Connecting Archives, Connecting People Project Guidelines for Creating a Person Database

Date of Report: 2023-10-30

Report Version: Version 01

English

Prof. Charlotte Roueche https://orcid.org/0000-0002-3606-2049
Dr. Nurdan Atalan Çayırezmez https://orcid.org/0000-0003-3708-5055
Gonca Özger https://orcid.org/0000-0002-5245-4771
Burcu Akşahin https://orcid.org/0000-0001-7266-6686
Christopher Jacques Dolphin https://orcid.org/0009-0001-0461-462X
Hüseyin Hakan Özdemir https://orcid.org/0000-0002-4957-6305

For more information; you are invited to contact Dr. Nurdan Atalan Çayırezmez
repository.manager@biaa.ac.uk

Table of Contents

1. INTRODUCTION

This document has been created to give information about the person taxonomy creation process for the BIAA and BILNAS, and to present the results of the "Connecting Archives, Connecting People" (CACP) project (01/11/2022-31/10/2023), funded by the AHRC/UKRI Imagining Futures Project.

1.1. Project Team

Team members

- Prof. Charlotte Roueche (Emeritus Professor, King's College London)
- Dr. Nurdan Atalan Çayırezmez (Digital Repository Manager, BIAA)
- Orhun Uğur (Digital Archivist, BIAA[1])
- Gonca Özger (Digital Repository Assistant, BIAA)
- Burcu Akşahin (Research Scholar, BIAA; University of York)
- Felicity Crowe (Archivist, BILNAS)
- Christopher Jacques Dolphin (Intern, BIAA)
- Zehra İlke Yıldız (Intern, BIAA)
- Fatma Betül Göğüş (Intern, BIAA)
- Pelin Su Uzuncagil (Intern, BIAA)

Consultants:

- Dr. Valeria Vitale (Lecturer, University of Sheffield and Honorary Collections Officer, BILNAS)
- Hakan Özdemir (Chief Digital Officer, Üsküdar University)
- Rainer Simon (Independent Software Developer)
- Drupart UK

1.2. Acknowledgements

- Imagining Futures Project and Team
- BIRI colleagues, BSR, BSA, CBRL and Konstantina Georgiadou
- Dr. Amara Thornton from the Beyond Notability Project
- Dr. Gabriel Bodard from Standards for Networking Ancient Prosopographies: Data and Relations in Greco-Roman Names (SNAP:DRGN or SNAP) Project
- Başak Tosun, Zafer Batık; Wikimedia Community User Group Turkey
- King's Digital Lab, King's College London
- Administrative support: Laura Paterson (BIAA) and Gülgün Girdivan (BIAA)
- Interns and Volunteers

---

[1] Since August 2023 he has held the position of Digital Repository Associate at the University of Exeter

2. PERSON DATA IN THE ARCHIVES

i: In archives, 'creator information' typically refers to the organisation (corporate body), individual, or family responsible for an entire body of material (Introduction to Archival Authority Records). An Archivist defines the individuals, organisations and families according to their role or relationship with specific archival material. This plays a significant role in the creation or the acquisition of materials for an archive (e.g., depositor of a collection).

However, a 'creator' can also be responsible for the intellectual or artistic content of a single item, for example, the writer of a letter, the collector of a specimen or the painter of a portrait. The relationship of the person to a specific item, such as architect, photographer, data contributor or curator can be selected from the relators list: Relator Code and Term List -- Term Sequence: MARC 21 Source Codes (Network Development and MARC Standards Office, Library of Congress), e.g., **Writer of added text [wat]:** A person, family, or organisation contributing to a non-textual resource by providing text for the non-textual work (e.g., writing captions for photographs, descriptions of maps).

This more granular 'creator information' makes archives more accessible and creates opportunities for making archives more egalitarian. During the digitalisation metadata can be gathered from various sources. Bibliographic standards such as MARC21 and RDA provide information and guidelines for archivists for identifying person data and creator information, e.g., complete example of RDA

ii: This basic structure has been developed over many generations of archival experience. The development of digital tools, and digitised archives are now presenting new possibilities, and new challenges. Most fundamentally, the disappearance of the limitations on space within an analog archive require the archivist to consider new criteria as to what to include - and to make those criteria very clear. Equally importantly, the possibilities of Linked Open Data require each archivist to think carefully about which elements of metadata are their direct responsibility - since they can easily refer across to other sources of information. Our aim has been to work through the implications of this new environment.

iii: With the semantic web, archivists link data to increase access rates for archives and facilitate the use of the material. GLAM (Galleries, Libraries, Archives and Museums) institutions are working to create open-access archives. Archivists and digital data curators are applying international metadata standards and using vocabularies to create FAIR (Findable, Accessible, Interoperable, Reusable) archives. This is done by identifying unstructured[2] person information in existing metadata fields and matching them to authority records in LCNAF, VIAF and Wikidata. Creating such matches is a time-consuming task and requires significant human resources. Small archives and museums have a limited capacity to take on this work.

iv. The new environment would seem to make such tasks even more demanding. Our aim, therefore, has been to use our shared experience both to set out guidance for best practice, and also to identify the bare minimum required to start building shareable resources.

---

[2] Structure vs unstructured data for more information see: Kansa, E (2022) slides: ARF-Field-School-Eric-Kansa-Data-2022-08-08

v: The background: For cultural heritage organisations, person data can be grouped into two categories: Ancient and modern people. Several prosopography[3] projects have been undertaken to develop person identification of ancient and modern people.

Standards for Networking Ancient Prosopographies: Data and Relations in Greco-Roman Names (SNAP:DRGN or SNAP) project worked on ancient people data and created a cookbook SNAP:DRGN Cookbook v. 1.0 (GoogleDoc) for ancient people prosopographies. See also the discussion of Factoid prosopography, by John Bradley https://www.kcl.ac.uk/factoid-prosopography/about

For modern people, the Beyond Notability Project (Beyond Notability: Re-evaluating Women's Work in Archaeology, History and Heritage, 1870 – 1950) explores the histories of women active in archaeology, history and heritage as revealed in the archives of the Society of Antiquaries of London and the Royal Archaeological Institute. You can visit their Wiki page for the project database.

The resource of person data in biocultural archives has been explored in works like People are Essential to Linking Biodiversity Data, Connecting Te Papa Natural History Collectors and Identifiers and Guidelines for Using Wikidata to Mobilise Information about People in Collections: A Paleontology Perspective, which develop the ideas of linked data and Wiki tools.

The Connecting Archives Connecting People (CACP) Project is the latest of the international cultural heritage projects granted funding by the AHRC/UKRI Imagining Futures Project. This funding has enabled collaboration between the BIAA and BILNAS. Our collaboration takes place within the framework of discussions (funded by the British Academy) between the various BIRI (British International Research Institutes) on how to make their various archive catalogues more interoperable; in this project, we seek to increase person visibility within institution archives and to link this data using Wiki Tools; our wider partnerships help us to consider the wide range of challenges confronting different kinds of archives.

2.1. Project aims and objectives

Archivists have traditionally focussed on identifying and describing creators (e.g., the creator/donor/historian/video-recorder/photographer/squeeze-maker) and prominent people who play a role within the archive (e.g., correspondents). The new environment described above enables and requires us to undertake a new approach, providing identities for all the persons mentioned, represented or appearing within an archive. This may seem a daunting task, but it is an essential element of increasing equity within the archive. Our aim is to share our experience and to provide guidelines which may make the task simpler.

This document draws on case studies from the experiences of both the BIAA and BILNAS preparing person databases for the Institutes' archives before and during the CACP project, as well as other projects on ancient and modern prosopographies, who kindly shared their know-how for this project. Also, the CACP project collected research from and sought out

---

[3] *a study that identifies and relates a group of persons or characters within a particular historical or literary context* (Merriam-Webster. (n.d.). Prosopography. In Merriam-Webster.com dictionary. Retrieved October 29, 2023, from https://www.merriam-webster.com/dictionary/prosopography)

the advice of directors of previous visibility projects within archives. With these resources CACP was able to identify several common issues that arise when dealing with person data including; notability criteria, anonymous or unidentified people, as well as data protection and human resource scarcity. Thus, this project has approached these issues using: Wikidata notability criteria, General Data Protection Regulation (GDPR), CARE Principles (Collective Benefit, Authority to Control, Responsibility and Ethics), FAIR Principles among other resources described in the following guidelines.

On September 27th, 2023 an online workshop was held in English that brought GLAM professionals together to discuss these issues and develop solutions. The volume of unidentified or unacknowledged people in the BIAA and BILNAS archives led the project to focus its efforts on identifying modern people, specifically those who created archival items. The following guidelines offers insights taken from that workshop as well as other workshops, consultations, and person data projects. Alongside these insights, these guidelines describes the technical process of creating a person list as well as creating a linked person database for GLAM organisations.

3. GUIDELINES

GLAM organisations handle different types of materials e.g., epigraphical collections, biocultural collections, photographic collections and archaeological collections. Thus, creating person data for modern people within the archive has its challenges. E.g., For archaeological projects, team members may be listed in official documents and matching these team members with the archival materials requires a human resource. For example, a trench form, notebook or drawing might only include the initials of its creator, so an archivist must search through archaeological legacy data to reconcile the initials with a full name.

These guidelines will give information about creating a modern person list for different types of archival materials, tables and excel spreadsheets provided to help cultural heritage professionals. Versions of these guidelines are provided in English, Turkish and Arabic.

3.1. Data Gathering - Create a basic person list

Steps #1-#7: Building the basic structure

Archivists can gather person data from different resources. If a database already exists, person data can be exported from that database. If a database has not been made, it can be created from resources in the archive. Archival material, photographs, documents, drawings etc should be checked to find the initials and names to be added to the list.

The key which makes all this possible is to determine the minimum data necessary to build a person list. There are many desirable categories, but the essential elements are:

- A name or a describer (e.g., Michael Jones or Anonymous 1).
- One or more references to item(s) in the Archive.

- A Uniform Resource Identifier (URI)[4]- a unique identifier for that person in your archive.

This structure allows the lead archivist or anyone in the team or the community to add desirable items to a person list.

Step #1 Create Basic Person List

Anyone who created or worked on archival material for your institution should be added to the Basic Person List. Suggestions for the Basic Person List can be seen in Table 1.

Table 1: Basic Person List and description showing roles, dates, and initials

Initials	Alternative Names[5]	Default Name	Description[6]	Notes
DHF	David H. French, D. H. French	David Henry French	Tille Höyük Excavation Director (1978-1990)	

Table 2: (Explanation of Table 1: Basic Person List)

Initials	Person initials	DHF
Alternative Names[7]	Alternative names	David H. French, D. H. French etc.
Default Name	The form most common in the archive	David Henry French
Description	Write clear, succinct sentences to describe the role of the person for all archive or individual item (if you have only one item)	e.g., Excavation director between 1978-1990 e.g., Photographer between 1978-1990 and worked for this project e.g., write only archive name Tille Höyük archive
Notes	Any uncertainties or queries are noted in this column to be cleared up by colleagues with more experience and knowledge on the history of the archiving institute.	

---

[4] For more information about linked open data and URI https://www.w3.org/DesignIssues/LinkedData.html

[5] Other Designation Associated with the person (RDA)

[6] Field of activity of the person (RDA)

[7] It can be two separate columns

Note: ! *Who to add to this Basic Person List*?

> *You must decide your criteria for making your list. Notability criteria* are important; expect to start with entries for persons easily identified. For example, contributions to archaeological work or archival work are good criteria. In order to be more inclusive, persons appearing in archival material (e.g., worker in the trench) should also be described; but some of these will be hard to identify. If you are cataloguing an item e.g., a photograph, which records many individuals, you may need to add a note that there are some people yet to be identified; at a later stage researchers may add entries for those people. When the *Basic Person List* is complete, the notability criteria have been determined, and all relevant descriptions of the persons have been added you have made a *Basic Person List*.

Note: ! *If there isn't enough information about the person, their relationship to archival material will be enough in the description area to meet notability criteria.*

Step #2: Create unidentified/anonymous person list

In order to increase individual visibility within archives and institutions it is imperative to begin keeping records of *anonymous* (or as of yet unidentified) people in the archives. Such anonymous people could include unidentified people in photographs, a person mentioned in a letter, or people who were not acknowledged but contributed to an excavation. This task requires a human resource and will be time consuming. The table below (Table 3) offers an example of how an anonymous person list could appear.

Table 3: (Unidentified/Anonymous Person Table)

Unidentified/Anonymous Person No	Description	Default Name	Connection to an archival item or project
Unidentified Person 1	Pottery Washer, Ahmet		Tille Höyük Excavation-1979
Unidentified Person 2	Guard, Hüseyin		Kibyra Excavation-1990
Anonymous Person 1	First man on the left in item #123 (a photograph)		Photograph Item #123

Step #3 Normalize the data (Data Cleaning)

After gathering the data, items may need to be cleaned manually or automatically by reducing data anomalies and redundancies. E.g., Duplicate terms should be identified and then merged or deleted from previously created datasets and databases.

Step #4 Create Person List

These basic elements will create a Person List for the archives. In order to enhance this follow Steps #4 to #7 (see Table 4), These provide the framework for enriching the metadata and adding authority links to external data sources. Step #8 will help to add authority links (optional) manually or semi-manually.

Step #5 Add Alternative Names

Go through the database and archive to identify any alternative names of the person being identified which are used in the archive documents.

Step #6 Give unique local identifier to person

Give a unique local identifier to the person being identified. Establishing a local identifier (e.g., biaa.person.00049; slsgazetteer.person.18 etc.) for this person creates a persistent record. This will be a URI for the database.

Step #7 Linking Person record with your archival material

Enrich the metadata for the person being identified from the archive. It is essential to link the person to all of the archival information, material or projects they may have been a part of. The archival resource needs to be matched with the person e.g., letter written by David H. French on this date, letter sent from David H. French to Alan Stirling Hall. This will normally be the number, but it can be a physical item number, digital collection number, project number as well as a URI from a digital repository or database.

Table 4: (Person List Table) for excel template see Appendix 1

Local ID[8] /URI	**Essential**	the number given to entry in your list: it will be URI	biaa.person.00001, slsgazetteer.000090
Default Name (as in the archive)[9]	**Essential**	Full name of the person	David Henry French
Connection to an archival item or project[10]	**Essential**	This can be information for archival material, physical archival item no, digital collection item no, project no and date can be added	Tille Höyük Project or Letter dated on 1979-12-28
Initials[11]	Desirable	Initials or different versions of the name used in the Archive	DHF
Alternative Names[12]			David French, D. H. French, David H. French
Description[13]	Desirable	The role of person or any other information related to person	BIAA Director (1968-1993)
Authority Link (VIAF)[14]	Desirable	URL	https://id.loc.gov/authorities/names/nr90015065
Authority Link (LCNAH)[15]	Desirable	URL	https://viaf.org/viaf/39225703
Authority Link (Wikidata)[16]	Desirable	URL	https://www.wikidata.org/wiki/Q29517525
Authority Link (ORCID)[17]	Desirable	URL	
Note	Desirable	Any uncertainties or queries are noted in this column to be cleared up by colleagues with more experience and knowledge on the history of the archiving institute.	E.g., Needs further work

---

[8] Identifier for the Person

[9] Preferred Name of the Person (RDA)

[10] Source consulted (RDA)

[11] Other Designation Associated with the person (RDA)

[12] Variant Name for the Person (RDA)

[13] Field of activity of the person (RDA)

[14] Identifier for the Person (RDA)

[15] Identifier for the Person (RDA)

[16] Identifier for the Person

[17] Identifier for the Person

3.2 Linked Open Data

Step #8: Linked Open Data. In Steps #1-#7 you established the basic internal structure for your data. The next steps explain how to connect your data to other sources.

Step #8 Linking Person record with name authority records

Enriching and linking the person data with international Name Authority Records or any other authorities can be done manually or semi-manually. Check the Wikidata record to be redirected to different name authorities (see Identifiers area in Wikidata). There is no hierarchy between the authorities but the most common ones are the Library of Congress, VIAF and FAST. Be aware that the person being identified may only appear in one authority. Link as many authority records as you consider useful: if someone has a Wikidata entry which already includes many of these links you may not feel it necessary to add any more. These authorities are simply suggestions but feel free to add different authorities depending on the region, country and research area of your institution or organisation. Sometimes no or only one authority link will be found, thus your record will be the only authority. In order to verify the person, matching the person information with relevant archival items or projects is essential.

Linking person information with name records can be done manually or semi manually. Different tools are welcome, however, feel free to use whatever is compatible or standard with the process of your institution or organisation. E.g., Wikipedia and Wikidata Tools or OpenRefine tools.

Step #8-Option #1 Find authority links manually

● Collect all the related links from name authority records: visit each website:

[1] Library of Congress Name Authority File (LCNAF)

[2] Virtual International Authority File (VIAF)

[3] Wikidata

Search the name and surname and find the person information of the person being identified. Be careful, sometimes records can be found with the same name. Find the exact person information that matches with the person in your database (an archivist needs to decide the VIAF link to take and should add information to the note area of the excel sheet when necessary) and retrieve the relevant identifier before establishing a permalink (URL).

For example, searching David French or David Henry French on the VIAF website will bring different results (Figures 1 and 2).

Figure 1: Searching David French will get these results

Figure 2: David Henry French search results on VIAF[18]
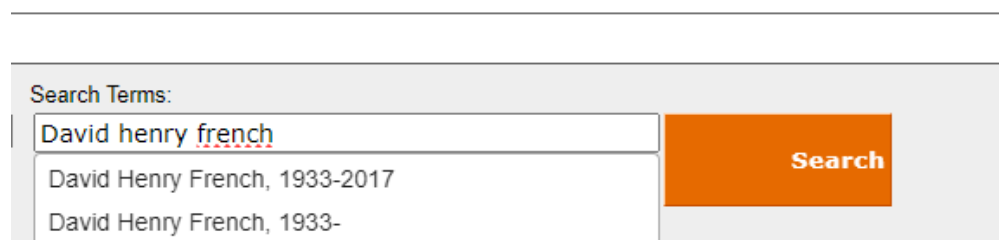
- Take the identifier and create a URL in the excel sheet by adding a URL (permalinks are important for the database) VIAF ID: 39225703 (Personal) Permalink: http://viaf.org/viaf/39225703

Step #8-Option #2 Using Wikipedia and Wikidata Tools

First create a Google account. Save the excel file (Person List) to Google Drive as a Google Spreadsheet. Google Spreadsheets enables automatic Wikidata link retrieval. Install "Wikipedia and Wikidata Tools[19]" extension for Google Spreadsheets.

Upload the *Person List* file to the google drive and convert it into a Google Spreadsheet. Open Google Spreadsheet and the "Default Name" column needs to be searched. After the Authority Link (Wikidata)[20] column create a new column named Wikidata ID. This column will help to compare the Wikidata ID's that may have been added manually. Or you may not have any Wikidata links. To initiate a search on Wikidata, use the following formula **WIKIDATASEARCH(B2)** to get the Wikidata ID's. In this formula, replace B2 with the cell containing the desired term. Below is Figure 3 demonstrating how to use this formula to search for individuals and obtain their Wikidata IDs.

---

[18] Usually VIAF merged to two records with same name. Sometimes they dont, you can send email to VIAF to suggest merging same items.

[19] For more information about you can visit **Wikipedia Tools for Google Spreadsheets**

[20] Identifier for the Person

Figure 3: Wikidata Search function in the Google Spreadsheet

After applying the formula Wikidata ID's will be produced. Then, write a formula (CONCATENATE) to get the Wikidata URL in a new column. See Figure 4 below.

Figure 4: Wikidata ID and Wikidata URL creation using Google Spreadsheet

Make sure the right person has been identified. Check the Wikidata URL and if it is not the person you are trying to identify then search Wikidata manually. When finished with the search and verification copy and paste the Wikidata ID's to the original Person List. Alternatively, download the file in an Excel format. Permalinks are essential for the database: e.g., Wikidata URL column.

Step #8-Option #3 Using Open Refine Tool

OpenRefine is an open-source tool designed for handling complex and messy data. It helps with tasks like data cleaning, format conversion, and the enhancement of data through integration with web services and external sources. It is recommended to read the documentation in OpenRefine https://openrefine.org/docs before creating a project.

One particularly useful feature is data reconciliation, which involves matching the datasets with external databases, such as Wikidata. Authority link information can be added automatically to the spreadsheet using OpenRefine. Read more here: https://openrefine.org/docs/manual/reconciling

- Reconciliation using OpenRefine Services: In order to reconcile the data with international authority links, it is important to note that some, but not all databases are

13

supported by OpenRefine's reconciliation services[21]: e.g., Wikidata and VIAF are supported. If the authorities have a reconciliation service in OpenRefine, follow the guidelines here: https://openrefine.org/docs/manual/reconciling and watch this video: https://www.youtube.com/watch?v=wfS1qTKFQoI

- The OpenRefine Reconciliation Services are not supported for some authorities e.g., LCNAF: In the OpenRefine reconciliation service for Wikidata, you have access to various statements and information, including Library of Congress authority ID. So, if the Library of Congress authority IDs are present in Wikidata, you can effectively use the Wikidata reconciliation service to reconcile and link data with LCNAF, even though LCNAF isn't directly supported by OpenRefine's reconciliation services. This makes Wikidata a valuable resource for achieving Library of Congress authority ID for reconciliation within your data management and integration efforts
Follow the guidelines here: https://github.com/mcarruthers/LCNAF-Named-Entity-Reconciliation

Reconciled data with authority links can be exported from OpenRefine.

If you want to publish your data online follow steps below

Step #9 onwards takes you to publication and beyond.

Step #9 Critical Data

It is crucial to analyse the data before publishing the person data online and sharing it with other portals and authorities.The Excel file needs to be checked before creating the final database for person data. Duplicate data needs to be deleted or merged.

Note: ! *If a person in your database is alive you need to consider getting consent to publish their information online. For legacy data, it is difficult to contact the person and get consent. Follow the archives legislation and privacy laws of your country as well as your institute policies (if any). Intellectual property and sensitivity issues must also be considered.*

Step #10 Publishing linked data

The data needs to be published online. See SNAP:DRGN Cookbook 1.0 "Scenario 1" section to do this. IT Service support may be useful as well. After this step; an URI for each person's record needs to be created. The Local Identifier will become an URI and Persistent Identifier. Displaying the person data publicly and for linking other authorities is important. Examples can be seen below (Figures 5-7):

Note: ! *Enriching and updating the person data with affiliation, occupation, university etc is time consuming and requires a human resource to maintain. The idea is to link the data with other international authorities to verify the record and enrich the metadata. So using VIAF and Wikidata for reaching publications and other information will reduce the time to create new information for each person. See Step #11.*

---

[21] you can find the supported reconciliation services from this link: https://reconciliation-api.github.io/testbench/#/).

An example from BILNAS can be seen Figure 5, https://slsgazetteer.org/person/18/ . The dates are the dates of attestations in the Archive: not essential, but desirable. Note that, the BIAA URI was added to James Crow's record.

Figure 5: BILNAS person example (James Crow)

Figure 6: BIAA person example (David Henry French)
https://digitalrepository.biaa.ac.uk/biaaperson00049

Figure 7: BSR person example (Frank Joseph Archer) https://archive.bsr.ac.uk/agents/people/143 = *Archer, Frank Joseph, 1912-1995 (British painter and engraver).*

Step #11 Publishing Data using Wikidata
More information can be found: https://paleo-data.github.io/guidelines-wikidata-people/

What to ask before sending the data to Wikidata:
 Is there an existing Wikidata item for the person?

 If Yes: Then you can publish your data in Wikidata following the steps detailed below.
 If No: Do you have another international authority eg. VIAF, LCNAH?

  If Yes: Then you can publish your data in wikidata following the steps detailed below.
  If No: You are the only source of this person. Then ask a question: Is the person alive?

   If Yes See Wikidata notability criteria https://www.wikidata.org/wiki/Wikidata:Notability
   If No: Provide source information to Wikidata as a statement adding publication, excavation directorship etc.

Step #11.1. Create Wikimedia account

Create a Wikimedia Account: Special:CreateAccount
How to create an account on Wikipedia: https://www.youtube.com/watch?v=1lYEZhUCJVA

With the Single User Login (SUL) system, this account can be used across all Wikimedia Foundation projects including Wikidata.

Step #11.2. Create Wikidata item
Go to Wikidata Page: https://www.wikidata.org/wiki/Wikidata:Main_Page

Step #11-Option #1: Create Wikidata item manually

Creating a Wikidata item manually is a structured process that begins by ensuring that an item doesn't already exist to represent the concept or entity you have in mind. Check to see if the Wikidata item already exists by manually searching the Wikidata page for the person to be identified.

If the item doesn't exist, begin the process of creating it.
Click on "Create a new Item" from the menu on the left side of the Wikidata interface.
For more information:
https://www.wikidata.org/wiki/Help:Items#Creating_a_new_item
https://www.youtube.com/watch?v=3CuUgJK4L78
https://paleo-data.github.io/guidelines-wikidata-people/

As mentioned earlier, creating Wikidata entries for individuals or GLAM institutions as properties becomes essential when there isn't a direct match with your dataset on Wikidata. However, manually creating Wikidata pages can be a resource-intensive and time-consuming task. Also, see Step 11.3 for getting Wikidata IDs automatically.

After creating basic information in the wikidata, editing is possible for items (see also https://paleo-data.github.io/guidelines-wikidata-people/). You can add statements, qualifiers and a references section. At this stage, add various properties and statements to enrich the data about the item, like the external id from a repository. See Step #11.3.

Step #11Option #2: Create wikidata items using Quick Statements Tool

Create wikidata item using template from the excel file and use the quick statement tool https://www.wikidata.org/wiki/Help:QuickStatements/en Go to the QuickStatements website (Quickstatements toolforge) and watch the video: How to use QuickStatements | The University of Edinburgh

QuickStatements allows for the mass creation and editing of items using a simple set of text commands. The process can be broken down into a few steps, which include accessing QuickStatements, preparing data, uploading and processing data, and reviewing and executing commands.

Note: ! It's important to note that to use QuickStatements with a Wikimedia account, the account needs to be autoconfirmed. This typically requires an account on Wikidata to be at least four days old and have made at least 50 edits.

Finally, a person list has been created and published online. The data may be linked to Wikidata as well.

Step #11-Option #3: Use OpenRefine to create Wikidata items

OpenRefine is not only a good application for cleaning messy data and reconciliation, but it is also a powerful tool for establishing Wikidata as a property. OpenRefine offers an efficient solution by enabling users to generate Wikidata pages in batches automatically. To utilise this feature, users need a Wikidata account. Additionally, updating specific items in Wikidata for a particular dataset can be managed through OpenRefine's 'Schema' template, a reliable service that also avoids creating duplicate entries. This template can be rerun as needed, and it will update only the items that haven't been uploaded yet. OpenRefine's user manual provides clear instructions on how to effectively use this tool (access the manual here: https://openrefine.org/). Watch this video: https://www.youtube.com/watch?v=wfS1qTKFQoI

Caution is essential during the data ingestion process via OpenRefine. Any mistakes made while structuring the data in OpenRefine may carry over into Wikidata entries. Therefore, when working with this tool, it's crucial to double-check the statements assigned as properties and their references.

Step #11-Option #4: Use API service

You may use API services to upload and download data. IT support may be needed for those in a repository. Remember that an institutional account for this API service is needed, because records from Wiki Services will be uploaded and downloaded.

Note: ! *Don't forget to create a verification mechanism in the system before taking data automatically from Wikidata.*

While the person records have been uploaded to Wikidata, you have the option to add your institute as an external identifier to Wikidata. In order to do this your institute needs to be a property in Wikidata.

Step #11.3. Become a property in wikidata
For more information: https://www.wikidata.org/wiki/Help:Properties

To be a property in Wikidata as an institute you have to fill out a property proposal form (https://www.wikidata.org/wiki/Template:Property_proposal) in Wikidata. You can also contact a Wikimedia User group from your country for more information.
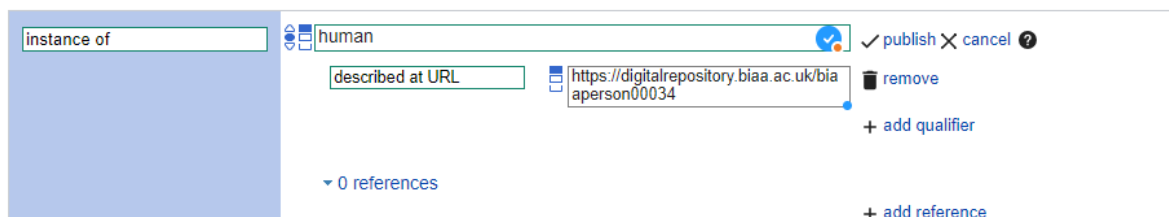
After approval you will be able to add your person Id as an Identifier to Wikidata see Figure 8.

Figure 8: BIAA Wikidata Property information https://www.wikidata.org/wiki/Property:P11940

If you don't want to be a property or you're not ready yet, you can still link your data with a wikidata record. You can use the statement "instance of" and choose "human" then add "described at URL" adding your website link (Figure 9).

Figure 9: Adding URL as statement

Step #11.4. Add external identifiers

Person Identifiers and stable URI information for person records can be found from several archives and in order to create multilinear interaction/connection between archives you can add external identifiers or provide your ID to other authorities. E.g., Wikidata. You can add other institutes' external ID's using statement function in Wikidata. For your institute to add your person URI to Wikidata you need to be property in the Wikidata.

Congratulations! You have completed a person list and published these person records online.

4. CHECKLIST

☐ Step #1: Create a Basic Person List

☐ Step #2: Create unidentified/anonymous person list (*Optional*)

☐ Step #3: Normalize the data (Data Cleaning/*Optional*)

☐ Step #4: Create Person List

☐ Step #5: Add Alternative Names (*Optional*)

☐ Step #6: Give unique local identifier to person

☐ Step #7: Linking Person record with your archival material

If you want to enrich your metadata with authority links follow steps below

☐ Step #8: Linking Person record with name authority records

    ☐ Option 1: Find authority links manually

    ☐ Option 2: Using Wikipedia and Wikidata Tools

    ☐ Option 3: Using OpenRefine Tool

If you want to publish your data online follow steps below

☐ Step #9: Critical Data

☐ Step #10: Publishing linked data (*Optional*)

☐ Step #11: Publishing Data using Wikidata (*Optional*)

    ☐ Step #11.1.: Create Wikimedia account

    ☐ Step #11. *Option 1:* Create Wikidata item manually

    ☐ Step #11. *Option 2:* Create Wikidata items using Quick Statements Tool

    ☐ Step #11. *Option 3:* Use OpenRefine to create Wikidata items

    ☐ Step #11. *Option 4*: Use API service

    ☐ Step #11.6.: Become a property in Wikidata

    ☐ Step #11.7.: Add external identifiers

5. WORKFLOWS

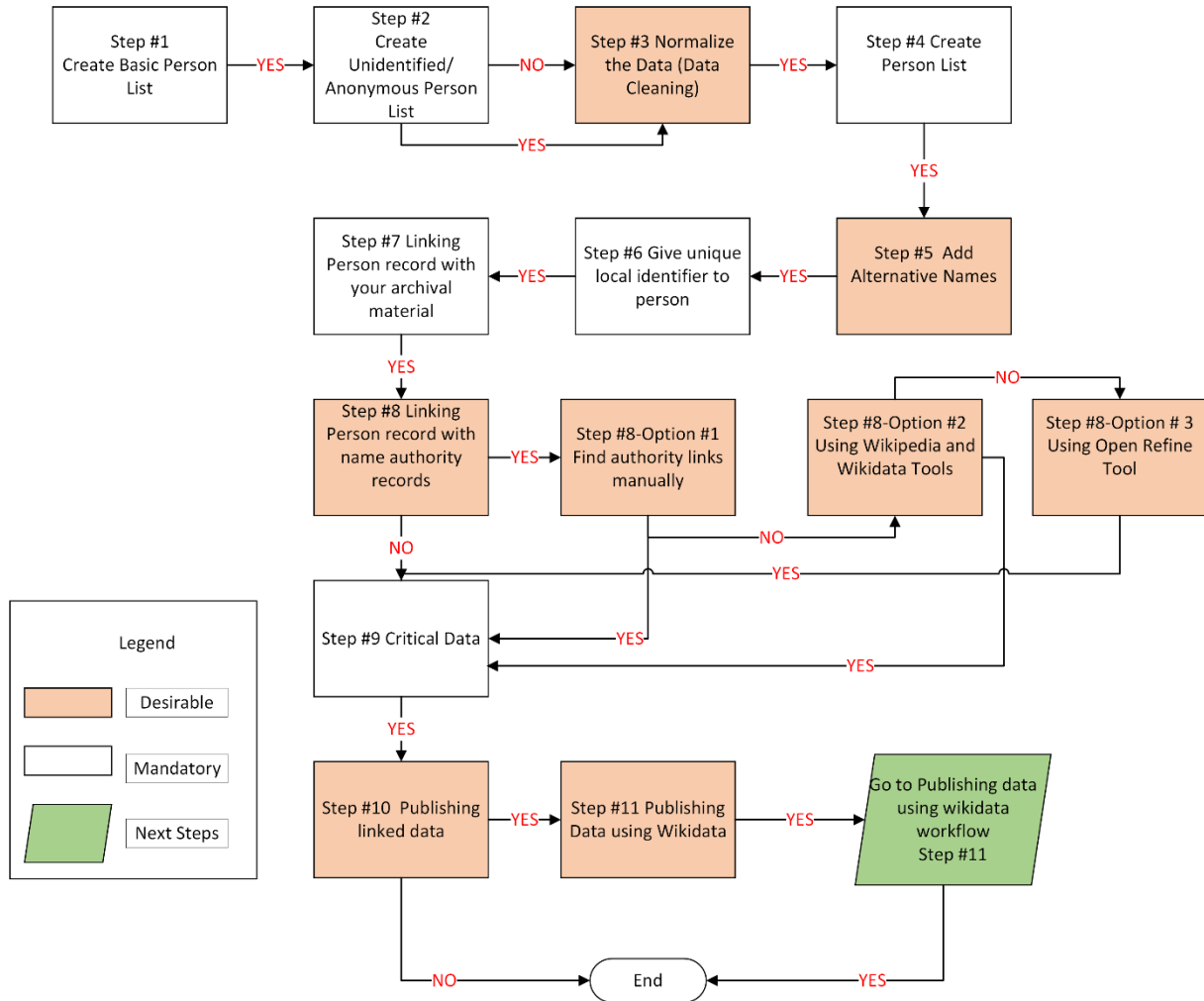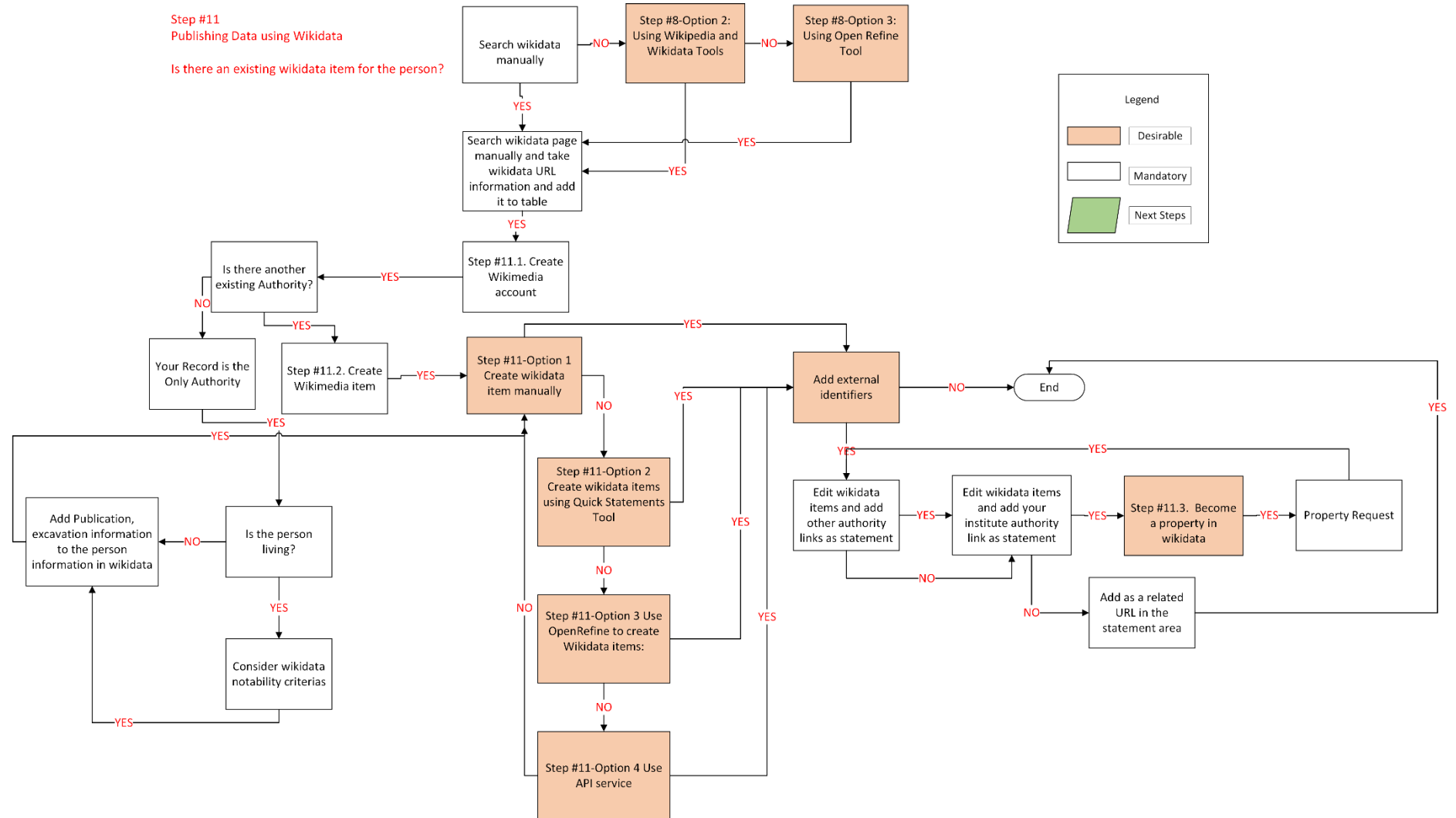*Figure 10: Creating Person List Workflow*

Figure 11: Publishing Person data information via wikidata

6. CONCLUSION

The value of Wikidata is enormous, particularly for enabling connections between different online archives, but it is still necessarily limited by notability criteria. Everyone in our archives will not meet those criteria - some may only appear once and not be identifiable. Although they may not meet the notability criteria for Wikidata, they do meet the criteria for local archives, and local communities. All such persons, however obscure, should be given an identifier.

The next steps for the CACP project will be to develop protocols and structures for the community sourcing of person data e.g. working with local communities. We hope that such communities might add information to our archives, and also consider building local person lists, with our support. ***Person visibility projects are resource-intensive and will require time as well as training to establish standards and practices amongst a diverse field of archival institutions.*** With this in mind, the BIAA and BILNAS view the Connecting Archives Connecting People (CACP) project as a part of an international process of increasing openness and visibility of persons in archives. Thus, CACP follows its peers and offers its hopes that its contribution to the equitizing process lays the groundwork for future visibility projects.

7. RESOURCES

This section has resource information related to linked data, person data, related projects etc.

Webpages:

Bodard, G., Cayless, H., Depauw, M., Isaksen, L., Lawrence, F., & Rahtz, S. (n.d.). *SNAP:DRGN Cookbook.* Retrieved October 31, 2023, from https://snapdrgn.net/about.html

Harloe, K., Thornton, A., Baker, J., Mahal, A. K., & Howard, S. (2021-2024). *Beyond 'Notability': Re-evaluating Women's Work in Archaeology, History and Heritage in Britain, 1870 – 1950.* Retrieved from https://beyond-notability.wikibase.cloud

IDEA. (2022). *IDEA: International (Digital) Dura-Europos Archive*. Retrieved October 31, 2023, from https://duraeuroposarchive.org/

Librarianship Studies & Information Technology. (2020, August 4). *Authority Record.* Librarianship Studies. Retrieved October 31, 2023, from https://www.librarianshipstudies.com/2016/02/authority-record.html

Library of Congress. (2022). *Authority Files for Cataloging Pictures: Common Choices*. Retrieved October 31, 2023, from https://www.loc.gov/rr/print/resource/228_authfile.html#name

MediaWiki. (2023, September 10). *Wikibase/DataModel/Primer.* Retrieved October 31, 2023, from https://www.mediawiki.org/wiki/Wikibase/DataModel/Primer

Meeus, S., Ariño, A. H., Bakken, T., Braun, P., Dillen, M., Endresen, D., Haston, E. M., Lowe, M., Meyke, E., Santos, J., von Mering, S., & Groom, Q. (2023). Who is Who in Natural History Collections?. Zenodo. https://doi.org/10.5281/zenodo.7781754

Book Chapters:

Keats-Rohan, K. (2007). Biography, Identity and Names: Understanding the Pursuit of the Individual in Prosopography. In *Prosopography Approaches and Applications: A Handbook* (pp. 139–181). Oxford: Occasional Publications UPR. Retrieved October 31, 2023, from www.researchgate.net/publication/304580535_Biography_Identity_and_Names_Understanding_the_Pursuit_of_the_Individual_in_Prosopography#fullTextFileContent

Verboven, K., Carlier, M., & Dumolyn, J. (2007). A Short Manual to the Art of Prosopography. In *Prosopography Approaches and Applications: A Handbook* (pp. 1, 36–69). Oxford: Occasional Publications UPR. Retrieved October 31, 2023, from prosopography.history.ox.ac.uk/images/01%20Verboven%20pdf.pdf

Articles & Blogs:

Bodard, G. (2021). Linked Open Data for Ancient Names and People. *ISAW Papers*. Retrieved October 31, 2023, from hdl.handle.net/2333.1/zs7h4fs8

Bodard, G., Cayless, H., Depauw, M., Isaksen, L., Lawrence, F., & Rahtz, S. (2017). Standards for Networking Ancient Person Data: Digital Approaches to Problems in Prosopographical Space. *Digital Classics Online,* 3(2). Retrieved October 31, 2023, from https://journals.ub.uni-heidelberg.de/index.php/dco/article/view/37975

Hawkins, A. (2021). Archives, Linked Data and the Digital Humanities: Increasing Access to Digitised and Born-Digital Archives via the Semantic Web. *Archival Science,* 22. Retrieved October 31, 2023, from link.springer.com/article/10.1007/s10502-021-09381-0

Neubert, J. (2017). Wikidata as Authority Linking Hub: Connecting RePEc and GND Researcher Identifiers. *ZWB.* Retrieved October 31, 2023, from zbw.eu/labs/en/blog/wikidata-as-authority-linking-hub-connecting-repec-and-gnd-researcher-identifiers

Ward, C. (2022). Excavating the Archive / Archiving the Excavation: Archival Processes and Contexts in Archaeology. *Advances in Archaeological Practice,* 10(2), 1–17. Retrieved October 31, 2023, from doi.org/10.1017/aap.2022.1

Wood, L. (2008). Who Are They? Unidentified Photographs in the Archives. *Ohio History Connection.* Retrieved October 31, 2023, from www.ohiohistory.org/who-are-they-unidentified-photographs-in-the-archives/

Adams, J., Brückner, H., & Naslund, C. (2019). Who Counts as a Notable Sociologist on Wikipedia? Gender, Race, and the "Professor Test". *Socius*, 5. https://doi.org/10.1177/2378023118823946

Martini, F. (2023). Notable enough? The questioning of women's biographies on Wikipedia. *Feminist Media Studies.* DOI: 10.1080/14680777.2023.2266585

Examples of Unidentified People in Archives:

Unidentified Nursing Sister. Dalhousie University Archives. Retrieved October 31, 2023, from https://findingaids.library.dal.ca/unidentified-nursing-sister

Unknown People. East Cleveland Image Archive. Retrieved October 31, 2023, from www.image-archive.org.uk/?cat=56

Videos:

Sunoikisis DC. (2023, June 15). SunoikisisDC SS 2023 Session 3. YouTube. Retrieved October 31, 2023, from https://www.youtube.com/watch?v=bCHNG9vbKJg&ab_channel=SunoikisisDC

Discussions:

Linked Open data related person data for several disciplines discussions are ongoing See: **https://linkedpasts.hcommons.org/symposium/**

Bodard, G. (2023) LOD People Activity, Home · DigiClass/LOD-People Wiki · GitHub And LOD People Projects
List Linked Person Data · DigiClass/LOD-People Wiki · GitHub

8. APPENDIXES

Appendix 1:

No	Document type	Title	Description	Author
1	Excel file	CACP Person List Template (in English)	biaa_PRJ240_RP1-1.xlsx	