Reviewer 1:

The manuscript documents an ongoing effort to establish a computational chemistry resource that provides both extensive data and associated tools for computation and analysis. Their open-source and community-driven approach is commendable, and the manuscript is publishable after the following points are taken into consideration.

1) The objectives of the project are ambitious and no doubt many previous attempts have floundered due to problems arising during their development. Despite this, there is little discussion on the challenges faced in the QCArchive project, what possible solutions were considered, and the reasoning behind the resolutions of these challenges. For example, can data from multiple programs be combined into one data set and, if so, how are the various program options (the defaults of which can vary widely across packages) treated? When can levels of theory from two different packages be considered equivalent? (SCF convergence thresholds, algorithms, DFT grids etc.) How are excited states handled?

Author Response: *Data from multiple programs can indeed be combined into one data set, and we have added the following statement on p.12 to indicate this:*

"For example, unique identifiers for given single point computation are molecule, method, basis, program, and keyword arguments create a unique computation."

Data is stored/retrieved/displayed based on the unique identifiers given above. Trying to reconcile the 'same' calculation using different software is beyond the scope of the project, and is one area where previous projects may have stumbled. Excited states are not currently handled, but can easily be added.

2) I downloaded the software and, although I was able to run the examples in the manuscript, it felt very much like I had been given a few fish and not taught how to fish. Perhaps the authors could include some examples of how to access help topics, and how to obtain listings of valid keywords and data identifiers. Without these the experience feels like groping around in the dark.

Author Response: *While continued usability of the package is one of our highest short-term priorities, in the meantime we have provided examples to help users get started here:*

`https://qcarchive.molssi.org/examples/`

3) An important aspect of any collection of data such as the MQCAS is data integrity. Provenance meta-data are included, which would allow for bulk data invalidation if a bug were to be discovered, but what about other potential problems such as SCF convergence to a saddle point? Figure 1. indicates a review process is carried out on contributions to the database, but no details are given. I think for a topic which is so central to the project, more discussion is required.

Author Response: *We agree fully with the reviewer that quantum chemical calculations are complex and frequently lead to unsought results. Thus, we have recently employed a review process by the* MQCAS *maintainers for data submitted sets. We have added the following on p.5:*

"Submitted datasets are reviewed by the MQCAS maintainers for integrity and often require comparisons to literature values to ensure consistency (where possible and appropriate) before accepting a contribution."

4) In Section 3.2 four software components are mentioned (page 7, line 25), yet five are summarised in the subsequent bullet points. Furthermore, only four of of the five are expanded on in the following sub-subsections (QCSchema is not covered).

Author Response: *We have corrected the number noted on p. 7 from "four" to "five". However, we have not provided further elaboration on* QCSCHEMA *in this manuscript because it involves a much larger number of contributors than the current author list and its complexity requires a separate article to be published later. To clarify this in the paper, we have added the following parenthetical statement to the* QCSCHEMA *bullet on p.8.:*

"(This important component, which has been developed somewhat separately from the rest of the QCARCHIVE package, will be decscribed in a later publication.)"

5) It would be interesting to know if there are any efforts being made toward integrating other chemical databases such as ChemSpider and whether or not an HTTP API exists or is being considered.

Author Response: *We have had several conversations with PubChem and Citrine about cross referencing data, but this has not yet been complete, and thus we feel it would be premature to comment on this in the manuscript. QCFractalinteracts exclusively over standard TCP/IP (HTTP/REST) protocols that can be accessed from any language as discussed briefly in Fig 4.*

Reviewer 2:

The authors provide an overview of a new archive for quantum-chemical data that is accessible for free for users. It offers a large variety of options, interfaces with many quantum-chemical programs and shows huge potential for further extension. This is without doubt a great initiative. The article is well written and easy to follow and has the potential for a large number of citations as soon as more users have become aware of this initiative.

The article can almost be published as is and I would just like to mention four points for a very minor revision.

1) Ultimately, it is up to the editor to decide on the article type, but as this article does not provide an in-depth review of a methodology published in the literature but revolves more around software for data mining, I feel that a more appropriate article type would be Software Focus.

Author Response: *This is clearly up to the editor, but we have no preference.*

2) Could the authors comment on their strategies to ensure that existing databases are monitored for updates? For instance, they use reaction databases such as ACONF, S22, etc. and in the past the reference values for such sets have been updated multiple times. In addition to having access to structures and low-level energies, it is also important for the CMS community to keep track of the currently most accurate reference data.

Author Response: *We agree, and we make a concerted effort to keep up to date on the literature and update the values where available and have also created the open gateway to adding new values. An example of this is the S22 dataset where currently the original benchmark values as well as the "S22a" and "S22b" have been added to the dataset.*

3) Ref. 63 is not the correct reference for the B3LYP functional.

Author Response: *We have removed the original Ref. 63 and added the following two corrrect references (recognizing, of course, that not all quantum chemical programs define this functional equivalently.)*

(63) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. J. Chem. Phys. 1993, 98, 56485652.

(64) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys. Rev. B. 1988, 37, 785789.

4) It is common to write "def2-SVP" instead of "def2-svp".

Author Response: *All instances have of "def2-svp" have been changed to "def2-SVP".*

Reviewer 3:

The review by Smith et al. describes the QCArchive project. This is an initiative by The Molecular Sciences Software Institute (MolSSI). QCArchive is an important resource that aims to organize, curate, and share quantum chemistry data for the benefit of the entire computational molecular sciences community. QCArchive consists of multiple fully open sources of SW modules that control specific functionalities. All are accessible via GitHub. I think this is an extremely useful review that gives an introduction about the basic functionality of QCArchive and provides short snippets of code and how to use it. This manuscript is well written and overall as a great fit for WIRES CMS.

My only minor comments would be:

1. For the benefit of users and readers, it would be nice is to supplement code snippets with links to the entire example/tutorial in the jupyter notebook.

Author Response: *We agree, and we have added to p. 12 the following link to an example:*

`https://docs.qcarchive.molssi.org/projects/QCFractal/en/stable/quickstart.html`

2. Is QCArchive unique? Are there any comparable efforts in the EU or perhaps in China? It would be great to complement conclusions with comparison (if any).

Author Response: *QCArchive is unique in the sense that the data it stores are subject to rigorous standards of units, representation, etc. However, there are certainly other existing efforts with overlapping goals, such as NOMAD and the Materials Project mentioned in the manuscript's introduction. However, to provide further context, we have added the following parenthetical with references to the introduction:*

"(e.g., ioChem-BD, the Pitt Quantum Repository, and the Computational Chemistry Comparison and Benchmark DataBase)"