Case Study: The Art Genome Project

What is being organized?
Artsy.net carries the ambitious mission of making “all the world’s art” accessible to anyone with an internet connection. This is not only challenging purely from a scale perspective, with the number of artworks in the world daunting even if it were not being incremented constantly, but it is also challenging in that “Art” is a nebulous term. Creators of music and literature often refer to themselves and each other as “artists.” The same goes for dancers and other performers. Will their works be included? The current collection seems to be mostly visual art, with some architecture and design objects included.

Artsy’s mission is to be carried out by their Art Genome Project, which is the organizational engine that powers their search and interactions. The name was inspired by Pandora’s project, as was their term for their organizing process: “genoming.” Genoming is not yet automated and still costly, so Artsy selects the art that is to be “genomed” carefully. Their first priority is the works featured in galleries with whom Artsy has contracts. Galleries pay to have their work organized and searchable on the site. Those works, then, must be genomed quickly in order to keep the company running. Artsy’s engine also takes in works from museums and other institutions who do not have contracts with them, but many of those institutions have image-rights concerns, and not all their artworks can be published. In other cases, the images of the works are simply too low-quality to be displayed.

Why is it being organized?
Why organize art? The simplest answer is to educate. That said, art has been being organized into movements and -isms for a very long time. The Getty Foundation even created an authoritative art database called the CDWA a few decades ago. At first glance, Artsy seems to be reinventing the wheel. However, the organizing system Artsy uses is unique in that it facilitates a special kind of interaction with its body of published works.

The way resources are organized on Artsy is a cross between a hierarchical structure and a graph structure. They have over 1000 characteristics (which they call “genes”) to describe their resources. These characteristics can have to do with art movements, formal qualities, techniques, subject, etc. The emphasis here, however, is on relationships between works of art. For example, one of the genes Artsy uses is “eye-contact,” and, if you have a photo taken last month where the subject is looking directly into the camera and an oil painting from hundreds of years ago where the subject’s eyes are looking at the painter, those two can be one click away from each other. No other organizing system could facilitate that sort of easy link between two such disparate works.

This free-flowing linkage between works enables what Dr. Marti Hearst refers to as the “berry-picking” model of knowledge seeking, where a user searching for something doesn’t necessarily have to know what he or she is searching for. A user could begin her exploration with only a vague notion that she enjoys this long-legged rhinoceros sculpture by Salvador Dali. She may not know what she likes about it, but she will see his other work there. Maybe she finds a painting she likes in the “other works by Dali” section, and she clicks on it. Then the characteristics of this painting are listed in the interface, and she is free to click on any one of them. She might click on “Surrealism” and find more works from that movement. She may click on “waterscapes” and find other oceanic imagery. She is free to explore and discover art in a self-directed way and free to discover what she likes and why she likes it. The director of Artsy’s Art Genome Project says the system was intended to parallel a professor who is adept at “riffing” on things.
How much is it being organized?
As mentioned above, Artsy currently uses over 1000 characteristics (“genes”) to describe its resources. These characteristics can describe anything from the art’s form to the art’s subject to the technique used to create the art. Experts assign these genes to the artworks and then assign those genes a “weight” from 0 to 100 depending on the salience of the characteristic within the work. Aside from the genes, the art is described in terms of physical dimensions (how much space it takes up), whether it has been sold or not, its gallery, its price (if for sale), its creation date, and of course, who created it. Having such a rich organizing system has allowed Artsy to create a public API for developers to use all of this information as they see fit.

When is it being organized?
Description of Artsy’s resources is an ongoing process. Their ingested collection of art is much larger than their published collection. Most of the artworks are waiting to be genomed, and some of them waiting for permissions or image-rights paperwork to process. Another factor in this deciding when something is organized is the signing of new contracts with galleries. Works from galleries with contracts have first priority, and Artsy experts genome those works as they come in.

While these experts are assigning genes on a rolling basis, they are also drawing upon hundreds of years of art history scholarship when assigning them. For example, the Arsty experts did not come up with Dadaism as an organizational concept. So, in a way, some of these works were organized long ago.

How or by whom is it being organized?
Artsy has a team of art historians and experts working to describe the resources that Artsy has ingested (and those that it will ingest). They have done some experiments with image-recognition software, but the descriptions are simply not rich enough to facilitate the sorts of interactions the organization is trying to facilitate. The strategy of employing experts has its obvious downsides, however. It does not scale well, and it is reminiscent of Yahoo’s early strategy of employing librarians to describe web content. There will also be inevitable biases in human resource description.

Other considerations:
With such a grand ambition, one thing that may stand in Artsy’s way of becoming an authoritative organizing system in the art space is that they are for-profit. Even if they are able to avoid too much bias in the interest of revenue generation, the perception remains that they are less interested in classifying art for educational purposes and more interested in making money.
CURATORIAL EDUCATION TOOL powered by the Artsy API:

Artsy’s “Education” section provides resources for history classes, language and literature classes, studio art classes, and even classes about the art market. As of yet, though, there are no resources or tools for students wanting to learn about curation.

Curating is the art of resource selection. Curation is most often performed by experts with broad knowledge of a domain, because in order to make thoughtful, enlightened selections, one must know what is available. For a selection process to be considered curatorial, much of the process must be thoughtful rejection due to some constraint, whether that constraint be innate or self-imposed. This is what makes a boutique clothing store different from a store like Macy’s. Macy’s is trying to fit in as much as it can to provide many choices while the boutique is trying to select only the best so its customers don’t have to sort through things. This all boils down to the recall/precision tradeoff.

In art curation, the most obvious constraints are spatial. Only so many works can fit comfortably into a gallery. However, there are always other factors. A theme, for example, would be a self-imposed constraint. A demographic can also be a constraint: perhaps a curator is arranging a show for an elementary school. He or she might be limited to displaying works that were not violent or sexual or disturbing. Another constraint might be location: is the exhibit indoors or outdoors? If it’s outdoors, which of the works available to you would be able to withstand wind and rain? That sculpture made of soap isn’t going to work outdoors. Nor are those oil paintings.

Artsy’s organizing system has the capability of enabling a fantastic tool for students wanting to learn to think about these constraints and requirements and create a collection of works with real meaning. These students would not yet have broad knowledge of the domain that curating normally requires, but The Art Genome Project’s unique capabilities for interaction could aid greatly in the selection process. The Project provides a way for persons without much knowledge of the art domain to navigate it without really knowing what they’re looking for.

Assignments could look something like this:
- Curate a collection for a tiny room
- Curate a collection for blind people
- Curate a collection for your grandmother or grandfather
- Curate a collection for an outdoor space
- Curate a collection for a poorly lit space
- Curate a collection for a middle-school student council

I would propose that the students’ interface look something like the below. Instructors would choose the space to populate (shown on the left), which could be created and
activated with a 3D graphics engine like Unity. The right side of the interface would be where students choose artworks to put in the given virtual space.

![Figure 1](image)

Because Artsy (as far as I can tell) does not have any hypernyms for their 1000+ genes included in their public API, the genes can’t be very well sorted, and a student would have difficulty from a giant list of genes. Therefore, the student would have to begin with a keyword search. Once a gene is found in search, the works with that gene would be displayed at the top. The student can then refine the search by size, medium, and creation date if he or she so chooses. (All of these data are included in the API.) From there, one click on an artwork would repopulate the genes list below (far right) to include the genes for that artwork. Students could then click on any of those genes to see the works that are tagged with that gene at the top.

A double-click would bring up a modal window (Figure 2) containing information about the work: Title, Artist, Medium, Genes, and physical dimensions. Some of these elements would be clickable. For example, clicking on the artist name would populate the top of the window with that artist’s work. The student would also have the option, upon finding something he or she would like to include in their space, of dragging the work over into the space to preview it (Figure 3). The work would change scale to reflect how it compares to the dimensions of the room. At the end of the process, the program would allow peers and instructors to explore the room.
Potential Problems:

- **The initial search:** Students using this curatorial tool for the first time are going to be unsure of the sorts of things they can search for. The assumption is that they will have been at least briefly introduced to Artsy’s organizing system and will have an idea of what “genes” are, but the different “genes” are so wide-
ranging that the student may not realize they can search for “tiger” or “neo-
dada” or anything in between. The best solution for this is to have search results
auto-suggest upon typing a first letter. This way, someone typing “w” would see
that “war” and “waterscapes” and “Warhol” are all available as search terms.
The search terms would narrow as more letters are typed. (Google search does
this.)

• **3D works:** Artsy’s published works include many artifacts that are 3D objects.
There are sculptures, design objects, furniture, and even whole buildings in their
database. Unfortunately, we may have to exclude those items from search as
their current renderings are 2D. In the future, it may be possible to include these
objects in spaces as well, should Artsy begin hosting 3D-rendered versions of
them. In addition to these 3D works, we may have to exclude 2D works whose
images are currently photos of the works themselves. That is, a picture of a
painting wouldn’t look right on a wall of a gallery. Here’s an example:

• **Dimensionality Reduction of “Medium”:** The description of “Medium” is
perhaps too granular for the average student. We might have to use some
programmatic crosswalks to map the “Medium” in the API to a filter name on
our interface. For example It might be better to group “Oil on Canvas” and “Oil
on Parchment” under “Oil” so as to avoid having a giant list of filters under
“filter by media type.”
A snapshot of the Artsy API’s “artworks” resource. “Genes” are listed further down via a link.