Enhancing enterprise software with KNIME

7th KNIME User Group Meeting, 12th-13th Feb 2014, Zurich
Dr. Fabian Bös
Certara = $\sum_{k=1}^{3} \text{(Tripos, SimCyp, Pharsight)}$

- **Certara** was formed through the integration of Tripos, SimCyp and Pharsight, creating a company that offers **Scientific Consulting** and **Software Products** with focus on model based approaches in Drug Discovery and Development.

- With these capabilities and by our internationally recognized and experienced team of scientists, **Certara** can help you to be more confident in the decisions you have to make in order to bring your products to the market: **Fast, Safe and Viable**.
Great to be here again!

• This is my 5th KNIME user group meeting, so what did I talk about the last four years?

2010: Introducing the Tripos Chemistry Extensions (TCE)
2011: TCE “free nodes”, basic cheminformatics tasks
2012: TCE “licensed nodes”, access to unique Tripos technology
2013: KNIME-based scoring functions in de novo design; Muse

2014: Something completely different …
Certara’s D360

• D360 is an enterprise-class solution that allows researchers to easily access data in immediately usable views.

• D360 helps analyze data in order to gain insight and understanding, to collaborate cross the organization with the goal to drive action and to move projects forward more effectively.
D360 Value Model

• Certara’s goal in developing D360 has been to provide the **best solution** tackling long existing problems in scientific decision making.

• D360 provides value in areas spanning from **efficiency** up-to **quality and excellence** in science.

• Feedback of our existing clients **confirm this vision**.
How KNIME and D360 can play together

- D360 can be extended by custom plugins, wrapping existing 3rd party or in-house developed code
- Configurable interface to web services
- REST API to core D360 functionality, e.g. annotations, query building etc.
Use case scenario

- Project to design multi-target kinase inhibitors
- New compounds are measured against panel of kinases, data in cooperate database
- Expert scientist to rank compounds using inhibition data, physchem properties and off-targets.
- Expert ranking to be made accessible to all user, e.g. research management

All subsequent examples use CHEMBL data for demonstration purpose.
The kinase panel dataset
Switching to KNIME

Data cleaning

Applying two predictive models

Reading dataset of scheduled query from D360 server as SDF

Pareto ranking of compounds
Switching to KNIME

Applying two predictive models

Data cleaning

Reading dataset of scheduled query from D360 server as SDF

Pareto ranking of compounds
### Switching to KNIME

#### Applying two predictive models

1. Data cleaning
2. Reading dataset of scheduled query from D360 server as SDF
3. Pareto ranking of compounds

[Table Output]

<table>
<thead>
<tr>
<th>Row ID</th>
<th>VEGFR</th>
<th>c-Abl</th>
<th>HCK</th>
<th>PI3Kalpha</th>
<th>PI3Kbeta</th>
<th>PI3Kdelta</th>
<th>PI3Kgamma</th>
<th>Mol_Wt</th>
<th>LogP</th>
<th>Activity (S)</th>
<th>Activity (I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row 15</td>
<td>32</td>
<td>5.8</td>
<td>176</td>
<td>1,350</td>
<td>150</td>
<td>293</td>
<td>315.758</td>
<td>3.126</td>
<td>inactive</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Row 16</td>
<td>21</td>
<td>8.8</td>
<td>113</td>
<td>550</td>
<td>20</td>
<td>58</td>
<td>331.757</td>
<td>2.332</td>
<td>inactive</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Row 17</td>
<td>14</td>
<td>9.2</td>
<td>187</td>
<td>613</td>
<td>65</td>
<td>209</td>
<td>287.292</td>
<td>2.472</td>
<td>inactive</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Row 18</td>
<td>295</td>
<td>15</td>
<td>652</td>
<td>3,220</td>
<td>173</td>
<td>123</td>
<td>292.338</td>
<td>2.757</td>
<td>inactive</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Row 19</td>
<td>6.4</td>
<td>5.7</td>
<td>347</td>
<td>1,240</td>
<td>58</td>
<td>283</td>
<td>327.356</td>
<td>3.589</td>
<td>inactive</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Row 20</td>
<td>52</td>
<td>47</td>
<td>175</td>
<td>350</td>
<td>46</td>
<td>340</td>
<td>287.292</td>
<td>2.727</td>
<td>inactive</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Row 21</td>
<td>410</td>
<td>190</td>
<td>11,000</td>
<td>50,000</td>
<td>12,000</td>
<td>13,000</td>
<td>398.474</td>
<td>2.257</td>
<td>inactive</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

[Output Data Table]
From KNIME back to D360

Writing back the ranking into an D360 annotation using REST API
From KNIME back to D360

Writing back the ranking into an D360 annotation using REST API
Want to know more?

- Unlocking the value of preclinical safety data
  - Live webinar, February 26th, 2014
  - Focus on how Certara’s D360 can help to unlock the value of preclinical data, by making datasets readily queryable and usable - for analysis in D360 or external tools, to generate and utilize models, and to generate visualizations and summary reports.

  - [http://goo.gl/ZPcVLJ](http://goo.gl/ZPcVLJ)

- Agile Access to Drug Discovery Data with D360
  - Recorded webinar, Nov. 2013
  - Focus on utilizing D360 to get the most out of drug discovery data by utilizing simple query building with integrated chemistry tools and analysis tools, dynamically linked data visualizations and easy to use annotations.

  - [http://goo.gl/3rthZW](http://goo.gl/3rthZW)
Thank you.

contact_us@certara.com