GSRS Development Update

16 Nov 2018

TYLER PERYEA
DANNY KATZEL
NIH/NCATS

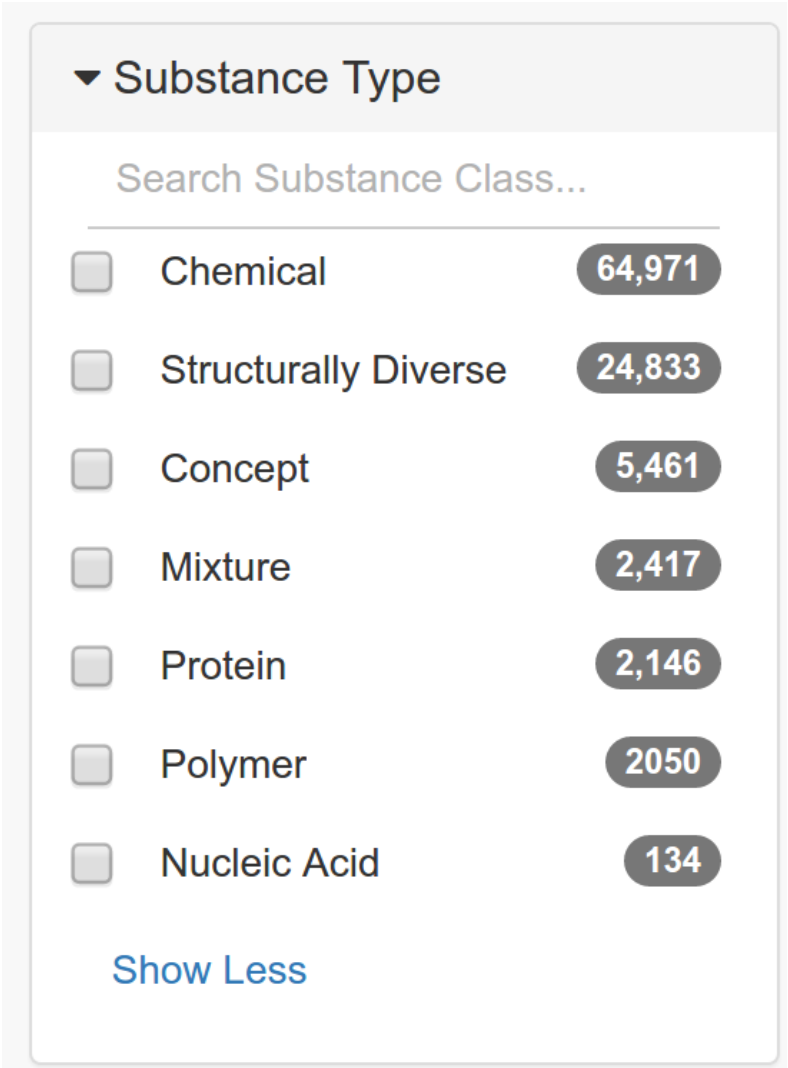
NCATS

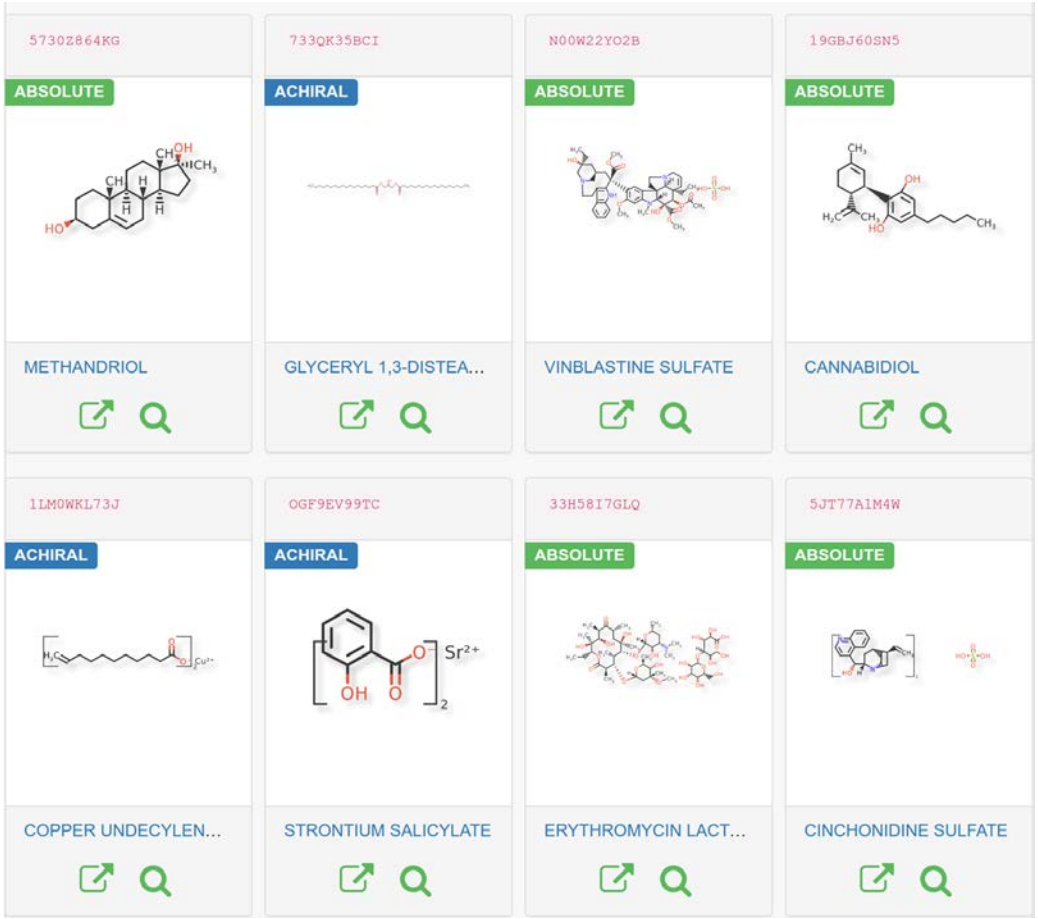
GSRS Status

High Level GSRS Core Status

- GSRS 2.3.1 in production at FDA (October 2018)
  - 2.3.1 public released
- GSRS 2.3.2 development (November 2018)
  - Pre-production at FDA
  - Pending public release soon (December 2018)
- GSRS 2.4 development (1st quarter 2019)
  - In development
  - Improved data exchange / communication

GSRS Status



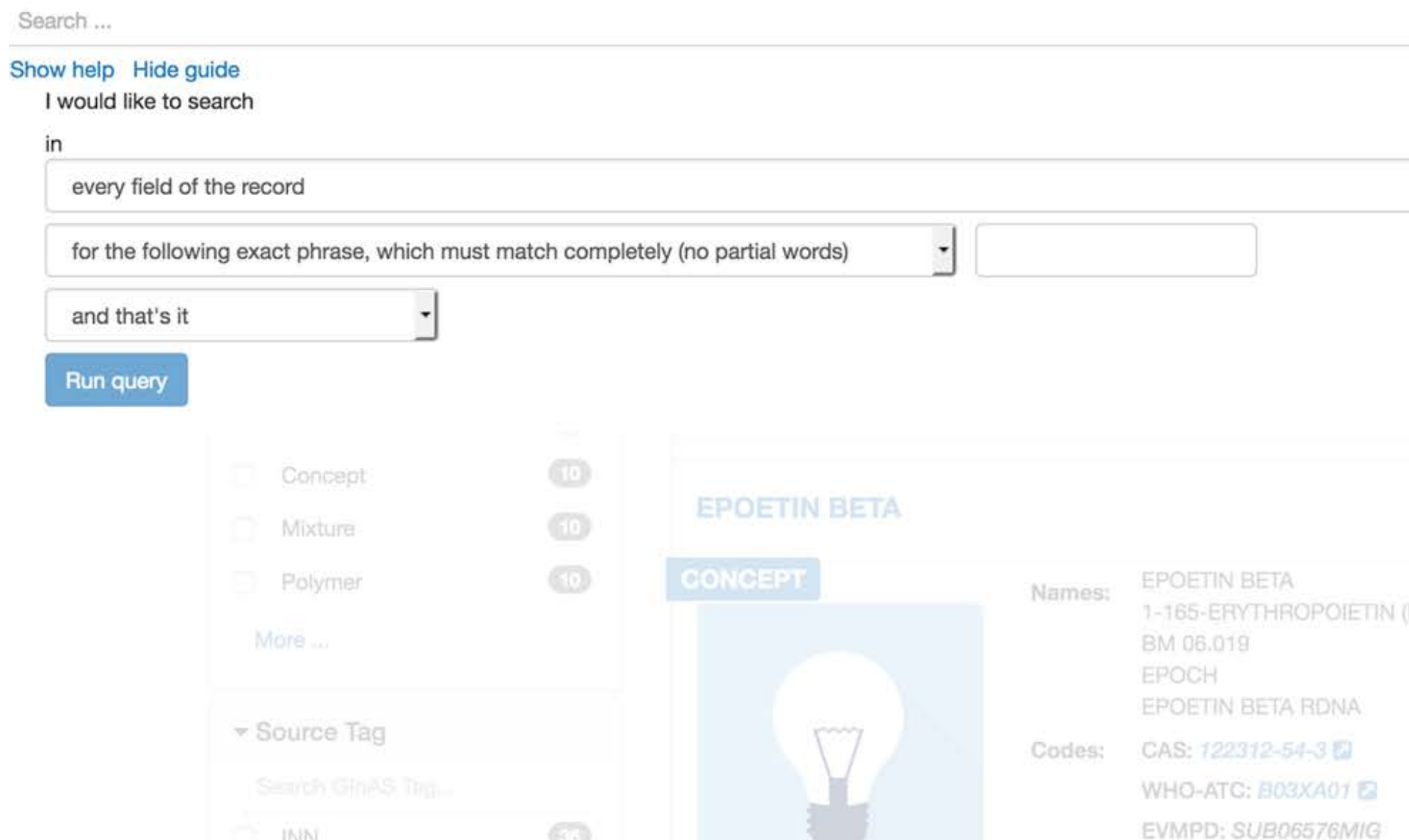


National Center for Advancing Translational Sciences

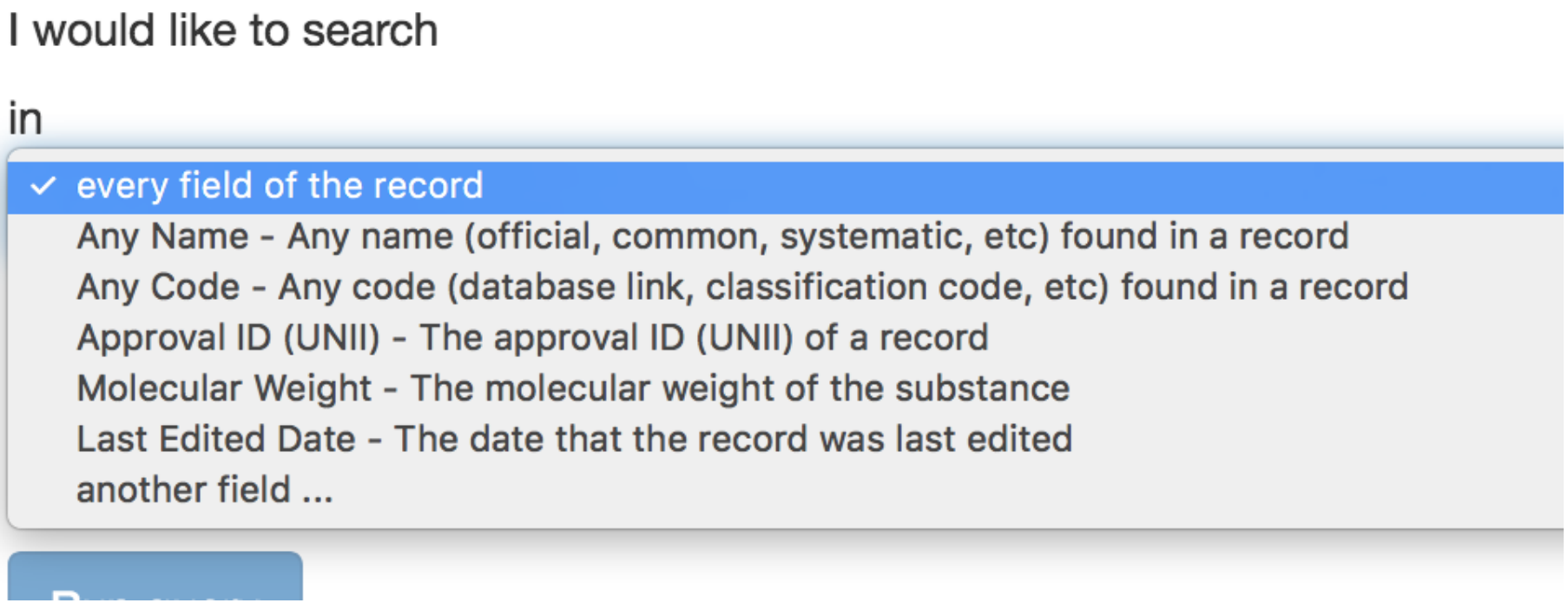
GSRS 2.3.1 Status

- **GSRS 2.3.1 Improvements**
  - Substance “Hierarchy” views
  - Additional Validation Rules
  - Improved Relationship Handling
  - Improved Structure Resolver
  - Configurable / modular Validation Engine
  - Porting of some FDA-specific enhancements to main code base
  - Bug fixes
  - Improved UI
  - New Guided Search

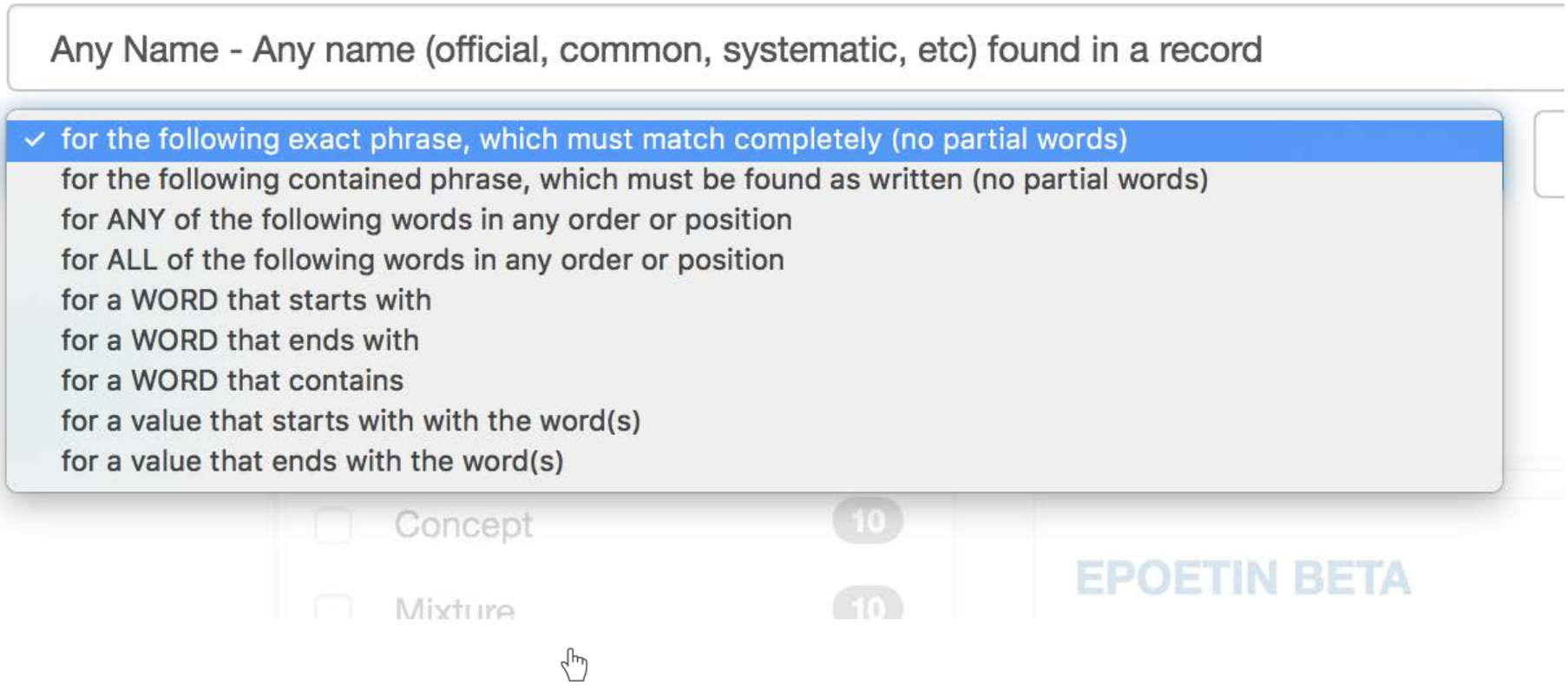
GSRS 2.3.1 New Search Guide



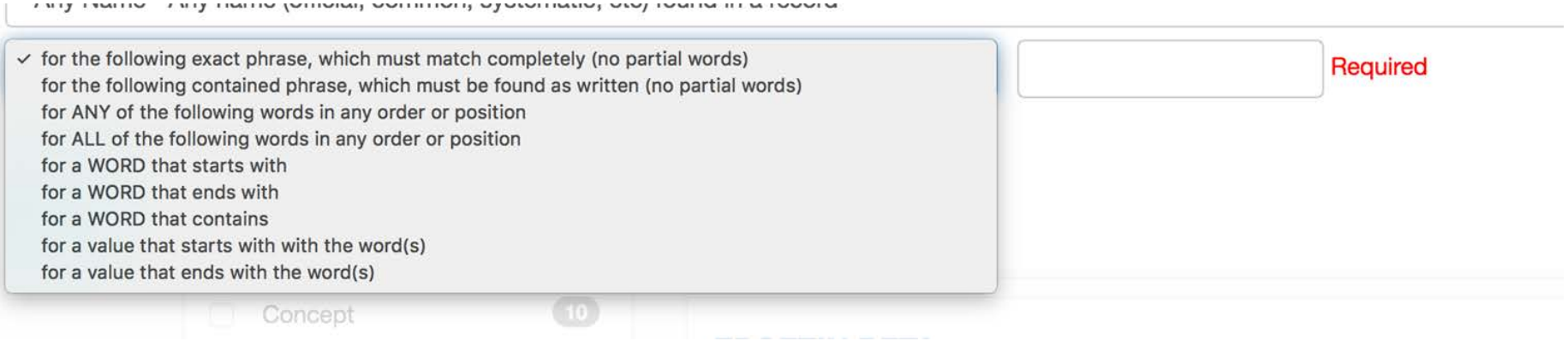
GSRS 2.3.1 New Search Guide



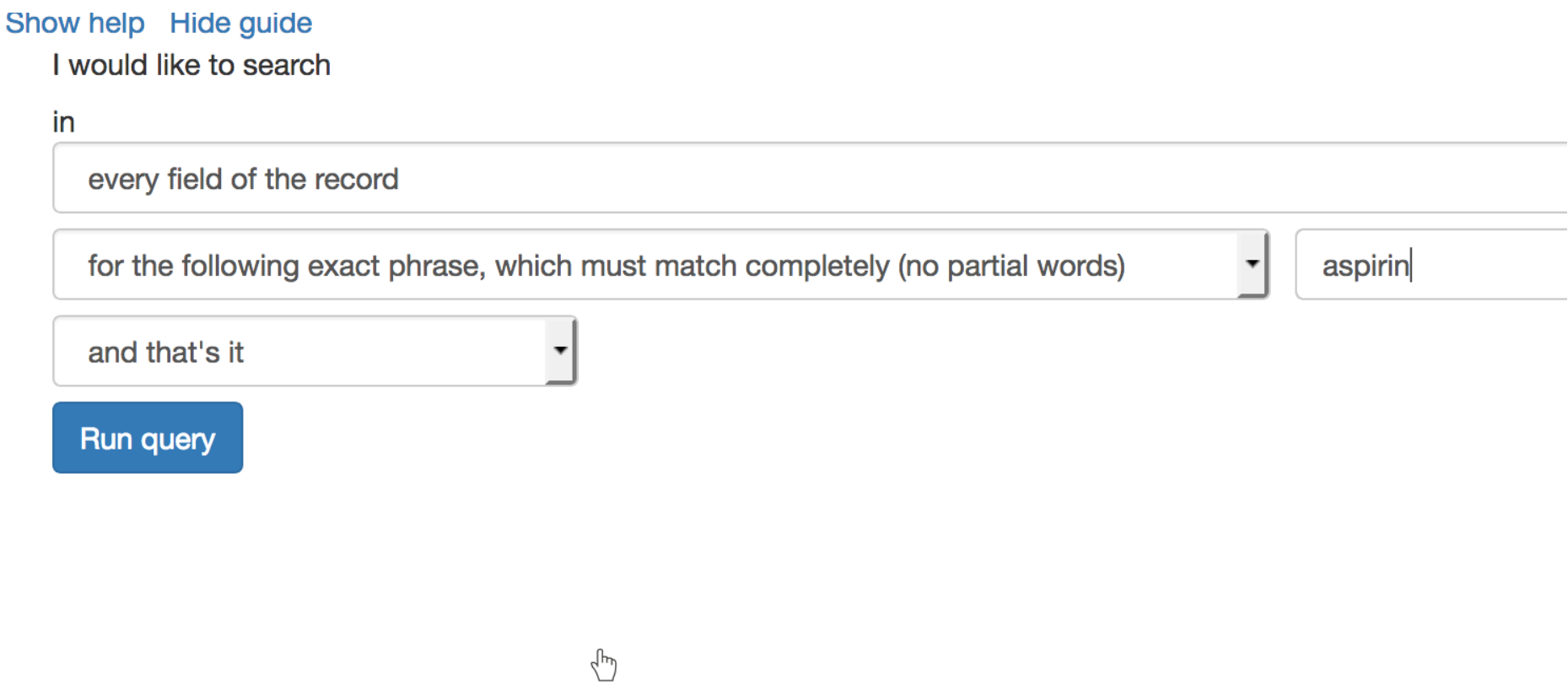
GSRS 2.3.1 New Search Guide



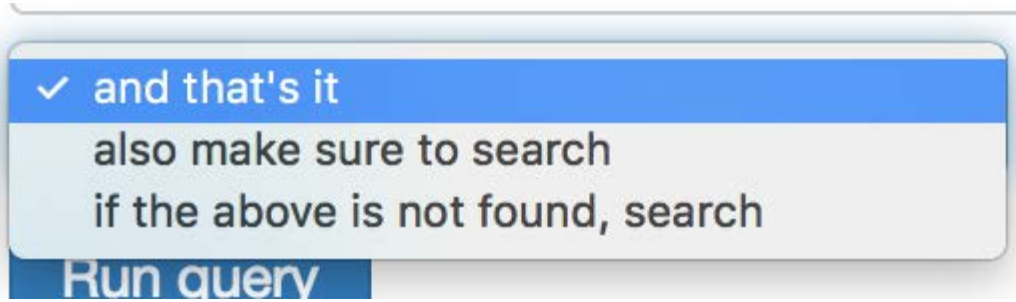
GSRS 2.3.1 New Search Guide

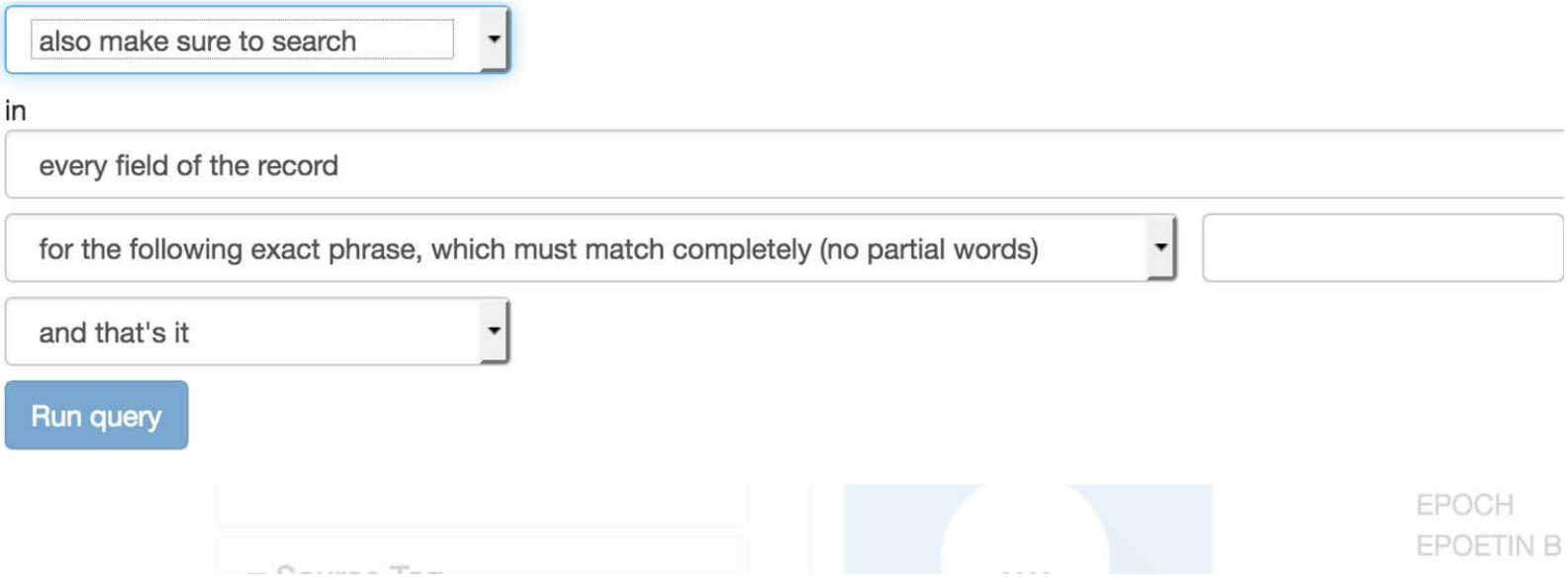


GSRS 2.3.1 New Search Guide



GSRS 2.3.1 New Search Guide





GSRS 2.3.1 New Search Guide

root_names_name:"^aspirin$" OR "^foo$"

Show help Hide guide

I would like to search

in

Any Name - Any name (official, common, systematic, etc) found in a record

for the following exact phrase, which must match completely (no partial words) | aspirin

if the above is not found, search

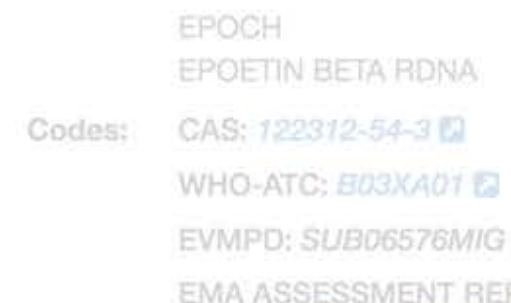
in

every field of the record

for the following exact phrase, which must match completely (no partial words) | foo

and that's it

Run query

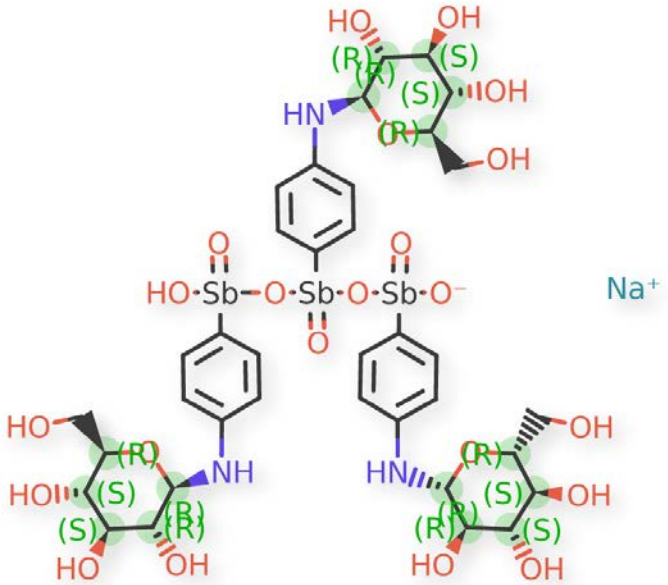


GSRS 2.3.2 Status

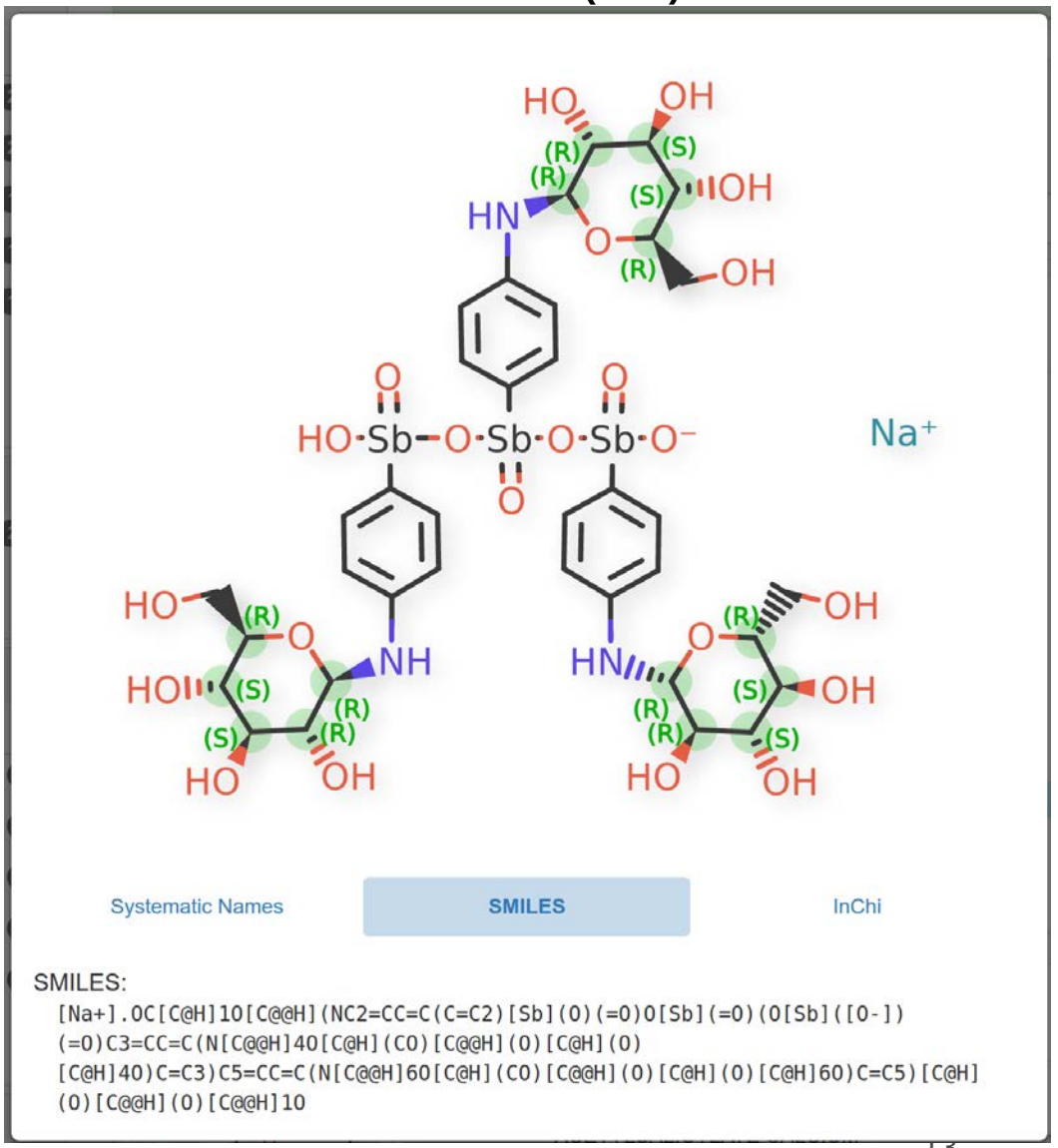
- **GSRS 2.3.2 Improvements**
  - Improvements to rendering of molecules
  - Further improvements to guided search
  - Further improvements to duplicate check / sequence searching
  - Ability to easily revert records for authorized users
  - Additional stability and data integrity checks
  - Improvements to hierarchy view

GSRS 2.3.2: Rendering

GSRS 2.3.1 (old)



GSRS 2.3.2 (new)



GSRS 2.3.2: Guided Searches

GSRS 2.3.1 (old)

(a little opaque and inaccurate)

Searched for:
root_structure_mwt *: [100 TO 10E50] AND aspirin (contains)*

GSRS 2.3.2 (new)

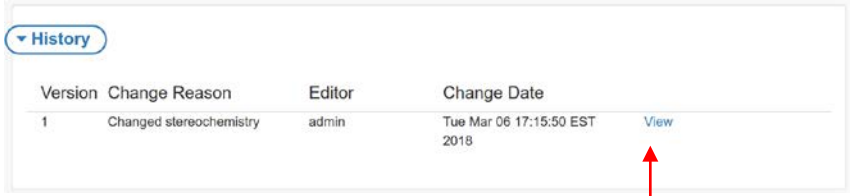
(more informative and accurate)

You searched *for the* ***Molecular Weight*** *field >* ***100*** *AND for any field = ...****aspirin****...* *show long explanation*

You searched *in the* ***Molecular Weight*** *field, for a value or values that must greater than* ***100****. Also making sure to search in any field, for a value or values that contains* ***aspirin*** *show short explanation*

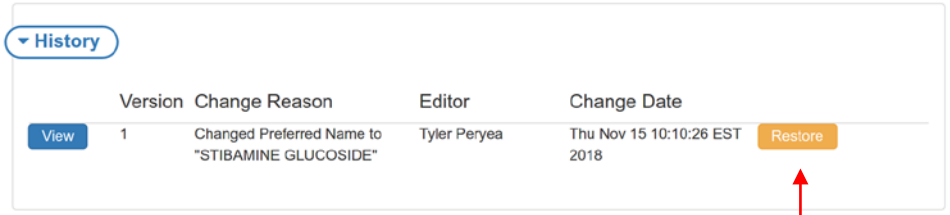
GSRS 2.3.2: Restoring Old Versions

GSRS 2.3.1 (old)



Super users can **view** old versions, but involved process to revert

GSRS 2.3.2 (new)



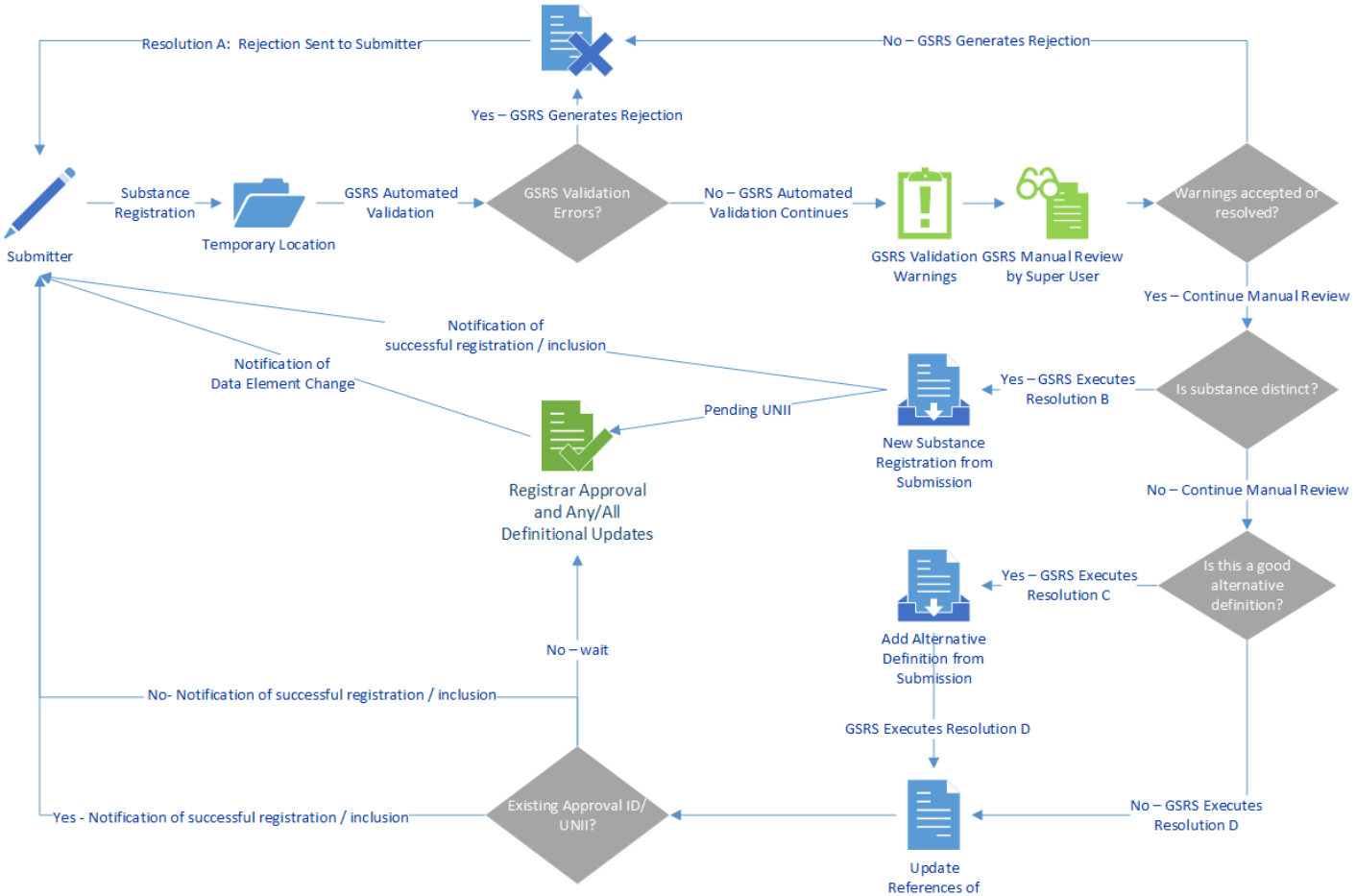
Super users can quickly **revert** records, after review

GSRS Status

On the horizon

- GSRS 2.4
  - Submissions / System-to-system communication
    - Establishment of requirements document *(complete)*
    - Basic support for accepting / staging / processing new and modified records from external sources *(1st quarter 2019)*
  - Simplified installation / launch
    - High-level feasibility report *(complete)*
    - Basic support *(1st quarter 2019)*

System Submission and Exchange Mechanism



- Green indicates existing functions/processes
- Blue indicates update required to GSRS (primarily to automate, manual processes already supported)

GSRS Status

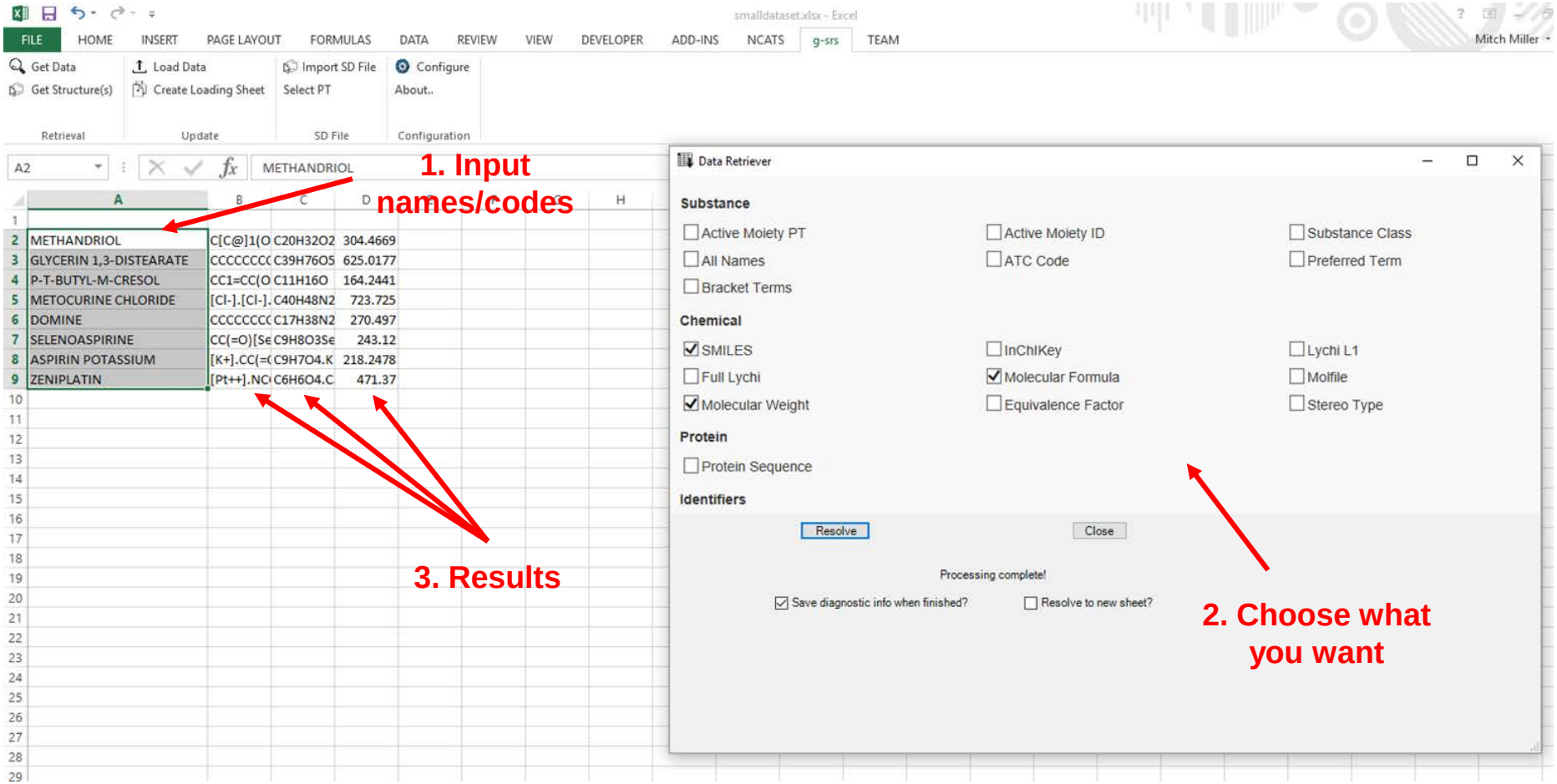
On the horizon

- GSRS 2.5+
  - Improvements to REST API / Systematic access
  - Upgrade of UI to modern web framework for client-side UI
    - Angular 6+
    - Websocket support

GSRS Tools

- **GSRS Find excel tool** *(currently in alpha release)*
  - Search, create, update, retrieve records and substance information in batches
  - Support for loading SD files
- **Stand-alone Java API library** *(1st quarter 2019)*
  - Easier programmatic access to data for integration with other applications / stand-alone tools
- **Stand-alone Javascript API library** (2nd quarter 2019)
  - Provides simple communication and data operations with running instance

GSRS Tools: GSRS Find



National Center for Advancing Translational Sciences

https://tripod.nih.gov/ginas/

GSRS Tools: Stand-alone Java API Library

```
try(GsrsRestClient client = DefaultGsrsRestClient.create("https://ginas.ncats.nih.gov/ginas/app")){

        GsrsSearchResult result = client.searcher().substructure("C1=CC2=CC=CC=C2C=C1");

                System.out.println("search results :");
                System.out.println("total : " + result.getTotal());
                System.out.println("top 10 substance names:");

                result.getUuids()
            .stream()
                        .limit(10)
                        .forEach(uuid->{
            try {
                Substance sub = client.fetchSubstance(uuid);

                System.out.println(sub.getName());
            }catch(IOException e) {}});


}
```

```
search results :
total : 1973
top 10 substance names:
NAPHTHALENE
POLYCYCLIC AROMATIC
HYDROCARBONS
PHENANTHRENE
ANTHRACENE
BENZ(A)ANTHRACENE
TRIPHENYLENE
CHRYSENE
BENZO(C)PHENANTHRENE
TETRACENE
DIBENZ(A,H)ANTHRACENE
```

National Center for Advancing Translational Sciences

https://tripod.nih.gov/ginas/

Acknowledgements

NCATS

Tyler Peryea
Noel Southall
Dac-Trung Nguyen
Ivan Grishagin
Dammika Amugoda
Jorge Neyra
Niko Anderson
Mark Williams
Tim Sheils
Chris LeClair
Paul Shinn
Qian Zhu
Christine Colvis
Mark Williams

FDA

Larry Callahan
Frank Switzer
Yulia Borodina
Ramez Ghazzaoui
Elaine Johansen
Ta-Jen Chen
Archana Newatia
Mary-Ann Slack
Alex Welsch
Sarah Stemann
Yoshiyuki Tokiwa

Medicines Evaluation Board

Herman Diederik
Marcel Hoefnagel
Joris Kampmeijer
Ciska Matai
Burt Kroes

US Pharmacopeial Convention

Fouad Atouf
Andrej Wilk
Tina Morris

Questions