Accepted version
Licence CC BY-NC-ND Please cite as:
Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

A network analysis using metadata to investigate innovation in clean-tech - Implications for energy policy

Alessandro Marra[1], Paola Antonelli[2], Luca Dell'Anna[3] and Cesare Pozzi[4]

Accepted version

Licence CC BY-NC-ND

Please cite as:

Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

Abstract

Clean-technology (clean-tech) is a large and increasing sector. Research and development (R&D) is the lifeline of the industry and innovation is fostered by a plethora of high-tech start-ups and small and medium-sized enterprises (SMEs). Any empirical-based attempt to detect the pattern of technological innovation in the industry is challenging. This paper proposes an investigation of innovation in clean-tech using metadata provided by

---

[1] University D'Annunzio. E-mail: amarra@unich.it (Corresponding author)
[2] University D'Annunzio. E-mail: paola.antonelli@unich.it
[3] Ltka. E-mail: dellanna@ltka.org
[4] University of Foggia. E-mail: cesare.pozzi@unifg.it

Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

CrunchBase. Metadata reveal information on markets, products, services and technologies driving innovation in the clean-tech industry worldwide and for San Francisco, the leader in clean-tech innovation with more than two hundred specialised companies. A network analysis using metadata is the employed methodology and the main metrics of the resulting networks are discussed from an economic point of view. The purpose of the paper is to understand specifically specializations and technological complementarities underlying innovative companies, detect emerging industrial clusters at the global and local/metropolitan level and, finally, suggest a way to realize whether observed start-ups, SMEs and clusters follow a technological path of complementary innovation and market opportunity or, instead, present a risk of lock-in. The discussion of the results of the network analysis shows interesting implications for energy policy, particularly useful from an operational point of view.

Keywords:

Clean-technology; Industry studies; Innovation; Network analysis; Energy policy; CrunchBase

1. Introduction

1.1. The clean-tech industry: definition and economic trends

The current unsustainable models of production and energy use have resulted in a vibrant market in clean-technology (clean-tech) centred on technological innovation and improved environmental performance. Clean-tech is defined as any product, service or process that delivers value using fewer resources and producing less pollution than current standards stipulate (Cooke, 2008). Therefore, any innovation that results in improved environmental performance falls under the clean-tech umbrella: recycling, renewable energy (wind-power, solar-power, biomass, hydropower, biofuels), information technology, green transportation, green buildings, electric motors, green chemistry, lighting, grey-water, and many other energy efficient appliances (Pernick and Wilder, 2007).

The transition to a more sustainable way to produce and use energy needs huge and risky investment in research and development (R&D), as clean-tech are capital-intensive and very unpredictable ventures in the early stages. Most innovation in the industry takes place in young high-tech companies, which are often characterized by large intangible assets, negative cash flow, technological uncertainty, and low liquidation value: this is the case of many promising and ambitious start-ups hitting the market with products and services that often impact on people energy consumption (Rasmussen et al., 2012; Pernick and Wilder, 2007; Walsh, 2012).

Clean-tech is already a large and increasing market. According to the UK Department for Business, Innovation & Skills (UK BIS, 2013), the market for low carbon environmental goods and services has been valued at more than US$5 trillion worldwide in 2012 and is forecast to grow at around 4,1% annually until 2016. This significantly outstrips global average economic growth projections of 2,2% to 3,3% over the same period (World Bank, 2013). Moreover, clean-tech has an impressive potential for job creation. The International Labour Organization estimates that a transition to a greener economy could yield a net gain of 60 million jobs (ILO, 2013). In the US, employment in the clean-tech industry represents 2,6% of the total workforce, that is 2.5 million private sector and 886,000 public sector jobs (US Bureau of Labor Statistics, US BLS, 2013). This is greater than in educational services, which employ 3.2 million, and represents about one-third of America's employment in manufacturing (11.5 million), and 40% of the financial services sector (7.8 million). In Germany, about 2 million people are employed in the clean-tech

Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

industry that is almost 5% of its total workforce (European Employment Observatory, EEO, 2013). In UK, about 940,000 people are employed in clean-tech, compared to about 213,000 in telecommunications (UK BIS, 2013).

The strong presence of start-ups and small and medium-sized enterprises (SMEs) in the industry and the increased perception of commercial prospects have drawn investors towards innovative clean-tech companies. Investment in clean-tech increased from nearly US$30 billion in 2005 to about US$160 billion in 2012. Investment at the global level rose by nearly 10% year on year. For example, investment for the first quarter of 2014 investment almost reached US$50 billion (Bloomberg New Energy Finance, BNEF, 2014a). The increase is striking compared to weak results achieved in 2012 and 2013. Investment has been driven largely by small solar projects in Japan and America, and emerging markets. Overall, investment in wind-tech has fallen, while in solar power, the largest share of renewable-energy spending, investment rose by 23% to US$27.5 billion (The economist, 2014). Recently, BNEF (2014a) forecast that US$5 trillion of an estimated US$7.7 trillion of global energy investment will be spent on clean-tech and renewables by 2030.

The industry has received increasing venture capital interest over the last few years. Interestingly, although in 2009 the economic recession led to a significant decrease in venture capital funding across many sectors, clean-tech was less affected and attracted 25% of venture capital investment worldwide and 20% of venture capital investment in the US alone (Baker 2010, The Economist 2014). According to Ernst & Young (2013), there are currently 1,400 clean-tech companies around the world, often spatially concentrated in clusters, which have raised US$27 billion in venture capital in the period 2006-2013. According to CBInsights (2014) products, services and technologies that attract the most interest are in renewable energy, with the 22 largest renewable energy investment funding in 2014 totalling more than US$2.5 billion. Beyond these capital-intensive ventures, energy efficient technologies (e.g., hardware and software for energy management systems to monitor and conserve energy use, efficient light fixtures, air conditioning and heating systems, low carbon building materials and increased data server efficiencies) have proven to be very successful and are expected to have significant gains in the future as their returns become more predictable.

R&D is the lifeline of the clean-tech industry and its current focus is on both improving the energy efficiency of existing processes and bringing new technologies into the market. In 2008, 70% of total R&D expenditure was allocated to energy-smart technologies such as efficient lighting and insulation, smart-grids, new batteries, and hybrid or electric vehicles. Since innovation is fostered by a plethora of high-tech start-ups and SMEs, which often need to enter horizontal joint ventures and vertical R&D agreements, any empirical-based attempt to detect the pattern of innovation in the clean-tech industry is challenging (Tan, 2010).

It has been estimated that start-ups make up over 90% of clean-tech in UK, which compares similarly to other sectors like ICT (UK BIS, 2010) or biotechnology, and differs from industries like mining where large companies dominate. Obtaining updated, detailed and reliable information about the pattern of technological innovation in this fast growing and highly fragmented market is fundamental for private sector players as well as policy makers and yet identifying driving products, services, technologies and complementarities in the industry is key to ensure its continued growth.

This paper proposes an investigation of innovation in clean-tech using metadata provided by CrunchBase, the world's most comprehensive database on high-tech companies. Metadata are tags, keywords and terms that help describe an item and that in the database reveal products, services and technologies driving innovation. Such information

Accepted version
Licence CC BY-NC-ND Please cite as:
Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

on the clean-tech industry worldwide and for San Francisco, the leader in clean-tech innovation with more than two hundred specialised companies, is represented by means of a network analysis and the main metrics of the resulting networks (at both node and network level) are discussed from an economic point of view. The purpose of the paper is to understand specifically specializations and technological complementarities underlying innovative companies, detect emerging industrial clusters at the global and local/metropolitan level and, finally, suggest a way to realize whether observed start-ups, SMEs and clusters follow a technological path of complementary innovation and market opportunity or, instead, present a risk of lock-in. The discussion of the results of the network analysis shows interesting implications for energy policy, which are particularly useful from an operational point of view.

1.2. A short literature review on clean-tech innovation

There is a growing body of studies on clean-tech in major scientific industry-specialized journals and academic reviews on technology, business strategy, management and entrepreneurship. To review the large number of studies, a Boolean search was carried out on most common scientific databases based multiple keywords (e.g., clean technology, green technology, innovation, clean-tech start-ups, environmental technology, renewable energy, green clusters, technological change, technological path, and so on).

Literature from a wide range of disciplines supports the view that the clean-tech industry is a catalyst for economic development, employment, innovation clusters and revitalization of cities (see, inter alia, Kammen et al., 2006; Roland-Holst, 2008; European Commission, 2009; Wei et al., 2010; Cai et al., 2011; Singh and Fehrs, 2011; Chapple et al., 2011; Carley et al., 2011; Yi 2013, 2014). Much attention has been given to the external and

Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

firm-specific factors that facilitate the founding and growth of clean-tech start-ups and SMEs (Bjornali and Ellingsen, 2014).

Policies have received consideration specifically for their effectiveness or influence on demand for clean-tech and firm innovation (Tsoutsos and Stamboulis, 2005; Nemet, 2009; Doblinger et al., 2013; Hoppmann et al., 2013). Some research has looked at how policies can facilitate the fastest possible rate to adopt clean-tech (Kaplan, 1999; Jacobsson and Lauber, 2006; Veugelers, 2012). Eyraud et al. (2013) showed that macroeconomic policies, in particular those enhancing GDP growth and lowering the cost of capital, are useful to stimulate green investments. However, not all government measures are successful to attract and/or maintain high investment rates.

There is widely accepted empirical research that investigates several and diversified conditions facilitating clean-tech start-ups and SMEs growth.

Walsh (2012) demonstrated that commercialization of innovation is influenced by two important market dimensions, renewable energy technology demand and eco-sophistication of the market. York and Venkataraman (2010) suggested that innovation that improves environmental performance can be best achieved if entrepreneurs are given incentives to focus on businesses that are ethically motivated and where innovation delivers superior environmental performance than current norms (see Meek et al., 2010 and Sine and Lee, 2009). Green et al. (1994) investigated 800 firms from different manufacturing sectors and identified the reorganization of R&D goals, recruitment of new skilled staff and search for collaboration with customers and suppliers as the drivers of environmental innovation. Tsoutsos and Stamboulis (2005) proposed that learning processes across the value chain, cooperation within the industry and flexible financing mechanisms are relevant factors. Hansen (2014) analysed the role of spatial proximity between firms in the clean-tech industry and showed that when firms seek partners to develop specific and innovative products, collaboration is difficult and costly, but often necessary (on partnership and collaborative projects in the clean-tech industry, see Tanţău and Chinie, 2013; Meyskens and Carsrud, 2013; Doblinger et al., 2013).

Research on technological pathways to moderate climate change has addressed their uneven spatial development: Theyel (2012) studied the location patterns of the US wind turbine manufacturing industry; Fornahl et al. (2012) investigated whether the new industrial path of offshore wind-energy emerged out of existing paths, mainly shipbuilding, in Germany; Cooke (2012) argued that the creation of new economic pathways in renewable energy technologies took place in Denmark and Sweden, where these have been created out of their previous path-dependent industrial and knowledge bases; Essletzbichler (2012) reviewed the regional and local geographies of energy transitions by comparing Denmark, Germany, Spain and the UK. Knight and Howard (2012) looked at the effects of government policies on generating new economic pathways in renewable energy technologies and focussed on the role of carbon pricing in developing a market-led response to low carbon energy innovation.

This research is a contribution to innovation studies applied to clean-tech, as it investigates the industrial pattern of innovation and suggests a way to realize whether observed companies and clusters follow a technological path of complementary innovation and market opportunity. More specifically, the paper allows the detection of emerging industrial clusters in a more detailed way than current methodologies and exercises that, limited by available information, indicate the existence and location of clusters only if immediately or statistically visible based on anecdotal evidence or standard industrial classifications (such as, National Classification of Economic Activities, NCA and Statistical Classification of Economic Activities in the European Community, NACE; for an overview of criticisms, see Martin and Sunley, 2003). Metadata reveal supplied products and

Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

services, specializations and technologies, production and R&D activities carried out by innovative companies and underlying interconnections between them allow to distinguish industrial clusters, typically meant as geographically proximate groups of companies, linked by commonalities (that is, sharing features or attributes such as markets or technologies) and complementarities (which, as known, come in many forms; the most obvious is when products and technologies complement one another in meeting customers' needs).

2. Method

2.1. Network analysis in economics and innovation studies

Network analysis is a well known and used methodology founded by sociologists and researchers in social psychology and then further developed in collaboration with mathematicians and statisticians to a point where it is currently employed in several disciplines, such as economics, marketing and industrial engineering. Clearly, its versatility across different sectors and complementarity with various disciplines are factors in its wide usage.

Network analysis has many potential uses within the field of economic analysis. The two key areas of application are the understanding of the way networks influence economic activity, such as R&D and patent activity, and its ability to reveal network formation and network influence (for an overview of network applications in economics: Bloch, 2004; Jackson, 2006, 2011).

Accepted version
Licence CC BY-NC-ND Please cite as:
Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

Several studies have investigated R&D and innovation activities by referring to network analysis. It is not a coincidence that Hägerstrand (1967), in the first pioneering work on the geographical dimension of knowledge with a focus on spatial aspects of innovation diffusion, examined the dissemination of innovations as a spatial process and hypothesized that geographic differences of human behavior must be analysed in terms of a network of interpersonal communications through which the knowledge diffuses. Cantner and Graf (2006) applied network analysis to describe the evolution of the innovators network in Jena, Germany between 1995 and 2001. Owen-Smith et al. (2002) looked at R&D cooperation and compared the organization of scientific research in the US and Europe by employing network analysis. Balconi et al. (2004) analysed inventor networks resulting from common team-membership in patenting, focusing on the specific role of academic inventors in different technological classes. Paci and Batteta (2003) investigated localized knowledge transfer and examined the technological networks represented by the flows of patent citations in different sectors.

In the present paper network analysis allows keeping up with emerging markets, products, services and technologies, emphasising technological complementarities and clusters formation in a rapidly evolving industry where is challenging to track companies R&D efforts and technological change.

2.2. The proposed methodology on metadata

Given available data, two perspectives of investigation could have been pursued:

- a network of clean-tech start-ups and SMEs, where links between start-up$_i$ and start-up$_j$ result from tags used for both of them, which is based on a two-mode matrix X_c, where rows represent the companies and columns represent the tags;

- a network of clean-tech tags, where links between tag_i and tag_j result from the co-existence of tag_i and tag_j in the same start-up, which is based on a two-mode matrix X_t, where rows represent the tags and columns represent the companies.

In the former case, the square matrix indicating the number of links between start-up_i and start-up_j would be called the adjacency matrix Ac, which is computed as the product of X_c and its transposed (X'_c). In this case, the proposed methodology would have been similar to the approach followed by Cantner and Graf (2006) in their investigation on the innovators network in Jena, where nodes represent innovators and links between innovators are formed whenever they patent in the same technological class. This approach is called "technological overlap" and consists in linking innovators by their technological knowledge: the more fields of research the innovators have in common, the closer they are related (the number of technological classes in which two actors both hold patents is meant as a necessary condition for cooperation as innovators share a minimum of common knowledge for understanding each other).

In the latter case, a network of clean-tech tags, the square matrix indicating the number of links between tag_i and tag_j would be called the adjacency matrix A_t, which again is computed as the product of X_t and X'_t. For the purpose of the study, the paper employs this second perspective, where the nodes of the network represent tags and the co-occurrence in one or more companies that use different tags is depicted through the edges of the network. For example tags A and B are linked in the network if these coexist in the same company and the weight is heavier if the number of companies in which the two tags coexist is larger. Therefore for tags A and B, the weight of the edge A-B is 5 since these coexist in five different companies and the weight of the edge A-C is 2 since these coexist in two different companies, and so on (Table 1).

Marra A., Antonelli P., Dell’Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

Table 1

Edges represent relationships and complementarities between two nodes. Heavier or more weighted edges represent the actual links between two tags over which innovative start-ups usually and in a combined way invest tangible and intangible resources and typically drive their business, production activities and R&D efforts. Weaker edges originating from a specified node also bring information and can be interpreted in terms of potential for horizontal and/or vertical expansion and collaboration based on less complementary or disruptive innovations.

To sum up, the proposed network analysis reveals links between tags (which are formed whenever these co-occur in the same company) and not between companies (which instead would be formed whenever these are active in the same market, are specialized in the same R&D domain, offer the same product, make use of the same technology). To visualize data, the open-source network analysis software package Gephi is used (see Bastian et al. 2009).

2.3. Data selection criteria

CrunchBase is the world’s most comprehensive database on high-tech companies, accessible to everyone through an application-programming interface (API). Founded in 2007, CrunchBase began as a simple crowd sourced database to track high-tech start-ups covered on TechCrunch (one of the most highly regarded blogs concerning technological innovation on the web). Today CrunchBase collects information on more than 200,000 profiles of start-ups and other companies and is maintained by tens of thousands of contributors.

Companies listed in CrunchBase are active all over the world in several high-tech industries including bio-tech, clean-tech, nano-tech, finance, hardware, software, mobile, e-commerce, and so on. For most start-ups and SMEs the database includes information such as the city of registration and operating offices, number of employees, category code, total money raised, number and timing of financing rounds and tags related to markets, products, services, technologies, and so on.

Data from CrunchBase is increasingly used in research (Block and Sandner, 2009, 2011; Waldner et al., 2012; Yuxian and Yuan, 2012; Werth and Boert, 2013; Adcock et al., 2013; Homburg et al., 2014; Marra et al., 2014). Block and Sandner (2009) tested the quality of information in CrunchBase by comparing the funding data in the US tech industry with the industry statistics published by the National Venture Capital Association (NVCA, a trade association representing the US VC industry): the number of deals in CrunchBase amounted to about 97% of the internet-related deals as reported by NVCA and a high and statistically significant Pearson correlation coefficient was found between the time series of new deals related to the two sources ($r=0.67$; $p<0.05$).

The observed dataset includes more than one thousand clean-tech companies founded between 2001 and 2013 and for which there was available data on company foundation, industry category code, tags, number of employees, funding information and office location. Further underlying the decision to have a large sample, data selection criteria were a reflection of the focus on innovative and young companies for which available information attached to them allows for the descriptive purpose of the study.

Some key economic performance indicators are illustrated below referring to a subset of 563 start-ups, all of which have complete information on financing and location. Firstly, the trend in the number of clean-tech start-ups and SMEs by year of foundation reveals that the number of founded companies rose substantially between 2007-2009. A minor decline

Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

in 2012 and 2013 is due to the delay with which the tech community updates CrunchBase (the date of data extraction through the API was the 30th of December 2013).

Secondly, the volume and annual trend in funding help to depict the dataset relevance. Data about the funding volume per year of investment is available and is consistent with other estimates (at least in terms of trend) since 2007 when CrunchBase was launched. Data from CBInsights (2014) showed that funding raised by clean-tech start-ups was US$4.6 billion in 2013, US$8.3 billion in 2012, US$13.7 billion in 2011, US$10 billion in 2010 and US$8.9 billion in 2009. For the observed subset of clean-tech start-ups, the total amount raised equals US$27.8 billion in the years 2009-2013 - the annual financing is above the threshold of US$4 billion.

As said above, in CrunchBase tags are keywords and terms providing company on markets, products, services, technologies, and so on. For example, there is a green building technology company listed in the database that currently supplies concrete product manufacturers in the US and Canada: this company offers concrete producers the ability to manufacture green concrete products without compromising on either quality or price. The tags that define this company are: “clean-tech”, “concrete”, “cement”, “manufacturing”, “green-building”, “materials”, “construction”.

Tags are a reliable source of information and consistent as other financial and geographic data provided in CrunchBase, which benefit not only from the wide number of contributors but also from the high quality of the attached technological blog. Sorting by relevance of tags, the most high-ranking terms are generic keywords and include clean-tech (1.3% of the total), energy (0.86%), solar (0.7%), environment (0.5%), sustainability (0.44%), recycling (0.3%), smart-grid (0.23%), electricity (0.19%), water (0.17%), solar panels (0.15%). Many other tags (such as lighting control, solar water purification systems, solar monitoring software, wind-power forecasting, micro wind-turbines, micro-grid technology, wind-farms, and so on) are more specialised terms and refer to emerging market niches, key technologies, innovative products and services driving change in the industry (hereinafter, sometimes called "technological hotspots").

To notice that available data do not allow to offer a dynamic perspective of investigation: the foundation date cannot be referred to as the moment in which metadata are attached to the company, since contributors often intervene later to update the company profile and insert new tags.

Cities are often interpreted as the nurseries of young high-tech companies. Looking at the percentage of clean-tech companies per city it emerges that San Francisco is at the top of the list with 7% of the start-ups in the dataset, followed by London (5,8%), New York (4,7%), Cambridge (4,6%), San Jose (3,3%), Austin (3,1%) and Boston (3,1%). California continues to serve as the epicentre of the US clean-tech market, maintaining a broad leadership role ranging from clean electricity deployment, energy efficiency (in particular, dominating in high-profile areas like electric/hybrid vehicle adoption, smart-meter installations, solar power capacity), and policy innovation and investment attraction. A major role is played by San Francisco, San Jose and San Diego: from 2006 to 2007 clean-tech grew in Silicon Valley by 94% (Nauman, 2007). San Jose aims to be the "home of clean-tech", but Boston and Austin are also competing for that title. Other states also investing in clean-tech include New Jersey, Ohio and Iowa. Also it is interesting to note that Iowa and South Dakota are rivalling the world's leading clean-tech nations for pre-eminence in many areas.

3. Results and discussion

3.1. The clean-tech network worldwide

The network of clean-tech companies worldwide has 2,261 nodes (or tags) and 12,941 edges. Aside from generic and very common terms (e.g., clean-tech, energy, environment, sustainability and renewable-energy), the network is represented by some major technological hotspots that well describe the areas catalysing the attention of innovative start-ups and SMEs: smart-grid, electricity, wind, batteries, recycling, design, photovoltaic, software, engineering and solar-energy.

The relevance of the hotspots in the network is widely supported by data on investments and financing from several reliable sources. According to the World Bank (2014), investment in renewable power and fuels was dominated by solar in 2013, which rose from US$12.1 billion in 2004 to US$113.7 billion in 2013, with a CAGR of 28%, higher than any other renewable energy source (see also BNEF, 2014b). McKinsey (2014) estimated that since 2006 global installations of solar panel have continued to rise by an average of 50% year on year. CBInsights (2014) found that global investments in the recycling sector increased from US$110 million to US$210 million in the period 2010-2013. Smart-grids and storage have become the new preferred high environmental performance energy systems driving the energy industry, with deployment of smart-grid technologies becoming a critical factor if states or cities want to be recognized as leaders in energy efficiency. Overall, there seems to be also a move towards chemicals, transportation and energy efficiency (due to investors favouring capital-light deals). The companies that may benefit most from this shift are the ones involved in water management, waste-to-energy and "clean-web" (software, applications and data analytics) technology, as also demonstrated by the European clean-tech market.

The main metrics used to describe the clean-tech network worldwide are the average degree of all the nodes in the graph and the average weighted degree, representing respectively the total number of edges connected to each node and the sum of the edge weights of all edges connected to all nodes. Values, 11.5 and 16.8, suggest that worldwide the clean-tech industry is extended, diversified and interconnected compared to other industries (e.g., biotech or healthcare). The network diameter is 7 and the modularity is 0.7. These, together with an average path length at 3.1, indicate the emergence of some aggregates or clusters of tags (Table 2, first column "Extended").

Table 2

The network has more than one hundred clusters, meant as groups of nodes more densely interconnected internally than with the rest of the network. At a high level, the most interesting ones are those concerned with solar, wind energy, fuel cell technology and "clean-web".

The "solar" cluster includes photovoltaic, solar financing, solar panels, smart-grid technology, electric systems, lighting control, solar water purification systems, solar renewable energy certificates and solar monitoring software. The "wind energy" cluster includes wind-turbines, wind-power forecasting, wind-electric, wind-energy patents, energy storage, micro wind-turbines, wind-tower raising systems, micro-grid technology and wind-farms. For "fuel cell technology" hotspots include direct-methanol fuel cells, proton-exchange and polymer electrolyte membrane fuel cells, membrane electrode assembly and gas diffusion electrodes. The "clean-web" cluster, which incorporates all aspects of web-based clean-tech business, has hotspots in software, energy efficient, solar, information technology, mobile-phone, storage, big-data and product stewardship.

Narrowing the extension of the network enables the identification of technological hotspots and detection of relevant complementarities between them. For a degree range set above

35, the resulting network has 68 nodes and 598 edges (Figure 1). Accordingly, through the weighted degree, it is possible to detect the most significant links, or technological complementarities, between couples of nodes. These are between clean-tech, renewable energy, solar, photovoltaic, smart-grid, wind, hydrogen, electricity, recycling, solar panel, green-buildings, led and lighting.

Figure 1

The resulting delimited network has an average degree that equals 17.6, while the average weighted degree is 50.4. The network diameter is at 3, the graph density equals 0.3 and the modularity is 0.2 (see Table 2, second column "Restricted"). According to these last metrics and the value of the average path length at 1.8, no relevant clusters emerge in the network: this is due to the relatively small dimension of the network compared to the extended one with no range degree. This suggests that the proposed methodology needs large sample (e.g., more than one hundred start-ups and SMEs and complete data on company profile) to produce valuable information. As shown in Section 4, a network analysis on small sample, because circumscribed to a metropolitan area, needs to be corroborated with other sources such as specialized blogs, reports or materials with focus on the observed geographical area and industry to check for the effective presence and relevance of emerging specializations and clusters.

To investigate the level of attractiveness of single nodes, the degree centrality (a measure of local centrality of the most visible nodes in the network) has been estimated. According to this measure, a node with a large degree is in direct contact to many other nodes and is immediately recognized as a hub. Central hotspots are solar, renewable-energy, power, energy-efficiency, smart-grid and electricity, recycling, battery and wind. The closeness centrality, instead, suggests that a node is central if it is linked with all the others, not only with first neighbours: accordingly, among the closest hotspots there are smart-home, air-conditioning, fuel-cells, carbon-offsets, led, data, chemicals, saas and hydrogen.

3.2. The clean-tech network in San Francisco

The above analysis provides investors, companies, incubators and policymakers with knowledge on emerging markets, services, products, technologies and underlying complementarities between these. If some technological hotspots emerge and strong connections result from a geographically based view, then the network analysis offers a tool to support decisions based on evidence of actual trends and production activities from local industry, specific identification of R&D strategies and vertical specialisations of existing local companies. This is crucial information for financial institutions and policymakers who want their actions and policies to be more targeted and effective in young and dynamic industries, such as clean-tech.

The analysis above has been replicated at the local level, for the metropolitan area of San Francisco. San Francisco is the leader in clean-tech innovation. With more than two hundred clean-tech companies, excellent state and local financial incentives for high-tech production and R&D, San Francisco is home to the country's largest concentration of venture capital investors in the industry.

The network linking clean-tech companies in San Francisco has 167 nodes and 837 edges: no similar network for cities has been found from the observed dataset in clean-tech (for example, the clean-tech network in New York has 157 nodes and 612 edges). Some major hotspots emerge from the clean-tech network in San Francisco: aside from common and generic terms, technological hotspots are given by solar-energy, solar panels, solar-financing, green building, and green transportation.

Accepted version
Licence CC BY-NC-ND Please cite as:
Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

Recently, San Francisco has promoted solar power installations through local incentive programs, providing rebates for solar installations, as well as workforce development incentives for installers. As result, with a generating capacity of 5 megawatts, San Francisco hosts the largest urban solar array in California. The city has also stimulated the private sector to adopt green building practices by streamlining the photovoltaic permit process, instituting a priority permitting program for private projects and establishing binding green building specifications. In 2011, San Francisco has been recognized as one of the largest markets for green building opportunities and among the top places in the sustainable buildings category.

The average degree equals 10 and the average weighted degree is at 16.8. The network diameter is 7; the graph density equals 0.1, and the modularity at 0.7. The average path length equals 3.1 (see Table 3, first column "Extended").

Table 3

The last mentioned metrics confirm that some clusters exist. More specifically, fourteen clusters emerge, among which one is related to "solar energy" (represented by hotspots such as solar panels, residential-solar, energy management, pay-as-you-go technology, solar hot water systems, energy analytics, solar-pool heating systems, photovoltaic systems, home-solar and wireless energy management), which in San Francisco shows strong connections with the real estate industry and residential buildings. Within this cluster it is interesting to identify companies focused on more established markets such as photovoltaic and solar panels and start-ups specialized in smart home technologies to control home's energy through digitally interactive tools and energy management and analytics.

Accepted version
Licence CC BY-NC-ND Please cite as:
Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

Another interesting cluster concerns the “carbon trading market” with tags focussing on the carbon network, carbon offsets, renewable energy certificates, carbon sequestration, carbon balancing programs, carbon credits and emission management software. In 2009 San Francisco International Airport was the first airport in the US to introduce a passenger offset program, managed by a local company, which allows them to calculate and reduce the carbon footprint of their air travel. Moreover, several companies offer a wide range of business greening services including consulting, ecological foot printing, and carbon offsetting.

Last, with emphasis on the electric vehicle specialization, the “smart-mobility” cluster uses terms such as electric cars, green transportation, lithium-ion batteries, electric motorcycles, green travel, hydrogen technology and electric vehicle battery. Several companies in San Francisco produce electric vehicles and related technologies and achieved technological advances in engineering and supplying high-performance powertrains, energy storage systems, drive systems and software, and electric and hybrid vehicles. Such innovation had ripple effects on local industry, inspiring a new wave of start-ups working to design and manufacture new efficient vehicles and storage battery systems, and to develop apps that locate available charging stations, inform a driver on the battery life, and so on.

With degree set above 13, the restricted clean-tech network in San Francisco has 25 nodes and 139 edges.

Figure 2

Accepted version
Licence CC BY-NC-ND Please cite as:
Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

The average degree equals 11.1, while the average weighted degree is 13.3. The network diameter is at 3, the graph density equals 0.5 and the modularity is 0.3 (Table 3, second column "Restricted"). According to this last metric, it emerges that still there is a strong presence of clusters or specializations in the network, in line with the value of the average path length (1.6) and graph density (0.5). According to the weighted degree, the most significant connections and complementarities are those between solar panels, solar-energy, clean-tech, sustainability and solar-financing.

The degree centrality suggests that most of the interesting hotspots are those related to green building, green-transportation, green-travel, green-computing, corporate-responsibility and green investing. According to the closeness centrality, central nodes are smart-grid, solar-financing, clean-energy, solar panels and solar installations.

High value of the centrality indicates not only that technological hotspots are well connected to other nodes but also that such markets, products, services and technologies represent very attractive areas from an R&D point of view because they can be adopted or integrated more readily into existing technological regimes and systems. In particular, the betweenness centrality of a node, which measures how often a node appears on shortest paths between nodes in the network, could provide key information to realise whether complementary innovation and market opportunity is potentially attached to a specific tag and, then, to the companies and clusters that lever on those technological hotspots. The underlying rationale is that the closer a single tag is to other products, services and technologies, the more the chance that this tag could benefit from technological progress and business opportunity related to the surrounding (that is, complementary) environment. Even if too often overlooked by companies, focused on disruptive or radical innovations and new opportunities, complementary innovations are supportive to current regimes as they also increase the value of existing products and services. An example is given by the technological advances originating in energy storage and the positive externalities that could arise for wind and solar technologies: storage increases the quality of generated power, by regulating the voltage, and the reliability of the power, which would be no longer dependent on weather conditions. Similarly, storage technologies could help to manage energy demand, exploiting excess generation stored during the night while reducing the need for extra generation during the day. Last, the technological complementarity between stationary fuel cells and electricity systems can reach its full potential once further infrastructure developments take place.

If clean-tech companies do not belong to industrial clusters made of nodes with high value of centrality, the risk of lock-in effects is higher since the probability to exploit a sufficient amount of complementary innovations and market opportunities is limited. This exercise might be meant as a valuable forecasting tool if centrality of all the nodes shaping clusters in San Francisco is calculated on the clean-tech network worldwide (Table 4).

Table 4

The results suggest that the tags in the solar cluster have the most technological complementary connections revealing it to be a versatile and interesting area in the near future. The overall value, weighted for the number of nodes, is 17.654, followed by the smart-mobility cluster (9.933) and the carbon trading market cluster (1.416). In recent years solar energy has experienced significant growth, as a result of lower prices of competitive and complementary technologies, more financing opportunities, larger public policy support, increasing market size and emerging innovative business models. The rapid cost decrease of photovoltaic modules and systems in the last few years opened new perspectives for using solar energy as a major electricity source, and solar plants and

Accepted version
Licence CC BY-NC-ND Please cite as:
Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

batteries are disruptive technologies for the power system. Complementary technologies and innovative business models are creating new applications and consumption approaches. For example, the emergence of customer-focused business and financing models creates a platform for several complementary energy services, such as energy storage and energy management systems. Moreover, as seen, the low carbon, predictable, and flexible characteristics of solar make it a promising complement to electric vehicles.

Investigating the co-existence between tags attached to all high-tech companies located in San Francisco listed on CrunchBase, it is possible to find some evidence to determine whether product and technological overlap takes place across industries and to what extent. In this case the network is between high-tech industries, where links between industry_i and industry_j result from the co-existence of a certain tag in both industries, which is based on a two-mode matrix X_S, where rows represent the industries and columns represent the tags. The resulting square matrix indicating the number of links between industry_i and industry_j gives rise to the adjacency matrix A_S, which is computed as the product of X_S and its transposed (X'_S). The emerging network between industries shows 593 edges, with an average degree of 32.1 and an average weighted degree of 34.6: clean-tech is an interconnected industry (in line with average degree), linked to most other sectors, in particular with software and web.

4. Conclusions and policy implications

4.1. Clean-tech network analysis overview

Accepted version
Licence CC BY-NC-ND Please cite as:
Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

As seen, clean-tech is a fast growing industry and a fertile area of innovation, for which it is difficult to detect the industrial pattern of innovation carried out by thousands of high-tech start-ups and SMEs all over the world.

This paper has proposed a network analysis to understand specifically specializations and technological complementarities underlying innovative companies, detect emerging industrial clusters at the global and local/metropolitan level and, finally, suggest a way to realize whether observed start-ups, SMEs and clusters follow a technological path of complementary innovation and market opportunity or, instead, present a risk of lock-in. The analysis allows appreciating technological complementarities underlying companies' production and R&D activities and where innovative start-ups drive their business. It could also reveal useful information for horizontal and/or vertical expansion and collaboration between companies emphasising the potential of unexploited connections (i.e., disruptive or radical innovations).

The results have given a clear snapshot of the innovation pattern in the clean-tech industry in the twenty first century and, hence, vertical specializations, prevailing R&D efforts and emerging clusters. As seen, the global technological network builds on existing major hotspots, such as solar, recycling, smart-grid, electricity and water: the relevance of the above listed hotspots is widely supported by evidence of where venture capital is invested.

At the global level, the most relevant clusters are those connected to solar (denoted by tags such as solar panel, photovoltaic, smart-grid technology, lighting control, solar-water purification systems and solar financing), wind-energy (represented by wind-turbines, wind-power forecasting, wind-energy patents, energy storage, wind-tower raising systems, micro-grid technology and wind-farms), fuel-cell technology (direct-methanol fuel-cells, proton-exchange and polymer electrolyte membrane fuel-cells, membrane electrode assembly, gas diffusion electrodes) and clean-web (characterized by tags such as big-

Accepted version
Licence CC BY-NC-ND Please cite as:
Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

data, software, energy efficient, solar, information technology, mobile-phone, storage, product stewardship).

San Francisco is specializing in solar energy (emerging tags include solar panels, residential solar, energy management, solar-hot water systems, energy analytics, solar-pool heating systems, photovoltaic systems, home-solar and wireless energy management), carbon trading market (with main tags including carbon offsets, renewable energy certificates, carbon sequestration, carbon balancing programs, carbon credits and emission management software) and smart mobility (with connections emerging between electric cars, green transportation, lithium-ion batteries, electric motorcycles, hydrogen technology and electric vehicle battery).

4.2. Energy policy implications

Since low-carbon economy needs huge investment and the development of new technologies, policies should be primarily directed to this goal not only by promoting their deployment (e.g. through incentives), but also by favouring the birth and the growth of innovative companies. The understanding of the technological pattern behind a specific industry is compelling to define more effective policies and boost economic growth.

The results of the network analysis have interesting policy implications. The analysis provides investors, companies, incubators and, above all, policymakers with a reliable (evidence based), detailed (based on descriptive tags) and geo-localized knowledge on new markets, products, and technologies. Findings are even more useful for geographically localised industries: if the local dimension is correctly identified in terms of technological hotspots, complementarities and clusters, the public sector is able to design and implement effective and supportive actions to local R&D and production activities.

Accepted version
Licence CC BY-NC-ND Please cite as:
Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

It is widely recognized in the US that sub-federal entities such as states and cities have long supported clean-tech growth: these entities compete with each other with ad hoc policies, such as funding and the establishment of environmentally friendly zones (Bürer and Wüstenhagen, 2009; Jung and Tyner 2014). It is US cities rather than Federal government that have been at the forefront of climate action and CO2 reduction. The US Conference of Mayors' Climate Protection Agreement has been signed by more than a thousand mayors nationwide with the purpose to meet the Kyoto Protocol emissions-reduction goal. More recently, several large US cities have become members of the global C40 Cities Climate Leadership Group, in which more than seventy-five metro areas around the world collaborate and share best practices and policies that aim to reduce emissions and use energy and other natural resources more efficiently. Such collaborations show that cities, often frustrated with federal or national policies, can exhibit leadership on their own, and to support the design of effective energy policies mayors and urban/industrial planners need detailed and local information on existing businesses and specializations, R&D investments and complementarities.

From an operational point of view, a data repository containing metadata would allow to bypass the implicit limits of too broad nomenclatures of economic activities or industry classifications. As well known, industrial clusters rarely conform to standard industrial classification systems, which fail to capture firm specializations. Using data that more accurately reflects the range of specialized economic activities carried out by companies under observation is crucial to define the industrial boundaries of clusters.

Having the means to signal advances in R&D, new products or prevailing technological trends, policy makers would have the ability to continually identify and direct active industrial players and financial institutions towards emerging businesses and specializations and/or products and technologies that show complementary innovation and

Accepted version
Licence CC BY-NC-ND Please cite as:
Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

business opportunity. Spatial barriers are still significant and business location choice is limited by the nature of R&D and innovation processes, which favour geographical concentration of high-tech and interdependent value adding activities and lead to clustering. Moreover, they would also be able to promote interaction between complementary technologies and support the creation/aggregation of core competences. Finally, policy makers could facilitate closer relationships between innovative companies, suppliers and clients, venture capitalists, large corporations and research laboratories involved in the clean-tech industry (Horwitch and Mulloth, 2010).

Experience and lessons learned in the US and several countries point to effective innovation systems, where public entities provide a real boost and deliver concrete actions to stimulate R&D spending and increase company size, total number of patents and quality of innovation. Mazzuccato (2011) emphasized that a targeted public innovation program is essential and referred to many of the most successful businesses in the US ICT market (such as Intel and Google): in this specific case the US Federal government funded the underlying innovation activities through a small number of specialist public institutions over many decades. To be effective requires targeting innovation and R&D efforts and financial resources in specific areas and directions, building the paths for new opportunities both within a given industry and cross-industry, brokering the interactions between public and private players involved in technological development and facilitating commercialization of R&D output (Block, 2008). On a practical level, a network analysis using metadata enables to focus on a smaller number of technology families, i.e. those where industrial clusters and innovation districts have significant competitive advantage and for which it is high the chance to get complementary innovation and market opportunity.

References

- Adcock, A.B., Lakkam, M., Meyer, J., 2013. CS 224W Final Report Group 37. Stanford University. (http://snap.stanford.edu/class/cs224w-2012/projects/cs224w-037-final.pdf).
- Baker, D.R., 2010. Funding for clean-tech firms plunges 33% in '09. SfGate.com. (http://articles.sfgate.com/2010-01-07/business/17470394_1_venture-funding-cleantechgroup- venture-capitalists).
- Balconi, M., Breschi, S., Lissoni, F., 2004. Networks of inventors and the role of academia: an exploration of Italian patent data. Res. Policy 33, 127–145.
- Bastian, M., Heymann, S., Jacomy, M., 2009. Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media.
- Bjornali, E.S., Ellingsen, A., 2014. Factors Affecting the Development of Clean-tech Start-ups: A Literature Review. Energy Procedia 58, 43-50.
- Bloch F. 2004. Group and network formation in industrial organization, In: Group Formation in Economics; Networks, Clubs and Coalitions. Demange G., Wooders M. (Eds.). Cambridge, UK: Cambridge Univ. Press.
- Block, F., 2008. Swimming against the current: The rise of a hidden developmental state in the United States. Politics & Society 36, 169-206.
- Block, J.H., Sandner, P., 2011. Venture capital funding in the middle of the year 2011: are we back to pre-crisis boom levels?. Strategic Change 20, 161-169.
- Block, J.H., Sandner, P.G., 2009. What is the Effect of the Financial Crisis on Venture Capital Financing? Empirical Evidence from US Internet Start-Ups. Venture Cap. An International J. Entrepreneurial Finance 11, 295-309.
- BNEF, 2014a. 2030 Market Outlook. Bloomberg New Energy Finance, New York. (http://bnef.folioshack.com/document/v71ve0nkrs8e0).

Marra A., Antonelli P., Dell'Anna L., Pozzi C. (2015). A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy. ENERGY POLICY, vol. 86, p. 17-26, ISSN: 0301-4215, doi: 10.1016/j.enpol.2015.06.025

- BNEF, 2014b. Global Trends in Renewable Energy Investment 2014. Bloomberg New Energy Finance, New York. (http://fs-unep-centre.org/publications/gtr-2014).
- Bürer, M.J., Wüstenhagen, R., 2009. Which renewable energy policy is a venture capitalist's best friend? Empirical evidence from a survey of international cleantech investors. Energy Policy 37, 4997-5006.
- Cai, W., Wang, C., Chen, J., Wang, S., 2011. Green economy and green jobs: The case of China's power generation sector. Energy 36, 5994-6003.
- Cantner, U., Graf, H., 2006. The network of innovators in Jena: an application of social network analysis. Res. Policy 35, 463–480.
- Carley, S., Lawrence, S., Brown, A., Nourafshan, A., Benami, E., 2011. Energy based economic development. Renew. Sustainable Energy Rev. 15, 282-295.
- CBInsights, 2014. Venture Capital Database 2014. (https://www.cbinsights.com/).
- Chapple, K., Kroll, C., Lester, T.W., Montero, S., 2011. Innovation in the green economy: an extension of the regional innovation system model?. Econ. Dev. Q. 25, 5-25.
- Cooke, P., 2008. Cleantech and an Analysis of the Platform Nature of Life Sciences: Further Reflections upon Platform Policies. Eur. Plan. Stud. 16, 375-393.
- Cooke, P., 2012. Transversality and Transition: Green Innovation and New Regional Path Creation. European Plan. Stud. 20, 817-834.
- Doblinger, C., Dowling, M., Helm, R., 2013. Does Public Policy Stimulate Firm Entrepreneurial Behavior and Innovative Activities?. Working paper. The 73d Academy of Management Annual Meeting, San-Antonio.
- EEO, 2013. Promoting Green Jobs Throughout the Crisis: A handbook of best practices in Europe. European Employment Observatory Review, Luxembourg. (http://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=7585&type=2&furtherPubs=yes).

- Ernst & Young, 2013. Cleantech matters - Global competitiveness. Ernst & Young, London.
- Essletzbichler, J., 2012. Renewable Energy Technology and Path Creation: A Multi-scalar Approach to Energy Transition in the UK. European Plan. Stud. 20, 791-816.
- European Commission, 2009. The impact of renewable energy policy on economic growth and employment in the European Union. Karlsruhe, 27 April.
- Eyraud, L., Clements, B., Wane, A., 2013. Green investment: Trends and determinants. Energy Policy 60, 852-865.
- Fornahl, D., Hassink, R., Klaerding, C., Mossig, I., Schröder, H., 2012. From the Old Path of Shipbuilding onto the New Path of Offshore Wind Energy? The Case of Northern Germany. European Plan. Stud. 20, 835-855.
- Green, K., McMeekin, A., Irwin, A., 1994. Technological trajectories and R&D for environmental innovation in UK firms. Futures 26, 1047–1059.
- Hansen, T., 2014. Juggling with Proximity and Distance: Collaborative Innovation Projects in the Danish Cleantech Industry. Econ. Geogr. 90, 375–402.
- Homburg, C., Hahn, A., Bornemann, T., Sandner, P., 2014. The Role of Chief Marketing Officers for Venture Capital Funding: Endowing New Ventures with Marketing Legitimacy. J. of Marketing Res. 51, 625-644.
- Hoppmann, J., Peters, M., Schneider, M., Hoffmann, V.H., 2013. The two faces of market support-How deployment policies affect technological exploration and exploitation in the solar photovoltaic industry. Res. Policy 42, 989-1003.
- Horwitch, M., Mulloth, B., 2010. The interlinking of entrepreneurs, grassroots movements, public policy and hubs of innovation: The rise of Cleantech in New York City. J. of High Technology Manag. Res. 21, 23-30.
- Howarth, N., 2012. Clean Energy Technology and the Role of Non-Carbon Price-Based Policy: An Evolutionary Economics Perspective. European Plan. Stud. 20, 871-891.

- IEA, 2014. Renewable energy market analysis and forecasts to 2020. International Energy Agency, Paris. (http://www.iea.org/Textbase/npsum/MTrenew2014sum.pdf).
- ILO, 2013. Report V: Sustainable Development, Decent Work and Green Jobs. International Labour Organization, Genèv. (http://www.ilo.org/wcmsp5/groups/public/ed_norm/relconf/documents/meetingdocument/wcms_207370.pdf).
- Jackson, M., 2006. The Economics of Social Networks, In: The Econometric Society, Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society. Blundell, R., Newey, W., Persson, T. (Eds.). Cambridge, UK: Cambridge Univ. Press.
- Jackson, M., 2011. An Overview of Social Networks and Economic Applications, In: Social Economics. Benhabib, J., Bisin, A., Jackson, M.O. (Eds.). Amsterdam: Elsevier.
- Jacobsson, S., Lauber, V., 2006. The politics and policy of energy system transformation - explaining the German diffusion of renewable energy technology. Energy Policy 34, 256-276.
- Johnson, D., Lybecker, K., 2009. Challenges to Technology Transfer: A Literature Review of the Constraints on Environmental Technology Dissemination. Colorado College Working Paper No. 2009-07.
- Jung, J., Tyner, W.E., 2014. Economic and policy analysis for solar PV systems in Indiana. Energy Policy 74, 123-133.
- Hägerstrand, T., 1967. Innovation Diffusion as a Spatial Process, University of Chicago Press, Chicago.
- Kammen, D.M., Kapadia, K., Fripp, M, 2006. Putting renewables to work: how many jobs can the clean energy industry generate?. RAEL Report. University of California, Berkeley. (http://rael.berkeley.edu/sites/default/files/very-old-site/renewables.jobs.2006.pdf).

- Kaplan, A.W., 1999. From passive to active about solar electricity: innovation decision process and photovoltaic interest generation. Technovation 19, 467-481.
- Marra, A., Luciani, M., Antonelli, P., 2014. Data analytics nel comparto high-tech statunitense: impatti, evoluzione e tutela della privacy. L’industria 35, 319-338.
- Martin, R., Sunley, P., 2003. Deconstructing clusters: chaotic concept or policy panacea?. Econ. Geogr. 3, 5-35.
- Mazzucato, M., 2011. The Entrepreneurial State. London: Demos.
- McKinsey, 2014. The disruptive potential of solar power. McKinsey Quarterly. (http://www.mckinsey.com/insights/energy_resources_materials/the_disruptive_potential_of_solar_power).
- Meek, W.R., Pacheco, D.F., York, J.G., 2010. The impact of social norms on entrepreneurial action: Evidence from the environmental entrepreneurship context. J. Bus. Venturing 25, 493-509.
- Meyskens, M., Carsrud, A.L., 2013. Nascent green-technology ventures: a study assessing the role of partnership diversity in firm success. Small Bus. Econ. 40, 739-759.
- Nauman, M., 2007. ‘Green’ VC funding sets record. San Jose Mercury News. (http://www.mercurynews.com/business/ci_7587942).
- Nemet, G.F., 2009. Demand-pull, technology-push, and government-led incentives for non-incremental technical change. Res. Policy 38, 700-709.
- Owen-Smith, J., Riccaboni, M., Pammolli, F., Powell, W.W., 2002. A comparison of U.S. and European university-industry relations in the life sciences. Management Science 48, 24–43.
- Paci, R., Batteta, E., (2003), Innovation Networks and Knowledge Flows across the European Regions, Contributi di Ricerca CRENoS 03/13, Cagliari.

- Pernick, R., Wilder, C., 2007. The Clean Tech Revolution: The Next Big Growth and Investment Opportunity, HarperCollins, New York.
- Rasmussen, E., Bulanova, O., Jensen, A., Clausen, T., 2012. The Impact of Science-Based Entrepreneurial Firms - a Literature Review and Policy Synthesis. Report 3, 154.
- Roland-Holst, D., 2008. Energy efficiency, innovation and job creation in California. Department of Agricultural & Resource Economics, UCB. (http://are.berkeley.edu/~dwrh/CERES_Web/Docs/UCB%20Energy%20Innovation%20and%20Job%20Creation%2010-20-08.pdf).
- Sine, W.D., Lee, B.H., 2009. Tilting at Windmills? The Environmental Movement and the Emergence of the US Wind Energy Sector. Administrative Science Q. 54, 123-155.
- Singh, V., Fehrs, J., 2011. The work that goes into renewable energy. Renewable Energy Policy Project, Washington, DC. (http://www.globalurban.org/The_Work_that_Goes_into_Renewable_Energy.pdf).
- Tan, X., 2010. Clean technology R&D and innovation in emerging countries—Experience from China. Energy Policy 38, 2916-2926.
- Tanţău, A.D., Chinie, A., 2013. Green Clusters as New Cooperation Strategy for Cleantech Companies, In: The Changing Business Landscape of Romania. Thomas, A.R., Pop, N.A., Bratianu, C. (Eds.). New York: Springer.
- The Economist, 2014. Global investment in clean energy. (http://www.economist.com/news/economic-and-financialindicators/21601272-global-investment-clean-energy).
- Theyel, G., 2012. Spatial Processes of Industry Emergence: US Wind Turbine Manufacturing Industry. European Plan. Stud. 20, 857-870.
- Tsoutsos, T.D., Stamboulis, Y.A., 2005. The sustainable diffusion of renewable energy technologies as an example of an innovation-focused policy. Technovation 25, 753-761.

- UK BIS, 2010. Low Carbon and environmental Goods and Services: an Industry Analysis. UK Department for Business, Innovation & Skills, London. (http://www.berr.gov.uk/assets/biscore/business-sectors/docs/10-795-low-carbon-environmentalgoods-analysis-update-08-09.pdf).
- UK BIS, 2013. Low carbon environmental goods and services (LCEGS) Report for 2011/12. UK Department for Business, Innovation & Skills, London. (https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/224068/bis-13-p143-low-carbon-and-environmental-goods-and-services-report-2011-12.pdf).
- US BLS, 2013. Employment Situation summary. US Bureau of Labor Statistics, Washington. (http://www.bls.gov/news.release/pdf/ggqcew.pdf).
- Veugelers, R., 2012. Which policy instruments to induce clean innovating?. Res. Policy 41, 1770-1778.
- Waldner, F., Zsifkovits, M., Heidenberger, K., 2012. Emerging Service-Based Business Models in the Music Industry: An Exploratory Survey. In: Snene (Ed.), Exploring Services Science. Springer-Verlag, Berlin Heidelberg, pp. 321-329.
- Walsh, P.R., 2012. Innovation Nirvana or Innovation Wasteland? Identifying commercialization strategies for small and medium renewable energy enterprises. Technovation 32, 32-42.
- Wei, M., Patadia, S., Kammen, D., 2010. Putting renewables and energy efficiency to work: How many jobs can the clean energy industry generate in the US?. Energy Policy 38, 919–931.
- Werth, J.C., Boert, P., 2013. Co-investment Networks of Business Angels and the Performance of their Start-up Investments. International J. of Entrepreneurial Venturing 5, 240-56.

- World Bank, 2013. Global Economic Prospects. World Bank, Washington. (http://siteresources.worldbank.org/INTPROSPECTS/Resources/334934-1322593305595/8287139-1371060762480/GEP2013b_full_report.pdf).
- World Bank, 2014. Renewables 2014 Global Status Report. World Bank, Washington. (http://www.ren21.net/portals/0/documents/resources/gsr/2014/gsr2014_full%20report_low%20res.pdf).
- Yi H., 2014. Green businesses in a clean energy economy: Analyzing drivers of green business growth in US states. Energy 68, 922–929.
- Yi, H., 2013. Clean energy policies and green jobs: An evaluation of green jobs in US metropolitan areas. Energy policy 56, 644–652.
- York, J.G., Venkataraman, S., 2010. The entrepreneur-environment nexus: Uncertainty, innovation, and allocation. J. Bus. Venturing 25, 449-463.
- Yuxian, E.L., Yuan, S.T.D., 2012. Where’s the Money? The Social Behavior of Investors in Facebook’s Small World. In: 2012 ieee/acm International Conference on Advances in Social Networks Analysis and Mining, 26-29 August 2012, Istanbul, Turkey.

Tables

Table 1: Row data (example)

	Name	**Tags**
1	Company$_1$	A, B, C, D
…	…	A, B, E, F
…	…	A, B
…	…	A, B, C, G, H, I
5	Company$_n$	A, B, L, M, N

Table 2: Descriptive statistics of the clean-tech network worldwide (extended and restricted view)

	Value	
Metrics	**Extended**	**Restricted**
Number of nodes	2,261	68
Number of edges	12,941	598
Avg. Degree	11.5	17.6
Avg. Weighted Degree	16.8	50.4
Network Diameter	7	3
Graph Density	0.005	0.3
Modularity	0.7	0.2
Avg. Path Lenght	3.1	1.8

Source: Own elaboration on Crunchbase (2013)

Table 3: Descriptive statistics of the clean-tech network in San Francisco (extended and restricted view)

	Value	
Metrics	**Extended**	**Restricted**
Number of nodes	167	66
Number of edges	837	566
Avg. Degree	10	17.2
Avg. Weighted Degree	16.8	50.7
Network Diameter	7	3
Graph Density	0.1	0.3
Modularity	0.7	0.2
Avg. Path Lenght	3.1	1.8

Source: Own elaboration on Crunchbase (2013)

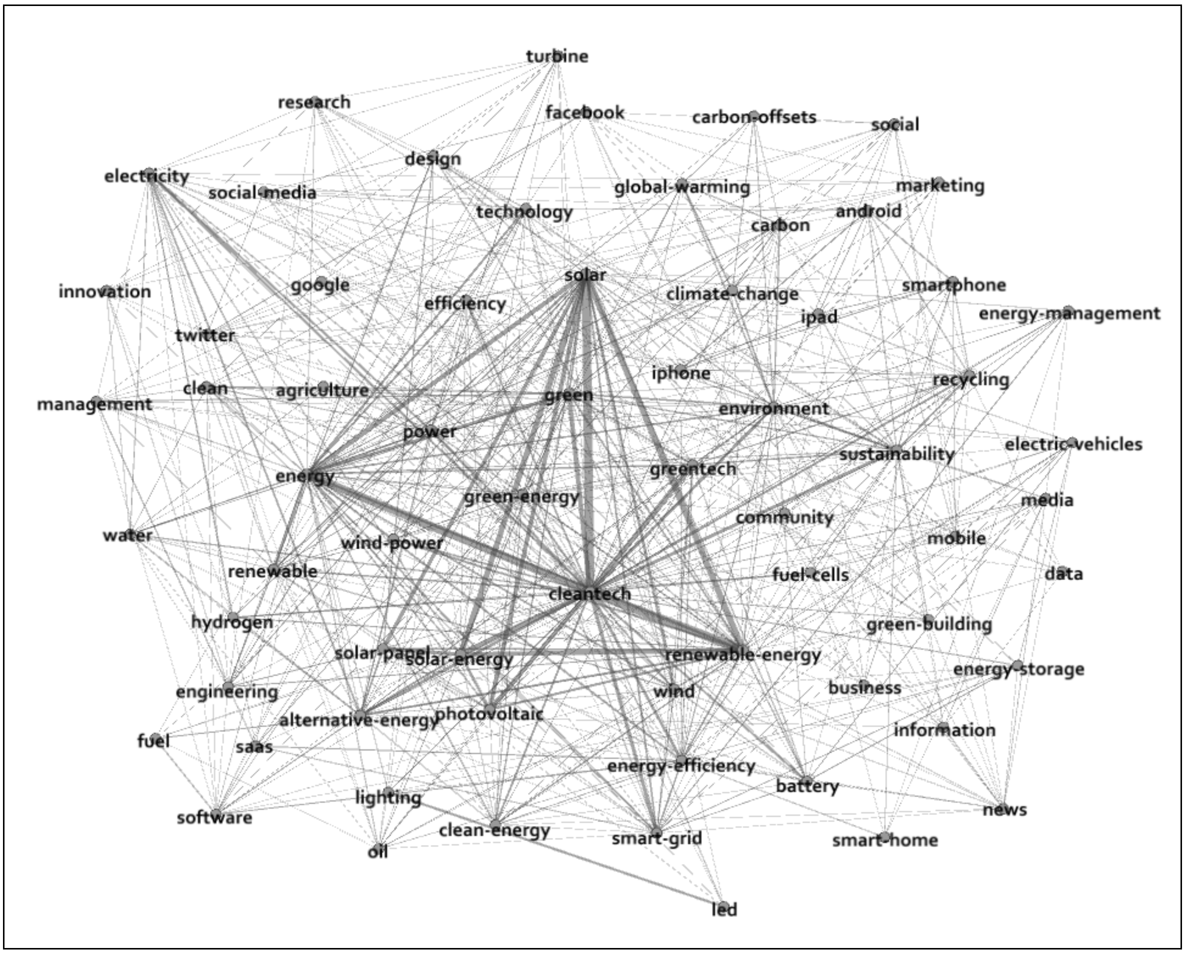
Table 4: The betweenness centrality (B.c.) at the global level for the high-ranking nodes in San Francisco clusters

solar energy		**smart mobility**		**carbon trading**	
nodes	**B.c.**	**nodes**	**B.c.**	**nodes**	**B.c.**
solar-energy	276.715	battery	64.877	carbon-offset	17.012
solar-panels	32.860	electric-vehicles	43.575	energy-credits	1.375
smart-lighting	27.267	fuel-cells	21.685	emissions-software	870
solar-installation	24.819	hydrogen	16.714	carbon-calculator	290
solar-electricity	18.197	green-transportation	6.635	carbon-credits	286

Source: Own elaboration on Crunchbase (2013)

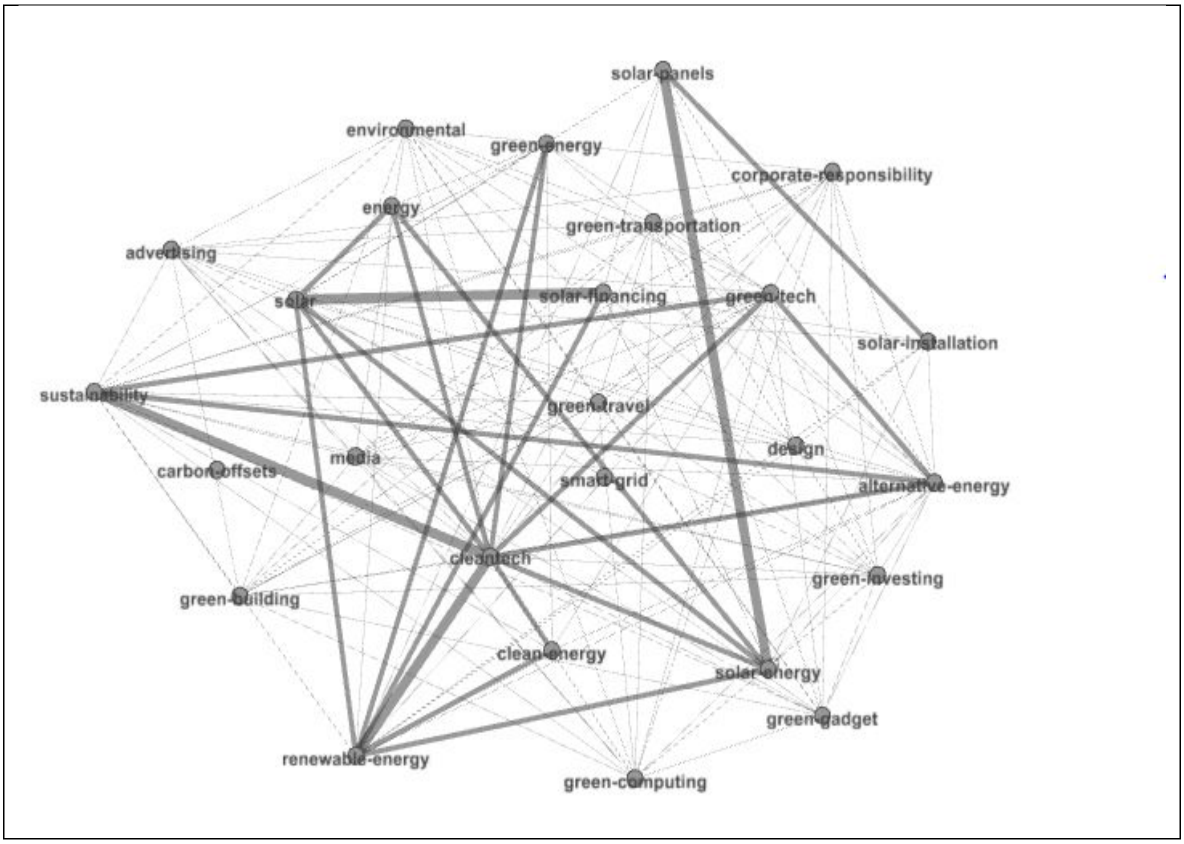
Figures

Figure 1: The clean-tech industry worldwide (restricted view)



Source: Crunchbase (2013)

Figure 2: The clean-tech industry in San Francisco (restricted view)



Source: Crunchbase (2013)