Open PHACTS
Open Pharmacological Space

Solving Use Cases

Description of practical workflows for addressing scientific questions and published use cases

Open PHACTS
Open Pharmacological Space

Published List of Use Cases

Number	sum	Nr of 1	Question
15	12	9	All oxidoreductase inhibitors active <100nM in both human and mouse
18	14	8	Given compound X, what is its predicted secondary pharmacology? What are the on and off,target safety concerns for a compound? What is the evidence and how reliable is that evidence (journal impact factor, KOL) for findings associated with a compound?
24	13	8	Given a target find me all actives against that target. Find/predict polypharmacology of actives. Determine ADMET profile of actives.
32	13	8	For a given
37	13	8	The curren in serine pr
38	13	8	Retrieve al structure (v
41	13	8	A project is compounds the target c level of the
44	13	8	Give me al
46	13	8	Give me th (disease)
59	14	8	Identify all

ELSEVIER

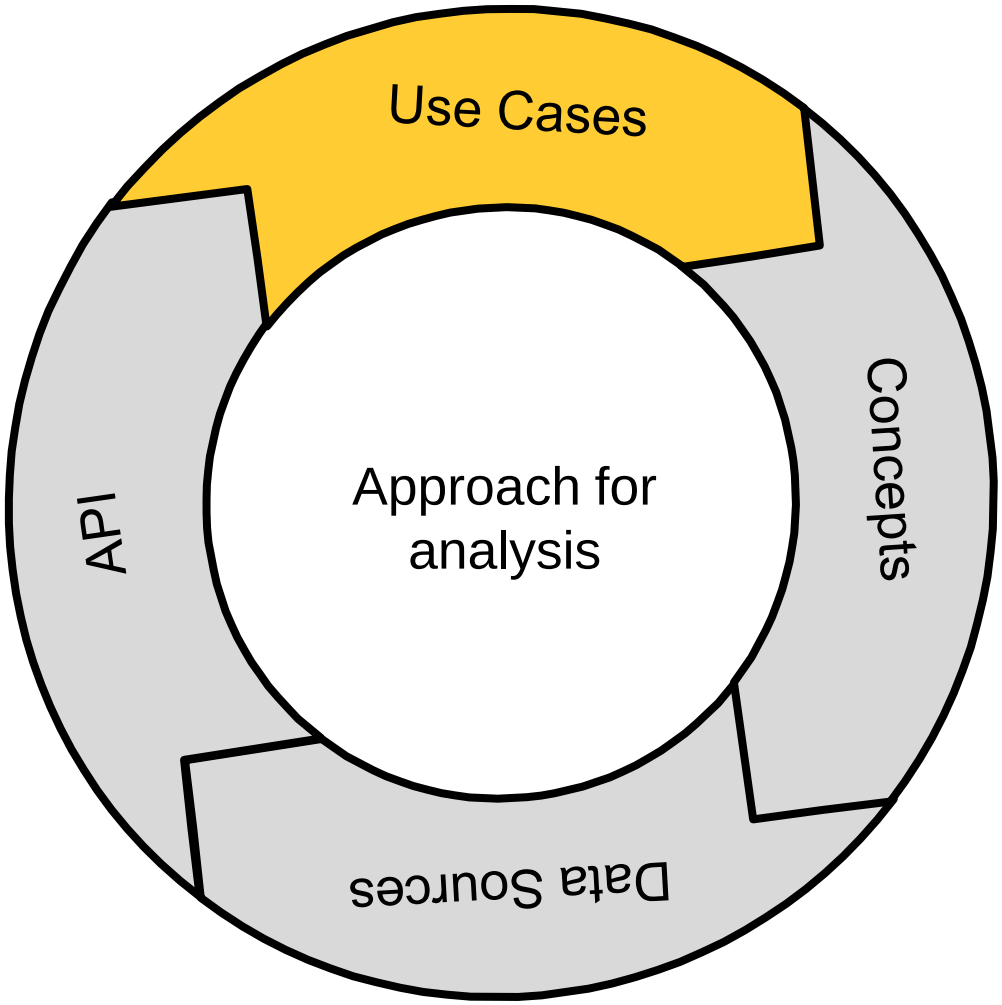
Drug Discovery Today

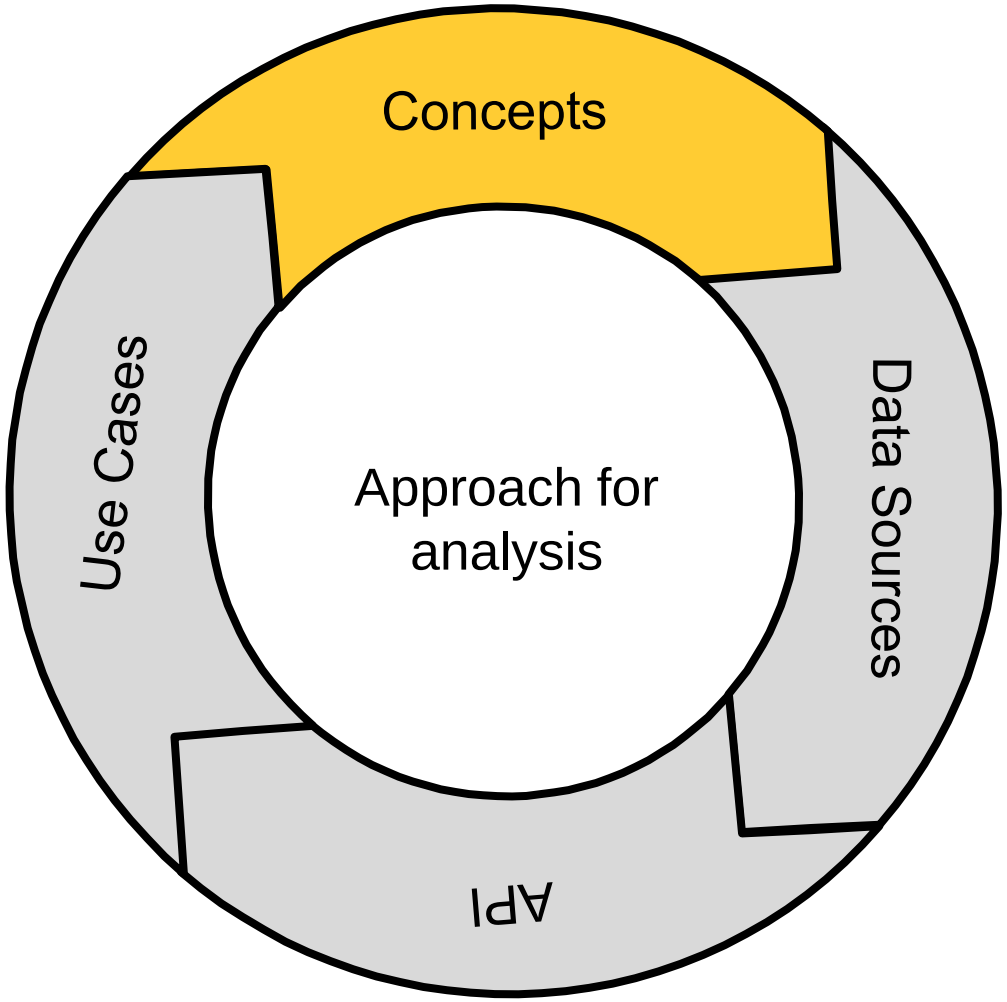
Volume 18, Issues 17–18, September 2013, Pages 843–852

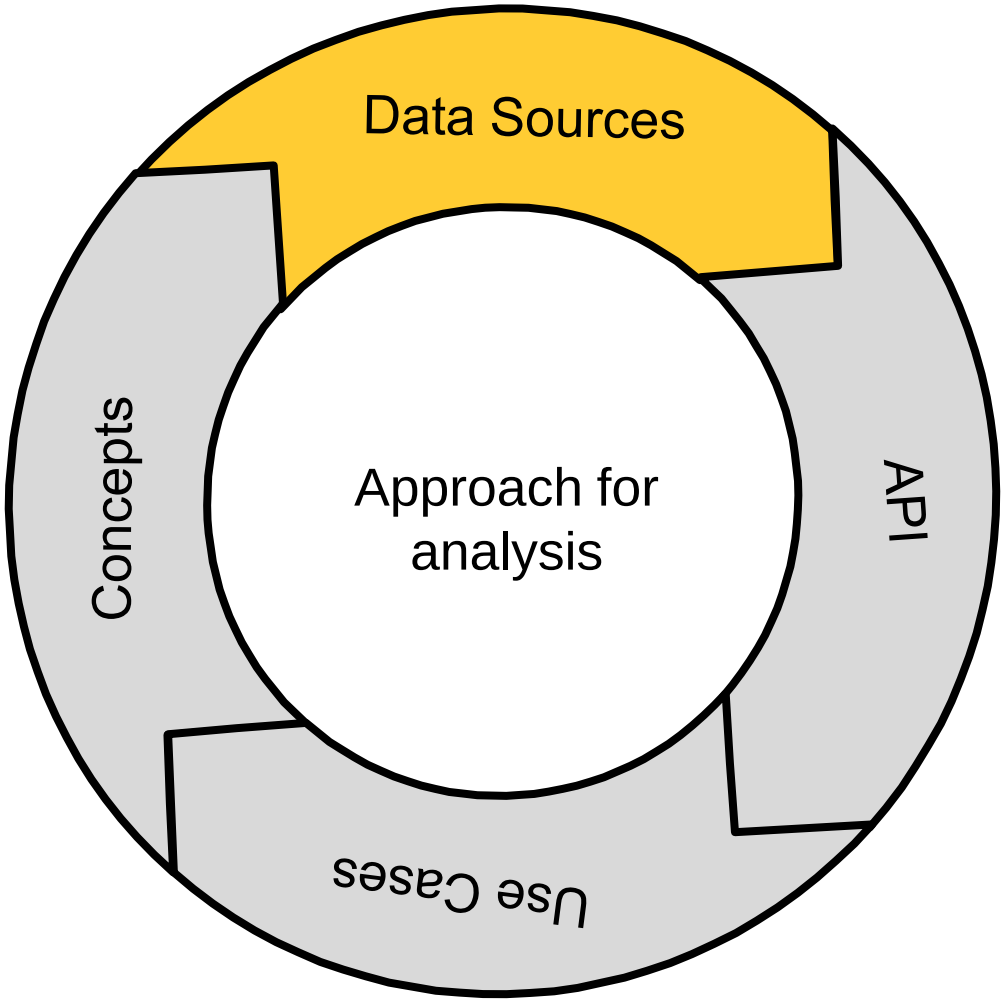
Review

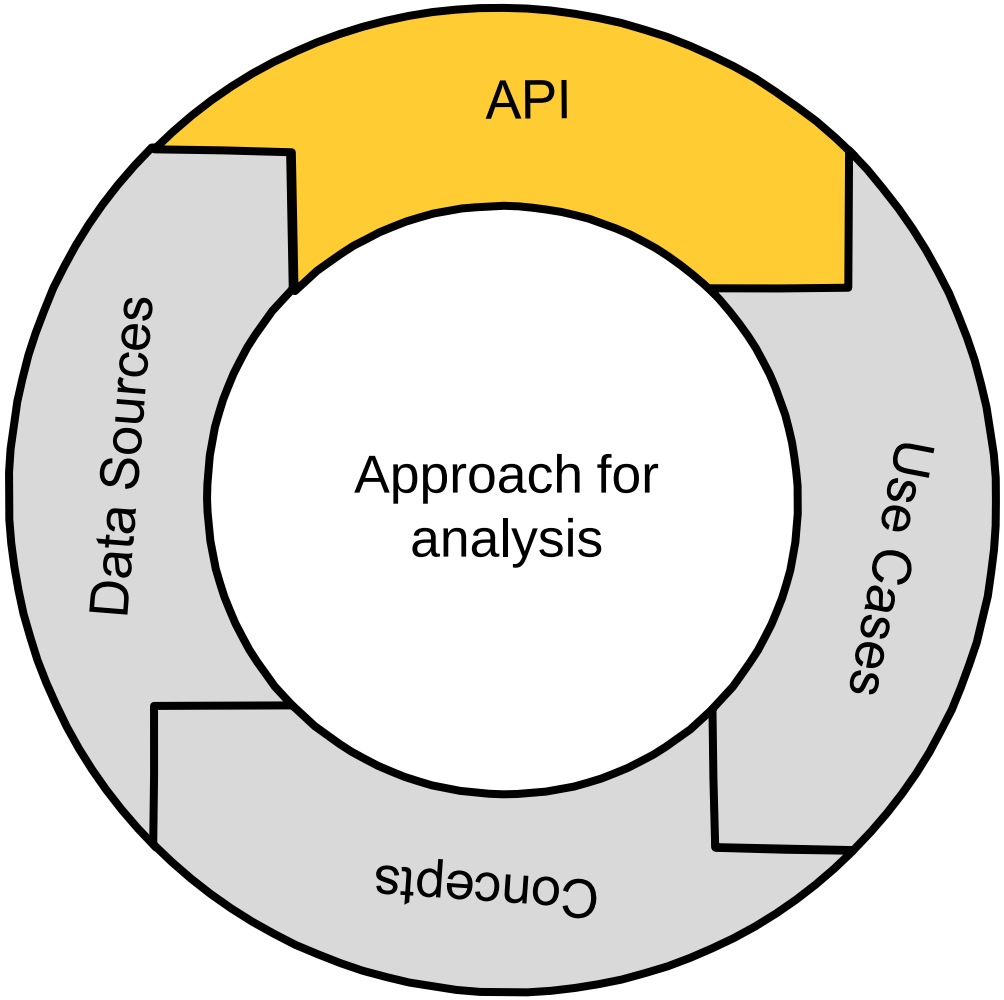
Scientific competency questions as the basis for semantically enriched open pharmacological space development

Kamal Azzaoui[1], Edgar Jacoby[14], Stefan Senger[2], Emiliano Cuadrado Rodríguez[3], Mabel Loza[3], Barbara Zdrazil[4], Marta Pinto[4], Antony J. Williams[5], Victor de la Torre[6], Jordi Mestres[7], Manuel Pastor[7], Olivier Taboureau[8], Matthias Rarey[9], Christine Chichester[10], Steve Pettifer[11], Niklas Blomberg[12, a], Lee Harland[13], Bryn Williams-Jones[13], Gerhard F. Ecker[4]


Use Cases
Concepts
Data Sources
API
Approach for analysis


Concepts
Data Sources
API
Use Cases
Approach for analysis


Data Sources
API
Use Cases
Concepts
Approach for analysis


API
Data Sources
Use Cases
Concepts
Approach for analysis

Open PHACTS
Open Pharmacological Sp

Use Cases

-- oxidoreductase inhibitors active <100 nM in human and mouse

-- active compounds on a given target with the relevant assay data

-- for a given compound, give the interaction profile with targets

-- for a given compound, summarize all ‘similar compounds’ and their activities

-- identify all known protein–protein interaction inhibitors

-- activity data in serine protease assays for molecules that contain substructure X

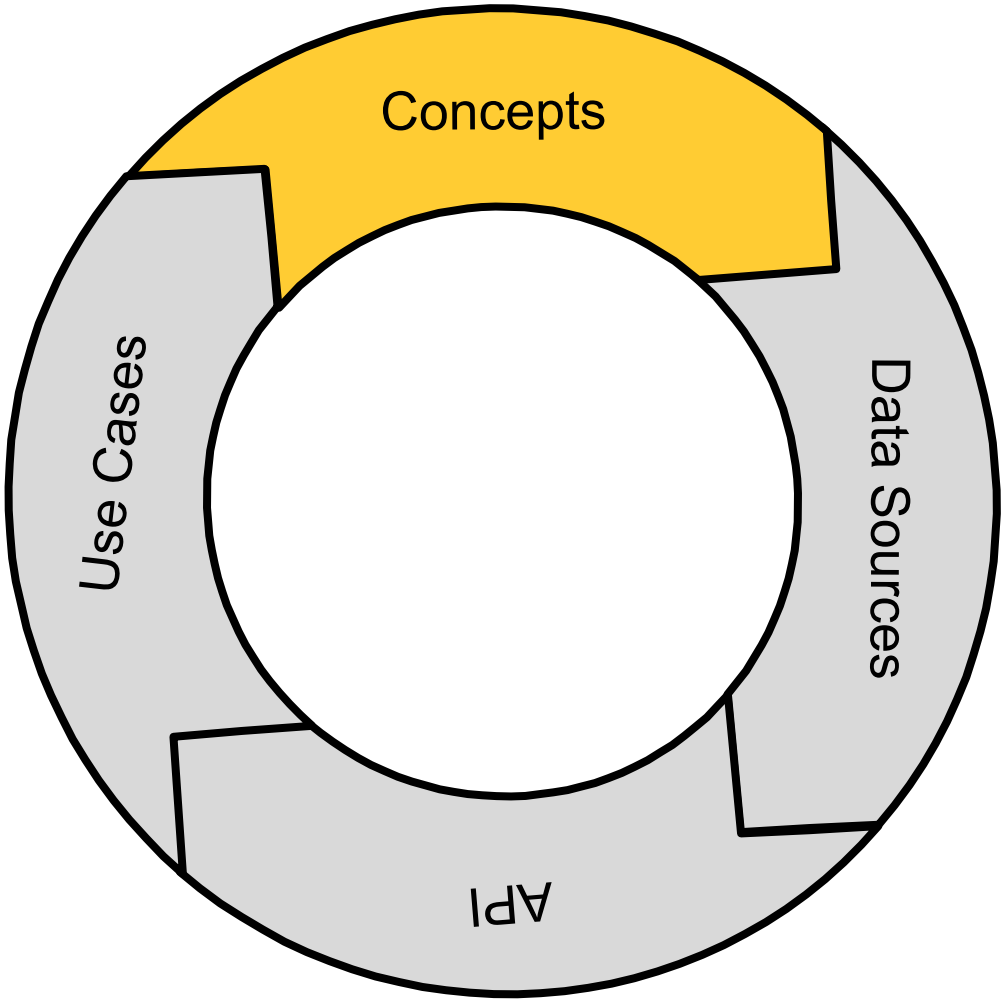
-- compounds assayed is at resolution of level of the target family

-- a given disease give all targets in the pathway and all compounds hitting them

-- targets that have been patented in the context of a specific disease

-- targets that have been screened by specific companies

-- all clinical data for ligands of specific targets


Concepts
Data Sources
API
Use Cases

Open PHACTS
Open Pharmacological Spa

Concepts

-- targets, activities, species

-- compounds, target, assays

-- compounds, targets

-- compound, ‘similar compounds’, activities

-- ppi inhibitors (compounds), targets

-- assay families, substructure

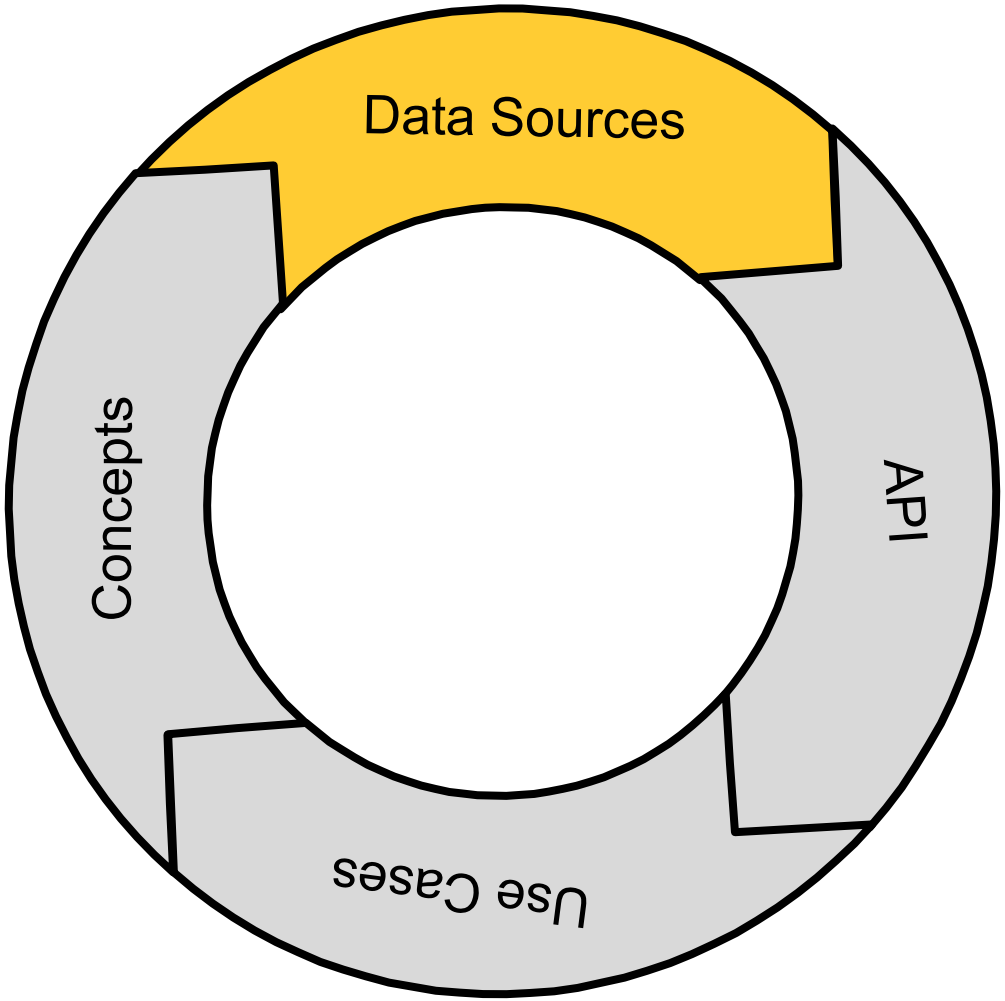
-- compounds, target families

-- compounds, target, pathways, diseases

-- targets, patents, diseases

-- targets screened by specific companies

-- clinical data, compounds, targets


Data Sources
API
Use Cases
Concepts

Open PHACTS
Open Pharmacological S

Data Sources

-- ENZYME

-- Chembl

-- UniProt

-- DrugBank

-- ConceptWiki

-- ChemSpider

-- structure search

-- substructure search

-- Chembl

-- UniProt

-- WikiPathways

-- GeneOntology

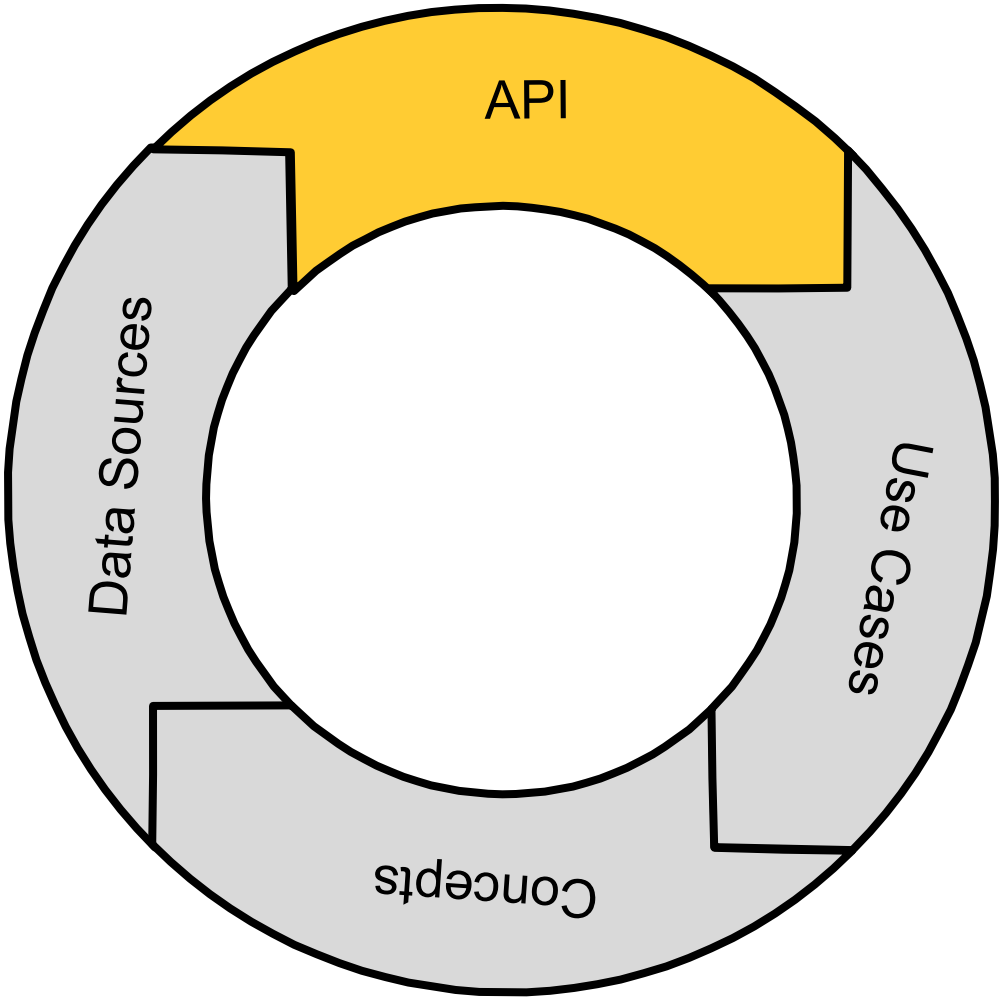
-- ChEBI

-- Chembl Target Classification

-- DisGeNet

-- neXtProt

-- FDA Adverse Event Reporting System


API
Use Cases
Concepts
Data Sources

Open PHACTS
Open Pharmacological

API

-- Target Class Pharmacology

Target Class: http://purl.uniprot.org/enzyme/1.-.-.-

*restrictions: homo sapiens|mus musculus, activity_type IC50, maxEx-activity_value 100, activity_unit nanomolar, activity_relation <=|=

-- Target Pharmacology Paginated

Target Pharm: Prostaglandin G/H synthase 2 (Homo sapiens)http://www.conceptwiki.org/concept/e75776b6-0433-419c-a15a-74729d8bf254

* Retrieve assay descriptions

-- Chemical Structure Similarity Search

*Cmpd Info: http://www.conceptwiki.org/concept/342a03eb-3311-49ac-8d6e-8bf9b605dab1

* Retrieve Smiles for similarity search: CC(=O)NC1=CC=C(C=C1)O

-- Target Information

Target Info: E3 ubiquitin-protein ligase Mdm2 (Homo sapiens) http://www.conceptwiki.org/concept/2e754389-d9d4-4c93-8046-f03722573946

•Restriction: TargetType ppi in Target pharm

• PPI from Intact in Target info

-- Target Classification

Members of target class: Serine Protease activity http://purl.org/obo/owl/GO#GO_0008236

*restriction :http://purl.uniprot.org/taxonomy/9606

*get pharm by target, take molecules

*substructure structure of unknown compare with results with molecules from pharm by target

-- Targets in Pathway

Target: E3 ubiquitin-protein ligase Mdm2 (Homo sapiens) http://www.conceptwiki.org/concept/2e754389-d9d4-4c93-8046-f03722573946

-- Disease Information

-- Expression Information

*Protein expressed in tissue

-- Off target Information

*Cmpd Info: Adverse events

Progress on Use Cases

- Most use cases can be solved with 1.3 API
- Some use cases are partially solved
- Not many need new data sources and API calls

EXTRA EXAMPLES

Procedure for Use Case

Example 1: Provide all activities for a given compound X, with targets annotated by gene (Compound -> target)

- needed API calls:
  - Compound Pharmacology Paginated
  - Map URL

Compound Pharmacology Paginated – parameters

- uri: needs a compound uri. Possible sources:
  - Map free text to a concept URL methods
  - SMILES, Inchi or InchiKey to URL methods
  - Chemical Structure search methods
  - known identifier from other sources: e.g. http://www.chemspider.com/2157, http://rdf.ebi.ac.uk/resource/chembl/molecule/CHEMBL25,
  - http://en.wikipedia.org/wiki/Aspirin
- app_id and app_key
- target_type: single_protein (if wanted)
- _pageSize (default 10, use with loop on _page, or set to all)

Compound Pharmacology Paginated - results


```
- items: [
    - {
        _about: http://rdf.ebi.ac.uk/resource/chembl/activity/CHEMBL_ACT_10834249,
        pmid: http://identifiers.org/pubmed/22386242,
      - hasAssay: {
            _about: http://rdf.ebi.ac.uk/resource/chembl/assay/CHEMBL2016143,
            description: "Inhibition of COX1 at 400 uM after 5 mins by spectrophotometric analysis",
          - hasTarget: {
                _about: http://rdf.ebi.ac.uk/resource/chembl/target/CHEMBL221,
                title: "Cyclooxygenase-1",
                targetOrganismName: "Homo sapiens",
                inDataset: http://www.ebi.ac.uk/chembl,
                type: http://rdf.ebi.ac.uk/terms/chembl#SingleProtein
            },
            inDataset: http://www.ebi.ac.uk/chembl
        },
      - activity_unit: {
            _about: http://qudt.org/vocab/unit#Percent,
            prefLabel: "%"
        },
        publishedRelation: "=",
        publishedType: "INH",
        publishedUnits: "%",
        publishedValue: 38.4,
        activity_relation: "=",
        activity_type: "Inhibition",
        activity_value: 38.4,
        inDataset: http://www.ebi.ac.uk/chembl
    },
```

Retrieving a protein/gene ID

- Use the Map URL API call
- Uri: any input uri, e.g. http://rdf.ebi.ac.uk/resource/chembl/target/CHEMBL221
- (Restrict to wanted targetUriPattern)
- Example results:
  - http://www.ncbi.nlm.nih.gov/gene/5742
  - http://www.uniprot.org/uniprot/P23219

Example 2: Provide all compounds assayed for target Y with target indicated by a gene (Target -> Compound)

- needed API call:
  - Target Pharmacology Paginated

- input parameter:
  - uri: protein or gene uri, e.g. http://www.ncbi.nlm.nih.gov/gene/5742

```
- hasAssay: {
      _about: http://rdf.ebi.ac.uk/resource/chembl/assay/CHEMBL762601,
      description: "Inhibitory concentration against Prostaglandin G/H synthase from isolated rabbit granulocytes",
      assayOrganismName: "Oryctolagus cuniculus",
    - hasTarget: {
          _about: http://rdf.ebi.ac.uk/resource/chembl/target/CHEMBL2094253,
          title: "Cyclooxygenase",
          targetOrganismName: "Homo sapiens",
          inDataset: http://www.ebi.ac.uk/chembl,
          type: http://rdf.ebi.ac.uk/terms/chembl#ProteinFamily
      },
      inDataset: http://www.ebi.ac.uk/chembl
  },
- hasMolecule: {
      _about: http://rdf.ebi.ac.uk/resource/chembl/molecule/CHEMBL19657,
    - exactMatch: [
        - {
              _about: http://ops.rsc.org/OPS73199,
              inDataset: http://ops.rsc.org,
              inchi: "InChI=1S/C14H17NO/c1-15-13-6-4-2-3-5-11(13)12-9-10(16)7-8-14(12)15/h7-9,16H,2-6H2,1H3",
              inchikey: "BAHOYABULMMMHV-UHFFFAOYSA-N",
              molweight: 215.291,
              ro5_violations: 0,
              smiles: "CN1C2=C(CCCCC2)C3=C1C=CC(=C3)O"
          },
        - {
              _about: http://www.conceptwiki.org/concept/9d0642ad-cd9a-4f29-a650-046bba7a12f2,
              inDataset: http://www.conceptwiki.org,
              prefLabel_en: "5-methyl-5,6,7,8,9,10-hexahydrocyclohepta[b]indol-2-ol",
              prefLabel: "5-methyl-5,6,7,8,9,10-hexahydrocyclohepta[b]indol-2-ol"
          }
      ]
  },
- activity_unit: {
      _about: http://www.openphacts.org/units/Nanomolar,
      prefLabel: "nM"
  },
  pChembl: 6.4,
  publishedRelation: "=",
  publishedType: "IC50",
  publishedUnits: "M",
  publishedValue: 4e-7,
  activity_relation: "=",
  activity_type: "IC50",
  activity_value: 400,
```

Example 3: For target X provide target family. (Gene -> Gene family)

- needed API calls:
  - Target Classifications

- parameters for classification:
  - tree: chembl, enzyme or go


```
  _about: http://rdf.ebi.ac.uk/resource/chembl/target/CHEMBL221,
  inDataset: http://www.ebi.ac.uk/chembl,
- hasEnzymeClassification: {
    _about: http://purl.uniprot.org/enzyme/1.14.99.1,
  - inDataset: {
      _about: http://purl.uniprot.org/enzyme,
      prefLabel: "Enzyme Classification"
    },
  - prefLabel: [
      "Prostaglandin-endoperoxide synthase",
      "Prostaglandin synthetase",
      "PG synthetase",
      "Prostaglandin synthase",
      "Prostaglandin G/H synthase"
```

Retrieving Pharmacology for all proteins with a given classification

- Target Class Pharmacology Paginated
  - uri: Classification uri e.g. http://purl.uniprot.org/enzyme/1.14.99.1

Filter annotation results by activity cut-offs

- Available filters:
  - Activity type / cutoff value / units combinations
  - pChembl cutoff values
  - activity relation filters (>, >=, =, <, <=)
  - Organism filters (target and assay organism)
  - Target type (e.g. single_protein, protein_family)

Activity type / cutoff value / units combinations

- possible values:
  - Activity types: e.g. Potency, GI50, IC50, Ki, …
  - cutoff values: number with the appropriate parameter
    - = activity_value
    - >= min-activity_value
    - > minEx-activity_value
    - <= max-activity_value
    - < maxEx-activity_value
  - Units for activity type: e.g. nanomolar, microgram_per_milliliter

pChembl cutoff values

- Definition: -Log(molar IC50, XC50, EC50, AC50, Ki, Kd or Potency)
  - e.g. IC50=10µM -> pChembl = 5
- Filters:
  - = pChembl
  - >= min-pChembl
  - > minEx-pChembl
  - <= max-pChembl
  - < maxEx-pChembl

Open PHACTS
Open Pharmacological Space

Substructure and Similarity search in Open PHACTS

- Uses ChemSpider search tools (Bingo from GGA)
- Chemical Structure Exact/Similarity/Substructure Search

Chemical Structure Exact Search	/structure/exact	GET
InchiKey to URL	/structure	GET
Inchi to URL	/structure	GET
Chemical Structure Similarity Search	/structure/similarity	GET
SMILES to URL	/structure	GET
Chemical Structure Substructure Search	/structure/substructure	GET

Chemical Structure Similarity Search

- selected parameters:
  - searchOptions.Molecule: a SMILES string
  - searchOptions.SimilarityType: 0: Tanimoto; 1: Tversky; 2: Euclidian
  - searchOptions.Threshold: value between 0 and 1
  - resultOptions.Count: number of results

Example 5: For a target Y, find pathway Z

- Needs API call:
  - Pathways by Target
    - Find an identifier for the target
      - http://www.conceptwiki.org/concept/2e754389-d9d4-4c93-8046-f03722573946
    - Additional: Map URL, possibly with the restriction http://identifiers.org/ncbigene/ or http://identifiers.org/ensembl/ to retrieve complementary identifiers
    - Additional: Count Pathways by Target

Open PHACTS
Open Pharmacological Space

Pathways by Target


```
 "_about": "http://rdf.wikipathways.org/Pathway/WP710_r67156",
   "identifier": "http://identifiers.org/wikipathways/WP710",
   "title": "DNA damage response (only ATM dependent)",
   "description": "This is the second pathway out of two pathways which
deals with DNA damage response",
   "hasPart": {
    "_about": "http://identifiers.org/ensembl/ENSG00000135679",
    "type": "http://vocabularies.wikipathways.org/wp#GeneProduct",
    "exactMatch": {
     "_about": "http://www.conceptwiki.org/concept/2e754389-d9d4-4c93-
8046-f03722573946",
      "prefLabel": "E3 ubiquitin-protein ligase Mdm2 (Homo sapiens)"
    }
   },
   "inDataset": "http://www.wikipathways.org",
   "pathway_organism": {
    "_about": "http://purl.obolibrary.org/obo/NCBITaxon_9606",
    "label": "Homo sapiens"
   },
```

Map URL with restriction on namespace

"_about": "http://www.conceptwiki.org/concept/2e754389-d9d4-4c93-8046-f03722573946",
"exactMatch": "http://identifiers.org/ncbigene/4193”,
"http://identifiers.org/ensembl/ENSG00000135679",

Example 5: For a pathway Z, find target Y

- Needs API call:
  - Get Targets for Pathway, textually names for pathways can be searched (Map free text to concept URL) or use WikiPathway URI directly:
    - http://www.conceptwiki.org/concept/ad73ac9e-f606-4ea9-b357-3eed6f837ca8
    - http://www.wikipathways.org/instance/WP710

Open PHACTS
Open Pharmacological Space

Get Targets for Pathway

"_about": "http://rdf.wikipathways.org/Pathway/WP710_r67156",
"title_en": "DNA damage response (only ATM dependent)",
"title": "DNA damage response (only ATM dependent)",
"hasPart": [
"http://identifiers.org/ncbigene/7476",
"http://identifiers.org/ncbigene/7481",
"http://identifiers.org/ncbigene/8061",…

List of genes/proteins in pathway