

Vyhledávání informací v projektu Malach

Příprava testovací kolekce

Pavel Pecina

pecina@ufal.mff.cuni.cz

Ústav formální a aplikované lingvistiky
Matematicko-fyzikální fakulta Univerzity Karlovy



Mixer, 12. duben 2006

Osnova

Historie projektu

Cíle a řešení projektu

Rozpoznávání řeči

Vyhledávání informací

CLEF evaluace

Testovací kolekce

Odkazy

Stručná historie projektu Malach

Multilingual Access to Large Spoken Archives

- 1993 Stephen Spielberg uvádí premiéru filmu Schindlerův seznam *během natáčení se setkává s lidmi, kteří přežili holokaust, zpracovává jejich příběhy*
- 1994 Založena nadace Survivors of the Shoah Visual History Foundation *s cílem zaznamenat a zpřístupnit svědectví lidí, kteří přežili holokaust*
- 1999 VHF vytvořila největší audiovizuální archív na světě *obsahující 52 000 výpovědí svědků holokaustu z 57 zemí*
- 2000 10 % nahrávek manuálně katalogizováno ze cenu 8 mil. dolarů *zpracování jedné výpovědi trvá ø 35 hodin (indexace, sumarizace, kontrola)*
- 2001 NSF požádána o udělení grantu na automatické zpracování celého archívu *hlavní řešitelé: University of Maryland, Johns Hopkins University, IBM*
- 2001 NSF uděluje grant, projekt Malach startuje *rozpočet 7,5 mil dolarů na 5 let, k řešitelům se přidávají UK a ZČU*
- 2006 předpokládané ukončení projektu

Archív

- ▶ spravován Shoah Foundation (nyní součást University of South California)
- ▶ vytvářen během let 1994–1999 za pomoci 2 300 tazatelů a 1 000 kameramanů
- ▶ obsahuje výpovědi 52 000 svědků holokaustu z 56 zemí ve 32 jazycích
- ▶ celkem se jedná o 117 000 hodin VHS záznamů, zdigitalizovaných 400 TB
- ▶ průměrná délka výpovědi je 2h 15min a náklady na její vytvoření 2 000 USD
- ▶ výpovědi ručně indexovány pomocí tezauru o 30 000 klíčových slovech
- ▶ část výpovědí katalogizována podrobněji
- ▶ on-line přístup k metadatům všech výpovědí a vzorku plným záznamům
- ▶ na vyžádání přístup ke všem nahrávkám pro vědecké a studijní účely

- ▶ 570 rozhovorů proběhlo v ČR (většina v češtině) za pomoci 38 tazatelů
- ▶ 4 500 svědectví podali lidé narození v ČR

Cíle a řešení projektu Malach

1. automatické rozpoznávání řeči
doslovný přepis všech rozhovorů
2. strojový překlad doménově specifického vícejazyčného tezauru
tazaurus vytvořen přímo pro doménu výpovědí během ruční katalogizace
3. automatické přiřazování metadat
přiřazování klíčových slov k úsekům výpovědí
4. prostředí pro vícejazyčné vyhledávání informací a prohledávání archívu
založené na (nedokonalých) výsledcích předchozích úloh



IBM T.J. Watson Center, New York
- rozpoznávání mluvené řeči (angličtina)



Center for Speech and Language Processing, JHU, Baltimore
- rozpoznávání mluvené řeči v ostatních jazycích
- čeština a další slovanské jazyky ve spolupráci s UK a ZČU



University of Maryland, College Park
- vyhledávání informací, prohledávání archívu, vytvoření testovací kolekce
- testovací kolekce pro češtinu ve spolupráci s UK

Rozpoznávání řeči

- ▶ doslovný přepis spontánní řeči (nezávislý na řečníkovi)
- ▶ záznamy z technického hlediska poměrně kvalitní
- ▶ řešení úlohy stěžuje jazyková kvalita (emoce, stáří, zdravotní stav, jazyk)
- ▶ specifickým problémem v češtině jsou hovorové výrazy a výslovnost

oběd	[o b j e d]	Osvětim	[o s v j e t i m]
	[o b j e t]		[v o s v j e t i m]
	[v o b j e d]		[o s v j e n ě i m]
	[v o b j e t]		[o z v j e t i m]

- ▶ výsledky měřeny na vzorku ručně přepsaných záznamů
– jako poměr chybně rozpoznaných slov

jazyk	WER (%)
angličtina	36.81
čeština	38.57
slovenština	34.49
ruština	45.75

Rozpoznávání řeči - ukázka

jméno: Hugo Pavel

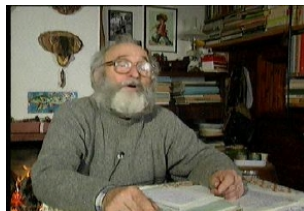
narození: 26.12. 1924

země původu: Československo

vyznání (pre): judaismus

vyznání (post): N/A

klíčová slova: hiding/death marches
underground/resistance



Rozpoznávání řeči - ukázka

jméno: Hugo Pavel

narození: 26.12. 1924

země původu: Československo

vyznání (pre): judaismus

vyznání (post): N/A

klíčová slova: hiding/death marches
underground/resistance



pane pavle začal jste historku o srncích a tatínkovi bez hvězdy jak to pokračovalo bylo pokračovalo to tím způsobem že tatínek si sundal hvězdu pan doktor jeřáb mu napsali skupinku na kladně to bylo báječný doktor ten a fandila náš tatínek se vydal na cestu na křivoklátsko aby upekla že sem se ... pochopitelně že strejda prošek s tím nechtěl nic mít za to byly krutý tresty za to se tenkrát popravovalo takže strejda prošek nepytlačil a bál se tady všude v lesích byli němci střílelo se ... a náš táta se vydal na tuhle cestu a ubytoval se mnou slůvko toho v roce sem opravdu podařilo u pytláčit – za pomoci legendární a volal na to byl pes vlčák s kterým dříve prošek nepytlačil a ten prostě každého sem se nepytlačil ...

Rozpoznávání řeči - ukázka

jméno: Hugo Pavel

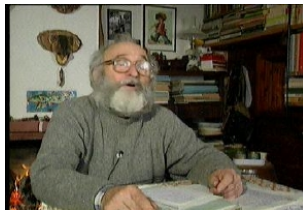
narození: 26.12. 1924

země původu: Československo

vyznání (pre): judaismus

vyznání (post): N/A

klíčová slova: hiding/death marches
underground/resistance



pane pavle začal jste historku o srcích a tatínkovi bez hvězdy jak to pokračovalo bylo pokračovalo to tím způsobem že tatínek si sundal hvězdu pan doktor jeřáb mu napsali skupinku na kladně to bylo báječný doktor ten a fandila náš tatínek se vydal na cestu na křivoklátsko aby upekla že sem se ... pochopitelně že strejda prošek s tím nechtěl nic mít za to byly krutý tresty za to se tenkrát popravovalo takže strejda prošek nepytláčil a bál se tady všude v lesích byli němci střílelo se ... a náš táta se vydal na tuhle cestu a ubytoval se mnou slůvko toho v roce sem opravdu podařilo u pytláčit – za pomoci legendární a volal na to byl pes vlčák s kterým dříve prošek nepytláčil a ten prostě každého sem se nepytláčil ...

Rozpoznávání řeči - ukázka

jméno: Hugo Pavel

narození: 26.12. 1924

země původu: Československo

vyznání (pre): judaismus

vyznání (post): N/A

klíčová slova: hiding/death marches
underground/resistance



pane pavle začal jste historku o srncích a tatínkovi bez hvězdy jak to pokračovalo bylo pokračovalo to tím způsobem že tatínek si sundal hvězdu pan doktor jeřáb mu napsali skupinku na kladně to bylo báječný doktor ten a fandila náš tatínek se vydal na cestu na křivoklátsko aby upekla že sem se ... pochopitelně že strejda prošek s tím nechtěl nic mít za to byly krutý tresty za to se tenkrát popravovalo takže strejda prošek nepytláčil a bál se tady všude v lesích byli němci střílelo se ... a náš táta se vydal na tuhle cestu a ubytoval se mnou slůvko toho v roce sem opravdu podařilo u pytláčit – za pomoci legendární a volal na to byl pes vlčák s kterým dříve prošek nepytláčil a ten prostě každého sem se nepytláčil ...

Rozpoznávání řeči - ukázka

jméno: Hugo Pavel

narození: 26.12. 1924

země původu: Československo

vyznání (pre): judaismus

vyznání (post): N/A

klíčová slova: hiding/death marches
underground/resistance



pane pavle začal jste historku o srcích a tatínkovi bez hvězdy jak to pokračovalo bylo pokračovalo to tím způsobem že tatínek si sundal hvězdu pan doktor jeřáb mu napsali skupinku na kladně to bylo báječný doktor ten a fandila náš tatínek se vydal na cestu na křivoklátsko aby upekla že sem se ... pochopitelně že strejda prošek s tím nechtěl nic mít za to byly krutý tresty za to se tenkrát popravovalo takže strejda prošek nepytlačil a bál se tady všude v lesích byli němci střílelo se ... a náš táta se vydal na tuhle cestu a ubytoval se mnou slůvko toho v roce sem opravdu podařilo u pytláčit – za pomoci legendární a volal na to byl pes vlčák s kterým dříve prošek nepytlačil a ten prostě každého sem se nepytlačil ...

Rozpoznávání řeči - ukázka

jméno: Hugo Pavel

narození: 26.12. 1924

země původu: Československo

vyznání (pre): judaismus

vyznání (post): N/A

klíčová slova: hiding/death marches
underground/resistance



pane pavle začal jste historku o srcích a tatínkovi bez hvězdy jak to pokračovalo bylo pokračovalo to tím způsobem že tatínek si sundal hvězdu pan doktor jeřáb mu napsali skupinku na kladně to bylo báječný doktor ten a fandila náš tatínek se vydal na cestu na křivoklátsko aby upekla že sem se ... pochopitelně že strejda prošek s tím nechtěl nic mít za to byly krutý tresty za to se tenkrát popravovalo takže strejda prošek nepytlačil a bál se tady všude v lesích byli němci střílelo se ... a náš táta se vydal na tuhle cestu a ubytoval se mnou slůvko toho v roce sem opravdu podařilo u pytláčit – za pomoci legendární a volal na to byl pes vlčák s kterým dříve prošek nepytlačil a ten prostě každého sem se nepytlačil ...

Rozpoznávání řeči - ukázka

jméno: Hugo Pavel

narození: 26.12. 1924

země původu: Československo

vyznání (pre): judaismus

vyznání (post): N/A

klíčová slova: hiding/death marches
underground/resistance



pane pavle začal jste historku o srcích a tatínkovi bez hvězdy jak to pokračovalo bylo pokračovalo to tím způsobem že tatínek si sundal hvězdu pan doktor jeřáb mu napsali skupinku na kladně to bylo báječný doktor ten a fandila náš tatínek se vydal na cestu na křivoklátsko aby upekla že sem se ... pochopitelně že strejda prošek s tím nechtěl nic mít za to byly krutý tresty za to se tenkrát popravovalo takže strejda prošek nepytlačil a bál se tady všude v lesích byli němci střílelo se ... a náš táta se vydal na tuhle cestu a ubytoval se mnou slůvko toho v roce sem opravdu podařilo u pytláčit – za pomoci legendární a volal na to byl pes vlčák s kterým dříve prošek nepytlačil a ten prostě každého sem se nepytlačil ...

Rozpoznávání řeči - ukázka

jméno: Hugo Pavel

narození: 26.12. 1924

země původu: Československo

vyznání (pre): judaismus

vyznání (post): N/A

klíčová slova: hiding/death marches
underground/resistance



pane pavle začal jste historku o srncích a tatínkovi bez hvězdy jak to pokračovalo bylo pokračovalo to tím způsobem že tatínek si sundal hvězdu pan doktor jeřáb mu napsali skupinku na kladně to bylo báječný doktor ten a fandila náš tatínek se vydal na cestu na křivoklátsko aby upekla že sem se ... pochopitelně že strejda prošek s tím nechtěl nic mít za to byly krutý tresty za to se tenkrát popravovalo takže strejda prošek nepytláčil a bál se tady všude v lesích byli němci střílelo se ... a náš táta se vydal na tuhle cestu a ubytoval se mnou slůvko toho v roce sem opravdu podařilo u pytláčit – za pomoci legendární a volal na to byl pes vlčák s kterým dříve prošek nepytláčil a ten prostě každého sem se nepytláčil ...

Rozpoznávání řeči - ukázka

jméno: Hugo Pavel

narození: 26.12. 1924

země původu: Československo

vyznání (pre): judaismus

vyznání (post): N/A

klíčová slova: hiding/death marches
underground/resistance



pane pavle začal jste historku o srcích a tatínkovi bez hvězdy jak to pokračovalo bylo pokračovalo to tím způsobem že tatínek si sundal hvězdu pan doktor jeřáb mu napsali skupinku na kladně to bylo báječný doktor ten a fandila náš tatínek se vydal na cestu na křivoklátsko aby upekla že sem se ... pochopitelně že strejda prošek s tím nechtěl nic mít za to byly krutý tresty za to se tenkrát popravovalo takže strejda prošek nepytláčil a bál se tady všude v lesích byli němci střílelo se ... a náš táta se vydal na tuhle cestu a ubytoval se mnou slůvko toho v roce sem opravdu podařilo u pytláčit – za pomoci legendární a volal na to byl pes vlčák s kterým dříve prošek nepytláčil a ten prostě každého sem se nepytláčil ...

Rozpoznávání řeči - ukázka

jméno: Hugo Pavel

narození: 26.12. 1924

země původu: Československo

vyznání (pre): judaismus

vyznání (post): N/A

klíčová slova: hiding/death marches
underground/resistance



pane pavle začal jste historku o srcích a tatínkovi bez hvězdy jak to pokračovalo bylo pokračovalo to tím způsobem že tatínek si sundal hvězdu pan doktor jeřáb mu napsali skupinku na kladně to bylo báječný doktor ten a fandila náš tatínek se vydal na cestu na křivoklátsko aby upekla že sem se ... pochopitelně že strejda prošek s tím nechtěl nic mít za to byly krutý tresty za to se tenkrát popravovalo takže strejda prošek nepytláčil a bál se tady všude v lesích byli němci střílelo se ... a náš táta se vydal na tuhle cestu a ubytoval se mnou slůvko toho v roce sem opravdu podařilo u pytláčit – za pomoci legendární a volal na to byl pes vlčák s kterým dříve prošek nepytláčil a ten prostě každého sem se nepytláčil ...

Rozpoznávání řeči - ukázka

jméno: Hugo Pavel

narození: 26.12. 1924

země původu: Československo

vyznání (pre): judaismus

vyznání (post): N/A

klíčová slova: hiding/death marches
underground/resistance



pane pavle začal jste historku o sncích a tatínkovi bez hvězdy jak to pokračovalo bylo pokračovalo to tím způsobem že tatínek si sundal hvězdu pan doktor jeřáb mu napsali skupinku na kladně to bylo báječný doktor ten a fandila náš tatínek se vydal na cestu na křivoklátsko aby upekla že sem se ... pochopitelně že strejda prošek s tím nechtěl nic mít za to byly krutý tresty za to se tenkrát popravovalo takže strejda prošek nepytláčil a bál se tady všude v lesích byli němci střílelo se ... a náš táta se vydal na tuhle cestu a ubytoval se mnou slůvko toho v roce sem opravdu podařilo u pytláčit – za pomoci legendární a volal na to byl pes vlčák s kterým dříve prošek nepytláčil a ten prostě každého sem se nepytláčil ...

Vyhledávání informací (Information Retrieval)

- ▶ hledání tématicky relevantních částí výpovědí *napříč jazyky* (cross-lingual)
- ▶ narozdíl od klasického IR nehledáme dokumenty
 - neznáme jejich hranice, celé výpovědi jsou příliš dlouhé
- ▶ varianty hledaných úseků:
 - ▶ předem daná ruční segmentace - *hledají se relevantní segmenty*
 - ▶ bez předchozí segmentace - *hledají se začátky a konce relevantních úseků*
 - ▶ bez předchozí segmentace - *hledají se pouze začátky relevantních úseků*
- ▶ pro každý rozhovor jsou k dispozici tyto data:
 - ▶ ručně přiřazená metadata
 - ▶ automaticky přiřazená metadata
 - ▶ automatické přepisy
- ▶ testování probíhá vůči předem dané sadě témat a ručně ohodnocené kolekci
 - pro **každé** téma známe relevantní segmenty ve **všech** rozhovorech
- ▶ testování probíhá v rámci CLEF

- ▶ Aktivita **DELOS Network of Excellence for Digital Libraries** jakožto součást 6. rámcového programu Evropské komise.
- ▶ Vytváří infrastrukturu pro testování, ladění a evaluaci systémů pro vyhledávání informací v evropských jazycích (v mono- i cross-linguálním prostředí).
- ▶ Vytváří standardní testovací sady dat opětovně použitelné pro vývoj a srovnávání těchto systémů (benchmarking).
- ▶ Každý rok se zaměřuje na různé úlohy. Pro rok 2006 např.:
 - ▶ Mono- and Cross-Language IR on Structured Scientific Data
 - ▶ Interactive Cross-Language IR
 - ▶ Multiple Language Question Answering
 - ▶ Cross-Language Retrieval in Image Collections
 - ▶ Cross-Language Geographical Retrieval
 - ▶ **Cross-Language Speech Retrieval**

CLEF: Cross-Language Speech Retrieval track

- ▶ V roce 2005 vytvořena testovací kolekce obsahující část výpovědí zpracovaných v projektu Malach
- ▶ témata v 5 jazycích, výpovědi v angličtině s automatickým přepisem, manuálně a automaticky přiřazenými klíčovými slovy (koncepty, data, geografická místa, jména lidí) a s ruční sumarizací.
- ▶ V roce 2006 kolekce obohacena o výpovědi v češtině zpracovávané na UK
- ▶ Předpokládaný rozsah české kolekce je 500 výpovědí a 25 témat
- ▶ Program:

Nyní Přihlášky do „soutěže“

15.4. Uvolnění ukázkových výpovědí a témata pro trénování

1.5. Uvolnění všech výpovědí a témat pro evaluaci

6.6. Odeslání výsledků všech účastníků

1.8. Zveřejnění výsledků evaluace

15.8. Odeslání příspěvků na workshop

20.9. 2-denní CLEF workshop v Alicante, Španělsko

Testovací kolekce: témata

- ▶ vznikla z reálných požadavků odborníků a zájemců o hledání v archívu
- ▶ z angličtiny přeložena do francouzštiny, němčiny, španělštiny a češtiny
- ▶ 75 témat použito pro anglickou kolekci
- ▶ očekáváme, že česká kolekce bude obsahovat 25 témat

title 1173: Dětské umění v Terezíně

desc Hledáme popis uměleckých aktivit dětí v Terezíně, jako např. hudby, divadla, malování, poezie a jiných psaných děl.

narr *Relevantní materiál by měl obsahovat diskuse o těchto aktivitách a to, jak ovlivnily přečkání holokaustu a následný život dětí. Zejména jsou žádané příběhy, ve kterých účastník rozhovoru uvádí příklady takových aktivit.*

Testovací kolekce: témata

- ▶ vznikla z reálných požadavků odborníků a zájemců o hledání v archívu
- ▶ z angličtiny přeložena do francouzštiny, němčiny, španělštiny a češtiny
- ▶ 75 témat použito pro anglickou kolekci
- ▶ očekáváme, že česká kolekce bude obsahovat 25 témat

title 1431: **Denní život v Terezíně**

desc Popište denní život v táboře Terezín.

narr *Anekdoty, příběhy nebo detaily o denním životě v táboře. Relevantní jsou příběhy, které se zmiňují o modlitbách, svátcích, volném čase vězňů a strukturách vnitřní vlády v táborech. Příběhy, které popisují neobvyklé události v životě vězňů, relevantní nejsou.*

Testovací kolekce: témata

- ▶ vznikla z reálných požadavků odborníků a zájemců o hledání v archívu
- ▶ z angličtiny přeložena do francouzštiny, němčiny, španělštiny a češtiny
- ▶ 75 témat použito pro anglickou kolekci
- ▶ očekáváme, že česká kolekce bude obsahovat 25 témat

title 1431: **Denní život v Terezíně**

desc Popište denní život v táboře Terezín.

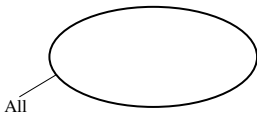
narr *Anekdoty, příběhy nebo detaily o denním životě v táboře. Relevantní jsou příběhy, které se zmiňují o modlitbách, svátcích, volném čase vězňů a strukturách vnitřní vlády v táborech. Příběhy, které popisují neobvyklé události v životě vězňů, relevantní nejsou.*

Testovací kolekce: určování relevance

- ▶ V ideálním případě je potřeb pro každé téma najít všechny relevantní úseky
- ▶ Není možné procházet a (poslouchat) všechny výpovědi
- ▶ Řešením je tzv. *search guided relevance assessment*
- ▶ Před ručním procházením je množina potenciálně relevantních výpovědí omezena pomocí fulltextového vyhledávání (první kolo)
- ▶ Získáme výsledky z reálného nasazení vyhledávacích systémů a projdeme jejich výsledky (druhé kolo)

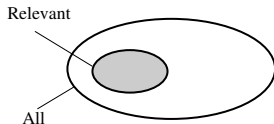
Testovací kolekce: určování relevance

- ▶ V ideálním případě je potřeb pro každé téma najít všechny relevantní úseky
- ▶ Není možné procházet a (poslouchat) všechny výpovědi
- ▶ Řešením je tzv. *search guided relevance assessment*
- ▶ Před ručním procházením je množina potenciálně relevantních výpovědí omezena pomocí fulltextového vyhledávání (první kolo)
- ▶ Získáme výsledky z reálného nasazení vyhledávacích systémů a projdeme jejich výsledky (druhé kolo)



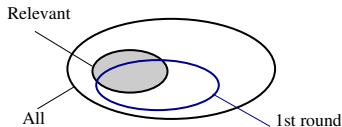
Testovací kolekce: určování relevance

- ▶ V ideálním případě je potřeb pro každé téma najít všechny relevantní úseky
- ▶ Není možné procházet a (poslouchat) všechny výpovědi
- ▶ Řešením je tzv. *search guided relevance assessment*
- ▶ Před ručním procházením je množina potenciálně relevantních výpovědí omezena pomocí fulltextového vyhledávání (první kolo)
- ▶ Získáme výsledky z reálného nasazení vyhledávacích systémů a projdeme jejich výsledky (druhé kolo)



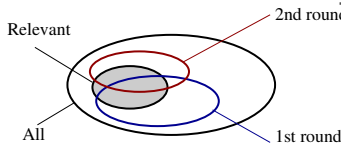
Testovací kolekce: určování relevance

- ▶ V ideálním případě je potřeb pro každé téma najít všechny relevantní úseky
- ▶ Není možné procházet a (poslouchat) všechny výpovědi
- ▶ Řešením je tzv. *search guided relevance assessment*
- ▶ Před ručním procházením je množina potenciálně relevantních výpovědí omezena pomocí fulltextového vyhledávání (první kolo)
- ▶ Získáme výsledky z reálného nasazení vyhledávacích systémů a projdeme jejich výsledky (druhé kolo)



Testovací kolekce: určování relevance

- ▶ V ideálním případě je potřeb pro každé téma najít všechny relevantní úseky
- ▶ Není možné procházet a (poslouchat) všechny výpovědi
- ▶ Řešením je tzv. *search guided relevance assessment*
- ▶ Před ručním procházením je množina potenciálně relevantních výpovědí omezena pomocí fulltextového vyhledávání (první kolo)
- ▶ Získáme výsledky z reálného nasazení vyhledávacích systémů a projdeme jejich výsledky (druhé kolo)



File Help

Search Assessment

PIQ

Keyword

Transcript Třezin

Exact match





196 interviews found. Displaying results 1-10

Efim Markelis-Dominin (ID: 31115, 1 segment, 0 point(s) already marked for this topic)

Locations: ...

Concepts: ...

Kunhuta Buresova (ID: 14932, 3 segments, 0 point(s) already marked for this topic)

Locations: ... Theresienstadt (Czechoslovakia : Ghetto), Czechoslovakia 1943, Czechoslovakia 1944 ...

Concepts: ... forced labor in the ghettos, means of adaptation and survival in the ghettos, forced labor: agriculture, stealing in the ghettos ...

Aizbeta Bernatova (ID: 30796, 2 segments, 0 point(s) already marked for this topic)

Locations: ... Czechoslovakia 1939 (Mar 15) - 1945 (May 9), Theresienstadt (Czechoslovakia : Ghetto) ...

Concepts: ... sharing food and drink in the ghettos, interaction with family members in the refugee camps ...

Hugo Pavel (ID: 13135, 6 segments, 0 point(s) already marked for this topic)

Locations: ... Theresienstadt (Czechoslovakia : Ghetto), Czechoslovakia 1943, Germany 1944, Germany 1945 (Jan 1 - May 7), ...

Concepts: ... housing conditions in the ghettos, food in the camps ...

Hana Trvrská (ID: 23005, 3 segments, 0 point(s) already marked for this topic)

Locations: ... Czechoslovakia 1941, Czechoslovakia 1942, Protivin (Czechoslovakia) ...

Concepts: ... awareness of deportations and/or transfers ...

Felice Kvapilova (ID: 23179, 2 segments, 0 point(s) already marked for this topic) - Hiding

Locations: ...

Concepts: ...

Eva Liskova (ID: 28592, 5 segments, 0 point(s) already marked for this topic) - Death Marches

Locations: ...

Concepts: ... cultural and social activities in the ghettos, clandestine activities in the ghettos, Appell in the ghettos, deportations, means of transport, ...

Margit Herrmannova (ID: 31114, 1 segment, 0 point(s) already marked for this topic)

Locations: ...

Concepts: ...

Anna Brynberg (ID: 13682, 1 segment, 0 point(s) already marked for this topic)

Locations: ...

Concepts: ...

Jiri Bures (ID: 20065, 3 segments, 0 point(s) already marked for this topic)

Locations: ... Theresienstadt (Czechoslovakia : Ghetto), Czechoslovakia 1943

Keyword Transcript

Contains



Germany 1944
 Germany 1945 (Jan 1 - May 7)
 Zossen-Wulkow bei Trebnitz (Germany :
 Concentration Camp)
 food in the camps

Germany 1944
 Germany 1945 (Jan 1 - May 7)
 Zossen-Wulkow bei Trebnitz (Germany :
 Concentration Camp)
 brutal treatment in the camps
 beatings
 Stuschka, Franz

Show: Locations Concepts People

nde my jsme bydlci a my jsme byli dobytovan ty nas mndajce po práci
 a tam už to bylo volně v zátoce výstaviště nemohlo být prostě první prostě prostě
 s tou svou otec byl tam pokusím o útek zaplatili ty lidi život protože chytli
 také popravili to je docela známá věc že byli pověšeni ty kosti dělal šli do
 koncentráku takže s- pokud ale kde nebyl problém ani utekl z terezína když budeme od
 | pavlovi když to hřídali četníci ty mladiství chodili na práci do zemědělství říkalo se
 tomu landwirtschaft to bylo za terezín žilo kolem terezína byli zelinářském poli pro terezín pro
 ss komandaturu takže nebyl problém otec terezína nebyl problém otec vulkova problém protože sem
 měl
 doma rodiče a ty si to odnesl v první řadě ... tady byla zodpovědnost vůči
 té rodině doufat ... ale nebyl problém ten spor že se tam mezi těma lidma
 udržet protože němci byli na stole prostě by vás udal jo zapálit ty otázky ...
 jak jste tam byli stravovali ... no tak stravovali žili jsme ovšem zažili jsme tam
 také strašný hlad ke konci o tom že jsme měli kůru ze stromu trávu se
 svou fotku tý se z kuchyně jsme vybíraly byla ta potom ke konci špatná doba
 chodili franta většinou potraviny z | terezína takový ty tvrdý potraviny a brambora takový ty
 věci to zřejmě ty ... esesáci tě zde scházeli tam takže to měl na starosti
 těch okolní vesnici od sedláků to vypadalo že jo to bylo to byl kuchařem byla
 taková že v kuchyni a tom také pak vaření to byl v terezíně takový to
 byli většinou tuřín spal to bylo celý vařil dost asi deset deka směli táborů ...
 filosof ještě pluku vyhodil nebo sem to vřdycy okolo pak sem se ty ty šlupky
 vod docentovi jednou za týden byla buchta s takovým krémem my jsme tomu říkali takové
 to bylo kafe že jo jsme vřdycy říkali až bude po válce se to buchtu
 udělal paninko to v životě dál potom korupci koupili to bylo také takový ... prostě
 ze čtyřech letech opakovala ke konci to proto že ke konci to dostala ke konci
 | to znamená ke konci když už potom byl rok čtyřicet šátek roku čtyřicet pět
 tak už to bylo zlý protože z východu se ztratili rusové a došlo k tomu
 já bych chtěl ještě ráda ještě k tomu i já sem včera četl tady podrobili
 výsledku z toho že frances počkej rakousku tma toho teda se jmenoval že on říká
 že vůbec se tam nějak provinili tím ale vo tom strašně vězně miátt ... já
 sem to už na floridu to jsme ještě jako nemluvíli ale i tam sem byl
 u něj strašně bydlel tam byl rodině se ozývaly na těch lidí víte koho tři
 z těch prostě že na tu práci nedělá dobře ale byli lidi který byli placený
 za den voskovec lakýrník to mlátil lidi otu rezka že ze zbraslavy z toho mlátili
 a oni mlátili tím potom také byli esesáci nemlátil nevrátil | akorát voda a potom
 byl mezi námi jeden vězeň pracoval v lodži který také hlásila ale ten jako fakt
 fackoval toho po válce popravý ale tenhle ten stužka ten prostě mlátil byl takový zámeck
 hrabova ten nebo když se škoda žili ve se by stromy tak mlátili těma kupy

11

2

3

4

5

6

7

8

11

Speaker Information



Hugo Pavel



Personal details

Date of birth Dec 26, 1924**Country of birth** Czechoslovakia**Other names** Not available**Other attributes/roles** Not available**Religious identity (PIQ)** none

- Hiding
- Underground/Resistance/Partisan
- Death Marches

Family details

None Available

Transcript

Contains



... kde my jsme bydleli a my jsme byli dobytými ty nás mluvil je po práci a tam už to bylo volně v zátoce výstaviště nemohlo být prostě první prostě prostě s tou svou otec byl tam pokusím o útek zaplatili ty lidi život protože chytli také popravili to je docela známá věc že byli pověšeni ty kosti dělal šli do koncentráku takže s- pokud ale kde nebyl problém ani utekl z terezína když budete od | pavlovi když to hřídali četníci ty mladiství chodili na práci do zemědělství říkalo se tomu landwirtschaft to bylo za terezín žilo kolem terezína byli zelinářském poli pro terezín pro ss komandaturu takže nebyl problém otec terezína nebyl problém otec vulkova problém protože sem měl

doma rodiče a ty si to odnesl v první řadě ... tady byla zodpovědnost vůči té rodině doufat ... ale nebyl problém ten spor že se tam mezi těma lidma udržet protože němcí byli na stole prostě by vás udal jo zapálit ty otázky ... jak jste tam byli stravovali ... no tak stravovali žili jsme ovšem zažili jsme tam také strašný hlad ke konci o tom že jsme měli kůru ze stromu trávu se svou fotku ty se z kuchyně jsme vybíraly byla ta potom ke konci špatná doba chodili franta většinou potraviny z | terezína takový ty tvrdý potraviny a brambora takový ty věci to zřejmě ty ... esesáci tě zde scházeli tam takže to měl na starosti těch okolní vesnici od sedláků to vypadalo že jo to bylo to byl kuchařem byla taková že v kuchyni a tom také pak vaření to byl v terezíně takový to byli většinou tuřín spal to bylo celý vařil dost asi deset deka směli táborů ... filosof ještě pluku vyhodil nebo sem to vždycy okolo pak sem se ty ty šlupky vod docentovi jednou za týden byla buchta s takovým krémem my jsme tomu říkali takové to bylo kafe že jo jsme vždycy říkali až bude po válce se to buchtu udělal paninko to v životě dál potom korupci koupili to bylo také takový ... prostě ze čtyřech letech opakovala ke konci to proto že ke konci to dostala ke konci | to znamená ke konci když už potom byl rok čtyřicet šátek roku čtyřicet pět tak už to bylo zlý protože z východu se ztratili rusové a došlo k tomu já bych chtěl ještě ráda ještě k tomu i já sem včera četl tady podrobili výslechu z toho že frances počkej rakousku tma toho teda se jmenoval že on říká že vůbec se tam nějak provinili tím ale vo tom strašně vězně mlátt ... já sem to už na floridu to jsme ještě jako nemluvíli ale i tam sem byl u něj strašně bydlel tam byl rodině se ozývaly na těch lidí víte koho tři z těch prostě že na tu práci nedělá dobře ale byli lidi který byli placení za den voskovec lakýrník to mlátil lidi otu rezka že ze zbraslavy z toho mlátili a oni mlátili tím potom také byli esesáci nemlátil nevrátil | akorát voda a potom byl mezi námi jeden vězeň pracoval v lodži který také hlásila ale ten jako fakt fackoval toho po válce popravý ale tenhle ten stužka ten prostě mláti byl takový zámek hraková ten nebo když se škoda žil ve se ty stromy tak mlátili těma lupavama

Keyword



Transcript

Contains



Theresienstadt (Czechoslovakia : Ghetto)
Czechoslovakia 1943
[housing conditions in the ghettos](#)

jako tato skupina která byla po tu dobu těch třech neděl v tom v těch veletřních boudách tak i v terezíně jsme se potom scházeli pokud to bylo možný protože podporu to nařízení že se nesmí říct že se mohou stýkat muži se ženami my jsme tam měli také děvčata v našem věku takže jsme se potom ještě v tom terezíně scházeli teda v tom terezíně jednak jak se ta ... do té no v praze společnosti která tam byla zavřená řekněme jak jste byli ubytováni a celá řada na vás udělalo dojmy vůbec toho tak ten první do jednoho dítěte to velice takový veliký odstup času ale | je to asi tak ty kteří byli mladý ty čtrnáctiletí patnáctiletí nevím do kolika let to bylo snad už ty šli do jugendheimu byly tam ty takzvané dětský domovy byly tam dětský domovy pro chlapce byly tam dětský domovy ... ovšem je třeba se na to dívat těch podmínkách ghetta tohle zní strašně vznešeně všechno to byly hromadný ubikace kde na takový místnosti jako je tady bydlelo třeba dvacet lidí to zní strašně vznešeně ale šli prostě děvčata sly zvlášť chlapi šli také zvlášť ten chlapecký jugendheim teda ten domov mládeže byl v hannoveru nahoru ale mě už bylo moc a našemu jirkovi také zase jsme tam byli přestáří tak my už jsme do těch dětských domovů nešli my jsme museli do mužských ubikací to znamená že já s naším jirkou a s ostatními z té naší skupiny jsme byli ubytováni v hannoveru ... na půdách hannoveru | ... a tam už byli před vámi nějaký jako bylo plně obsazený to bylo plně terezín byl plný v té době když jako přijeli třeba ty spolubydliči čekali jsme se přidali ... no tam se nějaký takový velký ale nenávisť nebo dvě zelenou něco na to na úplně normálně heledte tam byl takový velký pohyb vězňů když si uvědomíte že tam bylo v době toho kdy byl terezín plný čtyřicet tisíc lidí měste který mělo v míru tři tisíce obyvatel tak sem si dovedete představit jak tam byli jaký tam byly podmínky tam byly třípatrový palandy ... a na těch místnostech na těch půdách a řekla sem a teď ty kasárna v těch domech bydlel strašná spousta lidí ... a obměňovalo se to já sem to tam nezažil že jo ten odchod těch transportů a když chodily transporty tak ty lidi odešel syrový tam přicházely | nějakou činnost jste tam viděli nebo pracovního lágru museli jsme pracovat v terezíně aspoň ty co byli starší museli pracovat ne- nevím jak to měli organizoval malinko zrádná konkrétně já konkrétně jak sem říkal že sem pracoval u těch sedláků tyčí čili utíkal tam tak sem trošku jako přidělal u zedníků a tak tak sem se tam hlásil jako zedníky orchestr bylo takový pomocník zednický házeli si tam nebyly auto odveze nic nevěděli tak sem tam dělal zedničinu a potom když tam byl postavený stan na náměstí nevím jestli tu historii znáte tam byl postavený my jsme tomu říkali cirkus obrovské stan ... a němci tam zavedli výrobu tak jsme byli naverbovaný do tohoto stanu tam se pro frontu na východě kompletovaly takový soupravy do beden pro | motorový vozidla do těch velkých mrazů to znamená byla tam letlampy a byly tam takový věci který každý to motorový vozidlo mělo dostat tam

Keyword



Transcript

Contains



Theresienstadt (Czechoslovakia : Ghetto)
Czechoslovakia 1943
[housing conditions in the ghettos](#)

jako tato skupina která byla po tu dobu těch třech neděl v tom v těch veletřních boudách tak i v terezíně jsme se potom scházeli pokud to bylo možný protože podporu to nařízení že se nesmí říct že se mohou stýkat muži se ženami my jsme tam měli také děvčata v našem věku takže jsme se potom ještě v tom terezíně scházeli teda v tom terezíně jednak jak se ta ... do tě no v praze společnosti která tam byla zavřená řekněme jak jste byli ubytováni a celá řada na vás udělalo dojmy vůbec toho tak ten první do jednoho dítěte to velice takový veliký odstup času ale | je to asi tak ty kteří byli mladý ty čtrnáctiletí patnáctiletí nevím do kolika let to bylo snad už ty šli do jugendheimu byly tam ty takzvané dětský domovy byly tam dětský domovy pro chlapce byly tam dětský domovy ... ovšem je třeba se na to dívat těch podmínkách ghetta tohle zní strašně vznešeně všechno to byly hromadný ubikace kde na takový místnosti jako je tady bydlelo třeba dvacet lidí to zní strašně vznešeně ale šli prostě děvčata sly zvlášť chlapi šli také zvlášť ten chlapecký jugendheim teda ten domov mládeže byl v hannoveru nahoru ale mě už bylo moc a našemu jirkovi také zase jsme tam byli přestáří tak my už jsme do těch dětských domovů nešli my jsme museli do mužských ubikací to znamená že já s naším jirkou a s ostatními z té naší skupiny jsme byli ubytováni v hannoveru ... na půdách hannoveru | ... a tam už byli před vámi nějaký jako bylo plně obsazený to bylo plně terezín byl plný v té době když jako přijeli třeba ty spolubydliči čekali jsme se přidali ... no tam se nějaký takový velký ale nenávisť nebo dvě zelenou něco na to na úplně normálně heledte tam byl takový veliký pohyb vězňů když si uvědomíte že tam bylo v době toho kdy byl terezín plný čtyřicet tisíc lidí měste který mělo v míru tři tisíce obyvatel tak sem si dovedete představit jak tam byli jaký tam byly podmínky tam byly třípatrový palandy ... a na těch místnostech na těch půdách a řekla sem a teď ty kasárna v těch domech bydlel strašná spousta lidí ... a obměňovalo se to já sem to tam nezažil že jo ten odchod těch transportů a když chodily transporty tak ty lidi odešel syrový tam přicházely | nějakou činnost jste tam viděli nebo pracovního lágru museli jsme pracovat v terezíně aspoň ty co byli starší museli pracovat ne- nevím jak to měli organizoval malinko zrádná konkrétně já konkrétně jak sem říkal že sem pracoval u těch sedláků tyčí čili utíkal tam tak sem trošku jako přidělával u zedníků a tak tak sem se tam hlásil jako zedníky orchestr bylo takový pomocník zednický házeli si tam nebyly auto odveze nic nevěděli tak sem tam dělal zedničinu a potom když tam byl postavený stan na náměstí nevím jestli tu historii znáte tam byl postavený my jsme tomu říkali cirkus obrovské stan ... a němci tam zavedli výrobu tak jsme byli naverbovaný do tohohle stanu tam se pro frontu na východě kompletovaly takový soupravy do beden pro | motorový vozidla do těch velikých mrazů to znamená byla tam letlampy a byly tam takový věci který každý to motorový vozidlo mělo dostat tam

Odkazy na závěr

- ▶ Shoah Foundation
<http://www.usc.edu/schools/college/vhi/>
- ▶ Malach Project
<http://www.clsp.jhu.edu/research/malach/>
- ▶ Cross-Language Evaluation Forum
<http://www.clef-campaign.org/>
- ▶ Cross-Language Speech Retrieval track of CLEF
<http://clef-clsr.umiacs.umd.edu/>