**CIAF Extension to ANN, GNN, CNN, and Vectorized Models**

**1. Artificial Neural Networks (ANNs)**

**Key Insight:**

* In ANNs, each **node activation** can be tied to a **Dataset-Derived Key (DDK)**.
* Instead of exposing the weights (which are proprietary), CIAF signs **node activations**.

**Process:**

1. Dataset Anchor → generates DDK.
2. DDK → generates **NodeKeys** per node.
3. Node activations → encapsulated as **Node Provenance Capsules**.
4. Capsules → aggregated into a **Layer Merkle Root**.
5. Merkle Root → secured in the Training Snapshot.

**Benefit:**

* Regulators can verify which parts of the network fired, **without ever seeing the actual weights**.
* Mathematically sound within a **tolerance margin** (i.e., small numeric differences don’t invalidate verification).

**2. Graph Neural Networks (GNNs)**

**Key Insight:**

* GNNs are structured around **nodes and edges**.
* You proposed **not storing explicit edges**, since they can be **inferred** mathematically from the node activations.

**Process:**

1. Dataset Anchor + DDK signs each **graph node**.
2. Each node activation = provenance capsule.
3. **Edges** are reconstructed by inference:
   * Two activated nodes with shared provenance/time correlation → inferred edge.
4. Merkle Root → covers both node activations and inferred edge structure.

**Benefit:**

* Keeps the **graph structure private**.
* Still provides **verifiable integrity** of the node-level activity.

**Analogy:**

Like plotting dots on a paper (nodes). Two dots make a line (edge).  
You only need to prove the dots — the line can be inferred later.

**3. Convolutional Neural Networks (CNNs)**

**Key Insight:**

* CNNs work with **filters and feature maps**.
* Instead of tracing every convolution weight, CIAF can anchor **activation maps**.

**Process:**

1. Dataset Anchor → DDK → NodeKeys for filters.
2. Each **filter activation map** = provenance capsule.
3. Capsules → combined into a **Layer Merkle Root**.
4. Receipts include metadata: filter ID, activation strength, SHAP importance.

**Benefit:**

* Regulators/auditors can verify that the image/video was processed through a **real model snapshot**, not a spoofed CNN pipeline.

**4. Vectorized Datasets**

**Key Insight:**

* When data is vectorized (e.g., embeddings, token matrices), provenance must capture the **embedding state**.

**Process:**

1. Each embedding vector → hashed and signed as a capsule.
2. Capsule = vector hash + metadata (source, tokenization, dimension).
3. During inference, the receipt includes:
   * Input vector hash
   * Capsule linkage
   * SHAP / attribution vector
   * MAK signature

**Benefit:**

* Even if the raw text/image/audio is transformed into vectors, you can **trace back the exact embedding provenance**.

**5. Metadata on Inference for Traceback**

This was one of your most powerful points.

**CIAF Metadata for Inference Receipts includes:**

* **Model Info:** Name, version, vendor, compliance mode.
* **Dataset Anchor:** Links to the dataset root used in training.
* **Snapshot ID:** Cryptographic hash of the Training Snapshot.
* **Query Hash:** Fingerprint of the input (so the input cannot be swapped).
* **Prediction Result:** Output label + probability.
* **Explainability Data:** SHAP values or feature attributions.
* **Capsule Proofs:** Subset of the Merkle proofs needed to verify the query path.
* **Digital Signature:** MAK signature to confirm authenticity.

**Traceback Flow:**

1. Take an inference receipt.
2. Verify its MAK signature → ensures it belongs to the authorized model.
3. Follow the Snapshot ID → find the training snapshot.
4. Check the Snapshot → ties back to Dataset Anchor + capsules.
5. Verify the dataset anchor → check immutable log for anchored integrity.

**Result:**

* A verifiable chain: **Inference → Snapshot → Capsules → Dataset Anchor**.
* If any part is tampered with, the cryptographic verification fails.

**6. Why This Matters**

* **For ANNs:** Private weight verification.
* **For GNNs:** Edge inference reduces exposure while proving structure integrity.
* **For CNNs:** Provenance for visual pipelines, critical for deepfake defense.
* **For Vectorized Data:** Ensures embeddings (often overlooked) have lineage.
* **For Inference Traceback:** Anyone (auditor, regulator, platform) can verify authenticity and compliance without access to raw training data or model weights.

⚡ Put simply:

The CIAF system makes ANNs, GNNs, CNNs, and vector models **auditable, explainable, and tamper-proof** — all while preserving the privacy of proprietary architectures and training data.