Unveiling scientific communities about sustainability and innovation. A bibliometric journey around sustainable terms

Franceschini, Simone; Faria, Lourenco; Jurowetzki, Roman

Published in:
Journal of Cleaner Production

Link to article, DOI:
10.1016/j.jclepro.2016.03.142

Publication date:
2016

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Unveiling scientific communities about sustainability and innovation. A bibliometric journey around sustainable terms

Simone Franceschini1. Corresponding author – simonefranceschini@gmail.com
Lourenço G. D. Faria1
Roman Jurowetzki2

1Technical University of Denmark, Department of Management Engineering, Diplomvej 372, 2800 Kgs. Lyngby, Denmark
2Aalborg University, Department of Business and Management, IKE, Fibigerstræde 11, 9220 Aalborg, Denmark

Abstract: Literature about the relationship between innovation and sustainability has skyrocketed in the last two decades and new terms have appeared. However, only very few bibliometric analyses have reviewed some of these terms (eco-innovation, environmental innovation, green innovation, and sustainable innovation), and they concluded that such terms are mostly interchangeable. These findings surprise in light of the different positions shown in the innovation for sustainability debate. Our bibliometric analysis tracks meanings and communities associated with these four terms and indicates some overlaps, especially between eco-innovation and environmental innovation. However, we found relevant differences of meanings and communities that reflect the different positions in the innovation for sustainability debate.

Keywords: eco-innovation; environmental innovation; green innovation; sustainable innovation; bibliometric analysis; scientific trajectories

1. Introduction

The relationship between technology, innovation, and environment is an example of a widely contested topic because technological change has been considered both the source and the solution for many environmental issues related to anthropogenic activities (Hekkert et al., 2007). The root of academic discovery in this field began in the 1970s, when several authors discussed the feasibility of endless economic growth on a finite planet (Beckerman, 1974; Cole et al., 1973; Georgescu-Roegen, 1971; Meadows et al., 1972; Solow, 1973). The well-known idea of sustainable development (SD) was a milestone in this debate. Linking economic growth to the actual state of technology gave innovation a central role -- as the way to stretch the limits of economic growth within the availability of finite resources. One consequence of the SD debate was to settle the scientific agenda. This resulted in more scholars analyzing innovation through the lens of sustainability (Freeman, 1996). The approach also finds important applications in policy contexts, as in recent reports and manuals written by regional, national, and international organizations (Dutz and Sharma, 2012; O'Hare et al.,...

When contested positions exist, terms and languages may have a powerful role because they can be used to shape meanings and identify belongings to the different communities (Nicolini, 2012). Therefore, the comparison between concepts is crucial to define and explore the intellectual structure of a given scientific field, to access the influence and scientific impact of different journals, authors and geographic locations to each concept, to suggest future paths for the development. For this reason, we were surprised to find only few bibliometric analyses (Dias Angelo et al., 2012; Karakaya et al., 2014; Schiederig et al., 2012) that addressed the language dimension of the relationship between innovation and sustainability without finding relevant differences in the usage and meanings of different terms. More specifically, Dias Angelo et al. (2012) reviewed papers – over the last three years and only in the journals tied to organizational environmental management -- which contain the terms environmental innovation, green innovation and eco-innovation in titles or abstracts indexed in the ISI Web of Science (WoS) and Scopus. They found a predominance of environmental innovation, but not any difference in meanings. Karakaya et al. (2014) studied the diffusion of eco-innovation looking at eco-innovation, ecological innovation, green innovation, sustainable innovation and environmental innovation terms in Google Scholar. While the focus of Karakaya et al. is to identify the core disciplines and research streams of literature, they did not highlight any differences between these terms. Schiederig et al. (2014) identified and analyzed four main sustainable innovation terms (eco-innovation, environmental innovation, green innovation, and sustainable innovation) and concluded that the terms “can be used largely interchangeably” (p. 182), even though “sustainable innovation includes a social dimension as well as ecological dimension” (p. 188).

Such non-confictual view seems to stand in contrast with the richness of the positions in the sustainability debate. For instance, Rennings (2000) uses the terms eco-innovation and environmental innovations as synonymous, while Ekins (2010) makes a clear distinction between them. In addition, these three bibliometric reviews seem not to define a clear methodology to identify meanings and communities, leaving room for more advanced and detailed bibliometric analyses.

We performed an alternative bibliometric analysis that explicitly aimed to (i) disentangle the meanings and (ii) identify associated scholarly communities and discussions behind these same four terms. We utilized bibliographic data from WoS and a methodology that combined keywords analyses -- as a way to track meanings -- with community detection based on shared references.

Differently from the cited reviews, our results indicate that these four terms focus on different topics and partially identify different scientific communities. For example, sustainable innovation is preferred by communities dealing with complex system-oriented approach, especially the transition school of UK and The Netherlands. Green innovation is used by the management community, and it is very popular outside Europe. Eco-innovation has an important focus on eco-design and it has important overlaps with environmental innovation especially within specific communities – as for example – those studying evolutionary economics.

We also found a correspondence between journals and communities, and – very interesting – the use of the Journal of Cleaner Production as common platform of the different communities.

In conclusion, we confirm that terms and language are important concepts to understand different positions and meanings within different scientific sub-communities. The different importance and popularity of the scientific sub-communities can influence future policies for sustainability. For example, the growing popularity of the eco-innovation term may result in policies which focus on eco-design and eco-labels, whilst the sustainable innovation perspective may focus on policies which pursue wider societal changes (Franceschini and Pansera 2015).

The paper is organized as follows: Section Two briefly introduces the concepts of Kuhnian scientific communities and the discourse analysis approach to sustainability. Section Three presents the data and methodology used for our bibliometric analysis. Section Four presents the main results and discussions, and section Five outlines our main conclusions and potential future developments for this approach.

2. The Discourse Analysis about Innovation and Sustainability in a Kuhnian World

Before Kuhn, theorists of epistemology and science understood scientists as individual agents free from any social boundaries (Jacobs, 2006). Polanyi (1951), Royce (1968), and Fleck (1979) touched upon the notion of the scientific community, but it was Kuhn’s seminal work The Structure of Scientific Revolutions (1962) that popularized this topic (Jacobs, 2002). In Kuhn’s view, a scientific community consists of scientists who agree on specific paradigms about reality. Paradigms are ways in which scientists look at the world, and each paradigm consists of specific theoretical frameworks, puzzles to be solved, methodological processes, and potential solutions. These paradigms are the “theoretical hard core” of scientists who shape research programs (Lakatos and Musgrave, 1970).

Different scientific communities seek to gain popularity and reproduce themselves as they attract new members through specific processes of education, initiation, and selection in which students have been similarly educated and are thought to use the same language (Jacobs, 2006). Consequently, paradigms evolve and compete at any time, representing the progress of scientific knowledge. Paradigms and scientific communities are found in all research topics in which different ideologies, approaches, and interests exist. The existence of different scientific communities is crucial to solve complex problems through the continuous exposition and confrontation of parallel theories (Kornfeld and Hewitt, 1981) and, therefore, the advance of scientific research is intrinsically dependent on diversity (Popper, 1963).

The use of a common language defines the existence of--and draws the boundaries between--different paradigms and scientific communities. The use of language is a specific subject of study, called discourse analysis, which has become popular to address the relationship between science, technology, and society (Hajer and Versteeg, 2005). As Nicolini argued, discourse is “first and foremost a form of action” (2012, p. 189) through which each community tries to attach meaning to topics and influence other communities. Consequently, any discourse is a way to sustain specific social groups and cultures (Gee, 2010). Therefore, discourse analysis can be applied to study the dominant ideologies and values in the scientific world.
The comparison between concepts is important to define and explore the intellectual structure of a given scientific field (Dobers et al., 2000; Hill and Carley, 1999; Ramos Rodríguez and Ruiz Navarro, 2004), to access the influence and scientific impact of different journals, authors and geographic locations to each concept (Baumgartner and Pieters, 2003; Ingwersen, 2000), and to suggest future paths for the development of the many different branches within a field. It has been used largely to define concept-based scientific communities in many fields such as Strategic and operational management (Charvet et al., 2008; Ramos Rodríguez and Ruiz Navarro, 2004; Vokurka, 1996), corporate social responsibility (De Bakker et al., 2005), logistics and transportation (Kumar and Kwon, 2004), service innovation (Sakata et al., 2013), National Innovation systems (Teixeira, 2013) and even Innovation itself (Fagerberg et al., 2012).

Under the lens of discourse analysis, nature, innovation and sustainability are socially constructed and historically dependent concepts. As any social concepts, they are widely debated within scientific communities that carry different theoretical lenses, terms, and ideological values (Castro, 2004; Franceschini and Pansera, 2015; Garud and Gehman, 2012; Hopwood et al., 2005; Markard et al., 2012; Pansera, 2012; Rennings, 2000; Scoones, 2007).

The relationship between technological change and environment has been discussed at least since the early 1970s, when the first general discussions on the environmental impacts were conducted (Ehrlich and Holdren, 1972; Meadows et al., 1972). As the research field has evolved in the last decades, the scope of the innovation literature has widened in the last decades to include not only technical innovations (Freeman and Soete, 1997) but also organizational, marketing, institutional, and normative aspects (Fagerberg and Verspagen, 2009).

Such discussion was also incorporated in early evolutionary works (Freeman, 1984) and in the so-called Berlin school of environmental policy research, which came up with the related concept of ecological modernization (Christoff, 1996), focusing on a sociological, policy-oriented perspective. With the idea of sustainable development being formulated and presented in the late 1980s (Brundtland, 1987) and specific environmental targets being defined later through the Kyoto Protocol, many scholars from different backgrounds started to incorporate its premises in order orient their research fields towards the premises of the concept.

In the beginning of the 1990s the importance of sustainable development guidelines for technological change and growth was highlighted by business (e.g. Barrett, 1991; Elkington, 1994; Gladwin et al., 1995; Porter and Linde, 1995; Repetto, 1995; Welford, 1995), economics (Jacobs, 1993; Jaffe and Peterson, 1995; Jaffe and Stavins, 1995; Jorgenson and Wilcoxen, 1990; Tietenberg, 1990), and design (Keoleian and Menerey, 1994) literatures.

With such diverse roots, the literature about the relationships between innovation and sustainability is expected to show branching terms with differentiated attached values. Likewise, we could assume to find scholars with different understandings of the four terms, in opposition to the findings of the already existing bibliometric analyses. In fact, we found cases in which the terms were used interchangeably, as synonyms, and cases in which they had contrasting meanings.

In the mid-1990s, the incipient literature on sustainable development and technological change started to use specific terms such as eco-innovation and environmental innovation to refer explicitly to the innovations aiming at reducing environmental impacts, in the attempt of operationalizing the sustainable development
premises (Carraro and Siniscalco, 1992; Fussler and James, 1996; Green et al., 1994; A. B. Jaffe and Palmer, 1997; Johansson and Magnusson, 1998; Lanjouw and Mody, 1996; Pickman, 1998). The terms green innovation and sustainable innovation could also be found at this time, although their use was restricted to very few papers (e.g. Azzone and Noci, 1998).

Lately, Rennings (2000) stood out as one of the main references for the concept of eco-innovation and environmental innovation, using both interchangeably. His definition was widely cited and influenced subsequent works, many of which also made no distinction between the two terms (for example Arundel and Kemp, 2009; Hojnik and Ruzzier, 2015; Horbach et al., 2013; Triguero et al., 2013; Oltra et al., 2008; De Marchi, 2012). In another influential project, “Measuring Eco-innovation” (MEI), Kemp and Foxon (2007) explicitly stated, citing Rennings, that “often eco-innovation is used as a shorthand for environmental innovation” (p. 2).

In fact, many authors use two or more terms to refer to the same idea or concept: Hellström (2007) used eco-innovation as a synonym for “environmentally sustainable innovation” and also for sustainable innovation. Bernauer et al. (2007) stated, “The terms eco-innovation and green innovation are used synonymously for environmental innovation” (p. 3). Andersen (2010) and Pujari (2006) used green innovation and eco-innovation synonymously, and Halila and Rundquist (2011) used all four sustainable terms to refer to the same concept. Similarly, and more recently, Hojnik and Ruzzier (2015) stated that eco-innovation, ecological innovation, green innovation, and environmental innovation are interchangeable.

On the other hand, many scholars made some distinctions between these terms. For example, Kemp and Foxon (2007), Schiederig et al. (2012), Charter and Clark (2007) agreed that an explicit social positive aspect, besides economic and environmental gains, differentiates sustainable innovation from the other terms. Charter and Clark (2007) argue that, “although the two terms are often used interchangeably, eco-innovation only addresses environmental and economic dimensions while sustainable innovation embraces these as well as the broader social and ethical dimensions” (p.10).

Noteworthy, Ekins (2010) defined environmental innovation as “changes that benefit the environment in some way,” while eco-innovation is “a sub-class of innovation, the intersection between economic and environmental innovation” (p. 269). In other words, for him eco-innovation is related to both environmental and economic benefits, and environmental innovation is related only to the former. Therefore, the author made a clear, conceptual distinction between the two terms, contrasting with Rennings (2000) and subsequent works.

These examples demonstrate how complex it is to delineate these terms according to their existing, explicit definitions. This motivated us to define a methodology which allows to consistently identify the existence of different meanings and different communities.

3. Methodology

Our methodology is designed to disentangle the meanings and communities related to the different sustainable terms, as a way to understand the complexity involved in their use by scholars. We reviewed four sustainable terms eco-innovation, environmental innovation, green innovation, and sustainable innovation widely used in
the literature and applied a combination of content analysis techniques—which draw meanings from the
manifest content of language and communication (Baregheh et al., 2009) and community detection in
networks (Blondel et al., 2008). We narrowed the analysis to peer-reviewed, English-written journal articles,
gathered through WoS.

WoS data is considered the central source of information for extensive bibliometric exploration within the
social sciences (Liu et al., 2014). In fact, only the WoS data has the high level of curation, essential to our
analysis. To the best of our knowledge, WoS is the only bibliographic database that normalizes the cited
references for each article record across the whole collection. This feature allowed us to calculate pairwise,
bibliographic coupling and perform the community-level detection as explained in phase three of the analysis.

We extracted the full records for the analyzed articles, including cited references. The keywords at the center
of our analysis were the original, author-provided keywords, which exposed a high level of linguistic variation.
To prepare these terms for quantitative analysis, we applied a combination of manual consolidation and
algorithmic stemming, explained below in more detail. While a certain level of linguistic normalization is
essential to achieve comparability, we cannot completely exclude the possibility that changes in meanings
were introduced in the course of data preparation.

The restriction to leading peer reviewed journals results in smaller samples which can be regarded as
representative for the respective research areas (van Leeuwen 2006). The use of WoS limited the number of
analyzed articles, as the number of publication records is significantly larger in other bibliographic databases,
such as Scopus even using Google Scholar (GS) as for instance in (Schiederig et al., 2012) where several
thousand publications constitute the basis of the analysis. While GS is an excellent choice for literature
discovery, it contains all kinds of publications including working papers, conference papers, and even student
assignments and forged documents (Bornmann et al., 2008; Delgado López-Cózar et al., 2014; Giustini and
Boulos, 2013; Lasda Bergman, 2012). According to Kousha and Thelwall (2007), “it is likely that a significant
mass of non-refereed web documents which do not pass any ‘qualitative’ process are indexed by Google
Scholar, although some may be postprints or preprints of subsequently accepted refereed articles.” (p. 290).

The over-time development of publications within the different areas (Figure 1) shows differences but no
alarming signs of systematic bias of particular publication groups over time. Another indication for the validity
of the sample is the relative number of search results with a similar query but using a different database. The
results of this cross-check using the Scopus database resemble for the most the patterns found in the WoS
data. Eco-innovation and environmental innovation are similar in size and the two “larger groups” while there
are less hits for “sustainable innovation” and “green innovation”. In contrast to the WoS data, Scopus contains

---

2 Cross-check query on Scopus for each of the 4 sustainable terms, excluding hits for the three others. Results restricted
to journal articles from late 2014 in the subject matters "Business", "Engineering", "Energy", "Social Science”,
"Environmental Science", and "Economics”. Eco-innovation (169), environmental innovation (223), green innovation
(99), sustainable innovation (147)
more records for “sustainable innovation” than “green innovation”, which might indicate that our sample contains relatively little literature on the former term.

With this methodology, we are able to detect i) different meanings carried by the four sustainable terms; and ii) different scientific communities behind these terms. Meanings were detected by looking at co-occurrence patterns of keywords. More specifically, we analyzed the co-occurrence between each of the four sustainable terms when used as article keywords and other recurrent keywords. This technique was based on the idea that if a sustainable term is highly connected to specific keywords, these associations may be meaningful. In other words, if these sustainable terms are fully interchangeable, we would not expect to find any specific pattern of correlations because their use would be random. To evaluate the association of any of the keywords with each of the four sustainable terms, we used the term frequency inverse document frequency (tf.idf) statistic (Rajaraman and Ullman, 2011) which is often used as a weighting approach in information retrieval. The term frequency ($TF_{ij}$) measures the frequency (number of occurrences) $f_{ij}$ of a term (keyword) $i$ in a document $j$, normalized by the maximum number of occurrences of any term in the same document:

$$TF_{ij} = \frac{f_{ij}}{\max_k f_{kj}}$$  \hspace{1cm} (1)

If the term $i$ is the most frequent term in a document $j$, then $TF_{ij} = 1$. The inverse document frequency ($IDF_i$) measures how frequently the term $i$ occurs in a collection of documents, based on the total number of documents ($N$):

$$IDF_i = \log_2\left(\frac{N}{n_i}\right)$$  \hspace{1cm} (2)

Combining (1) and (2)--the term frequency and the inverse document frequency returns the final tf.idf equation (3):

$$tf.idf_{ij} = \frac{f_{ij}}{\max_k f_{kj}} \times \log_2\left(\frac{N}{n_i}\right)$$  \hspace{1cm} (3)

In our analysis, the “document” is comprised of keywords that appeared together with one of the four sustainable terms in the set of keywords in one article. The tf.idf counts the number of times a word occurs in a document, discounting for the overall generality of a keyword in the whole corpus. In this way, the importance of keywords (such as innovation) that are fairly general in the overall corpus is lowered, yet they are not excluded from the corpus as contextual stop words. In fact, having a keyword highly associated with all four sustainable terms did not indicate a specific association of the keyword with any of the sustainable terms. Using this relatively simple word co-occurrence and weighting approach, we were able to identify the keywords associated with each of the four terms and score them by their level of association.

Scientific communities were explored using the bibliographic information extracted during the analysis of meanings. For those articles, we focused on: i) the journal in which the paper was published, ii) the authors’ countries of origin, and iii) the cited references.

The data preparation and analysis was divided into three phases: Phase 1 included the preparation of the database of journal articles. Phase 2 analyzed the meanings of the sustainable terms looking at a) the co-occurrences between these sustainable terms used as keywords and other keywords, and b) the content of titles.
and abstracts of journals articles. Phase 3 consisted of the analysis of the scientific communities, looking at citations, authors, and journals.

**Phase 1** – We extracted a list of 473 items from Web of Science that were matched by a “topic search” for one of the following terms: eco-innovation, environmental innovation, green innovation, and sustainable innovation. From this first list, we selected the 400 items that contained keywords and citations in the WoS record, and, finally, the 196 papers that used one or more of those terms as keywords. These 196 journal articles contained 788 unique keywords that were grouped, by stemming or conceptual similarity, in 321 unique keywords for a total of 1,216 hits.

**Phase 2** – We applied the tf.idf analysis to the selected data to find patterns of correlation between the sustainable terms and the other keywords.

**Phase 3** – We investigated the community-level dimension by looking at journals, authors, and citation statistics. To construct the network, we first calculate a variation of the bibliographic coupling (BC) between each pair of papers in our corpus of 196 articles. The traditional BC indicator is calculated as

\[ w_{ij} = \frac{n_{ij}}{\sqrt{n_i \times n_j}} \]  

(4)

where the number of shared references between paper \( i \) and \( j \) is discounted by the tendency of the papers to cite. We propose to extend this measure by accounting for the general popularity of literature to be cited. The argument behind this extension is the following: a shared reference to a seminal paper that stands in the beginning of a larger academic discussion is probably a weaker indicator for communality between paper \( i \) and \( j \) as compared to a shared reference to a more specific and less cited empirical study. We use Newman’s (2001) collaboration index, which he developed to identify relationships between scholars from co-authorships. This index suggests that, for instance, the collaboration on a physics article with 10 authors is probably generating a weaker connection between the participating scientists than the joint authorship of a paper by 2 scholars. In order to include this extension, we changed the numerator from equation (4), assuming that

\[ n_{ij} = \sum_k \frac{\delta_i^k \delta_j^k}{n_k - 1} \]  

(5)

Where \( n_k \) is the number of citations that \( k \) receives and \( \delta_i^k \delta_j^k = 1 \) if papers \( i \) and \( j \) both cite \( k \). The final BC equation is, therefore:

\[ w_{ij} = \frac{\left( \sum_k \frac{\delta_i^k \delta_j^k}{n_k - 1} \right)}{\sqrt{n_i \times n_j}} \]  

(6)

---

$^3$ Extracted on the 13th August 2014.
Finally, we apply the established Louvain algorithm in order to identify communities in the network (Blondel et al., 2008).

4. Data Analysis

4.1. Unfolding meanings and the evolution of the four sustainable terms

In the first part of our analysis, we investigate the evolution of use of the four terms (as keywords) over time and make a detailed analysis of the bibliometric characteristics associated with each one of them using the $tf-idf$ as parameter. Since keywords are among the central elements of scientific papers – used to indicate their main topics – such an analysis is likely to provide insights into the changes on the use of these terms by the scientific community and their assumed meaning.

Figure 1 plots the cumulative counts of the four terms over time. $Eco$-$innovation$ and $environmental$ $innovation$ are the most used terms. $Environmental$ $innovation$ is the oldest term and its cumulative growth trend presents two clear breakpoints: 2000 and 2007. It seems to be the most established term among the four and it presented a stable growth after 2007. Despite having followed the growth of the other two “less popular” terms until 2009, the use of $eco$-$innovation$ dramatically increased after 2010 – becoming the most used since that year. The other two terms lag behind in popularity; $green$ $innovation$ was the most popular in 2013, which might suggest that it could catch up in the coming years. The use of $sustainable$ $innovation$ also increased after 2010, but it remains the least used among the selected terms.

Figure 1 Cumulative number of the four sustainable terms used as keywords over time

On a more detailed level, Table 1 shows the 10 most important keywords correlated--appearing as keywords with one of the four terms in the same paper--to each of the four sustainable terms, ranked according to their $tf.idf$ value. By associating these sustainable terms with complementary keywords, we are able to draw some preliminary differences between their use by the scientific community. For instance, scholars working with eco-design use mostly the term $eco$-$innovation$, while $environmental$ $innovation$ is used by scholars dealing
with regulatory and policy effects—worth mentioning is the presence of keyword “ecological modernis” as a reference to the “Ecological Modernization” school of policy research—and Porter-type competitive advantages derived from such innovations.

*Sustainable innovation*, on the other side, is a term used by scholars working with the more sociologic-driven approaches; these include actor network theory, user-driven innovations, and multilevel perspective. Finally, *green innovation* is related with management and competition issues, since its main correlated keywords are all related with such topics. There are similarities between the terms *eco-innovation* and *environmental innovation*, as both are correlated with keywords associated with quantitative modeling such as “triz”, “indic”, and “innovation survey”, and between *sustainable innovation* and *green innovation*, given that both present high co-occurrence with keywords related to management issues.

**Table 1 The 10-most important correlated keywords for each sustainable term.** The keywords are ranked according to tf.idf value.

<table>
<thead>
<tr>
<th>Environmental innovation</th>
<th>Sustainable innovation</th>
<th>Eco-innovation</th>
<th>Green innovation</th>
</tr>
</thead>
<tbody>
<tr>
<td>porter hypothesis</td>
<td>ant</td>
<td>ecodesign</td>
<td>competitive advantag</td>
</tr>
<tr>
<td>environmental regul</td>
<td>user-driven innov</td>
<td>triz</td>
<td>environmental manag</td>
</tr>
<tr>
<td>ecological modernis</td>
<td>partnership build</td>
<td>sustainab</td>
<td>corporate environmental manag</td>
</tr>
<tr>
<td>Sustainab</td>
<td>sustainable business model</td>
<td>indic</td>
<td>green supply chain manag</td>
</tr>
<tr>
<td>innovation survey</td>
<td>multilevel perspect</td>
<td>environmental polici</td>
<td>sustainable develop</td>
</tr>
</tbody>
</table>

In the next step, we calculated the association between the sustainable terms and journals. Among the 196 papers, we found 92 scientific journals that contained at least one article with one of the four terms as a keyword. Table 2 shows the three most popular journals for each sustainable term. *Journal of Cleaner Production* (JCP) ranks as the most important for all the sustainable terms, and it is the only one to be present—among the first three—in each of them, reinforcing its claimed transdisciplinary nature. *Eco-innovation* is a term appearing in a relatively higher number of journals (42), which may indicate that its increased popularity after 2010 was the result of its use by different communities. As for the keywords shown in Table 1, the journals associated with *sustainable innovation* and *green innovation* reinforce the hypothesis that these terms are mainly related with business and management issues when compared with the other two terms.

**Table 2 Most important journals.** Percentage was calculated as the number of occurrences of a journal on the number of articles in the sustainable term group.
We found 406 unique authors in our database. Table 3 shows the three most present authors of the four sustainable terms.

Table 3 The three most present authors. Numbers of publications for each author.

<table>
<thead>
<tr>
<th>Environmental innovation (119 authors)</th>
<th>Sustainable innovation (69 authors)</th>
<th>Eco-innovation (140 authors)</th>
<th>Green innovation (92 authors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rennings, K (6)</td>
<td>Partidario, PJ (2); Smith, A; Quist, J; Boons, F; Tukker, A; Evans, S; Lambert, J</td>
<td>Peiro-Signes, A (6)</td>
<td>Chen, YS (6)</td>
</tr>
<tr>
<td>Mazzanti, M (4)</td>
<td></td>
<td></td>
<td>Chang, CH (4)</td>
</tr>
<tr>
<td>Oltra, V (3)</td>
<td></td>
<td></td>
<td>Qi, CY (2); Tseng, ML: Zeng, SX</td>
</tr>
</tbody>
</table>

50 authors have more than one publication using at least one of the four sustainable terms as keywords. 36 of them always use the same keyword for all the publications, while 14 have used two different ones (no one has used three of four different keywords). Table 4 shows the number of authors by the use of the different sustainable terms as keywords.

Table 4 Number of authors for keywords. Numbers of authors using the different sustainable terms as keywords. Percentage shows the quota of authors - for each keyword - using only a keyword

<table>
<thead>
<tr>
<th></th>
<th>Eco</th>
<th>Env</th>
<th>Green</th>
<th>Sus</th>
<th>tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eco</td>
<td>16</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td>Env</td>
<td>10</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Green</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Sus</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Unique</td>
<td>53%</td>
<td>44%</td>
<td>71%</td>
<td>78%</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4 shows that about half of the authors that use eco-innovation or environmental innovation as keywords, they also use other sustainable terms as keywords. We found that the combination eco-innovation and environmental innovation is by far the most relevant, being used by 10 out of 14 authors. Thus, we found that eco-innovation is used by all the 14 authors using two sustainable terms as keywords, and that there are no combinations between two of the other three sustainable terms.
Table 5 shows the distribution of the sustainable terms according to main authors’ affiliation country. Overall, Germany is the country with most scholars, particularly addressing *environmental innovation*, which is not surprising given the tradition by German scholars to study the topics related with environmental policy and regulation; this includes the so-called Berlin School of environmental policy research that is linked to the term ecological modernization (Table 1). *Sustainable innovation* is used by scholars coming from English-speaking countries as well as The Netherlands, corroborating the results from Table 1, since the latter hosts many well-known academics working with the multilevel perspective and within technological transitions tradition. Again, this analysis shows some similarities between *environmental innovation* and *eco-innovation*, e.g. being more Europe-centered. In comparison, *green innovation* is a term used more often outside Europe, although the number of countries in which scholars refer to this term is overall low.

**Table 5 Most important countries.** Percentages were calculated as the number of occurrences of a country on the total number of papers using the sustainable term.

<table>
<thead>
<tr>
<th>Sustainable term</th>
<th>TF.IDF</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental innovation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sustainable innovation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eco-innovation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Green innovation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Finally, Table 6 lists the five most important references for the papers containing one or more sustainable terms as keywords, ranked by their *td.idf* value. References and citations are traditionally referred by the literature as indicators of interconnection between authors (Bornmann et al., 2008; Moed, 2005; Narin, 1976). Thus, looking at the central references in the four groups might indicate the association of the different sustainable terms with particular strands of literature.

In the case of *environmental innovation*, the most connected reference is the seminal paper by Porter and van der Linde (1995). This corroborates the results of Table 1, as it is the origin of the so-called Porter hypothesis. The other references are related to determinants of product and process environmental innovations. *Sustainable innovation* presents references that can mainly be associated with transition theories and systemic thinking, therefore also confirming the results of the co-word based analysis. Also here, *eco-innovation* shows similarities to *environmental innovation*, especially through shared referencing of works by Rennings and colleagues. In both cases, references point to determinants of eco-/environmental activities, especially in terms of structural and policy characteristics, and are therefore associated with ecological economics literature. Lastly, the term *green innovation* has, among its main references, papers linked to resource-based view, firm’s competences, and competitive advantages. Also this is in line with the results shown in Table 1.

**Table 6 Central references for each of the four sustainable terms.** The references are ranked according to their *tf.idf*.
Environmental innovation


Sustainable innovation


Eco-innovation


Green innovation


4.2. Sustainable terms at community-level: the cluster analysis results

The second part of the data analysis focuses on cluster identification in the citation network and analysis of detected communities (See Section 3). Starting from the 400 items with keywords and references, we obtained 10 major clusters—containing 367 items—with more than two papers. The network was constructed using the
bibliographic coupling between each pair of papers in our corpus as explained in the methodological section. The network was then clustered into communities of articles that show strong similarities in terms of shared citation patterns. Clusters with a high number of papers with one or more of the four sustainable terms are assumed to have a thematic association with the respective research field. The results are presented in the Table.

**Table 7** Communities related to each of the sustainable terms. The numbers highlighted indicate that the cluster has a high number of papers using that term as keyword.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Eco-innovation</th>
<th>Environmental innovation</th>
<th>Green innovation</th>
<th>Sustainable innovation</th>
<th>None of the terms among the keywords</th>
<th>Total terms</th>
<th>Percentage Sustainable terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>8</td>
<td>8</td>
<td>27</td>
<td>3</td>
<td>33</td>
<td>79</td>
<td>58%</td>
</tr>
<tr>
<td>B</td>
<td>7</td>
<td>31</td>
<td>2</td>
<td>-</td>
<td>33</td>
<td>73</td>
<td>55%</td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>8</td>
<td>3</td>
<td>16</td>
<td>33</td>
<td>65</td>
<td>49%</td>
</tr>
<tr>
<td>D</td>
<td>16</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>19</td>
<td>37</td>
<td>49%</td>
</tr>
<tr>
<td>E</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>17</td>
<td>25</td>
<td>32%</td>
</tr>
<tr>
<td>F</td>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>13</td>
<td>24</td>
<td>46%</td>
</tr>
<tr>
<td>G</td>
<td>5</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>23</td>
<td>65%</td>
</tr>
<tr>
<td>H</td>
<td>6</td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>8</td>
<td>20</td>
<td>60%</td>
</tr>
<tr>
<td>I</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>11</td>
<td>13</td>
<td>15%</td>
</tr>
<tr>
<td>J</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>50%</td>
</tr>
</tbody>
</table>

At first, *eco-innovation* is dominant in cluster D and F, *environmental innovation* in cluster B, *green innovation* in cluster A, and *sustainable innovation* in cluster C. Moreover, clusters G and H combine *eco-innovation* and *environmental innovation*, indicating that, for some papers sharing similar characteristics, these two terms are being used by the same communities. We have excluded clusters E and I from the discussion because of the low presence of papers with at least one of the four sustainable terms as keywords (respectively 32% and 15%) and cluster J because of its limited size.

For the eight selected clusters, Appendix 1 presents the most important keywords and references, and the following discussion will be based on these results. The *eco-innovation* term has been used in papers addressing issues related with the design of more environmental friendly technologies/products and the evolutionary dimension of environmental friendly innovation. We noticed that the scientific community, which focuses on eco-design, has widely used the term—as shown through both keywords and references from Cluster D—in which eco-design and sustainability are connected to the efficiency dimension of sustainability. The technical perspective of eco-innovation is also confirmed by the technical focus of two out of the three most relevant journals for that term—DYNA and Environmental Engineering and Management Journal; both journals focus on in the area of engineering, technology, and sustainability.

The regulation dimension related to the term is represented by the keywords and references from Cluster F and it is confirmed by the relative importance of authors such as Beise, Rennings, and Nelson. Cluster F includes also many works which theorize the evolutionary approach to innovation and how this approach may contribute to the diffusion of environmental technologies.
The term *environmental innovation* seems to be a mid to strong European placed term which focuses on the Porter’s hypothesis about the impact of environmental policies on competition of different sectors and industries and the determinants of innovation at the industry level, as shown in Cluster B. Clusters G and H demonstrate that, for some scientific communities, *eco-innovation* and *environmental innovation* have been used interchangeably. Cluster G relates the environmental dimension to the evolutionary economics theory, as represented by the central references to some of the most prominent scholars in this approach, including Nelson and Malerba. Cluster H focuses on the ecological modernization and industrial ecology/symbiosis.

The *green innovation* term represents the clearly delineated non-European, management-focused approach for innovation and sustainability, as seen through the strong affiliation with Cluster A. Its focus on the corporate dimension of sustainability is confirmed by the importance of the journals as well as main keywords used; almost all keywords related with some aspect of management and competitive conditions of firms. Lastly, the *sustainable innovation* term has a strong connection with the technological innovation system perspective and the transition approach. The works of Kemp, Hekkert, Bergek, and Geels are central within Cluster C.

### 4.3. Discussion

Both the analyses of keywords and communities generated compatible results, which allow us to draw some consistent remarks about the meanings and the use of the four sustainable terms by scientific communities. These remarks are summarized below.

- *Eco-innovation* and *environmental innovation* have been used interchangeably by some communities (Clusters G and H), especially those related with evolutionary economics, ecological modernization and industrial ecology/symbiosis. The interchangeability of these terms is also confirmed by the important presence of several authors using these two keywords (table 4). The case of Rennings is a remarkable example because he mainly uses *environmental innovation* as keyword, but his works are central references for the *eco-innovation* cluster F. However, the popularization of these terms occurred at different points in time, as eco-innovation became widely used only after 2010.

- Scholars dealing with eco-design strongly prefer to use the term *eco-innovation*, as indicated by the exclusivity of the community based within Cluster D. *Environmental innovation* is more strongly associated with regulatory aspects as well as scholars addressing the effects and determinants of such innovative activities (Cluster B). Both terms seem to be used mostly by European scholars.

- *Sustainable innovation* is a system-oriented term, especially related with scholars associated with the transition school (primarily emanating from The Netherlands and The UK) and complex systems. As these approaches carry a stronger sociological component, our analysis confirms the conclusions of Schiederig et al. (2012) regarding the difference between this term and the others.

- *Green innovation* is strongly related to management and competition objectives, as shown by the term’s strong association with Cluster A. It is also a term used mostly by scholars outside Europe.

- All the different communities share Journal of Cleaning Production (JCP) as the most central journal. Although the analyses show different meanings and communities, we identified such journals as the platform through which knowledge between different scientific communities is shared.
Finally, we can answer to our main research question: *Do the four sustainable terms carry different meanings?* We found some similarities—especially among *eco-innovation* and *environmental innovation* and in the use of JCP. However, such four terms carry different meanings and identify different scientific communities from different traditions, well representing the complexity and the differences in the debate about innovation for sustainable development. Based on these conclusions, we suggest avoiding considering such terms as synonymous, without first considering the context in which they are used.

5. **Conclusion**

We reviewed the peer-reviewed literature about the relationship between innovation and sustainability, looking at the different meanings of four sustainable terms: *eco-innovation, environmental innovation, green innovation,* and *sustainable innovation*. Based on our findings, we can conclude that these sustainable terms focus on different topics and are affiliated with different communities. However, we found that there are some similarities between the terms and the communities, especially in regard to the terms *eco-innovation* and *environmental innovation*. All publications also share a common publication--the JCP--which seems to act as a “hub” for these different communities.

The Kuhnian perspective is confirmed as a valid key to analyze the evolution of knowledge within the scientific community. Innovation for sustainability can be framed as a complex/contested notion in which different scientific sub-communities highlight different visions and interests. The birth of different terminologies can be explained by the richness of debate among scholars. New and old terms are continuously shaped, abandoned, and re-used to highlight continuity and discontinuity with other meanings and with previous branches of research.

The scientific popularity of the different terms may be expected to influence the development of policies for sustainable development. While some terms focus on eco-efficiency, eco-design and other specific eco-performances of any innovation, others may lead to wider societal policies which target the demand side included – for example - users’ values and ideologies. For this reason, we find the study of the evolution of terminology and meanings among the scientific community a relevant dimension to understand the overall societal debate about sustainability and the role of innovation.

The boundaries of our analysis offer opportunities that can be targeted by further research. First, we focused on the four sustainable terms used by Schiederig et al. (2012), but during our data analysis, we spotted other terms that may have specific meanings (and communities), such as *eco-efficient innovation, low-carbon innovation, innovation for sustainability, socio-ecological innovation,* and *externality reducing innovation,* among many others. These terms may provide additional knowledge about the evolution of the academic literature and of scientific communities.

Second, since we narrowed the analysis to the scientific peer-reviewed literature; we are not able to explain the societal roots of these terms beyond the scientific communities. The Kuhnian perspective emphasizes the connection between scientists and overall societal dynamics. Yet, our methodology requires standardized keywords and references which cannot be guaranteed if we considered *grey literature* (e.g. industrial magazines, news, and reports from private and public organizations. However, recent developments in natural language processing, such as entity extraction techniques, might allow us to draw on broader collections of
literature. Also more efficient normalization of references is gradually allowing for utilization by other larger academic publication databases (e.g. Scopus and Google Scholar).

Given these limitations and opportunities, future research can focus on understanding other remaining questions such as whether these terms and concepts originate within or outside the scientific community or such as the coevolution of these terms between the scientific community and other societal communities.
References


Giustini, D., Kamel Boulos, M.N., 2013. Google Scholar is not enough to be used alone for systematic reviews. Online J. Public Health Inform. 5. doi:10.5210/ojphi.v5i2.4623


Kousha, K., & Thelwall, M., 2007. Sources of Google Scholar citations outside the Science Citation Index: A comparison between four science disciplines. Scientometrics, 74(2), 273–294. doi:10.1007/s11192-008-0217-x


## 1 Appendix 1 – Cluster analysis’ results

### Cluster A (“Green Innovation”)

<table>
<thead>
<tr>
<th>#</th>
<th>Keywords</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>new product develop</td>
<td>Fornell, C., &amp; Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. <em>Journal of marketing research</em>, 382-388.</td>
</tr>
</tbody>
</table>

### Cluster B (“Environmental innovation”)

<table>
<thead>
<tr>
<th>#</th>
<th>Keywords</th>
<th>References</th>
</tr>
</thead>
</table>
### Cluster C ("Sustainable innovation")

<table>
<thead>
<tr>
<th>#</th>
<th>Keywords</th>
<th>References</th>
</tr>
</thead>
</table>

### Cluster D ("Eco-innovation 1")

<table>
<thead>
<tr>
<th>#</th>
<th>Keywords</th>
<th>References</th>
</tr>
</thead>
</table>

---
greenhouse gases
disassembly process planning.


extended producer

etied producer
<table>
<thead>
<tr>
<th>#</th>
<th>Keywords</th>
<th>References</th>
</tr>
</thead>
</table>

**Cluster H ("Eco-/environmental innovation 2")**

<table>
<thead>
<tr>
<th>#</th>
<th>Keywords</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
</tr>
</tbody>
</table>