The National Cancer Institute's Community Hub: A resource for collaboration and sharing of data, tools, and standards among the cancer research community

Presented at Gateways 2022, San Diego, USA, October 18–20, 2022. https://zenodo.org/communities/gateways2022/

Esmeralda Casas-Silva*
emi.casas-silva@nih.gov

Claire Stirm†
cstirm@ucsd.edu

Pascal Meunier†
pmeunier@ucsd.edu

Brian Cheung‡
brian.cheung@nih.gov

*National Cancer Institute, National Institute of Health
†San Diego Supercomputer Center, University of California, San Diego
‡NetImpact Strategies

***Abstract*—The National Cancer Institute Community Hub (NCI Hub) is a science gateway facilitating scientific knowledge exchange and networking to accelerate cancer research. The NCI Hub allows members to utilize collaborative online community spaces to meet and exchange knowledge. Additionally, members can publish resources including datasets and computational tools with digital object identifiers (DOIs), teaching materials, workshop materials, and much more. The NCI Hub is managed and funded by NCI's Center for Biomedical Informatics and Information Technology (CBIIT) in collaboration with Hubzero.**

***Index Terms*—cancer research, science gateways, working groups, NCI Hub**

I. Introduction

The National Cancer Institute Community Hub (NCI Hub) is a free, FISMA-compliant, secure public online collaborative science platform advancing NCI's mission to help people live longer, healthier lives by supporting cancer research and advancing scientific knowledge. Through the NCI Hub, users can create and manage groups and projects with members from around the world and leverage collaborative tools such as wikis, discussion forums, blogs, group calendars, and more. In addition, the NCI Hub allows customization of member roles and privacy settings for individual groups and resources.

Since its inception in 2014, NCI Hub has grown to include the following:

- 264 community groups
- 5,400+ user accounts
- 604 published resources
- 48 visualization tools
- 542 projects
- 105 publications

This paper will discuss the NCI Hub platform, group features, and illustrate several use cases.

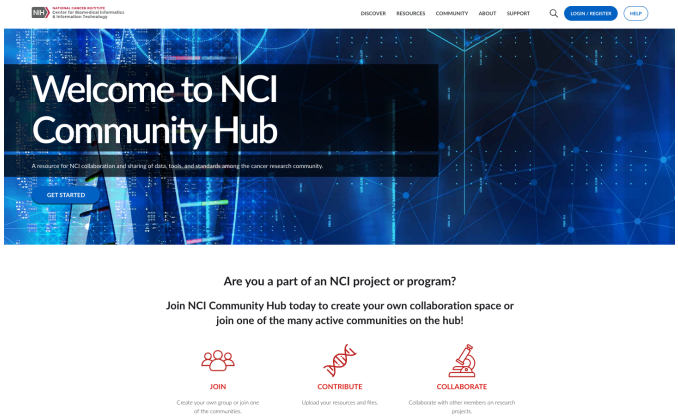

Fig. 1. NCI Community Hub Landing Page. (https://ncihub.org). The NCI Community Hub landing page welcomes members with a short introduction on how to join the NCI Community Hub, contribute, and collaborate on the science gateway.

II. NCI Hub Platform

The NCI Hub is built on the Hubzero® platform [1] and is hosted at the University of California San Diego's San Diego Supercomputer Center (SDSC). The NCI Hub is managed by NCI's Center for Biomedical Informatics and Information Technology (CBIIT) with support from members of the Hubzero and Sherlock Cloud Solutions and Services teams [2].

The NCI Hub allows each group 500 MB of storage space. Automatic alerts are set to notify the Hubzero team when 75% of disk space is used. A group, project, or tool may request additional storage when necessary. The Hubzero team assesses where disk space is being used and appropriately takes actions to clean up or expand disk space.

The NCI Hub is hosted on a Federal Information Security

Act (FISMA) compliant server. The U.S. Federal government requires that agencies or contractors working with the government demonstrate FISMA compliance for all information systems within their network [3]. FISMA sets minimum security requirements for information security protocols. Certification levels range from low to moderate to high, depending on risk assessment and the sensitivity of the information on a site. The NCI Hub is categorized as a FISMA-low system given the limited risk of severe impacts to organizational operations, assets, or individuals should the system be compromised.

FISMA certification can be an extensive process due to the controls and documentation required across various phases. As such, there were several technical, administrative, and management challenges related to this process. Technical challenges involved mapping implemented functionality and configuration to the various National Institute of Standards and Technology (NIST) controls. During this process, a partially missing functionality was discovered; consequently, configuration management plans and remediation plans, a.k.a. Plan of Actions and Milestones (PoAM), were created.

Additional challenges were found in clearly attributing responsibility and accountability for the implementation of the relevant NIST SP 800-53 security and privacy controls between all parties. The SDSC Sherlock Cloud's pre-existing security plan was used to guide control implementation. Given their experience implementing NIST SP 800-53 controls, the Hubzero team created an NCI Hub-specific contingency plan to explicitly identify responsibilities relying on Hubzero and Sherlock incident response plans.

Despite the work involved, achieving FISMA certification is a worthy goal for science gateway operators since compliance and certification increase data security. It assures funding agencies and grantees that due diligence is being made to protect their investments and endeavors in a manner that aligns with well-known industry standards. Additionally, FISMA certification opens broader opportunities for science gateways to work more closely with federal entities and a larger audience.

III. NCI Hub Audiences

The NCI Hub hosts an international audience, with most users in the United States and affiliated with academic institutions. Any visitor to the site can view publicly available resources, tools, and groups without an account. However, visitors must create an account to be able to join community groups, contribute data sets, and take advantage of the Hub's full functionalities.

Group interests span a wide range of topics from imaging to systems biology, data analysis, cancer health disparities, and more. Given the wide breadth of groups and topics, the NCI Hub uses tags to help users find content, events, and members. Tags may be created by any user registered on the Hub. The top five of the 2,600 tags currently created are nanoparticles, Informatics Technology for Cancer Research (ITCR) Initiative, Community Health Educators (CHE), nano working group, and nanotechnology. In addition, if an existing tag exists, the Hub will auto-populate a tag for a user as a suggestion to use the existing tag.

IV. NCI Hub Community Features

A vital feature of the NCI Hub is the number of built-in collaborative tools that allow users to connect easily and work together across institutions and affiliations. Group creation is a function available to all members. Through groups, members have contributed workshop materials, established mentor-mentee relationships to foster professional development, and established project working groups. Members can even export their content to other websites and brand it with their logo. Access to groups and group features are fully customizable, including:

- Roles and permissions – "manager" or "member" roles and permissions can be specified for each group member. These roles determine what permissions will be available to each group participant. The group creator is the default manager and is given access to manage all administrative aspects of a group. Management functionality includes approving new group members, assigning user roles, changing group visibility to restrict access or enable a public view of the group, changing available functionality for group members, and more. If the manager leaves a group, they may assign a new group manager.
- Visibility – Group visibility can be set to private or public and may be defined separately for individual group content and resources such as collections and calendars. This allows a group to have a select public-facing portion of their site to share their work and build interest with the community while keeping other content private.
- Membership – Membership can be defined as public, restricted, invite-only, or closed. The group manager controls group membership.

Several tools are available for NCI Hub groups:

- Communications – The NCI hub offers several ways to spark interaction between community members, including:
  - Wikis – Collections of informational articles written and edited by group members
  - Forums – Conversations and discussions on relevant topics created and curated within a group to allow for easy exchange of information between users.
  - Blogs – Members can post their thoughts, ideas, accomplishments, announcements, and more through a journal format. Content creators can link to other authors as well.
  - Announcements – Group-wide messages can be sent to members, along with automatic email notifications. Announcements include a sticky feature to keep announcements posted for as long as relevant.
  - Calendar – Organization and communication of group events is supported by personalized calendars providing the ability to import events from external calendars and add event registrations.

- Collections – NCI Hub content is grouped by topic, allowing users to follow and "collect" material made available publicly. Users can create their collections or follow those made by other members.
- Document and media storage – Members can upload, share, and store resources such as presentations, teaching materials, videos, data sets, and more. The NCI Hub provides a centralized location to load project data and gives members a shared drive for file storage. Users can also link data from outside sources such as Github, Google Drive, and AWS.
- Publications – Publications can be minted through DOIs (Digital Object Identifiers), enabling easy tracking, accurate citations, and increased discovery.

V. Sample use cases

Although each group uses the NCI Hub uniquely, we showcase a few examples below to illustrate how some groups have leveraged the NCI Hub's features to further their mission and collective goals.

A. Quantitative Imaging Network (QIN)

The Quantitative Imaging Network (QIN) is comprised of 22+ major academic sites promoting research and development of quantitative imaging methods to help predict response to therapies in clinical trial settings [4]. QIN has 73 NCI Hub members across the U.S. and Canada. QIN groups utilize the NCI Hub to support networking, cataloging and pushing source code onto the platform. As part of their process of developing cutting-edge imaging software, the QIN Group has created two project types:

- Challenges – Quantitative imaging algorithm competitions
- Collaborations – Collective imaging data collection and analysis across multiple sites

The QIN Group uses an open-source platform called QinLabs to compute and evaluate metrics for Challenges. QIN groups have even developed code within the NCI Hub. Users can make use of the repository to evaluate and optimize published tools. These successes led the QIN Network to systematically move their groups to the Hub to take advantage of this collaborative tool-building model.

B. NCI Center to Reduce Cancer Health Disparities (CRCHD) Supplements Café

NCI's CRCHD aims to help reduce cancer health disparities by fostering and supporting cancer research networks and building a more diverse cancer research workforce. To support their goals, CRCHD developed the NCI Hub super user group, the "CRCHD Supplements Cafe" [5]; the Cafe supports research trainees by providing a platform to meet, network, and find program and peer mentoring for research and career development.

The CRCHD Supplements Café serves 222 group members to date with over 58 pages of support content, including a Career Station with a job board and blog; career development and mentorship resources; a professional development workshop page; Professional Advancement Virtual Engagement Series (PAVES) recordings, and more.

Similar to the QIN group, the CRCHD group requires community members to join before accessing resources and participating in discussions. However, some areas, such as blogs and forums, are viewable to the public.

C. Informatics Technology for Cancer Research (ITCR) Initiative

ITCR is an NCI program supporting informatics technology innovation to improve the management and dissemination of cancer data and knowledge to accelerate cancer research [6]. ITCR currently has a membership of 181 on the NCI Hub with several subgroups. Group members are ITCR investigators who use the group to network and share NCI-funded technology work products with other investigators. The ITCR also makes available documents from investigator meetings, working groups, conferences, and more. In addition, presentations and materials from past annual meetings are available to the public on ITCR's NCI Hub site at https://ncihub.org/groups/itcr.

VI. Conclusion

The NCI Community Hub has grown to include over 5,000 members from around the world who employ the NCI Hub's vast tools and resources to network and exchange knowledge to support the advancement of cancer research and scientific discovery.

The NCI Hub continues to grow and search for new ways to meet community needs. We welcome any suggestions or comments at https://ncihub.org/about/contact.

Acknowledgments

We want to acknowledge the Hubzero team and the Sherlock Cloud Solutions and Services team for technical support of the NCI Hub.

References

[1] M. McLennan and R. Kennell, "Hubzero: a platform for dissemination and collaboration in computational science and engineering," *Computing in Science & Engineering*, vol. 12, no. 2, pp. 48–53, 2010.

[2] S. C. S. . Services. (2022) Sherlock cloud solutions services. [Online]. Available: sherlock.sdsc.edu

[3] C. I. S. Agency. (2022) Federal information security modernization act. [Online]. Available: https://www.cisa.gov/federal-information-security-modernization-act

[4] (2013) Qin portal. [Online]. Available: https://ncihub.org/groups/qinportal

[5] A. Lin. (2022) Crchd supplements cafe. [Online]. Available: https://ncihub.org/groups/crchdsupplements/overview

[6] Itcr. [Online]. Available: https://itcr.cancer.gov/