Linked4Resilience: Linked Open Data for Data-Centric Resilience of Damaged Cultural Properties and Infrastructures in Ukraine

Shuai Wang[1,*], Oleksandr Berezko[2], Eirik Kultorp[3], Maksym Stefashko[2,†], Sofiia Fedak[2,†], and Olena Denyshchuk[4,†]

1 Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, the Netherlands
2 Lviv Polytechnic National University, 12 Bandera Street, 79013 Lviv, Ukraine
3 Unaffiliated, Oslo, Norway
4 University of Amsterdam, 1012 WP Amsterdam, the Netherlands

Abstract
Datasets of reported damaged properties are vital for resilience efforts during and after wars. They are essential for the statistical analysis of damages and the distribution of funds and resources. This paper concentrates on these datasets to enhance the resilience of Ukraine's damaged cultural heritage sites and infrastructures as a result of the ongoing Russo-Ukrainian war. More specifically, we use a semiautomatic approach to enrich such datasets and a linked data approach to improve the quality of published datasets and facilitate their integration. We demonstrate our approach by using a well cited dataset published by the UNESCO for damaged cultural properties and the ScienceAtRisk project for damaged educational and research infrastructures. Our final datasets consist of 2,910 and 389 triples respectively. Finally, we demonstrate the use of our datasets with three use cases.

Keywords
Ukraine resilience, data-centric resilience, linked data, UNESCO, cultural heritage

1. Introduction

The Russo-Ukrainian war devastates all aspects of life in Ukraine and has already resulted in damages and loss of properties of diverse types in the country, including museums, cultural heritage, and research infrastructure. The influx of destroyed objects underscored the urgent need for systematic documentation and actions for resilience [1], [2]. The destruction of cultural

SCIA-2024: 3rd International Workshop on Social Communication and Information Activity in Digital Humanities, October 31, 2024, Lviv, Ukraine
* Corresponding author.
† These authors contributed equally.
Shuai.wang@vu.nl (S. Wang); oleksandr.l.berezko@lpnu.ua (O. Berezko); eirik@ekul.no (E. Kultorp); maksym.stefashko.sk.2023@lpnu.ua (M. Stefashko); sofiia.fedak.sk.2022@lpnu.ua (S. Fedak); denyshchukolena@gmail.com (O. Denyshchuk).
0000-0002-1261-9930 (S. Wang); 0000-0002-0664-4339 (O. Berezko); 0009-0005-5211-045X (E. Kultorp); 0009-0003-0984-9035 (M. Stefashko); 0009-0008-0560-4168 (S. Fedak); 0009-0006-4038-3028 (O. Denyshchuk)

properties[2] and infrastructures left many people unable to continue their work and access essential infrastructure, leading to a mass exodus [3]. UNESCO reported an estimated cost of over $1.26 billion for the restoration of public research infrastructure, including 1,443 buildings belonging to 177 public scientific institutions that have been damaged or destroyed from February 2022 to 2024 [4].

Multidisciplinary research has been conducted to record and study damaged cultural properties and infrastructures with a particular focus on the challenges of monitoring and documenting damage, as well as providing recommendations for minimising war's effects on them [1], [5]. Accurate and complete datasets on damaged cultural properties and infrastructures play a crucial role in estimating loss, policy-making, and the arrangement of funding and resources. However, the availability of these datasets and their documentation remains poor, with many hosting platforms and sites lacking the necessary structure and clarity [5]. Datasets are often closed to public use and available only upon request[3]. Additionally, there is no centralised platform where data and information can be gathered, further complicating efforts to manage the data effectively.

The United Nations is among the main publishers of datasets on the damages in Ukraine since the war. For example, satellite images were made publicly available by the United Nations Satellite Center (UNOSAT) and were used for the identification of damaged buildings and other assessments in the Kyiv Oblast [6]. As reported, the reliability of the ground truth data can be limited in the ongoing war zone [6]. Moreover, the frequency of updates and the resolution of images can also influence the results of assessments. In January 2024, UNESCO published a list of damaged cultural properties. The list has been updated monthly with new damages. As of 10th April 2024, the dataset includes 351 properties (see Figure 1). Despite that it has been cited and referenced by some articles [7], [8] and many websites (e.g. e.g. the Museums Association referred to their data[4], Wikipedia has a page with more enriched information based on the reported damage [6]) and articles [7], [8], the entries in the list are missing key information, such as location, date of damage, etc. Moreover, we found outdated spellings based on Russian forms, such as "Zaporizhzhya" (should be "Zaporizhzhia"). Some titles lack details, such as "Monument to workers killed in World War II (renovated in 2016)", "Building of the Research Institute of Venereology (building 1889)", and "Residential historical building in Kharkiv". The title "Music School – Sviatohirsk” refers to multiple sites when searching using Google Maps. The interpretation of some others, e.g. "St. Andrew's Church – Kharkiv", remain unclear with no corresponding geolocation found.

A solution to reduce ambiguity is to convert the unstructured data to linked data with the use of uniformed representation of cities, provinces, geolocation, etc. Our previous work5 focused on the conversion and integration of damaging events [5]. We demonstrated how using

[2] We align with UNESCO's use of the term ‘cultural properties’ to refer to immovable cultural property irrespective of its origin, ownership or status of registration (see the complete definition in Article 1 of the 1954 Hague Convention (https://ihl-databases.icrc.org/en/ihl-treaties/hague-conv-1954/article-1a) as well as facilities and monuments dedicated to culture, including memorials.

[3] For example, the datasets by the Ukrainian Heritage Monitoring Lab (https://www.heritage.in.ua/) are available upon request.

[4] https://www.museumsassociation.org/museums-journal/news/2024/02/unesco-verifies-damage-to-343-cultural-sites-as-war-in-ukraine-enters-third-year/

[5] Details about the past attempt and the extension described in this paper can be found on the website: https://linked4resilience.eu/.

Uniform Resource Identifiers (URIs)6 can provide a unique reference to events, cities, and provinces and remove the ambiguity of spelling mistakes and variances. We added geolocation to enable visualisation of damage in Ukraine. Moreover, queries can be performed on the resulting integrated data, enabling use cases for various purposes. In this paper, we extend the use of linked data technology in previous work to damaged cultural properties and infrastructures, and demonstrate data enrichment, conversion, and integration using two sources: the list of damaged cultural properties by UNESCO [9] and that by [10]. We provide details of the implementation and publication of the data, as well as use cases. Our data is published in an online triplestore that is accessible through its SPARQL endpoints. This increases the findability, accessibility, and usability. Due to limited access to the data, we were unable to include all reported damages. Instead, our paper aims to demonstrate a workflow that outputs high-quality linked data with rich geo-related information to report damaged cultural properties and infrastructures for their resilience.

The paper is organised as follows. First, we present how we annotated and enriched the data from two sources, the UNESCO webpage [4] and the ScienceAtRisk webpage [10] in Section 2. Details about the conversion of our annotated data to linked data and their integration and publication are included in Section 3. Section 4 provides three use cases. Finally, the discussion, the conclusion, and future work can be found in Section 5.

2. Data Annotation and Enrichment

In this section, we explain how we extracted data from the webpage of UNESCO [9] and ScienceAtRisk [10]. The data from these two websites were collected on 30th April 2024. We used the corresponding Wikipedia page [7] as a reference for geolocation and additional information. Google Maps was used for the geolocation and alternative names in English and Ukrainian for the damaged properties. Our team includes seven Ukrainian-speaking volunteers for manual data annotation and validation. Next, we provide more details about data annotation and how information on damaged properties was validated and enriched for our dataset.

As of 10 April 2024, UNESCO has verified damage to 351 cultural properties since 24 February 2022. The reported properties include 129 religious properties, 157 buildings of historical and/or artistic interest, 31 museums, 19 monuments, 14 libraries, and one archive. As mentioned in Section 1, the representation of these damaged properties is simply a line of description grouped by their oblasts (regions). For example, "Historic building of the regional youth center, formerly the Shchors Cinema – Chernihiv". Despite that being cross-checked, this data remains difficult to use. There could also be confusion regarding cases such as "Holy Dormition Church – Mariupol" where Google Maps[7] returned 8 results with names in English and Ukrainian. As illustrated, this demands a significant amount of manual work for annotating, checking, and filtering of items listed on the UNESCO website with careful manual cross-checking of news articles, posts on social media, etc. Next, we explain the guidelines for manual annotation and enrichment, as well as how entries were filtered out. Damaged properties on the webpages of both UNESCO and ScienceAtRisk were processed the same way.

[6] For example, Kharkiv Region is assigned http://sws.geonames.org/550558 in GeoNames and its multilingual names can be retrieved using this URI.
[7] This search was performed on 28th July 2024 using Google Maps in English in Amsterdam.

Step 1: Names. Given the ambiguity in the names of damaged properties, our volunteers took advantage of Google Maps to validate each name. It was noticed that the names on Google Maps may be more accurate or updated than the names on the UNESCO page. For this reason, we include the names on Google Maps as an alternative label sdo:alternateName. So were their Ukrainian names. The names were also checked against the Wikipedia page.

Step 2: Types of damaged properties. The volunteers specified a unique type of damage for each property: religious sites, buildings of historical and/or artistic interest, museums, monuments, libraries, archives, and educational properties. In cases where the volunteers are uncertain about the exact type, “other" is added as the type.

Step 3: Geolocation. For each property, the latitude and longitude in decimal format retrieved from Google Maps were taken as the geolocation of the damaged properties. The locations were checked against that of DBpedia as described in Section 3. The volunteers may use the news articles found to make the location more specific in the case of large or multi-site properties.

Step 4: Wikipedia in English and Ukrainian. The volunteers took advantage of Wikipedia pages in Ukrainian and English during annotation. We included the corresponding URLs as sources. Information on Wikipedia was also used to help the volunteers to validate the type of damaged properties and infrastructures in Step 2.

Step 5: Including DBpedia entries for validation. The volunteers prepared data for the steps in Section 3 by manually finding the corresponding URI in DBpedia. The names in English and Ukrainian were checked on DBpedia's Faceted Search & Find service[8]. If found, the corresponding DBpedia URI is then added to the annotation. For example, “Church of the Resurrection of Christ" has a corresponding entry dbr:Cathedral_of_the_Resurrection_of_Christ_Kyiv[9]. Information on DBpedia could be used for validation.

Step 6: Media reports and news articles. We included additional resources as references such as news articles, social media, etc. Social media is not ruled out in this process despite issues with its accuracy (see Section 5 for discussion). These resources were included for further manual examination and replication of our results and future work.

Step 7: Additional information. Additionally, we include the date of the first damage (if possible to infer from media and reports), other media reports, the year of construction, funding information, and the website. As a primitive work, we do not study the case of multiple damages. Our annotation is based on an incomplete search, so the actual first date of damage may be different.

We apply the following criteria when annotating and enriching the data. For an entry, if there are multiple items with the same name, and the location information does not help with uniquely identifying the damaged object, then the entry is excluded. A property is excluded if no location information can be found or its location changed multiple times. For example, the "Institute of Bio-Stem Cell Rehabilitation, Ukrainian Association of Biobank," has changed its address multiple times and remains unclear to the volunteers based on information online. Thus it was excluded from integration.

---

[8] https://dbpedia.org/fct/

[9] https://dbpedia.org/describe/?url=http%3A%2F%2Fdbpedia.org%2Fresource%2FChurch_of_the_Ascension%2C_Lukianivka&sid=67097

For cases where there are alternative names. We included these alternative names to facilitate searching functionality. For example, for the entry "Old church (Tserkva Heorhiyivska) in Zavorychi village of Kyiv region," the church is also called "St. George's Church." We, therefore, added its alternative name. Media reports, news articles, and other forms of evidence of damaged cultural properties and infrastructures were recorded. For places where there can be multiple locations, the best estimate is provided based on media reports, news, etc. We rely on the media reports for estimation of the date of damage. In some cases, the volunteers can only specify up to a month or year. If no reliable information was reported, we kept it blank. Given that the source dataset has been verified by UNESCO, damaged objects are included even if there is no corresponding media report.

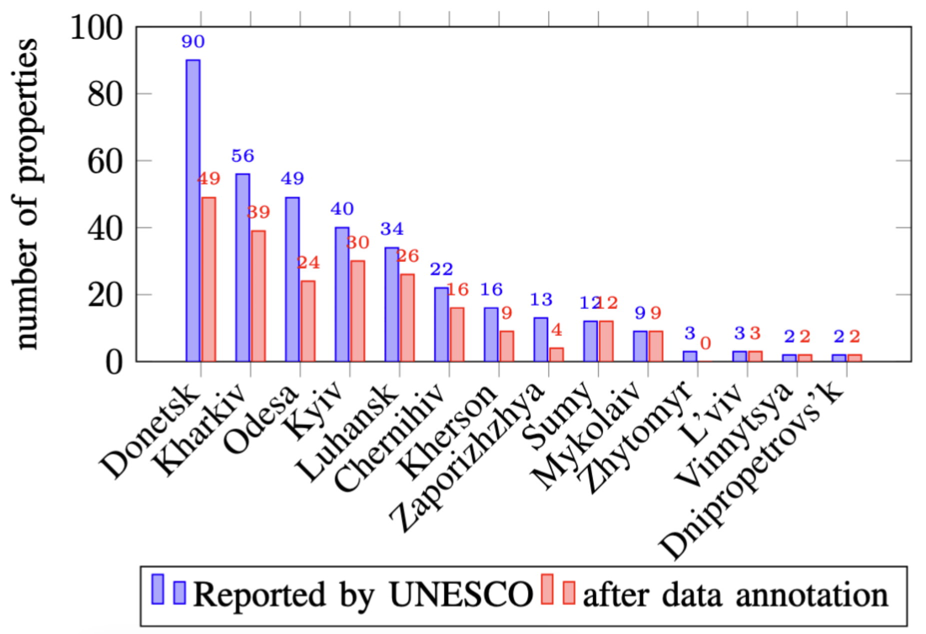

Figure 1: Damaged cultural properties in each region reported by UNESCO v.s. those included in our processed UNESCO dataset after filtering.

After the manual process, we include 211 (out of 351) damaged properties from UNESCO's list. Figure 1 illustrates the number of properties from the top 14 provinces with the most reported damage. After filtering, we include only those properties that meet our criteria (red bars). 40 out of 351 have corresponding Wikipedia pages in English, and 111 of them have Wikipedia pages in Ukrainian. Only 34 of them have corresponding DBpedia entries. We can find the year of construction or related information for 30 entries. 147 properties were at least associated with one media report. Together with Wikipedia pages in English and Ukrainian and other media reports, we can infer the date of the first damage for 225 properties. The process for ScienceAtRisk is similar, with one entry excluded[10].

[10] It was noticed by the volunteers that the location of "Institute of Bio-Stem Cell Rehabilitation, Ukrainian Association of Biobank" changed multiple times. We therefore decided to leave it out.

3. Data Conversion, Enrichment, and Publishing

We take a similar approach as our previous work [5]. For each damaged property, we assign an URI. We use *schema.org* for names and alternative names in English and Ukrainian. Each location is uniquely represented using OpenGIS as a *geo:wktLiteral*. In the cases where multiple damaging events exist, as a proof of concept of our approach, we include only the first date of damage. Moreover, we extend our vocabulary with properties such as *wasMentionedIn* for reference of media reports[11], and siteType for the type of damaged properties. We use *sdo:observationTime* for capturing damaged dates. Finally, we capture the Wikipedia pages used in English and Ukrainian by extending our vocabulary with *wikipediaEnglish* and *wikipediaUkrainian*. This results in three graphs: a graph with 2,910 statements on cultural properties and infrastructures and their related damaging media reports (UNESCO), a graph with 389 triples about that of damaged educational and research infrastructures in ScienceAtRisk, and a graph of 203 mapping statements (i.e. a linkset) between events and damaged objects between the two graphs as well as our previous work. Additionally, we have included an excerpt of the GeoNames dataset which we link to in our own data.

We implement data conversion and integration in TriplyDB, programatically with the TriplyETL package.[12] It uses a combination of client-side Javascript transformation functions and server-side SPARQL CONSTRUCT queries to transform the data. The code, raw data, annotation data, SPARQL queries, and backup files of the two sources of data (UNESCO's webpage and ScienceAtRisk) are available on Github.[13] Our data is published on TriplyBD[14] under the license CC BY-NC-SA 4.0. The SPARQL endpoint[15] can be used to retrieve information from the dataset. More description is available on the project website.[16] Next, we present three use cases to demonstrate the use of data in real-life scenarios.

4. Use cases

4.1. Use case 1: Visualisation of the geolocation of damaged cultural properties

A visualisation of damaged cultural properties and infrastructures provides an intuitive understanding of the location of damages. We plot our geo-annotated events on a map, making use of a SPARQL query, which produces data complying with TriplyDB's geo-renderer for SPARQL results. Figure 2[17] illustrates all damaged properties and infrastructures from the UNESCO and ScienceAtRisk graphs. It shows that the damage is concentrated along the frontlines and in the major cities. We take advantage of the interactive interface of TriplyDB: by clicking the points in the map, on the top right corner, it shows a small summary with the

---

[11] Our vocabulary uses the namespace https://linked4resilience.eu/vocab/.
[12] ETL standar for extract, transform, load. More about the package is available at https://docs.triply.cc/triply-etl/.
[13] https://github.com/LinkedData4Resilience/damaged_cultural_properties
[14] https://triplydb.com/linked4resilience/linked-4-resilience-2024
[15] The SPARQL endpoint using Virtuoso is available at https://api.triplydb.com/datasets/linked4resilience/linked-4-resilience-2024/sparql
[16] https://linked4resilience.eu/
[17] The figure is a screenshot of the visualisation available at https://triplydb.com/linked4resilience/-/queries/cultural-site-map-1

name and date of damage as well as a link to a full representation of the damaging events. Figure 2 shows that most of these damages are concentrated in eastern and southern Ukraine with some others around the Kyiv area. Based on our observation, the result is consistent with Figure 1. The corresponding SPARQL query[18] is below.

```
prefix geo: <http://www.opengis.net/ont/geosparql#>
prefix sdo: <https://schema.org/>
prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
select ?wkt ?wktLabel where {
  graph <https://linked4resilience.eu/graphs/cultural-site-damage-events> {
    ?culturalSite geo:asWKT ?wkt .
    { filter not exists {?culturalSite sdo:name ?name .}
      bind("(Name missing)" as ?name) .
    } union {?culturalSite sdo:name ?name .}
    {filter not exists {?culturalSite sdo:observationTime ?time .}
      bind("(unknown time)" as ?time) .
    } union { ?culturalSite sdo:observationTime ?time .}
    bind(
      strdt( concat(
        "<div><div><b>Name</b>: ",
        ?name, "</div><div><b>Damaged on</b>: ", ?time,"</div><a href=\"",
        "https://triplydb.com/linked4resilience/cultural-sites-
poc/browser?resource=",
        ENCODE_FOR_URI(str(?culturalSite)), "\">More info</a>",
        "</div>"), rdf:HTML
      ) as ?wktLabel)
  }
}
```

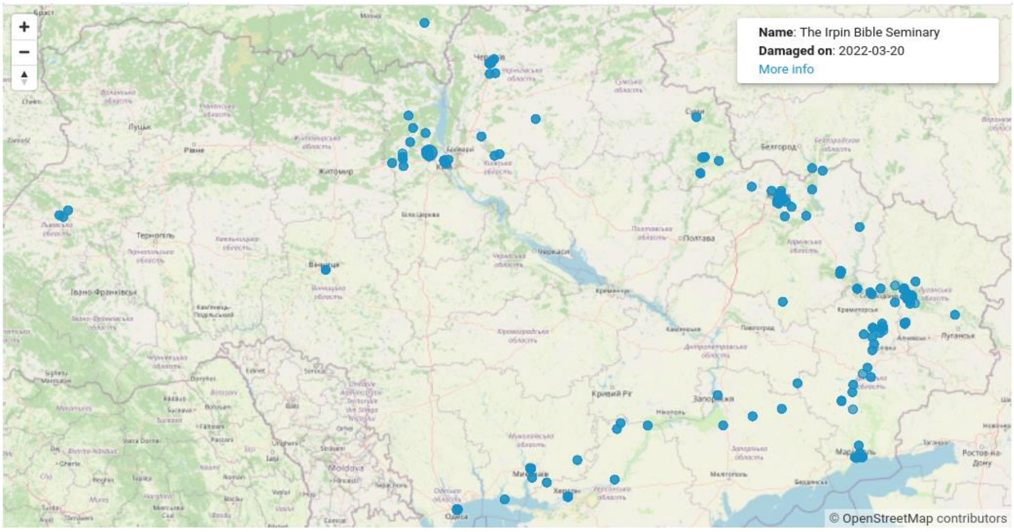

Figure 2: Use case 1: Visualisation of the geolocation of the damaged cultural properties in Ukraine.

[18] https://triplydb.com/linked4resilience/-/queries/cultural-site-map-1/1

4.2. Use case 2: Understanding the trend of damage

For each type of damage, the difference between months is an important indicator of the trend of targets as the war progresses. Figure 3 illustrates the trends for three types: museums, libraries, religious sites, and education and research (E&R) infrastructures. We can observe that the trends are similar except for March 2022, when significantly more damage was reported, especially that on religious sites.

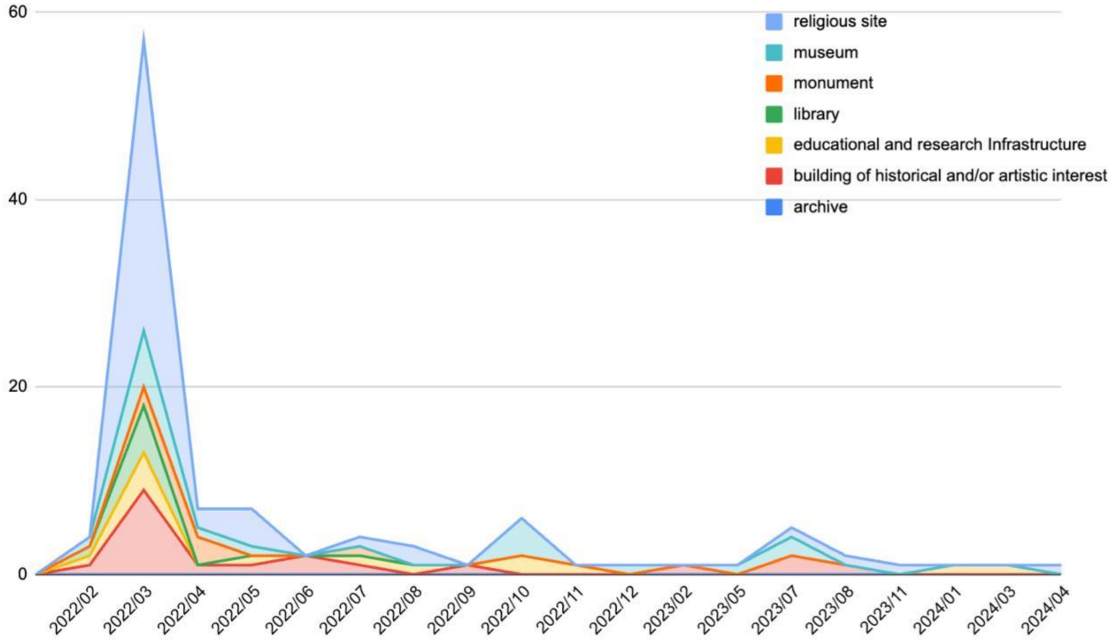

Figure 3: Trends of types of damaged properties as the war progresses in Ukraine in a stacked area chart. Months with where data do not present were omitted.

4.3. Use case 3: Associating with damaging events

Despite the fact that our previous work [5] is outdated, as a proof of concept, we attempt to provide a mapping. We found 193 damaging events for 58 damaged properties. We capture this by extending our vocabulary with *isCloseInLocationTo*. This mapping[19] could further enrich our data with details about the damage but requires some manual examination before use. However, validating the association of damaging events with the damaged sites and properties requires further manual examination.

5. Conclusion, and Future Work

This paper presents the annotation, enrichment, conversion, and integration of two datasets on cultural properties and infrastructures damaged during the war in Ukraine. The biggest two problems were the ambiguity in names and the lack of geolocation information. This shows the

[19] The mapping is hosted on TriplyDB in the same repository as others.

need to add semantics to the data for a unique and accurate representation. Among the 351 damaged cultural properties published by UNESCO, only 211 meet our criteria and were included. We demonstrated how to semi-automatically enrich it using multiple resources with the help of volunteers and improve its quality. The same approach was used for the ScienceAtRisk dataset. The dataset by UNESCO is being continuously extended. Thus, future work includes updating our dataset with new entries. Our approach can be extended to other datasets and the resulting integrated data can be aligned with other reported damage and relevant linked data. For example, our approach could also be extended to other domains such as occupied museum items. However, as the war progresses, the ownership could be arguable. Thus, the modelling could take advantage of some ployvocal solutions [11]. Given that our approach is heavy on manual work, future work includes exploring automation of our pipeline and validation with other external resources. Our dataset could also be used as a source to update the webpage corresponding to reported damages of cultural properties on Wikipedia [6]. For example, news articles and media reports that our volunteers found could be added as references to the corresponding damaged objects. The geolocation could be added and used to remove ambiguity. Names in Ukrainian (including their alternative spellings) could be added.

There are several limitations to our approach. Given that only 34 out of 351 have corresponding DBpedia entries, the reuse of information from DBpedia for validation and enrichment is limited. Our approach requires a significant amount of manual effort. Moreover, given the dynamic nature of the data during wartime, it is impossible to validate each data entry on-site. Thus, matching it with identified damages extracted using satellite images could be an alternative means of validation. We include reported damages from media, but not all of their trustworthiness has been verified. This makes the annotation highly dependent on the knowledge and interpretation of volunteers. In the current case, the involvement of Ukrainian volunteers, who understand the language and the situational context, was crucial.

Acknowledgements

The project received kind help from many volunteers: Tianyang Lu, Mariia Fomina (the University of Economics and Law, 'KROK'), and students from Lviv Polytechnic National University, including Taras Alfavitskyi, Roman Pukach, and Denys Hambotov. Their manual work is invaluable. We also greatly appreciate the help from Igor Potapov (University of Liverpool) for communication between the Linked4Resilience team and the Ukrainian universities. We thank Olexandr Konovalov (University of St Andrews) for coordination and looking for volunteers. Finally, we would like to thank Olga Polotska (National Research Foundation of Ukraine) for providing us information about ScienceAtRisk.

References

[1] Y. Kocak Bilgin and E. Hazarhun, "The Cultural Heritage Impact of The Russia-Ukrainian War," vol. 10, pp. 307-322, 2023.

[2] M. Sokil, Y. Syerov, and V. Boiko, "From Destruction to Digitization: Safeguarding Ukraine's Cultural and Archival Heritage in Wartime," 2024. DOI: 10.1007/978-3-031-59131-0_12.

[3] G. Sotnik, Y. Spinova, T. Sotnik, O. Polotska, O. Kysil, S. Krakovska, O. Diachuk, Y. Bezvershenko, O. Berezko, V. Beldavs, and A. Andrusevych, "How to help Ukrainian

scientists overcome Russia's invasion and advance sustainability," Proceedings of the National Academy of Sciences of the United States of America, vol. 120, no. 6, pp. e2219792120, 2023. DOI: 10.1073/pnas.2219792120.

[4] UNESCO, "Analysis of war damage to the Ukrainian science sector and its consequences," 2024. URL: https://unesdoc.unesco.org/ark:/48223/pf0000388803.

[5] M. Attar, S. Wang, R. Siebes, and E. Kultorp, "Converting and Enriching Geo-annotated Event Data: Integrating Information for Ukraine Resilience," in Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL '23), 2023. DOI: https://doi.org/10.1145/3589132.3625580.

[6] Wikipedia, "List of damaged cultural sites during the Russian invasion of Ukraine," 2022. URL: https://en.wikipedia.org/wiki/List_of_damaged_cultural_sites_during_the_Russian_invasion_of_Ukraine#References, accessed on 20 May, 2024.

[7] 7. Y. Aimaiti, C. Sanon, M. Koch, L. Baise, and B. Moaveni, "War Related Building Damage Assessment in Kyiv, Ukraine, Using Sentinel-1 Radar and Sentinel-2 Optical Images," Remote Sensing, vol. 14, article 6239, 2022. DOI: https://doi.org/10.3390/rs14246239

[8] S. Ivanysko, G. Kazakevych, and P. Shydlovskyi, "Cultural Heritage in the Russo-Ukrainian War: A Victim in the Conflict," Complutum, vol. 35, pp. 191-214, 2024. DOI: 10.5209/cmpl.95930.

[9] UNESCO, "Damaged cultural sites in Ukraine verified by UNESCO," 2024. URL: https://www.unesco.org/en/articles/damaged-cultural-sites-ukraine-verified-unesco.

[10] Assistance in reconstruction, "Scientific infrastructure damaged during the war," 2024. URL: https://scienceatrisk.org/infrastructures.

[11] S. B. A. Shoilee, V. D. Boer, and J. V. Ossenbruggen, "Polyvocal Knowledge Modelling for Ethnographic Heritage Object Provenance," in Proceedings of the 19th International Conference on Semantic Systems (SEMANTICS 2023): Knowledge Graphs: Semantics, Machine Learning, and Languages, vol. 56, pp. 127-143, 2023. DOI: https://doi.org/10.3233/SSW230010.