



The success of your drug discovery innovation efforts depends heavily on the speed and accuracy in which hypotheses can be generated and tested. The Katana Graph intelligence platform fulfills this essential need for life sciences companies in ways alternative solutions simply cannot match.

Katana Graph empowers researchers, scientists, computational biologists and immuno-therapists to design, synthesize, optimize and retarget molecules; design better, targeted, safer clinical trials, and much more. Katana Graph makes this possible by providing massively scalable data ingestion, integration, query, analytics and artificial intelligence (AI) inferencing at unrivaled scale and performance.

Katana Graph has extended its support for life sciences even further, with new, seamless integration with RDKit—the widely-used open-source software package to model, analyze and visualize chemical compound structures and properties.

Katana Graph offers pre-built integration with RDKit—for *in silico* analysis that enables faster, more efficient and cost-effective drug discovery than ever before:



Get started quickly using provided reference ChEMBL27 graph database and scripts to import ChEMBL27 dataset



Write openCypher queries to search ChEMBL27 and SureChEMBL databases quickly and efficiently for similar compounds using molecular fingerprints or molecular substructure matches



Enable researchers to select the most promising options *in silico* far more quickly and cost-effectively than wet lab assay tests



Empower scientists to expedite accurate acceptance or rejection of hypotheses for new molecules or new applications



Organize and enrich existing small molecule and macromolecule databases alike by combining data from various knowledge graphs

Katana Graph with RDKit integration creates efficiencies in complex cheminformatics pipelines, enabling users to write more concise and performant code, including openCypher queries, in a unified graph environment.

Katana Graph and RDKit in Action: Example

Life sciences organizations are increasingly shifting from cheminformatics pipelines, based on traditional, smaller scaled, less integrated relational data, to large, heterogeneous, fully integrated, graph-based systems. For example, performing cheminformatics tasks such as drug hypothesis generation studies currently requires complex pipelines working in multiple different platforms and environments.

Katana Graph provides an integrated graph-RDKit cheminformatics platform where complex cheminformatics workflows can be streamlined over large, heterogeneous, biomedical knowledge graphs.

ChemInformaticians can run a drug disease association query written in openCypher, extract a subset of SMILES representation of compounds, and perform specialized chemical searches to identify similar compounds in phase IV of clinical trials. Similarity searches allow the researchers to focus on most promising candidates in assay studies.

The following openCypher query (pseudocode example) runs a similarity search for a given desired compound (in red; see below) against a SMILES string in a ChEMBL27 graph, using the rdk_fingerprint and Tanimoto similarity function available as part of our RDKit Cartridge:

```
small_molecule_query=MATCH(cc:compound_structures)-[:has_molecule]-
(b:molecule_dictionary{max_phase:4, molecule_type:'Small molecule'})
WHERE cc.canonical_smiles <> 'None' and cc.canonical_smiles IS NOT NULL
return b.chembl_id, cc.canonical_smiles, rdk_fingerprint(cc.canonical_smiles),
rdk_tanimoto_similarity(cc.canonical_smiles,
'SMILES_string_for_desired_compound_goes_here')
ORDER BY rdk_tanimoto_similarity(cc.canonical_smiles,
'SMILES_string_for_desired_compound_goes_here') DESC"
```

About Katana Graph

The Katana Graph intelligence platform interoperates with its own graph database system, providing a single platform for:



Graph Queries (contextual search)



Graph Analytics (path finding, centrality and community detection)



Graph Mining (pattern discovery)



Graph AI & Deep Machine Learning (deep learning and prediction)

at unrivaled speed, using massive graphs, scalable to multiple clusters in any cloud environment.

Contact Katana Graph today! info@katanagraph.com