Awakening the Material in the Network of Things through Data-Driven Semantic Analysis

Background

Material in not just material or medium, especially in the art history of things, the history of artificial objects. However, material has been flatted and simplified during the process of archiving or metadating of objects as metadata of "material". This research explores the intricate relationship between materials and themes within a vast repository of 800,000 artificial objects around the world in the past more than 3000 years.

By revealing hidden narratives and cultural dimensions embedded in the metadata of objects, the study focuses on reshaping the concept of material in the art history by setting it in the network of things, including the images, metadata, AI annotated tags to complicating and enriching the meaning of material, not only as the material of physical objects but also the computational material, as Yuk Hui proposed in the *On the Existence of Digital Objects (*2016) Lighting on the interconnectedness and semantic networks of materials and themes, this research offers a fresh perspective on human civilization's artistic expressions.

Data Acquisition and Cleaning

This research involved the collection and curation of open data interfaces from an extensive selection of 76 museums, from which 9 highly renowned institutions were singled out for further investigation. The systematic acquisition of information primarily relied on a Python module known as Requests to construct HTTP requests and retrieve curated metadata about the museums' collections. The metadata encompassed structured textual fields, unstructured descriptive information, and imagery of the artifacts.

Museums, owing to their diverse collections of cultural artifacts and artworks, exhibition and display methodologies, exhibit a myriad of distinctions. These distinct facets are manifest in the disparate data requirements and digitization methods employed by each institution, thus engendering considerable variations in the standards and encoding conventions applied to metadata. This heterogeneity poses formidable challenges in achieving uniform data processing and curation. Therefore, during the data curation process, we conducted varying degrees of data cleansing on each of the nine museums.

By scrutinizing the unique identifiers assigned to the collections within each museum, we successfully eliminated redundant entries. Subsequently, meticulous attention was given to rectify instances of evident file format or encoding errors present in the collection data. Furthermore, after diligently consulting the API documentation provided by each museum, we supplemented missing field information for select artifacts and adeptly extracted crucial details from redundant formats.

To establish harmonization and consistency within the dataset, the semantic processing of collection information entailed the fusion of fields bearing analogous meanings across the metadata tables of the nine museums. Therefor, universally accepted metadata labels were thoughtfully assigned to these merged fields. Regarding pivotal label data, such as classification, material, technique, and temporal attributes, we executed a normalization procedure to achieve a unified and standardized set of label values across the entire dataset. Due to the disparate and heterogeneous nature of metadata among museums, and to circumvent excessive data sparsity, in addition to a unified information table, we curated individualized supplementary information tables for each museum. Each is tailored to its specific characteristics. For instance, while all nine museums share common fields such as name, continent, and source, the supplementary information table

for the Victoria and Albert Museum includes exclusive details such as styles, production, and summary. Processed textual information was separately recorded in XLSX files for expert perusal and MongoDB databases for ease of bulk operations. Image data was uniformly transformed into the JPG format using a Python module for image processing called OpenCV.

The resulting dataset encompasses a total of 821,189 entries, originating from museums spanning four continents, six countries, and represented in five different languages. Additionally, the dataset comprises 1,170,343 images. The curated collection data encompasses a diverse array of information, including the historical background, artistic value, cultural significance, and spatiotemporal origins of the cultural artifacts. The dataset can readily facilitate further research, and the data cleansing process has significantly enhanced the accessibility and retrievability of the subsequent data.
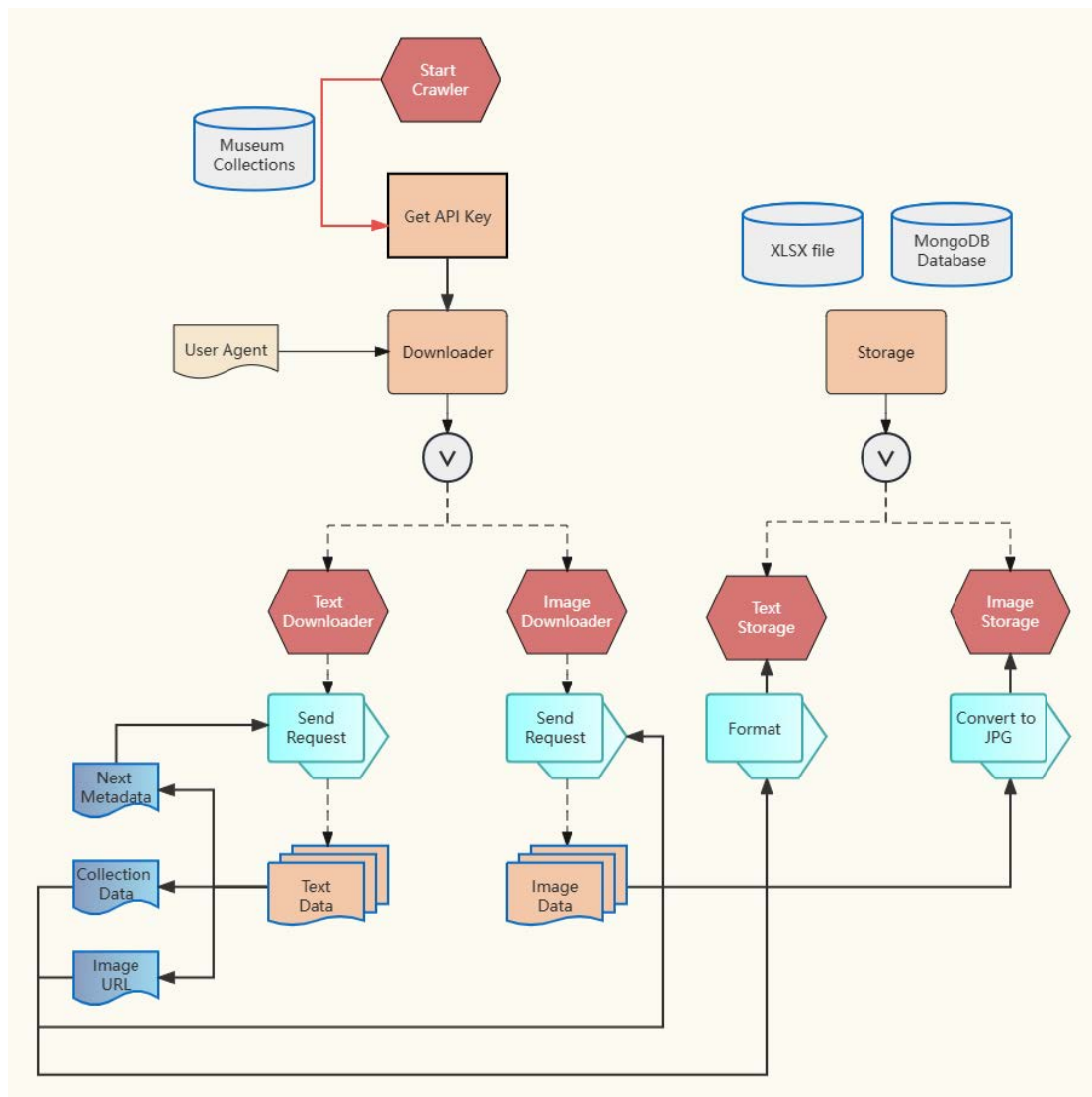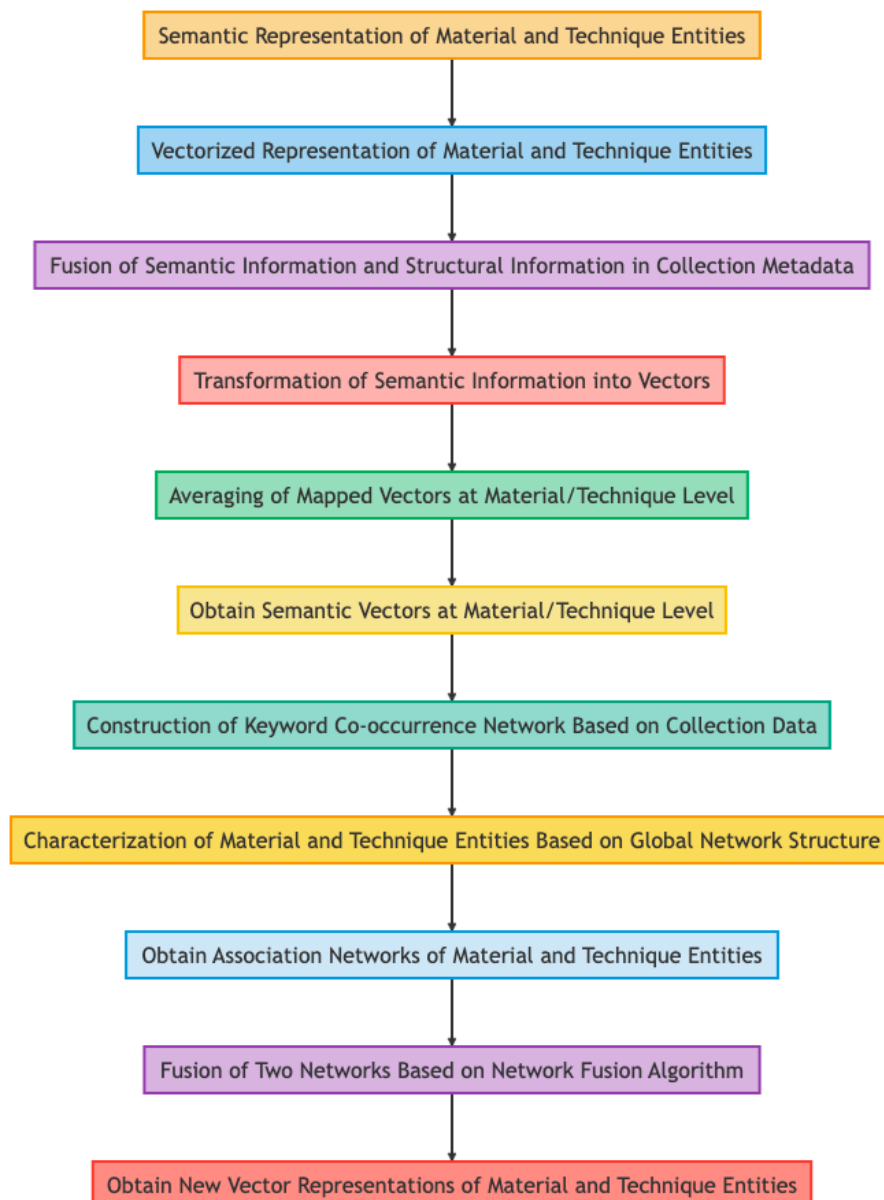

Process Flowchart for Museum Data Acquisition and Cleaning

Analysis Process

The research delves into the inherent content and narrative threads that bind diverse art collections. Combining Latent Dirichlet Allocation (LDA) for theme clustering and manual filtering guided by art historical expertise, shared themes and motifs transcend temporal

and spatial boundaries, transforming our understanding of the role of material in the artificial objects.

In the stage of semantic representation of material and technique entities, this research carries out vectorized representation of material and technique entities by fusing semantic information and structural information in the collection metadata. Firstly, the semantic information such as titles and descriptions of collections are transformed into vectors through textual representation learning, and all the mapped vectors are averaged at the material/technique level to obtain the semantic vectors at the material/technique level. Secondly, A keyword co-occurrence network, built upon collection keywords, themes, and categories, characterizes material and technique entities based on the global network structure. Finally, based on the semantic vectors and network structure information, the association networks of material and technique entities are obtained respectively, and the two networks are fused based on the network fusion algorithm to obtain the brand-new vector representations of material and technique entities respectively.

```
Semantic Representation of Material and Technique Entities
                            ↓
Vectorized Representation of Material and Technique Entities
                            ↓
Fusion of Semantic Information and Structural Information in Collection Metadata
                            ↓
Transformation of Semantic Information into Vectors
                            ↓
Averaging of Mapped Vectors at Material/Technique Level
                            ↓
Obtain Semantic Vectors at Material/Technique Level
                            ↓
Construction of Keyword Co-occurrence Network Based on Collection Data
                            ↓
Characterization of Material and Technique Entities Based on Global Network Structure
                            ↓
Obtain Association Networks of Material and Technique Entities
                            ↓
Fusion of Two Networks Based on Network Fusion Algorithm
                            ↓
Obtain New Vector Representations of Material and Technique Entities
```

Results

The findings offer insights into artistic expressions, e.g., artworks made of silk are intricately associated with themes of 'Family', 'Architecture', 'Religion', and 'Mythology'.

The identified keywords, such as 'black', 'border', 'design', 'gold', 'ground', 'kesi', 'marriage', 'pattern', 'ribbon', 'robe', 'round', 'scattered', 'silk', 'similar', 'sleeves', 'straight', 'swastika', 'tapestry', 'thread', and 'tissue', were obtained through text analysis of the ' Description' column in the dataset, collectively revealing specific material attributes and characteristics of artworks made of silk within the theme of "Religion and Mythology."
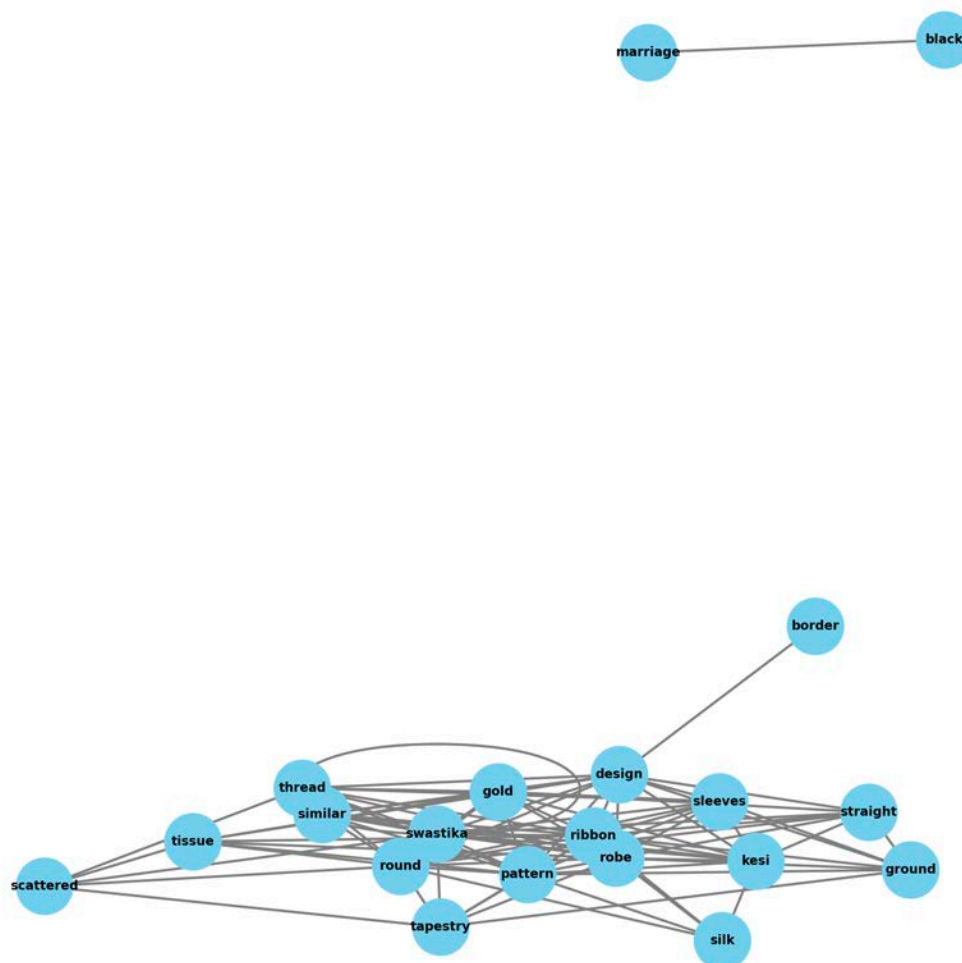
the visualized graph based on the co-occurrence matrix data of the identified keywords
('Design' is a central keyword as it co-occurs with many other keywords such as 'gold', 'ground', 'kesi', 'pattern', 'ribbon', 'robe', and 'round'. This suggests that the theme 'Religion and mythology' in silk artworks often incorporates various designs.

'Black' and 'marriage' have very limited connections with other keywords, suggesting that these terms may not play significant roles in the theme of 'Religion and mythology' in silk artworks, or they may

represent unique sub-themes that do not commonly intersect with other themes.

'Silk' is isolated, which could suggest that the term 'silk' (perhaps a descriptor for the medium of the artwork) was not often used in conjunction with other descriptive keywords in the context of the artworks' 'Religion and mythology' theme.

Some keywords like 'pattern', 'ribbon', 'robe', and 'round' form a tightly interconnected group with 'design', suggesting these elements often appear together and could represent a common motif or style in silk artworks with a 'Religion and mythology' theme.

The keywords 'thread', 'tapestry', 'tissue', and 'swastika' also appear to co-occur with several other keywords, indicating their importance in this theme. Their roles could be related to the medium (like 'thread', 'tapestry', 'tissue') or specific symbols (like 'swastika') used in these silk artworks.)

The presence of terms related to material composition ('silk', 'thread') emphasizes the fundamental materiality of these artworks, descriptive words ('black', 'gold', 'pattern', 'design') underscore the visual elements of silk artworks, indicating the presence of intricate patterns, rich colors, and artistic designs that contribute to the aesthetic appeal of the artworks within the specified theme, thematic motifs such as 'swastika', a religious symbol, suggest the incorporation of significant religious symbolism in silk artworks, further strengthening their association with religious contexts. The mention of 'tapestry' and 'robe' implies the utility and ceremonial significance of silk in religious rituals or mythological representations.

Here are some artworks related to the collection :

Main_ID	Museum	Title	category	origin	dynasty	date	material	description	technique
136246	the MET	Canonization scroll of Li Zhong	Paintings	—	Ming dynasty (1368 – 1644)	colophon dated 1641	ink color, and gold on silk	This painting, an unsigned work of a provincial atelier, is a rare example of popular religious art. Modeled on imperial commendation scrolls, with their ornamental borders of dragons chasing flaming pearls, the scroll illustrates the investiture of a local god into the Daoist pantheon. The first half of the scroll depicts the court of the Jade Emperor, supreme deity of popular Daoism. The enthroned emperor is flanked by four female attendants, four high officials, and two constellations of star deities, as well as guardians of the east and west, who are accompanied by their directional emblems, the dragon and tiger. Approaching this assembly is a file of haloed deities who have entered the sacred precincts through a pair of heavenly gates. The second half of the scroll depicts the entourage of the local god, including standardbearers with emblems of his status, a white charger, a palanquin with eight porters, guardians of the four directions, four attendants carrying incense burners, and figures bearing his sword and seals. This procession ends with a depiction of the local god, shown larger than life to emphasize his importance. In the sky above, a heavenly messenger appears bearing the scroll of the god's investiture. Following the painting is a transcription of the petition made to the Celestial Master of the Orthodox Unity sect, headquartered at Dragon-Tiger Mountain (Longhu shan) in Jiangxi Province, requesting that the local god, Li Zhong, be recognized and as a member of the pantheon. The petition is dated A.D. 1157, but both it and the painting are likely to be copies, probably created during the Ming dynasty. The terminus ante quem for the scroll is established by two appended colophons, both dated 1641.	Handscroll

Museum	Title	category	origin	dynasty	date	material	description		technique
the V&A	Panel	Silk (textile)	China (possibly Chaozhou, made)	—	1863 (made)	Silk (textile)	Hanging panel of red silk satin embroidered with silver thread and coloured silks in satin and stem stitches, with couched work. It is a celebratory hanging for Mr Huang's 90th birthday; a long inscription bearing the names of the donors in gold leaf is embroidered on crimson ground. Around it is a border of fret pattern and peonies and chou characters. Another border shows a scene on the top: an Emperor and Empress surrounded with attendants are seated by a shrine before which a small figure is kneeling in worship. The sides and bottom of this border is filled with depictions of figures in landscape or pavilions. A narrow band of fret pattern and peonies divides this border from the next, depicting birds including peacock, crane, mandarin duck and phoenix amidst flowering branches of prunus blossom, peony, magnolia, and lotus; at the bottom are three dogs of Fo. At the top is a five-clawed dragon in between two phoenixes with clouds in the background. From the top of the hanging suspend two narrow pendants decorated with Taoist letters, bats, citrus fruits and lotus blossoms.		Embroidered

Overall, the identified keywords enhance the materiality of silk as an artistic medium by elucidating its unique visual qualities, symbolic significance, and contextual relevance within the realm of "Religion and Mythology" themes. This contextualization facilitates a deeper understanding and appreciation of the distinctive artistic expressions embedded in silk artworks associated with religious and mythological subjects.

Intriguingly, this interdisciplinary approach enriches our appreciation of material and materiality in the global and shared cultural experiences that shape our collective artistic legacy. Ultimately, this research redefines materiality, illuminating the artistic process, the transcendent connections between materials and themes, and cultural narratives echoing through human civilization's annals.
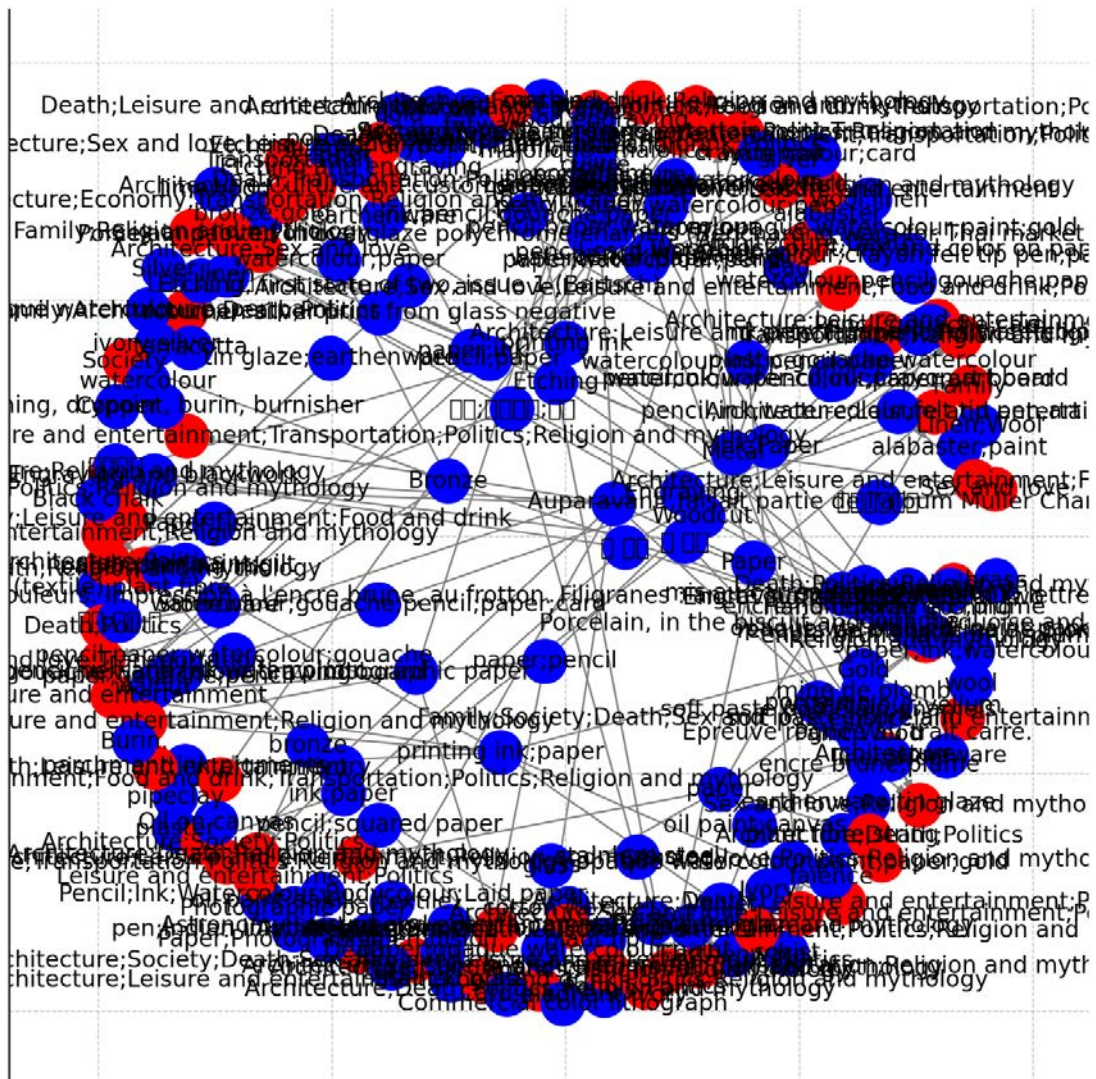
Here is a simplified network graph showing the connections between the top 10 materials and themes based on the sample dataset. The blue nodes represent materials and the red nodes represent themes. The edges between nodes represent associations between materials and themes.The network graph has been simplified for clarity, and only includes materials and themes that appear in at least 5 artworks.

Here is a simplified network graph showing the connections between the top 10 materials and themes based on the sample dataset.