HelioWeb: A resource for 21st century science

Authors: S. J. Schonfeld[1], A. K. Higginson[2], B. L. Alterman[3], M. S. F. Kirk[2,4]
Cosigners: J. L. Verniero[5], Y. J. Rivera[6], P. J. Wright[7], E. Lichko[8]

[1]Institute for Scientific Research, Boston College, [2]NASA Goddard Space Flight Center, [3]Southwest Research Institute, [4]Atmospheric and Space Technology Research Associates, [5]University of California, Berkeley, [6]University of Michigan, [7]W. W. Hansen Experimental Physics Laboratory, [8]University of Arizona

To contend with the ever-accelerating pace of discovery within Heliophysics, we must change our modes of communication and information exchange if we hope to meet the scientific challenges of 2050. To that end, we need a new resource for the internet age, ***a single portal through which all heliophysics knowledge is cataloged, interlinked, and discoverable***. Such a resource should facilitate scientific discovery, coordinate research efforts, develop collaborations, enable data and knowledge access, recruit and train new scientists, and educate the public. Through the proposed **HelioWeb**, we can **strengthen the existing heliophysics community while also making it more welcoming and accessible to new members**.

The state of heliophysics communication

Technology developments in the 21st century mean it is no longer sufficient to design an instrument, make a measurement, and publish the resulting data and interpretation. Today, teams of researchers with diverse expertise address interdisciplinary questions by coordinating and synthesizing multi-instrument data sets and complex physical models. Existing online resources are the foundation of modern research, but they do not enable collaboration and discovery to the extent possible in the modern interconnected age.

- **Data archives:** Many data archives exist, providing online access to the majority of modern and historical heliophysics data. However, these archives are typically disconnected, containing data from individual institutions (e.g. NSO, HAO, LMSAL, CSTR) or observation modes (e.g. solar images at the VSO and JSOC and solar irradiance at LISIRD). These archives provide their data in different formats with inconsistent tools for data access, sorting, discovery, and visualisation. This makes it difficult to find data without prior knowledge of a particular archive and data set.
- **Science papers:** The primary method of distributing scientific results is still the peer-reviewed, published paper. While these essential “units” of scientific information are available online (most searchable through the ADS) with references to previous work, they are rarely well linked or organized. This makes keeping up with scientific advancements difficult, and searching for work relevant to a particular topic tedious and time consuming.
- **Conferences and seminars:** Scientific conferences are the primary media through which the “heliophysics community” is actualized. It is during face-to-face meetings at community-wide conferences or institution hosted seminars that professional networks are built, most collaborations are developed, and the majority of scientific advancement is disseminated throughout the community. Yet these meetings raise serious barriers to

attendance[1], requiring significant investments of time and money. By over relying on in-person connections, the heliophysics community effectively limits networking and collaboration opportunities to short periods throughout the year.

These and other communication challenges combine to negatively impact the development of both science and scientists. Knowledge tends to be fragmented and siloed within organizations, subfields, and research groups, often passed down institutionally and through word-of-mouth. This can have significant adverse effects such as: creating inconsistent opportunities within the field, slowing the spread of research insights, and delaying or preventing the creation of collaborations. **To enable heliophysics research of the future, it is essential that the community embraces online communication modes, technologies, and opportunities.**

HelioWeb: the future of scientific communication

Heliophysics communication of the future will be enabled through HelioWeb: a single, open source, dynamic platform designed to connect researchers with each other, science results, and community resources. HelioWeb will create connections between existing resources, enabling users to easily navigate between and synthesize from disparate data, code, and knowledge sources. It will enable a heliophysics researcher to start with a science question outside their area of expertise, read a review-level discussion of current community understanding, discover related papers, find specific relevant data and model runs, and connect with experts interested in addressing that question, all within a matter of minutes without leaving HelioWeb. Importantly, this resource will not only bring science tools and information to researchers, but bring researchers closer to each other with tools that facilitate discussion and collaboration. As the field evolves, HelioWeb will grow through active and collaborative community engagement, with changes archived similar to modern code package development (e.g. PlasmaPy).

A project of this magnitude will require significant effort across the community backed by consistent funding support. At least three tiers of participation will be necessary for the platform's long term stability and growth while ensuring the contents are kept up-to-date with community advancement. HelioWeb must be supported with full-time staff who build, maintain, and expand its infrastructure, information, and services. To augment these permanent staff, rotator positions for active researchers, educators, software developers, hardware designers, etc. will ensure a close connection between the archives and the community. Finally, the creation and maintenance of such a sweeping database will require buy-in from the community at large to participate in HelioWeb by submitting papers, codes, projects, etc. and actively contributing to keep these resources up to date[2]. Where today, the presentation of science results at a conference represents the arrival of a young researcher to the field, in the future, the first contribution to HelioWeb will signal the entrance of a new researcher into the heliophysics community.

---

[1] Particularly for: early career researchers, researchers with young children, researchers at small or non-traditional institutions, researchers with disabilities, and any researcher without travel funding.

[2] These efforts must be explicitly funded, e.g. a fractional effort applied to all funded ROSES grants.

A proposed structure

Community input must guide HelioWeb's long term development. In the interim, Wikipedia provides a concept and template for initial conceptualization and testing. The important design characteristics are easy searchability, robust and smooth navigation between related topics, and growth through community contribution. Each topic (the solar cycle, magnetotail reconnection, etc.) should have its own dedicated page containing a review-level description as well as relevant archives of papers, data, models, and codes. Topics should link within and between pages[3] to form the connected knowledge web that enables discovery and new science.

As a first step, this format could simply include linked, review-level descriptions of science topics, datasets, models, etc. while directing users to existing external paper, code, and data archives. Just this would be an amazing tool to promote discovery and enable access to currently scattered heliophysics resources. However, to fully realize HelioWeb's power, these linked topics must evolve into fully integrated research tools that are supported by interoperable and reusable infrastructure. This will enable granular resource discovery beyond what is possible with current archives e.g. direct access to all data related to a specific coronal mass ejection within a single interface.

Building on existing resources

Standing up such a resource from scratch would be a herculean effort. Fortunately, existing community projects represent important first steps that can guide HelioWeb development.

- CDAWeb provides a robust data archive that aggregates and standardizes measurements from a wide range of sources, provides the data in a common form, and enables simple visualization to assist users to discover relevant data.
- The SunPy Federated Internet Data Obtainer (Fido) python packages enables uniform access to a wide range of solar data from a number of online archives.
- The CCMC hosts a wide range of models along with basic descriptions of their physics, outputs, and uses. It also runs customized models on request.
- The Python in Heliophysics Community (PyHC) embodies the possibilities of a future interwoven with HelioWeb. In addition to the important tools it contains, PyHC's deliberate mindset of collaborative development has organically created a community of users working together to address solar physics questions.
- The Living Reviews in Solar Physics (and other reviews) represent valuable snapshots of the current state of the field.

Each of these resources would be greatly enhanced if they contained elements of the others. **To realize HelioWeb, these and other resources must be integrated so they can be more easily discovered and shared throughout the heliophysics community.** This prioritization of interconnected knowledge (as applied to missions) exists within the NASA Office of the Chief Knowledge Officer whose mission is to ensure "practitioners have access to critical knowledge when they need it – now and in the future." By 2050, this access to critical knowledge, embodied by HelioWeb, should be realized not just in Heliophysics but across NASA disciplines.

[3] perhaps inspired by the hierarchy of the Unified Astronomy Thesaurus