**CSOF5 RP – Climate Data Analyst**  
**Interview Task**

**Data source:** <https://drive.google.com/file/d/1nJTr6luWErpXFCqfptBSysIk7jW9ktoR>

**Task 1:**   
The file included in the above link is raw climate model data. However, it is far from ready to share (even with colleagues!) due to the lack of metadata.

A verbose description of this dataset is:  
*Hourly near-surface air temperature* (in *Kelvin*) from the 1st Jan 2015, generated from the *CSIRO* model *CCAM (*version *r5262M)* for the *Australian Climate Hazards* project, over the domain known as *AUS-10i* (*12km* resolution).   
The regional climate model was driven using data from the *CSIRO-ARCCSS* global climate model *ACCESS-CM2* which performed the *CMIP6* experiment *SSP370*.  
The contact email for this dataset is *ccam@csiro.au*, and the calendar CCAM uses is *Proleptic Gregorian.*  
  
**Q.** Thinking about what kind of post-processing might be needed to make this file more machine-readable and human-interpretable, please write script (in whatever language you prefer) to amend the data file to add appropriate metadata according to the description above, with commenting in the script for other potential users.

**Task 2:**  
Version control of software is best practice in CSIRO, and a requirement for the projects in which this position will work.

**Q.** Create public repository containing the post-processing script you wrote in Task 1, and a copy of the amended CCAM data file. Include in this repository a README file with a brief description of your code, including its purpose, library sources, and an open software license.

**Task 3:**  
Let's imagine your post-processed data now needs to be published and made publicly accessible, to support a research paper that utilised this data. For argument’s sake, let’s assume that the full dataset to be published is a range of variables, over a much longer timespan, with a volume in the TB range.

**Q.** How might you go about achieving an open data publication for this dataset (i.e., what platforms & tools are required)? What principles might you follow? *(please include this answer in your script repository)*