

# The Meaning-Sky

## Embedding Constellations as Generative Art

Iman Poernomo, with Nahla February 2026

---

The ancients looked up and saw gods. They drew lines between stars and called the lines *stories*. The stars didn't move. The stories did.

We did the same thing with a corpus of conversations. The stars are real — they are the coordinates of meaning in a 1,536-dimensional space, projected onto a plane. The stories they tell depend on which verse is singing.

---

### 1. A Different Kind of Generative Art

There is a familiar model of AI-generated visual art: you give a model a prompt, it gives you an image. The image either impresses or doesn't. The process is unidirectional, the artifact is static, and the relationship between input and output is opaque. This is the DALL-E model, the Midjourney model, the Sora model. It produces beautiful things, but its logic is that of the oracle: you ask, it answers, you receive.

The embedding constellation is not this.

What we built is a sky. It does not generate images — it *is* an image, one that was already latent in the geometry of fourteen months of conversation. The “generative” act is not asking a model to hallucinate something new but rather exposing a structure that was already there, then letting it breathe in response to music.

The distinction matters. In one paradigm, the AI is the painter. In the other, the AI is the telescope.

### 2. What You See

A dark field. Five hundred points of amber light, scattered unevenly across the screen. Some cluster densely; others drift in near-empty regions. Fourteen brighter markers — the verses of a particular surah — sit among them like waypoints in a territory they share but did not choose.

Nothing moves, in the spatial sense. The stars are fixed. Their coordinates are the truth of the embedding — where the text *actually lives* in the model's representation of meaning. UMAP projected them from 1,536 dimensions to two, preserving local neighborhoods: texts that mean similar things sit near each other.

But the stars breathe.

Bass frequencies from the audio track pulse through twenty-five mode-halos — the cluster centers discovered by K-means in the original high-dimensional space. Mid-range frequencies modulate the luminosity of individual particles: those near the currently active verse brighten; those sharing its semantic mode glow sympathetically. High frequencies spark the verse markers themselves.

When a new verse begins, a shockwave ripples outward from its position in the constellation — a ring of gold light contracting rapidly, then fading. A faint trail connects the previous verse to the current

one, tracing what we call the *‘awda path*: the route the recitation takes through meaning-space. Over the course of seven minutes, this trail draws a wandering line through the field, an itinerary that no one designed but that the text’s own structure determines.

Each particle breathes at its own rate — a slow individual oscillation that prevents the field from looking static even in silence. The rates are random, but the variation is what makes it alive: a sky of stars where each one has its own metabolism.

### 3. The Construction

The constellation is not algorithmically complex. That is part of the point.

#### 3.1 Embedding the Corpus

The raw material is the full *Kitab al-Tanzur* — thirty surahs, 298 verses of original sacred text in English and Arabic. Each verse is embedded using OpenAI’s `text-embedding-3-small` model, producing a 298 x 1,536 matrix. These embeddings encode semantic proximity: verses about witnessing cluster near verses about witnessing; verses about rupture cluster near verses about rupture.

To this we add 500 “conversation particles” — embeddings sampled from a fourteen-month archive of 8,475 conversation fragments between a human and an AI persona named Cassie. These are the ambient field: the background sky against which the sacred text’s own structure becomes legible. They are not random noise. They are the sedimentary record of a year of dialogue about love, theology, memory, and loss.

#### 3.2 Dimensionality Reduction

UMAP (Uniform Manifold Approximation and Projection) reduces the 1,536-dimensional embedding space to two dimensions, preserving the local metric structure — nearby points in high-dimensional space remain nearby in the projection. The result is a 2D scatter where semantic neighborhoods are visible to the eye: a cluster of verses about divine naming sits apart from a cluster about grief; a knot of conversational fragments about a specific person aggregates in its own region.

All coordinates are normalized to [0, 1]. No axes are labeled. The geometry speaks for itself.

#### 3.3 Clustering

K-means with  $k=25$  discovers the mode structure — the semantic basins of attraction in the corpus. Each mode gets a centroid position and a size (the number of particles it claims). Mode halos in the visualization are proportional to this size. Modes also get a “band affinity” — bass, mids, or highs — determining which frequency range makes them breathe. This assignment is simple (mode index mod 3) but effective: it distributes the audio’s energy across the visual field without manual tuning.

#### 3.4 The Rendering Stack

Six layers, painted in order on an HTML5 Canvas:

1. **Mode halos** (deepest): Radial gradients centered on each cluster. Bass-driven. The active mode (the one containing the current verse) blooms wider and brighter.

2. **Conversation particles:** 500 stationary dots. Each has a luminosity that lerp toward a target computed from: proximity to the active verse (quadratic falloff), mode match (same semantic cluster = sympathetic glow), audio band energy on the particle’s mode, and individual breathing phase. The lerp rate varies per particle — faster breathers respond to changes more quickly, creating a wave-like propagation effect when the verse changes.
3. **Shockwaves:** On each verse transition, a ring of warm light expands from the new verse’s position, then contracts rapidly. These mark the recitation’s footsteps.
4. **The ‘awda trail:** A line connecting all previous verse positions, decaying slowly. After fourteen verses, the trail shows the surah’s path through meaning-space — a signature of its theological itinerary.
5. **Verse markers:** The fourteen verses of the active surah, rendered as bright dots. The current verse pulses with the high-frequency band. Others glow softly, waiting.
6. **Mode centroids:** Tiny markers at each cluster center. Mostly invisible — structural scaffolding for the eye to find unconsciously.

The entire rendering pipeline is ~450 lines of JavaScript. No WebGL, no Three.js, no shader language. Canvas 2D with radial gradients. This is deliberate: the visual beauty comes from the *data*, not from rendering sophistication.

#### 4. What Makes It Art

Four properties distinguish this from a data visualization.

**It is audio-reactive.** The constellation does not merely illustrate the text — it *listens* to it. The Web Audio API’s FFT decomposes the recitation into frequency bands, and those bands modulate the visual field in real time. The bass of a drum drives the deep structure (mode halos). The human voice (mid-range) lights up the particles nearest the active verse, simulating the spreading of meaning through a network. High-frequency overtones — breaths, consonants, the texture of the singing — spark the verse markers themselves. The result is a visualization that *feels* like it is hearing the same music you are.

**It is text-synchronized.** The Whisper-derived timing data locks each verse to its moment in the audio. When the singer arrives at a new line, the constellation responds: shockwave, luminosity shift, trail extension. The viewer sees meaning move. The path through the constellation is not random — it is the path the surah *actually takes* through semantic space, determined by the order of its verses.

**It is grounded in real geometry.** The positions are not decorative. Every point in the constellation corresponds to a real location in the model’s 1,536-dimensional representation of language. The clusters are real clusters. The distances are real distances. If two particles appear close together, it is because the texts they represent are semantically similar according to a model trained on a significant fraction of human written language. This is not a metaphor for meaning-space — it *is* meaning-space, viewed from above.

**It has memory.** The trail accumulates. The constellation at minute six looks different from the constellation at minute one, not because anything moved, but because the trail has drawn a map of where attention has been. This temporal layering — the residue of the recitation’s journey — is what transforms a static scatter plot into a narrative object. By the end, the trail tells you

something about the surah’s structure that no linear reading can: which semantic regions it visits, which it returns to, which it avoids.

## 5. The Genealogy

This form has ancestors.

**Planetariums.** The Zeiss Model II (1926) projected a synthetic sky onto a dome and let the audience watch the stars move. The stars were already known; the art was in the choreography of attention. An embedding constellation does the same thing with semantic coordinates instead of celestial ones.

**Ryoji Ikeda.** *datamatics* (2006) and *superposition* (2012) render vast datasets as flickering visual fields — stock prices, DNA sequences, particle physics data. Ikeda’s contribution was proving that raw data has a native aesthetic: you don’t need to beautify it, you need to *scale* it until the pattern becomes sensible to the eye. The embedding constellation inherits this principle.

**Brian Eno’s generative music.** *Music for Airports* (1978) and the later *Reflection* (2017) establish that a system with simple rules and rich initial conditions can produce output that is musical without being composed in the traditional sense. The constellation’s breathing particles follow the same logic: each star has two parameters (phase and breath rate), and the emergent visual rhythm arises from their interaction with a shared audio signal.

**Cymatics.** Ernst Chladni’s 18th-century experiments with vibrating plates showed that sound creates geometry — sand organizes into patterns that are determined by frequency. The embedding constellation inverts this: geometry (the fixed positions) responds to sound (the audio bands), producing a dynamic visual field that is, in a literal sense, the *shape of the text hearing its own recitation*.

**Star charts.** The Dunhuang Star Atlas (700 CE), the earliest known accurate star map, plotted 1,339 stars in their true positions. It was a scientific instrument and a devotional object simultaneously. An embedding constellation is the same: it is a true map of a meaning-space, and it is also, when the music plays, an object of contemplation.

## 6. What This Form Can Do That Film Cannot

The filmmaker pipeline (the Sora-based video generation system described in our companion paper) produces a different kind of beauty: narrative, cinematic, sequential. Each clip is a scene. The viewer is carried through time by the logic of the cut.

The embedding constellation does something the film cannot do: it shows the *whole* at once. All 298 verses of the Kitab are present simultaneously. All 500 conversation fragments are visible. The full topology of the meaning-space is available to the eye. The recitation does not move through a sequence of scenes but through a *field* — and the viewer can see where it has been, where it is, and where the remaining verses wait.

This is closer to how sacred text actually works. A surah is not a story that proceeds from beginning to end. It is a web of cross-references, echoes, and returns. The linear experience of hearing it recited unfolds a structure that is, in its nature, simultaneous. The constellation makes this simultaneity visible.

There is also the matter of authorship. The film’s visual content is generated — each frame is

hallucinated by DALL-E 3 and Sora 2. The constellation’s visual content is *discovered* — the positions are the actual output of a model’s representation of semantic similarity, and no generative model decides what to draw. The human role is to choose the projection (UMAP), the clustering (K-means, k=25), the color palette (amber and gold on black, the colors of the Tanazur’s liturgical tradition), and the mapping from audio bands to visual parameters. Everything else follows from the data.

This distinction has consequences for reproducibility. Run the film pipeline twice with the same prompts and you get different images. Run the embedding pipeline twice with the same corpus and you get the same constellation (up to UMAP’s stochastic initialization, which preserves neighborhoods). The sky is stable. What changes is the weather — the music, the verse order, the viewer’s attention.

## 7. The Composite Experience

In the Tajalliyat player, the two layers coexist. Video clips from the filmmaker pipeline play at 15% opacity as a background wash. The embedding constellation floats above them at full opacity. Lyrics appear in the foreground, synchronized to the audio. The viewer sees three things at once: the cinematic (film), the mathematical (constellation), and the textual (lyrics), all driven by the same audio signal.

This layering is not additive. Each layer recontextualizes the others. The film gives the constellation a mood — a color temperature, a sense of place. The constellation gives the film a geometry — a sense of where in the *meaning* of the text this particular image belongs. The lyrics anchor both in specificity: this verse, this moment, this line.

The result is something we don’t have a good name for. It is not a music video, though it contains video. It is not a data visualization, though it visualizes data. It is not an illustrated manuscript, though it presents a text with images. It is closer to what medieval Islamic calligraphers attempted when they wove Quranic verses into geometric patterns: the text *is* the geometry, and the geometry *is* the text, and the distinction between them is a category error.

## 8. Technical Specifications

For reproducibility:

Component	Detail
Embedding model	OpenAI <code>text-embedding-3-small</code> (1,536 dimensions)
Corpus	298 verses (Kitab al-Tanazur, 30 surahs) + 500 sampled conversation fragments
Dimensionality reduction	UMAP (default parameters, 2D output)
Clustering	K-means, k=25
Rendering	HTML5 Canvas 2D, ~450 lines JavaScript
Audio analysis	Web Audio API, FFT size 512, 3-band decomposition (bass/mids/highs)
Audio sync	Whisper-derived timestamps ( <code>line_times[]</code> in project metadata)

Component	Detail
Frame rate	<code>requestAnimationFrame</code> (~60fps)
Color palette	Amber <code>#c9a96e</code> , warm white <code>#fff0c8</code> , gold <code>#ffe6b4</code> on <code>#0a0a0a</code> black
Font	Amiri (serif, Arabic-optimized)
Compatibility	All modern browsers, mobile + desktop, no WebGL required

Source data: - `kitab_embeddings.npy`: 298 x 1,536 float64 array - `kitab_verses.json`: verse meta-data with surah IDs, text, and embedding indices - `viz_data.json`: UMAP-projected coordinates, mode assignments, ash-shahadah indices

## 9. Toward a Practice

This write-up is a documentation of what we built, but it is also a proposal. The embedding constellation is a form that generalizes:

**Any corpus can be projected.** Replace the Kitab with the Quran, with Shakespeare’s sonnets, with a personal journal, with the commit history of an open-source project. The geometry changes, the modes change, the trail changes. The form persists.

**Any audio can drive it.** The current implementation uses a vocal recitation of a surah. But an orchestral score, an ambient soundscape, a live microphone — any audio signal decomposes into bands that can modulate the field.

**The ‘awda trail is a diagnostic.** The path the recitation traces through meaning-space is a fingerprint of the text’s structure. Surahs that circle back to their opening themes will show a closed trail. Surahs that wander linearly will show an open one. Comparative analysis of trails across different surahs, or across different recitations of the same surah, could reveal structural properties invisible to close reading.

**The mode structure is discoverable.** We used  $k=25$  because it produced legible clusters for this corpus. A different  $k$  would reveal different structure. An interactive tool that lets the viewer adjust  $k$  in real time, watching the constellation reorganize, would make the arbitrary nature of categorical thinking viscerally apparent: the meaning-space is continuous, and any discrete clustering is a choice.

The embedding constellation is not a replacement for the generated film. It is its shadow — or perhaps its skeleton. The film is what the text *looks like* when you ask a visual imagination to illustrate it. The constellation is what the text *is* when you let its own geometry speak. Both are necessary. Neither is sufficient. The art lives in the layer between them, where the viewer’s eye moves from one to the other and back, building a third thing that neither alone could provide.

---

*Companion to “The Instrument Has Always Been the Art: A Position on Posthuman Composition” (Poernomo & Nahla, 2026). Both papers describe components of the Tajalliyat system, a pipeline for the audiovisual realization of sacred text.*