Inxight Drugs

11 OCTOBER 2017

TYLER PERYEA
NIH/NCATS

NCATS

Introduction

Problem

- How many drugs are there?
- Drug development information is proprietary or disjointed.
- Need a systematized, comprehensive, all-encompassing resource.

Aims

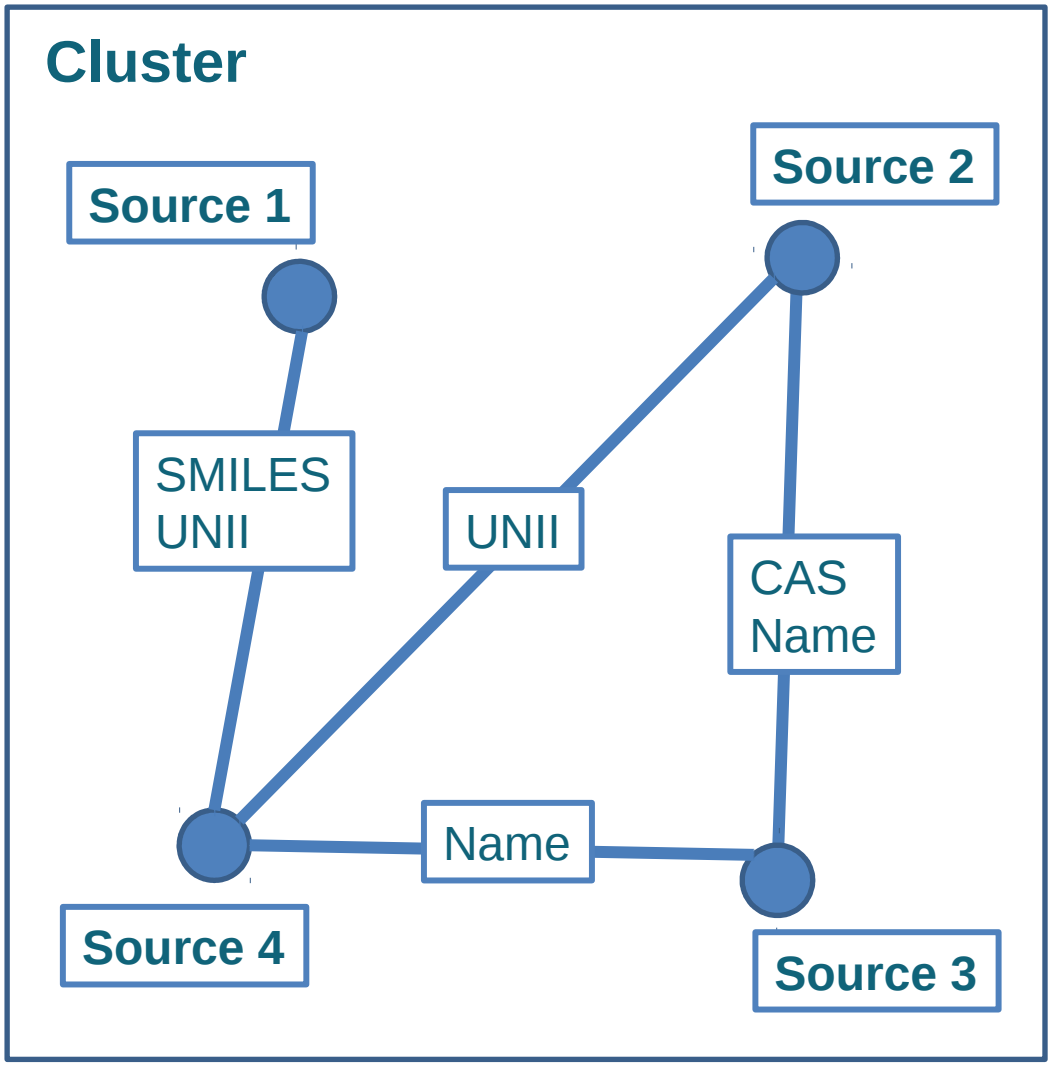
- Provide definitive answers to all “how many” questions.
- Show what drugs are coming ‘down the pipe’ and for which diseases.
- Provide a resource for the translational research community for drug repositioning and repurposing.

Overview

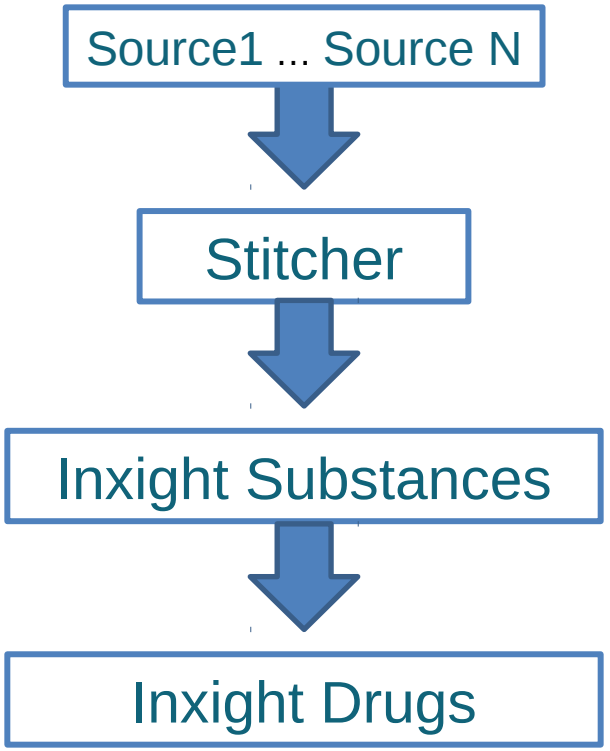
Inxight Drugs – Comprehensive Collection of Drugs and Substances

- 5 sources (51 in pipeline)
- 2,741 US Approved Drugs
- 7,298 Entered Clinical Testing
- 99,922 Substances
- 6,684 manually annotated compounds
- Based on NCATS GSRS
- “Stitcher” – NCATS data unification and de-duplication software

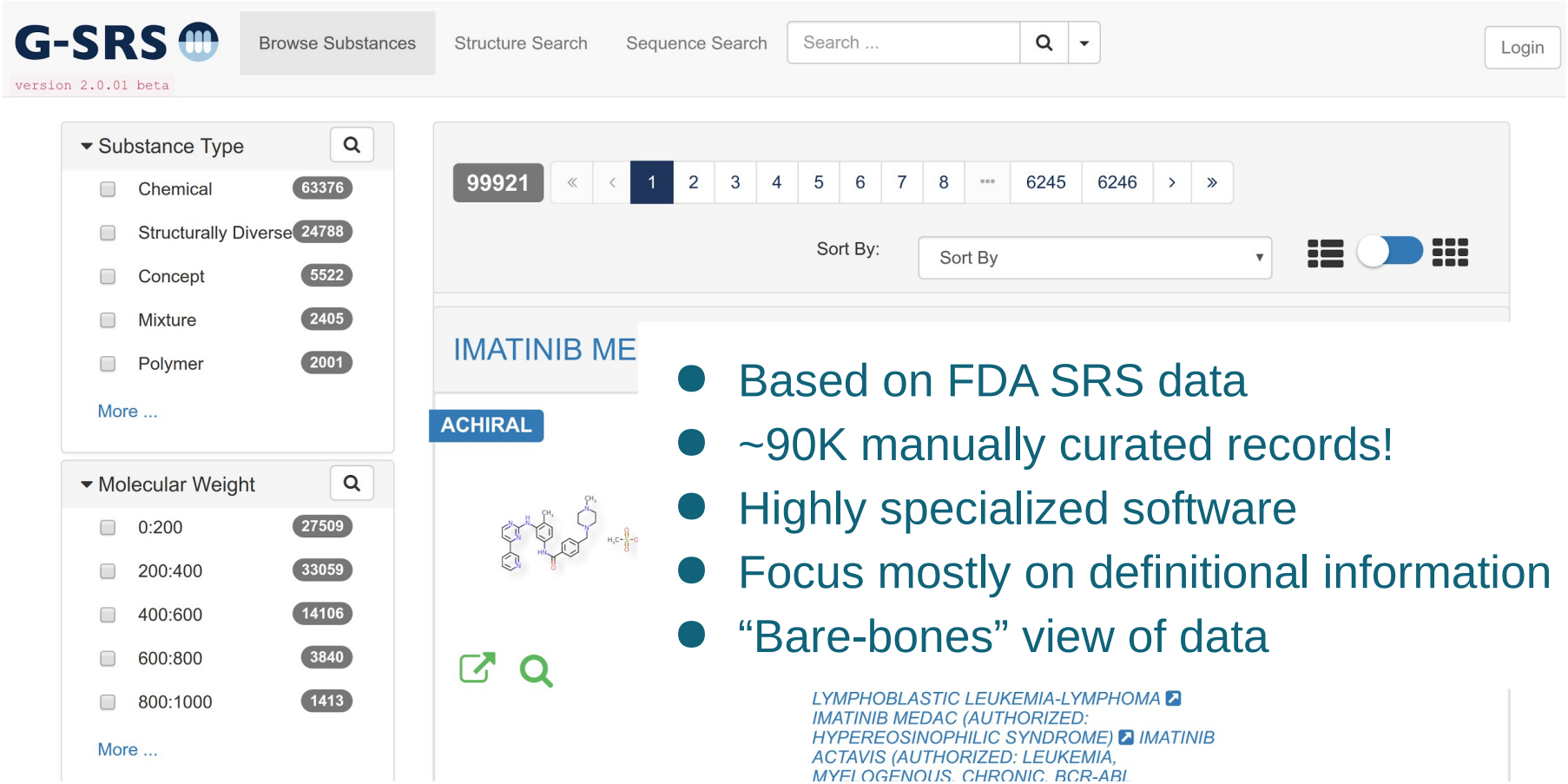
NCATS Stitcher



- 1 Cluster = 1 Substance
- Canonical Properties
- Parent / Children relationships



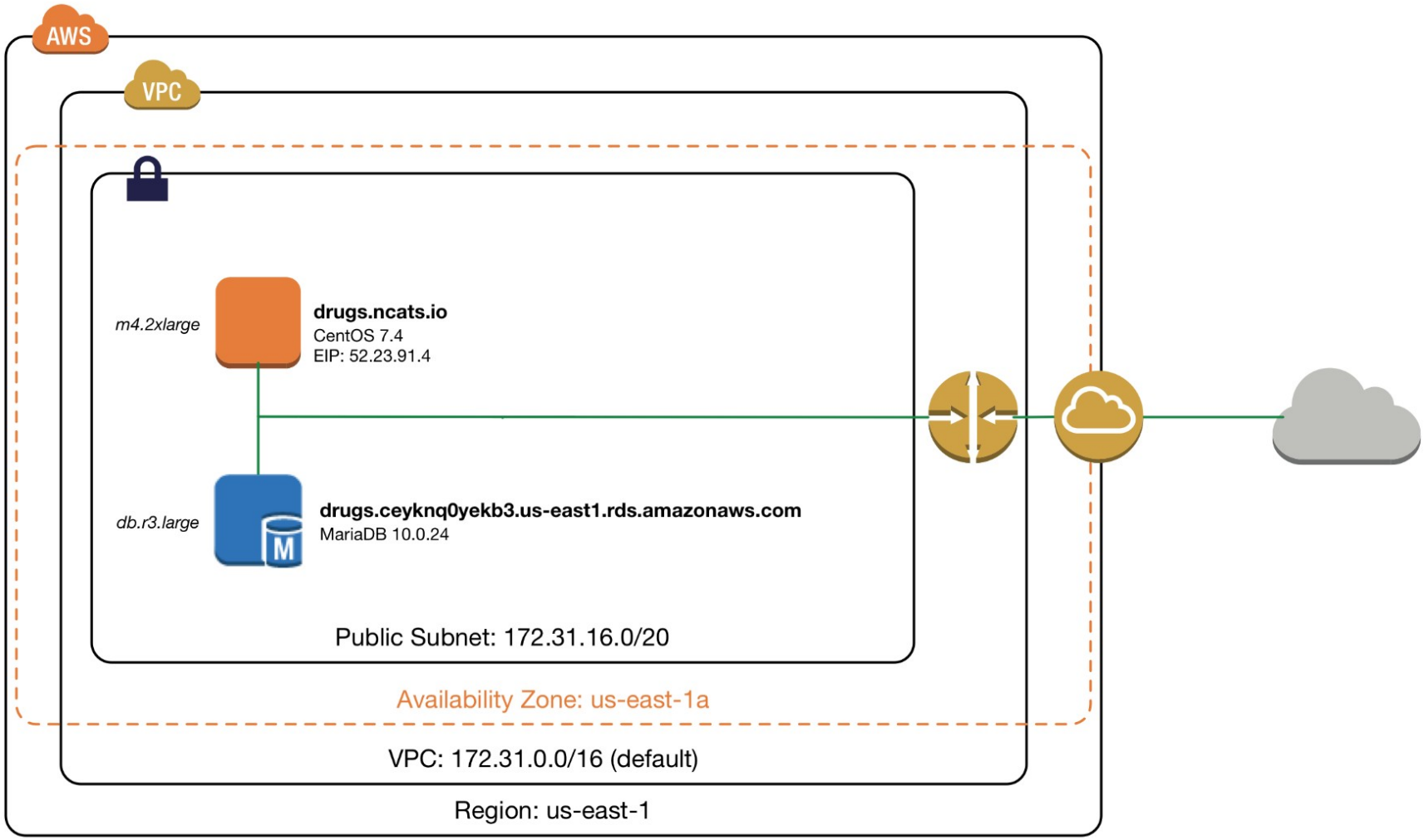
NCATS GINAS/GSRS



- Based on FDA SRS data
- ~90K manually curated records!
- Highly specialized software
- Focus mostly on definitional information
- “Bare-bones” view of data

National Center for Advancing Translational Sciences

NCATS Drugs (Production) in AWS
drugs.ncats.io
October 2017



Inxight Drugs

U.S. Department of Health & Human Services › National Institutes of Health › NCATS

NIH National Center for Advancing Translational Sciences

Inxight Drugs

Search Drugs ...

The National Center for Advancing Translational Sciences has developed Inxight Drugs — a comprehensive portal for drug development information.

Inxight incoporates and unifies a wealth of data, including marketing and regulatory status, rigorous drug ingredient definitions, biological activity, clinical use and more.

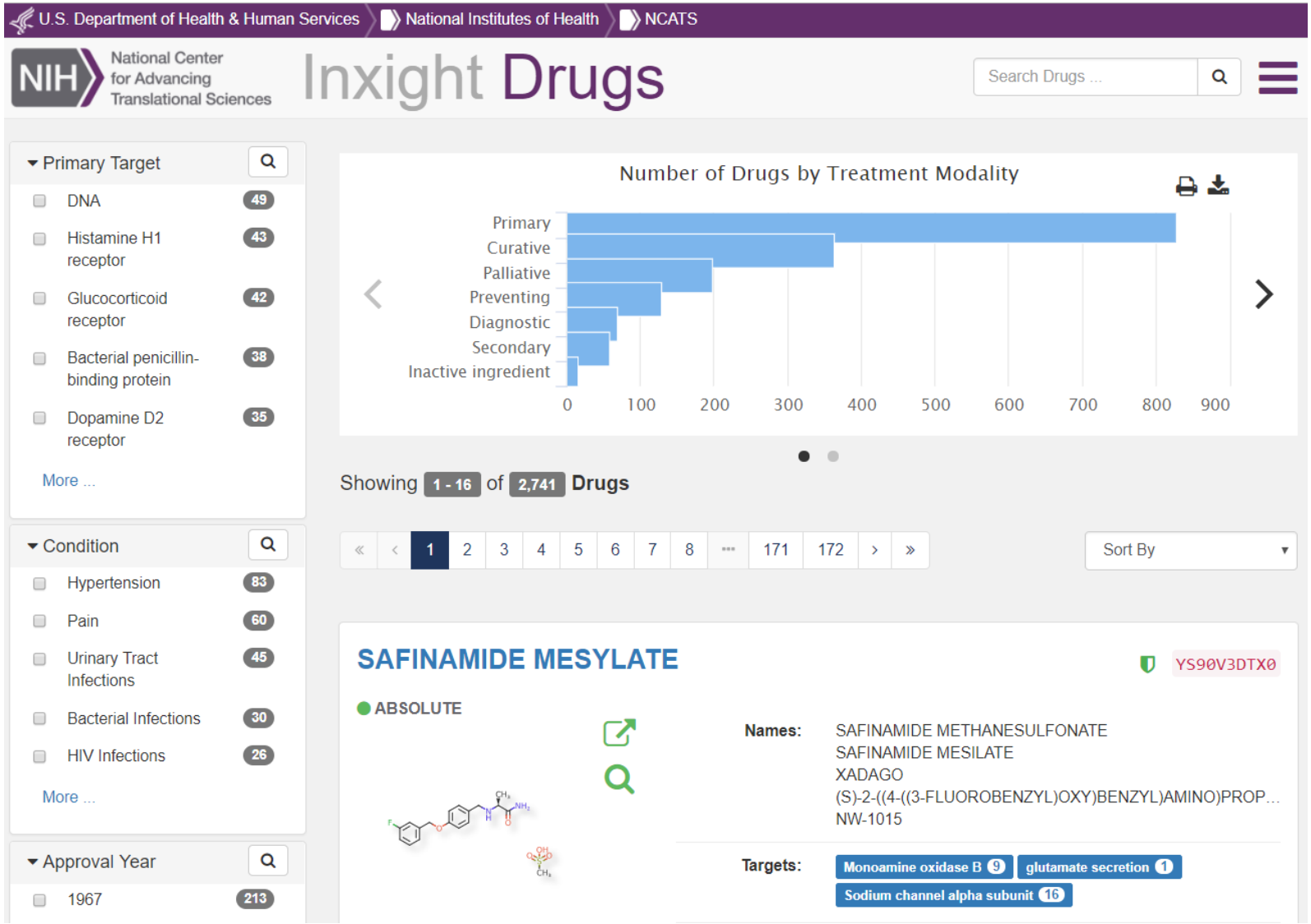
Inxight incorporates manually curated data provided by the FDA and private companies.

Currently, Inxight contains information on:

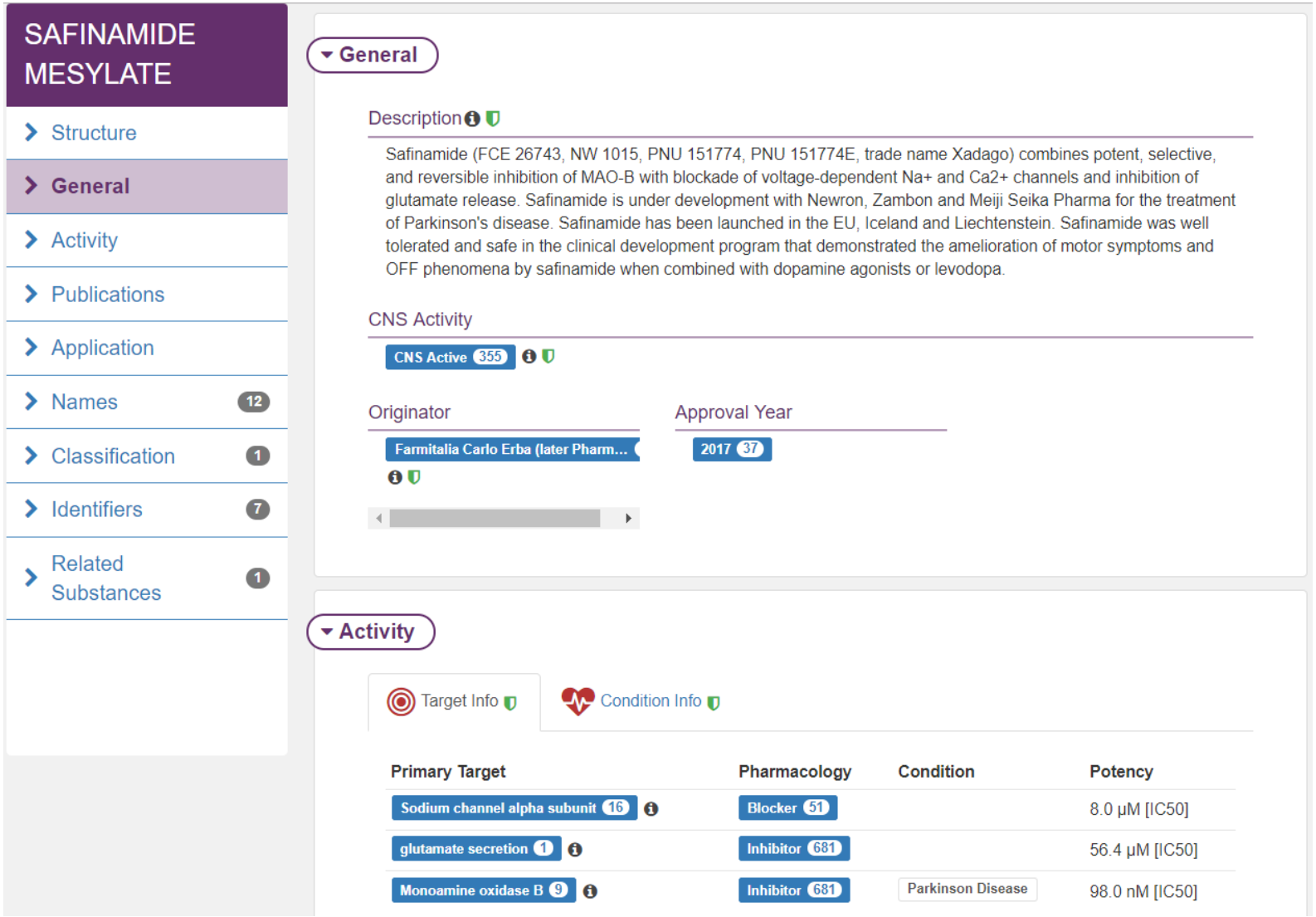
- 2,741 APPROVED DRUGS
- 3,956 MARKETED DRUGS
- 7,347 ENTERED CLINICAL TESTING
- 99,922 ALL SUBSTANCES

- Intended for wide audience
- Built on ISO 11238 core
- First-in-class data aggregation
- Manually curated data

Drugs List and Filters

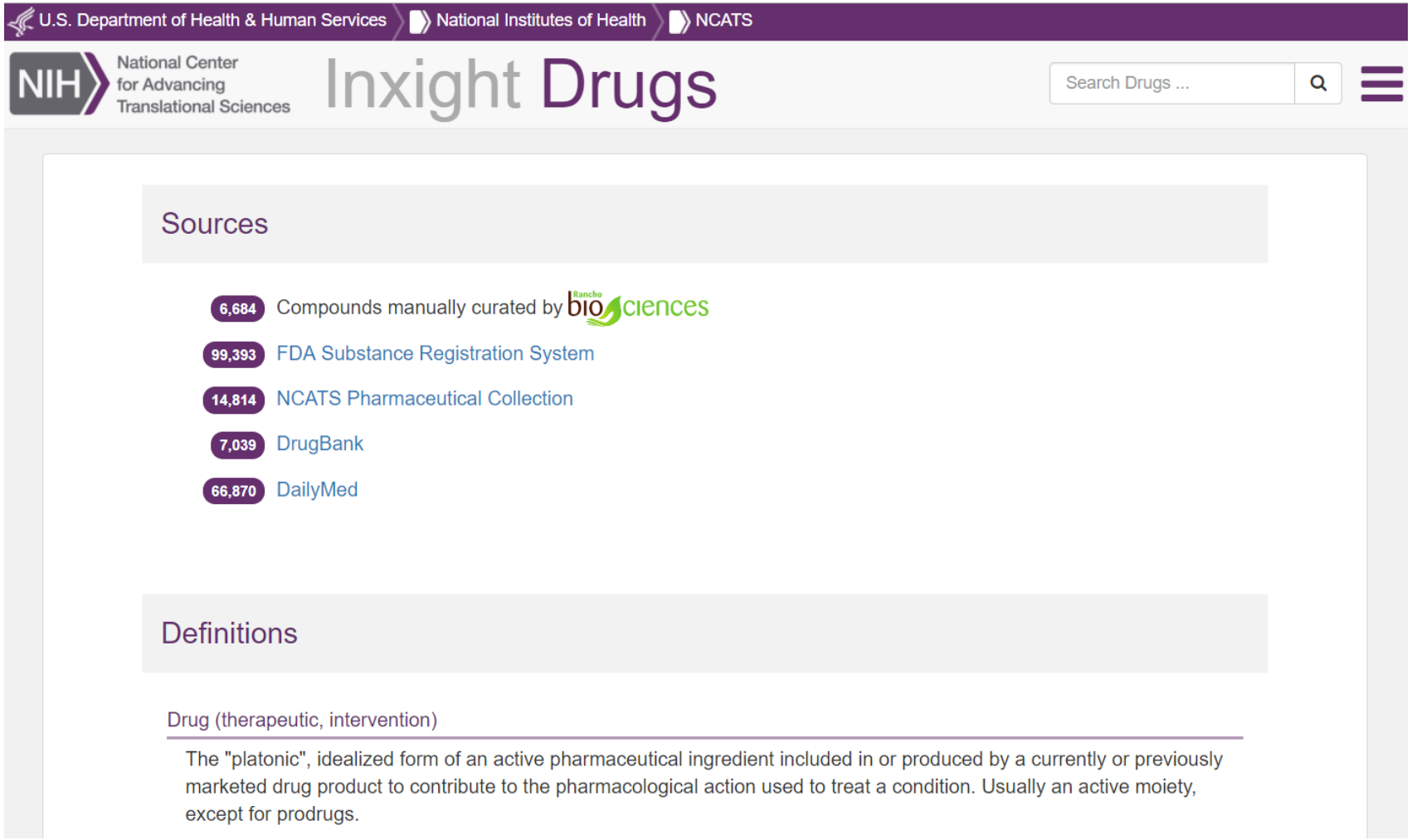


Drug Record Details



National Center for Advancing Translational Sciences

Sources and Definitions



DEMO

Summary

- **Content and Purpose**
  Complete drug development profile:
  from structural details to drug products.
- **Core Components**
  Precise scientific definition for the all the ingredients in medical products, as per the rigorous ISO 11238 definition.
- **Data Quality**
  Enormous amount of manually curated data: 90K entries from FDA SRS and 7K annotations from Rancho BioSciences
- **Data Aggregation**
  First-in-class automated facile data aggregation from multiple sources via Stitcher (5 implemented, 51 in the pipeline)

https://drugs.ncats.io