



## BIOMARKER COMMONS

### **SCan-Mark Explorer V1 Launched to Accelerate Cancer Biomarker Discovery and Support Personalized Medicine**

After 2 years of software product development funded by the National Cancer Institute (NCI) SBIR Program, Sophic recently announced the launch of the Sophic Cancer Biomarker Knowledge Environment, SCan-Mark Explorer V1. The SCan-Mark Explorer software is intended to accelerate research, drive drug discovery, support translational medicine, improve the accuracy of disease diagnosis, increase the effectiveness of treatments and support finding cures for cancer. SCan-Mark Explorer combines a growing database of manually curated breast, ovarian, colon, non-Hodgkin's lymphoma and melanoma biomarker targets with the power of BioXM, an integrated knowledge management system. Throughout the development project, Sophic collaborated with and had the benefit of a strong and active Scientific Advisory Board that included scientists from cancer research centers and hospitals, NCI labs and executives from the pharmaceutical industry who all provided invaluable feedback, advice and direction on the development of SCan-Mark Explorer.

During the two-year Phase II SBIR project, a team of Sophic scientists curated select full-text papers that were rich in clinical data and experimental evidence from human biospecimens. Sophic scientists identified and extracted details for up to 33 often hard-to-find Critical Data Elements (CDEs) from each paper. A CDE represents detailed evidence intended to improve a researcher's confidence that potential biomarkers identified in the literature will deliver value to the broad cancer community of basic researchers, biotech scientists, drug developers and clinicians. CDEs are integrated into highly computational Sophic Cancer Biomarker Objects (SCBOs) that allow scientists to quickly and easily find correlations between valid semantic and scientific relationships between CDEs, SCBOs, enrichment databases and clinical information. BioXM uses precise scientific data models and high-quality ontologies, thesauri and vocabularies that allow scientists to find and map complex biomarker knowledge networks. An intuitive, easy to use WIKI interface allows researchers to visualize networks of valid semantic and scientific relationships connecting SCBOs, pathways, near-neighbors, analyzed sequence data, mutations and drug compound information.

K. Stephen Suh, PhD., Director of the Tissue Bank and Genomics Program at the Juris Research Center, Hackensack University Medical Center, said: In addition to participating on the Sophic SBIR Scientific Advisory Board, my lab has collaborated with Sophic for over 5 years on a range of projects and we have published biomarker papers together in various cancer journals. We participated in the Beta Test for SCan-Mark and found the WIKI interface very easy to use which allowed bench scientists to access valuable, hard-to-find biomarker information. This saved my lab not only time and money, but also allowed us to accelerate our biomarker discovery projects. This ultimately lead to promising lymphoma, ovarian cancer, breast cancer biomarker targets. Many have been validated against cancer patient tissues and are still under investigation.

Michael Liang, PhD., the Sophic Principal investigator responsible for the SCan-Mark project, said: We considered the full range of biomarker papers with a range of data and various confidence levels but decided to curate only papers rich with clinically relevant information supported by strong evidence derived from human experiments. The Sophic team mined and selected papers for manual curation based on the quantity and quality of evidence in each paper and the value of the information to researchers. Sophic scientists studied each paper to identify and extract up to 33 CDEs that are the building blocks for SCBOs. To date, we have curated over 1,400 papers in order to extract these highly structured "building

blocks". The CDEs also function as information hooks that relate to other SCBOs and, more importantly, as bridges to valuable information stored in our bank of enrichment databases.

Patrick Blake, Sophic's President, said: Phase II of the SBIR Program gave us 2 years to develop SCan-MarK so we focused on quality information vs. quantity in our mining and curation efforts. Our strategy was to concentrate the time-consuming manual curation efforts to focus on five of the most deadly diseases: breast, ovarian, colon, Non-Hodgkin's lymphoma and melanoma. This allowed Sophic scientists to dig very deep into the full text papers to extract valuable details while in parallel, integrating information from over 30 enrichment databases that help scientists "connect the dots" between biomarker targets with quality information in widely used data sources. Scientists can now save time and avoid dead-end research projects by accessing rich biomarker information related to pathways, sequence, mutations, etc., in a single system. Our strategy from the beginning was to collaborate with the cancer community, so we are grateful for the support we received over the past 2 years from each member of our Scientific Advisory Board. It would have been impossible to develop a comprehensive, multi-purpose system like SCan-MarK Explorer without their help.

SCan-MarK Explorer is a software-as-a-service offering that will provide scientists with quarterly updates and enhancements to the BioXM software configuration. Sophic provides a 2-week free trial for scientists to evaluate a sample of the curated information and the power of BioXM on the company's website. Government and academic researchers can license SCan-MarK Explorer as a hosted, annual term service. SCan-MarK will be licensed for installation at commercial biotechs, pharmaceuticals and government/academic institutions. Sophic also provided a wide range of contract research services.