IntuitionLabs

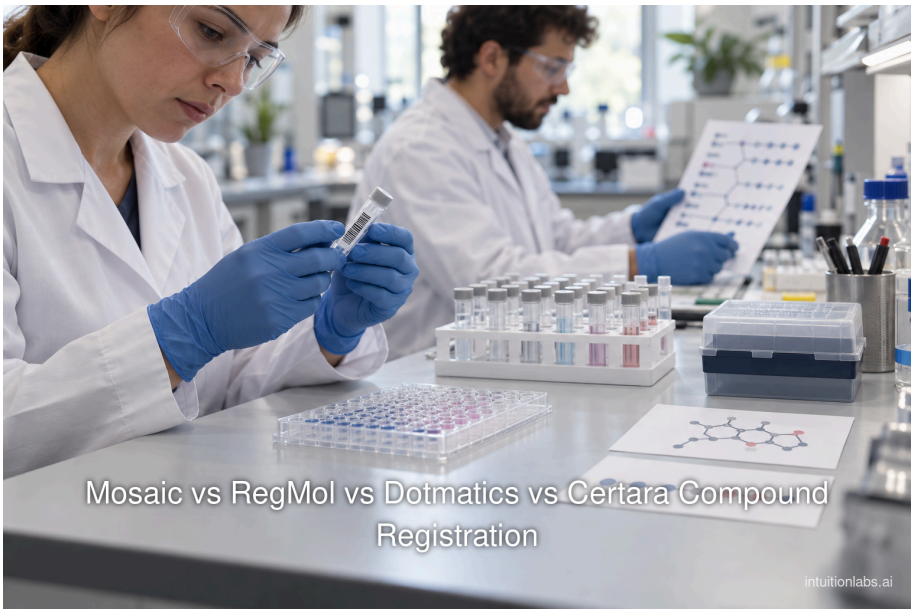
INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

Mosaic vs RegMol vs Dotmatics vs Certara Compound Registration

> Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-09-25 · compound registration · sample management · mosaic · regmol · dotmatics · certara · chemical registration · biological registration · laboratory inventory

Original article: https://intuitionlabs.ai/articles/mosaic-regmol-dotmatics-certara-registration-comparison



In this article

Executive Summary

Mosaic, RegMol, Dotmatics and Certara Compound Registration answer different parts of the same discovery data problem. A corporate registry decides whether a proposed chemical or biological entity already exists, assigns a durable identifier and records its relationships. A sample system accounts for physical material, its containers, location, quantity and movements. **Mosaic**, now a Cenevo product after the Titian and Labguru rebrand announced in July 2025, leads this comparison on documented sample operations: its published workflow spans tubes, racks, plates, orders and automated stores. Cenevo also describes Labguru as providing registration within its portfolio, so a procurement exercise should specify whether it means Mosaic alone or a Cenevo bundle. (www.cenevo.com) (www.titian.co.uk) **RegMol** explicitly covers chemical and biological registration, while Scilligence documents Inventory as a connected, separately named product. (scilligence.com)

Dotmatics publicly describes configurable chemical and biological registration, corporate and batch identifiers, entity lineage and an inventory capability with storage and aliquoting workflows. Its literature describes connections among registration, inventory and electronic lab notebook modules, but procurement must establish which components and interfaces are included in a particular offer. (www.dotmatics.com) **Certara** markets **Chemaxon Compound Registration**, following its 2024 acquisition of Chemaxon. It documents chemical normalization, parent/version/preparation hierarchy, duplicate decisions, a REST interface and several export paths. Its page also describes peptide registration with HELM; it should not be treated as proof of a general biological registry. (www.certara.com) (www.certara.com)

The practical selection is therefore an **architecture choice**, not a universal ranking. A centralized compound store should test Mosaic's request-to-dispatch and automation workflows; a team defining canonical identities across modalities should test RegMol or Dotmatics registration rules; a chemistry-led team with complex salts, stereochemistry and preparations should test Certara's documented hierarchy. The boundary can also be split: Mosaic explicitly describes linking its inventory to an external registry, including a CDD Vault example. (www.titian.co.uk) (www.titian.co.uk) Every shortlist needs a demonstration using its own structures, mixtures, sequences, barcodes, audit history and reconciliation rules.

Public evidence does not support a comparable four-vendor speed, accuracy or total-cost ranking as of September 25, 2026. The strongest usable numbers in this report serve different purposes: Cenevo reported **more than 150** Mosaic device and system integrations in its 2025 announcement, while Chemaxon documentation limits one search-page SD-file download path to **5,000 compounds**. Neither figure measures end-to-end migration performance. (www.cenevo.com) (docs.chemaxon.com) A 2018 PubChem study found that **44%** of structures passing its standardization were modified; this illustrates why migration teams must audit transformations, but it is not a vendor benchmark. (pmc.ncbi.nlm.nih.gov) Procurement should price the actual combination of registry, inventory, automation, connector, support and deployment components and run a reconciliation pilot before committing.

Introduction and Background

The phrase **compound registration software** often hides two different records. The first is the abstract entity: a structure or sequence, its parent form, assigned corporate identifier, aliases, and rules for deciding whether an incoming submission is new. The second is a physical object: a batch, vial, tube, aliquot, rack or plate well whose quantity and location can change. A physical container can be exhausted without deleting the entity, and a single entity can have many preparations, lots and containers. Chemaxon's documented parent, version and preparation model provides one concrete example of the former distinction; Mosaic's tube, rack and plate tracking illustrates the latter. (docs.chemaxon.com) (www.titian.co.uk)

A useful target architecture is a directed data flow: **scientist or contract research organization submission → normalization and uniqueness check → corporate entity ID → batch or lot → physical container and location → request, transfer and assay result**. The arrows indicate controlled exchanges, not one required vendor stack. The registry owns the identity decision and correction rules; inventory owns observed quantities and movements; an electronic lab notebook (ELN), laboratory information management system (LIMS) or assay system references stable IDs. Scilligence documents unique RegMol IDs and a separate product, Inventory, at product, lot and package levels. Dotmatics describes entity-batch-sample relationships within its platform. (scilligence.com) (scilligence.com) (www.dotmatics.com)

Conceptual data-flow diagram:

```
ENTITY LAYER       Submission → Normalize → Uniqueness decision → Corporate entity ID
                                                                  ↓
SAMPLE LAYER       Batch or lot → Container → Location → Request → Plate or assay
                         ↑                ↑
                   Registry reference   Movement and quantity history
```

The arrow between layers is the governed ID mapping. The diagram assigns decisions to layers, not to a particular vendor.

This report is for discovery informatics leads, compound-management scientists and biotech information-technology architects assessing **Mosaic versus RegMol versus Dotmatics versus Certara compound registration**. It compares documented public capabilities and turns gaps into tests, without treating a missing public statement as absent functionality. It also separates a vendor claim from independently measured evidence. As an adjacent consultancy, IntuitionLabs' first-party materials describe data integration and

governance services, not a compound registration product; that perspective informs the migration questions here and does not add a fifth software option. (intuitionlabs.ai) Its published Veeva X-Pages partnership concerns a different application surface and does not imply endorsement of any registry in this comparison. (intuitionlabs.ai)

The boundary matters because standardization changes identities. PubChem describes its own process as validating and determining a unique chemical structure, and the InChI Trust describes layers for connectivity, tautomerism, isotope, stereochemistry and charge. These sources do not dictate a corporate uniqueness policy; they show why a policy must say which distinctions survive normalization. (pubchem.ncbi.nlm.nih.gov) (www.inchi-trust.org) The same principle extends to macromolecules, where HELM is a representation of complex biomolecules, but a vendor's statement that it supports HELM does not identify the exact grammar version or every entity type. (pistoiaalliance.org)

Mosaic

Capabilities

Mosaic is the sample and logistics-centered contender. Titian and Labguru announced their rebrand as **Cenevo** in 2025 while keeping product names; Cenevo characterized Mosaic as sample-centric operations and Labguru as experimental inventory, registration and notebook work. (www.cenevo.com) The Mosaic page lists small molecules, reagents, DNA, proteins, antibodies and other sample types. That is evidence of **handling those samples**, not evidence that Mosaic applies every modality-specific uniqueness algorithm expected of an enterprise entity registry. (www.titian.co.uk)

Mosaic documents contents and location for tubes, racks and plates, storage hierarchies including freezers and automated stores, and amounts or concentrations. These are the operational fields a compound store needs when a plate map is revised or a tube is consumed. (www.titian.co.uk) (www.titian.co.uk) Its Inventory page says substance metadata may be used with or in place of a third-party registration system. The important procurement question is which system alone may create or amend the authoritative substance record, and how conflicts are detected. (www.titian.co.uk)

The published application family includes REST application programming interfaces (APIs), a dedicated connectivity route to outside substance registration and named assay connectivity with Genedata Screener. (www.titian.co.uk) (www.titian.co.uk) A documented CDD Vault pattern assigns registration to CDD Vault and inventory plus workflow to Mosaic; it is unusually clear evidence of an explicit registry-inventory boundary. For physical operations, the SampleBank demonstration describes dispensing, solubilizing and cherry-picking into a plate, with container-to-container transfer tracking. (www.titian.co.uk)

Adoption

Cenevo's rebranding announcement states **over 150 device and system integrations** for Mosaic, a vendor-reported count dated 2025. It is a breadth claim, not a published list of validated interfaces for an individual deployment. (www.cenevo.com) The Automated Store module documents pick-job creation and real-time inventory updates; Cenevo separately describes an interface that passes order and labware parameters to

HighRes Cellario OS. (www.titian.co.uk) (www.cenevo.com) These examples make Mosaic especially relevant where robotics and sample fulfillment are the main constraints, but the buyer should request a matrix by instrument model, software release and integration type.

Strengths and Limitations

Best-documented fit: requests, locations, containers and automation. Cenevo presents Foundation, Premium and Enterprise plans; Premium adds ordering, file-based liquid-handler exchange and automated stores, while Enterprise targets direct liquid-handler integration and larger automation scope. No public price in the reviewed material supports a cost comparison. (www.titian.co.uk) (www.titian.co.uk) The module description also claims a freeze-thaw audit record; a regulated deployment needs to verify configured retention, user attribution and export, since a product page is not independent validation. (www.fda.gov) Public Mosaic materials reviewed here do not establish a HELM version, salt normalization policy or the exact rule for assigning a canonical structure ID. That is an evidence gap to resolve in a demonstration, not a finding that the functions are absent.

RegMol

Capabilities

Scilligence RegMol is explicitly described as an entity-registration and testing-protocol system. Its product page names small molecules, peptides, oligonucleotides, antibodies and antibody-drug conjugates, and says users can search by chemical structure, **Hierarchical Editing Language for Macromolecules (HELM)** notation or sequence. (scilligence.com) (scilligence.com) The stated modality range is broader than a chemistry-only registration claim, but a project still needs to test its own sequence variants, conjugate definitions and monomer library. Scilligence says RegMol assigns unique IDs to chemical and biological entities; the adjacent Inventory product tracks material at product, lot and package levels. (scilligence.com)

Release notes give unusually concrete clues about rule behavior. RegMol 7.0 documents neutralization and salt stripping in parent and duplicate checks, an improved mixture uniqueness check, JavaScript Object Notation (JSON) input for API registration, and comma-separated values (CSV) or structure-data file (SDF) export from advanced search. These are versioned vendor statements, not a guarantee about a customer's current configuration. (scilligence.com) (scilligence.com) RegMol 7.4 adds sequence-based uniqueness for certain gene and ribonucleic acid (RNA) batches and an entity-parent history for virtual and non-virtual parents. (scilligence.com)

Adoption

Scilligence says its named applications can stand alone or operate together. RegMol 7.3 release notes describe project-code synchronization with Inventory and say some container controls in new installations depend on that integration. (scilligence.com) (scilligence.com) This matters when a buyer asks whether RegMol alone can control physical samples: the public record points to a defined integration boundary, and the commercial scope must be specified by quote. The 7.6.0 notes describe a **HELM2MOL** API endpoint and CSV downloads from Parent Explorer, useful probes for migration and downstream analytics. (scilligence.com)

Strengths and Limitations

RegMol merits a serious test when a laboratory wants one corporate identity service spanning chemistry and defined biologic modalities. Its public release trail is helpful for writing concrete acceptance cases, such as whether a salt is a version or new parent, whether a mixture is a duplicate, and whether an edited virtual parent retains traceable history. (scilligence.com) (scilligence.com) The named Scilligence Inventory layer should be scoped separately for location, reservation, aliquot and robot control; neither the product page nor these release notes establish that every such operation is licensed within RegMol. Public pages reviewed did not establish the exact HELM version, current deployment choices, a full export contract or an independent registration accuracy benchmark. A release note does state that one Flexviewer addition needs configuration and a nominal separate license cost, illustrating why module assumptions require line-item confirmation. (scilligence.com)

Dotmatics

Capabilities

Dotmatics publicly positions entity registration for biological and chemical research. It describes configurable entity subtypes and uniqueness rules, corporate IDs, batch and lot IDs, and lineage across entity, batch and sample records. (www.dotmatics.com) (www.dotmatics.com) The description is a strong basis for an identity-model workshop, particularly when the same program handles small molecules, proteins and conjugates. A Dotmatics page also lists DNA, RNA, peptides, proteins, antibodies and other entities, but the buyer must test the precise canonical definition for each entity type in its configured instance. (www.dotmatics.com)

Dotmatics separately documents inventory tracking for samples, reagents, plates, consumables and equipment, plus aliquoting and dispensing workflows. (www.dotmatics.com) It says the inventory layer defines a storage hierarchy and provides a location, date and quantity audit trail. (www.dotmatics.com) A platform release post says inventory connects out of the box with ELN, Data Discovery, Chemical Registration and Biological Registration modules. This documents a suite approach, while leaving commercial packaging and specific interface behavior for the proposal and demonstration. (www.dotmatics.com)

Adoption

The cited release post describes bidirectional BioRegister sequence search from Geneious Prime and SnapGene. (www.dotmatics.com) A separate RNA workflow article says candidate profiles can include a HELM string, InChI key, related entities and inventory information. (www.dotmatics.com) These are named integration and data-model examples, not a declaration that every available Dotmatics registration workflow accepts a particular HELM grammar version. The selection team should ask for a schema export and a round-trip demonstration from registration to inventory to assay results, including failed and corrected submissions.

Strengths and Limitations

Dotmatics has the broadest public **within-platform narrative** among the four in this review: the registration, inventory and notebook pages all describe connections. That is an editorial reading of the cited pages, not a measured feature score. (www.dotmatics.com) (www.dotmatics.com) It may reduce interface design work

when those modules are bought and configured together; that benefit must be tested against actual data flows and contract terms. Public materials retrieved here did not establish registration-specific API endpoints, bulk export details, deployment choices, HELM version or separate module prices. Unknown public data cannot become a zero score or a presumed product shortcoming. The procurement response should provide current documentation, a reproducible export and a demonstration with the buyer's own representative compounds and biologics.

Certara Compound Registration

Capabilities

Certara now presents this application as **Chemaxon Compound Registration** after acquiring Chemaxon in 2024. (www.certara.com) Its product page covers small molecules, salts, multicomponent compounds, polymers and peptides and describes duplicate checking across stereoisomers, tautomers, isotopes and salt forms. (www.certara.com) Peptides are described as registered using HELM notation, but the reviewed page does not specify the HELM version and does not establish the wider biological scope described for RegMol or Dotmatics. (www.certara.com) This makes the product particularly relevant when the identity-policy debate centers on complex chemical forms.

The technical documentation defines a **parent, version and preparation hierarchy**. Isotopic or charged versions and salt or solvate information sit at the version level; another preparation can be created under an existing version with a generated lot number. (docs.chemaxon.com) (docs.chemaxon.com) Such explicit semantics give migration designers a place to map source parent IDs, salt forms and batches. The documentation also states that amendments are audited and history remains available. That is a product behavior claim; any regulated use must test its own permissions, retention and inspection path. (www.fda.gov)

Adoption

Chemaxon documents a REST interface for registration and an SDF export route. Its download guidance describes CSV export and warns that structures saved as SMILES in CSV do not preserve enhanced stereochemistry. (docs.chemaxon.com) (docs.chemaxon.com) Its DSClient documentation describes downstream registration-tree data, with HTTP polling or ActiveMQ message transfer described as options. (docs.chemaxon.com) These details matter to teams feeding analytical stores or external inventory: a one-time CSV is not necessarily a faithful archival copy of every structure feature or audit event.

Strengths and Limitations

A strength of the public record is specificity about hierarchy, export and installation. The documentation names PostgreSQL or Oracle schemas for application setup, so the buyer can ask which deployment and operations model the vendor will supply. (docs.chemaxon.com) The reviewed download page limits one **search-page SDF download to 5,000 compounds**; it does not establish a whole-system throughput ceiling or bulk API limit. (docs.chemaxon.com) Public pages did not disclose price, separate DSClient licensing, a universal biologics registry or comparative speed results. Request a licensed-module schedule, an extract that preserves stereochemistry, and a full migration method before concluding that the documented export path covers the project.

Feature Comparison

Table 1 is a **requirements crosswalk**, not a product-performance ranking. Each cell describes public evidence reviewed as of **September 25, 2026**. “Confirm” means the cited material is insufficient to score a live deployment; it does not mean the capability is absent. The registry column concerns canonical identities, while the sample column concerns physical objects.

Decision axis	Mosaic	RegMol	Dotmatics	Certara Compound Registration
Primary documented role	Sample operations, with substance metadata and external registry connection (www.titian.co.uk)	Chemical and biological entity registration; Inventory is separately named (scilligence.com)	Configurable entity registration plus inventory within a wider platform (www.dotmatics.com)	Chemical and peptide registration with explicit hierarchy (docs.chemaxon.com)
Canonical identity and duplicates	Confirm structure rules and authoritative ID owner	Salt stripping, mixture check and sequence rules in release notes (scilligence.com) (scilligence.com)	Configurable uniqueness and corporate ID assignment	Stereochemical, tautomer, isotope and salt checks (www.certara.com)
Physical sample evidence	Tubes, racks, plates, locations and store jobs (www.titian.co.uk)	Separate Inventory product at product, lot and package levels	Samples, plates, aliquoting and storage hierarchy (www.dotmatics.com)	Preparation and lot hierarchy documented; confirm container operations (docs.chemaxon.com)
HELM evidence	Confirm support and version	HELM search and conversion endpoint, version unspecified	HELM string in an RNA workflow, version unspecified (www.dotmatics.com)	Peptide registration using HELM, version unspecified
Interfaces and extract	REST APIs and external registry integration (www.titian.co.uk)	JSON registration API and CSV/SDF search export in release notes	Named platform-module integration; obtain interface and export specification	REST API, SDF/CSV download and downstream DSClient path (docs.chemaxon.com) (docs.chemaxon.com)
Audit evidence	Vendor claims sample-lifecycle audit trail (www.titian.co.uk)	Entity-parent history in release notes	Vendor claims location/date/quantity trail	Documented audited amendments

The crosswalk shows why a product name alone is an inadequate request for proposal. **Mosaic plus a registry** is a coherent architecture when physical operations dominate, and the vendor's CDD Vault example demonstrates that split. **RegMol plus Scilligence Inventory** presents another explicit boundary. **Dotmatics** offers connected modules, while **Certara Compound Registration** has richly documented identity semantics and a downstream-data path. (www.dotmatics.com) (docs.chemaxon.com) None of these claims answers whether a specific licensed deployment meets local throughput, access-control or data-retention requirements.

Table 2 converts the comparison into **two separate evaluation scores** that a buyer can fill after a scripted pilot. The score is a local acceptance result, not a vendor score inferred from marketing pages. Use **0** only for an observed failure in the tested configuration, **1** for a manual workaround, **2** for a repeatable configured workflow, and **3** for a repeatable automated workflow with exportable evidence. Record **N/A** when a test has not run, and exclude N/A from the denominator. These are editorial rubric units, created for this article on September 25, 2026; they are not vendor measurements.

Test group	Registry score: evidence to collect	Physical-sample score: evidence to collect
Identity	Parent, salt, stereoisomer, mixture and sequence duplicate fixtures; assigned ID and correction history	Batch-to-container mapping; barcode uniqueness and container-parent links
Lineage	Parent/version/batch graph survives export and re-import	Aliquot, pool and plate-well ancestry survives transfer and re-labeling
Operations	Single and bulk submission queues, exception review, immutable ID handling	Request, reservation, pick, dispense, freeze-thaw and depletion events
Integration	ELN submission, registry API, full structure export, downstream change feed	Robot worklist, assay plate map, inventory API and location feed
Audit and recovery	Actor, timestamp, old and new values, reason and replay after correction	Actor, timestamp, quantity/location change and chain-of-custody replay

Report **registry score = observed registry points ÷ possible points on tests actually run** and **sample score = observed sample points ÷ possible points on tests actually run**, each with the number of completed tests. A product with many N/A results receives no overall score until coverage is adequate. This preserves a crucial distinction: absence of public documentation is a request for evidence, not an observed failure. FAIR's persistent-identifier and provenance principles support the choice to make ID mapping and lineage explicit acceptance criteria. (www.go-fair.org) (www.go-fair.org)

Pilot acceptance checklist

The evaluation team should record each result, its configuration and the exported evidence. These are proposed test steps, not vendor capability claims.

- **Parent identity:** Submit the same neutral structure twice.
- **Salt policy:** Submit a free base and its salt.
- **Stereochemistry:** Compare defined and undefined stereocenters.
- **Tautomer policy:** Compare alternate tautomeric representations.
- **Isotope policy:** Submit a labeled isotopologue.
- **Mixture policy:** Reverse component order in a mixture.
- **Polymer policy:** Vary repeat-unit and end-group descriptions.
- **Peptide policy:** Register a modified peptide and its parent.
- **Sequence policy:** Compare closely related sequence variants.
- **Conjugate policy:** Change payload, linker and attachment site.
- **Alias control:** Import partner identifiers and resolve collisions.
- **Correction route:** Amend an entity without erasing prior meaning.

- **Batch lineage:** Attach several preparations to one entity.
- **Lot lineage:** Reconcile supplier lot and internal batch IDs.
- **Container identity:** Scan duplicate and unreadable barcodes.
- **Quantity control:** Reconcile dispense, evaporation and depletion.
- **Concentration:** Record measured and nominal values separately.
- **Aliquot lineage:** Trace a daughter tube to its source.
- **Plate mapping:** Verify every well after cherry-picking.
- **Location history:** Move material across storage hierarchies.
- **Freeze-thaw:** Check event capture and review rules.
- **Reservation:** Prevent simultaneous requests from double allocating material.
- **Robot interruption:** Resume a partially completed transfer safely.
- **ELN submission:** Register from a notebook with stable references.
- **Assay exchange:** Return results with entity and plate identifiers.
- **API retry:** Repeat an integration request without duplicate creation.
- **Bulk export:** Rebuild structure, lineage and status outside the tool.
- **Audit export:** Recover actor, time, previous value and reason.
- **Access control:** Demonstrate reviewer and administrator permissions.
- **Reconciliation:** Account for every source entity and container.

The team can use the same challenge set in every demonstration, while allowing each vendor to show its configured method. A completed test needs a reproducible result and an exportable record; an unrun test remains N/A.

Performance and Benchmarks

There is **no public, controlled four-way benchmark** in the reviewed sources for registry accuracy, sample-order cycle time, migration throughput or total cost. Published vendor counts and limits should remain in their own units. Cenevo's **more than 150 integrations** is a vendor-reported count of device and system integrations from its 2025 announcement; the announcement does not specify a geography. Chemaxon's **5,000-compound** limit concerns one search-page SDF download operation described in documentation updated in August 2026; geography is not applicable to that software limit. Neither number can be divided into the other or used to rank registration speed. (www.cenevo.com) (docs.chemaxon.com)

A useful benchmark protocol begins with a fixed, labeled challenge set from the buyer's own estate. It should include neutral forms and salts, tautomers, isotopologues, stereoisomers, mixtures, polymers, peptides, sequences, conjugates and deliberately malformed entries. The exact expected outcome for each pair is approved before the vendors run it. The InChI specification's layered identity representation helps explain why a shared “duplicate” percentage is meaningless unless the layers and transformations are fixed. (www.inchi-trust.org) The InChI Trust also calls InChIKey *nearly* unique, a reminder to retain the full structure and policy rather than relying on a short identifier as an infallible primary key. (www.inchi-trust.org)

For physical workflows, measure request receipt to dispatch, pick accuracy, container traceability, plate-well lineage, concentration updates and recovery after a robot interruption. A sample stability study considered water, oxygen, freeze-thaw cycles and container material, so teams should define which state changes

actually matter to their assay context rather than counting every transfer as equivalent. (pubmed.ncbi.nlm.nih.gov) Mosaic documents automated pick jobs and freeze-thaw tracking; Dotmatics describes aliquoting and a location/date/quantity trail. These are testable feature claims, not comparative latency measurements.

Auditability needs a scenario, not a checkbox. The U.S. Food and Drug Administration's guidance for computerized systems used in clinical trials defines an audit trail by its ability to reconstruct events and says record changes should preserve original information. This is generic context, not a claim that the four discovery products are regulated clinical-trial systems. (www.fda.gov) A buyer can nonetheless test who changed a chemical parent, who approved a correction, which downstream IDs changed, and whether the old sample record remains intelligible. Certara documents audited amendments, Scilligence documents parent history, Dotmatics describes an inventory trail, and Mosaic claims a sample-lifecycle trail.

Data Analysis and Evidence

The most relevant **independent quantitative evidence** concerns structure transformation risk, not these vendors' relative performance. In a **2018 PubChem standardization analysis** of its public-substance data, researchers reported a **0.36% rejection rate**, that **44%** of passing structures were modified, and a **14.5% reduction** in unique structures after standardization. The denominator is PubChem substance structures, and geography is not applicable to the study population. Its algorithm and units differ from a private corporate registry, so the figures are warning signals for a migration design, not expected rejection or duplicate rates for Mosaic, RegMol, Dotmatics or Certara. (pmc.ncbi.nlm.nih.gov) (pmc.ncbi.nlm.nih.gov) (pmc.ncbi.nlm.nih.gov) PubChem's current standardization documentation still frames the task as validation and determination of a unique structure. (pubchem.ncbi.nlm.nih.gov)

One public reference resource illustrates scale without implying vendor capacity. EMBL-EBI's ChEBI page reported **more than 195,000 entries** when accessed on September 25, 2026. The unit is database entries and geography is not applicable; it is neither a market-size estimate nor a capacity benchmark for the products here. (www.ebi.ac.uk) ChEBI's downloadable SDF, ontology and SQL formats also illustrate why migration teams should retain structured crosswalks and external identifiers rather than flattening everything into a label field. (www.ebi.ac.uk) (www.ebi.ac.uk)

A practical **migration reconciliation** should be signed off as arithmetic and as identity policy. The following is a **hypothetical example**, created for this report on September 25, 2026; geography is **not applicable** and units are **source entity records and physical container records**. Suppose the source registry contains **12,000 entity records**. After an agreed normalization run, **11,520** are accepted, **180** are rejected for review, and **300** map to already accepted duplicates. These categories are mutually exclusive: **11,520 + 180 + 300 = 12,000**. The accepted rate is **96.0%**, rejection rate **1.5%**, and duplicate rate **2.5%**, each divided by the 12,000 source-entity-record denominator. These numbers are arithmetic examples, not measured vendor outcomes. The trial should save each original ID, submitted structure or sequence, normalized representation, reason code and target ID. A published discussion of selected chemical-identity layers supports testing alternative definitions of “same,” rather than assuming a single universally correct duplicate rule. (pmc.ncbi.nlm.nih.gov)

Now suppose a separate source inventory contains **20,000 physical container records**, of which **19,940** join to an accepted target entity or approved retained alias and **60** remain unmatched. The unmatched-container rate is **0.30%** of 20,000, and it must not be confused with the 1.5% entity rejection rate because the denominators differ. Set tolerances before the pilot, for example **zero unresolved active containers** in the cutover set and a separately approved queue for historical non-active material; these are proposed acceptance criteria, not industry benchmarks. Reconcile counts by status, site, temperature zone, plate and owner, then sample high-consequence lineages by scanning actual labels. InChI's nonproprietary identifier is useful as a cross-system comparison field, but a corporate ID and its history remain the authority for the migration. (iupac.org)

The strongest conclusion from the numbers is methodological: a “successful import” count alone can mask transformations, merged duplicates and orphaned physical material. A complete report displays source, accepted, rejected, merged and unmatched counts with denominators, provenance and exception ownership. The FAIR principles explicitly call for persistent identifiers and detailed provenance, consistent with retaining both the source-to-target crosswalk and the decisions behind it. (www.go-fair.org) (www.go-fair.org)

Implications and Future Directions

For a **virtual biotech**, the first decision is who owns canonical identity when synthesis, storage and assays are outsourced. A registry-centered procurement may be sensible if the company needs to compare material from several partners, while a lighter sample system may suffice when an external party already supplies authoritative IDs. RegMol's declared multi-modality scope and Dotmatics' entity-batch-sample hierarchy are starting points for the former. Mosaic's documented outside-registry integration is a starting point for the latter. The recommendation is conditional: test the data exchange with the actual contract research organization and require the ability to export every original and corrected ID.

For a **medicinal chemistry group**, the decisive cases are usually stereochemistry, salt forms, mixtures and repeat preparations. Certara documents parent/version/preparation semantics and duplicate checks across several chemical distinctions; RegMol release notes describe salt stripping and mixture checks. The team should agree whether a free base, hydrochloride salt and mixture are new parents, versions or preparations, then evaluate each system against that written policy. It should also test whether an SDF round trip preserves enhanced stereochemistry, since Chemaxon explicitly warns that one CSV/SMILES export does not. (docs.chemaxon.com) No generic ranking can resolve a lab's identity policy.

For a **centralized compound store**, the expensive error is often a wrong container, quantity, plate position or pick instruction. Mosaic's storage and automation documentation maps closely to those tasks, while Dotmatics' inventory pages show aliquoting and location audit functions. If corporate registration is already settled, the store can evaluate those physical workflows against existing registry integration rather than replacing the registry by default. If no registry exists, the project should include an explicit identity authority and uniqueness-policy workstream.

The implementation boundary should be contractually and technically explicit. Assign one system to create entity IDs, one to own physical container IDs, and named interfaces for batch creation, alias correction, plate maps and assay references. Specify whether the transfer is event-driven, API-based or scheduled; how idempotent retries work; how orphaned events are reported; and how full exports are produced. Certara's

DSClient documentation describes a downstream stream, Mosaic publishes REST APIs, and Scilligence has versioned registration API notes. (docs.chemaxon.com) IntuitionLabs' own data-engineering description includes pipelines, transformation and quality assurance, which is the appropriate advisory lens for testing this boundary rather than presenting the consultancy as a competing registration product. (intuitionlabs.ai)

Future evaluations should request **current** vendor evidence: licensed-module maps, exact HELM grammar and monomer support, bulk export coverage, deployment architecture, audit retention, and a tested interface inventory. The Pistoia Alliance's HELM materials demonstrate that versions matter, while an International Organization for Standardization biobanking standard and International Society for Biological and Environmental Repositories best practices address distinct quality contexts and should not be casually converted into product certifications. (pistoiaalliance.org) (committee.iso.org) (www.isber.org) The final choice should follow the organization's controlled challenge set and total architecture cost, with N/A kept visible until it has been tested.

Conclusion

There is no single winner for all compound-registration projects. Mosaic is best understood from its documented sample, container and automation workflows; RegMol from its chemical and biological identity rules linked to Scilligence Inventory; Dotmatics from connected registration and inventory capabilities; and Certara Compound Registration from its detailed chemical hierarchy, normalization and export documentation. (www.dotmatics.com)

The most useful procurement question is **which system is authoritative for each record**. Define the entity ID owner, batch or lot owner, physical container owner, and controlled exchange between them before comparing screens or subscriptions. Use the two-part registry and physical-sample scorecard only after running identical acceptance cases, and publish unresolved evidence as N/A. For migration, reconcile entity transformations and duplicate decisions separately from unmatched physical containers, preserving the old-to-new ID map and audit history.

As of September 25, 2026, public sources substantiate specific capabilities but do not provide comparable prices, accuracy measurements, throughput or complete module entitlements. The responsible decision is a scored pilot with vendor-supplied contract and technical evidence. A buyer that can reproduce identity decisions, trace a vial to a plate well, and restore its lineage after correction has a stronger basis for selection than any unqualified “best software” label.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://intuitionlabs.ai/articles/open-source-lims-functionalities
2. https://www.cenevo.com/press/labguru-and-titian-software-rebrand-as-cenevo
3. https://www.titian.co.uk/mosaic-module/automated-stores
4. https://scilligence.com/scilligence-rd-informatics-software-for-pharmaceutical-biotechnology-organizations/
5. https://www.dotmatics.com/capabilities/entity-registration

6. https://www.certara.com/brochure/certara-discovery-streamline-your-drug-discovery-process/
7. https://www.certara.com/compound-registration/
8. https://www.titian.co.uk/en-us/mosaic-module/connectivity
9. https://www.titian.co.uk/en-us/cdd-vault
10. https://docs.chemaxon.com/latest/compound-registration_download.html
11. https://pmc.ncbi.nlm.nih.gov/articles/PMC6086778/
12. https://docs.chemaxon.com/latest/compound-registration_overview.html
13. https://www.titian.co.uk/mosaic-sample-management
14. https://intuitionlabs.ai/services/data-engineering-bi
15. https://intuitionlabs.ai/about/veeva-partner
16. https://pubchem.ncbi.nlm.nih.gov/docs/standardization-service
17. https://www.inchi-trust.org/about-the-inchi-standard/
18. https://pistoiaalliance.org/project/helm-project/
19. https://www.titian.co.uk/en-us/mosaic-module/inventory-tracking
20. https://www.titian.co.uk/mosaic-modules
21. https://www.titian.co.uk/mosaic-samplebank-demo-video-transcript
22. https://www.cenevo.com/press/highres-and-cenevo-co-marketing-partnership
23. https://www.titian.co.uk/mosaic-subscriptions
24. https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/fda-bioresearch-monitoring-information/guidance-industry-computerized-systems-used-clinical-trials
25. https://scilligence.com/regmol/
26. https://scilligence.com/regmol-7-0-release-notes/
27. https://scilligence.com/regmol-7-4-release-notes/
28. https://scilligence.com/
29. https://scilligence.com/regmol-7-3-release-notes/
30. https://scilligence.com/regmol-7-6-0-external-release-notes/
31. https://www.dotmatics.com/capabilities/lab-inventory-management
32. https://www.dotmatics.com/blog/whats-new-in-dotmatics-platform-version-6-1
33. https://www.dotmatics.com/blog/cross-functional-informatics-solutions-for-rna-drug-discovery
34. https://docs.chemaxon.com/lts-platinum/compound-registration_api-documentation.html
35. https://docs.chemaxon.com/latest/dsclient/architecture.html
36. https://docs.chemaxon.com/latest/compound-registration_installation-requirements.html
37. https://docs.chemaxon.com/latest/dsclient/index.html
38. https://www.go-fair.org/fair-principles/?trk=article-ssr-frontend-pulse_little-text-block
39. https://www.inchi-trust.org/technical-faq/
40. https://pubmed.ncbi.nlm.nih.gov/12857383/
41. https://www.ebi.ac.uk/chebi

42. https://www.ebi.ac.uk/chebi/downloads
43. https://pmc.ncbi.nlm.nih.gov/articles/PMC11351014/
44. https://iupac.org/who-we-are/divisions/division-details/inchi/
45. https://pistoiaalliance.org/resource-library/helm-software-for-biopolymers/
46. https://committee.iso.org/standard/67888.html?browse=ics
47. https://www.isber.org/page/BPR

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore AI enablement, pharma and biotech software development, data engineering and BI, and Veeva services. Contact IntuitionLabs to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: https://intuitionlabs.ai