FUNG INSTITUTE
for ENGINEERING LEADERSHIP
UC BERKELEY COLLEGE OF ENGINEERING

CleanEnergyPatentMapper: Visualization of the sources of clean tech inventions

Guan-Cheng Li
Department of Electrical Engineering and Computer Sciences
UC Berkeley
guanchengli@eecs.berkeley.edu

Lee Fleming
Fung Institute for Engineering Leadership
UC Berkeley
lfleming@ieor.berkeley.edu

College of Engineering
University of California, Berkeley

Fung Technical Report No. 2013.06.17
http://www.funginstitute.berkeley.edu/sites/default/files/CleanEnergyMapper.pdf

June 17, 2013

The Coleman Fung Institute for Engineering Leadership, launched in January 2010, prepares engineers and scientists – from students to seasoned professionals – with the multidisciplinary skills to lead enterprises of all scales, in industry, government and the nonprofit sector.

Headquartered in UC Berkeley's College of Engineering and built on the foundation laid by the College's Center for Entrepreneurship & Technology, the Fung Institute combines leadership coursework in technology innovation and management with intensive study in an area of industry specialization. This integrated knowledge cultivates leaders who can make insightful decisions with the confidence that comes from a synthesized understanding of technological, marketplace and operational implications.

We wish to thank the United States Patent and Trademark Office and National Science Foundation for supporting this work, with grant #1064182, as well as the Coleman Fung Institute for Engineering Leadership. Code is available for all non-profit researchers, please contact the Fung Institute. All errors remain ours.

Abstract: Fossil fuel combustion continues to release increasing amounts of CO2 into the earth's atmosphere; this has created a greenhouse effect that has warmed the earth. In counteract this trend, a broad number of institutions have stepped up their investment in clean energy research and invention. To study and visualize these efforts geographically, we provide CleanEnergyPatentMapper, a tool that visualizes clean energy patents by technology type, inventing organization, and geography. This tool maps all U.S. clean technology patents by first inventor location across the world, from 1975-2012, and allows immediate retrieval of the mapped patent. Patents are categorized by technology: solar, wind, biofuel, geothermal, hydroelectricity, and nuclear, and by inventing source and original owner: large incumbent firms, small and medium incumbent firms, venture capital backed startups, government, academic, and unaffiliated (lone) inventor. CleanEnergyPatentMapper is at: http://funglab.berkeley.edu/cleantechx/.

Introduction

In response to global warming, many concerned actors have initiated or increased their efforts to discover better clean energy technologies. This can be seen in a marked increase in clean energy patenting, in the US and abroad, particularly in the mid 2000s (Ramana et al. forthcoming). To facilitate investigation of this phenomenon, we have built a tool that maps clean energy patents, by technology, original inventing organization, and geography. The tool can retrieve patent data from the map, including inventor name, application date, and organization.

We first identified clean technology patents in the US patent database, using a combination of manual and statistical methods (Younge and Paisner, 2013). Using the address of the first inventor on each patent, we then located the patent on a Google Map. We describe and provide examples of a variety of features. An appendix describes the code repository and documentation.

Identifying clean tech patents from the US patent record

We began to compile our dataset of all clean energy energy patents by searching for patents (from 1975-2012) with one of the following keywords: energy, power, combustion, turbine, petroleum, coal, solar, biofuel, fuel. We identified a total of 232,387 US patents that contained one or more of these keywords. We then developed six categories of patents (which we refer to as CTYPE *categories*): solar, wind, biofuel, geothermal, hydroelectricity, and nuclear. Based on smaller and hand-coded data sets provided by IP Checkups[1] for a set of patents in each of these categories, we trained the LIBLINEAR (Fan et al. 2008) tool and categorized the remaining patents. This increased the original sample size by a third and had a cross validation of 85.01% to the original training set patents. IP Checkups also manually review the newly classified patents for a total of 26,572 patents.

We also developed *assignee* or ATYPE categories: lone inventor, venture backed firms, small & medium incumbent firms, academic, governmental, large incumbent firms. Assignees are the owners of the patents upon issuance (patents can subsequently be bought and sold).

[1] http://www.ipcheckups.com

Geolocation of the patents

Each patent provides an assortment of information, including the inventor's address, which we can exploit in mapping the patent. We geo-encode the (city, state, country)-tuple of the first inventor of each clean tech patent and obtain the latitude and longitude coordinates of that patent. We visualize the geographic distribution of the patents on Google Maps.

The interactive map interface is coded using D3.js (a JavaScript library also named as D3 for Data-Driven Documents) and jQuery. The array of our clean tech patents is formatted into a dynamic graphical set of variables, which are then fed into D3 in form of JSON (JavaScript Object Notation) objects. The visualization is designed in order for the interface to be compatible and displayable across major browsers, including Chrome, Firefox, Opera, Safari, and IE 9/10. The filtering engine behind the visualization translates HTML (Hypertext Markup Language) objects into JavaScript variables and triggers the appropriate functions for the immediate calculation and the corresponding update of visual variables such as bar charts, circles, and icons of the map.

Description of features and use examples

We also designed interactive that enables focus, i.e., (1) year range of patents being granted, (2) ATYPE, (3) CTYPE, and (4) distance from the hub of a major city (San Francisco, Boston, Houston, Chicago). The web-based interface is aimed to encourage exploratory patent research. We step through a variety of examples below (the tool is at: http://funglab.berkeley.edu/cleantechx/).

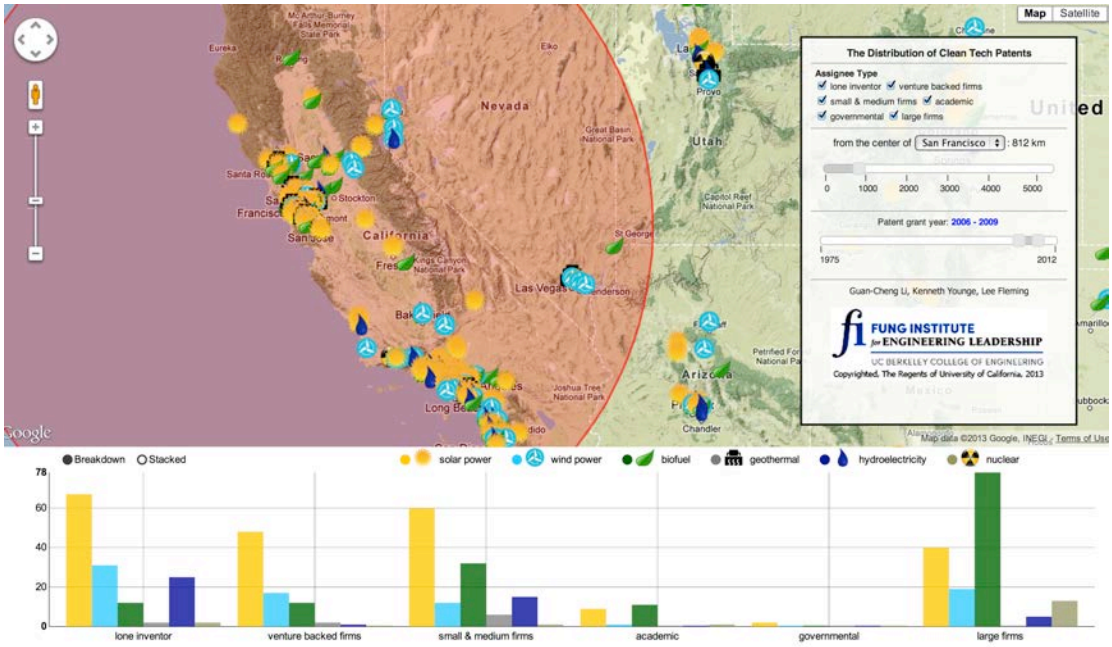

Figure 1: Patents within a given radius of a city can be identified. Note the radius can be changed, and the histogram will automatically update.

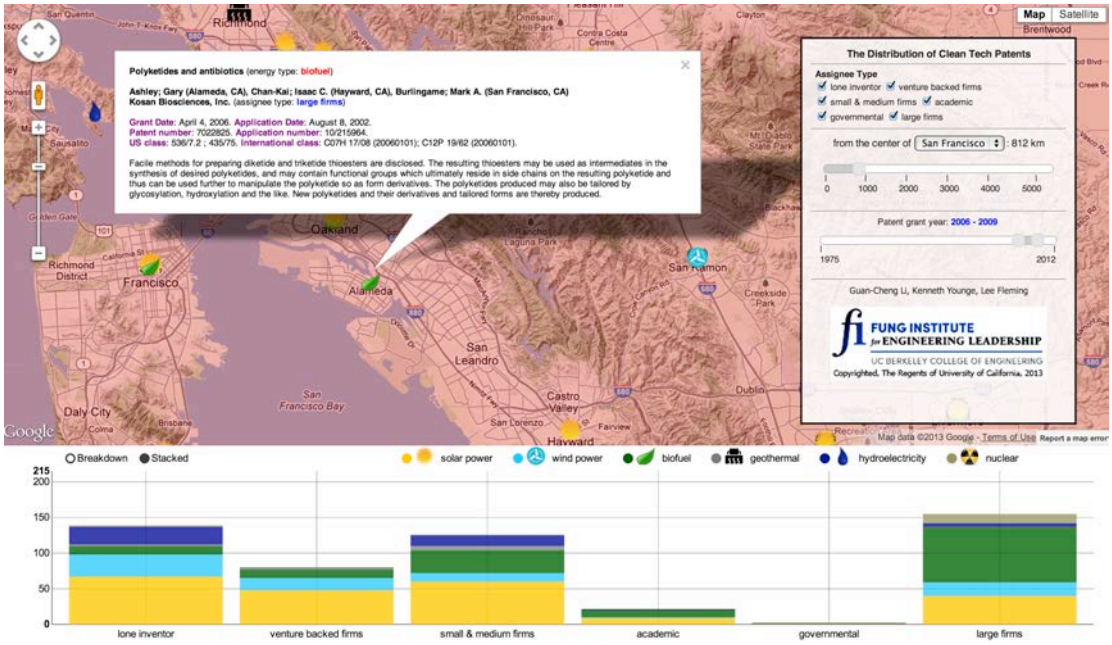

Figure 2: The tool allows zooming in for more geographical detail and clicking on the icon of a patent provides details. The histogram can be presented in two formats, "Breakdown" and "Stacked", which allows for easier comparison amongst ATYPEs.

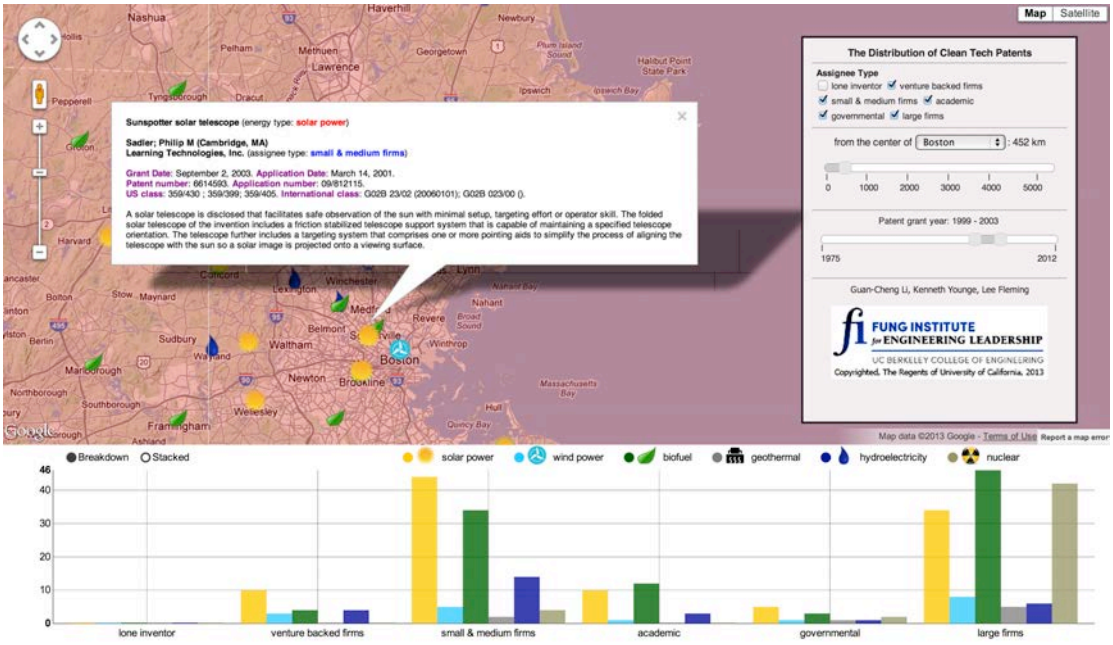

Figure 3: The years of analysis can be changed and different urban hubs investigated.

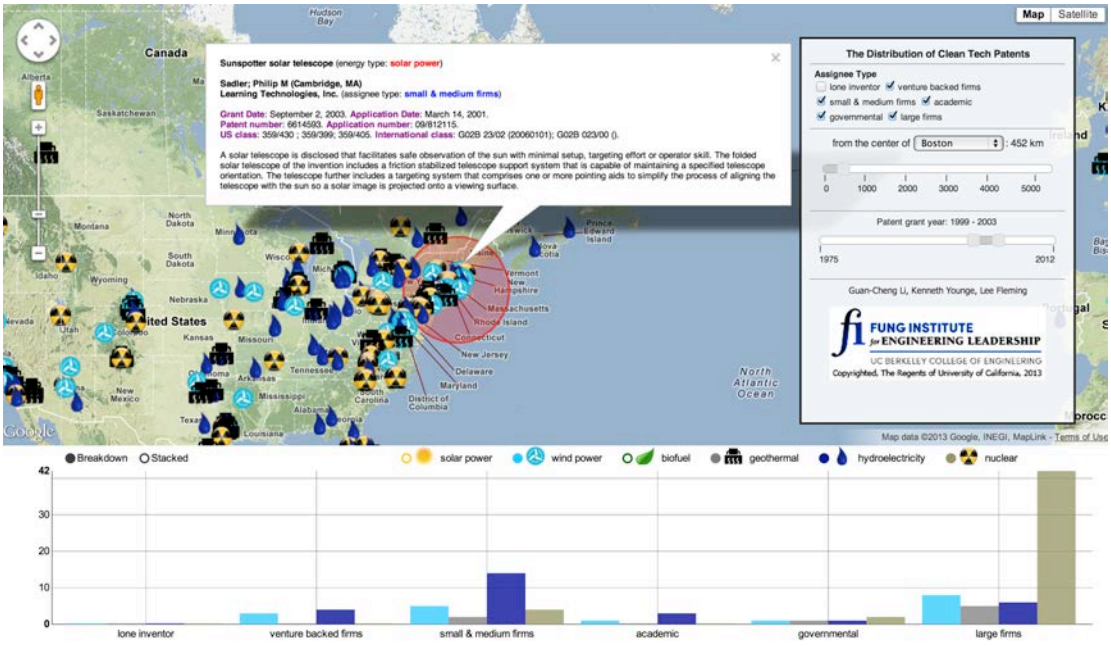

Figure 4: Different types of institutional sources of inventions can be highlighted: here the lone inventor is removed. Technology types can also be specified: here we uncheck two dominating CTYPEs: solar power and biofuel.

References

Nanda, R. and K. Younge, L. Fleming, (forthcoming 2014). "Innovation and Entrepreneurship in Clean Energy," *Rethinking Science and Innovation Policy*, NBER.

Younge and Paisner, (forthcoming 2013). Fung Institute Technical Note, "Clean Technology U.S. Patent Dataset, 1975-2012."

Fan, Rong-En and Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, Chih-Jen Lin: LIBLINEAR: A Library for Large Linear Classification. *Journal of Machine Learning Research 9*: 1871-1874 (2008).

Bostock, M., Ogievetsky, V., and Heer, J. D3: Datadriven documents, *IEEE Transaction on Visualization and Computer Graphics (Proc. InfoVis)*, 2011