Analysing Wikidata For Artistic Creation

Enrique Tabone

A dissertation submitted for the degree of
Master of Research (MRes) Digital Curation

School of Arts, Media & Creative Technology
University of Salford

May 2023

DECLARATION OF ORIGINALITY – CONDUCT OF ASSESSED WORK

ASSESSED WORK WHICH DOES NOT HAVE THIS FORM ATTACHED WILL NOT BE ACCEPTED

Research Degree Programme …………. Master of Research (MRes) Digital Curation ……………………………

Assessment Title ……………… Dissertation …………………………………………………………………………….

Title of the report ………. Analysing Wikidata For Artistic Creation ……………………………………..

………

………..

Family name of candidate …………………………TABONE…………………………………… (in BLOCK CAPITAL LETTERS)

Given name of candidate……………………………ENRIQUE……………………………….. (in BLOCK CAPITAL LETTERS)

ID number: …………… @00575087 ………………………

In presenting my Assessment 1: Interim Assessment / Assessment 2: Internal Evaluation (please delete as appropriate) I declare that I have read and understood the University Policy on Academic Misconduct (available at http://www.salford.ac.uk/about-us/corporate-information/governance/policies-and-procedures/browse-by-theme/2) and that:-

1. this work is my own
2. the work of others used in its completion has been duly acknowledged
3. I have been granted the appropriate level of Ethics Approval for my research

If you have a student support or reasonable adjustment plan, please indicate if you would like this to be

shared with your examiners ……………………………n/a……………………………………………………………..

Signature of candidate: ………………………………………………………………… Date: ……… 26/05/23 ………………

Please Note: If Ethics Approval is required you MUST submit the approval confirmation letter alongside your report.

In the event of any uncertainty, students should consult the members of their supervisory team.

Ethics Application: Panel Decision

ethics <ethics@salford.ac.uk>

To: Enrique Tabone <E.Tabone@edu.salford.ac.uk>
Cc: Toni Sant <A.Sant2@salford.ac.uk>

The Ethics Panel has reviewed your application: Visualizing Open Data for Artistic Practice
Application ID: 8434

The decision is: Application Approved.

If the Chair has provided comments, these are as follows:

You will no longer be able to edit your application in the system.

Link to the Ethics Application Tool: https://apps.powerapps.com/play/de0240e7-3d59-4974-849e-ba87d2541856?tenantId=65b52940-f4b6-41bd-833d-3033ecbcf6e1

Abstract

As an open database, Wikidata can be used freely for artistic exploration. It is not always easy to extract useful knowledge from the structured data held in Wikidata. SPARQL query language enables such extraction, to extract information and visualise data. Data visualization tools are largely used to generate charts, trees, and maps. There are also possibilities for artistic creation. This is the basis for data art.

My dataset for this project comes from Heritage Malta's collection of prehistoric female figurines. These are held at the National Museum of Archaeology in Valletta and the Ġgantija Archaeological Park in Gozo. Through Wikidata I have processed the raw data on this collection and applied FAIR data principles to it. This enabled the analyses of the visual descriptive data to ensure opportunities for open interpretations, including for feminist points of view that have frequently been disregarded, overlooked, or completely excluded. In 2020 a network of women involved in data science coalesced around the idea of data feminism. Their aim is to lead a cultural movement to promote equitable and gender-sensitive data systems for more inclusive decision making. This is also considered in the context of this project.

Complementing the visualisation of the dataset, I have also employed a technique of data sonification. This method involves the representation of aspects of the dataset in sound. Along with the written thesis, which provides a literature review for context, as well as a detailed account of my methodology, I have created a series of art objects to analyse the possibilities of working with this type of data in an artistic context. The practical work provides a method for researching new insights into the dataset, which otherwise remains a flat document containing basic information about very ancient objects that have frequently been overlooked for reasons that are also discussed in this thesis. Ultimately, this research project proposes ways to make otherwise forgotten archaeological museum objects more noticeable through artistic interpretation.

Table of Contents

LIST OF FIGURES

Chapter 1
INTRODUCTION

1.1   A general overview

This practice-based research project is designed to collect, analyse, and store data about a collection of prehistoric female figurines found in Malta so that it can be processed through data representation techniques towards artistic creation. As intended within this research project, the data representation aspect is both scientific and artistic in its approach. Scientific in the way it handles facts and figures associated with them and artistic in its ability to present creative experiences that enable diverse audiences to understand things they may overlook when presented to them otherwise.

Choosing to focus on this collection of prehistoric figurines from Malta from an artistic perspective presents a challenge in terms of established archaeological research on these objects. While Christina Biaggi (1994) describes some of the figures in this collection as lacking sexual characteristics while attributing the gendered term "Goddess" to them, in the present study the preferred terminology is that employed by Isabelle Vella Gregory (2005), which is significantly more measured in its deductions and labeling.

The overarching objective of this practice-based art research project relates to four other key concepts along with the choice of prehistoric female figurines in Malta. These are: open knowledge management through Wikidata, data visualization as a representational technique, data art as a form of creative expression, and a feminist perspective on the way the data is managed. The close interplay between these five elements, is carefully curated to ensure that whatever is produced as an outcome of this research project is original, in the sense that it relates to insights that have so far not been engaged systematically.

Data art provides opportunities to convey new experiences and knowledge through the analysis of datasets beyond what is afforded through scientific methods, such as conventional data visualisation or critical discourse analysis. This practice-based research project will show how metadata in a dataset pertaining to a specific collection held within a cultural institution, is analysed and curated to convey new experiences of the objects it documents. The decision to adapt all this in the context of an open-source database, which is Wikidata, immediately enables future reuse and further research opportunities with the data pertaining to the present research project. Wikidata is an open knowledge platform where anyone can contribute and improve

data in a method similar to the now ubiquitous Wikipedia; Wikidata is managed by the same organisation that administers Wikipedia, the Wikimedia Foundation.

A research method employed for this study involves collecting data from a museum collection and making it accessible through Wikidata in a way not previously available. This provides an opportunity to expose relatively hidden collections or even overlooked aspects of the better-known items from a particular collection. This is how Heritage Malta's collection of prehistoric female figurines and fragments, exhibited at the National Museum of Archaeology in Valletta and Ġgantija Archaeological Park in Gozo is approached here. Applying a feminist perspective when looking at this collection ensures a contemporary outlook, especially when this is combined with digital data analysis. The research process designed for this study thus involves digital curation and data science on the one hand, and artistic practice and creative expression on the other. As Michael Hohl puts it, "while scientists must strive for accuracy in their data visualisations, artists have more creative freedom." (2011, p.1041). This study therefore proposes to strike a delicate balance between scientific rigour and the dynamics of originality afforded by artistic practice. As such, this study provides an opportunity to explore ways to create a balance between accuracy and creativity for the purpose of sharing new sensorial experiences through this specific museum collection. This is intended to provide new insights to its audiences.

Data mining the collection, while aiming in find which aspects of the metadata connect these works together, provides an appropriate set of circumstances within which to share these findings through new artistic outputs. Understanding a collection through data about it can provide new ways to look at the objects within it. Within this practice-based research project, the dataset conveys information through art, by transforming the data into three experiences: video art, sculpture, and sound art.

1.2   Project aims and objectives

Why use data art to interpret a data set that can be visualized through conventional data visualization methods? This research question enables us to evaluate the art objects created through this research project, not as isolated art installations but potentially as models for the analysis and application of datasets to make new art objects. This study engages the chosen dataset through a creative process with the aim of discovering new perspectives on the

museum's collection of archaeological objects. This is presumed less likely to take place by simply observing, replicating or directly re-interpreting the objects individually.

As Escobar et al (2020) rightly point out, "many libraries, museums, and archives are currently exploring ways in which to publish their catalogues as Open Data" for a richer experience (p.1). This practice-based research project aims to enable new research opportunities not only through the same dataset developed and adapted for it but also on other datasets through the methodology designed for this study. There are several other opportunities to explore datasets, even within the same museums or other museums operated by the same umbrella organization, if not in the broader sector related to all this. This is not as bold an assertion when taken in the context of the way Wikidata works. This structured data platform enables connections between disparate collections brought together through links that are either not obvious or ones that are not necessarily easy to process without big data aggregation. One of Wikidata's more attractive features is that it is an open knowledge database that can be read by both humans and machines.

When data about museum collections is readily available, the public at large can discover things that remain hidden in museums, often even in reserve collections where only specialists can gain access through special permissions. Opening up knowledge about cultural heritage collections can enable new audiences to look for things that are not necessarily the ones presented prominently by institutional curators. This line of thinking, as presented here, requires crediting Nancy Duarte, who maintains that:

> As you explore data, you'll begin to formulate thoughts about what it's telling you. A point of view will emerge from the deep thinking. Sometimes, what you've uncovered will be blatantly self-evident to everyone and based 100 percent on the data. Sometimes, you will have to use a pinch of intuition and make some assumptions. Once you've taken a clear stance on what you've found, you're ready to construct a data point of view (2019, p.49).

In the present research project, what Duarte calls "a data point of view" is expressed primarily through an art installation intended as an opportunity for feminist data perspectives to be explored in ways that have not happened before.

1.3   Background

Working as a professional artist for over a decade, I came to data curation from my interest in art curation. I developed my first insights into data curation as a postgraduate student at the

University of Salford's Digital Curation Lab. My first practical experience with digital curation led me to SPARQL queries in Wikidata and subsequent data visualisations arose from that work. The RDF language quickly felt very familiar to me, probably because I am a visual learner and communicator. My initial experience with data analysis through data visualization came in 2020, when I worked on exploring gender gaps in the University of Salford's Art Collection. One of my aims with that project was to create better awareness of the works by women artists in that collection. I addressed this by collecting data and exploring ways to make that information accessible to a broader audience. This was instigated by the University of Salford's Art Collection drive to commission and collect more works by female and non-binary or queer artists.

This work on the art collection involved significant input on the reorganisation of a scant Excel spreadsheet containing a list of artists that formed part of this collection and little else. This was expanded into a structured data schema, which I then could start curating on Wikidata. This is how I explore the ways SPARQL enables data visualisation. This work helped the collection curators understand from where the artist whose work was in the university's collection come from geographically, enabling them to see the spread of locations represented in the collection, simply through the artists' place of birth. Other useful properties for which metadata was acquired include the date of birth of the respective artists, as well as the year when the work was acquired by the university. This immediately provided a way to see better where there were gaps in the timeline, such as when works by women artist were not collected. Mapping the data geographically and looking at timelines through the dataset instantly revealed insights into the collection that the curators previously did not have access to. I was particularly intrigued by the connections that were possible with data about the same artists in this collection in other collections across the UK or even beyond. This clearly spoke to one significant aspect of Digital Curation: ways of adding value to data.

That project enabled better awareness of women artists in the University of Salford Art Collection and in the process increased the visibility of women artists in the collection. In the process, the dataset enabled the identification of gaps through data visualisation and provided support for the acquisition of relevant works by women artists. Patterns and gaps tell stories about a collection. This dataset also provided a resource from which to help create or improve Wikipedia articles about women artists, as well as engage further with the collection more broadly, regardless of gender. Through the 2020 project on the university's art collection, I was

pleasantly surprised to see that other students on the Digital Curation course chose to use the methodology I had developed for this work on their own projects. I therefore saw that what I had created was not just a project that served itself but rather a method of inquiry that was easily adaptable for other contexts.

The university's art collection project helped me to investigate an interest in exploring the possibilities of how this type of data can be applied to a creative project. What are the possibilities of using digital curation skills and data visualization to make art, which either examines the power of data or uses data representations for aesthetic ends? This is how I arrived to the research question for the present study.

1.4   Structure

This document comprises five chapters, along with three appendices, and a list of references. The first of these five chapters is this introduction, which serves to give an overview of the research topic. The first section of this chapter gives a sense of the broader aspects of the selected topic, before details the general approach to the research questioin is addressed. The next section provides a glimpse into the thought process behind why this project was created, while also relating to the methods used to enact the methodology chosen for this research project. Taken together, these first two sections highlight the contribution that the present study aims to make within the field of data curation, especially through the domain of art making.

Chapter 2 consists of a review of the relevant literature for this research project. This is divided across the five key concepts on which it is built: Heritage Malta's collection of prehistoric female figurines, open knowledge through Wikidata, data visualization, data art, and a feminist perspective on data work. The combination of the project's areas of interest, as presented in the second chapter, provides a multi-dimensional approach to the subject matter of this research project.

The methodology for the project is discussed at length in chapter 3, with particular focus on the research methods chosen for this study. This research project stems from data analysis based on data that has been captured and processed specifically with a target goal of exploring artistic expression as a method of practice as research. A discussion of the data analysis is provided along with a consideration of Practice as Research (PaR) as a research method that aligns data

capture and processing with artistic expression, prior to addressing the ethical considerations for this research project.

The project outputs and research findings are the subject of chapter 4. Following an overview of the main practical outputs, a detailed account of the working process is presented. The main practical outputs from this research project are: an open dataset available through Wikidata, data representation work through data visualisation and data sonification tools, and the creation of original works of art. The working process is subdivided into nine steps that are generally aligned in clusters of three around each of the project's practice-based outputs. Selecting the topic and creating a hypothesis, leads on to the data gathering and processing phase. This eventually enables the data capture into Wikidata to create the dataset required for the completion of the present research project. The ensuing data analysis was also a three-step process, spanning from before the data visualisation work to the initial data processing, and on to the stage of refining the data to yield the desired design effects in the data visualisation outputs. Further processing of the dataset enabled the visualisation of the dataset towards the creation of data art from the same dataset, eventually leading to further artistic exploration via data sonification. Taken together, these outputs and findings constitute the new insights delivered through this practice-based research project.

Could data analysis of the prehistoric female figurines from Heritage Malta's collection help archaeologists do further research on this topic? Archaeologist Peter Ucko, the former director of the Institute of Archaeology at University College London established that it is "necessary to interpret each figurine, and each group of figurines, in their own right and to avoid generalizing from one figurine or group of figurines to all figurines in general" (1962, p.47). This, among other proclamations from established archaeologists, has helped create a respectful sensibility around the present research project towards the work of archaeologists. Still, in casual discussions about this research project with archaeologists, they insisted that artists have more freedom to interpret archaeological objects as they think they should be presented. Although this research project and its outputs may encourage re-use of the data and further research, one of the main aims of this research project is to make this collection more accessible through data visualisation, while raising the profile of this collection, as a feminist act.

Chapter 2
PROJECT CONTEXT: A LITERATURE REVIEW

To get a fuller picture of the overarching context of this practice-based art research project, it is essential to explore more closely five key concepts on which it is built. These are prehistoric female figurines in Malta, sharing open knowledge through Wikidata, data visualization, data art, and a feminist perspective. The ways in which each of these concepts overlaps with one or more of the others, relates to insights presented within the present project, in terms of exploring how systematic exploration can lead to further work beyond this study, in similar directions.

While it can be argued that this project is primarily about data art, it is essential to consider the theme or topic that has been chosen to inflect this work. The choice is not secondary because it informs the research methodology, as will be demonstrated later. This is why, the first context presented here relates directly to the theme or topic of the project rather than the working methods at the heart of the creative work.

2.1   Prehistoric female figurines in Malta

A significant number of prehistoric figurines, largely depicting humans and animals, were discovered and collected in the central Mediterranean island nation of Malta as early as the late nineteenth century and throughout the twentieth century (Pace, 1996). This followed the earlier identification of specific sites across the Maltese islands as prehistoric remains. The figurines were found in megalithic complexes and burial sites. The precise dating of figurines remains difficult, but many aspects of the iconography endure over the entire span of the period. Representations are made from clay, stone and bone, ranging from naturalism to complete abstraction, and vary from 1.2 cm to just over 3 m in height (Vella Gregory, 2005, p.336). All representations display a range of gestures and positions. They can be seated, standing or lying down. Arms and legs (where present) are placed in various positions. These objects are now held by Heritage Malta, which is the national agency responsible for guarding most of the country's tangible heritage and presenting it to the public. This research project is based on the two collections of female prehistoric figurines and fragments held at the National Museum of Archaeology in Valletta and the Ġgantija Archaeological Park in Gozo.

The style of Maltese anthropomorphic figurines has a particular form that is very distinct when compared with others from the Neolithic Mediterranean. Two particular forms stand out in Malta and Gozo. These are large figures, mostly seated, either naked breasted or fully naked pregnant figures. Nicholas Vella (2007) observes that some of the larger stone figures in this

collection look like they had removable heads, possibly interchangeable and for various possible ritual uses (p.63). It should be pointed out here that Vella's work is partly a commentary on a scholarly gathering that took place in Malta in 1985 to discuss the theme of Archaeology and Fertility Cult in the Ancient Mediterranean, the proceedings of which were published the following year.

Human beings tend to want to categorise items to organize things, objects, and ideas around them. Sometimes, this results in attempts to simplify our understanding of the world. A lack of categorisation can lead to greater levels of ambiguity in the interpretation of specific things, whether physical objects or abstract ideas. This seems to be the case for Heritage Malta's collection of prehistoric figurines and fragments, which have historically been overlooked unless simply categorised as statues representing ancient, predominantly obscure or unidentified deities. The gender of the "deity" is often also ambiguous.

The main themes arising from prehistoric female figurines that are relevant to the present practice-based art research project revolve around the artistic possibilities afforded by the historical interpretations of these objects. These interpretations are also embedded in the way these archaeological objects have been categorised, or not, by Heritage Malta. In the process, a significant shift in attention is emphasized; from "prehistoric deities" to representations of the female body.

2.2   Open knowledge through Wikidata

One way to approach any topic is to explore how it is represented in open knowledge contexts. This relates to open access and open data. In his book *Open Access*, Peter Suber (2012) provides a useful introduction to the basics of open access, explaining how it enables authors and readers of research to reach each other with greater ease than what is possible in closed or proprietary systems, wherein the prospect of financial profit often limits access to information and knowledge. The concepts of open knowledge and open data have also been discussed and articulated across different disciplines, such as information science and digital humanities. The key concepts are aptly examined through the theoretical foundation and practical implications of both open data and open knowledge. *Open Access and the Humanities* by Martin Paul Eve (2014) – which also contains a preface by Suber – provides a broad insight into this topic. Aspects considered by Eve include data governance, policies, ethics, open access to

knowledge, intellectual property rights, and the role of institutions in promoting openness and access to information.

However, for the most part, the literature that has been published over the past ten years or so rarely considers art making as a possible application for open data or open knowledge. One way this consideration will prove useful for the work presented in this research project is the way open data gives access to the contents of museum collections in ways that are not readily available otherwise. Open data can give access to knowledge that is otherwise held either within the museum itself (in labels, information panels, etc.) or in books or catalogues that need to be bought from the museum's gift shop or similar retail outlet. Many museums have attempted to move away from these former modes of knowledge sharing by providing information on their websites. However, making data available through open repositories such as Europeana[1] and Wikimedia Commons[2] gives audiences significantly greater access to the collections they're able to experience once inside the museum itself. The openness afforded by these open online platforms is further enhanced by the possibility for literally anyone to add metadata about each item held within the repository, or even expand the repository itself, subject to each set-up's own moderation procedures.

Open knowledge relates to data or information that "is free to access, use, modify, and share...subject, at most, to measures that preserve provenance and openness."[3] Over the past two decades this concept has been greatly popularised, primarily due to the success of Wikipedia and other open access projects associated directly with it. Wikimedia, the organisation behind Wikipedia and other open knowledge projects such as Wikidata, Wikimedia Commons, and Wiktionary, is just one of several advocates of open access to knowledge. Other proponents are likely also involved with the Open Knowledge Foundation, established in 2004 to promote open ideas.[4]

Open knowledge online platforms, particularly those operated by Wikimedia, provide an excellent solution to issues that frequently arise on the sustainability and longevity of digital media projects, especially in the cultural heritage sector. The FAIR data principles enable

---

[1] Available at http://www.europeana.eu accessed 22 May 2023.
[2] Available at http://commons.wikimedia.org accessed 22 May 2023.
[3] From the Open Knowledge Foundation's *Open Definition 2.1* available at https://opendefinition.org/od/2.1/en/ accessed on 22 May 2023.
[4] See https://okfn.org accessed 22 May 2023.

Wikimedia's platforms to focus on developing projects such as the ones mentioned earlier, safe in the knowledge that the data will be findable, accessible, interoperable, and reusable (Wilkinson et al, 2016). This is why Wikidata has been selected as the preferable choice for a data repository on which to build the present research project. It is open and accessible in the way it captures and links information through unique identifiers, making it findable by its application through a public domain CC-0 license. This also makes it easily reusable, particularly through the interoperable use of standard formats such as JSON and RDF.[5] Moreover, in the context of digital curation, it is useful to not that a Digital Preservation Technology Watch guidance note states that, "as digital preservation is an international activity, it is especially important that Wikidata is multilingual, and anyone can contribute to the project in their choice of language" (Thornton, 2021, p.2).

Wikidata is a collaboratively edited multilingual knowledge graph, operated through common source open data under a CC-zero license. In simple terms, Wikidata is an open database. Just like other Wikimedia projects, such as Wikipedia, anyone can add and edit data on it freely. It is designed to accommodate queries based on triples, such as how many prehistoric figurines in Malta are made of limestone or any other triple statement containing a subject, predicate, and object. These queries are run through SPARQL, which is a query language for databases in the Resource Description Framework (RDF) format.[6] All this will be discussed in greater detail in the ensuring chapter, detailing the working process for the present project. Significantly, "Wikidata feeds into the data used by digital assistants like Siri and Alexa," and "influences the knowledge panels of search engines like Google and Bing" (Strickler, 2021, pp. 301-302). Datasets become more useful and reusable when they are closely interlinked with other collections (cf. Escobar et al, 2019).

There's no denying that open knowledge works best by promoting transparency of data. The transparency ensures that key considerations for evaluating the quality and reliability of the data is more readily accessible. This is not to say that open knowledge systems do not possess their own problems, such as data completeness and reliability. It is thus essential to be careful when working with open datasets to ensure that the information contained within them, as well

---

[5] This is noticeable, for example, by Wikidata's inclusion in the FAIRsharing.org registry of knowledgebases and repositories of data and other digial assets, available at https://fairsharing.org accessed 22 May 2023.
[6] Wikidata has been available at http://www.wikidata.org since 2012.

as any additions or changes made to the data, are of a high quality, reliable, and relevant towards reflecting whatever the abstract information is meant to represent in material terms.

The data about prehistoric figurines and fragments in Malta shared on Wikidata can be queried in ways that enable new insights into this collection. As more data becomes available through Wikidata – for example, through references parsed from academic publications and exhibition catalogues – new insights into the ways objects relate to each other may also be discovered. In this way, for example, the collection in Gozo may be better cross-referenced with the one in Malta. Another example is the way knowledge gaps within a collection, including cataloguing errors, can be identified, enabling curators and collection managers to act on them as needed. One such gap in the prehistoric figurines collection involves the coordinates for the megalithic temples and other sites where the objects were originally found. Such data can enable a visual map from which to analyse possible patterns in terms of specific properties within the geographic area where the objects were found.

The power of Wikidata is particularly evident in its ability to connect items in the database together through properties that go beyond a specific concern such as a collection of artefacts within a museum. For instance, if an object in the museum inspires a poem, gets featured in a television documentary, or is used as part of a corporate logo, all these aspects can be linked together and queried through this dataset. As everything is connected, links to external databases are also possible, especially with persistent unique identifiers for each object in the collection. Through these connections, differences between different datasets can be investigated, while maintaining the potential to combine datasets and sources with existing knowledge stored in Wikidata. Thus, open data creates new possible understandings of the data sources. Such connections are more easily identified through data visualization.

2.3 Data visualization...and sonification

Structured data, available through platforms such as Wikidata, is useful in various ways. However, it is not always easy to extract useful knowledge from the data beyond the obvious, unless it is processed through data visualization tools. About fifty visualization tools have been

developed for Wikidata.[7] These data visualization tools are largely used to generate charts (e.g. radar charts, bubble charts, timelines, etc.), trees (e.g. clusters, radial trees, etc.), and maps. Most of them were built for specific projects, such as openArtBrowser, designed to show artworks by the same artist, movement, or motif,[8] or the Histropedia timeline generator, which uses data from Wikipedia and Wikidata to generate timelines from various historical perspectives.[9]

According to information designer David McCandless, "one of the potentials of [data] visualization is to bring the data down to earth" (cited in Edge, 2015). Visualizing data can help to provide a better understanding of the dataset. One example of this is the ability to show and see ways data items connect to each other via specific properties or how missing data on that same property within a dataset can show a gap in knowledge about the subject being visualized. This makes for a realistic insight into the possibility of discovering new information by connecting data. In other words, what the data is unable to communicate directly is also important and useful when attempting to understand datasets more holistically. McCandless' own work (2012) aptly demonstrates how even flat (two-dimensional) visualisations can portray information is simple creative ways. Through his body of work, we're able to trace a progressive development from earlier ideas relating to data visualisations, such as those presented a generation earlier by Edward R. Tufte, whose most design-oriented book, *Envisioning Information* (1990), is the last notable publication before the rise of computer-generated graphics, providing practical advice on how to communicate complex ideas through visual means. Tufte's earlier work remains in print through revised editions (cf. 2001) and is widely used in university courses on data visualisation.

Data represented in charts and diagrams have been around for a significantly long time. Aside from earning herself a place in history for revolutionising the nursing profession, Florence Nightingale also managed to change data visualisation forever in the nineteenth century. As a trained statistician, Nightingale intuitively understood the power of visual representation of information and used this as a tactic to convey statistics about her patients in exciting ways. She developed the polar area diagram, which is also known as the Nightingale rose diagram.

---

[7] For a current list of visualization tools on Wikidata see
https://www.wikidata.org/wiki/Wikidata:Tools/Visualize_data accessed 22 May 2023.
[8] See https://openartbrowser.org accessed on 22 May 2023.
[9] See http://histropedia.com/timeline/ accessed 22 May 2023.

This is still widely used in data visualisation to represent cyclic phenomena. Together with a team of experts, Nightingale worked to design a way to attract attention and engage readers through visual representation. Her technique involves the arrangement of information using line graphs, colour, scatterplots, and other graphic devices to emphasize particular aspects of the data. In her case, she presented her charts and diagrams with the intention of telling a story through her data. In the process, she showed contrasting data that compared army mortality with civilian mortality, visually presenting this data with a polar area diagram, which she designed specifically for this purpose (Small, 2001). Data visualisation historian Michael Friendly (2008) points out that Nightingale's work with information visualisation helped to reform the sanitary conditions of the army and effect changes in hospitals.

Another historical example, dating back to the nineteenth century is the well-known map of a cholera outbreak in London by the English physician John Snow. This is an early model of what is now commonly called data journalism. When Snow's visuals depicting the cholera outbreak were presented in contemporary newspapers, they often required about four paragraphs to explain to the readers how to interpret the visualisations. This was largely because the newspaper readers were not used to seeing data represented visually, even though it can easily be argued that this was done to enable a more immediate understanding of how the impact the disease was having on the city. In the mid-1800s, mapping data was still rather new and the myriad possibilities and combinations that have been developed since were not yet available to display and map data. Nevertheless, Snow's map offered a new way for data to be displayed, shared, and analysed. Tom Koch (2004) argues that, the urgency to uncover the cause of so many deaths drove Snow to experiment with new data visualization techniques to better understand the data. He recorded the location of each death in the city with simple markings on a street map of the district. Koch points out that, like Nightingale, Snow used a method of comparison towards making a more compelling case for sanitary reform. In his case, this involved cleaner water supplies to control the spread of cholera, which was revealed through visualising where the disease was having the greatest impact.

In thinking about ways to represent data, it is essential to move beyond visualisations. Another method that has been developed to communicate data is through sound. This is also known as data sonification. David Worrall provides a succinct overview of the topic, which he originally presented at the 24th International Conference on Auditory Display, which took place in Michigan in 2018. According to Worrall (2018), sound has been used to convey information

since before the modern era. He alludes to the first association of astronomy and music in the western culture of antiquity, better known as the "harmony of the spheres" through the philosopher Pythagoras, wherein the movement of celestial bodies are envisioned as a form of music. The sixteenth-century astronomer Johannes Kepler took this further with his cosmology, attributing the movements of the planets with intervals of sound. In Kepler's cosmology, the lowest sound is the minimum velocity of the orbit, creating the acoustics of the cosmos. Worrall highlights various models in what he calls a prehistory of sonification. One such model is by Italian Renaissance composer Franchino Gaffurio, who reproduced what he called "the perfect harmony of the heavens" through numeric data components to compose a celestial sonification. This harks back to Ptolemy's circular zodiac diagram to form arches and diameters creating musical proportions, which become a new musical experience. Numbers and symbols have long been associated with music in western society. Worral bring this closer to our time by discussing how the nineteenth and early twentieth century saw the inclusion of gestural and aesthetic forms in music through dance, a more creative representation of music. Dance is thus seen as a new artistic expression that can (re)create and express emotions. In 1957 the creation of *Illiac Suite* for string quartet was developed through a computer using a 'rule-base' system by Lejaren Hiller and Leonard Isaacson. This was the first example of what came to be called computer music. However, its significance in this context is that it was created through the sonification of data. The same tools used by Hiller and Isaacson where later adopted "in an attempt to obtain a better understanding or appreciation of relations in datasets of various sizes, dimensions and complexities – what is now called scientific sonification" (Worrall, p.180).

Information visualisation research in so-called GLAM institutions is on the rise after about three decades of digitised work and data collected from large cultural heritage collections. These are increasingly available on the web, making the web itself a large-scale repository of cultural assets, or at least a giant catalogue, and in the process making collections more accessible. This jointly raises interest in research with historical data from cultural sciences and humanities scholars (Windhager et al, 2019). Providing a focus on metadata which is stored privately within closed institutional databases, Windhager et al (2019) highlight the way open data can also benefit from added value through what Thomas Vanderwal termed folksonomy around twenty years ago.[10] The purpose of having basic open metadata is to remove barriers

---

[10] See Vanderwal's website on the coinage and definition of "folksonomy" at
https://www.vanderwal.net/folksonomy.html accessed 23 May 2023.

between experts and casual users who may identify objects within a cultural heritage collection differently than through formal methods. Experts encompass all people with professional or scientific interest in cultural heritage data, whereas casual users include a range of people from hobbyists to general museum visitors. In a system that affords the capture of contributions from casual users, the likelihood that the data is enhanced further is significantly greater than the dilution of whatever expert data is gathered and organised within the same dataset. In the practical work described in greater detail later on in this document, Wikidata is taken as one such open database that benefits greatly from the affordance of folksonomy.

According to Kerracher, Kennedy, and Chalmers (2014), offering multiple views through data visualisation helps "to maximise insight, balance the strengths and weaknesses of individual views, and avoid misinterpretation" (p.9). Their position engages serendipity as a method, even if not in a creative or artistic sense. Playful exploration is more clearly desirable in the perspective presented for what Thudt, Hinrichs, and Carpendale call 'the support of serendipitous discoveries through visualization' (2012).

Returning to the central topic of the creative project associated with the present research project, its necessary to observe that data representations – regardless of whether its data visualization or data sonification – enables us to look *into* a collection rather than *at* a collection. This is a data science point of view, where the aim is to discover relationships between items within the dataset while stitching them together through shared properties. Data humanist Giorgia Lupi (2017) claims that design serves the data, in the sense that it is the design you see first in any data visualization. With careful graphic design choices, the viewer should want to know more. This level of interpretation leads from data visualization into data art.

2.4   Data art

In a letter written by Pablo Picasso to Marius de Zayas in 1923, the famous artist claimed that:

> **"**We all know that Art is not truth. Art is a lie that makes us realize truth, at least the truth that is given us to understand. The artist must know the manner whereby to convince other of the truthfulness of his lies**"** (cited in Barr, 1939, p.10).

While he certainly did not have data art in mind when he wrote this, one hundred years ago, Picasso's often quoted statement works perfectly for the way works of art are made from data. Beyond artistic perspectives on truth, a description of data art that works well in this context is data driven art where the creator relies on the usage of a dataset. The artist must respect what the data shows and/or says while becoming an agent for messages embedded in the data itself. As an agent of the data, the artist attempts to convey emotions to the viewer, even if they potentially move far away from what is originally expressed, factually, by the data. As artist Laurie Frick proposes, in data art you "take your data back and turn it into something meaningful" (2015, n.p.). In this way, data art is not bound within the purview of artists but also of data scientists, data journalist, data visualizers, data analysts, and information illustrators, among others. Data art is not strictly a genre of art but rather a range of methods by which data is presented to diverse audiences. It should also be noted here that while most data art is also digital art, in some cases data art may extend into non-digital objects, such as the sculptures and musical scores created by Nathalie Miebach from weather datasets.[11]

Miebach came to data art when she was seeking a tactile experience for her studies in astronomy. In the process, she turned space data into visual art as a way to provide an experiential learning experience for herself and whoever encountered her art (Koren, 2015). From an artist's point of view, it is rather more dynamic to work with data that is constantly changing, with its potential surprises from whatever may happen in a natural way. There's something quite fresh and exciting in not knowing how the data will change and form the artwork over a specific period of time. Miebach's work is directly informed by weather data and temperature values she has selected to determine the way she weaves reeds as an integral part of her art. "Because these cycles change every day, you are working this grid in different ways," she says (cited in Koren, n.p.). Her datasets vary from floods, storms, and bush fires across different works of art that she has created, choosing to make her reed weavings both colourful and vibrant. While colour coding is a direct technique derived from data visualisation, sometimes it is evident that the artist intends her audience to lose themselves in the colours towards experiencing a flood or a fire more intensely than through the straight interpretation of scientific data. Subjectivity plays a significant part in all this, of course, as it takes aesthetic distance to elicit attraction to a sculpture derived from the visualisation of climate change or flooding that most likely involves unpleasant aspects and potentially catastrophic

---

[11] See this artist's website at https://www.nathaliemiebach.com accessed 23 May 2023.

consequences. With data art, the dataset can be overlooked by the audience towards an aesthetic experience rather than the actual meanings embedded in the data.

Artistic practice in data art relies on a dataset to convey emotions through specific readings of information held within the data. In this way, data art is not the mere visualization of the data so that it may be better understood but work that aims to elicit some sort of specific reaction from its audience. This applies regardless of whether the artist proposes aesthetic appreciation or an ethical position that involves some interplay with the data itself, if not a combination of both approaches. Either way, data visualization is a first step towards the creation of data art. Data art is data visualized or otherwise aesthetically represented to give a specific experience, or set of experiences, to its audiences. As outputs derived from data may be complex, it is the data artist's job to turn the visualizations, sounds, or whatever other form the outputs may take, into a work of art that can potentially convey specific emotions or messages to its audiences. This is where the artists technique becomes a crucial aspect for the best conceptual representation of the data they're working with and what they ultimately want to do with it. The material form selected by the artist is secondary to all this, and so data art can be presented in any material or form, digital or analogue, from print to sound, and sculpture or immersive environments that employ extended reality technology.

Collaborations between data scientists and artists are quite common in the creation of data art. For the purpose of illustration, I will cite two here. One by Jane Prophet dates back to the 2000s and involves interdisciplinary collaborations between mathematicians, scientists, and artists using rapid prototyping. This working method is contrasted with screen-based and online artworks in the artist's own writing about it (Prophet, 2010). Another considering works from around the same period suggests that physiological experience can contribute to the understanding of scientific data visualisation, especially through artistic approaches (Hohl, 2011). However, as with most cases, it is the scientists rather than the artists who tend to take the lead in writing up scholarly documents and/or theorising these collaborations.

There seems to be a gap in the literature detailing artists' experiences and perspectives in creating data art. Perhaps understandably, this is because artists tend to be visual thinkers who even in academic or scholarly contexts engage in practice as research rather than academic paper writing as the main output from their work. For example, David Bowen's *tele-present water* (2011) has been exhibited in Michigan and at the National Museum of Wroclaw in

Poland, however it has only been written about from an art reviewer's journalistic angle, along with other related works of data art by four other artists, exhibited separately: Janet Echelman, Rebecca Rutstein, Jill Pelto, and Nathalie Meibach (Delisle, 2016). Bowen's piece manifests data through an extraordinary mechanical sculpture within which he incorporates data into a live kinetic structure intended for audiences to "witness waves a world away." The data for this work comes from NOAA's National Data Buoy Center and through this sculptural work Bowen re-creates ocean waves movement from the data.[12]

Albeit in a non-scholarly context, for the most part, journalism is still a useful way to consider the works of artists creating data art. One such useful contribution is a Laura Sydell's report for National Public Radio in the United States on Laura Frick. The data Frick works ranges across a number of varied subjects such as identity, sleeping patterns, walking, or stress, to find new ways of looking at the human condition. Although there are clear patterns embedded in her works on these aspects of human life, she never proposes fixed meanings for her audiences. This is possibly because artists can choose to be deliberately ambiguous, if not multifaceted, to encourage their audiences to bring their own experiences to the work. Through such insights into the work of Frick, stylistic comparisons can also be drawn between data artists. For example, Sydell conjures up the idea of Piet Mondrian, who is best known for his abstractly bold linear colour compositions, drawing a spreadsheet, arguing that it would probably look more like the elaborately colourful work of Frick (Sydell, 2018). By extension, it would therefore look less like the data art works of Nathalie Meibach, even if it would likely be presented as a balanced and elegant pattern, it would still be rather rigid and flow less in its grid-like form.

Adam Nash (2014) considers the practical and theoretical considerations artists face when working with digital data. His focus is on the creation of real-time 3D audio-visual environments composed from digital data. His approach describes his own collaborative work called *Reproduction*, a networked immersive audio-visual work of art created with John McCormick in 2010. The artist draws on algorithmic information theory comparing this to the observational results he gathers from his own data art installation. Nash's insights are somewhat generic and relate more to the idea of what he terms virtual art than the type of data

---

[12] See the artist's website for images and a detailed description of this work:
https://www.dwbowen.com/telepresentwater accessed 23 May 2023.

art that relies on established datasets. This is particularly because of the performative nature inherent in *Reproduction*, which gathers data from its audience's motion and presence more centrally for its execution than it relies on a prescribed dataset that's of any specific significance outside the art installation itself.[13]

Data art can be seen to oversimplify complex datasets, sometimes emphasizing the artists' biases or even providing lack of adequate context (Bishop, 2023; Manovich, 2015; O'Neill, 2016). As outlined earlier, and by comparison, one of the significant critiques of open data involves data completeness, or rather the lack of completion. This, however, is not an issue that is unique to open data. It is still something that can easily trickle down into both data visualisation or sonification and data art. Uncollected data is a fairly common problem with many dataset. Artist Mimi Onuoha installation called *The Library of Missing Datasets* (2016) features a filing cabinet filled with folders precisely intended to highlight this matter.[14] This installation is also interactive as visitors can grab a file and read through it. The artist feels that this is data that should be collected and shared, mostly concerning social issues, as evidenced by a folder titled "Mobility in older adults with physical disability or cognitive impairments." Others include subjects such as "gender data gaps" and ones provoking the "who" question, as in "who would benefit from this data." Within the present study, such an approach is presented from a feminist perspective. This also stems, in part, from the description of collecting missing data such as for public monuments and sculptures found in the UK presented in Caroline Criado-Perez 2020 book *Invisible Women: Exposing Data Bias in a World Designed for Men*. This approach leads to data activism, which is primarily intended to help to collect data and instigate systematic change and/or monitoring of data gathering. The AirBeat community monitoring project that was created in 1998 is a good example of effecting change through data activism leading to the collection of better data regarding pollution through the action of citizens demanding cleaner air in a specific neighbourhood in Roxbury, Massachusetts.[15]

Generally speaking, there seems to be a tendency in data visualisation towards data art, where scientific rigour can sometimes be compromised for playful experiences. An example of this approach is evident in the Giorgia Lupi's *Dear Data* project developed with Stefanie Posavec

---

[13] See the artist's website for images and a detailed description of this work:
http://www.wildsystem.net/reproduction.html
[14] See the artist's website for images and a detailed description of this work: https://mimionuoha.com/the-library-of-missing-datasets accessed 23 May 2023.
[15] The AirBeat project website is available at https://www.airbeat.org accessed 23 May 2023.

wherein they sent weekly postcards to each other in a non-digital mode, consisting of a total of fifty-two hand-drawn data visualisations (Lupi & Posavec, 2016). These where mostly titbits of information collected each week about their lives, enabling them to get to know each other better. Subjects included whatever they noticed, emotions they felt or things they noted, such as every time they looked at a clock. Their project presents more of a personal experience and intimacy between two friends, even if it was eventually shared widely, particularly through the publication of a book about it all. In turn this become a story of friendship, as this type of data is the sort that friends tend to share on the phone or when they meet face to face.

Another significant point to make about data art, especially as it concerns the present research project, relates to artistic expression. Data art, as described here, must be developed from data, or at least somehow refer to data, even if intended only as the starting point in artistic practice. Data art, in this context, cannot rely solely on artistic expression. A work of art in this manner may provide an opportunity for its audiences to appreciate data on which it is based, even if not necessarily at first glance.

2.5   Feminist perspectives

Proposing a feminist perspective can be problematic. The feminist perspectives involved in the present research project is deliberately simplified while retaining an awareness that the very fact that a feminist perspective is required is a way to address a historical gap in the field. This is so because for centuries the female point of view has been too frequently relegated to an inferior position, or mere afterthought, in a broad range of circumstances. A feminist perspective can inflect the primary topic of this study, which involves the application of data art/data visualisation/data sonification from open data on a dataset concerning a collection of prehistoric figurines held by Heritage Malta.

In this practice-based research project on prehistoric female figurines, the data gathering work involves looking at Heritage Malta's collection to address potential knowledge gaps rather than to create gendered separation within the collection. While projects in open knowledge ecologies, such as those operated by Wikimedia, are most often focussed on addressing the gender gap (Hill & Shaw, 2013; Redi et al, 2020), the present project seeks to analyse visual descriptive data to ensure that this provides opportunities for open interpretations, including

from feminist points of view that have previously been disregarded, overlooked, or completely excluded.

Pace (1996), Vella (2007), and most other archaeological historians embrace established stereotypes about prehistoric women. Some archaeologists, such as Bonanno (2010) and Trump (2002), ignore potentially alternative narratives almost completely. This is even more evidently problematic when the work of most conventional archaeologists is read in the context of Cirotteau, Kerner, and Pincas' book *Lady Sapiens: Breaking Stereotypes About Prehistoric Women* (2022), which asserts the primacy of women in ancient societies by proposing that advanced civilization had sophisticated outlooks on art and power. This viewpoint needs to also be considered in the context of the palpable shift that has taken place in recent years in the art world's general narrative, particularly in the ways museum collections are being developed. This is clearly what drives the publication of factual data by the National Museum of Women in the Arts (NMWA) in Washington D.C. to highlight gender inequity in the art world.[16] This current shift is also reflected in books that aim to re-write histories of art. For example, Katy Hessel's book *The Story of Art Without Men* (2022) focuses on women artists who have often been overlooked or outrightly dismissed. In the process she provides a revised history of art from a feminist perspective. This book is also prominently available at major art museum giftshops in the United Kingdom, where this book was published in 2022. In an equitable society, gender data can be a powerful way to bring previously forgotten or untold stories to life, adding meaning to raise awareness about different perspectives on things that have been presented in a monolithic way for a very long time.

Nevertheless, feminist thought is not only about gender. In considering gender data, the intersectionality that comes along with it can also be acknowledged through data-driven decision making, which continues to shape our day-to-day realities. In 2020 a network of women involved in data science coalesced around the idea of data feminism. Their aim is to lead a cultural movement to promote equitable and gender-sensitive data systems for more inclusive decision making.[17] In their work with the Data Feminism Network, Catherine D'Ignazio and Lauren F. Klein point out that "intersectional feminists have keyed us into how

---

[16] The NMWA's factual data is available at https://nmwa.org/support/advocacy/get-facts/ accessed on 23 May 2023.
[17] See also the Data Feminism Network website available at https://www.datafeminismnetwork.org accessed 23 May 2023.

race, class, sexuality, ability, age, religion, geography, and more are factors that together influence each person's experience and opportunities in the world" (2020, p.14). The growing work of data feminism is therefore to uncover and reinforce existing inequalities around the world through data science. D'Ignazio and Klein suggest four ways to challenge power: collect, analyse, imagine, and teach. To collect missing data or institutional neglect. Analyse inequitable outcomes across groups and hold institutions accountable. Thy propose that we imagine our end point as co-liberation and empower newcomers to the field. This perspective guides the present research project, especially in terms of identifying gaps in Heritage Malta's collection, and data about it, which may lead to misinformation or incomplete approaches to the interpretation of objects displayed in the museum.

The main themes from the Data Feminism Network resonate well with the work of Caroline Criado-Perez (2020), who argues that data bias impacts most aspects of public life and decision-making processes. The key problem, according to Criado-Perez, is that women's experiences and perspectives are far too often overlooked or ignored. While acknowledging that some progress has been made, she points out that there is evidently still a clear need to create more equitable societies. Inclusive data collection and analysis can have a long-term impact. As with the work of D'Ignazio and Klein, it is easy to extend Cradio-Perez's arguments beyond gender and equality towards race, class, neurodiversity, and other aspects also in need of similar attention for better social representation. Feeding on many decades of feminist activism, Cradio-Perez proposes awareness through exposure as an essential step in redressing the bias, historical or otherwise, against women and marginalised others.

2.6   An initial conclusion on overlapping contexts

Data ethics and equality are at the heart of the context presented here. The perspectives provided through this literature review enables the project to have a number of significant touchstones from which specific explorations, particularly artistic research, can be conducted. For example, the Kerracher, Kennedy, and Chalmers (2014) article on the design space of temporal graph visualisations relates to the concept of serendipity, which as an artistic method provokes curiosity. It also supports playful exploration towards artistic creativity, even if this is more overtly presented by Thudt, Hinrichs, and Carpendale (2012). This is not too radical a departure because both the points of view expressed in these articles and the present project's

artistic goal are to explore ways to create more open, more diverse and engaging user experiences.

In his highly influential book *Art and Visual Perception*, Rudolf Arnheim proposes "the grasping of significant structural patterns" (1974, p.ix) as an artistic approach to reality. What are the well-balanced figurines in Heritage Malta's prehistoric collection depicting in their simplicity? Unless a cataloguing exercise has been performed thoroughly and each object documented appropriately, this observation can only be executed fully by spending time in the archaeology museum looking at each piece individually while annotating and comparing similar features. Looking at good high-resolution photographs – such as the macro photos taken by Daniel Cilia for Isabelle Vella Gregory's 2005 book – helps prepare you for the experience of seeing the actual object and alerting yourself for specific details. In Arnheim's words, "consistent shapes can create connections [...] as long as the given units indicate strongly enough a common pattern" (pp. 74-75). Photographs of the archaeological object give greater detail than drawings consist only of an outline of the object itself, with little other detail.

Another important element proposed by Arnheim is the concept of "overlapping" (pp. 106-107), which relates to the way hidden characteristics are rearranged and discovered. Psychological closure is an artistic freedom that enables the elimination of body parts from a figurative sculpture but still suggests that the missing bits are present without interfering with the visual concept. Or, in other words, "what is needed is shown and the unnecessary remains hidden" (p.107). There is no better way to describe these figurines in what they depict. The viewer of these ancient objects needs to remember that these are not pieces of modern art or realistic figurative sculpture that realistically aims to imitate human anatomy. These are prehistoric objects that were probably created to communicate ideas and concepts through a visual language, and possibly convey emotions too. Then again, to experience the objects closer to the way their creators intended them, audiences need to shed modern western ideas of art. Prehistoric art – if that is an appropriate term in this context – was evidently most likely created as a means to record and transmit information, in Arnheim's terms (p.129), rather then for ornamental purposes. Figurines and fragments full of imaginative forms, as Arnheim would have it is related to "the need to revive the old" and "reaffirms the truth" (p.141), if not an "attempt to reproduce an experience" (p.151). If the creator of the object has a good idea of human anatomy or, better still, if the creators are also women who can represent their own bodies in ways that a man could not easily see – such as looking down at their own torso –

what is seen in the process of creating these figurines is not the same as what you would see by beholding them as presented objects. The depiction of vulvic symbols and pregnancy take on a different meaning altogether as a woman's experience captured through material with "expressive qualities of a pattern of visual forces" (p.153).

The combination of the five areas of interest presented in this chapter provides a multi-dimensional approach to the subject matter of this research project, which aims to explore ways of reimagining prehistory female figurines from the Heritage Malta collection through data art from a feminist perspective.

Chapter 3
RESEARCH DESIGN: METHOLOLOGY AND METHODS

This chapter outlines the methodological aspects of relevance to this study, with particular attention to the empirical work underpinning it. After a brief discussion of the philosophical considerations for the research, a discussion on the various research design elements and empirical strategies follows. In outlining the study's methodology, this chapter provides insights into why this project matters, why it is necessary, and how these perspectives frame the choices made for the research design. Furthermore, the description of the research methods adopted for this project provides an understanding of the ways the methodology is enacted. Ultimately, the methodology and methods selected for this research project enable the findings presented in the next chapter.

3.1   Philosophical considerations for the research

A small number of philosophical considerations underpin the research design for this practice-based art project. A consideration for the prehistoric female figurines permeates the entire research project. In the previous chapter, the intricate connection between the archaeological discussions of these objects and aspects of feminist thought was discussed in a way that shows that they are perceived as two sides to the same coin, which is why they are presented as the first and last topics within the literature review. The choice to explore prehistoric female figurines from a contemporary art perspective, as a primary source for the creation of data art, necessitates an equally robust contemporary way of thinking. The choice of data feminism (D'Ignazio and Klein, 2020) and the revisionist shift in contemporary art history (Hessell, 2022) positions the work presented in this project in a way that is undeniably particular and original. The decision to focus on the collection of prehistoric female figurines held by Heritage Malta, enhances the choice of the philosophical underpinning because this approach is largely innovative and unexplored vis-a-vis this specific collection of archaeological artefacts.

3.2   Considering research methods

This project employs two research methods that are interlinked to yield the outputs described in the next chapter. The research presented here is built squarely on specific data that has been captured and analysed. In turn, this data is processed through visualisation and sonification tools towards the ultimate goal of exploring how artistic expression can yield original works of art based on essential elements within the dataset, all the while inflected by a feminist perspective as presented in the previous chapter.

An important point to make here is that while there were primarily two research methods outlined in the proposed research design for this project (data analysis and artistic expression) these have been brough together through a third method – that of practice as research – which is presented as a blending method for the other two. In research design terms, this is also a reflection on the implications of approaching the two primary methods as an application of mixed-methods (Creswell, 2014). In this way, the selected research methods combine qualitative and quantitative data collection with analytical techniques in a number of different ways. The dataset is captured into and queried from Wikidata in both a quantitative and a qualitative manner. Similarly, artistic practice derives qualitative outputs from quantitative data extracted from Wikidata.

According to John W. Creswell (2014), there are two main mixed-method strategies: sequential and concurrent. Sequential design takes either a qualitative or a quantitative starting point and then proceeds through an ordered sequence of data collection and analysis. The type of study that starts with quantitative data and later uses the collection and analysis of qualitative data to corroborate its findings reflects what Creswell calls a "sequential explanatory strategy." (p.215). By contrast, one that begins with qualitative data and then uses quantitative data reflects a "concurrent exploratory strategy" (p.216). The present study can thus be considered to apply a sequential exploratory strategy, even if the circulatory nature of artistic practice as presented here, implies some application of a concurrent exploratory strategy.

3.2.1   Data analysis: capture and processing – a digital curation method

The nature of this project necessitates data analysis. The capture and processing of this data is the digital curation aspect of this research project. It involves the capture and processing of metadata, in this case about each of the objects in Heritage Malta's collection of prehistoric female figurines. The analysis is subsequently performed primarily through data visualisation techniques. A data sonification exploration is also brought in later as part of the artistic experimentation stage; more on this in sections 3.2.2 and 3.2.3.

The data analysed in this practice-based research project needed to be captured and processed first. This is because at the start of the project there was no digital iteration of the metadata for Heritage Malta's collection of prehistoric female figurines available anywhere. The

organisation made the metadata held in their non-digital catalogue available for this project, understanding that this would be digitised as structured data through Wikidata and concurrently transcribed to an Excel spreadsheet.

The main purpose of applying data analysis as a research method in this project revolves around the possibility of identifying correlations. These findings are discussed in the next chapter. These relate to uncovering patters, trends, and insights that may not be immediately apparent as a clear goal in selecting this research method for this practice-based project. Data analysis can also help identify gaps, biases, or disparities within the data, contributing to a more nuanced understanding of the underlying issues (Tufte, 1990 and 2001; McCandless 2022).

Data analysis has been selected to be both the starting point and an integral part of the creative process within this research project. In some ways, this may seem inevitable with the creation of data art. In analysing datasets, you can uncover interesting relationships or even anomalies that can provide the initial elements needed for artistic interpretations. Through data analysis via visualisation techniques, the most relevant and compelling aspects in a dataset can be visually rendered and eventually applied towards the creation of data art.

Kieran Healy (2018) provides an excellent practical introduction to the creation of graphics from data. Going beyond the principles and practice of data visualisation, Healy discusses the methods that yield data representations through graphic designs that are faithful to the facts contained within the datasets on which they are based. Thus, following Healy, data visualisation is seen to be concerned with showing whatever may be within datasets, without necessarily analysing these visual representations towards specific ends other than to understand what information is contained within the data. By extension, it can be argued that Healy's approach to data visualisation attempts to expose any biases the creator of the visualisation may have. He certainly does not encourage the sort of subjective interpretation that leads to the creation of data art where, more often than not, the artist's own biases are the basis for the creation of the work.

3.2.2 Practice as research – explorations in data representation through visualisation and sonification

Experiential learning can be an integral part of a practice as research (PaR) method. For

example, Hohl (2011) proposes experiential learning through physiological experience as a method for comprehending scientific data through artistic approaches. It may also be argued that auto-didacticism can be another way where the practice itself becomes a research method. This can very much be the case for artists who teach themselves how a particular software package works, when their ultimate purpose is to develop their skill set to enable the creation of outputs derived from artistic expression. Although not intended as research outputs in the present study, I chose to model some of the Heritage Malta collection objects in clay. This was not an attempt to create replicas of the objects in the museum but rather to acquire experiential knowledge of the body forms through the physical materiality of clay. Vella Gregory points out that, "prehistoric women chose to embody in clay significant events in their lives: giving birth and possibly miscarriage. These representations also shed light on another aspect of women's lives – the transmission of knowledge and the creation of a community feeling among women as they spoke about and shared their experiences of their lived bodies" (2005, p.106). In this way, clay is used as an early knowledge sharing medium. It is also useful to remember that the Ishango Bone, which was discovered in 1960 in the Congo, is thought to be one of the earliest pieces of evidence of prehistoric data storage. It is believed that palaeolithic tribes may have used it to keep track of trading activities and supplies. While prehistoric societies are presumed to lack written texts, recording data to share knowledge is evidently a human concern that predates written histories as passed down through the centuries.

Fig. 1: Experiential learning exercise in clay (images below) developed as part of the present research project, intended to explore the shape and form of the prehistoric art object shown in the photo above by Daniel Cilia

PaR is a research method that integrates artistic practice with academic research. Seen this way it is the linking method between data analysis (as outlined in section 3.2.1) and artistic expression (discussed in the next subsection of this chapter). In the context of data visualisation and art derived from it, PaR is used here as a method to generate new insights on Heritage Malta's collection of prehistoric female figurines. This is performed in a way that is evidently transferable to other datasets and possibly even perspectives other than the feminist approach selected for the present project. This is particularly so because data feminism is only one of many positions that can inflect the analysis of data in a specific way.

The creation of audio-visual artefacts to be presented as an art installation is served directly by the means of inquiry and generation of insights associated with PaR as a research method. Robin Nelson's book published in 2013 is a very popular reference on PaR and his work is considered pioneering in this regard in the UK. While PaR is primarily associated with the performing arts, Nelson makes it clear that, from the perspective of PaR as a research method, this can be extended to artistic practice even beyond the performing arts. This may be seen as following on the work of Graeme Sullivan (2010) whose insights into visual arts are as much a touchstone on the subject as Nelson's, even if Sullivan's reference points are largely based on North American case studies and observations.

One useful point to gather from both Nelson and Graeme for this practice-based research project is the idea of the cyclical process of artistic experimentation, spanning from data analysis to visualisation creation, towards reflection and critical analysis, leading back to further data processing and creative expression. The essential quality of PaR is that it revolves around the artistic process and, more specifically, the knowledge generated through artistic practice. In turn, the artistic outputs can contribute to the research by providing new insights into ways complex relationships between research elements can be recontextualised to yield new understandings.

Furthermore, in the application of this research method to this practice-based research project, I came across the concept of data sonification, whereby datasets are represented though sounds rather than graphics. The possibilities that this offered for the creative process could not be overlooked as part of this project. The findings from this exploration are discussed in the next chapter.

3.2.3 Artistic expression – creation of data art and installation works for exhibition

Artistic expression is a broad research method, which relates very directly with practice as research (Bishop, 2023). It can be argued that artistic expression is only a research method within a framework of practice as research. Within this project, this assertion easily holds true. Nevertheless, there are some formal art techniques that cannot be overlooked.

The first of these is the way Arnheim (1974) discusses rules of visual grouping through what he calls the "principle of similarity" (p.67). In this way, shapes and patterns (including the physical material used to create an object) can be grouped by the way they resemble each other. By extension, Arnheim's principle of similarity can be related to a construct such as "similarity of location," even if this is probably better aligned with what he credits Gestalt psychologist Max Wertheimer calling the "rule of proximity or nearness" (p.67). The following illustration is designed to better illustrate this concept.
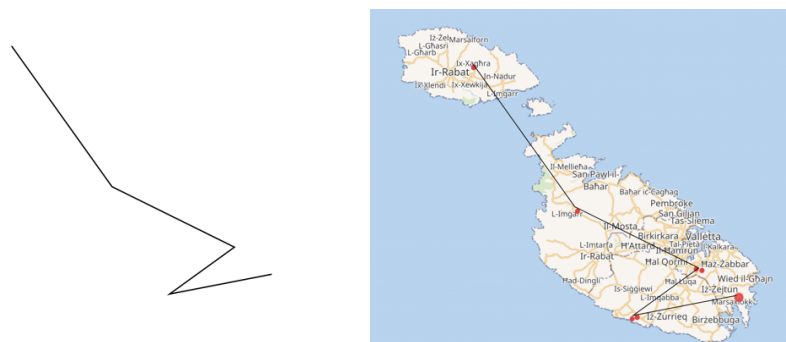
Fig. 2: A line drawing (on the left) is derived from joining location dots on a map showing the main sites associated with archaeological excavations across the Maltese islands that have yielded object held within Heritage Malta's collection of prehistoric female figurines

What appears to be a random angular line, if not an abstract shape, is based on a visual exercise creating a line joining the dots on a map of the main sites associated with discoveries of prehistoric female figurines in the Maltese islands. To be more precise, the map is generated through a Wikidata SPARQL query mapping the locations where Heritage Malta's prehistoric female figurines collection were found. The line is drawn following what Arnheim calls a "consistent shape" exercise (p.71), from a choice of possibilities with a spontaneous preference for an intrinsic nature of things to form a simple line structure. This follows from Arnheim's assertion that "the straight line is more identified than the irregular" (p.71).

The first work of art created through this research process is a sculpture made from transparent plexiglass rods, designed to be part of an art installation called *Naked Data*. The choice of transparent plexiglass in the sculpture created within the framework of this research project, can also be viewed in the context of the concept of transparency inherent in the ways open knowledge works through Wikidata and other Wikimedia platforms. Aside from the choice of material, the structural form of this sculpture builds on from Arnheim's observation that "the more consistent the shape of a unit, the more readily it will detach itself from the environment" (p.71). And thus at first sight the sculptural object within the *Naked Data* installation may seem less rigid and stiff than the data visualisation illustration on which the form is based, but detaches itself from readily.
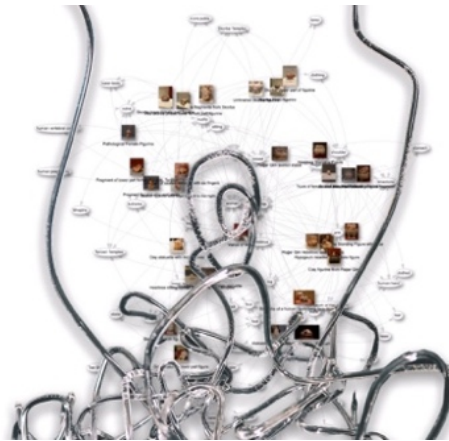
Fig. 3: Detail from *Naked Data* (2022) composed to show the contrast on stiffness and rigidity between data visualisation and data art intended to present the data in a less rigid mode

Another significant aspect within this method is known as "parallax", whereby seeking information is a playful exercise that is less goal oriented than regular data retrieval and/or analysis (Drucker, 2013). This is certainly a good way to describe the practice as research approach to artistic creation in data art. In a less goal-oriented data retrieval mode, such as an artistic/creative research-driven approach, the engagement is through curiosity and awareness. This is performed towards deriving an aesthetic pleasure, which is of the artist-researcher's choosing. In the process, the artist-researcher may develop skills "that are of interest to them but that they were unaware of prior to visiting" this line of inquiry (Chan, 2007). In this way it is also easy to see a connection with the notion of experiential learning, identified in the previous sub-section of this chapter as an integral part of the way practice as research is adopted as a research method in this study.

Data visualisation work extended into data art needs little additional narration or curatorial guidance for its intended audience. This is because data art by nature tends to tell its own story or stories, which are often open to interpretation by its audiences, possibly even through their own personal biases. Moreover, the data employed in data art is curated – even if only in the digital curation sense of working with and on the metadata as discussed in 3.2.1 above. Most data art is not concerned with haphazard data but rather with data that is curated to some degree or other. Even if elements of the data are randomly generated (which is not the case for the present research project), these are deliberately included in the data art experience intended by the artist through a more conventional curatorial process of inclusion and elimination.

This way of working is being proposed as a gateway towards exploring a specific topic – in this case, Heritage Malta's collection of prehistoric female figurines from a feminist perspective – rather than a fait accompli presented through an original work of art or series. The curatorial experience involved in this method provides the possibility to engage audiences through an innovative form of artistic expression. This presumably piques their interest enough to take a closer look at what they are experiencing as a work of art. One based on data derived from something presented rather matter of fact-like within an archaeological museum setting. The artist can attempt to narrate the collection in a way that is not too different from the prospect of seeking to inspire others to creatively engage with the material in artistic practice and design. Following Sanderhoff (2013), this involves telling a story by selecting and presenting objects from a collection in a purposeful manner, accompanying them with additional information, and even guiding audiences through and between exhibits (p.134). This is also because visual interfaces to digital cultural heritage collections tend to disrupt the pattern of search-centric interfaces afforded by the Internet, providing relatively static tableaus of objects and overviews, excluding the means for individual exploration, whether vertical (in depth), horizontal (relational), or tangential (random).

Artistic expression through a practice as research method, as presented here, can provide new insights for curators in terms of what they can do with their collections and the way they are displayed to their audiences. Rather than focusing only on screen displays within the museum, curators can create new experiences that can also be experienced off-site, providing opportunities for vertical, horizontal, or tangential explorations that are otherwise not possible within the static museum display experience. A specific example of this as it relates to the

present research project involves a way to potentially link up the objects displayed at the two museums more directly in ways that most visitors presently overlook, unless they happen to have visited both museums (which are on different islands) in close temporal proximity. There's also the possibility of presenting speculative data, even if this is an artistic rather than an archaeological approach. There is an opportunity for interesting engagement through the prehistoric dates when the objects were originally created or used, and the more recent (nineteenth- or twentieth-century) date when they were found as archaeological remains, even if these are not always known and there's often some uncertainty attached to such dating. In the present museum displays, such details are not always available to the on-site display viewers. Understandably, such data needs to be handled carefully. Kräutli and Boyd (2013) recommend that such uncertainties are not rendered invisible and that curators should explore ways to represent and visualise these uncertainties. This creates new opportunities for further research and discussion on the missing or uncertain data elements.

3.3   Technical considerations

For this particular project it is also essential to consider the technical aspects of the research work. The technical tools used for this project can be grouped into three: (a) Wikidata as a structured data platform along with data visualisation tools associated with it, (b) Adobe Illustrator, Photoshop, and video editing software used to create the video art component of the *Naked Data* installation, and (c) data sonification tools along with the Garage Band software for sound composition and sequencing.

Adobe Illustrator is the only application used for this project on which I had developed technical proficiency before embarking on the present research project. I was formally trained to use this software as an undergraduate student and developed advanced skills through my professional practice as an artist and designer in the ensuing years. The aspect of adopting an auto-didactic stance when it comes to technical packages that can enable research has already been discussed earlier, in 3.2.2. Although Wikidata is introduced to the PGCert students in Digital Curation at the University of Salford's Digital Curation Lab (a course I completed before embarking on this master's degree) the level of engagement with this open wiki that has developed the world's largest structured database is something I taught myself as part of my research training associated with the present practice-based project. As discussed earlier, serendipity is also a present element in the selected research method associated with the way

the data sonification piece was developed. In this case, I came across this technique as I was finalising the work on the *Naked Data* installation, before it was exhibited during the DRHA 2022 conference at Kingston University's Stanley Picker Gallery. The possibility of representing data through sound and not only visually, as per the techniques of data visualisation, is something that appeals to me greatly, especially in the context of data art through sonification.

Aside from all this, there's also the artistic technique that I have developed for working with plexiglass. This involves industrial processes that have been adapted for artistic ends. This is what enables the selected material, plexiglass, with which I have worked as an artist for about fifteen years before completing the present research project. This technique is worth mentioning here more for the purposes of comprehensive consideration of the practical elements of this research project than anything else. This work, which falls outside the scope of this research project, has been documented separately in a catalogue raisonnè of my entire body of work as a professional artist (cf. Sant, 2023).

3.4   Ethical considerations

This research project was submitted to the University of Salford for ethical consideration, which is an institutional requirement, even for projects that do not require approval regarding specific issues that have ethical implications. As such, this project was therefore submitted for consideration on ethical clearance rather than ethical approval.

The following text was submitted as the project description for the purposes of ethical clearance:

> Along with a conventional literature review on the key topics of this research project (i.e. open data, data art, critical perspectives) I will also be reviewing a number of notable works that employed data art created over the past years.
>
> The collection of prehistoric female figurines at the National Museum of Archaeology in Malta will provide the dataset for my research into open data. All open data will be based on information that the collection has already made public, through books, papers and through the Internet. No unpublished personal data will be gathered or used in the research project.
>
> The data will be processed through the Wikidata platform to enable further use under an open license for all structured data within this platform. Consent from the Museum will

be acquired on this ahead of any work taking place to ensure that no proprietary data issues arise later.

I will also be exploring a number of data visualisation tools that interface with Wikidata, as this is the main way that data from the project will be presented in the final submission.

The data sonification aspect of the project (described in sections 2.3 and 4.2.9) was added to the research design after ethical clearance was formally granted on 22 October 2022. Other than this, the project description submitted for formal ethical clearance squares up completely with the project as enacted.

The ethical clearance application requires researchers to indicate how the research question(s) outlined in the answer to "Project Description" will be addressed. The research methodology was described as follows:

The research methodology for this project involves processing open data from the National Museum of Archaeology in Malta through the Wikidata platform through visualization tools to enable the creation of an art object, which is designed to reveal information that is embedded within the data.

The methods applied will be primarily two: (1) open data processing, and (2) data art object making, based on the processed open data.

A key point that warrants emphasis here is the fact that all the data used for this project is either already openly available or will be made openly available by the art collection as part of this project, through a specific agreement with the art collection manager/s.

Once again this is in sync with the way the methodology of this research project is described in this chapter. There are two points to note here in terms of research development after the ethical clearance application was submitted and approved. The first point is that the second research method identified in this response does not exclude the development of data art through the process of data sonification, along with that of data visualisation. The other point is that while there were primarily two research methods outlined in the ethical approval application, these have been brough together through a third method – that of practice as research – which is presented as a bending method for the other two.

It is also worth noting, for the record, that the standard questions regarding the involvement of human participants, the potential of physical or psychological harm, distress or discomfort befalling any participant or researcher, the collection of personal data, activity involving work with human tissue, biological fluids or DNA samples, the use of animals and/or animal tissue, the use of hazardous substances, and non-human genetic resource, where all answered as

negative. The same applies for questions about potential ethical or political concerns, including potential for liability to the University from the research or ethical concerns about collaborator organisations.

3.5   Summary

This chapter started out by outlining philosophical and methodological issues relating to empirical research and, in particular, the main research methods employed for this study. This research project stems from data analysis based on data that has been captured and processed specifically for the ultimate goal of exploring artistic expression as an enactment of the PaR research method. A discussion of the data analysis was provided along with a consideration of PaR as a research method that aligns the data capture and processing with artistic expression, prior to addressing the ethics of this research. The results and findings of this study, or rather its outputs, are presented in the ensuing chapter.

Chapter 4
PROJECT OUTPUTS AND FINDINGS

An art-based research project needs to be evaluated through its outputs. The methodology outlined in the previous chapter and the methods applied in the process have yielded practical results – mainly the various outputs derived from the research project, detailed later on in this chapter – and findings that may be applicable to other projects. Thus, the project outputs are the results derived from the research work outlined thus far in this document.

4.1   The project outputs

Aside from this written document, this research project has been designed to yield a number of specific outputs. The first of these outputs is an open dataset available through Wikidata, detailing Heritage Malta's collection of prehistoric female figurines and fragments, which are permanently displayed at the National Museum of Archaeology in Valletta and the Ġgantija Archaeological Park in Gozo. Directly derived from this dataset are a series of data visualisations, presented in Appendix B at the end of this document, following examples of work on the development of the dataset. These formed the basis for the creative outputs, which is derived from the artistic practice associated with this project. This artstic output, entitled *Naked Data*, comprises an art installation made up of a sculptural object made from transparent plexiglass rods and a video art display, which includes a stop motion sequence signalling the creation of another art object, called *Ninfa*, a large inflatable female figure, which falls outside the scope of this research project. However, a further output has been attached to the *Ninfa* installation, and this consists of a sound recording produced through data sonification from the aforementioned dataset created specifically for this research project.

4.1.1   An open dataset available through Wikidata

The open dataset created specfically for this research project is now available in Wikidata, detailing the collection of prehistoric female figurines in the Heritage Malta collection held at the National Museum of Archaeology in Valletta and the Ġgantija Archaeological Park in Gozo. This dataset can be seen through a Wikidata query, either as a text-based table or organised visually through associated reference images.

For the purpose of reference, here is the SPAQRL query developed for query.wikidata.org:

```
SELECT DISTINCT ?item ?itemLabel ?itemDescription (SAMPLE(?image) AS
?image)  {
  ?item wdt:P17 wd:Q233;
    wdt:P31 wd:Q10855061.
  OPTIONAL { ?item wdt:P18 ?image. }
  OPTIONAL {
    ?sitelink schema:about ?item;
      schema:isPartOf <https://en.wikipedia.org/>.
  }
  SERVICE wikibase:label { bd:serviceParam wikibase:language
"[AUTO_LANGUAGE],en,fr,es,de,ru,it,nl,ja,zh,pl,cs". }
}
GROUP BY ?item ?itemLabel ?itemDescription
```

Adding the following constructor before the query yields the text-based table shown below:

```
#defaultView:Table
```

item	itemLabel	itemDescription	image
🔍 wd:Q1135588 73	Stone block with figure carving	Stone block with figure carving, Tas-Silġ	🖼 commons:National Museum of Archaeology Malta- Inventory number 14405.jpg
🔍 wd:Q1160359 65	Headless seated figurine crossed hands	Headless seated figurine with crossed hands, Xagħra Circle	🖼 commons:Headless seated figurine crossed hands, Xagħra Circle.jpg
🔍 wd:Q1160362 75	Red terracotta headless seated figurine, hands on thighs	Red terracotta headless seated figurine, Xagħra Circle	🖼 commons:Red terracotta headless seated figurine, Xagħra Circle.jpg
🔍 wd:Q1160369 55	Xagħra Stone Circle seated figurine with hands on stomach	Xagħra Circle seated figurine with hands on stomach	🖼 commons:Xagħra Stone Circle seated figurine with hands on the stomach.jpg
🔍 wd:Q1134128 42	Seated statue with raised neckline	Prehistory seated sculpture with raised neckline discovered at Ħaġar Qim in Malta	🖼 commons:Ħaġar Qim seated figurine with raised neckline.jpg
🔍 wd:Q1135594 25	Clay figurine from Ħaġar Qim	Clay female figurine from Ħaġar Qim temples	🖼 commons:National Museum of Archaeology Malta- Inventory number 21246.jpg
🔍 wd:Q1135639 40	Statuette of a human figure lying face down on bed	Statuette of a human figure lying face down on bed, National Museum of Archaeology Malta	🖼 commons:National Museum of Archaeology Malta-face down Clay figurine on a couch-Hal Saflieni-21247.jpg
🔍 wd:Q1135640 34	Low part of seated statuette with six fingers	Low part of seated statuette with six fingers, Ħaġar Qim, Malta	🖼 commons:National Museum of Archaeology Malta-21247.jpg

Fig: 4. Partial view of the dataset ViewTable query in Wikidata.
This is available to view in full at https://w.wiki/6kE9

Similarly, this constructor generates an image grid from the same dataset:
`#defaultView:ImageGrid`
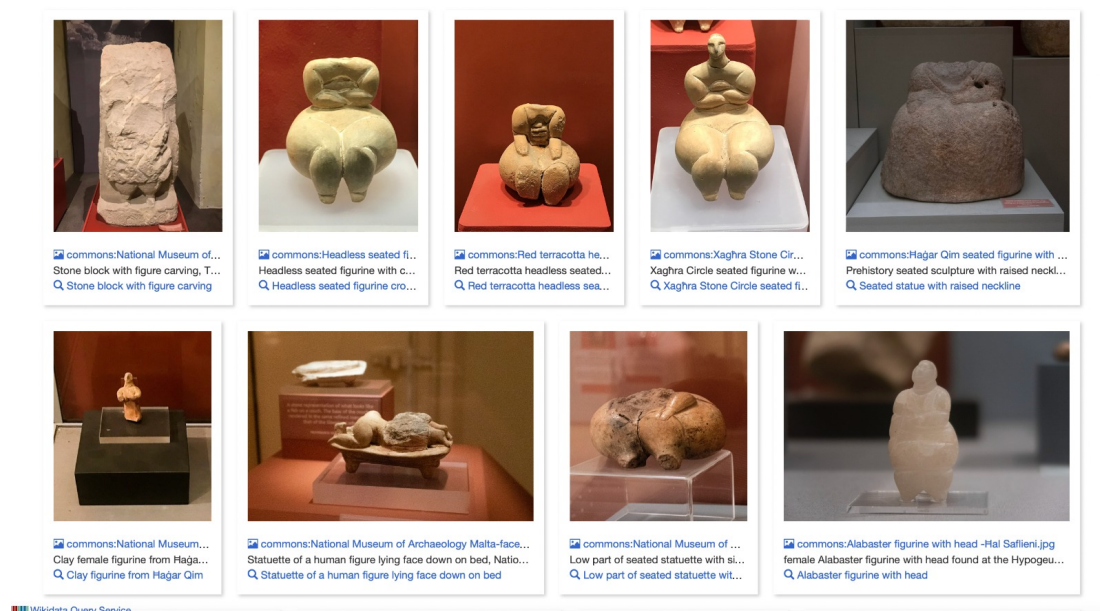
Fig: 5. Partial view of the dataset ImageGrid query in Wikidata.
This is available to view in full at https://w.wiki/6kEB

By virtue of the fact that this dataset is held within an open structured data platform under a CC-0 license (i.e. in the public domain, without any intellectual property rights reserved), it can be reused freely by anyone for whatever purpose. Elements from the dataset are already in use on Wikipedia, enabling the Wikidata item to be attached to encyclopaedia articles about specific items in the collection. At the time of writing (May 2023) this has not happened yet and only one Wikidata item (Q1809373) from this dataset is associated with a Wikipedia article: 'Slapende Dame van Malta' in the Dutch-language version of Wikipedia.[18] However, a similar example of this sort of thing – outside the dataset for Heritage Malta's collection – can be seen in the 'Venus of Willendorf' English-language Wikipedia article,[19] which is attached to the Wikidata item Q131397.[20] This is also available in another 64 languages. Another similar example can be seen in the 'Venus of Dolní Věstonice', which is attached to Wikidata item Q552113, linking it to Wikipedia articles about it in 35 languages.[21]

---

[18] See https://nl.wikipedia.org/wiki/Slapende_Dame_van_Malta
[19] See https://en.wikipedia.org/wiki/Venus_of_Willendorf
[20] See https://www.wikidata.org/wiki/Q131397
[21] See https://www.wikidata.org/wiki/Q552113

4.1.2 Data representation through data visualisation and data sonification

Next are the data visualisation outputs derived through the various tools outlined in the previous chapter and presented Appendix B at the end of this document. These constitute the second stage of the practical aspect of the research project. Perhaps, these may be considered less of an actual output in themselves but rather evidence of a working process intended to see how the dataset developed for this project can be represented visually. However, in turn, these served as the initial building blocks, so to speak, for the data art works created during the course of this research project. In any case, this is partly why they are presented as an appendix at the end of this document rather than embedded in the main body of the text itself, so as not to break the flow of the discussion and analysis of the overall work applied to this research project. Whenever it was appropriate to include data visualisation examples in the main body of the text to illustrate a specific point, this was done in the respective section, for example as presented in section 3.2.3 with figures 3.1 and 3.2.

It is still essential, however, to refer back to the outline of data representation techniques (i.e. data visualisation and data sonification) as presented in subsection 3.2.2 of the previous chapter. To put it another way, to get a full sense of the application of data visualisation and data sonification to the dataset developed through this project, it is essential to take the aforementioned chapter subsection and the Appendix B as the sum of the parts relating to the creation of data representation, which eventually led to the creation of the outputs described in the next subsection as original works of data art.

4.1.3 Original works of data art

The works of data art created through this research project can be largely divided into two: the *Naked Data* installation (exhibited during the DRHA 2022 Conference at the Stanley Picker Gallery in London in September 2022) and the sound recording presented along with the *Ninfa* installation (exhibited during the Prestorjha exhibition at Spazju Kreattiv in Valletta in March/April 2023). The latter work is only partially related to this study, as the installation and works presented during the 2023 exhibition are outside the scope of this research project.

The *Naked Data* art installation consists of a plexiglass structure made from 1-meter-long plexiglass transparent rods. The base section is bent into shapes echoing the linked data nodes in the knowledge graph shown in the video displayed behind the sculpture within the same installation. The vertical section of the sculpture is derived from predominant values displayed in the graphs extracted from the property 'depicts' (P180), which are also captured in the video art displayed behind it in the installation. The video art is based on data visualizations created from the Heritage Malta dataset developed specifically in Wikidata during the first few months of this research project. The creative process of this installation is described in subsection 4.2.8 below.

The digital sound art component within the *Ninfa* installation is composed from the dataset captured on Wikidata for Heritage Malta's collection of prehistoric female figurines. The first part of the sound recording is structured from artistic choices made to set a template from which the Wikidata items, properties, and values chosen to generate the remaining sound composition. The composition depends on the frequency of occurrence in the dataset and correlations between properties and values on different items within the selected dataset, creating rhythms, melodies, and harmonies accordingly. The creative process of this output is described in subsection 4.2.9 below.

4.2   The working process

To provide a fuller picture of the outputs created from this research project and discuss the results and findings of the research appropriately, a general overview of the working process is provided in this section. This spans from the selection of the research topic all the way to the creation and exhibition of the original data art objects created directly from the present research project, including the way the latter aspects of the research project led to a critical reflection on the entire working process. The salient points of the critical reflection can be found in the concluding chapter.

4.2.1   Selecting the topic and creating a hypothesis

Upon closer inspection of the prehistoric female figurines at the National Museum of Archaeology in Valletta, the objects on display can be imagined presented differently than the way the museum's curators have chosen. This idea prompted me to start looking *into* the

collection rather than just *at* the collection. A data science point of view seemed appropriate at this point, especially because physical access to the objects in the collection is understandably beyond what is afforded to the general public visiting the Heritage Malta museums. My first insights came through discovering correlations and stitching things together, allowing the data to provide a better understanding of the collection as a whole. Seeing patterns and connections between data that would otherwise be scattered randomly or lost in a long list, without any sorting criteria, came next. The ultimate goal of this research project is to visualise and process all this into a creative artistic experience.

Following ideas proposed by McCandless (2012), the possibilities of how one can explore a collection while being able to search through it to discover connections and understand them can be explored further. The visual language presented by the object, correlates to whatever other conceptual language we adopt to comprehend the meaning of things more broadly. This idea can possibly also be inverted, asking a hypothetical question on ways to understand the prehistoric female figurines visually through data about them. In turn these insights could be extended to artistic practice based on both the data and the visual images associated with it.

4.2.2   Data gathering and processing

Getting to the initial data entry stage posed its own challenges. The first hurdle came because neither the collection as a whole nor individual objects within it were listed on Wikidata. As noted earlier, one exception was item Q1809373 (The Sleeping Lady of Ħal-Saflieni) which only had two listed properties on Wikidata at that point, and existed through its association with the Dutch-language Wikipedia article on the topic. This item still needed to be curated appropriately on Wikidata with a new structure of properties and values on par with the other items in the same collection.

The data gathered for each item in the collection needed to be both quantitative (sizes, dates, etc.) and qualitative (given names, descriptions, interpretations, etc.). In asserting this need, it became immediately evident that as far as data visualisation techniques go, not all data properties and values within this collection were going to be equally suitable for that purpose. This is particularly so for qualitive data. This is also why it is essential to ensure that the data is captured and formatted consistently so that properties and values are congruent across a target dataset. This is perceived as curatorial practice, both as data curation and curation

through data. Data curation relates directly to the management of the data itself, while curation through data involves decisions on the dataset and adjustments to individual data entries towards a specific curatorial goal.

A request for digital data from the National Museum of Archaeology in early 2022 was met with an offer to be shown printed information, as there was no digital file or any digital dataset available about Heritage Malta's collection of prehistoric female figurines. The process of data analysis of the present research project therefore needed to start with the curation of a list of works. Initially, I was unsure about how many items would be included in the curation of the selected theme: prehistoric female figurines and fragments. An initial list was created through museum visits, documenting whatever is on display, and supplementing this with information from books and journal articles on Malta's prehistoric archaeology (for example: Pace, 1996; Trump, 2002; Vella Gregory, 2005; Malone & Stoddart, 2007). During my visits to Malta's National Museum of Archaeology in 2022 I noticed that the statues, figurines and fragments of interest to this study were somewhat overlooked within the museum. One notable exception is the so-called *Sleeping Lady of Ħal Saflieni*, which is showcased in a small dimly-lit room, to be admired by itself. No wonder it was the only one that had a Wikipedia article about it, albeit in Dutch, and an associated Wikidata item with basic properties and value captured ahead of the start of the present study. The *Sleeping Lady* was originally discovered in 1905 in the central chamber at the prehistoric site known as the Ħal Saflieni Hypogeum, which is one of only three UNESCO World Heritage Sites in Malta. This small statute shows a woman dressed only in a long-pleated skirt lying on a bed, sleeping on her right side. While a small number of pieces are presented in a prominent glass case in the same area of the museum, the rest of the collection is presented on shelves next to each other, inside an unevenly lit glass cabinet. Unless you deliberately position yourself to take a closer look at them, they can very easily be missed during a casual visit to the museum. This provided a greater sense of purpose to the development of this theme for the present research project.

In creating a comprehensive list of figurines and fragments, I made sure to remain respectful to the work of previous researchers who classified figurines as female, refraining from new subjective classifications if not already identified as such by any archaeologist endorsed by Heritage Malta. The creation of the initial list was based on site visits at both the National Museum of Archaeology in Valletta and the Ġgantija Archaeological Park in Gozo. Neither site had a digital list of what was held and displayed in the other museum. The printed information

provided to me in a file folder by the curatorial team in Valletta appeared to be copy of a document printed out from a computer. Although no official record confirmed this, it was inferred that the original database was held on a computer that no longer functions, and the paper-based print outs provided the only backup for the data held by Heritage Malta on the objects displayed at this museum. The metadata for each item included a scientific photo of each object, however, the images are not readily available for researchers outside the paper-based ledger held by the curatorial team. Non-flash photography of the objects as exhibited in the museum is allowed freely. The values captured for each object within these paper-based documents includes Inventory Number, Current Repository Number, Location, Title, Dimensions, Material, and Description. This information is often accompanied by handwritten notes or markings on the same sheet. I scanned the documents with my mobile phone, with the intention of capturing the inventory number, title, and location within the museum to supplement information that I had managed to gather from the relevant literature, especially Vella Gregory's 2005 book, which includes photos by Daniel Cilia. The inventory number is among the more important bits of metadata established by Heritage Malta for each object in its entire collection of holdings. In a couple of instances, the title was ambiguous or problematic and the data from the description field was useful to help disambiguate objects with very similar, if not identical, names. A generic title such as "seated headless figure" can make for an imprecise title. Some of the data had evident human error factors, such as m (for metres) instead of cm or incorrect material information, such as ceramic instead of stone.
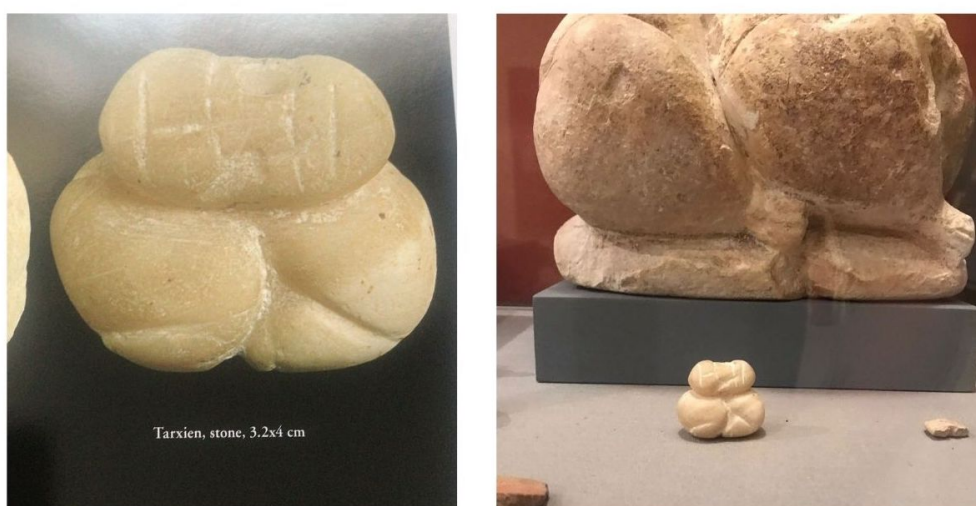
Fig. 6: 'Headless sitting human figurine' from Heritage Malta's collection - Daniel Cilia's photo from Vella Gregory's 2005 book and a snapshot of the object as exhibited in the museum
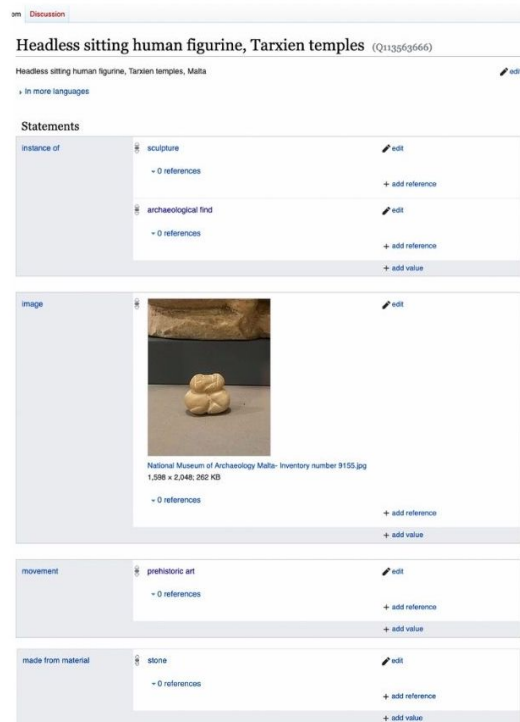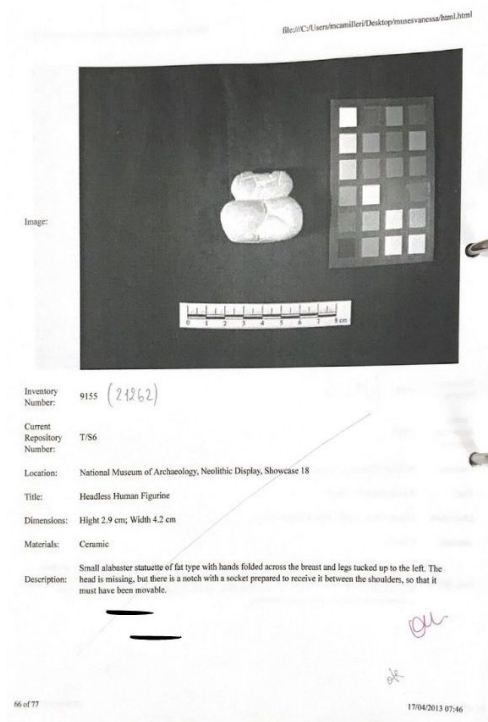
Fig. 7: 'Headless sitting human figurine' - Heritage Malta's data sheet showing inconstancy in the data relating to 'made from material', and the Wikidata item created through the present research project

4.2.3 Entering the gathered data into Wikidata

Wikidata is a free knowledge base that anyone can edit. It contains over 100 million data items. While anyone can access and edit it freely, any editing will be attributed to the user's IP address unless they register a user account first, and ensure they're logged in every time they edit. A user account is also useful for users to keep track of their own edits. The process of creating a new item is Wikidata normally starts by clicking the 'Create a new Item' link, which can be found on the left-hand side navigation menu. This action opens up a web page that contains a basic form that requires the user's language (which is pre-set to whichever language has been selected as the user's default language in their personal user settings), a label for the item, a description, and an option for aliases. Once the item has been created it is assigned a unique identifier, known as QID – Wikidata item unique identifiers have the format Qx, where x is a sequential number.

Once the item has been created, the user can allocate statements through properties and values. Some properties require qualifiers (such as language or start/end date) and an external reference can also be added to each property value. Each property has its own unique identifier, known as PID – Wikidata property unique identifiers have the format Px, where x is a sequential number. Property values normally consist of items already within the database. Both item and property unique identifiers are essential for use in SPARQL queries. Although not immediately of interest to the present research project, each Wikidata item can also be matched up with unique identifiers from other databases, which are also represented within the Wikidata codebase; for example, VIAF ID, ISNI etc. for published writers, or ISBN for books, to mention just two very common identifiers in this category. This provides an opportunity for authority control and disambiguation among items with similar labels.

The open structure of Wikidata makes it essentially a collaborative database, where data can be freely updated as needed. As the data within Wikidata is machine-readable, programmatic interactions can also be automated. While Wikidata is a web-based interface, making it human-readable while the back-end manages the linked data complexity, storing structured data. In this way, Wikidata serves as the central storage system for the structured data of other Wikimedia projects, such as Wikipedia, Wikivoyage, Wikisource, and Wikimedia Commons, among others.

Each item in Wikidata needs to have statements structured by a property and at least one given value. Each item has (or rather, can have) a consistency of statements. The first property required by Wikidata for items within it to correlate with others appropriately is 'instance of' (P31). This refers to that class of which an item is a particular example and member. The property for 'sub-class' (P279) may be used separately to denote a subset of that item. The Wikidata editing interface is designed in such a way as to suggest further properties for each item once the 'instance of' property has been entered. Properties are also automatically reordered according to an item type structure template, which is invoked automatically as the user entering data for a specific item publishes (the term used by Wikidata for saving data to the database) the specific information they have contributed or edited. This makes for a neater experience in the human-readable interface and is a greatly appreciated feature in Wikidata as the openness of the platform would otherwise make for a rather chaotic reading experience for ordinary human users. This feature also makes Wikidata more accessible to new data entry volunteers and newbies who may not have received any formal training in metadata basics,

such as some of the distinct types of metadata: descriptive, structural, administrative, reference, statistical, etc.
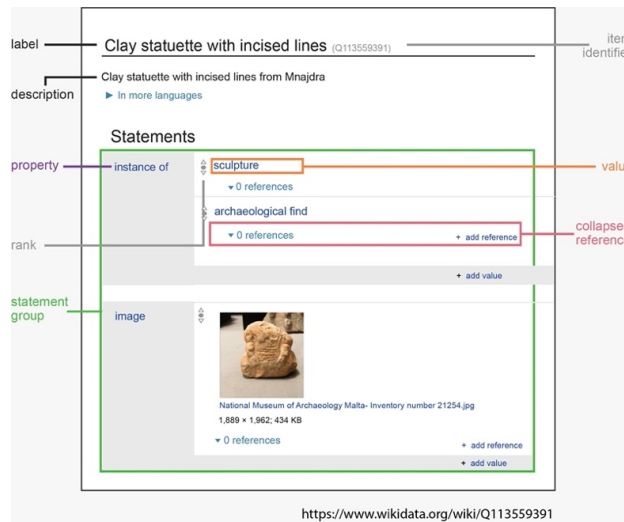
Fig. 8: Data model diagram for Wikidata, adapted from 'Datamodel in Wikidata.svg' (2016) on Wikimedia Commons by Charlie Kritschmar – CC-0

Once this basic structure is understood, it is relatively easy to comprehend a fuller item containing multiple properties and values within its statement group, such as the one shown here, which was created specifically for the present research project:

Fig. 9: Screenshot of item Q113559391 from Wikidata.org

The way to describe resources on the web, developed by the W3C is known as RDF (Resource Description Framework) triples. A triple is made of three parts: subject, predicate or feature, and object or value. Thus, an example from the present research project would be 'The Sleeping Lady of Ħal-Saflieni' (Q1809373) is an 'instance of' (P31) an 'archaeological find' (Q10855061). There can be multiple triples for each subject, and values can be subjects in other

triples. Triples can be simple or they can be complex, containing multiple instances of triples concatenated to each other through at least one common value.

Exploring data in a structured manner through Wikidata's query service involves the use of SPARQL, the RDF query language. However, the web interface also provides a query assistant for users who are less familiar with SPARQL. Knowing this query language is useful for anyone who wants to extract more specific results from the database. This also applied for basic data visualisation of datasets through basic functions such as colour coding. A basic query associated with the present research project has already been presented in section 4.1.1.

4.2.4   Analysing the data: before visualisation

The initial stage of data analysis in this type of research project is very rudimentary. Because the number of items and statements associated with them is manageable, it involves manual scouring to ensure consistency, catching human error, and getting a sense of the big picture captured in the dataset as a whole. This holistic method of data analysis readily leads to an observation of correlations through similarities, even before the dataset is fed into a data visualisation tool. This can also help in the selection of the most appropriate visualisation process to adopt, assuming that a basic understanding of the possibilities of which data visualisation tools can be applied to which end. All the while, keeping an open mind to the prospect of discovering answers to unasked questions is also useful as this potentially yields opportunities for further research.

In capturing the data gathered for the present project into the Wikidata platform, the main focus was on the visual representations of the female body, with special attention to the choice of material from which these figurines and fragments are made. Along with the essential metadata mentioned earlier, what was acquired at this stage were properties relating to self-evident representations of the female body through biological characteristics. In some cases, the figurines have previously been conventionally described as "fat" or having "non-realistic proportions" (see especially Trump, 2002; Bonanno, 2010). Interestingly, Heritage Malta's records show that such descriptors are now being replaced with terms like "corpulent." In this way, however, the data is also sex-disaggregated, and remains sex-disaggregated until gender justice has been achieved in the manner advocated by data feminists (D'Ignazio and Klein, 97-

123). The method developed for the present study can easily be scaled up to the reset of Heritage Malta's collection of prehistoric human and animal figurines and fragments.

The metadata pertaining to descriptions of the object is particularly useful from a creative exploration perspective. This translates to the 'depicts' property (P180) on Wikidata, which helps capture the formal qualities of some of these objects in ways that can sometimes be somewhat subjective. As described in 4.2.2 above, the approach selected for the present study aims to retain respectful capturing of the metadata provided by Heritage Malta endorsed archaeologists when classifying figures as women or fragments as parts of a woman's body. In collecting factual data, historic data needs to be kept accurately too as this enables the historiographic analysis of the way such objects are described and classified. Simiarly, the values associated with 'depicts' (P180) are the ones that are predominantly open to an artistic interpretation from a visual perspective.

4.2.5   Visualising the Wikidata: initial processing

An essential detail guiding the research process at this point questioned whether the data should drive the visualization. Such a question is most pertinent in the context of data art, or any other artistic manifestation that can arises out of a particular dataset. The question is one about ensuring that the source data remains recognisable as an interpretation of the dataset on which it is based. In this case, the dataset of 38 items from the Heritage Malta collection is rather small, but working on larger collections one can import metadata in batches either by using an application designed for machine readable data, such as OpenRefine[22] or a spreadsheet structure ingested through a bot like QuickStatements[23].

When a knowledge machine learns, whatever it learns goes into latent space. In other words, data is simply stored statically until it is queried. The stored data awaits input, requesting something to be done with it. As statistician Nate Silver famously explained in his 2012 book *The Signal and the Noise*, "The numbers have no way of speaking for themselves. We speak

---

[22] OpenRefine is a downloadable free, open-source tool for working with messy data: cleaning it; transforming it from one format into another; and extending it with web services and external data. Available at https://openrefine.org

[23] QuickStatements is a Wikimedia Cloud Services tool designed to batch-edit Wikidata. Available at https://quickstatements.toolforge.org

for them. We imbue them with meaning" (9). This is to say that our agency is what gives data, especially quantitate data, meaning beyond the facts.

The intention for the project at this stage was to capture each object within Heritage Malta's collection of prehistoric female figurines and fragments as an individual Wikidata item. Reading the initial contents of the dataset created in Wikidata through visualisation tools was not very useful. This is because most data visualisation tools are directed at creating graphic representations of quantitative data or statistics. Using the visualisation tools on datasets within Wikidata, provides the opportunity to compare a particular dataset about a specific collection of objects within a museum with another collection, even at another museum, provided this is already present within Wikidata. This comparative process provides an opportunity to better understand the relevant item structure within Wikidata, enabling updating and improvements on specific properties within the item's data. The presence of different item structures in a single repository enables users of Wikidata to query the metadata in a more flexible way.

As a Wikimedia project, Wikidata integrates images from Wikimedia Commons quite seamlessly through the 'image' property (P18).[24] I photographed all the objects in the Heritage Malta collection of prehistoric figurines and uploaded them to Wikimedia Commons to enable me to provide an appropriately licensed reference image for each object within the Wikidata item about it. These photographs provided an instant visual aid even when simply browsing through the dataset before the application of Wikidata queries and visualization tools.[25] It is in this empirical way that feminine body forms were eventually identified as an essential property for objects to be included in the final dataset adopted for application in the *Naked Data* art installation.

Once a dataset has been captured into Wikidata, open tools such as Knowledge Grapher can help users explore and navigate through the data.[26] LOD4Culture is a Spanish tool that extracts cultural heritage data from Wikidata and presents it through location maps.[27] The #defaultView:Graph constructor on SPARQL queries is very powerful, and it is one that has been applied extensively to the present research project. This enables relations between data

---

[24] P10 provides a 'video' property in Wikidata from Wikimedia Commons but this has not been used in the present research project.
[25] A grid containing all the images with links to the individual images on Wikimedia Commons as well as the individual Wikidata items is available here: https://w.wiki/6kEB
[26] Wikidata Knowledge Grapher is available at http://knowledgegrapher.toolforge.org
[27] LOD4Culture is available at https://lod4culture.gsic.uva.es

items to be visualised and explored interactive through a simple click of a mouse. When clicking on each node in the knowledge graph, the user is provides with a new set of triples instantly showing the correlations between items in the dataset.

For anyone who may find these tools a little daunting, the one to go to for visualising Wikidata queries is called Viz Query.[28] This tool can be found within the ToolForge platform provide by Wikimedia but it is not integrated into the Wikidata Query Service visualisation tool set. As it happens, this tool provides more user-friendly results when displaying an image grid from a dataset. No knowledge of SPARQL is required to use this tool. You only need to know the property value that to be queried. Its command line is as simple as a form is simple English, such as: "Select items where...it has a property (click to set property) that is (click to set item)"; wherein the clickable variables pull up a dropdown search menu directly from Wikidata for a property and item respectively.

Back to using the Knowledge Grapher: the 38 items within the Heritage Malta set, all but one of which have been created in Wikidata through the present research project, constitute a mere fraction of Heritage Malta's vast collection of cultural heritage artefacts. However, if these 38 items are queried along with all other items that are an 'instance of' (P31) an 'archaeological find' (Q10855061), 2,227 items show up in the relational knowledge graph. The experience of clicking on any of the various nodes and encountering other items is an similar to discovery by serendipity within a museum when viewing objects in an exhibited collection for the first time.

Fig. 10: The 2,227 items that have the property 'archeological find' in Wikidata, mapped across part of a world map, the mapping visualisation function in the Wikidata Query Service

---

[28] Viz Query is available at https://hay.toolforge.org/vizquery

Another example of how different collections can be linked together can be performed using a better-known archaeological find: the Venus of Mauern, which is from a collection outside Malta: the Bavarian State Archaeological Collection. If data items follow similar item structures better linking across datasets is enabled. Running SPARQL queries on Heritage Malta's collection of prehistoric female figurines and fragments, provides a better understanding of the collection as well as clusters of items that share properties. Removing parameters helps to better understand ways to visualise the collection. Writing a query that gathers the necessary information for a basic visual with a location is rather straightforward. This is available through the visualisation tools in the Wikidata Query Service and the location value can be added to the SPARQL query. A French tool called Datavizdev.fr provides a more sophisticated interface that produced more advanced visualisations on maps.

4.2.6   Refining the dataset in Wikidata

After the initial data entry stage on Wikidata, I started applying the data property 'made from material' (P186) for each object in the collection. Another significant bit of work I needed to perform was on the properties associated with location. There are three relevant properties here, each of which has significantly different values. These are: 'location' (P276) – i.e. , 'location of discovery' (P189) – as originally reported by archaeologists, and 'coordinate location' (P625) of the site of discovery.[29] Once again, this can be reported here as an exercise in data journalism, as such data is not consistently recorded by Heritage Malta. This particular way of working follows McCandless' (2012) approach wherein he encourages data analysts to focus on the relationship between facts, context, and connections that make information look beautiful when visualised through appropriate tools.

The main property I chose to visualise through a SPARQL query was the property 'made from material' (P186). Visualising this as a relational knowledge graph enabled me develop alternative ways to visualise it through Adobe Illustrator and Photoshop to produce an aesthetic that is more to my liking. I created Thus, a sunburst diagram created by visualising values associated with P186 to item labels in RAWGraphs 2.0,[30] was subsequently manipulated through Illustrator and Photoshop, the popular vector graphics editor and image editing

---

[29] This data can be viewed in this table generated through an appropriate SPARQL query: https://w.wiki/6kzN
[30] RAWGraph is an online data visualisation tool available at https://app.rawgraphs.io

application used widely by visual designers. The data immediately started showing itself in ways that are not obvious in a text-based table intended primarily to show a list of items. These basic visual representations of the data provided an opportunity for the data to be imagined from a creative, if not artistic, point of view. In this way, progressive iterations can be seen as a possible approach to the ways the data about these individual objects and this collection as an ensemble can be visualised through artistic interpretation. Here it is also useful to distinguish between infographics and diagrams. Infographics present information in an explicitly visual way, while a diagram is more implicit, possibly requiring further explanation.

Analysing visualised data has also helped me to see patterns that have resulted from sorting, grouping, and filtering the dataset. On an open data platform like Wikidata, datasets can be altered and adapted for different visualizations. While the primary concern in the context of this project is artistic, this is also applicable in developing enhanced approaches to collection cataloguing. It also demonstrates the possibilities embedded in looking at a dataset holistically, rather than a collection organised by title or the period in which the object was originally created, discovered by archaeologists, or acquired by the museum.

4.2.7 Visualising the Wikidata: further processing

Giorgia Lupi (2017) insists that design should serve the data in visualisations. This makes for design-driven data visualisations, where the design is the first thing you see in any visualisation of data. Lupi makes the case for graphic design choices that make the audience want to know more. Although data visualisations often aim to convey information to their intended audiences, they can also convey an experience.

Once the basic dataset was in place on Wikidata, I started exploring data visualisation tools that can either read SPARQL queries or the exported datasets from Wikidata. My aim at this stage was to find ways to arouse curiosity by the potential for artistic narration within the collection, from the dataset. With this dataset, mapping tools invoking 'location' (P276), 'location of discovery' (P189) and/or 'coordinate location' (P625) did not render visuals that were potentially as interesting, from a visual design perspective, as those involving the visualisation of the property 'made from material' (P186).

The data visualisations from the dataset created for this research project took various forms and iterations until something I felt was interesting enough to convey showed itself. Risking creative vulnerability, I let go of my preconceived direction, largely abandoning what I had planned; some of these earlier workings have been included in Appendix B at the end of this document for illustrative purposes. The data thus started informing the development of the illustration, which had taken on a new visual meaning in the chosen context. Using Adobe Photoshop to apply the material texture to the chart I created, I started to take a more playful approach with my exploration. I applied the tool to move shapes across the screen, following the direction of the form that was being shaped. David McCandless' TED talk immediately spring to mind here:

> "There's something almost quite magical about visual information. It's effortless. It literally pours in. If you're navigating a dense information jungle, coming across a beautiful graphic or lovely data visualization is a relief. It's like coming across a clearing in the jungle." (2010)

The moment I became aware what was happening to the illustration originally processed through RAWGraphs was a spectacular moment in my creative process. Along with the material textures I inserted into the chart, I had also applied the effect of light and shadows on the structure that emerged. Part of the stonelike design was easy (for me) to anthropomorphise into a face, without overthinking the tendency for pareidolia, which is the technical term for the human tendency to see faces and other meaningful shapes in random objects (cf. Arnheim, 1969). In this mode of looking at the illustration, the other three materials joined up easily to form a figurine-like body. The anthropomorphisation of the data visualisation was thus complete. The data did not need any more adjustments before I could proceed to the next stage: making data art.
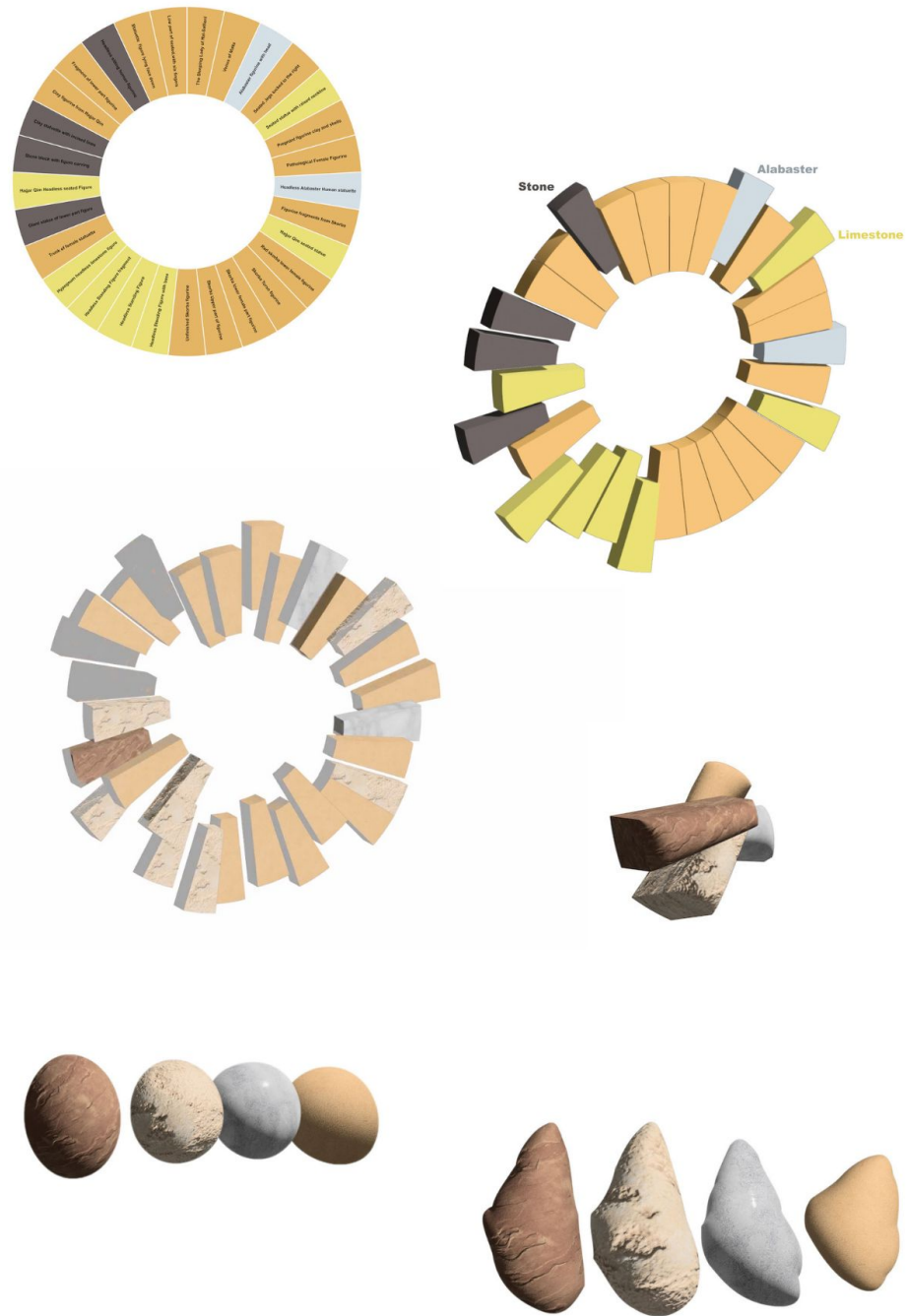
Fig.11a: The sunburst diagram produced through RAWGraphs 2.0 is manipulated through Adobe Illustrator and Photoshop in an artistic exploration of the four main types of material found the Heritage Malta's collection of prehistoric female figurines and fragments
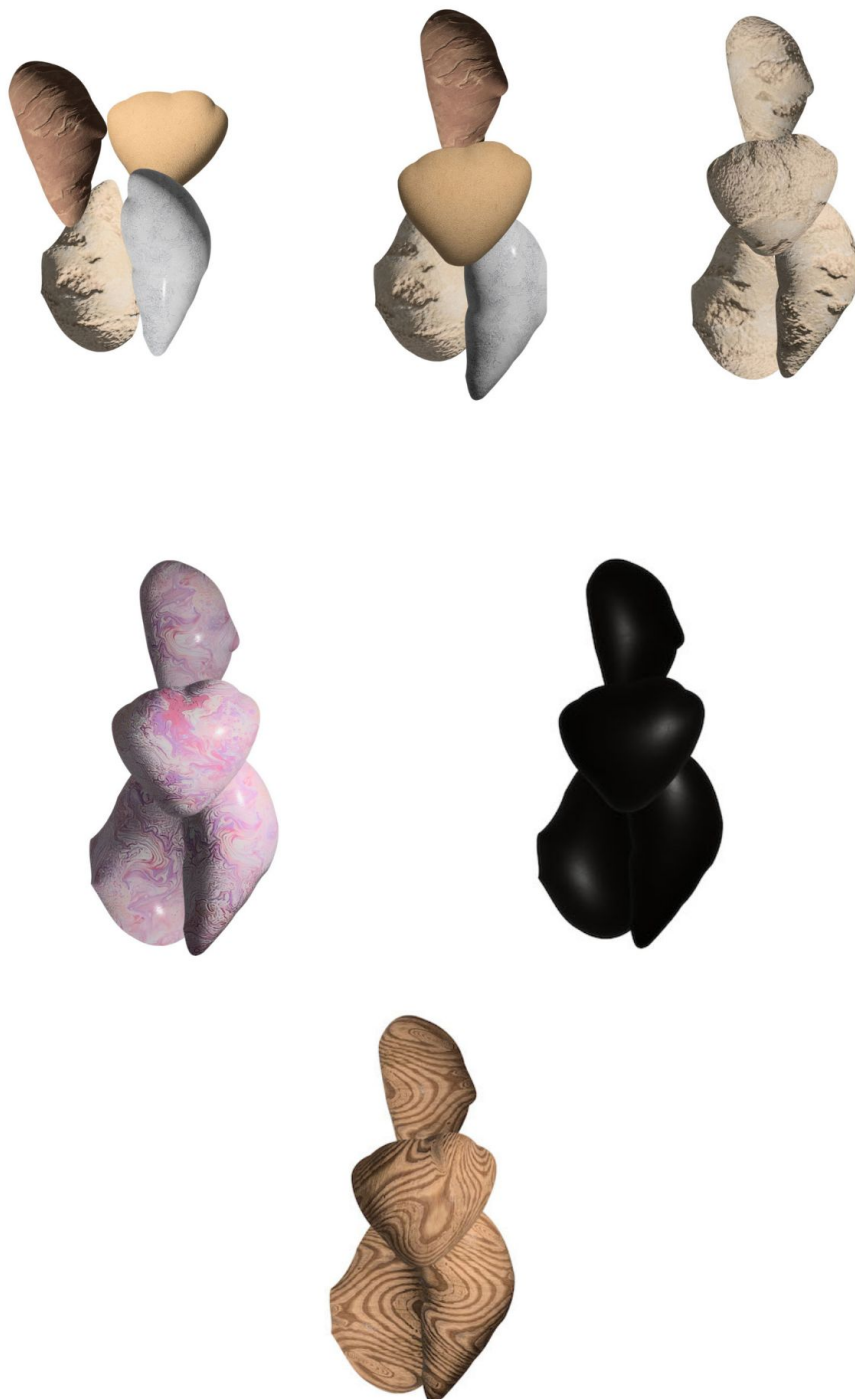
Fig.11b: The lumps or slabs of material represented digitally are arranged into what emerges to be an anthropomorphic figure to which textures are applied in Photoshop to explore the effect created Adobe Illustrator and Photoshop in an artistic exploration of the four main types of material found the Heritage Malta's collection of prehistoric female figurines and fragments

4.2.8   Visualising the dataset to make data art

Artists tend to see things that are not visible to everyone else. Data artists need to understand the dataset they are visualising before creating art from it. The art still needs to reflect the reality of the data, all the while giving the possibility of understanding and discovering new information by connecting data in ways that may not be obvious to the casual viewer. An artist's role in this process is to observe and tweak while being aware of the iterative changes, from data visualisation to an artistic creation. This is a type of artistic practice that produces art that does not only rely on the artist's expression but also on the creative interpretation of data. Some of the most impactful data art (including examples mentioned in Chapter 2) provide their audiences with an opportunity to better understand the dataset on which they are based, even if not at first glance. For the sake of visual clarity, sometimes it is important to exclude information in a query intended to render a particular data visualisation. Having all the essential data in place for the dataset selected for this research project on Wikidata, and having explored various ways to visualise this data the next step, my process involved generating a specific query to model my visual ideas accordingly. This was done even if it meant that aspects of the dataset would be excluded from the data art visualisation.

The first sequence in the *Naked Data* video is a screen recording of my own interaction with the knowledge graph I made from the full SPARQL query visualising all the objects in the dataset located at the National Museum of Archaeology in Valletta. In this relational knowledge graph visualisation, I show all the selected items in Heritage Malta's collection dataset linked together by location, material, and by what they depict. The main property value that links everything together is visualised in the middle of the graph. Understandably, this a central topic for this research project. The middle node connecting all this data together is the property value 'woman' (Q467), which is derived from the value attached to the 'depicts' property (P180). Every object in this dataset depicts a woman. This is a collection of female figurines and fragments. In this screen recording I show how pulling this symbolic representation for the value 'woman' across the screen brings everything else in the collection along with it. The 'woman' node pulls every other node behind it, each of which is also connected to other nodes.

I next chose to render the same SPARQL query from which I derived this relational knowledge graph into a pie chart, identified within RAWGraphs as a sunburst diagram, to show correlations and similarities in the dataset. Through a process of selection, the property 'made

from material' (P186) became the focal point of this visualisation. This instantly made very visible the materials the figurines in Heritage Malta's collection are made from: clay, stone, limestone, and alabaster. Once again, every item in the dataset was connected through these values. In the largest graphic in this visualisation exercise, each item is colour coded according to the material it is made from, identified by the figurine or fragment label, which is essentially the name of the item. Playing with 3D renditions of the sunburst diagram, I sequenced these into a snippet of stop motion-like animation. In this brief animation sequence, I wanted to illustrate the process the dataset was being put through in being animate from a csv file that was exported to create a basic pie chart that eventually transforms into an anthropomorphic figure, as explained earlier. The data visualisation diagram is given a new life by being consolidated into four slabs, each representing a specific material. Each of these slabs starts resembling different body parts from the prehistoric figurines. As the anthropomorphised figure emerges, the material becomes secondary and all four slabs a mixed down into oval object and one textured colour. It is as if these are four pieces of some malleable material that is about to be moulded to create a figure. This figure presented itself in the data visualisation as I attempted to make data art from it and I did not force the creation of the figure in any way. I merely rotated them and animated them to appear on top of each other. The stone slab became the face, the clay slab turned into the torso, while the alabaster and limestone slabs formed the lower body and legs. One by one, I changed the texture of these four digital object materials into one: limestone. For the final sequence, this figure was then transformed into a contemporary context, represented into modern materials, including marble, broken tiles, wood, fabrics and other such textures showing the large range of possibilities. Developing these possibilities into physical object was outside the immediate scope of the present study.

Fig. 12: A still from the final sequence in the video for the *Naked Data* art installation, showing the figurines created from the visualisation of the dataset in various alternative textures

My material of choice as a professional artist is transparent plexiglass. This has been consistently present in my body of work since 2009 (cf. Sant, 2023). Through this final image in the *Naked Data* video, I did not see this transparent material as a possible texture for the final figure that appears in the video. However, I was intrigued by the way the dataset from Heritage Malta's collection was visualised to show how everything correlates as a web of things with the 'woman' node at the centre. To my mind, this visual needed to be expressed through my own material: transparent plexiglass. What I created is a delicate nest of nodes connected together through a loose-knotted weblike structure. This has been created with fine, transparent, plexiglass rods. This object is not based on an artist's sketch. I only kept the information graph visualisation at hand as I formed the sculpture. I aimed to visualise this experience, by capturing the web of the linked data and some of the common anatomical features that depict values found in this collection of prehistoric figurines and fragments. These item property values include, hips, breast, shoulders and vulvic symbols. These are largely visible in the quasi-figurative vertical component of the sculpture emerging from the nest of nodes at the base.

This is how I created a free-standing sculpture that can be viewed from any side, embodying these features in a figure-like outline made from one plexiglass rod while the base of it echoes the shape of the information graph as a web, nest-like form made of various pieces of plexiglass rod. A data visualisation experience that presents the essence of the selected dataset has the sculpture shaped through the interaction of the material as it was formed by hand to visualise the prehistoric handmade objects through data art. An aesthetic effect produced by this sculpture is only visible when light is cast on it. In a direct beam of light it not only glows, because of the way transparent plexiglass deflects light, but the shadow it casts resembles a line drawing of the information graph rendered through the data visualisation. A set of images showing this installation as it was exhibited at the Stanley Picker Gallery in September 2022 can be found in Appendix C containing the research process portfolio, presented at the end of this document.

4.2.9   Creating data art with sound: exploring data sonification

Through the curation of the dataset to create the *Naked Data* installation, I became very familiar with Heritage Malta's collection of prehistoric female figurines and fragments, especially in

terms of the data held about it within the Wikidata repository. I soon came to realise that there was further potential for this dataset to be explored from an artistic angle, which could also provide further access through experiential options. It was at this point that I chose to consider data sonification.

Like data visualisation, data sonification is a process of data representation through aesthetic means. In this case, sound can be created from a dataset, attributing specific aspects of sound composition to specific properties and values associated with items within a selected dataset. One such example is *Listen to Wikipedia* (2013) by Hatnote, which is quite well-known among seasoned Wikimedia project editors and developers. This instance of data sonification, which has now been active for about ten years, enables you to listen to the sound of Wikipedia's 'recent changes' feed. The creators of *Listen to Wikipedia* describe its representational elements very succinctly on the Hatnote website:[31]

> "Bells indicate additions and string plucks indicate subtractions. Pitch changes according to the size of the edit; the larger the edit the deeper the note. [...] You may see announcements for new users as they join the site, punctuated by a string swell."

Part of my artistic practice entails setting challenges for myself to explore new materials and techniques. In this way, I chose to go beyond the visual representation of the dataset I curated for this research project by exploring data sonification. A basic knowledge of tempo and melody through my hobby of playing the kalimba gave me enough confidence to explore sound composition from a musical perspective.

Keeping in the general direction of the present research project, employing open-source applications whenever possible, I turned to TwoTone, which is an open-source web app that converts data into music.[32] A Wikidata query of all the prehistoric female figurines and fragments in the Heritage Malta collection held at both the National Museum of Archaeology in Valletta and Ġgantija Archaeological Park in Gozo, provided me with an exportable csv file, which is all that is need to import into TwoTone to get going. In the process of importing the dataset into TwoTone it is essential to identify the data source that is to be sonified. The property 'depicts' (P180) is appropriate for such an exercise. Selecting it as the main data source enabled the initial creation of the data sonification from this dataset.

---

[31] Listen to Wikipedia by Hatnote is available at http://listen.hatnote.com
[32] TwoTone is a free open-source web app to turn data into music. Available at http://twotone.io

The resulting data sonification file provided a sonic representation of the selected data from the Wikidata dataset. Each of the 227 sound triggers represents a data value within the selected property. Although TwoTone provides options for several combinations of sounds and filters of the data, I chose to keep it relatively simple. This exercise provided enough raw sound objects to proceed to manipulate them into a musical composition; the equivalent of the visual art outputs in the *Naked Data* installation. Here I was guided by David R Worral's explanation that, "It is useful to distinguish data sonifications made for the purpose of facilitating the communication or interpretation of relational information in data, and data-driven music composition, ambient soundscapes and the like—the primary purpose of which is personal expression and other broader cultural considerations, whatever they may be" (180). Some resulting sonifications may be more complex, or even more or less musical, than others, depending on the artist's choice of format manifestation.

The sound produced by the data imported into TwoTone was exported as a MP3 file enabling it to be manipulated in a digital audio workstation that has multiple tracks, which can be layered sounds towards a final sound production. My digital audio workstation of choice is GarageBand, which is freely available on my Mac laptop. As I was already familiar with the media production timeline interface from working with iMovie (within the same Mac productivity suite) the learning curve to adapt my existing skills was not steep. Although I did not design the sound file to accompany the *Naked Data* art installation I produced earlier, I kept the name Naked Data for this sound work, as the same creative principles were now being applied to a sound composition stemming from data sonification with data from the same dataset I had used earlier. The TwoTone file was used to trigger new musical sounds, such as an acoustic guitar with a sharp chorus effect applied to it. A rhythmic layer provides a heartbeat-like quality throughout the track, giving me a sense that this is data that relates to human figures. The variations in the rhythmic track come from the different depictions from the data about the figurines. I also experimented with voice labels for each of the P180 values (hips, legs, mouth, etc.) but felt it made the sound composition too busy, so I discarded the voice track in the final mix. This opens up the sound art to multiple interpretations by whoever hears it. The first part of the sound work is manipulated to create the aural palette I chose for this work. However, after a little while, the rest of the composition is produced by the dataset itself without further artistic manipulation. You are therefore literally listening to the dataset for Heritage Malta's collection of prehistoric female figurines and fragments.

The Naked Data sound file is available at https://on.soundcloud.com/Ex6J7

4.3   Summary of outputs and findings

The project outputs and findings have been outlined in this chapter. Following an overview of the three main practical outputs discussed in this written document – an open dataset available through Wikidata, data representation work through data visualisation and data sonification tools, and the creation of original works of art – a detailed account of the working process was presented. The working process has been subdivided into nine steps that are generally aligned in clusters of three around each of the project's practice-based outputs. Selecting the topic and creating a hypothesis led on to data gathering and processing, which eventually enabled the data capture into Wikidata to create the dataset required for the completion of the present research project. The ensuing data analysis was also a three-step process, from before the data visualisation work to the initial data processing, and on to the stage of refining the data to yield the desired design effects in the data visualisation outputs. Further processing of the dataset enabled the visualisation of the dataset towards the creation of data art from the same dataset, eventually leading to further artistic exploration via data sonification. Taken together, these outputs and findings constitute the new insights delivered through this practice-based research project.

Chapter 5
CONCLUSION

Why use data art to interpret a dataset that can be visualized through conventional data visualization methods? This question enables an evaluation of the practice-based research project. It does so not by looking at an isolated art installation but potentially by proposing a model for the application of datasets in the making of new art objects. The project engages with a specific dataset through a creative process to enable the learning of new things about a collection of museum objects that would probably not have been possible by simply observing, replicating, or interpreting them in isolation. The main topic for specific consideration on this project is inflected through a consideration of feminist perspectives on the collection. As shown in the written document produced as part of this research project, this guided the development of data art through data representations built through digital curation techniques.

One of the main objectives established for this project was to provoke new research opportunities not only on this dataset but also on other datasets through the method developed here. While it is easy to say that are numerous other opportunities for exploration, even within the same museum system, only time will tell whether this working model is actually picked up by other researchers. Ultimately, the way of working with Wikidata proposed here is to connect collections together through links that are either not obvious or ones that are not necessarily easy to process without the magnitude of data aggregation that is provided through this open knowledge database, which is both human readable and machine readable. In the context of this research project, this entailed transforming data into open knowledge and making it inclusive, while producing art that provides wider creative experiences.

With the dataset developed through this project readily available on Wikidata, things that remain relatively hidden in the museums can now potentially be discovered either before or after a visit to the venues where they are displayed. Opening up knowledge about cultural heritage collections can undoubtedly enable new audiences to look for things that are not necessarily the ones presented prominently by institutional curators. The method of artistic expression has been deliberately chosen for this project as fuel for flights of fantasy beyond ordinary imagination. The idea for this research project arose from an exercise to create data from art, which in turn triggered a curiosity towards the exploration on the possibility for the data from art enabling the creation of a new work of art. A data point of view (Duarte, 2019) is expressed in an art installation intended as an opportunity for feminist data perspectives (D'Ignazio & Klein, 2020) to be explored in ways that have not happened before. The main idea is to turn data visualisations from a dataset into an art object, driven by the aim to create

an experience for the exploration of the collection from a data perspective. This can potentially also make possible the discovery of connections that are otherwise not evident without curatorial intervention.

Wikidata is an open repository for structured data. The more galleries, libraries, archives, and museums connect their databases to this open web platform the more powerful the results of data queries become. A free and open knowledge base, such as that offered through Wikidata, has potential for data re-use. It freely allows the creation of new data-driven projects. This is an innovative way to analyse and transform data into new research projects, raise awareness on things that otherwise lack visibility, and use data visualisation (and data sonification) as materials for the creation of new art installations. The collection of prehistoric female figurines in Malta provided the dataset for this practice-based research into open data art. All open data was based on information that the collection had already made public, through books, papers, and through the Internet, such as Google Arts & Culture. No unpublished data and no personal data was used in this research project. The data was processed through the Wikidata platform to enable further use under an open (CC-0) license for all structured data within this platform.

At the start of this research project, I was not sure which collection I would be working with, to gather and organise data through the Wikidata platform. The initial idea was to work with a collection that holds works by women artists and possibly address the gender gap, while analysing the art in the collection and the artists who created the work. In the Heritage Malta collection that was eventually selected for this study, there is no information about who created the figurines, so the focus shifted to other aspects. This included considerations for things like locations associate with each object and the materials from which they are made. This made for the development of the initial purpose into an opportunity to showcase valuable aspects of the data through a critical perspective via a visualisation project that extends the factual information about the collection into a work of digital art, which also contains a physical component and sound. Developing a methodology that can be used with other collections, from prehistoric art to contemporary art, was also a goal of this research project.

Through this project I have also developed a working relationship with the context in which the collection is held. While this is not necessarily consequential for the present research project, or even for whatever I decide to do next, I think that it is valuable for the people working directly with the archaeological collections at the Heritage Malta museums. Heritage

Malta staff have informally expressed an interest in developing their Wikidata skills and capacity for using Wikimedia projects. One of the archaeologists has even suggested an away day for the Heritage Malta team to engage further with Wikidata, once this research project has come to a formal end within the university setting.

An established aim in the process of this project was to explore data visualisation techniques towards the creation of art objects through data processing. Once the dataset was refined through database queries, a creative freedom emerged, enabling the exploration of possibilities to make art from this particular dataset. This is how I came to the point of creating a work of sonic art from data sonification. Although this was not one of the intended outputs at the start of the research project, it is nonetheless a very concrete example of the broad possibilities for art making that emerge from engaging with a dataset in Wikidata the way I have proposed for this research project. Although sound art is a new way of working for me, I am aware that acquiring skills in music composition and/or sound production can enable the creation of more works that potentially more meaningful or impactful from a technical perspective.

The sound art piece I developed through the data sonification exploration within the framework of this research project has already been integrated in a solo art exhibition I presented in Malta in March 2023. This exhibition, entitled Prestorjha (a Maltese-language neologism equivalent to the feminist term Preherstory), builds on some of the insight I have gained through the work on this research project. However, because most of the works in the March 2023 exhibition were not derived from this research project, I believe that it is beyond the scope of this study to include much more information about that experience here. The part that matter relates to the sound art piece, which enabled me to present my insights into Heritage Malta's collection of prehistoric female figurines in an art setting outside the structure of this research project.

Although I knew I wanted to move away from a digital screen presentation, that is beyond the video art piece that comprises an element of the *Naked Data* installation I presented at the Stanley Picker Gallery in September 2022, my initial considerations where for the creation of a 3D printed art object, such as the figures that appear in the final sequence of the video art piece. Digital video is a noticeably more common medium used by data artist. Sculpture is less frequently created from data. Producing a handmade object in my preferred medium of transparent plexiglass gave me the opportunity to create a deeper artistic connection with the Heritage Malta collection of prehistoric objects, even if the direct connection is through

querying the dataset rather than mimicking the formal aspects of figurines or fragments displayed in the museum. The tactile aspects of handmaking an aesthetic object brough me closer to the creative impulse held within the prehistoric objects displayed in the museum. Through this connection I came to see that the final manifestation will always relate directly back to the collection and the dataset that represents it, even if this is not immediately evident to the audiences who see the new work of art.

As an open repository of structured data, Wikidata enables information within it to be organized meaningfully outside the database. Applying data visualization tools on a specific dataset within Wikidata, enabled the creation of data art. This is probably not what data scientists would expect as a follow up sentence, or action, after the first sentence in this paragraph. This research project has however been consistent in its aim to use data visualisation techniques to create art objects providing an experience for the exploration of the selected collection of archaeological objects from a data perspective.

This research project has enabled me to explore and develop new skills while mastering ones I already had. This is especially applicable to Wikidata and the use of SPARQL queries. This line of inquiry has enabled me to meet data science professionals who were fascinated by the way an artist engages with things they rarely see as related to art making. Conversely, I have also communicated with other artists who are keen to explore Wikidata and data representation techniques as a new way of making contemporary art. While I am certainly much more of an artist than I will ever be any type of data scientist, I have also developed a greater appreciation for the need to bring more science into art and more art into science.

Appendix A

Dataset samples

- Examples of data from Heritage Malta
- Wikidata sample entry

From Heritage Malta
Examples of Heritage Malta input towards the dataset now available on Wikidata

file:///C:/Users/mcamilleri/Desktop/musesvanessa/html1.html
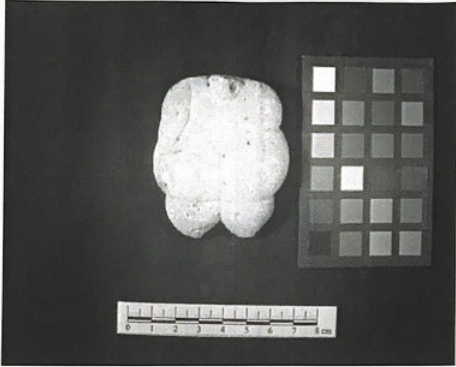
Inventory Number:	20525 (21236)
Current Repository Number:	S/S36
Location:	N/A *showcase 17. Location on online mives is correct.*
Title:	Alabaster Human statuette
Dimensions:	Height 6.1cm, Maximum width (at the hips) 5.5cm, Thickness (at the hips) 2.7cm, Thickness (shoulder blade) 1.5cm
Materials:	Stone, Probably Alabaster
Description:	Headless figurine of white alabaster. Biconical hole asymmetrically bored through neck. The arm positions are identical with SS37 (object ID 20521) (the left arm is crossed over the abdomen, the right hangs down) but the hands are better represented. The back is concave and there is no clear representation of breasts.

7 of 46   17/04/2013 08:08

file:///C:/Users/mcamilleri/Desktop/musesvanessa/html.htm
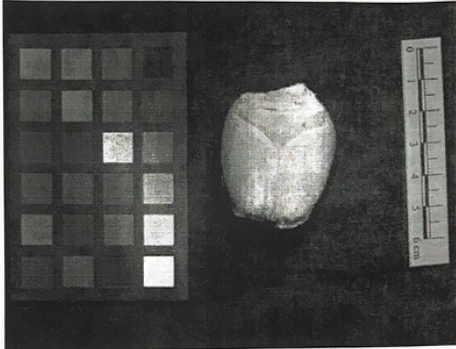
Inventory Number:	20530 (21294)
Current Repository Number:	S/P1005
Location:	N/A *showcase 20. (Location on online mives is corrected)*
Title:	Fragment of Lower Body of a Female Figurine
Dimensions:	Length 3.5cm, Width 2cm, Height 4.5cm
Materials:	Ceramic
Description:	Fragment of female statuette, comprising of just the lower body and thighs. It is carefully polished, with its material being very fine hard fired buff clay with an ochreous slip. The modelling is good and there is little sign of the usual distortion. The buttocks were cut off flat and the width of the upper body reduced by this process to 5 mms.

done.

32 of 46   17/04/2013 08:0

file:///C:/Users/mcamilleri/Desktop/musesvanessa/html1.html
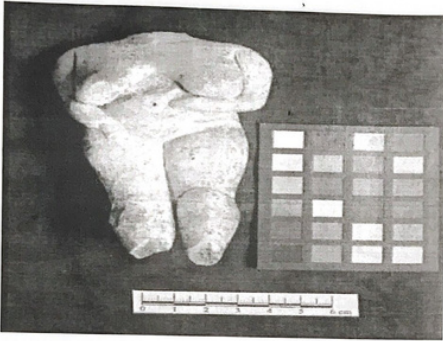
Inventory Number:	21295
Current Repository Number:	Q/P1000
Location:	National Museum of Archaeology, Neolithic Display, Showcase 20
Title:	Venus of Malta
Dimensions:	Length 7cm, Width 4.5cm, Height 13.3cm
Materials:	Ceramic
Description:	A headless standing female statuette consisting of a standing woman made from fine, hard fired buff clay with grey ochre. It seems to have had a polished self slip, and to have been lying in red ochre. The head is missing, broken off at some stage, while the left hand is folded across the abdomen beneath the pendant breasts and the right arm rests on the thigh. The legs are together and broken off at the feet. There is no distortion and modeling is extremely good, especially that of the back.

1 of 1   16/04/2013 15:21

file:///C:/Users/mcamilleri/Desktop/musesvanessa/html.htm
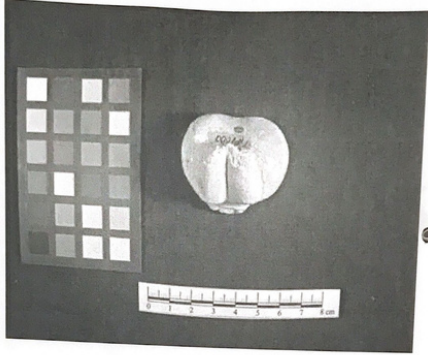
Inventory Number:	9142 (21239)
Current Repository Number:	T/P1000
Location:	National Museum of Archaeology, Neolithic Display, Showcase 18
Title:	Lower Half of a Human Corpulent Figure Statuette
Dimensions:	Height 2.9cm, Width 4.5cm
Materials:	Ceramic
Description:	Corpulent figure (lower part): Lower portion of a small corpulent figure, broken off at the waist and with feet missing. The buttocks are large and flattened at the back, which is highly polished. The lower limbs are fastened to the enormous thighs and come down straight in front. The figurine is of hard-fired grey-buff clay with thick self-slip.

OR W

53 of 77   17/04/2013 07:46

74

Sample Wikidata item

Wikidata item sample as developed for the present research project

Full entry available at https://www.wikidata.org/wiki/Q113559391

Clay statuette with incised lines (Q113559391)

Clay statuette with incised lines from Mnajdra

▸ In more languages

Statements

instance of	sculpture	✎ edit
	▾ 0 references	
		+ add reference
	archaeological find	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
image	National Museum of Archaeology Malta- Inventory number 21254.jpg 1,889 × 1,962; 434 KB	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
movement	prehistoric art	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
country	Malta	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
coordinate location	35°49'36"N, 14°26'11"E	✎ edit
	▾ 0 references	
		+ add reference
		+ add value

made from material	stone	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
depicts	woman	✎ edit
	▾ 0 references	
		+ add reference
	hand	✎ edit
	▾ 0 references	
		+ add reference
	incising	✎ edit
	▾ 0 references	
		+ add reference
	breast	✎ edit
	▾ 0 references	
		+ add reference
	hip	✎ edit
	▾ 0 references	
		+ add reference
	arm	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
collection	Heritage Malta	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
inventory number	21254 collection National Museum of Archaeology	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
exhibition history	National Museum of Archaeology	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
location of discovery	Mnajdra	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
width	6.2 centimetre	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
height	5.8 centimetre	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
thickness	3.1 centimetre	✎ edit
	▾ 0 references	
		+ add reference
		+ add value
		+ add statement

Appendix B

Data Visualisation

- SPARQL Query example from Wikidata Query Service
- Scatter Chart created for data exploration
- Area Chart created for data exploration
- Building your own graph
- Infogram
- Map Visualisation
- RAWGraphs
- From RAWGraphs to Data Art
- Data Art - Animation Test
- Knowledge Graphs

SPARQL Query
From Wikidata Query Service

Available at https://query.wikidata.org

Scatter Chart
Data visualisation option available through the Wikidata Query Service

Instance of : Item name

Depicts P180 colour coded according to the location found

Instance of : Item name

Depicts P180

Instance of : Item name
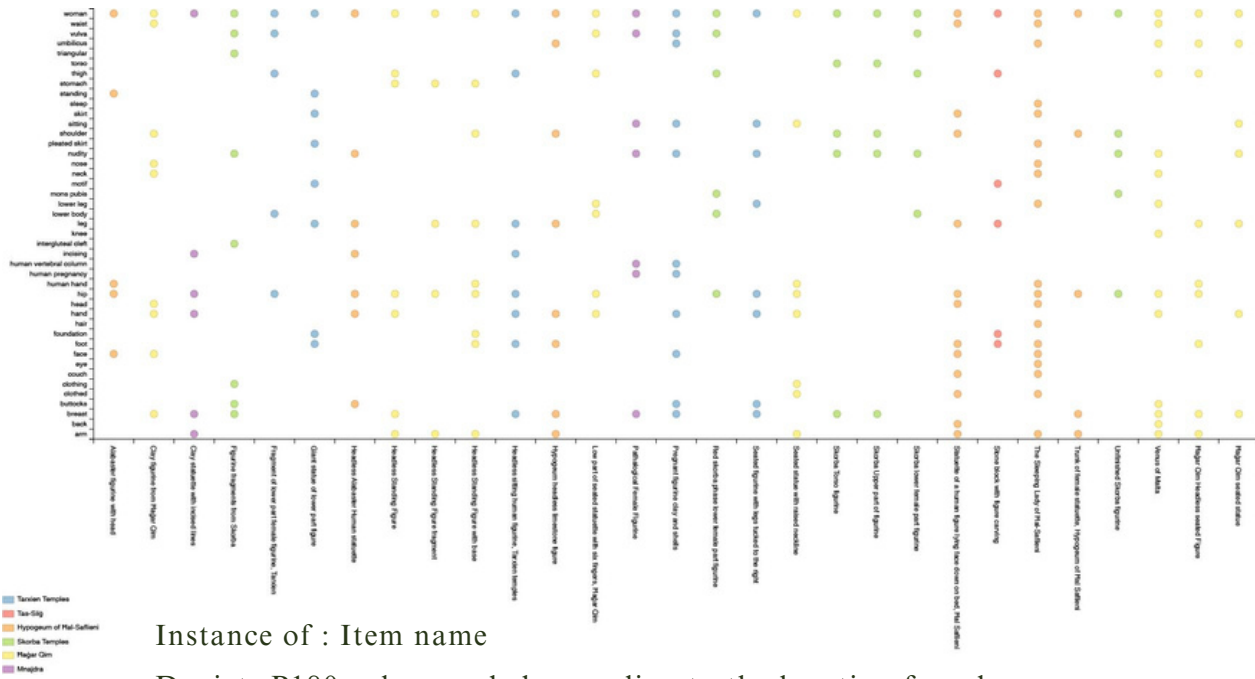
Depicts P180 colour coded according to the material
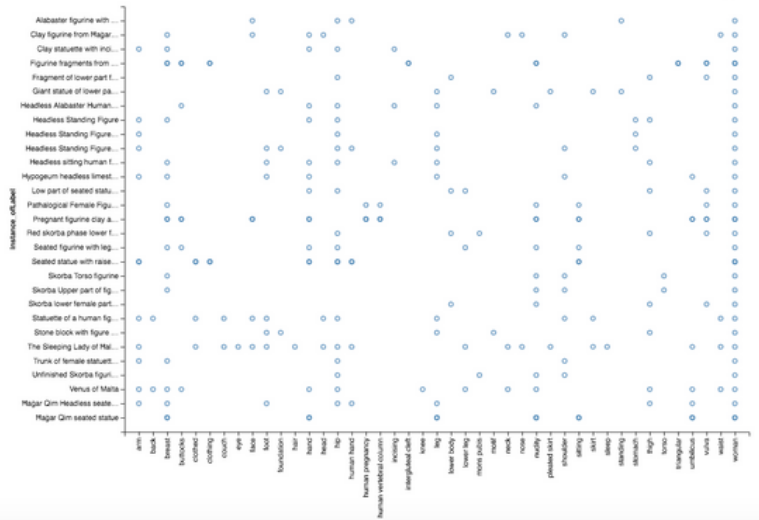
Area Chart

Data visualisation option available through the Wikidata Query Service
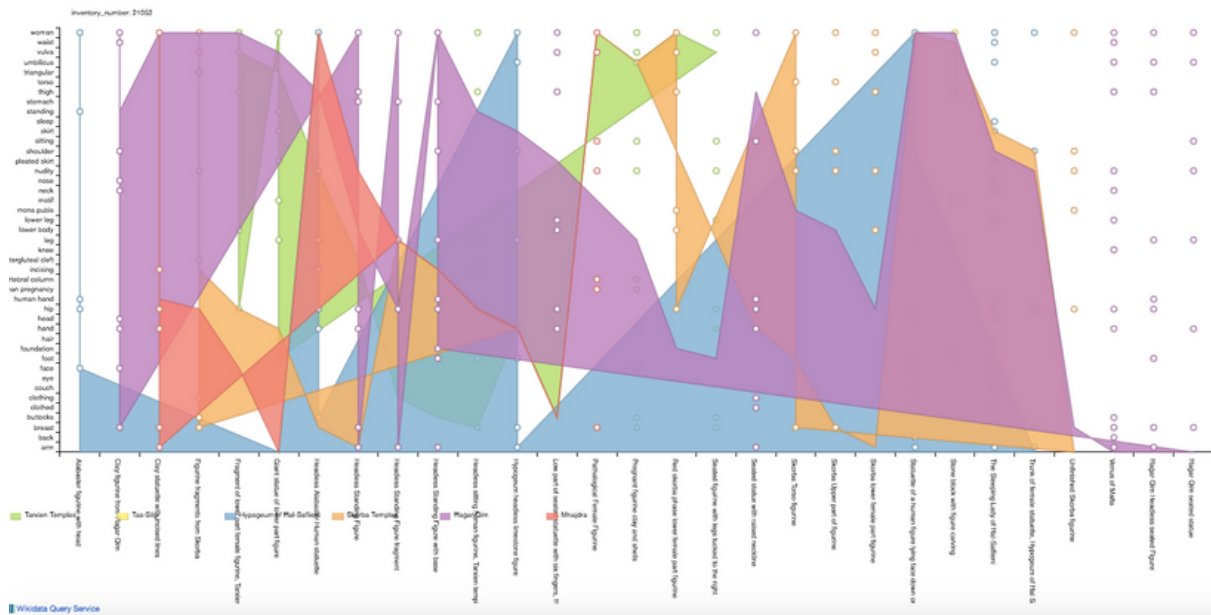
Item name linked with Depicts P180 and colour coded according to location found
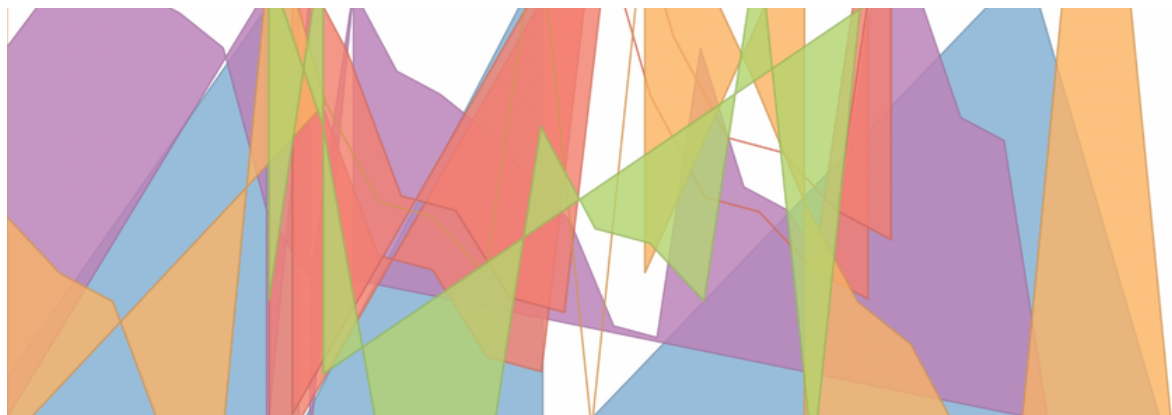
Removing annotations leaves abstract patterns, which can be manipulated artistically

The same patterns can be changed and movement can be recorded according to the selection of findings one chooses to visualise.

Building your own graph
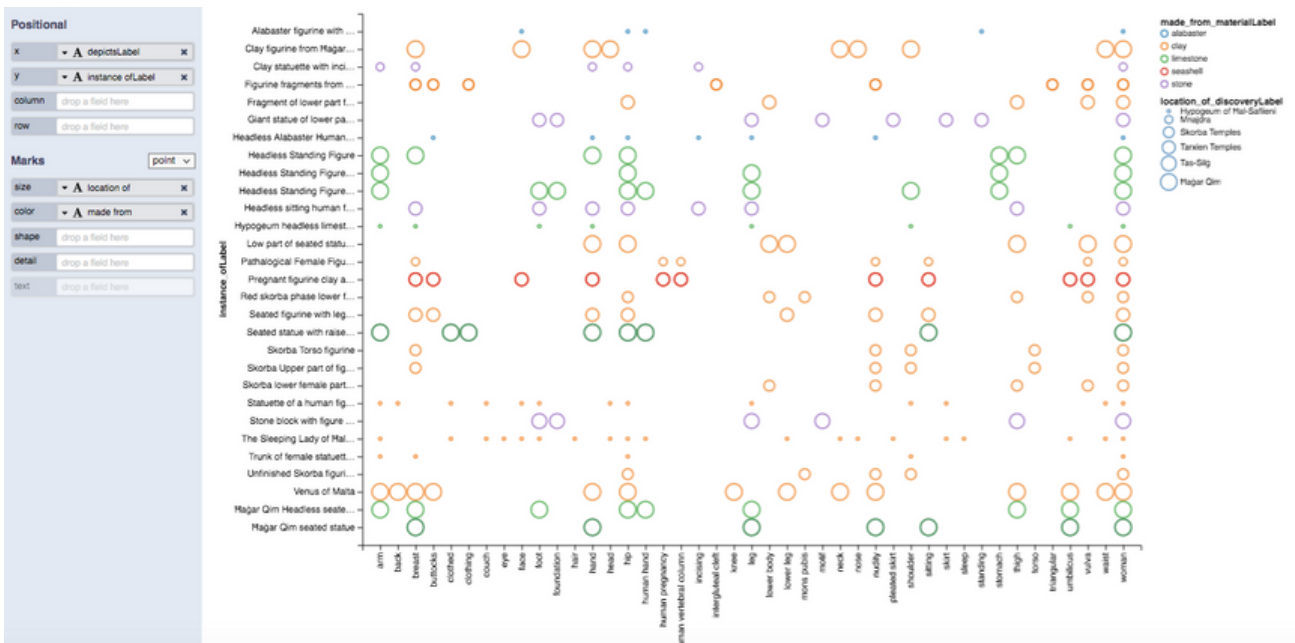
Through the Wikidata SPARQL Query Service you can also build your own graph while colour-coding items and assigning shapes to properties. The colours and shapes can give way towards artistic exploration beyond the meaningly data visualisation
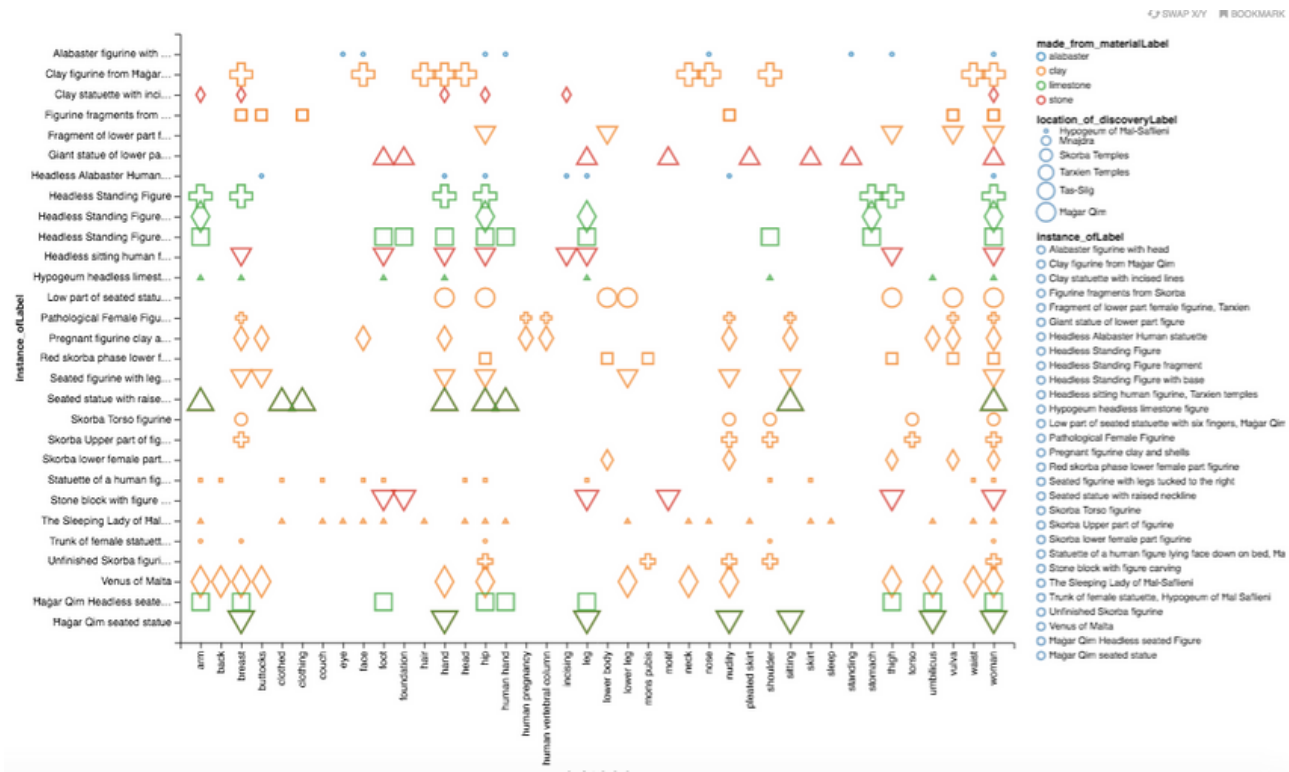
Infogram

Exploration of data visualisation through infographics.

Tool used: https://infogram.com

All these visualisations can be found in this link:

https://infogram.com/infographic-preherstory-in-malta-1h7v4pwyvke786k

By material
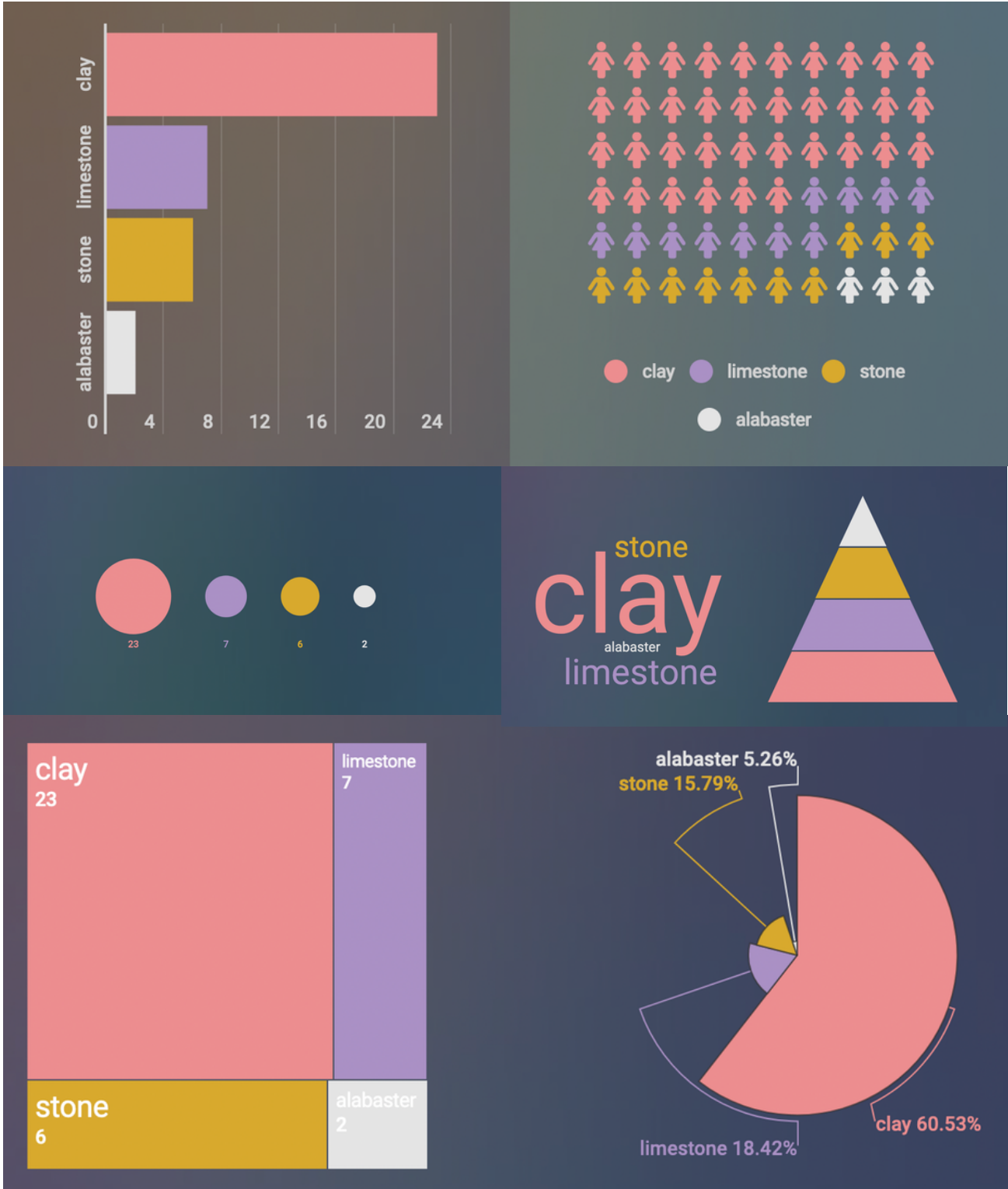
By location

location of discovery	Count of location of discovery
Xaghra Stone Circle	9
Hagar Qim	9
Hypogeum of Hal-Saflieni	6
Skorba Temples	6
Tarxien Temples	5
Mnajdra	2
Tas-Silg	1

9   9   6   6   5   2   1

● Xaghra Stone Circle   ● Hagar Qim   ● Hypogeum of Hal-Saflieni   ● Skorba Temples

● Tarxien Temples   ● Mnajdra   ● Tas-Silg

Tas-Silg 2.63%
Mnajdra 5.26%
Xaghra Stone Circle 23.68%
Tarxien Temples 13.16%
Skorba Temples 15.79%
Hypogeum of Hal-Saflieni 15.79%
Hagar Qim 23.68%
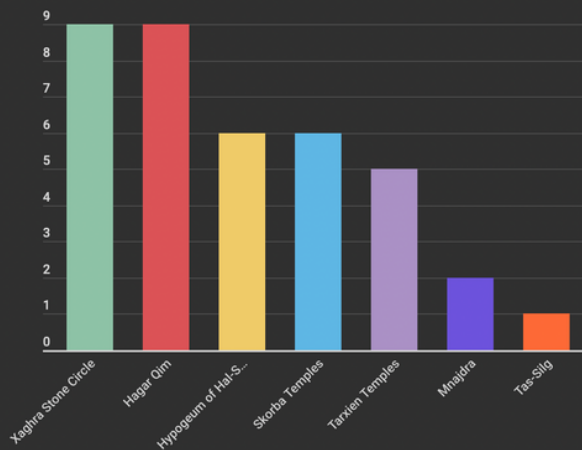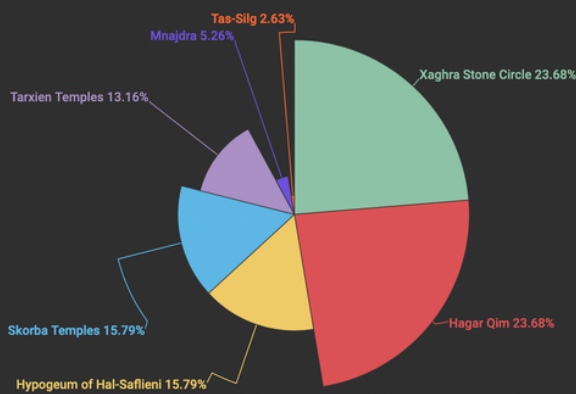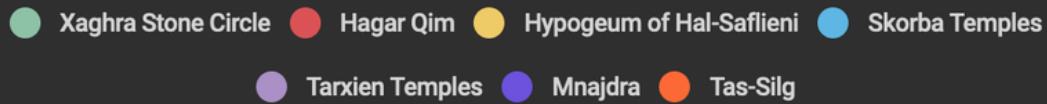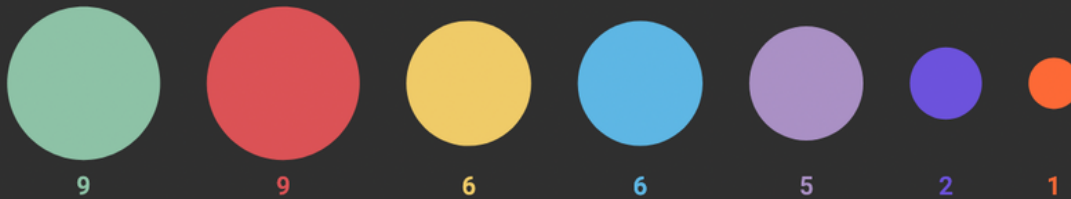
Skorba Temples
Hypogeum of Hal-Saflieni
Xaghra Stone Circle
Tas-Silg
Hagar Qim
Tarxien Temples
Mnajdra

By depicts
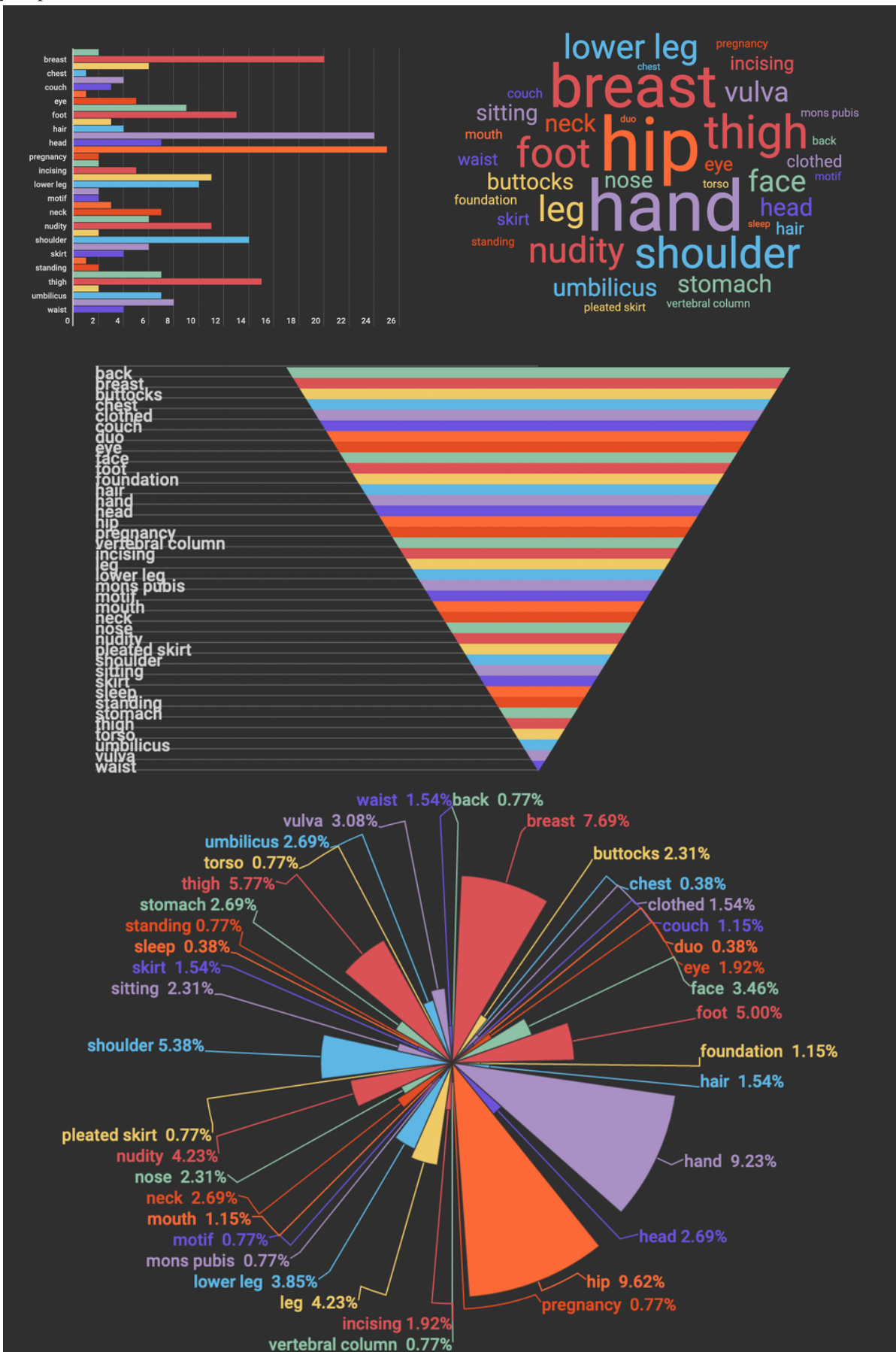
Map Visualisation

Using Datavizdev.fr

http://datavizdev.fr/QueryGraph/index.html?lang=en

RAWGraphs

Exploration of data visualisation through RAWGraphs.

Tool used: https://app.rawgraphs.io

Arc Diagram with RAWGraphs

Everything that depicts
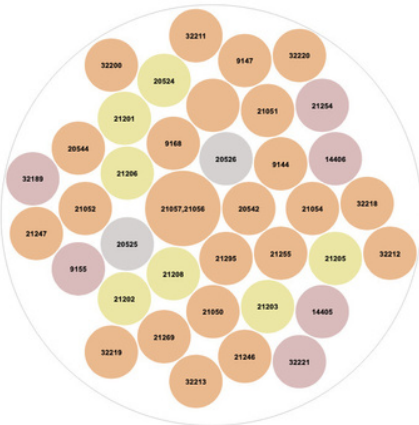
Material and inventory number
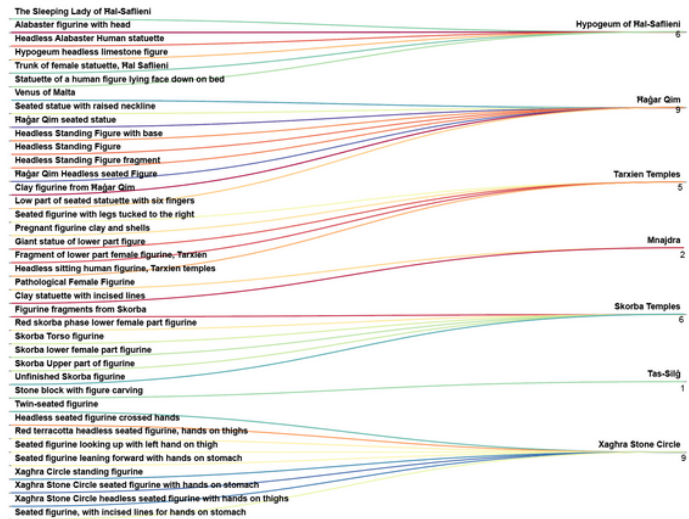
Instance of : Item name linked with location found

From RAWGraphs to Data Art
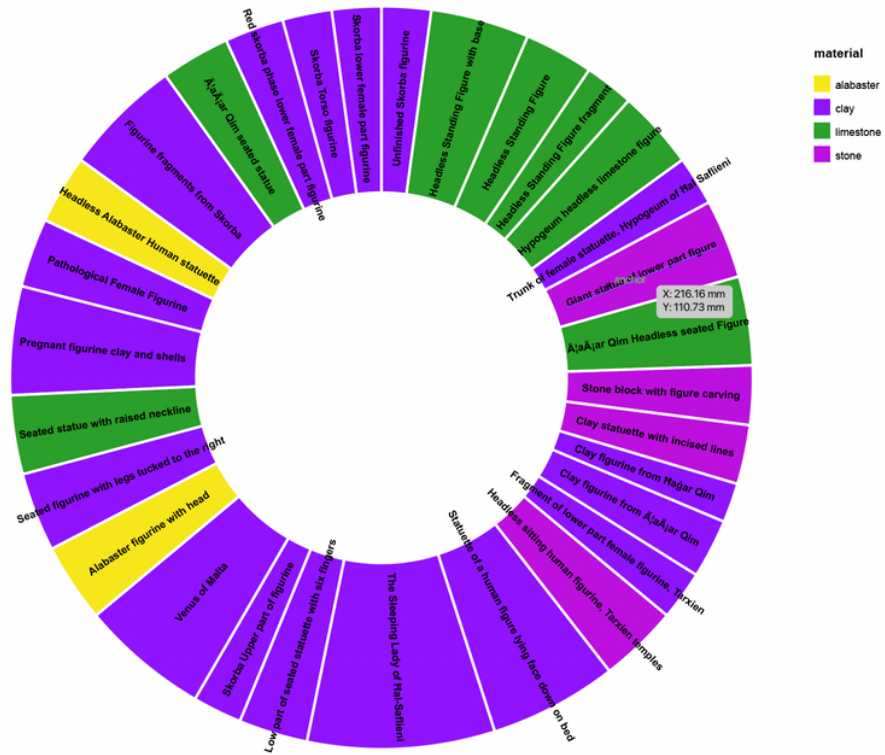
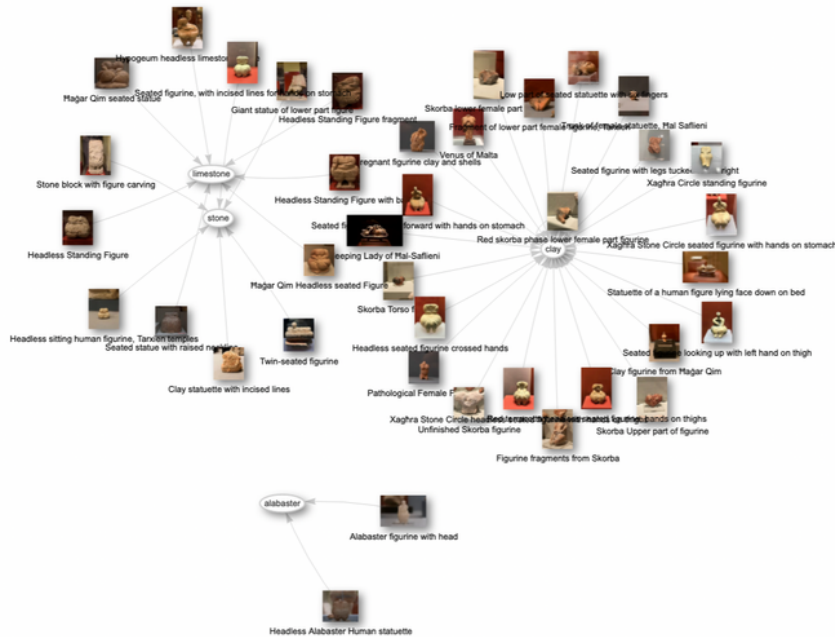# Data Art - Animation test
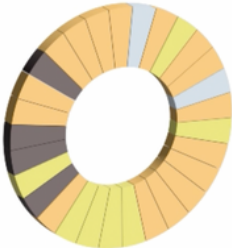
Clay

Limestone

Stone

Alabaster
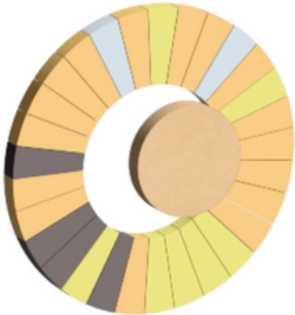
Knowledge Graphs

Generated from Wikidata Query Service
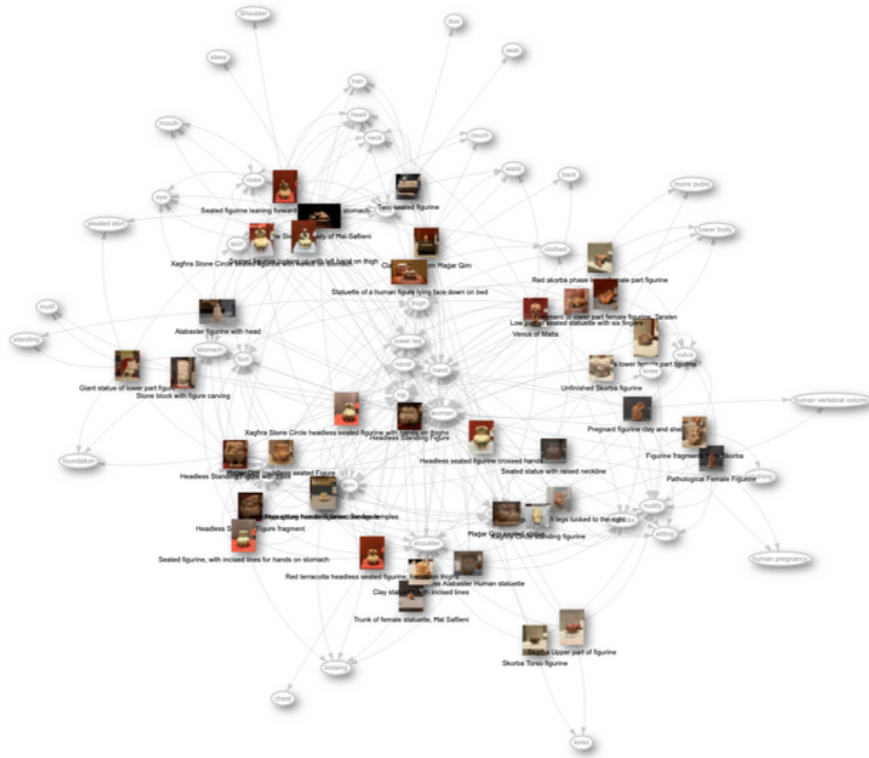
The Wikidata property 'archeological find' linked with what each item's 'depicts' values in Heritage Malta dataset

The Wikidata property 'archeological find' linked with what each item 'made from material' value in Heritage Malta dataset

The property 'archeological find' linked with what each item 'made from material' value across all of Wikidata
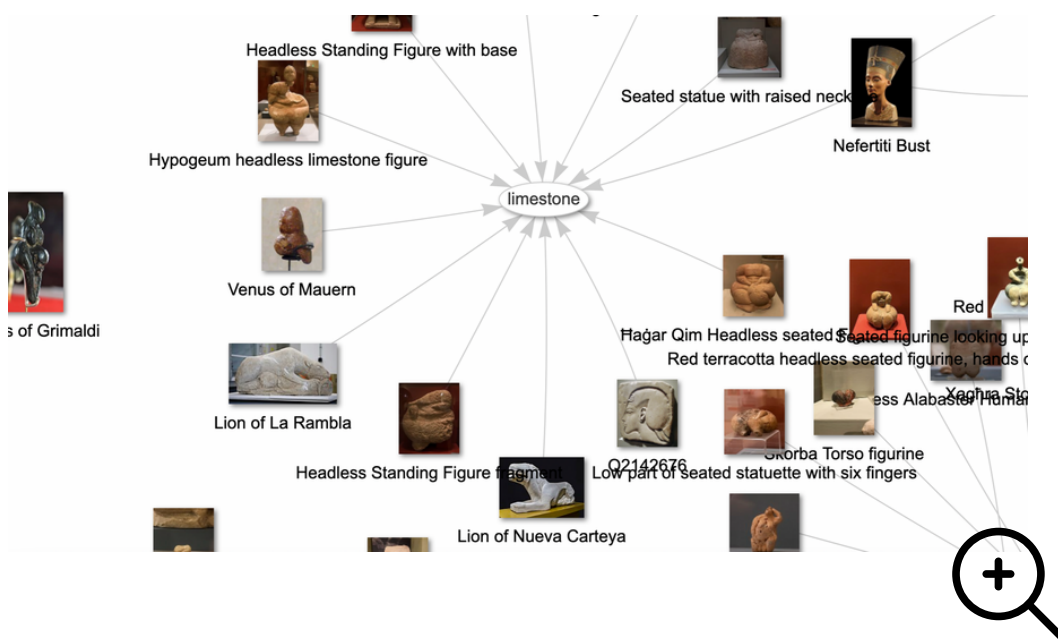
Appendix C

Naked Data portfolio

- Naked Data
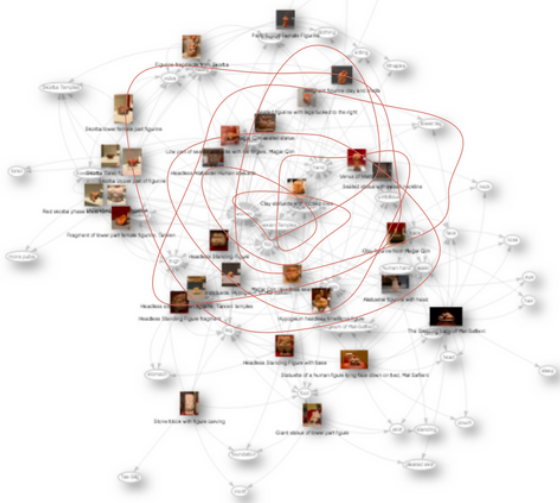- Links to Naked Data digital components

NAKED DATA

Visual documents from process to exhibition
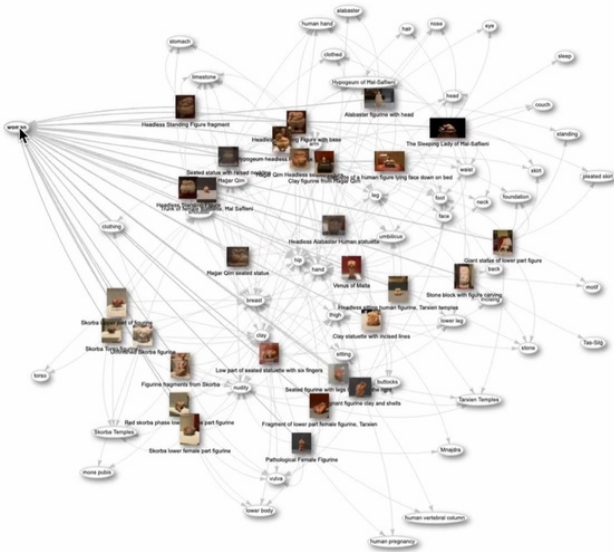
Still image from the Knowledge Graph sequence created for the first segment of the animation video.

The same query was used to develop the plexiglass sculpture.

Red line scribbled through Adobe Illustrator (left) to explore the dynamic flow of the knowledge graph before rendering this into the base of the plexiglass sculpture (right)
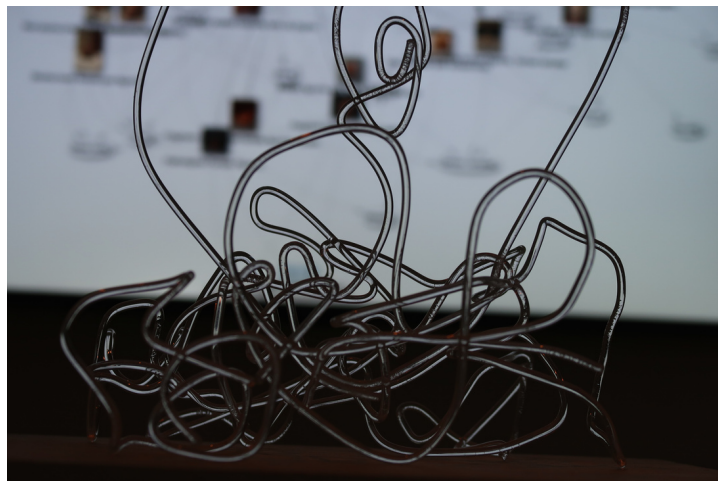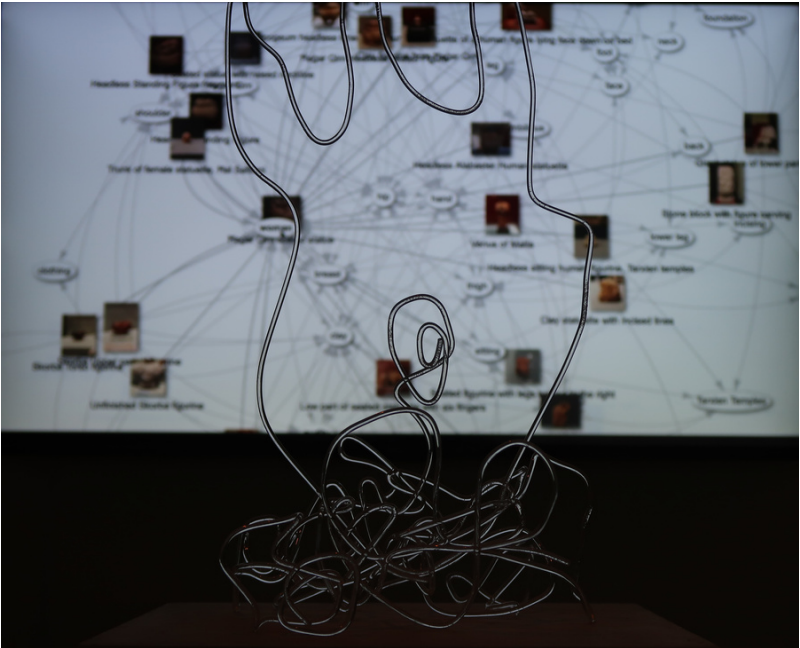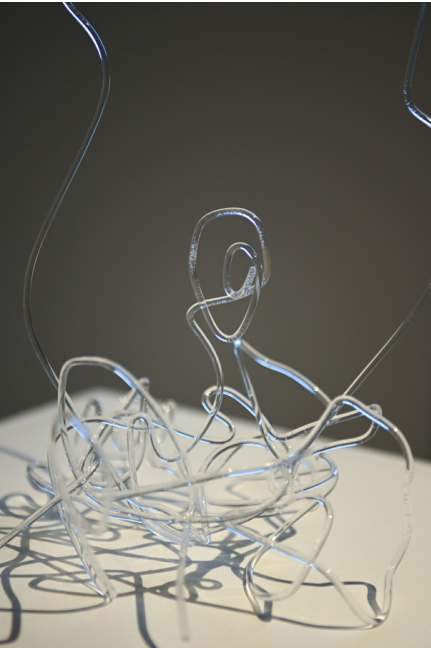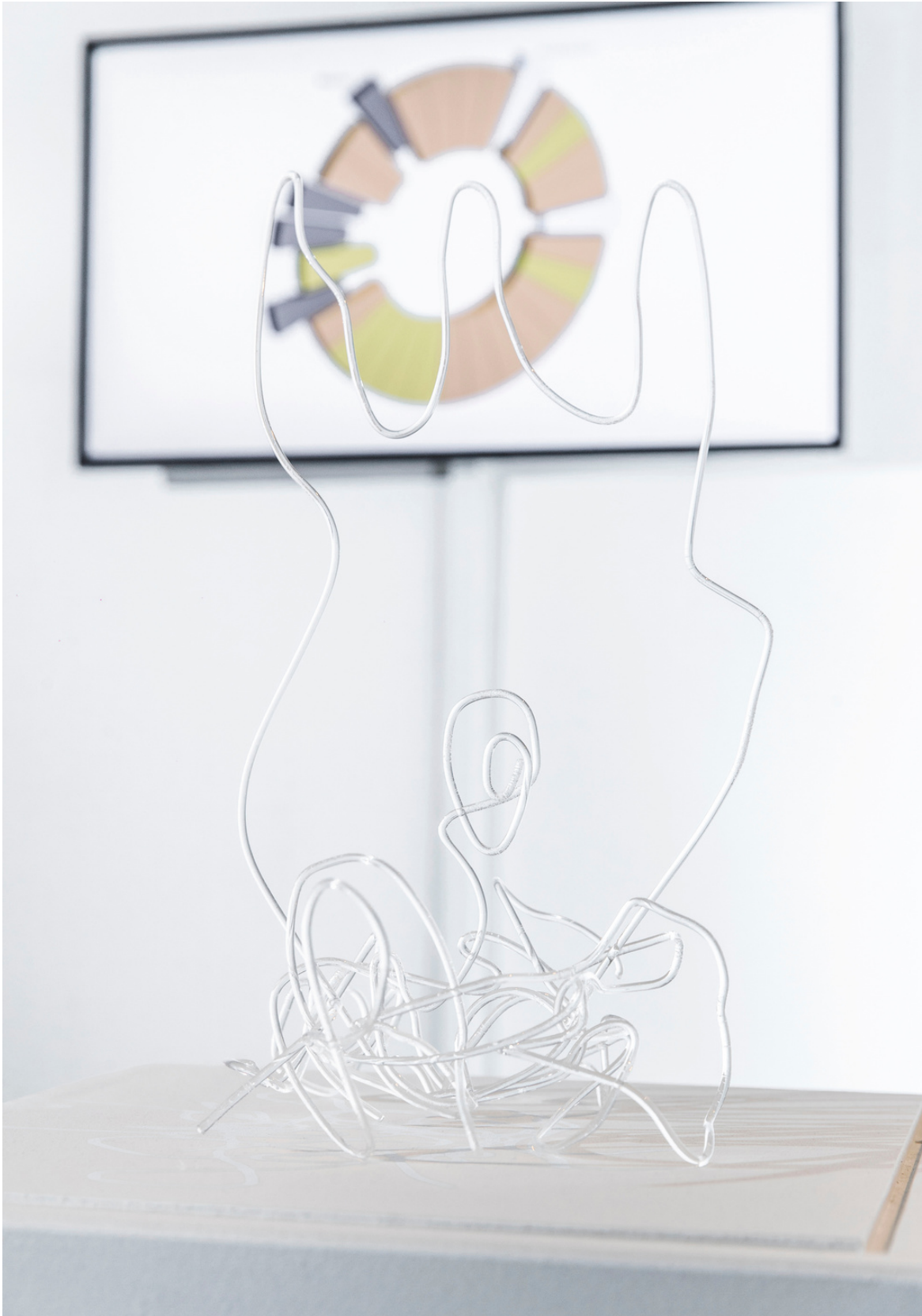
NAKED DATA

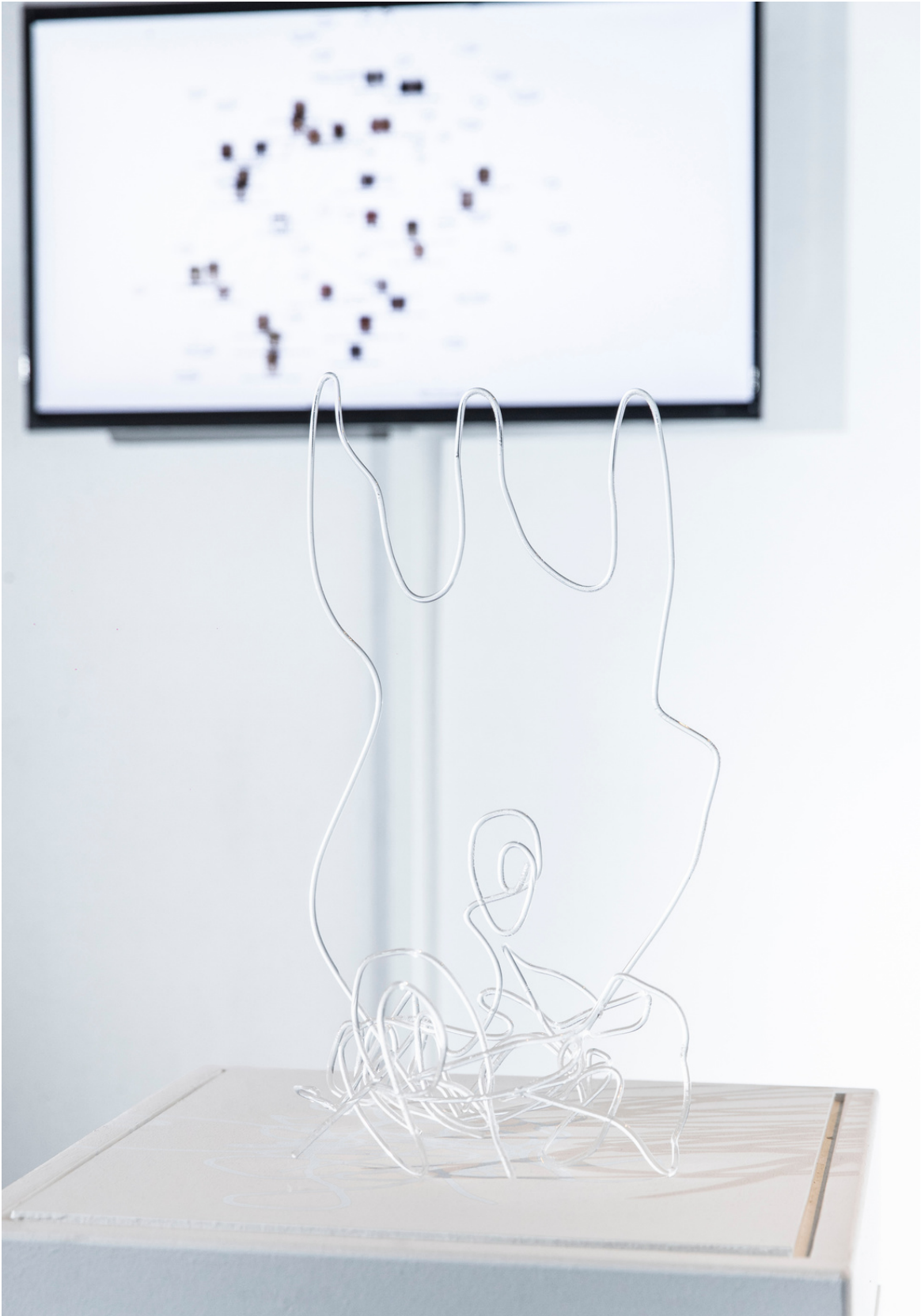As shown at DRHA 2022 exhibition at Stanley Picker Gallery

The plexiglass sculpture is shown with the video animation screened behind it

NAKED DATA

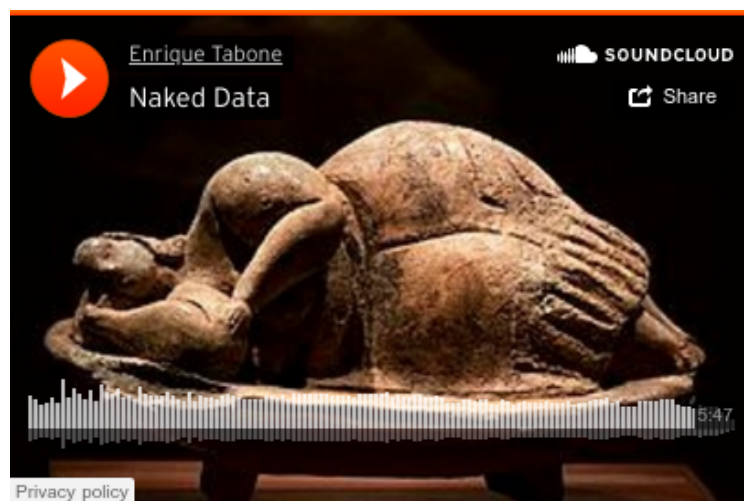# NAKED DATA

NAKED DATA

The *Naked Data* video art component is available to watch on YouTube via this link:

https://youtu.be/pGjNNnO4rzA

Naked Data sound is available to hear on soundcloud via this link:

https://on.soundcloud.com/7o2Sb

List of references

Arnheim, R. (1969). *Visual thinking*. Berkeley, CA : University of California Press.

Arnheim, R. (1974). *Art and visual perception: A psychology of the creative eye* (New version, expanded and rev. ed). University of California Press.

Barr, A. H. (Ed.). (1939). *Picasso: forty years of his art*. New York : The Museum of Modern Art.

Biaggi, C. (1994). *Habitations of the great goddess*. Manchester, CT : Knowledge, Ideas & Trends.

Bishop, C. (2023). *Information Overload*. https://www.artforum.com/print/202304/claire-bishop-on-the-superabundance-of-research-based-art-90274

Bonanno, A. (2010). Women and society in prehistoric and ancient Gozo. In J. F. Grima (Ed.), *60th anniversary of the Malta historical society: A commemoration* (pp. 61–72). The Malta Historical Society.

Chan, S. (2007). *Tagging and Searching–Serendipity and museum collection databases*. Museums and the Web. https://www.archimuse.com/mw2007/papers/chan/chan.html

Cirotteau, T., Kerner, J., Pincas, E., & Gaussein, P. (2022). *Lady sapiens: Breaking stereotypes about prehistoric women* (P. Hurd, Trans.). London: Hero.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed). Thousand Oaks, CA : SAGE Publications.

Criado-Perez, C. (2020). *Invisible women: Exposing data bias in a world designed for men*. London : Vintage.

Delisle, R. (2016, August 22). *Amplifying Data Through Art*. Hakai Magazine. https://hakaimagazine.com/article-short/amplifying-data-through-art/

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. Cambridge, Mass : The MIT Press.

Drucker, J. (2013). Performative Materiality and Theoretical Approaches to Interface. *DHQ: Digital Humanities Quarterly*, *7*(1).

Duarte, N. (2019). *Data story: Explain data and inspire action through story*. Oakton, Virginia : Ideapress Publishing.

Edge, A. (2015, May 8). *Data visualisation tips from Information is Beautiful*. Journalism.co.uk. https://www.journalism.co.uk/news/data-visualisation-tips-from-information-is-beautiful/s2/a565043/

Escobar, P., Candela, G., Trujillo, J., Marco-Such, M., & Peral, J. (2020). Adding value to Linked Open Data using a multidimensional model approach based on the RDF Data Cube vocabulary. *Computer Standards & Interfaces*, *68*, 103378. https://doi.org/10.1016/j.csi.2019.103378

Eve, M. P., & Suber, P. (2014). *Open access and the humanities: Contexts, controversies and the future*. Cambridge University Press.

Frick, L. (2015, May 15). Personal data beginning to feel less sinister. *Laurie Frick - Blog*. https://www.lauriefrick.com/blog/data_less_sinister

Friendly, M. (2008). The Golden Age of Statistical Graphics. *Statistical Science*, *23*(4), 502–535. https://doi.org/10.1214/08-STS268

Healy, K. (2018). *Data visualization: A practical introduction*. Princeton University Press.

Hessel, K. (2022). *The story of art without men*. Hutchinson Heinemann.

Hill, B. M., & Shaw, A. (2013). The Wikipedia Gender Gap Revisited: Characterizing Survey Response Bias with Propensity Score Estimation. *PLoS ONE*, *8*(6), e65782. https://doi.org/10.1371/journal.pone.0065782

Hohl, M. (2011). From abstract to actual: Art and designer-like enquiries into data visualisation. *Kybernetes*, *40*(7/8), 1038–1044. https://doi.org/10.1108/03684921111160278

Kerracher, N., Kennedy, J., & Chalmers, K. (2014). The Design Space of Temporal Graph Visualisation. *EuroVis - Short Papers*. https://doi.org/10.2312/EUROVISSHORT.20141149

Koch, T. (2004). The Map as Intent: Variations on the Theme of John Snow. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *39*(4), 1.

Koren, M. (2013, March 1). *Transforming Raw Scientific Data Into Sculpture and Song*. Smithsonian Magazine. https://www.smithsonianmag.com/science-nature/transforming-raw-scientific-data-into-sculpture-and-song-235223/

Kräutli, F., & Davis, S. B. (2013, July 1). *Known Unknowns: Representing Uncertainty in Historical Time*. Electronic Visualisation and the Arts (EVA 2013). https://doi.org/10.14236/ewic/EVA2013.16

Lupi, G. (2017, July 10). Data Humanism, the Revolution will be Visualized. *Medium*. https://medium.com/@giorgialupi/data-humanism-the-revolution-will-be-visualized-31486a30dbfb

Lupi, G., & Posavec, S. (2016). *Dear data*. London: Particular Books.

Malone, C., & Stoddart, S. (2017). Figurines of Malta. In T. Insoll (Ed.), *The Oxford Handbook of Prehistoric Figurines* (pp. 729–754). Oxford University Press.

Manovich, L. (2015). Data Science and Digital Art History. *International Journal for Digital Art History*, *No 1*. https://doi.org/10.11588/DAH.2015.1.21631

McCandless, D. (2010). *The beauty of data visualization*. TED Talks. https://www.ted.com/talks/david_mccandless_the_beauty_of_data_visualization

McCandless, D. (2012). *Information is beautiful*. London : Collins.

Nash, A. (2014). Interference Wave Data and Art. *Leonardo Electronic Almanac*, *20*(2), Article 2.

Nelson, R. (2013). *Practice as research in the arts: Principles, protocols, pedagogies, resistances*. New York : Palgrave Macmillan.

Odell, J., Lemus-Rojas, M., & Brys, L. (2022). *Wikidata for Scholarly Communication Librarianship*. https://doi.org/10.7912/9Z4E-9M13

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group.

Open Knowledge Foundation. (n.d.). *Open Definition 2.1—Defining Open in Open Data, Open Content and Open Knowledge*. Retrieved 18 May 2023, from https://opendefinition.org/od/2.1/en/

Pace, A. (Ed.). (1996). *Maltese prehistoric art, 5000-2500 BC*. Malta : Fondazzjoni Patrimonju Malti in association with the National Museum of Archaeology.

Prophet, J. (2010, February 1). *Models, Maquettes and Art Objects: Making Data Physical*. CAT 2010: Ideas before their time : Connecting the past and present in computer art. https://doi.org/10.14236/ewic/CAT2010.17

Redi, M., Gerlach, M., Johnson, I., Morgan, J., & Zia, L. (2020). *A Taxonomy of Knowledge Gaps for Wikimedia Projects (Second Draft)*. https://doi.org/10.48550/ARXIV.2008.12314

Sanderhoff, M. (2013). Open Images. Risk or opportunity for art collections in the digital age? *Nordisk Museologi*, *2013*(2), 131–146. https://doi.org/10.5617/nm.3083

Sant, T. (2023). *Enriqué Tabone, Catalogue Raisonné*. Malta : Kite Group.

Silver, N. (2012). *The signal and the noise: Why so many predictions fail–but some don't*. New York : Penguin Press.

Small, H. (2001). *Florence Nightingale: Avenging Angel*. London : Constable & Robinson.

Strickler, C. M. (2021). Mind the Wikidata Gap: Why You Should Care About Theological Data Gaps in Wikipedia's Obscure Relative, and How You Can Do Something About It. *Atla Summary of Proceedings*, 301–314. https://doi.org/10.31046/proceedings.2021.2978

Suber, P. (2012). *Open access*. Cambridge, Mass : MIT Press.

Sullivan, G. (2010). *Art practice as research: Inquiry in visual arts* (2nd ed). Thousand Oaks, CA : Sage Publications.

Sydell, L. (2018, February 25). An Artist Sees Data So Powerful It Can Help Us Pick Better Friends. *NPR*. https://www.npr.org/sections/alltechconsidered/2018/02/25/583682718/an-artist-sees-data-so-powerful-it-can-help-us-pick-better-friends

Thornton, K. (2021). *Wikidata for Digital Preservationists*. Digital Preservation Coalition. https://doi.org/10.7207/twgn21-19

Thudt, A., Hinrichs, U., & Carpendale, S. (2012). The Bohemian Bookshelf Supporting Serendipitous Discoveries through Visualization. *Proedings of the SIGCHI Conference on*

Human Factors in Computing Systems (CHI), 1461–1470.
http://dx.doi.org/10.11575/PRISM/30518

Trump, D. H. (2002). *Malta, prehistory and temples*. Malta : Midsea Books.

Tufte, E. R. (1990). *Envisioning Information* (First Edition). Cheshire, Conn : Graphics Press.

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed). Cheshire, Conn :
Graphics Press.

Ucko, P. J. (1962). The Interpretation of Prehistoric Anthropomorphic Figurines. *The Journal of
the Royal Anthropological Institute of Great Britain and Ireland*, *92*(1), 38.
https://doi.org/10.2307/2844320

Vella Gregory, I. (2005). *The human form in Neolithic Malta* (1. ed). Mata : Midsea Books.

Vella, N. (2007). From Cabiri to Goddess. In D. Borrowclough & C. Malone (Eds.), *Cult in
Context* (pp. 61–77). Oxford : Oxbow Books.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A.,
Blomberg, N., Boiten, J.-W., Da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes,
A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., …
Mons, B. (2016). The FAIR Guiding Principles for scientific data management and
stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

Windhager, F., Federico, P., Schreder, G., Glinka, K., Dork, M., Miksch, S., & Mayr, E. (2019).
Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges.
IEEE Transactions on Visualization and Computer Graphics, *25*(6), 2311–2330.
https://doi.org/10.1109/TVCG.2018.2830759

Worrall, D. (2018). Sonification: A Prehistory. *Proceedings of the 24th International Conference
on Auditory Display - ICAD 2018*, 177–182. https://doi.org/10.21785/icad2018.019