International Journal of Information Management Data Insights 4 (2024) 100287

Contents lists available at ScienceDirect

International Journal of Information Management Data Insights

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

Location factors and ecosystem embedding of sustainability-engaged blockchain companies in the US. A web-based analysis

Jan Kinne [a,b,c,*], Robert Dehghan [a,d], Sebastian Schmidt [a,e], David Lenz [a,f], Hanna Hottenrott [b,g]

[a] ISTARI.AI, Julius-Hatry-Str. 1, 68163 Mannheim, Germany
[b] Department of Economics of Innovation and Industrial Dynamics, Leibniz Centre for European Economic Research (ZEW), L7 1, 68161 Mannheim, Germany
[c] Center for Geographic Analysis at Harvard University, 1737 Cambridge Street, Cambridge, MA 02138, USA
[d] Department of SME Research and Entrepreneurship, Ehrenhof Ost, University of Mannheim, 68161 Mannheim, Germany
[e] Department of Geoinformatics – Z GIS, University of Salzburg, Schillerstr. 30, 5020 Salzburg, Austria
[f] Department of Statistics and Econometrics, Justus-Liebig-University, Licher Str. 64, 35394 Giessen, Germany
[g] Department of Economics & Policy, Technical University of Munich, Arcisstr. 21, 80333 Munich, Germany

A R T I C L E   I N F O

JEL classification:
Q56
R30
L86

Keywords:
Sustainability
Blockchain
Ecosystem
Location factors
Natural language processing

A B S T R A C T

While many digital technologies provide opportunities for creating business models that impact sustainability, some technologies, especially blockchain applications, are often criticized for harming the environment, e.g. due to high energy demand. In our study, we present a novel approach to identifying sustainability-focused blockchain companies and relate their level of engagement to location factors and entrepreneurial ecosystem embeddedness. For this, we use a large-scale web scraping approach to analyze the textual content and hyperlink networks of all US companies from their websites. Our results show that blockchain remains a niche technology, with its use communicated by about 0.6% of US companies. However, the proportion of blockchain companies that are committed to sustainability is significantly higher than in the overall firm population. Additionally, we find that such sustainability-engaged blockchain companies have, at least quantitatively, a more intensive embedding in entrepreneurial ecosystems, while infrastructural and socio-economic location factors hardly play a role.

1. Introduction

Understanding the sustainability potential of novel information technology is crucial for assessing its long-term impact on society. One of these emerging technologies gaining significant attention in recent years is blockchain which is among the most controversial digital technologies with regard to its environmental impact (Asongu, Agboola, Alola, & Bekun, 2020; Jones, Goodkind, & Berrens, 2020; Stoll, Klaaßen, & Gallersdörfer, 2019). In principle, blockchain can be used both in environmentally harmful ways (e.g. high emissions due to energy intensity) and in applications that reduce waste of natural resources, increase efficiency in inputs or improve efficiency in the distribution of products. In light of the debate on the (non-)sustainable use of blockchain technology, it seems crucial to explore whether and to what extent blockchain companies pursue sustainability goals and which factors contribute to using blockchain technology in a more sustainable context.

This study explores the use of blockchain technology by companies across the United States (US) in a sustainability context. Since retrieving data on such a specific topic from traditional databases is difficult (Kinne & Axenbeck, 2020), we relied on an innovative, web-based methodology. We identified companies with blockchain-based business models using web text mining and a deep learning approach. While the possible benefits (Kushwaha, Kar, & Dwivedi, 2021; Nair, Agrawal, Domnic, & Kumar, 2021; Patón-Romero, Baldassarre, Toval, Rodríguez, & Piattini, 2022; Simmonds & Bhattacherjee, 2012) and threats of information technology (Asongu et al., 2020; Jones et al., 2020) to the environment have been discussed, there is still a lack of large-scale studies on the adoption of information technology in sustainable applications. Therefore, we further identified companies employing sustainable blockchain applications such as in the areas of energy management, supply chain management, resource use, waste management, or the monitoring of natural disasters. To our knowledge, this is the first paper to examine the relationship between blockchain and sustainability for the entire US firm population.

While research on innovation ecosystems generally stresses the importance of location factors for the invention and adoption of new technologies and for the performance of companies using and diffusing

* Corresponding author at: Center for Geographic Analysis at Harvard University, 1737 Cambridge Street, Cambridge, MA 02138, USA.
E-mail address: jan.kinne@istari.ai (J. Kinne).

https://doi.org/10.1016/j.jjimei.2024.100287

Available online 21 September 2024

them (Asheim & Gertler, 2006; Williamson & Meyer, 2012), it remains an open question whether local milieus still matter for digital technologies. To answer these questions, we related the use of sustainable blockchain applications with the local business ecosystem and infrastructure. For this, we distinguished between (physical) infrastructure and the local business ecosystem in their role in the sustainable use of blockchain.

Our paper addresses the research gap regarding the unclear relationship between sustainability commitment and blockchain use in the US, as well as which location factors influence this relationship. To our knowledge, no studies have yet investigated this relation at the company level. Within the framework of this paper, we address the following research questions:

- **RQ1**: How many companies are using blockchain technologies in the US?
- **RQ2**: How important is the topic of sustainability for these companies?
- **RQ3**: What role do local location factors and the embedding of blockchain companies in corporate networks play in their sustainability alignment?

2. Theoretical background

In the following, we provide an overview over necessary theoretical background information regarding blockchain technologies and ecosystem embeddings.

2.1. Information systems and environmental sustainability

Technological progress is key for transforming business practices and consumer behavior (Aghion, Antonin, & Bunel, 2021). In particular, information technologies provide the opportunity to support the transition toward a more sustainable economy (Singh & Sahu, 2020; Wang, Chen, & Benitez-Amado, 2015; Wang, Luo, et al., 2015). Green IT, Artificial Intelligence (AI) and the Internet of Things (IoT), for example, offer solutions for reducing negative environmental impact through increasing resource-use efficiency (Kushwaha et al., 2021; Nair et al., 2021; Patón-Romero et al., 2022; Simmonds & Bhattacherjee, 2012). For instance, big data analytics can contribute to precision agriculture and the reduction of water waste or the use of fertilizers, pesticides and other pollutants (Dwivedi et al., 2022). Applications may also reduce fuel requirements and food waste during transportation. AI alone or in combination with other technologies (e.g. blockchain) has the potential to enhance business decisions and hence improve practices from input choice (including water and energy use) to supply chain management and waste reduction (Kshetri, 2018; Lin, Petway, Lien, & Settele, 2018; Nishant, Kennedy, & Corbett, 2020).

However, some information technologies are also often discussed in light of their potential negative environmental impact (Asongu et al., 2020; Jones et al., 2020). The main concerns relate to the high levels of energy consumption and related greenhouse gas emissions as well as the use of toxic disposal of devices used for operation (Murugesan & Benakanahally Lakshminarasaiah, 2022). Blockchain technology is probably the most controversially discussed among the newer information technologies (Asongu et al., 2020; Jones et al., 2020; Stoll et al., 2019).

2.2. Blockchain

Blockchain is a relatively new data storage technology first introduced to the public in 2008 (Nakamoto, 2008). This publication also presents the first use case of blockchain technology with the presentation of Bitcoin software, a peer-to-peer electronic cash system. Blockchain technology has applications in various industries, such as FinTech, public services, healthcare, and private sectors where it can radically change business models, organization managements, supply chains, payment processes, security of data and even whole markets (Abbas, Martinetti, Moerman, Hamberg, & van Dongen, 2020; Beck, Müller-Bloch, & King, 2018; Du, Pan, Leidner, & Ying, 2019; Notheisen, Cholewa, & Shanmugam, 2017; Rimba et al., 2020; Schinckus, 2020; Schmidt & Wagner, 2019; Treiblmaier, 2018). In financial services, for example, blockchain technology is being used to ensure secure, decentralized and transparent transactions and to enable cryptocurrency exchanges (Chen & Bellavitis, 2020). By tracking and verifying the journey of goods, blockchain technology can also help provide transparency and reduce fraud in supply chain management, which can be critical in industries such as agriculture and healthcare (Engelhardt, 2017; Kamble, Gunasekaran, & Sharma, 2020). Another application is in the energy sector, where blockchain technology can manage renewable energy certificates and facilitate peer-to-peer energy trading (Andoni et al., 2019). Since the technology is organized in a decentralized peer-to-peer network, all stakeholders share equal access to information, such as records and transaction history, rather than relying solely on a single authority, such as a government or bank, for validation and recording (Ali, Ally, Clutterbuck, & Dwivedi, 2020; Karafiloski & Mishev, 2017; Schinckus, 2020). After being validated by the entire network, additional information (e.g. a new transaction of a good), is added by complementing a new block to the unalterable blockchain through a cryptographic process within a database that is public and can therefore be accessed by any stakeholder (Schinckus, 2020). The validation process is also open to any actor within the network. The underlying validation procedure determines which actor in the network gets permission to validate a new block. Proof-of-work (POW) and proof-of-stake (POS) are two main validation approaches (Schinckus, 2020).

Because of its characteristics as a general purpose technology, blockchain is often compared to the importance of the internet (Schmidt & Wagner, 2019). Whereas the aim of the internet is to connect people all over the world, the aim of blockchain is to diminish risks and reduce inefficiencies, insecurity and uncertainty among firms that exchange goods or services by providing transparency within transactions (Beck et al., 2018; Beck, Stenum Czepluch, Lollike, & Malone, 2016; Karafiloski & Mishev, 2017; Nærland, Müller-Bloch, Beck, & Palmund, 2017; Notheisen et al., 2017; Schmidt & Wagner, 2019).

Despite the increasing attention to and rapid development of blockchain technology, especially since the 2008 global financial crisis (Schinckus, 2020), there are also barriers to adoption and diffusion such as high development costs and technological limitations (Babich & Hilary, 2020; Schmidt & Wagner, 2019; Treiblmaier, 2018). In addition, regulatory uncertainty plays an important role for the diffusion of blockchain technologies. Public opinion and policy skepticism regarding the potentially harmful environmental and societal impacts may play a role in the speed of adoption and the development of novel applications by private companies (Gökalp, Gökalp, & Çoban, 2022). A central question is, therefore, to what extent blockchain technology is used in applications that contribute to sustainable business and consumer practices.

It is widely discussed whether blockchain technology consumes too much energy to ever result in sustainable applications (Stoll et al., 2019). In this context, there is a lot of criticism of the POW validation approach in particular. Within the POW approach, every cryptographic problem that needs to be solved for validation is sent to all actors within the network to ensure a decentralized and safe structure. However, as only one actor is allowed to validate a new block, all others working on the problem consume energy for nothing. As the POW approach favors very efficient and fast miners, many of them team up and form mining pools, which can mainly be found in countries where energy costs are lower (Schinckus, 2020). A negative consequence is that countries like China, where 65% of such mining pools can be found, even increase their consumption of non-environmentally friendly resources like coal.

Researchers predict that just because of the trading of Bitcoin, the global temperature might increase by 2 °C by 2034 (Mora et al., 2018). There is currently no alternative, including POS, that offers a similar or equal level of transparency and security as the POW approach (Schinckus, 2020). Yet, not all blockchain applications are as energy-intensive as mining Bitcoin. Use cases such as SolarCoin and VerdePay even have the potential to reduce carbon emissions (Howson, 2019).

On the other hand, blockchain technology has the potential to radically alter the way contracts and financial transactions are conducted, increasing efficiency as well as financial and operational performance. Furthermore, applications may also have a positive impact on the environment by improving the sustainability of existing processes (Schinckus, 2020; Schmidt & Wagner, 2019). Some blockchain technology applications facilitate technology efficiencies which in total results in lower energy consumption (Sharma, Kumar, & Park, 2020). In addition, smart contracts based on blockchain technology have been shown to enable the trading of carbon credits, ultimately reducing corporate emissions, as well as optimizing energy distribution and consumption by decentralizing energy markets (Aitzhan & Svetinovic, 2016). Applications of blockchain technology also include use cases such as managing the energy-intensive tracking of product flows along the supply chain, and verifying the origin of inputs (Howson, 2019). This ensures that ethical and sustainable practices are followed along the product journey (e.g. fair payment to producers), promotes responsible sourcing, and limits fraud by reducing information asymmetry (Christidis & Devetsikiotis, 2016; Kshetri, 2017; Saberi, Kouhizadeh, Sarkis, & Shen, 2019).

Furthermore, blockchain technology applications could make a major contribution to sustainability endeavors by impacting at least 14 out of the 17 United Nations SDGs (Schinckus, 2020; UN, 2022). Among many use cases, blockchain technology can empower communities and their networks through its creation of trust and transparency, improve food trust, facilitate more efficient water management, reduce electricity consumption and improve energy efficiency through establishing high credibility and reduce fraud through transparent and unchangeable records (Blakstad & Allen, 2018; Friedman & Ormiston, 2022; Hwang et al., 2017; Sanderson, 2018; Schinckus, 2020; Sikorski, Haughton, & Kraft, 2017; Treiblmaier & Beck, 2019). Table A.2 in the appendix provides a detailed overview of the 17 SDGs and includes websites of blockchain technology companies found through our web mining approach whose business models focus on a particular SDG.

2.3. Ecosystems and infrastructure

Research shows that local characteristics such as the innovation ecosystem and infrastructure impact the regional innovation performance by facilitating and contributing to the adoption and diffusion of new technologies (Asheim & Gertler, 2006; Gschnaidtner, Dehghan, Hottenrott, & Schwierzy, 2024; Williamson & Meyer, 2012). Access to employees and financial resources as well as agglomeration benefits from the co-location with other companies or universities are among the key elements of a local ecosystem conducive to innovation (Czarnitzki & Hottenrott, 2009; Feldman, 1994). Therefore, companies tend to locate in close proximity to similar companies in order to use the established social and professional links (Stuart & Sorenson, 2003). Benefits from being located close to key suppliers and customers can also bring competitive advantages which contribute to the existence of innovation clusters (Berkes & Gaetani, 2019; Berliant, Reed III, & Wang, 2006; Carlino, Chatterjee, & Hunt, 2007; Shearmur, 2012). Innovation research, hence, has long stressed the impact of location factors for facilitating collaboration between different actors and thereby allowing positive knowledge spillovers. Certain location factors therefore facilitate the generation of inventions, such as the presence of universities and the availability of a high-skilled labor pool. Other location factors, such as physical or financial infrastructure, drive the diffusion of new technology by providing a basis for knowledge flows and application opportunities (Feldman, 1994). In cases of nascent technologies such as blockchain, local knowledge spillovers could play an important role. On the other hand, digital and decentralized technology could also rely less on location factors compared to technologies that rely more directly on location factors.

Previous research has also stressed the role of transportation infrastructure in driving innovation as it facilitates the mobility of human capital and the flow of goods across locations (Agrawal, Galasso, & Oettl, 2017). Recent research indeed documents that the physical layout of cities in the US affects innovation by influencing the organization of knowledge exchange (Roche, 2020). Another study shows that upgrades to infrastructure have an important impact on innovation, suggesting that a new bridge between Malmö (Sweden) and Copenhagen (Denmark) had a significant effect on the number of patents per capita in Malmö through the attraction of highly qualified workers (Ejermo, Hussinger, Kalash, & Schubert, 2022). It remains, however, an open question whether such physical infrastructure plays a relevant role in the adoption and diffusion of blockchain technology.

It could further be argued that none of these location factors work in isolation but that their co-occurrence results in specific local ecosystems. These are characterized by rather static factors such as population structure and natural resources as well as by dynamic factors that stem from networks and exchange between actors within the ecosystem. Some prior research investigated the role of ecosystems in the context of sustainable technologies. Studies based on individual cases of selected regions (i.e., two regions in Finland), selected company types (i.e., multinational enterprises), or theoretical considerations suggest that business ecosystems also matter for a sustainable context (Nylund, Brem, & Agarwal, 2021; Sotarauta & Suvinen, 2019; Yang, Chen, Du, Lin, & Lu, 2021). Another study shows that local knowledge stocks matter by illustrating that in German regions where the existing stock of environmentally related patents is already high, the probability that a company develops or adopts sustainable innovations is significantly higher (Horbach, 2020). Geographic proximity to other innovators has also been show to accelerate the time-of-adoption of sustainable technologies (Losacker, Horbach, & Liefner, 2022).

Larger-scale systematic evidence is, however, still scarce and we know very little about digital technologies such as blockchain. This highlights a gap in our current understanding of the role of specific individual location factors and local networks for digital technologies. In particular, while the link between infrastructure, ecosystems, and innovation, in general, is quite established, it is less clear whether and how individual location factors and networks determine sustainable blockchain adoption. On the one hand, in the case of blockchain technology diffusion, it could be argued that location factors should matter less because of its digital and decentralized nature. In principle, its use should, therefore, be less dependent on specific characteristics of the location in which a blockchain-using company is located. On the other hand, recognizing opportunities of blockchain applications may require exposure to other users or even direct knowledge exchange between current and potential users, i.e. knowledge spillover through connectedness. Because of its complex nature and the lack of established standard applications, the ecosystem may play an even larger role for blockchain than for other technologies. In the case of sustainable blockchain applications, it may be even more crucial that entrepreneurs are exposed to sustainable business practices more generally, which makes the discovery of sustainable applications of blockchain technology more likely. While the results from earlier research suggest that local characteristics may be decisive for new technologies and innovation more generally, there is currently no evidence that this also applies to new digital and decentralized technologies, such as blockchain. Therefore, the following analysis aims to shed light on whether the local ecosystem, as reflected in location factors and networks, matters for companies adopting sustainable uses of blockchain technology.

3. Materials and methods

In the following, we present our data and methodology. First, we address our company database before we explain how our web-based indicators were generated. Lastly, we describe additional data we used to cover infrastructural and socio-economic variables.

3.1. Basedata

As base data for all of our analyses, we used the ORBIS company database (as of February 2022). ORBIS is a proprietary database compiled by Bureau van Dijk, in which company data from various national providers are harmonized to achieve an almost global coverage with over 400 million included companies. For our analyses, we extracted all companies that were incorporated in the US and also had their postal address and web address (URL) included. Furthermore, we removed all URL duplicates from our subsample so that each URL was unique. After this filtering, approximately 5.76 million companies remained in the dataset. The postal addresses of the companies were then used to perform a house number-accurate geocoding via the OpenStreetMap (OSM)-based service *Nominatim*.

As the dataset also included some economically inactive companies, we were only able to retrieve the websites of 3.72 million companies (64.5% of the URLs queried) using our web scraping approach (cf. Section 3.2). This corresponded to a coverage of about 61% of all economically active companies in the US according to United States Census Bureau (2021). A study in Germany has shown that the coverage there is 46%, although this can vary greatly depending on the industry, size, age and region studied (Kinne & Axenbeck, 2020).

3.2. Webdata

Building on the URLs contained in our company base data, we used the cloud-based web scraping tool webAI, developed by ISTARI.AI, to retrieve company websites and download their textual content. We followed a query logic in which the input URL of a corporate website is retrieved first, and then subwebpages are queried using a simple heuristic (Kinne & Axenbeck, 2020). First, all internal hyperlinks to the subwebpages are identified on the landing page and then queried in descending length (number of characters in URL) to download texts and identify further internal hyperlinks. Prioritizing shorter URLs generally leads to 'top-level' information being downloaded first, i.e. '*/products*' is downloaded before '*/news/2022/january*'. Following this logic, up to a maximum of 25 subwebpages per company website were processed and their texts downloaded.

3.2.1. Web-based blockchain indicator

To infer a company's blockchain capacity from its website texts, we trained an Natural Language Processing (NLP) model and represented the output as a firm-level *blockchain intensity score*. This approach has been implemented and tested by Gschnaidtner et al. (2024) in an application of blockchain detection in Germany, Austria and Switzerland. We understand blockchain capacities in this context as products and services with integrated blockchain technology or personnel with blockchain-related skills. Our indicator reflects how prominently the topic of blockchain is communicated by the company on its own website and how it is portrayed as essential to its own business model. We assume that companies that serve blockchain-oriented business areas or offer related products and services generally communicate this on their web presence. The more central this topic is for the company, the more significant it is for its external communication. For example, a startup for integrating blockchain into supply chains communicates almost exclusively on the topic of blockchain, while a company that offers 'blockchain consulting' among many other topics only communicates about this technology to a limited extent. Our NLP model was trained to distinguish between communication that is related to offering its own products and services with integrated blockchain and pure information dissemination. An example of the latter would be the website of a regional newspaper reporting that a local incubator for blockchain startups opened recently.

In a first step, the downloaded texts of each company were searched for text paragraphs that deal with the topic of blockchain. For this, we relied on a simple, but extensive keyword search (cf. Table A.1). In addition to manual research, frequently occurring words were extracted from an extensive corpus of academic discussion papers on blockchain.

Based on the millions of paragraphs found through the keyword search, a random sample of 3500 paragraphs was drawn. Then, each of these paragraphs was randomly assigned to three out of twelve briefed annotators, who labeled the paragraph as either 'information' or 'know-how'. The labeled data was then used to train a NLP model, based on state-of-the-art multilingual transformer models and a domain adaptation training strategy. The trained model exhibited an accuracy of 0.95 when tested only on examples where all three human annotators unanimously assigned a category. In an extended test dataset, which also included "disputed" paragraphs where there was disagreement among the human annotators (2:1 decisions), the model still achieved an accuracy of 0.72.

Using this model, we then classified all paragraphs that contained at least one of our blockchain keywords. In doing so, the model determined whether own blockchain know-how was reported or only information on the topic of blockchain was communicated. In the next step, we counted the number of paragraphs that the model evaluated as 'know-how' for each company website. We then related this number to the total amount of text content on the website, thus, determining a *blockchain intensity* for each company. This intensity would be 0.0 for a company completely without any blockchain-related text. For the consulting company example described above, on the other hand, the value could be 0.25. The aforementioned startup could have a blockchain intensity of 3.8. The regional newspaper, on the other hand, would have an intensity of 0.0 because its website texts only represent blockchain-related information and not its own know-how. Companies with a very high blockchain intensity (e.g. above the 90th percentile) are those companies that have focused their business model on blockchain and, accordingly, mainly communicate on blockchain-related topics. Examples of such particularly committed companies are those that offer seminars exclusively in the area of blockchain (blockchaintrainingalliance.com), crypto mining companies (corescientific.com), crypto trading platforms (nexo.com, stably.io) or blockchain frameworks (stellar.org). Accordingly, very high blockchain intensity values for such companies are a desirable phenomenon for us.

Unlike simpler, binary classifications (e.g., blockchain YES/NO), this continuous score with no upper limit allowed us to distinguish between companies where blockchain is only a marginal topic and those for which it plays a central role. Similar models have already been employed to study 3D printing diffusion (Schwierzy et al., 2022), AI diffusion (Dahlke et al., 2024), the effect of the COVID-19 pandemic on firms (Dörr, Kinne, Lenz, Licht, & Winker, 2022), and sustainability in the US metal industry (Schmidt et al., 2022).

Since some of the blockchain companies were only identified because they have integrated cryptocurrency-based payment systems into their online stores, we additionally used information on the tech stack of the company websites. For this, we captured as a boolean variable whether companies have integrated any of the over 300 "e-commerce" technologies (e.g. Woocommerce, Shopify) or crypto-based payment systems (e.g. Bitcoin) into their website's tech stack.

3.2.2. Web-based sustainability indicator

In order to identify companies that are engaged in sustainability, we developed an NLP model in the same way as the blockchain model described above. The resulting web-based sustainability indicator has already been used in a study on greenwashing in the US metal industry, which also explains in more detail the concrete use of keywords, step-by-step model development, and application to companies in the US (Schmidt et al., 2022). Sustainability here refers only to the ecological dimension, i.e., to concepts such as circular economy, the energy transition, ecological agriculture, regenerative energy, efficient use of resources, reduction of emissions, or recycling. As with blockchain companies, we assumed that firms active in these or related areas usually communicate this on their websites. The more central this topic is for the company, the more significant it is for the company's external communication.

For this approach, we also first used a simple keyword search, working with a list of potentially sustainability-related search terms (e.g. 'emission', 'organic', 'circular'). This list was developed together with experts from the OECD and included around 1000 words from more than 20 Indo-European languages. After a labeling and training process, which was implemented in analogy to the blockchain model described above, the trained model was used for the evaluation of all paragraphs containing keywords related to sustainability. However, in this case, no distinction was made between 'know-how' and 'information', but whether or not the topics were actually related to sustainability in the desired context. An example of this is the English word 'environment', which would be a positive hit in the sense of a 'natural environment', but not in the sense of an 'investment environment' or 'working environment'.

The paragraphs assigned to the category 'environmental sustainability' were again counted at the company level and normalized over the entire website text length into a *sustainability intensity*.

3.2.3. Location and web-based ecosystem mapping

In order to measure the embeddedness of the companies in ecosystems, we used a location-based and a web-based approach. As argued above, the local ecosystem may be an important driver of technology adoption through knowledge spillovers. Such spillovers, however, are often very local and require direct exchange between agents to facilitate the transfer of tacit knowledge (Rammer, Kinne, & Blind, 2020). For the location-based approach, we utilized the exact geocoding of the companies in our base dataset and determined the number of neighboring companies (1 km radius around the company's location) for each blockchain company. Additionally, we distinguished these neighboring companies according to their status as 'sustainability-engaged' and 'not sustainability-engaged'.

For the web-based approach, we built a hyperlink network for all approximately 3.72 million corporate websites, where the edges represented the linkage of one firm to its partners, to approximate their interconnectedness. Another company became a partner of a blockchain company if the other company had included a hyperlink to the investigated blockchain company on its own website or vice versa. Hyperlinks can be considered as the "basic structural element of the internet" (Park & Thelwall, 2003) and creating, maintaining, or removing a hyperlink "may be viewed as acts of association, non-association or dis-association, respectively" (Rogers, 2013). Several studies have shown the significance of hyperlinks for uncovering firms' network relations (Axenbeck & Breithaupt, 2021; Heimeriks & Van den Besselaar, 2006; Kinne & Axenbeck, 2020; Vaughan, Gao, & Kipp, 2006). Nevertheless, this approach does not show all possible connections in a company network, as companies do not necessarily mention all their partners on their website.

In addition, we calculated the number of sustainability-engaged partners per blockchain company, i.e. partners with a sustainability intensity greater than 0.0. We also calculated the mean value of the sustainability intensities of all partners of the blockchain company under investigation.

3.3. Infrastructure and socio-economic data

For the quantification of hard and soft location factors, we mainly used two data sources: official statistical data and OSM. The latter is a project founded in 2004 in the United Kingdom with the aim of producing freely available, open, worldwide geodata. It is one of the most important projects of Volunteered Geographic Information (VGI) (Neis & Zielstra, 2014). OSM data always consists of a geometry (point, linestring, polygon) and associated so-called tags. These are key–value pairs that represent the properties of an object, e.g. *amenity=restaurant*. For this study, we downloaded a dataset for the US from https://download.geofabrik.de/ and transformed it with the help of *osm2pgsql* into a PostgreSQL/PostGIS database, on the basis of which our subsequent calculations were carried out.

We obtained information on motorway links and airports from OSM. For the variables 'distance_motorway' and 'distance_airport,' we calculated the distance from each firm to the nearest respective feature using the PostGIS function *ST_Distance*. Due to calculation constraints, we set the maximum distance to 50 km (or 100 km in the case of airports). If no suitable OSM feature was found within this radius, the variable value was set to the respective maximum value of 50 or 100 km. Additionally, we aggregated several OSM features in order to derive information about location factors for the following three categories: transport infrastructure, leisure, and culture. For this, we defined a matching radius within which we searched for OSM features in our database. We chose a radius of 1 km which has been found empirically as a significant threshold of walkability (Liao, van den Berg, van Wesemael, & Arentze, 2020). We then counted all the OSM features with the corresponding tags within this radius. The respective OSM tags for each variable can be found in Table A.3.

Most of the socio-economic variables for our analysis were derived from a dataset published by the Federal Communications Commission (FCC). It includes information on unemployment and internet availability, amongst many other variables. Additionally, rent data, i.e. the Zillow Observed Rent Index (ZORI), was acquired from Zillow, which is the self-proclaimed most important marketplace for real estate in the US. All this data was then merged based on the FIPS code of the respective counties. In a next step, each company was assigned the respective values of the county in which it is located for each variable. An overview of all the used variables can be found in Table 1.

Table 2 shows some descriptive statistics for the main variables of interest. Table 3 shows the correlations between these variables. Overall, there were only modest correlations between the exploratory variables and stronger positive correlations between some location factors and the sustainability intensity.

4. Results

In the following, we present our main findings. First, we show descriptive statistics and the geographical distribution of blockchain companies. In the second step, we present the results of our regression analyses.

4.1. Descriptive statistics

In total, we identified 22,847 blockchain companies, i.e. companies with a blockchain intensity greater than 0.0. This represented just over 0.6% of the approximately 3.72 million companies we examined. Fig. 1 shows the histogram of the blockchain intensity scores for these 22,847 companies. From the distribution, it can be seen that most blockchain companies had indeed a low intensity: The median of blockchain intensity was 0.18 and the mean was 0.39 (standard deviation 0.47). Only five companies had a value above 4.0, including the official Ethereum blockchain website (ethereum.org).

We aggregated the blockchain companies at the county level in order to assess their spatial distribution in relation to the overall firm

Table 1
Overview of variables used in analysis.

Variable name	Description	Source	Measure
Sustainability intensity	web-based intensity sustainability engagement	ISTARI.AI	≥ 0.0
Blockchain intensity	web-based intensity of blockchain engagement		
Partners' Sustainability Intensity	mean sustainability intensity of linked partners		
E-commerce	e-commerce plugin on website		Boolean
Cryptopay	cryptocurrencies payment plugin on website		
# Sustainable companies (1km)	sustainable companies within 1 km		count
# Partners	number of hyperlinked partners		
Poverty (percent)	population in poverty	FCC	%
Unemployment (percent)	population without employment		
Food insecurity (percent)	population without reliable source of food		
Physical inactivity (percent)	non-physical leisure activity adults		
Adult obesity (percent)	obese adults		
Broadband access (percent)	population with Broadband access		
Distance motorway	distance to nearest motorway link	OSM	km
Distance airport	distance to nearest airport		
Transport (count)	weighted count of local public transport stops		count
Recreational (count)	recreational amenities within 1 km		
Cultural (count)	cultural amenities within 1 km		
Leisure (count)	leisure amenities within 1 km		
Rent (2022)	Zillow Observed Rent Index (ZORI)	Zillow	index

Table 2
Summary statistics.

	mean	standard deviation	minimum	maximum
Sustainability_intensity	0.201	0.499	0	5.185
Blockchain intensity	0.392	0.474	0.066	4.574
Partners' Sustainability Intensity	0.148	0.217	0	3.520
E-commerce	0.2356	0.424	0	1
Cryptopay	0.0047	0.069	0	1
# Sustainability-engaged companies	2.840	1.731	0	7.749
# Partners	2.029	1.102	0	10.736
ln(Employees)	1.946	1.212	0	10.878
Poverty (percent)	14.406	4.890	3.200	38.500
Unemployment (percent)	7.125	1.702	0.900	27.700
Food insecurity (percent)	14.552	3.330	4.000	29.000
Physical inactivity (percent)	20.185	4.461	9.200	39.700
Adult obesity (percent)	24.340	4.882	12.000	43.700
Broadband access (percent)	95.555	8.796	0	100.000
Distance motorway	4.144	7.802	<0.001	50
Distance airport	9.334	5.882	0.060	56.769
Transport (count)	57.064	150.301	0	1,110.600
Recreational (count)	23.126	43.780	0	1,190.000
Cultural (count)	3.609	8.554	0	86
Leisure (count)	39.725	96.038	0	591.000
Rent (2022)	2,299.662	824.182	549.063	8,033.910
Observations	19,491			

Table 3
Cross-correlation table.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Sustainability_int.	1.000																				
Blockchain int.	0.017	1.000																			
Partners' Sus. Int.	0.362	−0.025	1.000																		
# Sus. companies	0.063	0.013	0.057	1.000																	
# Partners	0.177	0.041	0.275	0.163	1.000																
ln(Employees)	0.087	0.001	0.055	0.106	0.184	1.000															
E-commerce	−0.042	−0.009	−0.001	−0.067	−0.111	−0.027	1.000														
Cryptopay	0.030	0.089	−0.005	−0.009	−0.019	−0.025	0.122	1.000													
Poverty	0.009	0.002	0.002	0.140	0.014	0.006	0.037	0.016	1.000												
Unemployment	−0.006	−0.001	−0.005	0.035	−0.009	0.012	0.039	0.014	0.543	1.000											
Food insecurity	0.005	0.016	0.005	0.148	0.023	0.002	0.023	0.017	0.805	0.424	1.000										
Physical inactivity	−0.027	−0.006	−0.049	−0.375	−0.092	−0.009	0.049	−0.004	0.305	0.203	0.219	1.000									
Adult obesity	−0.038	−0.010	−0.040	−0.535	−0.112	−0.033	0.056	−0.002	0.241	0.152	0.220	0.776	1.000								
Broadband access	0.016	0.008	0.008	0.271	0.055	0.032	−0.031	−0.014	−0.107	−0.031	−0.038	−0.265	−0.299	1.000							
Distance motorway	−0.011	−0.006	−0.009	−0.311	−0.050	−0.039	0.041	0.016	0.004	0.018	−0.054	0.172	0.203	−0.430	1.000						
Distance airport	−0.010	0.004	0.011	−0.080	0.005	−0.029	−0.000	0.003	−0.151	−0.110	−0.060	−0.089	−0.075	0.011	0.018	1.000					
Transport	0.048	0.013	0.043	0.678	0.131	0.086	−0.051	−0.010	0.173	0.053	0.152	−0.216	−0.426	0.141	−0.146	−0.142	1.000				
Recreational	0.013	−0.003	0.021	0.461	0.088	0.024	−0.033	−0.004	0.126	0.074	0.096	−0.196	−0.310	0.137	−0.127	−0.101	0.486	1.000			
Cultural	0.050	0.008	0.042	0.634	0.133	0.071	−0.036	−0.006	0.163	0.024	0.158	−0.207	−0.372	0.129	−0.129	−0.064	0.750	0.400	1.000		
Leisure	0.052	0.014	0.049	0.723	0.146	0.078	−0.050	−0.005	0.152	0.020	0.146	−0.265	−0.460	0.150	−0.146	−0.114	0.884	0.463	0.844	1.000	
Rent (2022)	0.023	0.008	0.029	0.496	0.096	0.046	−0.033	0.016	−0.157	0.027	−0.177	−0.539	−0.756	0.282	−0.179	−0.016	0.408	0.295	0.340	0.424	1.000

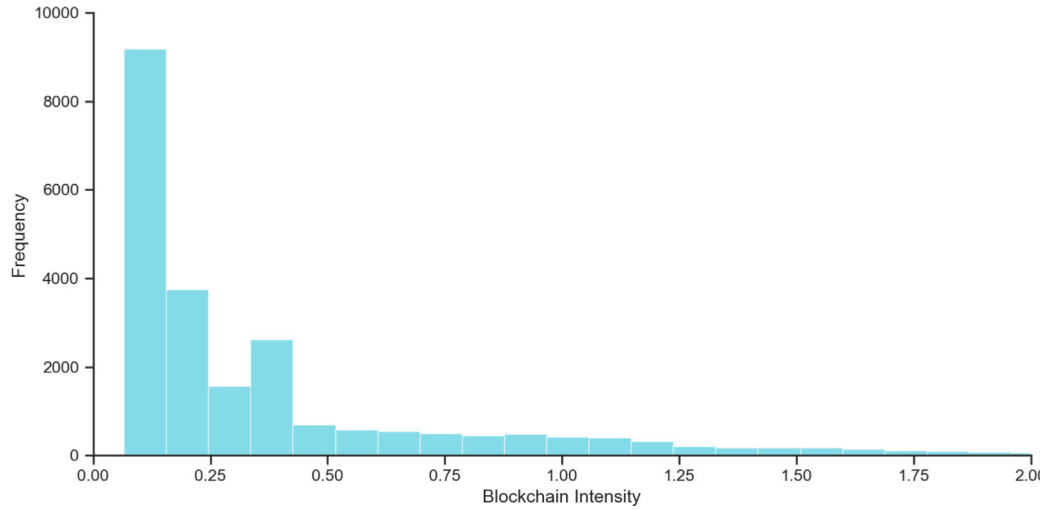

Fig. 1. Histogram of blockchain intensity for companies with blockchain intensity ≥ 0.00.

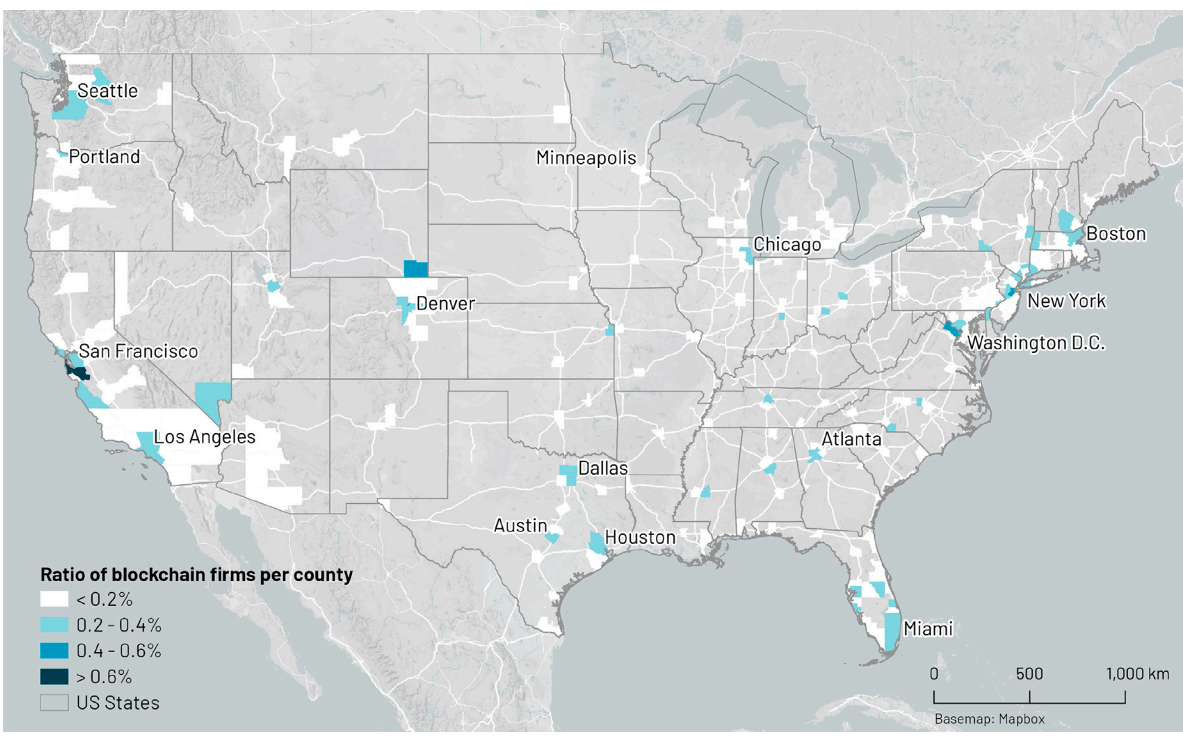

Fig. 2. Share of above-average (≥ 0.185) blockchain intensity firms in the overall firm population per county. Counties with less than 10 blockchain firms are excluded.

population in the US. Since many counties in the US are rather small and therefore contain few companies, we do not show counties with fewer than 10 blockchain companies in Fig. 2, which portrays the share of blockchain firms in the local firm population per county. For the remaining counties, we calculated the Global Moran's I statistic to check for spatial clustering in the geographic distribution.

The Global Moran's Index is a widely used indicator of spatial association and spatial autocorrelation which results in values between −1 and +1. The expected value of 0 indicates a random pattern. Negative values indicate negative spatial autocorrelation (a dispersed pattern), while positive values indicate positive spatial autocorrelation (a clustered pattern) (Ord & Getis, 1995). Like all measure of spatial autocorrelation, Moran's I requires hypotheses about the spatial relationships in the study area. The resulting spatial weights matrix is a formal representation of hypothesized interactions among spatial entities in the form of a weighting function. In the context of this study, the commonly used Queen contiguity method was used which only takes directly adjacent observations, i.e. neighboring administrative units, into account.

The I value of 0.38 (p-value: 0.001) indicated a significant and positive spatial autocorrelation, suggesting spatial clustering. Regions where blockchain seems to have higher relative importance were found in California (especially around San Francisco), on the East Coast (around Washington D.C., New York City and Boston) and in Florida (Miami, Orlando). Table 4 shows the ten counties and federal states that had the highest percentage of blockchain companies in the firm population and were also home to at least fifty blockchain companies. The two counties with the highest percentage were both located in California, which was also the state with the highest average percentage (excluding D.C.). Seven of the top ten counties were located on the East Coast, which was also reflected in the top ten states. The table additionally lists the top ten counties with the highest shares

Table 4
Top 10 list of counties and states with highest share of blockchain and sustainable companies in overall company population. Only counties with more than 50 blockchain firms are included.

	blockchain (state)	[%]	blockchain (county)	[%]	sustainability (county)	[%]
1	District of Columbia	1.23	San Francisco, CA	1.70	District of Columbia, DC	19.56
2	California	0.61	Santa Clara, CA	1.40	Boulder, CO	18.43
3	New York	0.60	New York, NY	1.30	Multnomah, OR	17.35
4	Delaware	0.59	San Mateo, CA	1.28	Chester, PA	15.86
5	Nevada	0.57	District of Columbia, DC	1.23	Marin, CA	15.74
6	Wyoming	0.57	Arlington, VA	1.16	Suffolk, MA	15.69
7	Virginia	0.56	Fairfax, VA	1.07	Denver, CO	15.52
8	Massachusetts	0.50	Loudoun, VA	1.03	Arlington, VA	15.18
9	Colorado	0.48	Suffolk, MA	0.99	Middlesex, MA	15.18
10	New Jersey	0.47	Middlesex, NJ	0.82	Alameda, CA	14.95

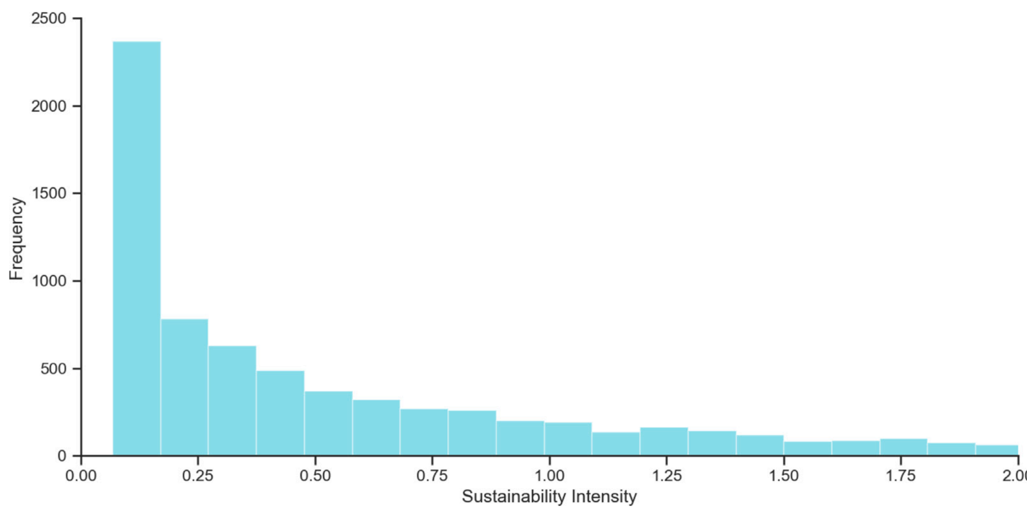

Fig. 3. Histogram of sustainability intensity for blockchain companies with sustainability intensity ≥ 0.00.

of sustainability-engaged companies. Three counties appeared in both categories (District of Columbia; Suffolk, MA; and Arlington, VA).

About 32.1% of the blockchain companies had a sustainability intensity greater than 0 and thus communicated a commitment to sustainability on their websites. Accordingly, this means that the proportion of sustainability-engaged blockchain companies was significantly higher than the 12.7% of sustainability-committed companies in the overall US company population. Fig. 3 shows the sustainability intensity distribution of these sustainability-engaged blockchain companies. Similar to the blockchain intensity, it can be seen that most companies had a low sustainability intensity. For the entire distribution, the mean was 0.21 and the median was 0.00. For the companies with sustainability intensity greater than 0 (7,351 companies), the mean was 0.65, the median was 0.35, and the standard deviation was 0.73. The highest score of 5.18 was achieved by a small company that describes itself as "the GREEN computer company".

Fig. 4 shows the scatterplots for sustainability intensity values of blockchain companies and selected location and ecosystem variables. Third-degree regression lines were also fitted to the data to illustrate the statistical relationships. This shows that primarily the ecosystem variables, especially the number of hyperlink partners (f) and their mean sustainability intensity (d), exhibited a strong correlation with the sustainability intensity of the blockchain companies. Likewise, the local ecosystem variables (b and c) showed a correlation with the sustainability intensity of the blockchain companies, albeit to a lesser extent. The striking inverted U shape of (e) was due to the fact that especially companies with only a single hyperlink partner showed a 100 % share of sustainable partners. Companies with such few partners were usually not sustainable, as (f) clearly shows. The remaining infrastructural and socio-economic location factors (g–l) showed hardly any correlation with the sustainability intensity of blockchain companies. It is important to keep in mind here, however, that we did not account for differences in the industry affiliation, company size, and the location of blockchain companies in these relationships. However, doing so was the purpose of the subsequent regression analyses.

Fig. 5 adds the dimension of sustainability to the mapping of blockchain firms. There were 18 counties (with more than ten blockchain firms) in which more than half of the blockchain companies were identified as sustainable, while only one county in Montana had 0% sustainable blockchain firms. The counties with the highest ratios were found in Vermont and Missouri. With 0.149, the Moran's I of sustainable blockchain firms was considerably lower than previously. Still, most of the areas with a high proportion of sustainability-engaged blockchain companies were identified on the East Coast.

Since we conducted our analyses at the company level, it was possible for us to make microgeographical statements on the topic of blockchain. We, therefore, want to illustrate the high granularity of our data using the example of Suffolk County in Massachusetts, one of the leading counties in both blockchain and sustainability. Fig. 6 shows the section of the county, where most blockchain companies were identified. The area corresponds to the central districts of Boston, particularly the Downtown and Back Bay areas. The companies were grouped into four categories based on their sustainability score divided into natural jenks.

Another important economic aspect in our analysis was the networking of companies. Fig. 7 shows a spatial representation of the identified online relationships. For this figure, we calculated the graph between a sample of blockchain firms and their sustainable partners. 29.8% (403,439) of the linked partners of the identified blockchain companies were sustainability-engaged. The average sustainability intensity of all blockchain company partners was 0.16. The map reveals

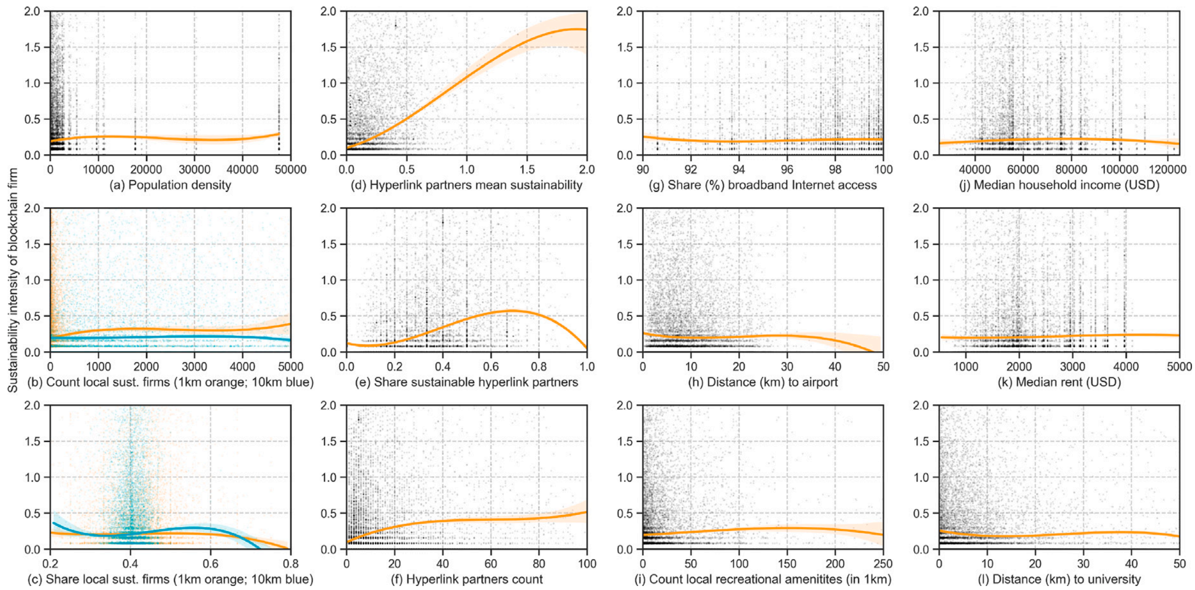

Fig. 4. Scatter plots and fitted regression lines of third order for sustainability intensity of blockchain companies and selected location and ecosystem factors.

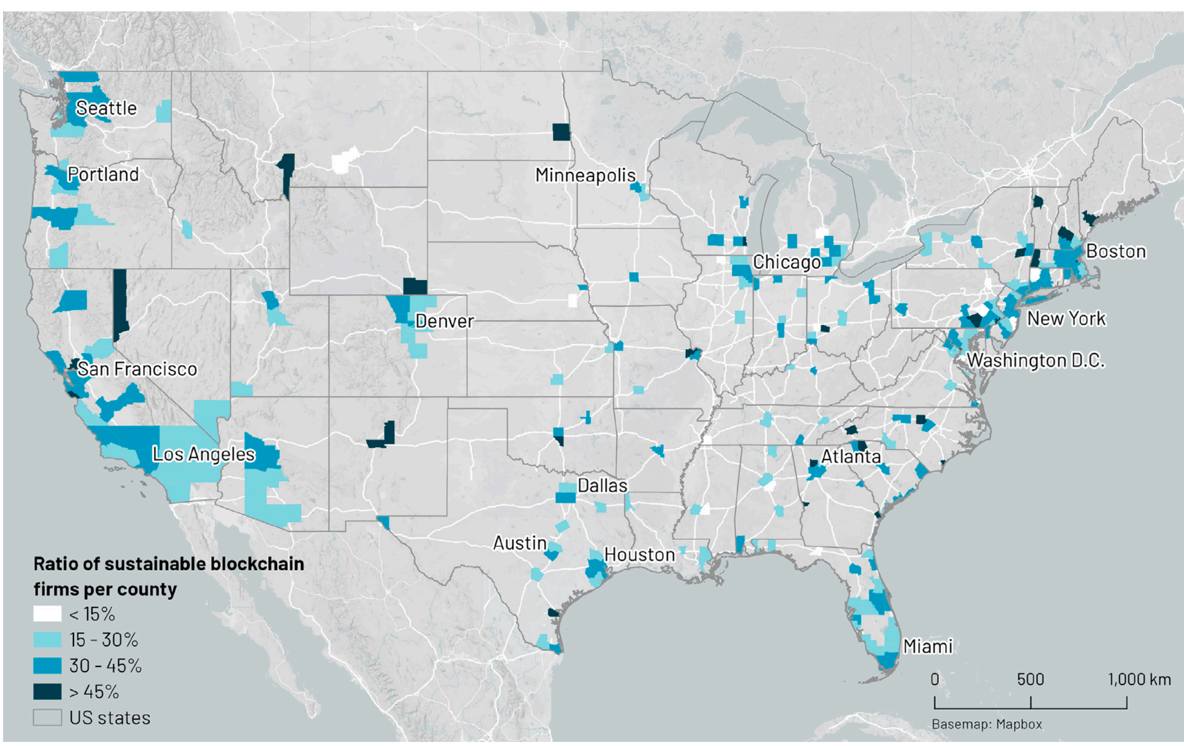

Fig. 5. Share of sustainable blockchain firms in overall firm population per county. Counties with less than 10 blockchain firms are excluded.

that there were strong connections across the entire US. In particular, the major agglomerations stood out here, with many edges falling on the Miami–Atlanta–Chicago, Washington D.C.–New York–Boston, and San Francisco–Portland–Seattle axes. However, there were also strong connections between West and East Coast, e.g. between Los Angeles and New York. Strikingly, there were very few connections in the northern border regions with Canada and in parts of the west central US.

4.2. Regression analysis and results

We estimated Ordinary Least Squares (OLS) models to identify the multivariate links between regional factors and a blockchain firm's sustainability intensity, while accounting for the sector, firm size and blockchain intensity. The sustainability score was used as dependent variable in each model. We also included state fixed effects in all models to account for differences in state-level predictors, e.g. varying environmental regulations across states. Due to missing values in the industry affiliation information and the number of employees in the ORBIS data base, the regression sample consisted of 19,491 unique companies.

Specification (1) presented in Table 5 included the main predictors capturing the local network links to sustainability-engaged companies, the number of hyperlinked partners, and the overall number of partners. Moreover, we included the indicators for whether a company

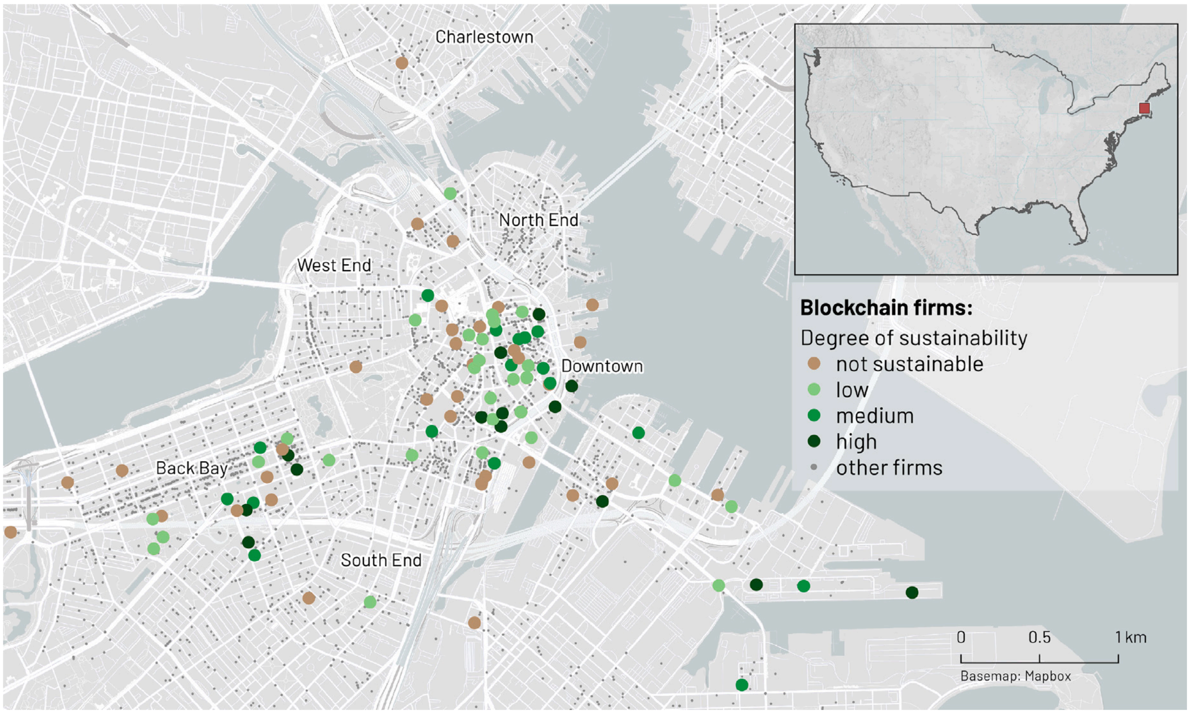

Fig. 6. Exemplary zoom-in map of blockchain firm locations in Suffolk County, MA (i.e. Boston).

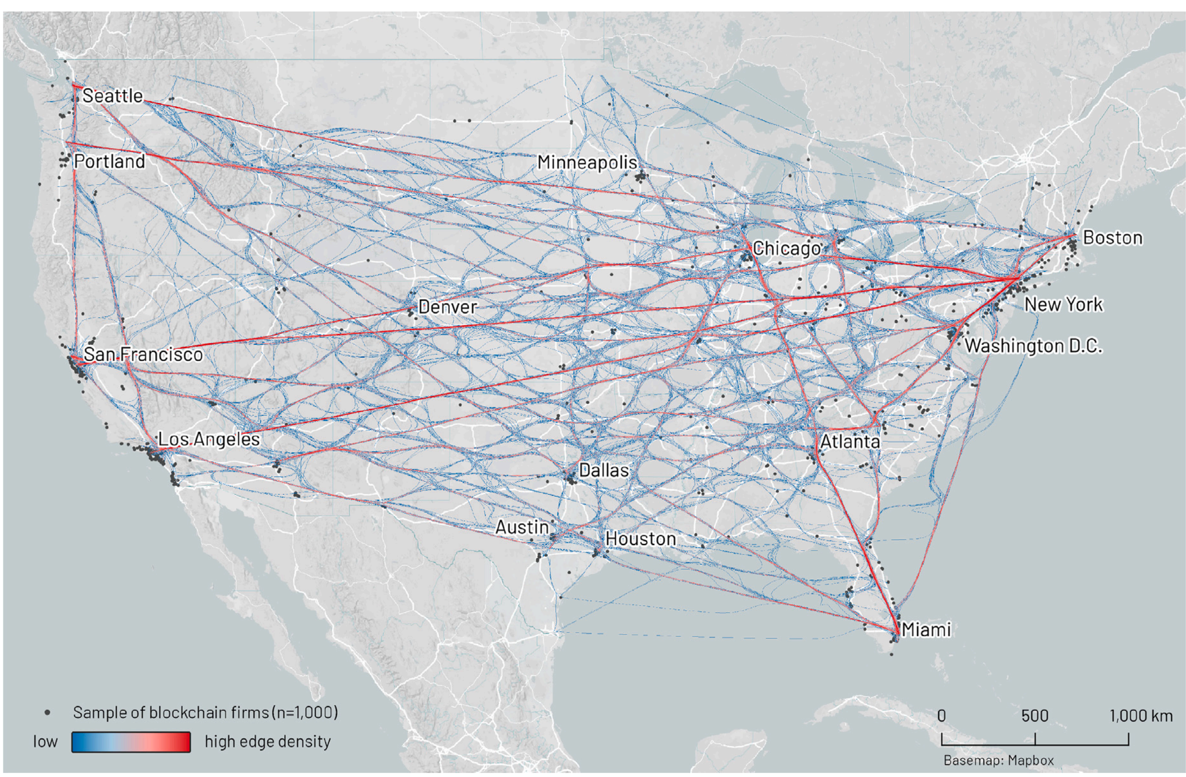

Fig. 7. Exemplary hyperlink network between a random sample of 1000 blockchain firms and their sustainable partner firms. Edge bundling technique has been applied in order to show high (red) and low (blue) density aggregated connections.

is active in e-commerce or simply offers crypto-pay options on their website. These predictors alone explained about 14% of the variance in the sustainability score. In all models, we additionally controlled for company size measured by the number of employees since larger companies generally scored higher on sustainability. The local network indicators were positive and statistically significant, indicating that being embedded in local ecosystems was related to a higher degree of sustainability commitment for blockchain companies. The higher the embeddedness of the company as measured by the number of partners, the higher was the firm's own sustainability intensity. One additional partner was associated approximately with a 0.3 increase in the sustainability score. Particularly the partners' sustainability intensity appeared

Table 5
Regression results for dependent variable: Blockchain companies' sustainability intensity.

	Ordinary Least Squares					Lewbel IV
	(1)	(2)	(3)	(4)	(5)	(6)
Partners' Sustainability Intensity	0.783***	0.785***	0.758***	0.797***	0.796***	2.365***
	(0.037)	(0.037)	(0.036)	(0.065)	(0.064)	(0.154)
Partners' Sustainability Intensity^2				−0.034	−0.033	−0.805***
				(0.068)	(0.068)	(0.218)
# Sustainable companies (1km)	0.007***	0.007***	0.009***	0.009***	0.010***	0.007
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.006)
# Partners	0.029***	0.029***	0.032***	0.031***	0.031***	−0.041***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.007)
ln(Employees)	0.005	0.005	0.006	0.006	0.006	−0.010
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)
ln(Employees)^2	0.003*	0.003*	0.002*	0.002*	0.002*	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
E-commerce	−0.042***	−0.041***	−0.047***	−0.048***	−0.048***	−0.078***
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)	(0.009)
Cryptopay	0.280***	0.267***	0.241***	0.242***	0.245***	0.286***
	(0.073)	(0.073)	(0.074)	(0.074)	(0.075)	(0.082)
Blockchain intensity		0.020**	0.021**	0.021**	0.021**	0.035***
		(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Poverty (percent)					0.001	
					(0.001)	
Unemployment (percent)					−0.004	
					(0.003)	
Food insecurity (percent)					−0.001	
					(0.002)	
Physical inactivity (percent)					0.001	
					(0.002)	
Adult obesity (percent)					−0.004**	
					(0.002)	
Broadband access (percent)					0.000	
					(0.000)	
Distance motorway					0.000	
					(0.000)	
Distance airport					−0.001	
					(0.001)	
Transport (count)					0.000	
					(0.000)	
Recreational (count)					−0.000***	
					(0.000)	
Cultural (count)					0.001	
					(0.001)	
Leisure (count)					−0.000	
					(0.000)	
Rent (2022)					−0.000*	
					(0.000)	
R^2	0.144	0.144	0.164	0.164	0.165	0.151
Industry Fixed Effects	No	No	Yes	Yes	Yes	Yes
State Fixed Effects	No	No	Yes	Yes	Yes	Yes

n = 19,491
All models contain a constant. Robust standard errors in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

to matter while controlling for the number of partners. A one-unit increase in the partner's sustainability intensity was associated with a 0.78 unit increase in the firm's own sustainability intensity.

Adding the blockchain intensity to the model in specification (2), we found a statistically significant correlation between a company's blockchain intensity and its sustainability score. Specification (3) was identical to the previous specification except that we now accounted for the sector of activity and the state. This increased the R^2 only marginally to 16.4% indicating that the previously included factors were of comparably higher importance or captured already much of the sector and state variation. Specification (4) also contained the second-order term of the average partner sustainability. Its coefficient turned out to be negative but statistically insignificant, indicating a mainly positive link to the degree of sustainability. Specification (5) further included the large set of county-level characteristics related to knowledge, accessibility, poverty, health, and quality of life as shown in Table 1. Adding these factors did, however, not increase the explanatory power of the model by much, since most factors were statistically insignificant. Importantly, the insights for the sustainable ecosystem measures remained unchanged through the inclusion of these additional regressors. It was striking that hardly any of the general location characteristics explained blockchain use in our study. While this finding may be less surprising for health, cultural, general economic, or leisure characteristics, it was not evident that transport infrastructure or digital infrastructure did not seem to matter after controlling for direct network factors. This could mean that the networks were driven by those factors and have no additional role besides this. Alternatively, these factors did indeed not matter in the specific case of sustainable blockchain use. These findings also confirmed the descriptive patterns shown in Fig. 4 and highlighted the role of network-related factors as compared to more general location factors.

In order to go beyond the analysis of correlation, we further estimated an instrumental variable (IV) model which allowed us to address the potential endogeneity of a company's partner's sustainability intensity. Moreover, some unobserved common drivers could explain both the partner's sustainability as well as the sustainability of the company itself. The number of sustainability-engaged companies in a company's neighborhood and the overall number of partners could also be endogenous. To address such endogeneity concerns, we would need to instrument the Partners' Sustainability Intensity (as well as the second-order term), the number of sustainability-engaged companies, and the total number of partners with instrumental variables that explained the endogenous measures, but not the sustainability commitment itself. Since such variables are typically hard to find, we employed a heteroscedasticity-based approach that generates suitable variables from the data (Lewbel, 2012). The advantage of this method is that it allows estimating causal coefficients for the main variables of interest if no external instruments are available (Baum & Lewbel, 2019). The approach requires a few assumptions to hold: First, that the model is linear, and second that there are indeed endogenous variables, i.e. unobserved factors affect both the explanatory measures and the primary dependent variables. We argue that this was indeed the case for our study. Ultimately, we may also assess whether the third condition is satisfied, which would indicate the presence of heteroscedasticity in the data. Note that these assumptions are not necessary for validity of the estimator but result in a higher confidence that the estimator is consistent (Baum & Lewbel, 2019). The White/Koenker test of homoscedasticity can be rejected for the Partners' Sustainability Intensity ($\chi^2 = 18.8$), the number of sustainability-engaged companies ($\chi^2 = 56.6$), as well as for the total number of partners ($\chi^2 = 269.6$).

We present the results from this Lewbel-IV model in column 6 of Table 5 using the same variables as in specification (4). The F-test (Cragg-Donald Wald F statistic) of excluded instruments was 19.96 and hence exceeded the critical value for 5% maximal IV relative bias (Stock & Yogo, 2005). This statistic shows that the excluded instruments were indeed relevant (i.e. jointly significant) for explaining the first stage on the IV model. The results from the IV model confirmed the positive link between partner sustainability intensity and a company's own sustainability intensity. However, the squared term was now statistically significant, indicating a non-linear relationship saturated at higher partner intensities. The coefficient for the number of sustainability-engaged companies in the same location was still positive but no longer statistically significant at the 10% level. For the number of partners, we found a negative sign, stressing the role of partners with sustainability expertise in the network rather than the pure size of the network. We re-estimated the models with a binary dependent variable as a final robustness test to address potential concerns regarding the uncertainty in measuring the sustainability intensity. With this additional approach, we addressed the concern that our sustainability score may not fully indicate the companies' degree of sustainability. However, the score may still allow us to classify companies into ones with a score larger than zero and others. Linear probability models (LPM) on binary outcomes further reduced the likelihood that some high scores could drive the results and that some companies may have low scores although they were actually more sustainable than those with higher scores. Table 6 shows the results, which are in line with the ones for the continuous score in terms of the main explanation factors. The main difference was that the coefficient estimates of the IV model (specification 6) in terms of size and significance were closer to the correlational models (specifications 1–5).

5. Discussion

This study provides initial insights into the diffusion of blockchain technologies in the US and the connection between blockchain and sustainability commitment in companies' external communication. For the entire US firm population (**RQ1** & **RQ2**), we found that blockchain is still a niche technology. While most blockchain companies showed a sustainability intensity score of zero and, accordingly, no demonstrated commitment to sustainability, about one in three blockchain companies could be classified as pursuing at least some sustainable activities. This was a much higher number than in the overall firm population. Based on a spatial analysis as well as a multivariate regression model at the company level, we found that the ecosystem measures derived from neighboring firms and hyperlink networks played a crucial role in the sustainability activities of blockchain companies (**RQ3**). In line with research on (local) knowledge spillovers (Czarnitzki & Hottenrott, 2009; Feldman, 1994; Rammer et al., 2020; Roche, 2020), we found that both being connected to other sustainability-engaged companies and being in close geographic proximity to a large number of other sustainability-engaged companies were key predictors of the sustainability degree of blockchain companies. The more sustainable a blockchain company's network partners were, the higher was its own sustainability score.

We also found that blockchain companies with a particularly strong focus on this technology tended to emphasize their commitment to sustainability and present it as central to their business model. This was possibly because such companies are more aware of the negative environmental impact and, therefore, address these consequences at least superficially with a high level of awareness. On the other hand, it is also possible that there are many application areas for blockchain technology in the area of sustainability that are addressed by companies with a strong blockchain focus. However, further research is needed to fathom this.

We observed a statistically significant, negative relation between the use of e-commerce plugins and the sustainability intensity, indicating that companies operating blockchain-based e-commerce were less focused on sustainability. We introduced this control variable to account, at least partially, for companies that use blockchain technology only as a means of payment in their online store. For the same reason, we also introduced another control variable for the use of dedicated cryptopay website plugins. However, this variable showed a significant positive impact on the sustainability intensity of blockchain companies, which is counter-intuitive at first. However, such plugins are still very rare, being used by less than 0.5% of all blockchain companies. Thus, this specific website technology could be more indicative of a particular type of blockchain company that is more concerned with sustainability than the average.

With respect to **RQ3**, infrastructural and socio-economic location factors seemed rather irrelevant for the sustainability engagement, but our results regarding the ecosystem embeddedness of blockchain companies suggested interesting correlations. More sustainability-engaged blockchain companies tended to be located in high-density areas in terms of overall and sustainability-engaged firm counts. They also tended to have more hyperlinked partners, which were more committed

Table 6
LPM regression results for dependent variable: Blockchain companies' sustainability probability (robustness test).

	Ordinary Least Squares					Lewbel IV
	(1)	(2)	(3)	(4)	(5)	(6)
Partners' Sustainability Intensity	0.432*** (0.018)	0.433*** (0.018)	0.410*** (0.018)	0.554*** (0.030)	0.555*** (0.030)	0.715*** (0.070)
Partners' Sustainability Intensity^2				−0.127*** (0.023)	−0.127*** (0.023)	−0.207*** (0.037)
# Sustainability-engaged companies	0.010*** (0.002)	0.010*** (0.002)	0.011*** (0.002)	0.011*** (0.002)	0.010*** (0.003)	0.015*** (0.005)
# Partners	0.062*** (0.003)	0.061*** (0.003)	0.064*** (0.003)	0.058*** (0.003)	0.058*** (0.003)	0.041*** (0.015)
ln(Employees)	0.002 (0.008)	0.002 (0.008)	0.006 (0.008)	0.005 (0.008)	0.005 (0.008)	−0.002 (0.008)
ln(Employees)^2	0.003** (0.001)	0.003** (0.001)	0.002* (0.001)	0.002* (0.001)	0.002* (0.001)	0.003*** (0.001)
E-commerce	−0.071*** (0.007)	−0.071*** (0.007)	−0.074*** (0.007)	−0.078*** (0.007)	−0.078*** (0.007)	−0.082*** (0.008)
Cryptopay	0.308*** (0.051)	0.306*** (0.052)	0.283*** (0.051)	0.286*** (0.052)	0.287*** (0.052)	0.303*** (0.051)
Blockchain intensity		0.003 (0.007)	0.003 (0.007)	0.004 (0.007)	0.004 (0.007)	0.006 (0.007)
Poverty (percent)					0.001 (0.001)	
Unemployment (percent)					−0.004 (0.003)	
Food insecurity (percent)					0.000 (0.002)	
Physical inactivity (percent)					−0.003 (0.002)	
Adult obesity (percent)					0.001 (0.002)	
Broadband access (percent)					0.001 (0.000)	
Distance motorway					−0.000 (0.000)	
Distance airport					0.000 (0.001)	
Transport (count)					−0.000 (0.000)	
Recreational (count)					−0.000 (0.000)	
Cultural (count)					0.000 (0.001)	
Leisure (count)					−0.000 (0.000)	
Rent (2022)					0.000 (0.000)	
R^2	0.098	0.098	0.116	0.117	0.118	0.104
Industry Fixed Effects	No	No	Yes	Yes	Yes	No
State Fixed Effects	No	No	Yes	Yes	Yes	Yes

n = 19,491.
All models contain a constant. Robust standard errors in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

to sustainability on average. These findings could already be observed in the descriptive statistics and were additionally confirmed by our regression analyses. We also found a positive correlation between the sustainability intensity of a blockchain company and its immediate geographic neighborhood. Here, many neighbors, and especially many sustainable neighbors, seemed to be related to the company's own level of sustainability. This could potentially be related to latent or deliberate spillovers. However, the topic of greenwashing, i.e. lip service about sustainability engagement, should also be mentioned here, which we were unable to address properly with our data basis. This distorted portrayal of sustainability engagement may have had an influence on all our analyses.

Table A.1
List of blockchain-related keywords.

Keyword		
aave	decentralized applications	pegged currency
altcoin	decentralized autonomous applications	permissioned ledger
altcoins	decentralized autonomous organization	proof of stake
atokens	decentralized exchange	proof-of-authority
binance	decentralized exchanges	proof-of-stake
bitcoin	decentralized finance	proof-of-work
blockchain	devcon	siacoin
byzantine fault	dexes	sidechain
cbdc	distributed ledger technology	smart contract
central bank digital currency	dogecoin	smart contracts
chainlink	dyor	smart legal contract
coinbase	eclipse attack	smart-contract
consensus algorithm	eosio	smart-contracts
consensus mechanism	erc-20	stablecoin
crypto	erc-721	stablecoins
cryptoassets	erc20	tezos
cryptocurrencies	ethereum	tokenomics
cryptocurrency	etherscan	total-value-locked
cryptoeconomics	genesis block	transaction block
cryptojacking	governance tokens	uniswap
ctokens	gwei	unspent transaction output
dappradar	hyperledger	usd coin
dapps	initial coin offering	usdc
de-fi	ipfs	usdt
decentralised applications	litecoin	utility tokens
decentralised autonomous applications	makerdao	utreexo
decentralised autonomous organization	mimblewimble	utxo
decentralised exchange	mining pool	ytokens
decentralised exchanges	multichain	zero-knowledge-proof
decentralised finance	non-fungible tokens	zk-snarks

All these results and interpretations must, of course, be understood as mere fact-finding, and correlations can, at most, be indications of causal relationships. However, a continuous update of our dataset will enable econometric time series analyses in the future, which may also be able to identify causal relationships. Such results would then be of particular value for evidence-based policy decisions, so that the use of a high-potential technology such as blockchain could be steered in a long-term sustainable direction.

Despite achieving unprecedented coverage, our purely web data-based approach still has limitations, since it is dependent on the content that is written and communicated on corporate websites. Therefore, relevant companies that do not communicate about sustainability or blockchain on their website, or partners of companies that are not mentioned in the form of hyperlinks, cannot be identified by our methodology. This means that we were unable to intercept a correlation concerning the entire US firm population, e.g. a general relation between sustainability commitment and company network characteristics, in our analysis. However, similar limitations also exist with traditional approaches such as surveys or patent analyses. Depending on the industry and geography, patent data only has very limited informative value if little or no patenting takes place (or can take place) there. This applies to the entire software industry in Europe, for example, where it is difficult or impossible to register patents on software. Surveys, on the other hand, can usually only be carried out on an extremely limited scale and results must then be extrapolated, which may produce correct (i.e. unbiased) results in the aggregate, but are inadequate at the individual level and especially in the context of microgeographic analyses, where complete information must also be available at the smallest spatial level.

In the best case, of course, data from several sources based on different data collection methods should be combined in order to achieve robust results. However, there is currently no alternative data available on sustainability-engaged blockchain companies in the US, so we inevitably had to focus on a purely web data-based approach in this study. Previous studies have shown that purely web data-based approaches can be used to achieve robust results: (Kinne & Lenz, 2021) has shown that information from high-quality but limited surveys can be extrapolated to the entire company population to generate valuable microdata. Furthermore, Mirtsch, Kinne, and Blind (2021) has demonstrated that information extracted from websites can be matched with information from proprietary databases for the use of standards, but can also provide additional information. Dörr, Kinne, Lenz, Licht, and Winker (2022) has shown that a web data-based approach can provide information on company performance that anticipates movements in individual company creditworthiness for more than a year. Of particular relevance is the study by Kinne and Axenbeck (2020) that uncovers that the (digital-savvy) companies examined in this study can be very well covered by web data-based approaches.

As blockchain is a rather procedural technology, it is conceivable that many companies use it but do not necessarily communicate it as central to their business model. The low blockchain intensities we found could also be influenced by this. Future research could combine our data with other databases that use a different approach to analyze companies to mitigate such biases (e.g. official corporate sustainability reports). A textual analysis and subsequent quantitative evaluation of such reports could further increase the informative value of this study. However, the best methodological approach for comprehensive data extraction from such reports still needs to be investigated.

Regarding our use of OSM data for operationalizing infrastructural location factors, we have to keep in mind that OSM data quality can vary greatly by region (Sehra, Singh, & Rai, 2014), which may impact our results, particularly when comparing companies in rural and urban areas. We also attempted to account for an array of soft location factors. However, there are some that we could not cover purely through OSM

Table A.2
Mapping of Sustainable Development Goals and exemplary use of blockchain technology.

SDG	Main goal (UN, 2022)	Related literature	Exemplary website (accessed June 2024)	Application
SDG 1	End poverty in all its forms everywhere	Schinckus (2020), Treiblmaier and Beck (2019)	www.skuchain.com www.stellar.org	Inventory tracking, provenance trail, decentralized blockchain network enabling low-cost payments
SDG 2	End hunger, achieve food security and improved nutrition and promote sustainable agriculture	Schinckus (2020)	www.skuchain.com www.damcogroup.com www.nisum.com	Supply chain tracking, provenance trail
SDG 3	Ensure healthy lives and promote well-being for all at all ages	Schinckus (2020)	www.damcogroup.com	Run of clinical trials based on smart contracts
SDG 4	Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all	–	www.stellar.org	Public blockchain open for developers
SDG 5	Achieve gender equality and empower all women and girls	–	www.stellar.org	Public blockchain open for developers, creating equitable access to the global financial system
SDG 6	Ensure availability and sustainable management of water and sanitation for all	Treiblmaier and Beck (2019)	www.waste2wear.com	Supply chain tracking, provenance trail tracking water usage in textile production process
SDG 7	Ensure access to affordable, reliable, sustainable and modern energy for all	Blakstad and Allen (2018), Hwang et al. (2017), Sanderson (2018), Schinckus (2020), Sikorski et al. (2017)	www.waste2wear.com www.kleangas.com	Supply chain tracking, provenance trail tracking water usage in textile production process
SDG 8	Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all	Blakstad and Allen (2018), Treiblmaier and Beck (2019)	www.10pearls.com www.stellar.org	Enable growth in emerging countries through services such as blockchain consulting
SDG 9	Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation	Schinckus (2020)	www.stellar.org	Build resilient infrastructure through open-source blockchain network
SDG 10	Reduce inequality within and among countries	Blakstad and Allen (2018), Schinckus (2020), Treiblmaier and Beck (2019)	www.10pearls.com	Support emerging countries by offering jobs in blockchain consulting
SDG 11	Make cities and human settlements inclusive, safe, resilient and sustainable	Schinckus (2020), Treiblmaier and Beck (2019)	www.waste2wear.com	Supply chain tracking, provenance trail, realtime industry data
SDG 12	Ensure sustainable consumption and production patterns	Schinckus (2020)	www.waste2wear.com www.kleangas.com	Supply chain tracking and provenance trail tracking used to recycle plastic
SDG 13	Take urgent action to combat climate change and its impacts	Blakstad and Allen (2018), Hwang et al. (2017), Sanderson (2018), Schinckus (2020), Sikorski et al. (2017)	www.skuchain.com www.waste2wear.com www.energyweb.org www.nori.com www.kleangas.com	Inventory control, real-time industry data, supply chain tracking
SDG 14	Conserve and sustainably use the oceans, seas and marine resources for sustainable development	Schinckus (2020)	www.waste2wear.com	Supply chain tracking and provenance trail tracking used to recycle plastic
SDG 15	Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss	Schinckus (2020)	www.waste2wear.com www.nori.com www.kleangas.com	Marketplace transparency facilitating carbon removals

(*continued on next page*)

Table A.2 (*continued*).

SDG	Main goal (UN, 2022)	Related literature	Exemplary website (accessed June 2024)	Application
SDG 16	Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels	–	www.skuchain.com www.stellar.org www.damcogroup.com	Public blockchain open for developers and blockchain development services
SDG 17	Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development	Blakstad and Allen (2018), Hwang et al. (2017), Sanderson (2018), Schinckus (2020), Sikorski et al. (2017)	www.waste2wear.com www.nori.com	Enhance partnerships through blockchain facilitated market places and more transparency in the supply chain

Table A.3
List of OSM tags used to extract relevant features for location factor operationalization.

Location factor	OSM tag	Location factor	OSM tag
transport count	highway=bus_stop amenity=bus_station railway=tram_stop railway=stop railway=station	leisure count	amenity=bar amenity=cafe amenity=fast_food amenity=pub amenity=restaurant amenity=nightclub
cultural count	amenity=cinema tourism=museum tourism=gallery amenity=theatre amenity=arts_centre building=church building=mosque building=synagogue building=temple	recreational count	leisure=park leisure=garden leisure=nature_reserve leisure=playground leisure=pitch leisure=stadium leisure=fitness_centre leisure=sports_centre leisure=swimming_pool leisure=golf_course
airport	aeroway=aerodrome	highway	highway=motorway_link
university	amenity=university		

data, e.g the perception of a location in terms of safety, which could also be a key location factor. To evaluate this would be possible, for example, by taking social media data into account (Santos, Silva, Ferreira Loureiro, & Villas, 2018). We obtained other soft location factors from official statistics data, most of which do not have the same high spatial resolution as our other location factors and could therefore also lead to distortions in our results.

Moreover, advances in machine learning and AI could be used to further increase model quality, including specific question answering based on RAG (Retrieval-augmented Generation) to identify, classify, and describe e.g. blockchain-related products from companies.

The results of this study could be seen as a starting point for potential policy implications. The encouragement of business collaboration stands out as particularly important for the promotion of sustainability and blockchain utilization. Consequently, it may be beneficial for policymakers to consider the development of initiatives that foster stronger networks and partnerships among blockchain companies, with a particular focus on those with a sustainability-engaged approach. This could entail the establishment of regional innovation and acceleration hubs, the investment of resources in local infrastructure, the provision of (financial) incentives for companies to establish or expand their operations in areas with high entrepreneurial density, and the support of platforms that facilitate knowledge sharing and collaboration. These platforms could bring stakeholders such as academic institutions, investors, companies, and non-governmental organizations together. Hence, policy interventions may be more impactful if they prioritize network-building over purely geographical considerations.

6. Conclusion

In this study, we presented a novel approach to identify sustainability-engaged blockchain companies. Furthermore, we correlated their sustainability commitment levels with location factors and their ecosystem embeddedness. For this, we used a large-scale web scraping approach to analyze the websites of all US companies via NLP and deep learning and also captured the hyperlink network between these websites. Our results showed that blockchain remains a niche technology (**RQ1**), with its use communicated by 22,847 companies (0.6% of all US companies). Of these blockchain companies, 32.1% were classified by our language models as having a commitment to sustainability (**RQ2**), which was much higher than in the overall firm population, suggesting that sustainability plays a more important role for blockchain companies. We were also able to identify regions where there are particularly many blockchain companies, especially in California and on the East Coast.

Our regression models showed that blockchain companies with an intensified focus on sustainability had, at least quantitatively, a more intensive embedding in entrepreneurial ecosystems, while infrastructural and socio-economic location factors hardly played a role (**RQ3**). Thus, these companies had more direct hyperlink partners which were more focused on sustainability themselves. In addition, more sustainability-engaged blockchain companies were located in regions with a high density of companies and within one kilometer of many other sustainability-engaged companies.

We interpreted these results as indicative of the high relevance of entrepreneurial ecosystem embedding for the sustainable adoption of the novel blockchain technology. We discussed local (knowledge) spillovers as possible drivers, but also learning, inspiration and imitation in the wider partner network. However, we also pointed out that our results might only be indications of causal relationships that need to be explored in future studies using our data in the form of time series analyses.

CRediT authorship contribution statement

Jan Kinne: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Robert Dehghan:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Sebastian Schmidt:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **David Lenz:** Validation, Software, Resources, Methodology, Investigation, Data curation. **Hanna Hottenrott:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The authors Jan Kinne, David Lenz and Robert Dehghan founded and are employed by ISTARI.AI. Sebastian Schmidt is also employee of ISTARI.AI. ISTARI.AI develops webAI, an artificial intelligence and scraping software for large-scale automated analysis of company related web data. This does not alter our commitment to sharing data and materials.

Acknowledgments

We would like to thank Jannik Reißfelder, Kerem Cerit, André Meynioglu, Pol Puig (ISTARI.AI), Julian Schwierzy, Elisa Rodepeter, Andreas Stömmer, Christoph Gschnaidtner, Gabriel Vollert, and Linda Bandelow (Technical University of Munich) for their assistance in identifying relevant keywords and labeling.

This paper was funded through the Austrian Research Promotion Agency (FFG) and the DigitalInnovationLayer project (FO999898902).

Appendix

See Tables A.1–A.3.

References

Abbas, Y., Martinetti, A., Moerman, J.-J., Hamberg, T., & van Dongen, L. A. (2020). Do you have confidence in how your rolling stock has been maintained? A blockchain-led knowledge-sharing platform for building trust between stakeholders. *International Journal of Information Management*, *55*, Article 102228, Impact of COVID-19 Pandemic on Information Management Research and Practice: Editorial Perspectives.

Aghion, P., Antonin, C., & Bunel, S. (2021). The power of creative destruction: economic upheaval and the wealth of nations. Harvard University Press,.

Agrawal, A., Galasso, A., & Oettl, A. (2017). Roads and innovation. *The Review of Economics and Statistics*, *99*(3), 417–434.

Aitzhan, N. Z., & Svetinovic, D. (2016). Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. *IEEE Transactions on Dependable and Secure Computing*, *15*(5), 840–852.

Ali, O., Ally, M., Clutterbuck, & Dwivedi, Y. (2020). The state of play of blockchain technology in the financial services sector: A systematic literature review. *International Journal of Information Management*, *54*, Article 102199.

Andoni, M., Robu, V., Flynn, D., Abram, S., Geach, D., Jenkins, D., et al. (2019). Blockchain technology in the energy sector: A systematic review of challenges and opportunities. *Renewable and Sustainable Energy Reviews*, *100*, 143–174.

Asheim, B. T., & Gertler, M. S. (2006). 291 The Geography of Innovation: Regional Innovation Systems. In *The Oxford handbook of innovation*. Oxford University Press,.

Asongu, S. A., Agboola, M. O., Alola, A. A., & Bekun, F. V. (2020). The criticality of growth, urbanization, electricity and fossil fuel consumption to environment sustainability in Africa. *Science of the Total Environment*, *712*, Article 136376.

Axenbeck, J., & Breithaupt, P. (2021). Innovation indicators based on firm websites-Which website characteristics predict firm-level innovation activity? *PLoS One*, *16*(4), e0249583.

Babich, V., & Hilary, G. (2020). Distributed ledgers and operations: What operations management researchers should know about blockchain technology. *Manufacturing & Service Operations Management*, *22*(2), 223–240.

Baum, C. F., & Lewbel, A. (2019). Advice on using heteroskedasticity-based identification. *The Stata Journal*, *19*(4), 757–767. http://dx.doi.org/10.1177/1536867X19893614.

Beck, R., Müller-Bloch, C., & King, J. L. (2018). Governance in the blockchain economy: A framework and research agenda. *Journal of the Association for Information Systems*, *19*(10), 1.

Beck, R., Stenum Czepluch, J., Lollike, N., & Malone, S. (2016). Blockchain - the gateway to trust-free cryptographic transactions. In *Twenty-fourth European conference on information systems (ECIS), Istanbul,Turkey, 2016* (pp. 1–14). United States: Springer Publishing Company, 24th European Conference on Information Systems ; Conference date: 12-06-2016 Through 15-06-2016.

Berkes, E., & Gaetani, R. (2019). The geography of unconventional innovation. Rotman School of Management Working Paper 3423143.

Berliant, M., Reed III, R. R., & Wang, P. (2006). Knowledge exchange, matching, and agglomeration. *Journal of Urban Economics*, *60*(1), 69–95.

Blakstad, S., & Allen, R. (2018). *FinTech revolution*: *vol. 121*, (p. 132). Cham, Switzerland: Springer.

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2007). Urban density and the rate of invention. *Journal of Urban Economics*, *61*(3), 389–419.

Chen, Y., & Bellavitis, C. (2020). Blockchain disruption and decentralized finance: The rise of decentralized business models. *Journal of Business Venturing Insights*, *13*, Article e00151.

Christidis, K., & Devetsikiotis, M. (2016). Blockchains and smart contracts for the internet of things. *IEEE Access*, *4*, 2292–2303.

Czarnitzki, D., & Hottenrott, H. (2009). Are local milieus the key to innovation performance? *Journal of Regional Science*, *49*(1), 81–112.

Dahlke, J., Beck, M., Kinne, J., Lenz, D., Dehghan, R., Wörter, M., et al. (2024). Epidemic effects in the diffusion of emerging digital technologies: evidence from artificial intelligence adoption. *Research Policy*, *53*(2), Article 104917.

Dörr, J. O., Kinne, J., Lenz, D., Licht, G., & Winker, P. (2022). An integrated data framework for policy guidance during the coronavirus pandemic: Towards real-time decision support for economic policymakers. *PLoS One*, *17*(2), 1–30.

Dörr, J. O., Kinne, J., Lenz, D., Licht, G., & Winker, P. (2022). An integrated data framework for policy guidance during the coronavirus pandemic: Towards real-time decision support for economic policymakers. *PLoS One*, *17*(2), 1–30. http://dx.doi.org/10.1371/journal.pone.0263898.

Du, W., Pan, S. L., Leidner, D. E., & Ying, W. (2019). Affordances, experimentation and actualization of FinTech: A blockchain implementation study. *The Journal of Strategic Information Systems*, *28*(1), 50–65.

Dwivedi, Y. K., et al. (2022). Climate change and COP26: Are digital technologies and information management part of the problem or the solution? An editorial reflection and call to action. *International Journal of Information Management*, *63*, Article 102456.

Ejermo, O., Hussinger, K., Kalash, B., & Schubert, T. (2022). Innovation in Malmø after the Øresund bridge. *Journal of Regional Science*, *62*(1), 5–20.

Engelhardt, M. A. (2017). Hitching healthcare to the chain: An introduction to blockchain technology in the healthcare sector. *Technology Innovation Management Review*, *7*(10).

Feldman, M. P. (1994). *The geography of innovation*: *vol. 2*, Springer Science & Business Media.

Friedman, N., & Ormiston, J. (2022). Blockchain as a sustainability-oriented innovation?: Opportunities for and resistance to blockchain technology as a driver of sustainability in global food supply chains. *Technological Forecasting and Social Change*, *175*, Article 121403.

Gökalp, E., Gökalp, M. O., & Çoban, S. (2022). Blockchain-based supply chain management: Understanding the determinants of adoption in the context of organizations. *Information Systems Management*, *39*(2), 100–121.

Gschnaidtner, C., Dehghan, R., Hottenrott, H., & Schwierzy, J. (2024). Adoption and diffusion of blockchain technology. *ZEW Discussion*, *24–018*, URL https://www.zew.de/en/publications/adoption-and-diffusion-of-blockchain-technology-1.

Heimeriks, G., & Van den Besselaar, P. (2006). Analyzing hyperlinks networks: The meaning of hyperlink based indicators of knowledge production. *Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics*, *10*, Nr. 10..

Horbach, J. (2020). *The importance of regional spill-over effects for eco-innovations in german start-ups: SEEDS working papers 1620*, SEEDS, Sustainability Environmental Economics and Dynamics Studies.

Howson, P. (2019). Tackling climate change with blockchain. *Nature Climate Change*, *9*(9), 644–645.

Hwang, J., Choi, M., Lee, T., Jeon, S., Kim, S., Park, S., et al. (2017). Energy prosumer business model using blockchain system to ensure transparency and safety. *Energy Procedia*, *141*, 194–198, Power and Energy Systems Engineering.

Jones, B. A., Goodkind, A. L., & Berrens, R. P. (2020). Economic estimation of Bitcoin mining's climate damages demonstrates closer resemblance to digital crude than digital gold. *Scientific Reports*, *12*, 14512.

Kamble, S. S., Gunasekaran, A., & Sharma, R. (2020). Modeling the blockchain enabled traceability in agriculture supply chain. *International Journal of Information Management*, *52*, Article 101967.

Karafiloski, E., & Mishev, A. (2017). Blockchain solutions for big data challenges: A literature review. In *IEEE EUROCON 2017 -17th international conference on smart technologies* (pp. 763–768).

Kinne, J., & Axenbeck, J. (2020). Web mining for innovation ecosystem mapping: A framework and a large-scale pilot study. *Scientometrics*, *125*, 2011–2041.

Kinne, J., & Lenz, D. (2021). Predicting innovative firms using web mining and deep learning. *PLoS One*, *16*, http://dx.doi.org/10.1371/journal.pone.0249071,.

Kshetri, N. (2017). Blockchain's roles in strengthening cybersecurity and protecting privacy. *Telecommunications Policy*, *41*(10), 1027–1038.

Kshetri, N. (2018). Blockchain's roles in meeting key supply chain management objectives. *International Journal of Information Management*, *39*, 80–89.

Kushwaha, A. K., Kar, A. K., & Dwivedi, Y. K. (2021). Applications of big data in emerging management disciplines: A literature review using text mining. *International Journal of Information Management Data Insights*, *1*(2), Article 100017.

Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, *30*(1), 67–80.

Liao, B., van den Berg, P. E., van Wesemael, P., & Arentze, T. A. (2020). Empirical analysis of walkability using data from the Netherlands. *Transportation Research Part D: Transport and Environment*, *85*, Article 102390.

Lin, Y.-P., Petway, J. R., Lien, W.-Y., & Settele, J. (2018). Blockchain with artificial intelligence to efficiently manage water use under climate change. *Environments*, *5*(3), 34.

Losacker, S., Horbach, J., & Liefner, I. (2022). Geography and the speed of green technology diffusion. *Industry and Innovation*, 1–25.

Mirtsch, M., Kinne, J., & Blind, K. (2021). Exploring the adoption of the international information security management system standard ISO/IEC 27001: A web mining-based analysis. *IEEE Transactions on Engineering Management*, *68*(1), 87–100. http://dx.doi.org/10.1109/TEM.2020.2977815.

Mora, C., Rollins, R. L., Taladay, K., Kantar, M. B., Chock, M. K., Shimada, M., et al. (2018). Bitcoin emissions alone could push global warming above 2 C. *Nature Climate Change*, *8*(11), 931–933.

Murugesan, S., & Benakanahally Lakshminarasaiah, M. (2022). A survey on blockchain-based student certificate management system. In *ICEGOV '21, Proceedings of the 14th international conference on theory and practice of electronic governance* (pp. 44–50). Athens, Greece: Association for Computing Machinery.

Nærland, K., Müller-Bloch, C., Beck, R., & Palmund, S. (2017). Blockchain to rule the waves-nascent design principles for reducing risk and uncertainty in decentralized environments. In *ICIS*.

Nair, R. S., Agrawal, R., Domnic, S., & Kumar, A. (2021). Image mining applications for underwater environment management - a review and research agenda. *International Journal of Information Management Data Insights*, *1*(2), Article 100023.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. (p. 9). http://dx.doi.org/10.1007/s10838-008-9062-0, URL https://bitcoin.org/bitcoin.pdf.

Neis, P., & Zielstra, D. (2014). Recent developments and future trends in volunteered geographic information research: The case of OpenStreetMap. *Future Internet*, *6*(1), 76–106,.

Nishant, R., Kennedy, M., & Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management*, *53*, Article 102104.

Notheisen, B., Cholewa, J. B., & Shanmugam, A. P. (2017). Trading real-world assets on blockchain. *Business & Information Systems Engineering*, *59*(6), 425–440.

Nylund, P. A., Brem, A., & Agarwal, N. (2021). Innovation ecosystems for meeting sustainable development goals: The evolving roles of multinational enterprises. *Journal of Cleaner Production*, *281*, Article 125329.

Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis*, *27*(4), 286–306. http://dx.doi.org/10.1111/j.1538-4632.1995.tb00912.x,.

Park, H. W., & Thelwall, M. (2003). Hyperlink analyses of the world wide web: A review. *Journal of Computer-Mediated Communication*, *8*(4), (Visited on 04/28/2021).

Patón-Romero, J. D., Baldassarre, M. T., Toval, A., Rodríguez, M., & Piattini, M. (2022). Auditing the governance and management of green IT. *Journal of Computer Information Systems*, *62*(5), 896–906.

Rammer, C., Kinne, J., & Blind, K. (2020). Knowledge proximity and firm innovation: A microgeographic analysis for Berlin. *Urban Studies*, *57*(5), 996–1014.

Rimba, P., Tran, A. B., Weber, I., Staples, M., Ponomarev, A., & Xu, X. (2020). Quantifying the cost of distrust: Comparing blockchain and cloud services for business process execution. *Information Systems Frontiers*, *22*(2), 489–507.

Roche, M. P. (2020). Taking innovation to the streets: Microgeography, physical structure, and innovation. *The Review of Economics and Statistics*, *102*(5), 912–928.

Rogers, R. (2013). Mapping public web space with the issuecrawler. In *Digital cognitive technologies*. John Wiley & Sons, Ltd, (Visited on 04/28/2021).

Saberi, S., Kouhizadeh, M., Sarkis, J., & Shen, L. (2019). Blockchain technology and its relationships to sustainable supply chain management. *International Journal of Production Research*, *57*(7), 2117–2135.

Sanderson, O. (2018). Chapter 20 - how to trust green bonds: Blockchain, climate, and the institutional bond markets. In A. Marke (Ed.), *Transforming climate finance and green investment with blockchains* (pp. 273–288). Academic Press.

Santos, F. A., Silva, T. H., Ferreira Loureiro, A. A., & Villas, L. A. (2018). Uncovering the perception of Urban Outdoor Areas expressed in social media. In *2018 IEEE/WIC/aCM international conference on web intelligence* (pp. 120–127).

Schinckus, C. (2020). The good, the bad and the ugly: An overview of the sustainability of blockchain technology. *Energy Research & Social Science*, *69*, Article 101614.

Schmidt, S., Kinne, J., Lautenbach, S., Blaschke, T., Lenz, D., & Resch, B. (2022). Greenwashing in the US metal industry? A novel approach combining SO_2 concentrations from satellite data, a plant-level firm database and web text mining. *Science of the Total Environment*, *835*, Article 155512.

Schmidt, C. G., & Wagner, S. M. (2019). Blockchain and supply chain relations: A transaction cost theory perspective. *Journal of Purchasing and Supply Management*, *25*(4), Article 100552.

Schwierzy, J., Dehghan, R., Schmidt, S., Rodepeter, E., Stoemmer, A., Uctum, K., et al. (2022). Technology mapping using WebAI: The case of 3D printing.

Sehra, S. S., Singh, J., & Rai, H. S. (2014). A systematic study of OpenStreetMap data quality assessment. In *2014 11th international conference on information technology: New generations* (pp. 377–381).

Sharma, P. K., Kumar, N., & Park, J. H. (2020). Blockchain technology toward green IoT: Opportunities and challenges. *IEEE Network*, *34*(4), 263–269.

Shearmur, R. (2012). Are cities the font of innovation? A critical review of the literature on cities and innovation. *Cities*, *29*, S9–S18.

Sikorski, J. J., Haughton, J., & Kraft, M. (2017). Blockchain technology in the chemical industry: Machine-to-machine electricity market. *Applied Energy*, *195*, 234–246.

Simmonds, D., & Bhattacherjee, A. (2012). Environmental sustainability in organizations: The information technology role. *AMCIS 2012 Proceedings 11*.

Singh, M., & Sahu, G. P. (2020). Towards adoption of Green IS: A literature review using classification methodology. *International Journal of Information Management*, *54*, Article 102147.

Sotarauta, M., & Suvinen, N. (2019). Place leadership and the challenge of transformation: Policy platforms and innovation ecosystems in promotion of green growth. *European Planning Studies*, *27*(9), 1748–1767.

Stock, J., & Yogo, M. (2005). Asymptotic distributions of instrumental variables statistics with many instruments. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, *6*, 109–120.

Stoll, C., Klaaßen, L., & Gallersdörfer, U. (2019). The carbon footprint of Bitcoin. *Joule*, *3*(7), 1647–1661.

Stuart, T., & Sorenson, O. (2003). The geography of opportunity: Spatial heterogeneity in founding rates and the performance of biotechnology firms. *Research Policy*, *32*(2), 229–253.

Treiblmaier, H. (2018). The impact of the blockchain on the supply chain: A theory-based research framework and a call for action. *Supply Chain Management: An International Journal*.

Treiblmaier, H., & Beck, R. (2019). *Business transformation through blockchain: vol. 1*, Springer.

UN (2022). Sustainable development. The 17 goals. https://sdgs.un.org/goals. (Accessed: 12 December 022).

United States Census Bureau (2021). Statistics of U.S. Businesses datasets. URL https://www.census.gov/programs-surveys/susb/data/datasets.html.

Vaughan, L., Gao, Y., & Kipp, M. (2006). Why are hyperlinks to business Websites created? A content analysis. *Scientometrics*, *67*(2), 291–300, Publisher: Springer & Akadémiai Kiadó.

Wang, Y., Chen, Y., & Benitez-Amado, J. (2015). How information technology influences environmental performance: Empirical evidence from China. *International Journal of Information Management*, *35*(2), 160–170.

Wang, X.-D., Luo, Z.-C., Liu, H., Zhao, N., Liu, M., Zhu, Y.-F., et al. (2015). Gold nanorod as saturable absorber for Q-switched yb-doped fiber laser. *Optics Communications*, *346*, 21–25.

Williamson, P. J., & Meyer, A. D. (2012). Ecosystem advantage: How to successfully harness the power of partners. *California Management Review*, *55*(1), 24–46.

Yang, Z., Chen, H., Du, L., Lin, C., & Lu, W. (2021). How does alliance-based government-university-industry foster cleantech innovation in a green innovation ecosystem? *Journal of Cleaner Production*, *283*, Article 124559.