Journal of Cleaner Production 306 (2021) 127054

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

Sustainability startups and where to find them: Investigating the share of sustainability startups across entrepreneurial ecosystems and the causal drivers of differences

Sarah Tiba , Frank J. van Rijnsoever *, Marko P. Hekkert

Innovation Studies, Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, the Netherlands

ARTICLE INFO

Article history:
Received 21 April 2020
Received in revised form
2 April 2021
Accepted 5 April 2021
Available online 16 April 2021

Handling editor: Dr. Govindan Kannan

Keywords:
Entrepreneurial ecosystems
Sustainable entrepreneurship
Sustainable development
Topic model
Startup

ABSTRACT

While it is well known where the best entrepreneurial ecosystems can be found in general, such information is missing in the case of sustainability entrepreneurial ecosystems. These are entrepreneurial ecosystems with high shares of startups that are active in industries which address the Sustainable Development Goals through innovation. Not knowing what the best sustainability entrepreneurial ecosystems are, inhibits sustainability startups to locate to regions with the most favourable supportive conditions. This paper addresses the above-mentioned gap by establishing a ranking of entrepreneurial ecosystems according to their share of sustainability enterprises. For this purpose, we have analysed the websites of 19,997 startups in the 28 largest entrepreneurial ecosystems according to the Startup Genome project using Latent Dirichlet Allocation. We find that the highest proportion of sustainability startups among these entrepreneurial ecosystems is located in Boston, followed by Houston, Seattle and Lagos, respectively. A qualitative comparative analysis of the causal patterns underlying our results reveals that high GDP in combination with either (1) high shares of female founders of startups or (2) high shares of non-religious people in the population induce entrepreneurial ecosystems with relatively high levels of sustainability enterprises. Our work aims to stimulate further scholarly interest in sustainability-oriented entrepreneurial ecosystems. Policy makers can use the causal drivers to increase the share of sustainability startups in their region.

1. Introduction

Today's sustainability challenges call for the help of radical innovators like those startup founders who are finding creative solutions to pollution, the unsustainable use of resources and the spread of diseases. In fields ranging from clean energy to health treatments, innovative startup firms often translate scientific findings into actionable solutions that can reach a global audience. The new business models established by startup firms are simultaneously breaking open existing arrangements, contributing to societal change (Geels et al., 2017) and helping to solve today's pressing societal and environmental challenges (The United Nations, 2017).

On a global scale, the industries in which startups operate and the degree to which they are able to develop are determined to a large extent by the conditions of the regions in which entrepreneurs and startups operate' (Acs et al., 2017; Alvedalen and Boschma, 2017). These regions and their conditions are called entrepreneurial ecosystems (EEs). The most celebrated EE is Silicon Valley, which is well-known for its high founding rates and many profit-generating startups (Berger and Brem, 2016; Brock, 2012; Startup Genome Project, 2017). Other famous examples include Tel Aviv, Boston, London and Berlin. Various lists have been published that rank the cities or, in some cases, the nations that are most favourable to startups in general, taking into account factors such as founding rates and average funding amounts (Ács et al., 2014; Global Entrepreneurship Monitor, 2017; Startup Genome Project, 2017). None of these rankings, however, evaluate the extent to which such EEs are favourable to 'sustainability' startups, those startups that can contribute innovative solutions to societal and environmental challenges such as those identified by the United Nations' 17 Sustainable Development Goals (SDGs).

The aim of this study is thus to assess the way in which levels of sustainability entrepreneurship differ between EEs and to identify

* Corresponding author.
E-mail address: f.j.vanrijnsoever@uu.nl (F.J. van Rijnsoever).

https://doi.org/10.1016/j.jclepro.2021.127054
0959-6526/

those factors that foster the development of sustainability entrepreneurial ecosystems. These are entrepreneurial ecosystems with high shares of startups that are active in industries which address the Sustainable Development Goals through innovation. Identifying such fostering factors may be particularly useful for policy-makers who wish to increase the proportion of sustainability startups in their EE, perhaps with the hope of creating a 'Silicon Valley of sustainability startups'. Entrepreneurial activity, of course, also takes place outside of such EEs. However, we focus in our study on locations with high concentrations of entrepreneurial activity in order to cover the largest portion of the startup community possible.

To achieve the outlined aims, we established a ranking of 28 prominent and well-known EEs around the world according to the share of their enterprises operating in industries that contribute to sustainability through innovation (as opposed to through primarily operations). The ecosystem ranking we have produced differs from other rankings of entrepreneurial ecosystems in that it focuses solely on each ecosystem's share of sustainability enterprises rather than those factors such as average funding, number of startups and growth rate that such rankings are usually based on. We established our list by analysing the websites of 19,997 startups in the 28 EEs (Startup Genome Project, 2017) using the text mining method of latent Dirichlet allocation (LDA) (Blei et al., 2003). Startups that place emphasis on any topic identified by the SDGs were considered to belong to one of the core industries that contribute to sustainable development through innovation and were categorised accordingly. The SDGs, which were collectively developed by 193 countries, numerous NGOs, businesses, scientists, local authorities, women, youth, and indigenous peoples (Nilsson et al., 2016) and adopted in 2015 by the United Nations, are arguably the leading international framework for addressing global challenges related to sustainability. In line with Horne et al. (2020), we determined the sustainability score of each EE according to the absolute number of sustainability startups and share of total startups in industries contributing through innovation to sustainability. Next, we conducted a fuzzy set qualitative comparative analysis (fsQCA) to identify causal patterns that might explain the emergence of relatively sustainability-focused EEs. The factors that were evaluated for their potential causal role on EEs' shares of sustainability startups were derived from a review of the relevant literature. Our approach to this process is further detailed in the Methods section below.

The remainder of this paper is organized as follows: we first introduce the key concepts used in our study, and then outline the methods we employed; we then discuss the share of sustainability startups dedicated to particular topics and contributing to specific SDGs; next, we present the shares of sustainability startups corresponding to the various EEs and show our ranking; finally, we then outline the results of the fsQCA and conclude with a presentation of policy recommendations based on these results.

2. Background

In this section we briefly introduce the main concepts used in this paper, namely sustainability entrepreneurship and entrepreneurial ecosystem (EE), beginning with the former.

2.1. Sustainability startups

The concept of sustainability entrepreneurship employed in this paper is based on a broad body of literature containing a wide range of similar terms and definitions; 'social entrepreneurship', 'ecological sustainability entrepreneurship' and 'responsible entrepreneurship' are just a few of the terms that are often used to describe different but related phenomena (Johnson and Schaltegger, 2020; Piwowar-Sulej et al., 2021; Terán-Yépez et al., 2020; Tiba et al., 2018). Definitions of '*social entrepreneurship*' tend to vary across the academic literature. Some definitions only identify non-profit enterprises as practitioners of social entrepreneurship, while others allow both non-profit and for-profit enterprises focusing primarily on social issues or environmental challenges to be counted as social enterprises (though the latter type of organisation is less often discussed in the social entrepreneurship literature; see e.g. Mair and Martí (2006)). Social entrepreneurship scholars are broadly unified, however, in defining social entrepreneurs as prioritising social (or environmental) value creation goals over economic ones (Bacq and Janssen, 2011; Dacin et al., 2010; Mair and Marti, 2006; Saebi et al., 2019; Terán-Yépez et al., 2020; Zahra et al., 2009). *Ecological sustainability entrepreneurship* literature, on the other hand, moves environmental factors centre-stage. This literature, however, tends to focus more strongly on enterprises that provide solutions to environmental challenges while aspiring to earn financial profits (Dean and McMullen, 2007; Gast et al., 2017; Hansen and Schaltegger, 2013; York and Venkataraman, 2010).

Thus, depending on the substrand of entrepreneurship literature, either social or environmental issues may be moved centre-stage (Tiba et al., 2018). However, scholars are also developing more integrated frameworks (Muñoz and Cohen, 2017; Muñoz and Dimov, 2015; Walley and Taylor, 2002), and studies how to integrate the SDGs in strategic management (Sullivan et al., 2018).

The question that is central to the debate occurring in these fields is what it means to be social or responsible. The most widely accepted framework that articulates both social and environmental values is that of the Sustainable Development Goals developed by The United Nations (2017). In this paper, the term 'sustainability entrepreneurship' will thus be used when referring to enterprises that address one or more of these goals. The UN asserts that companies in all industries have a responsibility to contribute to achieving the SDGs through practices such as the use of renewable energy and creation of fair working conditions. Some key industries, however, also have a responsibility to innovate and develop novel solutions where there currently are none, as in the field of health treatments. As startups are known for their ability to develop such innovative solutions, in this study we focus on startups in such sectors. In doing so, we choose to build upon the perspectives represented by all three streams in the literature discussed above and adopt a relatively broad focus; we thus consider firms to be sustainability startups not only if they prioritise social value creation over economic value creation, but also if they create social or environmental value irrespective of economic value creation goals.

Startups the risk of promoting one SDG or one aspect of sustainability at the expense of others (Muñoz et al., 2018; Nerini et al., 2019). Hence, some argue that a holistic perspective is needed to assess sustainable entrepreneurship, rather than a SDG (Muñoz et al., 2018; Muñoz and Cohen, 2017). For the present study, such a holistic approach is problematic. Start-ups are emerging companies that and are surrounded by uncertainty. This makes it problematic to ex-ante systematically assess the impact of these startups on all SDGs over time (Tiba et al., 2020).

2.2. Entrepreneurial ecosystems

While early studies of (sustainability) entrepreneurship frequently focused on different personal traits of startup founders (Davidsson and Wiklund, 2001; Harris et al., 2009; Zahra et al., 2009), recent years have seen an increase in investigations that treat entrepreneurial ecosystems (EEs) as important drivers of how entrepreneurship develops (Ács et al., 2014; Alvedalen and Boschma, 2017; van Rijnsoever, 2020). Following Stam, EEs may be understood as 'a comprehensive set of resources and actors, which have an important role to play in enabling entrepreneurial action' (Stam, 2014). This set of resources and actors includes but is not limited to financiers, incubators, accelerators, pools of qualified potential employees, other startups and a set of rules that govern and enable entrepreneurial and innovative action. Through their constant interaction, these factors establish and shape the EE (Pitelis, 2012). The most common unit of analysis for an EE is a city such as New York, Berlin or London, though countries such as Estonia or Sri Lanka are sometimes also used (Stam, 2015; Startup Genome Project, 2017).

Importantly, each EE is a site of intellectual cross-pollination between the ecosystem's participants, who share knowledge through both informal and formal networks such as incubators and accelerators (Bank et al., 2017; Eveleens et al., 2017; van Rijnsoever, 2020). As such, the ecosystem in which a startup is located is likely to have a strong impact on many of its characteristics, including its business model and the degree to which that business model is geared toward sustainability. Previous studies have demonstrated that the degree to which young entrepreneurs place importance on social value creation differs across EEs (Hechavarría, 2016; Hoogendoorn, 2016; Tiba et al., 2020). However, as these studies have all focused on social entrepreneurship, this paper complements previous research by extending its scope to include sustainability entrepreneurship as this concept is defined above. We hold that the sustainability-orientation of an EE may be assessed based on the share of the enterprises within it that can be identified as sustainability startups.

2.3. Factors for Sustainability

Various theories have been proposed as to what drives the emergence of entrepreneurial initiatives that address societal or environmental challenges in the context of an EE (Cohen, 2006; Neumeyer and Santos, 2018; Theodoraki et al., 2018; Volkmann et al., 2019). One such theory states that strong economic development in a region encourages such initiatives. This follows from the argument that in the presence of relative economic security (when the basic needs of livelihood are met), people have a greater capacity to care for others and for the environment (Gelissen, 2007), and in such communities a greater number of sustainability entrepreneurs and customers who emphasise social and environmental sustainability in their decisions is likely to exist. Furthermore, higher education levels in more developed economies may be credited with bringing about greater social and environmental awareness. This suggests that focus on sustainability can only arise when a society and its entrepreneurs can afford it, a notion that further ties the issues of social and environmental sustainability to that of economic sustainability.

The flipside of this argument is presented by proponents of the failure thesis, which posits that in the absence of government sustainability initiatives entrepreneurs are more likely to start sustainability enterprises (Hoogendoorn, 2016). When they fail to offer particular social goods or services, governments leave market gaps into which sustainability entrepreneurs enter in order to address the unmet needs of the population (Matsunaga et al., 2010). When governments place sufficient emphasis on the provision of public goods themselves, sustainability entrepreneurship is less necessary.

Turning to the personal characteristics of the entrepreneur, it has been shown that female founders on average place greater emphasis than men on social value creation as opposed to mere profit-making (Bosma et al., 2015; Hechavarría et al., 2017). This tendency is consistent with the strong social expectations placed on women to be altruistic and selfless, which is also reflected in women's greater likelihood to serve as volunteers (Hechavarria and Ingram, 2016; McAdam, 2013). While men outnumber women in the fields of both social entrepreneurship and commercial entrepreneurship, the difference is smaller in the case of the former (Hechavarria and Ingram, 2016). As a result, we would expect the share of enterprises in industries contributing to sustainability to be higher in ecosystems where the share of female founders is higher.

A second personal attribute that has been suggested to impact an individual's disposition towards social and environmental consciousness is religious belief. Previous studies have revealed that an individual's religiosity is positively associated with their tendency to engage in ethical behaviour in business contexts (Longenecker et al., 2004; Weaver and Agle, 2002). The underlying mechanism in this relationship is the way individuals bring values of selflessness and social responsibility embedded in religious beliefs into their work life; as a result, the, decisions they make are based not only on economic factors but also on considerations corresponding with their religious beliefs. We may thus expect that in places where religious belief is more common, a higher share of entrepreneurs will be religious and bring their respective religious values to their work, which will result in a larger share of and sustainability startups.

The extent to which these attributes are found to be present in the ecosystems we consider and can explain variations in the share of sustainability startups of different EEs is discussed in the part of the Findings and Discussion section titled 'Factors for Sustainability'.

3. Methods

3.1. Data

In order to assess the sustainability orientation of startups in different EEs we chose to consider the startups' websites as the primary data source. Websites serve to communicate information about a company's products, initiatives, values and missions to customers, potential investors and potential employees. If a startup places an emphasis on social or environmental sustainability, this will most likely be reflected in their website.

Scholars such as Lyon and Montgomery (2015), however, warn against the possibility that companies engage in greenwashing, a practice that involves representing one's environmental and social practices in media such as the company website in an overly positive light. This practice can cause individuals to form overly positive views of such companies. Nevertheless, we view websites as a valuable resource offering insight into the products and practices of startups, but remain conscious of the necessity to be wary of firms' potentially overstating the actual social and environmental impact of their initiatives. In addition, given that our level of analysis is the entire ecosystem rather than the individual firm, we have no reason to suspect that the share of enterprises that engage in greenwashing (if there are any that do so) differs significantly between ecosystems or that greenwashing practices differ significantly across sectors.

3.2. Data collection

Since we focus our analysis on the prevalence of sustainability startups in different EEs, the initial task for the collection of data was to select the ecosystems on which the study should focus. We chose to use initially the top 45 global entrepreneurial ecosystems identified by the Genome Startup Ecosystem Report (2017), as this report is the most authoritative in the field and covers all major entrepreneurial ecosystems. The ecosystems that were selected and the corresponding data for the number and proportion of sustainability startups is presented below in the Findings and Discussion section.

Following the selection of ecosystems, we identified all the startups registered in the CrunchBase (2016) database that were founded in one of the selected ecosystems during the six-year period between January 2012 and December 2017; doing so allowed us to limit our dataset to only startups, as has been suggested by Baum et al. (2000). Crunchbase is an online database that contains information on innovative enterprises ranging from startups to Fortune 1000 companies. It is the most comprehensive startup database, and using it allowed us to collect data on a significant number of enterprises in the selected ecosystems. Data found in Crunchbase is provided by companies themselves, by venture partners and by in-house data teams. While this ensures a high quality of data overall, some pieces of information will ultimately be fraudulent or missing from the database. Furthermore, not all enterprises located in a given EE are likely to be contained in Crunchbase, and the database displays a regional bias towards the US. We partly correct for this bias later in the study by considering the share of sustainability startups in each ecosystem rather than the absolute number. Including only those entries that contained a web domain, we initially collected data on 41,434 startups.

Following the collection of the Crunchbase data, we downloaded the websites of startups with up to 250 subpages listed in Crunchbase in order to obtain their text content. As several websites were not in operation or did not contain downloadable content (an eventuality that confirmed the lack of complete data as a drawback of using the Crunchbase dataset), we ultimately obtained website data on 24,103 startups. The number of firms identified in each ecosystem is shown in Fig. 1. Ecosystems with very low numbers of firms (n < 50) were excluded from the remaining stages of the study because these numbers were not sufficiently representative to allow inferences to be made about the wider ecosystem, leaving 23,838 websites of startups located in 37 EEs.

3.3. Data cleaning

In a later stage of the study, we determined the languages of each startup website's text using the *cld2* (Ooms and Sites, 2017) package in R (R Core Team, 2017). Of the 23,838 websites, 4366 were identified as having text that was not primarily (at least 80%) in English. This finding is plausible given the international growth aspirations of startups in general. Within these 4,366, multi-language websites, websites in which the same content was available in several languages including English were manually cleaned to only contain English text. Those 1312 websites with text in German, Dutch or French, the three languages other than English that the authors of this paper can claim proficiency in, were automatically translated to English using the Translate function in Google Sheets. We chose to only select websites in these languages, as doing so allowed us to judge the quality of the translations. Websites in other languages (primarily Spanish, Portuguese, Russian, Korean and Chinese) were discarded. While this may result in an over- or underrepresentation of activity in some EEs relative to others, we assume that there are no significant differences in the proportion of sustainability-oriented companies that have or do not have an English-language website. This step was necessary in order to prepare the corpus for later analysis, as latent Dirichlet allocation (LDA) (Blei and Lafferty, 2007) can only be performed when data is in a single language.

Within the remaining 20,843 websites we then identified and removed duplicates. Texts were further cleaned by removing punctuation, numbers and certain standard English stop words, which are words that occur in a high frequency in texts without contributing significantly meaning (such as 'the', 'and' and 'some'). In the next step, the remaining words were lemmatized using the programme *TreeTagger* (Schmid, 1995) and stemmed using the *stem Document* function implemented in the *tm* (Feinerer and Hornik, 2017) package in R (R Core Team, 2017). Lemmatisation is a process that produces the basic form of a given word ('be', for instance, is the basic word form of 'is', 'are' and 'was') (Toman et al., 2006); it can thus significantly reduce the size of a vocabulary and improve the speed of algorithms applied to a dataset. Stemming, meanwhile, reduces all inflections of a word to their stems, thus leading to a further shrinking of the vocabulary represented in a dataset. We

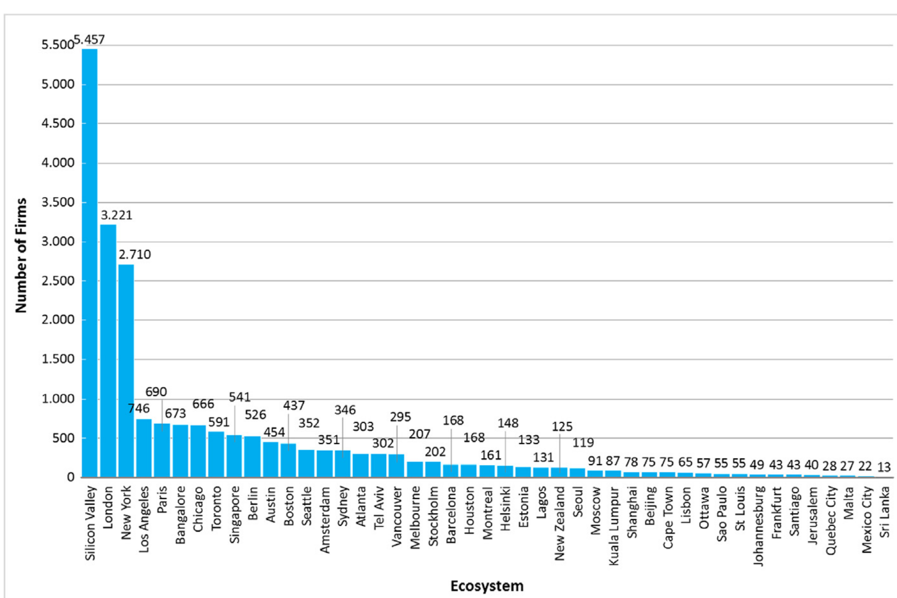

Fig. 1. Number of firms in Dataset per Ecosystem.

also removed all words that do not appear in at least five websites (Griffiths and Steyvers, 2004); doing this ensures that misspellings, addresses and enterprise names are removed from the dataset, as well as some remaining foreign words. These steps led to a further reduction of the number of startup websites to 20,626 with a vocabulary of 30,567 words.

In order to identify whether there are any additional corpus-specific stop words, or words that appear in most texts of the corpus and thus give little information about the distinct topics within them, we also generated a term frequency-inverse document frequency (tf-idf) (Salton and McGill, 1983) score for each word in each document. Variations of tf-idf are commonly used by search engines to score and then rank a document's relevance based on a user's query (Blei et al., 2003). The tf-idf score of a word indicates how important a word is to each document, with a higher score indicating that the word is frequently used in the document relative to its use in the corpus overall and a lower score indicating that it is highly frequent in the corpus overall. We found that setting the threshold for exclusion at the overall median tf-idf score (Grün and Hornik, 2017) had no impact on the vocabulary of the dataset, which indicates that previous data cleaning steps were successful at eliminating words that were less relevant to the corpus. The process by which the total number of websites contained in our initial dataset was reduced to the number finally presented in this paper, including steps described in the paragraphs that follow, is depicted in Fig. 2 below.

3.4. Latent Dirichlet Allocation

LDA is a generative probabilistic topic model proposed by Blei et al. (2003) which can be used to identify underlying topics in a large textual corpus such as the one described above (Guerreiro et al., 2016; Sugimoto et al., 2011). The model is based on the notion that texts are composed of a mixture of different topics (k), which are in turn made up of a distribution over words (w) (Blei et al., 2003). While the documents, or websites in our case, that constitute the corpus are known and observed, the topics are hidden or latent (Piepenbrink and Nurmammadov, 2015). One core advantage of LDA over other techniques is that it accurately treats

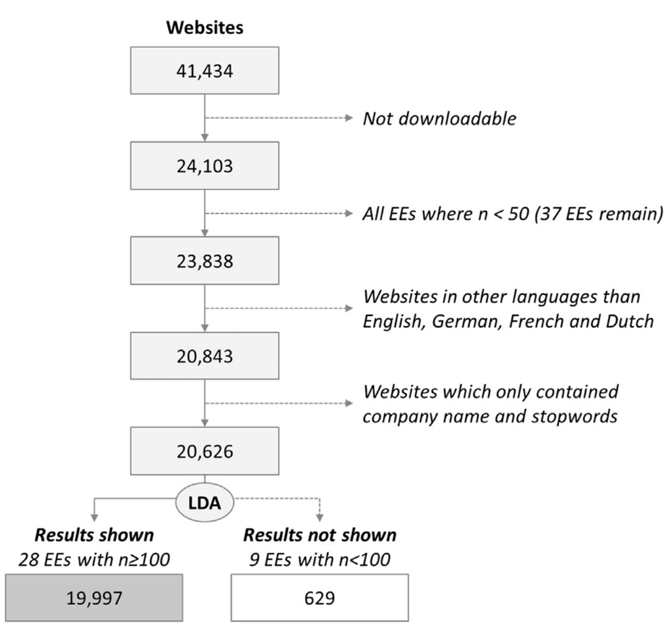

Fig. 2. Data cleaning steps and removal of websites.

texts as being constituted of a variety of topics rather than just one, which allows for relatively fine-grained modelling (Zheng et al., 2006). The modelling technique is based on the words that are observed in each text, and thus allows us to determine both the extent to which each topic is treated in a text and the words that are most likely to appear in association with each topic. This approach has been applied earlier to map sustainability startups (Horne et al., 2020; Leendertse et al., 2020).

For our study, the LDA was carried out in R with the package *topicmodels* (Grün and Hornik, 2017), which is based on the work of Blei et al. (2003), using the recommended default priors and Markov Chain settings. We used a Gibbs algorithm in which, as Griffiths and Steyvers (2004, p. 5229) explain, 'the next state [of a Markov Chain] is reached by sequentially sampling all variables from their distribution when conditioned on the current values of all other variables and the data'. In other words, the algorithm randomly assigns words to topics and then sequentially draws and reassigns each word in each document based on the assignment of all the other words in the document and the assignment of that word to topics throughout the corpus. With each draw, all other word assignments are held constant (Tiba et al., 2018).

The model is based on a number of simplifying assumptions. First, each document is treated as a 'bag of words' in which the order of words is inconsequential for the analysis (Blei et al., 2003; Grimmer and Stewart, 2013). The cleaned data for our study was thus transformed into a document-term matrix using the text mining package *tm* (Feinerer and Hornik, 2017) in R (R Core Team, 2017), which stores the corpus as a matrix of documents and frequencies of each word in each document. Second, it is assumed that the number of topics k is fixed and known. This is an input parameter of the LDA model. Choosing the correct number of topics is crucial, as it determines the granularity of the results and the fit of the model to the data, i.e. the accuracy with which the model describes the underlying data (Griffiths and Steyvers, 2004; Zheng et al., 2006). The best means for determining the correct number of topics, however, is a subject of debate among researchers. While some hold that the optimal number of topics can be determined by dividing the dataset into a training set and a hold-out set and then assessing the likelihood of the hold-out set on the trained model for different values of k (Li and McCallum, 2006), this technique has been demonstrated to not have predictive power when it comes to humans' ability to interpret the resulting topics (Chang et al., 2009).

We thus based our choice of the number of topics on our own human judgement, assessing selected numbers based on how clearly-defined the resulting individual topics were and how well each topic reflected the contents of the websites in which it was highly represented (Uys et al., 2011). Since we expected only a small subset of the topics present in the 20,626 startup websites to be relevant to our chosen focus area of sustainability, we adopted a two-step approach. First, we ran an initial LDA with the number of topics set to the relatively low value of 50, which produced less granular results than higher values would have, in order to identify overarching sustainability topics (see Table 1) and the websites within which these topics were prominent (Titov and Mcdonald, 2008). This model was ultimately chosen after three models with 25, 50 and 75 topics, respectively, were trained to assure the robustness of the model with 50 topics. The topics in the 25-topic model were found to lack detail, while the 75-topic model did not produce significant improvements compared to the 50-topic model.

When determining which topics are relevant to the SDGs and indicate that a startup operates in an industry that may address sustainability-related challenges through the development of innovative solutions, we consulted the SDGs (The United Nations, 2017) as well as additional resources provided by the UN such as

Table 1
Sustainability topics (First LDA).

Topic	Top Terms
Topic 1	bodi, use, skin, help, fit, health, sleep, exercis, healthi, weight
Topic 9	research, develop, diseas, cell, drug, cancer, clinic, studi, scienc, medic
Topic 13	woman, life, feel, person, peopl, love, want, like, help, look
Topic 19	communiti, organis, support, peopl, work, member, program, social, impact, help
Topic 21	home, child, care, famili, parent, kid, live, dog, love, help
Topic 24	patient, health, medic, care, healthcar, doctor, clinic, treatment, hospit, provid
Topic 34	energi, water, use, electr, product, power, solar, materi, industri, equip
Topic 38	student, learn, school, educ, cours, program, train, univers, teach, help

the SDG Compass (Bertazzi et al., 2015) and the SDG Industry Matrix (KPMG and United Nations Global Compact, n.d.). All topics related to areas where the potential exists for the development of novel solutions to existing problems were considered relevant. For example, the SDG Industry Matrix states that companies in the energy sector can contribute to the SDGs by developing and scaling ‘breakthrough technologies to accelerate the transition to a substantially higher share of renewable energy (solar, wind, hydro, geothermal and biomass) in the global energy mix’, a practice that clearly corresponds with Topic 34, as presented in Table 1.

In the second LDA that we implemented, a subset of 2246 startups (with a vocabulary of 10,611 words) was used in which topics related to sustainability were well-represented (suggesting that the startup is likely a sustainability startup); this allowed us to identify more granular sustainability-related topics and avoid the inclusion of unsuitable sub-topics and consequently unsuitable startups. In the results of the first LDA, for example, the topic related to energy (Topic 34) was highly represented in the websites of startups working on solar energy, but was also represented on the websites of startups in the oil and gas industry. The second LDA generated one topic corresponding to energy and another corresponding to the oil and gas industry, identified through the presence of terms such as *gas, oper, mine, oil* and *project*, and startups in which only the former was highly represented were counted as sustainability startups. Other topics that were excluded in this way were those related to home heating systems, apps, product design, and music and arts.

The number of topics for the second LDA was also set at 50 after training was conducted with several topic numbers between 30 and 70. We calculated models on the 2246 startup websites for these values using five different random seeds as suggested by Blei et al. (2003). In this process we ran 2000 initial iterations that we later discarded and 3000 iterations after, keeping only those models with the best log-likelihood. The model with 50 topics was subsequently selected, as it was the best-suited to the purposes of our analysis given that the resulting topics were easily interpretable with low word intrusion (i.e. few unsuitable words for a given topic) (Chang et al., 2009), and the topics represented the contents of the websites well.

Of these 50 topics, we identified 24 that are relevant to the SDGs of the UN (2017) (see Table 2) through an iterative process among the authors of this paper. We acknowledge that one's criteria for attributing the pursuit of particular SDGs to a given startup are likely to be ambiguous. We attempted to overcome this ambiguity by having several of the authors of this paper examine the attribution criteria and results of the other authors. The SDGs are widely accepted goals that aim to promote the betterment of society and the planet. While there are 17 goals in total, only 7 of these could be attributed to the startup firms in our dataset, as few of these firms mention topics related to the remaining 10 goals in the text of their websites. The number of topics related to the SDG corresponding to ‘Health and Well-being’, meanwhile, is quite high; this result may be explained by the wide variation and high degree of specificity of the vocabulary employed in different subfields of the medical sector.

After a reading of the websites with the highest proportions of text corresponding to each sustainability topic was conducted, a sustainability threshold was manually set at 10% (Krestel et al., 2009) such that all startups with websites having at least 10% of their text dedicated to one of the relevant topics were subsequently counted as sustainability startups. The figure of 10% represents the point at which the share of false positives for a 1% interval rose above 25% (in other words, for all websites with a sustainability topic share between 10% and 11%, less than 25% were false positives). For those startups that showed a high proportion of text dedicated to more than one of these topics, only the topic with the highest representation was counted.

The sustainability score for each EE was calculated by dividing the number of sustainability startups in that EE, using the method outlined above, by the total number of startups in the EE in our sample. As such, the ranking we have established of EEs by their sustainability scores differs from other EE rankings such as the Global Startup Ecosystem Report (Startup Genome Project, 2017) in that it takes into consideration only the share of sustainability startups in an ecosystem rather than various different indicators such as numbers of investors or local regulatory conditions.

The fact that the authors of this paper manually checked the startups' website contents also helped to ensure that the startups classified as being associated with SDG-related topics do in fact discuss these topics on their website. However, we recognise that the information provided on these websites cannot give a reliable indication of how these firms operate or the extent to which they contribute to the SDGs also through means such as using renewable energy or creating favourable working conditions. The LDA, rather, allows us to identify companies that operate in industries which have been identified as addressing the SDGs through innovation. We assume that within those industries the proportion of companies that do in fact develop innovative solutions to sustainability challenges is uniform across locations. Furthermore, though we cannot be certain that companies truly engage in the activities that they claim to on their websites, we assume that in most cases the information on the website of a startup corresponds with its activities.

We present the findings only for the 28 ecosystems which had more than 100 enterprises in our final sample. It is at this threshold that the dataset shows a clear gap and that the ratios in smaller ecosystems are likely to over-represent the relevance of a small number of sustainability startups. We use the larger dataset, however, to test the robustness of the fuzzy set qualitative comparative analysis (fsQCA) model presented in the next section.

3.5. Fuzzy set qualitative comparative analysis

Qualitative comparative analysis (QCA) is a set-theoretic

Table 2
Sustainability topics (second LDA).

Top Terms	Sustainable Development Goal
plant, water, lesson, grow, farm, soil, agricultur, produc, crop, survey	15 Life on Land
diseas, blood, diabet, heart, caus, condit, increas, chronic, stroke, medic	3 Health and Well-being
woman, pregnanc, health, babi, fertil, egg, birth, pregnant, reproduct, hormon	3 Health and Well-being
test, genet, dna, sampl, gene, result, sequenc, detect, clinic, genom	3 Health and Well-being
cannabi, medic, patient, effect, state, marijuana, product, cannabinoid, pain, research	3 Health and Well-being
fit, exercis, workout, train, bodi, muscl, athlet, strength, trainer, sport	3 Health and Well-being
clinic, develop, trial, drug, diseas, patient, studi, compani, treatment, pharmaceut	3 Health and Well-being
learn, student, school, teacher, educ, teach, game, classroom, scienc, lesson	4 Quality Education
child, parent, kid, famili, babi, year, time, love, toy, age	3 Health and Well-being
surgeri, treatment, pain, procedur, patient, medic, surgeon, surgic, clinic, scan	3 Health and Well-being
tutor, languag, english, learn, student, teach, level, lesson, onlin, teacher	4 Quality Education
communiti, member, event, support, organis, particip, peopl, group, social, share	10 Reduced Inequalities
cours, program, learn, code, train, develop, skill, student, project, design	4 Quality Education
care, home, live, senior, famili, assist, need, servic, caregiv, facil	3 Health and Well-being
exam, tooth, dental, question, prepar, smile, dentist, medic, cours, treatment	3 Health and Well-being
patient, health, care, medic, healthcar, clinic, hospit, physician, doctor, provid	3 Health and Well-being
cell, cancer, drug, develop, diseas, therapeut, target, therapi, clinic, technolog	3 Health and Well-being
food, diet, healthi, eat, bodi, weight, health, nutrit, meal, protein	3 Health and Well-being
doctor, eye, treatment, medicin, medic, test, vision, urgent, drug, onlin	3 Health and Well-being
energi, power, solar, system, electr, instal, cost, renew, save, generat	7 Affordable and Clean Energy
student, univers, school, educ, colleg, program, high, institut, degree, studi	4 Quality Education
donat, fund, campaign, organis, support, fundrais, nonprofit, chariti, donor, volunt	17 Partnerships for the Goals
sleep, therapi, feel, therapist, mental, stress, massag, life, treatment, anxieti	3 Health and Well-being
wast, sustain, market, carbon, recycl, busi, increas, emiss, reduc, environment	12 Responsible Consumption and Production

method of conducting systematic comparative analysis of medium-n cases in order to uncover causal patterns that explain the emergence of particular outcomes (Ide, 2015). As providing a detailed introduction to the method and its principles in this paper would carry us too far afield, the following paragraphs offer an elaboration of only those aspects that are relevant to the analysis at hand. Works by Ragin (2009), Schneider and Wagemann (2012) or Legewie (2013) may be consulted for an in-depth introduction to the relevant terminology.

In the present, in case, the outcome we aim to explain by employing QCA is the emergence of sustainability entrepreneurial ecosystems. In QCA one seeks to establish causal relations rather than correlations, as is commonly the goal in statistical methods with which the reader may be more familiar. The underlying idea of this method is that some phenomena cannot be explained by the additive effects of discrete, independent factors (in which case a regression analysis would be appropriate) but are rather the result of a specific kind of causal complexity. Causal complexity means that there may be multiple mutually non-exclusive explanations for an outcome, or that some conditions may only be effective if they appear with certain other conditions, and a particular outcome may occur when a given causally-related condition is present or not present.

QCA conceives of cases (in the present study, each ecosystem is a case) as either being a member or not being a member of a set, such as the set of EEs that have a high share of sustainability startups. Membership in a set, however, need not be characterised in such binary terms (0/1), such that an EE is classified as either sustainability-oriented or not sustainability-oriented. Membership in a set can rather be expressed as a matter of degree, just as EEs may be classified as more or less sustainability-oriented hence the term ‘fuzzy set’. By using fuzzy sets, fsQCA allows for more nuanced analysis than its antecedent QCA, although the 0.5 threshold which indicates whether a condition is more present than not maintains its significance (Ide, 2015). As only quantitative data were used in this analysis, values defining relative membership in each set were calibrated using natural numerical gaps found in the data considering all 28 cases to determine if a condition is rather present or absent in each case (see Tables 3 and 4 for an overview of the chosen thresholds and Tables 7 and 8 in the Appendix for both the raw and calibrated data matrix).

In the analysis fsQCA identifies whether an outcome (in this case, the emergence of a relatively sustainability-oriented EE) is a subset of one or more sets corresponding to a particular condition or a combination thereof (indicating that these are necessary conditions) or a superset of such sets (indicating that the corresponding conditions are sufficient conditions) or whether no set relations can be identified (Schneider and Wagemann, 2012), the latter indicating that no relationship between condition and outcome exists.

A first step in the analysis involves identifying a suitable outcome and potential causal conditions. The chosen outcome is described in the first two parts of the Findings and Discussion section below, while the potential causal conditions and their theoretical basis have already been described in the Background section above. The causal conditions selected for this study are grounded in a thorough review of relevant entrepreneurship literature in which they are presented as possible antecedents of sustainability entrepreneurship. While it is possible that there are other causal conditions besides the four we have selected, these conditions feature prominently in the literature we considered and present plausible causal links. As the number of conditions assessed in a QCA should be neither too high nor too low, for our core model we selected four conditions that we expected to influence the level of sustainability entrepreneurship found in an EE. Further potential conditions are introduced in the Sensitivity Analyses section below.

3.6. Sensitivity analyses

The robustness of the results presented in this paper is tested via exhaustive numeration. This is currently the most common approach to assessing the robustness of fsQCA results (Thiem et al., 2016); fsQCA results are considered robust by this method ‘if they involve similar necessary and sufficient conditions and if consistency and coverage are roughly the same across different model specifications’ (Schneider and Wagemann, 2012). We thus tested the robustness of the results for changes to the frequency threshold (row 2), to the number of cases analysed (row 3) and to the

Table 3
Operationalising sustainable EEs.

Outcome	Operationalisation	
Sustainability-oriented EE (sust)	0	Very low level of sustainability entrepreneurship as share of total entrepreneurship (≤5%)
	0.33	Low level of sustainability entrepreneurship as share of total entrepreneurship (>5% and ≤ 8%)
	0.67	Moderate level of sustainability entrepreneurship as share of total entrepreneurship (>8% and ≤ 12%)
	1	High level of sustainability entrepreneurship as share of total entrepreneurship (>12%)

Table 4
Possible causal conditions of EE sustainability.

Causal condition	Operationalisation	
GDP per capita[a] (gdpcap)	0	Low metropolitan GDP per capita (≤10,000 USD)
	0.33	Moderate metropolitan GDP per capita (10,001−30.000 USD)
	0.67	High metropolitan GDP per capita (30,001−70,000 USD)
	1	Very high metropolitan GDP per capita (>70.000 USD)
Share of female founders[a] (womenfound)	0	Low share of female founders (≤10%)
	0.33	Moderate share of female founders (>10% and ≤ 15%)
	0.67	High share of female founders (>15% and ≤ 20%)
	1	Very high share of female founders (>20%)
Religiosity of population[b] (religious)	0	Low share of believers in population (≤42%)
	0.33	Moderate share [b]f believers in population (>42% and ≤ 70%)
	0.67	High share of believers in population (>70% and ≤ 90%)
	1	Very high share of believers in population (>90%)
National social Expenditure (socexp)	0	Low social expenditure as share of national GDP (≤5%)
	0.33	Moderate social expenditure as share of national GDP (>5% and ≤ 15%)
	0.67	High social expenditure as share of national GDP (>15% and ≤ 20%)
	1	Very high social expenditure as share of national GDP (>20%)
Environmental performance[d] (epi)	0	Low value for Environmental Performance Index (≤70)
	0.33	Moderate value [d]or Environmental Performance Index (>70 and ≤ 80)
	0.67	High value for Environmental Performance Index (>80 and ≤ 86)
	1	Very high value for Environmental Performance Index (>86)
Early stage funding per startup[a] (funding)	0	Low average early stage funding per startup in ecosystem (≤200 k USD)
	0.33	Moderate average early stage funding per startup in ecosystem (>200 k and ≤ 400 k USD)
	0.67	High average early stage funding per startup in ecosystem (>400 k and ≤ 600 k USD)
	1	Very high average early stage funding per startup in ecosystem (>600 k USD)

[a] Source: Startup Genome Project (Startup Genome Project, 2017).
[b] Sources: Pew Research Centre (2017), London Census (2011), Israel Central Bureau of Statistics (2016), Statistik Berlin Brandenburg (2011), IFOP (2011), Statistics Singapore (2010), National Household Survey (2011), Australia Census (2016), Central Bureau voor de Statistiek (2016), Bangalore Census (2011), Instituto Nacional De Estadistica y Geografia (2011), Central Intelligence Agency (2017), Generalitat de Catalunya (2015), Population and Housing Census (2013), Zentralwohlfahrtstelle der Juden in Deutschland e.V. (2016), Statistics Finland (2016), Sreda (2012), Department of Statistics Stats NZ (2013), General Household Survey (2015).
[c] Sources: Organisation for Economic Co-Operation and Development (OECD) (2016), Channel News Asia (2017), World Bank (2016).
[d] '[The Environmental Performance[d]Index] provides a gauge at a national scale of how close countries are to established environmental policy goals'. ("Environmental Performance Index," 2018).

calibration of values of the outcome variable (row 4), as well as for the introduction of additional causal conditions (rows 5 and 6). The results of these sensitivity analyses are summarised in Table 5. The results reveal that in all sensitivity analyses the causal pathways were similar or identical. Furthermore, values for consistency and coverage remained roughly the same across the various model specifications; however, no model specifications have a combined consistency and coverage value as high as that of the core model. Ultimately, the results of the sensitivity analyses thus underscore the strength of the core model.

4. Findings and Discussion

4.1. Sustainability startups …

While previous comparative research on the sustainability of startups has primarily focused on the domain of social entrepreneurship (Griffiths et al., 2013; Hoogendoorn, 2016), this study adopts a broader view. Social entrepreneurs comprise a small subset of those entrepreneurs oriented first and foremost toward addressing particular social goals rather than generating profit, though definitional variance on the matter abounds (Dees, 1998). Sustainability startups, on the other hand, include also commercial

Table 5
Sensitivity Analyses fsQCA.

Row	Analysis	Solution formula	Cons.	Cov.
1	Main Analysis	gdpcap*(womenfound+~religious) → sust	0.90	0.88
2	Higher frequency cut-off (1->2)	gdpcap*~religious → sust	0.87	0.83
3	Nine additional cases included	gdpcap*(womenfound+~religious) → sust	0.84	0.81
4	Sustainability threshold = 0.1	gdpcap*(womenfound+~religious) → sust	0.90	0.76
5	Condition ecosystem funding incl.	gdpcap*(womenfound+~religious*funding) → sust	0.92	0.81
6	Condition epi included	gdpcap*womenfound*epi → sust	0.91	0.72

* = and; + = or; ~ = absence; → = sufficient for.

startups that address societal and/or environmental challenges. As our study concerns sustainability startups, it encompasses a potentially much larger and more diverse set of firms than previous studies; though such firms are oftentimes neglected in the literature, they are vital to achieving the SDGs.

Sustainability startups comprise 8.4% of the enterprises in our sample (see Fig. 3), the majority of these (56.2%) being startups that contribute to the 'Health and Well-being' goal of the SDGs. These health-oriented sustainability startups develop new pharmaceuticals and treatments or offer other health-related products, thereby working, as this SDG, to 'ensure healthy lives and promote well-being for all at all ages' (The United Nations, 2017). The high share of sustainability startups in this field may be explained to a large extent by the comparatively higher earnings potential of businesses in the medical and pharmaceutical sectors, as well as by the number of health issues in urgent need of solutions (though it must be noted that many such issues remain unaddressed because of the low earnings potential of developing solutions especially in niche markets). The goal of 'Health and Well-being' is also the SDG with the largest number of sub-topics, including care for children, families and the elderly and various topics related to the treatment of illnesses.

Companies active in the health sector are frequently criticised for implementing prohibitive pricing that does not allow the poorest individuals to obtain access to health treatments they urgently need, a practice that clearly goes against the SDGs' demand that health be improved and well-being be delivered to all. Our approach does not allow us to differentiate between those startups in the medical field that aim to increase access and those that do not. However, innovations within the health sector offer enterprises the potential to address prevalent sustainability challenges. The extent to which such innovations are then monetised is a different albeit very important question. It is not uncommon for new products to be prohibitively expensive and to become more affordable as innovation continues and patents expire. This initial pricing does not make the innovations less valuable. As this paper aims to determine and compare how prevalent particular industries are in different EEs and not to conduct an impact assessment of startups in those industries, we see the inclusion of health startups in our sample as valuable.

The second most prominent topic is 'Quality Education' (20.5%), covering various enterprises fostering learning at all ages. It is thus noteworthy that the vast majority of sustainability startups in our sample do not focus on environmental issues addressed by the SDGs. The three environmental SDGs represented in our dataset, 'Clean Energy' (6.8%), 'Life on Land' (2.6%) and 'Responsible Consumption' (2.2%), are addressed in sum by only 11.6% of the sustainability startups we identified. This finding confirms the results of previous studies which have found that companies engaged in responsibility-oriented activities place less emphasis on environmental than on social issues (Shnayder et al., 2015). One explanation offered for this is that environmental initiatives are often more difficult and more costly to implement than social ones. Furthermore, there is likely to be greater demand for products and services that improve customers' well-being by supporting, for example, their health or education than for green products, which may be more expensive than the products customers currently use or be less likely to facilitate a scalable business model (van der Linden, 2018). We assess ecosystem characteristics that may be conducive to an EE having a higher share of sustainability startups in the Factors for Sustainability section.

4.2. ... and where to find them

Our analysis indicates that the share of sustainability startups varies widely across the 28 EEs we investigated, which are among the most prominent in the world. See Figs. 4 and 5 for an overview of our findings. While in some ecosystems, a large share of startups emphasise sustainability, in others consideration for sustainability appears practically absent. With as many as 14.4% of the startups in Boston contributing to the SDGs, the EE of Boston tops our list of most sustainability-oriented EEs, immediately followed by Houston, where 14.3% are sustainability startups. The position of Boston in our list is in line with the 2017 SocEnt City report on social startups in the US, which found that Boston is the best location in the country for social enterprises (A Deeper Dive: Social Enterprise Ecosystems in the U.S., 2017). Meanwhile, the greatest absolute number of sustainability startups is found in Silicon Valley, which only ranks 11th in our index.

One particularly striking finding of our study is that the entrepreneurial ecosystems with the greatest share of sustainability startups are not those that usually top other ecosystem rankings such as the Global Startup Ecosystem Report (Startup Genome Project, 2017). Of the 10 most sustainability-oriented entrepreneurial ecosystems only Boston and Seattle also appear in the top 10 of the Global Startup Ecosystem Report ranking (startups that appear in the top 10 of the Genome ranking are marked with an asterisk in Fig. 4). This suggests that the methods that have thus far been employed to assess the quality of entrepreneurial ecosystems give little indication of the degree to which the startups that emerge in these entrepreneurial ecosystems are favourable to sustainability startups. Tel Aviv and Berlin, two of the most prominent and successful entrepreneurial ecosystems in the world (Startup Genome Project, 2017), even rank in the bottom quartile of our index (23rd and 26th out of 28, respectively). Meanwhile, of the 198 startups included in the dataset from Stockholm, we found only

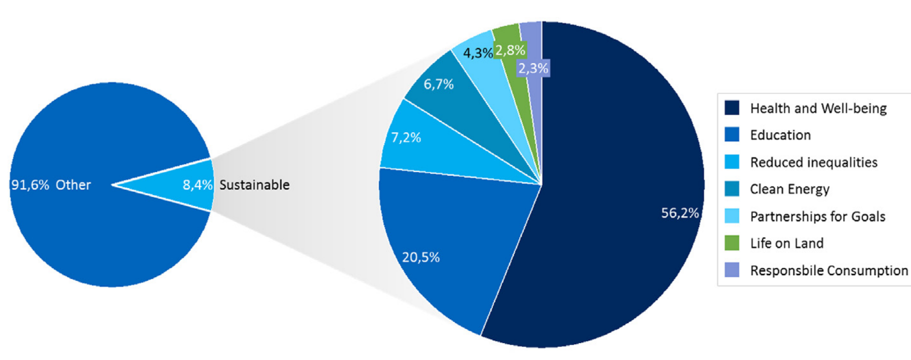

Fig. 3. Sustainability startups Overview.

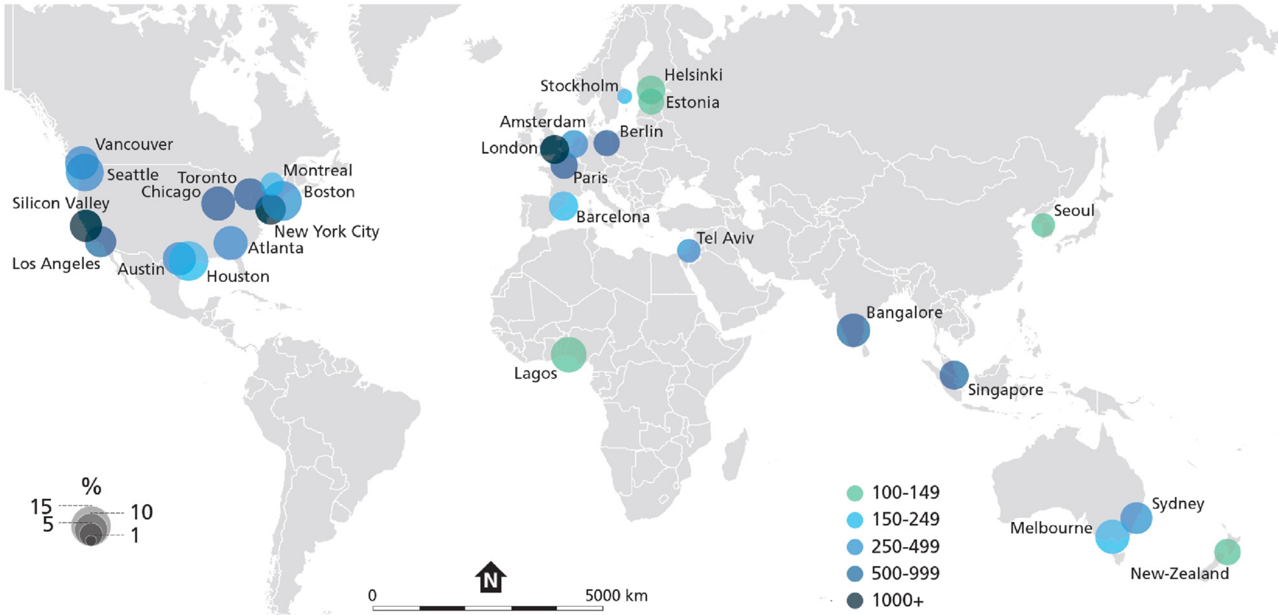

Fig. 4. Share of Sustainability Startups and Total Startups by Entrepreneurial Ecosystem – Global Overview. The size of each spere indicates the Share of Sustainability Startups, the color indicators the Total Startups in an Entrepreneurial Ecosystem. Made with Natural Earth under a CC0 license.

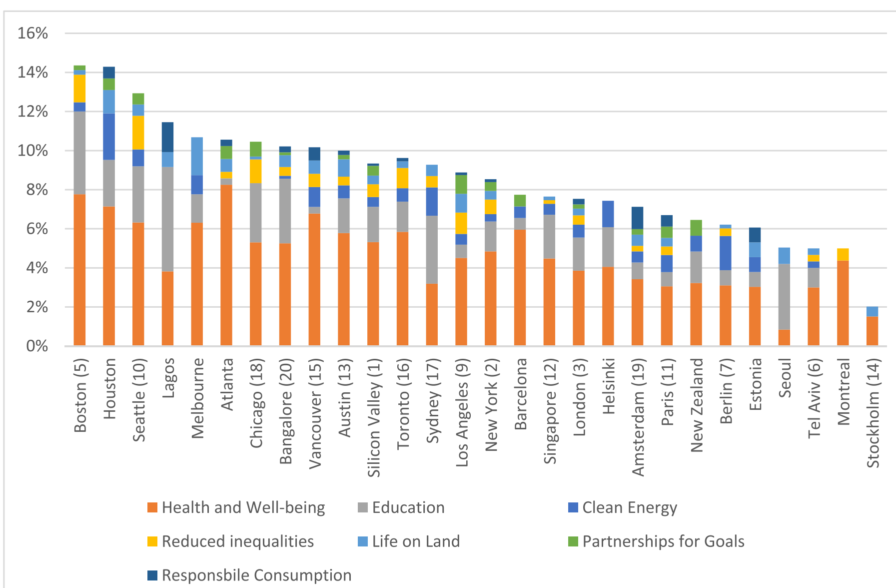

Fig. 5. Sustainability Score by Ecosystem. The brackets: (#) indicate the rank of the ecosystem in the top 20 of the Genome Startup Ecosystem Report 2017 – other ecosystems were not ranked.

four that prominently mention any of the identified sustainability topics on their website.

4.3. Factors for Sustainability

In order to better understand the conditions that may lead to regional differences in an ecosystem's share of sustainability startups, we have conducted an fsQCA, the principles of which are described above. Our choice of potential causal conditions draws on the theories that we presented in the Background section, which explain various potential causes for EEs having relatively high shares of sustainability startups.

GDP per capita was included in the model as a measure of regional economic development; higher GDPs are assumed to induce greater proportions of sustainability startups, as this causal relationship has been suggested by previous academic studies on the topic (Gerhards and Lengfeld, 2008). The amount of *social expenditure* was included for its importance to the failure thesis, as discussed in the Background section. The *share of female founders* in each ecosystem is the third condition included in the model. Finally, the *share of religious people* in the population was included as a measure of local religiosity; the presence of which was expected to induce higher shares of sustainability startups. In the first step of the fsQCA, the data was tested for the presence of a necessary condition for an EE having a high share of sustainability startups. This was done by testing the consistency value for each causal condition; this value is equal to the number of cases in which both the condition and the outcome are present divided by the number of cases in which the outcome is present, and thus denotes, the degree to which a causal condition leads to the outcome across the cases, for each causal condition (Ide, 2015). A condition was considered necessary if its consistency value was 0.9 or higher (Schneider and Wagemann, 2012). The presence of a high GDP per capita was the only causal condition that met the threshold for being considered a necessary condition (Schneider and Wagemann, 2012), with a consistency value of 0.91 and a coverage value of 0.75. The coverage value tells us the proportion of cases in which the outcome occurs that the condition is also satisfied – in the present case, the proportion of EEs having relatively high shares of sustainability startups that also have relatively high GDPs. The result indicates that a high GDP per capita in an EE tends to appear jointly with a high share of sustainability startups. None of the other causal conditions meet the 0.9 consistency threshold.

Two causal pathways were identified in the fsQCA that can explain the emergence of ecosystems with high shares of sustainability startups. Table 6 represents these two distinct causal pathways with each pathway's full solution formula and consistency and coverage values. First, a high GDP per capita in combination with a high proportion of female founders are sufficient conditions for a entrepreneurial ecosystem with a high proportion of sustainability startups to emerge. Second, a high GDP per capita in a location in combination with a low level of religiosity also appears sufficient to induce a sustainability entrepreneurial ecosystem. Of the 15 cases of sustainability entrepreneurial ecosystems that we identified, 13 can be explained by one of these two causal recipes with several cases being overdetermined, meaning that both causal pathways are satisfied (details can be found in Table 9 in the Appendix). This is indicated by the relatively low unique coverage values of the two sets of causal conditions, with 0.14 for gdpcap*womenfound and 0.05 for gdpcap*~religious; these figures denote the share of cases that can *only* be explained by the corresponding causal pathway. The two cases that are not explained by the satisfaction of either set of conditions are those of Lagos and Bangalore, which both have high shares of sustainability startups but low GDP figures. The solution model identified by fsQCA has relatively high consistency and coverage values of 0.90 and 0.88, respectively. The sensitivity analyses that were conducted demonstrate these results to be robust (see the Sensitivity Analyses part of the Methods section).

The first causal pathway links a high share of female founders in combination with high GDP to the emergence of high shares of sustainability startups. In 1993, Terpstra et al. suggested that observed differences between the orientation of men and women towards social value creation are due to differential conditioning (Terpstra et al., 1993). It is noteworthy that the same difference in orientation still exists today. Furthermore, the fact that an ecosystem with high shares of female founders must also have significant economic development (in particular, high GDP per capita) to produce sustainability-oriented ecosystems seems to correspond with the finding of Hechavarría et al. (2017) that female founders in post-materialist societies are more likely than men to become 'ecopreneurs'.

The second causal pathway establishes a relationship between the combination of the absence of high shares of believers and high levels of economic development with the emergence of sustainability-oriented EEs. That this combination of conditions should be a sufficient cause for the latter outcome is somewhat surprising, as previous business research, as described above, has found religious belief to be positively associated with ethical behaviour (Longenecker et al., 2004; Weaver and Agle, 2002). One possible explanation for our result may be that lack of religiosity functions as a proxy for high levels of education, as religiosity and education have been shown to be inversely correlated (Glaeser and Sacerdote, 2008) and education to be linked to an orientation towards sustainability (Weaver and Agle, 2002). Because lack of religiosity and high levels of education are, like GDP per capita, attributes of the general population rather than entrepreneurs in particular, their presence in an ecosystem can have a threefold effect, producing greater shares of sustainability-oriented entrepreneurs, investors and consumers and creating a sort of virtuous cycle of investment, production and consumption that promotes sustainability. When assessing the relationship between religion and education, however, we might need to distinguish between particular religions, as it has been shown that education levels vary significantly between one religion and another (Pew Research Center, 2016). In all of those EEs that rank highest in our sample in their share of sustainability startups, the most commonly-practiced religion is Christianity, which research has suggested is the religion whose average practitioners have the second-highest level of education after those of the Jewish faith (Pew Research Center, 2016). Evidently, the relationship between religiosity, education and the orientation toward sustainability among entrepreneurs is an issue that calls for more nuanced investigation.

Among the world's most prominent EEs, measures of GDP, shares of female founders and religiosity often differ significantly from differences to those found in EEs with the highest shares of sustainability startups. Silicon Valley, for example, has the highest GDP per capita of the EEs considered in the sample of this study, but places in the middle of the pack in measures of female founder share and religiosity. Chicago, on the other hand, has a slightly lower GDP per capita, ranking slightly above the median GDP of the 28 EEs we considered, but has the highest share of female founders of all the EEs considered with 34%, and ultimately ranks securely in the top 10 of EEs in terms of their share of sustainability startups. Chicago is the location of a flourishing female entrepreneurship scene that is fuelled by investment firms such as Invest Her, which focuses on investing in women-led ventures, and by several sustainability startups that have been launched by women (Henry,

Table 6
Intermediate solution for the Outcome of Sustainability-oriented Entrepreneurial EcosystemS.

Causal pathway	gdpcap*womenfound	gdpcap*~religious
Consistency	0.91	0.90
Raw coverage	0.74	0.83
Unique coverage	0.14	0.05
Example cases	Toronto, Houston, Chicago, Sydney	Seattle, Austin, Vancouver
Solution formula	gdpcap*(womenfound+~religious) → sust	
Solution consistency	0.90	
Solution coverage	0.88	

* = and; + = or; ~ = absence; → = sufficient for.

2018). Silicon Valley, meanwhile, maintains its reputation as a 'boys' club' (Coren, 2018) that is geared towards profit generation rather than sustainability (Berger and Brem, 2016; Brock, 2012; Startup Genome Project, 2017). Berlin, another ecosystem that usually ranks relatively highly in other EE rankings (Startup Genome Project, 2017), finds itself in the bottom quartile in our ranking. With a relatively low GDP per capita, Berlin appears to lack the level of economic strength an EE requires to develop a high share of sustainability startups. While its population is predominantly non-religious (at 63%, it is the second least-religious EE in the sample after Estonia), this condition alone does not appear to be sufficient to drive the emergence of sustainability startups.

5. Conclusion

Policy-makers around the world are keen to develop vibrant entrepreneurial ecosystems in their cities, and particularly ones which focus on sustainability. We found that the highest proportion of sustainability startups among entrepreneurial ecosystems is located in Boston, followed by Houston, Seattle and Lagos, respectively. A qualitative comparative analysis of the causal patterns underlying our results reveals that high GDP in combination with either (1) high shares of female founders of startups or (2) high shares of non-religious people in the population induce entrepreneurial ecosystems with relatively high levels of sustainability enterprises.

Policy-makers, can benefit from considering the results of our analysis. One promising path that economically developed regions might pursue in order to attract sustainability entrepreneurs is to support female founders, and particularly those involved in sustainability startups, through measures such as the creation of women-focused incubators. Such initiatives would both help to increase sustainability-focused entrepreneurial endeavours and improve gender parity in entrepreneurship overall.

Our research has several limitations. First, our lists of entrepreneurial ecosystems and the startups within them are not comprehensive due to the incompleteness of the Crunchbase database that was used to collect the underlying data for this study. However, Crunchbase is widely recognised as the most comprehensive startup database available. Furthermore, because it focuses on high-growth innovative startups, Crunchbase contains information on the most ambitious and globally-oriented startups in each ecosystem, and thus on those that are likely to contribute most to the achievement of the SDGs. We also acknowledge that whether or not the pursuit of a particular SDG can be properly attributed to a given startup is often ambiguous. We attempted to overcome this ambiguity by having several of the authors of this paper examine the others' attribution criteria and results. Future researchers may wish to delve deeper into the topic we have studied by seeking to identify not only sustainability startups but also *sustainable startups*, or those startups that emphasise sustainability in the ways they operate their businesses. Furthermore, including only six potential causal conditions (including sensitivity analyses), we may have omitted other conditions that also contribute to the emergence of sustainability-oriented EEs. However, as the set of conditions we have assessed has included causal factors that explain the outcomes of most of the cases we considered, we can conclude that several important factors are included in this set. Future research might examine an even larger number of EEs in order to possibly identify other leading ecosystems of sustainability entrepreneurship (pending, however, the development of improved databases on entrepreneurial ecosystems and their startups). In addition, we recommend further research on the drivers of such ecosystems' emergence and the possible interplay of these drivers.

CRediT authorship contribution statement

Sarah Tiba: Methodology, Conceptualization, Formal analysis, Writing – original draft, Visualization. **Frank J. van Rijnsoever:** Supervision, Conceptualization, Writing – review & editing. **Marko P. Hekkert:** Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank Allard van Mossel for his invaluable help with the Latent Dirichlet Allocation. Many thanks also to Maryse Chappin, Chris Eveleens and Julian Kirchherr for helpful conversations on the methods employed. Furthermore, we are grateful to Wouter Boon and Jarno Hoekman for their constructive comments on this piece of work.

Appendix

Table 7
Raw Data Matrix.

Ecosystem	Share sustainable enterprises	Share Women Founders	Environmental Performance Index (EPI)	Share of population religious	Early Stage Funding per Startup (USD k)	GDP/capita (USD k)	Social Expenditure (share of GDP)
SILICON	9,3%	18,0%	84,72	45%	762	81,4	19,3%
LONDON	7,5%	15,0%	87,38	79%	451	38,7	21,5%
BOSTON	14,4%	16,0%	84,72	47%	495	81,3	19,3%
TELAVIV	5,0%	8,0%	78,14	95%	509	41,4	16,1%
BERLIN	6,2%	13,0%	84,26	37%	483	26,3	25,3%
LA	8,9%	16,0%	84,72	59%	450	61,9	19,3%
SEATTLE	12,9%	13,0%	84,72	41%	332	81,4	19,3%
PARIS	6,7%	10,0%	88,20	72%	270	57,3	31,5%
SINGAPORE	7,6%	12,0%	87,04	85%	276	48,0	11,4%
AUSTIN	10,0%	12,0%	84,72	46%	410	57,5	19,3%
STOCKHOLM	2,0%	12,0%	90,43	72%	325	65,0	27,1%
VANCOUVER	10,2%	12,0%	85,06	51%	334	44,0	17,2%
TORONTO	9,3%	18,0%	85,06	79%	443	53,0	17,2%
SYDNEY	9,3%	22,0%	87,22	56%	254	64,0	19,1%
CHICAGO	10,5%	34,0%	84,72	64%	271	63,8	19,3%
AMSTERDAM	7,1%	14,0%	82,03	38%	144	45,9	22,0%

Table 7 (*continued*)

Ecosystem	Share sustainable enterprises	Share Women Founders	Environmental Performance Index (EPI)	Share of population religious	Early Stage Funding per Startup (USD k)	GDP/capita (USD k)	Social Expenditure (share of GDP)
BANGALORE	10,2%	10,0%	53,58	100%	229	5,2	1,5%
BARCELONA	7,7%	14,0%	88,91	67%	223	31,7	24,6%
ESTONIA	6,1%	16,0%	88,59	29%	272	17,7	17,4%
HELSINKI	7,4%	8,0%	90,68	75%	358	55,0	30,8%
MELBOURNE	10,7%	18,0%	87,22	69%	157	53,8	19,1%
NEWZEALAND	6,5%	21,0%	88,00	58%	301	39,8	19,5%
SEOUL	5,0%	16,0%	70,61	54%	174	26,5	10,4%
HOUSTON	14,3%	21,0%	84,72	64%	110	80,8	19,3%
LAGOS	11,5%	14,0%	58,27	100%	77,8	3,6	0,3%
MONTREAL	5,0%	12,0%	85,06	85%	123	38,0	17,2%
ATLANTA	10,6%	17,0%	84,72	66%	367	58,0	19,3%
NYC	8,4%	19,0%	84,72	61%	568	65,0	19,3%

Table 8
Calibrated Raw Data Matrix.

ecosystem	resp	womenfound	epi	religious	funding	gdpcap	socexp
SILICON	0.67	0.67	0.67	0.33	1	1	0.67
LONDON	0.33	0.33	1	0.67	0.67	0.67	0.67
BOSTON	1	0.67	0.67	0.33	0.67	1	0.67
TELAVIV	0	0	0.33	1	0.67	0.67	0.67
BERLIN	0.33	0.33	0.67	0	0.67	0.33	0.67
LA	0.67	0.67	0.67	0.33	0.67	0.67	0.67
SEATTLE	1	0.33	0.67	0	0.33	1	0.67
PARIS	0.33	0	1	0.67	0.33	0.67	1
SINGAPORE	0.33	0.33	1	0.67	0.33	0.67	0.33
AUSTIN	0.67	0.33	0.67	0.33	0.67	0.67	0.67
STOCKHOLM	0	0.33	1	0.67	0.33	0.67	0.67
VANCOUVER	0.67	0.33	0.67	0.33	0.33	0.67	0.67
TORONTO	0.67	0.67	0.67	0.67	0.67	0.67	0.67
SYDNEY	0.67	1	1	0.33	0.33	0.67	0.67
CHICAGO	0.67	1	0.67	0.33	0.33	0.67	0.67
AMSTERDAM	0.33	0.33	0.67	0	0	0.67	0.67
BANGALORE	0.67	0	0	1	0.33	0	0
BARCELONA	0.33	0.33	1	0.33	0.33	0.33	0.67
ESTONIA	0.33	0.67	1	0	0.33	0.33	0.67
HELSINKI	0.33	0	1	0.67	0.33	0.67	1
MELBOURNE	0.67	0.67	1	0.33	0	0.67	0.67
NEWZEALAND	0.33	1	1	0.33	0.33	0.67	0.67
SEOUL	0.33	0.67	0.33	0.33	0	0.33	0.33
HOUSTON	1	1	0.67	0.33	0	1	0.67
LAGOS	0.67	0.33	0	1	0	0	0
MONTREAL	0	0.33	0.67	0.67	0	0.67	0.67
ATLANTA	0.67	0.67	0.67	0.33	0.33	0.67	0.67
NYC	0.67	0.67	0.67	0.33	0.67	0.67	0.67

Table 9
Truth Table Main Analysis.

womenfound	religious	gdpcap	socexp	number	Sust	raw consist.	PRI consist.	SYM consist
1	0	1	1	10	1	0.909008	0.801980	0.801980
1	1	1	1	1	1	0.900151	0.602410	0.602410
0	0	1	1	4	1	0.879373	0.626866	0.626866
1	0	0	1	1	0	0.848714	0.000000	0.000000
1	0	0	0	1	0	0.848714	0.000000	0.000000
0	0	0	1	2	0	0.797583	0.000000	0.000000
0	1	0	0	2	0	0.790143	0.340000	0.407186
0	1	1	0	1	0	0.763345	0.198795	0.198795
0	1	1	1	6	0	0.620301	0.098214	0.098214
1	0	1	0	0				
1	1	1	0	0				
1	1	0	1	0				
1	1	0	0	0				
0	0	1	0	0				
0	0	0	0	0				
0	1	0	1	0				

References

A Deeper Dive, 2017. Social Enterprise Ecosystems in the U.S..
Ács, Z.J., Autio, E., Szerb, L., 2014. National systems of entrepreneurship: measurement issues and policy implications. Res. Pol. 43, 476–494. https://doi.org/10.1016/j.respol.2013.08.016.
Acs, Z.J., Szerb, L., Lloyd, A., 2017. Enhancing entrepreneurial ecosystems: a GEI approach to entrepreneurship policy. In: Global Entrepreneurship and Development Index 2017. Springer, pp. 81–91.
Alvedalen, J., Boschma, R.A., 2017. A critical review of entrepreneurial ecosystems research: towards a future research agenda. Eur. Plann. Stud. 25, 887–903.
Bacq, S., Janssen, F., 2011. The multiple faces of social entrepreneurship: a review of definitional issues based on geographical and thematic criteria. Enterpren. Reg. Dev. 23, 373–403. https://doi.org/10.1080/08985626.2011.577242.
Bank, N., Fichter, K., Klofsten, M., 2017. Sustainability-profiled incubators and securing the inflow of tenants – the case of Green Garage Berlin. J. Clean. Prod. 157, 76–83. https://doi.org/10.1016/J.JCLEPRO.2017.04.123.
Baum, J.A.C., Calabrese, T., Silverman, B.S., 2000. Don't go it alone: alliance network composition and startups' performance in Canadian biotechnology. Strat. Manag. J. 267–294.
Berger, A., Brem, A., 2016. Innovation hub how-to: lessons from Silicon Valley. Glob. Bus. Organ. Excell. 35, 58–70. https://doi.org/10.1002/joe.21698.
Bertazzi, P., (GRI), Lund Hansen, O., (UN Global Compact), Didden, M., (WBCSD), 2015. SDG Compass: The guide for business action on the SDGs. GRI, UN Global Compact, WBCD. https://sdgcompass.org/wp-content/uploads/2015/12/019104_SDG_Compass_Guide_2015.pdf.
Blei, D.M., Lafferty, J.D., 2007. A correlated topic model of science. Ann. Appl. Stat. 1, 17–35. https://doi.org/10.1214/07-AOAS136.
Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3, 993–1022.
Bosma, N., Schøtt, T., Terjesen, S., Kew, P., 2015. GEM Special Topic Report - Social Entrepreneurship.
Brock, D.C., 2012. From automation to Silicon Valley: the automation movement of the 1950s, Arnold Beckman, and William Shockley. Hist. Technol. 28, 375–401. https://doi.org/10.1080/07341512.2012.756236.
Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J.L., Blei, D.M., 2009. Reading tea leaves: how humans interpret topic models. In: Advances in Neural Information Processing Systems, pp. 288–296.
Cohen, B., 2006. Sustainable valley entrepreneurial ecosystems. Bus. Strat. Environ. 15, 1–14. https://doi.org/10.1002/bse.428.
Coren, M., 2018. Women founders and investors are finally breaking up the ultimate boys club in Silicon Valley — quartz at Work. Quartz. https://qz.com/work/1372702/female-founders-and-investors-are-finally-breaking-up-the-ultimate-boys-club-in-silicon-valley/.
CrunchBase, 2016. CrunchBase [WWW Document]. https://www.crunchbase.com/organization/500-startups#/entity. accessed 11.17.16.
Dacin, P.A., Dacin, M.T., Matear, M., 2010. Social entrepreneurship: why we don't need a new theory and how we move forward. Source Acad. Manag. Perspect. 24, 37–57.
Davidsson, P., Wiklund, J., 2001. Levels of analysis in entrepreneurship research: current research practice and suggestions for the future. Enterpren. Theor. Pract. 25, 81–100.
Dean, T.J., McMullen, J.S., 2007. Toward a theory of sustainable entrepreneurship: reducing environmental degradation through entrepreneurial action. J. Bus. Ventur. 22, 50–76. https://doi.org/10.1016/j.jbusvent.2005.09.003.
Dees, J.G., 1998. The Meaning of "Social Entrepreneurship." Environmental Performance Index [WWW Document], 2018.
Eveleens, C.P., Van Rijnsoever, F.J., Niesten, E.M.M.I., 2017. How network-based incubation helps start-up performance: a systematic review against the background of management theories. J. Technol. Tran. 42, 676–713. https://doi.org/10.1007/s10961-016-9510-7.
Feinerer, I., Hornik, K., 2017. tm.
Gast, J., Gundolf, K., Cesinger, B., 2017. Doing business in a green way: a systematic review of the ecological sustainability entrepreneurship literature and future research directions. J. Clean. Prod. 147, 44–56. https://doi.org/10.1016/j.jclepro.2017.01.065.
Geels, F.W., Sovacool, B.K., Schwanen, T., Sorrell, S., 2017. Sociotechnical transitions for deep decarbonization. Science (80-.) 357, 1242–1244.
Gelissen, J., 2007. Explaining popular support for environmental protection A multilevel analysis of 50 nations. Environ. Behav. 39, 392–415. https://doi.org/10.1177/0013916506292014.
Gerhards, J., Lengfeld, H., 2008. Support for European Union environmental policy by citizens of EU-member and accession states. Comp. Sociol. 7, 215–241. https://doi.org/10.1163/156913308X289050.
Glaeser, E.L., Sacerdote, B.I., 2008. Education and religion. J. Hum. Cap. 2, 188–215. https://doi.org/10.1086/590413.
Global Entrepreneurship Monitor, 2017. GEM 2016/2017–Global Report 2016/17. Global Entrepreneurship Monitor.
Griffiths, M.D., Gundry, L.K., Kickul, J.R., 2013. The socio-political, economic, and cultural determinants of social entrepreneurship activity. J. Small Bus. Enterprise Dev. 20, 341–357. https://doi.org/10.1108/14626001311326761.
Griffiths, T., Steyvers, M., 2004. Finding scientific topics. Proc. Natl. 101, 5228–5235.
Grimmer, J., Stewart, B.M., 2013. Text as data: the promise and pitfalls of automatic content analysis methods for political texts. Polit. Anal. 21, 267–297. https://doi.org/10.1093/pan/mps028.
Grün, B., Hornik, K., 2017. Topicmodels.
Guerreiro, J., Rita, P., Trigueiros, D., 2016. A text mining-based review of cause-related marketing literature. J. Bus. Ethics 139, 111–128. https://doi.org/10.1007/s10551-015-2622-4.
Hansen, E.G., Schaltegger, S., 2013. 100 per cent organic? A sustainable entrepreneurship perspective on the diffusion of organic clothing. Corp. Govern. 15, 583–598. https://doi.org/10.1108/CG-06-2013-0074.
Harris, J.D., Sapienza, H.J., Bowie, N.E., 2009. Ethics and entrepreneurship. J. Bus. Ventur. 24, 407–418. https://doi.org/10.1016/j.jbusvent.2009.06.001.
Hechavarría, D., 2016. The impact of culture on national prevalence rates of social and commercial entrepreneurship. Int. Enterpren. Manag. J. 12, 1025–1052.
Hechavarria, D.M., Ingram, A.E., 2016. The entrepreneurial gender divide. Int. J. Gend. Entrep. 8, 242–281. https://doi.org/10.1108/IJGE-09-2014-0029.
Hechavarría, D.M., Terjesen, S.A., Ingram, A.E., Renko, M., Justo, R., Elam, A., 2017. Taking care of business: the impact of culture and gender on entrepreneurs' blended value creation goals. Small Bus. Econ. 48, 225–257. https://doi.org/10.1007/s11187-016-9747-4.
Henry, Z., 2018. Women Entrepreneurship Report: Why This Midwestern City Has the Highest Concentration of Women Entrepreneurs in America. Inc. Mag.
Hoogendoorn, B., 2016. The prevalence and determinants of social entrepreneurship at the macro level. J. Small Bus. Manag. 54, 278–296. https://doi.org/10.1111/jsbm.12301.
Horne, J., Recker, M., Michelfelder, I., Jay, J., Kratzer, J., 2020. Exploring entrepreneurship related to the sustainable development goals-mapping new venture activities with semi-automated content analysis. J. Clean. Prod. 242, 118052.
Ide, T., 2015. Why do conflicts over scarce renewable resources turn violent? A qualitative comparative analysis. Global Environ. Change 33, 61–70. https://doi.org/10.1016/j.gloenvcha.2015.04.008.
Johnson, M.P., Schaltegger, S., 2020. Entrepreneurship for sustainable development: a review and multilevel causal mechanism framework. Enterpren. Theor. Pract. 44, 1141–1173.
PMG, United Nations Global Compact, n.d. SDG Industry Matrix | UN Global Compact..
Krestel, R., Fankhauser, P., Nejdl, W., 2009. Latent dirichlet allocation for tag recommendation. In: Proceedings of the Third ACM Conference on Recommender Systems. ACM. https://doi.org/10.1145/1639714.1639726.
Leendertse, J., Rijnsoever, F.J., Eveleens, C.P., 2020. The sustainable start-up paradox: predicting the business and climate performance of start-ups. Bus. Strat. Environ. 1–18. https://doi.org/10.1002/bse.2667.
Legewie, N., 2013. An introduction to applied data analysis with qualitative comparative analysis (QCA). Forum Qual. Soc. Res. 14. Art. 15.
Li, W., McCallum, A., 2006. Pachinko allocation: DAG-structured mixture models of topic correlations. In: Proc. Int'l. Conf. On Machine Learning, pp. 577–584.
Longenecker, J.G., Mckinney, J.A., Moore, C.W., 2004. Religious intensity, Evangelical Christianity, and business ethics: an empirical study. J. Bus. Ethics 55, 373–386.
Lyon, T.P., Montgomery, A.W., 2015. The means and end of greenwash. Organ. Environ. 28, 223–249. https://doi.org/10.1177/1086026615575332.
Mair, J., Marti, I., 2006. Social entrepreneurship research: a source of explanation, prediction, and delight. J. World Bus. 41, 36–44.
Matsunaga, Y., Naoto, Y., Okuyama, N., 2010. What determines the size of the nonprofit sector?: a cross-country analysis of the government failure theory. Voluntas 21, 180–201. https://doi.org/10.1007/s11266-010-9125-9.
McAdam, M., 2013. Female Entrepreneurship. Routledge, Oxford.
Muñoz, P., Cohen, B., 2017. Towards a social-ecological understanding of sustainable venturing. J. Bus. Ventur. Insights 7, 1–8.
Muñoz, P., Dimov, D., 2015. The call of the whole in understanding the development of sustainable ventures. J. Bus. Ventur. 30, 632–654.
Muñoz, P., Janssen, F., Nicolopoulou, K., Hockerts, K., 2018. Advancing sustainable entrepreneurship through substantive research. Int. J. Entrepreneurial Behav. Res. 24 (2), 322–332.
Nerini, F.F., Sovacool, B., Hughes, N., Cozzi, L., Cosgrave, E., Howells, M., Tavoni, M., Tomei, J., Zerriffi, H., Milligan, B., 2019. Connecting climate action with other sustainable development goals. Nat. Sustain. 2, 674–680.
Neumeyer, X., Santos, S.C., 2018. Sustainable business models, venture typologies, and entrepreneurial ecosystems: a social network perspective. J. Clean. Prod. 172, 4565–4579.
Nilsson, M., Griggs, D., Visbeck, M., 2016. Map the interactions between sustainable development goals: Mans Nilsson, Dave Griggs and Martin Visbeck present a simple way of rating relationships between the targets to highlight priorities for integrated policy. Nature 534, 320–323.
Ooms, J., Sites, D., 2017. Google's Compact Language Detector 2 [R package cld2 version 1.1].
Pew Research Center, 2016. Religion and Education Around the World.
Piepenbrink, A., Nurmammadov, E., 2015. Topics in the literature of transition economies and emerging markets. Scientometrics 102, 2107–2130. https://doi.org/10.1007/s11192-014-1513-2.
Pitelis, C., 2012. Clusters, entrepreneurial ecosystem co-creation, and appropriability: a conceptual framework. Ind. Corp. Change 21, 1359–1388. https://doi.org/10.1093/icc/dts008.
Piwowar-Sulej, K., Krzywonos, M., Kwil, I., 2021. Environmental entrepreneurship–bibliometric and content analysis of the subject literature based on H-core. J. Clean. Prod. 126277.
R Core Team, 2017. R: A Language and Environment for Statistical Computing.
Ragin, C., 2009. Redesigning Social Inquiry: Fuzzy Sets and beyond. University of

Chicago Press, Chicago.
Saebi, T., Foss, N.J., Linder, S., 2019. Social entrepreneurship research: past achievements and future promises. J. Manag. 45, 70–95.
Salton, G., McGill, M., 1983. Introduction to Modern Information Retrieval. McGraw-Hill.
Schmid, H., 1995. TreeTagger.
Schneider, C., Wagemann, C., 2012. Set-Theoretic Methods for the Social Sciences (Strategies for Social Inquiry), first ed. Cambridge University Press, Cambridge.
Shnayder, L., Van Rijnsoever, F.J., Hekkert, M.P.M.P., 2015. Putting Your Money where Your Mouth Is: why sustainability reporting based on the triple bottom line can be misleading. PLoS One 10, 1–23. https://doi.org/10.1371/journal.pone.0119036.
Stam, E., 2015. Entrepreneurial ecosystems and regional policy: a sympathetic critique. Eur. Plann. Stud. 23, 1759–1769.
Stam, E., 2014. The Dutch Entrepreneurial Ecosystem, Birch Research. https://doi.org/10.2139/ssrn.2473475.
Startup Genome Project, 2017. Global Startup Ecosystem Report 2017.
Sugimoto, C.R., Li, D., Russell, T.G., Finlay, S.C., Ding, Y., 2011. The shifting sands of disciplinary development: analyzing North American Library and Information Science dissertations using latent Dirichlet allocation. J. Am. Soc. Inf. Sci. Technol. 62, 185–204. https://doi.org/10.1002/asi.21435.
Sullivan, K., Thomas, S., Rosano, M., 2018. Using industrial ecology and strategic management concepts to pursue the sustainable development. Goals. J. Clean. Prod. 174, 237–246.
Terán-Yépez, E., Marín-Carrillo, G.M., del Pilar Casado-Belmonte, M., de las Mercedes Capobianco-Uriarte, M., 2020. Sustainable entrepreneurship: review of its evolution and new trends. J. Clean. Prod. 252, 119742.
Terpstra, D., Rozell, E., Robinson, R., 1993. The influence of personality and demographic variables on ethical decisions related to insider trading. J. Psychol. 127, 375–389.
The United Nations, 2017. The Sustainable Development Goals Report. New York City.
Theodoraki, C., Messeghem, K., Rice, M.P., 2018. A social capital approach to the development of sustainable entrepreneurial ecosystems: an explorative study. Small Bus. Econ. 51, 153–170.
Thiem, A., Spoel, R., Dus, A., 2016. Enhancing sensitivity diagnostics for qualitative comparative analysis: a combinatorial approach. Polit. Anal. 24, 104–120. https://doi.org/10.1093/pan/mpv028.
Tiba, S., van Rijnsoever, F.J., Hekkert, M.P., 2020. The lighthouse effect: how successful entrepreneurs influence the sustainability-orientation of entrepreneurial ecosystems. J. Clean. Prod. 264, 121616. https://doi.org/10.1016/j.jclepro.2020.121616.
Tiba, S., van Rijnsoever, F.J., Hekkert, M.P.M.P., 2018. Firms with benefits: a systematic review of responsible entrepreneurship and corporate social responsibility literature. Corp. Soc. Responsib. Environ. Manag. 26, 265–284. https://doi.org/10.1002/csr.1682.
Titov, I., Mcdonald, R., 2008. Modeling online reviews with multi-grain topic models. In: ACM (Ed.), Proceedings of the 17th International Conference on World Wide Web.
Toman, M., Tesar, R., Jezek, K., 2006. Influence of word normalization on text classification. In: Proceedings of InSciT, vol. 4, pp. 354–358.
Uys, J., Schutte, C., Van Zyl, W., 2011. Trends in an international industrial engineering research journal: a textual information analysis perspective. In: Proceedings of the 41st International Conference on Computers & Industrial Engineering. Los Angeles, California.
van der Linden, S., 2018. Warm glow is associated with low- but not high-cost sustainable behaviour. Nat. Sustain. 1, 28–30. https://doi.org/10.1038/s41893-017-0001-0.
van Rijnsoever, F.J., 2020. Meeting, mating, and intermediating: how incubators can overcome weak network problems in entrepreneurial ecosystems. Res. Pol. 49, 103884. https://doi.org/10.1016/j.respol.2019.103884.
Volkmann, C., Fichter, K., Klofsten, M., Audretsch, D.B., 2019. Sustainable entrepreneurial ecosystems: an emerging field of research. Small Bus. Econ. 1–9.
Walley, E.E., Taylor, D.W., 2002. Opportunists, champions, mavericks...? A Typology of Green Entrepreneurs. Greener Manag. Int. 38, 31–43.
Weaver, G., Agle, B., 2002. Religiosity and ethical behavior in organizations: a symbolic interactionist perspective. Acad. Manag. Rev. 27, 77–97.
York, J.G., Venkataraman, S., 2010. The entrepreneur–environment nexus: uncertainty, innovation, and allocation. J. Bus. Ventur. 25, 449–463. https://doi.org/10.1016/j.jbusvent.2009.07.007.
Zahra, S.A., Gedajlovic, E., Neubaum, D.O., Shulman, J.M., 2009. A typology of social entrepreneurs: motives, search processes and ethical challenges. J. Bus. Ventur. 24, 519–532.
Zheng, B., Mclean, D.C., Lu, X., 2006. Identifying biological concepts from a protein-related corpus with a probabilistic topic model. BMC Bioinf. 7 https://doi.org/10.1186/1471-2105-7-58.