Covid19 Drug Discovery Workflow Platform (CDD-WP)

Cloud Enabled AI Driven Workflows for Drug Discovery and Design

Context

- ✓Our **CDD-WP has been tested on 60% of the drug discovery work streams**; built on KNIME workflows with Python scripts running on MS Azure Cloud. We want to test the scale and execution speed of these workflows on an **HPC platform**

- ✓We intend to use multiple of AI/ML frameworks on identified molecules/compounds/protein structures to **increase the reliability and trust from these AI results**

- ✓Our objective is to **deploy and test to run CDD-WP as a service and leverage multiple AI/ML frameworks** on possible Covid19 structures/compounds to aid in drug design/discovery process

Credits from Existing Works

* Computer Aided Drug Design/Discovery

TeachOpen CADD

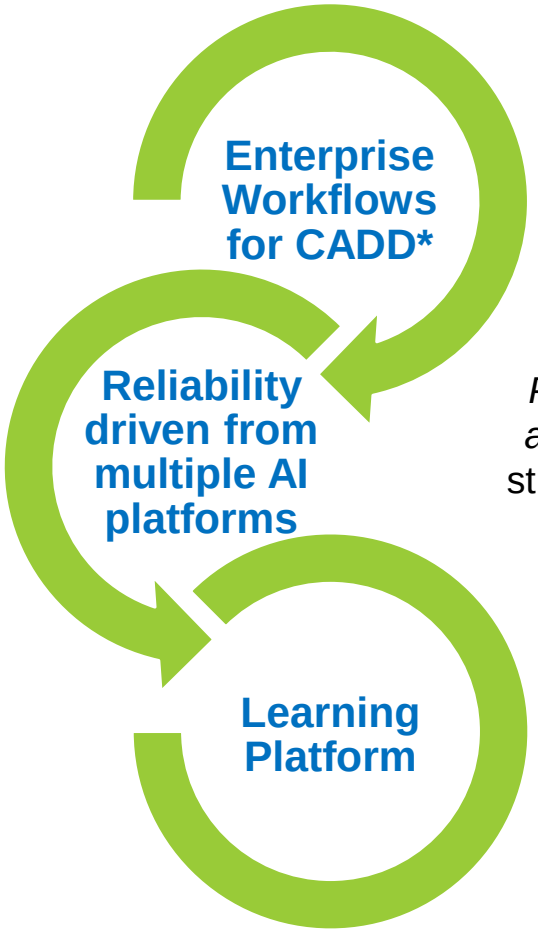
10+ generic workflows that are built for computer-aided drug design (CADD*) using KNIME that needs enterprise customization for Covid19

Google AlphaFold

5+ Covid19 Structure Predictions e.g. M_protein, Protein_3a, Nsp2, Nsp4, Nsp6 and PL-PRO C terminal

TCS

30+ compounds inhibiting the 3CL protease (3CL is essential for SARS-CoV-2's survival and replication in the host)

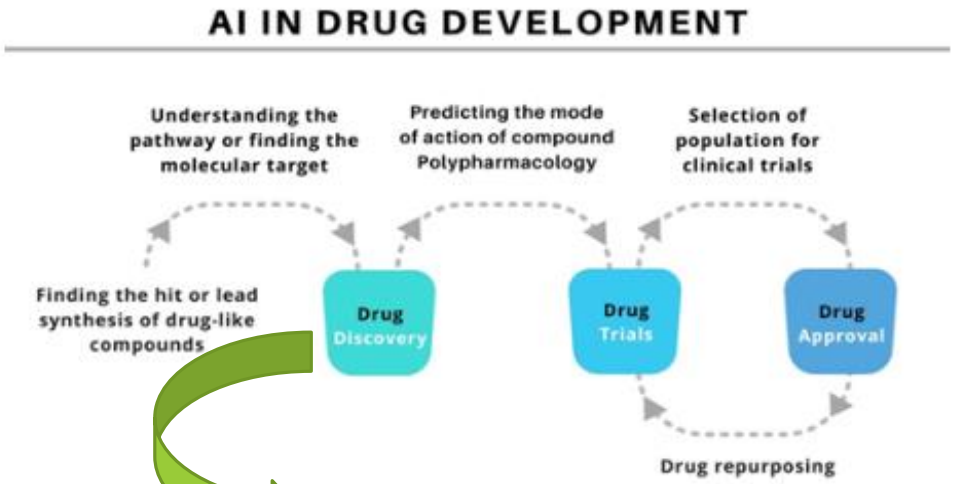


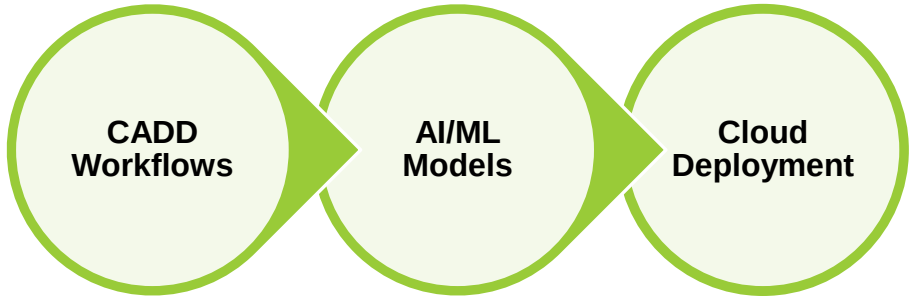
Ability to set up quick enterprise workflows for CADD* process

Platform should leverage already predicted protein structures from AI platforms like Google's AlphaFold

Scale the platform into a learning system to reduce biasness from identified compounds from multiple research studies

CDD-WP's Readiness

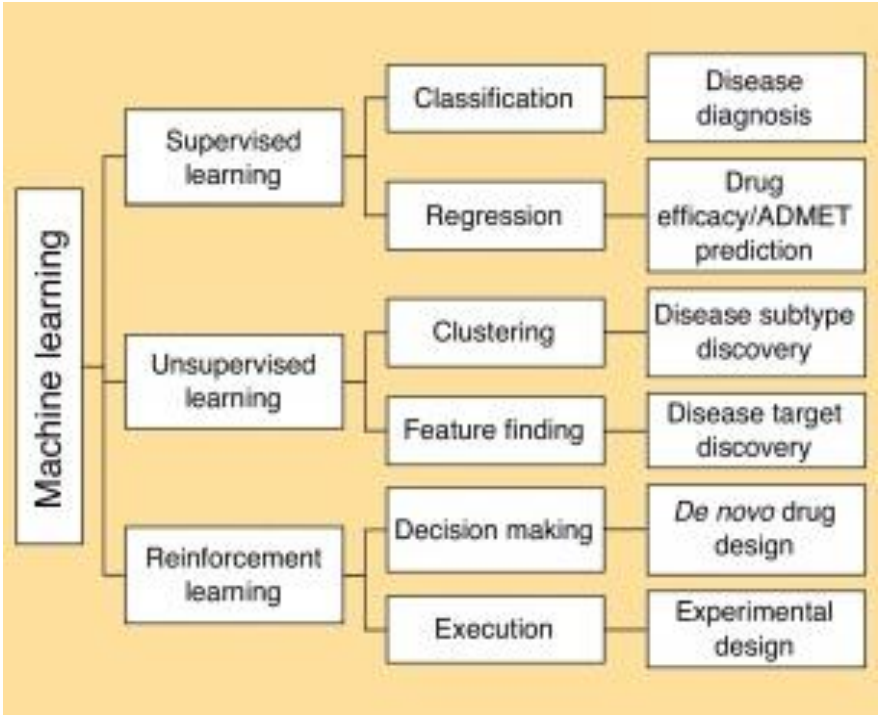




Run Drug Discovery as a Service

Sl No.	Steps	Digital Workflow ID	Workflow Available?	Python Scripts Available?	AI/ML Models Available?	Leveraging Existing Works?	Cloud Deployment Ready?
1	Compound data acquisition	CADD_W1	Y	Y	WIP	Y	Y
2	Molecular filtering: ADME and lead-likeness criteria	CADD_W2	Y	Y	Y	Y	Y
3	Molecular filtering: Unwanted substructures	CADD_W3	Y	Y	WIP	Y	Y
4	Ligand-based screening: Compound similarity	CADD_W4	Y	Y	Y	Y	Y
5	Compound clustering	CADD_W5	Y	Y	Y	N/A	Y
6	Maximum common substructures	CADD_W6	Y	Y	WIP	N/A	Y
7	Ligand-based screening	CADD_W7	Y	Y	Y	N/A	Y
8	Protein data acquisition	CADD_W8	Y	Y	WIP	Y	Y
9	Ligand-based pharmacophores	CADD_W9	WIP	Y	Y	N/A	WIP
10	Binding site similarity	CADD_W10	WIP	Y	Y	N/A	WIP
11	Structure-based CADD using online APIs/servers	CADD_W11	WIP	Y	Y	Y	WIP
11.1	Querying KLIFS & PubChem for potential kinase inhibitors	CADD_W12	WIP	Y	WIP	Y	WIP
11.2	Docking the candidates against the target	CADD_W13	WIP	Y	Y	Y	WIP
11.3	Visualizing the results and comparing against known data	CADD_W14	N/A	Y	N/A	N/A	WIP

CDD-WP's AI Models

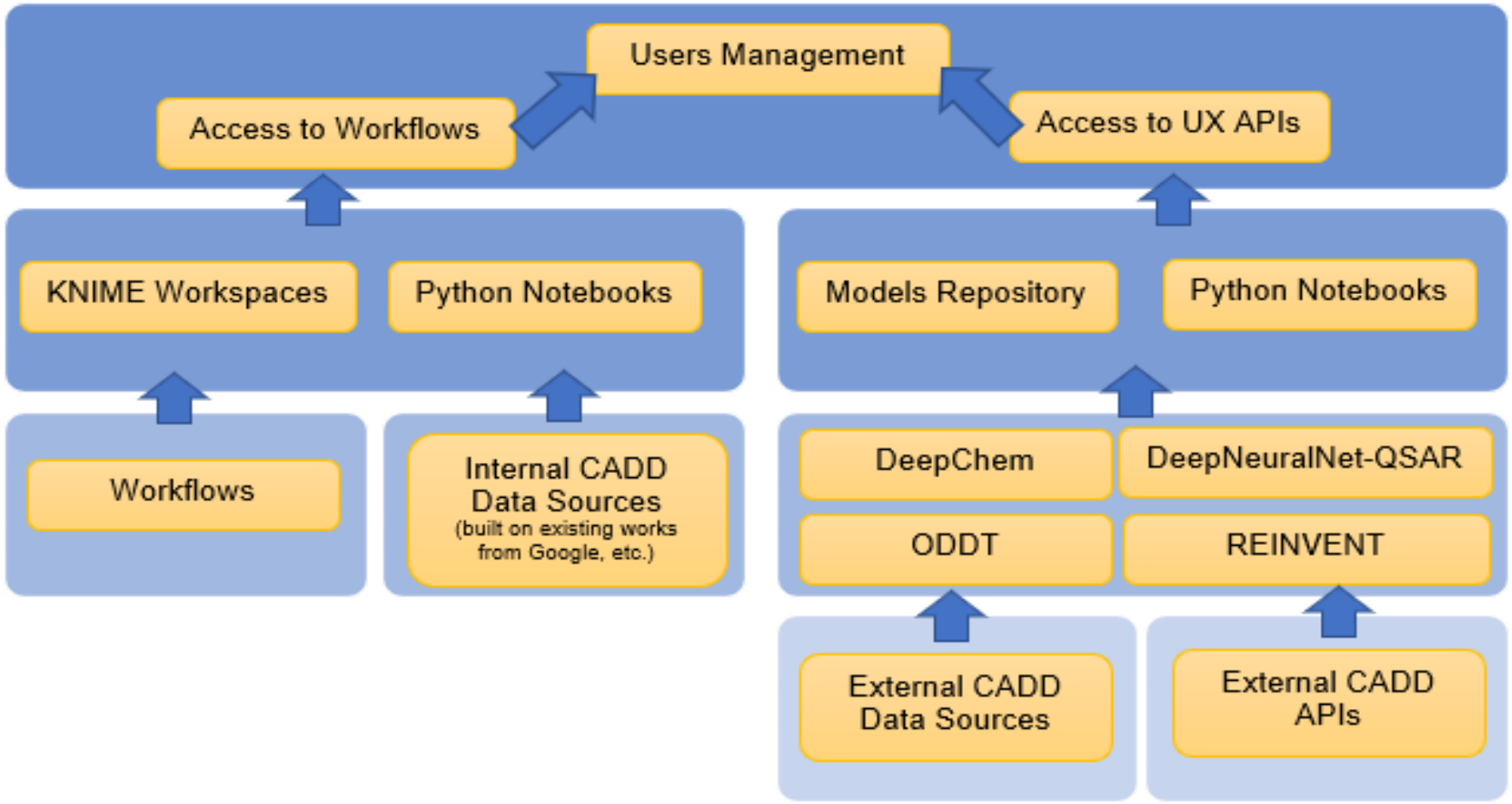


Drug Discovery Process	Drug Design Sub-Process	Applicable AI Models	Leverage Existing Works
Target Idenitification and Study	Prediction of Protein Folding	CNN	Google's AlphaFold
	Prediction of PPIs	FD/DCA	Google's AlphaFold
Hit Discovery	Drug Repurposing	Network Analysis	TCS
	Virtual Screening	SVM, AAE	TCS
	Activity Scoring	SVM, RF, CNN	TCS
Hit to Lead	QSAR	DNN	DeepNeuralNet-QSAR
	De Novo Drug Design	AAE, VAE	REINVENT, DeepChem
Lead Optimization	Evaluation of ADME/T properties	CNN	DeepChem, Open Drug Discovery Toolkit

We will be leveraging the below mentioned AI platforms and data sources in addition to ChEMBL/PubChem/PDB –

a) DeepChem
b) Open Drug Discovery Toolkit (ODDT)
c) REINVENT
d) DeepNeuralNet-QSAR
e) Google's AlphaFold (for 5 predicted protein structures)
f) TCS AI Virtual Screening (for 31 compounds)

CDD-WP's Functional Stack



CDD-WP's Benefits

* Computer Aided Drug Design/Discovery

1

Plug-in to any CADD* Process

Integrated CADD Workflows with base codebase available for implementation

2

Integrates with 3+ CADD AI Platforms

Deploy multiple AI/ML frameworks/models during an implementation cycle to improve the reliability score

3

CADD Workflows as a Service

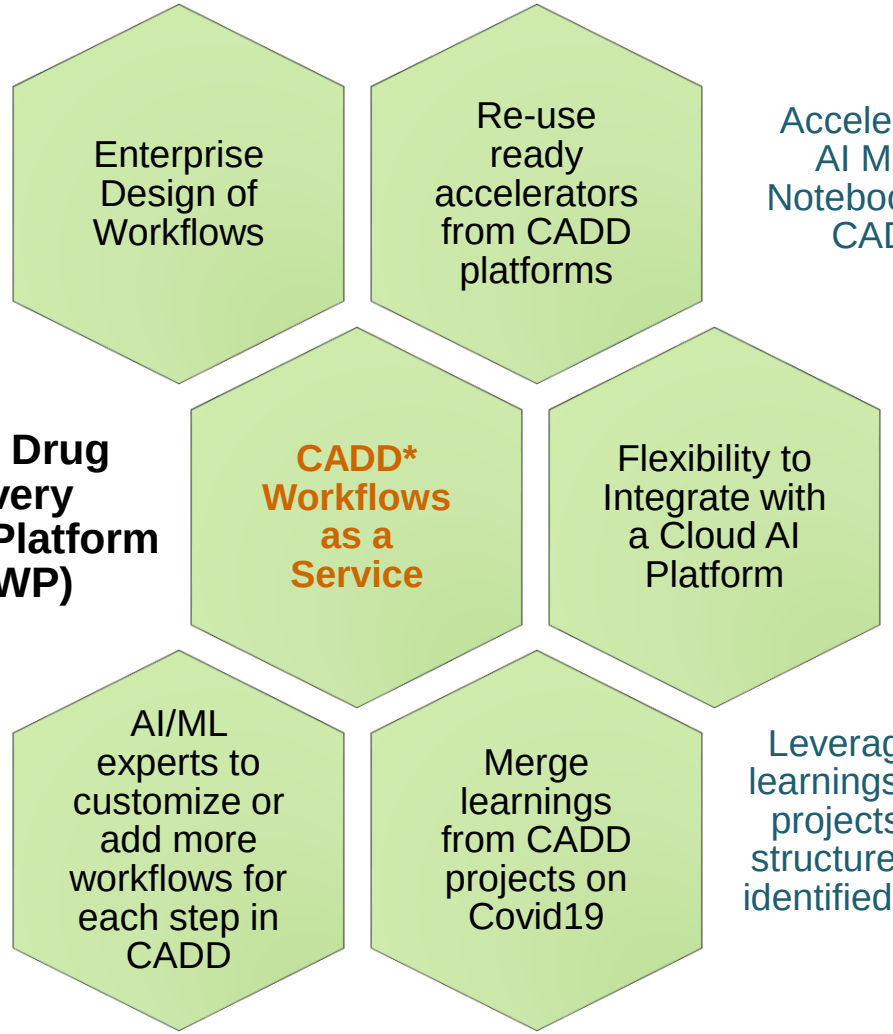
Data Management as a Service on identified compounds/molecules/ structures

- ✓ Helps to deliver the **Proof of Concept (POC) in 2-3 weeks**
- ✓ Easily scales up for enterprise adoption by using the **industry best practices in CADD as an integrated workflow platform**
- ✓ Improves reliability by leveraging **multiple CADD AI Platforms**

CDD-WP's Enterprise Scale

* Computer Aided Drug Design/Discovery

Workflows can be easily configured as per the CADD process for any research organization



Accelerators in form of AI Models, Python Notebooks from existing CADD platforms

CDD-WP can be easily extended and integrated with MS Azure AI/ML Platform

AI/ML experts to engage who have worked with cloud data lakes and enterprise workflow models

Leverage available learnings from CADD projects on protein structure predictions, identified compounds, etc.

Thank You

Greenojo provides Automation, Analytics and AI solutions to enterprise customers

For RFPs, Solutions and Sales/Partner enquiries, connect us at - *sales@greenojo.com*

Sales Offices
Houston, TX, USA |Burlington, ON, Canada |Dubai, UAE | Lagos, Nigeria

Delivery Offices - India
Bhubaneswar, Odisha |Hyderabad, Telangana | Trivandrum, Kerala