GOSTAR

Global Online Structure Activity Relationship Database

The extensive drug development programs of Pharma and Biotech companies comprising, early drug discovery, translational research, clinical trials and post-marketing studies, culminate with tremendous data generation. These vast datasets are increasingly being used to power AI/ML based drug discovery that enables rapid and informed decision making, thus saving the significant time and costs of traditional drug discovery.

Consequently, structured and high-quality datasets covering structure activity relationships (SAR), physicochemical properties, and ADMET are essential for predictive modelling and AI/ML driven drug discovery programs.

GOSTAR

Best-in-class SAR knowledgebase with analysis-ready datasets

Excelra's **Global Online Structure Activity Relationship Database (GOSTAR)** is the largest repository for structured SAR content. Available as a one-stop data source, GOSTAR allows researchers to thoroughly explore the known chemical space around a target of interest. The database is manually curated by our scientific teams who excerpt and enrich datasets from functional assays, *in vitro* and *in vivo* studies. A variety of small molecule activities encompassing SAR, physicochemical, metabolic, ADME and toxicological profiles are captured in GOSTAR.

GOSTAR for *In silico* Drug Discovery

Powering an array of AI/ML based applications

Target validation

GOSTAR contains large datasets expressing quantitative binding interactions of targets with known compounds, experimentally determined using *in vitro* and *in vivo* (enzymatic and cell-based models) techniques. Using machine-learning techniques, AI/ML modellers can generate druggability models and predict unknown mechanisms involved in either activation or inhibition of the targets of interest.

- Datasets to identify assays for validating drug targets
- Datasets around chemical probes (Reference compounds)
- Datasets associated with target mechanism of action

Hit identification

GOSTAR has a repository of millions of compounds, whose chemical structure is linked to biological, pharmacological and therapeutic information. AI/ML modellers can use this information to build and validate virtual screening models to discover hits.

- Content for virtual screening (Structure and analog-based approaches)
- Known and diverse chemical space around target of interest
- Novelty evaluation using vast SAR space
- Readily accessible fragment and scaffold libraries

Lead optimization

SAR, ADMET and physicochemical datasets within GOSTAR can be used to perform virtual lead optimization studies. ADMET datasets especially are useful in building and validating predictive models that can be used to optimize pharmacokinetic properties of lead compounds.

Focused SAR datasets to build models to optimize affinity/potency of leads

Novelty evaluation and expansion of the lead series

In vivo (disease models) datasets to build safety and efficacy models

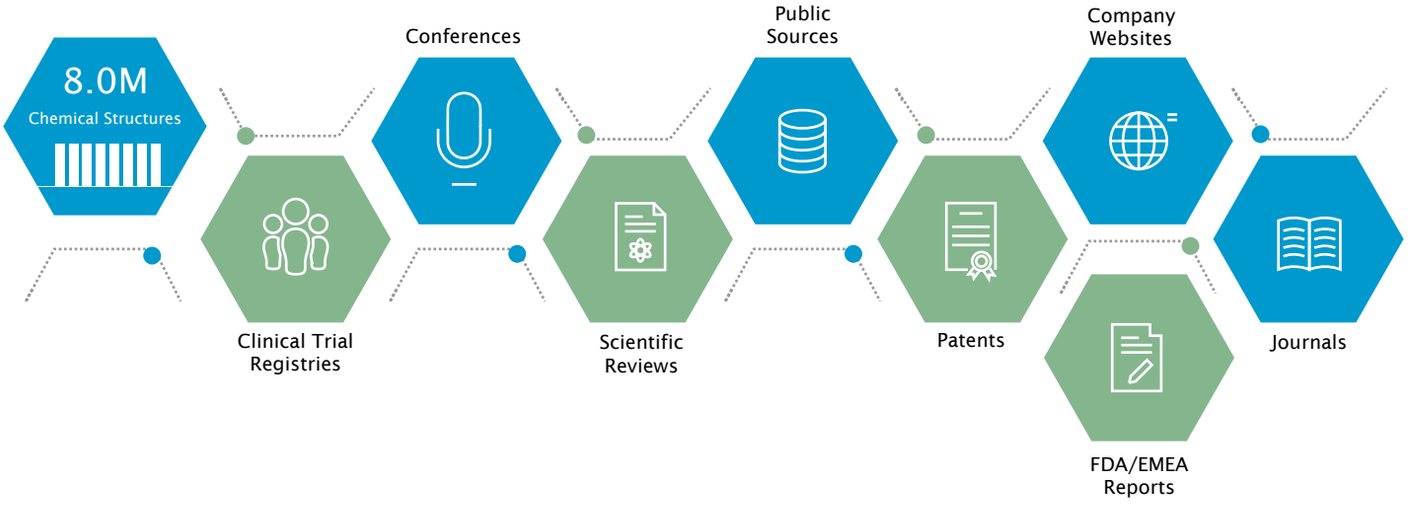
Drug repurposing

GOSTAR supports AI/ML modelers in exploring new therapeutic uses for already approved drugs or clinically studied compounds. Data scientists can mine vast SAR content represented by quantitative interactions of preclinical compounds, clinical candidates and approved drugs with known target space for novel relationships. The systematically captured polypharmacology content provides a data rich platform to explore novel associations and also to investigate safety and effectiveness of repurposed candidates.

GOSTAR Data Sources

Creating a knowledgebase from diverse sources

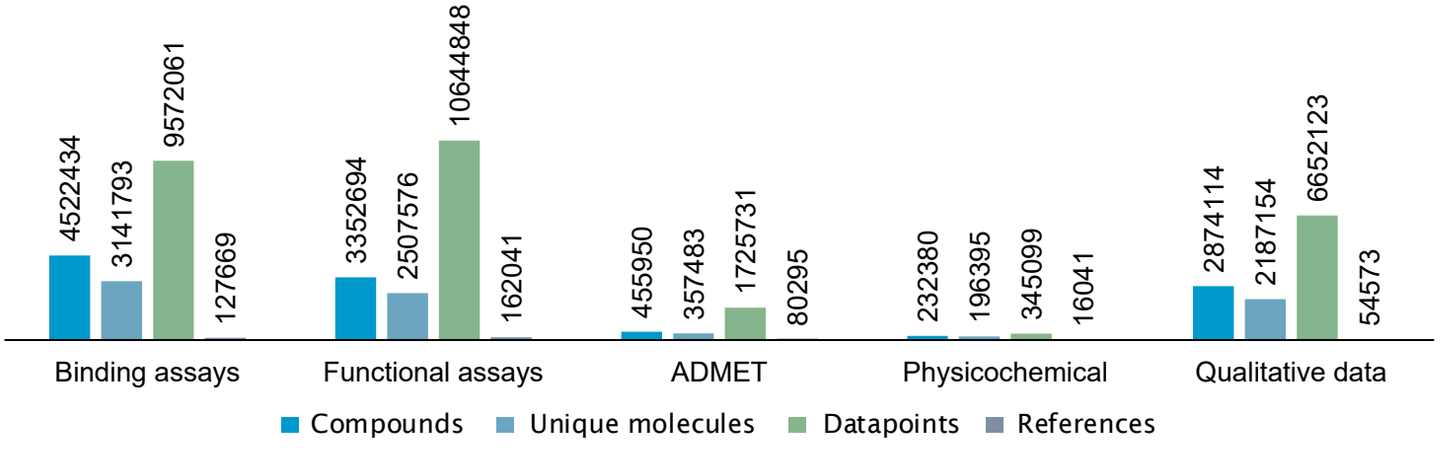
The GOSTAR database is integrated with a wealth of SAR information from journals, patents and a multitude of other disparate public sources. Heterogeneous and unstructured data reported in diverse sources is transformed into a structured relational database.



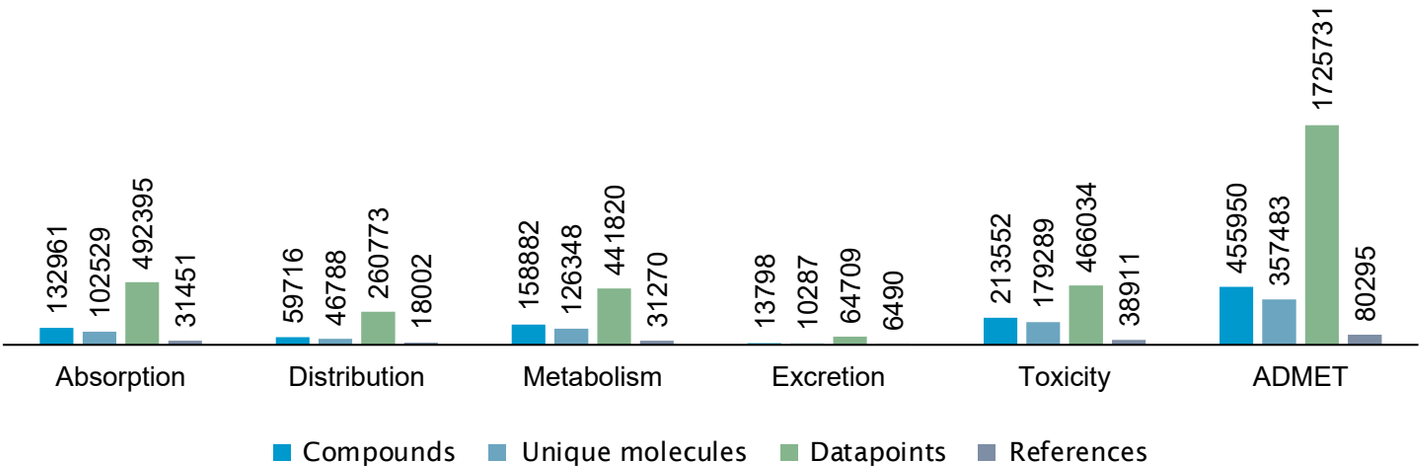
GOSTAR Data Coverage

A 360-degree view of small molecule intelligence

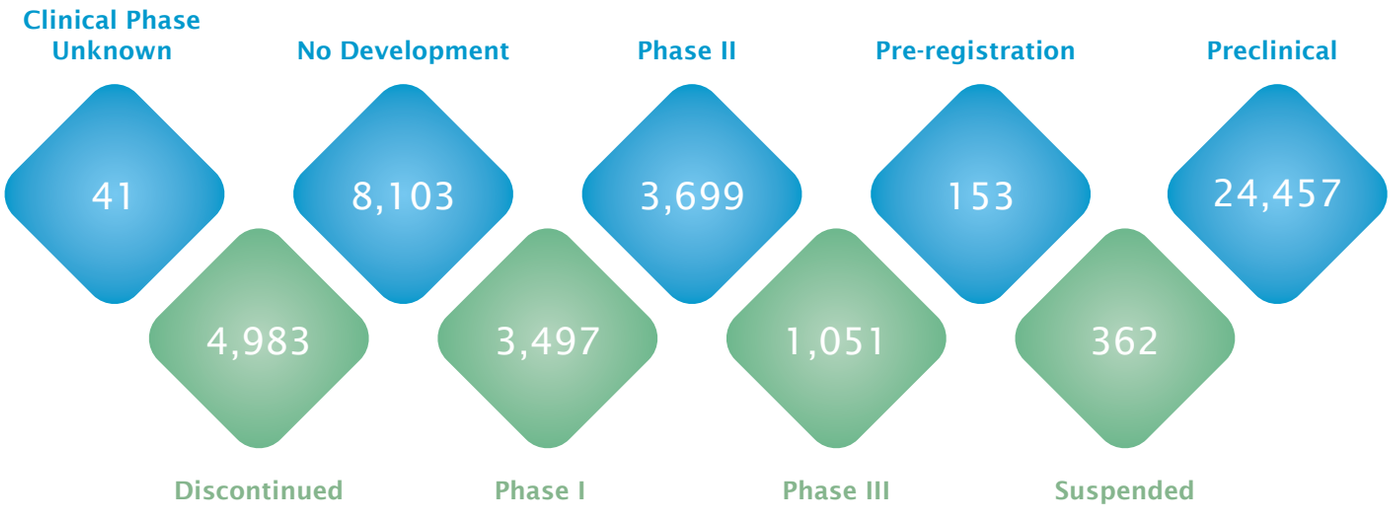
SAR data



ADMET data



Clinical compounds data



Approved drugs data

5,444
Launched

192
Withdrawn

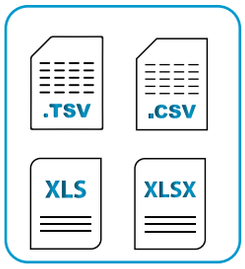
Custom Data Curation Services

Databases built for bespoke customer requirements

In addition to GOSTAR being made available as an off-the-shelf database, Excelra also offers customized data curation solutions catered to specific customer needs. Our services include custom content extraction, normalization, and analysis of diverse datasets in the biochemical space that includes but is not limited to; chemical structures, SAR data, protein data bank (PDB) information, chemical reactions, Markush structures and much more. We also create datasets with industry-accepted ontologies that can be seamlessly integrated with client-end platforms, analytical tools, protocols or methodologies.

Data delivered in flexible file-formats

Flat Files



Hierarchical Files

Database(s)

ORACLE

MySQL

PostgreSQL

Semantic format

GOSTAR Quality

3 level QC process certified by QMS-ISO standards

Enabled with process driven curation methodologies, NLP augmented human intelligence and a 3-level Quality Control process, the GOSTAR database is best in class and analysis ready.

> **"10 of the Top 20 pharma companies utilize GOSTAR to support their drug discovery programs"**

Our Clients

- Pharmaceuticals
- Biotech
- Research and Academia
- AI/ML Driven Drug Discovery Companies
- Contract Research Organizations

For more information, visit https://www.gostardb.com/gostar/

For any business queries, or to schedule a meeting with our scientific team, write to us on: marketing@excelra.com

excelra

About Excelra

Excelra's data and analytics solutions empower innovation in life sciences across the value chain from discovery to market. The Excelra Edge comes from a seamless amalgamation of proprietary curated data assets, deep domain expertise and data science. The company's multifaceted teams harmonize and analyse large volumes of disparate unstructured data using cutting-edge technologies. We galvanize data-driven decisions to unlock operational efficiencies to accelerate drug discovery and development. Over the past 18 years, Excelra has been the preferred data and analytics partner to over 90 global clients, including 15 of the top 20 large Pharma companies.

marketing@excelra.com

www.excelra.com