

Data-driven Semantic Modeling of Nam June Paik’s Artwork, A Tribute to John Cage : A LIDO-CIDOC CRM Case Study*

Kim, Hyun-sook **

Table of Contents

1. Introduction
2. Framework and Context of A Tribute to John Cage
 - 2.1 Methodological Approach: Metadata, LIDO Schema and CIDOC CRM
 - 2.2 Context of A Tribute to John Cage
3. Dataset Analysis from MoMA, Mplus, and ZKM
4. Data Integration and Transformation
 - 4.1 Data Alignment with LIDO Schema: MoMA, Mplus, and ZKM
 - 4.1.1 Object Wrap
 - 4.1.2 Event Wrap
 - 4.2 Mapping LIDO to CIDOC CRM
 - 4.3 Result: Querying using SPARQL
5. Extended Semantic Modeling using LIDO RelationWrap and CIDOC CRM
6. Limitations and Discussion
7. Conclusion

Abstract

본 연구는 백남준 “A Tribute to John Cage”작업의 의미론적 모델링에 초점을 맞춘다. MoMA, Mplus, ZKM의 해당 데이터셋을 기반으로, 온톨로지 모델인 CIDOC CRM과 중간 스키마인 LIDO Schema를 활용해 의미론적 매핑과 변환 과정을 수행하였다. 이후, SPARQL 쿼리로 결과를 도출하였으며, 이 전 과정은 Python으로 구현되었다. 이를 통해 본 연구는 이산되어 있는 백남준 데이터셋의 의미론적 통합 가능성을 탐구함과 동시에, 현대미술 데이터의 의미론적 구조화 가능성과 확장성을 살펴보는 데 의의를 지닌다.

This study focuses on the semantic modeling of Nam June Paik’s *A Tribute to John Cage*. The research conducted the end-to-end process, including data integration, semantic mapping, and transformation, followed by querying through SPARQL. In terms of semantically mapping and transformation, the standard controlled vocabularies such as LIDO Schema and CIDOC CRM were used. Ultimately, this study aims to explore the potential for semantic integration of Nam June Paik’s dispersed datasets and examine the extensibility and structuring of semantic data in the field of contemporary art.

Keywords: Semantic Modeling, CIDOC CRM, LIDO Schema, Metadata, Nam June Paik

* This research paper is a re-edited version of part of the master’s thesis at Friedrich-Alexander-Universität in Germany. The related description is available on GitHub(<https://github.com/Hyunsook0113/masterproject>).

**1st Author: annakim2327@gmail.com

1. Introduction

In contemporary art, the format, content, and modes of collaboration have become increasingly diverse. Nam June Paik, who continues to influence contemporary artists today, exemplified this diversity in the 20th-century art scene. Through cross-genre experimentation, collaboration with artists from diverse disciplines, and pioneering use of emerging technologies, he expanded the boundaries of artistic practice. However, the diversity of his artistic practice and the fact that his archival materials and works are held by numerous institutions worldwide make it difficult to accurately understand the context of his work. To understand Paik's complex artistic world, it is therefore necessary to connect information distributed across multiple institutional repositories.

Within the field of cultural heritage research, linked data principles developed under the Semantic Web paradigm have led to ontology models such as CIDOC CRM, which provides a structured framework of classes and properties for modeling events, actors, and objects. While prior studies have primarily applied CIDOC CRM to the description of cultural heritage events and images, transforming heterogeneous institutional metadata into this high-level ontology remains challenging. To address this gap, LIDO was developed as a middle-layer schema to bridge institutional datasets and semantic ontology models.

Based on this background, the present study examines the complete semantic modeling workflow, encompassing institutional data, LIDO, and CIDOC CRM, and resulting in SPARQL queries. Accordingly, this research investigates how heterogeneous institutional datasets can be integrated and semantically aligned, as well as the interpretive and structural issues that arise in this process, in order to enable the semantic linking of dispersed artworks by Nam June Paik.

This study focuses on *A Tribute to John Cage* (1973/1976) as a representative case to explore the semantic modeling of Paik's artworks. The work is analyzed within the broader context of Paik's German-period experimentation and his interactions with artists such as John Cage. Datasets were collected from three repositories—The Museum of Modern Art (MoMA) in the USA, Mplus in Hong Kong, and ZKM in Germany—and analyzed with respect to differences in metadata elements, terminologies, and data structures.

The objective of this research is not the construction of a complete system, but the development of a methodological approach. Ultimately, this study seeks to contextualize Nam June Paik's artworks by applying semantic data modeling standards, particularly using data from institutional repositories. Accordingly, this paper emphasizes the overall workflow rather than providing detailed technical descriptions of specific rules, structures, or syntax.

2. Framework and Context of *A Tribute to John Cage*

2.1 Methodological Approach: Metadata, LIDO Schema and CIDOC CRM

Metadata has historically been developed as a means to identify, manage, and retrieve information resources. Standards such as Dublin Core were designed primarily to ensure interoperability and consistency in resource description, particularly in bibliographic and information management contexts. These metadata schemas are generally structured around an object-centered model, in which descriptive elements are organized within a single record associated with a resource.¹

While such metadata frameworks have significantly improved information sharing in digital environments, they are limited in their ability to express the complex and dynamic nature of cultural heritage data. Artworks and cultural objects are not static entities; rather, they emerge through a series of activities such as creation, exhibition, movement, institutional acquisition, and collaboration. Their meanings are constructed through relationships among people, places, and events. Traditional metadata schemas, however, typically represent these aspects in parallel fields rather than as explicitly defined semantic relationships, making it difficult to interconnect between

entities.

From the late 1990s, the emergence of the Semantic Web introduced a shift from data representation as isolated records toward data as machine-readable meaning structures. In this new paradigm, information is no longer simply stored but is structured in a way that allows reasoning, linking, and inference across datasets. Ontologies became central to this approach, as they provide formal conceptual models that define not only entities but also the relationships among them.²

Within this semantic framework, CIDOC CRM was developed as a domain ontology for cultural heritage documentation. Rather than providing a fixed set of descriptive elements, CIDOC CRM introduces an event-oriented conceptual structure in which cultural heritage is modeled through abstract entities such as places, time-spans, and activities. This allows complex information—such as provenance, use, institutional history, and artistic collaboration—to be represented as interconnected knowledge structures rather than as isolated descriptive fields.³

The relationship between metadata and CIDOC CRM, therefore, is not one of linear technical development but of conceptual transformation. Converting metadata into CIDOC CRM is not a matter of simple format migration; it is a process of semantic reinterpretation in which descriptive elements are restructured according to an ontology-based model. Institutional metadata thus becomes content that is re-modeled within a higher-level conceptual system, enabling integration across heterogeneous datasets.

However, while CIDOC CRM offers a powerful semantic framework, its abstract structure presents challenges for direct implementation using existing institutional metadata systems. To address this gap, an intermediate schema was introduced to facilitate the translation between operational metadata formats and the ontological model.

LIDO (Lightweight Information Describing Objects) was developed for this purpose as a middle-layer schema that aligns institutional object-oriented metadata with the event-centric logic of CIDOC CRM. It enables organizations to preserve their practical data structures while preparing their data for semantic transformation. By operating between XML-based metadata and RDF-based ontology models, LIDO functions as a mediation layer that supports data normalization, interoperability, and semantic integration.⁴

2.2 Context of *A Tribute to John Cage*

Nam June Paik was a central figure in the Fluxus movement, an international avant-garde network organized in the 1960s by George Maciunas to challenge conventional distinctions between art and everyday life through experimental practices encompassing music, performance, and visual media. Even before his formal involvement in Fluxus, Paik was deeply engaged in experimental culture during the late 1950s in Cologne, where institutions and informal spaces such as Galerie 22 in Düsseldorf and Mary Bauermeister's atelier in Cologne fostered interdisciplinary collaboration among artists associated with Fluxus and the ZERO Group.

Paik participated in John Cage's course at the 1958 Darmstadt Summer Course for New Music, an event that marked a turning point in his artistic trajectory. He later recalled that Cage's ideas on chance, silence, and performance had a profound influence on his thinking about music and art. In 1959, following this encounter, Paik presented *Hommage à John Cage: Music for Tapes and Piano* at Galerie 22. According to contemporary accounts, the performance culminated in a dramatic finale in which Paik violently attacked and ultimately destroyed the piano. In 1960, at Mary Bauermeister's atelier, he staged another provocative action by cutting off John Cage's necktie.⁵

Within this historical context, *A Tribute to John Cage* (1973/1976) can be understood as a condensed articulation of Paik's long-standing engagement with Cage's ideas and experimental networks. Produced in commemoration of Cage's 60th birthday, the work assembles interviews, festival footage, performances, and commercial imagery.

A Tribute to John Cage opens with a scene featuring a roughly made robot barely able to walk through the

streets of New York, before transitioning to a sequence in which John Cage narrates. This is followed by an interview-style conversation between Alvin Lucier and Russell Connor. Their discussion centers on the question, “Who is John Cage?”, while the two speakers intentionally stammer, as if experiencing a digital lag, which foregrounds the gap between intended meaning and spoken language. Fragmentary images from the Woodstock Festival are interspersed throughout the work. Around the ten-minute mark, footage appears of Cage performing his work *4'33"* (1952) at Harvard Square. The performance begins with the closing of the piano’s fallboard and concludes with its reopening, echoing the original structure of the piece. A Pepsi commercial then appears, taken from *Waiting for Commercials* (1966-1972/1992) by Jud Yalkut and Nam June Paik, held in the collection of MoMA.⁶ The video subsequently returns to the Harvard Square performance.

Later sequences include *TV-Bra* featuring Charlotte Moorman, which was later developed into various forms, conceptually and materially linked to works such as *TV Bra for Living Sculpture*, *TV Cello*, and *TV Bed* (1977) at ZKM.⁷ Several image collages and noise compositions follow, before the work returns to the interview with Lucier and Connor, which concludes the piece.

Conceptually, the point of departure for this work is *4'33"*, Cage’s famous composition, which replaces traditional musical content with environmental sound and incidental noise over a duration of 4 minutes and 33 seconds. Although no original video recording of *4'33"* exists, Cage’s writings describing the work are held by MoMA.⁸ Also, one sequence in which a piano is violently destroyed directly recalls Paik’s earlier performance *Hommage à John Cage: Music for Tapes and Piano* (1959).

Namely, this video work lies at the intersection of diverse interrelated elements. To explore how these relationships and contextual structures can be effectively implemented, the next section examines the work using actual datasets from three repositories—MoMA, Mplus, and ZKM—with a focus on their differences and distinctions.

3. Dataset Analysis from MoMA, Mplus, and ZKM

The datasets for *A Tribute to John Cage* were collected from three repositories: MoMA in the USA, Mplus in Hong Kong, and ZKM in Germany. Each institution provides access through different methods and data formats. MoMA makes its collection data publicly available via the Museum of Modern Art GitHub repository, where datasets are released and regularly updated in CSV and JSON formats.⁹ Mplus provides access through a token-based GraphQL API that returns hierarchical JSON data, which were converted into flat CSV structures for analysis in this study.¹⁰ In the case of ZKM, internal datasets were shared in CSV and XLSX formats, however, public distribution of the original dataset is not permitted. Therefore, the metadata presented in this paper for ZKM is derived from the institutional website rather than from the internal dataset.

Table 1. Metadata in MoMA, Mplus, and ZKM

	MoMA	Mplus	ZKM
Object Metadata	'Title', 'Artist', 'ConstituentID', 'ArtistBio', 'Nationality', 'BeginDate', 'EndDate', 'Gender', 'Date', 'Medium', 'Dimensions', 'CreditLine', 'AccessionNumber', 'Classification', 'Department', 'DateAcquired', 'Cataloged', 'ObjectID', 'URL', 'ImageURL', 'OnView', 'Circumference (cm)', 'Depth (cm)', 'Diameter (cm)', 'Height (cm)', 'Length (cm)', 'Weight (kg)', 'Width (cm)', 'Seat Height (cm)', 'Duration (sec.)'	id sortNumber publicAccess objectNumber classification title titleOther displayDate displayDateOther beginDate endDate dimension dimensionDetails medium creditLine constituents images color search cloudinary	Artist/ Artist Group, Title, Year, Category, Format, Material/ Technique, Dimension/Duration, Contributors, Collection

Table 1 summarizes the metadata structures provided by each repository. The metadata elements are related to the object level. In the case of Mplus, classification information is structured into two components: area and category. The field “dimensionDetails” includes measurement attributes such as width, height, depth, unit, element, and rank. Information about artists can be accessed through the constituents field, which contains both identifiers and names. Artist metadata further includes attributes such as “id”, “name”, “nameOther”, “alphaSortName”, “displayBio”, “gender”, “beginDate”, “nationality”, “type”, “roles”, “isMaker”, “objectCountPublic”. MoMA similarly provides separate datasets for artworks and artists, which are linked through the ConstituentID field. Artist records in MoMA include metadata elements such as “ConstituentID”, “DisplayName”, “ArtistBio”, “Nationality”, “Gender”, “BeginDate”, “EndDate”, “Wiki QID”, “ULAN”.

When retrieving data for *A Tribute to John Cage*, differences among the three repositories become apparent, as shown in Figures 1, 2, and 3. In the CreditLine field, Mplus specifies the holding institution, whereas MoMA describes the method of acquisition.

In the case of MoMA, Classification and Department are recorded separately as “Video” and “Media and Performance,” respectively. In contrast, Mplus assigns the terms “moving image” and “video” within its classification area and category fields. The Mplus dataset also provides multilingual name descriptions in English, Chinese, and alphabetical formats (Figure 2). Furthermore, Paik’s role is generally defined across multiple categories, including artist, archive creator, and subject. ZKM stands out for its detailed recording of contributors and their roles, such as listing Charlotte Moorman as Darsteller/in (performer) (Figure 3).

They provide different information regarding the duration: ZKM records the runtime as 29 minutes and 15 seconds, while MoMA and Mplus describe it as 29 minutes and 2 seconds.

```
{'Unnamed: 0': {20: 95276},
'Title': {20: 'A Tribute to John Cage'},
'Artist': {20: 'Nam June Paik'},
'ConstituentID': {20: '4469'},
'ArtistBio': {20: 'American, born Korea. 1932-2006'},
'Nationality': {20: 'American'},
'BeginDate': {20: '1932'},
'EndDate': {20: '2006'},
'Gender': {20: 'male'},
'Date': {20: '1973/1976'},
'Medium': {20: 'Video (color, sound)'},
'Dimensions': {20: '29:02 min.'},
'CreditLine': {20: 'Gift of the artist'},
'AccessionNumber': {20: '540.1992'},
'Classification': {20: 'Video'},
'Department': {20: 'Media and Performance'},
'DateAcquired': {20: '1992-02-15'},
'Cataloged': {20: 'Y'},
'ObjectID': {20: 122335},
'URL': {20: 'https://www.moma.org/collection/works/122335'},
'ImageURL': {20: nan},
'OnView': {20: nan},
'Circumference (cm)': {20: nan},
'Depth (cm)': {20: nan},
'Diameter (cm)': {20: nan},
'Height (cm)': {20: nan},
'Length (cm)': {20: nan},
'Weight (kg)': {20: nan},
'Width (cm)': {20: nan},
'Seat Height (cm)': {20: nan},
'Duration (sec.)': {20: 1742.0}}
```

Figure 1. *A Tribute to John Cage* at MoMA (Source: MoMA GitHub)

```
{'constituent_id': {39: 359},
'constituent_name': {39: 'Nam June Paik'},
'constituent_nameOther': {39: '白南準'},
'constituent_alphaSortName': {39: 'Paik, Nam June'},
'constituent_displayBio': {39: 'American, born Korea. 1932-2006'},
'constituent_gender': {39: 'Male'},
'constituent_beginDate': {39: 1932},
'constituent_nationality': {39: 'American'},
'constituent_roles': {39: 'Artist, Archive Creator, Subject'},
'constituent_type': {39: 'Individual'},
'object_id': {39: 20001},
'object_number': {39: '2021.50'},
'object_title': {39: 'A Tribute to John Cage'},
'object_dimension': {39: 'duration: 29 min. 2 sec.'},
'object_displayDate': {39: '1973, re-edited 1976'},
'object_medium': {39: 'single-channel digital video (colour, sound)'},
'object_creditLine': {39: 'M+, Hong Kong'},
'object_classification_area': {39: 'Moving Image'},
'object_classification_category': {39: 'Video'},
'constituent_vitalPlaceActor': {39: 'Korea'},
'constituent_endDate': {39: '2006'}}
```

Figure 2. *A Tribute to John Cage* at Mplus (Source: Mplus)

Artist / Artist group	Nam June Paik
Title	A Tribute to John Cage
Year	1973
Category	Video
Format	Analog video
Material / Technique	U-Matic, color, stereo
Dimensions / Duration	00:29:15
Contributors	John Cage, Darsteller/in Micky Lemle, Schnitt Cinish Bhargava, Schnitt Pat Kane, Schnitt John McKnight, Schnitt Tom McCorley, Schnitt David Hutton, Schnitt David Crane, Schnitt Richard Teitelbaum, Darsteller/in David Behrmann, Darsteller/in David Tudor, Darsteller/in Maryanne Amacher, Darsteller/in Charlotte Moorman, Darsteller/in
Collection	ZKM Zentrum für Kunst und Medien

Figure 3. *A Tribute to John Cage* at ZKM (Source: ZKM)

4. Data Integration and Transformation

4.1 Data Alignment with LIDO Schema: MoMA, Mplus, and ZKM

This section describes the process of representing institutional metadata using the LIDO(Lightweight Information Describing Objects) Schema. LIDO is structured into descriptive metadata, which captures object-related information, and administrative metadata, which supports collection management and documentation processes. It also follows a hierarchical structure in which wraps function as higher-level containers and sets are nested within them.(Figure 4) This study focuses primarily on the descriptive components of LIDO, as they are central to modeling artistic content and contextual information.

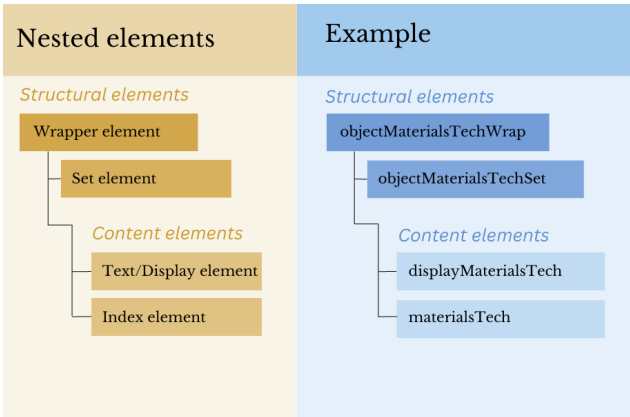


Figure 4. Wrap and Set construction in LIDO (Source: LIDO-Primer)

While LIDO provides a structured framework for metadata representation, it is designed to operate in conjunction with external controlled vocabularies, supporting thesaurus relationships, multilingual expressions,

and interoperability in the cultural heritage domain.¹¹ To enhance semantic precision and interoperability across institutions, standardized vocabularies such as the AAT(Art & Architecture Thesaurus) by Getty and identifiers from Wikidata are linked to LIDO elements in this research. The complete description code is available on GitHub.

4.1.1 ObjectWrap

In LIDO, defining type is a fundamental requirement for structuring object data consistently and semantically. Object-related information is primarily organized into two structural components: “objectClassificationWrap” and “objectIdentificationWrap”. The former is responsible for defining the conceptual type of the object, while the latter contains identifying and descriptive attributes such as title, measurement, and material or technique.

The most basic classification is established through “objectWorkType”, which specifies the conceptual category of the work. LIDO provides a controlled vocabulary for object types, and within this scheme, the material or technique (lido00789) is identified. Once the type is defined, an appropriate value must be assigned. In this case, video art was selected as the “objectWorkType” and was linked to the AAT via SKOS. This enables multilingual extensions through skos:prefLabel and skos:altLabel, as seen in Figure 5.

```
<lido:objectWorkTypeWrap>
  <lido:objectWorkType lido:type="http://terminology.lido-schema.org/lido00789">
    <skos:Concept rdf:about="https://vocab.getty.edu/aat/300102067">
      <skos:prefLabel xml:lang="en">video art</skos:prefLabel>
      <skos:prefLabel xml:lang="de">Video Kunst</skos:prefLabel>
      <skos:altLabel xml:lang="en">media art</skos:altLabel>
      <skos:altLabel xml:lang="de">Medienkunst</skos:altLabel>
    </skos:Concept>
  </lido:objectWorkType>
</lido:objectWorkTypeWrap>
```

Figure 5. ObjectWorkType

Measurements for media artworks require special treatment, as traditional LIDO measurement types are often insufficient for time-based works. Since duration is the primary measurable property for video art, the “measurementType” was extended using the Wikidata property Duration(P2047). Duration values were modeled as grouped pairs of units and values, following the conventional time structure of hours, minutes, and seconds. Each temporal unit was assigned its own “measurementSet”, ensuring that duration values remain machine-readable and semantically distinguishable.

For material and technique descriptions, AAT controlled vocabulary URLs were applied. In the case of Mplus, the displayed term “single-channel digital video (colour, sound)” was semantically decomposed into three distinct AAT concepts: single-channel video, color, and sound art. Each component was encoded individually to ensure that it remained semantically traceable and interoperable. Similarly, in the case of ZKM, the components U-matic, color, and stereo were treated as separate entities. At MoMA, the terminology used in the field of material, namely Video (color, sound), is more abstract in terms of the video component, as it can be ambiguous between the medium itself and a specific classification as video art. Therefore, color and sound were assigned as individual conceptual components. Additionally, each AAT term was directly linked via SKOS, enabling the displayed terms to be formally associated with standardized concepts.

4.1.2 EventWrap

The eventWrap structure was divided into two major event types: Production and Acquisition. The Production event focuses on the creator and creative context of the work, while Acquisition records the collecting institution and the date of collection.

In the Production event, the most central component is the description of the artist. LIDO provides a structured mechanism to represent personal and biographical information in a semantically organized manner. The element “nameActorSet” enables the expression of multilingual names, allowing the artist’s identity to be represented across different language contexts. Nationality is mapped using “nationalityActor”, the place of birth is expressed through “vitalPlaceActor”, birth and death dates are recorded in “vitalDatesActor”, and gender is represented via “genderActor”. Together, these elements provide a comprehensive profile of the artist associated with the production event.

LIDO also allows spatial information to be described using GML (Geography Markup Language), which records geographic coordinates in terms of longitude and latitude. Furthermore, hierarchical relationships between places can be expressed using “partOfPlace”, enabling the modeling of nested geographic structures from city to country level. For example, John Cage’s birthplace, Los Angeles, can be modeled in a hierarchical structure as Los Angeles – California – United States, with each place classified according to its corresponding classification such as City, State, and Country. Practically, this relationship can be understood as a reverse-nested structure, therefore the structure follows as: [Los Angeles [California [United States – Country] State] City].

A notable feature of the ZKM dataset is its representation of multiple contributors and their corresponding roles. Each participant involved in the production event is assigned a specific function, such as film editor or performing artist, which can be formally encoded using the “roleActor” element, additionally extended using AAT. This approach allows for a more detailed modeling of collaborative artistic production.

In the case of Acquisition events, the owning institution is modeled as an actor as well. When a corresponding identifier is available in the Getty ULAN (Union List of Artist Names), the institutional name can be linked using the ULAN URL as actorID. For example, MoMA is linked via <https://vocab.getty.edu/ulan/500303609>, and ZKM via <http://vocab.getty.edu/ulan/500300790>. This use of controlled identifiers enhances interoperability and enables institutional entities to be consistently referenced across datasets.

4.2 Mapping LIDO to CIDOC CRM

At the core of the transformation, each artwork described in LIDO is modeled in CIDOC CRM as an instance of “E22:Man-Made Object”. Object classification information is represented as typological assignment rather than literal values. For example, object types are modeled as instances of “E55:Type” and linked to the artwork via the property, “P2:has type”. Controlled vocabulary references from AAT and Wikidata are preserved as identifiers, allowing classification terms to serve as interoperable semantic nodes across datasets.

Titles are modeled as explicit appellations using “P102:has title”. Measurements are interpreted as structured “E54:Dimension” entities. Each dimension is associated with a unit, “E58: Measurement Unit” and a value, enabling consistent representation of size, weight, or other quantifiable attributes.

Materials and techniques are represented as “E57:Material”, connected to the artwork through “P45:consists of” and classified via “P2:has type” using AAT terms. This allows each material or technique to exist as a semantic entity that can be queried or grouped across contexts.

Production events are modeled as instances of “E12:Production”, linking the artwork through “P108:was produced by”. Artists and contributors are represented as “E39:Actor”, including identifiers (E42) and appellations (E41). In the case of multilingual names, it accommodates multiple language expressions using “rdfs:label xml:lang”, ensuring that individuals and institutions can be consistently referenced. Variant forms can also be represented using “P139:has alternative form”.

Acquisition events are represented as “E8:Acquisition”, connected to the artwork via “P24i:changed ownership through”, making changes of ownership explicit historical events.

This overall modeling description is also available on GitHub. Through this approach, all LIDO metadata is

transformed into a structured, relational semantic model in CIDOC CRM, where artworks, people, institutions, materials, and events are explicitly represented and interconnected.

4.3 Result: Querying using SPARQL

This section presents the results through SPARQL queries, based on the transformed data from the previous section. In particular, it focuses on missing information in the contributor field, which is available in ZKM but not in MoMA or Mplus, as well as on inconsistent information in measurements. Therefore, these issues are examined in relation to complementarities across repositories.

```
sparql_query = """
SELECT ?value ?unit_label ?type_label
WHERE {
  ?object a crm:E22_Man-Made_Object ;
         crm:P43_has_dimension ?dimension .

  ?dimension crm:P90_has_value ?value ;
             crm:P91_has_unit ?unit ;
             crm:P2_has_type ?type .

  ?unit rdfs:label ?unit_label .
  ?type rdfs:label ?type_label .

  FILTER(LANG(?unit_label) = "en")
  FILTER(LANG(?type_label) = "en")
}
"""

29 minute(minutes) 2 second
29 minute(minutes) 2 second
29 minute(minutes) 15 second
```

Figure 6. Duration

The first point is the representation of differences in duration across MoMA, Mplus, and ZKM: MoMA and Mplus record the duration as 29 minutes and 2 seconds, whereas ZKM indicates a duration of 29 minutes and 15 seconds. As shown in Figure 6, each of these values can be gained. However, the provenance of the information cannot be traced.

Secondly, the representation of information regarding the list of participating artists differs across repositories: ZKM provides a detailed list of participating artists, whereas MoMA and Mplus do not. Through the integration of datasets across these repositories, the missing data in MoMA and Mplus can be filled in, as illustrated in Figure 7.

Furthermore, ZKM provides the roles of each artist. When this aspect is queried, the query is structured as shown in Figure 8. In this case, the output shows every possibility of pairs. This situation is the so-called cross-product problem, which generally refers to combining every element of one set with every element of another. For example, if Set A = {a1, a2} and Set B = {b1, b2}, the result is {a1, b1}, {a1, b2}, {a2, b1}, {a2, b2}, representing all possible pairs. Therefore, when a query variable is not correctly linked to its related element, such a structure generates all possible combinations rather than the intended relationships.

```
sparql_query = """
SELECT DISTINCT ?actor_name WHERE {

    ?production a crm:E12_Production .
    ?production crm:P14_carried_out_by ?actor .

    ?actor crm:P1_is_identified_by ?id .
    ?id a crm:E41_Appellation ;
        rdfs:label ?actor_name .
    FILTER(LANG(?actor_name) = "en")
}
"""
query_result = g.query(sparql_query)

for row in query_result:
    print(f"Actor: {row.actor_name}")
```

Actor: Nam June Paik
 Actor: Cage, John
 Actor: Micky Lemle
 Actor: Cirish Bhargava
 Actor: Pat Kane
 Actor: John McKnight
 Actor: Tom McCorley
 Actor: David Hutton
 Actor: David Crane
 Actor: Richard Teitelbaum
 Actor: David Behrmann
 Actor: David Tudor
 Actor: Maryanne Amacher
 Actor: Charlotte Moorman

Figure 7. Contributing Artists

```
sparql_query = """
SELECT DISTINCT ?actor_name ?role_name WHERE {
    ?production a crm:E12_Production .

    ?production crm:P14_carried_out_by ?actor .
    ?actor crm:P1_is_identified_by ?actor_id .
    ?actor_id a crm:E41_Appellation ;
        rdfs:label ?actor_name .
    FILTER(LANG(?actor_name) = "en")

    ?production crm:P14.1_in_the_role_of ?role .
    ?role crm:P1_is_identified_by ?role_appellation .
    ?role_appellation a crm:E41_Appellation ;
        rdfs:label ?role_name .
}
"""
query_result = g.query(sparql_query)

for row in query_result:
    print(f"Actor: {row.actor_name}, Role: {row.role_name}")
```

Actor: Nam June Paik, Role: performing artists
 Actor: Nam June Paik, Role: film editors
 Actor: Cage, John, Role: performing artists
 Actor: Cage, John, Role: film editors
 Actor: Micky Lemle, Role: performing artists
 Actor: Micky Lemle, Role: film editors
 Actor: Cirish Bhargava, Role: performing artists
 Actor: Cirish Bhargava, Role: film editors
 Actor: Pat Kane, Role: performing artists
 Actor: Pat Kane, Role: film editors
 Actor: John McKnight, Role: performing artists
 Actor: John McKnight, Role: film editors
 Actor: Tom McCorley, Role: performing artists
 Actor: Tom McCorley, Role: film editors
 Actor: David Hutton, Role: performing artists
 Actor: David Hutton, Role: film editors
 Actor: David Crane, Role: performing artists
 Actor: David Crane, Role: film editors
 Actor: Richard Teitelbaum, Role: performing artists
 Actor: Richard Teitelbaum, Role: film editors
 Actor: David Behrmann, Role: performing artists
 Actor: David Behrmann, Role: film editors
 Actor: David Tudor, Role: performing artists
 Actor: David Tudor, Role: film editors
 Actor: Maryanne Amacher, Role: performing artists
 Actor: Maryanne Amacher, Role: film editors
 Actor: Charlotte Moorman, Role: performing artists
 Actor: Charlotte Moorman, Role: film editors

Figure 8. Role of Each Artist

5. Extended Semantic Modeling using LIDO RelationWrap and CIDOC CRM

Based on the context of *A Tribute to John Cage* in Section 3, this part supplements missing information in institutional datasets using the LIDO metadata standard.

In LIDO, additional information about individual objects can be recorded in relatedWorksWrap, which is placed under objectRelationWrap at the same hierarchical level as objectIdentificationWrap. Relationships between objects are defined via relatedWorkRelType, using controlled vocabulary provided by LIDO.

In this part, the relationship “is related to” (lido:00263) was used to link John Cage’s 4’33”(<https://www.moma.org/collection/works/163616>) to *A Tribute to John Cage*, while “has part” (lido:00573) was employed to associate one of the multiple footage components, Nam June Paik & Jud Yalkut’s *Waiting for Commercials*(1966–1972/1992) via <https://www.moma.org/collection/works/120523>. These works are stored at MoMA, where each work has a unique identifier and a corresponding URL. For convenience in RDF representation, the MoMA URLs were used in the RDF dataset.

When converting the LIDO records to CIDOC CRM, the “is related to” relationship can be mapped to multiple properties. Considering both semantic and logical relationships between entities, it was mapped to “P15:was influenced by”, whereas “has part” was mapped to “P5:consists of”. The corresponding LIDO and CIDOC CRM descriptions are available on GitHub as well.

Based on these datasets, SPARQL queries were executed to retrieve related works. The results, shown in Figures 9 and 10, demonstrate that the context of the artwork, *A Tribute to John Cage*, can be extended to other related works.

```
query = f"""
SELECT ?relatedWork ?title
WHERE {{
  <{aTribute}> crm:P15_was_influenced_by ?relatedWork .
  ?relatedWork crm:P102_has_title ?title .
}}
"""

for row in g.query(query):
    print(f'{sub} is influenced by URI: {row.relatedWork} , titled "{row.title}"')
```

<https://www.moma.org/collection/works/122335> is influenced by URI: <https://www.moma.org/collection/works/163616> , titled "4'33"

Figure 9. Relationship Representation with P15

```
query = f"""
SELECT ?relatedWork ?title
WHERE {{
  <{aTribute}> crm:P5_consists_of ?relatedWork .
  ?relatedWork crm:P102_has_title ?title .
}}
"""

for row in g.query(query):
    print(f'{sub} consists of URI: {row.relatedWork} , titled "{row.title}"')
```

<https://www.moma.org/collection/works/122335> consists of URI: <https://www.moma.org/collection/works/120523> , titled "Waiting for Commercials"

Figure 10. Relationship Representation with P5

6. Limitations and Discussion

In terms of LIDO as a middle-level schema derived from CIDOC CRM, its mapping is structured and standardized. From this perspective, transforming and extending metadata through LIDO and CIDOC CRM generally poses no significant difficulties. However, as observed in Section 5 regarding relatedWork, interpretive

variability in the mapping process from LIDO to CIDOC CRM can occur. Specifically, the researcher's interpretation during mapping may affect semantic expansion in semantic alignment, considering the defined relationship entities and properties in CIDOC CRM.

Most importantly, collaboration between institutions and technical experts is essential. First, a major limitation concerns the accessibility of data for linked data implementation. To contextualize Nam June Paik's works and the associated contexts described in Section 2—for example, the Darmstadt Music Festival Program from the Darmstadt Music Festival Archive or Baumeister's interview from the Städel Museum—open-access or exportable datasets are essential. However, many repositories do not provide their data in open formats or with export functions, which limits integration and reuse in distributed research environments. From a technical perspective, datasets should ideally be accessible through standardized web-based interfaces. Without such infrastructure, reliable retrieval and cross-system interoperability remain constrained.

Second, differences in the usage of standards can be observed. In practice, even if institutions provide metadata and controlled terms, internal guidelines or documentation for their usage may exist, but controlling the diversity in application is inherently difficult. From a technical standpoint, it is therefore necessary to develop systems that can semantically analyze variably interpreted metadata-term pairs and automatically assign them during import processes. Additionally, issues arising from query results—such as provenance tracking—could be addressed using tools like Named Graphs or Prov-O in combination with collaborative workflows.

Furthermore, language and accessibility present additional challenges. For instance, the Korea Arts Management Service(KAMS) provides multilingual terminology and descriptions in Korean, Japanese, Chinese, and English, yet these resources cannot be incorporated into the integrated dataset due to non-reusable URLs.¹²(Figure 11) Addressing such limitations requires coordinated technical development and institutional collaboration to ensure interoperability and proper inclusion of multilingual resources.

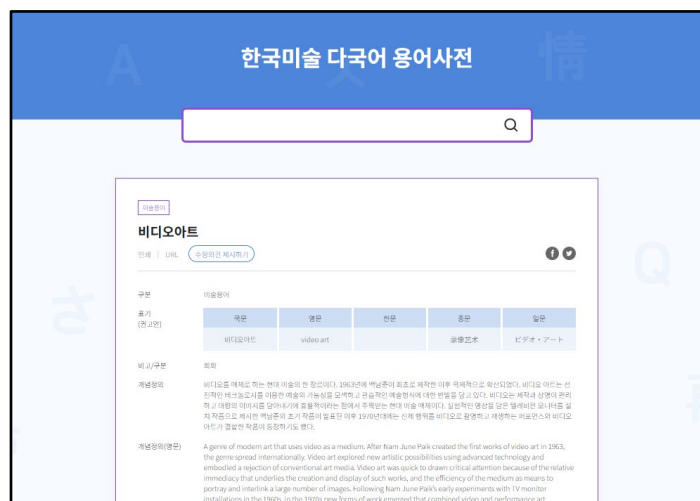


Figure 11. Multilingual Dictionary of Korean Art

7. Conclusion

This case study demonstrates how linked data can be applied to datasets from multiple repositories in the art domain. In the case of Nam June Paik's works, the multilingual nature of resources and the global dispersal of archival materials and artworks have made it difficult to gain a comprehensive understanding of his artistic world. Semantic modeling was proposed as an appropriate methodology to address these challenges, involving the mapping of institutional data to CIDOC CRM through the LIDO schema and querying the results using SPARQL.

During this process, the necessity of robust infrastructure—particularly collaboration between institutions and

technical experts—was identified, underscoring the importance of coordinated efforts for successful data integration.

Building on this semantic modeling approach, this research will further explore how Nam June Paik's various works, including *A Tribute to John Cage*, and archival materials dispersed across the world can be semantically linked, as well as investigate additional relationships within the Fluxus movement and connections across contemporary experimental art, while addressing technical challenges.

¹ Baca, M. (Ed.). (2016). *Introduction to Metadata* (3rd ed.). Getty Publications

² Berners-Lee, T., Hendler, J., & Lassila, O. (2001). "The Semantic Web." *Scientific American*, 284(5), 34–43.

³ CIDOC CRM Version 7.1.3. (https://cidoc-crm.org/html/cidoc_crm_v7.1.3.html)

⁴ LIDO-primer (<https://lido-schema.org/documents/primer/latest/lido-primer.html>)

⁵ <http://www.medienkunstnetz.de/works/hommage-a-cage/>

⁶ <https://www.moma.org/collection/works/120523>

⁷ <https://zkm.de/de/werk/tv-bra-for-living-sculptures-tv-cello-tv-bed>

⁸ <https://www.moma.org/collection/works/163616>

⁹ <https://github.com/MuseumofModernArt/collection>

¹⁰ <https://api.mplus.org.hk/en/documentation>

¹¹ It was developed by the ICOM-CIDOC LIDO Working Group, based on the standards such as CIDOC CRM, the ontological framework, and CDWA (Categories for the Description of Works of Art). LIDO version 1.0 was officially presented in 2010. Subsequently, LIDO version 1.1 was released in 2021.

¹² <https://www.gokams.or.kr/visual-art/art-terms>

References

Baca, M. (Ed.). (2016). *Introduction to Metadata* (3rd ed.). Getty Publications.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). "The Semantic Web." *Scientific American*, 284(5), 34–43.

CIDOC CRM Version 7.1.3. https://cidoc-crm.org/html/cidoc_crm_v7.1.3.html

Getty Vocabulary AAT. <https://vocab.getty.edu/aat>

Hanhardt, J. G., D. Ronte, M. Nyman, & D. A. Ross (Eds.). (1982). *Nam June Paik*. Whitney Museum of American Art.

Herzog, G., S. Rennert, & B. Jacobs (2005). "Nam June Paiks frühe Jahre im Rheinland". (ZADIK Ed.). *Sediment. Mitteilungen zur Geschichte des Kunsthandels* (Heft 9).

Iddon, M. (2013). *New Music at Darmstadt: Nono, Stockhausen, Cage, and Boulez*. Cambridge University Press.

LIDO Primer. <https://lido-schema.org/documents/primer/latest/lido-primer.html>

MoMA GitHub. <https://github.com/MuseumofModernArt/collection>

Mplus Developer Portal. <https://api.mplus.org.hk/en/documentation/about>

ZKM. <https://zkm.de/de/zkm-archiv>

Zerbib, D., Decker-Phillips, E., Lee, Y., Neuburger, S., Rothfuss, J., von Wiese, S., & Daniels, D. (2010). *Nam June Paik* (S. Rennert, S.K.Lee, Museum Kunst Palast, & Tate Liverpool, Eds.). Hatje Cantz Verlag.

데커, 에디트 & 리비어, 이르멜린 엮음. (2018). 백남준: 말에서 크리스토폰까지 (임왕준 외 역). 백남준아트센터.

레너르트, 수잔네 외. (2022). 백남준과 미디어아트, 그 시작: 1958-1963 독일 라인지역에서 (전선자 역). 성균관대학교출판부.

한국미술 다국어 용어사전(Multilingual Dictionary of Korean Art Terminology). <https://www.gokams.or.kr/visual-art/art-terms>