Cancer Informatics Collaboration and Computation: Two Initiatives of the U.S. National Cancer Institute

Authors: Ishwar Chandramouliswaran[1], Juli D. Klemm[1], Tanja Davidsen[1], Anthony R. Kerlavage[1], Warren Kibbe[1]
Author affiliations: [1]National Cancer Institute, Center for Biomedical Informatics and Information Technology
Presenting author email address: Ishwar.chadramouliswaran@nih.gov
URL for project website: www.nciphub.org,
http://cbiit.nci.nih.gov/ncip/nci-cancer-genomics-cloud-pilots/nci-cloud-initiative
Code URL:
Open Source License:

One of the objectives of the National Cancer Institute's Center for Biomedical Informatics and Information Technology is to focus on innovative programs and technologies to support the use of informatics in cancer research and encourage open science. We have several initiatives underway that are designed to enhance community driven software development, democratize access to cancer data, tools & standards as well encourage innovative approaches that improve access to software applications & build scientific communities that enable deep collaboration around specific research questions.

At BOSC 2013, we presented NCI's open development initiative to share open-source code for cancer informatics software applications through GitHub at https://github.com/ncip. This presentation describes two follow-on initiatives and how these initiatives can help researchers engaged in cancer informatics collaborate and contribute to NCI's open science efforts.

NCIP Hub: One of our goals is to enable cancer researchers to create community driven, adaptive, and collaborative environments that promote the exchange of research ideas and resources. To address this we have established the NCIP Hub, using the open source HUBzero Platform for Scientific Collaboration, to provide this online collaboratory for the cancer informatics community. The intent is to empower community members to both contribute and use software tools, data, standards, or other relevant digital assets to an ever-growing research and educational resource. The work products become discoverable and citable, and their impact can be measured. We hope that this will contribute to the creation of a 'community impact score' based on data sharing, algorithm sharing, software sharing, discoverability, annotation, and of course use and reuse. Individuals engaged in cancer informatics can become a member and contribute at www.nciphub.org.

The Cancer Genomics Cloud Pilots: The purpose of the Cancer Genomics Cloud Pilots is to support the development of a new model for computational analysis of biological data that can address the required computational capacity for storage, analysis, and discovery. In this new model the data repository is co-located with computational capacity that can be accessed either via a web interface or via an Application Programming Interface (API) while ensuring data security. Developers of analytical software applications will also be able to bring new tools to the data, and researchers will be able to bring their own data to the cloud to analyze in the context of TCGA data. The Cancer Genomics Cloud has the potential to democratize access to NCI-generated genomic data and provide a less resource intensive and more cost-effective way to perform computational analysis for the cancer research community while enabling reproducibility of the analyses. The three pilot systems being developed by the Broad Institute, the Institute for Systems Biology and Seven Bridges Genomics are expected to be available for community use and evaluation in early 2016. Stay informed by following http://cbiit.nci.nih.gov/ncip/nci-cancer-genomics-cloud-pilots/nci-cloud-initiative.