The CDD data-platform – How does CDD benefit the Drug Discovery Community?

1. What is CDD?
2. How is CDD different from other databases or methods to store and manage data?
3. What are CDD’s benefits for scientists?
4. How does the CDD community impact and evolve the CDD platform? CDD’s Open Innovation model.

1. What is CDD?

CDD is a data-management platform designed to organize and manage scientific data, particularly chemical structures and biological results, assay data. CDD is a “cloud” application serving scientists since 2004. Originally geared towards the academic scientific community (very easy to use and very affordable), the community grew to include scientists from large Pharma companies, small Biotechs, non-profit organizations, academic labs, screening centers and many more.

CDD has three distinct products that play together seamlessly. “CDD Vault” stores data securely and privately, visible only to the Vault user(s). “CDD Collaborate” provides tools for CDD Vault users to collaborate securely by exchange designated Intellectual Property (IP) between Vaults. “CDD Public” allows CDD Vault users to publish their data to the entire community if and when desired, and also includes a growing body of curated data that is not easily available anywhere else or only with limitations. CDD’s data collections can be viewed and mined together, so that Vault data can be compared to Collaborative and Public data.

2. How is CDD different from other databases or methods to store and manage data?

There is a large variety of data management systems available and many offer the ability to manage scientific data. This can be as simple as Excel spreadsheets, or as complex as the large traditional in-house systems big Pharma typically has implemented, or home-grown self programmed solutions which talented individuals created to address their groups’ specific needs. While the complex in-house systems include endless tools and are scalable, they are also expensive. Non-commercial self programmed software, while experienced as very cheap, often lacks scalability, sustainability, high security needed in Drug Discovery, and resources like active customer support.
3. **CDD combines all advantages of scalability, highest security, rich tools, and professional customer support with a very low price tag** (low cost of ownership) – as it resides in the “cloud”, major investments are shared by the growing user community.

**Advantages of the cloud:**

- **CDD is a cloud application, and thus purely web-based.** There is no overhead of maintaining or supporting your own hardware or software.
  - Secure log-in into your CDD Vault from any computer- PC or MAC, using IE, Firefox, or Safari.
  - No need for you to buy and maintain a server, spend money and resources to maintain the application, or to ensure security – our experts will do all that for you.
  - No implementation hassle. CDD works right “out of the box”, login and go.
  - Secure, scalable, fast, easy to use, and inexpensive.
  - Enables scientific collaborations: Real time sharing without the need to have IT expertise (VPN not required).

**Security:**

- **IP in Drug Discovery is extremely sensitive.** CDD has a multi-page detailed security whitepaper that can be provided upon request. The key points are:
  - All traffic between your browser and the server is encrypted
  - CDD implements IP recognition for your computer, upon request.
  - CDD runs the server at a professional Colocation facility that specializes in hosting. The facility is monitored, guarded, with extra protection against intrusion, fire, earthquake and other hazards, has hardware and software firewalls and special procedures to permit only designated staff access to the server.
  - Staff screening, security policies, regular staff education on security.
  - Regular nightly backups (encrypted) with a redundant location.
  - Fine-grained user privileges to allow only approved users full database control.
  - Audit trail.
  - Sophisticated Collaboration features with approval trails prevent accidental data sharing.
  - Much more, please request the security whitepaper.

**Benefits for Scientists in Biology, Chemistry, Drug Discovery**

- CDD helps companies and labs to develop drugs faster and at lower costs. CDD helps scientists save valuable time, money, and intellectual property, by organizing precious research results securely, and enabling data mining and visualization:
  - Identify hits and lead compounds.
  - Identify potential toxicology issues and potential ADME issues EARLY.
  - Identify issues in acquired libraries.
Avoid redundant screening, which is expensive, futile and happens frequently.
Keep your data when scientists are leaving your group, and educate new scientists on the work done by providing them access.
Utilize simple yet powerful and sophisticated data mining tools to understand your data: visualization, statistics, structure-based queries, multi-step Boolean queries are at your fingertips.

- CDD connects molecules with targets, chemists with biologists, Academia with Industry, scientists with collaborators. CDD enables the power for a network towards your Drug Discovery goal.

- Share results in real time with a colleague or collaborator.
- Consolidate variable data sources into a single repository.
- Upload compound structures with MOL, SMILES, or IUPAC structure formats.
- Upload biological results from SD or CSV files.
- Migrate your existing databases easily from SD or CSV files.
- View structure - activity relationships (SAR) quickly, as soon as you import your data.
- Share only the subset of data a collaborator needs to see.
- Receive the data you are interested in from your collaborators.
- Mine data across collaborations to avoid redundant work, or to identify hits and leads, ADMET issues.
- Use our growing body of public data to compare structures and data with your own results.
- Share your published insights with the CDD community in one easy step, by publishing data from your Vault to CDD Public.

- CDD is scalable for large labs, collaborative consortiums, big datasets, global collaborations
- CDD’s professional customer support is available to help, train, guide, inspire and motivate users.

4. **The CDD community guides the development of the database.** CDD’s open innovation model allows customers to invest in prioritizing feature development that ultimately benefits the entire community. Our development methods and dedication to customer needs, allow us to continuously evolve and improve the product.

- CDD is very easy to use and has a short learning curve.
- The GUI uses well known Web 2.0 navigation and the system can be used without any database knowledge.
- One major release per month, and additional minor releases ensure rapid evolution.
- CDD uses modern methodologies, Ruby on Rails, Agile Development, Pair Programming and automated testing to ensure highest quality implementations.
## CDD – Industrial strength at an Academic price

<table>
<thead>
<tr>
<th>Benefit</th>
<th>CDD</th>
<th>Excel</th>
<th>“traditional in-house system”</th>
<th>Self built</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalable</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>High security</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Audit trail</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>1*</td>
</tr>
<tr>
<td>Fine grained user management</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Professional customer support</td>
<td>✓</td>
<td>2*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enabling collaborations</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Ease of access</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>No infrastructure set-up needed</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>6*</td>
</tr>
<tr>
<td>Ease of use</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Sophisticated data mining, structure queries, Boolean, visualization</td>
<td>✓</td>
<td>3*</td>
<td>✓</td>
<td>4*</td>
</tr>
<tr>
<td>Regular release schedule, new features</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total low cost of ownership</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>5*</td>
</tr>
</tbody>
</table>

1* Some systems provide an audit trail
2* available but often not purchased
3* Excel while not being a relational database offers a rich data analysis environment plus plug ins can be purchased to address specific needs
4* depending on the purpose of the system they may exhibit advanced mining towards a specific need
5* to count: Hardware, other software, the human resources to maintain and document and implement, run basic security etc. may be overall low if low service level or high if high service level.
6* local installations require implementation for the users