

PHOTODANCE

How do you make thirty years of photos understandable?

THE MEANING MACHINE

how a photo becomes meaning



swipe to see how →

THE PROBLEM

A Photo Is Just a Grid of Numbers

THE MEANING MACHINE



To a computer, a photo is only three grids of numbers — red, green, blue. Who's in it, where, whether it mattered: none of that is in the pixels. Recovering it takes a dozen small steps.

FOR FREE

Laid Out by Time and Color



An EXIF tag gives the capture time and date, and sometimes the location. Using just that, plus the dominant color from each photo's own pixels, thirty years lay out as a timeline you can already **select, filter, and browse**. No model has run yet.

A Dozen Stages, In Order



From here it's a sequence. Each stage leaves behind something the next one leans on — duplicates before quality, quality before faces, faces before captions.

A Very Big Shoebox



227,000+ photos pour in from Amazon, Google, Apple, and Lightroom — already tagged with faces, albums, and ratings. PhotoDance keeps all of it as a free first guess.

Many of Them Are Copies

ENRICHMENT · THIN THE PILE

• Dedup • • • • • • • • • •

176,786

~ 27,092 re-saves — the same pixels in a new file (a JPG of a HEIC). Crops and edits we keep.

exact duplicates

backups of backups, re-saves



23,746 exact copies → keep one

the same photo, re-saved

a JPG of a HEIC — same pixels, new file

**same fingerprint → merge**

different fingerprints → keep both

27,092 re-saves merged · crops & edits kept

bursts

ten frames of one moment



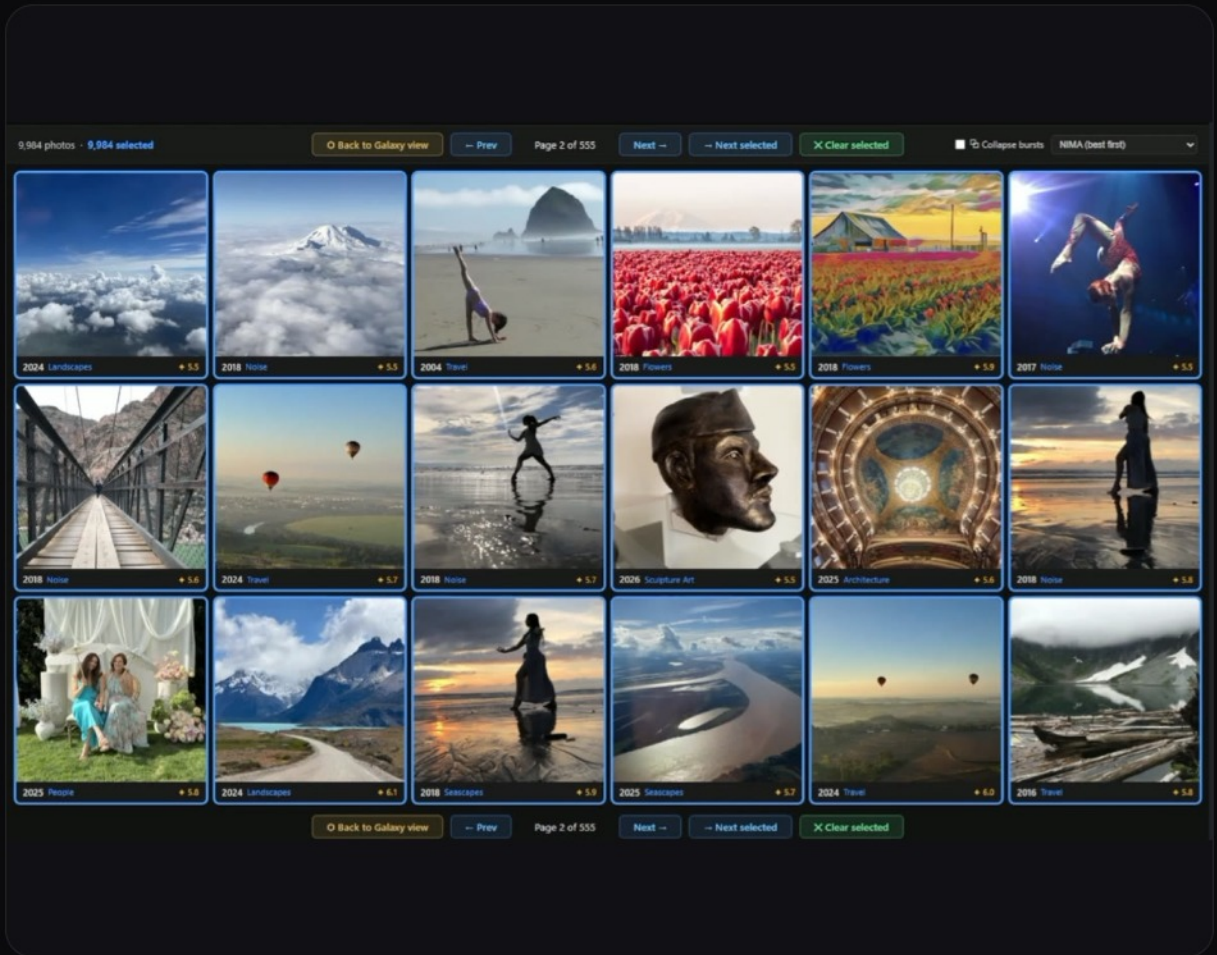
Exact duplicates, HEIC→JPEG re-saves caught by a perceptual fingerprint, and bursts folded to their sharpest frame. A quarter-million files become **~160,000**.

Scoring Every Photo



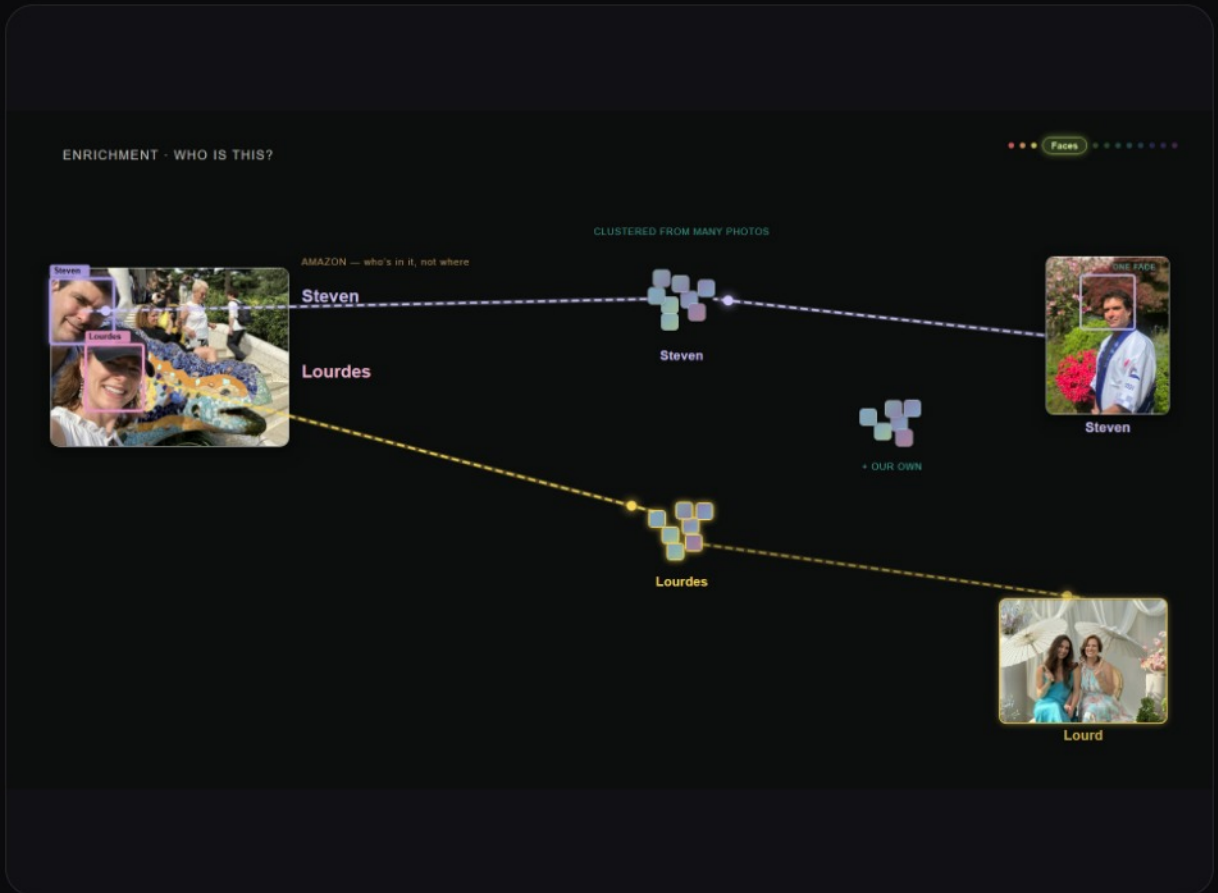
A stack of models scores every photo — MUSIQ, NIMA and MANIQA for technical quality, a LAION predictor and ArtiMuse for aesthetics, plus sharpness, bokeh and more. They feed one **ever-evolving model of our own taste**, whose only target is the ratings we've given over the years.

Surfacing the Best



Sorted by that taste, the collection's best rise to the top — the striking, the sharp, the well-composed — out of a hundred and sixty thousand. Nothing is ever thrown away; the score only decides what to **surface first**.

Who's in the Picture

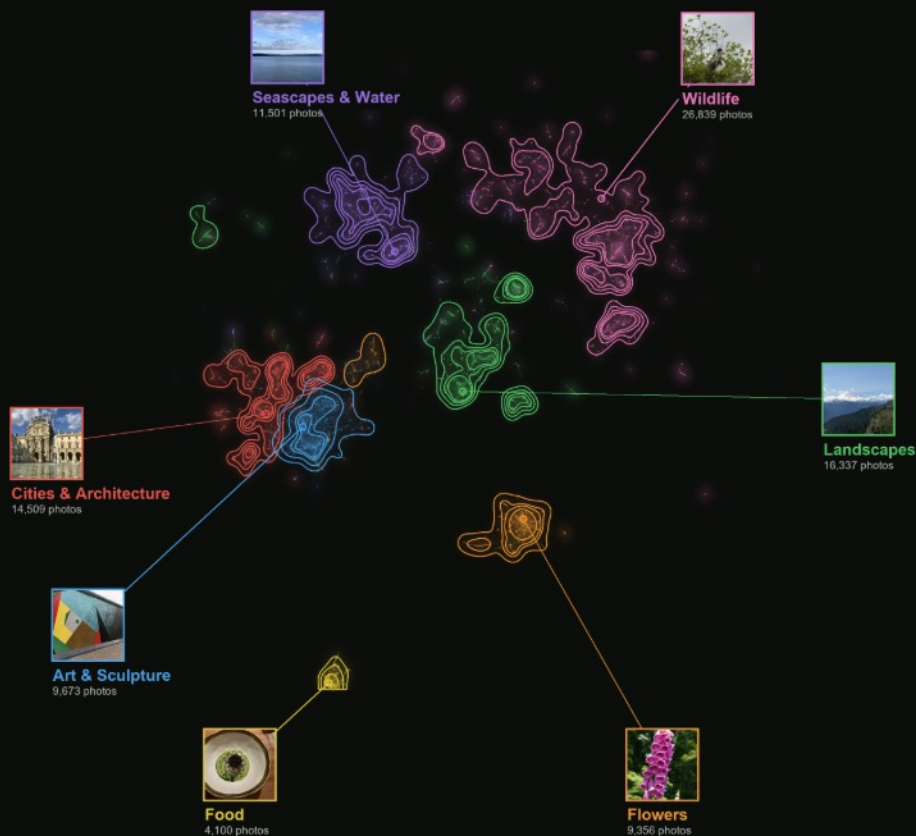


RetinaFace finds every face; ArcFace turns each into a 512-number signature that groups into people. The names we'd already confirmed in Amazon Photos seed the guesses — and from each certain name, the identity spreads outward to the rest of the collection.

A Space of Meaning

Curated collection - content regions

92,315 photos - ViT-H-14 - UMAP (label-supervised) - soft density field - nested contours at 35/55/75/90%



OpenCLIP (ViT-H-14) places every photo in a 1,024-dimension space where nearby means similar. Seascapes, cities, and wildlife sort themselves — and words land there too, so **you can search it**.

A Caption That's True

ENRICHMENT · THE ENRICHED CAPTION

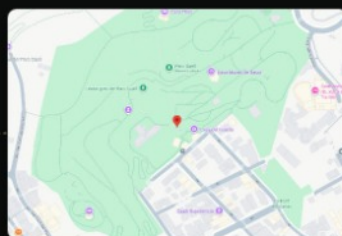


Lourdes

Steven

EXIF · WRITTEN BY THE CAMERA

WHEN 2018-07-12 16:41
CAMERA iPhone X
GPS 41.4137, 2.15285

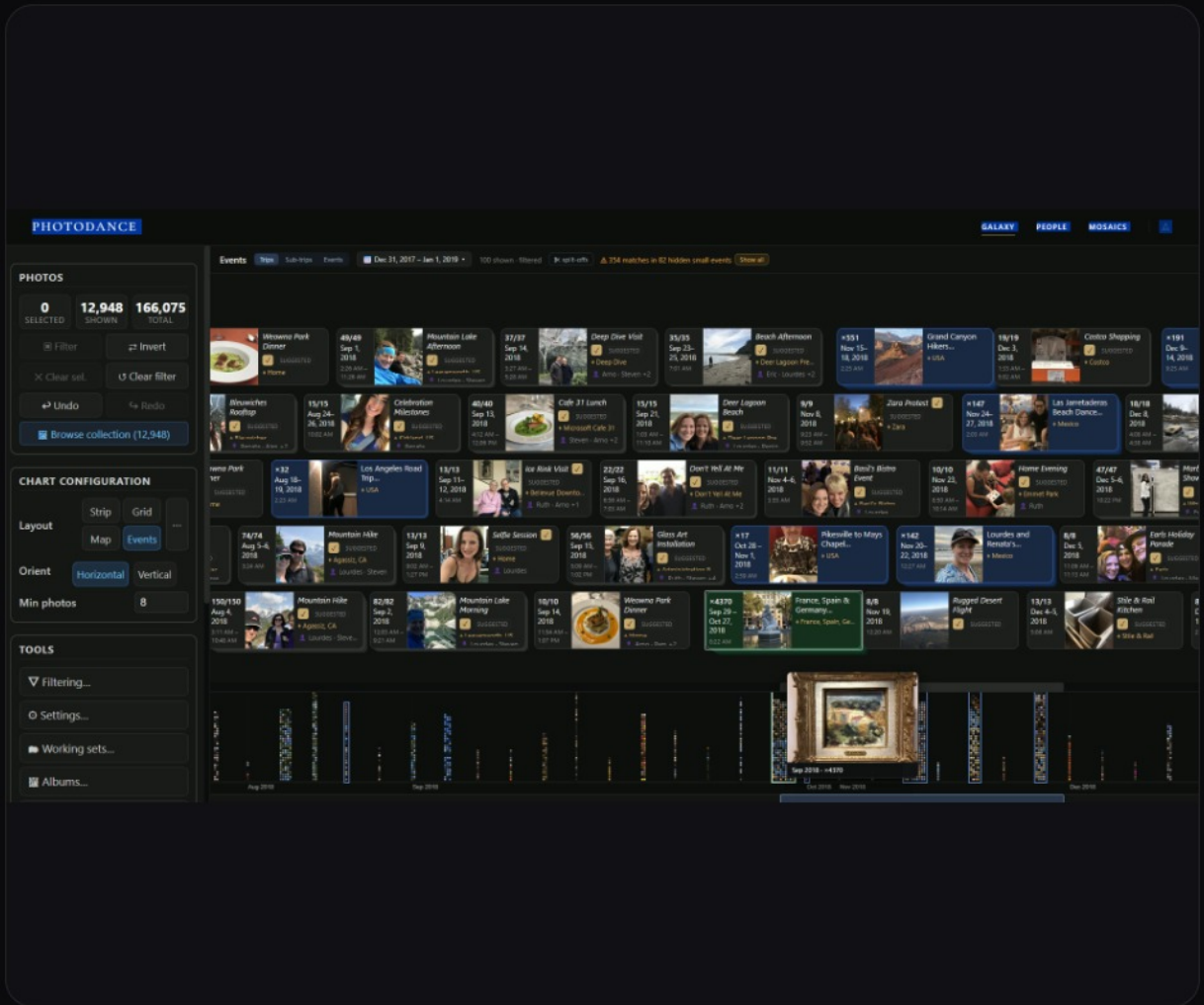


Park Güell · Barcelona

Lourdes and Steven are smiling as they pose in front of a colorful mosaic sculpture at Park Güell in Barcelona, with other visitors in the background.

A vision-language model writes a first description from the pixels alone. A second pass grounds it — looking the location up to a real place name and folding in the faces already resolved — so the caption finally lands on **who and where**.

Folded Into Events



Gaps between timestamps fold shots into named events, outings, and trips — “Snowy Forest Fun,” a birthday, a road trip. A journal, not a folder.

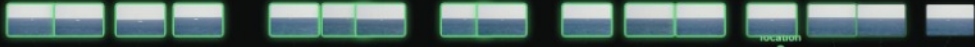
Borrowed Locations

ENRICHMENT · IT FLOWS BETWEEN PHOTOS

● ● ● ● ● ● ● ● ● ● Propagate ● ● ●

Maui — one afternoon, whales from the shore

Nikon
no GPS



3:00 PM

3:30 PM

4:00 PM

4:30 PM

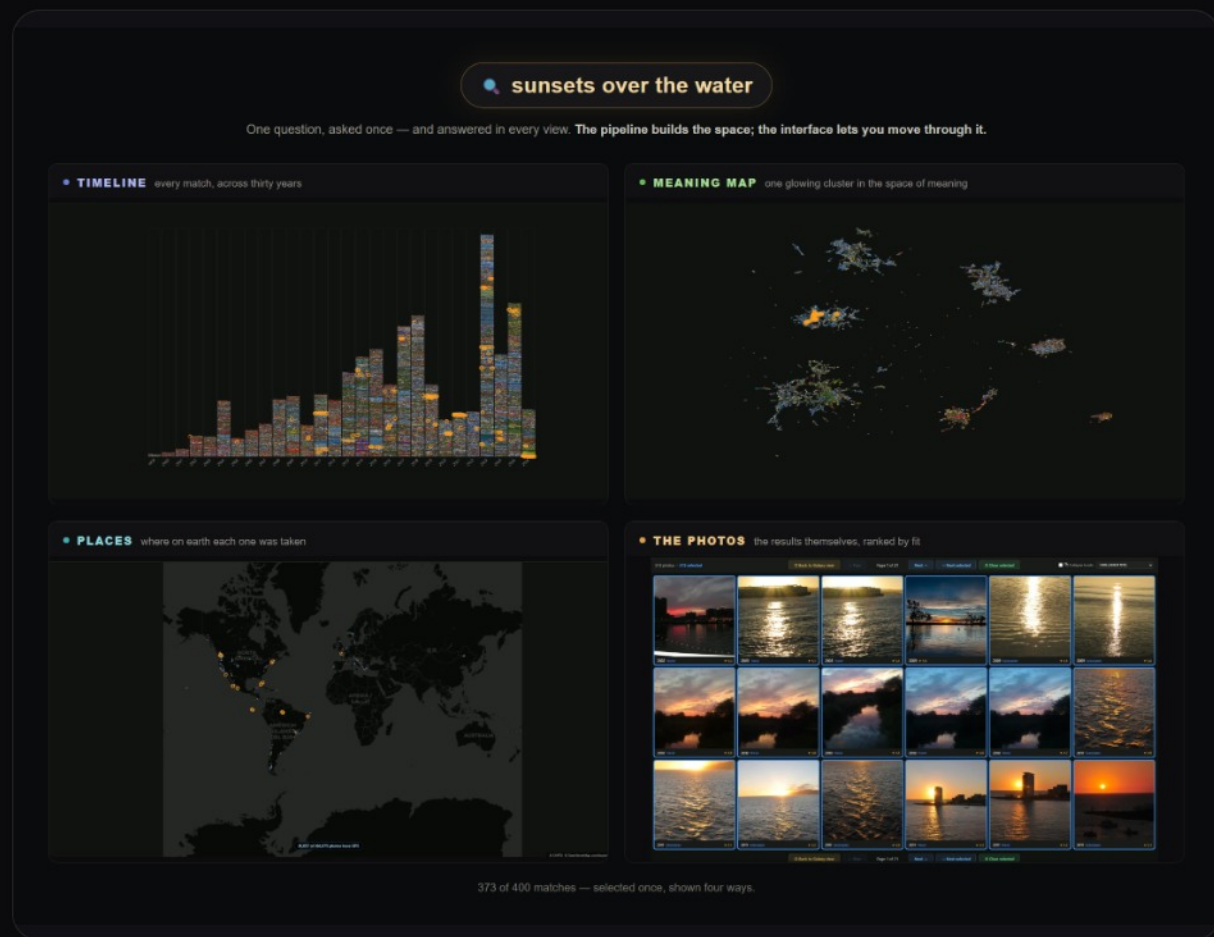
iPhone
has GPS



690 of 786 whale shots placed — from just a handful of iPhone anchors

Our early DSLRs — Nikons and Canons — recorded no GPS; a phone shot minutes away did. Those located photos lend their coordinates to neighbors in time, putting **8,905 photos** on the map that had none.

Ask Once, See It Everywhere



Ask “sunsets over the water” and the answer appears in every view at once — timeline, meaning-map, world-map, and photos. The same 373, **selected once, seen four ways.**

PHOTODANCE

From overwhelming to understandable.

No single model does it. A dozen of them, tied together by an interface built to use them all at once, bring meaning back to a collection that was simply too big to face. Now you can search it, browse it, and move through it just by asking.

► **Watch the full narrated film**

youtu.be/VRE8-3zHbio

interactive write-up · steven-drucker.com/##/Featured/meaning-machine