CANCER NETWORK MINER
DEMONSTRATING THE POWER OF DATA SCIENCE

Booz Allen Hamilton and Omnity have partnered to demonstrate the power of data science in addressing the USPTO's Cancer Moonshot Challenge. Together we developed the NIH / USPTO Cancer Network Miner to allow the National Institutes of Health (NIH), other funding institutions, and cancer researchers to see how research efforts are linked into networks of similar and complementary research. Our tool shows how individual research institutions can serve as a lynchpin between related groups of researchers, and how knowledge and discoveries are disseminated through research communities. This will help NIH managers better understand the return on investment from research grants by revealing otherwise hidden conceptual connections shared between patents, research grants, and clinical trials.

THE CANCER RESEARCH LANDSCAPE: A SELF-ASSEMBLING COMMUNITY ANALYSIS

After examining data provided by the USPTO, we noticed that many patents were supported by one or more NIH grants, and that many grants supported numerous patents. Using these relationships, we built a network graph where each node is a patent, and each edge is a grant linking two patents together. In the graph, each patent appears only once, while a single grant may appear many times as it supports numerous patents.

The graph shows how patents and grants self-assemble into a series of clusters, or research communities, which can represent a range of organizational scale, from a sprawling research network to a single patent and grant pairing. In all, 1,248 distinct clusters were identified.

The strength of the relationship between linked patents is described by the thickness of the lines connecting them. The line thickness represents the degree of semantic similarity between patent documents, based on data provided by Omnity. Omnity's semantic similarity is developed by examining related concepts between documents based on shared, unique words. Figure 1 is a single cancer network cluster showing the relationships between 6 research institutions and 4 NIH Institutes. The thickness of the lines show the semantic similarities between patents.

This framework allows stakeholders to rapidly detect where the greatest return on investment exists within an otherwise complex landscape of patent families. Rolling a cursor over a patent brings out a pop-up box with information about
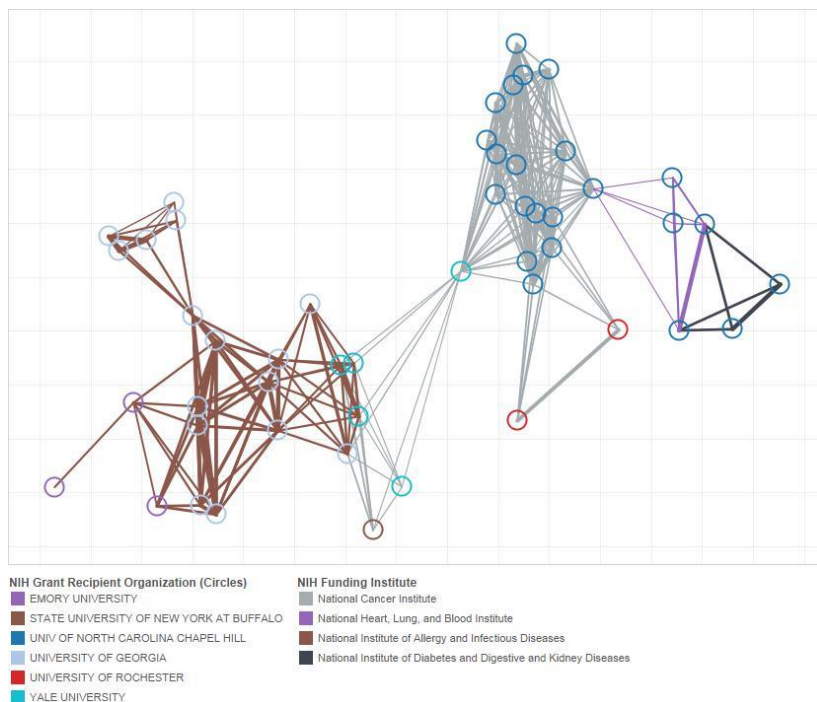

NIH Grant Recipient Organization (Circles)
- EMORY UNIVERSITY
- STATE UNIVERSITY OF NEW YORK AT BUFFALO
- UNIV OF NORTH CAROLINA CHAPEL HILL
- UNIVERSITY OF GEORGIA
- UNIVERSITY OF ROCHESTER
- YALE UNIVERSITY

NIH Funding Institute
- National Cancer Institute
- National Heart, Lung, and Blood Institute
- National Institute of Allergy and Infectious Diseases
- National Institute of Diabetes and Digestive and Kidney Diseases

Figure 1: A single cancer network structure

Booz | Allen | Hamilton

Cancer Network Miner Overview

the cluster (identification number, total funding, and dollars per patent produced), and information about the patent (grant recipient organization, patent title, abstract and patent ID).

Our tool also allows the user to drill down into information on each patent within the network graph. Clicking on a patent node opens a second browser window to reveal a semantic analysis of patents, research grants, and clinical trials most closely related to the patent. This Omnity-developed linguistic capability allows users to follow connections through shared concepts. Users can identify how patents, grants and clinical trials have evolved over time to further understand the research landscape.

A FRAMEWORK FOR ACTIONABLE INSIGHTS

Our tool provides stakeholders with valuable, actionable information. Understanding which patents are linked to the greatest amount of research funding and the largest clinical trials, USPTO management can then expedite examination of these patent classes in cancer research to reduce their time to market and accelerate their impact. Additionally, stakeholders within the research community can better understand what cancer technologies have been most impactful. Accelerating patent prosecution will provide industry with more certainty about their patent applications, reducing speculation about the business value of their investments while increasing confidence in the protection of their intellectual property.

The USPTO can further use our solution to demonstrate to key stakeholders at the NIH the value that cancer research provides to the scientific and medical community. Much of the innovation related to cancer is interdisciplinary in nature – our solution illuminates the research domains that are driving innovation to help identify highly productive domains. Our dashboard can connect researchers across research institutes with one another. These connections can drive broader engagement and greater information sharing with grantees, whose individual innovations may be synergistically associated with a cluster or a funding network.

> To view the Cancer Network Miner, please send an email to
> CancerNetworkMiner@bah.com
> to receive a link to the site

We believe that the connections uncovered in our solution provide the USPTO and its partners with a framework to rapidly and efficiently address several key aspects of public policy formulation that will help accelerate our collective objective of eradicating cancer. Through better allocation of its own resources, facilitating connections within the research community, and demonstrating the positive externalities of cancer research, the USPTO can drive the Cancer Moonshot initiative to a successful conclusion.

FOR MORE INFORMATION

Ezmeralda Khalil
Principal
khalil_ezmeralda@bah.com

Manish Kumar
Senior Associate
kumar_manish@bah.com

Jesus Jackson
Chief Technologist
jackson_jesus@bah.com

Andrew Weller
Senior Lead Scientist
weller_andrew@bah.com

Booz Allen Hamilton Proprietary and Business Confidential