Geography of Innovation and Sustainability Transitions

GEIST Working Paper series

New methodological inroads to regional path development - Epistemological reflections on the contribution of semantic network analysis

GEIST Working Paper No. 2024(05)

Bernhard Truffer

Environmental Social Science Department, Swiss Federal institute of Aquatic Science and Technology, Switzerland

Faculty of Geoscience, Utrecht University, Utrecht, the Netherlands

Cite as: Truffer (2024) New methodological inroads to regional path development - Epistemological reflections on the contribution of semantic network analysis. *GEIST – Geography of Innovation and Sustainability Transitions*, *2024*(05), GEIST Working Paper series.

New methodological inroads to regional path development

Epistemological reflections on the contribution of semantic network analysis

Bernhard Truffer

Environmental Social Science Department, Swiss Federal institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland; bernhard.truffer@eawag.ch

Faculty of Geoscience, Utrecht University, Utrecht, the Netherlands

Abstract

Evolutionary thinking has provided very potent explanations for regional industrial path development in the past decade. Recent commentators argued for extending the originally rather narrow focus on preexisting knowledge stocks to include institutional dimensions, system resource build up, and the agentic shaping of industrial pathways. On an epistemological level, such conceptual enlargements require the bridging of quantitative variance explanations and qualitative process explanations, which few scholars have successfully managed to do. In the present paper, I will argue that developments in the rapidly expanding field of semantic network analysis might improve rigor in qualitative process reconstructions and by this be more easily relatable to established quantitative approaches in evolutionary economic geography. Semantic networks enable the systematic reconstruction of higher order analytical constructs based on the analysis of statements and actions of actors as reported in collections of text documents. More specifically, we will introduce the socio-technical configuration analysis (STCA) method, recently developed in the scholarly field of sustainability transition studies, and show how it can inform regional path development research. An illustrative empirical case will analyze the path development dynamics in the German federal state of Baden-Württemberg, a global leader in automobile manufacturing, in the wake of the global challenge of electric cars. I conclude with wider ramifications of semantic network approaches for economic geography research and how it can be leveraged in mixed method designs.

Key words: Regional path development; semantic network analysis; socio-technical configuration analysis; evolutionary economic geography

1 Introduction

One of the key scholarly ambitions of economic geographers is to explain economic development of cities, regions, and countries. A classical research focus is how regional industry portfolios provide jobs and economic prosperity. More recently, the explanatory ambition has expanded to include broader public value concerns such as investing in more sustainable technologies, protecting natural resources or providing equal development opportunities to citizens (Castaldi and Binz 2024; Coenen, Hansen, and Rekers 2015; Tödtling, Trippl, and Desch 2022).

Over the past twenty years, evolutionary perspectives have gained a strong standing in this field, building on insights from evolutionary economics and processes empathizing the effects of path dependency (Boschma and Frenken 2006; Martin and Sunley 2006). While the evolutionary research program has proven highly successful, recent commentators have criticized it as an essentially static and deterministic framework (Martin 2010) and have called for including a wider array of knowledge forms, institutional structures and agency into the list of explanatory factors (Hassink, Isaksen, and Trippl 2019; Grillitsch, Asheim, and Trippl 2018; Martin and Sunley 2015; Bathelt and Storper 2023). In order to accommodate for such broader explanatory ambitions, these commentators asked for extending the research program to include insights from research on innovation systems, political economy, strategic management, institutional sociology, political science and transition studies (Grillitsch, Asheim, and Trippl 2018; Martin 2010; Binz, Truffer, and Coenen 2016; Boschma et al. 2017; Baumgartinger-Seiringer et al. 2022; Gong et al. 2022; MacKinnon et al. 2019).

On a methodological level, this extension led to a division among the self-identified proponents of the evolutionary economic geography approach. The original authors developed a strong methodological program constructing sophisticated variance oriented explanatory models based on standardized data sets on products, skills or patents, which are analyzed by means of quantitative regression models to prove consistent correlations between prior related technological capabilities and the emergence of new industries (Boschma 2017; Balland et al. 2022). The more process-oriented researchers mostly argued for in-depth qualitative case studies to reconstruct individual regions' engagement in the shaping of new industrial pathways (Steen 2016; Grillitsch, Asheim, and Trippl 2018; Tödtling and Trippl 2018). The latter hold epistemological preference for process explanations over variance based explanations (Geels 2022; Sotarauta and Grillitsch 2023). There have been recurrent calls for bridging these two cultures (Hassink, Isaksen, and Trippl 2019; MacKinnon et al. 2019). The most integrated conceptual framework was perhaps introduced by Martin and Sunley (2015) calling for a "developmental evolutionary economic geography" perspective. However, so far, these attempts have not yet translated into a unifying methodological paradigm.

In the following sections, I will argue that implementing such an enlarged conceptual agenda will require new methodological approaches that enable the bridging between variance and process-oriented explanations. As this bridge may be built from both sides, I will elaborate how conventional qualitative process reconstruction approaches may be further developed for that purpose. Recent developments in semantic network analysis methods suggest productive inroads (Segev 2022) as they enable to address the emergence of system structures, while also accounting for the agency of a diversity of actors. Based on a set of qualitatively derived data structures, it supports the systematic categorization and aggregation of terms into higher order constructs, which may represent mechanisms of variation, selection, and retention as well as phenomena of emergence, development and agency. More specifically, I will report on the recently developed method of socio-technical configuration analysis (STCA), which has proven potent in retracing dynamics of socio-technical configurations (Heiberg, Truffer, and Binz 2022) and was recently extended to address core economic geography research questions (Truffer et al. in print).

To illustrate the potential contribution of this methodological approach, I will reinterpret the case of the German federal state of Baden-Württemberg (BW), a global leader in the automobile industry, which had to reconsider its industrial pathway due to the emergence of a disruptive alternative: the electric car (Gong and Truffer 2024). I will show how this emerging external challenge was first received with considerable reservation, leading to all sorts of conflicts among formerly harmoniously cooperating regional actors. Ultimately, a new industrial path was endorsed as a guiding vision of the regional innovation system resulting from the confluence of external events, a reassessment of the regional resource base, the construction of specific institutional structures and conscious system building efforts.

We will first review recent conceptual developments in the field of evolutionary economic geography, and then introduce semantic network methods with their potential contribution to explain phenomena in economic geography. Illustrations on how this approach may support the analysis of a case like the BW automobile pathway renewal will be the core of section 4. In the conclusions, I will argue for the broader relevance of this methodological approach for future economic geography research on industrial path development.

2 Recent theorizing of regional path development

One of the most successful explanatory frameworks for explaining regional industrial development builds on evolutionary processes and mechanisms (Martin and Sunley 2006). According to this approach, industrial success depends on the availability of prior related knowledge stocks that enable the continuous innovation and adaptation of regional companies in an increasingly volatile global market (Boschma and Frenken 2006; Balland et al. 2022). Core mechanisms of the evolutionary perspective encompass variation, selection and retention of promising new technologies, products or services. A key mechanis is path dependency, which states that success will depend on resource stocks that had been developed in the region earlier (Martin and Sunley 2006) and which will be reconfigured over time depending on the needs and activities of newly entering industries in the region (Gong et al. 2024). The more related existing resource stocks are with the needs of a newly arriving industry, the more likely it will be that it will prosper there (Boschma 2017).

This impressive research program has been able to explain a large share of regional industrial development pathways over the past two decades (Boschma et al. 2017). However, critique has been voiced more recently that a narrow focus on variation, selection and retention risks resulting in a rather static, punctuated equilibrium framing of industrial development (Martin 2010; Garud, Kumaraswamy, and Karnøe 2010). Essentially, the perspective implies that pathways will develop in a deterministic way once initial conditions are set. Empirical evidence however suggests that regional development is under constant reconsideration by its constituting actors. It is influenced by institutional structures that actors will maintain or transform in interest-driven processes and guided by visions and expectations about preferred goals, which guides technological and institutional strategies (Bathelt and Storper 2023; Martin and Sunley 2015; Hassink, Isaksen, and Trippl 2019; Grillitsch, Asheim, and Trippl 2018; Steen 2016; Gong and Truffer 2024; Bækkelund 2021; Garud, Kumaraswamy, and Karnøe 2010). Furthermore, core resources that are necessary for successful industrial activities are not "just there" in a region, but will only emerge from the interplay between the different companies and government departments, research institutes and universities, and even marginal actors like NGOs, consultancies or cultural activists (Martin and Sunley 2015; Gong et al. 2024). Therefore, evolutionary explanations of path development should more proactively conceive of *paths as processes*, embracing emergence and strategic agency of diverse regional actors and their interactions and move towards a "developmental" economic geography (Martin and Sunley 2015; Benner 2023; Martin 2010; Miörner 2022; MacKinnon et al. 2019).

The epistemological and methodological requirements for implementing such an enlarged path-as-process research agenda are however formidable. While the original program has been successful by analyzing large, standardized data sets covering export products, skills or patents (Hidalgo et al. 2007; Neffke et al. 2014; Li and Neffke 2024), a process-oriented approach will usually rely on qualitative in-depth case studies to retrace how path transformation strategies developed over time and highlight which actors leveraged which strategies to shape the emerging pathway (Steen 2016). The recently proposed methodology of path tracing by Sotarauta and Grillitsch (2023)exemplifies this approach. Commentators have repeatedly called for integration of the two methodological approaches (Hassink, Isaksen, and Trippl 2019; MacKinnon et al. 2019). However to date, there have been very few attempts to actually provide an integrated methodological approach, not the least because of deep ontological and epistemological differences that are underpinning the different perspectives (Rutten 2020; Goertz and Mahoney 2012).

3 A new methodological perspective based on semantic networks

Despite these deeper challenges, I will argue that searching for methodological middle ground is still worthwhile. One of the key challenges separating qualitative and quantitative approaches is related to the databases on which the different research traditions build. Quantitative approaches require sufficiently large, standardized data sets to run average effects regression models, while qualitative, case-based approaches are more interested in understanding, alternative sets of conditions that may cause specific outcomes in a given case or in a set of empirical cases (Goertz and Mahoney 2012). Bridging the two inroads may perhaps need more than just aligning data sets because they represent fundamentally different cultures in doing research with different explanatory ambitions, different perspectives on causality and different methods of inference (Rutten 2020; Garud, Kumaraswamy, and Karnøe 2010). However, scholars from both sides may reflect on how they could develop their approaches further to enable fruitful exchanges.

A particularly promising approach for methodological outreach from the qualitative process reconstruction side is semantic network analysis, sometimes also called topic network or linguistic network approaches. As an established tool of artificial intelligence (AI) research, it enables the identification of higher order constructs from combinations of basic terms concepts identified in collections of documents like newspapers, meeting protocols, or social media posts (Drieger 2013; Hain et al. 2023). Coding texts requires context knowledge to identify the theoretically inspired basic concepts and their combination into higher order constructs.[1] When systematically applied across extensive sets of documents, the resulting codes may be aggregated into relational data structures and analyzed by standardized, quantitative indicators to reveal patterns of association. This two-pronged approach enables researchers to move back and forth between aggregated patterns and the underlying context rich information from the cases. It therefore does not compromise on the specific strengths of qualitative methodologies, while providing aggregated indicators that might more easily be related to structures derived from quantitative approaches.

3.1 Use contexts of semantic networks

In its most basic definition, semantic networks are graphical representations of sets of terms connected by relations of causality, membership, or similarity, which jointly generate meaning in texts. Semantic networks were first proposed as a tool in analytical philosophy in the middle of the 19th century but gained widespread application in AI from the 1970ies onward (Lehmann 1992). The first

[1] The approach to distinguish basic concepts from higher order constructs is very well established in the qualitative methodology literature even though the preferred terms may vary depending on the respective school (Gioia, Corley, and Hamilton 2013; Strauss and Corbin 1998).

generation of AI tools built on concatenations of logical truth statements. But such a "flat" approach to knowledge representation quickly led to nonsensical outcomes, due to insufficient appreciation of the contexts in which these truth statements would be valid (Lehmann 1992). Semantic networks provide a means to map out such contexts. Relations may be aggregated into networks either based on deep *a priori* knowledge about specific characteristics of the terms, or more inductively by retracing the usage of the terms in conversations or documents (Knees, Neidhardt, and Nalis 2024). Topological relationships represented by the networks of terms are derived from calculating bilateral similarities among their characteristics through the co-reference in text segments attributed to specific actors. Clusters of closely related terms may then be interpreted as representing higher order constructs like frames, narratives, preference structures, institutional logics or socio-technical configurations (Heiberg, Truffer, and Binz 2022).

A vivid illustration for the potency of this approach is provided by a core feature of platform-based business models of companies like Amazon, X, TikTok or Spotify, offering customers suggestions for consumption items that most likely match their preferences.[2] Algorithms on which such suggestions rely are known as recommender systems (Ricci, Rokach, and Shapira 2022; Knees, Neidhardt, and Nalis 2024). The core assumption of recommender systems is that preferences for music records, books, holiday destinations, etc. are not randomly distributed within a population but rather aggregate around particular bundles of characteristics representing "genres", or consumption styles. In the music sector, salient characteristics may be folk, rock, jazz, or classics, or may include the historical period of recording, specific bands, national background or combinations of instruments. By either explicitly classifying encompassing databases of large numbers of products (content-based coding) or by analyzing actual choices by customers (collaborative filtering), semantic networks may be constructed that represent a revealed "preference space" for the specific product category in a population. Clusters of products in this space indicate specific preference profiles, consumption styles, or "tastes", which enable to position new customers to the most similar "corner" in the space and suggest further items in the "vicinity" of their first choices. Semantic networks therefore enable to map out the higher order construct of preference spaces for items like music, books, or holiday apartments across large set of customers.

Semantic networks have found relevance also beyond business applications, informing a variety of social science disciplines (Segev, 2022). The brunt of recent method development builds on large document collections like social media posts or web-scraping. Higher order constructs are derived by applying large language models, which are increasingly complemented by AI algorithms (Hain et al. 2023). These methods have in particular found their inroad into innovation studies, to inform firm level innovation management (Shim et al. 2022; Antons et al. 2020), forecasting directionalities within specific innovation fields (Baskici, Atan, and Ercil 2019; Fu, Sarpong, and Meissner 2022) or even for analyzing dynamics of entire research fields, as for instance in innovation system research (Rakas and Hain 2019). In geography and entrepreneurship studies, large text collections have recently been proposed as alternative data sources to standardized databases (von Bloh et al. 2020; Ozgun and Broekel 2021). But the connection with semantic networks has remained rather anecdotal. Some studies have started to leverage data from social media streams (Sevin 2014), content enhanced geographical information system based methods (Liu et al. 2024) or engaged in very specific analyses like food geographies in Korean primary school text books (Cho 2022).

This relative neglect is surprising because the general methodology of social network analysis gained strong and increasing coverage in major strands of economic geography theorizing like in relational

[2] A well-known form is the message "users that have bought this book also liked the following ones" after having ordered a book on Amazon or playlists of music recordings proposed on Spotify based on the download history of a new subscriber.

and evolutionary economic geography over the past two decades (Glückler and Doreian 2016; Glückler and Panitz 2021; Bathelt and Gluckler 2003; Uitermark and van Meeteren 2021). The dominant focus in this literature is however on material exchange relations among actors (trade, collaboration, competition, etc.), while revealing meaning structures through semantic networks has not received much attention. A recent exception to this rule is a recent proposal by Eckhardt and Glückler (2024) who presented argumentation analysis as a semantic network based methodology to analyze controversial innovations. A more established, even though implicit exception, is also the work on complexity and related variety prominent in evolutionary economic geography and trade economics, which proposed concepts like product spaces, industry spaces or technology spaces (Schetter et al. 2024; Balland et al. 2022). These spaces, like the above discussed preference spaces, represent higher order constructs revealing deeper and perhaps more difficult to observe causal factors like capabilities and professional skills. Hidalgo et al. (2007) for instance constructed relative economic competitiveness of countries by using export products as an indirect measure that reveals underlying unique knowledge and institutional resources in the respective country, which provide the competitive edge. Schetter et al. (2024) characterized this indirect revelation approach as a "phenotypical" assessment of explanatory factors, compared to a "genotypical" approach, which would try to measure causal factors directly. Research found that explanatory variables derived from such networked data spaces were of substantially superior explanatory power compared to most former, "flatter" alternatives (Hidalgo et al. 2018). Product, technology, and industry spaces are similar to semantic networks more than to material exchange networks among actors or countries. But relatedness studies or review papers of spatial social network analysis have so far not made such distinctions (Ter Wal and Boschma 2009; Uitermark and van Meeteren 2021; Wu et al. 2022).

3.2 Revealing socio-technical structures by means of semantic networks

We qualify the established way of analyzing semantic networks based on word frequencies over massive collections of texts as an *extensive variant* of this family. They come with specific promises and challenges, which we cannot elaborate here in any detail (Hain et al. 2023). Instead, I want to elaborate on a more *intensive* form, which is based on conventional qualitative coding of texts by knowledgeable researchers. While this approach has often been criticized as being limited in terms of the volume that can be handled and for suffering from several kinds of interpretative biases, we maintain that human coding is much better able to appreciate context information for identifying complex higher order constructs than what can typically be derived from isolated (or weakly embedded) individual words in large language models. Regarding analysts' biases, qualitative researchers can be reflexive about their assumptions and make them transparent (Gioia, Corley, and Hamilton 2013; Eisenhardt 1989), while large language models often remain black boxes in terms of what exactly leads to the identification of higher order constructs, as for instance topics in topic modelling. We cannot settle this dispute here but will rather elaborate how an intensive form of semantic network analysis may provide promising new methodological perspectives on the path-as-process challenge.

There are several precursors of intensive forms of semi-quantitative text analyses in the innovation systems and political science literature. For instance, *event analysis* was introduced for understanding the development of particular technological innovation system structures by counting relevant events reported in professional magazines that promoted or hindered the development of an emerging socio-technical configuration (Suurs and Hekkert 2009). In economic geography, Strambach and Pflitsch (2020) proposed a transition topology approach which builds on tracing events in a similar way. Sotarauta and Grillitsch (2023) proposed the method of path tracing that has the ambition to reconstruct paths by focusing on actors and events. It builds however entirely on the qualitative reconstruction of narratives. These methods remain limited to counting events over time and abstain

from deriving higher order constructs. A more configurational approach was recently introduced to political sciences in the form of discourse network analysis (Leifeld and Haunss 2012). The focus here is on reconstructing political constituencies arguing in favor or against specific policy proposals and how the framing of these proposals changes to maximize support for either outcome (Markard, Rinscheid, and Widdel 2021). In the center of interest of this method are (virtual) actor coalitions more often than the shaping of the policy proposals.

We argue that both earlier approaches are only of limited usefulness for informing economic geography research as they do not relate to material interventions or are not oriented at identifying higher order constructs. To fill this gap, the method of socio-technical configuration analysis (STCA) was developed in the field of transition studies recently (Heiberg et al. 2022). Essentially, STCA aims at reconstructing coherent configurations of technological and institutional elements that represent alternative innovation options in a field. These configurations are derived from reported statements and activities of different actors in this field. The epistemological assumption is that assembling reported statements and actions of different actors in targeted text collections (newspaper articles, government protocols, company reports, etc.) will reveal the underlying configurations, which specific actors take for granted or prefer, if often only implicitly. Like the rather simply structured recommender systems discussed above, here the assumption is that actors will refer to specific aspects of multi-dimensional higher order constructs that correspond to their deeply held believes about how the world is structured. Clusters of such aspects may therefore reveal higher order constructs like worldviews, institutional logics, socio-technical configurations or, as we will show, regional industrial pathways. In general, actors will rarely be reported as covering a full set of those different characteristics in any single newspaper article. However, when aggregating those reports that cover a specific actor type over several documents and in a given time span, we may expect that interrelated sets of characteristics may reveal the higher order construct. These configurations will emerge in the respective semantic networks as clusters or sub-graphs of tightly interconnected terms. The proximities between two terms may be derived from counting how many times actors were associated to these two terms across all documents in the collection, relative to the total count of either of the two terms.[3]

Constructing these networks proceeds in four main steps (Miörner et al. 2022). First, the analyst must identify a suitable document stock that covers the relevant phenomenon in sufficient detail over a relevant timespan.[4] Most STCA applications so far were based on collection of newspaper articles because they provide continuous and reliable coverage of salient topics that the editors of a newspaper deemed to be of interest to a wider public in a sector, country, or region (Yap, Heiberg, and Truffer 2023; Heiberg and Truffer 2022; Heiberg, Truffer, and Binz 2022; Miörner, Heiberg, and Binz 2022; Frank, Feola, and Schäpke 2024). In a second step, the analyst will code relevant articles identifying the main reported actors and the topics they have been associated with either in the form of statements or as actions. Third, cross tabulating the number of coded text segments for actors and

[3] A range of similarity measures exist in the literature (Miörner et al. 2022). The one described in the above statement is known as the Jaccard similarity, depicting the ratio of all actors that were associated with both terms, divided by the number of actors that have been associated with either of the terms. Jaccard similarity varies between zero and one, zero meaning, no actor was simultaneously associated to both terms and one meaning that all actors in the document stock were always associated with both terms simultaneously.

[4] Selecting appropriate document collections is of course a key decision with deep implications for the whole analysis. A key selection criterion is whether the documents are likely to cover the different components of a higher order construct. When retracing the emergence of a new socio-technical configuration such a solar PV panels entering a particular region, local and regional newspaper articles might be an appropriate document base. Retracing the dynamics of regional innovation system structures instead would probably need to be based on a broader collection of company reports, government commission protocols, communications of regional intermediaries, reports by research institutes, etc.

topics enables the calculation of proximity matrices depicting either how consistently two terms have been associated across all the actors (resulting in a matrix with bilateral similarities between topics) or alternatively, how similar the topic profiles are across all actors (resulting in an actor similarity matrix measuring congruence in worldviews of actors). Fourth and finally, these similarity matrices may be depicted as two-dimensional maps or networks by means of layout algorithms such as multi-dimensional scaling. These graphical networks enable the identification of salient higher order constructs by interpreting clusters or configurations of actors or terms and their evolution over time.[5]

An illustrative example is reported in Heiberg et al. (2022), who retraced the dynamics among two alternative socio-technical configurations to cope with water scarcity in different parts of the world. The analysis builds on a collection of 576 articles in a set of 800 global, English language newspaper articles referring to urban water infrastructure related news coverage between 2010 and 2018, retracing developments before, during and after major droughts that ravaged large parts of the US, South-Africa, and India between 2014 and 2016.[6] The different actors were found promoting essentially one of two socio-technical configurations: a) a large-scale approach based on storage dams, centralized water and wastewater treatment plants and state-owned water utilities, and b) onsite water treatment and use, managed by local companies and prosumers. The study retraced how the decentralized configuration moved increasingly center stage in the US, indicating an increasing legitimation of the novel, disruptive technology. These developments differed across countries, depending on pre-existing infrastructures, and institutional preconditions. The semantic network approach revealed in this case the two socio-technical configurations of urban water management as higher order constructs. Retracing the development over time enabled to measure changes in the dominance of competing configurations and indicate early signs of an emerging socio-technical transition.

Meanwhile, the method has been applied to reconstruct a variety of socio-technical transition dynamics based on a different types of document stocks like world bank data bases, UN meeting protocols or expert interview transcripts (Lesch, Miörner, and Binz 2023; Yap, Heiberg, and Truffer 2023; Miörner, Heiberg, and Binz 2022; Heiberg and Truffer 2022).[7] Explicit application to economic geography research questions has however been scarce. Recently, Truffer et al. (in print) elaborated how STCA inspired approaches could inform economic geography research in at least four ways: i) reconstructing struggles over the directionality of regional industrial pathways, ii) tracing dynamics of asset profiles in individual regions following the arrival of new industries, iii) explaining the where of new industry location in multi-region settings and iv) assessing multi-scalar resource configurations that drive regional path development. An interesting insight of this analysis is that the qualitative semantic network approach is structurally equivalent to the quantitative, complexity and relatedness approaches that have been served as a backbone for a lot of evolutionary economic geography research in recent years (Hidalgo et al. 2007; Neffke, Henning, and Boschma 2011; Balland et al. 2022).

We can therefore conclude this review on semantic network applications by formulating the expectation that a methodological approach like STCA may potentially provide a systematic way to retrace processes of regional path development emphasizing systemic interdependencies, emerging resources, and strategic agency. I will in the following elaborate how this could be achieved for the first of the four potential application domains in economic geography, namely retracing the struggles

[5] A mathematically more explicit presentation of the approach is provided in Truffer et al. (in print)

[6] Coding resulted in 51 concept codes and 568 actor codes, which were attributed to 1'578 text segments, 524 of which related to the centralized configuration and 154 to the decentralized, modular configuration.

[7] In general, one must be very careful when mixing different types of documents in one single analysis, such as newspaper articles and meeting protocols. Selection criteria, modes of recording and imagined audiences will differ for each document type and hence real-world phenomena will be reported following different rationales.

over the dominant industrial pathway in a region and how existing structures and interests of actors came into play when originally trying to prevent but increasingly also shape a new industrial pathway.

4 Regional path dynamics in a leading automobile manufacturing region

STCA was originally developed as a tool to retrace socio-technical transitions. While they have become a more and more salient challenge for environmental and industrial policy, their impacts are increasingly becoming relevant for cities, regions, and countries (Markard, Raven, and Truffer 2012; Tödtling, Trippl, and Desch 2022). The most elaborate recent historical example of such transitions has happened in the electricity sector over the past forty years replacing a socio-technical regime focused on centralized fossil and nuclear power generation, high voltage grids and public utilities, by a new one that is predominantly organized around renewable energies, decentralized grids, and prosumers. In consequence, regions who have been the leaders in let's say the nuclear power industry may have to consider how this socio-technical transition impacts their economic prospects and which of the formerly important regional assets and resources might still be relevant for maintaining a strong position in the future (Gong and Truffer 2024). Socio-technical transitions can therefore take an imminent and potentially dramatic relevance for peripheral as well as for currently dominating industrial regions worldwide.

4.1 Coding newspaper articles documenting the emergence of a new pathway

To illustrate the potential contribution of semantic network methods to inform regional path reorientation processes, I will elaborate on the case of the German federal state of Baden-Württemberg (BW) building extensively on the analysis presented by Gong and Truffer (2024). It will retrace the struggle over the particular automobile manufacturing pathway as it unfolded between 2010 and 2023.[8] While the Gong and Truffer (2024) analysis builds on an extensive campaign of 37 expert interviews in BW, we will focus here on the complementary STCA, which was based on a selection of 170 German language newspaper articles in the Nexis-Uni database that matched with the terms automobile, electromobility and future of the region. The corresponding coding process led to 22 actor codes and 37 concept[9] codes, which were applied to 517 res. 795 text segments (see tables 1 and 2 for an overview on the actual concept and actor codes).

The coding proceeded in several iterations. First, basic concepts were defined in an inductive way, identifying emergent topics and actors covered in the articles. After coding subsets of articles, the coding tree was reconsidered, codes were merged or split up depending on how often they had appeared across the texts and how well they were relating to *a priori* set concepts derived from the theoretical literature in transition studies or economic geography. This process was reiterated several times until the coding tree converged into a robustly interpretable set of coding dimensions and specific codes both in terms of concepts and actors. This procedure is o different from the

[8] Application of STCA to this empirical case was presented in a couple of earlier publications, highlighting different aspects of the transformation and applying different graphical presentations. Gong and Truffer (2024) provide the deepest empirical and conceptual analysis based on an extensive set of expert interviews. The STCA part of that paper mostly served as an additional illustration. Truffer et al. (in print) has a strong methodological focus and used the case to illustrate the data structures that enable the reconstruction of pathways. A reduced earlier version of the present argument will be published as a book chapter. However, only in the present paper elaborates the whole process through STCA with the aim to Illustrate its epistemological foundations.

[9] Instead of thitherto used words like terms, elements, items, or topics, I will use "concept code" from here on to refer to the thematic content coded in the documents alongside the different actors. This is congruent with the technical term used in the earlier literature introducing STCA (see Miörner et al. 2022; Heiberg, Truffer, and Binz 2022).

methodological literature on qualitative case study research (Gioia, Corley, and Hamilton 2013). Table 1 represent the actor codes that resulted from this process. We see that a broad variety of actor types were covered in the newspaper articles ranging from car companies, suppliers, industry associations, civil society, governments at different levels as well as the main political parties. In the coding process, actor types were aggregated in a way to represent rather homogenous views on the topic, while singling out dominating actors. This is the reason why companies like Daimler (car manufacturing) or Bosch (electric appliances) received explicit actor codes, while other car companies (Audi, Porsche) or supplier firms were lumped together in categories of "other" car companies, res. suppliers. Frequency counts across the actors indicate that the government of BW as well as Daimler received most coverage in the data set (60 res. 34 times coded).

C#	Actor type	Explanation	Actor category	Count	C#	Actor type	Explanation	Actor category	Count
1	Media	Statements not attributed to any actor	Media	17	12	Utilities	Energy, water and public transport providers	Utilities	20
2	AFD	Right wing populist	Political party	1	13	Other car comp	Prosche and other brands	Car companies	23
3	FDP	Market liberal	Political party	6	14	Daimler	Mercedes-Benz group	Car companies	34
4	SPD	Social democratic	Political party	6	15	Other supp	Suppliers of car parts	Suppliers	12
5	CDU	Conservative	Political party	8	16	Bosch	Eletric components company	Suppliers	15
6	Green	Environment	Political party	9	17	Research Institutes	Academic and usiness	Research	17
7	Users, citizens, NGOs	Non-company actors	Civil Society	4	18	Universities	Universities	Research	12
8	CS and DB	Carsharing and national rail	Mobility companies	12	19	EU	European Union	Government	5
9	Ind Associations	Industry associations	Associations	11	20	National gov	German federal level	Government	21
10	Trade unions	Trade unions	Associations	17	21	BW gov	Federal state government	Government	60
11	EMBW Consultants	Intermediaries	Consultancies	21	22	Local and city gov	Local governments	Government	21

Table 1: List of the 22 actor codes. They were grouped to represent sufficiently coherent perspectives on the electrification of the car. Larger codes were split up if there were dominant individual actors like Daimler among the car companies. "C#" refers to the code number. "Count" indicates the number of text segments that had been coded respectively.

Table 2 depicts the concept codes, which are grouped along leading themes. The first code (C#1) assembles all statements of actors claiming that the current mobility system is not sustainable and therefore will need to undergo a fundamental transformation. Secondly, reports about actor statements and actions that were defending the superiority of the ICE trajectory are listed from C#2 to C#4, ranging from a strong assertion that the ICE trajectory will remain dominant for the next decades, to a more defensive attitude that the ICE drivetrain will at least not disappear anytime soon and by critical assessments of the EM drivetrain lacking maturity (*EM niche weak*). Arguments in favor of the electromobility trajectory are assembled under C#5 to C#9, encompassing the range from rather lukewarm endorsements (C#5) to statements about EM soon becoming the new dominant design in the car industry (C#6). A particularly important event was the Dieselgate scandal (C#9), in the aftermath of the US EPA proving in 2015 that German car manufacturers had consistently cheated in CO_2 emission tests of their diesel cars. Codes C#10 - C#12 refer to calls for different forms of policy support. C#13 – C#21 contain references to the regional innovation system (RIS) such as challenges (C#18 – C#20), strengths and opportunities (C#13 – C#15, C#21) and leading imaginaries for the old (C#16) and the emerging (C#17) technological trajectory. C#22 and C#23 relate to labor market challenges. The remaining codes (C#24 – C#37) refer to specific innovation and investment activities in all sorts of contexts relevant for a more sustainable automobile sector.

The ultimate concept code list represents a collection of basic concepts that can be aggregated into four main higher order constructs: two are related to the support and critique of the main opposing socio-technical configurations of the ICE and the EM drivetrain. A third one emerged in the context of the environmental movement, which more fundamentally questioned the current mobility trajectory and asked for a stronger promotion of non-motorized and public forms of transport. The fourth construct pivots around the battle over the future regional resource base associated with any of the alternative technological trajectories. This relates to different elements of this resource base, its current strengths and need for change represented by the RIS codes. Each of the configurations consist of diverse technological and institutional elements. The ultimate code list therefore represents a rather

encompassing list of knowledge, technology, formal and informal institutions that one would expect to be mobilized in battles over competing regional pathways regarding initial resource profiles, emerging systemic structures (resources, preferences, and priorities) and agents' strategies (innovations, policies, investments).

C#	Code Name	Short description	Category	Count
1	Transp unsust	Current mobillity form is not sutainable	EM	13
2	ICE dominant future	ICE will be dominant paradigm for decades	ICE	20
3	EM Niche weak	EM will take a long time to mature	ICE	18
4	ICE not dissapear	ICE cars will remain relevant	ICE	19
5	EM relevant option	EM has to be considered	EM	19
6	EM dominant future	EM is the new dominant paradigm	EM	20
7	End of ICE	ICE will have to be discontinued	EM	12
8	Transition needed	Fundamental transformation of region needed	EM	18
9	Diesel gate	Delegitimation of ICE trajectory	EM	3
10	Support policies	Government support for EM needed	Policy	24
11	Research prog and startu	Active funding of EM technology development	Policy	18
12	Demonstr and clusters	Real world implementation needed	Policy	20
13	Strong BW RIS	RIS is abel to cope with challenges	RIS	11
14	Joint RIS development	Need for a cooperative approach to transformation	RIS	26
15	Cross-sector synergies	Regional industrial base can leverage synergies	RIS	6
16	Cradle imaginary	BW has been the inventor of the ICE car	RIS	15
17	Future mob imaginary	A new imaginary rooted in BW is needed	RIS	26
18	BW is lagging	BW is risking to fall behind international competitors	RIS	14
19	Disruptive dynamics	Challenges will lead to disruptions in regional assets	RIS	15
20	Intern competition	International competition is getting more fierce	RIS	21
21	Intern collaborations	Increasing collaboration with international partners	RIS	8
22	Impacts on LabMarket	Labor market will suffer from transition	LaborMarket	24
23	Invest in knowl base	Focused support of skills and capabilities needed	LaborMarket	23
24	Sustainable transport	Sustainable froms of mobility developed	Innovation	12
25	Integr mob concepts	New mobility forms will integrated different means of transport	Innovation	12
26	ICT and KI	ICT and KI will help coping with the transformation	Innovation	14
27	Electric cars dev	Electric cars innovation	Innovation	13
28	Grid capacity	Problems with fueling Ecars on the grid	Infrastructure	17
29	Battery dev	Battery development activities	Innovation	16
30	Charging infra	Need for investment into chargin infrastructure	Infrastructure	34
31	Manufact and infrainvest	Investement into manufacturing and infrastructure	Infrastructure	25
32	New Busin models	New business models in mobility needed	Innovation	10
33	Instit Change	Change in regulatory conditions and cultural attitudes	Innovation	12
34	Fuelcells and Hydrogen	Innovation in fuel cells and hydrogen	Innovation	14
35	Carsharing	Carsharing as a sustainable mobility form	Innovation	8
36	Auton cars	Autonomous cars development	Innovation	5
37	Efuels	E-fuels development	Innovation	4

Table 2: List of the 37 concept codes. "C#" refers to the code number in the text. "Code Name" refers to the node labels in the networks of figures 1-3. "Short description" provides the meaning of the code, and "Category" refers to the major themes that these codes address: being supportive to the ICE drivetrain, the EM future, references to assets of the regional innovation system (RIS), concerns about developments in the labor market (LaborMarket) or references to specific infrastructures or innovation activities. "Count" indicates the number of times the code had been attributed to text segments.

4.2 Identifying higher order constructs

Before engaging in a detailed discussion of the semantic networks, a methodological note on how to identify higher order constructs may be appropriate.[10] Visually, topics or concept codes are represented as nodes of different sizes, shapes, and colors, representing salient characteristics of these concepts like the number of times they were coded, or which theme they represent (see table 2). The positioning of these nodes in a two-dimensional plane follows the principle that concepts, which were often co-mentioned by actors will be put in closer vicinity to each other, while those that were rarely or never co-mentioned will be separated into different "corners" of the plot. The corresponding layout algorithms will optimize the distribution of nodes in a two-dimensional plane according to these principles, but they will achieve this only imperfectly.[11] To consider actual similarity values, we must

[10] The approach of identifying higher order constructs is very well-known from quantitative method textbooks. Principal component or other factor analytical methods enable the identification of higher order – sometimes also called not directly observable – variables. The "meaning" of these constructed factors must be interpreted by knowledgeable researchers based on how these factors correlate with the originally measured variables. The approach presented here is equivalent to factor analysis with the only difference that the originally measured variables are here the coded basic concepts and instead of factor loadings, we interpret vectors of similarity scores. The method is also aligned with standard grounded theory based coding (Gioia, Corley, and Hamilton 2013).

[11] A wide variety of layout algorithms exist in the network analysis literature and are available in network presentation software packages like *Visone,* which we used in the following. The layout algorithm is called backbone layout, which provides a robust graphical clustering of concepts.

additionally look at the links between any pair of nodes. In the below figures, we chose the thickness of a line and its color saturation (between white and dark red) as being both proportional to the calculated similarity values, in our case the Jaccard similarity index ranging between 0 and 1.

Higher order constructs may now be identified by tracing clusters or sub-graphs of nodes that are strongly connected within and have weaker links outside the cluster. We visually identify such coherent clusters by encircling the respective nodes with a closed line.[12] Putting these enclosing ellipses will in general not be clearcut. There will always be links that connect across different clusters. There may also be concepts that sit between two clusters, belonging to both, so to speak. However, this fuzziness in the setting of system boundaries must not be considered a problem. As we will see, such situations may also indicate potential boundary objects that might lead to the merger of formerly distinct constructs in later development phases (for a very nicely fitting analysis in Dutch agricultural transiton, see Runhaar et al. 2020).

Higher order constructs are in general not the same as concept categories, which the coder will identify as overarching themes in the coding process (for instance those listed in the category column in table 2, and which provide the different colors of nodes in figures 1-3). The higher order constructs we are after typically consist of nodes from different concept categories, constituting identifiable "configurations" of elements. For the case of socio-technical configurations these higher order constructs refer to technological and institutional elements that need to be coherently combined to result in "configurations that work" (Rip and Kemp 1998). When aiming at identifying regional industrial pathways, the relevant configurations will consist of references to technological alternatives and regional institutional and capability related resources (Truffer et al. in print). If higher order constructs of interest were institutional logics, then configurations of value concerns, interest positions and legitimation procedures would rather be the constituting elements (Heiberg and Truffer 2022). There is of course an intimate relationship between the conceptual focus of a particular research interest and the kind of documents that will most likely enable to reveal these constructs (see also footnote 3). Obviously, identifying such higher order constructs is a deeply interpretative act executed by individual researchers. Different analysts would potentially draw different ellipses. However, in our experience the degree of cross-analyst variation has been rather limited. It is not unlike the well-known challenge of assuring cross-coder reliability in qualitative text coding, or the above-mentioned interpretation of factors in a principal component analysis (see footnote 6 as well as Gioia, Corley, and Hamilton (2013)).

4.3 Retracing the struggle over future regional industrial pathways

We may now proceed with analyzing the path development dynamics in some detail. Overall, the three time periods follow an "emotional" cycle pattern of denial, experimentation, and endorsement of the new EM trajectory and the corresponding demise of the ICE trajectory.[13] Figures 1-3 represent the semantic concept networks for the three time periods on the right side and the corresponding frequency distributions of actors across categories of topics to the left side. In the first period stretching from 2010 to 2014, the EM challenge was largely considered a minor nuisance for the

[12] We will proceed here with purely visual inspection for the identification of subgroups. Of course, the field of network analysis provides a large variety of community detection algorithms to identify such sub-graphs, but they all come with their respective strengths and weaknesses (Javed et al. 2018). While these approaches may be very useful and even necessary when following an extensive approach to semantic networks, we argue here that interpretative transparency is a strong asset when applying the intensive forms, as we do here.

[13] We ran several robustness tests regarding how to periodize the time dimension by constructing networks for a variety of time slices. The current periodization is the one that resulted in the most distinct network structures and are also most easy to argue for in terms of crucial events that changed the general attitude of the regional actors towards the emerging electromobility option.

dominant ICE-centered regional pathway. 2015-2019 depicts the confluence of the collapse of the defensive position of car manufacturers around the Dieselgate scandal, increasingly imminent policy pressures from the EU to lower CO_2 emissions and a more critical view of government actors towards the willingness of the automobile industry to engage in a future oriented technological trajectory. Finally, 2020 – 2023 represents a situation where the EM trajectory became taken for granted and where the actual shaping of a new regional industrial pathway was endorsed. The higher order construct that shaped the regional path formation indicate attitudes of regional actors regarding strengths and weaknesses of the current RIS and activities around three technological trajectories: ICE, EM and a generic sustainable transport direction. The frequency bars on the left of the figure indicate how actors engaged in the different story lines.

The first period spanning from 2010 to 2014 (figure 1) shows a clear cluster of nodes in the center of the plot (*red ellipse*) around the codes of *EM being a relevant option*, which we would interpret as the way BW actors embraced the newly emerging industrial pathway of electromobility. In this period, the EM cluster connects also statements about the EM technology still being very immature and lacking a proper market (*EM niche weak*), the ICE not disappearing anytime soon (*ICE not disappear*) and substantial policy support needed (the three yellow circles). The core symbolic reference for actors to support this rather critical assessment is the cradle of the automobile trope (*cradle imaginary*), which essentially states that the ICE car had been historically invented in the region[14] and therefore BW did not have to fear any competitor in the future (more extensively, see Gong and Truffer 2024). This core configuration is separated from a second cluster of nodes (*yellow ellipse*) connecting statements about the ICE not losing its leading position in the dominant technological core of the car industry and referring to high costs of any deviation from that trajectory for the region (*Impacts on LabMarket*). The associated higher order construct may be interpreted as representing the extant industrial pathway and concluding that no major deviation in that path is necessary.

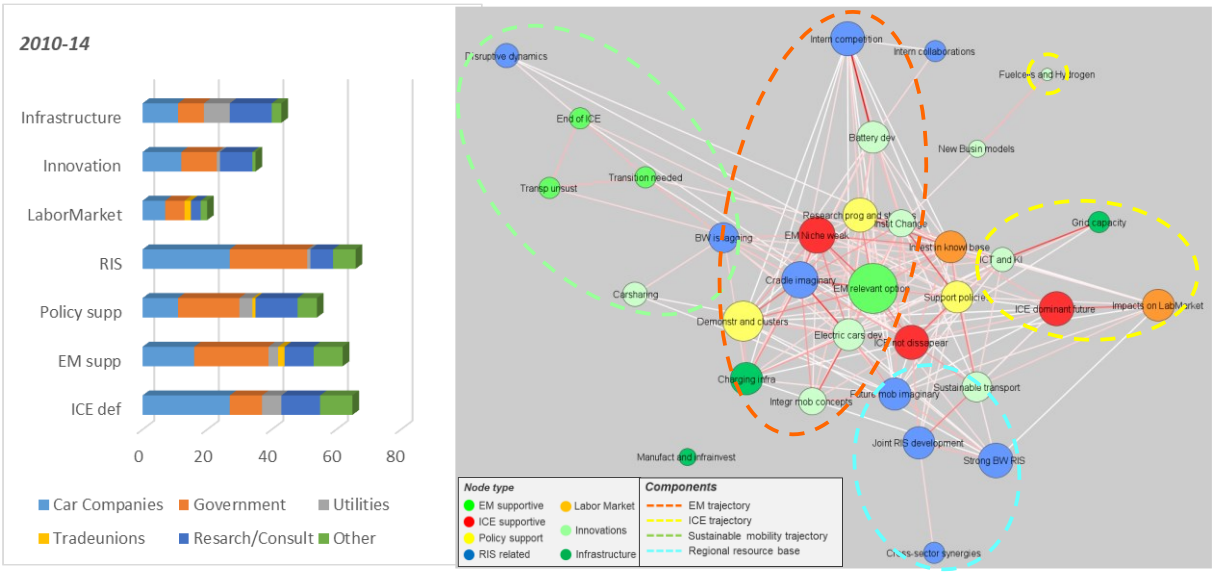

Figure 1: Topic networks (right) and frequency plots (left) of actors and codes for the period 2010 – 2014. Size of the nodes proportional to the number of times the concept had been coded. Thickness of links proportional to proximity between concepts (Jaccard similarity). Layout by backbone algorithm in Visone and deletion of all links with a value lower than 0.33. A legend for the graphical elements is provided at the bottom of the network. The categories of the frequency plots are aggregates of

[14] These were famously two Baden-Württembergers Carl Benz and Gottlieb Daimler who defined the first standards for the internal combustion engine car and are widely considered as its inventors.

concept and actor codes as identified in tables 1 and 2. While there are some similarities between these categories and the higher order constructs, they are based on different rationales.

The frequency bars to the left of figure 1 show that ICE supportive statements and activities were mostly reported from actors of the car sector (manufacturers and suppliers), while the EM supportive concepts were more strongly associated with government actors. Both actor types were equally referring to their belief in the strength of the region and that policy support was needed to cope with this challenge. A third cluster (*green ellipse*) refers to the more radical call for a transformation towards non-motorized mobility forms, which was strongly pushed by the green party. We interpret this cluster as representing a third industrial pathway, one that would be based on non-motorized mobility options with a strong priority on climate protection. Finally, the fourth cluster (*blue ellipse*) assembles all sorts of references to the regional resource base, regional actor strategies and regional cultural elements. We interpret this as representing the regional innovation system and its ability to absorb the challenges associated with the emerging socio-technical alternatives. This RIS concepts are however not very strongly interrelated in the first period and nodes mostly express the conviction that the regional resource base is still strong (*Strong BW RIS; Joint RIS development*) and that there is perhaps need for a new and broader imaginary (*Future mob imaginary*), which will accommodate for all future requirements.

Summarizing, we may say that during this first period the engagement of the regional actors with the theme of electromobility was dominated by a toning-down attitude, where the actors assure each other that nothing dramatic is likely to happen. Regional actors also seem to agree that the new technological trajectory of EM needs to be taken seriously, but they remain confident that the transformation will not be disruptive for the extant regional development pathway, requiring a path extension at most (Grillitsch, Asheim, and Trippl 2018).

In the second period spanning from 2015 to 2019, we see a fundamental shift in the overall network configuration (figure 2). The central cluster of nodes that represents the emerging EM trajectory (*red ellipse*) is now populated by more proactive statements and actions. Core to this cluster are new topics referring to the *Dieselgate* and new statements expecting EM to become the dominant future trajectory in the industry (*EM dominant future*). But also, beyond discursive acts, we see a strong increase in reports about actual investments in EM manufacturing and infrastructure buildup, indicating a more proactive embracement of the new trajectory. Furthermore, concerns about impacts on the labor market have given way to a more proactive view of needing to invest in capability structures required for the new technological trajectory (*Invest know base*). The defensive cluster representing the extant ICE pathway (*yellow ellipse*) also gets stronger referring to limitations of the electricity grid for accommodating masses of electric vehicles, which would limit the size of future EM market potential, as well as expectations towards *E-fuels, Fuel cells and Hydrogen* that would still provide a strong role for the ICE drivetrain in the future.

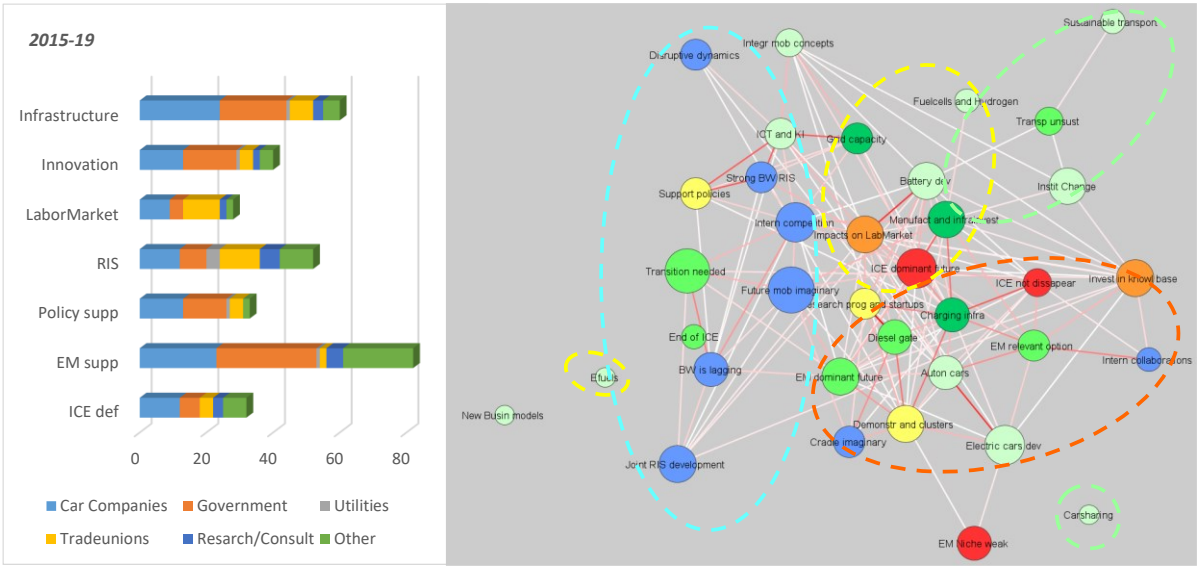

Figure 2: Topic networks (right) and frequency plots (left) of actors and codes for the period 2015 – 2019. Graphical elements as in figure 1.

In terms of the actor distribution across the different coding themes, we see a more proactive representation of car companies in EM supportive topics as well as in infrastructure and manufacturing related activities. Concerns about the regional innovation system structure (*blue ellipse*) are now much more variegated stating that a fundamental *transition is needed*, that the impact on the RIS will be disruptive (*disruptive dynamics*), that the region is lagging its international competitors (*BW lagging*) and explicit statements that the ICE train has no future (*End of ICE*). The core that connects the different clusters in this period is a new imaginary, which for lack of a more specific term was named "the future of mobility made in Baden-Württemberg" by the government (*Future mob imaginary*). Finally, the sustainable transport trajectory (*green ellipse*), which promoted solutions beyond the electrification of the car, seems to have been marginalized compared to the earlier phase.

Summarizing the developments in this period, we see that largely motivated by external events, actors start to reassess the competing technological trajectories, disengaging partially with the ICE pathway and endorsing the new EM pathway much more proactively. The RIS cluster also shows that implications on the regional development pathway are taken up much more proactively, considering the option of a genuinely new path that needs to be developed by investing in new assets and symbolic infrastructures.

Finally in the last period from 2019 to 2023 (figure 3), the dominant impression that the semantic network conveys is the endorsement of the EM trajectory, as being increasingly aligned into a new regional development pathway. The two clusters of the EM trajectory (*red ellipse*) and the RIS resource base (*blue ellipse*) are rather tightly interconnected through nodes like *EM will be the dominant future*, *manufacturing and infrastructure investment*, *investment in knowledge base* and the need for coordinated RIS development (*Joint RIS development*). The opposing trajectories of the ICE car (*yellow ellipses*) and sustainable mobility (*green ellipse*) have substantially weakened. The belief that the *ICE remaining dominant* has even moved to the periphery of the regional pathway space and so has the *cradle imaginary* as the symbolic anchor term for the ICE pathway. In terms of actor distribution in the frequency bars, we see quite a substantial decrease in the coverage of the automobile industry and a much stronger engagement by utilities, and local governments. In a nutshell, we may interpret this endpoint of the analysis as showing the establishment of a new EM pathway, which does not fully replace the old ICE pathway but starts to reshape the regional resource base in a way to make BW able to cope with the requirements of the new technological trajectory.

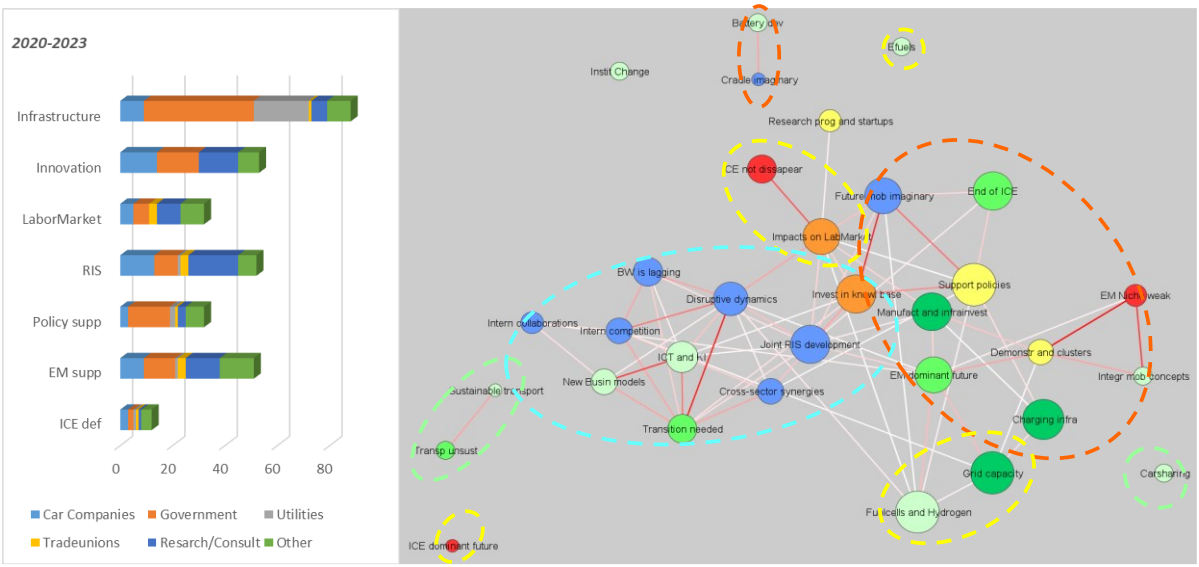

Figure 3: Topic networks (right) and frequency plots (left) of actors and codes for the period 2020 – 2024. Graphical elements as in figure 1.

4.4 Conceptualizing different forms of path development

Overall, the translation of coded documents into semantic networks by means of STCA has illustrated how higher order constructs like socio-technical configurations or regional resource profiles could be systematically retraced. Configurations of concepts in the three periods showed an initial hands-off, even defensive engagement with the new technological trajectory suggesting at most some minor path extension (Grillitsch, Asheim, and Trippl 2018). Over time, this led to the endorsement of a quite radically new regional pathway requiring substantial modification of regional assets (Trippl et al. 2020). The analysis shows furthermore that original capability related resources were indeed relevant for the ability of the region to engage with the new pathway. Related knowledge and capabilities played an important role for embracing the new mobility related options as exemplified by labor market related concerns.

However, this extant resource base did not always play a supportive role for the transformation. Incumbent actors mobilized substantial opposition and even feints to maintain their established technological trajectory (Bækkelund 2021). We also saw that new institutions emerged from the RIS dynamics (e.g. the future mobility imaginary), that visions and expectations played an important role (Steen 2016) represented by believes in the future dominant technological core, and that agency of specific actors made a decisive difference in branching off into a new pathway (Grillitsch and Sotarauta 2020; Grillitsch, Asheim, and Trippl 2018). The specific approach of STCA therefore enabled an explicit and transparent identification of core mechanisms generating systemic resources, new institutional structures and new actor strategies, all elements that are mandatory for a path-as-process perspective.

5 Conclusions and outlook

We started this paper by stating that a path-as-process view on regional industrial development requires an extended set of methodological tools and perspectives and we saw that semantic network may indeed be up to this task. The empirical elaboration of the case of BW illustrates how the emergence of a new pathway can be retraced using a systematic approach based on an in-depth qualitative analysis of a set of documents. A limitation of our empirical analysis is however that we cannot classify the BW case as an actual success. The jury is still out whether the new development pathway will be able to compete with the increasingly sturdy Chinese car companies and newcomers like Tesla. The retracing of configurational dynamics is primarily descriptive, and we cannot identify,

which of the mechanisms or resources would be critical for future success of the new pathway. This would need a more encompassing analysis like the one presented in Gong and Truffer (2024). However, adequate description is always the first step in academic research and needs to be reflected by solid conceptual reflection (Berends and Deken 2021), especially when higher order concepts have to be identified (Gioia, Corley, and Hamilton 2013). According to these criteria, we were able to show how regional paths, technological trajectories, emerging institutional structures, and visionary strategies of regional actors were interrelated and we provided a coherent account of systemic interdependencies and their dynamics.

Regarding the bridging of qualitative process and quantitative variance explanations, we were only able to elaborate on the qualitative side of this equation. The systematic identification of higher order constructs by means of text coding could be extended to encompass broader documents stocks as exemplified by the extensive forms of semantic network analysis. Community detection algorithms combined with statistical social network indicators would enable to identify core concepts such as resources that get mobilized in a struggle over the shaping of pathways. The systematic retracing of development trajectories in different regions and sectors would also enable comparative analyses supported by methods like qualitative comparative analysis, to identify of causal relationships (Rutten 2020).

An expected benefit of a semantic networks inspired approach is to increase the versatility of data structures on which economic geographers could draw. In particular, we believe that the conjoint analysis of standardized data from conventional databases with large text collections, as advocated for by von Bloh et al. (2020), would greatly profit from a semantic network approach. Finally, I would like to reiterate an interesting result of the earlier analysis by Truffer et al. (in print) that the STCA approach is formally equivalent to the established dominant methodological approach of relatedness and complexity perspectives on evolutionary economic geography (Hidalgo et al. 2007; Boschma, Balland, and Kogler 2014; Schetter et al. 2024). This opens for very promising opportunities to combine standardized data with more qualitative text coding approaches. We are not arguing here that such a combination will be seamless, due to diverging ontological and epistemological assumptions that undergird these different approaches (Goertz and Mahoney 2012; Rutten 2020). We believe however that a methodological inroad based on semantic networks opens new bridges where scholars can start to engage in mixed-method analyses essential to embrace a path-as-process or a developmental evolutionary perspective on regional path development.

Acknowledgment

This paper builds on many exchanges over the years with our research group Cirus at Eawag, which has been a highly productive and inspiring place for jointly designing and continuously improving STCA as a novel research methodology. I gratefully acknowledge very helpful exchanges on this paper with our current PhD students Djamila Lesch, Muhil Nesi, Max Hoos and Till Beer, as well as with Christian Binz, Johan Miörner, Alexia Dubuis, Xiao-Shan Yap and Huiwen Gong on diverse conceptual, empirical, and methodological aspects of STCA that led to the present epistemological reflections.

References

Antons, D.; E. Grünwald; P. Cichy; and T.O. Salge. 2020. The application of text mining methods in innovation research: current state, evolution patterns, and development priorities. *R&D Management* **50:329-351**.

Bækkelund, N.G. 2021. Change agency and reproductive agency in the course of industrial path evolution. *Regional Studies* **55:757-768**.

Balland, P.-A.; T. Broekel; D. Diodato; E. Giuliani; R. Hausmann; N. O'Clery; and D. Rigby. 2022. The new paradigm of economic complexity. *Research Policy* **51:104450**.

Baskici, C.; S. Atan; and Y. Ercil. 2019. Forecasting of innovation in the light of semantic networks. *Procedia Computer Science* **158:443-449**.

Bathelt, H. and J. Gluckler. 2003. Toward a relational economic geography. *Journal of Economic Geography* **3:117-144**.

Bathelt, H. and M. Storper. 2023. Related Variety and Regional Development: A Critique. *Economic Geography* **99:441-470**.

Baumgartinger-Seiringer, S.; L. Fuenfschilling; J. Miörner; and M. Trippl. 2022. Reconsidering regional structural conditions for industrial renewal. *Regional Studies* **56:579-591**.

Benner, M. 2023. Revisiting path-as-process: agency in a discontinuity-development model. *European Planning Studies* **31:1119-1138**.

Berends, H. and F. Deken. 2021. Composing qualitative process research. *Strategic Organization* **19:134-146**.

Binz, C.; B. Truffer; and L. Coenen. 2016. Path creation as a process of resource alignment and anchoring – Industry formation for on-site water recycling in Beijing. *Economic Geography* **92:172-200**.

Boschma, R. 2017. Relatedness as driver behind regional diversification: a research agenda. *Regional Studies* **51:351-364**.

Boschma, R.; P.-A. Balland; and D.F. Kogler. 2014. 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* **24:223-250**.

Boschma, R.; L. Coenen; K. Frenken; and B. Truffer. 2017. Towards a theory of regional diversification. *Regional Studies* **51:31–45**.

Boschma, R.A. and K. Frenken. 2006. Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *Journal of Economic Geography* **6:273-302**.

Castaldi, C. and C. Binz. 2024. The normative turn in innovation geography. *Progress in Economic Geography*.

Cho, C.-K. 2022. Semantic network analysis on the geography of food as represented in textbooks: for alternative food education for young people. *Cogent Social Sciences* **8:2148417**.

Coenen, L.; T. Hansen; and J.V. Rekers. 2015. Innovation Policy for Grand Challenges. An Economic Geography Perspective. *Geography Compass* **9:483-496**.

Drieger, P. 2013. Semantic Network Analysis as a Method for Visual Text Analytics. *Procedia - Social and Behavioral Sciences* **79:4-17**.

Eckhardt, Y. and J. Glückler. 2024. Cryptocurrencies, a controversial innovation? Unpacking argumentation analysis in economic geography. *Progress in Economic Geography***:100032**.

Eisenhardt, K.M. 1989. Building Theories from Case Study Research. *The Academy of Management Review* **14:532-550**.

Frank, L.; G. Feola; and N. Schäpke. 2024. Assessing regime destabilisation through policy change: An analysis of agricultural policy in the United Kingdom during Brexit. *Environmental Innovation and Societal Transitions* **50:100810**.

Fu, P.; D. Sarpong; and D. Meissner. 2022. Recalibrating, reconfiguring, and appropriating innovation: a semantic network analysis of China's mass innovation and mass entrepreneurship (MIME) initiatives. *The Journal of Technology Transfer* **47:1506-1523**.

Garud, R.; A. Kumaraswamy; and P. Karnøe. 2010. Path Dependence or Path Creation? *Journal of Management Studies* **47:760-774**.

Geels, F.W. 2022. Causality and explanation in socio-technical transitions research: Mobilising epistemological insights from the wider social sciences. *Research Policy* **51:104537**.
Gioia, D.A.; K.G. Corley; and A.L. Hamilton. 2013. Seeking Qualitative Rigor in Inductive Research:Notes on the Gioia Methodology. *Organizational Research Methods* **16:15-31**.
Glückler, J. and P. Doreian. 2016. Editorial: social network analysis and economic geography—positional, evolutionary and multi-level approaches. *Journal of Economic Geography* **16:1123-1134**.
Glückler, J. and R. Panitz. 2021. Unleashing the potential of relational research: A meta-analysis of network studies in human geography. *Progress in Human Geography* **45:1531-1557**.
Goertz, J. and G. Mahoney. 2012. *A tale of two cultures: qualitative and quantitative research in the social sciences*. Princeton: Princeton University Press.
Gong, H.; C. Binz; R. Hassink; and M. Trippl. 2022. Emerging industries: institutions, legitimacy and system-level agency. *Regional Studies* **56:523-535**.
Gong, H. and B. Truffer. 2024. Changing from within: the interplay between imaginary, culture and innovation system in regional transformation. *GEIST – Geography of Innovation and Sustainability Transitions Working Paper series.* **2024-02**.
Gong, H.; Z. Yu; C. Binz; and B. Truffer. 2024. Beating the Casino: Conceptualizing an Anchoring-based Third Route to Regional Development. *Economic Geography* **100:107-137**.
Grillitsch, M.; B. Asheim; and M. Trippl. 2018. Unrelated knowledge combinations: The unexplored potential for regional industrial path development. *Cambridge Journal of Regions, Economy and Society* **11:257-274**.
Grillitsch, M. and M. Sotarauta. 2020. Trinity of change agency, regional development paths and opportunity spaces. *Progress in Human Geography* **44:704-723**.
Hain, D.; R. Jurowetzki; S. Lee; and Y. Zhou. 2023. Machine learning and artificial intelligence for science, technology, innovation mapping and forecasting: Review, synthesis, and applications. *Scientometrics* **128:1465-1472**.
Hassink, R.; A. Isaksen; and M. Trippl. 2019. Towards a comprehensive understanding of new regional industrial path development. *Regional Studies***:1-10**.
Heiberg, J. and B. Truffer. 2022. Overcoming the harmony fallacy: How values shape the course of innovation systems. *Environmental Innovation and Societal Transitions* **42:411-428**.
Heiberg, J.; B. Truffer; and C. Binz. 2022. Assessing transitions through socio-technical configuration analysis – a methodological framework and a case study in the water sector. . *Research Policy* **51:104363**.
Hidalgo, C.; P.-A. Balland; R. Boschma; M. Delgado; M. Feldman; K. Frenken; E. Glaeser; C. He; D. Kogler; A. Morrison; F. Neffke; D. Rigby; S. Stern; S. Zheng; and S. Zhu. 2018. The Principle of Relatedness: Proceedings of the Ninth International Conference on Complex Systems, 451-457.
Hidalgo, C.A.; B. Winger; A.L. Barabási; and R. Hausmann. 2007. The product space conditions the development of nations. *Science* **317:482-487**.
Javed, M.A.; M.S. Younis; S. Latif; J. Qadir; and A. Baig. 2018. Community detection in networks: A multidisciplinary review. *Journal of Network and Computer Applications* **108:87-111**.
Knees, P.; J. Neidhardt; and I. Nalis. 2024. Recommender Systems: Techniques, Effects, and Measures Toward Pluralism and Fairness. In *Introduction to Digital Humanism: A Textbook*, ed. H. Werthner; C. Ghezzi; J. Kramer; J. Nida-Rümelin; B. Nuseibeh; E. Prem; and A. Stanger, 417-434. Cham: Springer Nature Switzerland.
Lehmann, F. 1992. Semantic networks. *Computers & Mathematics with Applications* **23:1-50**.
Leifeld, P. and S. Haunss. 2012. Political discourse networks and the conflict over software patents in Europe. . *European Journal of Political Research* **51:382-409**.
Lesch, D.; J. Miörner; and C. Binz. 2023. The Role of Global Actors in Sustainability Transitions – Tracing the Emergence of a Novel Infrastructure Paradigm in the Sanitation Sector. *Environmental Innovation and Societal Transitions* **49:100787**.
Li, Y. and F.M.H. Neffke. 2024. Evaluating the principle of relatedness: Estimation, drivers and implications for policy. *Research Policy* **53:104952**.

Liu, Y.; S. Wang; X. Wang; Y. Zheng; X. Chen; Y. Xu; and C. Kang. 2024. Towards semantic enrichment for spatial interactions. *Annals of GIS* **30:151-166**.

MacKinnon, D.; S. Dawley; A. Pike; and A. Cumbers. 2019. Rethinking Path Creation: A Geographical Political Economy Approach. *Economic Geography* **95:113-135**.

Markard, J.; R. Raven; and B. Truffer. 2012. Sustainability transitions: An emerging field of research and its prospects. *Research Policy* **41:955-967**.

Markard, J.; A. Rinscheid; and L. Widdel. 2021. Analyzing transitions through the lens of discourse networks: Coal phase-out in Germany. *Environmental Innovation and Societal Transitions* **40:315-331**.

Martin, R. 2010. Roepke Lecture in Economic Geography—Rethinking Regional Path Dependence: Beyond Lock?in to Evolution. *Economic Geography* **86:1-27**.

Martin, R. and P. Sunley. 2006. Path dependence and regional economic evolution. *Journal of Economic Geography* **6:395-437**.

———. 2015. Towards a Developmental Turn in Evolutionary Economic Geography? *Regional Studies* **49:712-732**.

Miörner, J. 2022. Contextualizing agency in new path development: how system selectivity shapes regional reconfiguration capacity. *Regional Studies* **56:592-604**.

Miörner, J.; J. Heiberg; and C. Binz. 2022. How global regimes diffuse in space – Explaining a missed transition in San Diego's water sector. *Environmental Innovation and Societal Transitions* **44:29-47**.

Miörner, J.; B. Truffer; C. Binz; J. Heiberg; and X.-S. Yap. 2022. Guidebook for applying the Socio-Technical Configuration Analysis method. . *GEIST – Geography of Innovation and Sustainability Transitions Working Paper series.* **2022-1**.

Neffke, F.; M. Hartog; R. Boschma; and M. Henning. 2014. Agents of structural change.The role of firms and entrepreneurs in regional diversification. *Papers in Evolutionary Economic Geography, Utrecht University* **14.10**.

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

Ozgun, B. and T. Broekel. 2021. The geography of innovation and technology news - An empirical study of the German news media. *Technological Forecasting and Social Change* **167:120692**.

Rakas, M. and D.S. Hain. 2019. The state of innovation system research: What happens beneath the surface? *Research Policy* **48:103787**.

Ricci, F.; L. Rokach; and B. Shapira. 2022. Recommender Systems: Techniques, Applications, and Challenges. In *Recommender Systems Handbook*, ed. F. Ricci; L. Rokach; and B. Shapira, 1-35. New York, NY: Springer US.

Rip, A. and R. Kemp. 1998. Technological Change. In *Human choice and climate change - Resources and technology*, ed. S. Rayner and E.L. Malone, 327-399. Columbus.

Runhaar, H.; L. Fünfschilling; A. van den Pol-Van Dasselaar; E.H.M. Moors; R. Temmink; and M. Hekkert. 2020. Endogenous regime change: Lessons from transition pathways in Dutch dairy farming. *Environmental Innovation and Societal Transitions* **36:137-150**.

Rutten, R. 2020. Comparing causal logics: A configurational analysis of proximities using simulated data. *Zeitschrift fur Wirtschaftsgeographie* **64:134-148**.

Schetter, U.; D. Diodato; E. Protzer; F. Neffke; and R. Hausmann. 2024. From Products to Capabilities: Constructing a Genotypic Product Space. *Growth Lab Working Paper,* **Harvard University Kennedy School of Government**.

Segev, E. 2022. Semantic network analysis in social sciences. London: Routledge.

Sevin, E. 2014. Understanding Cities through City Brands: City Branding as a Social and Semantic Network. *Cities* **38:47–56**.

Shim, H.; Y. Kim; G. Minton; D. Nan; Y. Kim; H. Lee; S.H. Lee; and J.H. Kim. 2022. Network analysis of open innovation in the era of the Fourth Industrial Revolution. *Asian Journal of Technology Innovation* **30:538-558**.

Sotarauta, M. and M. Grillitsch. 2023. Path tracing in the study of agency and structures: Methodological considerations. *Progress in Human Geography* **47:85-102**.
Steen, M. 2016. Reconsidering path creation in economic geography: aspects of agency, temporality and methods. *European Planning Studies* **24:1605-1622**.
Strambach, S. and G. Pflitsch. 2020. Transition topology: Capturing institutional dynamics in regional development paths to sustainability. *Research Policy* **49:104006**.
Strauss, A. and J. Corbin. 1998. *Basics of Qualitative Research*. Thousand Oaks: Sage Publications.
Suurs, R.A.A. and M.P. Hekkert. 2009. Cumulative causation in the formation of a technological innovation system: The case of biofuels in the Netherlands. *Technological Forecasting and Social Change* **76:1003-1020**.
Ter Wal, A.L.J. and R.A. Boschma. 2009. Applying social network analysis in economic geography: Framing some key analytic issues. *Annals of Regional Science* **43:739-756**.
Tödtling, F. and M. Trippl. 2018. Regional innovation policies for new path development – beyond neo-liberal and traditional systemic views. *European Planning Studies* **26:1779-1795**.
Tödtling, F.; M. Trippl; and V. Desch. 2022. New directions for RIS studies and policies in the face of grand societal challenges. *European Planning Studies* **30:2139-2156**.
Trippl, M.; S. Baumgartinger-Seiringer; A. Frangenheim; A. Isaksen; and J.O. Rypestøl. 2020. Unravelling green regional industrial path development: Regional preconditions, asset modification and agency. *Geoforum* **111:189-197**.
Truffer, B.; C. Binz; J. Miörner; and X.-S. Yap. in print. Bridging the methodological divide: Inspirations from semantic network analysis for (evolutionary) economic geography. *Progress in Economic Geography*
Uitermark, J. and M. van Meeteren. 2021. Geographical Network Analysis. *Tijdschrift Voor Economische En Sociale Geografie* **112:337-350**.
von Bloh, J.; T. Broekel; B. Özgun; and R. Sternberg. 2020. New(s) data for entrepreneurship research? An innovative approach to use Big Data on media coverage. *Small Business Economics* **55:673-694**.
Wu, L.; Q. Peng; M. Lemke; T. Hu; and X. Gong. 2022. Spatial social network research: a bibliometric analysis. *Comput Urban Sci* **2:21**.
Yap, X.-S.; J. Heiberg; and B. Truffer. 2023. The emerging global socio-technical regime for tackling space debris: A discourse network analysis. *Acta Astronautica* **207:445-454**.

For further information about the
GEIST Working Paper series, see:
www.geist-wp.com

Environmental Social Sciences
Eawag - Swiss Federal Institute of Aquatic Science and Technology
Contact person: Johan Miörner
Überlandstrasse 133, 8600 Dübendorf, Switzerland
Tel.: +41-587-656493
E-Mail: johan.miorner@eawag.ch
https://www.eawag.ch/en/department/ess/

Department of Geography and Regional Research
University of Vienna
Contact person: Michaela Trippl
Universitätsstraße 7/5/A0528, 1010 Vienna, Austria
Tel.: +43-1-4277-48720
E-Mail: Michaela.trippl@univie.ac.at
https://humangeo.univie.ac.at/