Exemem: Privacy-Preserving Network Effects on Vectorized Data

Explained Simply

Tom Tang

March 4, 2026

1 What's the Problem?

Imagine you take a photo at a concert. Other people at the same concert also take photos—and some of them accidentally capture your face. Those photos now live on strangers' phones and computers. You have no way to know they exist.

Today, the only way to find out is to use a centralized service like Google or Facebook that scans everyone's photos with facial recognition. But that means one company sees *everyone's* photos and knows *everyone's* face. You solve one privacy problem by creating a much bigger one.

Exemem takes a different approach. It lets you discover that strangers have photos of you—without anyone learning who those strangers are, and without any central authority seeing all the photos.

2 The Big Idea: Fingerprints, Not Files

Every file—a photo, a document, a voice recording—can be turned into a **fingerprint**: a short numerical summary that captures what the file is "about" without containing the file itself.

Two photos of the same person's face produce similar fingerprints. Two documents on the same topic produce similar fingerprints. But you can't turn a fingerprint back into the original file. It's a one-way process.

In Exemem, everyone's fingerprints are **public**. Anyone can search them. But the actual files stay **private**, protected by Fold DB's access controls—trust distance, cryptographic keys, and payment gates that determine who can see the real data.

3 Disguises: One Per Entry

Here's where it gets interesting. When you publish a fingerprint, you don't attach your name to it. Instead, each fingerprint gets its own unique **disguise**—a one-time code that can't be traced back to you.

This is like dropping a note in a suggestion box, except you use a different handwriting style for every note. Even if someone collects all the notes, they can't tell which ones came from the same person.

- You publish 50 photo fingerprints. Each one gets a different disguise. To everyone else, they look like 50 unrelated entries from 50 different people.
- You can always prove any entry is yours (you have a master key that generated all the disguises). But nobody else can figure that out.
- Two entries sitting right next to each other in a search result might belong to the same person or to two different people. There's no way to tell.

4 An Example: "Do Strangers Have Photos of Me?"

Alice wants to know if anyone in the network has photos of her face.

Step 1: Search. Alice creates a fingerprint of her own face and searches the network: "show me all fingerprints that look like this face."

Step 2: Results. She gets back 14 matches. Each result shows:

- A disguise code (like `0xA7f3...`)—unique to that entry
- How similar the fingerprint is to hers (98%, 95%, 91%, etc.)
- Basic info (photo, video frame, date)

What Alice knows: 14 images of her face exist in the network.

What Alice doesn't know:

- Who owns any of them. Every entry has a different disguise.
- Whether any two belong to the same person. Maybe 14 people each have one photo. Maybe one person has all 14. She can't tell.
- What the photos actually look like. She only sees the fingerprint match, not the photo itself.

Step 3: Request access. Alice picks one of the 14 entries and says: "can I see the actual photo?" The request goes to the owner's Fold DB access controls. If the owner's rules allow it—maybe Alice is trusted enough, or maybe she pays a small fee—she sees the photo. If not, she sees nothing.

Even after seeing several photos, Alice still can't tell which entries belong to the same person. Each disguise is independent.

Step 4: Request removal. Alice can send a message to any entry's disguise: "the person in this photo asks you to remove it." The owner gets the message. Whether they act on it is up to them, but the request is permanently recorded.

Step 5: Monitor. Alice can re-run her search next week, next month, next year. If new photos of her face appear, she'll see new entries with new disguises.

5 Two Layers

Exemem has two separate layers, each with its own rules:

Layer 1: The Public Catalog. All fingerprints and basic metadata are public. Anyone can search them. No permission needed. This is where discovery happens. Every entry has its own unique disguise that can't be linked to other entries or to real people.

Layer 2: The Private Vault. The actual files (photos, documents, recordings) are protected by Fold DB. To see the real file, you need to pass the owner's rules: trust distance, cryptographic keys, payment. This is where access control happens.

The catalog tells you something exists. The vault controls who gets to see it.

6 Why More People Makes It Better

6.1 Better Answers

If only 100 people use the network, Alice's face search might miss photos that exist on other people's devices. If 100 million people use it, she's much more likely to find every photo of her face that's out there.

More people = more fingerprints = more complete answers.

6.2 Stronger Privacy

This is the surprising part. In most networks (Facebook, Instagram), more users means *less* privacy—more data to mine, more connections to trace. In Exemem, it's the opposite.

When Alice finds an entry with disguise `0xA7f3...`, the owner could be any of the network's users. With 1,000 users, that's 1,000 possibilities. With 1 million users, that's 1 million possibilities. And because every entry has a different disguise, you can't narrow it down by grouping entries together.

More people = bigger crowd to hide in = harder to figure out who owns what.

7 Other Examples

7.1 Music: "Did Someone Copy My Song?"

A musician publishes fingerprints of her tracks. She searches the network and finds an entry with a 96% similarity match, uploaded 6 months before her release. She can't see who uploaded it. She can request access to hear the audio, or file a timestamped claim referencing the fingerprint.

7.2 Medical: "Who Else Has My Condition?"

A patient publishes fingerprints of their symptom profile. A researcher searches for rare disease signatures and finds 47 matching entries. She can't tell how many patients they represent. She can reach out through each entry's disguise to propose clinical trial participation—without knowing who anyone is unless they choose to reveal themselves.

7.3 Security: "Has Anyone Seen This Attack?"

A security team publishes fingerprints of a cyberattack pattern. Other teams search and find 7 matching entries. They can't tell which or how many organizations were hit, but they know the threat is widespread and can share defenses through the disguises.

7.4 Research: "Is Anyone Working on the Same Thing?"

A scientist publishes fingerprints of her research findings. She discovers 23 entries with highly similar research fingerprints. She can reach out to each disguise to explore collaboration—without knowing whether those 23 entries represent 23 research groups or 3.

8 What About Someone Trying to Cheat?

8.1 Can Someone Figure Out Who Owns an Entry?

Each entry has its own unique disguise. There's no username, no email, no account to trace. The system stores no IP addresses or session data. The most an attacker can do is guess based on the content of the fingerprint itself (e.g., "these photos were all taken in the same city"). But that only gives probabilistic guesses—never certainty—and the larger the network, the more people could plausibly be the owner.

8.2 Can Someone Reconstruct a File from Its Fingerprint?

Fingerprints are one-way: you can't turn them back into the original file. Research has shown that very rough approximations are sometimes possible (e.g., generating a blurry face from a face fingerprint), but the actual photo—the specific image, with all its details—cannot be recovered.

8.3 Can Someone Link Two Entries Together?

Not by their disguises—every entry has a unique one. An attacker might notice that two fingerprints are similar ("both are photos taken at the same location"), but they can't confirm they belong to the same person. Similar content doesn't prove shared ownership.

9 Summary

- Every file gets a **fingerprint**—a numerical summary that captures what it's about but can't be reversed into the original.
- All fingerprints are **public**. Anyone can search them.
- Each fingerprint gets its own unique **disguise**. No two entries can be linked to the same person.
- The owner can always prove an entry is theirs, but nobody else can figure it out.
- The actual files stay **private**, protected by Fold DB's access controls.
- More users means **better answers** (more complete discovery) and **stronger privacy** (bigger crowd to hide in).
- You can discover that strangers have photos of you, request access, request removal, and monitor for new appearances—all without anyone learning who anyone else is.