LUND UNIVERSITY

CIRCLE

CENTRE FOR INNOVATION RESEARCH

Not all twins are identical: the digital layer of "twin" transition market applications

Milad Abbasiharofteh, Lukas Kriesch

Papers in Innovation Studies no. 2024/16

Not all twins are identical: the digital layer of “twin” transition market applications

Milad Abbasiharofteh[1*], Lukas Kriesch[2]
[1]Department of Economic Geography, Faculty of Spatial Sciences, University of Groningen, Landleven 1 9747 AD Groningen the Netherlands.
[2] Department of Geography, Justus Liebig University Giessen, Senckenbergstr. 1, 35390 Giessen, Germany.
* Corresponding author
Milad Abbasiharofteh, m.abbasiharofteh@rug.nl
Lukas Kriesch, Lukas.Kriesch@geogr.uni-giessen.de

Abstract

A twin (a joint green and digital) transition aims to facilitate achieving the Green Deal goals. The interplay between regional capabilities and twin transition market applications remains understudied. This research utilizes Large Language Models to analyze web texts of more than 600,000 German firms, assessing whether their products contribute to the twin transition. Our findings suggest while AI capabilities benefit the twin transition market applications, clean technological capabilities play a significant role only in highly specialized regions. To facilitate future research and informed policymaking, we provide open access to our developed dataset and AI tools (i.e., the TwinTransition Mapper).

Keywords: Twin transition, TwinTransition Mapper, digital layer, technological capabilities

JEL Codes: O30, O31, C81, C88

1. Introduction

Climate change and related societal issues have made green transition a priority in policy circles. The European Commission recently emphasized that a combined green and digital transition (a twin transition, hereafter) is essential for achieving the Green Deal goals (Muench et al., 2022). The concept of the twin transition has attracted a lot of attention as policymakers aim to ensure that the ongoing green and digital transitions are aligned and mutually reinforcing (Bachtrögler-Unger, Balland, Boschma, & Schwab, 2023). Providing informed policy direction can create "windows of opportunity" for a just transition (European Commission, 2021; Lema, Fu, & Rabellotti, 2020). However, this term has recently emerged in the policy world. Therefore, pressing open questions still need to be answered to allocate resources effectively. For instance, there is little empirical evidence on the fundamental assumption of the twin transition that two separate innovations—green and digital—must co-exist within the same regions, leading to twin transition market applications. Thus, it is unclear whether massive concurrent investments in clean and digital technologies will translate into more sustainable and digital market applications, especially in less-developed regions.

Regional studies and green transitions literature substantially inform policies and improve our understanding of how regions specialize in new economic activities (Boschma, 2016; Boschma, Balland, & Kogler, 2015; Boschma, Miguelez, Moreno, & Ocampo-Corrales, 2023; Frenken, van Oort, & Verburg, 2007; Neffke, Henning, & Boschma, 2011). This extant literature suggests that regions diversify into new activities that require similar capabilities to those already required by existing specializations. Capabilities refer to a broad set of organizational routines for developing, integrating, utilizing, and adapting internal and external resources (Teece, Pisano, & Shuen, 1997). Boschma (2024) defines regional capabilities as knowledge, institutions, and networks. So far, empirical investigations primarily focus on technological capabilities and, to a lesser extent, on scientific ones (Balland & Boschma, 2022; Boschma, Heimeriks, & Balland, 2014). Far less attention has been paid to the application or market side (Tödtling, Trippl, & Frangenheim, 2020). Breznitz (2021) criticizes this one-sided focus on technologies that may nudge policymaking into "technological fetishism". That is the excessive focus on high-tech solutions to drive economic growth, ignoring the real needs of many regions, particularly peripheral ones outside major technology hubs. Binz and Castaldi (2024) propose a normative turn in the geography of innovation research that besides focusing on technological innovation, considers market dynamics and applications. Similarly, policy circles have started to discuss regional capabilities in market applications recently, as explicitly indicated in the recent Future of European Competitiveness report (European Commission, 2024, page 2):

> "*The problem is not that Europe lacks ideas or ambition. We have many talented researchers and entrepreneurs filing patents. But innovation is blocked at the next stage: we are failing to translate innovation into commercialisation…*"

Another open question is the role of regional context in the twin transition. Rodríguez-Pose and Bartalucci (2024) argue that specialized regions build on their existing related capabilities and adopt green economic activities, whereas less specialized ones lag behind. Previous research has often ignored context's critical relevance, concluding that transition occurs anywhere via similar processes (Coenen, Benneworth, & Truffer, 2012). Recently, several journals' special issues (e.g., Faggian, Marzucchi, & Montresor, 2024) and publications shed new light on the regional dimensions of a twin transition. These contributions study a twin transition's scientific, technological, and policy aspects have been studied recently (Bachtrögler-Unger et al., 2023; Damioli, Bianchini, & Ghisetti, 2024; Brueck, Losacker & Liefner, 2024; Cicerone, Losacker & Ortega-Argilés, 2024). However, no study investigated the twin transition market applications. In the recent editorial of Regional Studies' special issue on the twin transition, Faggian et al. (2024) suggest future research on twin transition by indicating that "*... scant are the works that focus on the actual adoption of twin innovation into single products, processes or services. This appears to be a higher class of twinning that may be characterized by different opportunities and challenges for economic agents and territories alike, which we hope future works may investigate.*" (page 5). Thus, this explorative study seeks to answer the following questions: Do regions' clean and artificial intelligence (AI) technological capabilities positively associate with regional twin transition market applications? Do regions' clean and AI technological capabilities equally associate with twin transition market applications?

To answer the research questions, we approximate whether firms' products contribute to the twin transition. Recent studies scraped firms' web text coupled with advanced machine learning techniques to extract information about goods and services of geolocated firms (Abbasiharofteh, Kinne, & Krüger, 2023; Abbasiharofteh, Krüger, Kinne, Lenz, & Resch, 2023; Nathan & Rosso, 2022; Kriesch 2023). We build on this recent approach and use Large Language Models (LLMs) to develop and validate machine learning algorithms (i.e., the TwinTransition Mapper) that classify the complete population of 600,000 German firms with at least one website (i.e., the digital layer). The TwinTransition Mapper suggests whether each firm provides goods and services related to green and AI products (the algorithms are publicly available at https://bit.ly/3AMgiYs).

By aggregating this information at the regional level, we approximate the rate of regions' twin transition market applications. Relying on the web text and machine learning techniques has several advantages. First, it is not too far-fetched to argue that firms update their website instantly to showcase their new products. Thus, the TwinTransition Mapper can receive market signals much faster than trademark data, subject to a time-consuming examination process (Abbasiharofteh, Castaldi, & Petralia, 2022). Second, this approach includes a substantial share of firms with at least 25 employees compared to firms filing for patents and trademarks (Kinne & Axenbeck, 2020). It is crucial for this study because ecologically innovative firms are often overlooked by trademark data (Block, Lambrecht, Willeke, Cucculelli, & Meloni, 2023). Third, the TwinTransition Mapper can be scaled up and utilized in other contexts in contrast to survey-based techniques. Lastly, the application of our algorithms is not limited to web text data. Researchers can potentially use them in regional studies to classify green and AI job postings, social media content, digitized news archives, and patent and trademark textual data.

By estimating a set of beta and spatial regressions, our results suggest while regions benefit from AI technological capabilities, the association between clean technological capabilities and twin transition is not straightforward. Particularly, capabilities in clean technologies are positively associated with twin transition only in highly specialized regions. Our findings cast doubt on the policy sentiment that investing heavily in clean technologies in less-specialized areas would automatically trigger eco-friendly and smart local economies.

The remainder of this paper is structured as follows. Section 2 reviews relevant literature on how regions diversify into new economic activities. Section 3 discusses our methodology, including a detailed account of how we developed the TwinTransition Mapper and the digital layer of twin transition applications. Section 4 reports the results and discusses the main findings concerning twin transition policies. Section 5 concludes the paper and suggests several avenues for further research.

2. Place-based pathways to a green and twin transitions

Marshall (1890), in his *Principles of Economics,* discussed the crucial micro-foundations of agglomeration externalities: sharing, matching, and learning (Marshall, 1890). A plethora of empirical studies suggest that "knowledge-sourcing" does not fully occur within a company and benefits from inter-firm relations (Cohen & Levinthal, 1990; Frigon & Rigby, 2022; Von Hippel, 1987). These knowledge-sourcing relations bridge short distances and remain mostly local, perhaps due to the cost and complexity of absorbing tacit knowledge (Audretsch & Feldman, 1996; Storper & Venables, 2004). Evolutionary economic geographers extended the work of Marshal (1890) and other innovation models like Jacobian diversification externalities (Jacobs, 1969), learning regions (Morgan, 1997), regional innovation milieu (Maillat, 1995), and regional innovation systems (Asheim & Gertler, 2006). Evolutionary scholars argue that the emergence of new technologies in regions is not random, and it is often a result of combining existing knowledge and materials (Arthur, 2009; Boschma, 2016; Hidalgo, 2015).

The seminal work of Hidalgo et al. (2007) provides a framework for researchers to investigate the process of regional diversification. The principle of relatedness framework suggests that a region diversifies in new industries, technologies, and occupations that require capabilities similar to those activities already in the region (Hidalgo et al., 2018). Accordingly, scholars view regional diversification as a branching process in which new activities emerge in a region that has already been home to related ones (Boschma et al., 2015; Boschma, 2016; Frenken & Boschma, 2007).

This place-based approach attracted attention in sustainability transition studies. Truffer & Coenen (2012) and Truffer et al. (2015) criticized the early research in sustainability transitions, which ignored spatial dimensions. As a result, more recently, there have been conceptual works of economic geographers and transition studies scholars investigating the geography of green transitions (Boschma, Coenen, Frenken, & Truffer, 2017; Truffer et al., 2015). A surge of empirical studies followed these scholarly works. For instance, an investigation of 95 European regions suggests that green diversification in regions is linked to related capabilities (Santoalha & Boschma, 2021). Similarly, Montresor & Quatraro (2020) show that green technologies influence future specialization in a path-dependent process. Grashof & Basilico (2024) also studied green diversification along two dimensions: the economic strength of regions and the characteristics of the regional knowledge base. Their result suggests that regions can successfully diversify into green technologies if they have specialized in related technological capabilities unless they are specialized in dirty industries. Van den Berge et al. (2020) found that cleantech knowledge production thrives in regions with related technological bases. They also show that fossil fuel specializations neither hinder nor promote cleantech knowledge production. However, some fossil technologies provide inputs for cleantech knowledge production, with certain organizations transitioning into cleantech.

Another strand of scholarly works investigates the impact of Key Enabling Technologies (KETs) on regional diversifications (Janssen & Abbasiharofteh, 2022; Montresor & Quatraro, 2017). The relevance of KETs (e.g., advanced materials and manufacturing) is perhaps rooted in the general-purpose nature of them. Goldfarb et al. (2023) studied 21 emerging technologies and showed that AI technologies related to machine learning and data science are among general-purpose technologies. Cockburn et al. (2018) suggest that AI technologies like robotics and deep learning not only have the potential to significantly enhance the efficiency of current economic activities but also may have a more profound effect by acting as a new general-purpose "method of invention".

More related to a green transition, empirical research suggests that KETs pave the way for transitions to sustainable technologies (Montresor & Quatraro, 2020). Among KETs, scholars study the impact of artificial intelligence (AI) on the green transition. Earlier research identified AI as a general-purpose technology revolutionizing various sectors and substantially contributing to sustainable development (Goldfarb et al., 2023; Petralia, 2020; Sachs et al., 2019). Another study provides empirical evidence of a positive association between firms' investment in AI and the adoption of environmental innovations (Montresor & Vezzani, 2023). Cicerone et al. (2023) showed that AI knowledge embedded in local firms promotes the green-tech specialization of regions if local firms have had a prior green-tech specialization.

Besides studies investigating the impact of AI and clean technological capabilities, several studies discuss the relevance of average regional specialization in a green transition. Hansmeier and Losacker (2024) detailed why specialization brings about more possibilities for a green transition. Specialized regions have favorable conditions for a green transition, such as human capital, infrastructure, and networks. One study reveals that regions with the strongest green technological capabilities are predominantly located in central and western Europe (Barbieri et al., 2023). This study identified a strong correlation between non-green and green regional fitness, suggesting that the underlying knowledge capabilities in these regions demonstrate significant complementarities.

Empirical evidence on the driving forces of a twin transition is far less than those of a green transition. A recent study investigates regions' green and digital knowledge scientific bases and shows that such regions are more likely to introduce new and better quality twin knowledge (Damioli et al., 2024). Using patent filings, Bachtrögler-Unger et al. (2023) investigated the technological aspect of a twin transition. They found that the contribution of structurally weak regions to a twin transition is marginal. Such structurally weak regions have diversification possibilities only in low-complexity green technologies.

To our knowledge, no study has investigated the link between capabilities in clean and AI technologies and the twin transition market applications. Researchers recently found an interplay between scientific and technological, and technological and market spaces (Castaldi & Drivas, 2023; Catalán, Navarrete, & Figueroa, 2022). Given the scarcity of empirical research on the relationship between technological capabilities and twin transition market applications, we adopt an exploratory approach in this paper. We develop three propositions based on the arguments and empirical findings discussed above.

Proposition 1. Regional capabilities in clean technologies embedded in local firms foster successful twin transition market applications.

Proposition 2. Regional capabilities in AI technologies embedded in local firms foster successful twin transition market applications.

Proposition 3: Regions' specialization positively associates with successful twin transition market applications.

3. Data and empirical setting

3.1. Mapping the digital layer of the twin transition

The 'big data revolution' has created new opportunities for alternative data sources. The digital layer consists of the web text of firms coupled with their geographic location (Abbasiharofteh, Krüger et al., 2023). Websites serve as platforms for companies to showcase their products (Gök, Waterworth, & Shapira, 2015). Recently, the use of firms' web text has attracted attention because it provides up-to-date information about the products of firms and represents a much greater share of firms and sectors compared to patent and trademark data (Abbasiharofteh, Kinne et al., 2023; Ashouri, Hajikhani, Suominen, Pukelis, & Cunningham, 2024; Bottai, Crosato, Domenech, Guerzoni, & Liberati, 2022; Hajikhani et al., 2022; Tranos, Carrascal-Incera, & Willis, 2022, Kriesch & Losacker. 2024). This subsection discusses how we created the digital layer and developed the TwinTransition Mapper to classify firms with AI and eco-friendly products (see Figure 1).

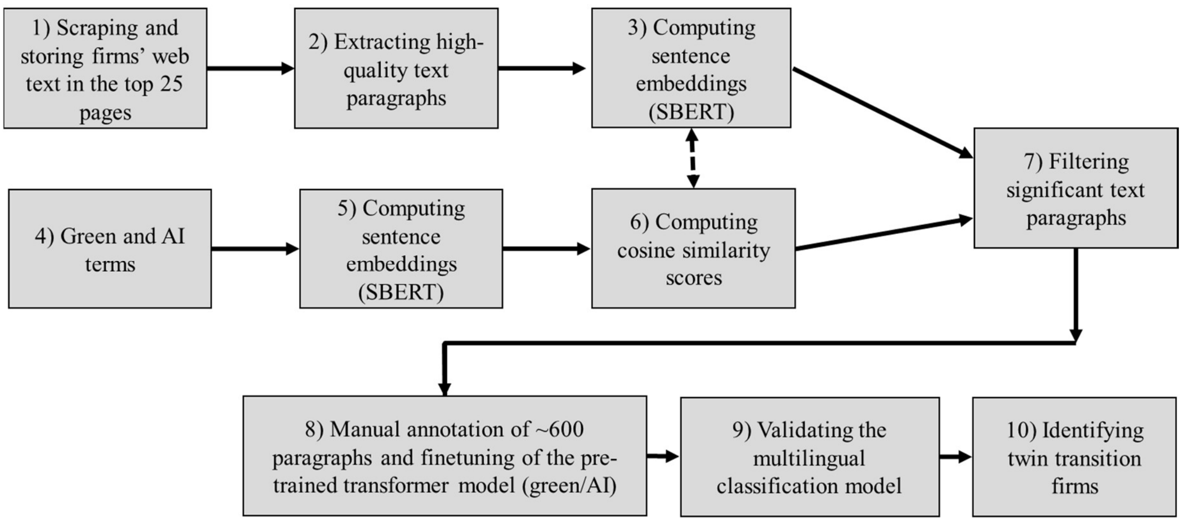

Figure 1. A schematic visualization of the twin transition digital layer.

We scrapped website texts from 678,381 German firms in 2023 (1). We scraped 9,601,260 subpages from the firm websites and extracted 106,046,851 paragraphs. We geocoded the firms based on the information in their imprint using a finetuned named entity recognition model (Kriesch, 2023).

Next, we cleaned the scraped web texts (2). We segmented the text into paragraphs and removed unnecessary content, such as menus, headers, footers, boilerplates, and advertisements. Moreover, we applied slightly modified versions of the quality filtering heuristics developed by Rae et al. (2021). These heuristics have been effective in preparing data for large language model training and are suitable for preparing our dataset for further analysis (see Appendix A). After these pre-processing steps, our dataset consists of 44,221,656 paragraphs.

To prepare the website texts for semantic filtering, we computed sentence embeddings of the extracted paragraphs (3). We used the “intfloat/multilingual-e5-large” model. This model was trained on more than 1 billion text pairs from various multilingual text resources like Wikipedia, news articles, research papers, and community discussions. The model supports 100 languages (Wang et al., 2024).

We utilized AI and Green terms (4) identified in trademark research as benchmarks for AI and eco-friendly goods and services (EUIPO, 2021; OECD, 2021). Using these terms reflects market realities because trademark examiners develop these terms based on trademark applications (for a review, see Abbasiharofteh et al., 2022). We used 71 AI terms and 43 eco-friendly terms. We provide the full list in Appendix B.

Similar to Step 3, we computed the embeddings of the green and AI terms (5).

We calculated cosine similarities between the vector embeddings of web text paragraphs and those of green and AI terms (6). Paragraphs are identified as twin transition related (7) if their cosine similarity values are greater than the 99th percentile of the cosine similarity values. This procedure resulted in 10,120,892 paragraphs.

Our observation suggests a substantial share of false positives. To remedy this situation, we manually annotated paragraphs to finetune two pre-trained transformer models (8). One model was designed to predict whether a text reflects AI products, while the other predicts whether a text reflects capabilities in clean technologies. We used the “intfloat/multilingual-e5-large” model as the baseline for both models. We employed an active learning approach to finetune this model effectively while minimizing the need for extensive manual annotations (Schröder & Niekler, 2020). This approach involved using the baseline model to assess a random selection of paragraphs and iteratively focusing on those with the lowest prediction certainty. This iterative annotation process conserves time and resources by selecting data points most likely to enhance the model’s predictive accuracy. For training the models, we utilized the SetFit framework (Sentence Transformer Finetuning), known for being a sample-efficient finetuning framework (Tunstall et al., 2022).

We applied both models on the filtered paragraphs from step (7). A firm was flagged as contributing to the twin transition if at least one paragraph was classified as reflecting the firm’s AI goods or services, and at least one other or the same paragraph was classified as reflecting the firm’s capabilities in clean technologies. We refer to the joint use of these two finetuned algorithms as the TwinTransition Mapper. To validate the TwinTransition Mapper (9), we randomly selected 500 web text paragraphs and manually annotated whether the web texts showcase a firm’s goods or services related to twin transition. Table 1 demonstrates the widely used four performance metrics (accuracy, precision, recall or sensitivity, and F1 score[1]).

[1] Accuracy measures the proportion of true results (both true positives and true negatives) among the total number of cases examined, indicating the overall correctness of the model. Precision calculates the ratio of true positive results to the total predicted positives, reflecting the model’s ability to correctly identify only relevant instances. Recall (Sensitivity) measures the ratio of true positive results to all actual positives, indicating the model’s ability to identify all relevant instances. The F1 score is the harmonic mean of precision and recall, providing a single metric that balances both concerns.

Our tests suggest that the TwinTransition Mapper scores high across all these metrics, suggesting this tool's reliability. The F1 score of 90% indicates a well-balanced combination of precision and recall (Manning, Raghavan, & Schütze, 2008).

Performance Metrics	Accuracy (95% CI)	Precision	Recall (Sensitivity)	F1 Score
Scores	0.897 (0.8668, 0.9223)	0.984	0.837	0.904

Table 1. The performance metrics of the TwinTransition Mapper.

The TwinTransition Mapper identified 23,819 German firms with AI-related and eco-friendly products (10). Figure 2 shows the web text of a twin transition firm in Aachen. Figure 3 shows the share of twin transition firms across 402 Nuts3 regions in Germany. Regions with a high share of twin transition firms are notably large independent cities like Darmstadt, Jena, Karlsruhe, Munich, and Frankfurt. These cities likely benefit from agglomeration economies and local institutions that drive innovation in AI and sustainability. Darmstadt, known as the "city of science," boasts a strong research and tech ecosystem. Karlsruhe, home to the Karlsruhe Institute of Technology, fosters growth in AI and eco-friendly industries through advanced research and supportive infrastructure. Jena, historically focused on optics and photonics, has evolved into a high-tech hub for biotechnology and optical technologies.

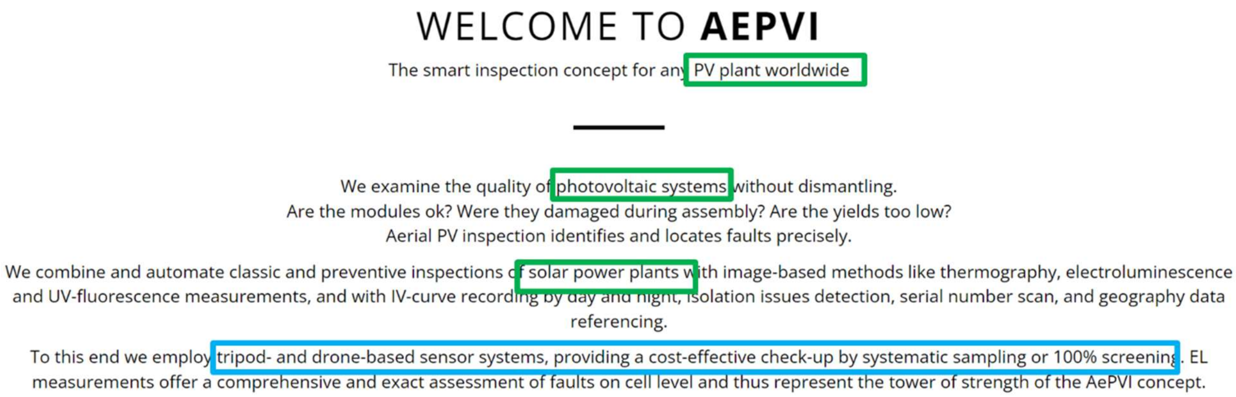

Figure 2. An example of a twin transition firm's web text.
Note: The terms highlighted in green and blue exemplify green and AI terms captured by the TwinTransition Mapper.

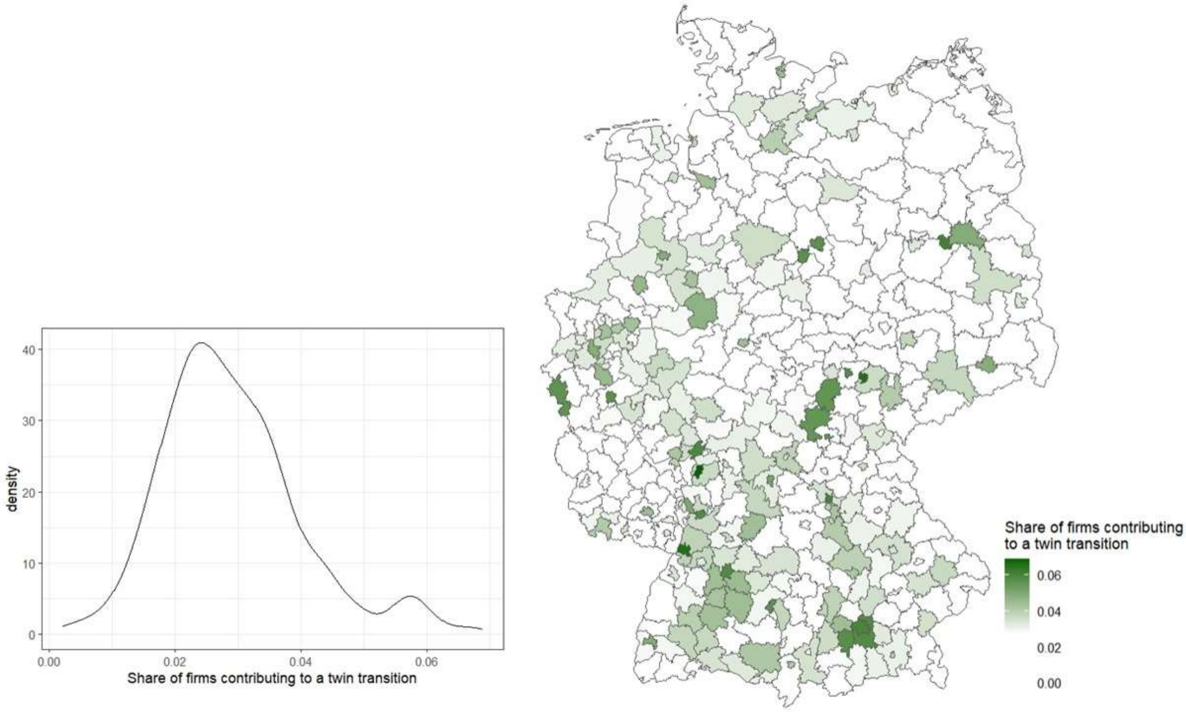

Figure 3. The statistical and spatial distribution of twin transition firms.

3.2. Variables

Using the results of the TwinTransition Mapper, we can approximate the extent to which firms in Nuts3 regions provide goods and services related to the twin transition in 2023. The dependent variable is *TwinTransition,* the number of twin transition firms divided by the total number of firms in the same spatial unit (see Figure 3).

Innovation scholars have extensively utilized patent databases to investigate technological capabilities and knowledge diffusion (Abbasiharofteh & Broekel, 2020; Bettencourt, Lobo, & Strumsky, 2007; Castaldi, Frenken, & Los, 2015; Jaffe, 1986; Jaffe, 1993). We used the OECD REGPAT database (version: 2023) which includes information on the location of inventors, filing date, and technologies (CPC codes) used in each technology (OECD, 2008).

To create variables approximating regional AI and clean technological capabilities, we selected patents filed between 2015 and 2022 with at least one CPC code (4-digit level) related to clean and AI technologies. The OECD ENV-TECH classification (Haščič & Migotto, 2015) suggests a list of clean technologies widely used to capture clean technological capabilities (e.g., Santoalha & Boschma, 2021). For AI technologies, we selected patents that include at least one of the eight CPC codes listed in Table 2. The rationale for this selection is that patents with these CPC codes entail AI terms (see supplementary materials) much more often than others. Accordingly, the variables *CleanTech* and *AI* are the relative number of patents with clean and AI technologies compared to the national level (i.e., location quotient).

Type	CPC codes	Labels
Cleantech	Y02A	Technologies for adaptation to climate change
Cleantech	Y02B	Climate change mitigation technologies related to buildings
Cleantech	Y02C	Capture, storage, sequestration or disposal of greenhouse gases
Cleantech	Y02D	Climate change mitigation technologies in information and communication technologies
Cleantech	Y02E	Reduction of greenhouse missions, related to energy generation, transmission or distribution
Cleantech	Y02P	Climate change mitigation technologies in the production or processing of goods
Cleantech	Y02T	Climate change mitigation technologies related to transportation
Cleantech	Y02W	Climate change mitigation technologies related to wastewater treatment or waste management
AI	G06N	Computer systems based on specific computational models
AI	G05B	Control or regulating systems in general; functional elements of such systems; monitoring or testing arrangements for such systems or elements
AI	G06K	Digital computers in which all the computation is effected mechanically
AI	G06T	Image data processing or generation
AI	G10H	Electrophonic musical instruments
AI	B60G	Vehicle suspension arrangements
AI	F05B	Indexing scheme relating to machines or engines other than non-positive-displacement machines or engines
AI	F16H	Gearing

Table 2. A list of clean and AI technology codes.

The variable *Specialization* is a proxy for each region's specialization level. This variable is based on the method developed by Hidalgo et al. (2007) and is often used in the related and unrelated diversification literature. In particular, we measured the related density of each CPC code (4-digit level) within regions and calculated the average related density for each region. Earlier research used this method to estimate regions' average specialization (Abbasiharofteh, Kogler, & Lengyel, 2023; van der Wouden & Rigby, 2019).

We created two control variables. The dummy variable *East* controls the fundamental socioeconomic differences between eastern and western German regions (Abbasiharofteh & Broekel, 2020; Fritsch & Graf, 2010). This variable takes the value of one if a region is one of the former DDR *Bundesländern* and takes the value of zero otherwise. *PopDensity* denotes the population density of German regions retrieved from Eurostat. Table 3 provides the descriptive statistics and Pearson correlation coefficients of the variables.

	N	Mean	St. Dev.	Min	Max	1)	2)	3)	4)	5)	6)
1) TwinTransition	402	0.029	0.011	0.002	0.069	1.00					
2) CleanTech	402	1.076	0.797	0.000	7.639	-0.11	1.00				
3) AI	402	0.762	0.676	0.000	6.553	0.27	0.07	1.00			
4) Specialization	402	0.203	0.069	0.034	0.375	0.42	-0.26	0.13	1.00		
5) East	402	0.192	0.394	0	1	-0.15	-0.01	-0.07	-0.38	1.00	
6) PopDensity	402	5.631	1.106	3.591	8.459	0.54	-0.05	0.21	0.31	-0.24	1.00

Table 3. The descriptive statistics and Pearson correlation coefficients of the variables.

4. Results and discussion

The dependent variable is bounded between zero and one, significantly violating the key assumptions of linear modeling. To remedy this predicament, statisticians developed beta regression models designed for dependent variables that represent rates, proportions, or concentration indices (Cribari-Neto & Zeileis, 2010). The advantage of the beta regression model is its ability to accommodate left- or right-skewed density shapes, depending on the combination of parameter values. Cribari-Neto and Zeileis (2010) formally define the beta density as:

$$f(y;p,q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{(p-1)}(1-y)^{(q-1)} \quad , \; 0 < y < 1, \tag{1}$$

where $p,\ q > 0$ and $\Gamma(.)$ is a gamma function. Ferrari and Cribari-Neto (2004) suggest an alternative parameterization by setting $\mu = p/(p+q)$ and $\Phi = p+q$:

$$f(y;\mu,\Phi) = \frac{\Gamma(\Phi)}{\Gamma(\mu\Phi)\Gamma((1-\mu)\Phi)} y^{\mu\Phi-1}(1-y)^{(1-\mu)\Phi-1} \quad , \; 0 < y < 1, \tag{2}$$

with $0 < \mu < 1$ and $\Phi > 0$[2].

Table 4 shows the results of beta regression models. We introduced the variables of interest in a stepwise manner. Model 4 represents the full model. The Akaike information criterion (AIC) values suggest that the goodness of fits of models do not substantially differ, with the full model providing the best fit.

Contrary to our theoretical argument, we observe a negative association between capabilities in clean technologies and the share of twin transition companies in German regions in Model 1, and the coefficient does not remain statistically significant in the full model. This finding does not align with *Proposition 1*. This observation points toward a potentially more complex process of translating technological solutions into the market. It is plausible that clean technology capabilities may have been developed recently, and their effects on eco-friendly products require much longer time and might not yet be observable. Moreover, we speculate that regional context is crucial in technology-market relations. For instance, differences in regional policies, regulations, or market demand for eco-friendly products can have hampering or catalytic effects.

[2] We conducted beta regressions with the betareg R-package.

Conversely, the results suggest that AI technological capabilities are positively associated with the share of regions' companies contributing to the twin transition. It supports *Proposition 2*. We conjecture that although clean tech may fundamentally transform sectors over a long time period, AI technologies often improve operational efficiency through data analytics, optimization, personalization, and automation. This efficiency may more quickly lead to AI-related and eco-friendly products.

	Dependent variable: TwinTransition			
	(1)	(2)	(3)	(4)
CleanTech	-0.06**			-0.03
	(0.02)			(0.02)
AI		0.08***		0.08***
		(0.02)		(0.02)
Specialization			2.01***	1.91***
			(0.25)	(0.26)
East	-0.07	-0.07	0.03	0.03
	(0.05)	(0.05)	(0.04)	(0.04)
PopDensity	0.17***	0.16***	0.15***	0.14***
	(0.01)	(0.02)	(0.01)	(0.01)
(phi)	302.71***	307.83***	347.29***	360.72***
	(21.53)	(21.89)	(24.67)	(25.62)
Constant	-4.40***	-4.48***	-4.78***	-4.74***
	(0.09)	(0.09)	(0.09)	(0.10)
Observations	402	402	402	402
AIC	-2619.95	-2626.59	-2673.95	-2684.85
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 4. Results of beta regressions.

The level of specialization of regions is positively associated with the dependent variable. This finding supports *Proposition 3* and earlier research that specialized regions are better off, perhaps because they are well-endowed with human capital and needed institutions and networks (Rodríguez-Pose & Bartalucci, 2024). These results further motivate an investigation of the effects of specialization as a contextual factor on the relationship between AI and clean technological capabilities and the share of twin transition firms. We dichotomized the three variables to specifically distinguish between highly specialized regions that host a high level of AI and clean technological capabilities and the rest. This dichotomization makes interpreting the coefficients of interaction terms easier, and we avoid multicollinearity issues in the models. Accordingly, *Specialization (Dummy)* takes the value of one if the specialization of a region is greater than the 75th percentile of the *Specialization* vector. Otherwise, it takes the value of zero. Similarly, we introduced the dichotomized version of *CleanTech* and *AI*. Table 5 shows the results of beta regressions with and without interaction terms.

The results suggest that the coefficients of the original variables and their dummy versions are consistent, except for the ones of *CleanTech* that are negative and significant. The models with interaction terms, *CleanTech (Dummy)* × *Specialization (Dummy)*, suggest that cleantech capabilities positively associate with a share of twin transition firms only in highly specialized regions and benefit from a high level of capabilities in clean technology. Alternatively, we estimated models including *CleanTech* × *Specialization (Dummy),* and the corresponding coefficient was not statically significant, suggesting only regions with high-level specialization and capabilities in clean technologies enjoy the growth in the number of twin transition firms. The marginal effects plots (Figure 4) visualize the interaction between *CleanTech* and *AI*, and *Specialization* on *TwinTransition*.

	Dependent variable: TwinTransition		
	(1)	(2)	(4)
CleanTech (Dummy)	-0.08^{**}	-0.12^{***}	-0.08^{**}
	(0.04)	(0.04)	(0.04)
AI (Dummy)	0.13^{***}	0.13^{***}	0.10^{**}
	(0.04)	(0.04)	(0.04)
Specialization (Dummy)	0.17^{***}	0.14^{***}	0.13^{***}
	(0.04)	(0.04)	(0.04)
CleanTech (Dummy) × Specialization (Dummy)		0.19^{**}	
		(0.09)	
AI (Dummy) × Specialization (Dummy)			0.13^{*}
			(0.08)
East	-0.06	-0.06	-0.06
	(0.04)	(0.04)	(0.04)
PopDensity	0.15^{***}	0.15^{***}	0.15^{***}
	(0.01)	(0.01)	(0.01)
(phi)	328.44^{***}	331.92^{***}	330.82^{***}
	(23.34)	(23.59)	(23.51)
Constant	-4.43^{***}	-4.41^{***}	-4.42^{***}
	(0.09)	(0.09)	(0.09)
Observations	402	402	402
AIC	-2647.91	-2650.03	-2648.87
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table 5. Results of beta regressions with and without interaction terms.

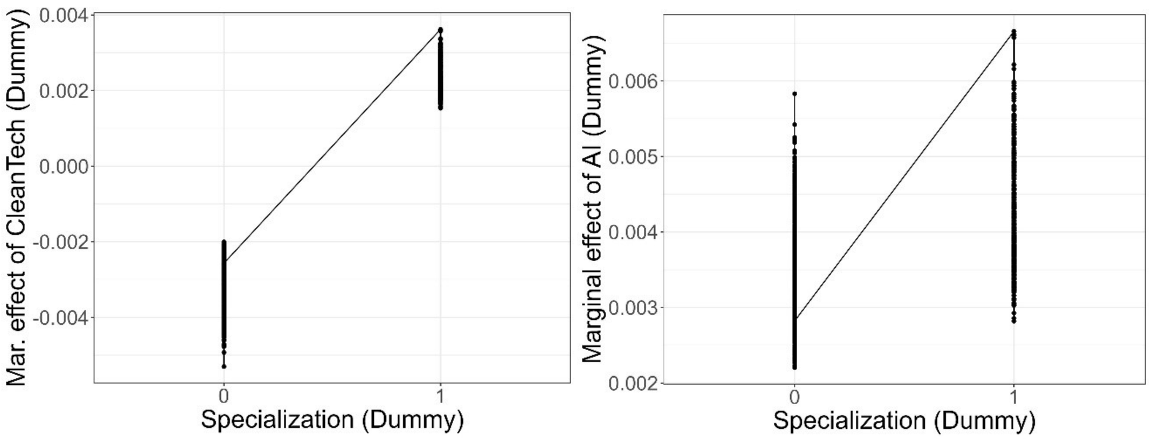

Figure 4. Marginal effects of *CleanTech* and *AI* on *TwinTransition* across specialization levels.

To ensure the robustness of our results, we retrieved TED (Tenders Electronic Daily) data on European public procurement[3]. For each region, we counted procurements that were categorized under the broad themes of "environment", "sustainability", and "ecology". This measure can be a proxy for policy demand pull stimulating environmental innovation. We also created variables based on a list of regions eligible for funding from the European Regional Development Fund to make a distinction between central and peripheral regions[4]. This variable is strongly correlated with the variable *East,* given that regions on the eastern side are relatively less developed. The results remained robust after the inclusion of these variables[5].

Finally, Figure 3 suggests spatial clustering of the dependent variable (*Moran's I* statistic suggests a weak positive spatial autocorrelation: 0.069, p-value<0.001). We created a parametrized inverse distance-based weights (*W*) matrix[6]. We estimated all spatial models suggested by Elhorst (Elhorst, 2010). The Spatial Durbin model (SDM) and Spatial lag of *X* model (SLX) provided the best goodness of fit. The SDM model is formally defined as:

$$y_{r,t} = \phi + \rho \sum_{j=1}^{N} w_{ru}\, y_{u,t-1} + x_{r,t-1}\beta + \theta \sum_{u=1}^{N} w_{ru}\, x_{u,t-1} + c_r + \alpha_{t-1} + \varepsilon_{r,t-1}, \quad (3)$$

where $\sum_u w_{ru}\, y_{u,t-1}$ represents the endogenous interaction effects of the dependent variable $y_{r,t}$ with the temporally lagged dependent variables $y_{u,t-1}$ in neighboring regions. ρ is the response parameter for these endogenous interaction effects. Accordingly, $\sum_u w_{ru}\, x_{u,t-1}$ denotes the exogenous relation between the dependent variable $y_{r,t}$ and the temporally lagged independent variables in neighboring regions. θ denotes a set of parameters for these exogenous interaction effects. The SLX model is nested in the SDM model ($\rho = 0$ in Equation 3) (Elhorst, 2014). Table 6 provides coefficients of direct and indirect effects (effects of surrounding regions). While the results align with those of beta regression models, we found no spatially exogenous interaction effects of clean and AI technological capabilities and no spatially endogenous interaction effects of the dependent variable (ρ: 0.064433, p-value: 0.52947, in the SDM model).

[3] https://ted.europa.eu/TED/main/HomePage.do (accessed: 08.08.2024)
[4] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32021D1130 (accessed: 08.08.2024)
[5] The results of these models are available upon requests from the authors.
[6] A parametrized inverse distance-based weights matrix, compared to the contiguity matrix, it is not biased by how administrative boundaries are defined.

Dependent variable: TwinTransition		
	SDM	SLX
Direct effects		
CleanTech	-0.0006	-0.0006
AI	0.0018***	0.0018***
Specialization	0.0577***	0.0579***
East	0.0006	0.0007
PopDensity	0.0051***	0.0051***
Indirect effects		
CleanTech	0.0002	0.0002
AI	0.0012	0.0014
Specialization	0.0027	0.0019
East	0.0000	0.0000
PopDensity	-0.0044***	-0.0042***
Observations (spatial units)	402	402
AIC	-2668.749	-2669.586

Table 6. Results of spatial regressions with direct and indirect effects.
Note: *** denote significance at the 0.05 level. SDM denotes Spatial Durbin model, and SLX denotes spatial lag of X model.

All in all, our results cast doubt on the policy approach that heavy investments in clean technologies will, at least in the short run, trigger AI-related and eco-friendly products. AI and clean technologies differ and, as a result, are differently associated with the share of green transition firms. Our results echo Bachtrögler-Unger et al. (2023), who recommend that policymakers avoid policies supporting clean technologies for which their region lacks related capabilities.

Our results show a varying association between clean and AI technological capabilities and the twin transition of regional markets. This finding perhaps reflects that these two technologies are different: the former being driven by policy demand-pull and the latter predominantly by market pull. Our results remind future policy measures that clean and AI technologies (and perhaps digital technologies) are not identical twins. This distinction appears to be largely overlooked in twin transition policy reports and working papers (e.g., Maucorps, Römisch, Schwab, & Vujanović, 2023).

The current (2021-2027) European Cohesion Policy allocates €545 billion to make Europe greener and more inclusive[7]. The Research and Innovation Strategies for Smart Specialisation (RIS3) is a core component of the EU Cohesion Policy because it strengthens strategic programming and the efficiency of the European Structural Investment Funds (European Commission, 2023). This policy framework helps regions identify and build on their strengths, fostering innovation, economic growth, and competitiveness tailored to local contexts (Foray, 2014). The new generation of Smart Specialisation (Smart Specialisation Strategies for Sustainability–S4) aims to link smart specialization with sustainable development goals. However, to our knowledge, this policy does not consider the cross-space technology-market relationship and its potential to achieve the goals of S4 by triggering market applications in structurally weak regions.

[7] See: https://ec.europa.eu/regional_policy/policy/how/future-cohesion-policy_en (accessed: 13.08.2024)

Considering our results, it may be a more effective allocation of resources to invest in strengthening AI technological capabilities in less specialized regions. This policy can trigger a low-tech application of high-tech (AI) solutions for place-based sectoral problems instead of overshooting by aiming to fully transform such sectors into clean high-technology sectors in the short run. This approach aligns with the argument of Coenen and Morgan (2020), who suggest focusing on problems to prioritize place-based strategies. Shifting attention to the "geography of problems" provides great potential for "ordinary people" and shapes local markets (Bailey, Pitelis, & Tomlinson, 2023). Such problem-driven policies leverage the "chimney" effects rather than relying on top-down dissemination and more effectively facilitate a just transition across European regions (McCann & Soete, 2020; Sachs et al., 2019).

5. Conclusions

This study has taken an exploratory approach to assess the association between AI and clean technological capabilities of regions and twin transition market applications. Our study aims to lay the groundwork for future hypothesis development in this emerging field. Our findings highlight that AI technological capabilities generally benefit the twin transitions of regional markets. However, the impact of clean technological capabilities is positive only in highly specialized regions. Our results challenge the assumption that heavy investment in clean technologies alone will drive twin transition market applications in less-specialized places. Of course, we do not claim to unveil causal relations and discuss the direction of relations between relevant factors.

We acknowledge several limitations of our study. This study considered specialization to approximate regional context. We recognize that specialization captures only one aspect of regional contexts. Future studies should broaden this view by investigating the role of social acceptance, local institutions, quality of government, and administrative capacity in successful twin transition market applications (Bachtler, Polverari, Domorenok, & Graziano, 2024; Bachtrögler-Unger et al., 2023; Barbero et al., 2023).

It is important to note that we focused on AI and not digital technologies because digital technologies are broad, and identifying firms with digital products poses an empirical challenge for this study. Similarly, AI includes various technologies such as machine learning, natural language processing, computer vision, and robotics. Future studies must investigate how specific AI technologies can help solve place-specific green transition problems. For instance, regions with a substantial share of agricultural activities are most vulnerable to the twin transition (Maucorps et al., 2023). Such regions can benefit from developing AI technologies like precision farming and predictive analytics, leading to a 60% reduction in fertilizers and energy use (Bohnsack et al., 2022). Also, AI-driven drones enhance the monitoring of soil health and water usage.

Our study recommends that policymakers monitor and support both regions' technological and market application capabilities for a just twin transition. We hope to assist more informed policymaking by providing open access to the TwinTransition Mapper (available at https://bit.ly/3AMgiYs). The algorithm can be scaled up to classify a larger firm's population. Moreover, the use of our algorithms extends beyond web text data. One can potentially apply them to identify green and AI texts in other types of textual data: job postings, digitized news archives, and patent and trademark filings. Policymakers can use this AI tool dynamically to create the first European twin transition observatory. This twin transition observatory will help Europe to become climate-neutral by providing real-time inputs to investigate place-based twin transition trajectories and assess the impact of transition policies. One can easily merge the output at the regional level with other standard statistics provided by national and European statistical offices (e.g., Eurostat).

References

Abbasiharofteh, M., Castaldi, C., & Petralia, S.G. (2022). *From patents to trademarks: towards a concordance map* (Final report). European Patent Office (EPO).

Abbasiharofteh, M., & Broekel, T. (2020). Still in the shadow of the wall? The case of the Berlin biotechnology cluster. *Environment and Planning a: Economy and Space, 53*(1), 73-94. 10.1177/0308518X20933904

Abbasiharofteh, M., Kinne, J., & Krüger, M.F.C. (2023). Leveraging the digital layer: The strength of weak and strong ties in bridging geographic and cognitive distances, *24*(2), 241–262.

Abbasiharofteh, M., Kogler, D.F., & Lengyel, B. (2023). Atypical combinations of technologies in regional co-inventor networks. *Research Policy, 52*(10), 104886.

Abbasiharofteh, M., Krüger, M., Kinne, J., Lenz, D., & Resch, B. (2023). The Digital Layer: Alternative Data for Regional and Innovation Studies. *Spatial Economic Analysis,* 1–23. 10.1080/17421772.2023.2193222

Arthur, W.B. (2009). *The nature of technology: What it is and how it evolves / W. Brian Arthur*. London: Allen Lane.

Asheim, B., & Gertler, M. (2006). The geography of innovation: Regional innovation systems. In J. Fagerberg, D.C. Mowery, & R.R. Nelson (Eds.), *The Oxford Handbook of Innovation* (pp. 291–317). Oxford University Press.

Ashouri, S., Hajikhani, A., Suominen, A., Pukelis, L., & Cunningham, S.W. (2024). Measuring digitalization at scale using web scraped data. *Technological Forecasting and Social Change, 207*(5), 123618.

Audretsch, D., & Feldman, M.P. (1996). R&D spillovers and the geography of innovation and production. *American Economic Review, 86*(3), 630–640.

Bachtler, J., Polverari, L., Domorenok, E., & Graziano, P. (2024). Administrative capacity and EU Cohesion Policy: Implementation performance and effectiveness. *Regional Studies, 58*(4), 685–689.

Bachtrögler-Unger, J., Balland, P.-A., Boschma, R., & Schwab, T. (2023). *Technological capabilities and the twin transition in Europe: Opportunities for regional collaboration and economic cohesion*. Bertelsmann Stiftung.

Bailey, D., Pitelis, C.N., & Tomlinson, P.R. (2023). Place-based industrial and regional strategy – levelling the playing field. *Regional Studies, 57*(6), 977–983.

Balland, P.-A., & Boschma, R. (2022). Do scientific capabilities in specific domains matter for technological diversification in European regions? *Research Policy, 51*(10), 104594.

Barbero, J., Christensen, M., Conte, A., Lecca, P., Rodríguez-Pose, A., & Salotti, S. (2023). Improving Government Quality in the Regions of the EU and its System-Wide Benefits for Cohesion Policy. *JCMS: Journal of Common Market Studies, 61*(1), 38–57.

Barbieri, N., Consoli, D., Napolitano, L., Perruchas, F., Pugliese, E., & Sbardella, A. (2023). Regional technological capabilities and green opportunities in Europe. *The Journal of Technology Transfer, 48*(2), 749–778.

Bettencourt, L.M.A., Lobo, J., & Strumsky, D. (2007). Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size. *Research Policy, 36*(1), 107–120.

Block, J.H., Lambrecht, D., Willeke, T., Cucculelli, M., & Meloni, D. (2023). Ecological Patents and Trademarks as Indicators of Ecological Innovation. *Available at SSRN: Https://ssrn.Com/abstract=4483035 or Http://dx.Doi.Org/10.2139/ssrn.4483035*.

Boschma, R., Balland, P.-A., & Kogler, D.F. (2015). Relatedness and technological change in cities: The rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. *Industrial and Corporate Change, 24*(1), 223–250.

Boschma, R. (2016). Relatedness as driver of regional diversification: A research agenda. *Regional Studies, 51*(3), 351–364.

Boschma, R., Coenen, L., Frenken, K., & Truffer, B. (2017). Towards a theory of regional diversification: Combining insights from Evolutionary Economic Geography and Transition Studies. *Regional Studies, 51*(1), 31–45.

Boschma, R., Heimeriks, G., & Balland, P.-A. (2014). Scientific knowledge dynamics and relatedness in biotech cities. *Research Policy, 43*(1), 107–114.

Boschma, R., Miguelez, E., Moreno, R., & Ocampo-Corrales, D.B. (2023). The Role of Relatedness and Unrelatedness for the Geography of Technological Breakthroughs in Europe. *Economic Geography, 99*(2), 117–139.

Bottai, C., Crosato, L., Domenech, J., Guerzoni, M., & Liberati, C. (2022). Unconventional data for policy. In *Conference on Information Technology for Social Good* (pp. 338–344). New York, NY, USA: ACM.

Breznitz, D. (2021). *Innovation in real places: Strategies for prosperity in an unforgiving world / Dan Breznitz*. New York, NY: Oxford University Press.

Castaldi, C., & Drivas, K. (2023). Relatedness, Cross-relatedness and Regional Innovation Specializations: An Analysis of Technology, Design, and Market Activities in Europe and the US. *Economic Geography, 99*(3), 253–284.

Castaldi, C., Frenken, K., & Los, B. (2015). Related Variety, Unrelated Variety and Technological Breakthroughs: An analysis of US State-Level Patenting. *Regional Studies, 49*(5), 767–781.

Catalán, P., Navarrete, C., & Figueroa, F. (2022). The scientific and technological cross-space: Is technological diversification driven by scientific endogenous capacity? *Research Policy, 51*(8), 104016.

Cicerone, G., Faggian, A., Montresor, S., & Rentocchini, F. (2023). Regional artificial intelligence and the geography of environmental technologies: Does local AI knowledge help regional green-tech specialization? *Regional Studies, 57*(2), 330–343.

Cicerone, G., Losacker, S., Ortega-Argilés, R. (2024). Regional diversification into green and digital economic activities – The case of UK Local. MESCoE Discussion Paper No. 2024-07.

Cockburn, I., Henderson, R., & Stern, S. (2018). *The Impact of Artificial Intelligence on Innovation*. Cambridge, MA: National Bureau of Economic Research.

Coenen, L., Benneworth, P., & Truffer, B. (2012). Toward a spatial perspective on sustainability transitions. *Research Policy, 41*(6), 968–979.

Coenen, L., & Morgan, K. (2020). Evolving geographies of innovation: Existing paradigms, critiques and possible alternatives. *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography, 74*(1), 13–24.

Cohen, W.M., & Levinthal, D.A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly, 35*(1), 128.

Cribari-Neto, F., & Zeileis, A. (2010). Beta Regression in R. *Journal of Statistical Software, 34*(2).

Damioli, G., Bianchini, S., & Ghisetti, C. (2024). The emergence of a 'twin transition' scientific knowledge base in the European regions. *Regional Studies,* 1–17.

Elhorst, J.P. (2010). Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis, 5*(1), 9–28.

Elhorst, J.P. (2014). *SpringerBriefs in Regional Science*. *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels* (1st ed. 2014). Berlin Heidelberg: Springer Berlin Heidelberg; Imprint: Springer.

EUIPO. (2021). *Green EU trade marks: Analysis of goods and services specifications, 1996-2020*. Alicante: European Union Intellectual Property Office.

European Commission. (2021). *The just transition mechanism: Making sure no one is left behind*. https://ec. europa.eu/info/strategy/priorities-2019-2024/european-green- deal/finance-and-green-deal/just-transition-mechanism-en.

European Commission. (2023). *EU Cohesion Policy: European Structural and Investment Funds supported SMEs, employment of millions of people and clean energy production*. https://ec.europa.eu/commission/presscorner/api/files/document/print/en/ip_23_2462/IP_23_2462_EN.pdf.

Faggian, A., Marzucchi, A., & Montresor, S. (2024). Regions facing the 'twin transition': Combining regional green and digital innovations. Regional Studies, 1–7.

Ferrari, S., & Cribari-Neto, F. (2004). Beta Regression for Modelling Rates and Proportions. *Journal of Applied Statistics, 31*(7), 799–815.

Foray, D. (2014). From smart specialisation to smart specialisation policy. *European Journal of Innovation Management, 17*(4), 492–507.

Frenken, K., & Boschma, R.A. (2007). A theoretical framework for evolutionary economic geography: Industrial dynamics and urban growth as a branching process. *Journal of Economic Geography, 7*(5), 635–649.

Frenken, K., van Oort, F., & Verburg, T. (2007). Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies, 41*(5), 685–697.

Frigon, A., & Rigby, D.L. (2022). Where do capabilities reside?: Analysis of related technological diversification in multi-locational firms. *Regional Studies, 56*(12), 2045–2057.

Fritsch, M., & Graf, H. (2010). *How General Conditions Affect Regional Innovation Systems - The Case of the Two Germanys*. Jena: Jena Economic Research Papers 2010 - 054 JENA.

Gök, A., Waterworth, A., & Shapira, P. (2015). Use of web mining in studying innovation. *Scientometrics, 102,* 653–671.

Goldfarb, A., Taska, B., & Teodoridis, F. (2023). Could machine learning be a general purpose technology?: A comparison of emerging technologies using data from online job postings. *Research Policy, 52*(1), 104653.

Grashof, N., & Basilico, S. (2024). Divergent paths for green transition: Regional disparities in technology diversification. *Regional Studies, 86*(3), 1–17.

Hajikhani, A., Pukelis, L., Suominen, A., Ashouri, S., Schubert, T., Notten, A., & Cunningham, S.W. (2022). Connecting firm's web scraped textual content to body of science: Utilizing microsoft academic graph hierarchical topic modeling. *MethodsX, 9,* 101650.

Hansmeier, H., & Losacker, S. (2024). Regional eco-innovation trajectories. *European Planning Studies,* 1–22.

Haščič, I., & Migotto, M. (2015). *Measuring Environmental Innovation Using Patent Data* (Vol. 211). OECD.

Hidalgo, C.A., Balland, P.-A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., . . . Zhu, S. (2018). The Principle of Relatedness. In A.J. Morales, C. Gershenson, D. Braha, A.A. Minai, & Y. Bar-Yam (Eds.), *Springer Proceedings in Complexity. Unifying Themes in Complex Systems IX: Proceedings of the Ninth International Conference on Complex Systems* (pp. 451–457). Cham: Springer International Publishing.

Hidalgo, C.A., Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science (New York, N.Y.), 317*(5837), 482–487.

Hidalgo, C.A. (2015). *Why information grows: The evolution of order, from atoms to economies*. London: Penguin Books.

Jacobs, J. (1969). *Vintage books: Vol. 584. The economy of cities*. New York: Vintage Books.

Jaffe, A.B. (1986). Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *American Economic Review, 76*(5), 984–1001.

Jaffe, A.M.T.H.R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics, 108*(3), 577–598.

Janssen, M.J., & Abbasiharofteh, M. (2022). Boundary spanning R&D collaboration: Key enabling technologies and missions as alleviators of proximity effects? *Technological Forecasting and Social Change, 180*(7), 121689.

Kinne, J., & Axenbeck, J. (2020). Web mining for innovation ecosystem mapping: A framework and a large-scale pilot study. *Scientometrics, 28*(4), 443. 10.1007/s11192-020-03726-9

Kriesch, L. (2023). Web Mining und Natural Language Processing als methodisches Komplement in der Wirtschaftsgeographie.

Kriesch, L., & Losacker, S. (2024). Bioeconomy firms and where to find them. *REGION*, *11*(1), 55-78.

Lema, R., Fu, X., & Rabellotti, R. (2020). Green windows of opportunity: Latecomer development in the age of transformation toward sustainability. *Industrial and Corporate Change, 29*(5), 1193–1209.

Maillat, D. (1995). Territorial dynamic, innovative milieus and regional policy. *Entrepreneurship & Regional Development, 7*(2), 157–165.

Manning, C.D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.

Marshall, A. (1890). *Principles of Economics*. London: MacMillan.

Maucorps, A., Römisch, R., Schwab, T., & Vujanović, N. (2023). The Future of EU Cohesion - Effects of the Twin Transition on Disparities across European Regions. . *Wiiw Research Report No. 467. Vienna Institute for International Economic Studies, in Cooperation with Bertelsmann Stiftung.*

Mccann, P., & Soete, L. (2020). *Place-based innovation for sustainability*. Luxembourg: Publications Office of the European Union.

Montresor, S., & Quatraro, F. (2017). Regional Branching and Key Enabling Technologies: Evidence from European Patent Data. *Economic Geography, 93*(4), 367–396.

Montresor, S., & Quatraro, F. (2020). Green technologies and Smart Specialisation Strategies: A European patent-based analysis of the intertwining of technological relatedness and key enabling technologies. *Regional Studies, 54*(10), 1354–1365.

Montresor, S., & Vezzani, A. (2023). Digital technologies and eco-innovation. Evidence of the twin transition from Italian firms. *Industry & Innovation, 30*(7), 766–800.

Morgan, K. (1997). The Learning Region: Institutions, Innovation and Regional Renewal. *Regional Studies, 31*(5), 491–503.

Muench, S., Stoermer, E., Jensen, K., Asikainen, T., Salvi, M., & Scapolo, F. (2022). *Towards a green and digital future*. Luxembourg: Publications Office of the European Union.

Nathan, M., & Rosso, A. (2022). Innovative events: Product launches, innovation and firm performance. *Research Policy, 51*(1), 104373.

Neffke, F., Henning, M., & Boschma, R. (2011). How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. *Economic Geography, 87*(3), 237–265.

OECD. (2008). *OECD Science, Technology and Industry Working Papers* (2008/02).

OECD. (2021). *Artificial intelligence companies, goods, and services: A trademark based analysis* (2021/06). OECD Science, Technology and Industry Working Papers 2021/06.

Petralia, S. (2020). Mapping general purpose technologies with patent data. *Research Policy, 49*(7), 104013.

Rodríguez-Pose, A., & Bartalucci, F. (2024). The green transition and its potential territorial discontents. *Cambridge Journal of Regions, Economy and Society, 17*(2), 339–358.

Sachs, J.D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., & Rockström, J. (2019). Six Transformations to achieve the Sustainable Development Goals. *Nature Sustainability, 2*(9), 805–814.

Santoalha, A., & Boschma, R. (2021). Diversifying in green technologies in European regions: Does political support matter? *Regional Studies, 55*(2), 182–195.
Storper, M., & Venables, A.J. (2004). Buzz: Face-to-face contact and the urban economy. *Journal of Economic Geography, 4*(4), 351–370. http://dx.doi.org/10.1093/jnlecg/lbh027
Teece, D.J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18*(7), 509–533.
Tödtling, F., Trippl, M., & Frangenheim, A. (2020). Policy options for green regional development: Adopting a production and application perspective. *Science and Public Policy, 47*(6), 865–875.
Tranos, E., Carrascal-Incera, A., & Willis, G. (2022). Using the Web to Predict Regional Trade Flows: Data Extraction, Modeling, and Validation. *Annals of the American Association of Geographers, 13*(1), 1–23.
Truffer, B., & Coenen, L. (2012). Environmental Innovation and Sustainability Transitions in Regional Studies. *Regional Studies, 46*(1), 1–21.
Truffer, B., Murphy, J.T., & Raven, R. (2015). The geography of sustainability transitions: Contours of an emerging theme. *Environmental Innovation and Societal Transitions, 17*(4), 63–72.
van den Berge, M., Weterings, A., & Alkemade, F. (2020). Do existing regional specialisations stimulate or hinder diversification into cleantech? Environmental Innovation and Societal Transitions, 35(1), 185–201.
van der Wouden, F., & Rigby, D.L. (2019). Co-inventor networks and knowledge production in specialized and diversified cities. *Papers in Regional Science, 98*(4), 1833–1853.
VonHippel, E. (1987). Cooperation between rivals: Informal know-how trading. *Research Policy, 16,* 291–302.

6. Appendix A. Text quality filtering

We adapted the text quality filtering heuristics established by Rae et al. (2021) to better suit the nuances of the German language, implementing the following modifications to the filtering process:

- Paragraphs are excluded if the average word length falls outside the 3-to-12-character range.
- Paragraphs are eliminated if the ratio of symbols to words exceeds 0.15.
- Paragraphs are discarded if they contain fewer than two stop words from either English or German.
- Any paragraph composed entirely of uppercase letters is also removed from consideration, as this often signifies non-standard text or spam.

7. Appendix B. Green and AI terms

8. Table B.1. AI-terms.

action recognition	evolutionary algorithm	object detection
artificial intelligence	expert system	optical character recognition
artificial neural network	face recognition	pattern recognition
association rule	feature engineering	predictive analytics
autoencoder	fingerprint recognition	probabilistic
automatic number plate recognition	fuzzy logic	random forest
autonomic computing	generative adversarial network	recommender system
autonomous vehicle	genetic algorithms	reinforcement learning
bayesian networks	gesture recognition	robotics
brain computer interface	gradient boosting	sensor fusion
classifier	image recognition	sentiment analysis
clustering	independent component analysis	speech recognition
cognitive computing	inductive logic programming	supervised learning
collaborative filtering	k-means	support vector machine
collision avoidance	logistic regression	swarm intelligence
computational intelligence	machine learning techniques	symbolic computation
computational pathology	machine translation	text mining
computer vision	meta learning	topic model
connectionism	motion planning	transfer learning
conversational interface	multi-agent system	unsupervised learning
cyber physical system	multi-objective optimization	virtual agent
data mining	natural language processing	word2vec
decision model	neural turing machine	xgboost
emotion recognition	neuromorphic computing	

9. Table B.2. Green-terms.

bio	renewable energies	heatpump
circular	energy saving	sustainable packaging
clean alternative	biomass	smart farming
climate change	biobased	waste management
low emissions	biodegradable	green buildings
corporate responsibility	renewable resources	sustainable materials
eco	bioenergy	electric vehicle
emissions	wind energy	carbon capture
enviroment	solar energy	biocides
esg reporting	photovoltaics	fuels from waste
waste	marine energy	biotechnology
organic	water power	natural fibres
sustainable	hydropower	plants
biologic	battery technology	
recycling	water treatment	