CCDC

advancing structural science

What's Up

Customer Update Webinar

23rd July 2020

CCDC

advancing structural science

Today’s presenters

Suzanna Ward

Head of Database

Clare Tovee

Senior Scientific Editor

Peter Wood

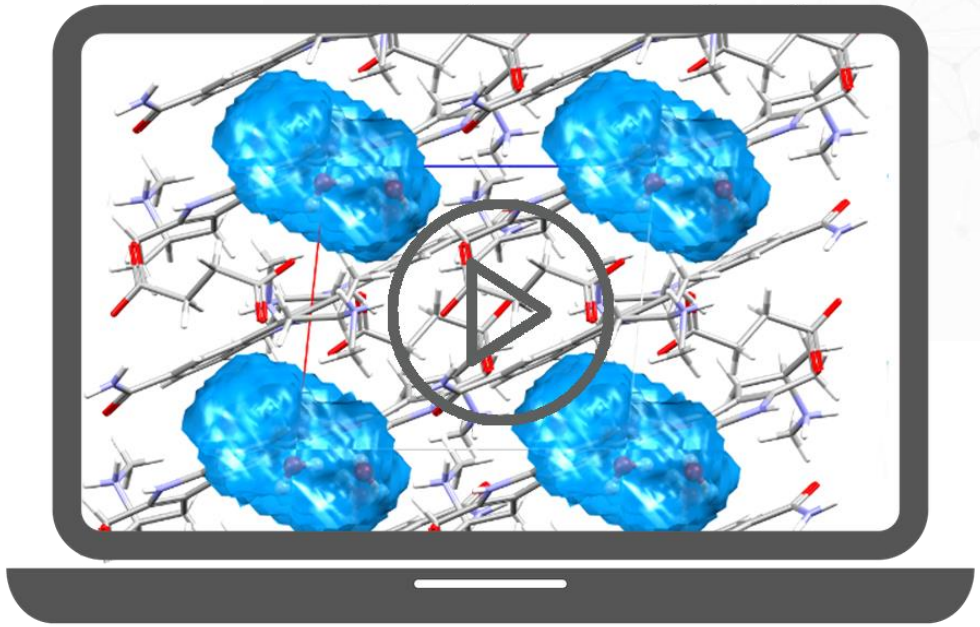
Senior Product Manager

CCDC

Overview

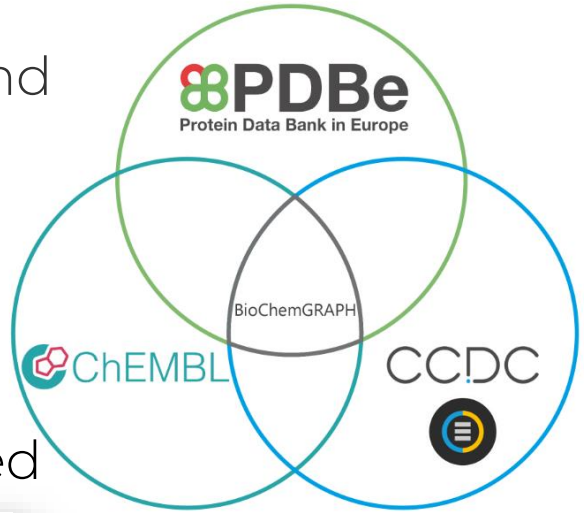
In this webinar we will discuss:

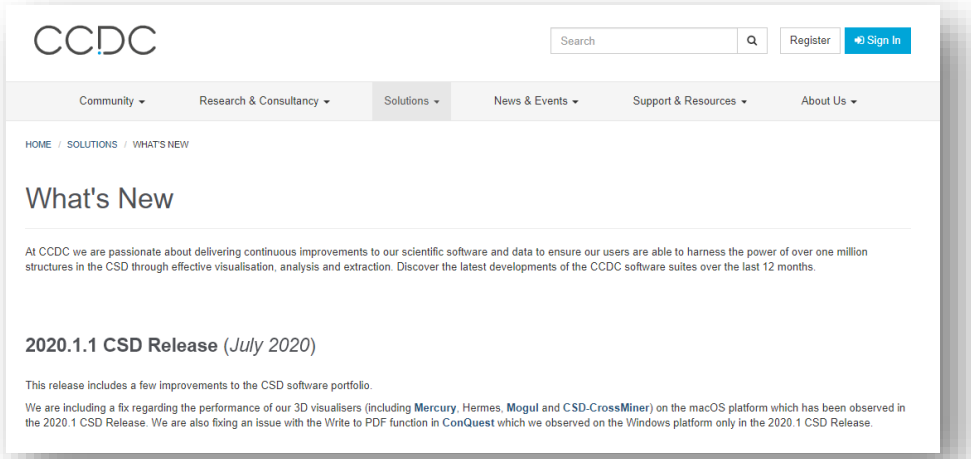
- Latest updates and news
- Behind the scenes of CSD data releases
- CSD Pipeline Pilot component collection
- Q&A: the floor is yours



Latest updates and news from CCDC

- **BioChemGRAPH** collaboration between PDBe, ChEMBL and CCDC
- **2020.1.1 CSD Release** is available
- **CSD Licence academic renewals** – contact admin@ccdc.cam.ac.uk by 30th September for a discounted renewal.




CCDC

Search | Register | Sign In

Community | Research & Consultancy | Solutions | News & Events | Support & Resources | About Us

HOME / SOLUTIONS / WHAT'S NEW

What's New

At CCDC we are passionate about delivering continuous improvements to our scientific software and data to ensure our users are able to harness the power of over one million structures in the CSD through effective visualisation, analysis and extraction. Discover the latest developments of the CCDC software suites over the last 12 months.

2020.1.1 CSD Release (*July 2020*)

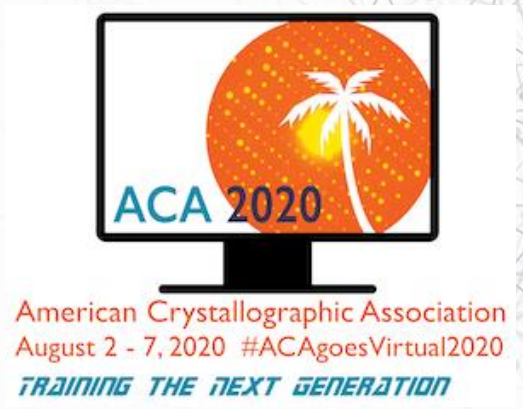
This release includes a few improvements to the CSD software portfolio.

We are including a fix regarding the performance of our 3D visualisers (including **Mercury**, Hermes, **Mogul** and **CSD-CrossMiner**) on the macOS platform which has been observed in the 2020.1 CSD Release. We are also fixing an issue with the Write to PDF function in **ConQuest** which we observed on the Windows platform only in the 2020.1 CSD Release.

CCDC

Latest updates and news from CCDC >Events

- American Crystallographic Association – 2-7th August
  - 7th Aug 12:20 PM
  - Yinka Olatunji-Ojo: Translating hands-on activities to virtual resources for broader scientific engagement



- Fall ACS Meeting – 16th-20th August
  - On demand - CINF - Moving Chemistry from the Lab into the Open
  - Ian Bruno: Publishing crystal structure data – keeping up with the times

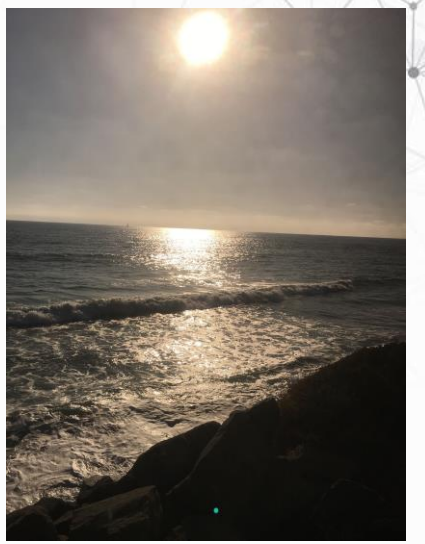
- E-Workshop: MX raw image data formats, metadata and validation – 22nd August 2020
  - Natalie Johnson: The Gold Standard for SCD - what's needed in addition to what we have

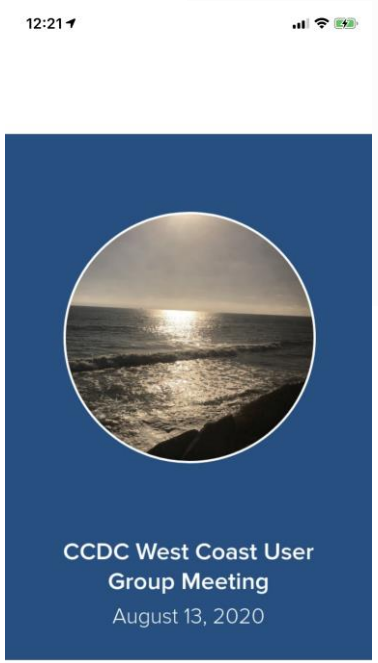
- Advancing the CCDC/FIZ Karlsruhe collaboration
  - Joint survey – launching soon
  - Joint webinar planned for 27th August

CCDC

Latest updates and news from CCDC >Events

- Virtual **West Coast UGM** – 13th August
  - Registration is open! https://www.ccdc.cam.ac.uk/News/Events
  - Download the event app



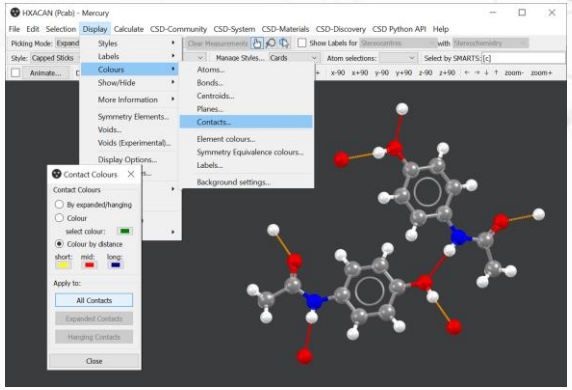


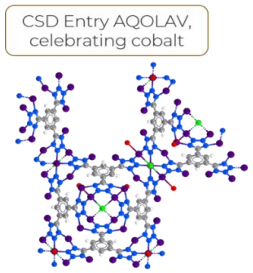
Latest updates and news from CCDC >Home learning

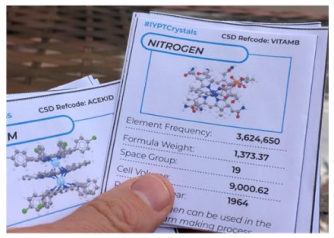
- New CCDC Home Learning Collection
  - IYPT in Crystals Battlecards
  - Lego, chocolate, polymorphs
  - Identikit of common substances

- Follow our new #TopTipTuesday

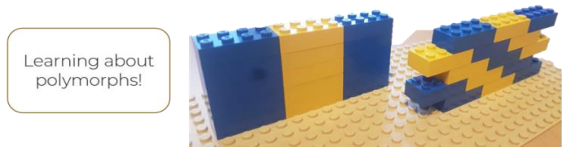
- Watch our new How To videos

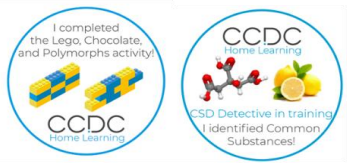


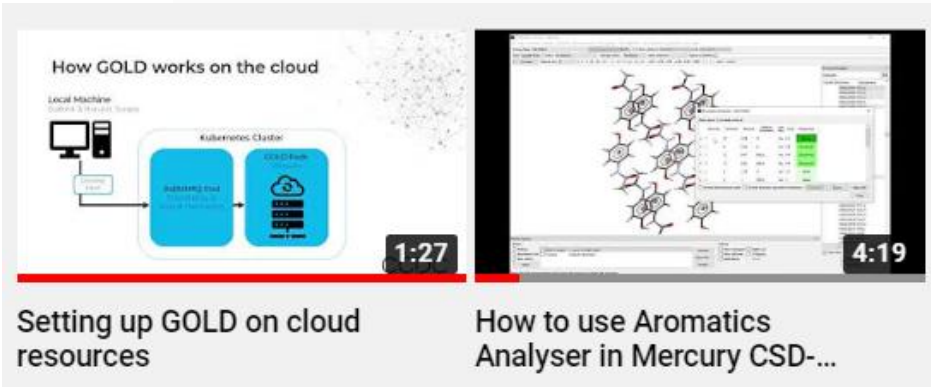




The IYPTCrystals Battlecards







CCDC

Latest updates and news from CCDC >CSD

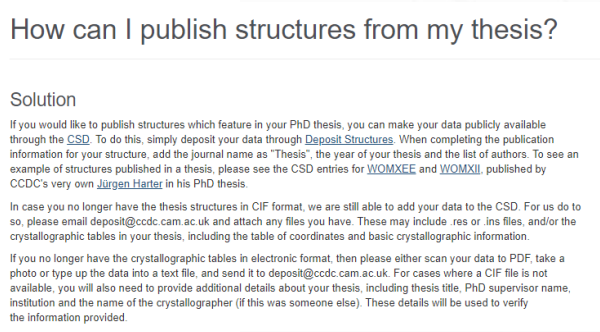
- Do you have unpublished data in:
  - a thesis
  - and/or non electronic format?

- We can help you share it!

- Last year Jake Bowden helped share >500 structures in this way

- This year our summer student, Cameron Wilson, is working to help add and convert your data

- Check out our FAQ or email deposit@ccdc.cam.a.uk



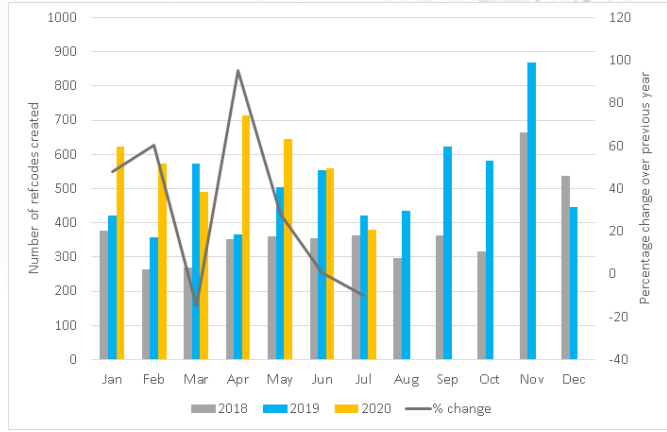
How can I publish structures from my thesis?

Solution

If you would like to publish structures which feature in your PhD thesis, you can make your data publicly available through the CSD. To do this, simply deposit your data through Deposit Structures. When completing the publication information for your structure, add the journal name as "Thesis", the year of your thesis and the list of authors. To see an example of structures published in a thesis, please see the CSD entries for WOMXEE and WOMXII, published by CCDC's very own Jürgen Harter in his PhD thesis.

In case you no longer have the thesis structures in CIF format, we are still able to add your data to the CSD. For us do to so, please email deposit@ccdc.cam.ac.uk and attach any files you have. These may include .res or .ins files, and/or the crystallographic tables in your thesis, including the table of coordinates and basic crystallographic information.

If you no longer have the crystallographic tables in electronic format, then please either scan your data to PDF, take a photo or type up the data into a text file, and send it to deposit@ccdc.cam.ac.uk. For cases where a CIF file is not available, you will also need to provide additional details about your thesis, including thesis title, PhD supervisor name, institution and the name of the crystallographer (if this was someone else). These details will be used to verify the information provided.



Thesis data summer project

We've been running a summer initiative this year to help you get your thesis data shared and acknowledged. Here's our summer student Jake Bowden helping CEO Jürgen Harter add structures from his PhD thesis. Discover how to publish structures from your thesis here.

CCDC

CSD Data Releases

Behind the scenes

Dr Clare Tovee

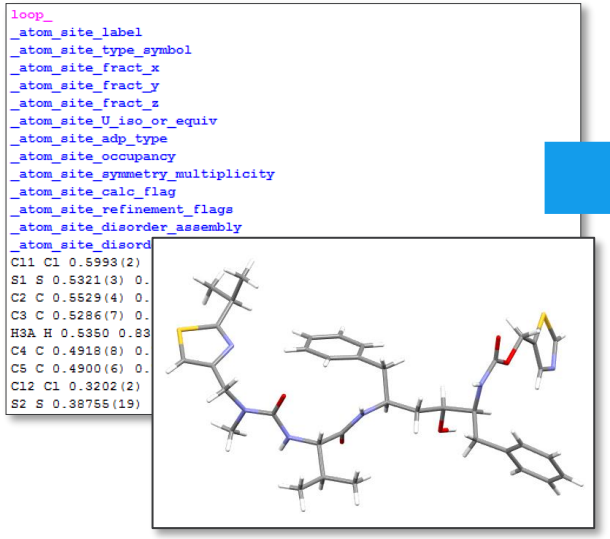
Senior Scientific Editor

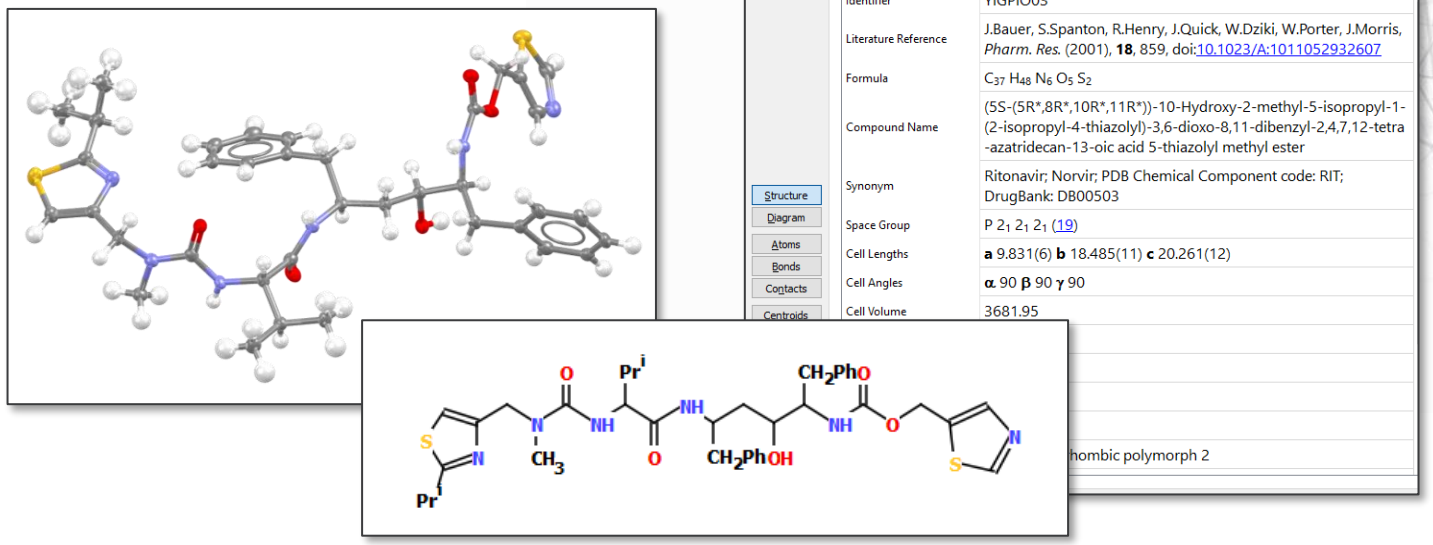
CCDC

How do we curate the CSD?

Deposited CIF

CSD Entry





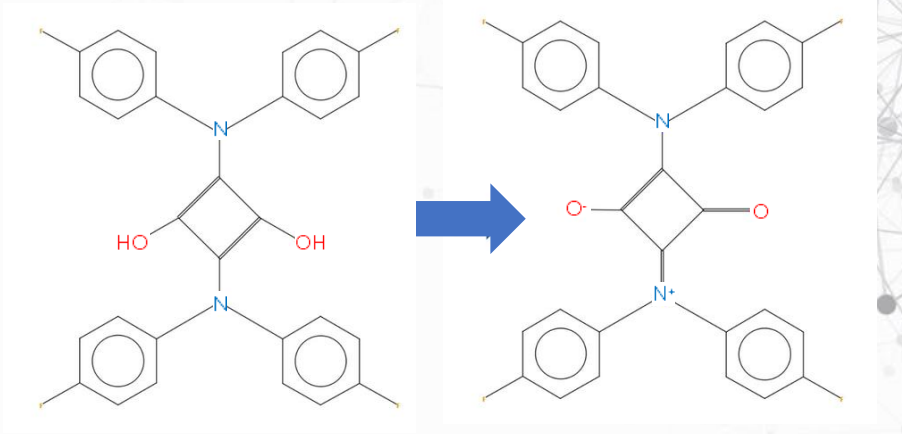
- Assignment of a chemically meaningful representation is determined using data in the CSD and manual curation.
- Important for data discovery, re-use, mining, analysis and interoperability

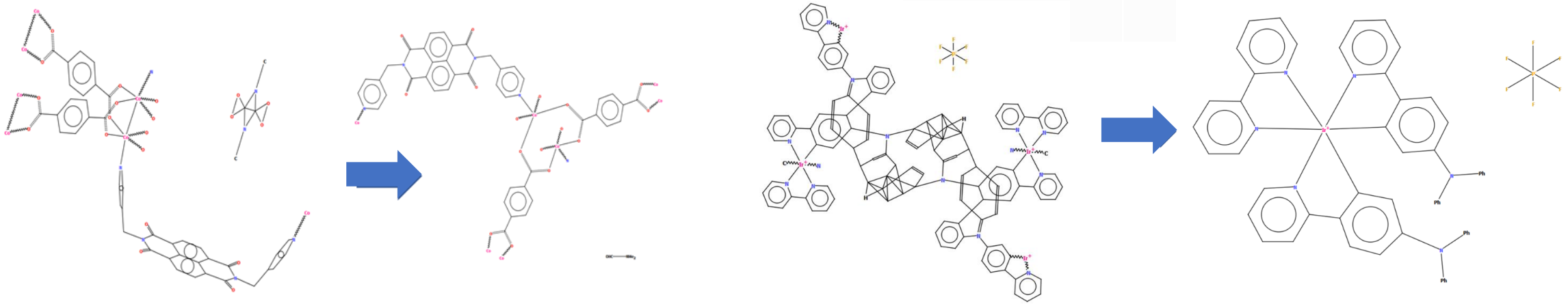
Bruno *et al*, *Acta Crystallogr. Sect. B Struct. Sci*. 67, 333–349 (2011)

CCDC

A look at what editors do

- Each entry looked at by expert Scientific Editors
- Automation focuses editorial efforts
- Manual validation of automated chemical interpretations improves automated methods





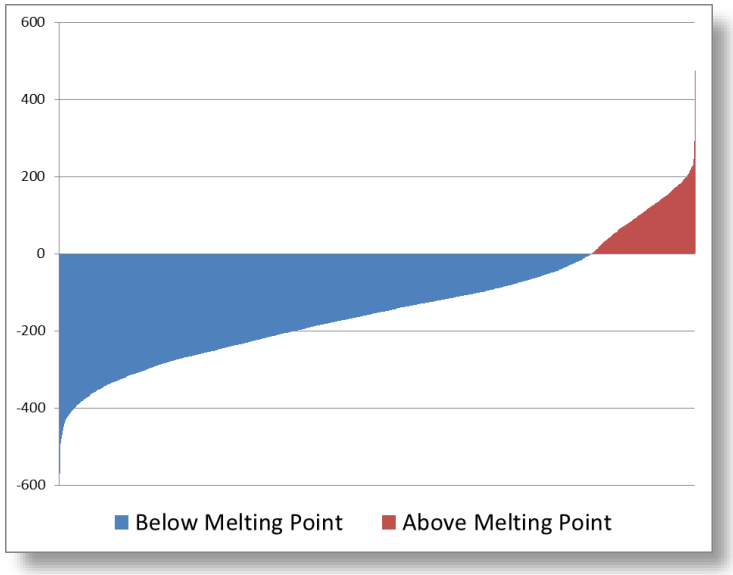
https://www.ccdc.cam.ac.uk/Community/blog/CSD-data-curation-the-human-touch/

CCDC

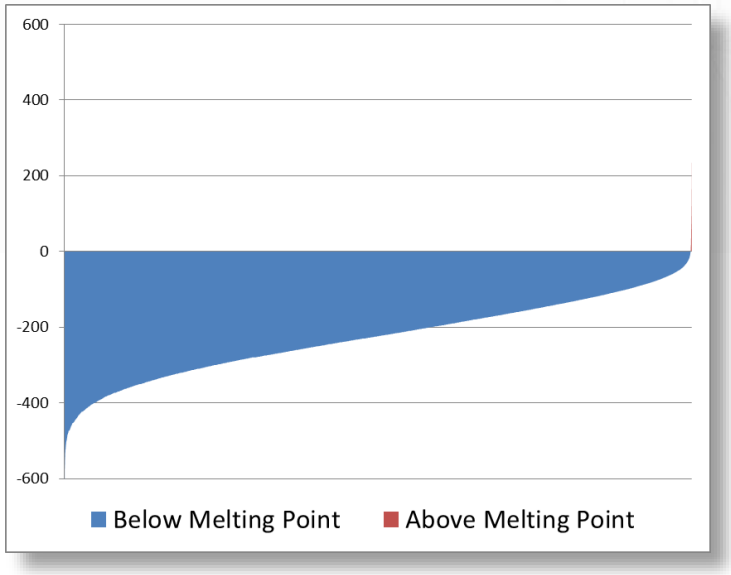
What checks are done before an update?

>170,000 Melting Points

Study Temperature relative to Melting Point
Before additional CCDC validation

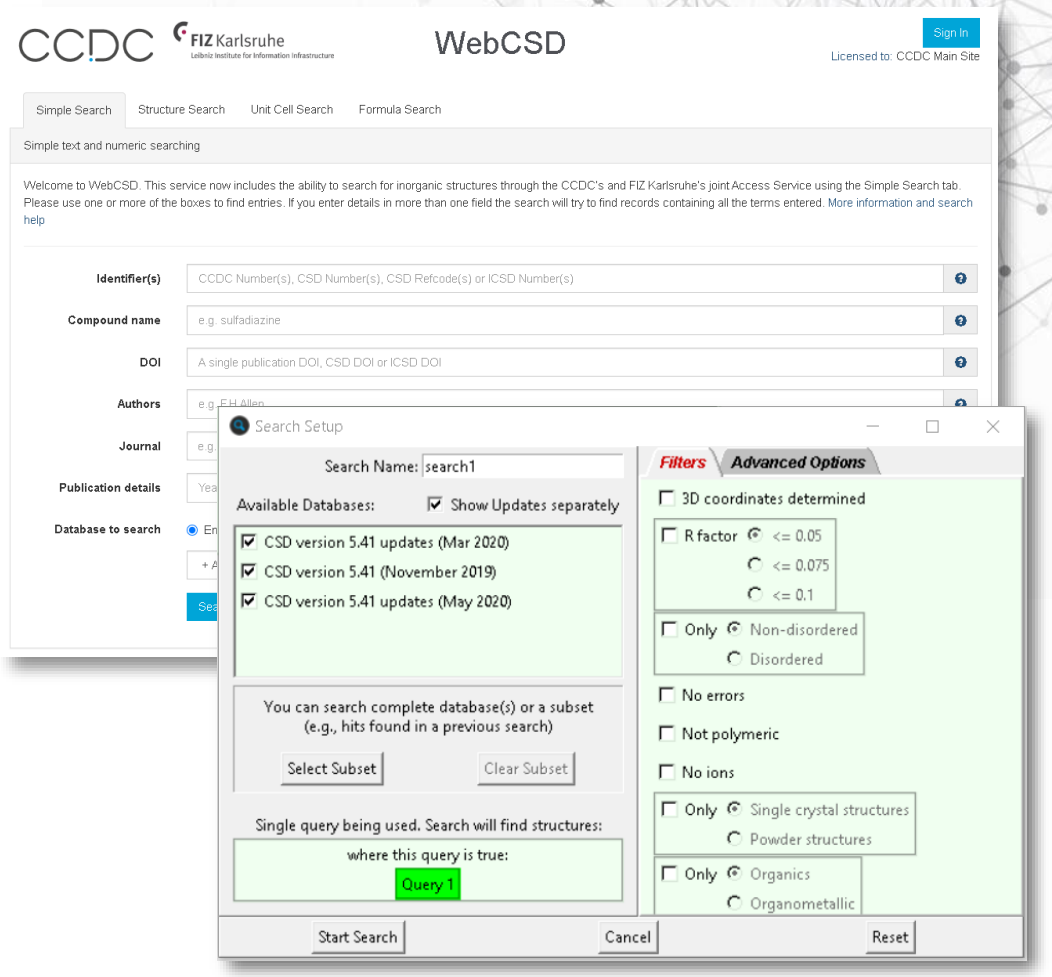


Study Temperature relative to Melting Point
After additional CCDC validation



How often are the CSD data updates?

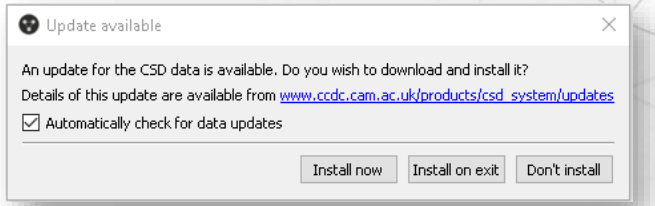
- Access Structures and WebCSD
  - Up to the minute updates available
- Desktop software and our API
  - Currently quarterly
  - Why quarterly?
    - Frequency enables you to stay up to date without continual messages about new updates!
    - Where possible/appropriate we align software releases with data releases

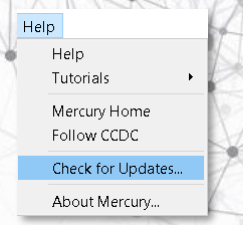


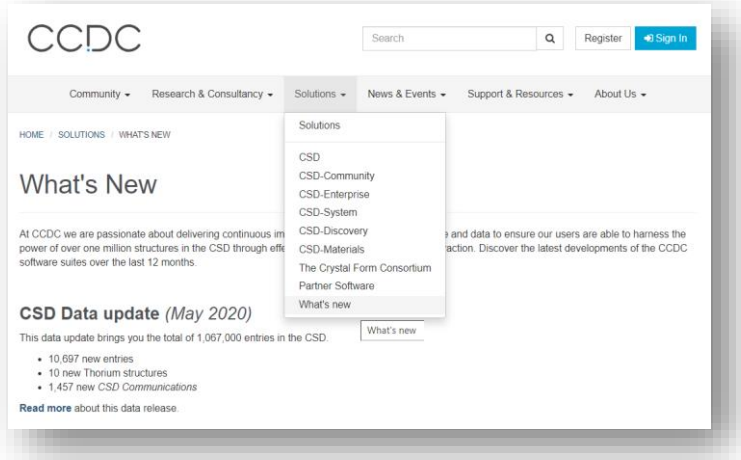
CCDC

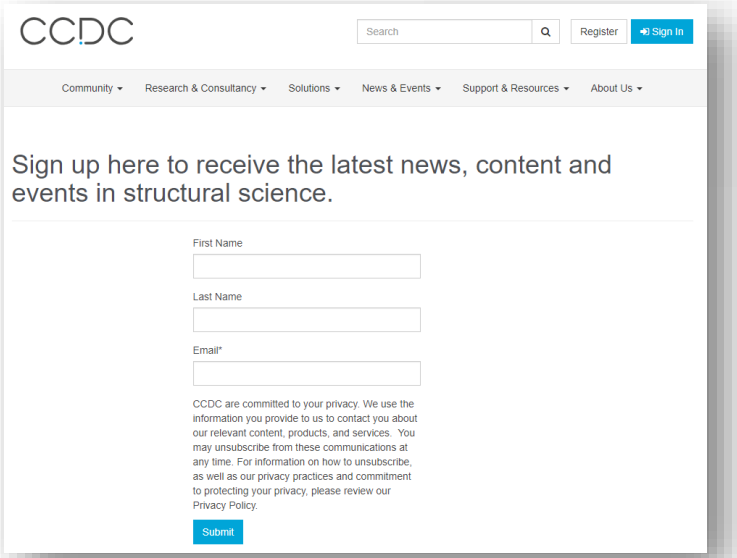
How will you know there is a new update?

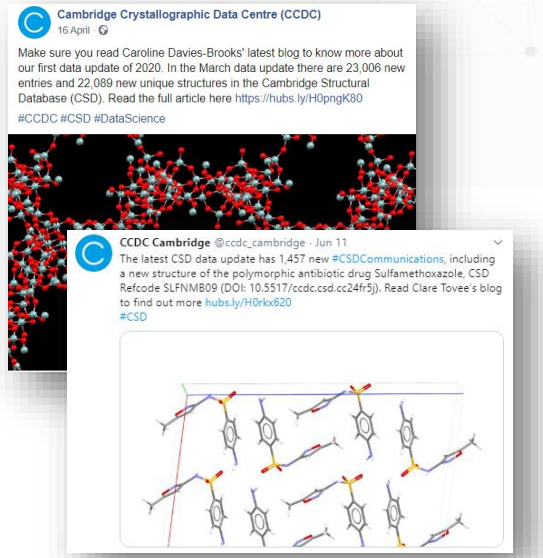
- Through our software suites
- On our website – Solutions > What's new
- Sign up to our newsletters
- Follow us on social media









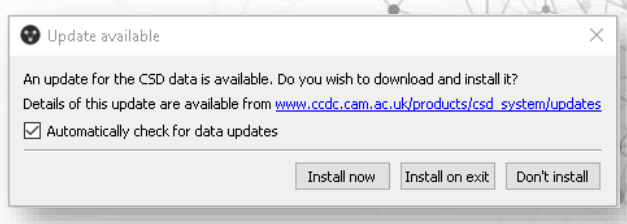


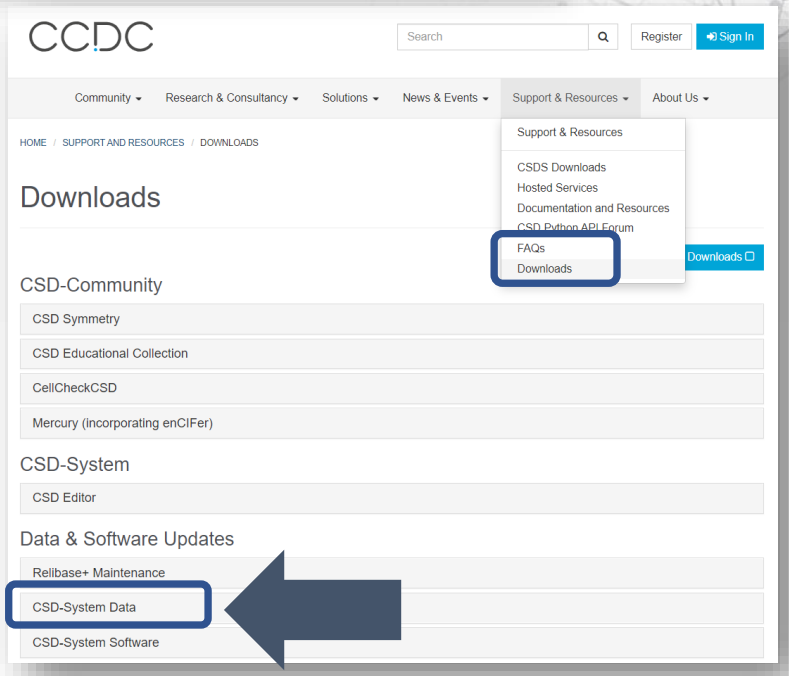
https://www.ccdc.cam.ac.uk/solutions/whats-new/
https://www.ccdc.cam.ac.uk/News/newsletters/

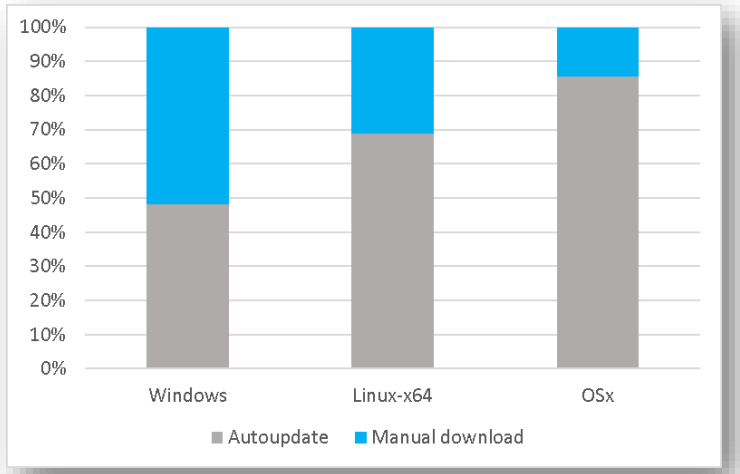
CCDC

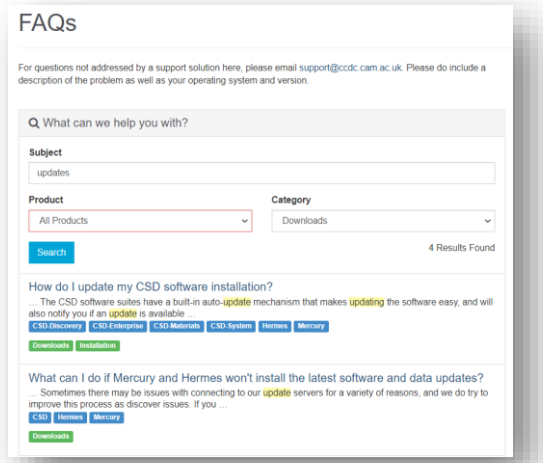
How do you get the new updates?

- Auto-update mechanism built into CSD software
  - Simplifies data and software updates
  - Notifies you when there are new updates available
- Manually from our downloads page
- Our FAQs provide more details and options





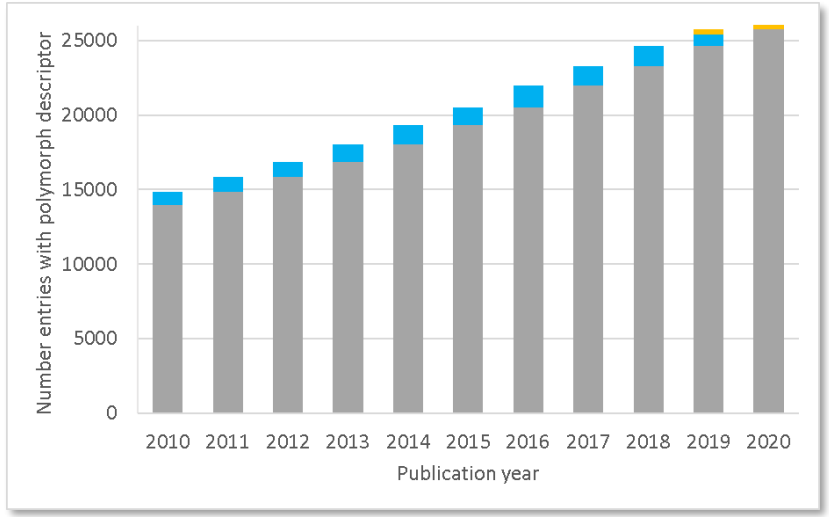


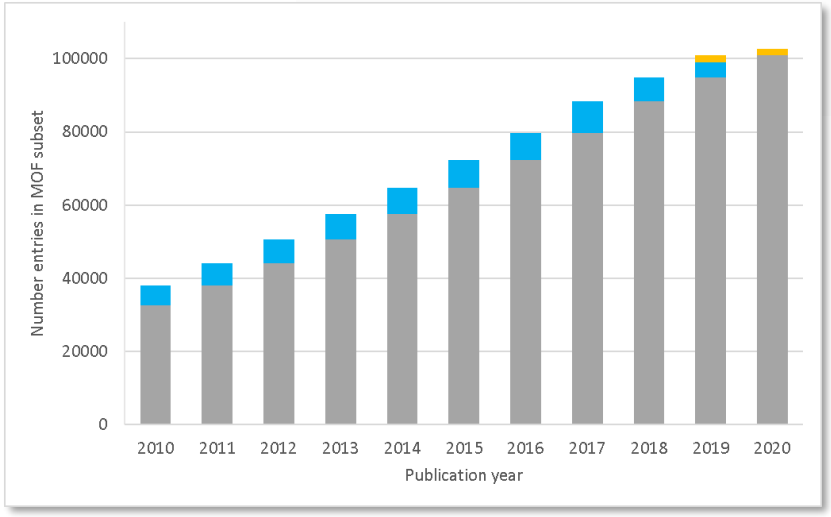


CCDC

What are the benefits of updating the CSD?

- Comprehensive look up
  - Access to the latest structures and associated literature
  - What is included in a data update?
    - All data that has been fully curated into the CSD since the last release

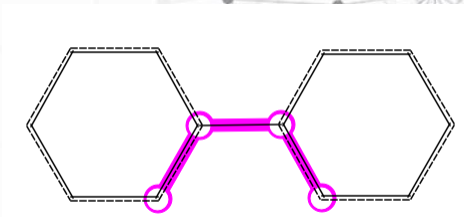


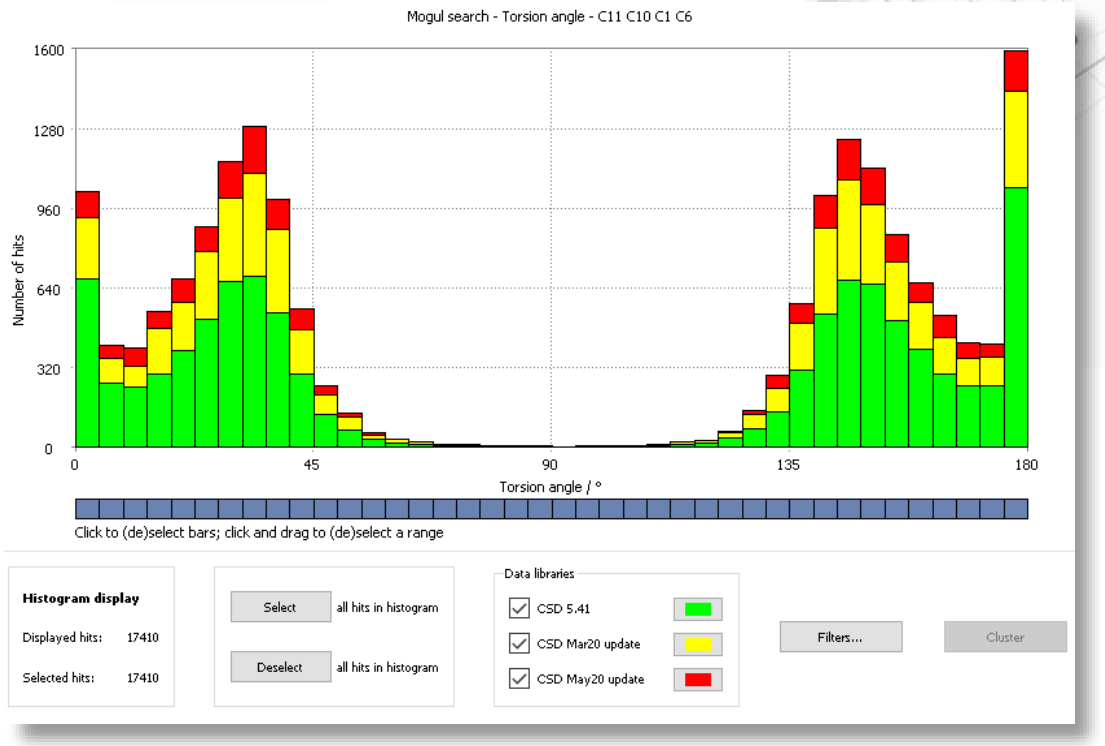


CCDC

What are the benefits of updating the CSD?

- Ability to generate new insights from the data
  - Can see emerging trends or novel discoveries in the very latest data
  - Diversity and depth of data and structures increases
  - Software and analysis that utilises the data can improve
  - Confidence in the data and insights and knowledge from the data can grow

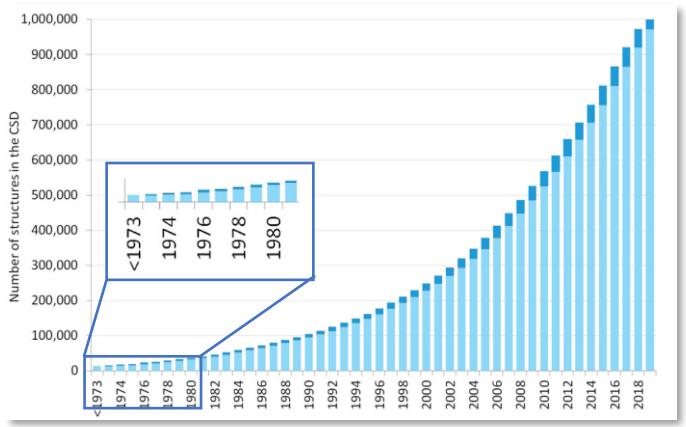




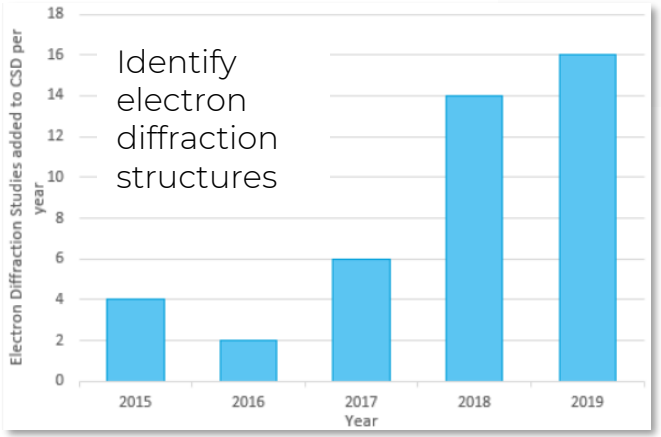
What are the benefits of updating the CSD?

Annual targeted improvements and enhancements to existing entries for improved data integrity, consistency, discoverability and help CSD users to better select data fit for purpose

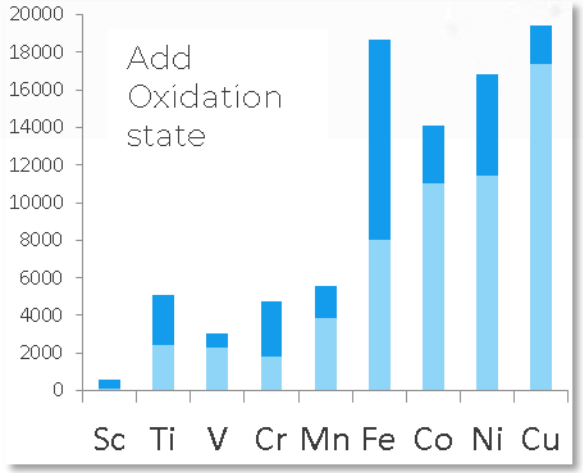
Improve standardisation of early CSD entries



Ensure completeness of experimental metadata



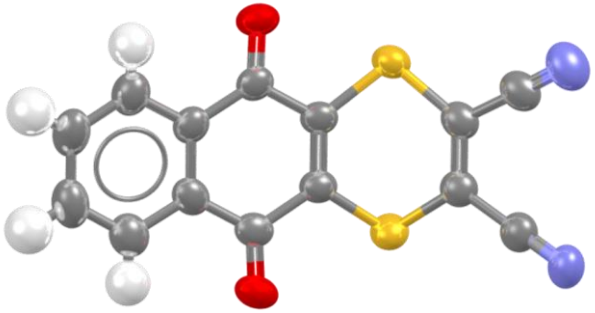
Enrichment of data



CCDC

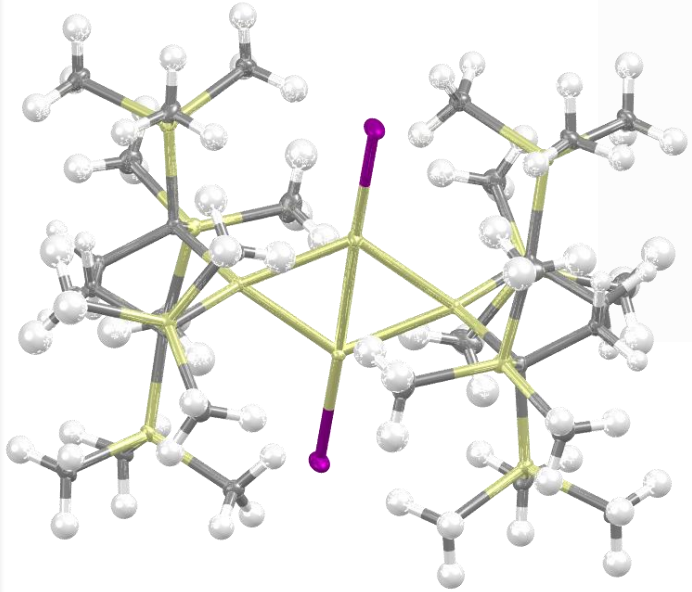
What you get in the updates…

Previously uncrystallisable single crystal structures using new techniques



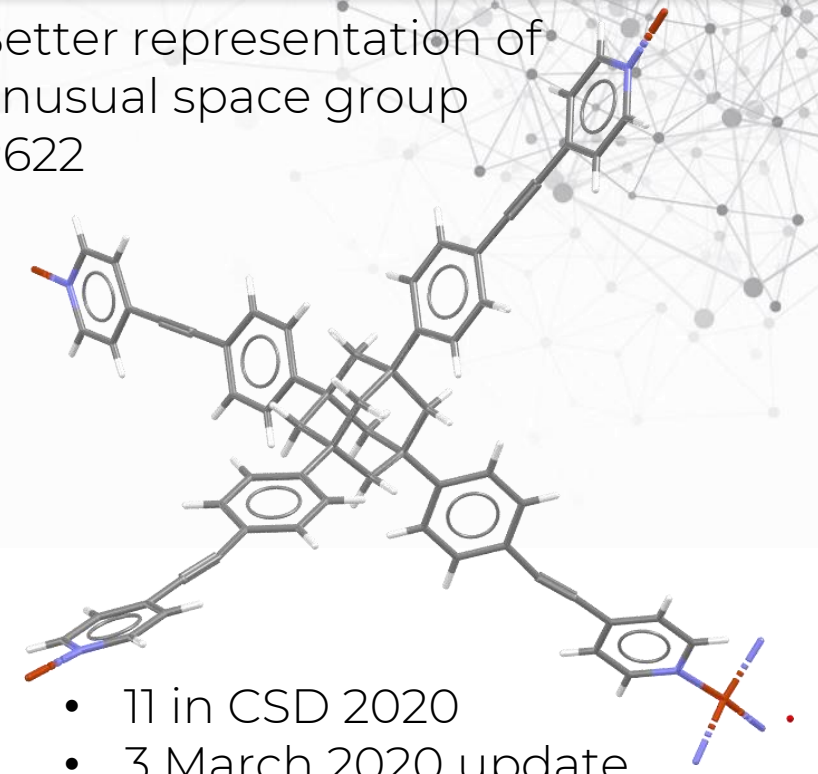
Structure of agricultural fungicide Dithianon CSD-CURHOP

Rare bonding example of Si-Si single π-bond without σ-bond



Structure of CSD-DURTES

Better representation of unusual space group P622

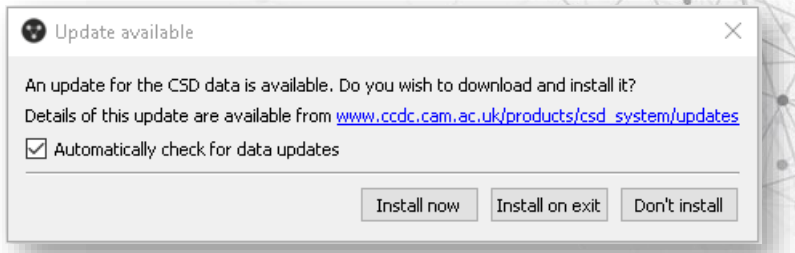


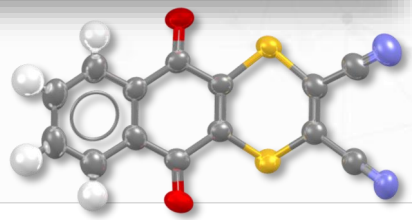
- 11 in CSD 2020
- 3 March 2020 update
- 10 in next update

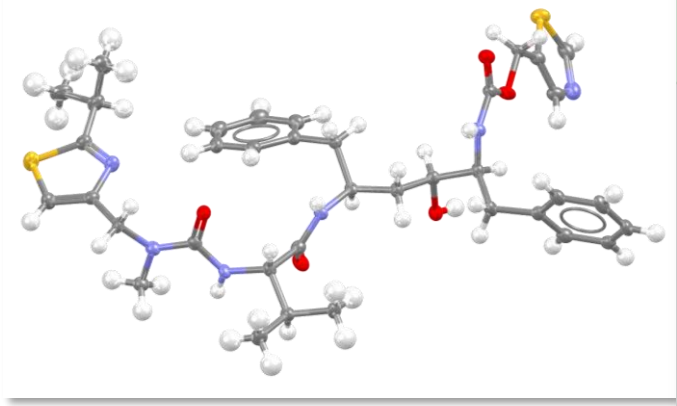
Structure of CSD-XUKMUO

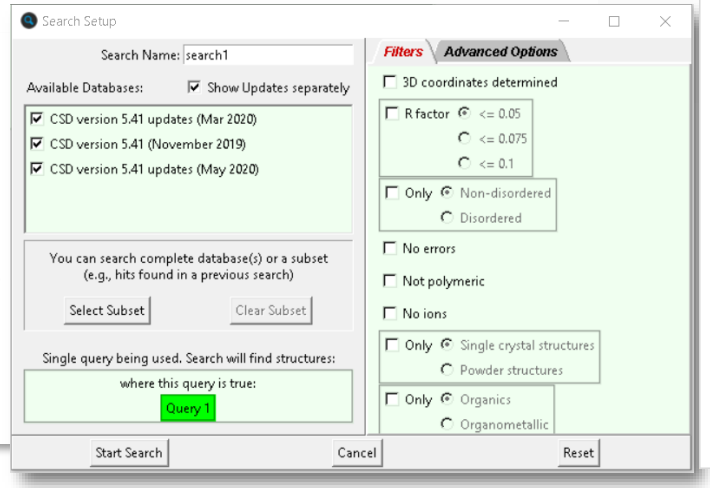
Summary of CSD data releases

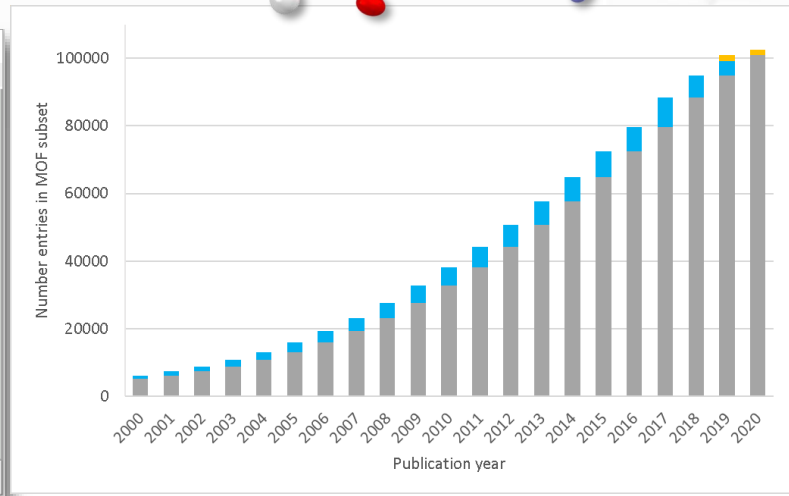
- How we create the CSD
- When and how you can get the update
- What the benefits are of updating











CCDC

CSD Pipeline Pilot component collection

Custom analysis without coding

Dr Peter Wood

Senior Product Manager

CCDC

Overview

- The CSD Pipeline Pilot component collection was launched in the 2020.1 CSD Release (April 2020) and allows you to build custom workflows for analysing CSD structural data without writing code.

- This allows research to be done faster and more efficiently, as well as making it much simpler for users to get to grips with the concept of building their own custom analyses.

- This component collection allows Pipeline Pilot users to access CSD Python API functionality without the need to write Python.

CCDC

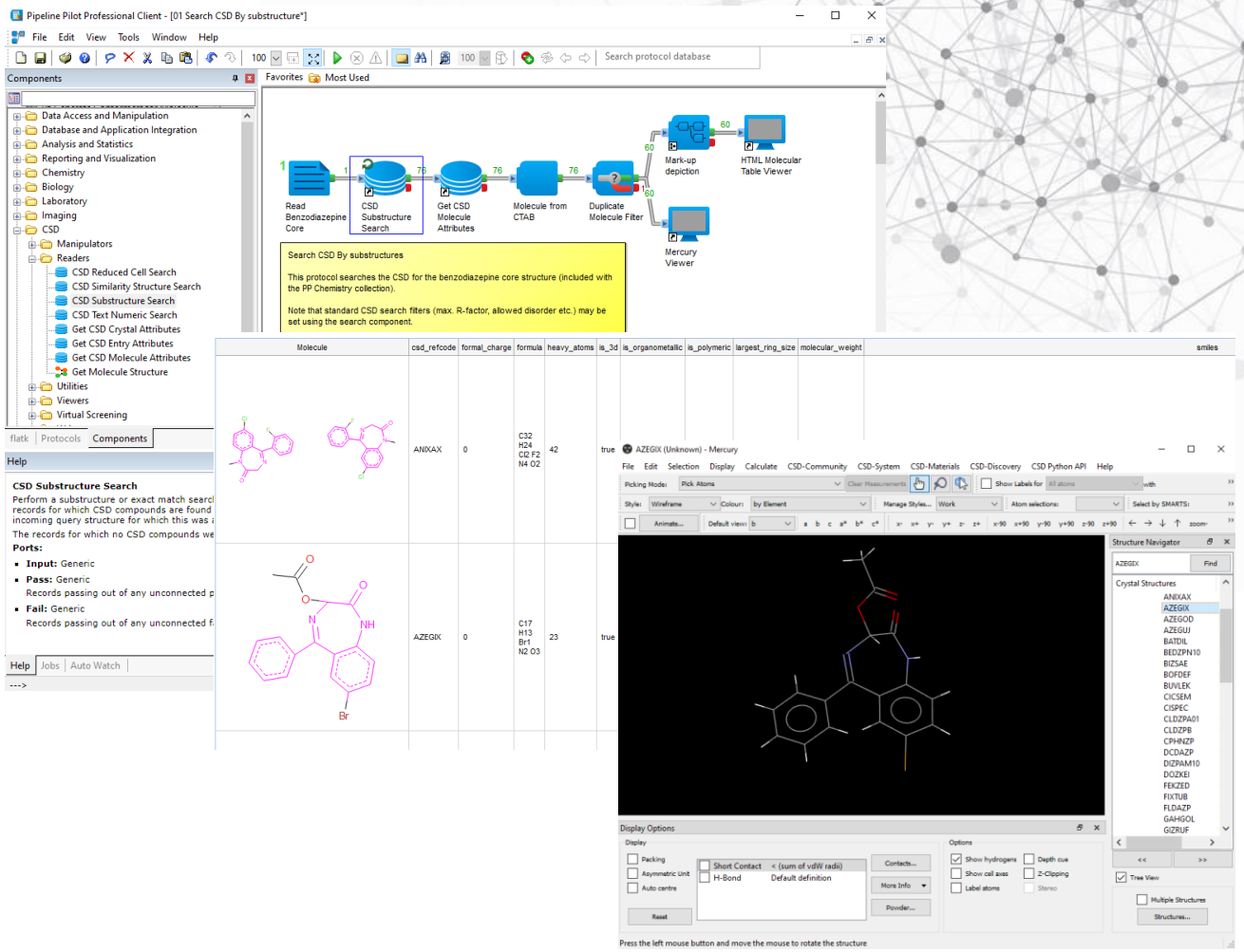
What's in the collection?

- This collection includes components for core tasks such as:
  - CSD data access
  - Text and numeric searching
  - Substructure and similarity searching
  - Conformer generation
  - Virtual screening
  - Visualising structures and data (via Mercury and Hermes)
- There is also a suite of example protocols which you can use as a starting point and develop further as required, including via the CSD Python API.

CCDC

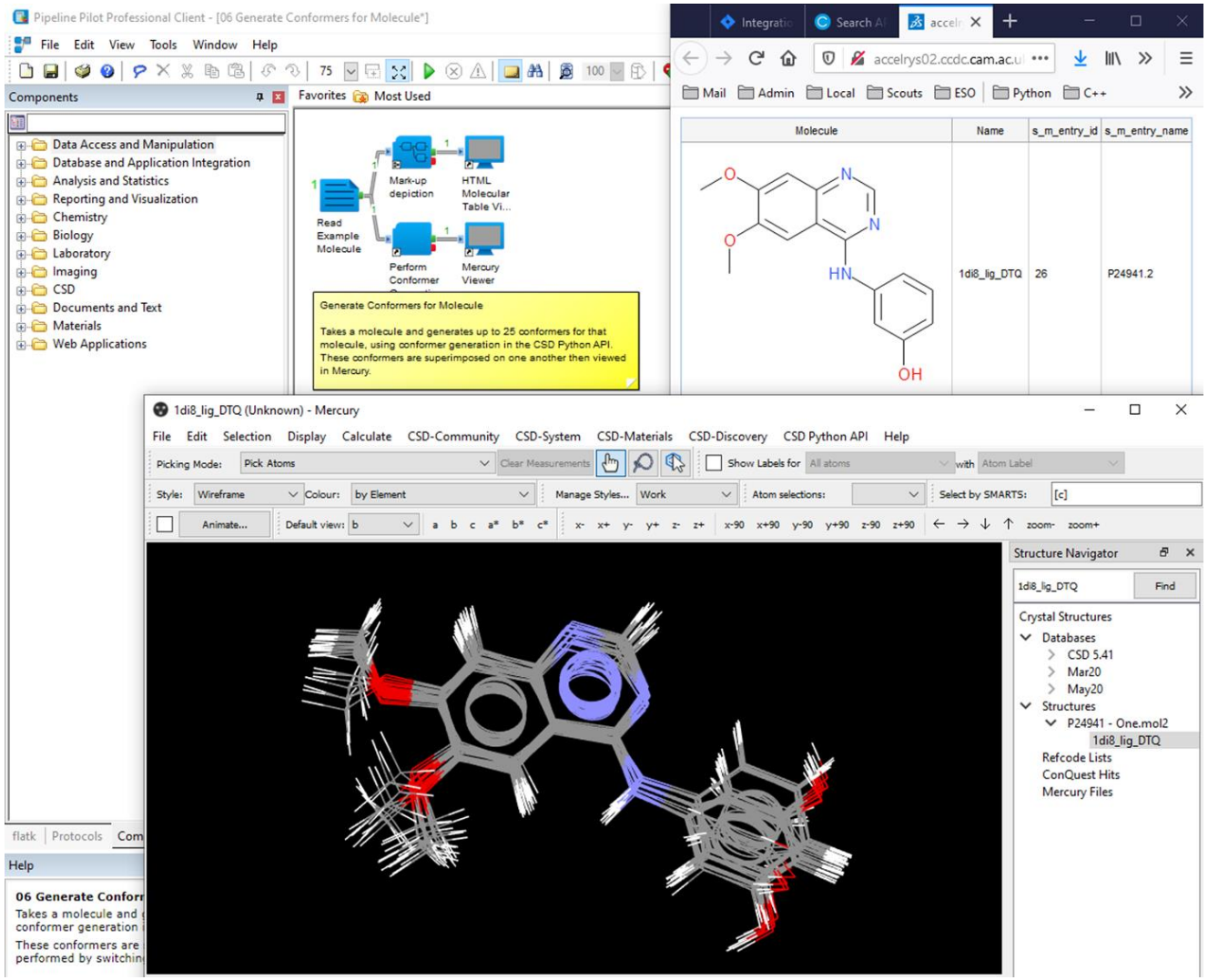
Case Study 1: Searching by structure

- Components available to search the on substructure or similarity.
- Standard CSD filters available in the search components.
- Connect up searches to bespoke analyses of the results.
- Customise what fields to include in the results.
- Visualise results through standard molecular table views, or CCDC's visualisers - Mercury and Hermes.



CCDC

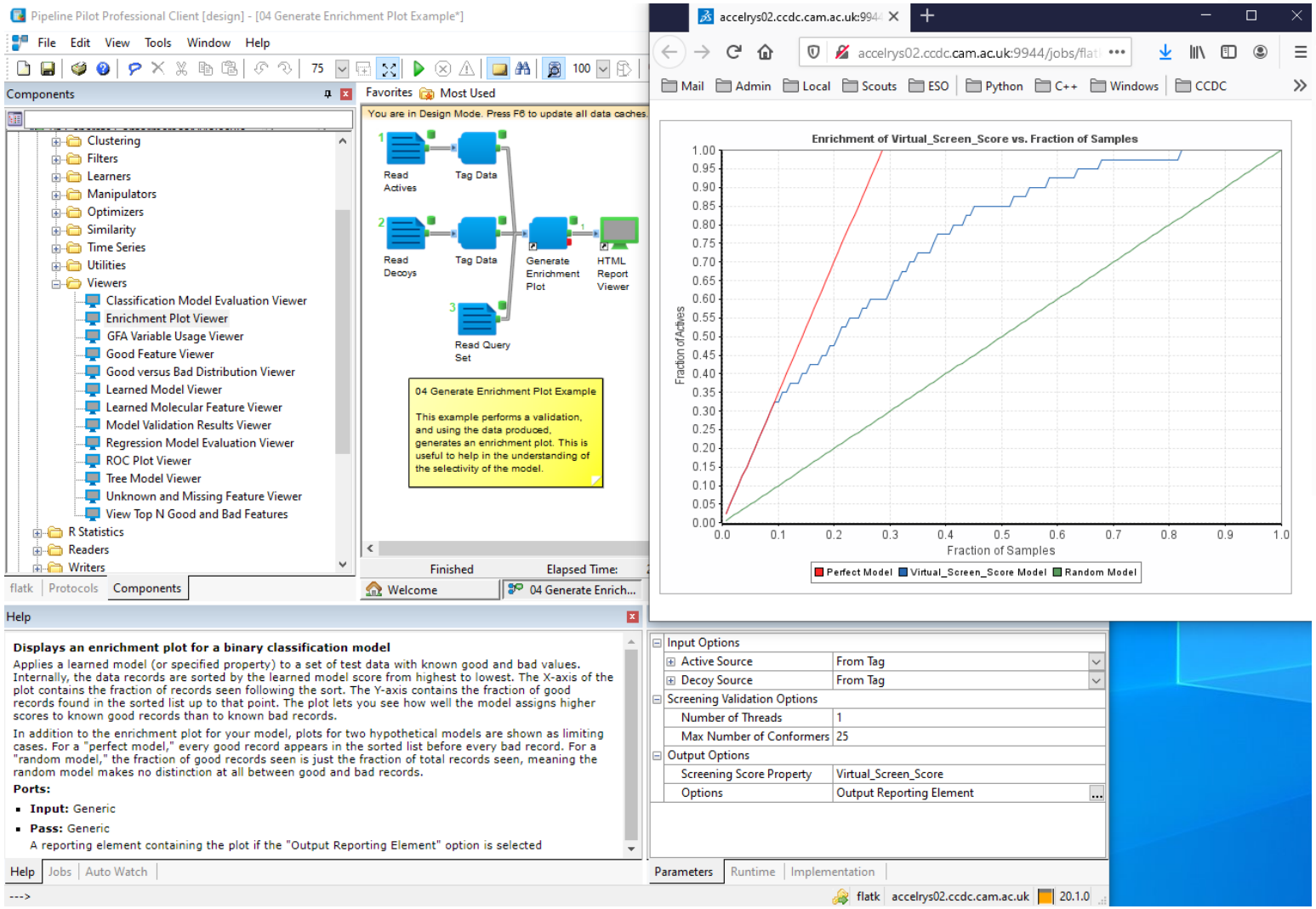
Case Study 2: Conformer generation



- Another key function that you may want to build into your protocols.
- This component allows you to harness the CCDC's effective knowledge-based conformer generation.
- You can plug this into any of your existing workflows.

CCDC

Case Study 3: Virtual screening

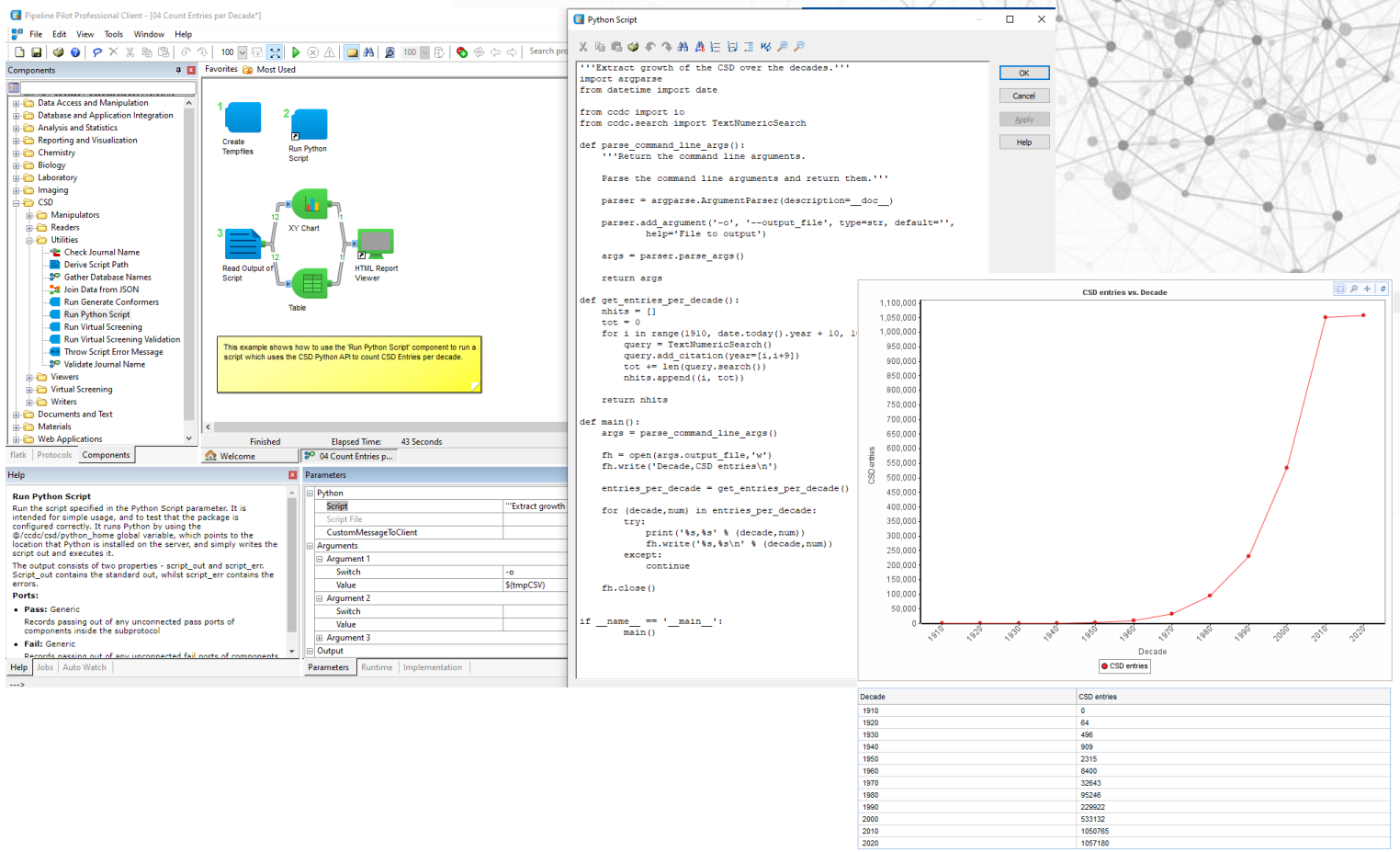


- We've put together a set of virtual screening components and example workflows.
- Here you can see visualisation of an enrichment curve.

CCDC

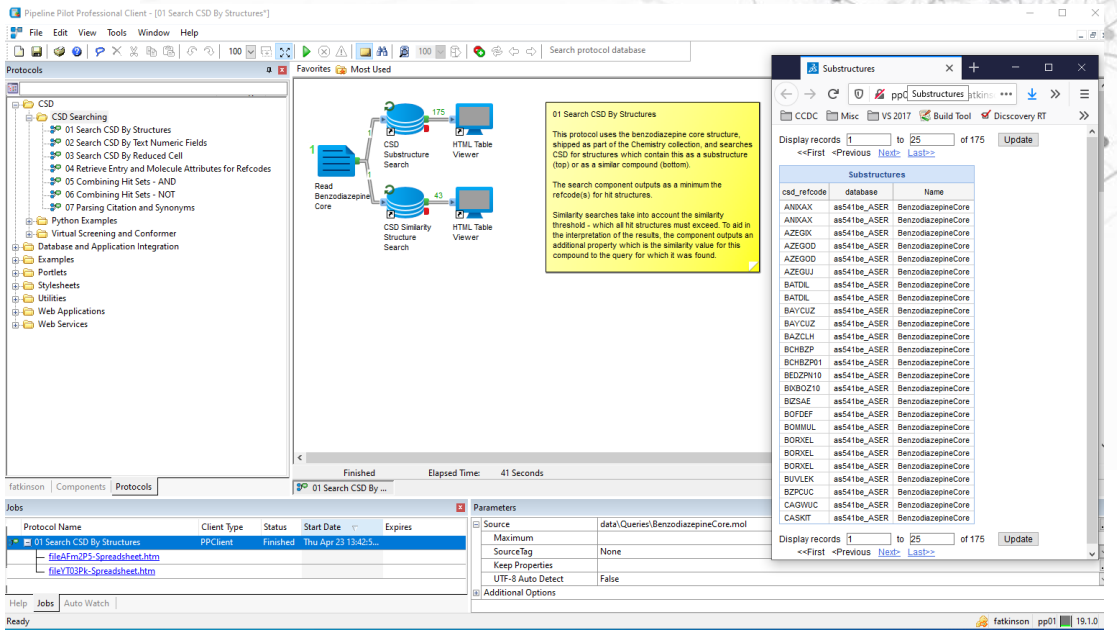
Case Study 4: Custom components

- There's plenty of opportunity to go further with this.
- Use the "Run Python Script" component for custom analyses.
- Access the CSD Python API directly, but in the context of Pipeline Pilot.



Practicalities

- You will need a licensed version of the CSD Python API installed.
- You will need access to Pipeline Pilot itself to make use of the Component Collection.
- Once you have these, you can download the CSD Pipeline Pilot component collection from our website, install and get going!



https://www.ccdc.cam.ac.uk/support-and-resources/csdsdownloads/

Summary

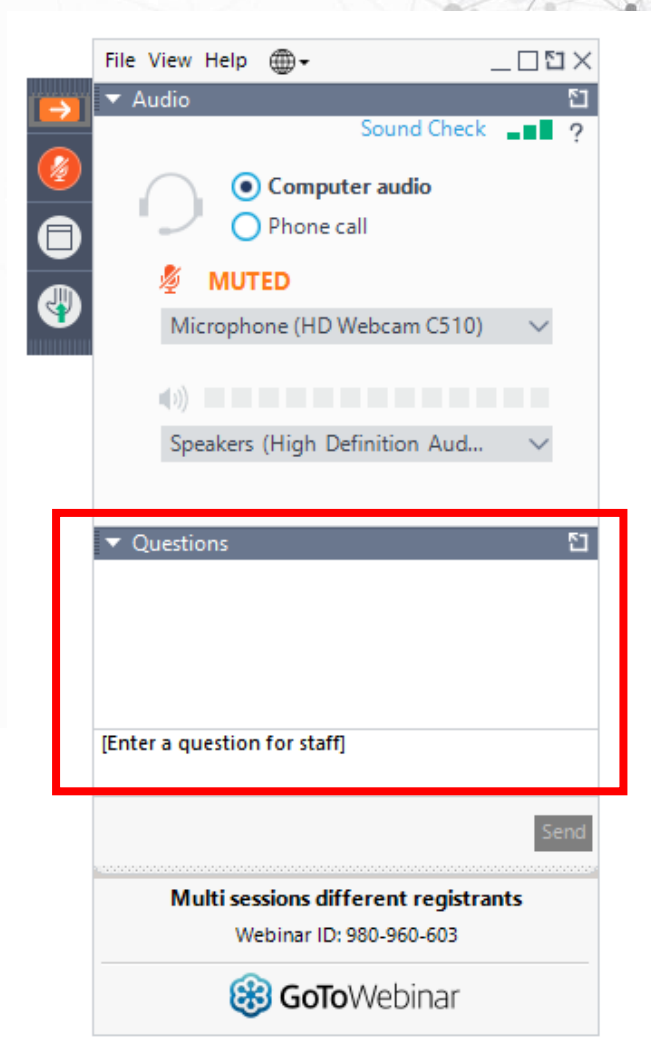
- The CSD Pipeline Pilot component collection is available now for you to create custom workflows harnessing Biovia's Pipeline Pilot platform and CCDC's data & structural science software.
- We've highlighted a couple of case studies today, but the component collection enables many different structural analyses and processes.
- Please do let us know what you think of the CSD Pipeline Pilot component collection as well as any requests you have for components to include in the future via support@ccdc.cam.ac.uk

https://www.ccdc.cam.ac.uk/support-and-resources/csdsdownloads/

CCDC

Q&A

- Type your questions in the box as shown

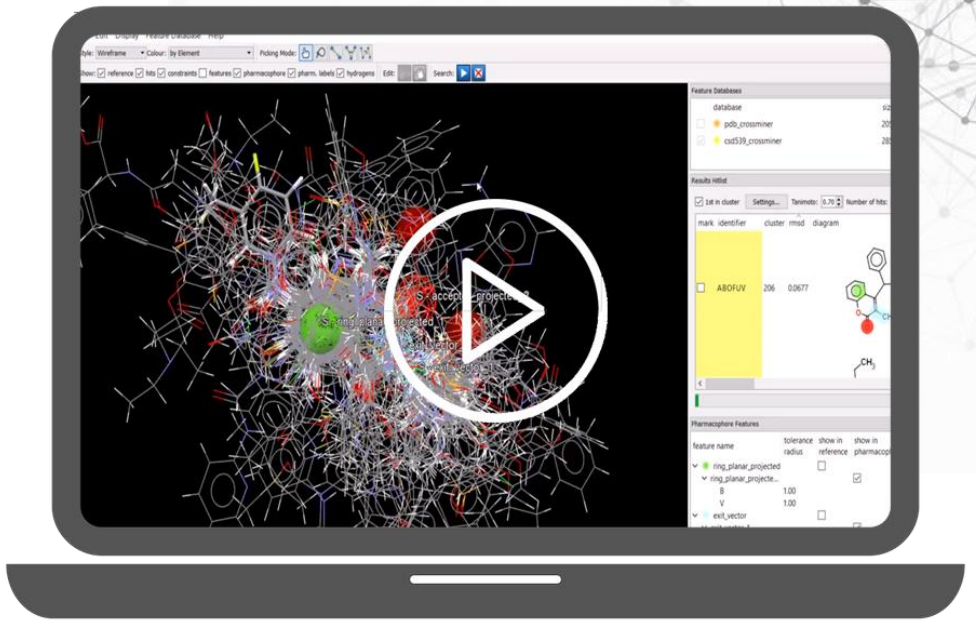


CCDC

Next What's Up Webinar

- Next webinar: September 24th
- Send us your ideas and news

  hello@ccdc.cam.ac.uk



Thank you

hello@ccdc.cam.ac.uk

The Cambridge Crystallographic Data Centre
12 Union Road, Cambridge CB2 1EZ, United Kingdom
Registered Charity No. 800579

CCDC