# Day95 ChromaDB CRUD and Metadata

October 2, 2025

### 1 Vector Databases CRUD with ChromaDB

In this notebook, we will learn how to perform CRUD operations in ChromaDB:

- $C \rightarrow Create (Add documents)$
- $R \rightarrow Read$  (Query documents)
- U → Update (Modify documents/metadata)
- $D \rightarrow Delete$  (Remove documents)

We will do this with and without metadata.

Metadata helps us add extra info (author, year, category) so we can filter results easily.

## 2 Setup & Imports

We first import chromadb and create a client to connect to ChromaDB.

We also set up an **embedding function** (SentenceTransformer model) that converts text  $\rightarrow$  vectors.

### 3 Create a Collection

We create a new collection named "news\_crud".

Think of a collection like a table in SQL – it stores documents, embeddings, and metadata.

```
[2]: # Delete old collection if it exists

try:
          client.delete_collection("news_crud")
except:
          pass

# Create new collection
collection = client.create_collection(
          name="news_crud",
          embedding_function=ef
)

print(" Collection created")
```

Collection created

## 4 Add Documents (Create)

We insert documents into the collection. Each document must have a **unique ID**.

```
[3]: collection.add(
    documents=[
        "Elon Musk founded SpaceX in 2002.",
        "Apple just released iPhone 16 Pro.",
        "Virat Kohli scored a century yesterday."
        ],
        ids=["doc1", "doc2", "doc3"]
)

print(" Added 3 documents")
```

Added 3 documents

# 5 Query Documents (Read)

We query the database.

The query is converted into an embedding  $\rightarrow$  compared with stored vectors  $\rightarrow$  top matches returned.

```
[4]: results = collection.query(
          query_texts=["Who founded SpaceX?"],
          n_results=2
)

print(" Query results:")
print(results)
```

```
Query results: {'ids': [['doc1', 'doc2']], 'embeddings': None, 'documents': [['Elon Musk founded SpaceX in 2002.', 'Apple just released iPhone 16 Pro.']], 'uris': None, 'included': ['metadatas', 'documents', 'distances'], 'data': None, 'metadatas': [[None, None]], 'distances': [[0.23058748245239258, 0.9184808731079102]]}
```

## 6 Update Documents

Updating = re-adding with the **same ID**. This overwrites the old document.

```
[5]: collection.add(
    documents=["Apple unveiled iPhone 17 Ultra with AI features."],
    ids=["doc2"] # same ID → overwrites
)
print(" Updated doc2")
```

Updated doc2

### 7 Delete Documents

We can delete documents using their IDs.

```
[6]: collection.delete(ids=["doc3"])
print(" Deleted doc3 (Virat Kohli news)")
```

Deleted doc3 (Virat Kohli news)

#### 8 Peek Collection

We check the current state of the collection.

```
7.50400275e-02, -3.88929248e-02, -1.51681295e-03,
 1.55410804e-02, 1.68378535e-03, -4.15757895e-02,
 1.99168641e-02, -2.87700929e-02, -2.23909467e-02,
-4.95715775e-02, 5.07577732e-02, -7.44896829e-02,
 1.79165751e-02, 6.92087635e-02, 2.33935454e-04,
-3.85729223e-02, 3.35584171e-02, -4.27402072e-02,
 3.35543938e-02, -1.17368344e-02, -1.08973064e-01,
 4.62819748e-02, 5.83663583e-02, -9.13591385e-02,
 4.48593684e-02, -7.01928586e-02, 2.47093998e-02,
8.00517425e-02, -6.65008742e-03, 3.59858433e-03,
-2.75719408e-02, -3.75631191e-02, -2.71984115e-02,
-3.01757343e-02, 1.93904955e-02, 8.09062645e-03,
-2.94163227e-02, -1.25438552e-02, -4.69481647e-02,
 2.56717429e-02, 1.02677301e-01, 5.43194003e-02,
 5.48535474e-02, 8.66503455e-03, 7.37245567e-03,
-9.22469608e-03, -9.27797034e-02, 3.55629362e-02,
-4.73680208e-03, 2.27553975e-02, 2.20977403e-02,
4.94795432e-03, -4.73573804e-02, -3.18449810e-02,
3.94190382e-03, 7.34087229e-02, 4.51874034e-03,
-2.70922203e-02, -1.50354495e-02, -2.13407017e-02,
-7.40479082e-02, 5.04463390e-02, -9.65139922e-03,
3.35048102e-02, -1.02009512e-02, 9.79748927e-03,
 5.67497909e-02, 3.05281132e-02, 1.86557397e-02,
-3.02823205e-02, -1.28541850e-02, -6.80522481e-03,
 3.78404255e-03, -5.22648580e-02, -1.04009375e-01,
-7.58859934e-03, -5.95393404e-02, 3.86113599e-02,
-1.23298885e-02, 3.81133854e-02, -5.78163639e-02,
3.50763164e-02, 7.90607482e-02, -7.09173530e-02,
 5.65729253e-02, -4.36585620e-02, -6.20657727e-02,
-1.10835321e-02, -2.00411249e-02, -2.18510106e-02,
-6.03525005e-02, -3.21414938e-34, -4.81008254e-02,
-1.47964442e-02, 2.83824746e-02, 6.26882240e-02,
-4.22883853e-02, -1.55396410e-03, 1.23385768e-02,
 6.84267795e-03, -1.35190502e-01, -3.43633071e-02,
-1.26960464e-02, 3.88799831e-02, 4.36825566e-02,
7.28901103e-03, 7.68791139e-02, -3.78630720e-02,
1.68981086e-02, -6.88661411e-02, 4.12552170e-02,
-7.90467393e-03, 2.19666474e-02, -8.64581093e-02,
 5.14613166e-02, -4.25460562e-02, 2.06542108e-02,
-1.25869326e-02, 2.12451033e-02, -6.42742068e-02,
2.26212367e-02, 9.87633131e-03, -6.87154233e-02,
 6.17425842e-03, -9.19150189e-04, 3.08582652e-02,
-3.26274894e-02, 9.17291194e-02, -1.41495029e-02,
-3.87055054e-03, 1.10848313e-02, -5.22843301e-02,
8.43378250e-03, 2.17624232e-02, -2.28052847e-02,
-8.78386572e-02, 2.70195063e-02, 6.89066723e-02,
-6.06571101e-02, -7.47965649e-02, 1.82350293e-01,
```

```
-1.57465860e-02, 2.45522168e-02, 2.46237814e-02,
1.48193045e-02, -4.74377275e-02, 5.05112484e-02,
-5.66205867e-02, -5.16199358e-02, 4.86528836e-02,
-1.90962777e-02, 5.10772020e-02, -5.87589145e-02,
2.57837102e-02, 9.29058064e-03, 1.31600991e-01,
-4.25595455e-02, 3.04558426e-02, 1.42830908e-02,
-7.89400470e-03, -4.90817092e-02, 7.93539807e-02,
 3.78813855e-02, -9.49774403e-03, -6.32541031e-02,
-5.30567877e-02, -3.74306515e-02, -6.63133562e-02,
-7.26773292e-02, -4.88017052e-02, -6.98020756e-02,
1.18351236e-01, 5.73537797e-02, 1.93497557e-02,
1.00355871e-01, 1.27617382e-02, -4.68207598e-02,
-8.89546610e-03, -1.26389805e-02, 1.08405203e-01,
 3.19386870e-02, 6.09389879e-02, -4.01449241e-02,
-5.37318289e-02, -1.62809789e-02, 1.36028631e-02,
-1.04922481e-01, -1.19312103e-34, 6.63314341e-03,
-6.16270909e-03, 8.06681514e-02, 3.79861221e-02,
7.21341581e-04, 7.36981034e-02, 6.55902177e-02,
-1.71400439e-02, 1.38352374e-02, 3.44624519e-02,
 6.71225414e-02, 5.27862050e-02, 2.61463225e-02,
-2.79419888e-02, 4.88017164e-02, 5.40834256e-02,
1.89413186e-02, -7.28053786e-03, -5.95930368e-02,
-1.06895953e-01, 7.62247965e-02, 5.12820715e-03,
-3.95314917e-02, 9.01894271e-03, -3.83367576e-02,
5.23726596e-03, 9.75697041e-02, 3.14175114e-02,
-5.93121685e-02, -4.09577526e-02, -1.14457227e-01,
4.72517265e-03, -2.80826651e-02, 9.29773748e-02,
 6.22360148e-02, 3.37061062e-02, -7.33620673e-02,
 5.39366826e-02, -3.69075942e-03, -4.93501313e-02,
-3.37257683e-02, 4.04293500e-02, -3.54434624e-02,
 8.68120417e-02, 7.29232933e-03, -4.31827530e-02,
 4.50918078e-02, 1.84965953e-02, 1.87082856e-03,
 1.09035103e-02, -8.08037296e-02, -4.27628607e-02,
 4.83494401e-02, -8.47228169e-02, 1.33162588e-02,
 5.87711371e-02, -9.14607793e-02, 8.32657441e-02,
 6.45337924e-02, -2.92467289e-02, 3.70134562e-02,
5.69142913e-03, 2.57742368e-02, 8.37730151e-03,
-5.74605875e-02, 4.02464904e-03, 1.60865614e-03,
 3.27997655e-02, -1.19893953e-01, 5.92170982e-03,
1.34708315e-01, 4.06239592e-02, -1.54181257e-01,
-1.97475520e-03, -6.39682040e-02, 5.92381805e-02,
-5.49767800e-02, 1.97432972e-02, -5.14510423e-02,
-5.97673235e-03, 7.84410983e-02, 3.12912799e-02,
 2.34079845e-02, 8.75096470e-02, 9.51785129e-03,
8.17615911e-02, 9.22470167e-03, -4.85752635e-02,
 2.47754529e-02, -1.41891399e-02, 1.34668965e-02,
-3.77505012e-02, -4.42278460e-02, 3.75909507e-02,
```

```
-3.67990471e-02, -1.65793921e-08, -7.21116588e-02,
-2.45400574e-02, -4.05419953e-02, 6.04146579e-03,
 6.35924786e-02, 4.00829688e-02, 9.27718654e-02,
 9.91322380e-03, 2.67034513e-03, 3.25934240e-03,
-3.07063144e-02, 2.24776042e-04, -2.74589602e-02,
 2.67863944e-02, -1.19700832e-02, -2.17179079e-02,
 2.50600204e-02, 6.36171401e-02, 6.19343109e-03,
 5.28733842e-02, -1.78733133e-02, 5.38698398e-02,
-2.62769079e-03, 5.42795099e-02, 8.70027859e-03,
-1.69487030e-03, -3.47254537e-02, 5.54371765e-03,
 8.91656950e-02, -2.72664847e-03, -9.40545928e-03,
-4.10718024e-02, -7.18520731e-02, -3.34933437e-02,
-9.85983163e-02, -1.35398551e-03, -6.77860528e-02,
 7.52707645e-02, -9.28485300e-03, -3.47929150e-02,
 8.13034084e-03, 1.05949737e-01, 1.10669620e-01,
-2.80007683e-02, -1.20649762e-01, 5.74429147e-02,
-1.07077584e-01, -1.25881150e-01, -1.53067419e-02,
-1.91945527e-02, 1.80230197e-02, 5.22620305e-02,
-5.30586094e-02, -1.50395390e-02, -4.75955531e-02,
-6.28250614e-02, -7.82314166e-02, 3.27615836e-03,
-1.77231096e-02, -7.10740909e-02, -2.86174007e-02,
-1.28306812e-02, -5.66757564e-03, 4.47214395e-02],
[ 1.23450570e-02, 6.78339601e-02, 5.57274222e-02,
-6.84980899e-02, 6.22243024e-02, -4.60883565e-02,
-7.96756335e-03, 2.07603946e-02, -1.05419466e-02,
 9.73602608e-02, 1.39922574e-01, 6.00863472e-02.
 2.57938937e-03, 1.08845770e-01, 5.41697927e-02,
 3.10288947e-02, 7.55666867e-02, -5.18820621e-02,
-3.52383219e-02, -1.07099386e-02, -5.32791503e-02,
-6.57008663e-02, -6.31552190e-03, 2.00158246e-02,
 6.32711574e-02, 4.98204259e-03, -8.25411305e-02,
-5.71174175e-02, -2.58350149e-02, 1.41775347e-02,
 3.02940421e-03, 4.15867195e-02, 7.83431754e-02,
 1.87551919e-02, -1.74901797e-03, -1.28214076e-01,
 3.23123368e-03, -1.37197860e-02, 6.16704896e-02,
-5.37485778e-02, -4.33199946e-03, 3.06663476e-02,
-4.15527113e-02, 7.33570382e-02, 4.25991043e-02,
 3.71911749e-02, 3.87898162e-02, 4.65319827e-02,
 4.59120460e-02, -1.97161213e-02, 6.52253702e-02,
 2.78766993e-02, -1.58002768e-02, 2.99429726e-02,
-4.38195281e-02, -1.98514871e-02, 3.86946439e-03,
-2.56795846e-02, 5.97121976e-02, 5.26282936e-02,
-2.56813504e-02, -8.63686353e-02, -6.15844093e-02,
 5.77740036e-02, 1.53752947e-02, -3.11981930e-05,
-4.34254762e-03, -2.97026895e-02, -5.36748320e-02,
-4.33563022e-03, 3.72829437e-02, 2.81155426e-02,
 3.68244536e-02, 4.76190075e-03, 2.54074335e-02,
```

```
3.71475033e-02, 2.90910341e-02, -7.64475146e-04,
 9.09575634e-03, 2.45322939e-02, 1.68711841e-02,
-5.16440049e-02, 3.31596769e-02, -4.79323417e-02,
 8.92870594e-03, 1.27324760e-02, 4.09457870e-02,
-3.13302912e-02, -3.55907269e-02, -6.06361702e-02,
-5.68488613e-02, 3.60390358e-02, -2.19257139e-02,
 1.42018870e-02, -2.16919556e-02, -4.30057049e-02,
 6.57396251e-03, -1.12971917e-01, -4.88190241e-02,
 1.57552473e-02, 3.99821438e-02, 1.28458776e-02,
 1.91281848e-02, -2.97569837e-02, 2.70119868e-02,
-6.64231554e-02, 1.08675575e-02, 4.19291109e-02,
8.65105074e-03, 1.28938824e-01, 8.71828422e-02,
-6.49212627e-03, -3.06797568e-02, -2.21334845e-02,
 5.58814667e-02, 1.99367199e-02, -8.49770829e-02,
1.18030518e-01, 8.54870901e-02, -1.58398300e-02,
-2.42747311e-02, 5.23062563e-03, 1.17219158e-03,
-3.87075841e-02, -8.54969919e-02, -3.28681134e-02,
9.90477763e-03, -2.25653351e-33, 6.58606691e-03,
4.58517410e-02, -1.85659807e-02, 1.67959929e-02,
-1.08339463e-03, -8.90919641e-02, 3.36533897e-02,
 6.29609525e-02, -3.17987353e-02, -8.18077624e-02,
-8.73722956e-02, -1.35211989e-01, -7.75673091e-02,
-4.09358516e-02, 8.37595761e-02, -6.36302903e-02,
-2.83142272e-02, 2.93465480e-02, 4.00296748e-02,
4.20304984e-02, -1.69813223e-02, -2.01956201e-02,
-7.67723075e-04, 7.26209069e-03, -1.01762507e-02,
2.94022541e-02, -1.56248631e-02, 5.07941190e-03,
4.49935123e-02, -1.41537180e-02, -9.42832306e-02,
7.58541673e-02, -4.24509170e-03, -3.48096676e-02,
-3.09047685e-03, 1.90341994e-02, 5.62662631e-02,
-9.18762386e-02, 7.30671510e-02, -2.63980068e-02,
-3.03057432e-02, -5.94921783e-03, -2.91407779e-02,
-6.84791477e-03, 1.21592954e-01, 2.96909492e-02,
-1.09275393e-01, -1.37844468e-02, 4.58296724e-02,
1.50030619e-02, -5.16599976e-02, 5.65072559e-02,
 8.34455248e-03, -9.44065750e-02, -6.15249798e-02,
-1.91789810e-02, -2.22514272e-02, -2.63368133e-02,
 8.50797817e-02, 4.36861366e-02, -1.91312209e-02,
 2.15450451e-02, -9.55054089e-02, -7.40828551e-03,
-7.66536966e-02, 5.45860268e-03, 1.91688612e-02,
-3.63680720e-02, -6.10825494e-02, 6.70821145e-02,
-3.58861536e-02, 4.01191376e-02, 3.22726555e-02,
-4.41241376e-02, 7.24015664e-03, 5.25749177e-02,
-4.09784280e-02, -9.15353894e-02, -9.17970482e-03,
3.24928947e-02, -1.59345791e-02, 5.74486665e-02,
-1.21760340e-02, 4.73094992e-02, 4.25198637e-02,
-3.79754454e-02, 1.89667307e-02, -2.49061491e-02,
```

```
7.41068572e-02, 1.09108500e-01, -7.03269169e-02,
-5.88515354e-03, -6.83640912e-02, 2.59987600e-02,
-2.42941193e-02, 4.73829145e-34, 4.54927469e-03,
-8.24521482e-02, -3.80077325e-02, -2.23266445e-02,
 3.71455029e-02, -7.45194405e-02, -1.59240700e-02,
5.37019856e-02, -5.37057780e-02, -1.83510613e-02,
 3.10776774e-02, -2.51092259e-02, 5.88086760e-03,
 9.42167267e-02, 1.41496677e-03, 3.81573220e-03,
-5.85224852e-02, -3.33878100e-02, -3.33163440e-02,
 3.47713865e-02, 5.71409101e-03, -2.82860231e-02,
7.97075927e-02, 6.13541715e-02, -2.64884066e-02,
-1.13138102e-01, 4.11813930e-02, 3.72657478e-02,
-4.12159525e-02, -2.52681561e-02, 7.96678811e-02,
-1.25984266e-01, -4.17161398e-02, 4.62167487e-02,
 2.35813558e-02, 3.56163718e-02, -3.69236208e-02,
-6.81017116e-02, 7.27936998e-02, -5.91539331e-02,
 3.95907238e-02, 4.65467572e-02, 3.15101929e-02,
 5.64673841e-02, 3.13490704e-02, 1.74114313e-02,
 6.09599128e-02, 3.65415700e-02, 3.29949334e-02,
 2.78062299e-02, -2.10567918e-02, 5.79359615e-03,
-2.46730745e-02, -4.50993404e-02, -8.94956812e-02,
-6.31225482e-02, 5.19337226e-03, -2.10509673e-02,
-2.77335066e-02, 1.34402784e-02, 4.07174081e-02,
-9.92844906e-03, -2.12341957e-02, -1.20803528e-01,
 6.77018566e-03, -1.05191432e-02, -3.53237055e-02,
-4.08614576e-02, -6.15176484e-02, 1.46740332e-01,
-1.77230090e-02, -1.44065991e-02, -6.72442839e-02,
-5.61573468e-02, 1.21265156e-02, 1.38349421e-02,
-4.40530851e-02, 3.40211205e-02, -5.01014143e-02,
-5.17557152e-02, 5.00885248e-02, 1.13820508e-01,
3.20511051e-02, -1.14666475e-02, 3.99977192e-02,
 2.32814029e-02, 3.29314359e-02, -3.13459970e-02,
1.80011056e-02, 5.82638569e-02, -4.98324484e-02,
-9.20754485e-03, -5.79370931e-02, -1.39935007e-02,
-1.16736844e-01, -1.43602499e-08, 5.05364798e-02,
 3.50800864e-02, 9.20597985e-02, -3.19862203e-03,
1.79321188e-02, -3.35250571e-02, 7.40948915e-02,
-6.73980266e-02, 1.32078677e-01, -2.93290783e-02,
 1.32912276e-02, -3.52250487e-02, -1.43077344e-01,
3.40433121e-02, 1.54904172e-01, 1.90075878e-02,
-3.48132364e-02, 8.82844925e-02, 2.01872159e-02,
-4.96340841e-02, 9.09835193e-03, 5.75696714e-02,
 9.60084423e-03, -1.13307154e-02, 4.79853265e-02,
-1.41017297e-02, -1.33626244e-03, 5.77185256e-03,
-1.78817566e-02, -3.11494861e-02, 4.64287214e-02,
-2.55541597e-02, 1.00688497e-02, 1.00694504e-03,
 1.96027048e-02, -2.57852040e-02, -3.17684561e-02,
```

```
7.56627470e-02, 9.89454314e-02, -5.81632182e-02, 7.04216631e-03, -3.12876552e-02, -1.11836335e-02, -8.75091925e-03, 1.22697335e-02, 2.83134133e-02, -2.02137064e-02, -7.75565878e-02, -7.04303430e-03, 6.09177314e-02, -1.40175289e-02, -3.43281813e-02, 6.83143884e-02, -2.18277294e-02, 6.32874444e-02, 1.42031517e-02, -3.62779683e-04, -6.10524192e-02, -1.08616399e-02, 4.41971561e-03, 1.31539032e-01, -5.70743009e-02, 6.69227317e-02, 3.03978585e-02]]), 'documents': ['Elon Musk founded SpaceX in 2002.', 'Apple just released iPhone 16 Pro.'], 'uris': None, 'included': ['metadatas', 'documents', 'embeddings'], 'data': None, 'metadatas': [None, None]}
```

## 9 Metadata Example

#### 9.1 Create a Collection with Metadata

Now let's use metadata (category, author, year). Metadata helps filter documents beyond semantic meaning.

```
[8]: # Delete old collection if exists
try:
    client.delete_collection("news_metadata")
except:
    pass

collection_meta = client.create_collection(
    name="news_metadata",
    embedding_function=ef
)

print(" Metadata collection created")
```

Metadata collection created

#### 9.2 Add Documents with Metadata

We attach metadata along with each document.

```
[9]: collection_meta.add(
    documents=[
         "Elon Musk founded SpaceX in 2002.",
         "Apple just released iPhone 16 Pro.",
         "Virat Kohli scored a century yesterday."
    ],
```

Added documents with metadata

### 9.3 Query with Metadata Filter

We filter search results by metadata. Example  $\rightarrow$  find "Apple iPhone" only inside **tech** category.

```
[10]: results = collection_meta.query(
        query_texts=["Apple iPhone"],
        n_results=2,
        where={"category": "tech"}  # filter
)

print(" Query results (filtered by category=tech):")
print(results)
```

```
Query results (filtered by category=tech): {'ids': [['doc2']], 'embeddings': None, 'documents': [['Apple just released iPhone 16 Pro.']], 'uris': None, 'included': ['metadatas', 'documents', 'distances'], 'data': None, 'metadatas': [[{'year': 2024, 'category': 'tech', 'author': 'Alice'}]], 'distances': [[0.39052993059158325]]}
```

#### 9.4 Update Metadata

Re-adding with the same ID allows us to update metadata too.

```
[11]: collection_meta.add(
          documents=["Apple unveiled iPhone 17 Ultra with AI features."],
          ids=["doc2"],
          metadatas=[{"category": "tech", "author": "Alice", "year": 2025}]
)
print(" Updated doc2 with new metadata")
```

Updated doc2 with new metadata

### 9.5 Delete by Metadata

We can delete all docs from a certain category. Example  $\rightarrow$  delete all "sports" news.

```
[12]: collection_meta.delete(where={"category": "sports"})
print(" Deleted all sports news")
```

Deleted all sports news

# 10 Summary

- Add (Create)  $\rightarrow$  collection.add()
- Read (Query)  $\rightarrow$  collection.query()
- $\mathbf{Update} \rightarrow \mathbf{re}\text{-add}$  with same  $\mathbf{id}$
- $\bullet \ \ Delete \rightarrow \texttt{collection.delete(ids=[...])} \ \ \mathrm{or} \ \ \texttt{collection.delete(where=\{...\})}$
- Metadata  $\rightarrow$  enables category/year filtering

This extends **Day94 notebook** by adding CRUD operations and metadata support.