



A Sneak Peek at the ChemAxon Compound Registration Service

Jonathan Lee

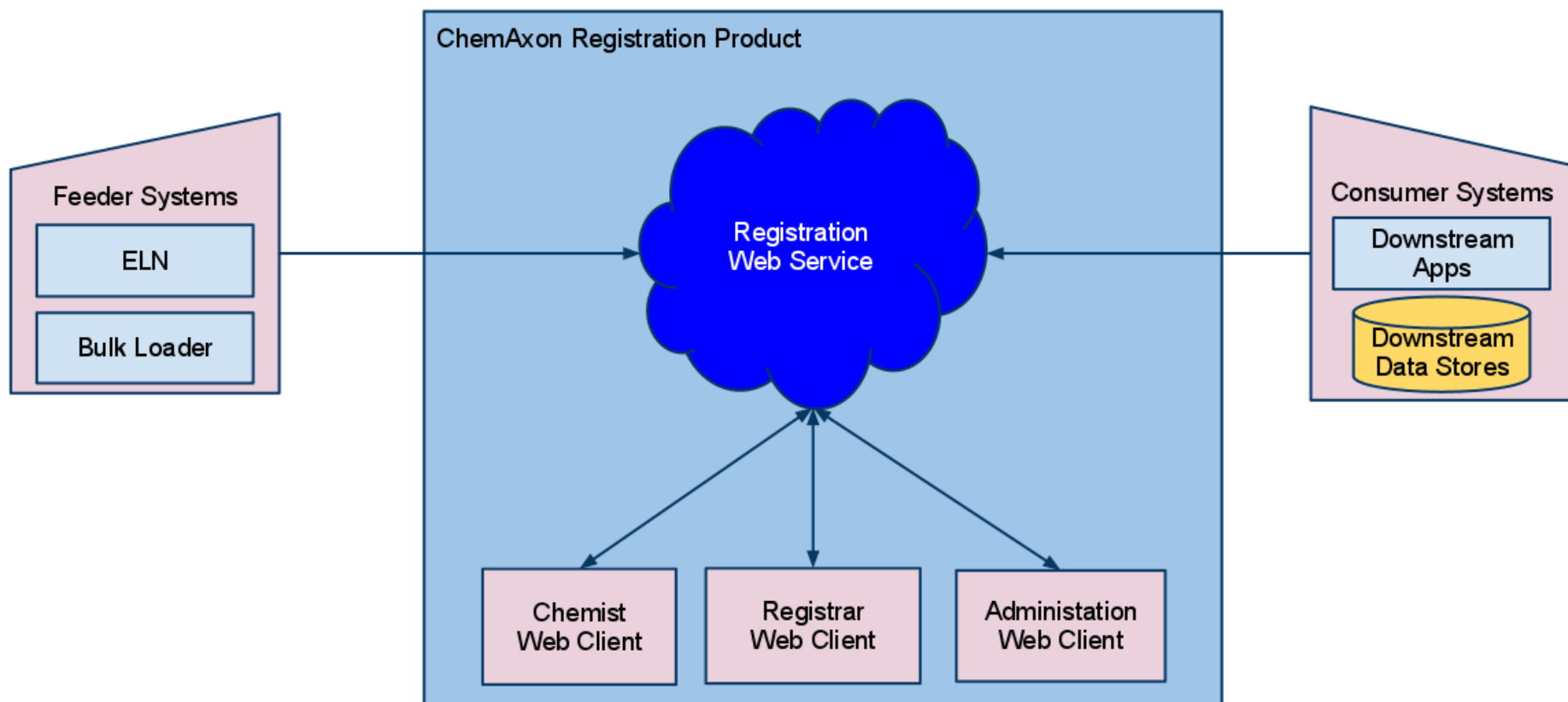


ChemAxon
Solutions for Cheminformatics

Introduction

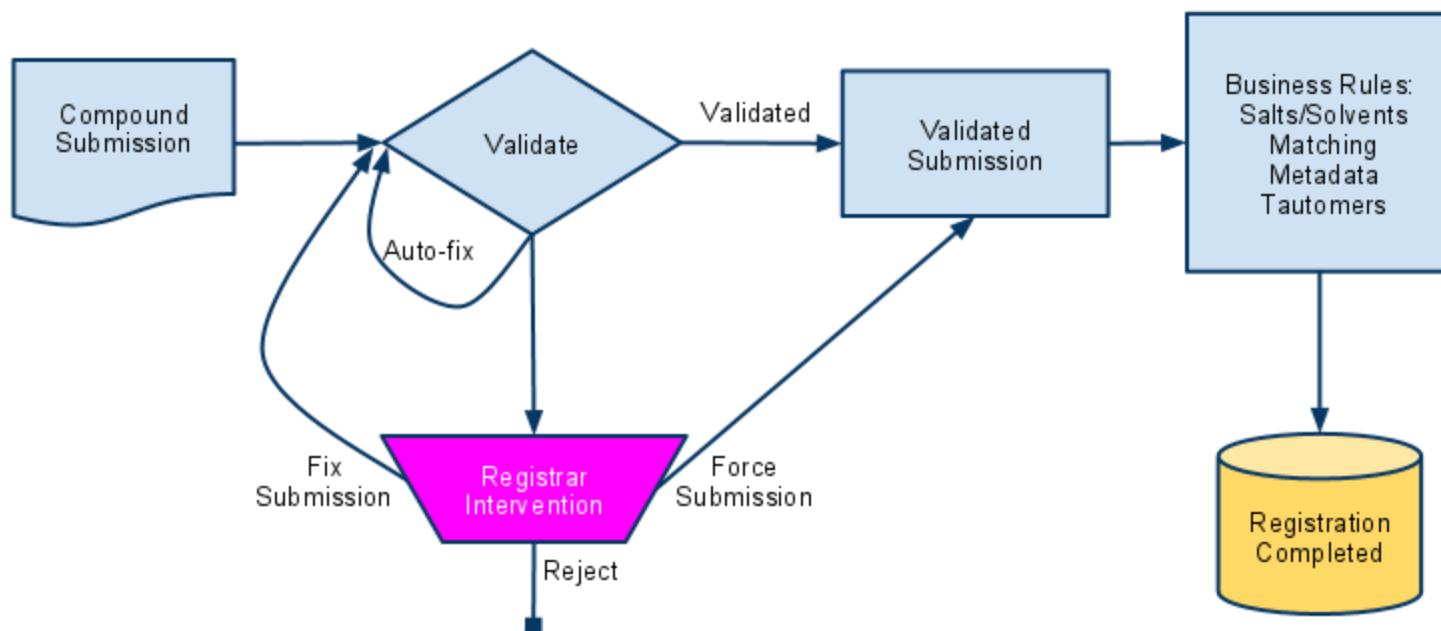
- ChemAxon and GlaxoSmithKline formed a collaboration to replace and improve the current GSK registration service.
- The ChemAxon Registration Service is a service module responsible for the submission and administration of compounds into a registration data store.

Architecture



Submission Workflow

Basic Workflow for Compound Submission



Technology Details

- Compound Input Format
 - Structural Information in MOL V3000
 - Chemically Significant Metadata (e.g. stereo isomer ranges, textual description)
- JChem Base
- ChemDraw and Marvin Compatible
- Oracle, MySQL Database
- Tomcat Web Container
- AJAX Web Client
- Service Oriented Architecture

Major Features

- Automatic Validation and Fix
 - Standardizer
 - Structure Checker
- Single and Multicomponent Compounds
 - Alternates, Formulations, Mixtures
 - Small Markush, Indefinite Locants
- Salt/Solvent Handling
- Automatic Salt/Solvent Split
- Registry Search Client (Duplicate, Taut., 2D)

Major Features

- Amendment of Registered Compound (Structure, Salt, Metadata, Hierarchy)
- Audit History of Amendments
- Administration
 - User and Roles
 - Salt/Solvent Dictionary
 - Standardizer and Structure Checker Config Files
- Role-Customizable Amendment and Access

Chemistry Features

- Tautomer Recognition
- Stereochemistry Intelligence
 - Double Bonds
 - Tetrahedral Stereo
 - Enhanced Stereo
 - Basic Allenes
 - Syn/Anti, Endo/Exo

Screenshots

Registration Client

[Search]

Registration User Interface - Submissions

SubmissionId	Source	UserId	Created	SubmissionType	LnbRef	Status
157	REGISTRAR	DEFAULTUSER	2011 05 09 00:00:00+02:00	AutoRegister	N1234-12-8	UnknowError
144	REGISTRAR	DEFAULTUSER	2011 05 09 00:00:00+02:00	AutoRegister	N1234-12-8	UnknowError
150	REGISTRAR	DEFAULTUSER	2011 05 09 00:00:00+02:00	UpdateStructure		UnknowError
152	REGISTRAR	DEFAULTUSER	2011 05 09 00:00:00+02:00	UpdateStructure		UnknowError
			2011 05 09 00:00:00+02:00	UpdateStructure		UnknowError
			2011 05 09 00:00:00+02:00	UpdateStructure		UnknowError
			2011 05 09 00:00:00+02:00	UpdateStructure		UnknowError
			2011 05 09 00:00:00+02:00	AutoRegister	N12345-123-2	LotIdDuplicated
			2011 05 09 00:00:00+02:00	AutoRegister	N12345-123-4	InvalidSalt
			2011 05 09 00:00:00+02:00	RegisterLot	N11111-10-1	LnbRefDuplicated

Query ✖

Identifier

UserId

Source

Date From

Date To

1/2 10

Future Plans

- Marvin Input Format
- Robust Migration and Bulk Loading tools
- Pluggable Compound ID Generation
- Configurable Hierarchy Levels
- More stereochemistry features
 - All allenes and cumulenes
 - Hindered rotamers (atrop)
 - Fischer/Haworth projections
 - Helicenes

Registration Service Team

ChemAxon

- Edvard Buki
- Jonathan Lee
- Roland Molnar
- Diana Nagy
- Akos Papp
- Attila Tajti

GlaxoSmithKline

- Matt Jones
- Paul Wallace
- Matt Biggs
- Charlie Wilkins
- Rama Bhamidipati
- Ian Mawer