12 Data-driven transitions research

Methodological considerations for event-based analysis

Fjalar J. de Haan, Alfonso Martínez Arranz, and Wouter Spekkink

1 The case for data-driven transitions research

We argue that serious progress can be made in transitions studies if a more data-driven approach to empirical analysis is adopted. By 'data-driven' we mean a thorough-going empiricism, where data, as many as can be mustered, lead rather than just support inference and hypotheses. Obviously, all empirical research worth its salt will relate its hypotheses to empirical data. What we suggest is a thoroughly systematic attitude towards this. An approach that enables anyone to verify exactly which data are adduced to corroborate what hypothesis – and to perhaps disagree on the basis of those data and propose rival hypotheses. We argue for an approach that facilitates dealing with many data, across cases and using statistical methods where appropriate, an approach amenable to meta-analyses. These are all issues of *methodology* and this chapter is explicitly about methodological considerations. We are, in this regard, theoretically agnostic.

This, of course, does not mean we do not have theoretical and philosophical convictions and it certainly does not mean we think transitions research should be theory averse at all. What we are after is an intensely empiricist methodology, that *postpones* theoretical framing and interpretation as much and as long as possible. The outcome of applying such a methodology should be a practical, shareable and cumulative collection of data that can then be used to test theories and hypotheses of whatever ilk. A sort of 'separation of powers' between the empirical and the theoretical, as it were. Whether one thinks such a separation is possible in principle or not, we think it is something to strive for.

What would such a data-driven approach practically look like? We envision that transition analysts would query a large database of empirical data coming from many cases. These data would be recorded in a standardised way, whilst retaining the detail of the original sources and providing references to them. The analyst would try to identify patterns and the conditions under which they occur ('do all backlashes look the same?', 'is government action relevant for niche formation?'). Hypotheses would be tested straightforwardly by searching for corroborating, contradicting and correlating data ('how often do we find scaling up following legislative change?'). This is clearly not the way transitions research is commonly carried out. Not only is there no such database but moreover it seems contrary to the way data are used in transitions research.

There is, however, no shortage of empirical work done in this field, which further convinces us of the feasibility of an approach such as we propose. For instance, transitions research relies heavily on case study narratives for its underpinning and the field has amassed a considerable body of empirical data in this form. Meta-analysis of the existing empirical literature may be an appropriate way to bootstrap data-driven transitions research, whilst enabling a much-needed re-examination of basic assumptions. A meta-analytic study of this kind was carried out by one of us (Martínez Arranz, 2017), more on this in Section 4.2. Another way is to enhance a common form of empirical research in the field: the study of innovation projects. One of us has been building on an approach where qualitative data on such projects are bracketed and recorded in standardised data sets (cf. Van de Ven and Poole, 1990), after which they are coded to make them amenable to visualisation and analysis with, for example, graph-theoretical tools (Spekkink, 2015). This approach will be discussed further in Section 4.1.

We are not just preaching what we practice though. We observe that data-driven methods have been successfully adopted in areas of research that share several challenges sometimes thought of as typical for transitions, e.g., complex chains of causes and effects, qualitative data, historical and geographical contingency and so forth. In this regard, we think of fields like medicine and psychology, which also share the ubiquity of case research but with an emphasis on replicable conclusions and where resolving conflicting conclusions is a key concern. Meta-analysis, one of the approaches we advocate, is a common data-driven method in these fields.

Data-driven methods necessitate a different way of engaging with empirical material, as well as thinking differently about explanation and theory building. Both, in a fashion, need to be disaggregated. From cases to their component facts and from theoretical frames to mechanisms[1]. We are not simply advocating for comparative case research or $N > 1$ case studies. We are pointing out that case studies may suggest promising generalisations and hypotheses, but that any *particular* case exemplifying them is nothing but precisely that: an example — which may well be unique. The phenomenon that 'theoretical frameworks' accrue more and more 'nuance' as they are 'tested' on more and more case studies is therefore a clear indicator of a problematic relationship with data (as well as questionable theorising[2]).

Data-driven research is not new, nor is it confined to particular areas of inquiry. However, of recent, the term 'data-driven' has gained popularity and is being used in much the same context as the term 'big data'. While our conception is certainly akin to this popular usage, in particular in its emphasis on statistics and postponing of theorising, we are not suggesting an 'algorithms-to-the-rescue' attitude but a rather more old-fashioned and common-sense approach: the Baconian method[3]. What is clear, at least to us, is that computational tools are opening up new ways of doing research on complex societal issues such as transitions. This is not just a matter of increased raw computing power, it is also the potential of new ways to organise research and sharing data inspired by the modus operandi of the free and open source software community.

2 The data of transitions research

What are the data of transitions research? And where to procure them? To get a better idea of this, we will look into what kind of science transitions research would be, if it were one. Transitions research would be a *historical science*, in the sense Diamond (1997, Epilogue) used it. This is not to say that all transitions have now passed, nor that we can or should not study ongoing or future transitions. It means that most questions in transitions research are of the form: 'why did things unfold in this particular way?' or 'how did this situation come about?'. Though their subject matter is at first blush wildly different, this reveals transitions research to be of the same family as evolutionary biology, geology, cosmology, historical sociology – that is, the historical sciences which include history itself and, for example, historical sociology in as far as they are scientific.

This suggests that the data of transitions research ought to be descriptions of the transitioning system over time. Let us call such a description, in line with the literature, a *pathway*. For the purposes of our discussion we will distinguish two main approaches to describing pathways. The first approach describes a pathway as a discrete sequence or continuous progression of *system states*, where each system state is described, for example, in terms of a set of key variables. The second approach describes pathways as sequences of *events* that capture episodes in which a system changes from one state to another, in some qualitatively or quantitatively appreciable way. While the first approach focusses on describing the qualities of the system over time, the second approach focusses on describing the *changes* to that system over time.

In principle these approaches are equivalent in terms of the information they carry about a transition pathway. If the state of the system is known at *some* time on the pathway, one can reconstruct the system-state description over the *entire* pathway by cumulatively applying the changes each event describes. Conversely, an event sequence can be distilled from a system-state description by aggregating appropriate clusters of change into events. For the mathematically inclined, the former approach represents pathways with the system state, S, as a discrete or continuous function of time, that is, St or $S(t)$. The latter approach represents pathways in terms of a difference or differential equation of the system state as a function of time, like $\frac{\mathrm{d}S(t)}{\mathrm{d}t} = f(t)$ or $\Delta S_t = f(t)$. Clearly, if some initial – or, in fact, any – condition is known, the former can be completely reproduced with the latter by integrating or summing the equations. In other words, the representations are equivalent.

Our distinction parallels Abell's (2004; 2007) distinction between *states of the world* (system states in our language) and *actions* that transform elements in states of the world (events in our language). Like Abell (2007), we suggest that, although both system states and events can be considered together in analysis, most often our interest will be primarily in understanding what changes have led from one system state to another, without necessarily wanting to explicitly consider each intervening system state that lies on this path. Moreover, this seems more parsimonious, as an event description would only mention those aspects (i.e., key variables) that mattered in that event, whilst others that did not change or followed some uninteresting trend can be ignored. In fact, it allows one to forego the exhaustive identification of key variables altogether[4]. Finally, an orientation on events also puts the focus of analysis more explicitly on how changes occur (Van de Ven and Huber, 1990), which is a key question in transitions research. In summary, in most cases we would want to describe our pathways in terms of events.

3 Methodological considerations regarding events

At the heart of our methodological proposition is the event as the 'unit' of empirical data. Therefore, some words on what we consider an event to be and how to recognise a description of one seem appropriate. An event is a change in some state of affairs occurring within a bounded period of time. Casually speaking, things are not the same after an event. The definition allows events to be of any duration *and* instantaneous.

Our definition of event leaves open a host of methodologically relevant issues, which we will discuss shortly. Amongst these are: what matters about an event, that is, what should one record about an event? What about agency, i.e. who is responsible for the event? Can events be part of other events? And, moreover, which events matter? Or, similarly, how heavily does one event count relative to another? Especially these latter two point are slippery. Lest these issues be used as arguments against a data-driven, event-based analysis of the kind we propose, it should be noted that these issues are of course also faced by the narrative analyst. In narratives, however, the choices and rationales of the analyst are typically implicit, if not unconscious. At least in a data-driven, event-based analysis such choices are transparent and can be the object of scrutiny. Indeed, the importance of certain events may then be established as a *conclusion* of the analysis, rather than remaining an unquestionable assumption.

Thus, we propose to view transition pathways as sequences of events, though we do not exclude at all the possibility that events occur in parallel, so perhaps we should be speaking of bundles, or even networks of events instead, as the sequences may interact or cross over as it were. Such a view has two characteristics which we think are advantages, namely that it is

- rather similar to the way narrative case studies are usually presented, which provides opportunities for *meta analysis*, and
- there is no advance commitment to a choice of variables, leaving identification of key variables as a potential research finding.

With the choice for an event-based perspective made, we can now explore a number of methodological considerations. What is involved, practically, in doing data-driven transitions research from an event-based perspective. Where to find events? How to record them? And, importantly, what to do with them? We propose a methodological frame consisting of two main parts:

1 Cataloguing of events
2 Categorisation of events

The aim of the first is to establish, with as little interpretation or application of theory as possible, a database of events. Such a database could be fed from primary research done in a particular project or built from a corpus of pre-existing literature. Ideally, a shared and openly accessible, online database would emerge — a sort of transitions data commons[5]. We will refer to such a database as an *event catalogue*, or simply catalogue. What we have in mind is very close to the event data sets described by Van de Ven and Poole (1990). The second is where actual analysis occurs. Categorising events is using events as data points, the equivalent of interpreting measurements. The categorisation of event data can be part of a theory-building approach, where categories are gradually constructed and refined through systematic comparison of event data (Langley, 1999), but categorisation can also entail sorting events into premade categories, (e.g., ‘entrepreneurial activity’, ‘scaling up’), or directly tallying them as corroborating or refuting some hypothesis. Though the term ‘categorising’ may sound somewhat simplistic, all acts of theorising on the basis of data are instances of categorising events.

In the following, we will outline our methodological considerations and suggestions for building event catalogues. These considerations are based on our personal experience as well as informed by literature. After that, in Section 4, we will discuss several approaches that make use of an event catalogue. Some of this is again reporting on our own experience, while some of it describes work by others or new avenues.

3.1 Cataloguing events

The events in the catalogue should be what we could call *statements of historical fact*. This means two things: Firstly, it means that the catalogue should only contain events that either no-one would dispute to have actually happened[6], or that could be checked to verify[7] whether they happened or not, for example by simply checking the sources cited. Secondly, it means that the events should contain as little interpretation as possible, that is, one should be able to accept the described event as a historical fact without having to accept some additional theory or to hold particular values. We realise that especially this latter criterion may lead to meta-analytic paralysis if taken at philosophical extremes. We do, however, maintain that in practice it is mostly straightforward to distinguish between *events proper* (i.e. statements of historical fact) and *interpreted* *events*. Van de Ven and Poole (1990) make a similar distinction between *incidents* and *events*, where the former is a bracketed, factual description of an occurrence of interest, and the latter is a theoretical construct that we might arrive at through a theory-informed interpretation of incidents.

We said that an event is a change in some state of affairs. That is, something ought to have happened for an event to have occurred. To jest, the plaque in Columbus, Ohio, USA (Wikipedia, 2018) with the text 'On this site in 1897 nothing happened.' does not commemorate an event. Note that periods of stasis can in this fashion still be informative as their lack of events can positively be noticed. We explicitly allowed events to be instants (e.g., New Year) as well as anything longer. Thus, a historical transition as a whole may count as an event. This brings us to another important aspect of events, which is that they can be *nested*. Similar to the idea that a set of basic actions can together constitute a more complex action (Danto, 1965), we suggest that a set of 'smaller' events can constitute a more complex event.

The ability to view events as composed of other events introduces a natural way of identifying levels of analysis. One may seek explanations on a given level of analysis, based on the nature of the question or on practical considerations, like the level at which data are available. Analysing events into their constituent 'sub' events in itself can be enlightening. For example, the approach outlined in this chapter largely revolves around attempts to increase our understanding of transitions (which are deeply complex events) by looking at the smaller events that compose them. Obviously, there are limits to the added utility of more and more finely grained events – we do not expect many events at the level of molecular interactions to be particularly interesting for the purposes of understanding the course of a transition. In practice, the available data sources impose a lower limit to the level of analysis and higher levels can be accessed by aggregating events into complex events.

To speak of an event proper, some temporal delineation is necessary, though in practice sources are often vague or ambiguous on this matter. For example, the statement '*E* happened in 1977' could mean anything from *E* occurring at some instant in that year to *E* taking that entire year. This may require the keen judgement of the analyst to interpret – or additional background checks. In addition to the temporal aspect, there are other aspects that can be part of a proper event. For example the spatial aspect (whatever happens, happens somewhere, though a spatially contiguous requirement seems too strong[8]). Depending on the purposes of the analysis, the precise relative timing of events may be of crucial importance, for example, if the priority of some action or other leads to a different conclusion as to the responsibility for some event. Or, it may matter very little and the approximate order in which the events took place is sufficient. Either way, again the data sources usually impose a lower limit to the precision of temporal delineations.

If we agree that history is shaped by human action, there will often be an agency aspect to an event as well. Indeed, Poole et al. (2000) define events as what entities do or what happens to them. In the context of transitions research, identifying the key agents in a sequences of events may well be the principal aim of the analysis. Thus we are interested in the aspect of agency of events. The notion of agency introduces some additional methodological issues, for two main reasons:

- Agency is bound up with the notion of causation. Causation is some thing one may wish to establish, based on analysis. Therefore, to avoid circularity, it should not be *assumed* in order to establish agency in relation to events.
- What entities one allows as agents is not exactly uncontroversial. While common sense (and the law) would restrict agency to adult humans who are *compos mentis*, some scholars would maintain that technological artifacts or trees are agents. This would mean that identifying agents related to events entails a theoretical commitment, making the event an *interpreted* event, whereas we want to avoid interpretation until the categorisation stages of analysis.

Nevertheless, in building a catalogue of events based on a corpus of source data, one most likely wants to catalogue the agents as well. What we propose is an approach to cataloguing events that is as agnostic about matters of causation and theory as possible. We propose that in addition to, but separately from, the event proper, one catalogues the agents *associated* with the event *according to the sources*. To avoid second-guessing who ultimately caused something, associating agents with events should be on the basis of sources describing them as if they were necessary conditions for the event – though sources obviously need not use this terminology. Where the agency of an event cannot be resolved from the sources, one could associate with the event an 'unknown' agent, or 'forces of nature', depending on what the case may be.

Compare in this regard the uncontroversial 'James Watt improved the Newcomen steam engine near the end of the 18th century', with the more problematic, perhaps even dubious 'near the end of the 18th century, James Watt helped bring about the Industrial Revolution'. This latter statement does not describe an event proper, it is a (rather heavily) *interpreted* event. It could however be the *conclusion* of the analysis, or it may be taken as a hypothesis to be tested. Note, however, that from a text fragment such as the latter an analyst *can* safely catalogue Watt's technological innovation as an event, and, moreover, safely catalogue Watt as an associated agent for this event. Distinguishing between various classes of agents can be a straightforward and informative exercise that would be part of the event categorisation side of the analysis.

Summarising this into an event-cataloguing procedure we yield the following:

- Collect from the raw research data (if primary research) or from the original sources (if relying on existing literature) those parts that describe proper events, statements of historical fact.

- Distill, from those fragments, the temporal aspect of each event. That is, for each event, determine the beginning and end point in time (taking these to coincide for instantaneous events).
- Describe the change in the state of affairs that each event effected. In other words, describe what happened and keep it matter-of-fact. This should, at most, amount to a rewording or synopsis of what the sources say, refraining from interpretation or judgement.
- Identify the actors associated with each event. Again, this does not mean an inquiry into who *ultimately* caused an event. Record who did what was done in the event in as far as the source reveals it[9].

We keep emphasising that the event catalogue as such should steer clear from adding a layer of interpretation. The catalogue should be as close to a structured data version of the original sources as possible. The main aim of the data-driven approach is to facilitate transparent, systematic, scientific enquiry — any interpretation, conclusion or inference should be based on the data in the catalogue so it can be traced back to the analyst and the sources used. Note that this requires that the events in the catalogue are appropriately referenced, so that they also may be scrutinised. One may wish to keep a companion catalogue of source fragments, for maximum transparency and ease of reference.

4 Approaches to transitions event analysis

In this section we will discuss a variety of approaches to transitions analysis based on an event catalogue or database of the sort we have been proposing. The section is divided into two parts, one part focussing on work that some of us have actually done (the *What We Have Done* sections) and the other part focussing on work that is in progress, done by others in different contexts, or speculative (the *What Could Be Done* sections).

Perhaps as a clarification in advance, the sections on *What We Have Done* do not present work that was based on the methodological considerations outlined earlier, nor do these considerations follow directly from the experiences we present there. What we aim to provide is examples of data-driven, event-analysis based work that we engaged with ourselves. The *What Could Be Done* parts that follow discuss approaches, techniques or technologies that we think can play an important role in actually performing data-driven event analyses. The selection we discuss is no doubt woefully incomplete but these are approaches that we are personally particularly interested in or already working with. For an overview of various other ways in which one might 'work' with event data, see for example Langley (1999).

4.1 What we have done (i) – event network analysis

One of us has been involved in the development of a methodology that makes use of the kind of event data sets that we envisioned in the previous section

(Spekkink, 2015; Spekkink and Boons, 2016; Boons et al., 2014; Spekkink, 2016). This methodology was inspired primarily by the methodological advances that came out of the Minnesota Innovation Research Programme (Van de Ven and Poole, 1990; Poole et al., 2000), by Abell's (1987; 1993) theory and method of Comparative Narratives, Heise's (1989) Event Structure Analysis (also see Griffin, 1993) and by Abbott's (1988; 2001) thinking on event-based methods.

A core idea underlying the methodology is that social processes can be usefully modelled as networks of events. The networked nature of the events arises from the observation that social processes often consist out of multiple 'streams' of events that sometimes occur in parallel, sometimes diverge into multiple streams, and sometimes converge onto a single stream (Van de Ven et al., 2008; Van de Ven, 1992). To capture this complexity, we make use of *event graphs*[10] in which social processes are visualised and analysed as directed a-cyclic graphs. In these graphs nodes represent events, and the edges represent relationships between events. This approach is thus especially useful for answering research questions about patterns of relationships between events. For example, Spekkink and Boons (2016) used this approach to examine how regional collaborations on sustainable industrial cluster development emerged at the intersection of collections of smaller projects that initially unfolded independent from each other. In this use case, the approach was instrumental in tracing the collaborations to their various building blocks.

The types of 'event catalogues' that we discussed earlier are a useful starting point for the application of the methodology, which originally also starts out with the creation of event data sets that capture relevant occurrences in the process of interest in the form of chronologically ordered, bracketed, qualitative descriptions. These data are then coded with qualitative data analysis software[11] to identify attributes of events and *relationships* between events. The attributes capture theoretically meaningful aspects of events, and can be used, for example, to classify events into different types, to identify actors involved in the events, to identify the locations where events occurred etc. The relationships between events capture how events are connected to each other into complex sequences. These relationships can be of various types. For example, we may define a relationship that captures how events *shape the conditions* under which later events occur (similar to Abell's (1987) paths of social determination), or we may define a relationship that captures how events occur *in response to* events that happened earlier (similar to Schatzki's (2016) chains of action). As in other forms of network analysis, offering a clear definition of the type of relationship under consideration is one of the main tasks in the construction of event graphs, and it will have important consequences for the type of event graph that is produced.

What types of analysis then, does the reconstruction of social processes in event graphs enable? Since event graphs are essentially a network representation of processes, one obvious candidate is the use of network analysis. For example, Spekkink and Boons (2016) used community detection algorithms to identify *modules* in the processes they investigated, and examined *paths* of relationships in the network as one step in the identification of 'emergent linkages' between events. In another example, Mu and Spekkink (2018) used the *outdegree* of events as a measure of their importance in bringing about new events. The use of network statistics possibly also paves the way for the development of statistical models used in other types of network analysis. For example, exponential random graph models Lusher et al. (2013) might be useful for examining the tendencies of different types of events to connect with each other, while also taking into account other characteristics of the events, such as their time and place of occurrence, and the actors involved in them. However, networks of events have some peculiarities, such as the temporal ordering of the nodes, and their a-cyclic nature, which have to be taken into consideration when specifying and interpreting models, for which there are currently few precedents.

We also strongly encourage to always make the step back from examination of abstract patterns in event graphs, derived through quantitative analysis, to an interpretation of the underlying qualitative data. Event graphs may be useful in bringing out patterns that otherwise remain 'buried' in the rich qualitative data, but a deeper understanding of these patterns typically requires going back to the qualitative descriptions of events to flesh out these abstract patterns.

4.1.1 Methodological reflections

Event network analysis has some commonalities with the data-driven approach introduced in this chapter, primarily with regards to how events are defined, and how they are recorded in data sets. However, there also important differences to highlight. In previous applications of event network analysis (e.g., Spekkink and Boons, 2016; Mu and Spekkink, 2018) decisions about what events to include in the data sets were already from the outset heavily influenced by the particular research questions and theoretical perspective of the analysts. Consequently, the resulting data sets are inevitably biased in what events are included and (perhaps more importantly) what events are excluded, or ignored. Even if the analysts are careful in recording *events proper* rather than *interpreted events* (see Section 3.1), their particular interpretations are implicit in the boundaries of the data set. The data-driven approach, by actively postponing interpretation, encourages the inclusion of a much wider range of events, making the data suitable as an empirical basis for tests of a wider range of theories.

In event network analysis, the particular interpretations of the analysts play an even stronger role in the stages that follow data collection. The fact that event attributes and relationships between events are identified through qualitative coding means that they capture, first and foremost, the analysts' interpretation of the data, rather than an objective account of how the process unfolded (cf. Heise, 1989; Griffin, 1993)[12]. However, this is generally true for any narrative account (including case descriptions) of social processes. The advantage of explicitly coding for attributes and relationships between events is that the choices made by the analyst in this regard are made transparent. Coding the event data in the ways described above makes them amenable to various types of quantitative (network) analysis. However, in performing these analyses one should always take into consideration that it is a subjective account of a social process that is being analysed, and not the social process 'an sich'.

The comparison with the data-driven approach reaches its limits here, because the qualitative coding of events is already a step beyond the cataloguing of events. Instead, most of the steps of event network analysis can be understood as a particular approach to *categorising events*, perhaps with the addition of *categorisation of relationships* between events, which is something we have not explicitly discussed as part of the data-driven approach (see Section 3). Ideally, the data collection stage of event network analysis is made redundant by the creation of event catalogues through a data-driven approach. It is replaced with the task of sampling relevant events from the wider catalogue. The added benefit of this approach is that the analyst has to justify not only decisions about what events to include, but also why other events in the catalogue were not included. Everything that follows this stage (the qualitative coding of events, the creation of event graphs, the summary of analysis in a narrative etc.) would still be imbued with the particular interpretations of the analysts (various checks of inter-coder reliability are available to compensate somewhat for this). However, since the catalogue of events is available to a wide audience, possibilities to scrutinise these interpretations improve, and comparisons with the interpretations of others (possibly based on other types of analysis) are easier to make as well.

4.2 What we have done (ii) – meta-analysis

Another one of us undertook meta-analysis of existing literature aimed at deriving 'lessons from the past' (Martínez Arranz, 2017). Meta-analysis is a versatile, resource-saving research tool that, in its simplest definition, relies on primary analyses to produce a secondary analysis. In other words, it re-uses and re-appraises the abundant good work carried out by a large number of scholars and institutions to elicit new insights (Borenstein et al., 2011). Given our considerations above, regarding positive parallels to other sciences like medicine and the large number of single-case transition studies, the use of meta-analysis in transitions research has been scarce to say the least: an inclusive search in Scopus for 'meta-analysis' and 'technolog* transition*' within social science, energy or environment journals yields exactly *two* directly relevant results: an article by Raven et al. (2016) and the above contribution by one of us. Wiseman et al. (2013) also appears but its evaluation of *future* transition strategies as published in government reports is only tangentially related to our work here, which deals with historical facts.

At a fundamental level, Raven et al. (2016) concur with us that large *N* studies are required to move the field of transitions forward. In their work they test a number of propositions against six low-carbon technology cases. The empirical section of the article is based on original research by the authors and, in that sense, it is a comparative analysis rather than a meta-analysis proper. Nevertheless it certainly is a big step forward from the usual focus on one or two transitions or technologies. A downside of their approach is that the original sources are not mentioned when the cases are used and compared, nor is the raw data available, e.g., as a table in an appendix or online. This is not to say we do not trust their findings (we do), rather, we and other researchers are now not in a position to verify those findings, nor to use their data to test other hypotheses. Making 'uninterpreted' data available is an integral part of the approach we advocate.

For its part, the meta-analytical work in Martínez Arranz (2017) more closely combined the two efforts that we have described in this article: on the one hand, there was a focus on extracting (and *abstracting*) elements from sources, although focusing on 'factors of change' rather than 'events'. These were identified for each of the 34 transitions analysed. On the other hand, there was categorisation by grouping these factors of change into more abstract categories that denote the motivations behind actions, for example: health & lifestyle, economic opportunity, economic necessity, ideals, etc. Despite the larger number of cases, issues remain around the validation of both the abstractions and the categories, which prompted some of our reflections here.

4.2.1 Methodological reflections

There has been criticism of meta-analysis when reduced to a mathematical/statistical tool, but the requirement for systematic and consistent analysis in the face of multifarious evidence only emerges reinforced after such criticism, even if with a different name (Stegenga, 2011). We are not married to the label of meta-analysis or to the statistical procedures used in medicine or other fields. Meta-analysis is merely one example of a data-driven tool that has important teachings for our event catalogue approach.

Amongst the well-known difficulties in meta-analysis are: how to select 'good' primary studies (including avoiding journals' publication bias towards 'success' and 'good stories'), how best to summarise findings (e.g., what metric is valid across all cases), whether meta-analysis can truly provide evidence for causality, and yet some others (Aguinis et al., 2011). Below are some of these methodological reflections adapted specifically to our endeavour in transitions research.

BIASES IN THE SELECTION AND CONTENTS OF SOURCES

Meta-analysis involves selecting a corpus of texts that *adequately* describe transitions. This is shared with any attempt at creating an event-catalogue. Biases in the selection of sources are therefore just as important as the procedures for extracting and analysing the 'events'. This has naturally been one of the earliest concerns of meta-analysis (Aguinis et al., 2011).

We concur with Lustick (1996) that any account has its own specific purposes (from confirming a theory to re-telling a story). Accounts of technological change that focus exclusively on quantifiable elements (as many economists tend to do (Fouquet, 2016) are by definition excluding certain explanatory events from our data set. By the same token, other accounts could be overloading it with contingent socio-cultural detail under the premise that some sociological construct *C* 'matters' (Healy, 2017). See, for instance, Geels (2011).

Accumulating 'events that made a difference' and that demonstrably took place during transitions should smooth out such *content bias* (e.g., too much discussion of raw material prices or of gender relations) in the aggregate. However, this potential bias may require us to be aware and record the theoretical background of inputs into the data set to avoid systematic bias (Lustick, 1996), such as the early MLP emphasis on heroic niche stories (Genus and Coles, 2008).

HOW TO ANALYSE ACROSS SOURCES

The flip side of allowing a wide enough input is the difficulty of comparability. We may ponder, firstly, are all socio-technical transitions from all sectors comparable? That is, how do we avoid *type bias*? Secondly, are *all* events recorded for *any* transition equivalent? How do we avoid *scale bias*? To pick one example, we can ponder whether events in the transition from hydropower to fossil-generated electricity in Mexico (Jano-Ito and Crawford-Brown, 2016) should count the same as that from alcohol to petroleum feedstocks in British chemicals (Bennett, 2012).

The first question may be easier to tackle as the number of sectors seems relatively manageable: chemicals, electricity generation, ground transport, sanitation, water supply, etc. Our database can incorporate a coding by sector, and each analysed independently or merged as needed. This can eventually help identify where and how transitions have common or different patterns across sectors in a much more reliable way. Martínez Arranz (2017) suggests that differences may be traced back to measurable differences in terms of incumbents. This leads to an answer to the second question, it should be possible to take recourse to relevant indicators, such as market share of new technologies and overall energy or emissions share of the sector concerned to ascertain what importance (weighting) to attach to each transition.

APPLYING LESSONS FROM THE PAST

An often overlooked aspect of meta-analysis with relevance for our event cataloguing efforts is that any 'lessons from the past' may not be applicable today. Firstly, key events may not be reproducible at will. Our database could contain entries for all three types of elements in Martínez Arranz's (2017) to differentiate between points of action for future policy-makers and the macro-trends that will condition them:

Indicators – reflect scientifically/systematically acquired information either about completely natural events (e.g., earthquakes), or human-related impacts (e.g., from the emissions of a particular factory to its oil supply figures). The transitions database may only be required to store rather local/not easily available data; for example, global CO_2 concentrations or oil production are widely and easily found for any given period.

Unintentional pressures – are fully purposeful activities not specifically targeting the socio-technical regimes under scrutiny (macro-political trends, e.g., liberalisation; macro-economic trends, e.g., information revolution, or trends in the perception of the natural world, e.g. environmentalism). These may affect regimes directly or through induced intentional pressures.

Intentional pressures – are deliberately initiated in order to induce change in or maintain the socio-technical regime under scrutiny. For example, the Arab oil embargo or the Soviet blockade of Berlin clearly targeted the energy regimes of the West. Similarly, as climate policy deliberately attempts to tackle climate change.

Secondly, the events that triggered transitions in the past may not have the same effects in these 'confusing times' (Castells, 2011). There seems to be a consensus that a new period started some time in the late 1980s. It has been variously called fifth Kondratiev wave (Freeman et al., 2001), 'second/late/liquid modernity' (Beck, 1986; Castells, 2000; Bauman, 2012). Regardless of its moniker and the exact start date, the characteristics of today's political economy have significant implications for technological transitions. We observe consolidated global supply chains, preoccupation with global ecological deterioration, a telecommunications revolution and an increasingly multipolar and ambiguous world order. An often asserted (and ignored) implication is that a Manhattan or Apollo Project for X or Y sustainable technologies cannot count on the concentration of manufacturing capacity, technological know-how, and single-mindedness of national will of the mid-20th century USA (Popp and Newell, 2012; Delina and Diesendorf, 2013; Mowery et al., 2010; Hargadon, 2010).

Considering that many of the events that may be included in our proposed catalogue will stem from periods where a completely different set of circumstances applied than in the present, the *interpretation of findings* should always be done with strict reference and contrast to the current context.

4.3 What could be done

4.3.1 Databases

Much in this chapter revolves around the notion of the event catalogue, which is in essence a database of events. What we mean by 'database' here more specifically is what is usually meant by it in the context of computing – the software implementation of a database or a database management system. Clearly, analysis of an event catalogue would be greatly empowered if it were accessible in the form of a database that can be queried programmatically. We preluded on this desirability in Section 3.1. The kind of data-driven approaches we are proposing benefit from databases in these (and likely other) ways:

Exportability / data-exchange – Once in electronic data format, events can be exported in various useful formats and, moreover, imported by other researchers to combine into larger data sets. Even a badly maintained database is better than a bunch of spreadsheets or publications with notes.

Scalability – Manually, scaling up an analysis from across a dozen to across a dozen million events is an insurmountable increase in work. If these data are in a database, the query will look basically the same. The increase in resolution that could be obtained by working with large volumes of data may well change the nature of the research altogether. Because of the previous characteristic of databases, building such large data sets is feasible.

Manipulability – This applies both to the data themselves and to the analytical tools one can apply to them. With regard to the former, this mostly refers to the ability to clean and organise data. Using script and similar tools for this greatly expedites these processes. With regard to the latter, a vast range of statistical and visualisation tools are available, often as free and open source software, that readily converse with databases. Moreover, writing one's own software to do so is fairly straightforward.

To further expand on the last point, numerous database management systems are available, in the free[13] and open source world the SQL-based offerings range from the simple, self-contained SQLite to the industrial strength PostgreSQL, amongst various others. So-called NoSQL options, which do not make use of relational database principles, abound also. Popular languages for data analysis, such as before R and Python (again free and open source) have built up lively communities of contributors developing and sharing massive amounts of software resources.

4.3.2 Ontologies

In computer science, the term 'ontology' has a precise meaning that is somewhat distinct from the philosophical usage. An ontology, in this sense, is a formal classification of the concepts and entities in some domain of discourse – in our case, transitions research. A database built on an ontology is referred to as a *knowledge base*. The difference between a 'normal' database on the one hand and an ontology or knowledge base on the other hand may not be intuitive to the uninitiated, but it is significant. The aid of an ontology would potentially allow a richer analysis than one just based on an event catalogue, but building an ontology for a large domain of discourse like transitions is far from trivial.

Ontologies provide, at the very least, a shareable vocabulary. Splendiani et al. (2014) remark that the use of ontologies has gained a 'critical dimension' in some areas of science (e.g., the life sciences) because of their 'convergence of disciplines', something which should also appeal to our field also. Ontologies are of course more than glorified dictionaries and they can be useful for transitions analysis in several ways: (1) as the frame of data- and knowledge bases (see e.g., Horrocks 2008 for a presentation of this), which then (2) enables improved meta-analyses; (3) as a basis for statistical and computer-based analyses; and (4) to support modelling exercises (e.g., van Dam and Lukszo (2006).

In other fields of research, ontologies and knowledge bases are successfully being used, notably bio-informatics and medicine, but also by the United Nations Development Programme, for example in their UNSPSC which codifies products and services (UNDP 2016; see e.g. Noy and McGuinness 2001). Closer to transitions studies, ontologies have been developed and used to represent and classify socio-technical systems, notably by van Dam in the context of the development of agent-based models (e.g., Keirstead and van Dam 2010; van Dam and Lukszo 2006; van der Sanden and van Dam 2010).

Referring back to our discussion of looking at transitions as changes in system states versus sequences of events – perhaps ontologies can play an important role in data-driven approaches based on the former, as they can provide a formal framework to capture what such a system state would encompass.

4.3.3 Extracting event data

One of the main tasks in cataloguing events is to extract event data from primary and/or secondary sources of data. This can simply be a matter of reading through the sources of data, and writing event descriptions whenever relevant events are found. This is how one of us has built event data sets in the past (see Section 4.1). This approach is easy enough to carry out, but it is also relatively sensitive to the analyst missing potentially important details, and to the analyst's particular understanding of what is important in the first place. Poole et al. (2000) suggest that the reliability of the process of recording event data can be improved by involving multiple researchers in the identification and description of events. However, this means that multiple researchers have to go through the same sources of data, which is an inefficient approach, and it may not even be feasible if the amounts of data to be processed are very large. In addition, if we wish to approach the construction of event data sets as a collective effort, it would be necessary to impose a clear set of standards on how events are to be identified and recorded, in order to ensure consistency in the way that event data are recorded.

It is therefore important to explore possibilities for the development of annotation schemes that can be used by researchers to explicitly identify events in primary and secondary sources of data, as well as relevant attributes of events (e.g., time, place, actors). Annotation schemes that capture most of the details of events that are important to us are currently already being developed in the field of text mining (e.g., Pustejovsky et al., 2011). Indeed, manually coding sources of data with such annotation schemes is still very time consuming. However, if many people are making contributions to shared data sets using the same annotation scheme, this becomes a feasible approach. A next step in this development would be to develop and apply tools that are able to assist in, or even fully automate the annotation of sources of data. Some work is being done in this area as well (Zhang et al., 2019), but this is still in an early stage, and there are plenty of difficulties to overcome, such as being able to distinguish between events that are relevant and events that are irrelevant, and being able to recognise events that are described across different sources as the same. In the short term, it is unlikely that we can rely on automated event extraction for the creation of event data sets. Nonetheless, we believe that these developments hold promise for the development of robust approaches to the creation of event data sets. It is worthwhile keeping an eye on these developments in our thinking about event extraction, to ensure the possibility for synergies 'down the road'.

5 Conclusion

We conclude this chapter briefly. We believe we have made a clear case for data-driven transitions research, in particularly in the form of event-based analysis. We have argued the desirability and possibility of data-driven transitions research and pointed out examples of other fields grappling with similar challenges that are successfully employing data-intensive methods. We discussed a range of methodological considerations about event-based analysis and provided examples of approaches we have worked with as well as tools and techniques we think may be fruitfully applied. As a final comment we pose the question: If not data-driven, then driven by what?

Notes

1 For example in the sense of Hedström's (2005) 'social mechanism', though the notion of mechanism is common in more recent philosophy of science, see, for example, Salmon (1990).

2 Nuance may appear a desideratum in a theory, but this is misguided. Nuance entails embracing exceptions, rather than precisely delineating the applicability of concepts and hypotheses. Healy (2017) calls this the 'nuance of the conceptual framework', describing it as the 'ever more extensive expansion of some theoretical system in a way that effectively closes it off from rebuttal or disconfirmation by anything in the world' and 'an evasion of the demand that a theory be refutable'.

3 In the Novum Organum, Bacon (1620) suggests a straightforward tabulation of empirical observations and to infer from their presence, absence and correlation in relation to a hypothesis of interest whether it should be accepted. Moreover, '[i]n forming our axioms from induction (…) we must observe, whether it confirm its own extent and generality by giving surety, as it were, in pointing out new particulars'. In other words, that our concepts and theories should not only accurately describe the phenomena they were inspired by, but also predict or explain other phenomena.

4 See also our observations about ontologies in Section 4.3.2.

5 This should preferably employ a Creative Commons share-alike, copyleft type licencing scheme (e.g., CC BY-SA https://creativecommons.org/licenses/by-sa/4.0/). Such a scheme would empower researchers to share their data for two reasons: (1) they can rest assured that their work would have to be acknowledged when used by others as a matter of law, and (2) any work based on it would have to be made openly available under the very same conditions. This would be in direct emulation of large collaborative, free and open source software projects such as the Linux kernel. Linus Torvalds, the project's initiator credits the role of the copyleft licencing (here the GPLv2) as follows at LinuxCon 2016: 'I really think the license has been one of the defining factors in the success of Linux because it enforced that you have to give back, which meant that the fragmentation has never been something that has been viable from a technical standpoint' https://www.zdnet.com/article/linus-torvaldss-love-hate-relationship-with-the-gpl/.

6 Unless perhaps as part of some philosophical debate. What we mean is a *common sense* understanding of what 'actually happened'.

7 Here again, we are not inviting a debate about the (im)possibility of verification in principle.

8 Where does a lunar eclipse happen? (Think about it.) Or more pertinent to this methodology, where did the Cuban Missile Crisis happen? In Washington? In Moscow? In the Caribbean sea, or in all of them somehow? Where does a cyber attack happen? Is the internet a place?

9 We should call this Mulisch' Rule of Guilt: *whatever is done, is done by those who did it.* In *De Aanslag* (1982), Mulisch describes how his protagonist gets stuck tracing who is ultimately responsible for the summary execution of his parents. Their execution was a Nazi retaliation in response to the assassination – with which they had nothing to do — of a local collaborating policeman. Was the resistance to blame for their execution because they assassinated the policeman? Were their neighbours to blame, who relocated the corpse to their doorstep? After three decades, the protagonist's back-then neighbour resolved the issue for him. His parents were killed by their killers, they were responsible for their deaths.

10 The methodology also makes use of other graph-theoretical concepts and visualisations to visualise networks of relationships that are enacted in social processes, and to visualise co-occurrences of, for example, attributes of events. For the sake of brevity, these are not discussed in detail here.

11 One of the authors has developed a dedicated software suite for this that will be released in the near future.

12 Of course, as with other forms of analysis that rely on qualitative coding, we may ask multiple coders to code the same data set, check for inter-coder reliability, and via this route attempt to achieve some degree of inter-subjectivity

13 'Free' as in 'freedom', not as in 'free beer' is the catchphrase often used by Richard M Stallman (RMS) and the Free Software Foundation to clarify that free software is not about price. See also https://www.gnu.org/philosophy/free-sw.html.en and https://en.wikipedia.org/wiki/Free_software_movement.

References

Abbott, Andrew (1988). Transcending General Linear Reality. *Sociological Theory 6* (2), 169–186.

Abbott, Andrew (2001). *Time Matters: On Theory and Method*. Chicago: University of Chicago Press.

Abell, Peter (1987). *The Syntax of Social Life: The Theory and Method of Comparative Narratives*. Oxford, Angleterre: Clarendon.

Abell, Peter (1993). Some aspects of narrative method. *The Journal of Mathematical Sociology 18* (2-3), 93–134.

Abell, Peter (2004, January). Narrative Explanation: An Alternative to Variable-Centered Explanation? *Annual Review of Sociology 30*, 287–310.

Abell, Peter (2007). Narratives, Bayesian Narratives and Narrative Actions. *Sociologica. Italian Journal of Sociology Online* (3), 1–21.

Aguinis, Herman, Charles A. Pierce, Frank A. Bosco, Dan R. Dalton, and Catherine M. Dalton (2011). Debunking Myths and Urban Legends About Meta-Analysis. *Organizational Research Methods 14* (2), 306–331.

Bacon, Francis (1620). *Novum Organum: Or True Suggestions for the Interpretation of Nature* (The Project Gutenberg EBook of Novum Organum, from the P.F. Collier & Son, New York, 1902 Edition. Release Date: June 15, 2014 [EBook #45988] ed.). Project Gutenberg.

Bauman, Zygmunt (2012). *Liquid Modernity*. Cambridge, England, UK and Malden, MA, USA: Polity Press.

Beck, Ulrich (1986). *Risikogesellschaft: Auf Dem Weg in Eine Andere Moderne*. Frankfurt am Main: Suhrkamp.

Bennett, S. J. (2012). Using past transitions to inform scenarios for the future of renewable raw materials in the UK. *Energy Policy 50*, 95–108.

Boons, Frank, Wouter Spekkink, and Wenting Jiao (2014). A Process Perspective on Industrial Symbiosis: Theory, Methodology, and Application. *Journal of Industrial Ecology 18* (3), 341–355.

Borenstein, Michael, Larry V Hedges, Julian PT Higgins, and Hannah R Rothstein (2011). *Introduction to Meta-Analysis*. Chichester, England, UK: John Wiley & Sons.

Castells, Manuel (2000). Materials for an exploratory theory of the network society. *The British Journal of Sociology 51* (1), 5–24.

Castells, Manuel (2011). *The Rise of the Network Society: The Information Age: Economy, Society, and Culture*, Volume 1. Chichester, England, UK: John Wiley & Sons.

Danto, Arthur C (1965). Basic Actions. *American Philosophical Quarterly 2* (2), 141–148.

Delina, L. L. and M. Diesendorf (2013). Is wartime mobilisation a suitable policy model for rapid national climate mitigation? *Energy Policy 58*, 371–380.

Diamond, Jared (1997). *Guns, Germs and Steel: A Short History of Everybody for the Last 13,000 Years* (2005 ed.). London: Vintage Books.

Fouquet, Roger (2016). Historical energy transitions: Speed, prices and system transformation. *Energy Research & Social Science 22*, 7–12.

Freeman, Christopher, Francisco Louçã, and Francisco Louçã (2001). *As Time Goes by: From the Industrial Revolutions to the Information Revolution*. Oxford University Press.

Geels, Frank W. (2011, June). The multi-level perspective on sustainability transitions: Responses to seven criticisms. *Environmental Innovation and Societal Transitions 1* (1), 24–40.

Genus, Audley and Anne-Marie Coles (2008). Rethinking the multi-level perspective of technological transitions. *Research Policy 37* (9), 1436–1445.

Griffin, Larry J. (1993). Narrative, Event-Structure Analysis, and Causal Interpretation in Historical Sociology. *American Journal of Sociology 98* (5), 1094–1133. bibtex: Griffin1993.

Hargadon, Andrew (2010). Technology policy and global warming: Why new innovation models are needed. *Research Policy 39* (8), 1024–1026.

Healy, Kieran (2017, June). Fuck Nuance. *Sociological Theory 35* (2), 118– 127.

Hedström, Peter (2005). *Dissecting the Social*. Cambridge, New York, Melbourne: Cambridge University Press.

Heise, David R. (1989). Modeling event structures. *The Journal of Mathematical Sociology 14* (2-3), 139–169.

Horrocks, Ian (2008). Ontologies and databases. *Presentation at the Semantic Days 4*.

Jano-Ito, M. A. and D. Crawford-Brown (2016). Socio-technical analysis of the electricity sector of Mexico: Its historical evolution and implications for a transition towards low-carbon development. *Renewable and Sustainable Energy Reviews 55*, 567–590.

Keirstead, James and Koen Haziël van Dam (2010, May). A comparison of two ontologies for agent-based modelling of energy systems. In *Proceedings of the First International Workshop on Agent Technologies for Energy Systems (ATES 2010)*, https://repository.tudelft.nl/islandora/object/uuid:6685b86e-6586-4680-9102-58fbb1efae62, pp. 21–28.

Langley, Ann (1999). Strategies for Theorizing from Process Data. *The Academy of Management Review 24* (4), 691–710.

Lusher, Dean, Johan Koskinen, and Gary Robins (Eds.) (2013). *Exponential Random Graph Models for Social Networks*. Cambridge: Cambridge University Press.

Lustick, Ian S. (1996). History, Historiography, and Political Science: Multiple Historical Records and the Problem of Selection Bias. *The American Political Science Review 90* (3), 605–618.

Martínez Arranz, Alfonso (2017, May). Lessons from the past for sustainability transitions? A meta-analysis of socio-technical studies. *Global Environmental Change 44*, 125–143.

Mowery, David C., Richard R. Nelson, and BR Martin (2010). Technology policy and global warming: Why new policy models are needed (or why putting new wine in old bottles won't work). Report, Nesta.

Mu, Rui and Wouter Spekkink (2018). A running start or a clean slate? How a history of cooperation affects the ability of cities to cooperate on environmental governance. *Sustainability (Switzerland) 10* (6).

Mulisch, Harry (1982). *De Aanslag*. Amsterdam: De Bezige Bij.

Noy, Natalya F and Deborah L McGuinness (2001). Ontology development 101: A guide to creating your first ontology. Technical report, http://protege.stanford.edu/publications/ontology_development/ontology101.pdf.

Poole, Marshall Scott, Andrew H. Van De Ven, Kevin J. Dooley, and Michael E. Holmes (2000). *Organizational Change and Innovation Processes: Theory and Methods for Research*. Oxford: Oxford University Press.

Popp, David and Richard Newell (2012). Where does energy R&D come from? Examining crowding out from energy R&D. *Energy Economics 34* (4), 980–991.

Pustejovsky, James, Robert Gaizauskas, and Beth Sundheim (2011). The TIMEBANK Corpus. In *Time and Event Recognition for Question Answering Systems Workshop*, Volume 30, pp. 647–656.

Raven, Rob, Florian Kern, Bram Verhees, and Adrian Smith (2016, March). Niche construction and empowerment through socio-political work. A meta-analysis of six low-carbon technology cases. *Environmental Innovation and Societal Transitions 18*, 164–180.

Salmon, Wesley C (1990). *Four Decades of Scientific Explanation*. Pittsburgh, PA, USA: University of Pittsburgh Press.

Schatzki, Theodore R. (2016). Keeping Track of Large Phenomena. *Geographische Zeitschrift 104*, 4–24.

Spekkink, Wouter (2015, July). Building capacity for sustainable regional industrial systems: An event sequence analysis of developments in the Sloe Area and Canal Zone. *Special Volume: Support your future today! Turn environmental challenges into opportunities. 98*, 133–144.

Spekkink, Wouter (2016). *Industrial Symbiosis as a Social Process: Developing Theory and Methods for the Longitudinal Investigation of Social Dynamics in the Emergence and Development of Industrial Symbiosis*. Ph. D. thesis, Erasmus University Rotterdam.

Spekkink, W. A. H. and F. A. A. Boons (2016, November). The Emergence of Collaborations. *Journal of Public Administration Research and Theory 26* (4), 613–630.

Splendiani, Andrea, Michele Donato, and Sorin Drăghici (2014). Ontologies for Bioinformatics. In N. Kasabov (Ed.), *Springer Handbook of Bio-/Neuroinformatics*, pp. 441–461. Berlin, Heidelberg: Springer Berlin Heidelberg.

Stegenga, Jacob (2011). Is meta-analysis the platinum standard of evidence? *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 42* (4), 497–507.

UNDP (2016, December). UNSPSC. http://www.unspsc.org/. GS1 US™ for the UN Development Programme (UNDP).

van Dam, Koen Haziël and Zofia Lukszo (2006, October). Modelling Energy and Transport Infrastructures as a Multi-Agent System using a Generic Ontology. In *Proceedings of the 2006 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, The Grand Hotel, Taipei, Taiwan, pp. 890–895.

Van de Ven, Andrew H. (1992). Suggestions for Studying Strategy Process: A Research Note. *Strategic Management Journal 13* (Special Issue: Strategy Process: Managing Corporate Self-Renewal), 169–191.

Van de Ven, Andrew H. and George P. Huber (1990, January). Longitudinal Field Research Methods for Studying Processes of Organizational Change. *Organization Science 1* (3), 213–219.

Van de Ven, Andrew H., Douglas E. Polley, Raghu Garud, and Sankaran Venkataraman (2008). *The Innovation Journey*. Oxford: Oxford University Press.

Van de Ven, Andrew H. and Marshall Scott Poole (1990, August). Methods for Studying Innovation Development in the Minnesota Innovation Research Program. *Organization Science 1* (3), 313–335. bibtex: Vande-VenPoole1990.

van der Sanden, Maarten C. A. and Koen Haziël van Dam (2010, November). Towards an ontology of consumer acceptance in socio-technical energy systems. In *Proceedings of the 3rd International Conference on Infrastructure Systems and Services: Next Generation Infrastructure Systems for Eco-Cities (INFRA)*, Shenzhen, China.

Wikipedia (2018, February). Clintonville, Columbus, Ohio. *Wikipedia, The Free Encyclopedia*. Page Version ID: 823516451.

Wiseman, John, Taegen Edwards, and Kate Luckins (2013). Post carbon pathways: A meta-analysis of 18 large-scale post carbon economy transition strategies. *Environmental Innovation and Societal Transitions 8*, 76–93.

Zhang, Hao, Frank Boons, and Riza Batista-navarro (2018). Whose Story Is It Anyway? Automatic Extraction of Accounts from News Articles. *Journal of Information Processing and Management 56* (5), 1837–1848.