



PAM Pokrajinski
arhiv
Maribor

Moderna
arhivistika

Časopis arhivske teorije in prakse
Journal of Archival Theory and Practice

Letnik 5 (2022), št. 2 / Year 5 (2022), No. 2

Maribor, 2022

Prejeto / Received: 30. 06. 2022

1.04 Strokovni članek

1.04 Professional Article

<https://doi.org/10.54356/MA/2022/VGLG3975>

PROJECT “ENRICH EUROPEANA +” AND THE STATE ARCHIVES IN ZAGREB: TRANSCRIBATHON EVENT

Sonja GALINA

State Archives in Zagreb, Croatia

sonja.galina@daz.hr

Abstract:

This paper will present the project "Enriching Europeana through citizen science and artificial intelligence - unlocking the 19th century" in which the State Archives in Zagreb is included. The project combines citizen science campaigns and artificial intelligence (AI) in facilitating access to 19th-century handwritten documents not only for researchers but also for students, amateur historians, and the general public. This paper will explain the organization of the project, like the divisions into activities as well as the desired outcomes of each of them. It will also explain the role of all the partners who are involved in the project. As for the State Archives in Zagreb, this paper will describe in detail all the tasks that the State Archives is involved in. The main focus will be on representing the organization of the Transcribathon event. The Transcribathon event serves as a crowdsourcing element of the project. The paper will exhibit elements of the Transcribathon event, the main goals of the event, the targeted group, and the materials on which the transcriptions will be done. Also, it will highlight the importance of the visibility of the archive and its materials for educational and scientific purposes.

Key words:

Europeana, Archive, Transcription, Crowdsourcing, Transcribathon

Izvleček:

Projekt Enrich Europeana+ in Državni arhiv v Zagrebu : dogodek Transcribathon

Avtorica predstavlja projekt »Bogatenje Europeane s pomočjo znanja državljanov in umetne inteligence – razkrivanje 19. stoletja« (krajše Bogatenje Europeane+), pri katerem sodeluje tudi Državni arhiv v Zagrebu. Projekt združuje zajemanje znanja državljanov in umetno inteligenco z namenom omogočanja lažjega dostopa do rokopisnih dokumentov 19. stoletja, ne samo za raziskovalce, ampak tudi za študente, ljubiteljske zgodovinarje in splošno javnost. Avtorica bo predstavila organizacijo projekta, npr. porazdelitev aktivnosti oziroma nalog, in željene rezultate vsake od nalog. Prav tako bo pojasnila vlogo partnerjev, vključenih v projekt. Podrobneje bo predstavila naloge, ki jih je v okviru projekta prevzel Državni arhiv v Zagrebu. Glavni poudarek bo na opisu organizacije dogodka Transcribathon, njegovih elementov, glavne ideje, ciljnih skupin in pomembnosti arhiva ter njegovega gradiva za izobraževalne in znanstvene namene. Transcribathon igra znotraj projekta vlogo elementa »vpreženja množic« (crowdsourcing). Udeleženci dogodka bodo študenti Filozofske fakultete v Zagrebu. O dogodku bodo obveščeni po Facebooku in drugih družbenih omrežjih.

Ključne besede:

Europeana, arhiv, transkripcija, crowdsourcing, Transcribathon

1. Introduction

For the first time ever the State Archives in Zagreb is involved in a project with Europeana and other European partners. The State Archives was previously included in smaller projects with our neighboring countries, but this is the first project that involves European funding, regulations, strict milestones, and deadlines. The project that I’m referring to is named “Enriching Europeana through citizen science and artificial intelligence - unlocking the 19th century” or for short “Enrich Europeana +”.



Picture 1: Europeana logo

The project started on April 1st, 2021, and was supposed to end on September 30th, 2022. Because of technical and other issues, the Project Management Board of the “Enrich Europeana +” project decided that we should submit a request for prolongation. The proposed prolongation would take another 4 to 6 months. At the time of writing this paper, we were still waiting for the final decision. The project is estimated at € 1,228,755 of which 75 % is co-financed by CEF - Telecom financing and the other 25% is co-financed by partners.

A total of 9 partners are involved in the project:

1. AIT Austrian Institute for Technology GmbH from Austria
2. Stichting Europeana from Netherlands
3. Facts & Files Historisches Forschungsinstitut Berlin Drauschke - Schreiber Partnerschaftsgesellschaft from Germany
4. Instytut Chemii Bioorganicznej Polskiej Akademii Nauk from Poland
5. READ-COOP SCE from Austria
6. The Provost, Fellows, Foundation Scholars, and the other members of the Board, of the College of the Holy and Undivided Trinity of Queen Elizabeth near Dublin from Ireland
7. Dublin City Council from Ireland
8. Uniwersytet Wroclawski from Poland and
9. State Archives in Zagreb from Croatia

The lead role in this project has AIT Austrian Institute for Technology GmbH from Austria (Grant Agreement, 2020). Other project partners also have their roles, but more on that later in the paper.

This project is based on 19th century manuscripts. The manuscripts are collected from different content providers involved in the project. These manuscripts were chosen because of their historical significance, the importance for the territory from which they are coming and their ability to bring the past closer. Working on said manuscripts has its difficulties. The main problem is the handwritten texts in different writing styles, often too hard for quick skimming. So they take up a lot more time to go through them and consequently are not used so much for scientific and educational purposes. As they are not used so much, these manuscripts can often be overlooked in researching different topics. Project “Enrich Europeana +” aims to resolve these problems. Providing new content that is important and valuable for a certain area or country represents a fundamental task in resolving the said problems. Project partners with the role of content providers have the assignment to digitize important manuscripts and collections that will be used for transcription activities with the help of artificial intelligence and the general public involved in citizen science activities.

2. Project goals

The project has three main goals. They will be achieved through all of the project activities. One of the goals is the development of innovative AI solutions for automating the transcription of manuscripts and their semantic enrichment. This will be achieved through further development of the Transcribathon platform (Grant Agreement, 2020). Transcribathon is a crowdsourcing platform that allows users to work on historical sources and become an important element in the process of digitizing them. Even though the Transcribathon platform was developed during a previous project, there is still room for improvement and growth (<https://transcribathon.eu/>). The second goal of the project is the development of new services based on innovative AI solutions for achieving transcription and enrichment automation. This goal needs new materials on which the development could be done and tested. For that reason, the project will aggregate over 100,000 new documents in Europeana using existing national aggregators i. e. Digital Repository of Ireland (DRI), and the Uniwersytet Wroclawski from Poland. The aggregated documents will be used to achieve the third goal, organizing crowdsourcing events where the aggregated documents will be transcribed and enriched. Enrichments obtained at the crowdsourcing events will be merged with the original collections and reloaded in Europeana, thus increasing the quality of the collection's metadata (Grant Agreement, 2020).

3. Project activities

As I mentioned before, the project and its main goals are divided into 5 project activities. Those are:

1. Content aggregation & data exchange infrastructure
2. Automated transcription and enrichment services
3. Innovation and scalability of Transcribathon Platform
4. Dissemination and citizen science crowdsourcing campaigns
5. Project management and sustainability planning

The first project activity, Content aggregation & data exchange infrastructure, includes the aggregation of new, high-quality content into Europeana that complies with the Europeana Publishing Framework. The content that is already ingested into Europeana has to be verified regarding the quality of content and its metadata. If the content and metadata are not of the right quality and tier, they must be improved. Other tasks in this activity include the exchange of data and standardization and the validation of metadata.

The second activity, Automated transcription and enrichment services, includes HTR (Handwritten Text Recognition) service, semantic enrichment, topic discovery, and transcription and translation of metadata. The main task of this activity is to develop innovative solutions for automating the process of transcription and semantic enrichments. To achieve this task we are using 19th-century manuscripts which will be automatically transcribed using the HTR technology. The same manuscripts will be used for content analyses and automatic topic detection in the text.

The third activity is Innovation and scalability of the Transcribathon platform. It integrates the work of AI into the user interface of the Transcribathon platform. This integration intends to improve the user interface and user experience. Also, it will develop new and improved enrichment-related activities on the Transcribathon platform. The activity consists of the development and integration of the HTR web plugin, the standardization of the Transcribathon API, the improvement of the user experience within the Transcribathon tool, and the incorporation of the Transcribathon within the Europeana CSP.

The fourth activity is Dissemination and citizen science crowdsourcing campaigns. The main focus of this activity is reaching out to the public and cultural heritage institutions. In doing so, these groups and institutions are familiarized with the Transcribathon platform and all the activities around it such as Transcribathon events, workshops, and webinars. This activity also involves dissemination and editorial activities such as promoting this project at various congresses, conferences, and workshops, preparing blog posts for Europeana and Europeana Pro websites, and so on. The last part of this activity is assessing the impact of different parts of the project as well as the project in its entirety.

Project management and sustainability planning are the fifth and last project activities. These activities coordinate and monitor all the stated activities as well as the timely delivery of milestones. They include project management, financial management, regular participation in Project Management Board meetings and any other necessary meetings, and enabling the long-term sustainability of the Transcribathon platform. For the long-term sustainability of the platform, a sustainable business model will be developed (Grant Agreement, 2020).

4. Role of the partners

Every project partner has a role within the project. The following partners are partaking in the first activity Content aggregation & data exchange infrastructure: Uniwersytet Wroclawski, Dublin City Council, State Archive in Zagreb, Stichting Europeana, The Provost, Fellows, Foundation Scholars and the other members of Board, of the College of the Holy and Undivided Trinity of Queen Elizabeth near Dublin and Instytut Chemii Bioorganicznej Polskiej Akademii Nauk. In the second activity, Automated transcription and enrichment services are included the following partners: Facts & Files Historisches Forschungsinstitut Berlin Drauschke - Schreiber Partnerschaftsgesellschaft, Stichting Europeana, AIT Austrian Institute of Technology

GmbH and READ-COOP SCE. In the third activity, Innovation and scalability of the Transcribathon platform, are included Facts & Files Historisches Forschungsinstitut Berlin Drauschke - Schreiber Partnerschaftsgesellschaft, Stichting Europeana, AIT Austrian Institute of Technology GmbH, READ-COOP SCE, Instytut Chemii Bioorganicznej Polskiej Akademii Nauk and State Archives in Zagreb. In Dissemination and citizen science crowdsourcing campaigns, the fourth activity, the participants are Uniwersytet Wrocławski, Dublin City Council, Facts & Files Historisches Forschungsinstitut Berlin Drauschke - Schreiber Partnerschaftsgesellschaft, The Provost, Fellows, Foundation Scholars and the other members of Board, of the College of the Holy and Undivided Trinity of Queen Elizabeth near Dublin, Instytut Chemii Bioorganicznej Polskiej Akademii Nauk, State Archives in Zagreb, Stichting Europeana, AIT Austrian Institute of Technology GmbH, and READ-COOP SCE. All of the partners partake in activity number 5, Project management and sustainability planning (Connecting Europe Facility, 2020).

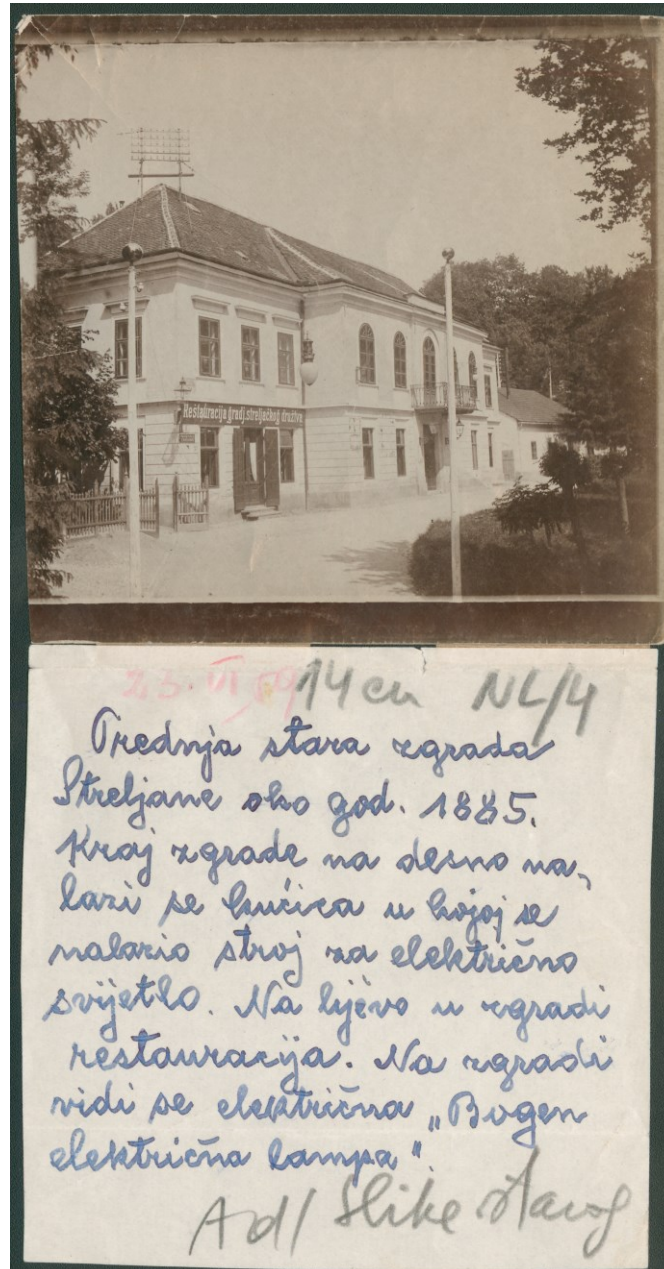
4.1 Detailed role of the State Archives in Zagreb

As was stated in the previous paragraph, the State Archives in Zagreb partakes in activity numbers 1, 3, 4, and 5. In this part of the paper, we will give a more detailed overview of the State Archives' involvement in those activities.

For the Activity 1 State Archives in Zagreb, as a content provider, has selected the following funds, collections, and documents:

1. HR-DAZG-857 Collection Ulčnik Ivan, 1848./1930. – 1745 records
2. HR-DAZG-839 Archival fond Dragutin Hirc, 1844./1921. – 141 records
3. HR-DAZG-1 Archival fond Government of the Free Royal City of Zagreb, Records of the city government meetings (Poglavarstvo slobodnog kraljevskog grada Zagreba, Zapisnici sjednica Poglavarstva), 1800./1845. – 4 records
4. Selection of letters, 1832./1903. – 8 letters from archival fonds HR-DAZG-823 Family Deželić no. 1260, no. 1695, HR-DAZG-838 Guteša Ilija no. 42, HR-DAZG-845 Collection Makanec Alfred no. 982, no. 2168, no. 2568, HR-DAZG-852 Radić Stjepan no. 133, no. 161
5. HR-DAZG-24 Archival fond Zagreb City Government, Supply Department II-3 (Gradsko poglavarstvo Zagreb, Opskrbni odjel II-3), Consumer cards, 120,621 records

We selected important archival fonds and collections to represent the 19th century in Zagreb. For example, Collection Ulčnik Ivan contains photographs, paper cutouts, and notes written by Ivan Ulčnik describing certain persons, events, buildings, streets and other vital parts of the City of Zagreb from the 14th century to the mid-20th century. He often wrote short notes on the back sides of the pictures to describe them and correctly record them. It was those handwritten notes, alongside the theme of the materials, that made this collection acceptable for the project.



Picture 2: Photograph and a handwritten note about the shooting range building in Zagreb (HR-DAZG-857 Collection Uličnik Ivan no. 1330)

As part of Activity number 3, the State Archives in Zagreb takes a minor role as a link between the project and the company Bonsai Tech from Zagreb and the Department of information and communication sciences at the Faculty of Humanities and Social Sciences in Zagreb. Both Bonsai Tech and the Department are the main reason why the State Archives in Zagreb is a part of this project. We started a project with Bonsai Tech and the Department of information and communication sciences using the Consumer cards. In this previous project, the Consumer cards were used to develop new software for recognizing different elements in structured texts as are the Consumer cards. For the next phase of the project, a larger crowdsourcing platform was needed. At the same time, the "Enrich Europeana +" project was in its beginnings. The work that was done on Consumer cards was presented to the project organizer who accepted the Consumer cards, and later other materials, to be a part of the "Enrich Europeana +" project.

Grad Zagreb

A.

Ulica, trg ili ina oznaka Levenska ceste

REG. BROJ:

101200

Kbr. 119

Prezime i ime podnosioca prijave:

Rebarkovića Zrinka

Prezime i ime i stan (kućevlasnika):

Rebarkovića Zrinka

Levenska ceste br 119

Potrošačka prijavnica
za kućanstva i samce – samice.

POTPISANI OVIM MOLIM, DA MI SE IZDA POTROŠAČKA ISKAZNICA, TE PODJEDNO IZJAVLJUJEM POD ODGOVORNOSTU IZ ČL. 18 ST. 1 NAREDBE O RASPODJELI (RACIONIRANJU) ŽIVOTNIH NAMIRNICA, DA SE U MOJEM KUĆANSTVU HRANE SLIJEDEĆE OSOBE:

REDNI BROJ	PREZIME I IME	ODGINA ROĐENJA	ODNOS PREMA PODNOSIOCU PRIJAVE ODN. STARIŠINI	ZANIMANJE	MJESTO RADA
1	RABADŽIJA LjUBIGA	1913	suprug	kućevlasnik	
2	RABADŽIJA ĐURO	1908	suprug	radnik	lokalni upravitelj Blijas i šah
3	RABADŽIJA STANKO	1936	sin	dijete	
4					
5					
6					
7					
8					
9					
10					

PRIRAST: (ispunjava vlast)

1					
2					

ODPAD: (ispunjava vlast)

1					
2					

UJEDNO IZJAVLJUJEM POD ISTOM ODGOVORNOSTU, DA U MOJEM KUĆANSTVU POSTOJE SLIJEDEĆE ZALIHE ŽIVOTNIH NAMIRNICA U KILOGRAMIMA ODNOSNO LITRAMA:

PŠENICA		RAŽ		JEČAM		KUKURUZ		TJESTE- NINE	JESTIVO ULJE	MAST	SOL	ŠEĆER	KAVA	SAPUN ZA PRANJE	PETRO- LEJ
ZRNO	BRAŠNO	ZRNO	BRAŠNO	ZRNO	BRAŠNO	ZRNO	KLIP	BRAŠNO							
/	5 kg	/	/	/	30 kg	/	/	10 kg	/	10 kg	/	/	/	/	/
P O V E Č A N J E															

Živežne namirnice nabavljat ću:

U radnji

Dragica Kobilica

ulica

Rebar br 21

DRŽAVNI ARHIV U ZAGREBU

ZAGREB,

13. Veljače

1941.

POTPIS PODNOSIOCA PRIJAVE:

Rebarkovića Zrinka

OVA PRIJAVNICA STOJI DIN 076 I NE SMIJE SE SKUPLJE PRODAVATI.
OBRAZAC K ČL. 2 ST. 8 NAREDBE O RASPODJELI (RACIONIRANJU) ŽIVOTNIH NAMIRNICA OD 27. SIBJEČNJA 1941.

Picture 3: Example of a Consumer card (HR-DAZG-24 Archival fond Zagreb City Government, Supply Department II-3, no. 313151)

As part of Activity number 4, the State Archives in Zagreb presents the “Enrich Europeana +” project at different conventions, conferences, and meetings, writes about the project and its activities and presents the work that has been done so far. Also, it has an obligation to organize a Transcribathon event with a particular group. More on the Transcribathon event later.

The fifth activity includes a representative of the State Archives in Zagreb at every Project Management Board meeting, as well as at monthly meetings for certain Activities. In this way, we inform our project partners about our progress, work that has been or needs to be done, any problems that occurred, and so on (Grant Agreement, 2020).

5. Transcribathon event in Zagreb

As I mentioned before, State Archives in Zagreb has an obligation to organize a Transcribathon event in Zagreb. For the event, we will be using the Transcribathon platform. Transcribathon is a crowdsourcing platform that allows its users to explore different historical materials such as letters, diaries, postcards, pictures, and so on. The materials were collected by various cultural heritage institutions, digitized, uploaded to Europeana, and then used for different activities on the Transcribathon platform. Those activities include reading, transcribing, tagging, and adding geographical tags for every individual document that is available on the platform. With those additions, the documents have increased searchability and visibility as well as gained a better potential for future use in education and research. Signing up on the platform is free. At the time of writing this paper, there was uploaded more than 372 000 documents, transcribed more than 45 million characters, and added 122 924 enrichments to the documents. The development of this platform is one of the “Enrich Europeana +” project goals. The aim is to increase the scalability and functionality of the Transcribathon platform to become a service that the cultural heritage institutions would use to organize their crowdsourcing campaigns and to use cultural heritage items from repositories in a new way, gaining new audiences and improving multilingual metadata. With that in mind, we begin to organize our own event.

The Transcribathon event can be organized either online or in person. We decided to opt for a physical event. But we came to a problem immediately. The earthquake that struck Zagreb in March 2020 left the building of the State Archives in Zagreb in a devastating shape. All of the main storage spaces were not fit to hold the very valuable materials anymore. All of the holdings were transferred to the archive's second location. The working space for the beneficiaries is safe to use, but it is not well equipped. The Transcribathon event needs a larger number of computers with a stable internet connection. At the moment, we have only two computers in the reading room designated for beneficiaries. So we turned to our partner in the previous project, the Faculty of Humanities and Social Studies in Zagreb, primarily to the Department of information and communication sciences. We will organize this event together. The event was first planned for May 2022, but we were not able to deliver our materials via the Croatian national aggregator which is in the final phase of its development. For that reason, we decided to postpone the event for Autumn 2022, most likely in November. The event will be 5 days long with a preparation week beforehand. The preparation week will include informational videos on how to work on the Transcribathon platform and a short overview of the materials that will be transcribed. We plan to film a short promotional video about the event and make a PR campaign. Also, we plan to use our social platforms, such as Facebook and Instagram, to promote the event among our desired audience. Our participants are students at the Faculty of Humanities and Social Studies in Zagreb. We hope to get as many students from different departments as

possible such as the Department of information and communication sciences, the Department of history, the Department of art history, language students, and so on.

The Faculty will secure us with a satisfactory workspace such as the computer room in the library as well as the large lecture room D7. In lecture room D7 we plan to have the opening of the Transcribathon event. The idea is to have representatives of the project partners talk about different aspects of their involvement in the project. So we would have a lecture about the Transcribathon platform, its development, and new features on the platform, the "Enrich Europeana +" project, and Europeana itself. We also plan to have a lecture about the materials that the State Archives in Zagreb are providing for the event to put them in the right context. For the transcription part of the event, we plan to use the computer room in the Faculty's library. Transcriptions are made by using the Transcribathon platform for which we need an internet connection. The students will be able to use the computer room in the library as much as they want or need, but it is not entirely necessary for them to use it. They can log in to their accounts from the comfort of their homes or dorms and continue to transcribe the materials there.

The Zagreb Transcribathon event will have a competitive note. The students will be divided into groups of a maximum of five participants. Every group is competing with each other regarding the number of transcriptions, quality of their transcriptions, and the number of enrichments. The participants will have three days to get as much transcription as possible. At the end of the event, student groups will give a short presentation on the materials they had transcribed, things they have learned, and comment on the Transcribathon event. The transcriptions, enrichments, and the presentations will be judged by a three-person committee who will have the last saying regarding the winner group. The winners will be awarded appropriate gifts.

For the transcriptions at the event, we are preparing the Collection of Ivan Ulčnik, a selection of letters, the Collection of Dragutin Hirc, and Records of the City Government Meetings. These collections were chosen because they are a part of the century the project is based on. The collection of Ivan Ulčnik was selected because of the main theme throughout the collection - the city of Zagreb, its buildings, events, and people through the centuries. Ivan Ulčnik collected various materials, photographs, newspaper articles, postcards, etc. He often wrote notes by hand about the theme of a certain document. They could include a short resume of a person, a description of some event that happened in Zagreb, facts about some street, or any other information that was important for Ivan Ulčnik. Transcription of this collection will improve the existing metadata and promote it for educational and promotional purposes. Collection of Dragutin Hirc contains autobiographical memoirs, manuscripts for the book *Stari Zagreb* published posthumously in two volumes, personal documents, newspaper cutouts for his book, and others. His writing is consistent and very neat so it is a good choice for transcription activities. The topics in his memoirs and manuscript are also very appealing to a larger audience which is why this collection was added to the Transcribathon event. The selection of eight letters (archival fonds HR-DAZG-823 Family Deželić, HR-DAZG-838 Guteša Ilija, HR-DAZG-845 Collection Makanec Alfred, and HR-DAZG-852 Radić Stjepan) is based upon the historical figures who wrote those letters. There are letters from Marija Radić, wife of Stjepan Radić, Croatian politician, Josip Juraj Strossmayer, Croatian bishop and politician, Marija Jurić Zagorka, Croatian writer and journalist, and others. The selected Records of the City Government Meetings have all the information about the decisions that were made during the first four years of the 19th century. They are written in Latin as Latin was the official language of the administration. For transcribing these records, we hope to get students of Latin and other languages as well as Latin enthusiasts involved. It would be a great opportunity for them to test their knowledge and skills in transcribing something old but still relatable. We hope to get as

many transcriptions and enrichment as possible at the Transcribathon event. But we are open to the possibility of organizing another event or enabling further transcriptions of our materials by promoting the Transcribathon platform.

The main goals for the Transcribathon event are engaging a large number of students, introducing new technologies to professors and students, and showing them the opportunities for using them for educational purposes. Also, the event will help students to develop new skills such as learning how to work on a crowdsourcing platform, improving their reading of handwritten texts, contextualizing archival materials regarding the period they were created, and so on. The State Archives in Zagreb hopes to gain new enrichments for the materials which would ensure more accessibility and searchability and consequently more usability of its materials. Organizing an event on this scale promotes the State Archives in Zagreb as a forward-thinking cultural heritage institution, interested in new technologies and their applicability. It also shows the State Archives as an institution that tends to use its own materials for educational purposes, not just in the traditional way with students coming to our institution but using new technologies and bringing the materials to them. With this project, the State Archives in Zagreb hopes to gain acknowledgment for its work and the possibility of joining in new and exciting projects with Croatian, European, or global partners.

6. References

Connecting Europe Facility 2014-2020 : TELECOM CALLS FOR PROPOSALS 2020 : APPLICATION FORM, PART A, General information.

Grant Agreement under the Connecting Europe Facility (CEF) – Telecommunications Sector : Agreement No INEA/CEF/ICT/A2020/2288325.

Transcribathon platforma: <https://transcribathon.eu/> (13. 6. 2022.)

POVZETEK

PROJEKT ENRICH EUROPEANA+ IN DRŽAVNI ARHIV V ZAGREBU: DOGODEK TRANSCRIBATHON

Sonja Galina

Državni arhiv v Zagrebu, Hrvatska

sonja.galina@daz.hr

Državni arhiv v Zagrebu sodeluje v projektu »Bogatenje Europeane s pomočjo zbiranja znanja državljanov in umetne inteligence – razkrivanje 19. stoletja« ali krajše »Bogatenje Europeane +«. Gre za skupni projekt Državnega arhiva v Zagrebu in osmih partnerjev iz Evrope: AIT Austrian Institute of Technology GmbH (Avstrija), Stichting Europeana (Nizozemska), Facts & Files Historisches Forschungsinstitut Berlin Drauschke - Schreiber Partnerschaftsgesellschaft (Nemčija), Instytut Chemii Bioorganicznej Polskiej Akademii Nauk (Poljska), READ-COOP SCE (Avstrija), The Provost, Fellows, Foundation Scholars and the other members of Board, of the College of the Holy and Undivided Trinity of Queen Elizabeth near Dublin (Irska), Dublin City Council (Irska) in Uniwersytet Wrocławski (Poljska). Projekt se je začel 1. aprila 2021 in se bo končal v začetku leta 2023.

Glavni cilj projekta je kombinirati znanje državljanov in umetno inteligenco (AI) ter omogočiti dostop do rokopisnih dokumentov iz 19. stoletja. Ti rokopisi so pogosto težje berljivi, preveč poškodovani za uporabo ali pa preprosto nepoznani. Aktivnosti v projektu bodo raziskovalcem, študentom, zgodovinarjem in drugim zainteresiranim posameznikom omogočile uporabo teh zapisov na bolj dostopen način. Projekt naj bi k Europeani dodal več kot 100.000 dokumentov, razvil AI-rešitve za avtomatizacijo transkripcije rokopisov in organiziral dogodke za javnost za promocijo platforme Transcribathlon. Le-ta bo služila transkripciji in prevodu, kakor tudi semantičnim dodatkom. Projekt je razdeljen na pet aktivnosti: zbiranje vsebin in priprava infrastrukture za izmenjavo podatkov; avtomatska transkripcija; inovativnost in obvladljivost Transcribathlon platforme; načrtovanje diseminacije in trajnosti. Državni arhiv v Zagrebu je aktivno vključen v štiri.

Aktivnosti diseminacije in zbiranja znanja množic nalagajo projektnim partnerjem izvedbo digitalizacije pomembnih rokopisov, ki bodo transkribirani na več dogodkih. Eden takšnih bo jeseni 2022 v Zagrebu na Filozofski fakulteti. Udeleženci bodo študenti različnih oddelkov. Pred glavnim dogodkom bodo študenti obiskovali predavanja, ki jih pripravljajo projektni partnerji in drugi priznani udeleženci. Študentje bodo razdeljeni v skupine po pet, imeli pa bodo tri dni časa, da prepišejo kolikor le mogoče veliko dokumentov. Dokumenti, ki bodo predmet omenjene obdelave, so: Zbirka Ivana Ulčnika, izbor pisem, Zbirka Dragutina Hirca in Zapisniki sej mestne uprave. Skupina študentov, ki bo transkribirala največ dokumentov, bo zmagala in osvojila nagrado. Pred dogodkom in po njem bo izvedena tudi anketa o učinkih dogodka samega.