

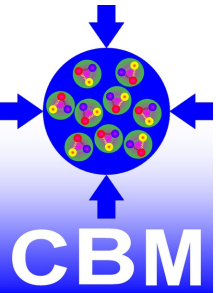
Workshop on Open Science



FAIR — Facility for Antiproton and Ion
Research in Europe



GSI - Darmstadt, 2023-10-20 14:40-15:00



CBM and Open Science

Dr. Eoin J. Clerkin

tel: (GSI ext) 2028
email: e.clerkin@gsi.de

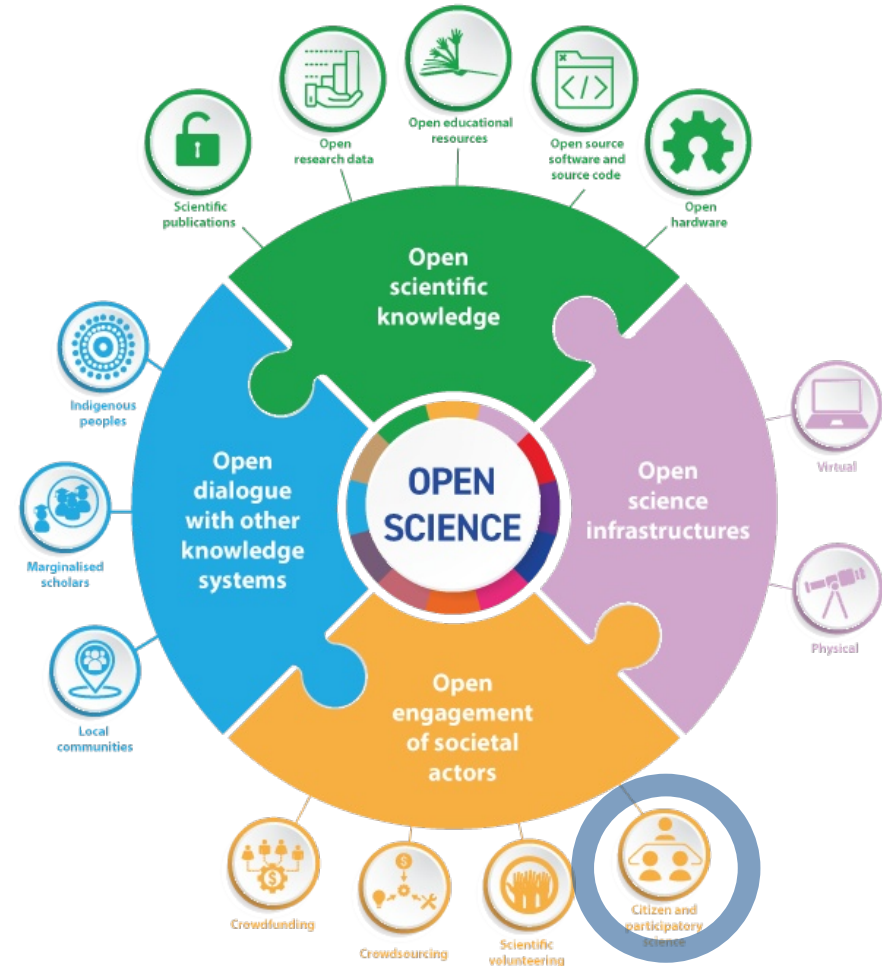
KBW 4.006
Facility for Anti-proton and
Ion Research in Europe
Planckstraße 1
64291 Darmstadt

eurizon

European network
for developing new horizons for RIs

"Open Science" Intro. & Motivation

- What is Open Science?
“Open science is the movement to make scientific research (including publications, data, physical samples, and software) and its dissemination accessible to all levels of society, amateur or professional.” [Wikipedia 2023]
- Kinda the default position anyway even for the scientist who only shares deliberations at the end, publication in close scientific journals.
- Much of development in fields, whether they be engineering, science or software development, occur in the “Open” and “Sharing” paradigm.
- Why this presentation?
Give a case study of Open Science in action. In particular for the CBM experiment based at FAIR.



UNESCO-Open science-pillars

Citizen Science

A personal story



- I am a citizen scientist of ESA-associated 72 strong TOTAS team, tasked with inspecting and confirming Asteroid detections. Many of our most active members have no official scientific background but gain much experience over a number of years. TOTAS has been operating for the last 12 years, and I became involved approximately 5 years ago. Our detection stats are shown on the right.
- Astronomical photography from the Teide Observatory in Tenerife need to be inspected for potentially life treating Asteroids A fully integrated system presenting easy to view imagery, semi-automatic algorithmic detection, record keeping, and reporting has been mostly development and maintained by Matthias Busch.



Teide Observatory Tenerife

- Most observable asteroids in our Solar System are in the main asteroid belts between, of which we have collectively reported 175000 "movers" of which more than 4000 are now tracked by the Minor Planet Center (MPC). However, our main purpose is seeking potentially hazardous asteroids - Near Earth Objects (NEO).

90 Asteroids (numbered discovery)
22 Asteroids (NEO discovery)
2770 Asteroids (new designation)
522 Asteroids (principal designation - old r
426 Asteroids (principal designation, lost v
800 Asteroids (lost designation)

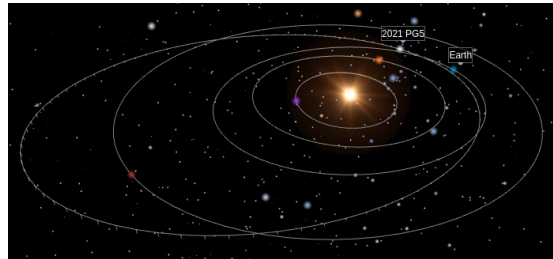
20970 Movers (discovery candidates - OG
8869 Movers (manually created discovery
174737 Movers (confirmed as real and de
54433 Movers (confirmed as real and not
76213 Movers (manually created)
1900 Movers (uncertain)
304052 Movers (rejected)
297 Movers (unconfirmed and not locked)

134 Observing Runs
395 Nights
831 Regions
77954 Fields
298623 Images
139716 Asteroids
789464 Positions
155802 Positions (not yet reported)
938 Astrometry E-Mails
838 Designations E-Mails
58 Measurers

- On August 11th 2021, I noted an unusually fast object in stellar photographs taken the night before 2021-08-10 23:16.

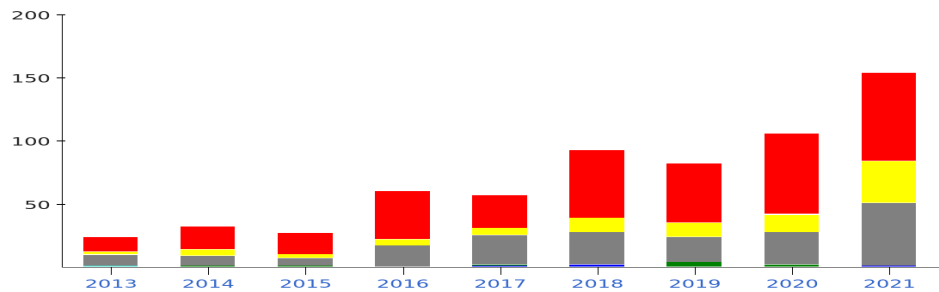


- Within hours, telescopes in Hawaii and Eastern Australia were locked on to the region of the sky for more ephemeris. Close approach but no impact was determined.



NEO OG21145 aka 2021PG5
Closest Approach: 0.0017 Au
2021-08-15 10:00 CEST

- Within 20 minutes of me reporting the potential treat. This was flagged as an object of high concern in the MPC.



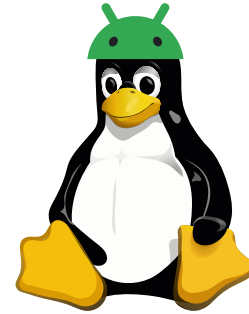
- The NEO had its own short lived conspiracy theory online which I enjoyed. It was the closest visiting object for almost 1 week.
- [Family Story] Made more special as the Christening ceremony of daughter coincide with its closest approach. She considers it her asteroid.

Open Software and Science Probably a typical Researcher!



- Linux Operating systems for last 20 years with Android on my phone.
- Regular reader of wikipedia, a collaborative encyclopedia, powered by open source software.
- Libre Office, which is open source office software was used to make this presentation.
- My research career began in Astronomy using 2MASS (Two Micron All-Sky Survey), after releasing their data to the public domain, after short embargo if memory serves me correctly.
- Publication used to feel wrong with all the documents to sign to give away copyright. Happily this changed more than a decade ago.

Android



TEX



LibreOffice
The Document Foundation



The CBM collaboration united but with different interests



According to the [CBM collaboration database](#) checked last night, there are 322 full collaboration members in some 59 full and associate institutes.

Germany	India	China Poland	Japan	CzechRep Hungary Romania Ukraine	Egypt, France Korea
18	15	6	3	2	1

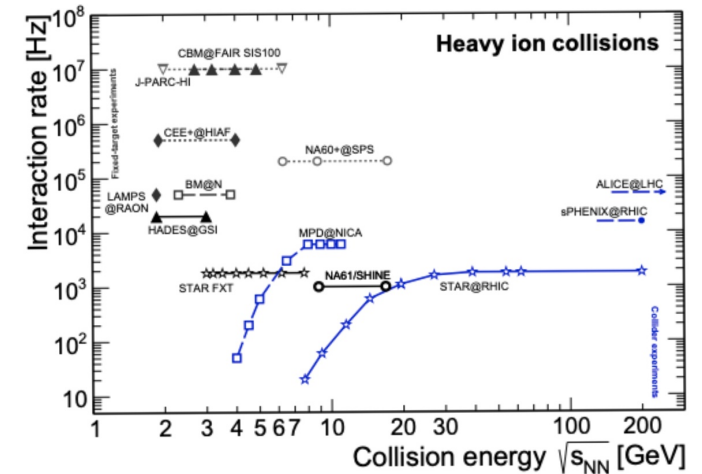
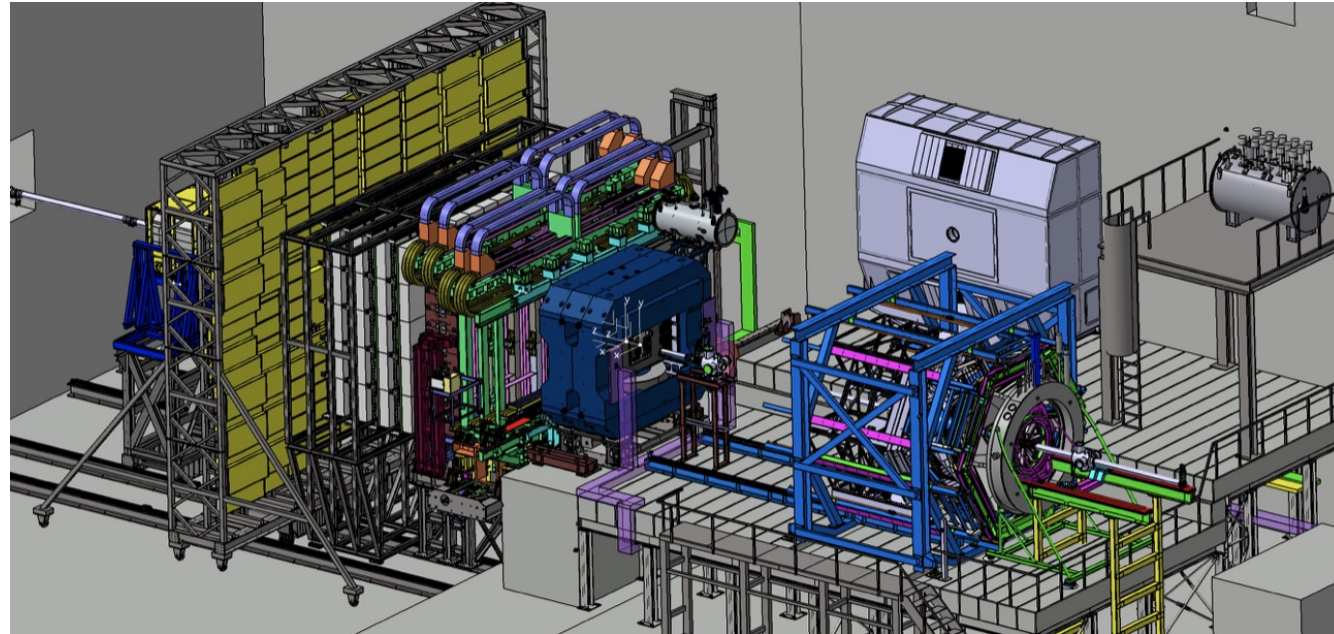
Our “Open Science” policy needs to work in the interest of our

- 1) Our physics goals.
- 2) Our Scientists.
- 3) Their institutes.
- 4) Tax payers of our respective nations.

At the 41st Collaboration Week at TU Darmstadt

- In some ways, we are already an “Open Collaboration”, Research departments and Institutes are welcome to apply to become members. (with only some exceptions)
- Not for me to speak for the CBM collaboration, nor to formulate CBM’s Open Science Policy. But I do intend to collect opinions internally, increase the priority of Open Science in the list of discussions and were possible advocate that a comprehensive policy should be formulated by us.

The CBM experiment



Interaction rates achieved by existing and planned heavy-ion experiments as a function of center-of-mass energy (taken from Tetyana Galatyuk. "Future facilities for high μ B physics." In: Nucl. Phys. A 982 (2019), pp. 163–169)

- A major technological challenge for the CBM experiment is the processing of large amounts of data from our 8 or more planned detectors.
- An 'Open Science' policy for the CBM experiment, is most likely, to be primarily differentiated
 - Software
 - Data processing
 - Analysis software
 - Data
 - Simulated data
 - Real Experimental data

CBM data

An “Open Science” discussion.

- Beam is still 5 years away. Why should we waste time formulating a plan now!
- Discussion on its handling is still in a nascent stage. For us, when and if to share real experimental data is a high interest to many in the collaboration.
- The raw data volume per typical CBM runtime (2 months per year) is estimated to be about 5 PB even at a moderate archival rate of 1 GB/s.

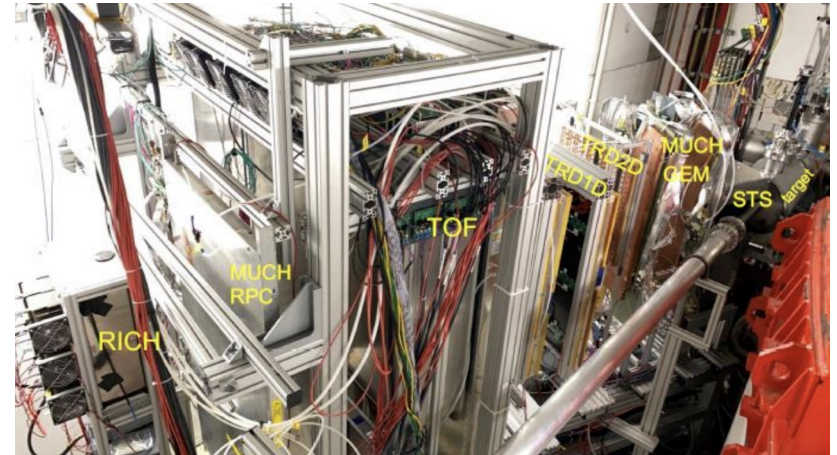
Real Data Today

- Mini-CBM is our test experiment using SIS18 beam
 - Alignment and configuration
 - High Rate Detector Tests
 - DAQ development and Readout
 - Λ excitation function measurement
- Real data from the mCBM experiment, may be used to test implementation of an “Open Science” policy.

setup	hadron	electron	muon	dark
avg. int. rate/1/s	5×10^6	1×10^5	5×10^6	0
	GB/s	GB/s	GB/s	GB/s
BMON	–	0.2	–	0.0
MVD	–	5.0	–	3.5
STS	101.8	9.1	101.8	7.2
MUCH	–	–	37.3	7.5
RICH	–	1.6	–	0.8
TRD	207.6	9.3	24.8	4.3
TOF	42.7	1.0	9.9	0.1
PSD	–	0.3	–	0.0
Sum	352.1	26.4	173.9	23.3

Average total data rates sent from the detectors to the FLES stage. The stated rates include raw data messages and overhead.

Applying a contingency factor of 1.5, we arrive at a minimum bandwidth requirement of 600 GB/s.



Real mCBM in realtime data registered to an ESCAPE datalake



- CBM and FAIR participated in a week long Data Acquisition Challenge in November 2021 for the EU funded Horizon 2020 ESCAPE.

- As a proof-of-concept, simulated data from CBM and real data from mCBM. Real time interacting and data registration methods were showcased.

- “Real and simulated CBM data interacting with an ESCAPE datalake”
E.Clerkin, P.-N.Kramps, P.-A.Loizeau, M.Szuba, CBM Progress Report 2021



Real and simulated CBM data interacting with an ESCAPE datalake
E. Clerkin¹, P.-N. Kramps², P.-A. Loizeau³, and M. Szuba⁴
¹FAIR, Darmstadt, Germany; ²FAIR, Darmstadt, Germany; ³FAIR, Darmstadt, Germany; ⁴FAIR, Darmstadt, Germany

Introduction
ESCAPE is an ongoing ESCAPE Horizon 2020 project, whose aim is to investigate, design and share data large scale FAIR astronomical and particle physics data with the open science paradigm. The project is a part of the ESCAPE challenge, which aims to demonstrate the CBM experiment, which is currently under construction at FAIR, in interaction with the FAIR and ESCAPE datalake, which are distributed repositories of real and simulated data for analysis. This is supported by other work packages of the ESCAPE project. For information on ESCAPE, visit: 1, 2, 3 and references therein may be consulted.

We describe the specifics of CBM data interacting with ESCAPE data and present a technical overview of the data registration and analysis pipeline. The CBM experiment is currently under construction at FAIR, and will be ready for operation by the end of 2025. The CBM experiment is currently under construction at FAIR, and will be ready for operation by the end of 2025. The CBM experiment is currently under construction at FAIR, and will be ready for operation by the end of 2025.

mCBM data registration
In order to demonstrate the registration of mCBM data to the ESCAPE datalake, a proof-of-concept was developed. This was a simplified pipeline for early testing, but will be replaced by a more sophisticated one in the future. During the ESCAPE challenge, the registration of mCBM data to the ESCAPE datalake was demonstrated. This was a simplified pipeline for early testing, but will be replaced by a more sophisticated one in the future. During the ESCAPE challenge, the registration of mCBM data to the ESCAPE datalake was demonstrated.

Conclusion
The registration of mCBM data to the ESCAPE datalake was demonstrated. This was a simplified pipeline for early testing, but will be replaced by a more sophisticated one in the future. During the ESCAPE challenge, the registration of mCBM data to the ESCAPE datalake was demonstrated.



CBMROOT ← A interesting licence case

”Open Software”



- Our main simulation, and reconstruction software suite, CBMROOT, has collaboratively been authored close to 100 authors.
- Initial development was conducted under with no licencing policy. It was deemed that the institutes, who paid the wages of the developers, were the fair copyright holder, so the institute of the first author of each file in our code base was listed as owned by their institute. Some authors had dual associations.



docker



```
/* Copyright (C) 2018-2021 GSI Helmholtzzentrum fuer Schwerionenforschung, Darmstadt
SPDX-License-Identifier: GPL-3.0-only
Authors: Florian Uhlig [committer], Alexandru Bercuci */
```

- As a Collaboration, permission from each copyright holder was sought to allow the release of our entire source code under

• General Public License v3 (GPL-3)

Can	Cannot	Must
Commercial Use	Sublicense	Include Original
Modify	Hold Liabile	State Changes
Distribute		Disclose Source
Place Warranty		Include Licence and ICopyright
Use Patent Claims		Include Install Instructions

- Our source code as well as installation instructions are publicly available online via <https://git.cbm.gsi.de/computing/cbmroot>



REDMINE

flexible project management



GitLab

▼ Things that will happen soon ▼



ESCAPE
OSSR

Open-source Scientific Software
and Service Repository



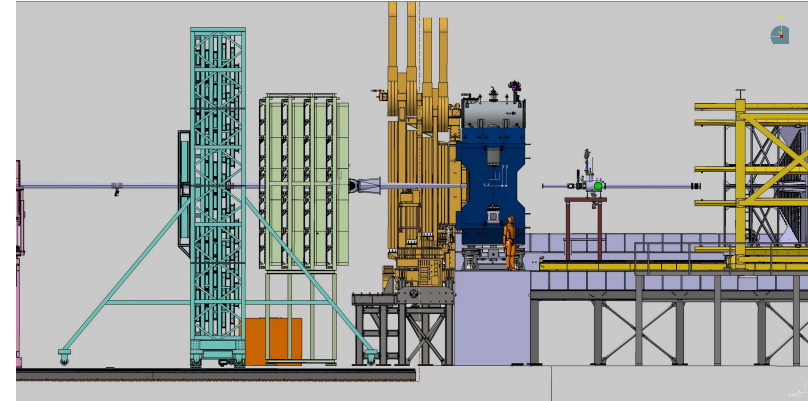
Open Science – CBM Case Study

A CAD Engineer's Perspective

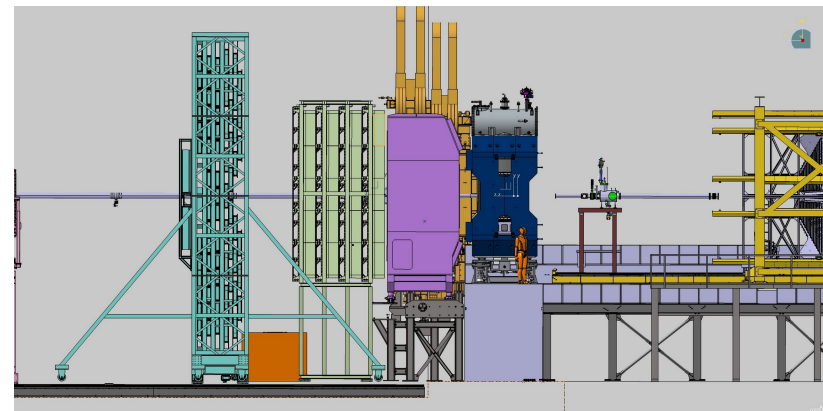


- CBM and FAIR contain highly specialised information from an engineers perspective with considerations of high radiation hardness and low material budgets. Some commonality for space sector.
- Designs and reports would be highly useful to engineers as a cross check for the design of future experiment designs. [cost saving]
- Each detector, sometimes parts and sensors are designed and/or constructed by our home institutions.
- Current designs not presently openly available
 - No clear guidance -- personal consequences.
 - Technology office -- patents not possible if publicly distributed.

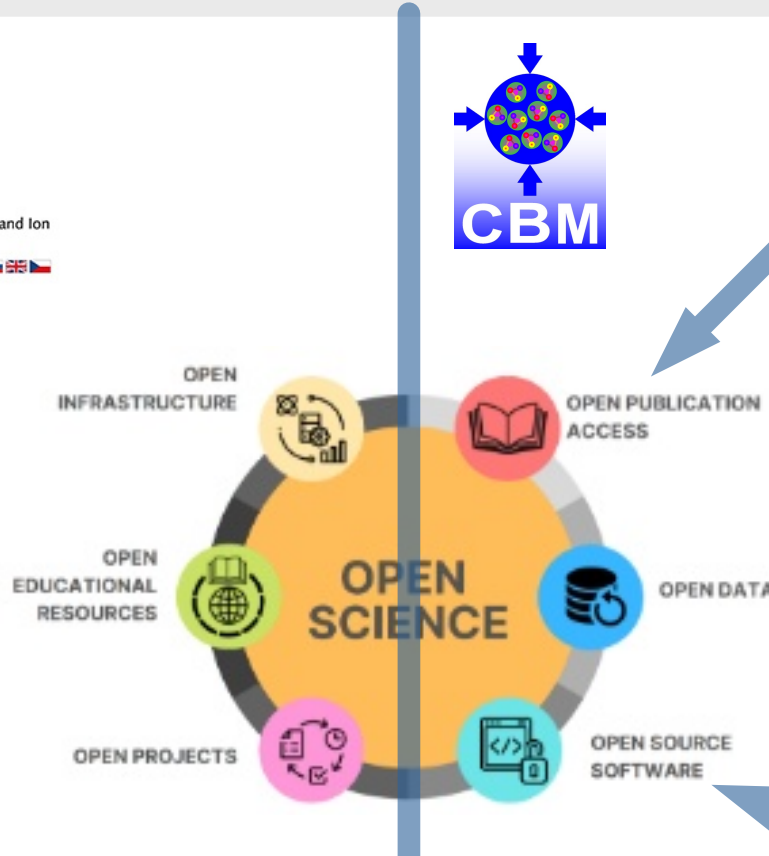
LMVM experimental setup



Electron experimental setup



“Open Science” CBM summary slide



May be a “fait accompli”

- CBM members want to publish results in an open as possible and has some obligations at national level.
- Use of services like ArXiv in case of closed journals.

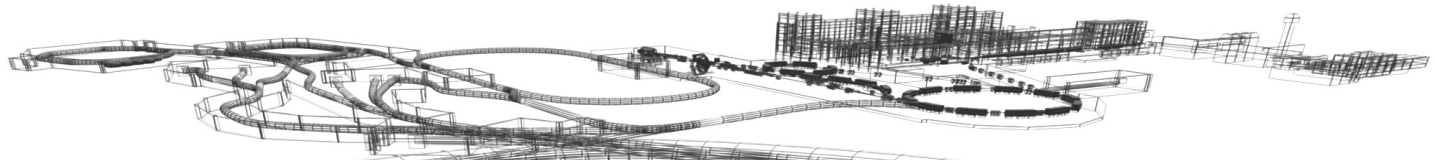
Under active discussion

- Realisation of need to formulate a policy.
- Should not be forced into a specific policy.
- Testing an “Open Science” policy in mCBM.

Software has been made open source

- Shared publicly online. Will engage with OSSR to increase “Findability”.
- May be some discussion around auxiliary Analysis software .

CBM may be more limited in its role here.



Thank you for listening!

**Questions or Comments
Welcomed.**

