

## Summary Workshop Wikidata: How to insert media artworks and collections?

### Case Study – Project Mediakunst op Wikipedia

**Presented by:** Sanneke Huisman and Hanno Lans

**Moderated by:** Sandra Fauconnier

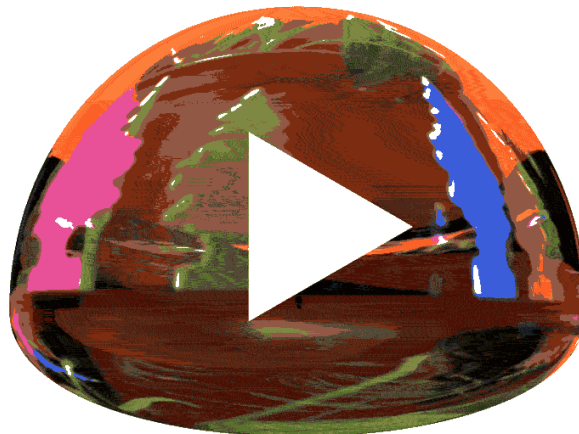
Friday 18-03-2022

Workshop Summary by Joost Dofferhoff

In March 2022 LIMA held its annual Transformation Digital Art Symposium, which was again held online due to the coronavirus pandemic. The symposium spanned two days from March 18<sup>th</sup> to March 19<sup>th</sup> and was focused on the theme of collaborative knowledge building and sharing. As knowledge becomes more readily available online it becomes imperative to collaborate in developing ways to share it as well as the responsibilities that come with it. LIMA believes that media art requires new collaborative approaches to sharing and caring for it. Through this symposium it hopes to highlight the networks that support it and exchange knowledge that can address its continued preservation.

 Mediakunst.net

powered by **LIMA**



### Case Study – Project Mediakunst op Wikipedia

[Mediakunst op Wikipedia \(2021-2023\)](#), a project initiated by LIMA, aims to generate visibility for media art on Wikipedia by writing 500 wikipedia articles and we make metadata accessible as linked open data. As part of the collaborative project, a total of 2000 media artists were entered last year in Wikidata, the reference database under Wikipedia from the collections of LIMA, Van Abbemuseum, Stedelijk Museum Amsterdam, the Cultural Heritage Agency and Frans Hals Museum. The next step is to make the media artworks easier to find and to make them visible on the Mediakunst.net platform

via a database link. This workshop discusses how the media artworks of (museum) institutes worldwide can be connected to each other so that their oeuvres are more easily accessible and a picture that transcends institutions is created. The workshop discusses the role of Wikidata as a connecting, 'open' and common database with the participants in the workshop, and how this data can be optimally connected with the collections and individual institutions.

The workshop, which was moderated by Sandra Fauconnier (Project Director at OpenRefine), started by questioning why and how Wikipedia and Wikidata could be used for media art collection. Sanneke Huisman (Curator at LIMA and project leader of Mediakunst op Wikipedia) stated that LIMA initiated the project because of the constatation that media art and media artists are underrepresented on Wikipedia. Existing biographies on the websites of collections often are not up to date as institutions do not have the time available to tend to them. Sharing this task of writing biographies with the Wikipedia community can help institutions keep their information up to date while at the same time expanding the visibility of their collections.

Starting up a Wikidata link with a collection asks for an investment of time, a good model and a certain level of expertise. For this purpose LIMA took Wikimedian in Residence Hanno Lans on board. Hanno Lans was in charge of incorporating the 2000 artists belonging to the five collections into Wikidata; the decentralized database of information that lies beneath Wikipedia. Before a Wikipedia article can be written, data and sources need to be inserted to Wikidata. One can then reference these sources in Wikipedia when writing an article. Wikidata, as opposed to other databases, is a so-called Secondary Database, this means that Wikidata does not host any of their own information but is an open source platform to which other (primary) databases and other information can be linked. It can therefore be said that Wikidata is a database of information that already exists elsewhere. What makes Wikidata even more valuable is the fact that all this information can be queried; meaning that graphs and calculations can be made on all the information that is available which can spark new and exciting revelations.

Other guests that were invited to speak on their experiences with Wikidata and Wikipedia were Andreas Kohlbeckerde (ZKM), BartMagnus (Meemoo) and Lozana Rossenova (Digital Designer, Researcher)

### **Interactive workshop**

Participants of the workshop were encouraged to actively think with the theme of the applied use of Wikidata in the field of media art. This participation was hosted on an Etherpad which is an open source collective real-time editor in which people can comment on what is happening when it's happening. The link to the etherpad that was used during the workshop can be found [here](#). The first question that was put up for audience input was to come up with positive and negative aspects of using Wikidata for (media) art collection.

The cross-institutional overview of the possibility of uniting information sources, for example by linking catalog and database identifiers, turned out to be the most important and popular aspect that the participants were excited for. Other positive aspects included increased visibility and up-to-date biographies or the fact that using Wikidata could help your institutions comply with the FAIR (findability, accessibility, interoperability and reusability) principles.

Regarding the negative aspects of Wikidata the participants seemed almost unanimous on two points; the fact that using Wikidata together with your own database creates extra work as it does not replace work on the collection's own database and the possibility for vandalism and possibly harmful information that can be added to the database and articles. The latter point started a discussion among the group on the theme of contested information. Some participants mentioned being hesitant to start working together with a community led database such as Wikipedia because they were not yet willing to give certain topics out of their control. In opposition to this Sandra Fauconnier mentioned that the abundance of opinions on Wikidata can also be considered a strong point of the platform as it is a place where these contested opinions can live next to each other. However, even after this opposition, it was said that some institutions would still curate what topics and what information they would and would not delegate towards the Wikimedia community.

### **Data Modeling**

The workshop continued on the topic of data modeling, linked open-data and the incorporation of complex works into a database such as Wikidata. Marina Abramovic and Ulay's video artwork *AAA-AAA (1978)* was presented by Sanneke Huisman as a complex work for incorporation into Wikidata. This video artwork, which was originally a live performance, was recorded in Amsterdam and in Liege and subsequently distributed as a video artwork. Next to this the video's of the work exist in different formats like Standard Definition (SD) and High Definition (HD). Works like this summon questions such as how to define the type of the work: is this documentation, an artwork or a registration? And how to incorporate this into Wikidata; is it considered 1 work with multiple versions? Are all versions considered independent works? For works like this a data model should be carefully created because the way it is incorporated into Wikidata changes the way that other people can interact with this information.

Jan Robert Leegte's *Scrollbar Composition (2000)* was presented by Lozana Rossenova (Digital Designer, Researcher) as an example of a work that she created a data model for. The complexity of this work exists in its many iterations and the fact that all these iterations exist simultaneously on the internet. In 2001, "untitled[scrollbars]" was accessioned by Rhizome's ArtBase online archive, where an archived variant (in HTML), as well as a WARC (Web ARChived) file are preserved. Lozana and Rhizome decided that for this data model they would make a different page for every single instantiation but ultimately link all these pages together with the 'instance of' and the 'has variant' functions of Wikidata (For all functions used in this data model; [click here](#)). This function makes sure that all iterations can simultaneously exist as their own artwork while being part of an overarching 'main' artwork. The artworks keep their independence while being part of a series. These variants are described with Linked Open Data thanks to the Wikibase software. Similar approaches can be taken in Wikidata, but the data modeling work will need to also fit within the larger schema of Wikidata, which can be more challenging than in the case of Rhizome's ArtBase where a dedicated data model for net art works was developed. To see how Rhizome has incorporated *Scrollbar Composition* into their ArtBase system with Wikidata; [click this link](#).

### **Fitting Information**

An important aspect of data modeling is making the decisions what does and what doesn't belong in such a model. For this query the Etherpad was used once more to ask the participants which

information or metadata, in the case of complex artworks, is fitting for a shared repository such as Wikidata? Opposed to this the audience was also asked what information was not fitting.

Out of many suggestions the main things people found fitting was version history, technical dependencies and memories. Version history was deemed fitting information because of the ability to structure a data model, as in the case of *Scrollbar Composition* (2000), according to the different versions that were made and the overarching artwork that belongs to them. Technical dependencies also were amongst the most fitting as they shape how an artwork can be entered into a data model. Lastly, memories of the artwork were pointed to as fitting metadata. Memories of the artworks by the public could have added value and add information such as feelings that an institution or collection can not add themselves.

Information such as the details of different versions, technical dependencies and even the memories that the public have of these particular artworks. An interesting addition that some participants deemed worthy was the inclusion of artist interviews. It was noted that complete artist interviews are often not possible to share because of the possible sensitive content. Therefore it was agreed that it would possibly be better to incorporate a reference to the existence of such an interview into Wikidata. In this way the public is notified about the existence and can look for it themselves.

Information that was not deemed fitting included detailed information about instantiations or manifestations of the work; especially in institutional collections. Other points included server locations in the case of net art or conservation actions that were undertaken by institutions or more private information such as acquisition price, insurance price or detailed loan information.

Bart Magnus (Meemoo) was curious as to why not many of the participants included genre information in the list of fitting information. Many stated that genre is still quite contested in the field of media art. On the other hand, not adding genre qualifications to an artist or a work on Wikidata will not increase its findability. Some participants added that there is a fear of entering wrong information and the possible debates this might provoke. Participants did agree that as much information as possible is usually a good thing for the link-ability and findability of an artist / work. Lozanna adds to this that more institutions should embrace messiness and relinquish purity in their data. Communities online will always use and play with your data. Therefore institutions should best accept that having a perfect data model will never exist and agree that a well-filled but messy database might profit everyone.

### **WikiProject Media Art**

The final few minutes of the workshop focused on how to structure Wikidata Projects. Hanno Lans, the WikiMedian in Residence at LIMA, has created a blank project page on Wikidata entitled [Project Media Art](#) for all platforms that wish to use Wikidata for their collections to collaboratively work on. Hanno's ambition with this WikiProject is to work on and improve all things media art on Wikidata, but in particular to be the main hub for all media art metadata.

Workshop participants were asked to ponder what would be necessary for such a big project to succeed and what would be needed for so many institutions to work together on shared data? Many agreed that before even starting filling in the content of such a project a common practice should be found. Another item that institutions should find common ground on would be the use of shared

identifiers on Wikidata. Linking collections together would be made a lot easier if one Wikidata identifier was used across all these collections. This shared identifier can co-exist next to the collection specific identifier that the collection already uses. Opposed to this the use of different properties per institution was suggested. Different properties would be useful to still distinguish collections from each other and to make this project go as broad as possible.