◢ Suggested Shard Schema

```json
{
  "id": "shard_4521",
  "type": "semantic",
  "name": "career_direction_conflict",
  "generated_by": "embedding_cluster",
  "created_at": "2025-05-03T00:00:00Z",
  "confidence": 0.94,
  "linked_memory_ids": ["memory_001", "memory_003", "memory_019"],
  "summary": "Cluster of entries related to questioning AI career direction",
  "centroid_embedding": [0.123, 0.456, ...]  // optional for fast similarity
}
```

<u>Centroid Embedding</u>: The average vector (the Center)
of all the embeddings in a shard
It represents the center of gravity for the shard

```less
[Memory A] → embedding: [0.1, 0.5, ...]
[Memory B] → embedding: [0.2, 0.6, ...]
[Memory C] → embedding: [0.12, 0.55, ...]

→ All belong to Shard X

[Shard X]
- type: semantic
- name: burnout_loop
- memory_ids: [A, B, C]
- centroid_embedding: average([A, B, C])
```

Now at query time, instead of scoring query vs A, B, and C:

→ Score query vs `centroid_embedding` of Shard X

If it's relevant → *then* look inside the shard at individual memories.