

Abbasiharofteh, Milad; Hermans, Frans

Article — Published Version

Using archetypal analysis to derive a typology of knowledge networks in European bioclusters

Regional Studies

Provided in Cooperation with:

Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Halle (Saale)

Suggested Citation: Abbasiharofteh, Milad; Hermans, Frans (2025) : Using archetypal analysis to derive a typology of knowledge networks in European bioclusters, Regional Studies, ISSN 1360-0591, Taylor & Francis, London, Vol. 59, Iss. 1, pp. 1-17,
<https://doi.org/10.1080/00343404.2024.2430354> ,
<https://www.tandfonline.com/doi/full/10.1080/00343404.2024.2430354>

This Version is available at:

<https://hdl.handle.net/10419/315645>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>

Using archetypal analysis to derive a typology of knowledge networks in European bioclusters

Milad Abbasiharofteh^a  and Frans Hermans^b 

ABSTRACT

The shift to a bioeconomy necessitates radical innovations to move from a fossil- to a bio-based economy, with bioclusters playing a crucial role by fostering collaboration among a wide range of co-located actors to find solutions. Through a novel methodological approach in regional studies (i.e., archetypal analysis), this study discusses the specificities of bioclusters and analyses the relational properties within European bioclusters' knowledge networks. Results suggest four archetypal knowledge networks. It is rare to observe the transition of biocluster networks from one archetype group to another. The findings provide a data-driven bioclusters typology, offering insights into designing informed place-based policies.

KEYWORDS

bioclusters; knowledge networks; inventive activities; archetypal analysis; unsupervised learning

JEL D85, O30, O31, Q57


HISTORY Received 28 September 2022; in revised form 29 October 2024

1. INTRODUCTION

The urgency for green innovations, particularly at the regional level, has become increasingly apparent in light of global environmental challenges and the need for sustainable development (Gibbs & O'Neill, 2017). The bioeconomy has come up as a concept promoted by different governments worldwide as a potential pathway for this transition (Stark et al., 2022). The bioeconomy encompasses using renewable biological resources to extract essential components for materials, chemicals, and energy. These raw materials and their associated waste streams are then processed to yield value-added products, including but not limited to food, feed, bioplastics, pharmaceuticals and bioenergy. The overarching objective of the bioeconomy lies in reducing our current reliance on carbon-based sources derived from fossil fuels by replacing them with renewable sources of carbon rooted in photosynthesis. Apart from the shift away from fossil fuels, the bioeconomy also promises to contribute to regional sustainable development, generate employment in high-tech sectors, and promote the creation of green innovations (McCorrick & Kautto, 2013). By capitalising on regional assets, such as biodiversity, agricultural lands, and natural resources, regions can harness the potential of the


bioeconomy and create new development paths (Refsgaard et al., 2021). Nevertheless, the intricacies of this process remain unclear, underscoring the need for a closer examination of the bioeconomy and how it influences innovation processes at the regional level.

In this paper, we focus on the role of bioeconomy clusters, or bioclusters for short, that are essential tools for governments to promote the regional bioeconomy and to establish, promote, and strengthen economic collaboration, learning, and innovation processes within particular regions (Hermans, 2018). The idea is that they possess the potential to integrate local resources, knowledge, and networks into biobased innovations that result in sustainability effects having impacts beyond their local level (Ayrapetyan et al., 2022; Kamath et al., 2022). As such, the growing attention to bioclusters falls into a broader trend in cluster research to take the role of clusters in the generation of sustainable, green or eco-tech innovations more into account (McCauley & Stephens, 2012; Njøs et al., 2017; Sedita & Blasi, 2021). The literature on industrial clusters has paid much attention to the underlying network characteristics of clusters and other territorial innovation systems (Abbasiharofteh, 2020; Abbasiharofteh & Maghssudipour, 2024; Hermans, 2020). This strand of literature discusses that clusters

CONTACT Milad Abbasiharofteh  m.abbasiharofteh@rug.nl; m.abbasi.h@gmail.com

^aDepartment of Economic Geography, Faculty of Spatial Sciences, University of Groningen, Groningen, the Netherlands

^bDepartment of Structural Change, Leibniz Institute of Agricultural Development in Transition Economies, Halle, Germany

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/00343404.2024.2430354>

differ based on their level of development and sectoral, institutional, and technological compositions. That said, studies that investigate the knowledge networks within these green-tech, or sustainable clusters, are sparse. Our research question is, therefore: What are the key characteristics of bioclusters' knowledge networks?

With this paper, we aim for two different contributions: a methodological and an empirical one. First, we want to illustrate a relatively new and, we think, promising tool for the further development of network studies within regional studies and economic geography archetypal analysis of networks (D'Esposito et al., 2010). To answer our research question, we utilised a methodological framework based on an unsupervised clustering algorithm to create benchmarks for analysing the structural properties of knowledge networks in biocluster research. Our study marks the first application of this method in regional studies, despite its common use in other scientific disciplines ranging from biology to astronomy (Chan et al., 2003; Gimbertat-Mayol et al., 2022). Our approach has broader applicability making it suitable for future regional studies that aim to address classification problems to answer theoretically informed research questions.

Second, our research makes an empirical contribution by first discussing key characteristics of bioclusters' knowledge networks and their specificities related to innovation processes in the bioeconomy. Subsequently, we systematically identify and track European bioclusters over time, 'from the bottom-up', adding insights to the ongoing debate surrounding this concept. Creating a data-driven typology of bioclusters enabled us to distinguish bioclusters based on their place- and path-dependent attributes. In this paper, we discuss how this typology helps policy-makers to design bioeconomy policy in the future that is place-sensitive and address the specific needs of bioclusters.

The paper is structured as follows. Section 2 discusses explicitly some of the characteristics of innovation processes within the bioeconomy and bioclusters and draws implications for the resulting knowledge networks and their evolution. Section 3 introduces our methodology with the specific machine learning approach we applied. Section 4 presents and discusses the results and suggests how future studies can take our study as a point of departure to classify other types of clusters. Section 5 concludes the paper and discusses potential policy implications concerning the success of place-based innovation policy measures.

2. INNOVATION AND KNOWLEDGE NETWORKS IN INDUSTRIAL CLUSTERS

The interest in the causes and effects of agglomerations and clusters goes a long way back to the work of Marshall (1890), who observed the geographical co-location of firms in different industrial sectors in 19th-century Britain and explained them through the effects of labour market pooling, supplier specialisation and knowledge spillovers. Since the 1990s clusters have become a popular tool for

governments to strengthen regional economic competitiveness and innovation (McCauley & Stephens, 2012; Porter, 1998). With innovation being an important goal of many cluster policies, interest has grown into the characteristics of the knowledge networks within agglomerations and clusters and how these change over time (Glückler, 2007; Powell et al., 2005). This type of research investigates the spatial distributions of regional economic processes not in isolation, but instead they consider the dynamic and interconnected nature of economic activities and uses a lens of social relationships, networks and interactions (Bathelt & Glückler, 2003; Boggs & Rantisi, 2003; Oinas & Malecki, 2002). This research investigates how network characteristics of an industrial cluster influence both the performance of the individual organisations within the cluster as well as its overall functioning (Belussi et al., 2010; Breschi & Malerba, 2005; Karlsson et al., 2005).

With the 'relation turn', the literature on industrial clusters has developed and applied several different methods: analysis tools, models, and procedures to aid in these investigations: social network analysis (SNA) (Giuliani & Pietrobelli, 2011; Ter Wal & Boschma, 2009), (adapted) gravity models, quadratic assignment procedures (QAP) (Broekel et al., 2014; Simensen & Abbasiharofteh, 2022), and statistical network models such as exponential random graph models (ERGMs) and stochastic actor-oriented models (SAOMs) (Abbasiharofteh, 2020; Hermans, 2020; Maggioni et al., 2007). In this section, we will discuss three topics. First, we discuss some of the current network approaches in regional studies and economic geography, and how they relate to some of the methodological toolboxes in use. We will shortly introduce the archetypal analysis and compare it with existing methods. We will specifically discuss the context of our empirical case study: biocluster and their main overlap and differences with other sectoral industrial clusters. We will end this section with a discussion on how the characteristics of bioeconomy will likely influence also the type of innovation processes and their associated knowledge networks within bioclusters.

2.1. Analysing knowledge networks in industrial clusters: methods and applications

Network analysis is a methodological approach used to study the structures of networks, focusing on nodes (actors) and edges (relationships) to understand their interactions and dynamics. Previous research has applied network analysis to industrial and innovation clusters to understand interactions, collaborations, and dynamics within clusters (Broekel et al., 2021). Early network studies focused especially on how the structural position of an actor within the cluster knowledge network would determine its economic and innovative performance. This type of research looks at the influence of structural holes (Ahuja, 2000; Walker et al., 1997) or the position of a firm within the core or the periphery of the cluster (Giuliani, 2013). Later, other models investigate the characteristics of the complete cluster network: how

these networks were formed under the influence of micro-level mechanisms of tie formation between actors, for instance, through reciprocity, homophily or triadic closure. Often, these micro-level processes were then linked to different forms of proximity: geographical, institutional, social, cognitive, and organisational and their interplay (Janssen & Abbasiharofteh, 2022; Kabirigi et al., 2022). Other studies looked into how these networks changed over time in studies of network evolution (Menzel et al., 2017; Nicotra et al., 2013; Ter Wal, 2013), or a combination of these two: how different forms of proximity influence network evolution over time (Baland, 2012; Lazzeretti & Capone, 2016).

Although network theory and network tools have become increasingly popular in cluster and agglomeration studies, there is no consensus on the effects of different types of network constellations on a cluster's innovative performance. What is considered beneficial in some instances can become detrimental in other cases: new and emerging industrial clusters versus old and established industrial clusters, sectoral specifics of clusters (high tech versus low tech), the type of innovation being pursued: incremental or radical and explorative versus exploitative. It is therefore important to analyse the features of a sectoral cluster, in our case bioclusters. This would also align with Hermans (2020), who calls for more reflection on a cluster's internal and external characteristics in the setup of a network study. In a similar vein, Glückler and Panitz (2021) argue for the use of 'informed pre-specification', which they define as 'actively integrating the already existing substantive understanding of a particular research context into the specification of a formal network model' (p. 19).

In this paper, we utilised archetypal analysis to study knowledge networks in clusters, starting from this 'informed pre-specification'. The archetypal analysis involves the identification of original patterns or models in multivariate data. In other words, the archetypal model aims to find 'pure types' in the data and determine how each datum point is a combination of configurations represented by those pure types (Cutler & Breiman, 1994; Eugster & Leisch, 2009). The flexibility of this method allows for classifying a wide range of data types ranging from texts to cities, to spatial diffusion patterns, and to networks. In this context, the problem is identifying several hypothetical points (archetypes) within a set of multivariate data based on *theoretically informed indices* (Cutler & Breiman, 1994). Network scientists use this technique to reduce the dimensionality of network data and classify them based on network attributes (Porzio et al., 2008; Ragozini et al., 2017). This technique is relevant for investigating cluster networks for two reasons. First, the model is flexible, enabling us to define network indices that are theoretically relevant for specific types of clusters. In this paper, we will illustrate the approach in the case of bioeconomy clusters, but the method would also be adaptable for use in other industrial clusters. Second, the results show how each data point is a mixture of identified pure types. The latter is particularly important for understanding the evolutionary trajectories of clusters by uncovering

the underlying structures and dynamics (see the method section for a more detailed account of the archetypal analysis method). In the following section, we will undertake informed pre-specification to investigate biocluster knowledge networks.

2.2. Green innovations in the context of the bioeconomy and bioclusters

In this section, we first analyse some of the particularities of knowledge networks in bioclusters. It is important to note that although the bioeconomy and bioclusters are often associated with positive policy goals as sustainable development, agricultural development, and eco-innovations (Wilde & Hermans, 2021), in reality, many sectors related to the bioeconomy carry a heavy pollution load. Not all bio-based innovations can be called sustainable, and some can even have serious detrimental sustainability effects and trade-offs at other scales and levels (Ayrapetyan & Hermans, 2020; Bröring et al., 2020). Therefore, Kamath et al. (2022) argue that many bioclusters need to make a sustainability transition before contributing to the overarching goals of green innovation. Based on the earlier work of Markusen (1996) and Iammarino and McCann (2006), Hermans (2021) identifies four different types of bioclusters and their potential sustainability effects in their localised production chains from raw materials to production and use (Table 1).

Table 1 explores the sustainability of various bioclusters, highlighting their environmental impacts and innovation dynamics. First, agricultural agglomeration and green chemistry bioclusters depend heavily on biomass production, often using imported biomass, which can cause significant environmental pressures. Marshallian bio-districts, which can localise their production processes, may achieve self-sufficiency and thus have a closer link to local natural resources. Life science clusters, on the other hand, rely minimally on biological sources, making biomass sourcing less critical. At the same time, bioclusters servicing global markets can create substantial local environmental pressures due to their concentrated production. Green chemistry clusters can theoretically enhance sustainability through circular 'cradle-to-cradle' designs based on industrial ecology principles (Ayrapetyan et al., 2022).

The typology of bioclusters shows that bioclusters are a broad topic where innovation processes can vary according to the type of knowledge base: doing, using and interacting (DUI) and science, technology, innovation (STI) (Jensen et al., 2007). Clusters using DUI knowledge, typical of agricultural agglomerations and bio-districts, often lack resources for high-tech innovations. Environmental innovations usually come from government-induced public-private partnerships and scale through local knowledge transfer (Hermans et al., 2019). On the other hand, green chemistry and life science clusters rely on STI knowledge for high-tech innovations. Regional universities play a key role in training specialists, partnering in R&D, and generating spin-offs. These clusters' transformative potential lies in developing environmentally

Table 1. Theoretical biocluster types and networks.

Biocluster type	Examples	Links to resource base	Production and manufacturing	Goal of innovation in sustainability transition
Agri-food clusters (agricultural agglomerations)	Horticultural clusters, wine clusters, intensive animal husbandry areas	Sector specific: <ul style="list-style-type: none"> • Crops and dairy: mainly local • Intensive animal husbandry: global 	Local negative environmental effects at cluster level	Optimising efficiency of production (through NARS)
Green chemistry clusters	Biorefineries, green chemistry (industrial biotech), paper and pulp clusters	Sector specific: <ul style="list-style-type: none"> • Paper and pulp: mainly local • Biofuel and bioenergy: global 	Local negative environmental effects at cluster level	<ul style="list-style-type: none"> • Improve efficiency of conversion • Industrial ecology
Marshallian bio-districts	Design, fashion, leather, wood construction, and building	Medium to high dependence on local inputs, often linked to regional branding	Risk of outsourcing environmentally polluting activities (textiles, leather)	<ul style="list-style-type: none"> • Demand-driven product design for ecofriendly consumers • Localisation of production chains and agro-ecology
Life science clusters	Pharmaceuticals and medicine (red biotech), cosmetics	Low	Low	<ul style="list-style-type: none"> • Innovation as a goal: new products and process incubator • Risk of rebound effect

Note: NARS, National Agricultural Research System.

Source: Adapted from Hermans (2021), used with permission.

friendly production methods and new products, though life sciences' sustainability impacts are initially more economic and social.

2.3. Pre-specification of biocluster knowledge networks

In this paper, we will focus on the types of bioclusters that rely more on codified forms of knowledge: life science clusters and green chemistry clusters. This focus simplifies the complexity associated with the wide variety of bioclusters in theory and practice and aligns with the nature of our patent-based database (see the methodology section). Drawing on Golembiewski et al. (2015) and Van Lancker et al. (2016), we identify four key factors that characterise sustainable bio-based innovation processes and are especially relevant for forming knowledge networks in these clusters.

First, a transition to the bioeconomy would require radical changes in the existing business models and supply chains. Even innovations, such as so-called 'drop-in chemicals', that focus on the manufacturing of known chemical building blocks, like (bio)methane and (bio)ethanol, will require radical changes in their supply chains as refineries make the transition towards biorefineries and switch from a single source of input (crude oil) that is

available all year around to a large number of smaller suppliers and where weather and climate effects have to be taken into account (Birch & Calvert, 2015). The switch towards these new value chains may be challenging and gradual, requiring a focus on different forms of connectivity as this switch involves a move from well-established relationships in the fossil fuel industry to fragmented, shorter connections in new value chains (Wilde & Hermans, 2024).

Second, the bioeconomy encompasses a complex knowledge base derived from diverse fields, including life sciences, biotechnology, agronomy, food science, social science, nanotechnology, renewable energy, waste management, information and communication technologies (ICT), and engineering (Golembiewski et al., 2015). The complexity of the knowledge base necessitates integrating diverse knowledge types within our sample, starting from a broad range of patents rather than a narrowly defined knowledge base. In addition, this broad knowledge base also requires specific attention to the issue of multidisciplinary, as many innovations occur on the interface of specific knowledge bases in different forms of regional diversification (Boschma et al., 2017).

Third, policy and institutional effects in the bioeconomy also shape preferred networking types. Established

industries, especially those related to fossil fuels and the chemical industry, restrictive policies, and high commercialisation costs lead to significant barriers due to high switching costs and a lack of quality standards (Wilde & Hermans, 2021). Sectors with restrictive policies and high commercialisation barriers are more likely to develop hierarchical knowledge networks. These networks are characterised by a few dominant nodes (key institutions or companies) controlling the flow of information, leading to slower knowledge dissemination but potentially more controlled and high-quality innovation outputs. However, sectors with flexible policies, supportive environments, and efforts to reduce adoption barriers may cultivate small-world networks. These networks have numerous interconnections among nodes, promoting rapid and wide-reaching knowledge transfer and fostering innovation through frequent interactions and collaborations. These networks can be characterised by their small-worldliness.

Finally, it is important to note that we expect a mix of established and emerging industrial clusters, with some clusters emerging around completely net technologies and other bioclusters that are actually in different transition phases, for instance, moving from ‘brown to green’, or ‘path upgrading’ (Tripl et al., 2020). Following Menzel and Fornahl’s (2010) cluster evolution theory, the industrial networks within a cluster only develop in one of the later stages, meaning that some European bioclusters might still be in an early development stage, and one can still expect the connectedness of their networks. This argument also resonates with other empirical findings that show collaborative networks often start as fragmented and ‘fluid’ (Fritsch & Zoellner, 2020). This means for our knowledge networks, we are particularly interested in the issues of (1) connectivity, (2) hierarchy and (3) small-worldliness. We will further elaborate on this in the methodology section.

3. METHODOLOGY

3.1. Construction of bioclusters and their knowledge networks

Inspired by the work of Marshall (1890) and Porter (1998), one can describe an industrial cluster as a geographical concentration of firms and affiliated organisations, such as universities, within one or multiple interconnected industries. These entities are linked through input–output relationships, informal interactions, cooperative local networks, and labour mobility. An empirical definition of a cluster is daunting (Martin & Sunley, 2003). More recent conceptual and empirical works suggest that detecting clusters should go beyond finding a geographical concentration of economic activities and consider inter-organisational relations (Janssen & Abbasiharofteh, 2022; Delgado et al., 2016). Following these studies, we have used patents to map the geographical concentration of bioeconomy collaborative relations. In other words, we define bioclusters from a network perspective as regions with a significantly higher number of inventive ties than surrounding regions. The assumption

is that this number also correlates with input–output relations, informal interactions, cooperative local networks, and labour mobility.

While one should be aware of problems associated with using this type of data (Archibugi & Planta, 1996; Fritsch et al., 2020), the assumption is that joint inventive activities are accompanied by knowledge exchange and mutual learning. We used the Organisation for Economic Co-operation and Development’s (OECD) REGPAT database.¹ This database provides information regarding the geographical location of inventors and applicants at the NUTS3 level, International Patent Classification (IPC) technological classes, and respective filing and grant dates for each patent. Concerning our broad definition of a cluster it is important to note that patents are always filed by one or several inventors, and those individuals can be associated with a business, but also a university, research institute, legal firm or government agency. As such, patents are a source of public and private organisations.

Since this study focuses on identifying and analysing bioeconomy clusters, we should first systematically define bioeconomy related patents. Technology codes included in patent documents provide valuable information about utilised technology in inventions. Kriesch and Losacker (2024) employed advanced machine learning techniques to identify bioeconomy patent abstracts. Their research suggests that several agricultural and food-related technology classes are associated with an exceptionally high share of bioeconomy patents. Accordingly, we followed the extensive literature review of Frietsch et al. (2016) that identified the key IPC technology codes for the bioeconomy (see Appendix A in the supplemental data online). Interestingly, the suggested technologies align with the findings of Kriesch and Losacker (2024). For instance, both studies identified bioeconomy technologies such as new plants or processes for obtaining them; plant reproduction by tissue culture techniques (A01H), and organic fertilisers, for instance, fertilisers from waste (C05F).

We selected patents that include technology codes associated with biotechnology (biotech) and at least one of 12 following bioeconomy categories: (1) agriculture and forestry, (2) pulp and paper, (3) machines cartons boxes printing, (4) genetic engineering, (5) landscape management, (6) food, (7) proteins, (8) biofuels (including biofuels for transport), (9) biomass, (10) biomaterials, (11) marine, (12) animals livestock management, and (13) food-related household appliances. Using this approach, we filtered 568,883 patents (17% of patents) at the aggregate level. Our approach resonates with our conceptual focus on interdisciplinary inventive activities that underline a set of activities in which biotech contributes to other bioeconomy technologies by providing cutting-edge technology solutions.

Next, we projected two-mode (inventor-by-project) collaborative knowledge networks to one-mode (inventor-by-inventor) networks. We used the information on inventors’ geographical position (NUTS2 European regions) to create intra- and interregional networks.² Following a common approach, the data were disaggregated

into six five-year time-windows: (1) 1986–90, (2) 1991–95, (3) 1996–2000, (4) 2001–05, (5) 2006–10 and (6) 2011–15 (Abbasiharofteh et al., 2023a; Li et al., 2014; Menzel et al., 2017; Ter Wal, 2014). For the sake of simplicity, we address each time-window with the last year of each time-window.

As a cut-off point for the cluster region, we defined the 90th percentile of the density for each time-window (see Appendix B in the supplemental data online). However, it is essential to note that defining cluster based on administrative boundary such as the NUTS classification system, is criticised due to problems associated with the so-called modifiable area unit problem (Brenner, 2017; Scholl & Brenner, 2014). To tackle this problem, we included collaborators outside bioclusters as a part of bioclusters' networks if there is at least one collaborative tie between them. Inspired by the relational turn in economic geography, this implies that knowledge transfer relations are not limited to the administrative boundaries of a given cluster. Thus, we allocated an interregional collaborative connection to more than one NUTS2 region if they are home to inventors who developed a given patent (Bathelt & Glückler, 2003).

3.2. Stylised characterisation of cluster knowledge networks

Based on our review of innovation processes within the bioeconomy we identified three elements as particularly salient: hierarchy (as displayed in scale-free networks), small world networks and connectivity. The scale-free networks of Barabási and Albert (1999) and the small-world networks of Watts and Strogatz (1998) belong to the most well-known network structures in science. These two networks can have both positive and negative effects on innovation outcomes. The scale-free, centralised networks of Barabási and Albert (1999) possess many 'structural holes'. Bridging a structural hole provides individuals with novel and non-redundant information (Burt, 1992, 2005). The skewed degree distribution of scale-free networks, with only a few nodes that possess the vast majority of connections, results in a star-like network configuration that increases the likelihood of actors receiving new information. If we define innovation as the novel combination of existing knowledge, then such a scale-free network could potentially be associated with more innovations due to the efficiency of a centralised network in knowledge circulation (Albert et al., 2000).

Another significant network pattern discussed in the literature that holds potential for fostering innovation is the 'small-world' structure, as introduced by Watts and Strogatz (1998). This type of network exhibits a high level of clustering (transitivity) and short average path length between its nodes. Scholars in the field of social sciences have emphasised the importance of clustering as a form of social embeddedness resulting from shared past experiences. Clustering, in turn, facilitates trust formation, limits opportunistic behaviour, and reduces transaction costs (Coleman, 1988; Granovetter, 1985; Uzzi, 1997).

Besides clustering, the second critical element is connectivity within the network. Connectivity plays a critical role in how knowledge is transferred among individuals and organisations because highly fragmented knowledge networks do not provide the required knowledge transfer channels (Fleming et al., 2007). Connectivity can work in two ways. As clusters develop, it is expected that their knowledge networks will become more connected over time (Menzel & Fornahl, 2010), however as we have argued in the theoretical section in the switching process from established industries towards more biobased production, the connectivity of a cluster might at the first time show a decrease of connectivity over time.

Figure 1 brings the two connectivity and degree distribution elements into a single framework. We can identify four quadrants based on the network hierarchy (scale-free versus small-world) and connectivity (low versus high density).

It is plausible that complex knowledge networks entail structural properties of both small-world and scale-free networks at the same time. For instance, an ill-structured core-periphery knowledge network could show small-worldness at the core (i.e., a high degree of clustering and short average path length) and scale-freeness in the periphery where poorly connected nodes are connected to nodes at the core (Vicente, 2017).

3.3. Using archetypal analysis to create a knowledge network typology

Investigating the structural properties of knowledge networks requires using multiple network indices as proxies for various aspects of networks. Although network indices might theoretically capture different structural properties of a network, they are often correlated. The archetypal analysis method suggested by Cutler and Breiman (1994) is an unsupervised machine learning technique that reduces the dimensionality of data by grouping observations into a set of archetypes using a data-driven approach. Instead of looking for 'typical' observations (cluster centres), it seeks extremal points in multidimensional data. We used this method to benchmark bioclusters knowledge networks (Porzio et al., 2008; Ragozini et al., 2017). We describe this method below in two steps: (1) selecting a set of parameters that describe the structure of networks; and (2) defining a set of archetypes that maximise within-archetypes commonalities and between-archetypes differences.

3.3.1. Selecting a set of network parameters

We have operationalised the four network types of Figure 1 by several indicators. Table 2 presents a list of network measures and corresponding methods that reflect the extent to which the observed networks are scale-free, small-world and connected. We refrained from including distance-based indices (e.g., the average path length) in the analysis because the observed bioclusters collaboration networks are often fragmented.

To control for the clustering problem associated with the projection of bipartite networks, we applied the so-

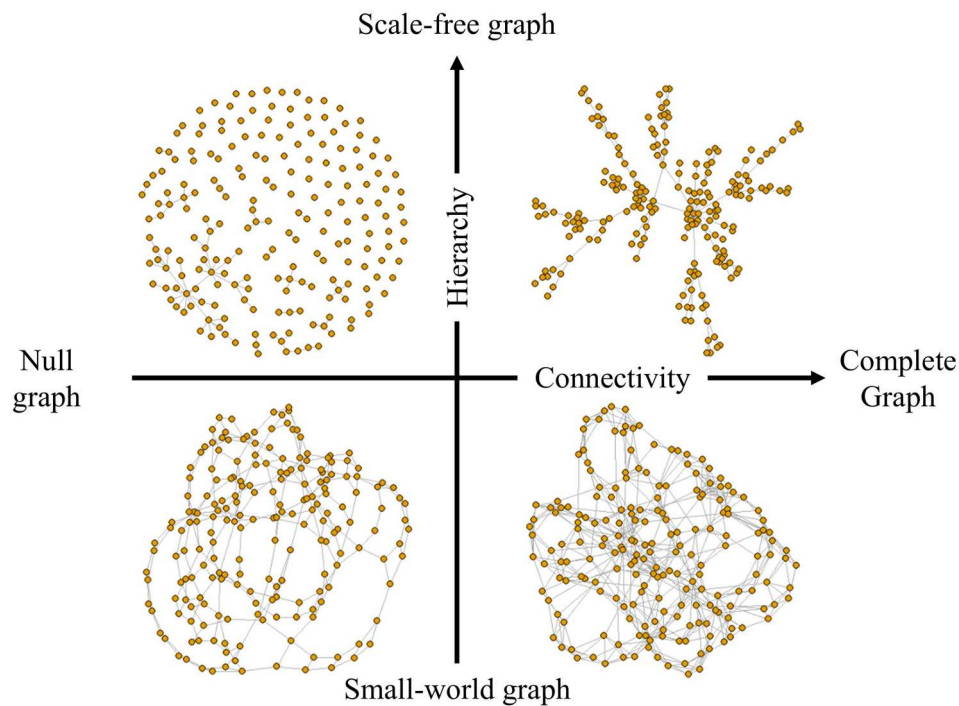


Figure 1. Stylised examples of knowledge network types with a similar number of nodes.

Note: Randomly removing half of the ties leads to a highly fragmented network with a scale-free structure, whereas the small-world network is not split into multiple components.

called ‘network of places’ (Abbasiharofteh et al., 2023a; Lucena-Piquero & Vicente, 2019; Pizarro, 2007). This clustering problem occurs when working with the projection of bipartite networks. This clustering becomes problematic because team size increases over time (Van der Wouden, 2020) and varies across different technologies (Broekel, 2019), causing a different degree of clustering and biases cluster-based measures for structural properties. To address the projection bias and reduce the impact of clustering caused by network projection, we utilise the concept of ‘structural equivalence’, which has its roots in social network theory. This concept was developed by Lorrain and White (1971) and Burt (1987). In this approach, nodes within a network are considered structurally equivalent if they share identical relationships. This equivalence enables them to access similar resources within the network (Gnyawali & Madhavan, 2001; Stuart & Podolny, 1996). Figure 2 illustrates three networks. The visualisation on the left-hand side shows a bipartite network in which inventors A, B and C collaborate on developing a patent, and inventors C and D are involved in developing another patent. This network can be projected to a one-mode network shown in the middle, which provides a high degree of clustering due to projecting a two-mode network. The method of networks of places (Figure 2, right) enables us to project a two-mode network without increasing the degree of clustering in the network. Earlier research shows that the new network (i.e., a network of places) is not biased towards different team sizes over time and across technologies. It represents the structural properties of original networks (Abbasiharofteh et al., 2023a). Our analysis is based on the empirical

investigation of networks of places (for the sake of consistency, we use the term knowledge networks or collaborative networks).

The second problem is associated with the fact that some network indices are size variant. This implies that estimated indices of collaborative networks with different sizes need to be normalised and then used as inputs for the archetypal analysis method. To normalise, we plotted each network index (corresponding to a biocluster in a given time-window) on the y -axis and the size of the network (node number) on the x -axis. We calculated the residuals of the non-parametric regression fit (the distance between each dot and the blue spline in Figure 3, left panel). In this way, the network size does not strongly affect the normalised coefficients (Figure 3, right panel).

3.3.2. Defining a set of archetypes

Based on the previous step, each network can be presented as a set of numerical vectors (i.e., normalised network indices). The archetypal analysis method is an unsupervised machine learning technique to define distinct categories by minimising the within-group and maximising the between-group squared errors in each mixture of identified archetypes in the multivariate Euclidean space (Cutler & Breiman, 1994). Formally, Ragozini et al. (2017) take a set of n networks $\Omega = \{x_i\} \ i = 1, \dots, n, x_i \in \mathcal{R}^p$ and a division $C = (C_1, \dots, C_k)$ of Ω in K groups, and define an internal similarity as $R(x_i, C_b), x_i \in C_b$, and external dissimilarity as $D(x_i, \bar{C}_b), x_i \notin C_b$. A mixing function $\Phi(\cdot)$ combines two internal similarity and external dissimilarity

Table 2. Overview of network indices for describing the structure of bioclusters knowledge networks.

Structural property	Measure	Method	Variable
Small-worldness	Community number	Number of identified communities using the multilevel modularity optimisation algorithm	<i>COMMUNITY</i>
	Share of inter-community ties	Number of ties connecting communities divided by the total number of ties using the multilevel modularity optimisation algorithm	<i>INTERCOM</i>
	Modularity	Modularity scores	<i>MODULARITY</i>
	Transitivity	Clustering coefficient (Wasserman & Faust, 1994)	<i>TRANSITIVITY</i>
Scale-freeness ^a	Network centrality	GINI coefficient for actors' degree centrality (Giuliani, 2013)	<i>CENTRALITY</i>
Connectivity	Share of the largest component	Number of nodes in the largest component divided by the total number of nodes	<i>COMPONENT</i>
	Share of isolates	Number of isolated nodes divided by the total number of nodes	<i>ISOLATES</i>
	Density	Number of ties divided by the number of possible ties (Wasserman & Faust, 1994)	<i>DENSITY</i>

Note: ^aWe used other network centrality indices and coreness (*K*-core decomposition) as a proxy for scale-freeness. We refrain from including these indices in the analysis as these indices are highly correlated with the GINI coefficient for actors' degree centrality.

functions:

$$T(x_i, C_b) = \Phi(R(x_i, C_b); D(x_i, \overline{C_b})) \quad (1)$$

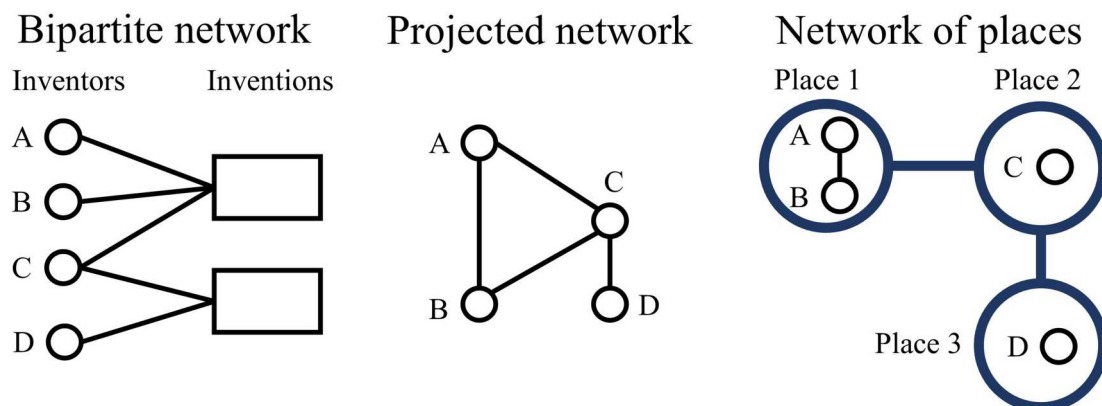
A set of archetypes $A = (a_1, \dots, a_k)$ is then defined as:

$$A = \{a_b \in \mathcal{R}^p \mid a_b \argmax T(x_i, C_b), b = 1, \dots, K\} \quad (2)$$

This equation provides multiple solutions suggesting a different number of archetypes with corresponding residual sum of squares (RSS) coefficients that show the ability of defined archetypes to categorise networks into distinct groups (for technical details, see Ragozini et al., 2017). Using the elbow plot, one can select the minimum number of archetypes that capture the optimal degree of variance in the observed data. Finally, one can use a fuzzy membership function and estimate a 'purity parameter' to maximise the internal similarity of archetypes. Then, networks are assigned to a specific archetype when its distance in the archetypal space is closer to one archetype and far from the other. They

are not assigned a particular archetype if they have a similar distance to two or several defined archetypes.

To provide an example of the above-discussed method and to test the reliability of the archetypal analysis method in grouping networks based on their structural properties, we simulated 60 networks with three distinct structural properties. We used the archetypal analysis method to group them based on network indices described in Table 1. More specifically, we simulated 20 small-world networks, 20 scale-free networks and 20 random networks and randomly removed between 10% and 50% of their ties to create different degrees of connectivity. Figure 4 shows that this technique provides a reliable method of grouping networks based on structural properties. A1 represents an archetype for small-world structural properties, A2 for scale-free ones, and A3 and A4 for random networks. Four random, five small-world and one scale-free networks are assigned to none of the four archetype groups, perhaps due to their low connectivity. In contrast, the method is highly accurate in assigning networks to each group. That is, the number of false-positive cases is zero.

**Figure 2.** The method of networks of places.

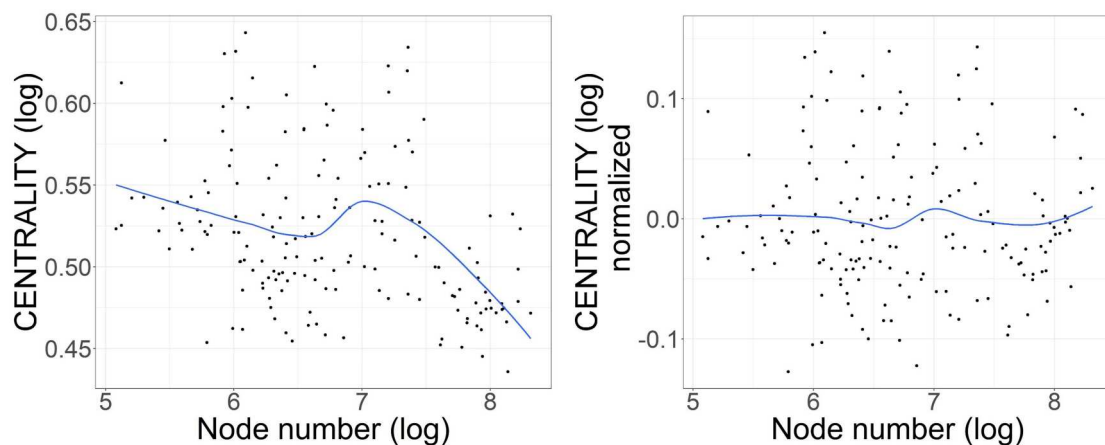


Figure 3. Normalising the centrality index using a non-parametric regression fit.

For the empirical analysis, we used several R packages ‘igraph’ by Csardi and Nepusz (2006) for simulating networks and estimating network indices, ‘Places’ by Lucena (2017) for defining networks of places, and ‘Archetypes’ by Eugster and Leisch (2011) for the archetypal analysis.

4. RESULTS

We identified 38 bioclusters between 1986 and 2015 in seven European countries: Germany, France, Italy, the Netherlands, Denmark, Belgium and the UK, with Germany having the highest share. Figure 5 and Appendix C in the supplemental data online show the distribution of bioclusters in Europe.

Using the archetypal analysis method, we identified four groups of bioclusters with distinct structural properties. It is worth mentioning that the archetypal analysis algorithm gives four coefficients between 0 and 1 that represent the position of a given archetype (ARCH) in an archetypal space. For instance, a network with a value

equal to 1 for ARCH1 and close to 0 for the other three ARCHs perfectly represents the structural properties of ARCH1. Table 3 shows the name and coefficients of typical biocluster networks. Most regions, however, position themselves between two ARCHs in the archetypal space (see Appendix D in the supplemental data online). The result suggests that about one-third of observed networks can be grouped in one of the four categories, whereas the rest of biocluster knowledge networks resemble the structural properties of two or several defined categories. Figure 6 answers our research question: there are four different archetypes of knowledge networks in bioclusters in Europe. Figure 7 demonstrates the structural properties of each defined archetype. We briefly describe the structure of each archetype.

- *ARCH1: Densely connected and moderately small-world.* Knowledge networks in this archetype are well-connected, shown by a higher density, a greater share of the largest component, and a lower share of isolates.

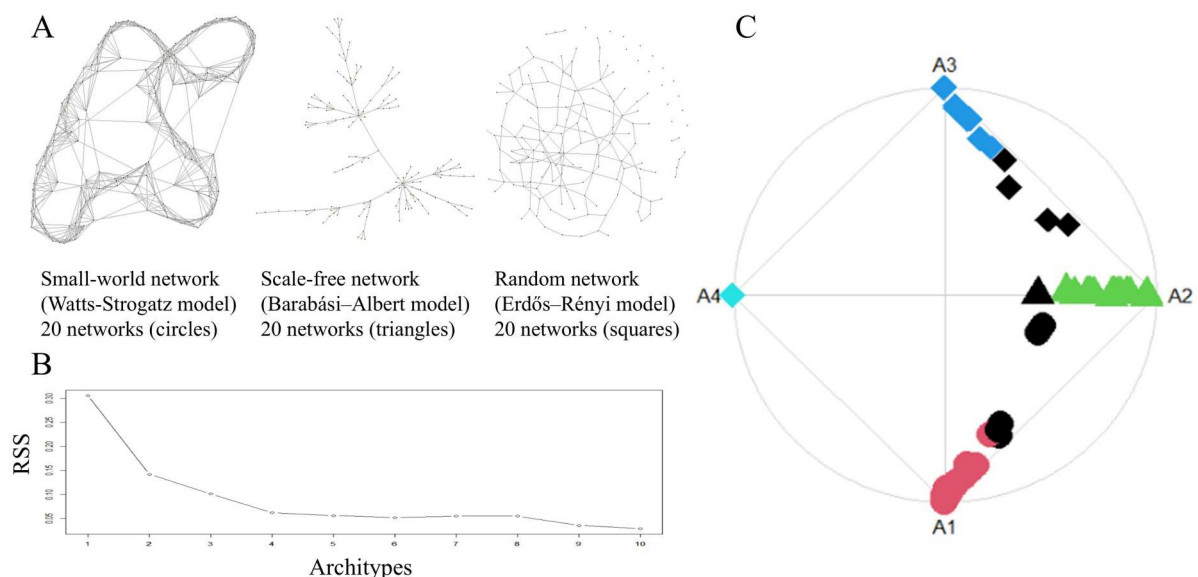


Figure 4. Three network structures (A) and estimated residual sum of squares (RSS) for different numbers of archetypes (B) and four defined archetypes (colour-coded) using the archetypal analysis method (C).

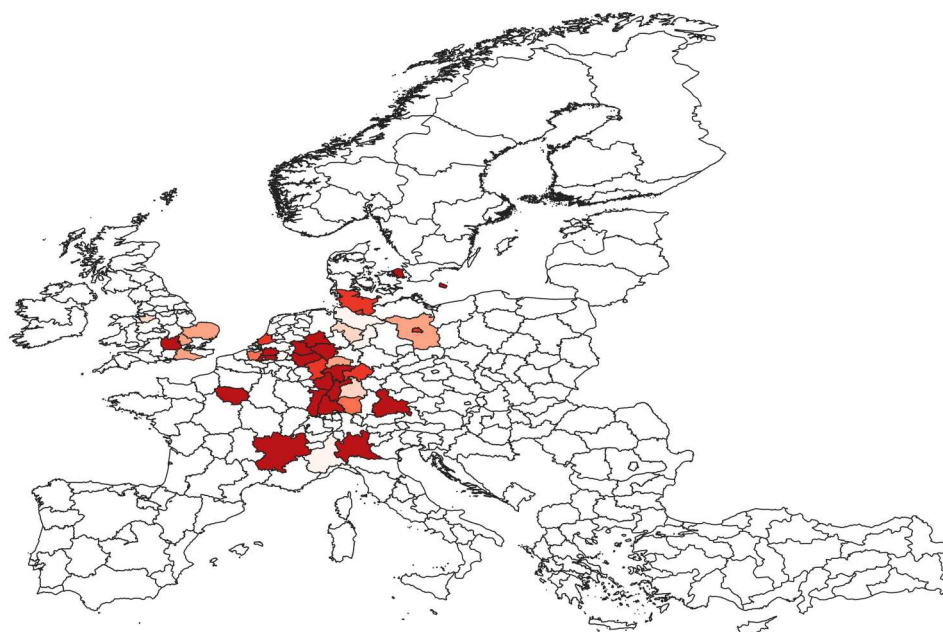


Figure 5. Geographical distribution of bioclusters in Europe.

Note: A darker colour represents regions identified more often as biocluster in the six time-windows between 1986 and 2015.

Also, a greater share of inter-community ties and a lower centrality show that this group of networks has characteristics associated with small-world networks. However, low modularity and transitivity suggest a moderate degree of small-worldness. Roughly 10% of knowledge networks represent the structural properties of this archetype. A typical example of this archetype is the Tübingen region in 2015.

- *ARCH2: Highly fragmented.* The main characteristic of the second group of networks is the lack of connectivity captured by a large share of isolates, a small share of the largest component, and low density. Contrary to the previous archetype, we observed a skewed degree distribution and a structure resembling scale-free networks. However, it does not perfectly follow Barabási and Albert's (1999) model due to low connectivity. About 4% of knowledge networks are assigned to this archetype group. The region Berkshire, Buckinghamshire and Oxfordshire in 2000 is a typical example of this archetype.
- *ARCH3: Moderately disconnected and modular.* Contrary to the other three groups that show a strong tendency towards only a set of structural properties, this archetype displays a moderate degree of almost all structural properties. However, networks associated with this

archetype are highly modular and are not dense and transitive. Only 2% of knowledge networks represent this structure. The Bedfordshire and Hertfordshire region in 1995 perfectly represents the structural properties of this archetype.

- *ARCH4: Transitive islands.* The last defined archetype shows a high degree of transitivity, modularity, and many communities. While this suggests a high degree of clustering, this does not represent a small-world network because clustered nodes are weakly connected, captured by an extremely low share of inter-community ties and a low share of the largest component. Interestingly, this archetype represents the largest archetype (about 20%). The Freiburg region in 2005 is a typical example of this archetype. The ARCH4 group could represent rather specialised regions than the networks assigned to the other archetype groups. It is plausible that inventors in the specialised clusters based on similarity in their technological portfolios and lack of inter-community collaboration reduce the capability of bioclusters to introduce patents with unconventional technology combinations (Abbasiharofteh et al., 2023a). Our observation aligns with the work of Fritsch et al. (2020), showing that co-inventive networks are

Table 3. Name and coefficients of regions perfectly representing the structural properties of each ARCH.

NUTS2	Typical region	Time-window	ARCH category	ARCH			
				ARCH1	ARCH2	ARCH3	ARCH4
DE14	Tübingen	2015	1	1.00	0.00	0.00	0.00
UKJ1	Berkshire, Buckinghamshire and Oxfordshire	2000	2	0.00	0.93	0.07	0.00
UKH2	Bedfordshire and Hertfordshire	1995	3	0.00	0.00	1.00	0.00
DE13	Freiburg	2005	4	0.00	0.00	0.01	0.99

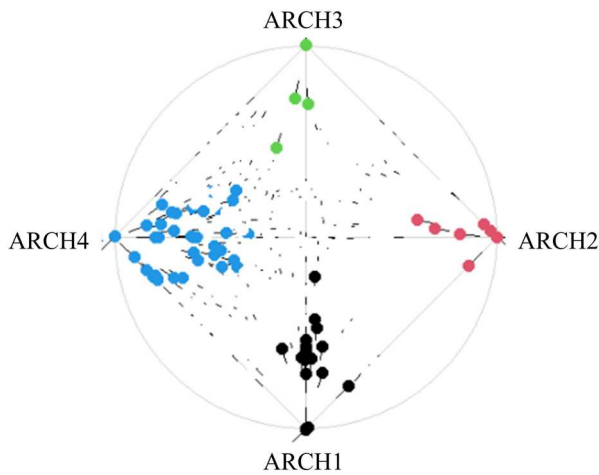


Figure 6. The position of each region (dots) in a given time-window in the archetypal space.

Note: Four archetypes are specified by four colours and regions that do not belong to a specific archetype are excluded.

highly fragmented and, therefore do not perfectly follow the structural properties suggested by the preferential attachment and small-world models.

Looking at the distribution of the archetypes among the four theoretically defined network types, it seems that most biocluster knowledge networks tend to have structural properties associated with small-world networks (e.g., transitivity) compared with properties associated scale-free and random ones.

5. DISCUSSION

This paper has aimed to investigate the key characteristics of bioclusters' knowledge networks. We have argued that the role of bioclusters in the transition to a bioeconomy, involves radical supply chain changes, diverse knowledge bases, and varying policy environments. As clusters are in varying stages of development in this transition process, their knowledge networks may also vary: established industries may develop hierarchical networks with controlled information flow, while sectors with supportive policies and reduced barriers are likely to form small-world networks with rapid, wide-reaching connectivity. This underscores the need for integrating diverse knowledge types and fostering transdisciplinary collaboration to facilitate innovation. We conceptually argued that the structural properties of bioclusters' networks might vary along connectivity and hierarchy dimensions. In this section, we will discuss our results in terms of their theoretical and empirical contributions and their policy implications.

5.1. Development of archetypes over time

We found four types of archetypes for biocluster networks, with only one archetype that shows some significant level of connectivity (ARCH1: densely connected and moderately small-world). This is the archetype that would facilitate the interdisciplinary knowledge exchange due to its structural properties. The other three archetypes share the characteristic that their connectivity is low and links between these different inventors or groups of inventors (as is the case of ARCH4: transitive islands) are missing.

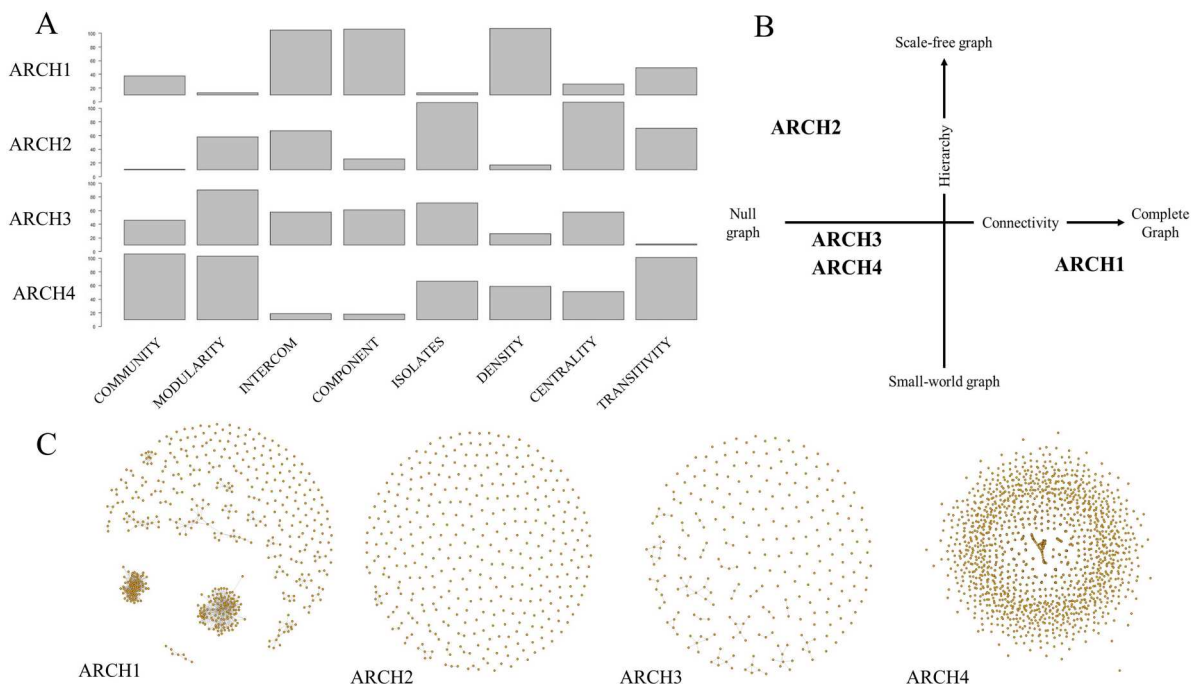


Figure 7. Biocluster knowledge networks: (A) the structural properties of each identified archetype; (B) the position of each archetype on the connectivity–hierarchy graph; and (C) the typical cluster networks for each archetype.

Note: For larger visualisations, see Appendix E in the supplemental data online.

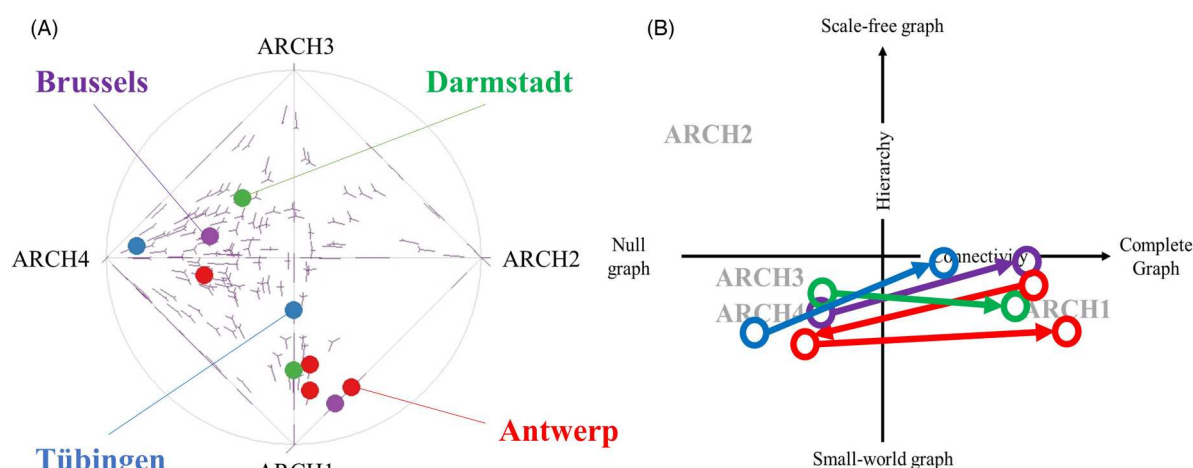


Figure 8. The position of regions transitioning from ARCH4 into ARCH1 in the archetypal space (A) and the connectivity–hierarchy graph (B).

Note: The demonstrated transitions in network structure took between 10 and 20 years.

It is also remarkable that the hierarchical network types are only found in one other archetype: ARCH2.

The question is how these archetypes develop over time and whether hierarchical cluster networks associated with fossil fuel and chemical cluster networks change over time towards networks that resemble more small-world characteristics.

To investigate this possibility, we checked which regions managed to break from structural properties associated with fragmented knowledge networks and transitioned into a more connected one (Figure 8). This investigation concludes that the archetypes are rather stable over time and that biocluster networks only slightly move in the archetypal space over time. It is rare to observe a transition of biocluster networks into ARCH1 in the past 30 years. Co-inventorship collaborative ties connected invention islands in only four regions (i.e., Brussels, Antwerp, Darmstadt and Tübingen). This observation implies that there is some evidence that bioclusters over time can become more connected and that regions can go through an evolutionary path that starts with an increase in inventive activities (ARCH2/3 → ARCH4), a few regions, however, manage a transition into a relatively cohesive network structure (ARCH1) that facilitates the combination of diverse knowledge created and trapped in different inventors' communities. However, given the very few cases where we observed this, we have to be careful not to draw any far-reaching conclusions. Our findings also suggest that collaborative ties among invention islands are fragile, and they may get dissolved and consequently cause a transition from ARCH1 into ARCH4. The evolutionary trajectory of the Düsseldorf region provides a typical example for this line of argument: ARCH2 (1990) → ARCH4 (1995) → ARCH1 (2000) → ARCH4 (2005).

5.2. Policy lessons: biocluster knowledge networks' path-dependency and specific needs

This brings us to our second important discussion point. Even though cluster networks can shift from one

archetype to the next, these are rare cases (four bioclusters out of 38). This observation means that in the majority of cases, the characteristics of the knowledge networks have been very stable over time.

This path dependency has important implications for policymakers because new collaborative ties will not emerge out of nowhere but must be actively stimulated. Depending on the characteristics of the knowledge network of a biocluster, our results suggest some specific policies that could fit the archetypes best (Table 4).

Biocluster policies that encourage the creation of formal and informal networks and platforms within the bioclusters to facilitate collaboration, knowledge sharing, and joint projects are particularly suited to biocluster types with highly fragmented or disconnected inventors (such as archetypes 2 and 3). These policy measures should then encourage the emergence of connected subgroups and 'islands' as they are found in archetype 4. For ARCH4 some specific policies should further increase connectivity between those islands. To this end, biocluster policy may target 'missing links' in a biocluster network instead of aiming at blindly increasing network density in bioclusters (Vicente, 2017). Finally, for those bioclusters in the ARCH1 category, more general cluster policies could be kept in place to ensure the knowledge network's resilience within the biocluster.

5.3. Advantages and disadvantages of archetypal analysis and future research

The use of archetypal analysis is essential because compared with other techniques (e.g., principal component analysis), this model provides interpretable results where data points are represented as mixtures of extreme types (archetypes) and less sensitive to outliers (Cutler & Breiman, 1994; Eugster & Leisch, 2009). Further, like principal component analysis, one can use archetypal analysis for dimensionality reduction, helping to summarise and compress regional data into several representative archetypes.

Table 4. Overview of relevant interventionist policy measures based on the structural properties of bioclusters.

Biocluster network type	ARCH2 (highly fragmented) and ARCH3 (moderately disconnected and modular)	ARCH4 (transitive islands)	ARCH1 (densely connected and moderately small-world)
General aim	Cluster policies that increase density and connectedness of knowledge networks	Cluster policies that target links between specific interdisciplinary subgroups	Cluster policies that stimulate the resilience of existing interdisciplinary ties
Examples	Establishment of formal and informal networks and platforms within the bioclusters to facilitate collaboration, knowledge sharing and joint projects, for instance through allocating funding for joint research and development (R&D) projects	Supply chain mapping and ecosystem mapping and analysis: conducting periodic analyses of the cluster ecosystem helps identify gaps, and opportunities to connect previously unconnected value chains and firms	Recognise and reward existing multidisciplinary collaborations
	Promote proximity and accessibility for face-to-face interactions	Promote cross-pollination of ideas through diverse physical and virtual spaces, mentorship programmes, cultivate curiosity	Create redundancy and adaptability within knowledge networks through multiple suppliers and strategic partnerships
	Encouraging social engagement and informal contacts within the cluster	Create temporary interdisciplinary proximity through events such as fairs or conferences	Train employees with versatile skills, create adaptable manufacturing processes, and maintain shared resources to quickly respond to disruptions

In this paper, we have utilised the archetypal analysis of bioclusters and theoretically motivated and selected a set of network indices to provide a better understanding of bioclusters knowledge networks. It is, however, important to note that a broader range of cluster studies can benefit from this analytical technique. The power of this method should be coupled with a set of theoretically informed indices. In this paper, we discussed the fact that bioclusters are known for a wide range of actors from various fields. As a result, we investigated bioclusters in relation to knowledge networks' connectivity and hierarchy (Figure 1). This conceptual framework may change if one investigates clusters in different stages of their life cycle or sectoral composition. For instance, if a study aims to understand clusters' capabilities in breaking from a lock-in situation, it should focus, among others, on extra-regional connectivity and restructuring the existing knowledge network (Abbasiharofteh, 2020). The application of the archetypal analysis is not limited to classifying clusters' network indices. For instance, researchers can use information on regions' share of new ecofriendly products and services to identify distinct trajectories in forming regional lead markets in eco-innovation (Beise & Gemünden, 2004).

The archetypal analysis method is not free of limitations, and discussing some words of caution seems vital. One issue is that the archetypal analysis algorithm depends on starting conditions, and the solutions that are provided may not be unique. Therefore, we strongly recommended that researchers replicate models multiple times to ensure the stability of the results. Another limitation is that this

model is computationally expensive because the algorithm provides solutions based on multi-step convex least-squared problems (Eugster & Leisch, 2009). This feature makes the algorithm not attractive for users analysing 'big' regional data without access to high-performance computing.

Besides the advantages and disadvantages of using the archetypal analysis method, we acknowledge several limitations in this paper. Using patent data as a common practice may bias our results towards regions hosting large companies filing more patents. Future studies may build on different data sources to map and investigate the structural properties of knowledge networks. Recent studies utilising unconventional data sources such as trademarks, web-text, and product launches of companies provide valuable input for mapping and analysing knowledge networks (Abbasiharofteh et al., 2022, 2023b; Nathan & Rosso, 2022).

To define bioclusters, we followed a relational turn in economic geography and refrained from an oversimplified definition of clusters (i.e., a mere concentration of economic activities). That said, our approach limits clusters by administrative boundaries. This approach suffers from the modifiable areal unit problem (2014). We encourage future studies to take the work of Scholl and Brenner (2014) as a point of departure and use archetypal analysis to investigate clusters not necessarily surrounded by administrative boundaries.

We acknowledge that our analysis is exploratory and based on several conjectures on the association between knowledge network structural properties and

interdisciplinarity. We cannot claim that a specific network archetype is positively associated with radical innovation driven by interdisciplinary collaborative teams. This limitation suggests a follow-up empirical work aiming at investigating this potential relationship.

Moreover, although we observed that a few regions managed to change their knowledge exchange network substantially, the question of what factors triggered such a transition remains open. Thus, we encourage qualitative works to better understand the driving forces of structural change in fortunate biocluster networks identified by this study. Finally, the empirical evidence on the relationship between bioclusters' network properties and the increase in interdisciplinarity is missing. Therefore, a new line of research may consider investigating the association of four defined groups of structural properties and interdisciplinary knowledge creation.

6. CONCLUSIONS

In this paper, we have provided a better understanding of the key characteristics of knowledge networks in European bioclusters. Based on an unsupervised clustering algorithm, we created benchmarks for the structural properties of biocluster knowledge networks based on patent data. To the best of our knowledge, this method has been used for the first time in regional studies. Our approach is not limited to the field of bioeconomy and network data. Thus, this work can be used as a point of departure for future regional studies that aim at tackling classification problems to answer theoretically informed research questions.

Our results show four different archetypes of knowledge networks in European bioclusters. Of these four archetypes, only one archetype can be associated with favourable characteristics for knowledge exchange between a wide range of actors in the form of high connectivity between actors. Our study shows that archetypes are stable and that only a small share of clusters' networks may transition into other forms over time. These results show a high degree of path dependency in biocluster network evolution and encourage investigating the key factors driving network evolution in a few regions in the future.

Our results suggest that biocluster policies should consider the structure of a knowledge network and, in some cases, target and promote collaborative ties that create missing links among innovation islands.

ACKNOWLEDGEMENTS

The authors acknowledge the support of the Trafobit project, 'The Role and Functions of Bioclusters in the Transition to a Bioeconomy™ (031B0020) under the call, 'Bioökonomie als gesellschaftlicher Wandel™ of the German Federal Ministry of Education and Research (BMBF).

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

FUNDING

This work was supported by the Bundesministerium für Bildung und Forschung [grant number 031B0020].

NOTES

1. For a detailed overview, see OECD (2008).
2. We opted for the NUTS2 level because large cities and metropolitan areas such as London might be divided into several NUTS3 regions.

ORCID

Milad Abbasiharofteh  <http://orcid.org/0000-0001-9694-4193>

Frans Hermans  <http://orcid.org/0000-0003-3394-9012>

REFERENCES

- Abbasiharofteh, M. (2020). Endogenous effects and cluster transition: A conceptual framework for cluster policy. *European Planning Studies*, 28(12), 1–24. <https://doi.org/10.1080/09654313.2020.1724266>
- Abbasiharofteh, M., Castaldi, C., & Petralia, S. G. (2022). *From patents to trademarks: Towards a concordance map (Final report)*. European Patent Office (EPO).
- Abbasiharofteh, M., Kogler, D. F., & Lengyel, B. (2023a). Atypical combination of technologies in regional co-inventor networks. *Research Policy*, 52(10), 1–17. <https://doi.org/10.1016/j.respol.2023.104886>
- Abbasiharofteh, M., Krüger, M., Kinne, J., Lenz, D., & Resch, B. (2023b). The digital layer: Alternative data for regional and innovation studies. *Spatial Economic Analysis*, 10, 241–262. <https://doi.org/10.1080/17421772.2023.2193222>.
- Abbasiharofteh, M., & Maghssudipour, A. (2024). Driving forces behind relational knowledge sourcing in clusters: Single- and multilevel approaches. *Journal of the Knowledge Economy*, 28(3), 1. <https://doi.org/10.1007/s13132-023-01714-x>
- Ahuja, G. (2000). Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly*, 45(3), 425–455.
- Albert, R., Jeong, H., & Barabási, A. L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794), 378–382. <https://doi.org/10.1038/35019019>
- Archibugi, D., & Planta, M. (1996). Measuring technological change through patents and innovation surveys. *Technovation*, 16(9), 451–519.
- Ayrapetyan, D., Befort, N., & Hermans, F. (2022). The role of sustainability in the emergence and evolution of bioeconomy clusters: An application of a multiscale framework. *Journal of Cleaner Production*, 376(3), 134306.
- Ayrapetyan, D., & Hermans, F. (2020). Introducing a multiscale framework for biocluster research: A meta-analysis. *Sustainability*, 12(9), 3890.
- Balland, P.-A. (2012). Proximity and the evolution of collaboration networks: Evidence from research and development projects within the Global Navigation Satellite System (GNSS) industry. *Regional Studies*, 46(6), 741–756.
- Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512. <https://doi.org/10.1126/science.286.5439.509>

- Bathelt, H., & Glückler, J. (2003). Toward a relational economic geography. *Journal of Economic Geography*, 3(2), 117–144. <https://doi.org/10.1093/jeg/3.2.117>
- Beise, M., & Gemünden, H. G. (2004). Lead markets: A new framework for the international diffusion of innovation. In K. Macharzina, M. Oesterle & J. Wolf (Eds.), *Management international review* (pp. 83–98). Gabler.
- Belussi, F., Sammarra, A., & Sedita, S.R. (2010). Learning at the boundaries in an ‘open regional innovation system’: A focus on firms’ innovation strategies in the Emilia Romagna life science industry. *Research Policy*, 39(6), 710–721.
- Birch, K., & Calvert, K. (2015). Rethinking ‘drop-in’ biofuels. *Science & Technology Studies*, 28(1), 52–72.
- Boggs, J. S., & Rantisi, N. M. (2003). The ‘relational turn’ in economic geography. *Journal of Economic Geography*, 3(2), 109–116.
- 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.
- Brenner, T. (2017). *Identification of clusters – An actor-based approach* (Working Papers on Innovation and Space No. 2.17). Philipps-Universität Marburg.
- Breschi, S., & Malerba, F. (2005). *Clusters, networks, and innovation*. Oxford University Press.
- Broekel, T. (2019). Using structural diversity to measure the complexity of technologies. *PLoS One*, 14(5), e0216856.
- Broekel, T., Bolland, P.-A., Burger, M., & van Oort, F. (2014). Modeling knowledge networks in economic geography: A discussion of four methods. *Annals of Regional Science*, 53(2), 423–452.
- Broekel, T., Lazzaretti, L., Capone, F., & Hassink, R. (2021). Rethinking the role of local knowledge networks in territorial innovation models. *Industry & Innovation*, 28(7), 805–814.
- Bröring, S., Laibach, N., & Wustmans, M. (2020). Innovation types in the bioeconomy. *Journal of Cleaner Production*, 266(6), 121939.
- Burt, R. S. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology*, 92(6), 1287–1335. <https://doi.org/10.1086/228667>
- Burt, R. S. (1992). *Structural holes: The social structure of competition*. Harvard University Press.
- Burt, R. S. (2005). *Brokerage and closure: An introduction to social capital*. Oxford University Press.
- Chan, B. H. P., Mitchell, D. A., & Cram, L. E. (2003). Archetypal analysis of galaxy spectra. *Monthly Notices of the Royal Astronomical Society*, 338(3), 790–795. <https://doi.org/10.1046/j.1365-8711.2003.06099.x>
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 95–120. <https://doi.org/10.1086/228943>
- Csardi, G., & Nepusz, T. (2006). The *igraph* software package for complex network research. *InterJournal, Complex Systems*, 1695.
- Cutler, A., & Breiman, L. (1994). Archetypal analysis. *Technometrics*, 36(4), 338–347. <https://doi.org/10.1080/00401706.1994.10485840>
- D’Esposito, M. R., Ragozini, G., & Vistocco, D. (2010). Exploring data through archetypes. In H. Locarek-Junge & C. Weihs (Eds.), *Classification as a tool for research* (pp. 287–298). Springer.
- Delgado, M., Porter, M. E., & Stern, S. (2016). Defining clusters of related industries. *Journal of Economic Geography*, 16(1), 1–38.
- Eugster, M. J. A., & Leisch, F. (2009). From Spider-Man to hero – Archetypal analysis in R. *Journal of Statistical Software*, 30(8), 1–23.
- Eugster, M. J. A., & Leisch, F. (2011). Weighted and robust archetypal analysis. *Computational Statistics & Data Analysis*, 55(3), 1215–1225.
- Fleming, L., King, C., & Juda, A. I. (2007). Small worlds and regional innovation. *Organization Science*, 18(6), 938–954. <https://doi.org/10.1287/orsc.1070.0289>
- Fritsch, R., Neuhäusler, P., Rothengatter, O., & Jonkers, K. (2016). *Societal grand challenges from a technological perspective – Methods and identification of classes of the International Patent Classification IPC* (Discussion Papers in Innovation Systems and Policy Analysis No. 53). Fraunhofer ISI.
- Fritsch, M., Titze, M., & Piontek, M. (2020). Identifying cooperation for innovation – A comparison of data sources. *Industry & Innovation*, 27(6), 630–659. <https://doi.org/10.1080/13662716.2019.1650253>
- Fritsch, M., & Zoellner, M. (2020). The fluidity of inventor networks. *Journal of Technology Transfer*, 45(4), 1063–1087.
- Gibbs, D., & O’Neill, K. (2017). Future green economies and regional development: A research agenda. *Regional Studies*, 51(1), 161–173.
- Gimbernat-Mayol, J., Dominguez Mantes, A., Bustamante, C. D., Mas Montserrat, D., & Ioannidis, A. G. (2022). Archetypal analysis for population genetics. *PLoS Computational Biology*, 18(8), e1010301. <https://doi.org/10.1371/journal.pcbi.1010301>
- Giuliani, E. (2013). Network dynamics in regional clusters: Evidence from Chile. *Research Policy*, 42(8), 1406–1419. <https://doi.org/10.1016/j.respol.2013.04.002>
- Giuliani, E., & Petrobelli, C. (2011). *Social network analysis methodologies for the evaluation of cluster development programs*. Inter-American Development Bank IDB Publ.
- Glückler, J. (2007). Economic geography and the evolution of networks Johannes. *Journal of Economic Geography*, 7(5), 619–634. <https://doi.org/10.1093/jeg/ibm023>
- Glückler, J., & Panitz, R. (2021). Unleashing the potential of relational research: A meta-analysis of network studies in human geography. *Progress in Human Geography*, 45(6), 1531–1557.
- Gnyawali, D. R., & Madhavan, R. (2001). Cooperative networks and competitive dynamics: A structural embeddedness perspective. *The Academy of Management Review*, 26(3), 431. <https://doi.org/10.2307/259186>
- Golembiewski, B., Sick, N., & Bröring, S. (2015). The emerging research landscape on bioeconomy: What has been done so far and what is essential from a technology and innovation management perspective? *Innovative Food Science & Emerging Technologies*, 29(1), 308–317.
- Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3), 481–510. <https://doi.org/10.1086/228311>
- Hermans, F. (2018). The potential contribution of transition theory to the analysis of bioclusters and their role in the transition to a bioeconomy. *Biofuels, Bioproducts and Biorefining*, 12(2), 265–276. <https://doi.org/10.1002/bbb.1861>
- Hermans, F. (2020). The contribution of statistical network models to the study of clusters and their evolution. *Papers in Regional Science*, 100(2), 379–404. <https://doi.org/10.1111/pirs.12579>
- Hermans, F. L. P. (2021). Bioclusters and sustainable development. In S. R. Sedita, & S. Blasi (Eds.), *Rethinking clusters: Place-based value creation in sustainability transitions*. Springer.
- Hermans, F., Geerling-Eiff, F., Potters, J., & Klerkx, L. (2019). Public-private partnerships as systemic agricultural innovation policy instruments – Assessing their contribution to innovation system function dynamics. *NJAS: Wageningen Journal of Life Sciences*, 88(1), 76–95.
- Iammarino, S., & McCann, P. (2006). The structure and evolution of industrial clusters: Transactions, technology and knowledge spillovers. *Research Policy*, 35(7), 1018–1036. <https://doi.org/10.1016/j.respol.2006.05.004>
- Janssen, M. J., & Abbasifarofteh, 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. <https://doi.org/10.1016/j.techfore.2022.121689>

- Jensen, M. B., Johnson, B., Lorenz, E., & Lundvall, B. Å. (2007). Forms of knowledge and modes of innovation. *Research Policy*, 36(5), 680–693. <https://doi.org/10.1016/j.respol.2007.01.006>
- Kabirigi, M., Abbasiharofteh, M., Sun, Z., & Hermans, F. (2022). The importance of proximity dimensions in agricultural knowledge and innovation systems: The case of banana disease management in Rwanda. *Agricultural Systems*, 202(3), 103465.
- Kamath, R., Sun, Z., & Hermans, F. (2022). Policy instruments for green-growth of clusters: Implications from an agent-based model. *Environmental Innovation and Societal Transitions*, 43, 257–269. <https://doi.org/10.1016/j.eist.2022.04.003>
- Karlsson, C., Johansson, B., & Stough, R. (2005). *New horizons in regional science. Industrial clusters and inter-firm networks*. Edward Elgar.
- Kriesch, L., & Losacker, S. (2024). A global patent dataset for the bioeconomy. *Papers in Innovation Studies*, 2024(08), 1–15.
- Lazzeretti, L., & Capone, F. (2016). How proximity matters in innovation networks dynamics along the cluster evolution. A study of the high technology applied to cultural goods. *Journal of Business Research*, 69(12), 5855–5865.
- Li, G.-C., Lai, R., D'Amour, A., Doolin, D. M., Sun, Y., Torvik, V. I., Yu, A. Z., & Fleming, L. (2014). Disambiguation and co-authorship networks of the US patent inventor database (1975–2010). *Research Policy*, 43(6), 941–955. <https://doi.org/10.1016/j.respol.2014.01.012>
- Lorrain, F., & White, H. C. (1971). Structural equivalence of individuals in social networks. *The Journal of Mathematical Sociology*, 1(1), 49–80. <https://doi.org/10.1080/0022250X.1971.9989788>
- Lucena, D. (2017). *Places: Structural equivalence analysis for two-mode networks R package v0.2.0*.
- Lucena-Piquero, D., & Vicente, J. (2019). The visible hand of cluster policy makers: An analysis of aerospace valley (2006–2015) using a place-based network methodology. *Research Policy*, 48(3), 830–842. <https://doi.org/10.1016/j.respol.2019.01.001>
- Maggioni, M. A., Nosvelli, M., & Uberti, T. E. (2007). Space versus networks in the geography of innovation: A European analysis. *Papers in Regional Science*, 86(3), 471–493.
- Markusen, A. (1996). Sticky places in slippery space: A typology of industrial districts. *Economic Geography*, 72(3), 293.
- Marshall, A. (1890). *Principles of economics*. Macmillan.
- Martin, R., & Sunley, P. (2003). Deconstructing clusters: Chaotic concept or policy panacea? *Journal of Economic Geography*, 3(1), 5–35.
- McCauley, S. M., & Stephens, J. C. (2012). Green energy clusters and socio-technical transitions: Analysis of a sustainable energy cluster for regional economic development in central Massachusetts, USA. *Sustainability Science*, 7(2), 213–225.
- McCormick, K., & Kautto, N. (2013). The bioeconomy in Europe: An overview. *Sustainability*, 5(6), 2589–2608. <https://doi.org/10.3390/su5062589>
- Menzel, M.-P., Feldman, M. P., & Broekel, T. (2017). Institutional change and network evolution: Explorative and exploitative tie formations of co-inventors during the dot-com bubble in the research triangle region. *Regional Studies*, 51(8), 1179–1191. <https://doi.org/10.1080/00343404.2016.1278300>
- Menzel, M.-P., & Fornahl, D. (2010). Cluster life cycles – Dimensions and rationales of cluster evolution. *Industrial and Corporate Change*, 19(1), 205–238.
- Nathan, M., & Rosso, A. (2022). Innovative events: Product launches, innovation and firm performance. *Research Policy*, 51(1), 104373. <https://doi.org/10.1016/j.respol.2021.104373>
- Nicotra, M., Romano, M., & Del Giudice, M. (2013). The evolution dynamic of a cluster knowledge network: The role of firms' absorptive capacity. *Journal of the Knowledge Economy*, 45(6), 425.
- Njøs, R., Jakobsen, S.-E., Wiig Aslesen, H., & Fløysand, A. (2017). Encounters between cluster theory, policy and practice in Norway: Hubbing, blending and conceptual stretching. *European Urban and Regional Studies*, 24(3), 274–289. <https://doi.org/10.1177/0969776416655860>
- Oinas, P., & Malecki, E. J. (2002). The evolution of technologies in time and space: From national and regional to spatial innovation systems. *International Regional Science Review*, 25(1), 102–131.
- Organisation for Economic Co-operation and Development (OECD). (2008). *The OECD REGPAT Database* (OECD Science, Technology and Industry Working Papers No. 2008/02). https://www.oecd-ilibrary.org/science-and-technology/the-oecd-regpat-database_241437144144
- Pizarro, N. (2007). Structural identity and equivalence of individuals in social networks: Beyond duality. *International Sociology*, 22(6), 767–792. <https://doi.org/10.1177/0268580907082260>
- Porter, M. E. (1998). Clusters and competition: New agendas for companies, governments, and institutions. In M. E. Porter (Ed.), *On competition* (pp. 197–299). Harvard Business School Press.
- Porzio, G. C., Ragozini, G., & Vistocco, D. (2008). On the use of archetypes as benchmarks. *Applied Stochastic Models in Business and Industry*, 24(5), 419–437. <https://doi.org/10.1002/asmb.727>
- Powell, W. W., White, D. R., Koput, K. W., & Owen-Smith, J. (2005). Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. *American Journal of Sociology*, 110(4), 1132–1205.
- Ragozini, G., Palumbo, F., & D'Esposito, M. R. (2017). Archetypal analysis for data-driven prototype identification. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 10(1), 6–20. <https://doi.org/10.1002/sam.11325>
- Refsgaard, K., Kull, M., Slätmo, E., & Meijer, M. W. (2021). Bioeconomy – A driver for regional development in the Nordic countries. *New Biotechnology*, 60, 130–137. <https://doi.org/10.1016/j.nbt.2020.10.001>
- Scholl, T., & Brenner, T. (2014). Detecting spatial clustering using a firm-level cluster index. *Regional Studies*, 50(6), 1054–1068. <https://doi.org/10.1080/00343404.2014.958456>
- Sedita, S. R., & Blasi, S. (2021). *Rethinking clusters: Place-based value creation in sustainability transitions*. Springer.
- Simensen, E. O., & Abbasiharofteh, M. (2022). Sectoral patterns of collaborative tie formation: Investigating geographic, cognitive, and technological dimensions. *Industrial and Corporate Change*, 31(5), 1–36. <https://doi.org/10.1093/icc/dtac021>
- Stark, S., Biber-Freudenberger, L., Dietz, T., Escobar, N., Förster, J. J., Henderson, J., Laibach, N., & Börner, J. (2022). Sustainability implications of transformation pathways for the bioeconomy. *Sustainable Production and Consumption*, 29, 215–227. <https://doi.org/10.1016/j.spc.2021.10.011>
- Stuart, T. E., & Podolny, J. M. (1996). Local search and the evolution of technological capabilities. *Strategic Management Journal*, 17(S1), 21–38. <https://doi.org/10.1002/smj.4250171004>
- Ter Wal, A. (2014). The dynamics of the inventor network in German biotechnology: Geographic proximity versus triadic closure. *Journal of Economic Geography*, 14(3), 589–620. <https://doi.org/10.1093/jeg/lbs063>
- Ter Wal, A. L. J. (2013). Cluster emergence and network evolution: A longitudinal analysis of the inventor network in Sophia-Antipolis. *Regional Studies*, 47(5), 651–668.
- Ter Wal, A., & Boschma, R. A. (2009). Applying social network analysis in economic geography: Framing some key analytic issues. *Annals of Regional Science*, 43(3), 739–756. <https://doi.org/10.1007/s00168-008-0258-3>
- Trippel, M., Baumgartinger-Seiringer, S., Frangenheim, A., Isaksen, A., & Rypestøl, J. O. (2020). Unravelling green regional industrial path development: Regional preconditions, asset modification and agency. *Geoforum*, 111(1), 189–197.

- Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*, 42(1), 35–67. <https://doi.org/10.2307/2393808>
- Van der Wouden, F. (2020). A history of collaboration in US invention: Changing patterns of co-invention, complexity and geography. *Industrial and Corporate Change*, 29(3), 599–619. <https://doi.org/10.1093/icc/dtz058>
- Van Lancker, J., Wauters, E., & Van Huylenbroeck, G. (2016). Managing innovation in the bioeconomy: An open innovation perspective. *Biomass and Bioenergy*, 90, 60–69. <https://doi.org/10.1016/j.biombioe.2016.03.017>
- Vicente, J. (2017). Network failures and policy challenges along the life cycle of cluster. In D. Fornahl, & R. Hassink (Eds.), *The life cycle of clusters: A policy perspective* (pp. 56–75). Edward Elgar.
- Walker, G., Kogut, B., & Shan, W. (1997). Social capital, structural holes and the formation of an industry network. *Organization Science*, 8(2), 109–125.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684), 440–442. <https://doi.org/10.1038/30918>
- Wilde, K., & Hermans, F. (2021). Innovation in the bioeconomy: Perspectives of entrepreneurs on relevant framework conditions. *Journal of Cleaner Production*, 314, 127979. <https://doi.org/10.1016/j.jclepro.2021.127979>
- Wilde, K., & Hermans, F. (2024). Transition towards a bioeconomy: Comparison of conditions and institutional work in selected industries. *Environmental Innovation and Societal Transitions*, 50, 100814.