// NO.24-023 | 04/2024

DISCUSSION PAPER

// JAN KINNE, ROBERT DEHGHAN, SEBASTIAN SCHMIDT,
DAVID LENZ, AND HANNA HOTTENROTT

Location Factors and Ecosystem Embedding of Sustainability-Engaged Blockchain Companies in the US. A Web-Based Analysis

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

Jan Kinne[1,2,3,*], Robert Dehghan[1,4], Sebastian Schmidt[1,5], David Lenz[1,6], Hanna Hottenrott[2,7]

1 ISTARI.AI, Julius-Hatry-Str. 1, 68163 Mannheim, Germany
2 Department of Economics of Innovation and Industrial Dynamics, Centre for European Economic Research (ZEW), L7 1, 68161 Mannheim, Germany
3 Center for Geographic Analysis at Harvard University, 1737 Cambridge Street, Cambridge, MA 02138, USA
4 Department of SME Research and Entrepreneurship, L9 1-2, University of Mannheim, 68161 Mannheim, Germany
5 Department of Geoinformatics – Z_GIS, University of Salzburg, Schillerstr. 30, 5020 Salzburg, Austria
6 Department of Statistics and Econometrics, Justus-Liebig-University, Licher Str. 64, 35394 Giessen, Germany
7 Department of Economics & Policy, Technical University of Munich, Arcisstr. 21, 80333 Munich, Germany

*Corresponding author: jan.kinne@istari.ai

Abstract

While many digital technologies provide opportunities for creating business models with an impact on 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 identify 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 sustainable blockchain firms is significantly higher than in the overall firm population. Additionally, we find that blockchain companies with an intensified focus on sustainability have, at least quantitatively, a more intensive embedding in entrepreneurial ecosystems, while infrastructural and socio-economic location factors hardly play a role.

Keywords: sustainability, blockchain, ecosystem, location factors, natural language processing
JEL classification: Q56, R30, L86

1 Introduction

Understanding the degree of sustainability of novel information technology is crucial for assessing its impact on society. This study explores the use of blockchain technology by companies across the United States (US) in a sustainability context. Retrieving data on such a specific topic from traditional databases is very difficult, which is why we relied on an innovative, web-based methodology. Using web text mining, we identified companies with blockchain-based business models and investigated their degree of sustainability. We further analyzed the role of the local business ecosystem and infrastructure in adopting 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.

We focused on blockchain because it is currently among the most controversial digital technologies with regard to its environmental impact (Stoll et al., 2019; Asongu et al., 2020; Jones et al., 2020). In principle, blockchain can be used both in environmentally harmful ways (e.g. high emissions as a result of 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.

Research on innovation ecosystems stresses the importance of location factors for the invention and adoption of new technologies and for the performance of companies using and diffusing them (Asheim & Gertler, 2006; Williamson & Meyer, 2012). While the possible benefits (Simmonds & Bhattacherjee, 2012; Kushwaha et al., 2021; Nair et al., 2021; Patón-Romero et al., 2022) 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. It is also an open question whether local milieus still matter for digital technologies. To answer these questions, we distinguished between (physical) infrastructure and the local business ecosystem in their role for the sustainable use of blockchain.

In our paper, we address the research gap that the relationship between sustainability and blockchain use in the US is unclear. To our knowledge, there is no information on this at the company level yet. 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?

The paper is structured as follows: In the second chapter, we present the theoretical background of the study, focusing on blockchain technology and the ecosystems of technology-savvy companies. In the third chapter, we present our data basis and methodology, focusing on the generation of web-based indicators. After presenting our descriptive results (e.g. in the form of choropleth maps), we present the results of our regression analyses. Finally, we critically discuss our results and methods.

2 Theoretical background

In the following, we will give 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 et al., 2021). In particular, information technologies provide the opportunity to support the transition toward a more sustainable economy (Wang et al., 2015b; Wang et al., 2015a; Singh & Sahu, 2020). Green IT, Artificial Intelligence (AI) and the Internet of Things (IoT) indeed offer solutions for reducing negative environmental impact through increasing resource-use efficiency (Simmonds & Bhattacherjee, 2012; Kushwaha et al., 2021; Nair et al., 2021; Patón-Romero et al., 2022). 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 et al., 2018; Nishant et al., 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 (Stoll et al., 2019; Asongu et al., 2020; Jones et al., 2020).

2.2 Blockchain

Blockchain is a relatively new data storage technology that was first introduced to the public in 2008 (Nakamoto, 2008). This publication also included 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, organisation managements, supply chains, payment processes, security of data and even whole markets (Du et al., 2019; Schinckus, 2020; Notheisen et al., 2017; Beck et al., 2018; Treiblmaier, 2018; Schmidt & Wagner, 2019; Abbas et al., 2020; Rimba et al., 2020). 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 (Schinckus, 2020; Ali et al., 2020; Karafiloski & Mishev, 2017). 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 the 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 (Schmidt & Wagner, 2019; Notheisen et al., 2017; Karafiloski & Mishev, 2017; Beck et al., 2018; Beck et al., 2016; Nærland et al., 2017).

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 (Schmidt & Wagner, 2019; Babich & Hilary, 2020; 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 et al., 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 that blockchain technology consumes too much energy to ever result in sustainable applications (Stoll et al., 2019). There is a lot of criticism of the POW validation approach in particular because of its high energy consumption. 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 like China, Iran and US 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 applications facilitate technology efficiencies which in total result in lower energy consumption (Sharma et al., 2020). Moreover, blockchain-based smart contracts have been shown to enable trading in carbon credits and thereby eventually reduce emissions of companies. Blockchain applications also include use cases such as managing the energy-intensive tracking of product flows along the supply chain, as well as verifying the origin of inputs (Howson, 2019).

Furthermore blockchain applications could make a major contribution by impacting at least 14 out of the 17 United Nations SDGs (Schinckus, 2020; UN, 2022). Among many use cases, blockchain can empower communities and their networks through its creation of trust and transparency, improve food trust, facilitate more efficient water management, and reduce electricity consumption and improve energy efficiency through establishing high credibility and reduce fraud through transparent and unchangeable records (Blakstad &

Allen, 2018; Treiblmaier & Beck, 2019; Schinckus, 2020; Sanderson, 2018; Sikorski et al., 2017; Hwang et al., 2017).

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; Williamson & Meyer, 2012; Gschnaidtner et al., 2024). Access to employees and financing, 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 (Feldman, 1994; Czarnitzki & Hottenrott, 2009). Therefore, companies tend to locate in close proximity to similar companies in order to use the established social and professional links (Stuart & Sorenson, 2003). Innovation research hence has long stressed the impact of location factors for facilitating knowledge spillovers and collaboration. Particularly, transportation infrastructure plays a crucial role in innovation as it facilitates the mobility of human capital and the flow of goods (Agrawal et al., 2017). This finding is in line with more recent research that documents that the physical layout of cities in the US affects innovation by influencing the organization of knowledge exchange (Roche, 2020). A recent 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 et al., 2022).

While the link between infrastructure, ecosystems, and innovation, in general, is quite established, it is less clear whether it also applies to the adoption of technologies in a sustainable context. 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 infrastructure and business ecosystems matter also for a sustainable context (Yang et al., 2021; Sotarauta & Suvinen, 2019; Nylund et al., 2021). Larger-scale systematic evidence is, however, still scarce. Another study shows 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). The time-to-adoption of technologies in a sustainable context further indicates that geographic proximity to other innovators accelerates their adoption by firms, and regions specializing in green technologies experience faster diffusion within the same area (Losacker et al., 2022).

While these 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. The following analysis therefore aims to shed light on the question of whether the local ecosystem matters for the sustainable use of blockchain.

3 Materials and Methods

In the following, we will 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 United States of America 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, using 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 (see next section). This corresponded to a coverage of about 61% of all economically active companies in the US according to the US Census Bureau. 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 the startup 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. In total, this corresponded to about one terabyte of text data.

3.2.1 Web-based blockchain indicator

To infer a companies' blockchain capacity from their website texts, we trained a NLP model and represented the output as a firm-level *blockchain intensity score*.We understand blockchain capacities in this context as products and services with integrated blockchain technology or personnel with blockchain-related skills. Our indicator reflected how prominently the topic of blockchain was communicated by the company on its own website and how it was portrayed as essential to its own business model. We assumed 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 related to offering 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 sections (paragraphs) that deal with the topic of blockchain. For this, we relied on a simple, but extensive keyword search (cf. Table A1). 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 3,500 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 proprietary 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 a 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. 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), the effect of the COVID-19 pandemic on firms (Dörr et al., 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 tech stack.

3.2.2 Web-based sustainability indicator

For the identification of companies engaged in sustainability, we developed a NLP model in the same manner 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 (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. This list was developed together with experts from the OECD and included around 1,000 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 eco-system 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 et al., 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 (1km 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 used the hyperlinks downloaded in the web scraping step to map out a network between the companies under study and to quantify the interconnectedness of the blockchain companies. 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 disassociation, respectively" (Rogers, 2013). Several studies have shown the significance of hyperlinks for uncovering firms' network relations (Heimeriks & Van den Besselaar, 2006; Vaughan et al., 2006; Kinne & Axenbeck, 2020; Axenbeck & Breithaupt, 2021). 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.

We built a hyperlink network for all approximately 3.72 million corporate websites, where the edges represented the linkage of one firm to their partners. 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.

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. In total, we thus recorded 1,355,433 partners for 26,905 blockchain companies.

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*. OSM data can be accessed in different ways, e.g. via APIs or dedicated websites. 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 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 50km (or 100km 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 100km. 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 1km which has been found empirically as a significant threshold of walkability (Liao et al., 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 A3.

Table 1. Overview of variables used in analysis.

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

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.

4 Results

In the following, we present our main findings. First, we show descriptive statistics and the geographical distribution of blockchain companies. In a 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. Figure 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).

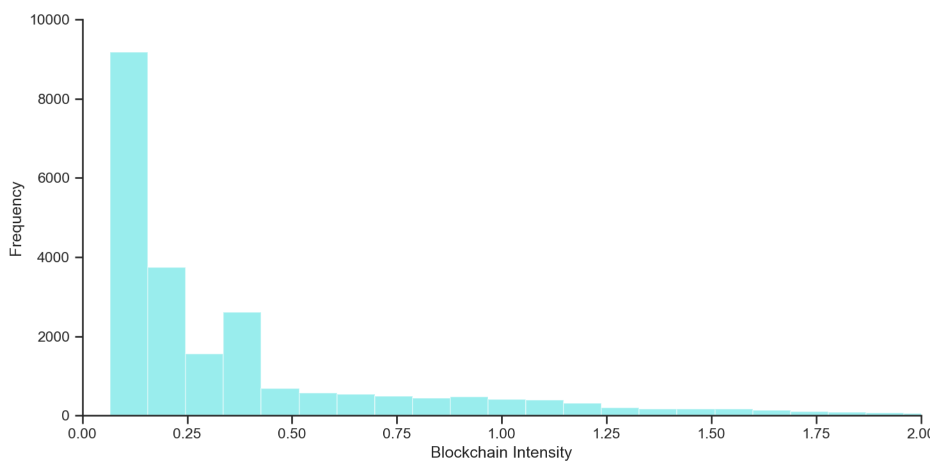

Figure 1. Histogram of blockchain intensity for companies with blockchain intensity ≥ 0.00.

We aggregated the blockchain companies at the county level in order to assess their spatial distribution in relation to the overall firm 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 Figure 2, which portrays the share of blockchain firms in the local firm population per county. For the remaining counties, we calculated the Moran's I statistic to check for spatial clustering in the geographic distribution. The 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 2 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

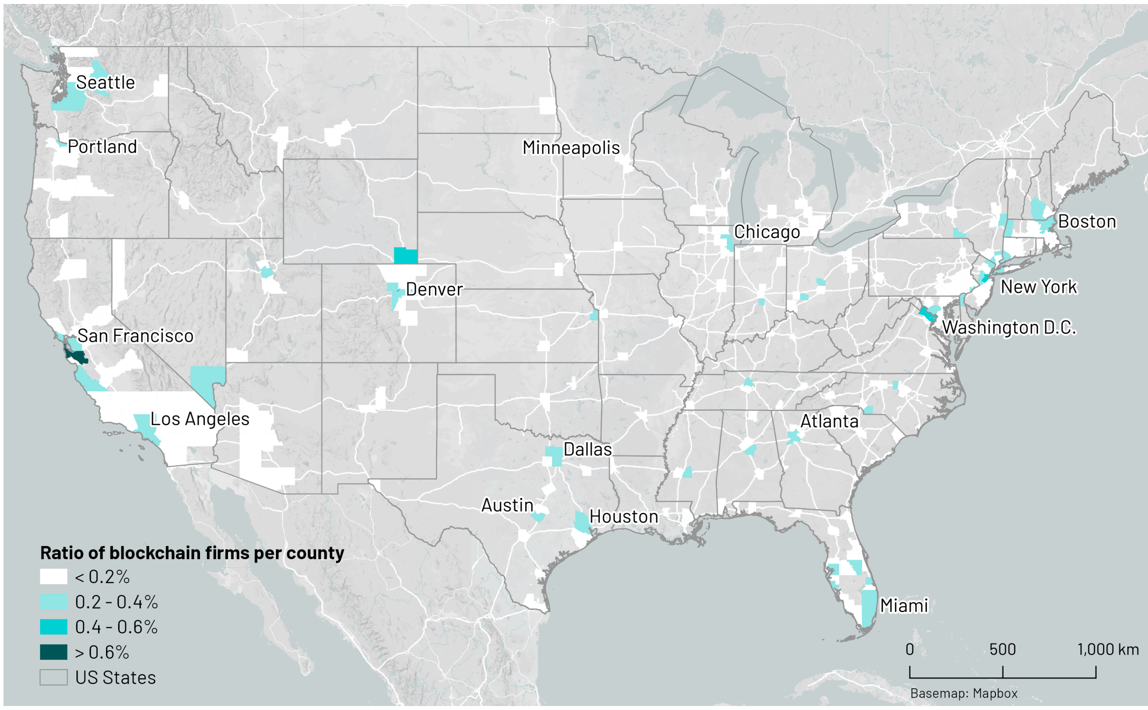

Figure 2. Share of above-average (≥ 0.185) blockchain intensity firms in the overall firm population per county. Counties with less than 10 blockchain firms are excluded.

Coast, which was also reflected in the top ten states. The table additionally lists the top ten sustainability counties. Three counties appeared in both categories (District of Columbia; Suffolk, MA; and Arlington, VA).

Table 2. 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

Using our sustainability intensity, we further divided the identified blockchain companies into those that were focused on sustainability and those that were not. About 32.1 % of the blockchain companies had a sustainability intensity greater than 0 and thus communicated a focus on or commitment to sustainability on their websites. Accordingly, this means that the proportion of sustainable blockchain companies was significantly higher than the 12.7 % of sustainability committed companies in the overall US company population. Figure 3 shows the sustainability intensity distribution of these sustainable 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".

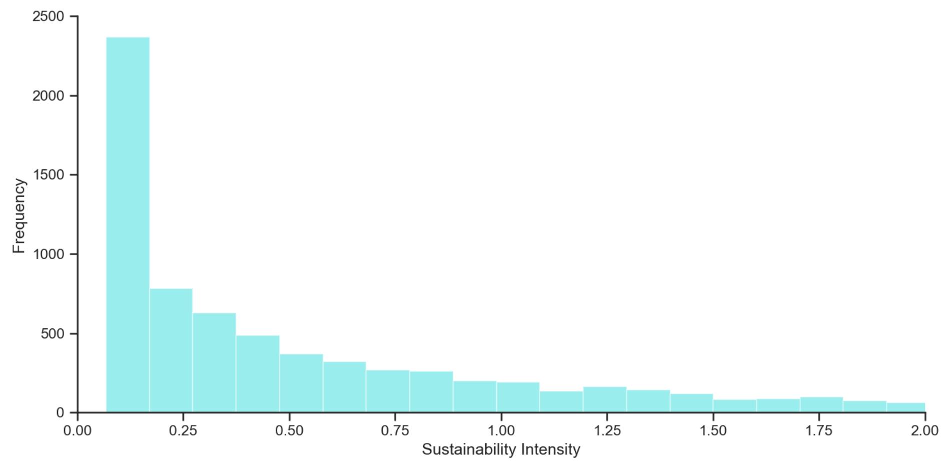

Figure 3. Histogram of sustainability intensity for blockchain companies with sustainability intensity ≥ 0.00.

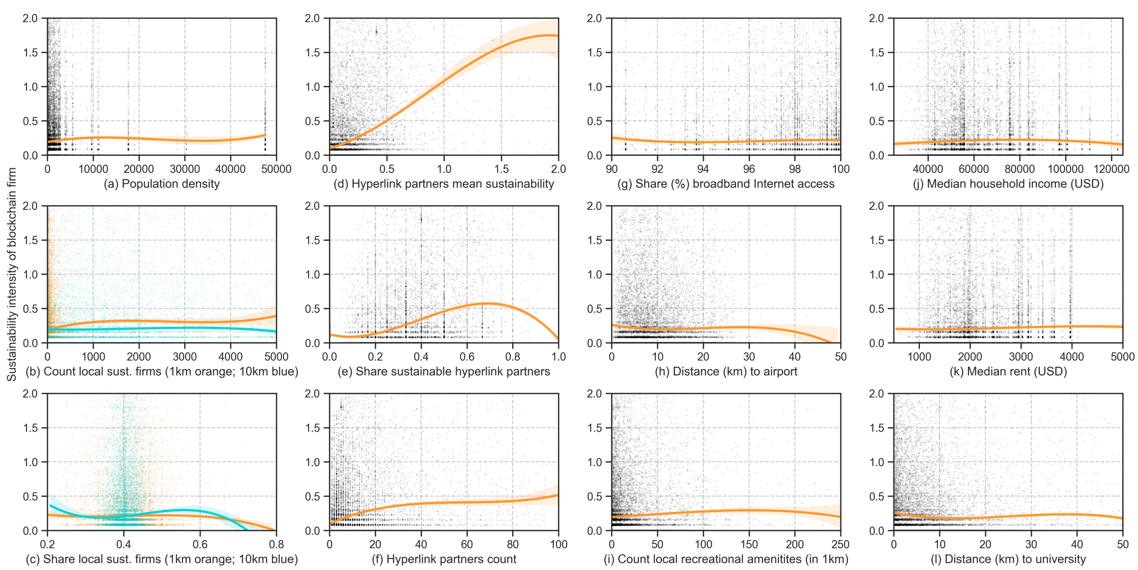

Figure 4. Scatter plots and fitted regression lines of third order for sustainability intensity of blockchain companies and selected location and ecosystem factors.

Figure 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) to (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.

Figure 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 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 sustainable blockchain companies were identified on the East Coast.

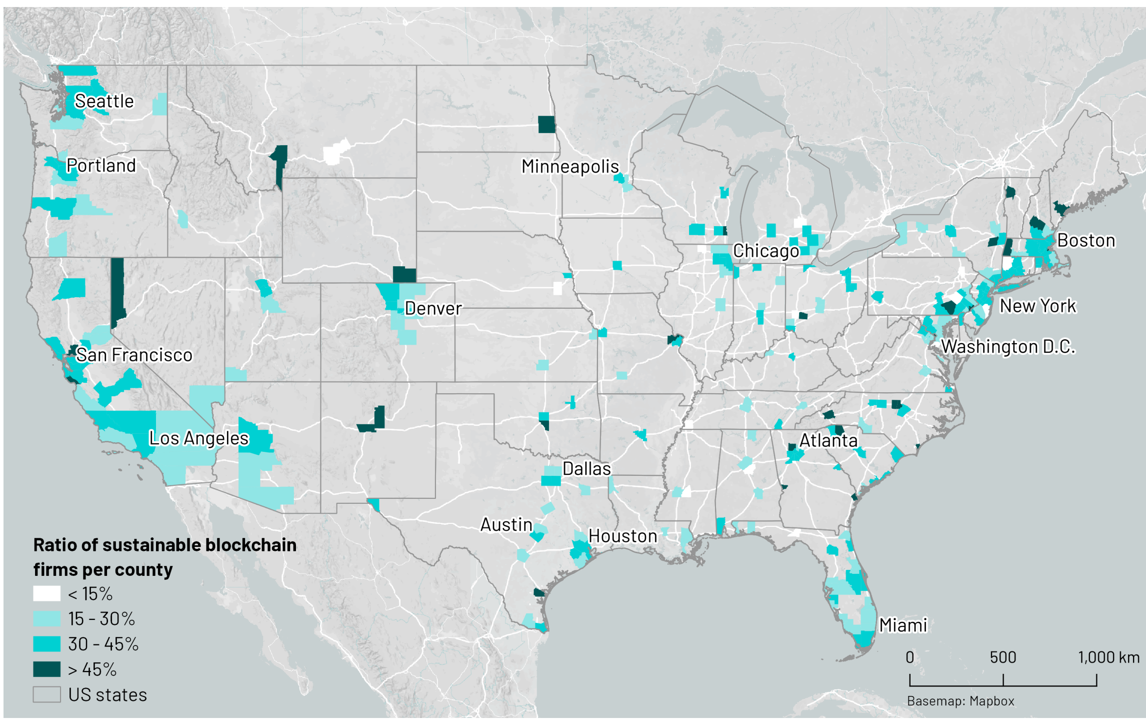

Figure 5. Share of sustainable blockchain firms in overall firm population per county. Counties with less than 10 blockchain firms are excluded.

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. Figure 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.

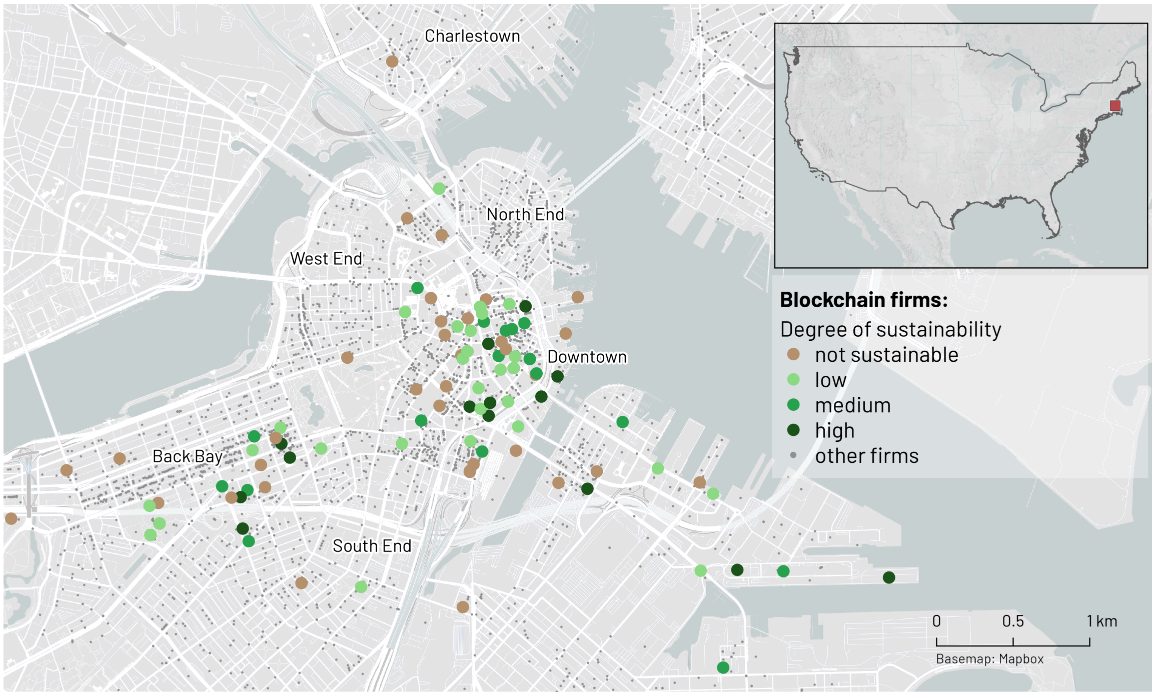

Figure 6. Exemplary zoom in map of blockchain firm locations in Suffolk County, MA (i.e. Boston).

Another important economic aspect is the networking of companies. Figure 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 that there were strong connections across the entire US. The major agglomerations in particular 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 3 included the main predictors capturing the local network links to sustainable companies, the number of hyperlinked partners, the overall number of partners as well as control variables for the size of the company measured in employees. Moreover, we included the indicators for whether it is an e-commerce company or simply uses crypto-pay options in the website. These predictors alone explained about 14% of the variance in the sustainability score. The local network indicators were positive

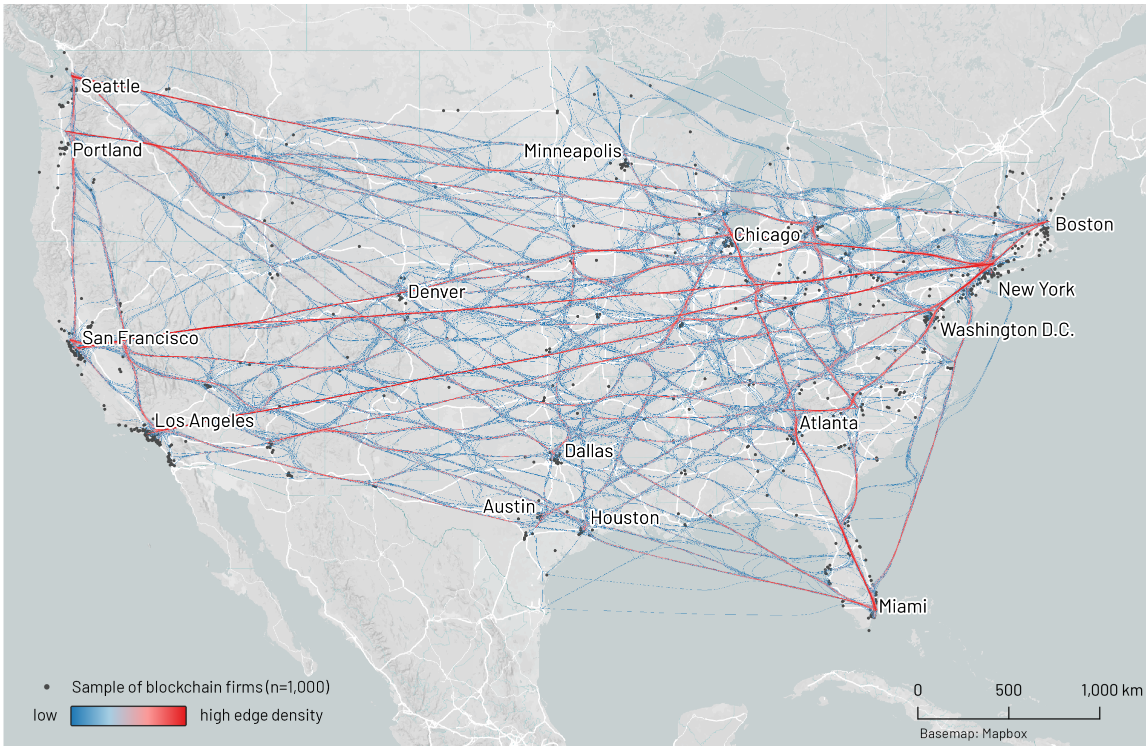

Figure 7. Exemplary hyperlink network between a random sample of 1,000 blockchain firms and their sustainable partner firms. Edge bundling technique has been applied in order to show high (red) and low (blue) density aggregated connections.

and statistically significant, indicating that being embedded in local ecosystems was related to a higher sustainability degree of blockchain companies. Particularly, the partners' sustainability degree appeared to matter while controlling for the number of partners. 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. We controlled for size since larger companies generally scored higher on sustainability. 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%. Specification (4) contained also 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 sustainability degree. Specification (5) further included county-level characteristics related to knowledge, accessibility, poverty, health, and quality of life as shown in Table 1. Adding these factors did not increase the explanatory power of the model by much, since most factors were insignificant. Importantly, the insights for the sustainable ecosystem measures remained unchanged through the inclusion of these additional regressors. These findings also confirm the descriptive patterns shown in Figure 4.

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 sustainable companies in a company's neighborhood and the overall number of partners could also be endogenous. To address such endogeneity concerns, we represented the Partners' Sustainability Intensity (as well as the second-order term), the number of sustainable companies and the total number of partners with instrumental variables that explained the endogenous measures, but not the sustainability 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 testing whether the main coefficients of interest switch signs once they are instrumented with exogenous regressors. We present the results from this Lewbel-IV model in column 6 of Table 3 using the specification 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). The IV model results 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 sustainable companies in the same location was still positive, but no longer statistically significant at the 10% level.

5 Discussion

In light of the debate on whether blockchain technology can be used for sustainable purposes, this study offers first insights. Table A2 in the appendix also provides an overview of the SDGs (as described in chapter 2.2) and includes sample websites of blockchain companies that we identified through our web mining approach and whose work aligns with the respective SDG. With respect to the entire US firm population (RQ1 & RQ2), we found that blockchain is still a niche technology and while indeed 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, which is much higher 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 (Feldman, 1994; Czarnitzki & Hottenrott, 2009; Roche, 2020; Rammer et al., 2020), we found that both being connected to other sustainable companies and being in close geographic proximity to a large number of other sustainable 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 due to the fact that 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 actually 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

Table 3. 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 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

variable showed a significant positive impact on the sustainability intensity of blockchain companies, which is counterintuitive 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 actually 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 sustainable blockchain companies tended to be located in high-density areas in terms of overall and sustainable firm counts. They also tended to have more hyperlinked partners, which were more sustainable on average. These findings could already be observed in the descriptive statistics and were additionally confirmed by our regression analyses. These results could be an indication that companies that are integrated into large networks with sustainable partners have incentives to implement blockchain technologies in a sustainability context themselves. Possible explanations could be that this creates better sales opportunities for their products with sustainability-focused customers or that the necessary know-how for their own sustainable products is sourced from the network. At the same time, however, it is of course also possible that this is pure lip service and that a "green image" is created on the basis of peer pressure (i.e. greenwashing).

The same considerations applied to the interpretation of the 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. If these variables were understood as proxies for being embedded in a local ecosystem that allows a company to experience latent spillovers, then the same considerations regarding peer pressure, imitation and know-how transfer could be applied here. Combining all these considerations with the fact that larger companies (with more employees) exhibited a higher sustainability intensity on average, this might indicate that sustainable blockchain applications were adopted by companies that have larger, existing operations already in place.

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.

Our investigation already showed the potential of web-based studies of this kind when compared against patent data based studies or other traditional methods. Despite achieving unprecedented coverage, our approach 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. Here, future research could combine our data with other databases that use a different approach to analyze companies (e.g. official corporate sustainability reports). 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 et al., 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 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 et al., 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.

6 Conclusion

In this study, we presented a novel approach to identify sustainability-engaged blockchain companies. Furthermore, we correlated their sustainability 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 natural language processing and 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 is 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 sustainable blockchain companies were located in regions with a high density of companies and within one kilometer of many other sustainable 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.

Acknowledgement 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 labelling.

References

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

Aghion, P., C. Antonin, and S. Bunel (2021). *The Power of Creative Destruction: Economic Upheaval and the Wealth of Nations*. Harvard University Press. ISBN: 9780674258686.

Agrawal, A., A. Galasso, and A. Oettl (2017). "Roads and innovation". In: *Review of Economics and Statistics* 99.3, pp. 417–434.

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

Asheim, B. T. and M. S. Gertler (2006). "291 The Geography of Innovation: Regional Innovation Systems". In: *The Oxford Handbook of Innovation*. Oxford University Press. ISBN: 9780199286805.

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

Axenbeck, J. and P. Breithaupt (2021). "Innovation indicators based on firm websites-Which website characteristics predict firm-level innovation activity?" eng. In: *PloS One* 16.4, e0249583.

Babich, V. and G. Hilary (2020). "Distributed Ledgers and Operations: What Operations Management Researchers Should Know About Blockchain Technology". In: *Manufacturing & Service Operations Management* 22.2, pp. 223–240.

Beck, R., C. Müller-Bloch, and J. L. King (2018). "Governance in the blockchain economy: A framework and research agenda". In: *Journal of the Association for Information Systems* 19.10, p. 1.

Beck, R., J. Stenum Czepluch, N. Lollike, and S. Malone (2016). "Blockchain - The Gateway to Trust-Free Cryptographic Transactions". In: *Twenty-Fourth European Conference on Information Systems (ECIS), Istanbul,Turkey, 2016*. 24th European Conference on Information Systems ; Conference date: 12-06-2016 Through 15-06-2016. United States: Springer Publishing Company, pp. 1–14.

Blakstad, S. and R. Allen (2018). "FinTech Revolution". In: *Cham, Switzerland: Springer* 121, p. 132.

Czarnitzki, D. and H. Hottenrott (2009). "Are local milieus the key to innovation performance?" In: *Journal of Regional Science* 49.1, pp. 81–112.

Dörr, J. O., J. Kinne, D. Lenz, G. Licht, and P. Winker (2022). "An integrated data framework for policy guidance during the coronavirus pandemic: Towards real-time decision support for economic policymakers". In: *PLOS ONE* 17.2, pp. 1–30.

Du, W., S. L. Pan, D. E. Leidner, and W. Ying (2019). "Affordances, Experimentation and Actualization of FinTech: A Blockchain Implementation Study". In: *The Journal of Strategic Information Systems* 28.1, pp. 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". In: *International Journal of Information Management* 63, p. 102456.

Ejermo, O., K. Hussinger, B. Kalash, and T. Schubert (2022). "Innovation in Malmø after the Øresund Bridge". In: *Journal of Regional Science* 62.1, pp. 5–20.

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

Gökalp, E., M. O. Gökalp, and S. Çoban (2022). "Blockchain-based supply chain management: understanding the determinants of adoption in the context of organizations". In: *Information systems management* 39.2, pp. 100–121.

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

Heimeriks, G. and P. Van den Besselaar (2006). "Analyzing Hyperlinks Networks: The Meaning of Hyperlink Based Indicators of Knowledge Production". In: *Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics, ISSN 1137-5019, Nr. 10, 2006* 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". In: *Nature Climate Change* 9.9, pp. 644–645.

Hwang, J., M.-i. Choi, T. Lee, S. Jeon, S. Kim, S. Park, and S. Park (2017). "Energy Prosumer Business Model Using Blockchain System to Ensure Transparency and Safety". In: *Energy Procedia* 141. Power and Energy Systems Engineering, pp. 194–198.

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

Karafiloski, E. and A. Mishev (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. and J. Axenbeck (2020). "Web mining for innovation ecosystem mapping: a framework and a large-scale pilot study". In: *Scientometrics* 125, pp. 2011–2041.

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

Kushwaha, A. K., A. K. Kar, and Y. K. Dwivedi (2021). "Applications of big data in emerging management disciplines: A literature review using text mining". In: *International Journal of Information Management Data Insights* 1.2, p. 100017.

Lewbel, A. (2012). "Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models". In: *Journal of Business & Economic Statistics* 30.1, pp. 67–80.

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

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

Losacker, S., J. Horbach, and I. Liefner (2022). "Geography and the speed of green technology diffusion". In: *Industry and Innovation* 0.0, pp. 1–25.

Mora, C., R. L. Rollins, K. Taladay, M. B. Kantar, M. K. Chock, M. Shimada, and E. C. Franklin (2018). "Bitcoin emissions alone could push global warming above 2 C". In: *Nature Climate Change* 8.11, pp. 931–933.

Murugesan, S. and M. Benakanahally Lakshminarasaiah (2022). "A Survey on Blockchain-Based Student Certificate Management System". In: *Proceedings of the 14th International Conference on Theory and Practice of Electronic Governance*. ICEGOV '21. Athens, Greece: Association for Computing Machinery, pp. 44–50.

Nærland, K., C. Müller-Bloch, R. Beck, and S. Palmund (2017). "Blockchain to Rule the Waves-Nascent Design Principles for Reducing Risk and Uncertainty in Decentralized Environments." In: *ICIS*.

Nair, R. S., R. Agrawal, S. Domnic, and A. Kumar (2021). "Image mining applications for underwater environment management - A review and research agenda". In: *International Journal of Information Management Data Insights* 1.2, p. 100023.

Nakamoto, S. (2008). "Bitcoin: A Peer-to-Peer Electronic Cash System". DOI: 10.1007/s10838-008-9062-0. URL: https://bitcoin.org/bitcoin.pdf.

Neis, P. and D. Zielstra (2014). "Recent developments and future trends in Volunteered Geographic Information research: The case of OpenStreetMap". In: *Future Internet* 6.1, pp. 76–106. ISSN: 1999-5903.

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

Notheisen, B., J. B. Cholewa, and A. P. Shanmugam (2017). "Trading real-world assets on blockchain". In: *Business & Information Systems Engineering* 59.6, pp. 425–440.

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

Park, H. W. and M. Thelwall (2003). "Hyperlink Analyses of the World Wide Web: A Review". en. In: *Journal of Computer-Mediated Communication* 8.4. (Visited on 04/28/2021).

Patón-Romero, J. D., M. T. Baldassarre, A. Toval, M. Rodríguez, and M. Piattini (2022). "Auditing the Governance and Management of Green IT". In: *Journal of Computer Information Systems* 62.5, pp. 896–906.

Rammer, C., J. Kinne, and K. Blind (2020). "Knowledge proximity and firm innovation: A microgeographic analysis for Berlin". In: *Urban Studies* 57.5, pp. 996–1014.

Rimba, P., A. B. Tran, I. Weber, M. Staples, A. Ponomarev, and X. Xu (2020). "Quantifying the cost of distrust: Comparing blockchain and cloud services for business process execution". In: *Information Systems Frontiers* 22.2, pp. 489–507.

Roche, M. P. (2020). "Taking Innovation to the Streets: Microgeography, Physical Structure, and Innovation". In: *Review of Economics and Statistics* 102.5, pp. 912–928.

Rogers, R. (2013). "Mapping Public Web Space with the Issuecrawler". en. In: *Digital Cognitive Technologies*. John Wiley & Sons, Ltd. (Visited on 04/28/2021).

Sanderson, O. (2018). "Chapter 20 - How to Trust Green Bonds: Blockchain, Climate, and the Institutional Bond Markets". In: *Transforming Climate Finance and Green Investment with Blockchains*. Ed. by A. Marke. Academic Press, pp. 273–288.

Santos, F. A., T. H. Silva, A. A. Ferreira Loureiro, and L. A. Villas (2018). "Uncovering the Perception of Urban Outdoor Areas Expressed in Social Media". In: *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pp. 120–127.

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

Schmidt, C. G. and S. M. Wagner (2019). "Blockchain and supply chain relations: A transaction cost theory perspective". In: *Journal of Purchasing and Supply Management* 25.4, p. 100552.

Schmidt, S., J. Kinne, S. Lautenbach, T. Blaschke, D. Lenz, and B. Resch (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". In: *Science of The Total Environment* 835, p. 155512.

Schwierzy, J., R. Dehghan, S. Schmidt, E. Rodepeter, A. Stoemmer, K. Uctum, J. Kinne, D. Lenz, and H. Hottenrott (2022). *Technology Mapping Using WebAI: The Case of 3D Printing*.

Sehra, S. S., J. Singh, and H. S. Rai (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., N. Kumar, and J. H. Park (2020). "Blockchain technology toward green IoT: Opportunities and challenges". In: *IEEE Network* 34.4, pp. 263–269.

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

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

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

Sotarauta, M. and N. Suvinen (2019). "Place leadership and the challenge of transformation: policy platforms and innovation ecosystems in promotion of green growth". In: *European Planning Studies* 27.9, pp. 1748–1767.

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

Stoll, C., L. Klaaßen, and U. Gallersdörfer (2019). "The Carbon Footprint of Bitcoin". In: *Joule* 3.7, pp. 1647–1661.

Stuart, T. and O. Sorenson (2003). "The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms". In: *Research policy* 32.2, pp. 229–253.

Treiblmaier, H. (2018). "The impact of the blockchain on the supply chain: a theory-based research framework and a call for action". In: *Supply chain management: an international journal*.

Treiblmaier, H. and R. Beck (2019). *Business transformation through blockchain*. Vol. 1. Springer.

UN (2022). *Sustainable Development. The 17 Goals*. https://sdgs.un.org/goals. Accessed: 12.12.2022.

Vaughan, L., Y. Gao, and M. Kipp (2006). "Why are hyperlinks to business Websites created? A content analysis". en. In: *Scientometrics* 67.2. Publisher: Springer & Akadémiai Kiadó, pp. 291–300. (Visited on 04/28/2021).

Wang, X.-D., Z.-C. Luo, H. Liu, N. Zhao, M. Liu, Y.-F. Zhu, J.-P. Xue, A.-P. Luo, and W.-C. Xu (2015a). "Gold Nanorod as Saturable Absorber for Q-switched Yb-doped Fiber Laser". In: *Optics Communications* 346, pp. 21–25.

Wang, Y., Y. Chen, and J. Benitez-Amado (2015b). "How information technology influences environmental performance: Empirical evidence from China". In: *International Journal of Information Management* 35.2, pp. 160–170.

Williamson, P. J. and A. D. Meyer (2012). "Ecosystem Advantage: How to Successfully Harness the Power of Partners". In: *California Management Review* 55.1, pp. 24–46.

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

Appendix

Table A1. 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

Table A2. Mapping of Sustainable Development Goals and exemplary use of blockchain technology.

SDG	**Main goal (UN, 2022)**	**Related literature**	**Exemplary website** (accessed December 2022)
SDG 1	End poverty in all its forms everywhere	Schinckus, 2020; Treiblmaier and Beck, 2019	`www.skuchain.com` `www.stellar.org`
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`
SDG 3	Ensure healthy lives and promote well-being for all at all ages	Schinckus, 2020	`www.damcogroup.com`
SDG 4	Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all	-	`www.stellar.org`
SDG 5	Achieve gender equality and empower all women and girls	-	`www.stellar.org`
SDG 6	Ensure availability and sustainable management of water and sanitation for all	Treiblmaier and Beck, 2019	`www.waste2wear.com`
SDG 7	Ensure access to affordable, reliable, sustainable and modern energy for all	Sanderson, 2018; Schinckus, 2020; Blakstad and Allen, 2018; Sikorski et al., 2017; Hwang et al., 2017	`www.grcooling.com` `www.waste2wear.com`
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`
SDG 9	Build resilient infrastructure, promote inclusive and sustainable industrialisation and foster innovation	Schinckus, 2020	`www.stellar.org`
SDG 10	Reduce inequality within and among countries	Schinckus, 2020; Blakstad and Allen, 2018; Treiblmaier and Beck, 2019	`www.10pearls.com`
SDG 11	Make cities and human settlements inclusive, safe, resilient and sustainable	Schinckus, 2020; Treiblmaier and Beck, 2019	`www.waste2wear.com`
SDG 12	Ensure sustainable consumption and production patterns	Schinckus, 2020	`www.ripple.com` `www.waste2wear.com` `www.kleangas.com`
SDG 13	Take urgent action to combat climate change and its impacts	Sanderson, 2018; Schinckus, 2020; Blakstad and Allen, 2018; Sikorski et al., 2017; Hwang et al., 2017	`www.skuchain.com` `www.waste2wear.com` `www.energyweb.org` `www.nori.com` `www.kleangas.com`
SDG 14	Conserve and sustainably use the oceans, seas and marine resources for sustainable development	Schinckus, 2020	`www.ondiflo.com` `www.waste2wear.com`

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.ondiflo.com www.waste2wear.com www.nori.com www.kleangas.com
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
SDG 17	Strengthen the means of implementation and revitalise the Global Partnership for Sustainable Development	Sanderson, 2018; Schinckus, 2020; Blakstad and Allen, 2018; Sikorski et al., 2017; Hwang et al., 2017	www.waste2wear.com www.nori.com

Table A3. 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		

Download ZEW Discussion Papers:

https://www.zew.de/en/publications/zew-discussion-papers

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

//
IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim
ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de