

Webscraping and the Wayback Machine

Jan Youtie, GeorgiaTech

gefördert durch die Gaul-Stiftung

Datum: 23.09.2014

Zeit: 9.00-12.00h

Ort: Campus Süd, Gebäude 20.13, Raum 109

Abstract: The Development of Innovative Green Goods SMEs: Exploring the Mix of Micro-Level Relationships Underlying Growth: While broad frameworks of industry, government and university ("Triple Helix") collaborations promise to be conducive to innovation and economic development at macro-levels, at the micro-level of the firm it should not be assumed that such relationships are pre-programmed or automatic. Each firm will negotiate and develop its own set of relationships with other innovation system actors, based on its capabilities and strategies. To understand the dynamics of relationships particularly from the perspective of small and medium-sized enterprises SMEs, this study probes the micro-level dynamics and impacts of industry, governmental, and university relationships. The empirical focus is on the green goods sector, which has received much attention for its potential for economic re-generation, in addition to its energy and environmental benefits. At the same time, the promise of "green jobs" for example, has not always necessarily lived up to near-term hopes for large scale economic renewal through the green economy. to understand what is behind the dynamics of the green sector, this study looks at a subset of US small and medium sized enterprises (SMEs) involved in green goods manufacturing. The paper uses an explicit definition of green goods that is keyword based, building on government industry classifications and patent-oriented definitions but operating at a more disaggregated level.

The work is guided by a set of hypotheses that above average growth in green goods companies is positively associated with the extent of "triple helix" linkages to other industries, universities, and government agencies and research organizations. A hallmark of this study is the use of websites as unobtrusive, unstructured data sources to complement commonly used information from business databases, patents, and publications. Information about company websites in the 2008-2011 timeframe is obtained through scraping of keywords and places to represent geographic and sectoral characteristics; these characteristics are regressed on sales growth, with controls for region, scale, and application areas included. The results suggest that micro-level Triple Helix linkages are positively associated with growth of green goods companies, a finding which emphasizes the importance of coordination in this domain.

Homepage von Jan Youtie: <http://www.planning.gatech.edu/people/jan-youtie>

Veranstalter: Lehrstuhl für Wirtschaftspolitik; Förderung durch die Gaul-Stiftung