UGM Boston Oct. 4th 23

Chemaxon

Project team data visualization and decision support solutions: core requirements and must-have features

John McNeil
Life Sciences R&D
Informatics Strategist

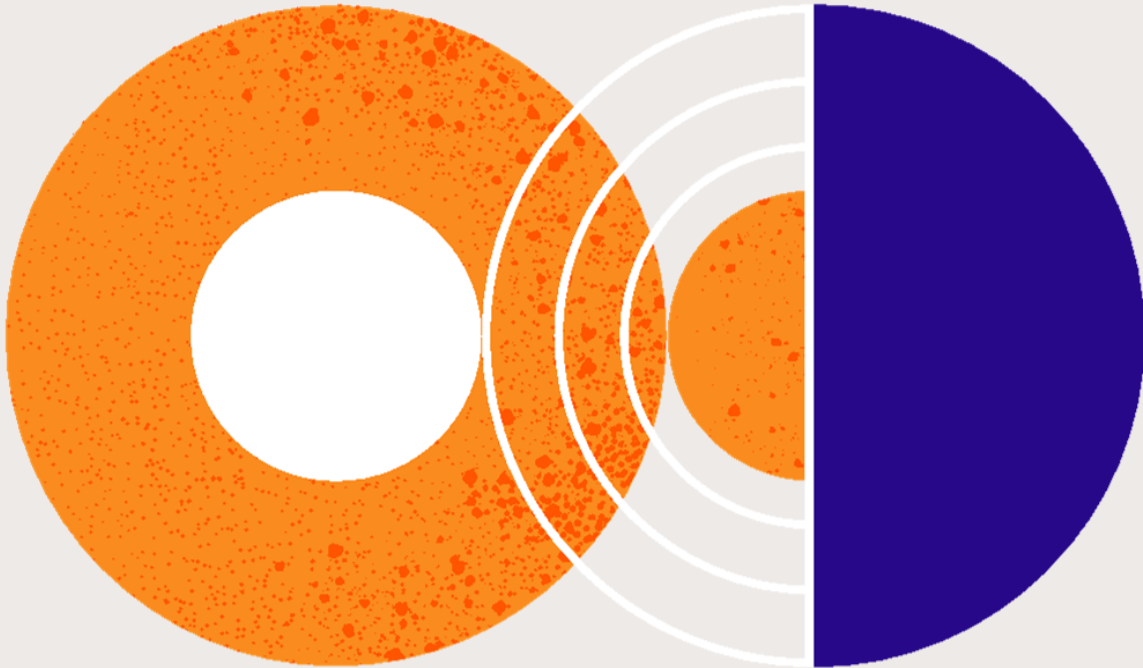
Project Team Data Visualization and Decision Support Solutions

UGM Boston Oct. 4th 23

Chemaxon

Core requirements and must-have features

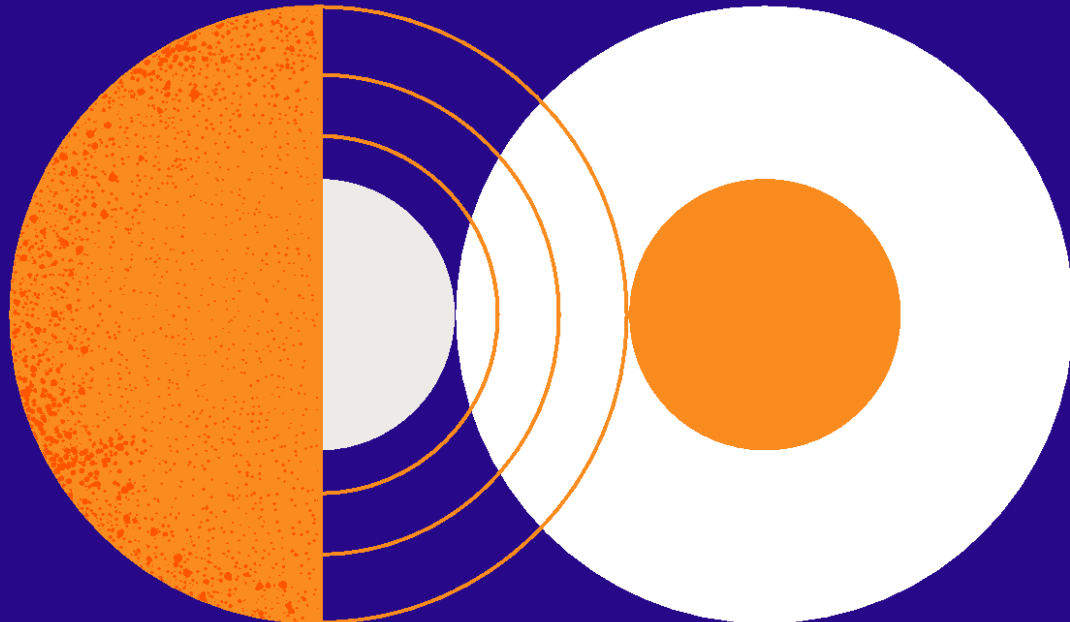
John A. McNeil



UGM Boston Oct. 4th 23

Chemaxon

Organizing Requirements



Science to Inform Decisions

All the effort and expense of research is only useful when it generates data that drives good decisions

Project Teams Members are Specialists

- Each discipline, like Medicinal Chemistry, Assay Biology, and Pharmacology, requires specialist analysis tools
- Project teams and members also need specialized project-analysis tools
- Each team member will need both discipline-specific tools and project team tools

Project Team Roles and Goals

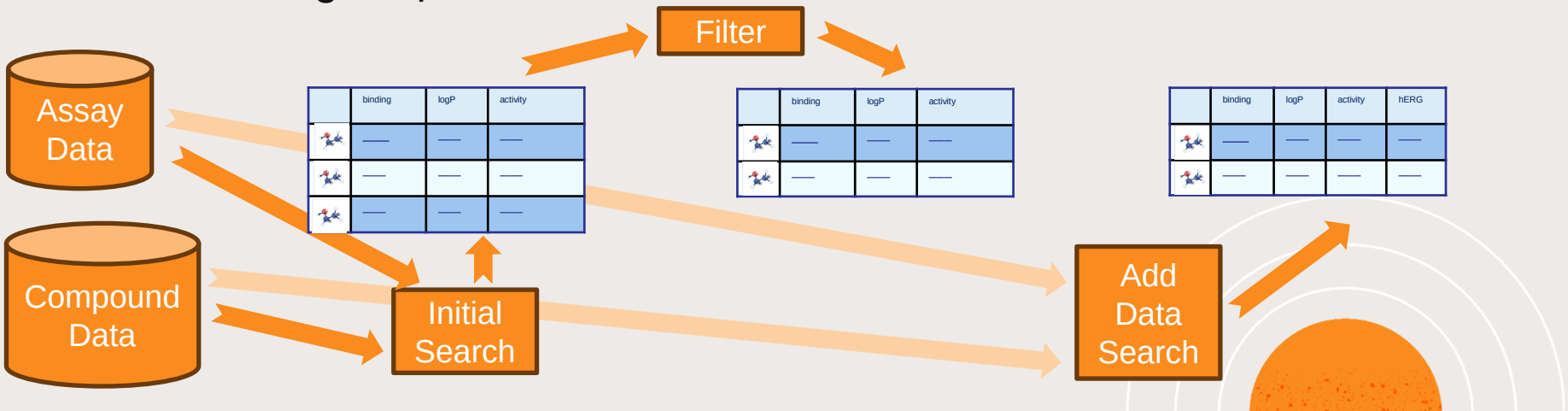
Small molecule drug discovery examples

- **Project leader goals**
  - Track program execution
  - Prioritize compounds to synthesize and test
  - Determine if compounds meet requirements to clear a project gate
- **Project team member**
  - Select the best datasets for making project decisions
  - Reach a shared understanding of current project information
  - Understand which aspects of a compound's performance as a drug must be improved to meet project goals
  - Collaboratively recommend compound synthesis and assay priorities
- **Medicinal chemist goals**
  - Understand how compound structure impacts all aspects of its performance as a drug
  - Understand if specific batches of the same parent structure behave differently
  - Propose new compounds that improve specific qualities or overall performance
- **Pharmacologist goals**
  - Understand how compound structure affects pharmacology
  - Understand if specific salt forms of the same parent structure behave differently
  - Verify data are properly curated and presented to the project team with sufficient context to correctly interpret it
- **Computational chemist goals…**

Each goal may require specialized software tools and data drill-down

Search vs Filter

- **Initial search requirements are not the same as filtering**
  - But they often overlap – e.g. search by substructure then filter by substructure
- Commonly missed requirement: Filter by assay condition for late-stage experiments

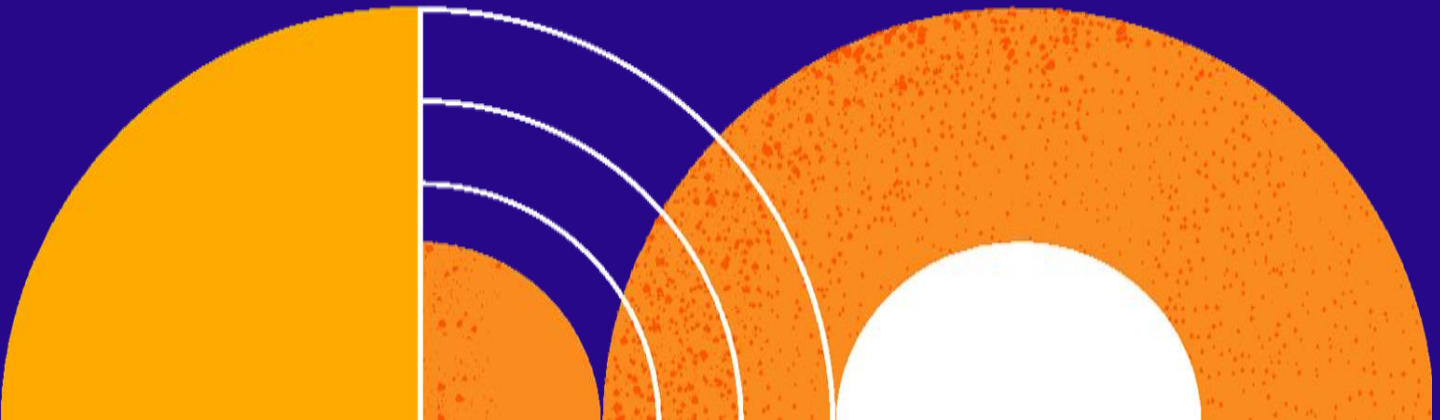


UGM Boston Oct. 4th 23

Chemaxon

Derived Requirements

Commonly missed requirements



Row Aggregation

“Average” results for the “same” assay and compound

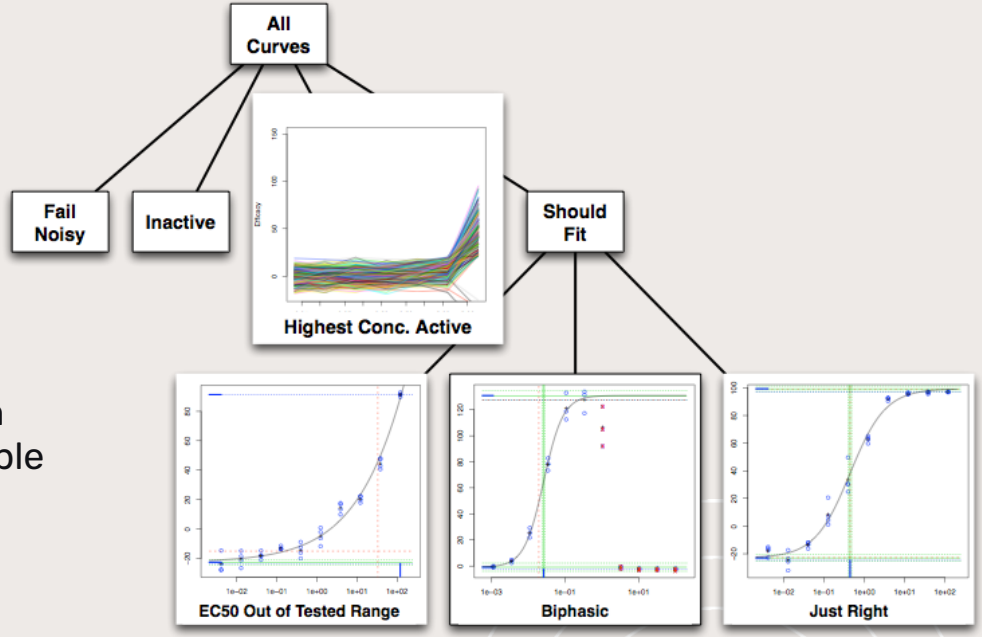
- Same “Compound” depends on the purpose
  - Same structure averaged for multiple lots is common, but sometimes you care about lot variability
  - Common aggregations use-cases
    - By lot
    - By parent
    - By salt
  - Less common use cases
    - By formulation
    - By assay condition
- “Averaging” must be mathematically correct for each endpoint type
  - Different data types require different operations, log avg, avg, etc.
  - Only average comparable* assays

*This topic could be a whole presentation

Dose-Response Curves

In early research, clean sigmoids are the exception

- Many of these curve categories fit results require ‘<‘ or ‘>’ in the EC50
  - The tools must sort numerically with operators
- Plots must be rendered for scientific interpretation
- Aggregated rows require overlaid curves
- Curves must be rendered in 30-40ms each to have good scrolling performance of a table with a few thousand rows of data and 3-4 curve data columns
  - Because of overlays, not all curves can be pre-rendered or cached



Assay and Endpoint Aliasing

Consensus is not always possible

Problems

- Assay names vary by CRO, lab, and scientist
- The same assay may be a primary-screen for one project but a counter-screen for another
- Endpoint names may not be standardized
- There is always project history and team needs evolve
- Some team members want the assay type in the name, while others need assay conditions

Solutions

- Allow assay name aliases
  - The name shown in report is not the same as assay name in the database
  - Ideally, alias by project
- Same for endpoint names
- Configurable assay trees
  - Ideally multiple user-selectable trees
    - by condition or by assay type

Collaborative Requirements

- Access controls
  - By project, by lot, & by assay
- Shared project team reports
  - Repositories of institutional knowledge
    - e.g. Just because data has the project target, does not mean the project team currently uses it for decision making
  - Evolve as the program progresses
    - e.g. Initially solubility is not important, but it will become important
  - Templates
    - Read, edit, clone

Team A Visible

	Lot Project	logP	Target A Activity	Target B Activity
	A	—	—	—
	A	—	—	—
	All	—	—	—
	B	—	—	—
	B	—	—	—

Team B Visible

	Lot Project	logP	Target A Activity	Target B Activity
	A	—	—	—
	A	—	—	—
	All	—	—	—
	B	—	—	—
	B	—	—	—

“And you may ask yourself, well, how did I get here?”*

Understanding the context of past decisions

- How do we learn from prior projects?
- What if we could restore our collaborative decision space to the exact state as when we discontinued a structure series, or approved a candidate?
  - The same sorting, filtering, analysis tools, but on historic data

*From “Once in a Lifetime” by Talking Heads and Brian Eno

Summary

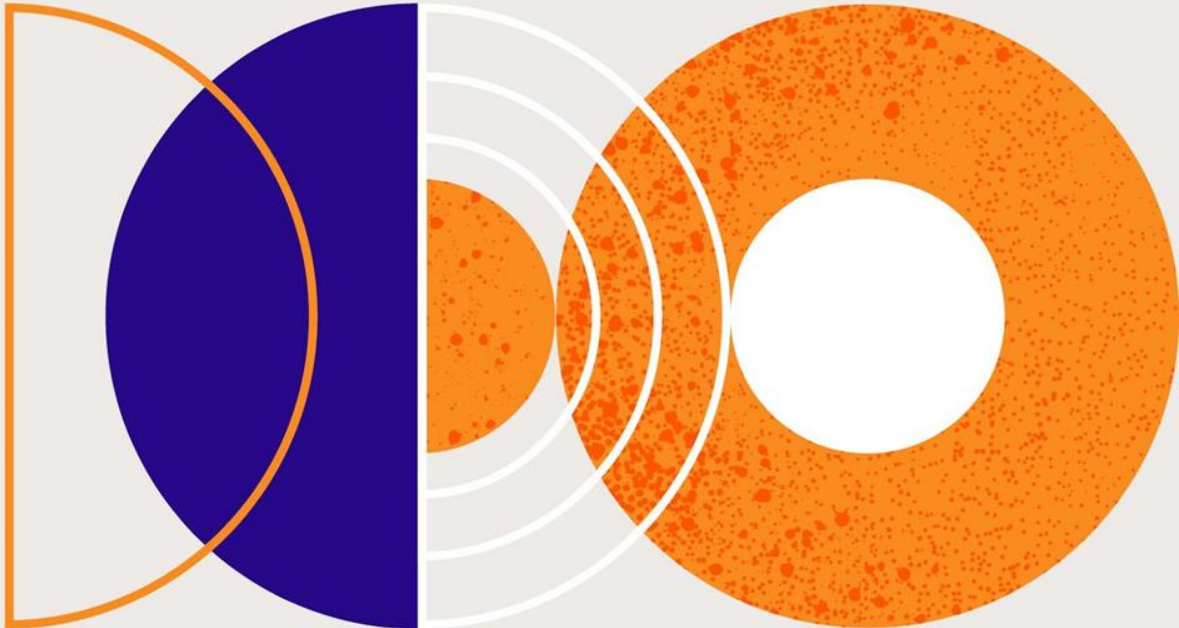
Intelligence Amplification

- Our job as informatics providers is to deliver tools that
  - Stuff as much information as possible into project team member's big, wet, neural nets
  - Reduce distracting or confusing information
  - Eliminate mental gymnastics to remember caveats and context
  - Support a project from inception to clinic
  - **Help project teams make excellent decisions and brilliant leaps of inductive logic**

UGM Boston Oct. 4th 23

Chemaxon

Thank you

John A. McNeil

john@jamcneil.com