CFIS Data Access and Collaboration: Rules & Guidelines

Rev. 2: Mar. 2, 2017

Rev. 3: August 3, 2017



This document was created and is maintained by the Canada-France Imaging Survey (CFIS) PIs and the CFIS Steering Group, with further input from the CFIS Collaboration at large upon the initial consultation. It was drafted in consultation with the SAC chair, the CFHT head of science operations, and the CADC lead.

The document will remain open for modifications over the course of the collaboration lifespan to adapt to unforeseen conditions.

Accessing the CFIS standard products outside of the CFIS Collaboration means endorsing the relevant rules described in Section 1.

Accepting to be part of the CFIS Collaboration and accessing the CFIS core and advanced products from the collaboration means endorsing all the rules and guidelines found herein.

The CFHT-MegaCam Large Program Canada-France Imaging Survey (CFIS, Feb. 2017 to Jan. 2020) is a legacy survey for the Canadian and French scientific communities.

The CFIS science Collaboration itself is a joint C+F community (mainly composed of the 2016 CFIS proposal co-Is) whose principal aims are to do great science and to ensure the production of a legacy quality data set for the public release (Feb. 2021).

Since there are more than 100 collaborators in the CFIS Collaboration and the final dataset is large and unique, it is necessary to have a set of high level rules to ensure the smooth running of the collaboration. The spirit of this document is to maximize communication across the collaboration and minimize top down management, while ensuring the collaboration members are appropriately credited for work and results stemming from the CFIS dataset. What follow describes the nature, organisation, management, and collaborative principles of the CFIS.

Sections:

| Standard dataset and access (for C&F not in the CFIS Collaboration) | p.2 |
|---|--|
| Collaboration core & advanced dataset | p.3 |
| Collaboration dataset access | p.3 |
| Scientific organization | p.4 |
| Publications | p.5 |
| Follow-up proposals | p.6 |
| Steering Group | p.7 |
| | Standard dataset and access (for C&F not in the CFIS Collaboration) Collaboration core & advanced dataset Collaboration dataset access Scientific organization Publications Follow-up proposals Steering Group |

1. STANDARD DATASET AND ACCESS: (for C&F users not in the CFIS Collaboration)

- a. The CFIS standard data that are openly available to all members of the C&F communities (scientists based in C&F institutions) consist of the regular CFHT MegaCam products: raw data and the individual detrended images by CFHT (Elixir). Access by any member of the C&F community (including those in the CFIS Collaboration) to this standard dataset at CADC implies registration through CFHT. To do so, interested scientists must email the alias cfiscf@cfht.hawaii.edu, using their home institution email and including "CFIS registration" in the subject line. The message must include an endorsement of the specific rules defined in this section 1 of the present "CFIS Data Access & Collaboration: Rules & Guidelines" document.
- b. The raw and detrended images are proprietary within the C&F communities up to January 31, 2021, and thereafter released globally by CADC in the usual way.
- c. CFIS C&F registered users agree to protect the proprietary rights of the Canadian & French community by not exporting proprietary images to persons other than CFIS C&F registered standard users, or by making public the data in digital form.
- d. Once the standard proprietary CFIS data have been processed to the point of published scientific analysis, these advanced measurements can be shared with non C&F communities as part of a pointed science collaboration. Publication implies documented analysis procedures and a scientific analysis that peer review accepts. The bulk of the intellectual effort in producing the advanced products must reside in the C&F communities, such that they are not just the result of running a non-C&F code in a C&F institution. The CFIS PIs must be consulted when anticipating such arrangements.
- e. Users of CFIS standard data should include an acknowledgement in all publications as follows (with the standard CFHT-MegaCam acknowledgement text):

This work is based on data obtained as part of the Canada-France Imaging Survey, a CFHT large program of the National Research Council of Canada and the French Centre National de la Recherche Scientifique. Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA Saclay, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii.

The rest of this present Rules and Guidelines document pertains exclusively to the CFIS Collaboration and its own products as described below.

All members of the C&F community (scientists based in C&F institutions) are welcome into the CFIS Collaboration as long as they commit to the rules and guidelines of this whole document. Interested scientists who did not sign the original proposal ought to contact the CFIS PIs via email (using their home institution email) with a brief description of their interests and anticipated involvement in the CFIS.

2. COLLABORATION CORE & ADVANCED DATASET:

a. The CFIS Collaboration core dataset will be made available as quickly as possible after each observing run, and consists of the following. The entire analysis takes place at CADC, starting over from the MegaCam raw data using the pixel layer of CFHT's Elixir (FLIPS) and Elixir-LSB, in order to produce optimally detrended frames internally to the CFIS. Those individual

frames are then fed into CADC's MegaPipe to finely calibrate the images (astrometry & photometry) and produce stacks, catalogs, masks, and quality control products. The public release (2021, considering the 3/2/1 proprietary period) will consist *at least* of the raw data and this core dataset (individual calibrated CFIS frames and MegaPipe products).

- b. Members of the CFIS Collaboration will naturally produce through various scientifically motivated investigations their own set of products, either advanced, complementary or parallel to the core dataset. Members of the CFIS Collaboration commit to share their products with the entire CFIS Collaboration (through the CFIS CADC repository) in a timely manner although there is no fixed schedule for those products. The well documented efforts stemming from a published work could become part of the 2021 public release.
- c. The CFIS core dataset includes all data obtained previously as part of the Luau observing campaign (2015 2016, PIs Ibata and McConnachie). It also includes the public data acquired as part of the usual MegaCam calibration routine by CFHT (flat-fields, bias, misc. photometric calibration images) gathered over the lifetime of the survey.
- d. CFIS data does not include by default data obtained through follow-up proposals to the main CFIS dataset.

3. COLLABORATION DATASET ACCESS:

- a. Only members of the CFIS Collaboration are able to access and work with data produced by the CFIS Collaboration while the data is proprietary.
- b. An exception is where a Memorandum of Understanding (MoU) has been signed between an external party and the CFIS Collaboration (with an endorsement from the CFIS PIs & Steering Group). In this case, access to CFIS Collaboration products by the external party is under the rules described by the MoU, and can be unique for each case.
- c. CFIS Collaboration rights extend to C&F institutions students and postdoctoral research associates directly associated to a CFIS Collaboration member. If a scientist leaves to a non-C&F institution, the CFIS Collaboration membership may be extended by the PIs for those who have made a significant contribution (typically those who have spent at least 50% of their research time on CFIS).
- d. Access to CFIS data produced by the CFIS Collaboration will be handled by CADC, with a mirror at CCIN2P3 to ease data retrieval in France.

4. SCIENTIFIC ORGANIZATION:

- a. The CFIS Collaboration scientific organization will consist of a collection of Science Projects (SP) of various sizes leading their own scientific research. A Science Project can range from a specific limited topic (leading to one paper) to a broader theme involving a large effort but still tailored to lead to a specific suite of papers. The CFIS Science Projects are expected to self-organise into the most efficient model for producing front-line science, adhering and being consistent to the principles, guidelines and rules described in this primary document.
- b. These Science Projects are primarily intended as a way of initiating collaborations on specific topics and developing plans for projects/papers by being able to include in the discussions, from the outset, all members of the CFIS Collaboration that have interests in these science

areas. They are the primary teams responsible for doing science with CFIS data. The lead of each Science Project is by default the contact point for the CFIS PIs.

- c. It is adequate for the CFIS Collaboration to operate with competitive Science Projects as there is no reserved science
- d. All proposals for a new Science Project must be presented by email to the CFIS PIs and the Steering Group, and will be accepted on the condition it is done well in advance of the analysis and drafting of the resulting paper and that it complies with the present set of rules. As new Science Projects are formed, the CFIS internal pages will be updated and announcements will be circulated to the entire CFIS Collaboration via a monthly newsletter. To sign up to a Science Project, a CFIS Collaboration member has to contact the lead person who reserves the right to refuse a new collaborator (contentious cases can be brought up to the CFIS PIs and Steering Group).
- e. Members must self-identify early on, and keep an up-to-date status, of their interests and intentions for each Science Project on the CFIS internal pages ("CFIS Collaboration: Who is doing what?"). This reference will be key to the healthy scientific life of the CFIS Collaboration. Regular newsletters (typically bi-monthly) will be circulated by the CFIS PIs that will include a current summary of all the Science Projects and participants active in the collaboration.
- f. It is expected that the structure of these teams will evolve as the survey proceeds, and that members of the Science Projects will be best placed to identify who does what and with whom, especially with respect to leading the science that will result in publications.
- g. It is anticipated that numerous students will work on CFIS data as part of a graduate thesis. The CFIS PIs and Steering Group strongly encourage student involvement in CFIS and are highly supportive of their involvement in, or leadership of, Science Projects. In such cases, it is essential that the CFIS Steering Group is informed of the student involvement, and that up-to-date information on the thesis topic and specific research projects involving the student is maintained on the CFIS Internal pages. Projects involving students will be highlighted in the newsletters. All CFIS collaborators will be cognizant of student involvement in projects in which they are involved or which may overlap with their research interests.
- h. Science Projects may need to involve people who are not currently part of the CFIS collaboration in order to fill gaps in expertise. In such cases, an email should be sent to the CFIS PIs that clearly identifies the roles of these people in the Science Project.

(i) If data access is not required by these individuals, then this should be stated in the email to the CFIS PIs. In such cases, the lead of the Science Project will ensure participation of the individuals is in a manner consistent with the rules and guidelines of this present document.

(ii) If limited access to the data is required by these individuals (i.e., to use a specific data product for a specific purpose towards the ultimate goal of contributing to a publication), then the lead of the Science Project should contact the CFIS PI and describe the reasons why data access is required and the anticipated scope of the data access. The CFIS PIs will determine whether limited data access is granted for the specific purposes described, and for a limited time interval ultimately leading to publication of results in a manner consistent with the CFIS publication guidelines.

(iii) If full data access is required for a long duration, the reasons should be described in an email to the CFIS PIs. The CFIS PIs will determine, in consultation with the Steering Group, whether to extend them full membership rights, with the explicit expectation that the individuals then adhere to the rules and guidelines of this present document. Generally, membership will only be granted to non-C and F members when it is clear that the prospective new member(s) will bring expertise or other value-added benefits not otherwise available in the collaboration, and where it is clear that extending membership is to the scientific advantage of the CFIS collaboration (e.g., by enabling new projects and publications, or vastly increasing the competitiveness of CFIS in a science area).

5. PUBLICATIONS:

- a. CFIS Collaboration publications are considered those publications that are based wholly or partly on CFIS proprietary data generated by the CFIS Collaboration.
- b. As an intended legacy type survey, there is no automatic right to anyone to authorship on CFIS publications. Instead, those who contribute significantly to a given paper can be granted authorship. "Contributing" could be any of: proposal writing, data reduction, analysis, interpretation, writing, etc, as long as it is deemed a significant contribution to the project.
- c. It is expected that some members have and will contribute more than others to the project and/or specific papers, and will therefore deserve to be on a paper when they so request. For the purposes of clarity, these members include individuals or groups responsible for acquiring and generating high level (collaboration-wide) data products such as images, stacks or object catalogs that are widely used. At the level of specific Science Projects, these members also include individuals or groups responsible for the creation of advanced products that may be used by the Science Project (e.g., galaxy shape catalogues for weak lensing, stellar metallicity catalogues, etc.).
- d. It is expected that there will be a few overview papers that will discuss the broad goals of CFIS, including details of data reduction. These papers will be treated as an exception to this guideline and authorship will be offered to the entire CFIS Collaboration, who must individually confirm that they want to be listed as a coauthor.
- e. It is expected that a paper will be written by a core group of authors who agree among themselves how they are to be ordered after the lead author, possibly followed by one or more CFIS Collaboration members, as appropriate. These other authors are to be listed in alphabetical order. If there is disagreement as to whom should be included as authors, the CFIS PIs should be consulted by the lead author.
- f. All CFIS papers to be submitted will be first posted on the CFIS internal pages once the draft is at a stage where it could be submitted to the journal. A notice will be circulated to the entire collaboration. CFIS Collaboration members who are not already a co-author have a 3-week period from date of circulation to make a contribution significant enough to the draft and make the case to the lead author for authorship. The lead author will give full consideration to coauthors' inputs to the paper. Note that it is generally expected that all people who are interested in the science will already have been identified as coauthors through their participation in the relevant Science Projects (Section 4), and this present guideline is intended as a failsafe mechanism. As such, extensions to the 3 week period will not in general be considered.
- g. The title of all papers arising out of the use of CFIS Collaboration dataset should include either the full survey name "Canada-France Imaging Survey" or its acronym "CFIS". All papers will acknowledge the help of the CFHT staff in taking the data. Appropriate credit to the CFIS Collaboration for running the survey and producing the data set is expected along the usual required acknowledgment text for MegaCam and the CFHT funding agencies (*c.f.* CFIS internal pages for the relevant texts).
- h. The CFIS PIs with the CFIS Steering Group will act as a form of light-handed editorial board, for example, ensuring the publications comply with the present rules and are using the most up-to-date data. It will have final right of approval over all CFIS publications based on the CFIS Collaboration dataset to be submitted to peer-reviewed journals, which they will convey to the lead author during the 3-week pre-submission period. The default answer will be positive and

no publications will be blocked based on scientific analysis or interpretation; rather, consideration will only be given to respecting the collaboration rules and guidelines, and ensuring the integrity of the dataset (e.g., in case of fundamental misinterpretation of catalogues or incorrect versions of the dataset being used).

- i. All articles for publication in peer-reviewed journals should be posted to the arXiv. It is left to the lead authors, in consultation with their coauthors, to decide if the first posting of an article should occur after submission or acceptance of the paper to the journal.
- j. Any CFIS member presenting a poster, talk, communication, etc. is encouraged to report it on the CFIS internal pages, as well as posting afterward on those pages a PDF of their presentation or proceeding.
- k. Confidentiality regarding unpublished results is expected from the CFIS Collaboration. Consultation should be made with the science lead in situations where unpublished material is going to be shown.

6. FOLLOW-UP PROPOSALS:

- a. Follow-up proposals stimulated in whole or in-part by CFIS can be submitted by any member of the CFIS Collaboration.
- b. Any follow-up proposal that is submitted must be shortly documented and the proposal posted to the CFIS internal pages and an announcement circulated to the collaboration via the email exploder. Any member of the CFIS Collaboration who wants to be included as a co-I should contact the relevant PI with a justification.
- c. The purpose of this guideline is to ensure that all CFIS Collaboration members have the opportunity to be involved in proposals stemming from the CFIS dataset to which they have contributed. However, data obtained as part of the follow-up proposal are not CFIS data and are not subject to CFIS oversight.
- d. Inclusion of non-members in CFIS follow-up proposals is at the discretion of the relevant proposal PI, but does not in any way imply any kind of access to the CFIS dataset. This must be made clear to the non-members at the outset of the project.
- e. If access to the CFIS dataset will be required by non-members as part of the follow-up proposal, then this will first be discussed between the CFIS Steering Group and the proposal PI.

7. STEERING GROUP:

- a. The CFIS Steering Group is advisory to the CFIS PIs. The CFIS Steering Group will meet regularly (typically via monthly telecons) with the CFIS PIs, to discuss any and all aspects of the CFIS Collaboration, especially with respect to management, strategy and internal communications. As a whole, the role of this management group is to ensure the CFIS Collaboration inherits the best possible data set to produce exciting science in a fruitful community environment, and to ensure the rules are respected by all.
- b. When considered useful or necessary, other CFIS Collaboration leads or members will be asked to participate in a Steering Group meeting. Any issues pertaining to the CFIS Collaboration that members want addressed should be directed to the CFIS PIs or any member of the CFIS Steering Group (with respect to their responsibilities, list below) who will propose an agenda item for the next meeting.
- c. The CFIS Steering Group (CFIS-SG <<u>cfis-sg@cfht.hawaii.edu</u>>) consists of:

- PIs: Jean-Charles Cuillandre (F) & Alan McConnachie (C)
- Science leads: Mike Hudson (WIQD), Rodrigo Ibata (LUAU)
- Data management: David Schade
- Canadian members at large: Michael Balogh, Ray Carlberg
- French members at large: Vanessa Hill, Yannick Mellier