

CCP4 Cloud

*for macromolecular structure determination
online*

Eugene Krissinel

CCP4, Research Complex at Harwell, RAL, UK
eugene.krissinel@stfc.ac.uk

November 22, 2022

- CCP4 stands for "**Collaborative Computational Project No. 4 in Protein Crystallography**"
- One of several CCPs set up in the UK to advance and support scientific software developments
- CCP4 is not a single group in RCaH-Rutherford but a world-wide collaboration of developers and users
- CCP4 was formally set up in **1979** to support collaboration between researchers working on MX software in the UK, and to assemble a comprehensive collection of software to satisfy the computational requirements of the relevant UK groups
- CCP4 was originally supported by the UK Science and Engineering Research Council (SERC), and is now supported by the Biotechnology and Biological Sciences Research Council (BBSRC).
- The project is coordinated at the **Scientific Computing Department of STFC UK**

● Core Group @ RAL/Harwell (Oxford)

- ~ maintain and support software suite
- ~ application and infrastructural software developments
- ~ collaborate with Diamond on beam line deployment of CCP4 software and joint developments
- ~ educational outreach
- ~ maintain CCP4 resources such as the CCP4 bulletin board

● MRC/LMB (Cambridge)

- ~ data processing software (Mosflm, Aimless)
- ~ refinement software (Refmac)
- ~ model building (Coot)

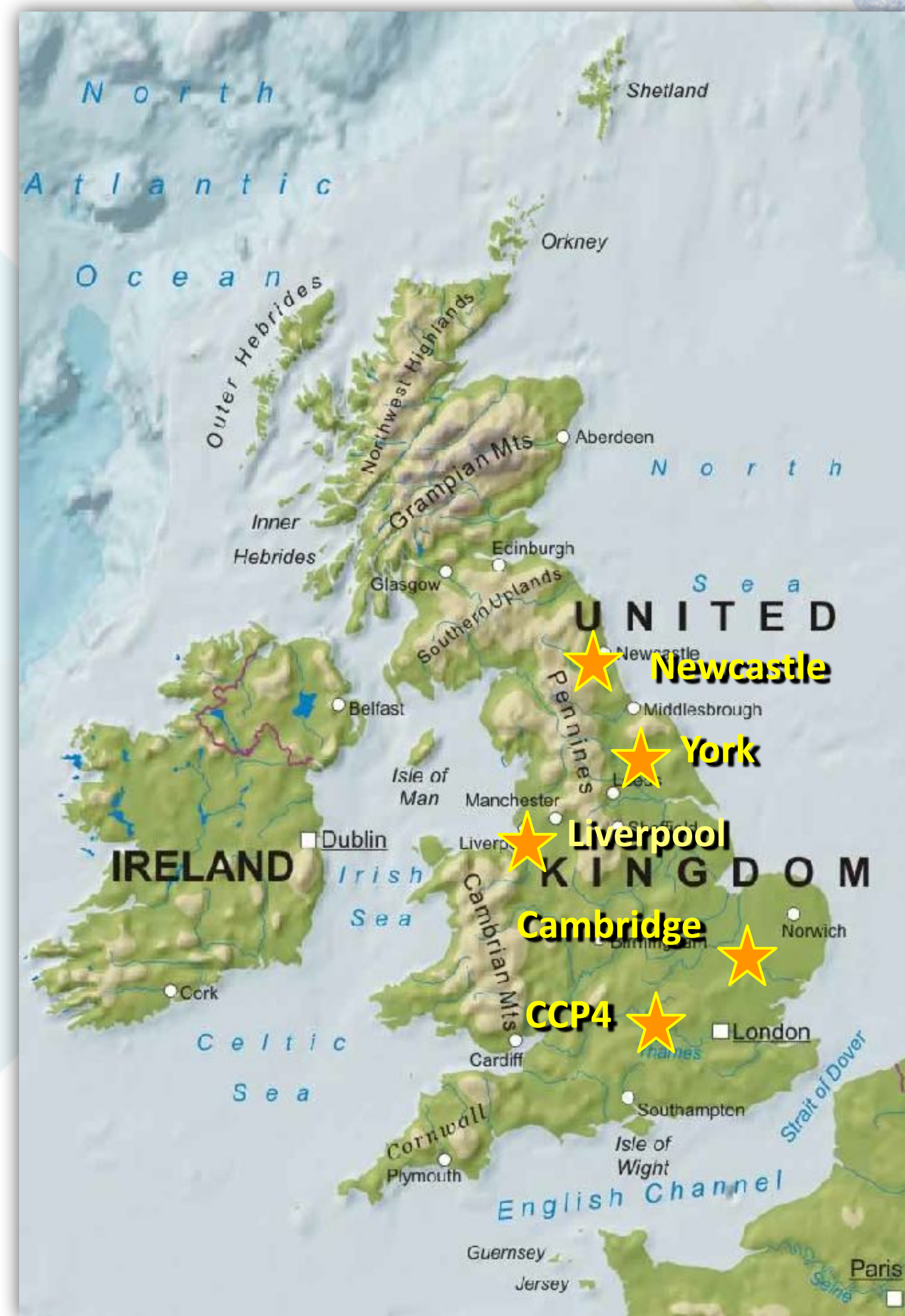
● Phaser group (University of Cambridge)

● University of York, Newcastle University

- ~ Software developments (CCP4mg, Modelcraft, Privateer)
- ~ CCP4 GUI-2 developments

● Associated projects

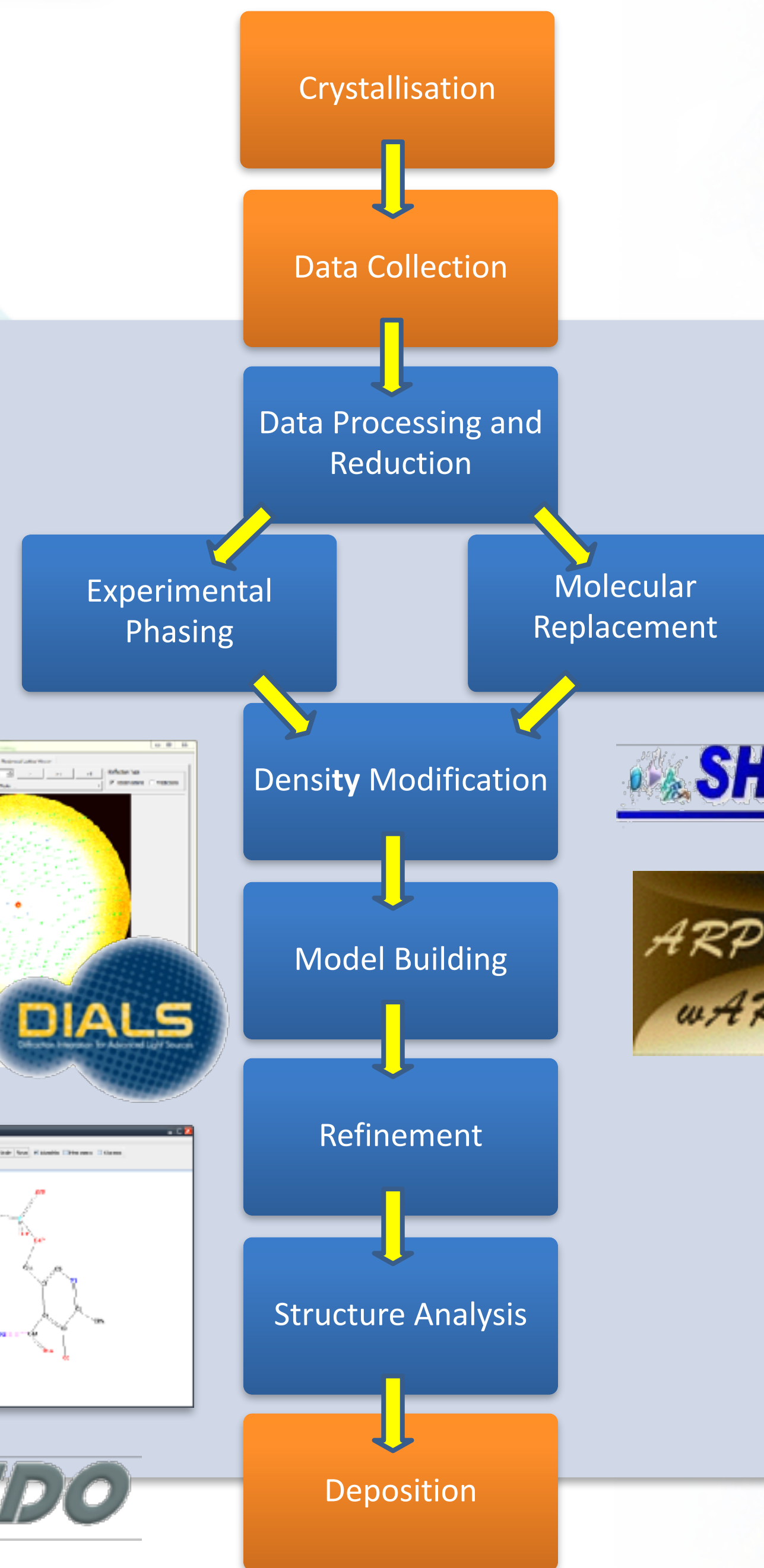
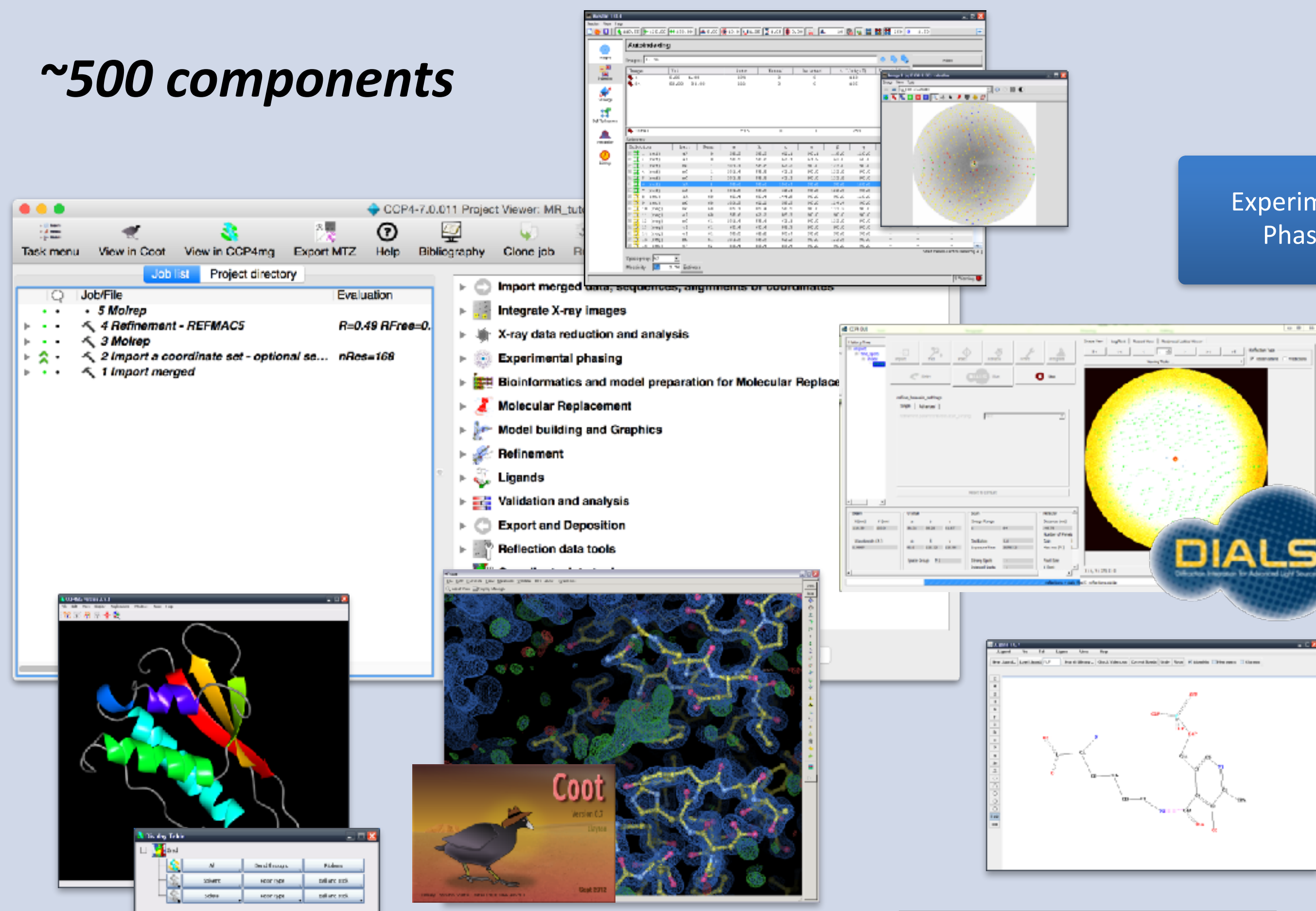
- ~ Experimental Phasing from **Leiden** (Crank-2)
- ~ Arcimboldo MR Software (**Barcelona**)
- ~ Ab-initio and automatic MR, University of **Liverpool** (AMPLE)
- ~ Arp/wArp (EMBL-**Hamburg**)
- ~ SHELX (**Goettingen** and **Barcelona**)
- ~ PDB-Redo (NKI, the **Netherlands**)





Probably the largest MX Suite

~500 components



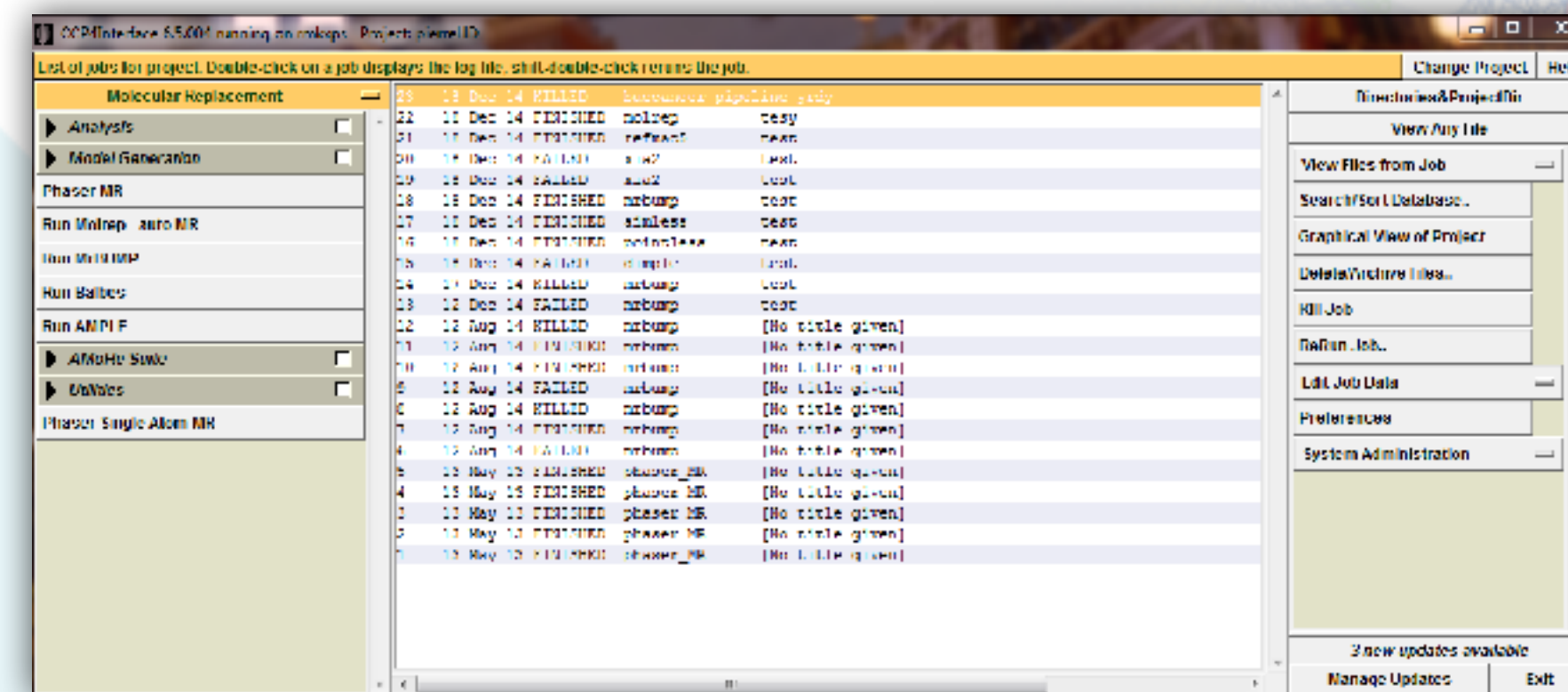
SHELX

7.3
ARP
wARP

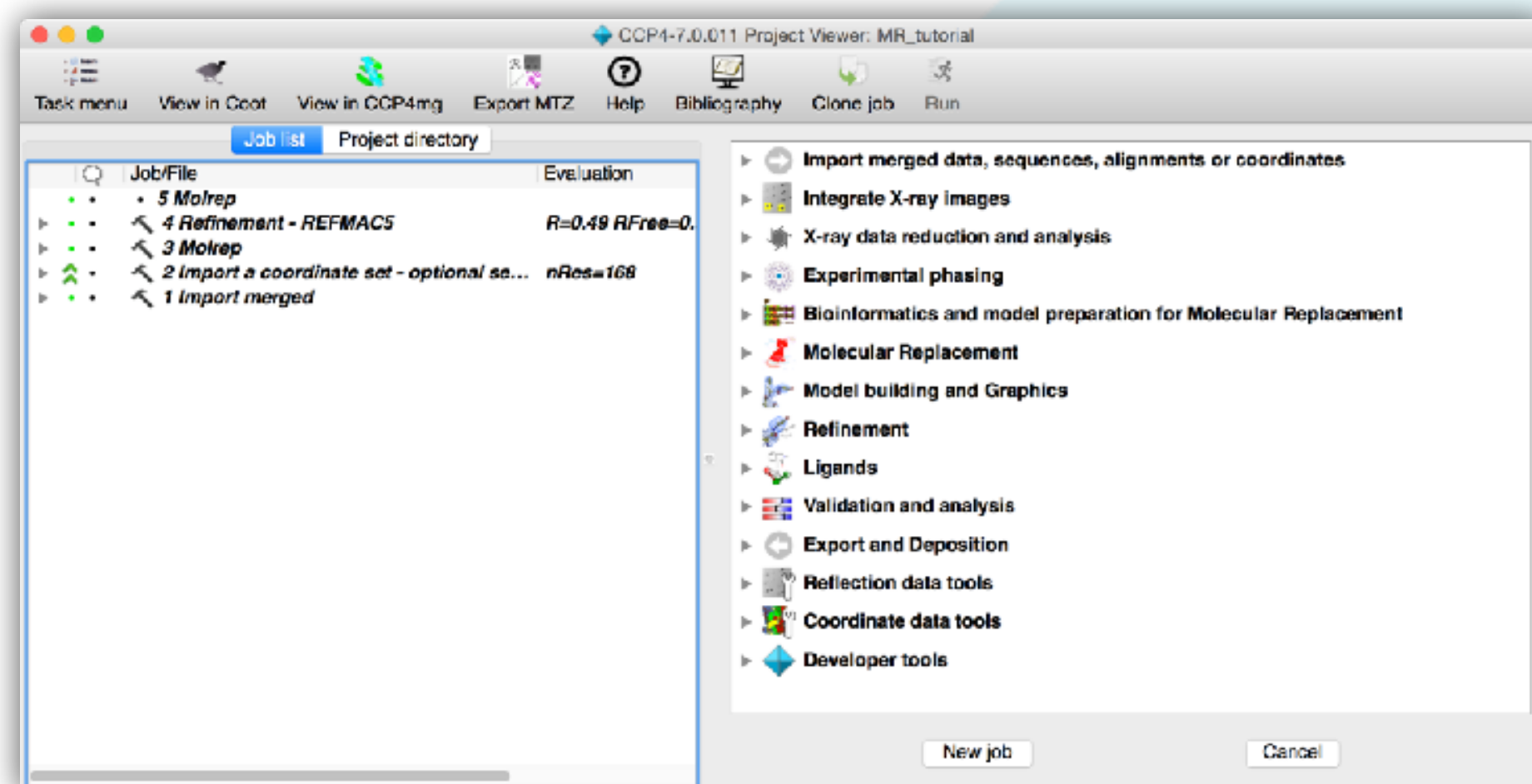
PDB_REDO

Three interfaces to the CCP4 Software Suite

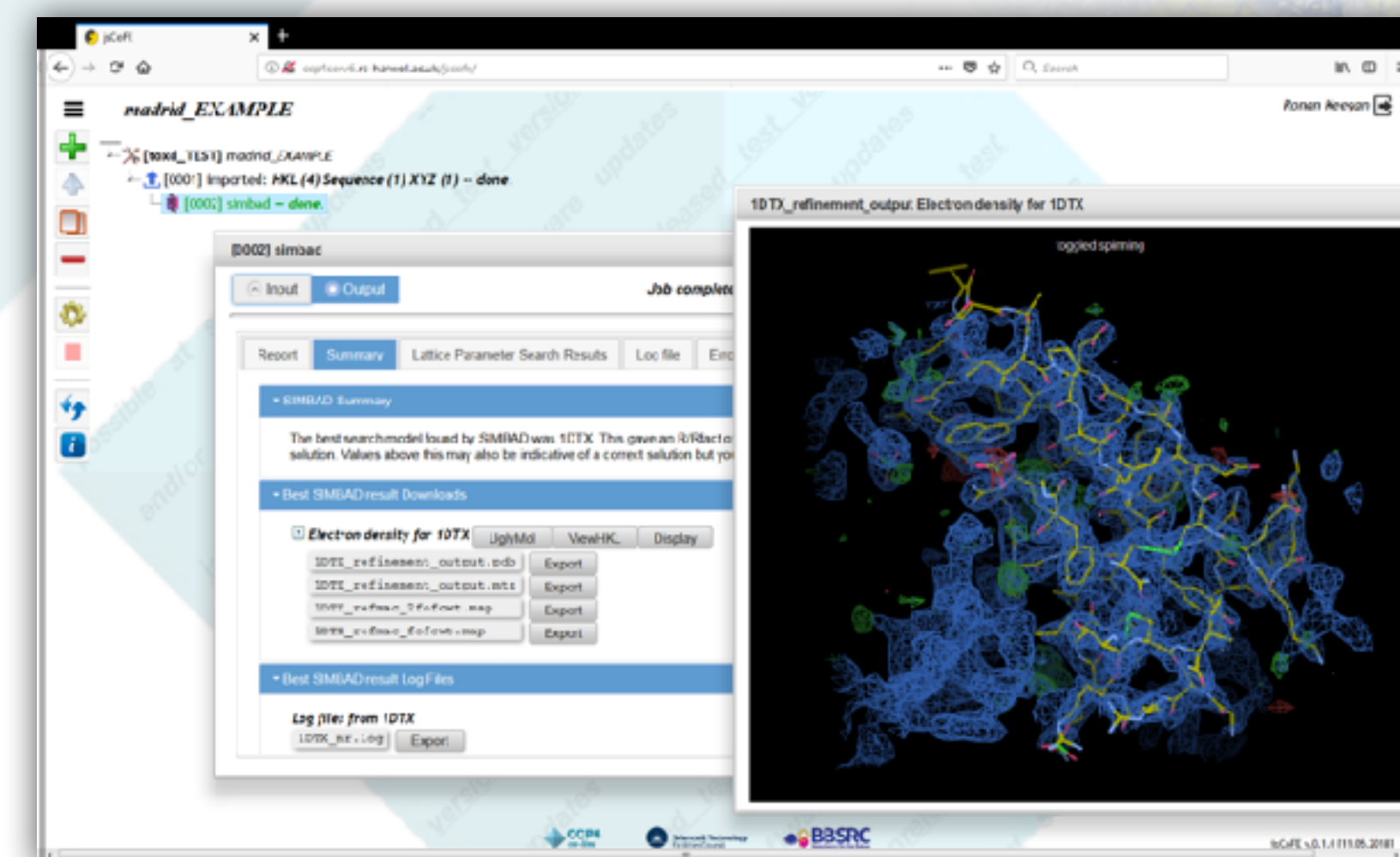
1. **CCP4i** – original interface developed around 2000 - *discontinued*



2. **CCP4i2** – new graphical desktop interface (2016)



3. **CCP4 Cloud**, browser based (2019)



Growing software complexity

- ~ Supporting wide variety of computing platforms is difficult
- ~ Full installation with 3rd party databases and software is difficult

Growing software demands

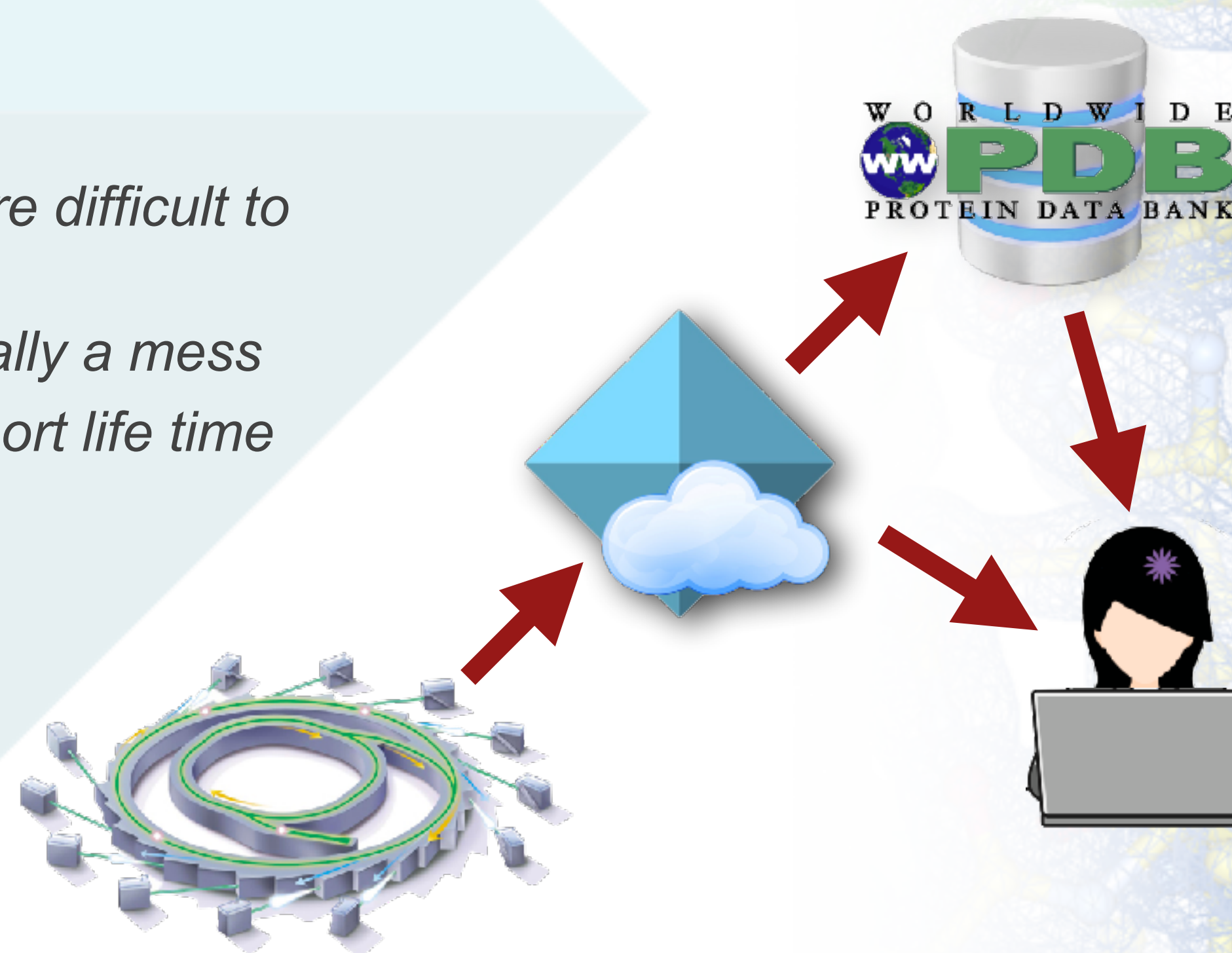
- ~ Modern automatic methods require more CPU and memory than most local setups can afford

Data logistics and team working

- ~ Growing volumes of data from modern sources are difficult to handle locally
- ~ File exchange in distributed collaborations is usually a mess
- ~ Local data and project archives have usually a short life time

Security

- ~ Increasingly more difficult to distribute for modern systems and corporate environments
- ~ Cloud solutions are safer, getting preferential in industry



A system for solving X-ray structures online using common web-browsers

Main features

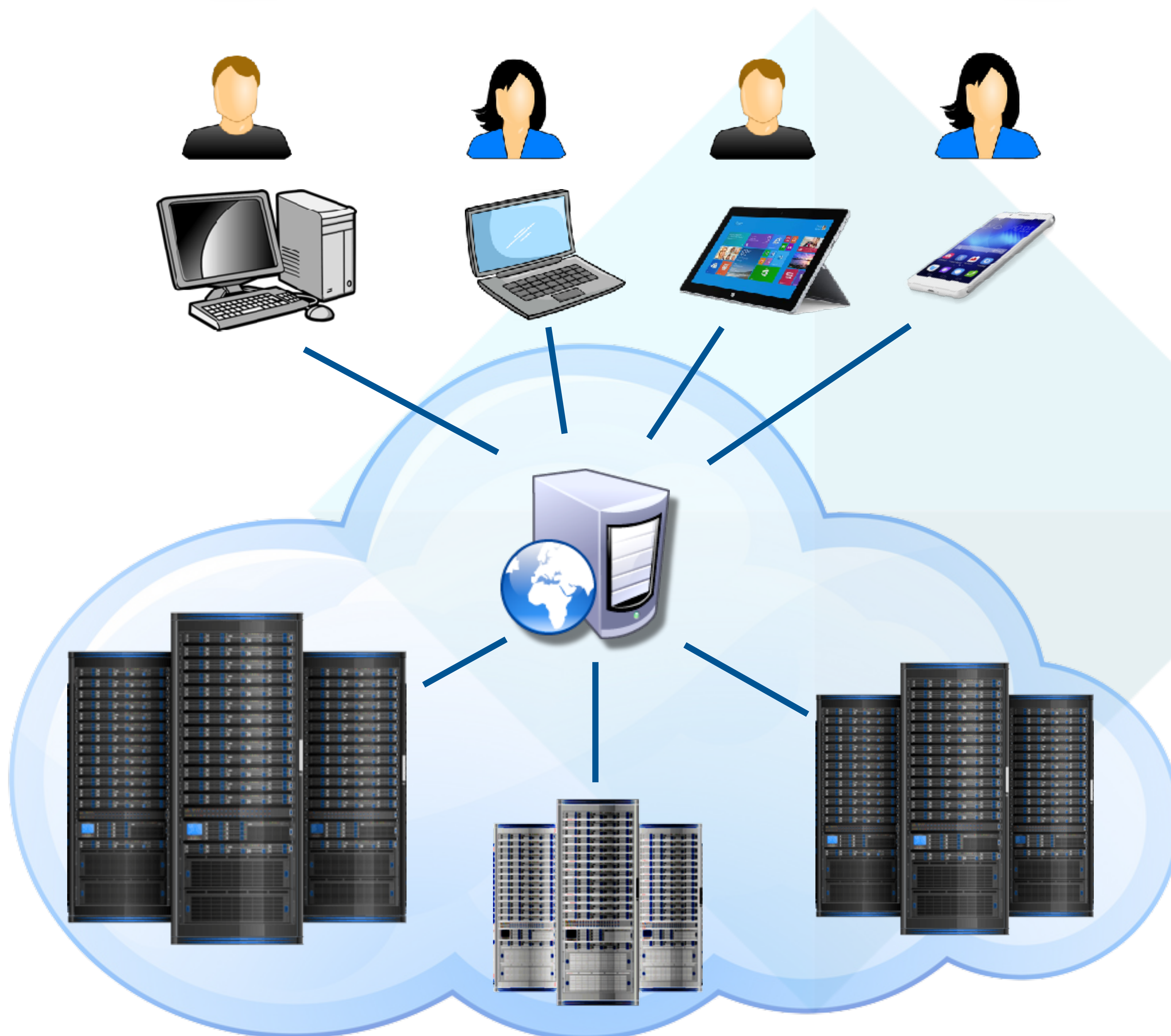
- ~ Software, resources and data as a **service**: go-and-use
- ~ **Cross-platform**: can be used on all Windows, Linux, Mac OSX, tablets and smartphones
- ~ Rich **project development** functionality
- ~ **All stages** of structures solution: from image processing to PDB deposition
- ~ Integrates access to **web-resources** such as PDB and AFDB
- ~ Facilitate **team work** by sharing projects in real time
- ~ Can be run **locally**
- ~ Can be installed in a **lab, institute or firm**
- ~ Highly configurable and adaptable to using **mixed distributed computational resources**
- ~ Integrated documentation and tutorials

Current limitation

- ~ Local installation of CCP4 is required for using Coot and some other desktop applications

Availability

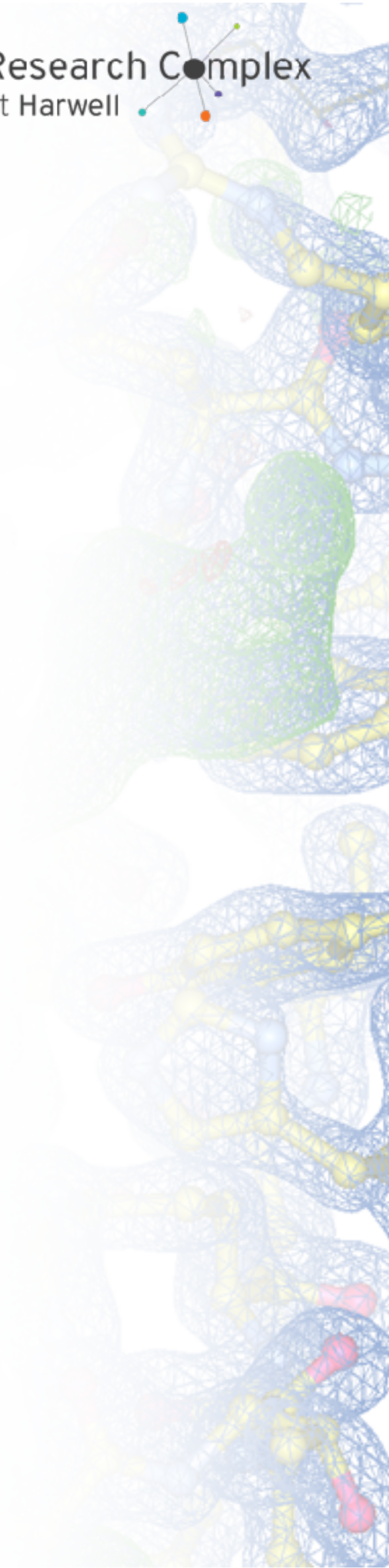
- ~ Publicly available at <https://cloud.ccp4.ac.uk>
- ~ Part of standard CCP4 distribution package starting from 7.1 release series (since June 2020)



*work anywhere
with anything*

*get power from
the Cloud*

Look around CCP4 Cloud



Start right!

CCP4 Cloud icons in CCP4 7.1+ setup



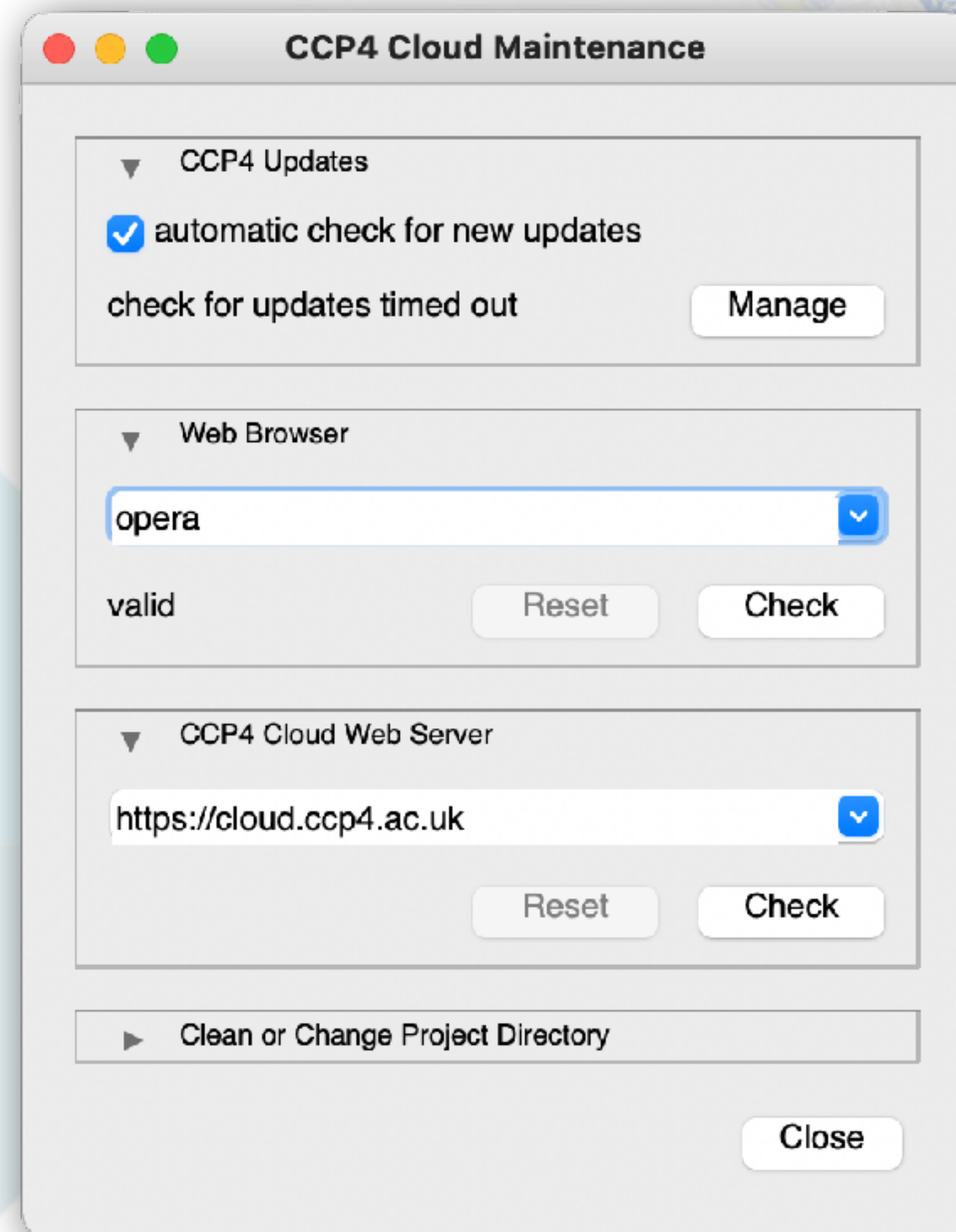
CCP4 Cloud
on your machine
(no internet)



CCP4 Cloud
on remote servers



CCP4 Cloud
configurator/
maintainer



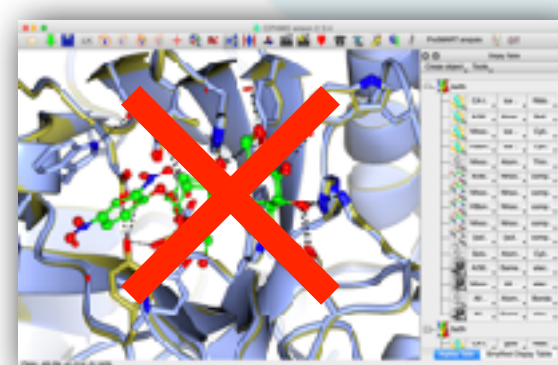
If started from direct URL <https://cloud.ccp4.ac.uk>:



No Coot



No DIALS



No CCP4mg

Only direct URL
(no local CCP4)



CCP4 Cloud Login

 CCP4-Harwell

<https://cloud.ccp4.ac.uk/>

Login name: `sbgrid-demo`

Password:



Login



Forgotten password



Registration

Check CCP4 Cloud [roadmap](#) for new users

[Privacy Statement](#)



 Did you know? Crashed Coot task does not mean that your work is lost! Read [here](#).

CCP4 Cloud

localhost:51253

SBGrid Demo

0%0%:0%

My Projects

Open

Add

Rename

Delete

Export

Import

Join

Tutorials

Help

ID	Name	R _{free}	Disk (MBytes)	CPU (hours)	Date Created	Last Opened
Your List of Projects is currently empty. Use "Add" button to create a new Project; "Import" button for importing a project exported from CCP4 Cloud; "Join" button for joining project shared with you by another user; or "Tutorials" button for loading tutorial/demo projects.						

Powered by CCP4 v.7.1.018

CCP4 on-line

UK

iris

CCP4 Cloud v.1.7.001 [26.11.2021]

CCP4 Cloud

localhost:51253

My Projects

SBGrid Demo

My Account

Toggle fullscreen

Log out

Add

Rename

Delete

Export

Import

Join

Tutorials

Help

	R _{free}	Disk (MBytes)	CPU (hours)	Date Created	Last Opened
--	-------------------	---------------	-------------	--------------	-------------

Your List of Projects is currently empty.

Use "Add" button to create a new Project;
"Import" button for importing a project exported from CCP4 Cloud;
"Join" button for joining project shared with you by another user;
or "Tutorials" button for loading tutorial/demo projects.

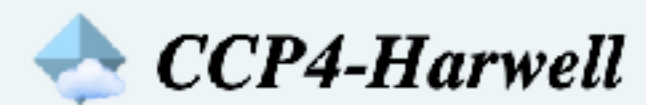
Powered by CCP4 v.7.1.018

CCP4 on-line

iris

CCP4 Cloud v.1.7.001 [26.11.2021]

My Account



https://cloud.ccp4.ac.uk/

User name:

SBGrid Demo

E-mail:

eugene.krissinel@stfc.ac.uk

Login name:

sbgrid-demo

Password:

password (old or new)

Retype password:

confirm password

Licence agreement:

academic

Feedback agreement:

accepted (1)

Software authorisations:

manage

Save changes

Delete my account

Preferences*

On login, go to:

list of projects

Initial size (% of browser window):

	width	height
Job dialogs:	125	85
Viewers:	140	97

☐ Use project name as default file prefix

☒ Guide through data import

☒ Send end-of-job notifications

for jobs taking longer than

24

hours

* Save changes for Preferences to take an effect

CCP4 Cloud

localhost:51253

SBGrid Demo

My Account

CCP4-Harwell

https://cloud.ccp4.ac.uk/

User name:

SBGrid Demo

E-mail:

eugene.krissinel@stfc.ac.uk

Login name:

sbgrid-demo

Password:

password (old or new)

Retype password:

confirm password

Licence agreement:

academic

Feedback agreement:

accepted (1)

Software authorisations:

manage

Save changes

Delete my account

Preferences*

Software Authorisations

CCP4 Cloud provides access to 3rd party software, not covered by CCP4 license. In order to use this software, a user needs to be authorised by the corresponding software provider. This page gives a summary of your authorisations and allows you to send authorisation requests. Note that authorisation request will redirect you to the provider's web-site. Any data requested by software provider will be processed at provider's end and will not be stored in CCP4 Cloud.

Software:

GΦL

Global Phasing Ltd. Software Suite

Provider:

GΦL

Global Phasing Ltd.

Authorisation:

not authorised

request authorisation

Close

CCP4 on-line

UK

iris

Powered by CCP4 v.7.1.018

CCP4 Cloud v.1.7.001 [26.11.2021]

CCP4 Cloud

localhost:51253

SBGrid Demo

0% 0%:0%

My Account



CCP4-Harwell

https://cloud.ccp4.ac.uk/

User name:

SBGrid Demo

E-mail:

eugene.krissinel@stfc.ac.uk

Login name:

sbgrid-demo

Password:

password (old or new)

Retype password:

confirm password

Licence agreement:



academic

Feedback agreement:



accepted (1)

Software authorisations:



manage



Save changes



Delete my account

Preferences*

On login, go to:

list of projects

Initial size (% of browser window):

	width	height
Job dialogs:	125	85
Viewers:	140	97

☐ Use project name as default file prefix

☒ Guide through data import

☒ Send end-of-job notifications

for jobs taking longer than

24

 hours

* Save changes for Preferences to take an effect

Resource	Used	Quota	%%
Storage (MBytes)	0	5000	0%
CPU 24h (hours)	0	24	0%
CPU 30d (hours)	0	240	0%

User stats:

Total storage used (MBytes) 0

Total jobs run 0

CPU lifetime used (hours) 0

Powered by CCP4 v.7.1.018



CCP4 Cloud v.1.7.001 [26.11.2021]

CCP4 Cloud

localhost:51253

SBGrid Demo

0%0%:0%

My Projects

Open

Add

Rename

Delete

Export

Import

Join

Tutorials

Help

ID	Name	R _{free}	Disk (MBytes)	CPU (hours)	Date Created	Last Opened
----	------	-------------------	---------------	-------------	--------------	-------------

Your List of Projects is currently empty.

Use "Add" button to create a new Project;
"Import" button for importing a project exported from CCP4 Cloud;
"Join" button for joining project shared with you by another user;
or "Tutorials" button for loading tutorial/demo projects.

Powered by CCP4 v.7.1.018

CCP4 on-line

UK

iris

CCP4 Cloud v.1.7.001 [26.11.2021]

CCP4 Cloud

localhost:51253

SBGrid Demo

My Projects

Open

Add

Rename

Delete

Export

Import

Join

Tutorials

Help

ID	Name	R _{free}	Disk (MBytes)	CPU (hours)	Date Created	Last Opened
----	------	-------------------	---------------	-------------	--------------	-------------

Add New Project

ID

mdm2

Name

MDM2 Protein with bound Nutlin Molecule

Initialisation mode

Autostart

Standard

Hop on

Simplified mode for starting with one of automatic workflows featuring structure solution scenarios in CCP4 Cloud, and optimised list of essential tasks for structure completion (can be switched to the full list).

Standard mode for operating all CCP4 Cloud tasks manually, with access to the full list of tasks by default.

Quick start from phased (possibly partially built and refined) structure or heavy-atom substructure, for further refinement and model building, with the optimised list of essential tasks for structure completion (can be switched to the full list).

Add Project

Cancel

Powered by CCP4 v.7.1.018

CCP4 on-line

iris

CCP4 Cloud v.1.7.001 [26.11.2021]

Demo Drive: Automatic Structure Solution



Why Automate?

Avoid the routine

- ~ *Similar targets → similar data → similar actions*

Save personal time

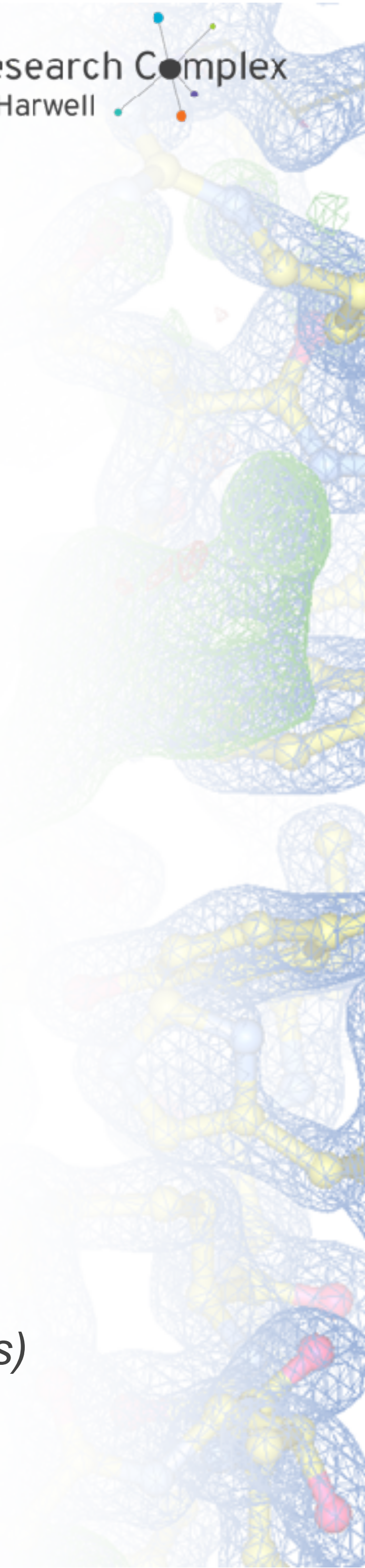
- ~ *“Start-and-forget”*

Mitigate incompetence

- ~ *Decision making in complex projects and workflows (**strategy**)*
- ~ *Choose the right tool for the task (**tactics**)*
- ~ *Cope with multiple options (**ambiguity**)*
- ~ *Choose right parameters (**expertise**)*

The result

- ~ *Works in standard cases (no crystal pathologies, good diffraction, close templates)*
- ~ *May give further ideas even if fails*



Automation Level 1: Pipelines (super-scripts)

Image Processing

xia2 Xia-2

Experimental Phasing

Crank2 Crank-2

Molecular Replacement

BALBES



MoRDa



MrBump



Simbad



Slicendice



Arcimboldo

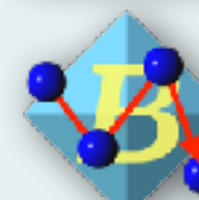
Model Building



Buccaneer



Arp/wArp



CCP4Build



Nautilus



Modelcraft

Refinement and Validation



LORESTR

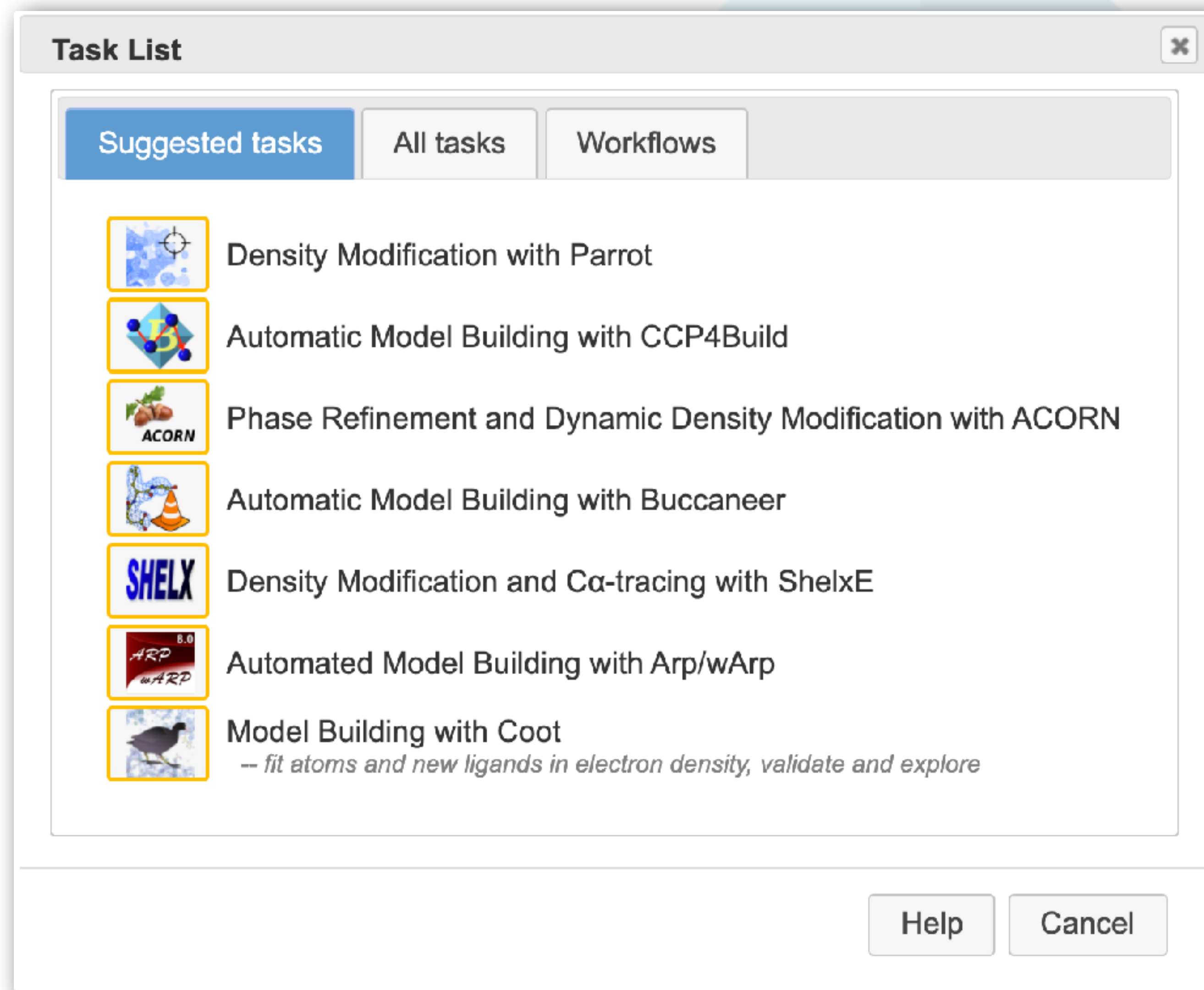


DIMPLE



Zanuda

Automation Level 2: Project Navigation



Intelligent next step suggestions based on in-built logics of project development and analysis of user actions

CCP4 Cloud learns your habits as you go ...

Automation Level 3: Project Development

Automatic development of
structure solution projects
based on available data
and their properties

Task List

Workflows

Essential tasks

Automatic Workflows

Each workflow will run a series of tasks, see details [here](#).

Workflows for starting a Project

**aMR** Automatic Molecular Replacement with MrBump or MrDa
-- data import, ASU definition, automatic MR, refinement, ligand fitting and PDB deposition

**sMR** Simple Molecular Replacement with Search Model
-- data import, ASU definition, phaser MR, refinement, ligand fitting and PDB deposition

**aEP** Automated Experimental Phasing with Crank-2
-- data import, ASU definition, automatic EP, refinement, ligand fitting and PDB deposition

**aDPL** Dimple Refinement and Ligand Fitting
-- data import, Dimple pipeline, ligand fitting, refinement and PDB deposition

Workflows for using within a Project

**aREL** Automated refinement and ligand fitting
-- refinement, ligand fitting and PDB deposition, starting from a phased structure

Autostart mode

Help

Cancel

[gamma] Gamma

```
graph TD
    A["simple-MR:[0001] Simple MR workflow -- imported HKL (1), Sequences (1), XYZ"] --> B["simple-MR:[0002] asymmetric unit contents -- 1 molecule in ASU, Solv=41.7"]
    B --> C["simple-MR:[0003] prepare MR model(s) from xyz -- 1 model(s) generated"]
    C --> D["simple-MR:[0004] phaser MR -- Nsol=1 LLG=5030 TFZ=46.4 R=0.2755"]
    D --> E["simple-MR:[0005] buccaneer -- Compl=100.0% R=0.2669 Rfree=0.3"]
    E --> F["simple-MR:[0006] arpwrap -- R=0.2234 Rfree=0.257"]
    F --> G["xyz simple-MR:[0007] xyz utils -- waters removed"]
    G --> H["simple-MR:[0008] refmac5 -- R=0.2642 Rfree=0.3170"]
    H --> I["simple-MR:[0009] fit waters -- Nwaters=65"]
    I --> J["simple-MR:[0010] refmac5 -- R=0.2369 Rfree=0.301"]
    J --> K["simple-MR:[0011] refmac5 -- R=0.2378 Rfree=0.300"]
    K --> L["simple-MR:[0012] refmac5 -- R=0.2331 Rfree=0.299"]
    L --> M["simple-MR:[0013] deposition -- package prepared"]
    M --> N["simple-MR:[0014] Automated Workflow"]
```

CCP4 Cloud

localhost:51253

SBGrid Demo

My Projects

0%5%:1%

OpenAddRenameDeleteExportImportJoinTutorialsHelp

ID	Name	R _{free}	Disk (MBytes)	CPU (hours)	Date Created	Last Opened
----	------	-------------------	---------------	-------------	--------------	-------------

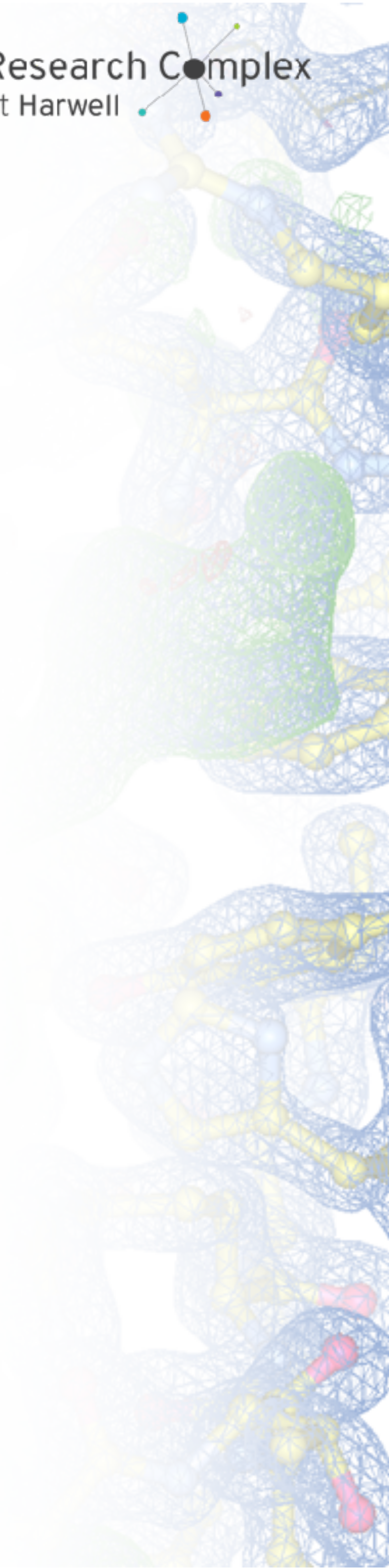
Powered by CCP4 v.7.1.018

CCP4 on-line

iris

CCP4 Cloud v.1.7.001 [26.11.2021]

Introduction in CCP4 Cloud Projects





MDM2 Protein bound to Nutlin 3A



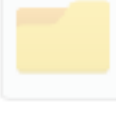
Add new job



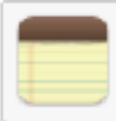
Clone selected job



Delete selected job



Collapse long branch in folder



Add remark



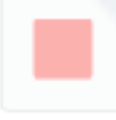
Highlight current branch



Toggle multiple job selection



Open job



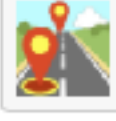
Stop job



Reload project



Help



Structure solution roadmap

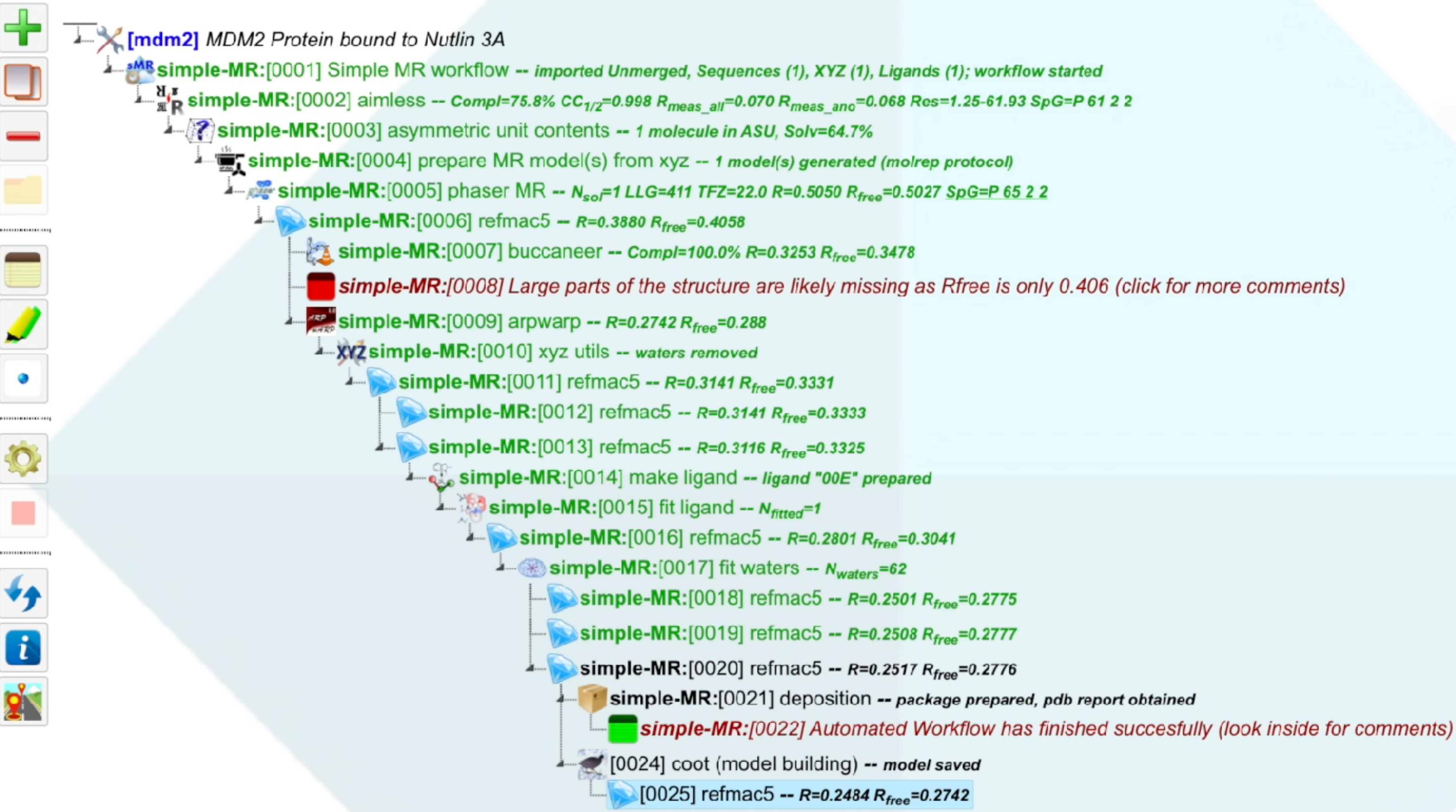
Data transfer

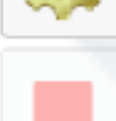
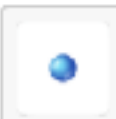
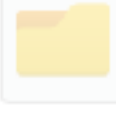
Branching





≡ *MDM2 Protein bound to Nutlin 3A*



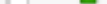
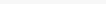
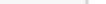
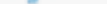
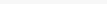
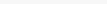
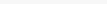


MDM2 Protein bound to Nutlin 3A



My Projects

SBGrid Demo

 Open
  Add
  Rename
  Delete
  Export
  Import
  Join
  Tutorials
  Help

ID	Name	R _{free}	Disk (MBytes)	CPU (hours)	Date Created	Last Opened
mdm2	MDM2 Protein bound to Nutlin 3A	0.2556	380.4	2.4482	2022-01-21	2022-01-25

- Open
- Rename
- Delete
- Export
- Share

Share Project

×

Share Project "mdm2"

The following users:

login1,login2,...

can join this project and work on it simultaneously.

* Full (comma-separated) list of users with access to the project must be given. In order to unshare project with a user, remove their login name from the list.

Apply

Cancel

CCP4 Cloud

localhost:51253

My Projects

73%0%:0%

Eugene Krissinel

OpenAddRenameDeleteExportImportJoinTutorialsHelp

ID	Name	R _{free}	Disk (MBytes)	CPU (hours)	Date Created	Last Opened
mdm2-1	MDM2	0.287	114.7	0.1252	2022-01-15	2022-01-24
cd44	CD44: Auto-EP with Crank-2		0	0	2022-01-12	2022-01-12
[sw-2022]:gamma-af	Solving Gamma protein with alphafold models	0.993	125.5	1.7625	2022-01-06	2022-01-19
[paul.bond]:bugs	Testing for bugs		26.2	0.2239	2021-12-10	2022-01-04
[maximeoef]:project_Jari	JW05	0.5625	740	74.912	2021-12-09	2021-12-13
rnase-script	DIMPLE from script			0.3548	2021-11-24	2021-11-26
gamma	Gamma			0.9806	2021-11-18	2022-01-20
docs	Documentation			0	2021-11-16	2021-11-16
T2_001	Creating Molecular Replacement			0.0005	2021-11-01	2022-01-12
[fabl.sm]:hGST	hGST		33.7089		2021-09-28	2021-10-13
[Karthi]:006	ROQ_112		6.783		2021-09-27	2021-11-12
CatV-PCI3-1	CatV-PCI3 Rescued			0.0381	2021-09-23	2021-09-24
[ed.lowe]:cd44-rec	CD44 recovery project			0.1015	2021-09-09	2022-01-12
MR-prac	Molecular Replacement Practical			0.0043	2021-08-20	2021-10-03
test	Installation Test			13.8841	2021-08-16	2022-01-18
[nicholls]:HemeTest	HemeTest			0.0761	2021-08-10	2022-01-21
[Afshan]:MBRWT	2.37_AUTOSELECT	0.1984	9639.9	16.2544	2021-08-10	2022-01-12
bl	Beta-lactamase		127.8	0.295	2021-07-28	2021-08-24
gdemo	Gamma demo to check parallel workflowing	0.2452	711.5	6.2086	2021-07-13	2021-11-15
[bwithd]:Demo	Demo	0.241	570.2	2.8235	2021-07-11	2021-12-16
[andrey_aps]:mdm2-2	mdm2	0.2915	296	1.4473	2021-07-02	2021-09-22
masha	Masha's data	0.5164	3432.7	164.3537	2021-06-24	2021-07-24
pisa	PISA		64.6	0.0062	2021-06-24	2021-06-24
SodM_Fraglites_060521	SodM_Fraglites_060521 initiated by script		10.9	0.0006	2021-06-01	2022-01-11

Join Shared Project

Join Shared Project

Select a project to join:

sbgrid-demo:[mdm2] "MDM2 Protein bound to Nutlin 3A"

Join

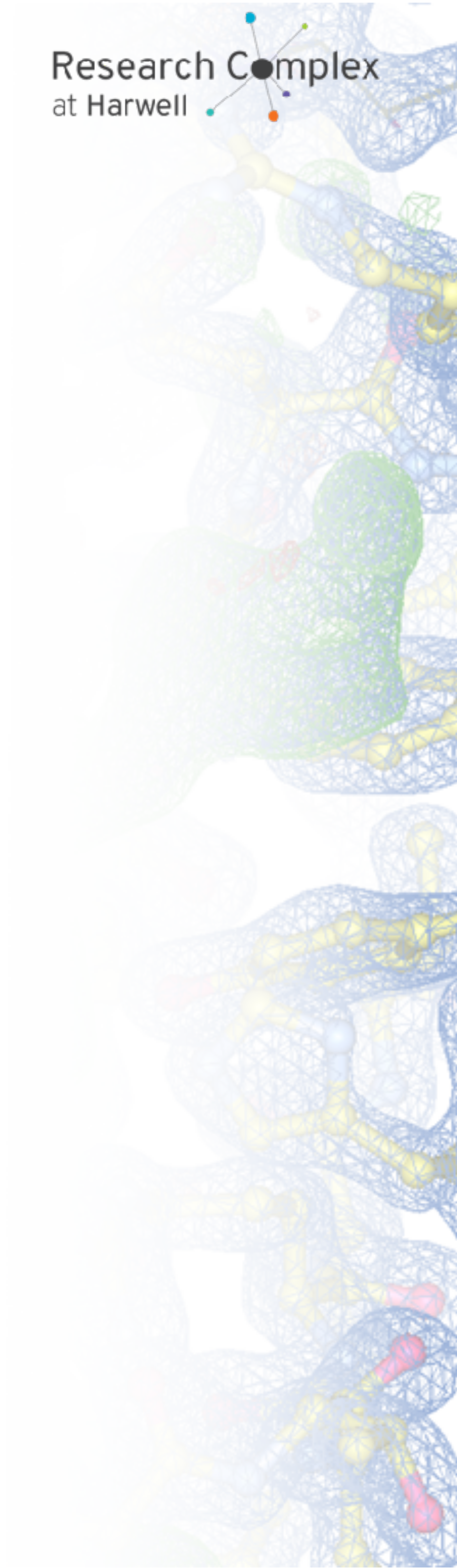
Cancel

CCP4 on-line

iris

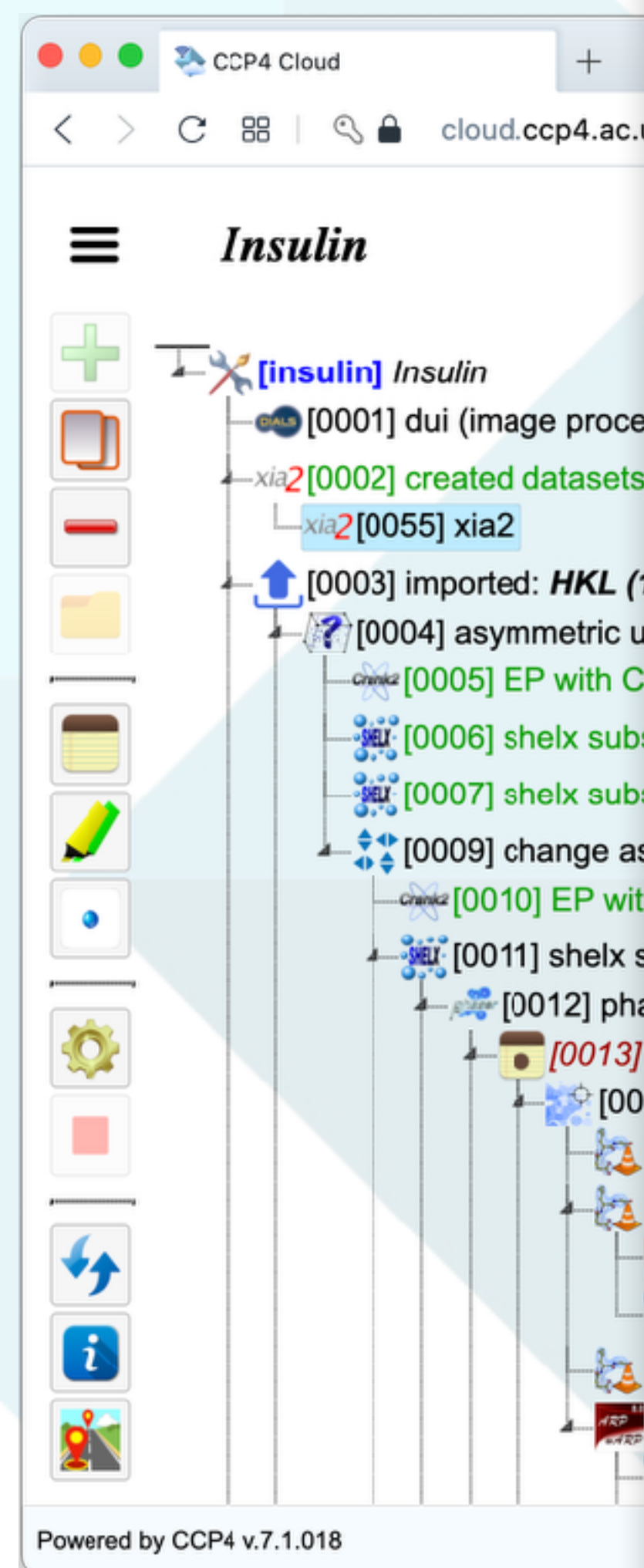
CCP4 Cloud v.1.7.001 [26.11.2021]

Final Remarks



Scope of Tasks

- ✓ Image Processing
- ✓ Data Scaling and Merging
- ✓ Molecular Replacement
- ✓ Experimental Phasing
- ✓ Density Modification
- ✓ Model Building
- ✓ Ligand fitting
- ✓ Refinement
- ✓ Validation
- ✓ PDB Deposition



Task List

Suggested tasks

All tasks

Workflows

▸ Data Import (5)

▸ Data Processing (4)

▸ Asymmetric Unit and Structure Revision (0)

▸ Automated Molecular Replacement (1)

▸ Molecular Replacement (0)

▸ Fragment-Based Molecular Replacement (0)

▸ Experimental Phasing (0)

▸ Density Modification (0)

▸ Refinement and Model Building (0)

▸ Coot (0)

▸ Ligands (1)

▸ Validation, Analysis and Deposition (0)

▸ Toolbox (1)

Total 75 tasks
and counting

Help

Cancel

Development is largely finished, but not completed ...



- new tasks are being added on demand
- most of functionality is developed as addition to existing tasks, e.g.,
 - *no task for Patterson maps - they are calculated as part of data import*
 - *no task for peak finding - done automatically at every refinement*
- the overall scope is equivalent to desktop solutions



~



Online Help



11.1. REFMAC5

**Early stages (e.g. straight after MR)**

- Run many cycles (up to 200 with jelly body restraints) of refinement

Medium stages – during model building

- 10-20 cycles
- Optimise additional parameters like NCS, TLS, etc
- Once your Rfree is lower than 40%, turn on twinning if twinned

 Warning

DON'T USE TWIN REFINEMENT IF YOUR Rfree IS HIGHER THAN 40%!

- Play with geometry weight parameter

Final stages of refinement

Back

Forward

Return

Detach

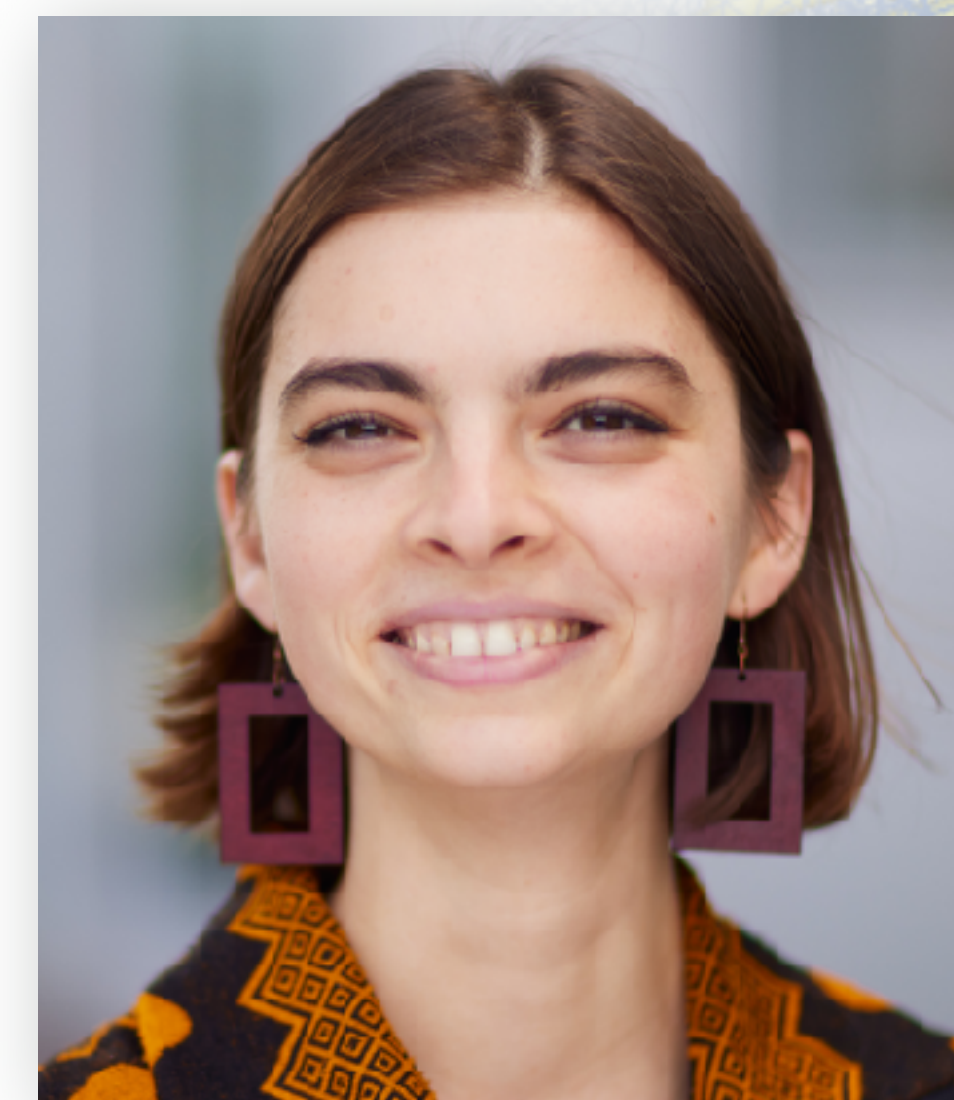
Close

- Task Reference and User's Guide
- Integrated
- Cross-linked
- Demo projects and data
- **Tutorials**
- **Designed for self-learning and using at Schools, Workshops and Courses**
- Open for additions and corrections



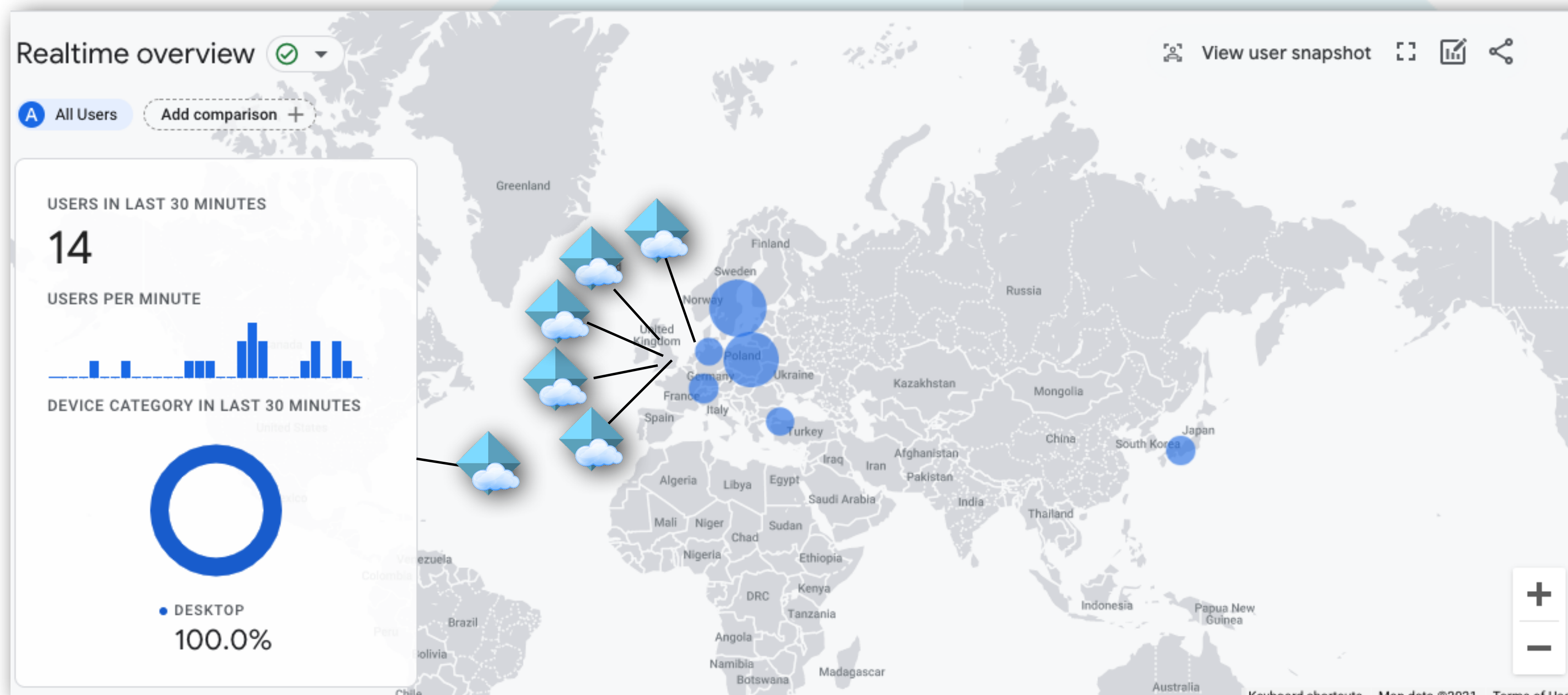
GitLab

*Contribution by
Maria Fando*



CCP4 Clouds are many

- Main CCP4 Cloud instance at CCP4-Harwell from 2018:
 - Over 3,000 user accounts
 - Over 70,000 jobs/year
- 5 CCP4 Cloud instances at partner sites, including industrial sector



RAL-Harwell

EMBL-Hamburg

Crick-London

Uni Newcastle

Uni Exeter

Pharma USA

To Cloud or Not To Cloud?

Why to Cloud? Easy:

Megaflops, Gigabytes, nothing to setup (go-and-use),
no maintenance, data backup,
24/7 from any device and anywhere! Forever!

* and **free** for you (CCP4 Cloud only)

* and it is **green** (save our environment!)



Today, it is trickier to say why not ...

Many thanks to:

CCP4, STFC & RCaH

Fantastic work environment, support and dissemination



**CCP4 Collaboration,
CCP4 School hosts and
CCP4 developers**

Contribution of task reports, tutorials, general support and valuable feedback



**~300 test users
Worldwide**

Trial use and feedback on development versions of CCP4 Cloud



Andy Purkiss

Francis Crick Institute, London

Grzegorz Chojnowski

EMBL-Hamburg

Arnaud Basle

Newcastle University

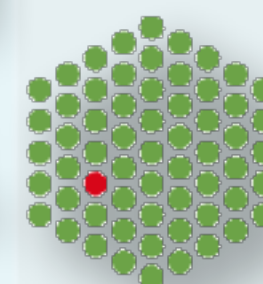
Michael Isupov

University of Exeter

Setup and maintenance of CCP4 Cloud instances in their home labs



EMBL



European Molecular
Biology Laboratory

**Biotechnology and Biological
Sciences Research Council
(BBSRC) UK**

*Research grant
BB/L0070317/1
(2014-2019)*



Developers and contributors

STFC, CCP4, Harwell, UK:

**Andrey Lebedev, Oleg Kovalevskiy,
Charles Ballard, Ville Uski,
Ronan Keegan**

RAS, Puschino, Russia:

Maria Fando

MRC/LMB, Cambridge, UK:

Robert Nicholls

EMBL-EBI, Hinxton, UK:

John Berrisford

Uni Leiden, The Netherlands:

Navraj Pannu, Pavol Skubak

Global Phasing Ltd, Cambridge, UK:

Marcin Wojdyr, Clemens Vonnrhein

Uni York, UK:

Stuart McNicholas

Uni Liverpool, UK:

Adam Simpkin, Jens Thomas

Uni Birmingham, UK:

Christopher Oliver