

The FIR-Data Gap and how the SOFIA Data Center (SDC) can help

Bernhard Schulz, Aaron Bryant, Christian Fischer, Nadine Fischer, Jonas Früh, Benjamin Greiner, Michael Hütwohl, Christof Iserlohe, Thomas Keilig, Alfred Krabbe, Bastian Knieling, Karsten Schindler, Andreas Siggelkow, Rainer Strecker, Manuel Wiedemann, Oliver Zeile
SOFIA Data Center (SDC) / Deutsches SOFIA Institut (DSI), University of Stuttgart, Germany

Abstract

During 783 scientific flights, SOFIA, the Stratospheric Observatory for Infrared Astronomy of the space agencies of Germany and the United States, DLR and NASA, respectively, collected a considerable volume of scientific data, that is available through the Infrared Science Archive at IPAC. After the end of flight operations in September 2022, with only one year of ramp-down time for Science Mission Operations remaining, only a limited data reprocessing of SOFIA Observing Cycles 5 to 9 (2017 to 2022) could be achieved. In order to complete the job and ensure a full exploitation of the considerable investment in SOFIA, the Institute of Spaceflight Systems of the University of Stuttgart, the home of the German SOFIA Institute (DSI), submitted a proposal to DLR to create a SOFIA Data Center (SDC) with a more comprehensive plan for a five-year post operational phase. The proposal was accepted and funded and the project started in July 2024.

This contribution will make a brief assessment of the history and outlook of Far-Infrared (FIR) astronomical observatories, and discuss the SDC's plans to perform a more comprehensive reprocessing of all stored SOFIA data. We will use upgraded data processing pipelines, with improved pointing reconstruction, derived from guide camera imagery and new corrections for atmospheric water vapor and instrumental effects. The science data - re-ordered by astronomical criteria - will be stored along with the technical and operational engineering data of the telescope in a searchable, VO-compatible, publicly accessible archive.

The newly formed SDC draws from the experience of the former DSI and seeks collaboration with scientists and engineers, that still hold knowledge of SOFIA operations and instrumentation. Once the SDC work is complete after 5 years, the archive will be transferred to the new German Astrophysics Center (DZA) as a long-term home. We expect to work with the entire astronomical community to promote the exploitation and publication of valuable Mid- and Far-Infrared data, establish a powerful scientific resource for the post-mission phase, and bridge the gap until the next space- or stratospheric FIR observatory becomes available.

Aims of the SOFIA Data Center

- Creating a powerful scientific resource for the time after the mission
- Bridging the expected 10-20 year gap in regular new FIR data availability
- Ensuring the quality of the SOFIA data archive and its documentation to secure its scientific productivity for the long term
- Enabling future scientific data exploitation without the need for expert knowledge of SOFIA subtleties
- Completing the work within 5 years with subsequent transfer to DZA
- Staying in sync with the current SOFIA data repository hosted within IRSA

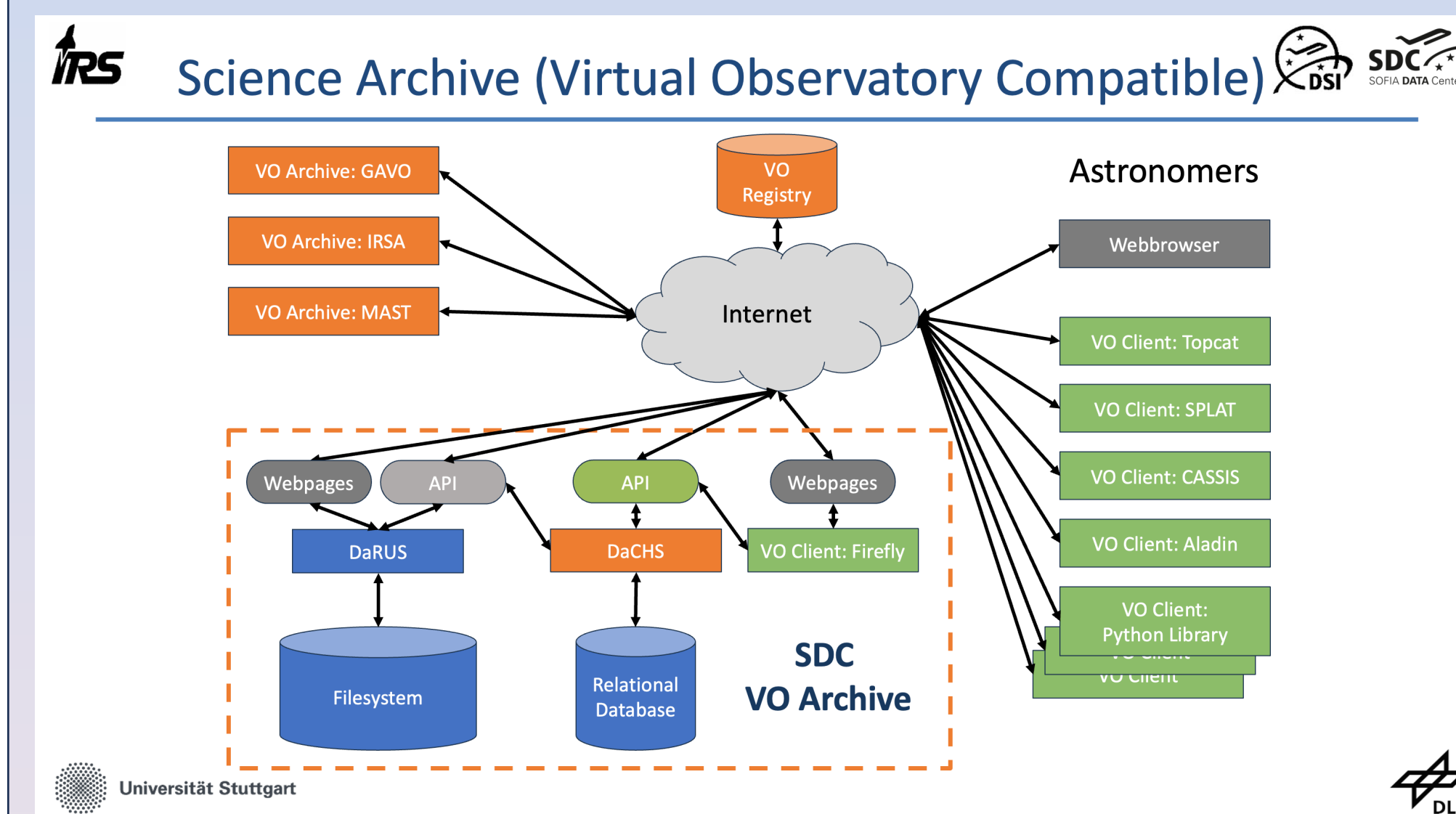


SOFIA during flight testing with chase plane (Credit: NASA/DLR)

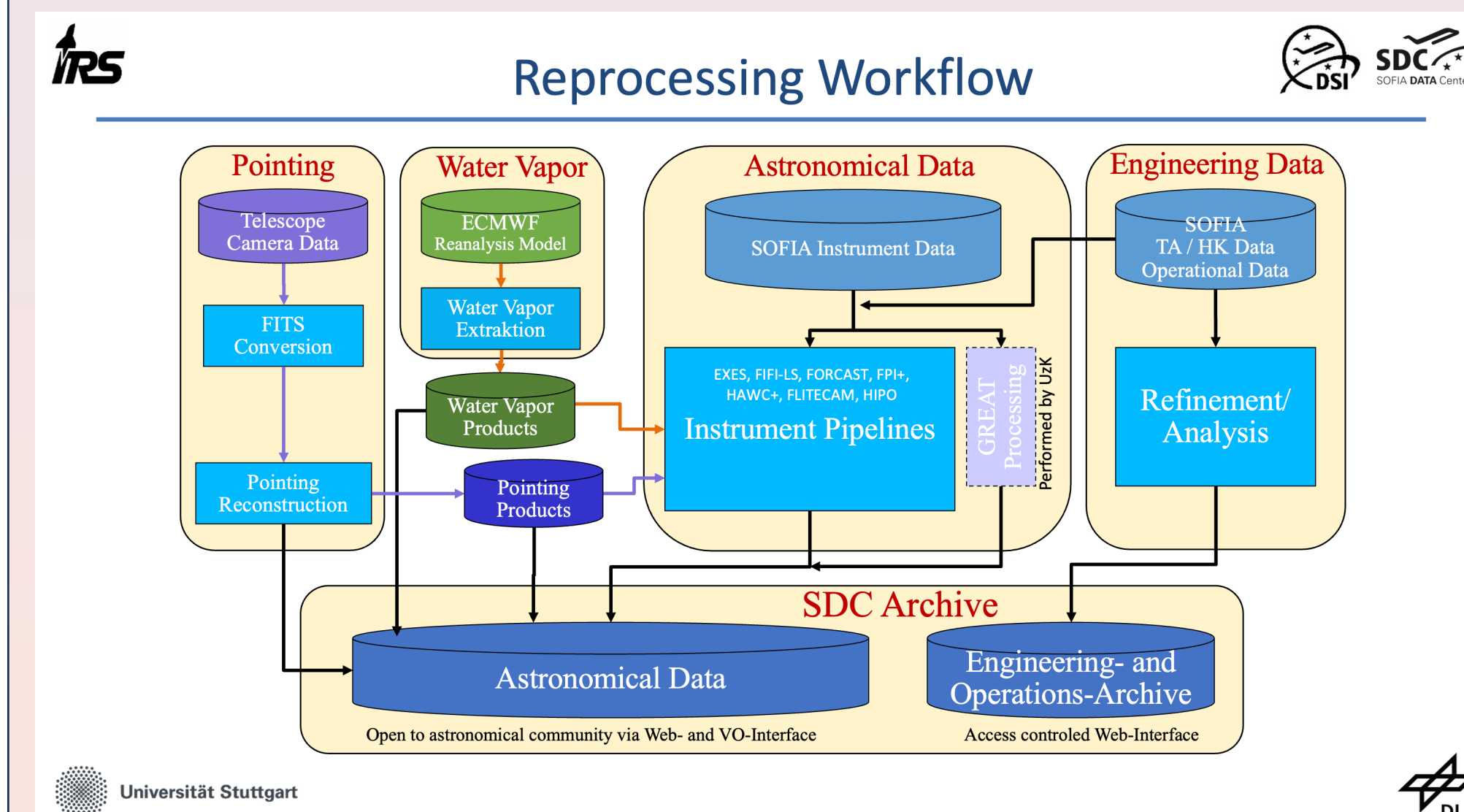
Main Tasks

- Update of metadata to enable reprocessing of Cycles 0 - 4
- Create an automatic reprocessing environment allowing for an iterative approach
- Reprocess instrument data (GREAT instrument will be separately done in Cologne)
- Iteratively implement additional pipeline improvements
- Develop supporting software tools
- Create a SOFIA Science Archive hosting raw and processed data
- Reorganize science data rather by astrophysical criteria than operational ones
- Create an Operations- and Engineering-Archive
- Update documentation where necessary
- Support scientists through webinars, conferences and direct advice
- Support engineering research by interested parties

VO-Archive and Data Reprocessing



The SDC Archive is planned to be VO compatible to enable a broad exposure of the data to science searches and minimize the need for new software development through use of available VO-client and server software. DaCHS (Demeitner 2014) will be used to create the VO-archive and provide an API to the internet. The Firefly software suite (Roby et al. 2013) will be used to access this API as a VO-Client and create the archive webpages.

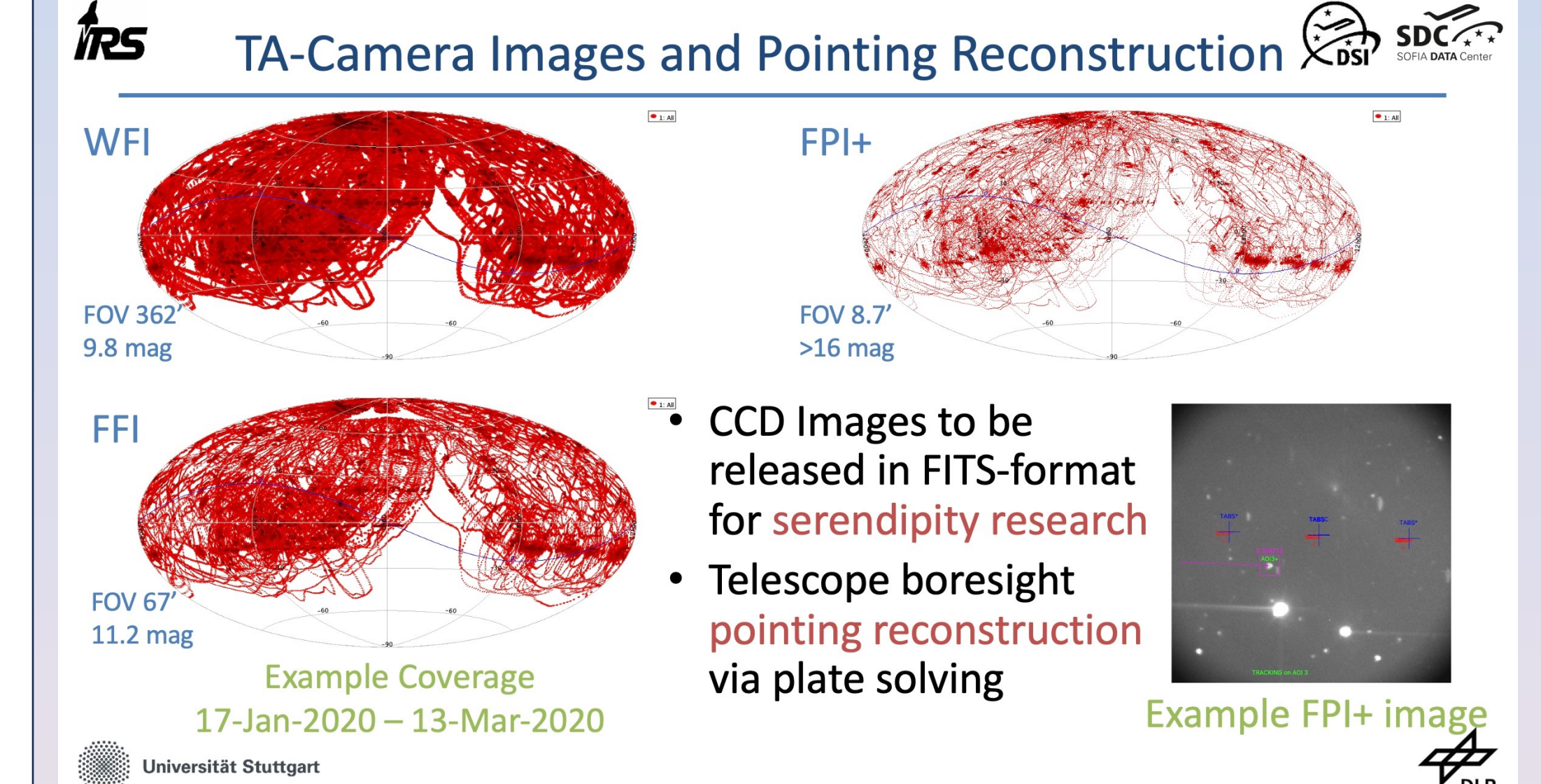


The SDC will build an automatic reprocessing system that can fully process a set of related observations into a final higher level product without user interaction. It will be using the instrument pipelines originally constructed by the SMO as well as additional new modules that integrate improved pointing information, corrections for water vapor and other improvements into the processing. The necessary pipeline parameters will be stored in the database for each group and can be refined as our understanding of the data improves. The grouping of related observations will be driven by scientific considerations. Groupings that arose only from operational constraints will be eliminated.

Summary

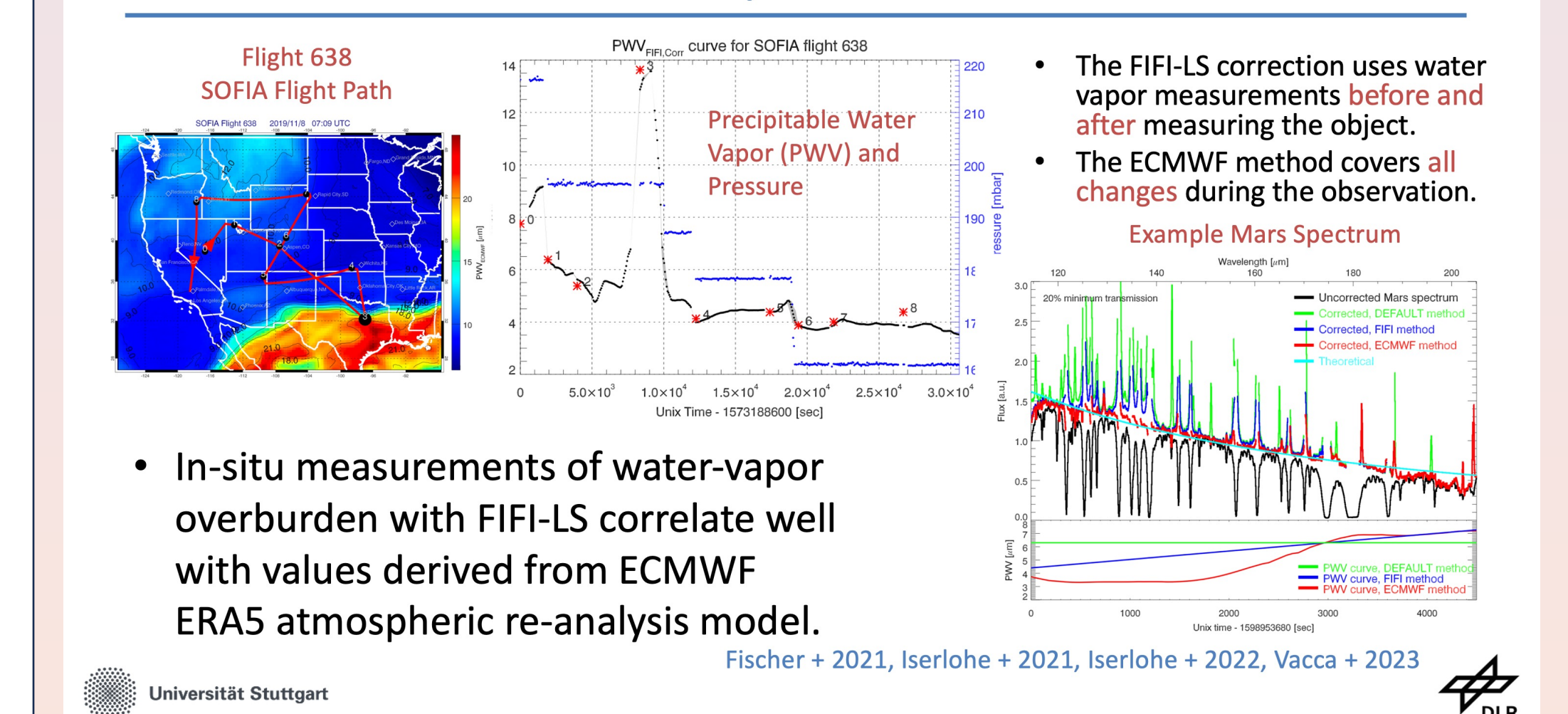
- The SOFIA Data Center (SDC) at the University of Stuttgart will provide an adequate post-processing and archival phase for SOFIA.
- The active phase for the SDC is 5 years.
- Within this period the data will be substantially improved, be placed into a well searchable and documented online-archive, and scientists will receive expert support in understanding residual instrumental effects and finding valuable, yet unpublished data.
- At the end of this period the archive contents are planned to be transferred into a permanent archive at the DZA.

Improved Pointing and Atmospheric Water Vapor Correction



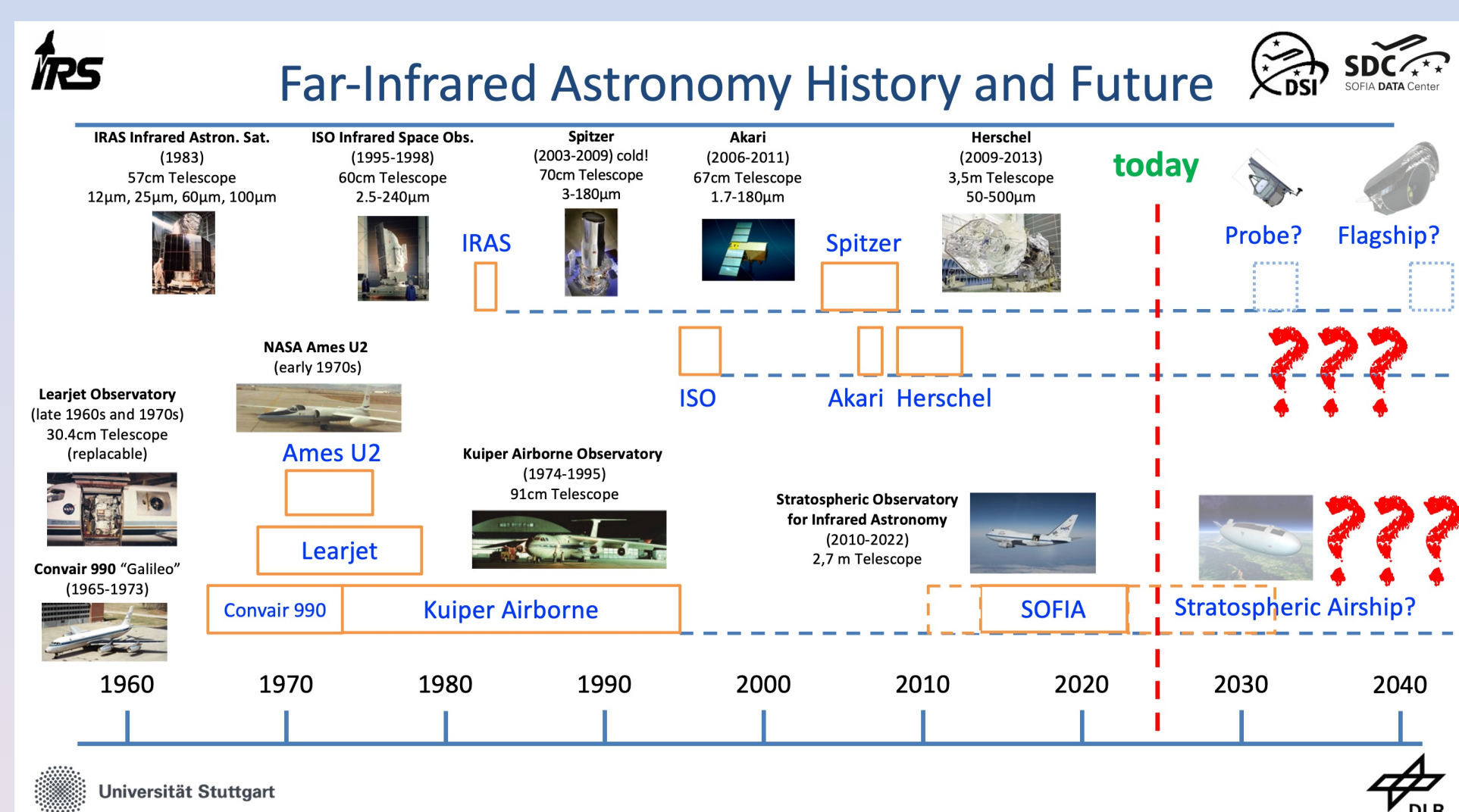
The three CCD cameras WFI, FFI, and FPI+ (Pfüller et al. 2018) of the telescope assembly (TA) recorded images through all science flights. The SDC will provide the images with FITS headers, publicly release them and generate detailed telescope pointing reconstruction products for every flight that can be used by the data processing pipelines.

Water Vapor Calibration



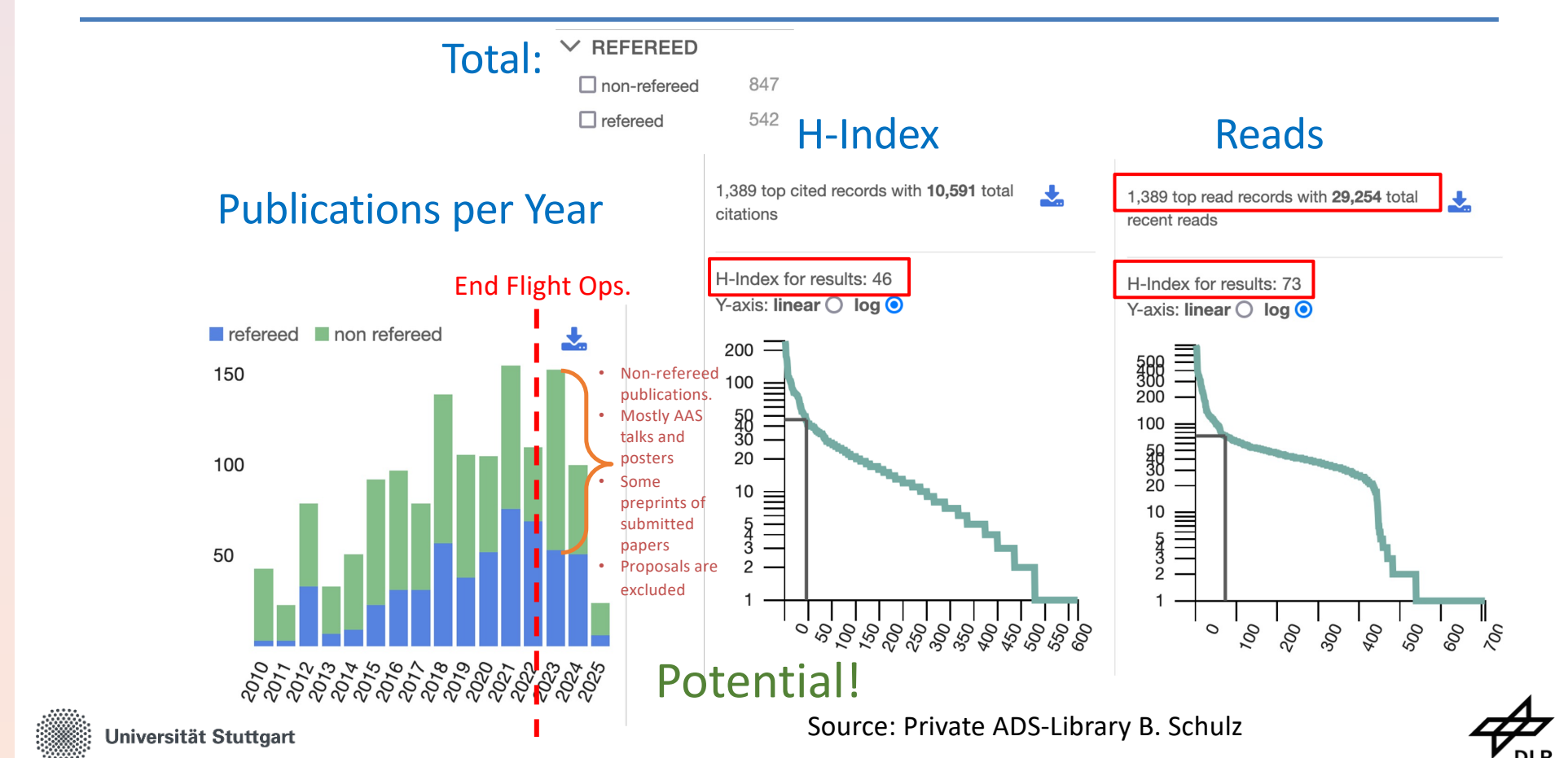
A new method to calibrate the atmospheric transmission mainly due to its water vapor content is based on a global atmospheric reanalysis provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). The method was validated using dedicated, in-flight FIFI-LS observations. It can provide corrections for all data from all SOFIA flights even after the fact (Fischer et al. 2012, Iserlohe et al. 2021, Iserlohe et al. 2022, Vacca et al. 2023). In particular, data from instruments with low to intermediate spectral resolution will benefit from an implementation of this method into their corresponding processing pipelines.

Motivation

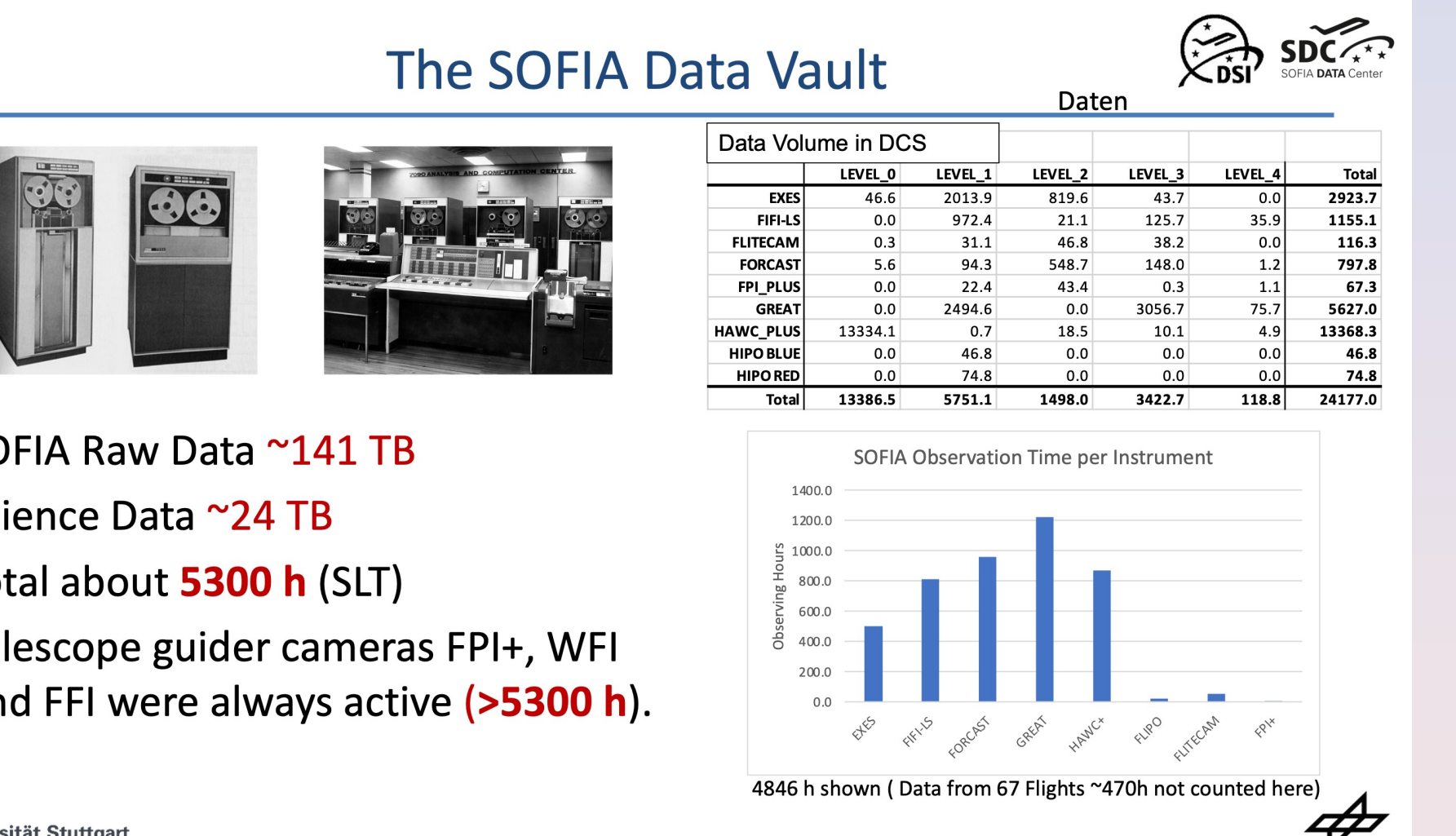


With its beginnings in the early 60s, the young field of infrared astronomy went through an exciting time with a number of very successful stratospheric and space-based observatories. However, the premature end of SOFIA, the last regularly scheduled observatory sensitive to the FIR, after less than half of its design lifetime of 20 years, leaves the field with an uncertain future. The earliest realistic chance for an infrared space mission is the NASA Probe mission, that is currently being competed for by an infrared and an X-ray mission.

SOFIA Publication Status 7.3.2025



SOFIA's yearly publication statistics was steadily rising while there were observing flights and support for scientists through the Science Mission Operations Center (SMO). A considerable fraction of the data yet waits for publication. Active work on data improvement and advertising the archive contents among the scientific community will help fully exploit the investment made in this mission.



SOFIA's eight science instruments collected a considerable amount of data during her operational lifetime. The data comprises the instrument science data, instrument and observatory housekeeping data and all images taken in parallel by the three telescope guider cameras. After observing flights ceased, there was only limited time and funding to perform data reprocessing. These activities focused on Cycles 5 - 9 and ended after one year.

References

- Young, E. T., Becklin, E. E., Marcum, P.~M., et al. 2012, ApJL, 749, L17
- Roby, W., Wu, X., Ly, L., et al. 2013, Astronomical Data Analysis Software and Systems XXII, 475, 315
- Demeitner, M. 2014, Astronomical Data Analysis Software and Systems XXIII, 485, 309
- Clarke, M., Vacca, W. D. & Shuping, R.~Y. 2015, Astronomical Data Analysis Software and Systems XXIV, 495, 355
- Pfüller, E., Wolf, J., Wiedemann, M. 2018, Journal of Astronomical Instrumentation, Volume 7, Issue 4, id. 1840006
- Fischer, C., Iserlohe, C., Vacca, W., et al. 2021, PASP, 133, 055001
- Iserlohe, C., Fischer, C., Vacca, W. D., et al. 2021, PASP, 133, 055002
- Iserlohe, C., Vacca, W. D., Fischer, C., et al. 2022, PASP, 134, 085001
- Vacca, W. D., Iserlohe, C., Shenoy, S., et al. 2023, PASP, 135, 085001