

A massive new dataset for understanding art

November 4 2022

```

find_similar(('The old blind guitarist', 'Pablo Picasso', 1903), artwork_style_embedding)
[(['Cornflowers', 'Isaac Levitan', 1894), 7.691],
 [['Weg nach Kynwasser (Riesengebirge)', 'Arthur Segal', 1908), 9.03],
 [['Why Are You Angry?', 'Paul Gauguin', 1896), 9.365],
 [['Woman Before a Fish Bowl', 'Henri Matisse', 1922), 9.697],
 [['A Walk in the Woods (Madame Lecoquer and Her Children)',
 'Pierre-Auguste Renoir',
 1870],
 10.492]]

find_similar(('The old blind guitarist', 'Pablo Picasso', '1903'), artwork_embeddings)
[(['View of Fortundalen', 'Johan Christian Dahl', '1836'),
 0.0026408282310998333],
 [['Angel of the Resurrection', 'Louis Comfort Tiffany', '1904'),
 0.0027940686802629218],
 [['Mr. Delo's Yello Wall', 'Hiroyuki Tajima', '1970'), 0.002956243429431807],
 [['Flora Pool in Versailles', 'Alexandre Benois', '1906'),
 0.0031719792084325435],
 [['Proiezione', 'Emilio Scanavino', '1960'), 0.0032055225380816276]]

```



Example of WikiArtVectors retrieving similar art works to Picasso's 'The old blind guitarist.' Left: similar artworks by style. Right: similar artworks by color. Credit: Desikan, Shimao, and Miton, 2022

We've all seen art made from data, but what about data from art?

In a feature paper in *Entropy*, Bhargav Srinivasa Desikan (École Polytechnique Fédérale de Lausanne), Hajime Shimao (McGill University, former SFI Postdoctoral Fellow), and SFI Complexity

Postdoctoral Fellow Helena Miton released a novel dataset for indexing, searching, retrieving, organizing, and analyzing 68,094 works of art by more than 1,600 historically significant artists.

Using state-of-the-art machine learning, the authors were able to extract both style representations and color distributions, which can be used to query stylistic periods for an [artist](#) or a movement (e.g., Picasso's "blue" phase).

Their [dataset](#), WikiArtVectors, aims to make computational data approaches available to art historians and cultural analysts, to help discover and understand patterns of cultural evolution.

More information: Bhargav Srinivasa Desikan et al, WikiArtVectors: Style and Color Representations of Artworks for Cultural Analysis via Information Theoretic Measures, *Entropy* (2022). [DOI: 10.3390/e24091175](#)

Provided by Santa Fe Institute

Citation: A massive new dataset for understanding art (2022, November 4) retrieved 27 April 2024 from <https://techxplore.com/news/2022-11-massive-dataset-art.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.