**Norwegian artist names**
CIDOC 2018 Heraklion, Crete, Greece

---

**Norwegian artist names**

**Authority list of artists in Norwegian art collections**

Yvonne Brenden Hansen, Dag Hensten, Gro Benedikte Pedersen and Magnus Bognerud

Nasjonalmuseet for kunst, arkitektur og design
Pb. 7014 St. Olavs plass, N-0130 OSLO
yvonne.hansen@nasjonalmuseet.no

**Keywords:** Name authorities, digitisation, from paper to digital, data cleaning, data reconciliation, interoperability, linked data.

---

**Abstract**

How do you best transform a paper-based publication into a living online resource? This is the theme of an ongoing project at The National Museum of Art, Architecture and Design in Norway, supported by the Arts Council Norway. The goal of the project is to create, publish and maintain an authority list of Norwegian artists, architects, designers and craftsmen. The list will in part be based on the Norsk Kunstnerleksikon [Encyclopaedia of Norwegian Artists], published 1982-86 and subsequently digitised in 2013. With the help of other public collections in Norway, we aim to make the new resource as complete as possible, and available in both human- and machine-readable formats. By doing this we hope to ease the digitisation process and contribute to better data quality in Norwegian online collections. Although the original paper publication contains biographical texts as well as lists of exhibitions, education, travels, publications and more, the data in the new authority list will be constrained to a set of core biographical data. It will however carry references to online biographical resources such as Norsk Kunstnerleksikon (NKL), Wikidata, Union List of Artist Names (ULAN) and Virtual International Authority File (VIAF).

In our paper we will discuss the process of defining the scope and setting constraints for the list, how to enrich and reconcile existing data, as well as strategies to ensure other institutions contribute both as content publishers and end users.

---

1 **Introduction**

Man has forever been seeking to accumulate knowledge and of finding ways to impart that knowledge to his or her descendants. From oral tradition to the great libraries of the ancient world, from Pliny’s *Naturalis Historia* to Wikipedia we have been gathering knowledge and crafted constructs to disseminate it. The printed encyclopaedia is perhaps the quintessential manifestation of such an obsession. However educated and refined, Man’s faculties come up short when it comes to holding the wisdom of the world, and a library, however vast and exhaustive was inaccessible to most, but a book (or set of books) was something that could, at least in principle, be universally attainable. This was the realisation and fervour at the spread and industrialisation of the printing press. This ‘encyclopaedic vision’, as the historian Richard Yeo (2001) has called it, was the idea that one could collate and create “a work containing the collective knowledge of a community” that could be spread, consulted and also, if the worst came to pass it could also serve as a “summary of intellectual accomplishments to be reactivated at a later stage […] to be put together again if all other books were lost” (3). As shown in this quote by Denis Diderot, co-editor of one of the first modern encyclopaedias, this undertaking was viewed with the loftiest of ambition:

> “Indeed, the purpose of an encyclopaedia is to collect knowledge disseminated around the globe; to set forth its general system to the men with whom we live, and transmit it to those who will come after us, so that the work of preceding centuries will not become useless to the centuries to come; and so that our offspring, becoming better instructed, will at the same time become more virtuous and happy, and that we should not die without having rendered a service to the human race.” (Diderot and d’Alembert 2002)

Diderot viewed his (and his co-editor d’Alembert’s) contribution as a “universal dictionary” a distillation of all the worlds specialist and thematic dictionaries, lexica and vocabularies into one single work (albeit divided into
seventeen volumes). The modern encyclopaedia would not only be a source of knowledge, but a hub of knowledge.

The production and spread of encyclopaedias, both universal and thematic, would continue with undiminished energy into the twentieth century, published in every conceivable language and relating to any imaginable topic. Most of these were in Richard Yeo’s (2001) words “seeking to record knowledge, not lives” (16), they did not include history, people and biographies as this was the domain of a specific sub type of reference work – the historical, or biographical, dictionary. One such effort was the Norsk Biografisk Leksikon [Norwegian Biographical Encyclopaedia] which was published in nineteen volumes between 1923 and 1983 (with a revised second edition published between 1999 and 2005). In its preface it is stated that:

“The Norwegian Biographical Encyclopaedia shall contain biographies of all Norwegians that in one way or another have made themselves more than commonly known in the country – not merely in a single region – without consideration to whether or not their ventures have been deserving or not; there lies in the selection no other considerations than the biographical importance, none towards his ventures more or less fortunate or unfortunate character.” (Bull, Krohg and Gran 1923)1

This naturally also included the arts, but as an all-encompassing biographical reference work there was scarce space for lesser known or obscure artists, designers or architects. Also, as seen from the long timespan between the volumes in the first edition, many artists and craftsmen were excluded due to the timing of the period of their activity. The initiative thus arose to create a separate publication for Norwegian artists. Work commenced in 1977 as a joint effort by the National Gallery (Nasjonalgalleriet) and the Directorate for Cultural Heritage (Riksantikvaren) through funding from the Research Council (Norsk forskningsråd) and the Arts Council (Norsk Kulturråd). From 1982 to 1986 four volumes of the Norsk Kunstnerleksikon (Østby 1982-1986), were published containing more than 3000 articles, long and short.

Fig.1 Norsk Kunstnerleksikon 1982-1986

1 Original text: “[...] skal Norsk Biografisk Leksikon indeholde biografier av alle nordmænd som paa en eller anden maate har gjort sig mere almindelig kjendt i landet – altsaa ikke bare i et enkelt distrikt – uten hensyn til om virksomheten har været fortjenstfuld eller ei; der ligger i utvalget ingen anden vurdering end av den biografertes betydning, ikke av hans virksomhets mer eller mindre heldige eller uheldige karakter.”
The idea was that the encyclopædia would be a “living” reference work, continuously updated, and thoughts of a digital edition were already present in the early 90’s. But as often happens, funding dried up and the involved institutions did not have the resources needed to keep the work going. Hence the content was never updated, and revised editions never saw the light of day.

That is, in a sense, until the digital revolution of the 2000’s. Norway was going digital, arts and culture had to follow, and funding was suddenly available for “all” projects that included digitisation. Funding was provided by (the now defunct) Archive, Library and Museum Authority (ABM-utvikling or Statens senter for arkiv, bibliotek og museum) and the Arts Council. The project was completed in two phases: First from 2007-2012, work was done to transfer and transform all data from the paper source to a semantically structured machine-readable database, published as RDF (Data Nasjonalmuseet n.d). When this work was concluded it was decided that a human readable web-interface would be desirable as well and an agreement was reached with the online edition of the Store Norske Leksikon [the Great Norwegian Encyclopædia] (SNL 2018) to publish the Norsk Kunstnerleksikon within their own publishing framework. The online version (NKL.SNL n.d) was unveiled in 2013. As many others will probably have experienced though, getting funding for a digitisation project is one thing, getting funding for editorial work, maintenance and updates is quite another.

![Fig.2 Screenshot of the web interface of Norsk Kunstnerleksikon](image)

As such, the content still invariably remains the same as it was when the last paper based volume was released in 1986. The date of death of persons deceased since 1986 have been added, but the biographical texts are untouched, and no new entries have been made to the encyclopædia. Unfortunately, neither the National Museum (who is now the rights holder to Norsk Kunstnerleksikon) nor the Store Norske Leksikon have resources to continue the editorial work. In contrast to what some hoped, digitising does not make your material self-sufficient, as Georg Kjell (2017) writes: “Putting a paper-based work online is not a solution to this

---

2 A SPARQL endpoint is available (Data Nasjonalmuseet n.d). This dataset is not updated.
challenge, it is the start of a solution. Web publishing in a framework which facilitate the update of content is a prerequisite of that work, but the content demands just as much professional and editorial processing to be as good as it did when it was on paper.\footnote{Original text: “Å legge papirverk på nett er ikke en løsning på denne utfordringen, det er begynnelsen på en løsning. Nettutgivelse i et rammeverk som muliggjør oppdatering av innhold er en forutsetning for arbeidet. Men innholdet krever like mye faglig og redaksjonell bearbeidelse som det trengte på papir for å være bra.”} In many instances the information required is already readily accessible from other sources, but it must be collated. How then, do you go about solving this problem, if a solution exists at all? How can you update and maintain an encyclopaedia, or any information really, when that requires the exact thing you do not have – resources?

In a way this conundrum is nothing new, a century ago, with the creation of ever more editions and variants of encyclopaedias and the like, some came to conclude that in a way one was back to square one; with knowledge from differing fields, cultures and languages collected in different publications, how would one source out the most relevant ones? What was required was some sort of central, universally accessible deposit. The writer H.G. Wells (1938, 63) stated it in this way:

“It is dawning upon us, we lay observers, that this work of documentation and bibliography, is in fact nothing less than the beginning of a world brain, a common world brain. What you are making me realise is a sort of cerebrum for humanity, a cerebral cortex which (when it is fully developed) will constitute a memory and a perception of current reality For the entire human race. Plainly we have to make it a centralised and uniform organisation but […] it need not have any single local habitation […] In these days of destruction, violence and general insecurity, it is comforting to think that the brain of mankind […] can exist in numerous replicas throughout the world.”

In this day and age of the world wide web, the internet, have we not created this “world brain”? Perhaps not centralised and uniform, but more infinite than even Wells could have conceived of. Herein lies the challenge, not of producing content – the world brain already has the content, but of structuring that content and being able to retrieve it at the right time. Accordingly, we will not be creating a new edition of the Norsk Kunstepnymesik, rather we will create an authority list of artists, architects, designers and craftsmen present in Norwegian public arts collections. An index of names linked to further information elsewhere. With the help of other public collections in Norway, we aim to make the new resource as complete as possible, and available in both human- and machine-readable formats. Although the original paper publication contains biographical texts as well as lists of exhibitions, education, travels, publications and more, the data in the new authority list will be constrained to a set of core biographical data. A connection of open information from disparate sources linked through name authorities. For this task funding has again been provided by the Arts Council Norway.

2 Background

Over the last decade The National Museum has invested considerable effort into improving the digital management of its art collection information. Working in the professional area of digital collection management, we aim to provide consistency, clarity and context for our digital catalogue - and are driven by the intention of ‘Create Once Publish Everywhere’ (COPE).\footnote{A concept strategy made well known to the museum community e.g. by Nick Poole, as chair of the Collections Trust in Britain. He describes the maxim as born – “coined in a blog post by Daniel Jacobson, Director of Application Development for National Public Radio in the US – and some powerful new work being done by the British Museum as part of their ResearchSpace project" (Poole 2016).} Like many other museums in the 2010s, we worked with specifying new conditions for digital information and analysing new types of metadata and layers of structure. During our different digitising projects, we benefited from best practices and deliverables from others, like the CIDOC committee’s standards and guidelines, the Getty’s Online Scholarly Catalogue Initiative or EU Program funded projects like Digitising Contemporary Art (DCA) and Terminologies as a Service (TaaS).

In 2012 the Arts Council Norway funded the national terminology platform KulturNav (kulturnav.org), initiated by the software company Kulturt\textsuperscript{3}. KulturNav aimed to facilitate the management and delivery of

\textsuperscript{3} Kulturt is a provider of a collection management system and other IT infrastructure for museums in Norway and Sweden. Originally a unit within the public ALM authorities in Norway, it was later transformed into a private company and is now
shared terminology, and thereby paved the way for establishing new online authority lists in the Norwegian museum sector. Together with KulturIT and the art museum community The National Museum started working with standardising (art history) terminology like object types, materials and techniques. We linked as much as possible to online resources (like The Art & Architecture Thesaurus) and published our lists on the web as authorities. Regarding artist names, one of the fundamental provenance information elements: the creator identity⁶, we had no more than linked the names of foreign (non-Norwegian) painters and sculptors in our collection to VIAF and created an authority list of our museum’s preferred spelling of them. Throughout the museum’s digitisation projects (2008-2018), the task of establishing a creator identity authority list for Norwegian artists was put on hold. We waited for resources and better technical solutions and hoped for the important work of the scholarly prestigious Norsk Kunstnerleksikon to be continued. Until the start of this project we had no up-to-date national, nor internal, authority list of identities to use for the digitised works of artists, architects, designers and craftsmen in our collections. The artist encyclopaedia has been our only authority list of Norwegian artists names, even though we never technically integrated it with our collections management system.

Starting our project, we wanted to concentrate on the areas that the Norsk Kunstnerleksikon is not covering. Since the encyclopaedia never has been linked to other name identity authorities we wanted to link all identities, both coming from the encyclopaedia and from this project, to already established authorities like ULAN and VIAF. During this decade of emphasising the digitisation and publishing of the collections, the expectations regarding scope, function and format of a list of artist names has changed in various ways.

3 Scope
How to define the scope of our authority list? As starting point, we had a list of more than 28 000 disorderly documented name entries in our internal collection management system. The names represented persons, companies, institutions, factories and (even) mythological figures, and was the accumulated result of 20 years of cataloguing. The list had been created by everyone who worked with our collections that had access to creating new records. Many of the records were also a result of digitisation of old catalogue cards. We needed to streamline the content, but what information elements had to be included in our project, and which would be redundant? Our first structural decision was narrowing down the concept Artist from the encyclopaedia’s definition to: ‘any person with a known name who has created works collected by The National Museum of Art, Architecture and Design in Norway’.

In comparison to ULAN’s broader definition of Artist, we excluded workshops, families, groups of artists, but also anonymous and unknown artists - and entries where we could not identify a proper name. A further decision to exclude companies, groups and institutions (hereafter Corporate entities⁷) from our working material was based mostly on our project timeframe, but also to focus on a uniform data model that could be applied to all entries we were going to work with. We had a year to complete our project with resources of 1.0 full time equivalent. To work with the list of corporate entity names also seemed more time-consuming than person names. Gathering information about corporate entities tends to be more challenging than person names, because they might evolve over time due to finances and ownership. We decided to establish the list type “Corporate entity” as a project of its own and plan to extend our authority list later.

Our initial list comprised approximately 14 000 person records. This was a somewhat messy list of variable quality with many duplicates and inconsistent structure, due to the down-prioritising of maintaining the name table in our collection management system. At first, we were only focusing on the Norwegian artists, as the project initially was to establish a controlled authority list of Norwegian artists, but the scope was soon

---

⁶ Ref. Categories for the Description of Works of Art (CDWA) ‘4.1.3 Creator Identity’.
⁷ We use Corporate entity in the same meaning as Corporate body in Cambridge dictionary: ‘an organisation such as a company or government that is considered to have its own legal rights and responsibilities.
https://dictionary.cambridge.org/dictionary/english/body-corporate. We thereby exclude persons from the definition.
expanded to include all artists in the collection independent of nationality. It made sense to exclude the criteria ‘Norwegian’ for the interoperability with published authoritative name lists later. We can then use our preferred spelling of all names, and within a longer timeframe contribute to making other authority lists more complete. We believe that maintaining a list of artists represented in our collections will be useful for other third-party users both nationally and internationally. For the sake of our own cataloguing workflow it makes perfect sense. For those specifically requiring it, an authority list of Norwegian artists only would still be readily available as a subset, or technically a query, to the full list. Or a list of French painters in our collection, for that matter.

After having decided on which names we could include, we continued considering which information elements from the encyclopaedia and which ones from our own database to include. Even though we had a strong wish to add mini biographies, we soon understood that, due to the resources and timeframe, we could not. We wanted to prioritise the production of a long-time lacking resource with core information elements, that could provide the basis for linking to both additional biographical information or any other desirable or necessary information in the future. Which then, are considered the core elements related to a person’s identity? Do we include relationships, associative relationships, preferred terms, pseudonyms? And what about uncertain information? After discussions we decided on a minimal set of fields, see field list in fig.3. These were similar to managing editor at Getty Vocabulary Program, Patricia Harpring’s (2010, 102) recommended core elements.

Fig.3

Record type: Person
Name: Munch, Edvard
Name Preference: preferred
Name Language: Norwegian
Name Source: Authority list of artists in Norwegian art collections
Name type: Birth name
Birth Date: 1863
Death Date: 1944
Birth Place: Løten, Hedmark, Norway
Death place: Oslo, Oslo, Norway
Nationality: Norway
Nationality preference: preferred
Gender: Male
Life role: Artist
Life role preference: preferred

This information would identify the artists, and it would be possible to differ them from each other. Gender was included as a core element, since this plays a role in the museum’s annual governmental reports (e.g. how many works by female or male artists are acquired or exhibited per year and so on). Birth and death place have also been maintained fields in our collection management system since it is a conventional part of exhibition labels in the National Museum, and we therefore decided to include them. We then added editorial rules (e.g. required, optional and repeatable fields). We established rules for how to choose a preferred term and which variant term to include (e.g. to use the transliteration as found in the most authoritative of available sources). We standardised format, syntax and character sets and discussed available authorised sources for each field.

We decided to depict ambiguous information in repeatable fields, with one preferred entry and other possible values for nationalities, name spellings and life roles. Uncertain bibliographic information e.g. for birth dates is not depicted in the conventional manner (circa/ possibly/ either-or), since support for this was unfortunately lacking in the original data structure, and we had no time to add this. This is a shortcoming of the authority list that we hope to correct by adding information about uncertain dates later.
3.1 Method and tools
Knowing what elements to include, we could then export them. The data from our collection management system could be retrieved with an SQL query that joined the name table with data from other tables such as roles and had checks such as ensuring the persons were creators of objects in the collection. The results of this query were then exported to a flat list in Excel. We kept the data unchanged in our database both as a security save but also to be able to keep working in the database during the project. In our spreadsheet of artists, each row consisted of one person and each column had one type of information about the person. Some of these columns were easily defined, others not. One example is the Life role column. Our collection comprises fine art, architectural, design and craft objects collected throughout our institution’s history. Hence, not all persons in our list are artists by profession. In similar publications in Kulturnav.org the persons are all identified with “artist” as a life role. We wanted to define this further and specify whether they were painters, architects, designers, engraves etc. Firstly, because our dataset already had this kind of detail on life roles, it would not make sense to lose this information on the way, and secondly it could prove useful in the process of identifying the right person in other web resources such as VIAF, ULAN and Wikidata. Records with similar names could be checked for match against other sources by including information such as birth or death date and the profession column to clarify the correspondence of the two records. We allowed for multiple values in the life role field to give a fuller biographical overview. As an example of multiple life roles, we also have authors who is the author of a text, but also the painter of the illustrations in a book, this person would then in our list be called author and painter.

After determining which columns to include about each artist we started editing the list in the data transformation application OpenRefine. OpenRefine is a free, open-source software, that can be downloaded and run locally. This program was a perfect fit for us. By using the very powerful sorting, filtering and facet tools, we could get an overview of our dataset so far, and start identifying problem areas that would need special attention. An example was the column Nationality, this information was from our collection management system a free text field. This meant a lot of inconsistent spelling. By using text facets, we could easily clear up this mess (See fig.4). This method was used in almost every column to check spelling.

![Fig.4](image.png)

**Fig.4.** A print screen showing how the text facets were ordered by count to see how many artists were from each country and shows also where we cleared up errors such as wrong spelling and missing capital first letter.

By using the text facets, we could also cluster similar cells, e.g. if the column included cells that had the values ‘Canada’, ‘canada’ and ‘cannada’, openRefine’s cluster tool identified these as similar and we could change them all to what we preferred – Canada (See fig.5). After using these tools in every column, the list became more and more complete and consistent. We could now search for cells with missing values.

---

8 OpenRefine - formerly Google Refine - is a powerful tool for working with messy data: cleaning it; transforming it from one format into another; and extending it with web services and external data. (Openrefine n.d.)
Reconciling our data with other online resources

To start filling in missing values we reconciled our data with other online resources, viaf.org, norskkunstnerleksikon.no, kulturnav.org and wikidata.org. For both VIAF and Wikidata there are existing reconciliation web services that can be queried from within OpenRefine. We decided that matching against external web resources would only be carried out automatically, by automated query processes using the reconciliation tool in OpenRefine, and that any manual mapping was outside the scope of our project. Starting with viaf.org, the search found a lot of our person records, but some were uncertain matches. We decided on how high the score had to be to automatically link our record to a VIAF record, the ones with scores too low were manually checked for correctness. We found links to about 50% of our person records. Some were found in several resources and others just in one. This taught us that that it is important to link to different kinds of web resources. We saw that some of the web resources had many more Norwegian artists than others and that some of the resources had more detailed information about international artists. The spelling of names also made the linking a bit more troublesome, as some resources would only match if there was exact correspondence on spelling. The Norwegian letters ø, æ, å also made some difficulties with matching our artist names to external resources, but for most resources this was solved with troubleshooting string encoding. For the other resources we had to spend some time working out the queries to send to the APIs of the different web services. The SPARQL endpoint for the RDF dataset of the Norsk Kunstnerleksikon was particularly tricky, passing a lengthy url-encoded SPARQL query to the endpoint through the GREL\textsuperscript{9} syntax of OpenRefine’s reconciliation tool.

Enriching our dataset with information from the linked records

Having linked as many records as possible we could now harvest data from the links. We used wikidata.org to collect core information such as birth, death, gender and nationality. This would complete our list even further, but also worked as a cross-check to see if our information corresponded with the info from Wikidata. External data was initially stored in separate columns to enable automated column comparisons in GREL. This was particularly useful on birth and death dates. In a great number of cases we were able to improve a year value in

---

\textsuperscript{9} Some of OpenRefine’s functionality requires defining queries, value comparisons and more using GREL (Github 2015).
our original dataset with a full date (D-M-Y) from Wikidata. Some data did not correspond and we had to do an extra check to see if it was the same person and to correct either our data or ignore the data from Wikidata. This was done based on which data we could find other places and by which number. This meant that if the data from Wikidata was only found in Wikidata and our “original” info was found on several websites, we would keep our “original” info and discard the info from Wikidata but keep the link. In some cases, there would be discrepancies for instance in birth and death year that were small enough to indicate that these records were still referring to the same persons. These discrepancies occurred almost exclusively for years before the twentieth century, and not surprisingly more frequently the further back in time. The differences could be identified automatically by setting a tolerance of e.g. +/- 2 years for date matching, then manually evaluated and addressed.

After we had gathered core elements from Wikidata, we found an error in some of the dates. Suspiciously many people were born or died on the 22nd of December during the sixteenth century. After some additional searching to find these dates we had to conclude that some data corruption had occurred in the process of enriching years with day and month from Wikidata. We decided to remove these dates from our list. In this vast dataset we had to make some pragmatic choices to complete it, even if this resulted in removing some good data. All the artists born in the sixteenth century now have no date of birth or death, just the year that was already in our original dataset.

As we kept supplementing our dataset with a growing number of links to matching entries in external web resources, it became clear to us that finding these links was just as important as finding exact and correct bits of information to enrich our own core data. This was at first somewhat troubling to us as museum workers, but we realised that the links have such great future potential to provide more information on a person. The link was just as important as making sure that every bit of information we collected was correct.

**When do we stop?**
Filling in the blank cells was a time-consuming process, but a very important one. Many of our artists are not found in the different online resources but can be found on other websites. This is the case for many Norwegian contemporary artists, but also a lot of the older Norwegian artists. Not surprisingly, we were unable to find information on all persons. At some point we had to stop and accept that there will be many records with some missing data elements. Our authority list, when finished consisted of almost 10 000 people records. As we aim to make this a collaboration between museums in Norway we hope that others will contribute later and fill in the blanks where they happen to have information. Regional museums may have more detailed info on some of their local artists than we have found. The long-term success of the project will depend on collaboration between museums and willingness to share information with each other.

**4 Function**
At the same time as we worked on narrowing our scope of content to a minimum, the possibilities concerning options for technical structure and web interoperability seemed to maximise. To be able to decide on this, we first had to make a clear definition of the intended function(s) of this list. We have had the purpose of our authority list in the back of our minds throughout the years and have had to revise our view several times. Our list of names was initially made by, and only meant for, cataloguers internally in the museum. After publishing the collections on the national cultural heritage portal Digitalmuseum.no in 2010, the list at least had to be made available to all cataloguers in Norwegian museums who wanted to share content on this portal. In the following years, the user demand for an authority name list as a web resource increased proportionately with the growing number of collections being published online. As we then took up this project again in 2017, it was a long-established fact that authority lists like these are created for retrievers across different types of collections and continents and for the referencing semantic web technology. By publishing a web resource authority, we have the potential to reach and supply all interested web parties. The audience had expanded from scholars who used the analogue archive, to including anyone browsing artists names and art collections on the internet. At the time of writing this paper, we are about to publish our controlled vocabulary as an authority list. Nevertheless, some unsolved matters remain to be analysed before we do so.
4.1 Fit for purpose?
As a museum we value scholarly conventional, updated and reliable information. And as the responsible museum for this list, we want, like in the “Lessons learned” of the OSCI, (The Getty Foundation 2017, Lesson 1), to underline the importance of our content standards to be trusted also in the digital environment. Going back and summing up our rationalised scope: We now have quality controlled descriptive metadata in a list shaped by the character of our collections’ content. It is categorised by international standards, but also by organisational preferences (such as the information units: gender, birth/death place).

For our authority list to contribute to improving knowledge about Norwegian art, or art in Norwegian collections, we must consider a variety of format steps. We asked ourselves again: For further use, are our data good enough? With the COPE-maxim in mind: Are our data fit for reuse in 2018? We experienced what information scientists Seth Van Hooland and Ruben Verborgh state in Linked Data for Libraries, Archives and Museums: “it is sometimes complex to find consensus on whether data are of good or bad quality” (2014, 73). Several scholars have over the last years analysed the quality of data and the utility of web services within the area of cultural heritage. Concerning our authority list as a web resource, we had two main issues that we involved in our further examination of quality and usefulness. Our first issue concerned our narrow scope of descriptive metadata. As associate professor of visual culture Nina Lager Vestberg (2013) points out: “[…] an online database may just as easily close off avenues of inquiry as open them up, thereby limiting rather than expanding the number of potential questions to be asked of the archive” (487). She analyses how archival order ‘affect the production of knowledge and meaning’, by comparing methods used in digitising web archives in the photographic collections of the Warburg Institute and the Courtauld Institute of Art in London. Vestberg exemplifies how sticking to the conventional cataloguing structural manner might exclude the expanded audience that the web service could have been intended for: “it should come as no surprise that an image resource designed by iconographers will tend to be of use chiefly to other iconographers” (484). Hence, in our case, adding descriptive metadata to our authority list was already out of our scope. We wanted to make the list available for linking to further content, and therefore went on evaluating how to structure it as linked data. Our second issue thereby targeted the more technical side of interoperability, as we concentrated on adding layers of metadata encoding to our authority list. Strategies for implementing linked data are published by CIDOC’s Statement on Linked Data identifiers for museum objects (2013). And suggesting answers and practical guiding to the reuse-challenge, also in the sense of machine-readable formats, is done e.g. by above mentioned Hooland and Verborgh (2014). They provide expeditiously explanations on how quality metadata can be applied to existing descriptive metadata, and how this is key to better access to content.

4.2 Question of Format
Having these issues in mind, and being aware of how rapidly data formats are evolving around us: How can we make good choices to meet the expectations of a name authority to be available both for humans and machines to link, systematise and make more valuable? We concluded that the importance of achieving interoperability with international standards is of greater importance than ever. And to achieve interoperability, we found that it was decisive to apply editorial- and structural rules of standard vocabularies where possible. Unlike the ULAN, our authority list is not a thesaurus. We only have one level and all terms belong to the same class (person). Our list nevertheless includes same-as-references (links to VIAF, ULAN, etc.) and preferred terms (name spelling, nationality and life roles). Considering the topics of technical format/ metadata encoding, we aspire to add: multi-language support for all text fields, controlled text, date and numeric formats and stable URI/ permalinks. We also have expected to map to a concept reference model, preferably CIDOC CRM[11]. Once published to KulturNav, the artist list will be available for use in the KulturNav API as RDF. KulturNav’s framework has a

---

[10] Descriptive metadata, meaning content in the definition from DCA: “Descriptive metadata is used to identify and describe collections and related information resources, e.g., artist, title, location, date […]” (DCA 2012, 8) and based on the CDWA standard (Baca and Harpring 2009).

[11] “The CIDOC CRM is intended to promote a shared understanding of cultural heritage information by providing a common and extensible semantic framework that any cultural heritage information can be mapped to. It is intended to be a common language for domain experts and implementers to formulate requirements for information systems and to serve as a guide for good practice of conceptual modelling. In this way, it can provide the "semantic glue" needed to mediate between different sources of cultural heritage information, such as that published by museums, libraries and archives.” (CIDOC CRM n.d)
number of entities, one of which is the person entity with the following data structure which matches our own data structure quite well:

**Entity type: Person**

**Property type:**

- `person.firstName` (First name), TEXT, 0:1
- `person.middleName` (Middle name), TEXT, 0:1
- `person.lastName` (Last name), TEXT, 0:1
- `entity.alternativeName` (Alternative name), TEXT, 0:M
- `person.title` (Title), ENTITY / Person.Title, 0:M
- `person.birth` (Birth), ENTITY / Person.Birth, 0:1
- `person.death` (Death), ENTITY / Person.Death, 0:1
- `person.lifeRole` (Life role), ENTITY / Person.LifeRole, 0:M
- `person.gender` (Gender), ENTITY_REFERENCE / Concept, 0:1
- `person.nationality` (Nationality), ENTITY_REFERENCE / Place, 0:M
- `person.biography` (Biography), TEXT, 0:1
- `person.parentOf` (Parent Of), ENTITY_REFERENCE / Person, 0:M
- `person.childOf` (Child Of), ENTITY_REFERENCE / Person, 0:M
- `person.siblingOf` (Sibling of), ENTITY_REFERENCE / Person, 0:M
- `person.memberOfGroup` (Member of), ENTITY_REFERENCE / Group, 0:M
- `agent.owns` (Owned), ENTITY_REFERENCE / Agent, 0:M
- `person.occupation` (Occupation), ENTITY / Person.Occupation, 0:M
- `person.education` (Education), ENTITY / Person.Education, 0:M
- `agentOrdered` (Has ordered), ENTITY_REFERENCE / NavalVessel, 0:M
- `agent.constructed` (Has constructed), ENTITY_REFERENCE / NavalVessel, 0:M
- `agent.scraped` (Has scrapped), ENTITY_REFERENCE / NavalVessel, 0:M
- `agent.participatedIn` (Participated in), ENTITY / Agent.Participation, 0:M
- `person.marriage` (Marriage), ENTITY / Person.Marriage, 0:M
- `person.residence` (Residence), ENTITY / Person.Residence, 0:M
- `person.memberOfMembership` (Membership), ENTITY / Person.Membership, 0:M
- `person.award` (Award), ENTITY / Person.Award, 0:M
- `person.appointment` (Appointment), ENTITY / Person.Appointment, 0:M
- `person.actorIn` (Actor in), ENTITY_REFERENCE / Organization, 0:M
- `person.established` (Has established), ENTITY_REFERENCE / Association, 0:M

KulturNav’s provider KulturIT is currently working on improved and extended use of standards such as CIDOC-CRM, DCAT-AP and Schema.org (Kulturnav 2018 n.d.). Metadata expert Tony Gill (2016) explains in “Metadata and the Web”, the different data structure, -format and -value standards, tools and methods for encoding metadata in a machine-readable form - and protocols for distributed search and metadata harvesting. He argues that: “By using these various components in intelligent and appropriate ways in order to provide access to the rich information content generated by libraries, archives, and museums, it should become possible to build a global Semantic Web of digital cultural content and integrated search tools to help users find the content they are seeking” (Gill 2016). These museum community’s recommendations also harmonise with the new guidelines for public data availability, published by the Norwegian Agency for Public Management and eGovernment (DIFI), where the following is one of their highlighted points:

> “Data must be available in machine readable formats. In addition, the formats should be standardised. This ensures good interactivity with other information (interoperability) and does not impose unnecessary limitations on what the information can be used for in the future. Examples of machine readable and standardised formats are CSV, XML, JSON and RDF serialisations such as RDF / XML, JSON-LD, and Turtle.”

(DIFI 2018, subsection 8)

Our goal by standardising all authority list information components is to achieve sustainability and be equipped to adjust to new evolving data formats in the future. Yet, some decisions about sustainability, richness and progress remain to be made for the future of our authority list.

5 Future plans

Management and further plans

As we have seen, the major shortcoming of Norsk Kunstnerleksikon is that it is not a living resource. At the National Museum in Oslo we have been using the online RDF version of the encyclopaedia to harvest and display biographies of artists in our collection online (Nasjonalmuseet Collection n.d.)\(^{13}\). We are seeing more and more that a number of biographies do not represent the individual artist well enough, and in many cases more recently updated biographies can be found elsewhere, such as on Wikipedia or in the online version of the Store Norske Leksikon. An online artist list resource with links to external resources is a great tool for harvesting the best biographies – or even giving the public the choice between different sources, like the Museum of Modern Art in New York does (Romeo 2016).

Our authority list of artists needs to be managed and edited in the future so that it is a living resource. One way of keeping the information updated could be to automatically harvest new or revised data through the links to external resources. This could be timesaving for the museum, but we have still not decided on whether to do this. One of the problems with automatically changing the information is of course adding errors and that correct information gets “lost”. This is a discussion that we have not completed yet. A more likely scenario is an extension of the joint museum community effort. The National Museum intends to involve and encourage other national art, design and architecture collections to make use of, and contribute to keeping the list updated and supplementing it, through the editorial interface and workflow of KulturNav.org. In addition to this, one could envisage a long-term and semi-automated crowdsourcing project where the public are encouraged to contribute through an interface in the National Museum’s online collection (or other museums’ online collections or Digitalmuseum.no). Through an interface the public would be able to report errors, supply missing information, report unrecorded deaths and so on. These submissions can find their way to the editorial workflow of KulturNav for manual check and submission by the editorial team.

As a start for future sustainability, the National Museum in Norway we commence to take on the editorial task and will encourage the Norwegian museum community network to participate in keeping the authority list up to date. We have established an editorial team and hope that before the authority list is published, that people representing the triangle finances (the Arts Council and other financing sources) infrastructure (KulturNav and IT providers) and content (collection cataloguers, the public) have committed to a strategy. If we manage, we would ensure that the authority list can live on and evolve to cover the long-awaited authority list that can ensure consistency, clarity and context in a professional way for digitised Norwegian art.

KulturNav as a platform for shared authority lists and terminology in Norwegian and Swedish cultural heritage has been slowly gaining momentum over the last couple of years. It has taken time for end users in the museums to see how such a platform can interoperate with other museum tools such as collection management systems, online collections and similar, and how it can provide value to both internal workflows and data quality. It has also taken time to establish an understanding among interested users that the success of any such platform must be a community effort. There is a huge potential for shared authority lists in Norwegian museums. So far, many lists have been smaller local initiatives with a clearly defined and limited scope, for instance a list of designers working for a particular porcelain factory. As our authority list of artists in the National Museum of Art is published to KulturNav, we plan to organise several local workshops for the Norwegian museum sector to engage other museums in the effort to improve and contribute to the list. Such workshops will not only evangelise about the benefits of a joint authority list for artists, but must also provide detailed training on how to contribute, editorial workflows and policy.

---

\(^{13}\) Nasjonalmuseet’s collection contains around 400 000 art, architecture and design objects. More than 42 000 objects are currently available online, and the number increases as we publish new objects every month.
Bibliography


Norwegian artist names
CIDOC 2018 Heraklion, Crete, Greece


Verborgh, Ruben, and Max De Wilde. 2013. Using OpenRefine: The essential OpenRefine guide that takes you from data analysis and error fixing to linking your dataset to the Web, Birmingham: Packt Publishing.


