

De l'encodage à la publication

Les éditions en ligne de la *European Holocaust Research Infrastructure* et leur chaîne éditoriale

Sarah Bénéière

Master "Technologies numériques appliquées à l'histoire"

Soutenance du mémoire (25/09/2023)

Contexte du stage

EHRI

- *European Holocaust Research Infrastructure*
- Infrastructure de recherche sur la Shoah
- Financement européen (FP7, Horizon 2020, Horizon Europe)
- Publication d'éditions thématiques en ligne de documents d'archives

Inria

- *Institut national de recherche en informatique et automatique*
- Équipe coordonnée par Laurent Romary
- Développement de méthodes et outils informatiques pour faciliter la recherche archivistique sur la Shoah

Missions

Objectifs principaux

1. Rédaction d'une ODD exhaustive et documentée des documents EHRI
2. Automatisation de la chaîne éditoriale par la création de *scripts* d'encodage et de transformation
3. Mise en place d'une plateforme de publication TEI Publisher pour les éditions EHRI

Tâches confiées

- Rédaction de l'ODD : documentation et schéma de validation
- Réflexion sur l'automatisation de l'encodage et création de *templates*
- Rédaction d'un cahier des charges pour l'application TEI Publisher et création de maquettes

Enjeux

Encodage des documents

- Analyser l'encodage des éditions publiées
- Tenir compte des capacités et envies des éditeur·ice·s du projet
- Proposer un nouveau modèle d'encodage : ODD et *templates*

Méthode de publication

- Analyser les éditions en ligne publiées avec Omeka
- Définir les besoins du site à partir des attentes des éditeur·ice·s
- Transposer ces besoins à une application TEI Publisher et proposer une alternative adaptée

Résultats

Rappel des objectifs

1. Rédaction d'une ODD exhaustive et documentée des documents EHRI
2. Automatisation de la chaîne éditoriale par la création de *scripts* d'encodage et de transformation
3. Mise en place d'une plateforme de publication TEI Publisher pour les éditions EHRI

Livrables

- ODD
- Schéma de validation RelaxNG

Travail préparatoire

- **Script de recherche** dans le corpus (Python)
- **Tableau récapitulatif** auquel peuvent s'ajouter de nouvelles éditions

Total number of documents in the collection

Comments on the use of the tag and its attributes

Tag Name	header	body	Number of Occurrences				List of @Attributes used in EHRI editions	ODD EHRI	Comment(s)	Files Involved
			Gezette Elucht (104)	Early Testimonies (119)	Diplomatic Reports (72)	Nisko (40)				
ab	<input type="checkbox"/>	<input checked="" type="checkbox"/>				6	@type="dinkus"	<input checked="" type="checkbox"/>	Use <metamark> instead (see Encoding Guidelines).	EHRI-NISKO-19400102_DE.xml
abstract	<input checked="" type="checkbox"/>	<input type="checkbox"/>	97	113		3		<input checked="" type="checkbox"/>	If the title of the document is not explicit enough, a good practice is to add an <abstract> tag and a summary of the document's content.	
author	<input checked="" type="checkbox"/>	<input type="checkbox"/>	208	293		120		<input checked="" type="checkbox"/>		
availability	<input checked="" type="checkbox"/>	<input type="checkbox"/>	104	119		40		<input checked="" type="checkbox"/>	Must be included in the Diplomatic Reports collection.	
bibl	<input checked="" type="checkbox"/>	<input type="checkbox"/>	135	119	72	36	@type	<input checked="" type="checkbox"/>		
biblScope	<input checked="" type="checkbox"/>	<input type="checkbox"/>	10				@unit	<input type="checkbox"/>	Use <bibi> tag instead.	EHRI-BF-19390202_DE.xml EHRI-BF-19380316_DE.xml EHRI-BF-19380408_DE.xml EHRI-BF-19380418_DE.xml EHRI-BF-19380421c_DE.xml EHRI-BF-19380723_DE.xml EHRI-BF-19380830b_DE.xml
birth	<input checked="" type="checkbox"/>	<input type="checkbox"/>	1				@when	<input type="checkbox"/>	Including the <birth> tag in the ODD is not relevant.	EHRI-BF-19380911_DE.xml
body	<input type="checkbox"/>	<input checked="" type="checkbox"/>	104	119	72	40		<input checked="" type="checkbox"/>		
catRef	<input checked="" type="checkbox"/>	<input type="checkbox"/>	117			37	@target	<input checked="" type="checkbox"/>	Guidelines: <catRef> (category reference) specifies one or more defined categories within some taxonomy or text typology. Values for the @target attribute include: "Schmuggel" (smuggling), "Schicksale" (fate/destiny), "Flüchtlingspolitik" (Refugee Policy), "IKG Wien" (Israelitische Kultusgemeinde Wien), "Berichte" (reports), etc. Values must be normalized (capital letter at the beginning of German nouns).	

Tag's place of occurrence in the XML tree

List of the files containing (or missing) the tag

Fonctionnement du tableau récapitulatif

Résultats

Rappel des objectifs

1. Rédaction d'une ODD exhaustive et documentée des documents EHRI
2. Automatisation de la chaîne éditoriale par la création de *scripts* d'encodage et de transformation
3. Mise en place d'une plateforme de publication TEI Publisher pour les éditions EHRI

Livrables

- ***Templates*** pour l'encodage des fichiers (métadonnées et entrées dans les index)

```
<place type="" xml:id="">
  <placeName xml:lang="en"/>
  <placeName xml:lang="{original language}"/>
  <location>
    <geo/>
    <country key="{country ID}"/>
  </location>
  <idno type="geonames"/>
  <idno type="wikidata"/>
  <note/>
</place>
```

Template d'une entrée dans l'index de lieux

Résultats

Rappel des objectifs

1. Rédaction d'une ODD exhaustive et documentée des documents EHRI
2. Automatisation de la chaîne éditoriale par la création de *scripts* d'encodage et de transformation
3. Mise en place d'une plateforme de publication TEI Publisher pour les éditions EHRI

Livrables

- **Cahier des charges** avec une correspondance entre les critères des éditions EHRI et les *web components* proposés par TEI Publisher
- **Maquettes** de l'application

EHRI Digital Editions



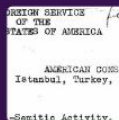
Bordered Escape

This edition gathers documents testifying of the increasingly restrictive refugee policy in Czechoslovakia, following the "Anschluss."



Early Holocaust Testimony

This edition gathers accounts of the persecution of Jews from the Nazi takeover of power in Germany (1933) to the Eichmann Trial (1961).



Diplomatic Reports

This edition gathers reports on the persecution and murder of European Jews.



The Nisko Deportations

This edition gathers documents on the deportation of thousands of Jews to Nisko am San (Poland).

This application was developed as part of the EHRI project with the help of TEI Publisher.



Maquette de la page d'accueil de l'application avec TEI Publisher

Métadonnées

Transcription

Mots-clés

< Ray Johnson on the reaction of the... Alexander H. on the signature of... >

John F. Montgomery on the "Hungarian Jew Bill" in 1939

[METADATA](#) [DOCUMENT TEXT](#) [MAP](#) [REFERENCES](#)

METADATA

6416-08-16005-1-102

1939-05-19 | Budapest

National Archives and Records Administration (RG 24 - Records of the Department of State relating to the internal affairs of Hungary; 1915-1945, Montgomery, Del. 1947 (Jews)), Box 844.6274-7-161. Original in English.

DOCUMENT TEXT

[ENGLISH](#)

Text from page2

The Department will note that the law was modified substantially. Primarily, the class of Jews to be considered officially as Christians was enlarged but its economic restrictions were increased, thus creating a sort of second-class citizenship. This represents the real compromise made between the two houses of Parliament. Also, the idea of placing the Jews in a distinct category for political purposes was discarded. Instead, the Jews retain free active and passive suffrage, but the conditions of their citizenship were made more severe. Agricultural property may be taken from the Jews, their right to acquire real estate is limited, and the restriction of their employment is greatly increased. Jewish representatives complain bitterly against the special police restriction by which the law was amended. It appears that, in practice, Jewish youth will suffer the most.

MAP

Toggle Fullscreen 13



REFERENCES

United States

Issued 2 years ago

The United States of America were neutral during the first two years of the Second World War. They were brought into the conflict by the Japanese attack on Pearl Harbor on 7 December 1941 and became one of the decisive belligerent nations, drawing Japan to the Pacific and heavily contributing to the war effort against the Axis powers, e.g. during the island-hopping in South Africa, Italy and France. American forces invaded German territory early in 1945. On the eve of the Second World War...

94 repeating document available

[View this item on the Internet Portal](#)

National Archives and Records Administration

NARS | National Archives

Revised 16 days ago

1939-05-19 | Budapest, Hungary, Poland, United States of America

[View this item on the Internet Portal](#)

Fac-similé

Carte

Présentation verticale avec Omeka

B. E., female, on the deportation of Jews from Nagyvárad to Auschwitz-Birkenau in 1944 and her slave labour in other Nazi concentration camps

ADD

View

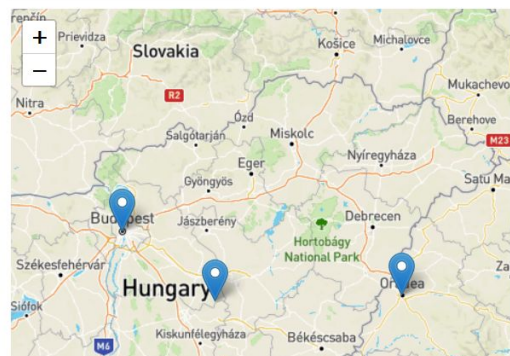
Translation

Fent nevezett előadja a következőket:

Nagyvárad on kb. 25-30000 zsidó lakott, kereskedők, sok **orvos, ügyvéd**, iparos, stb. nagyrészt jómódú emberek. Férjem műbútorasztalos volt, nagyon szépen keresett, raktárában a **deportálás** idején, akkori érték szerint 30000 pengő értékű bútor volt. Saját lakásunk és berendezésünk 100.000 pengő értéket tehet ki. Az általános rendelkezések szerint április 5-én kezdődött a **gettóba költöztetés**. A város egy részét, kb. 25 utcát jelöltek **gettónak** és kocsikon, **csendőrök** vittek oda bennünket. Mindent, amit akartunk magunkkal vihattunk a **gettóba**. Benn a gettóban nem volt szabad járkálnunk. Egy **szobában** kb. 12-15 **személy** lakott, a földön aludtunk, nagyon szűk helyen. Népkonyha **élelmezett** bennünket. Néhányan **dolgoztak** a

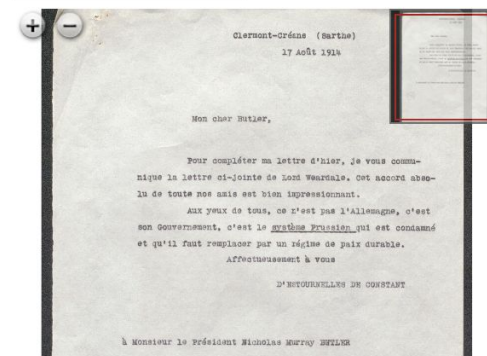
View

Map



View

Facsimile



Exemple d'affichage d'un document sur l'application en cours de développement

Perspectives au-delà du stage

Pour l'encodage

- Homogénéisation des éditions déjà publiées
- Conformité des futures éditions

Pour la publication

- Migration vers TEI Publisher
- Exploration de l'outil d'annotation intégré à TEI Publisher

Merci pour votre attention !