

COMMUNITY EDUCATION TO FOSTER FACILITY ENGAGEMENT:

Schools and Data Workshops to Enhance the Scientific Return of DKIST

G. Cauzzi

National Solar Observatory, 3665 Discovery Drive, Boulder, CO, USA

gcauzzi@nso.edu

Synopsis

The National Science Foundation's Daniel K. Inouye Telescope (DKIST) has recently commenced operations, and soon data will become available to the community at large. With this whitepaper, we advocate for a continuing program aimed at informing and educating the US solar community on high-resolution solar imaging, spectroscopy, and polarimetry from the ground, in order to properly exploit the incredible scientific opportunities provided by this facility. We use the example of a recent series of Schools and Workshops organized by the National Solar Observatory to derive suggestions on how to best structure this program.

Introduction

The recently inaugurated Daniel K. Inouye Telescope (DKIST) of the National Science Foundation is the largest optical solar telescope in the world. Built to observe the solar atmosphere and the inner corona at unprecedented spatial resolution, with high photometric and spectropolarimetric sensitivity, DKIST commenced scientific operation in early 2022 with the so-called “Operation Commissioning Phase”, aimed at testing the full life cycle of telescope operations, from proposal submission to execution of observations in service mode, to delivery of Level 1 data (instrument calibrated) to Principal Investigators. More details are available in Rimmele et al. (2020); Alexov et al. (2020); and Davey (2021).

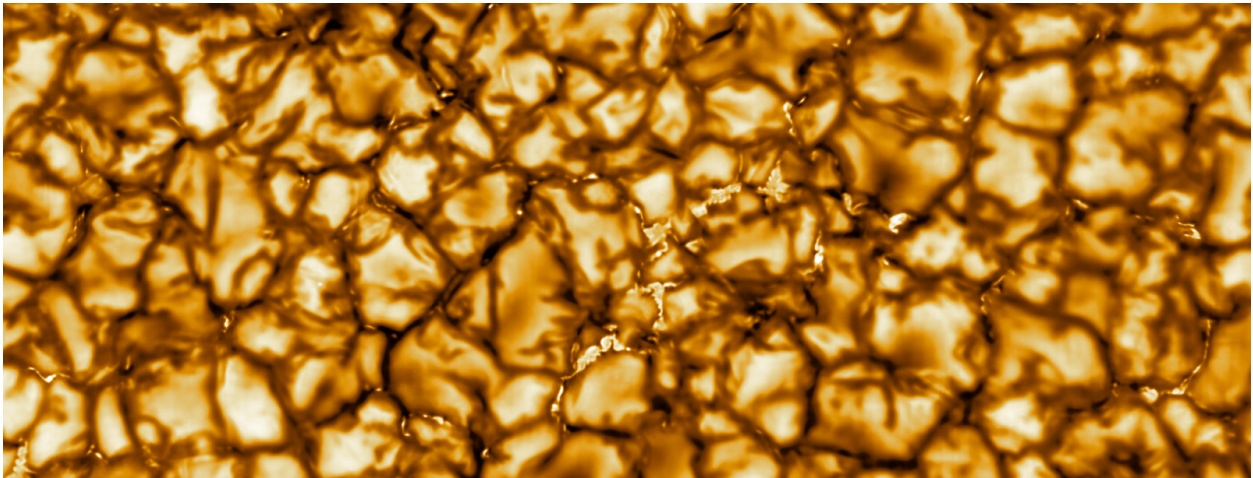


Figure 1: Portion of the DKIST first light image, acquired in the continuum at 789 nm (<https://nso.edu/inouye-solar-telescope-first-light/>). The area covered is 24 x 19 Mm on the Sun, with details at the diffraction limit of 30 km.

While the DKIST first light data (Figure 1) and subsequent images are striking, these tend to obscure the amount of processing that goes into generating the highest resolution images. Simple broadband images are also conceptually much simpler than the complex 5-D spectropolarimetric data produced by other instruments. Many aspects of the data to be produced by DKIST will be novel for a large part of the US solar community. Most researchers are currently more acquainted with high energy observations and diagnostics such as coronal emission in EUV, soft- and hard-X-rays (Yohkoh/SXT, Hinode/XRT, RHESSI, SDO/AIA); UV and EUV spectroscopy (SOHO/SUMER, Hinode/EIS, IRIS); and in-situ measurements of the solar wind and heliosphere (PSP). In addition to the wavelength differences, it is important to note that all these measurements originate from space-based facilities. Even the most widely used visible-light instrument, SDO/HMI, is a space-based instrument that provides a constant dataset already reduced to physical parameters. There are significant and important differences between these well-used datasets and the outputs from the various DKIST instruments. Among the important

observational issues that need to be considered when handling and, more crucially, interpreting DKIST observations, we highlight:

- **Effects due to the Earth’s atmosphere:** changes in transparency, differential refraction, atmospheric turbulence (seeing), scattered light. These affect spatial resolution, photometric and polarimetric sensitivity, characterization of motions, as well as our capability to properly observe coronal features;
- **Optically thick visible and IR diagnostics:** most of the spectral diagnostics observed on-disk by DKIST are optically thick, which implies the need of a proper treatment of radiative transfer, including non equilibrium effects (non local thermodynamic equilibrium; ionization equilibrium; chemical equilibrium; time-dependent ionization);
- **Emphasis on polarimetry:** this requires a proper understanding of the mechanisms that lead to the creation of polarized light (scattering polarization; Hanle, Zeeman, Paschen-Back effects), as well as an understanding of the complex calibrations required to provide reliable I, Q, U, V spectra within an extended field of view (FOV);
- **Inversion techniques:** most of the important physical parameters in the solar atmosphere (macroscopic velocity, magnetic field, density, temperature) are not immediately derivable from the observed spectra, but need to be inferred from so-called “inversions”, i.e. from solving the (ill-posed) inverse problem of how the observed radiation is created in the presence of such physical conditions. While these techniques are a powerful analysis tool for much of the DKIST observations, there is still only limited expertise in this area among US solar physicists (cf. also the whitepaper “Spectroscopic and Polarimetric inversions” by Reardon et al.).

Finally, because of the fundamental constraints on the available amount of usable time (due to the day-night cycle, as well as weather and seeing conditions), there is more competition for observing time than is typical for space-based satellites. DKIST is operated like other large astronomical facilities, e.g. Keck, ALMA, Gemini, with observing time awarded on the basis of annual or bi-annual proposals and proprietary periods for data (six months). DKIST also has a very flexible way of configuring instruments and the observing schemes, that allows PI’s to define innovative and unique “experiments” tailored to their specific scientific questions. Such a flexibility however betrays a significant complexity, and proposals need to clearly state the connection between the proposed observational approach and the scientific return (why is the defined scheme an “optimal” approach). Since the number of configuration options is so large with DKIST, even compared to a relatively flexible space-based instrument like IRIS or Hinode/SOT, the proposing scientists are responsible for properly devising their observational strategy. This means they must become familiar with more instrumental details and the types of parameters that can be derived with different approaches.

For these reasons, *we advocate for a continuing program aimed at informing and educating the US solar community on high-resolution solar imaging, spectroscopy, and polarimetry from the ground, in order to properly exploit the incredible scientific opportunities provided by DKIST.* This is consistent with the recommendations of the Decadal Survey in Astronomy and Astrophysics 2020 (Astro2020): “...*A significant investment is needed in the people who will train the community on the new products and tools*” (Section H, page 3; Pathways to Discovery in Astronomy and Astrophysics for the 2020s, <https://nap.nationalacademies.org/catalog/26141/pathways-to-discovery-in-astronomy-and-astrophysics-for-the-2020s>), and will represent a low-cost investment on the scientific return of a major NSF facility.

In envisioning the evolution of such a program of engagement and training, we build on the experience gained in the past few years, as detailed below.

The example of the NSO “DKIST Data Training workshops”

Having recognized the needs outlined in the previous Section, over the last three years the National Solar Observatory (NSO) has organized and conducted several “DKIST Data Training Workshops”, with the overarching goal of starting to prepare the US solar community for the wide variety of data expected from DKIST. A summary of the five workshops can be viewed at: <https://nso.edu/ncsp/ncsp-workshops/>. As a complementary activity, NSO and the High Altitude Observatory (HAO) jointly organized two intensive schools (each two weeks long) on “Solar Spectropolarimetry and Diagnostics Techniques”, with emphasis on the theoretical foundations of polarimetry and radiative transfer. (The second school was held in August 2022, see: <https://www2.hao.ucar.edu/events/workshop/spectropolarimetry-2022>).

The DKIST Data Training Workshops addressed a variety of topics, including the nature of ground-based solar observations, the analysis of high-cadence imaging data, spectro-polarimetric observations in various regimes, and common data reduction and interpretation techniques. Emphasis has been given to advantages and limitations offered by different diagnostics. All workshops lasted 4-5 days, and followed a similar structure, with 1-2 hour long lectures on a specific issue, followed by extensive hands-on exercises. To this end we used existing data from several ground-based telescopes, similar to what DKIST will provide, and made large use of Python notebooks as a way to guide the exercises. All activities were recorded and made publicly available, so that not only the original participants but also the overall community can revisit them at will.

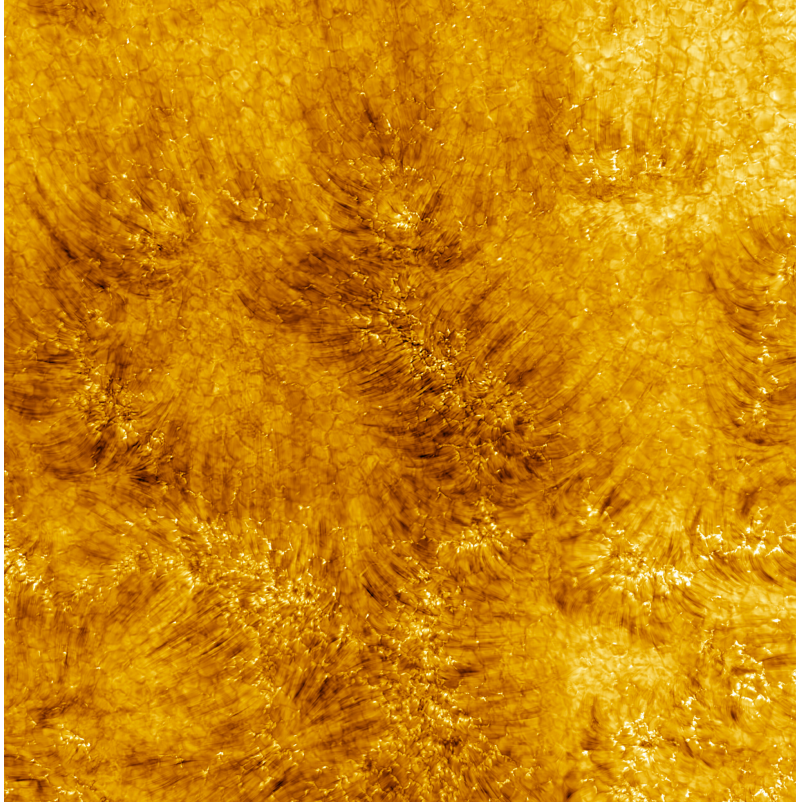


Figure 2: A stunning view of the solar chromosphere, as observed with the H β filter of the VBI instrument on DKIST, on June 3, 2022. The overall field of view is over 80 Mm on the side, with a resolution of 18 km. During the DKIST Data Training Workshop #4 (“Introduction to Chromospheric Diagnostics”, https://nso.edu/ncsp/ncsp-workshop/intro_to_chromosphere/), we discussed for example the formation and dynamics of the fibrils extending from the photospheric magnetic (bright) elements.

While open to everyone, the DKIST Data Training Workshops have been tailored to an early-career audience of PhD students and recent postdocs, with the declared aim of cultivating a new cadre of researchers that can take full advantage of the unprecedented volume of diverse observations expected from DKIST. To date, the experience of the attendees has been overwhelmingly positive. On average, about 40 young US researchers have participated in each Workshop, with many of them participating in multiple events. Most importantly, over 90 unique US participants have attended, from most of the US institutions (mainly Universities) where solar research is conducted. Post-attendance surveys have confirmed that the students were very appreciative of the format and contents of the workshops, and, most interestingly, that about 50% of them partially steered their research towards topics and methods presented during the workshops.

A continuation of this experience in the coming years, revisiting previous themes for new students, as well as offering new topics adapted to the evolution of the data types generated by DKIST, will be a highly effective means to broaden the quantity and quality of community participation in planning and analyzing observations with the telescope.

Pandemic lessons

Some of the DKIST Data Training Workshops were held during the COVID pandemic, and as such their format evolved to a fully virtual experience. While initially dreaded (mostly by the teachers), this format has drawn many positive remarks from the participants. As expected, the virtual setting has been detrimental to forming strong personal relationships among the students, but it has also opened up more possibilities, including allowing the participation of “non-traditional” attendees (such as international students as well as more experienced researchers) who otherwise would not have had the resources or time to attend in person. It also greatly facilitated the organization of the workshops, eliminating many of the logistical problems (while introducing a few technical ones) and greatly reducing the associated costs for travel and lodging. This format could also allow organizers to better reach out to, and promote involvement of, underrepresented groups in the field, such as students at non-traditional or teaching-only institutions. Of note, we learned that recording everything (including exercises) and making it available in almost real time is extremely useful to ensure that people fully participate in the workshops’ activities, as they can revisit the lectures at will and at their own pace.

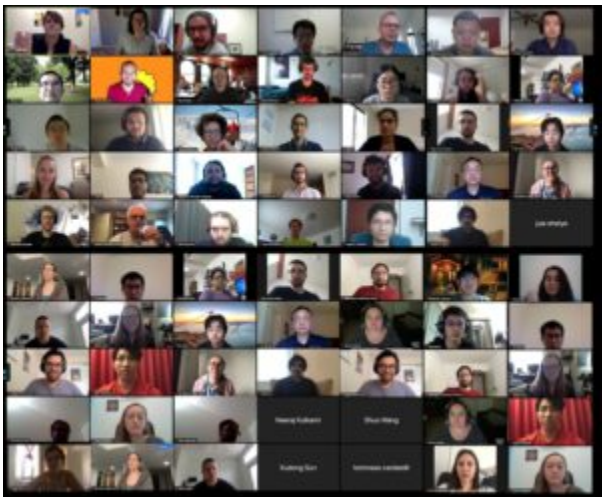


Figure 3: Left: participants to one of the virtual Data Training Workshop in 2021s. Right: H. Uitenbroek of NSO teaching in person at the 2nd “Solar Spectropolarimetry and Diagnostics Techniques” school, August 2022.

A judicious mixture of in-person and remote workshops will represent an excellent way forward, balancing the advantages of in-person training and networking (especially important for the younger generation) with the need to contain the related costs and broaden participation among diverse groups. As an example, on-line activities could focus on proposal preparation, or basic diagnostic techniques (Milne-Eddington inversions, Weak Field Approximation; etc., see e.g. DKIST Data Workshops 3 & 4) while in-person gatherings could be used for actual data analysis, or complex inversions.

Moving forward

DKIST has commenced operations, and soon increasing volumes of data will become available to the community at large (initial proprietary periods may start to expire at the start of 2023). Data Training Workshops addressing different aspects of **real data** and their interpretation will be an extremely effective way to engage US researchers with the facility, and guarantee that the large NSF investment in it will quickly start to produce exciting scientific returns. To highlight parts of the the Astro2020 recommendations that are particularly relevant here:

“NSF and the National Science Board should consider actions that would preserve the ability of the astronomical community to fully exploit the Foundation’s capital investments in ALMA, DKIST, LSST, and other facilities.”

And: “[T]he data alone are not enough. A significant investment is needed in the people who will develop the theoretical framework, build the archives, develop the software, **train the community on the new products and tools**, create and implement the computational methods, make the vital laboratory measurements, and analyze the data to produce these transformative results.” [emphasis ours]

We thus advocate for a consistent, continuing program, with stable funding, aimed at informing and educating the US solar community on the handling and interpretation of high-resolution solar imaging, spectroscopic, and polarimetric data provided by DKIST.

Past DKIST Training Workshops have been funded with an ad-hoc, time-limited NSF grant that encompassed other, more extensive activities, including the creation of tools for so-called Level 2 data production (see <https://nso.edu/ncsp/>). Based on this experience, we found that participant costs for an in-person, one-week school or workshop accommodating ~25-30 persons would typically be around \$50k, while for an on-line workshop these costs would be essentially zero. In either case, there are additional costs to cover the time for lecturers to prepare and deliver their presentations, which may be on the order of \$40k for four lecturers who need one to two weeks each to prepare curated datasets and software tasks for the students to follow. An additional \$10k would be needed in either case for recording, organizing, and distributing the videos. Therefore, for a mix of two virtual and two in-person workshops per year, each tailored to a different, topical subject, the annual costs would be around \$300k, or approximately \$3k per participant on average. This is a small fraction of the operating costs of the whole DKIST facility. This makes it an effective investment if it can produce even limited improvements in facility usage or publication rates. In order to quantify the effectiveness of these workshops, in addition to the formal participant surveys, we will also track proposal submission and success rates for workshop participants as well counts of DKIST publications from those attendees.

Another means to reach a large swath of the community at relatively low cost and effort is to organize mini-workshops as an extension of other relevant solar or heliophysics meetings (e.g. SPD, AGU, AAS, etc.), for example one day before or after the full meeting. Such an approach has been successfully employed by other missions or projects.

As we move forward with these training activities, we also envision having some of the workshops address issues of joint data analysis with other facilities, such as ALMA, PSP, Solar Orbiter, ngGONG, MUSE, or other facilities. This would both enable exciting joint science, and help build engagement with scientists initially working primarily with those other resources. In addition, we think this same model could be followed by other new facilities that provide new observing paradigms or unfamiliar data types to the community, such as FASR or COSMO. Collaborations with those groups in tailoring their training and engagement efforts would be useful by sharing lessons learned and best practices.

References

Alexov, A., Tritschler, A., Rimmele, T., Marshall, H., Parraguez, A., 2020, “An Overview of DKIST Operations Commissioning and Tools”, in *Astronomical Data Analysis Software and Systems XXIX*. ASP Conference Series, Vol. 527, 435

Davey, A., 2021, “The Daniel K Inouye Solar Telescope (DKIST) Data Center: Cyber-infrastructure for a next generation ground based solar telescope”, AGU Fall Meeting 2021, held in New Orleans, LA, 13-17 December 2021, id. SH53A-02..

Rimmele, T. R., Warner, M., Keil, S. L., Goode, P. R., Knölker, M., Kuhn, J. R., et al. (2020). The Daniel K. Inouye Solar Telescope - Observatory Overview. *Sol. Phys.* 295, 172. doi: [10.1007/s11207-020-01736-7](https://doi.org/10.1007/s11207-020-01736-7).