Cambridge Healthtech Media Group

www.bio-itworld.com

Bio·IT World

Indispensable Technologies [illegible], and Clinical Trials

JULY | AUGUST 2008 • VOL. 7, NO. 6

Kenneth Buetow, Ph.D.; National Cancer Institute

Best Practices 2008

SPECIAL REPORT: Complete coverage of the 2008 Best Practices Awards. 18

Connecting the Cancer Community caBIG™ Time

BY LAURIE WIEGLER

Through a federated framework, the **National Cancer Institute's** (NCI) Cancer Biomedical Informatics Grid, better known as caBIG™, provides an interoperable method of connecting disparate health care providers and offering a common language for furthering development of cancer research tools and better patient outcomes.

EDITOR'S CHOICE

Kenneth Buetow, Ph.D., associate director for bioinformatics and information technology at NCI, believes the Best Practices Editors' Choice award will draw more attention to caBIG™, a three-year effort to connect people, organizations, and data through IT: a worldwide cancer web.

caBIG™ has been designed to further medicine's potential through an open source network, with Buetow and team optimistic that the open model grid will overcome silos that have impeded cancer research. "Open" is actually "operationally defined," he says. NCI's technical infrastructure for caBIG™ has limitations. caBIG™ is open to anyone who wants to use it, says Buetow, but just as anyone can use the Internet, "there can be things that anybody can access; but, designed into the very DNA of caBIG™ is the capacity to have restricted access."

While caBIG™ sprung from a cancer network, its open source model means that it's helping other disease models. For example, the National Heart, Lung, and Blood Institute's Cardio-Vascular Research Grid has also adopted the platform.

Of this, Buetow is particularly enthusiastic. "caBIG™ has the capacity to be as open as any user can or would share information." The technical applications and infrastructure "are open to anyone who wants to use it," says Buetow. "We do not discriminate," Buetow stresses. "Whether you're an academic; whether you're in government; whether you're in industry." But without the initial focus on cancer, none of this would have been possible. "Our initial efforts were pretty heads down in the area of trying to work with our NCI-designated cancer centers," he says. "We needed to have a place where this could work."

caBIG™ provides three collections of tools to support molecular medical research, including semantic data standards and infrastructure to support sharing of heterogeneous data. The caBIG™ collections are designed to facilitate clinical trials, tissue banking, imaging and integrative cancer research, a clinical trials compatibility framework, and a data sharing and security framework.

The first collection of tools is designed to manage biospecimens, microarray, and sequence data; store clinical and *in vivo* imaging data; and provide the caBIG™ backbone. The second collection offers tools to promote adverse event management and study participant registration and management, for example. The last "framework" touts policies tailored toward facilitating evaluation of data sensitivity.

Kenneth Buetow, Ph.D.; National Cancer Institute

DAVID LEIFER (2)

The community currently includes over 50 cancer centers and numerous government organizations and nearly 900 individuals. Their work is organized by so-called "work spaces," virtual teams that develop software in specific areas of interest and then share them with the group. As with many areas of genetic research these days, caBIG™ suffers from its own success. "As you might speculate, there is a just a tremendous amount of data [that results from this]," says Buetow.

"And because what we are doing on this project is being comprehensive, it's not only generating large volumes of sequencing [data], [but] it's also generating volumes of copy number ...data on gene expression, epigenetic information and [information that is] tied to very well-characterized biospecimens as well as clinical data," Buetow further explains.

Right now, Buetow says the "coolest thing" is a new portal that debuted in April. Launched for The Cancer Genome Atlas (TCGA), this portal shows how doctors are assimilating tissue data from cancer patients to determine the disease's potential genetic causes. The portal integrates information from TCGA pilot projects to explore the molecular underpinnings of cancer. One of caBIG™'s principal aims is to pinpoint molecular pathways that make some people more susceptible to a cancer "gene" than others.

A priority is that the initiative remains vendor neutral. "We are trying to be agnostic as to individual platforms, so we actually have interfaces to most of the large-scale data producers." Buetow does acknowledge, however, holding "discussions with all of the major players" which have included conversations "myriad" bioinformatics and software development companies. •

To learn mote about caBIG™, visit http://cabig.cancer.gov