

Data Quality, Standardization & Chemical Registration — Hands-on

Create **trustworthy, searchable, and compliant** chemical datasets. This hands-on module teaches **structure standardization** (salts/tautomers/protonation/valence), **ID generation** (SMILES/InChI), **duplicate detection & merge rules**, and **registration workflows** (business rules, lots/batches, audit trails) that power discovery, QSAR, and regulatory reporting. You will ship reproducible QC/QA pipelines aligned to **FAIR** principles.

Data Quality, Standardization & Chemical Registration

[Help Desk](#) · [WhatsApp](#)

Session Index

[Session 1 — Structure Standardization Foundations](#) [Session 2 — Identifiers, Matching & Dedupe](#)
[Session 3 — Registration Workflows & Audit](#) [Session 4 — Mini Capstone: QC/QA & Registry](#)

Session 1

Fee: Rs 18800 [Apply Now](#)

Structure Standardization Foundations

Normalization strategies

[salt/solvent stripping](#) [tautomer & charge
normalization](#) [valence & aromaticity models](#)

Stereochemistry & isotopes

[chiral flags/unspecified centers](#) [E/Z normalization](#)
[mixtures & polymers \(basics\)](#)

Pipelines & configs

standardizer rulesets **unit tests & golden sets**
reproducible seeds/versions

Session 2

Fee: Rs 21800 Apply Now

Identifiers, Matching & Dedupe

Canonical IDs & encodings

SMILES (canonical/isomeric) **InChI & InChIKey layers**
registration-friendly hashes

Equivalence & dedup rules

tautomer-insensitive matching **stereo-sensitive**
options **salts/mixtures handling**

Ingestion & validation

vendor SDF/CSV imports **schema checks & CV terms**
error buckets & fixes

Session 3

Fee: Rs 24800 Apply Now

Registration Workflows, Governance & Audit

Registry design & business rules

parent/lot/batch hierarchy **merges/splits &**
supersession **status & lifecycle states**

Governance & compliance

audit trails & versioning **GHS/REACH tags (light)**
privacy/IP considerations

LIMS/ELN integration

API/ETL patterns **delta loads & snapshots** **error**

handling & alerts

Session 4

Fee: Rs 28800 Apply Now

Mini Capstone: QC/QA Pipeline & Mini-Registry

Build an ingest → standardize → dedupe pipeline

Theory + Practical

Register parents/lots, generate IDs & audit entries

business rules & merges **governance checks** **QC report export**

Deliverables

QC/QA summary (PDF/HTML) **registry sample (CSV/JSON)** **configs/notebooks for reruns**