Mein Lehrer in Schwerin wollte das gar nicht zulassen Soziologie, aber ich hatte schon geschrieben an die Uni, hatte schon ein Angebot, das gab es, und insofern konnte ich da einfach anfangen. Bei der Aufnahmeprüfung hier in Berlin wurde ich gefragt, wie das denn mit dem Ost West Vergleich finde, also sehe, da dachte ich, oh, das hier ist nicht meine Uni und habe ziemlich viel Leder gezogen, dass man das nicht vergleichen kann.

Sprecherin 1 : Alles hat Stärken und Schwächen. Dafür wurde ich genommen für diese Antwort. Also die Soziologie der DDR Kleiner Leute war doch relativ kritisch, hatte mit den Wurzeln zu tun, geht auf polnische Wissenschaftler zurück, die marxistische Soziologie generell. Die Geschichtsschreibung zur DDR ist ja eh auch eine Geschichtsschreibung der Sieger der Wiedervereinigung, deswegen gibt es viele Dinge, die nicht im Mainstream der Interpretation liegen, aber die in der Einzelbiografie und da bin ich, glaube ich, keine Ausnahme nochmal ganz anders gelesen werden können als die offizielle Lesart. Na ja, für mich war das also eine ganz tolle Kindheit, weil, erst wenn man älter ist, kann man das wertschätzen, normal, aber aus heutiger Perspektive kann ich das sehr wertschätzen.

Sprecherin 1 : dass diese osteuropäisch bäuerlichen Werte, zum Beispiel, es gab nie ein Thema mit Geld, es war Kleingeld da, aber es natürlich nicht viel da. Dieses Kollektive, also das war ja auch ein Pendant in der Großfamilie, in der bäuerlichen, es gab großes Körperbewusstsein in osteuropäischen einzelstehenden Gehöften, weil sonst hätte man nicht überlebt, und all das im Sozialismus möglich, also diese Kollektivierung entsprach eigentlich auch den Werten meiner Großfamilie, nur dass es immer darum in eine gewissen industriellen Produktionsgesellschaft kam, um sie die Wertung der

Obrázek 13: Uživatelské rozhraní archivního systému Aureka.AI
 [https://www.aureka.ai/solution-for-archives-libraries-museums]

4.6 Případová studie: 100 studentských revolucí

Závěrečná případová studie vzájemně kombinuje jednotlivé technologie, které byly podrobně popsány v předchozích kapitolách. Cílem je analyzovat veškeré rozhovory, které byly nahrány a zveřejněny v rámci orálně-historického projektu *100 studentských revolucí*. K tomu jsou použity NLP nástroje pro tokenizaci a lemmatizaci textu, pro identifikaci pojmenovaných entit a pro analýzu sentimentu. K tomu jsou opět, stejně jako v předešlých příkladech, použity skripty napsané v programovacím jazyku Python. Cílem je provést analýzu česky vedených rozhovorů, a to efektivně, levně (ideálně zcela zdarma), jednoduše a bezpečně. Předstírejme, že rozhovory nebyly zveřejněny na webu a že z bezpečnostních důvodů rozhovory nechceme odesílat k analýze na cizí servery, čímž by hrozilo riziko jejich odcizení a zneužití.

V prvním kroku je tedy nutné nejprve nahrávky rozhovorů stáhnout z platformy YouTube, kde jsou uloženy. K tomu použijeme softwarovou knihovnu PyTube, kterou stáhneme do počítače vykonáním příkazu v příkazové řádce.

```
1. pip install pytube
```

Poté spustíme Python skript, který veškeré nahrávky stáhne. Jelikož pro analýzu nepotřebujeme video nahrávky, postačí nám pouze jejich audio stopy (ve formátu .mp4).

```
1. from pytube import YouTube
2. recordUrls = ["agQ8N10_Dz8", "CH5HNGLfUWU", "BfSn8TLZeTA"] # ...zkráceno (celkem 261 nahrávek)
3. for record in recordUrls:
4.     yt = YouTube("https://www.youtube.com/watch?v=" + record)
5.     audio = yt.streams.filter(only_audio=True).first()
6.     audio.download()
```

Nyní je nutné pořídit přepis stažených zvukových nahrávek. K tomu použijeme službu Whisper¹⁶⁰, která, jak již bylo zmíněno, na lokálním počítači funguje bez připojení k internetu. V tomto případě je tedy bezpečná, zcela zdarma a zároveň při použití největšího dostupného jazykového modelu poskytuje poměrně velmi kvalitní přepis česky mluvených nahrávek. Pro stažení knihovny opět použijeme příkazovou řádku.

```
1. pip install -U openai-whisper
```

Nyní, když máme nahrávky, je načase udělat jejich přepisy do textové podoby pomocí následujícího Python skriptu.¹⁶¹

```
1. import whisper
2. import os
3.
4. model = whisper.load_model("large")
5.
6. folder_path = './Rozhovory/'
7. # Prohledání všech souborů ve složce
8. for file_name in os.listdir(folder_path):
9.     # Kontrola formátu souboru – chceme přepisovat soubory ve formátu MP4
10.    if file_name.endswith('.mp4'):
11.        file_path = os.path.join(folder_path, file_name)
12.
13.        audio = whisper.load_audio(file_path)
14.        audio = whisper.pad_or_trim(audio)
15.        mel = whisper.log_mel_spectrogram(audio).to(model.device)
16.
17.        # Rozpoznání jazyka
18.        _, probs = model.detect_language(mel)
19.        print(f"Rozpoznán jazyk: {max(probs, key=probs.get)}")
20.
```

160 Podrobnější informace dostupné z: <https://openai.com/research/whisper> a také z: <https://github.com/openai/whisper>

161 Python skript s podrobnějšími informacemi se nachází například zde: <https://github.com/openai/whisper>

```

21. options = whisper.DecodingOptions()
22. # Přepis nahrávky
23. result = whisper.decode(model, mel, options)
24.
25. # Vypsání výsledku přepisu
26. print(result.text)
27.
28. # Uložení přepisu do textového souboru
29. with open(file_name + ".txt", "w", encoding="utf-8") as txt:
30.     txt.write(result.text)

```

Skript pro každou nahrávku vygeneruje textový soubor obsahující přepis.

```

1. import os
2.
3. # Cesta ke složce, kde jsou uloženy .txt soubory
4. source_directory = "/Rozhovory/"
5. # Cesta a název výstupního souboru
6. output_file_path = "/Prepisy.txt"
7.
8. # Inicializace seznamu pro uložení obsahu všech souborů
9. all_files_content = []
10.
11. # Procházení všech souborů ve složce
12. for filename in os.listdir(source_directory):
13.     if filename.endswith(".txt"): # Zpracování pouze souborů ve formátu TXT
14.         with open(os.path.join(source_directory, filename), 'r', encoding='utf-8') as input_file:
15.             # Přidání obsahu vstupního souboru do seznamu
16.             all_files_content.append(input_file.read())
17.
18. # Otevření výstupního souboru v režimu pro zápis a zápis všech souborů najednou
19. with open(output_file_path, 'w', encoding='utf-8') as output_file:
20. # Přidání nového řádku mezi soubory
21.     output_file.write("\n".join(all_files_content))

```

Na konci této fáze máme vygenerovaný textový soubor, který obsahuje všechny přepisy všech stažených rozhovorů.

Tabulka 4: Statistika automatických přepisů rozhovorů

Celkem narátorů / narátorek	90
Celkem nahrávek rozhovorů	261
Délka nahrávek ¹⁶²	10 hodin 20 minut
Počet přepsaných slov	78 609
Počet přepsaných znaků	433 756
Doba přepisu	46 hodin

Jak již bylo řečeno v kapitole věnované automatickým přepisům rozhovorů, generování automatického přepisu je velmi náročné na výpočetní výkon počítače. V této případové studii je to velmi patrné, když se pro nahrávky, které v souhrnu trvají 10 hodin a 20 minut, generoval jejich automatický přepis celkem 46 hodin. Jaký je zde rozdíl oproti manuálně pořízenému přepisu profesionální přepisovatelky (či přepisovatele), která dokáže mnohem kvalitnější transkript pořídit ve srovnatelném čase? Je zde nutné vzít v úvahu několik faktorů.

Počítač je schopen pracovat (kromě spotřeby elektrické energie)¹⁶³ 24 hodin denně zdarma. Dlouhá doba přepisu je determinována tím, že přepis byl generován na kancelářském laptopu a pomocí procesoru¹⁶⁴, který není pro takovéto výpočetní operace vhodný. Pokud by byl přepis pořízován pomocí (i výkonově slabší) grafické karty, byla by doba přepisu řádově menší. Pokud by byla použita výkonná grafická karta, kterou disponují například herní počítače, doba přepisu by se opět velmi výrazně zkrátila. V neposlední řadě má na výkon obrovský vliv i výběr jazykového modelu. V tomto případě byl použit model Whisper Large, který poskytuje nejlepší výsledky, ale například v porovnání s modelem Small pracuje 6krát pomaleji. V porovnání s modelem Base je pomalejší 16krát a oproti nejmenšímu modelu Tiny dokonce 32krát.¹⁶⁵ Nutno však podotknout, že transkripty pořízené pomocí modelů menších, než je model Large, jsou pro česky vedené rozhovory a nepřiliš kvalitní nahrávky z tohoto projektu téměř nepoužitelné. Při zachování kvality lze menší modely použít pouze pro kvalitní nahrávky v angličtině.

162 Každá z nahrávek má různou kvalitu i délku. Nejkratší nahrávky mají pouze kolem deseti vteřin, zatímco ty nejdelší se blíží deseti minutám.

163 V případě kancelářského notebooku je spotřeba elektrická energie poměrně nízká. Notebook střední cenové třídy spotřebuje při průměrném provozu 5 hodin denně přibližně 27 kWh za rok (~128 Kč v roce 2023), viz <https://www.elektrina.cz/pocitac-versus-notebook-zjistete-jak-se-lisi-naklady-na-elektrinu>.

164 Přepis byl prováděn pomocí procesoru Intel® Core™ i7-8750H CPU @ 2.20GHz.

165 Viz <https://github.com/openai/whisper>.

Nyní zjistíme, která slova a termíny se v rozhovorech nejčastěji vyskytují. Na rozdíl od knihovny SpaCy, jejíž použití bylo ukázáno v kapitole věnující se NLP nástrojům, zde je pro česky vedené rozhovory (resp. jejich česky psané přepisy) nutné použít knihovnu, která dokáže pracovat s českým jazykem. V tomto případě je použita knihovna UDPipe, pro kterou je dostupný český jazykový model¹⁶⁶. Pro stažení knihovny opět použijeme příkazovou řádku.

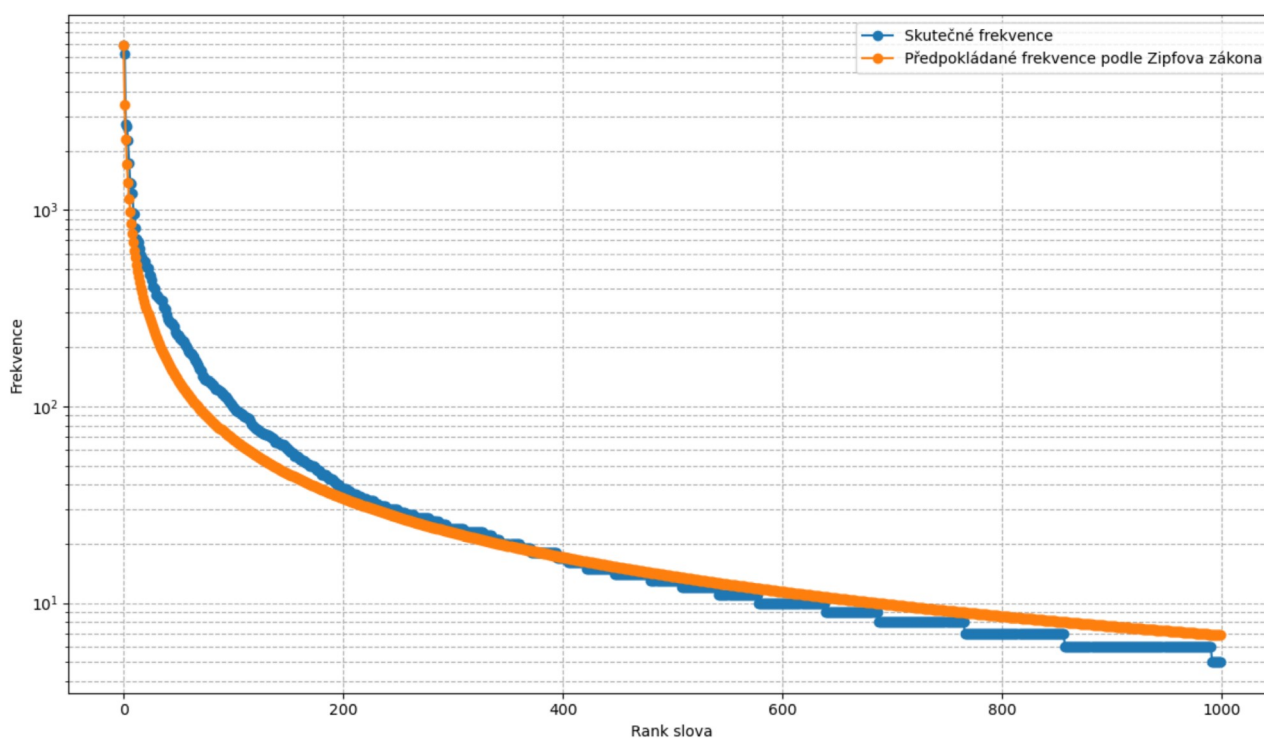
```
1. pip install ufal.udpipe
```

Následně již můžeme spustit následující skript.

```
1. from ufal.udpipe import Model, Pipeline
2. from collections import Counter
3.
4. # Načtení českého modelu
5. model_path = "./czech-pdt-ud-2.5-191206.udpipe"
6. model = Model.load(model_path)
7. pipeline = Pipeline(model, 'tokenize', Pipeline.DEFAULT, Pipeline.DEFAULT, 'conllu')
8.
9. # Načtení textu
10. with open('./Prepisy.txt', 'r', encoding='utf-8') as file:
11.     texts = file.read()
12.
13. # Zpracování textu
14. processed = pipeline.process(texts)
15.
16. # Definice nevýznamových slovních druhů k odfiltrování
17. ignore_pos = {'PUNCT', 'PART', 'CCONJ', 'SCONJ', 'PRON', 'DET', 'ADP', 'ADV'}
18.
19. # Vytvoření slovníku pro frekvenci slov a slovních druhů, které nejsou ignorovány
20. word_freq = Counter()
21. for line in processed.strip().split("\n"):
22.     if not line.startswith('#') and len(line.split("\t")) > 3:
23.         lemma, partOfSpeech = line.split("\t")[2:4]
24.         if partOfSpeech not in ignore_pos:
25.             word_freq[lemma] += 1
26.
27. # Získání 100 nejčastěji se vyskytujících slov
28. common_words = word_freq.most_common(100)
29.
30. # Vypsání 100 nejčastěji se vyskytujících slov
31. for word, freq in common_words:
32.     print(f"{word}: {freq}")
```

166 UDPipe je česká NLP knihovna, vyvinutá na Ústavu formální a aplikované lingvistiky - MFF UK. Podrobnější informace a knihovna ke stažení jsou dostupné zde: <https://lindat.mff.cuni.cz/repository/xmlui/handle/11234/1-3131>

Tento skript byl spuštěn ještě podruhé, ale tentokrát bez filtrování bezvýznamových slov (předložky, spojky). Výslednou četnost slov jsem pomocí Python knihovny *NumPy Zypf*¹⁶⁷ zanesl do grafu a porovnal s křivkou určenou podle *Zipfova zákona*¹⁶⁸. Skutečnost, že jsou křivky relativně podobné, nepřímo dokazuje, že automatický přepis a lemmatizace jednotlivých slov proběhla úspěšně. Pokud by některý z těchto algoritmů generoval zcela nesmyslné výsledky, viditelně by se to projevilo na podobě tohoto grafu.



Obrázek 14: Zastoupení slov v textu a předpokládaná frekvence podle Zipfova zákona

Po odfiltrování bezvýznamových slov získáme následující seznam nejčastěji se vyskytujících slov v rozhovorech.

Tabulka 5: Četnost nejčastěji se vyskytujících slov v rozhovorech

být	6207	mezník	117	den	64	student	45
mít	972	první	108	žít	63	změna	44
člověk	542	dobrý	106	problém	62	práce	44
rok	356	generace	99	svoboda	59	stejný	43
myslit	347	mladý	94	hodnota	59	země	43
vědět	294	dát	93	situace	58	znamenat	43

167 Podrobnější informace dostupné z:

<https://numpy.org/doc/stable/reference/random/generated/numpy.random.Generator.zipf.html>

168 Zipfův zákon vyjadřuje vztah mezi pořadím slov přirozeného jazyka uspořádaných podle frekvence jejich výskytů. Viz například SPÍCHAL, Luděk. *Zipfův zákon a další mocninné zákony* [online]. 2021 [cit. 2024-03-18]. Dostupné z: https://www.researchgate.net/publication/351841475_Zipfuv_zakon_a_dalsi_mocninne_zakony

říkat	279	dva	91	bavit	56	smysl	42
moci	264	druhý	91	různý	56	představit	41
věc	254	svět	87	osobní	56	cesta	40
říci	231	listopad	87	zkušenost	54	názor	40
muset	216	způsob	87	rodina	53	cítit	39
chtít	202	celý	81	možnost	53	rodič	38
dělat	199	mluvit	80	vysoký	53	schopný	38
dítě	191	dostat	78	zajímat	52	řešit	38
jít	188	společnost	77	řád	51	spokojený	37
život	165	důležitý	77	obrovský	50	určitý	36
škola	152	začít	75	vnímat	50	pamatovat	36
velký	152	revoluce	70	podstata	48	špatný	36
vidět	142	další	69	umět	47	snažit	36
jiný	140	spousta	69	starý	47	politika	36
jeden	136	změnit	68	dokázat	47	učit	36
doba	134	teďka	66	chodit	46	sociální	35
pocit	132	strana	66	zajímavý	45	otázka	35
přijít	125	fakt	64	čas	45	politický	35
stát	121	udělat	64	jasný	45	měnit	34

1. `pip install textblob`
2. `pip install googletrans==4.0.0-rc1`

Následující kód rozdělí texty na jednotlivé věty, každé větě přiřadí index sentimentu. Ten se pohybuje v intervalu $< -1,1 >$, přičemž hodnota (-1) představuje absolutně negativní větu, (0) neutrální a hodnota (1) větu absolutně pozitivní.

```

1. from textblob import TextBlob
2. from googletrans import Translator, LANGUAGES
3. import nltk
4. import numpy as np
5. from nltk.tokenize import sent_tokenize
6.
7. translator = Translator()
8. nltk.download('punkt')
9.
10. with open('/Prepisy.txt', 'r', encoding='utf-8') as file:
11.     texts = file.read()
12.
13. sentiments = []
14. sentences = sent_tokenize(' '.join(texts))
15. for sentence in sentences:
16.     # Překlad věty do angličtiny
17.     text_en = translator.translate(sentence, dest="en").text
18.     # Analýza sentimentu přeloženého textu
19.     sentiment = TextBlob(text_en).sentiment
20.     sentiments.append(sentiment.polarity)

```



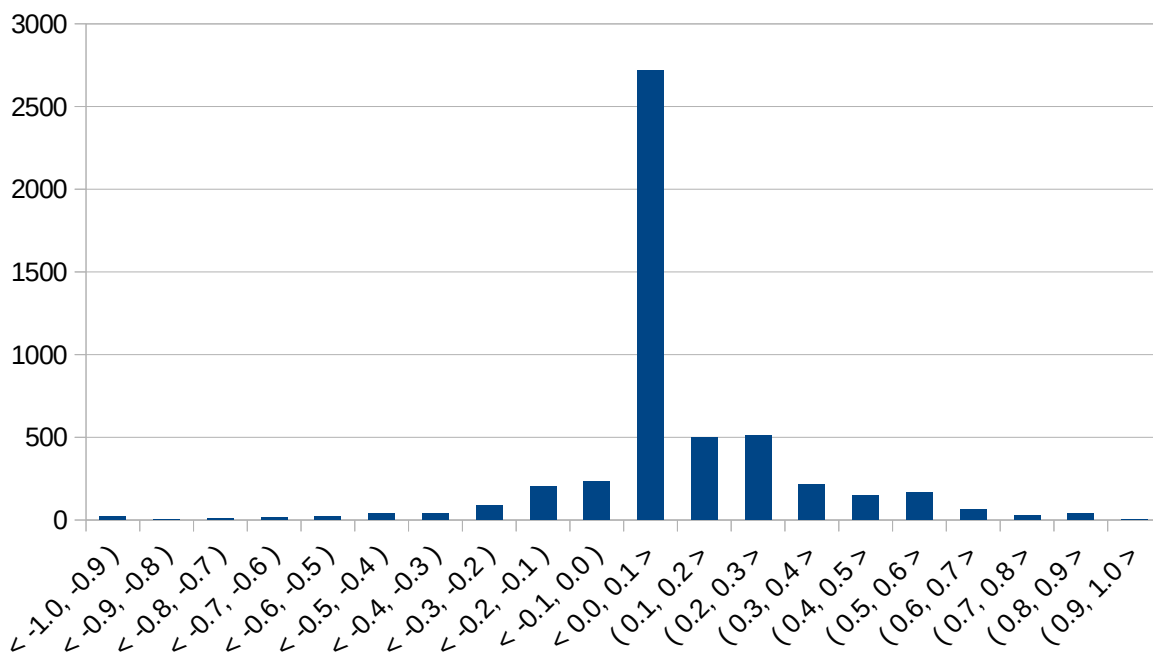
```

21.
22. intervals = np.arange(-1, 1.1, 0.1)
23. counts = np.histogram(sentiments, bins=intervals)[0]
24. interval_counts = {f"{intervals[i]} to {intervals[i+1]}": count for i, count in enumerate(counts)}
25. print(interval_counts)

```

Tabulka 6: Počet vět podle indexu jejich sentimentu

Sentiment	Interval	Počet vět	Počet vět
Pozitivní	(0.9, 1.0 >	2	657
	(0.8, 0.9 >	42	
	(0.7, 0.8 >	28	
	(0.6, 0.7 >	62	
	(0.5, 0.6 >	164	
	(0.4, 0.5 >	146	
	(0.3, 0.4 >	213	
Neutrální	(0.2, 0.3 >	513	4254
	(0.1, 0.2 >	499	
	< 0.0, 0.1 >	2718	
	< -0.1, 0.0)	233	
	< -0.2, -0.1)	201	
	< -0.3, -0.2)	90	
Negativní	< -0.4, -0.3)	37	146
	< -0.5, -0.4)	38	
	< -0.6, -0.5)	23	
	< -0.7, -0.6)	14	
	< -0.8, -0.7)	9	
	< -0.9, -0.8)	2	
	< -1.0, -0.9)	23	



Obrázek 15: Počet vět podle indexu jejich sentimentu

Zde jsou některé ukázky vět, které byly ohodnoceny nejvyšším indexem sentimentu (nejpozitivnější věty).

- *Protože to opojení tou svobodou, to je pocit prostě k nezaplacení.*
- *Takže pro mě je to taková zásadní zkušenost a samozřejmě, když dnes vidím padat berlínskou zeď, tak je to pro mě velmi dojemný okamžik.*
- *To jsou asi ty nejlepší. Nejlepší, nejlepší mezníky.*
- *Ale samozřejmě ty pocity, to je velmi dobré, z toho, kolik lidí se sešlo, jaká tam byla nálada, jaké tam byly proslovy, včetně těch požadavků.*
- *Perfektně připravená prostě, jako vy.*

Zde jsou naopak ukázky vět s nejnižším hodnocením sentimentu.

- *To je strašně bezradné.*
- *Ta výuka dějepisu, ta historie je naprosto příšerná.*
- *To je strašně zneužitelný, je to strašná hrůza.*
- *A to jsme seděli v aule a na chodbě se ozval strašný jekot.*

Dále je možné v textu identifikovat pojmenované entity, o kterých již byla zmínka. Identifikaci je možné provést opět pomocí některé z NLP knihoven, je zde ovšem velký problém s podporou

českého jazyka a identifikací typicky českých entit (jmen osob, míst, organizací, dat). Tyto entity musí být obsaženy a anotovány ve zdrojových jazykových korpusech, které konkrétní služba využívá. Proto se mi jeví pro tuto analýzu vhodný český online nástroj NameTag¹⁶⁹. Nástroj disponuje jednoduchým webovým rozhraním a po nahrání zdrojového textu vypíše přehlednou tabulku identifikovaných pojmenovaných entit. Analýza přepisů rozhovorů ukázala, že v rozhovorech se nejčastěji objevují zeměpisná označení Evropa (25x), Praha (18x), rok 1989 (18x jako 1989 a 11x pouze jako 89) a v nejrůznějších kontextech se tu opakuje číslovka 20 (11x).

Mezi dalšími pojmenovanými entitami se dále s různou frekvencí objevují například:

- vlastní jména (*Václav Havel, Miloš Forman, Jack Nicholson, Jiří Kolář, Boris Jeltsin, Karel Kryl, Saša Vondra, Masaryk, Josefus Flavius, Zeman, Andrej Babiš, Regan, Gorbačov, Martin Mejstřík, Monika Pajerová, Martin Benda, Harry Potter, Kennedy, Palachův, Egon Bondy, Stanislav Gross, Šimon Pánek, Martin Klíma, Igor Chaun, Veronika Žilková, Petr Kotyk, Daniel Kroupa, Radim Palouš, Jan Opletal*),
- názvy geografických míst (*Čechy, Maďarsko, Paříž, Finsko, Švédsko, Británie, Norsko, Finsko, Rakousko, Francie, Jeruzalém, Československo, Národní třída, Ukrajina, Morava, Amerika, Bamberk, Karlův most, Albertov, Polsko, Kuba, Londýn, Bratislava, SRN, Jugoslávie, Etiopie, Austrálie*),
- data (*listopad, listopad 89, 1981, 2000, 1991, 17 listopad, 1918, 2013, 1848, 1968, 1938, 1994, 1993, 1995, 2001, 2000, 2006, 2002, 2010*),
- organizace (*Google, McDonald, Svaz československo-sovětského přátelství, ROH, StB, Facebook*).

Na závěr se vrátím k příkladu, který byl zmíněn v první části práce¹⁷⁰. Představme si, že o projektu *100 studentských revolucí* nevíme zhruba nic. V archivu se našlo pouze těchto 261 nepopsaných, nepřepsaných, neanalyzovaných nahrávek rozhovorů k nedokončenému projektu. Jedná se o desítky hodin nejrůznějších rozhovorů. Použitím několika těchto nástrojů a skriptů popsanych v této kapitole si lze poměrně snadno a levně vytvořit dobrou představu o tom, co nahrávky obsahují.

Mezi nejčastěji se vyskytujícími slovy najdeme například termíny jako *mezník, změna, mladý, generace, listopad, společnost, revoluce, strana, svoboda, hodnota, student, změna, politika*,

169 Nástroj NameTag je vyvíjen na Ústavu formální a aplikované lingvistiky - MFF UK. K analýze je možné využít český jazykový korpus (Czech Named Entity Corpus 2.0), obsahující 8993 česky psaných vět, které obsahují 35 220 manuálně anotovaných českých pojmenovaných entit, viz <https://ufal.mff.cuni.cz/cnec>. Webové rozhraní pro ovládání nástroje je dostupné z: <http://lindat.mff.cuni.cz/services/nametag/>

170 Příklad byl zmíněn v kapitole 4.3.3 Metadata.

politická. Časté používání těchto slov směřuje k tématu 17. listopadu 1989 poměrně jasně. Po identifikaci pojmenovaných entit může být tato prvotní domněnka téměř s jistotou potvrzena. Data a místa¹⁷¹ spojená s událostmi listopadu 1989, jména „*studentských vůdců revoluce*“¹⁷², postav Občanského fóra¹⁷³ a tehdejších světových politiků¹⁷⁴ se zde také neustále opakují.

Z ostatních pojmenovaných entit je ale patrné, že rozhovory se nezaměřují výhradně na několik listopadových dnů roku 1989. Nechybí ani výrazný historický exkurz do minulosti. Padají zde jména studentských martýrů Jana Palacha a Jana Opletala a důležité historické mezníky české a československé historie posledních dvou století¹⁷⁵. Zároveň jsou ale v rozhovorech zmíněna jména Miloš Zeman, Bamberk, McDonald, Jugoslávie, které si spojujeme spíše až s 90. lety dvacátého století. Zmíněn je ale také Andrej Babiš, Google, Facebook a zároveň roky 2000, 2001, 2002, 2006, 2010, z čehož lze vyvodit, že rozhovory nemohly být pořizeny bezprostředně po událostech listopadu 1989, ale až po více než 20 letech od těchto událostí. V případě, že jsou například pro potřeby našeho výzkumu relevantní pouze některá konkrétní jména osob, lze si v seznamu entit ověřit, zda je některá z těchto osob v rozhovorech vůbec zmíněna.

Další závěry lze vyvodit z analýzy sentimentu. Jelikož je naprostá většina rozhovorů zcela bez emočního zabarvení a ty zbylé jsou relativně rovnoměrně zastoupeny mezi pozitivními i negativními kategorie, lze také nepřímou usuzovat, že rozhovory nebyly nahrávány „ve víru událostí“, ale až s dostatečným odstupem od popisovaných událostí.

Pokud by ani tato základní analýza nestačila, lze také libovolnému jazykovému modelu poskytnout přepis celého rozhovoru a požádat o sepsání shrnutí, nebo provedení jiné libovolné akce nad textovým vstupem. Pro analýzu textu lze použít některý z populárních velkých jazykových modelů, ale také například dříve zmíněnou aplikaci GPT4All¹⁷⁶, která umožňuje používat některé jazykové modely (například Mistral, Hermes, Wizard, Orca) bezpečně na lokálním počítači, bez připojení k internetu a zcela zdarma. Opět postačí software pouze příkazem v příkazové řádce nainstalovat do počítače.

2. `pip install gpt4all`

171 Listopad, 1989, 89, 17 listopad, Albertov, Národní třída, Československo.

172 Saša Vondra, Martin Mejstřík, Monika Pajerová, Martin Benda, Šimon Pánek, Martin Klíma, Igor Chaun.

173 Václav Havel, Radim Palouš.

174 Bush, Gorbačov, Regan.

175 1848, 1918, 1938, 1968.

176 Podrobnější informace například z: <https://medium.com/@kharatmoljagdish/using-openai-whisper-python-library-for-speech-to-text-dda4f558fccc>

Ovládat chatovací nástroj lze pomocí přehledného grafického rozhraní, nebo opět pomocí jednoduchého Python skriptu.

```
1. from gpt4all import GPT4All
2.
3. # Načtení souboru s přepisem rozhovoru k analýze
4. with open('/Prepis.txt', 'r', encoding='utf-8') as file:
5.     texts = file.read()
6.
7. # Načtení požadovaného jazykového modelu
8. model = GPT4All("wizardlm-13b-v1.2.Q4_0.gguf")
9. # Spuštění promptu
10. output = model.generate("Udělej shrnutí následujícího rozhovoru: " + texts)
11. # Vypsání výsledku
12. print(output)
```

5 Závěr

V této diplomové práci jsem se zabýval použitím moderních informačních technologií a nástrojů umělé inteligence v oboru orální historie. Práce ukázala, že využití AI může značně přispět k efektivitě a kvalitě práce orálních historiků tím, že zjednodušuje sběr, analýzu, a archivaci orálně-historických dat. Dále práce poskytla ucelený pohled na možnosti, které digitální éra historikům nabízí. Moderní technologie mohou usnadňovat práci historiků ve všech jednotlivých fázích orálně-historického výzkumu: během přípravy projektu, přípravy před rozhovorem, během nahrávání, při zpracování a analýze a také ve fázi archivace.

Tato práce také poukázala na význam bezpečnosti a ochrany dat v digitálním prostředí a zdůraznila význam digitalizace dat, mimo jiné v rámci iniciativy Open Science, která usiluje o otevřený a přístupný vědecký výzkum. Digitalizace orální historie v souladu s principy Open Science rozšiřuje možnosti sdílení, uchování a analýzy historických dat, což má pozitivní dopad na akademickou komunitu i širší veřejnost.

Vzhledem k rychlému vývoji v oblasti umělé inteligence a informačních technologií je nezbytné, aby orální historici a výzkumníci v humanitních oborech tyto trendy sledovali a adaptovali nové nástroje do své praxe, aby mohli plně využít jejich potenciál pro rozšíření poznání v oboru orální historie. Lze očekávat, že AI technologie se budou nadále zlepšovat a poskytovat lepší, a především spolehlivější možnosti pro podporu historického výzkumu.

Integrace AI nástrojů a IT služeb v orální historii není pouze technologickou inovací, ale může být také významným přínosem pro historické vědy, otevírá nové cesty pro výzkum a poskytuje nové nástroje pro lepší porozumění spleťtité mozaice lidských příběhů.

Použité prameny a literatura

AUREKA. *AI Cataloging Assistant*. Aureka - The AI Research Assistant [online]. 2023 [cit. 2024-01-13]. Dostupné z: <https://www.aureka.ai/solution-for-archives-libraries-museums>

BIEVER, Celeste. *ChatGPT broke the Turing test — the race is on for new ways to assess AI*. Nature [online]. Springer Science and Business Media, 2023, 619, 686-689 [cit. 2024-03-09]. Dostupné z: doi:<https://doi.org/10.1038/d41586-023-02361-7>

BOYD, Doug, 2013. *OHMS: Enhancing Access to Oral History for Free*. The Oral History Review. [online]. 40(1), 95-106. DOI: 10.1093/ohr/oht031. Dostupné z: <https://www.tandfonline.com/doi/full/10.1093/ohr/oht031>

BOYD, Douglas. *Announcement: Integrating OHMS and Aviary*. Digital Omnium [online]. 2023 [cit. 2024-01-06]. Dostupné z: <https://digitalomnium.com/announcement-integrating-ohms-and-aviary/>

CARPENTER, Layne. LAUPUS HEALTH SCIENCES LIBRARY. *Oral History Digitization Project*. East Carolina University [online]. 2019 [cit. 2024-01-19]. Dostupné z: <https://hsl.ecu.edu/2019/08/15/oral-history-digitization-project/>

COPELAND, B. J. *Artificial intelligence*. *Encyclopedia Britannica* [online]. 2024 [cit. 2024-02-04]. Dostupné z: <https://www.britannica.com/technology/artificial-intelligence>.

EKMAN, Paul. *Basic emotions*. In: DALGLEISH, Tim a POWER, Mick J. (eds.). *Handbook of Cognition and Emotion*. John Wiley & Sons, 1999, s. 45–60.

FRISCH, Michael. *Oral history and the digital revolution: Toward a post-documentary sensibility*. In: PERKS, Robert a THOMPSON, Alistair (eds.). *The Oral History Reader*. 2nd ed. London, UK: Routledge, 2006, s. 102–114.

GRAF, M., SCHMIDT, C. a KOEHLER, J., 2018. *Improving Robust Speech Recognition for German Oral History Interviews Using Multi-Condition Training*. In: *Speech Communication; 13th ITG-Symposium*. Oldenburg, Germany. s. 1–5. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/8578034>

HLAVÁČEK, Jiří. *Covido, ergo zoom! Metodologické, technické a etické aspekty distančně vedených rozhovorů během pandemie Covid-19*. *Memo*. 2021, 11(1), 57–72.

HLAVÁČEK, Jiří. *Digitální historie: teoretická a metodologická východiska pro výzkum soudobých dějin 2.0*. 2024. Soudobé dějiny. 2024, 31(2) – aktuálně v tisku (rukopis studie poskytnutý autorem).

HLAVÁČEK, Jiří. *Metodické postupy archivace orálně-historických pramenů v digitálním věku*. Archivní časopis. 2019, 69(4), 389–401. ISSN 0004-0393.

HUGHES, Alex. BBC Science Focus. *ChatGPT: Everything you need to know about OpenAI's GPT-4 upgrade* [online]. 2023 [cit. 2024-01-06]. Dostupné z: <https://www.sciencefocus.com/future-technology/gpt-3>

JAILLANT, Lise a Arran REES. *Applying AI to digital archives: trust, collaboration and shared professional ethics*. *Digital Scholarship in the Humanities* [online]. 2023, 2023-06-01, 38(2), 571–585 [cit. 2024-03-17]. ISSN 2055-7671. Dostupné z: doi:10.1093/llc/fqac073

JENTSCH, Patrick a Stephan PORADA. *From Text to Data: Digitization, Text Analysis and Corpus Linguistics*. In: *Digital Methods in the Humanities: Challenges, Ideas, Perspectives* [online]. 2021, s. 89–128 [cit. 2024-03-10].

MORTON, Sunny Jane. *5 Voice-Recorder Apps for Oral History Interviews*. Family Tree Magazine [online]. 2021 [cit. 2024-01-13]. Dostupné z: <https://familytreemagazine.com/resources/apps/voice-recorder-apps/>

ORAL HISTORY ASSOCIATION. *Archiving Oral History*. Oral History Association [online]. 2019 [cit. 2024-01-19]. Dostupné z: <https://oralhistory.org/archives-principles-and-best-practices-overview/>

ORAL HISTORY ASSOCIATION. *Oral History Association response to developments in Boston College case*. Oral History Association [online]. 2014 [cit. 2024-03-09]. Dostupné z: <https://oralhistory.org/2014/05/05/oral-history-association-response-to-developments-in-boston-college-case/>

PESSANHA, Francisca a SALAH, Almila Akdag. 2021. *A Computational Look at Oral History Archives*. *Journal of Computer and Cultural Heritage*. [online]. 15(1), Článek 6 (únor 2022), 16 stran. DOI: 10.1145/3477605. Dostupné z: <https://doi.org/10.1145/3477605>

SON, Y. a kol., 2020. *World Trade Center responders in their own words: predicting PTSD symptom trajectories with AI-based language analyses of interviews*. *Psychological Medicine*. [online]. 53, 918–926. Dostupné z: <https://doi.org/10.1017/S0033291721002294>

SPANG-WILLIS, D., 2020. *AI: An opportunity to reimagine an old-world anew*. Oral History Master of Arts. [online]. Dostupné z: <http://oralhistory.columbia.edu/blog-posts/an-opportunity-to-reimagine-an-old-world-anew>.

UNESCO. *An introduction to the UNESCO Recommendation on Open Science*. Open Science [online]. 2022 [cit. 2024-02-04]. Dostupné z: <https://doi.org/10.54677/XOIR1696>

UNIVERZITA KARLOVA. *Perzistentní identifikátory*. Centrum pro podporu open science [online]. aktualizováno 17. 10. 2022 [cit. 2024-03-03]. Dostupné z: <https://openscience.cuni.cz/OSCI-193.html>

UNIVERZITA KARLOVA. *FAIR data*. Centrum pro podporu open science [online]. [cit. 2024-03-02]. Dostupné z: <https://openscience.cuni.cz/OSCI-64.html>

ÚSTAV PRO SOUDOBÉ DĚJINY AV ČR. *O projektu. 100 studentských revolucí* [online]. 2019 [cit. 2023-11-30]. Dostupné z: <http://www.studenti89.usd.cas.cz/studenti/>

VANĚK, Miroslav a MÜCKE, Pavel. *Třetí strana trojúhelníku: teorie a praxe orální historie. 2.*, přepracované a doplněné vydání. Orální historie a soudobé dějiny. Praha: Univerzita Karlova v Praze, nakladatelství Karolinum, 2015. ISBN 978-80-246-2931-5.

WIKIMEDIA FOUNDATION. *Announcing Wikipedia's most popular articles of 2023*. Wikimedia Foundation [online]. 2023 [cit. 2024-01-06]. Dostupné z: <https://wikimediafoundation.org/news/2023/12/05/announcing-wikipedias-most-popular-articles-of-2023/>

Seznam obrázků

Obrázek 1: Predikce vývoje velikosti trhu umělé inteligence od roku 2022 do roku 2032. Dostupné z: [https://www.precedenceresearch.com/artificial-intelligence-market].....	1
Obrázek 2: Deepfake fotografie papeže Františka v bílé nafukovací bundě měla na sociálních sítích více než 20 milionů shlédnutí. [https://www.nytimes.com/2023/04/08/technology/ai-photos-pope-francis.html].....	12
Obrázek 3: Ukázka vytváření deepfake videa záměnou obličejů narátorů pomocí programu FaceFusion.....	13
Obrázek 4: „A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. She walks confidently and casually. The street is damp and reflective, creating a mirror effect of the colorful lights. Many pedestrians walk about.“, prompt, Sora, OpenAI, https://openai.com/sora	15
Obrázek 5: „Beaten students, Velvet revolution 1989 National street Czechoslovakia Prague“, prompt, Leonardo.Ai, Leonardo interactive, 6.1.2024, https://leonardo.ai/	15
Obrázek 6: „Sametová revoluce v Československu 1989“, prompt, DALL-E, verze 3, OpenAI, 6.1.2024, https://labs.openai.com/	16
Obrázek 7: „Velvet revolution 1989 Václav Havel Czechoslovakia Prague“, prompt, Leonardo.Ai, Leonardo interactive, 6.1.2024, https://leonardo.ai/	16
Obrázek 8: „Náhled mobilní aplikace s virtuální rekonstrukcí historických postav, Tomáš Garrigue Masaryk“, prompt, DALL-E, verze 3, OpenAI, 6.1.2024, https://labs.openai.com/	17
Obrázek 9: Původní zvukové spektrum rozhovoru.....	30
Obrázek 10: Upravené zvukové spektrum rozhovoru.....	30
Obrázek 11: Výstup analýzy sentimentu z fotografie pomocí služby DeepFace.....	47
Obrázek 12: Vybrané záběry z rozhovoru s Klárou Jirovcovou Pospíšilovou pro účely analýzy emocí.....	48
Obrázek 13: Uživatelské rozhraní archivního systému Aureka.AI [https://www.aureka.ai/solution-for-archives-libraries-museums].....	53
Obrázek 14: Zastoupení slov v textu a předpokládaná frekvence podle Zipfova zákona.....	58
Obrázek 15: Počet vět podle indexu jejich sentimentu.....	61

Seznam tabulek

Tabulka 1: Přehled mobilních aplikací pro nahrávání rozhovorů.....	28
Tabulka 2: Přesnost přepisů pomocí různých AI nástrojů.....	38
Tabulka 3: Výsledky analýzy sentimentu pomocí služby DeepFace.....	48
Tabulka 4: Statistika automatických přepisů rozhovorů.....	56
Tabulka 5: Četnost nejčastěji se vyskytujících slov v rozhovorech.....	58
Tabulka 6: Počet vět podle indexu jejich sentimentu.....	60