Scheller, Fabian; Johanning, Simon; Bruckner, Thomas

Research Report
A review of designing empirically grounded agent-based models of innovation diffusion: Development process, conceptual foundation and research agenda

Beiträge des Instituts für Infrastruktur und Ressourcenmanagement, No. 01/2019

Provided in Cooperation with:
Institute for Infrastructure and Resources Management, University of Leipzig

Suggested Citation: Scheller, Fabian; Johanning, Simon; Bruckner, Thomas (2019) : A review of designing empirically grounded agent-based models of innovation diffusion: Development process, conceptual foundation and research agenda, Beiträge des Instituts für Infrastruktur und Ressourcenmanagement, No. 01/2019, Universität Leipzig, Institut für Infrastruktur und Ressourcenmanagement (IIRM), Leipzig

This Version is available at:
http://hdl.handle.net/10419/191981

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
A review of designing empirically grounded agent-based models of innovation diffusion: Development process, conceptual foundation and research agenda

Fabian Scheller, Simon Johanning, Thomas Bruckner

01/2019
Fabian Scheller, Simon Johanning, Thomas Bruckner

A review of designing empirically grounded agent-based models of innovation diffusion: Development process, conceptual foundation and research agenda
A review of designing empirically grounded agent-based models of innovation diffusion: Development process, conceptual foundation and research agenda

Fabian Schellera, Simon Johanninga, Thomas Brucknera

aInstitute for Infrastructure and Resources Management (IIRM), University Leipzig

Abstract
Modeling the diffusion of innovations is a very challenging task, as there are various influencing factors to consider. At the same time, insights into the diffusion process can help decision makers to detect weak points of potential business models. In the literature, various models and methodologies that might tackle this problem are presented. Between these, empirically grounded agent-based modeling turned out to be one of the most promising approaches. However, the current culture is dominated by papers that fail to document critical methodological details. Thus, existing agent-based models for real-world analysis differ extensively in their design and grounding and therefore also in their predictions and conclusions. Additionally, the selection of modeling aspects seems too often be ad hoc without any defendable rationale. Concerning this matter, to draw on experiences could guide the researcher. This research paper seeks to synthesize relevant publications at the interface of empirical grounding, agent-based modeling and innovation diffusion to provide an overview of the existing body of knowledge. The major aim is to assess existing approaches regarding development procedure, entity and dynamics consideration and theoretical grounding to suggest a future research agenda. This might lead to the development of more robust models. According to the findings of this review, future work needs to focus on generic design, model coupling, research consistency, modular testing, actor involvement, behavior modeling, network foundation, and data transparency. In a subsequent step and based on the findings, a novel model approach needs to be designed and implemented.

Keywords: Innovation diffusion models, Agent-based models, Empirically grounded models, Data driven models, Literature review

1. Introductory remarks

1.1. Problem statement

Product or service innovation constitutes an effective means for organizations to create and maintain a competitive advantage. In this sense, it is important to understand how market actors engage with and adopt innovations, since, even good innovations may fail or diffuse at a slow rate [1]. For many companies, it is hard to predict how innovations will diffuse in the dynamic environment, resulting in uncertainty about whether an innovation is fit to become a sustainable business model.

This may be to a large part because of the adoption of these innovations by intended target groups are not always assured, and as [2] shows, it does not just depend on the qualities of the innovation. Instead, it takes place within a complex social system, in which the diffusion of the respective innovations depend on many factors and mechanisms [3]. Business models and innovations need to encompass the dynamics of the market setting by including the mental structures, such as personal characteristics, behavioral attitudes as well as conscious and subconscious purchase decisions, of market participants in general and of customers in particular. As [4] points out, “[...] the diffusion of innovation paradigm postulates that markets are in fact dominated by social influences [...]”

Thus, decision makers responsible for these innovations are confronted with making informed decisions about complex matters [5]. Insights into the diffusion of innovations can help to detect weak points of potential business
models and innovation marketing. Particularly quantitative models of innovation diffusion analysis that account for the complexity of the modeled system might assist decision makers in the investigation of potential measures and in the development of effective strategies. One promising approach for this is to employ empirically grounded agent-based models [5]. In this context, innovation diffusion can be seen as the analysis of the spread of an innovation [11]. An agent-based model is a model where entities are modeled individually, as autonomous, social, reactive and proactive agents [6]. Autonomous decision strategies in accordance with their personal objectives describe the procedure of taking an action depending on several conditions. To model the strategies of the heterogeneous agents realistically, it is necessary to collect and analyze an extensive amount of empirical data to derive a theory for grounding [7]. In this context, an agent needs to be theoretically and empirically grounded [8].

In the last years, empirically grounded agent-based models have lost its niche character and gained importance as a valuable methodology for describing diffusion processes [5]. Thereby, they are specially applied to reflect real market issues. Thus, papers with real-world case studies to support decision makers are increasing [9]. "In the spirit of modern complexity science, these models have the potential to reproduce and explain complex non-linear diffusion patterns observed in the real world as the result of relatively simple local micro-level interactions.” [4].

Methodology regarding model design and model grounding is a crucial element while developing empirically grounded agent-based models to investigate diffusion processes. Reasons are scientific comparability and robust results. However, existing models differ strongly in their design and their grounding and therefore also in their predictive power [10], resulting in a lack of ” [...] a clear foundation of agreed-upon approaches and libraries that offer a baseline for problem solutions that characterize other modeling fields.” [11]. The selection of modeling aspects is often ad hoc [12] without any defendable rationale [9]. Besides, the current culture is also dominated by papers that fail to document critical methodological details [9]. Concerning this matter, to draw on experiences could guide the researcher in the context of designing, modeling, and application.

1.2. Research objective

This research paper seeks to synthesize relevant publications on the interface of empirically grounded agent-based innovation diffusion models to provide an overview of the existing body of literature, with a focus on specific models and practical applications. The aim is to shape the understanding of robust modeling, to bring together techniques, to tackle challenges as well as to propose future research of empirically grounded agent-based modeling of innovation diffusion. In this context, the present paper assesses the following research questions:

- Which procedure is useful in terms of developing empirically grounded agent-based models of innovation diffusion?
- Which model entities and model dynamics need to be addressed while investigating innovation diffusion processes based on empirically grounded simulations?
- Which methods assist in modeling formalization theories as well as in collecting parametrizations data to empirically ground agent-based models?
- Which challenges occur during modeling empirically grounded agent-based models of innovation diffusion and what solutions are resolved?

Answering these research questions contributes to the theory in several ways. First, the review of existing papers demonstrates the status quo and constitutes a basis for further targeted research. A comprehensive collection of papers with practical application of such models does only partially exist. In this sense, this research extends the review of [5] in particular. Guiding principles of more general reviews [13] [14] [15] [16] [17] [18] [19] [20] are also continued with respect to innovation diffusion processes. Second, a mutual understanding of designing empirically agent-based simulations for analyzing diffusion processes of technologies improves theoretical comparability of research. By doing this, the paper on hand also complies with the raised issues. The findings are intended to serve as the foundation for the researcher to systematically develop models. In addition, the paper provides an overview of usage and configuration possibilities regarding the practical application. This can guide professionals to reconsider traditional decision-making approaches.
1.3. Research structure

The paper is organized as follows: Section 2 comprises the theoretical foundations of the proposed research domains and introduces the concepts of innovation diffusion modeling and agent-based modeling. Section 3 describes the underlying investigative methodology of this review. Section 4 presents findings of the literature review regarding the procedure, the application domain as well as the theoretical and empirical grounding of existing empirically grounded agent simulations with respect to innovation diffusion. Section 5 discusses these findings and suggests future research topics. Section 6 concludes with a recapitulation of the main findings, followed by an argument of limitations as well as subsequent research.

2. Conceptual background

For a systematic review of existing research papers, relevant domains need to be conceptualized first. An overview of the broached research domains (RD) and the relevant intersection is given in Figure 1. For this, section 2.1 introduces the concept of innovation diffusion. Section 2.2 covers the concept of agent-based models, while section 2.3 explains the theory of empirically grounded models.

![Figure 1: Overview of research domains](image)

2.1. Innovation diffusion modeling

Although the roots of innovation diffusion research lie in the late 19th century, studying the diffusion of innovations can generally be traced back to the seminal study of Ryan and Gross in the 1940s in rural sociology about the diffusion of hybrid corn. This might be due to the fact that the study advanced theoretical exploration of the diffusion
process.” [22], or because of the study ”[...] was driven by scholarly interest in the relative influence of economic versus social factors in the adoption of a technological innovation.” [22].

A fundamental aspect of innovation diffusion Ryan and Gross identified was the interpersonal communication between farmers. ”The hybrid corn study established diffusion as essentially a social process. A farmer typically adopted the innovation because of interpersonal communication with other farmers who already had adopted it [...]” [22]. Through this social process ”[...] subjective evaluations of an innovation spread from earlier to later adopters rather than one of rational, economic decision making.” [22].

Rogers defines the diffusion of an innovation as ”[...] the process in which an innovation is communicated through certain channels over time among the members of a social system.” [1]. This definition exemplifies the four major elements, namely innovation, communication channels, time, and the social system. Each one of these elements is identifiable in nearly every diffusion research or diffusion campaign [1]. In other words, diffusion can be seen as a ”special kind of communication in which the messages are about a new idea. [...] Diffusion is a kind of social change, defined as the process by which alteration occurs in the structure and function of a social system.” [1]. Thereby, ”[a]n innovation is an idea, practice or object that is perceived as new by an individual or another unit of adoption.” [1]. The perception of newness matters, but not the absolute newness as described by [23]. ”Adoption is a decision to make full use of an innovation as the best course of action available. Rejection is a decision not to adopt an innovation.” [1]. The units of adoptions could be individuals, households, institutions or other entities. A summary of operational definitions of presented key concepts are outlined in Table 1.

<table>
<thead>
<tr>
<th>Concept</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product innovation</td>
<td>“Innovation is an idea, practice, or object perceived as new by an individual or other unit</td>
</tr>
<tr>
<td></td>
<td>of adoption. It can also be an impulse to do something new or bring some social change.”</td>
</tr>
<tr>
<td></td>
<td>[1]. In this work, a perception of newness matters, but not the absolute newness [23].</td>
</tr>
<tr>
<td>Innovation adoption</td>
<td>“Adoption is a decision to make full use of an innovation as the best course of action available. Rejection is a decision not to adopt an innovation.” [1]</td>
</tr>
<tr>
<td>Innovation diffusion</td>
<td>Innovation diffusion is “the process by which an innovation is communicated through certain channels over time among the members of a social system.” [1]</td>
</tr>
</tbody>
</table>

Starting from the 1960s, innovation diffusion processes have been investigated using models, which aim at empirical generalizations of prototypical diffusion patterns at aggregate levels [4]. As Ryan and Gross stress, social contacts, social interaction and interpersonal communication are important influences on the adoption of new behaviors. Kiesling emphasizes that innovations are not evaluated objectively, but instead, the dynamic formation of attitudes and subjective perceptions are transmitted through communication at disaggregated levels [4].

A large amount of aggregated innovation diffusion models are refined versions of the Bass model [24], a parsimonious, aggregated innovation diffusion model, based on models of epidemiological spread. The Bass model as described in [24] is based upon the assumption that the timing of initial purchases is related linearly to the number of previous buyers. The goal of the model is to develop a theory of timing of initial purchases for new classes of products.

Despite their popularity, these aggregate models have several shortcomings [3][4]. The most fundamental shortcomings of aggregated innovation diffusion models are their assumption of a homogeneous population. Furthermore, aggregate models cannot differentiate between the social network of one potential adopter and the other, so they have to impose the assumption of a fully connected social network. Additionally, these models require information about events they ought to predict, and lack predictive power.

To overcome these limitations, many approaches employ disaggregate models, most notably agent-based models. Disaggregated models are models that avoid aggregating model entities individually. They focus on micro behavior instead of macro behavior and are grounded in complexity science. In contrast to macro simulations, where the entire system is described directly and ‘phenomenologically’, societal phenomena of interest are modeled bottom-up based on the underlying processes. The phenomena then emerges from the behavior and micro-level interactions of the agents.
2.2. Agent-based modeling

As the name suggests, agent-based models are conceptualized from the perspective of disaggregated units, so-called agents or actors\(^1\) instead of modeling the system on the aggregate level. As noted in [6], no single universally accepted definition of an agent exists. Instead Wooldridge [6] enumerates abilities actor entities need to exhibit in order to be called agents. In their general definition, agents need to exhibit four abilities: autonomy, social ability, reactivity, and pro-activeness. In this, autonomy is the ability to act without being directly controlled or manipulated by humans or others, as well as having some control over their actions and internal state. Social ability means the use of an agent-communicative language to interact with other agents, where reactivity represents the perception of and response to their environment. Finally, proactiveness is the ability to take initiative in goal-directed behavior instead of solely responding to stimuli [6]. Most crucially, agent-based models allow for modeling heterogeneity of potential adopters. Since this approach describes actors on the level of their entity, actors can be designed differently from one another. The characterization of agents not only considers the interaction between agents but also puts interaction with an immersion in an environment at the heart of the models [11]. A summary of operational definitions of presented key concepts is outlined in Table 2.

Table 2: Key concepts of agent-based modeling

<table>
<thead>
<tr>
<th>Concept</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software agent</td>
<td>“An agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives.” [6].</td>
</tr>
<tr>
<td>Multi-agent model</td>
<td>“A multi-agent system is one that consists of a number of agents, which interact with one another […] [T]he agents in a multi agent system will be representing or acting on behalf of users or owners with very different goals and motivations.” [25].</td>
</tr>
<tr>
<td>Model procedure</td>
<td>Model procedures pose step-wise guidelines for the designing and modeling of complex systems in terms of scientific purposes. This covers activity lists, building blocks, structural items, best practices, design choices, methodological issues as well as functional protocols and frameworks.</td>
</tr>
<tr>
<td>Model component</td>
<td>Model components represent the functional elements of complex systems. Master categories are model strategies, model entities and model dynamics [4].</td>
</tr>
</tbody>
</table>

2.3. Empirically grounded modeling

The level of detail with respect to the data and information incorporated into agent-based simulations vary from "Picasso" to "Photograph" models [26]. "An obvious difference relates to the representation of space, ranging from empty and simple artificial landscapes [...] to very detailed, realistic representations of the environment.” [26]. At first, more abstract models (so-called "Picasso" models) were widely used to show general mechanisms rather than to make exact predictions.

In this context, different aspects as actor heterogeneity can enter the agent-based innovation diffusion models through different values of characteristics such as income or preference [26], various sources of knowledge [27], different types of agents that differ in decision rules and interaction with other agents and the environment [26]. This is due to the fact that the behavior of the modeled agents ”[…] can be empirically informed using a combination of different kinds of data (e.g. qualitative and quantitative) and data collection methods […] that support multiple

---

\(^1\) Many publications from the sociological, ecological or socio-economic perspective use the term agent to refer to these units, publications in computer science often try to avoid this term, since it might be confused with the concept of software agents, and rather use the term actor. In this publication, they will be used interchangeably, and where the distinction between actors and software agents is meaningful, this will be made explicit.
approaches to represent actor decision making in an agent (e.g., heuristic decision trees, utility functions).” [11]. Through this they ”[…] can go beyond the typical representation of a population or average individual in EBMs [equation based models] and capture the heterogeneity of individual actors, their characteristics, and decision-making structures.” [11]. Thus, empirically grounded agent-based models of innovation diffusion can reproduce and explain complex non-linear diffusion patterns observed in the real world as a result of simple local micro-level interactions [4].

A summary of operational definitions of presented key concepts is outlined in Table 3.

Table 3: Key concepts of empirically grounded modeling

<table>
<thead>
<tr>
<th>Concept</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grounded theory</td>
<td>Grounded theory is defined as “discovery of theory from data systematically obtained from social research.” [7]. The derived constructs constitute the grounding of the models. In this work, it refers to theoretical grounding as well as empirical grounding.</td>
</tr>
<tr>
<td>Theoretical grounding</td>
<td>Theoretical grounding describes the characterization of the model. It “aims at surfacing the intended model as an artifact: qualifying its contours and interfaces.” [8].</td>
</tr>
<tr>
<td>Empirical grounding</td>
<td>Empirical grounding describes the parametrization of the model. It “aims at connecting model and target system, through giving values to the set of parameters in order to enable simulation.” [8].</td>
</tr>
<tr>
<td>Micro-level approach</td>
<td>The micro-level approach describes a “bottom-up” or “microscopic” modeling [13]. “Rather than describing the whole system directly and phenomenologically, macro-scale dynamics in [system models] are emergent phenomena that arise from micro-level interactions between agents when the model is executed.” [5].</td>
</tr>
<tr>
<td>Case-based applications</td>
<td>Case-based applications “have an empirical space-time circumscribed target domain. […] The goal […] is to find a micro-macro generative mechanism that can allow the specificity of the case […]” [28]. They are usually built “to provide forecasts, decision support, and policy analysis […]” [5].</td>
</tr>
</tbody>
</table>

3. Research methodology

A comprehensive literature review demonstrates a solid foundation about a particular object of study and identifies crucial methodological insights as well as recommending valuable further research aspects [29]. While keeping in mind the research questions and the focal domains, this research follows a procedure of [30]. According to this, the five phases mark the process of systematically reviewing literature:

- Definition of the review scope,
- Conceptualization of the topic and associated research domains,
- Carrying out the literature search and documenting the search process,
- Evaluating and synthesizing the retrieved literature,
- Presentation of findings and derivation of the research agenda.

According to the proposition of [31] and the adoption of [29] or [29], a review might be characterized based on six characteristics each having a different number of categories. While the characterization does not provide immediate answers to the research questions, its application helps to be aware of the focus [30].

This review concentrates on research methods of empirically grounded agents of innovation diffusion. Additionally, the research also deals with practical applications to detect differences in methodology between areas of application.
The review puts its main emphasis on critically analyzing existing papers to identify guiding principles and to justify future research. To guarantee a neutral representation no specific perspective is taken.

Furthermore, this review covers a representative sample of articles. It is limited to samples of articles representing also other articles, instead of explicitly considering the entire corpus of existing papers [30]. The only papers of interest are journaled articles since they typically have been peer-refereed and represent completed research. In the process, journals which seemed to be the most likely to publish related work with respect to the research domains have been selected. In addition, the relevance of the retrieved studies has been determined based on the review of abstracts or introductory content.

The research findings are structured along the stated research issues. The audience addressed by this review is comprised of specialized scholars interested in the design and modeling of empirically grounded agent systems. The underlined categories in Table 4 summarize the presented selection.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus</td>
<td>research outcomes</td>
</tr>
<tr>
<td>Goal</td>
<td>integration</td>
</tr>
<tr>
<td>Perspective</td>
<td>neutral representation</td>
</tr>
<tr>
<td>Coverage</td>
<td>exhaustive</td>
</tr>
<tr>
<td>Organization</td>
<td>historical</td>
</tr>
<tr>
<td>Audience</td>
<td>specialized scholars</td>
</tr>
</tbody>
</table>

The review is guided by the research questions as stated in Section 1. By this, relevant research papers need to be situated at the intersection of empirical grounding, agent-based modeling, and innovation diffusion modeling. The conceptualization of these key concepts is outlined in Section 2 as requested by [21]. An overview of relevant operational definitions is given in Table 3, Table 2, and Table 1. Additionally, attention is given to research papers with highly specific models. In other words, papers should be based on real-world applications and thus provide forecasts and policy analysis as a managerial basis for decision-making. Duplicates are avoided by concentrating on journal papers. For systematic structuring and for subsequent derivation of a sound modeling procedure, various methodological papers have also been consulted. These make it easier to organize and to expand on the findings. Most of the papers have been identified during the concept-based search.

A general concept-based search helped to specify domains as well as to select journals. Most prominently, google and google scholars were used to discover papers initially. Keywords used as selection criteria within the documents’ titles, abstracts, and keyword list were a combination of “agent-based modeling”, “innovation diffusion”, “empirically grounded agents” or “multi-agent systems” (in the preparation phase the following possible search terms were identified according to the research domains: empirical agent-based model, decision theory agent-based model, social agent-based model, decision algorithms development for an agent-based model, computational sociology and agent-based modeling, modeling human decision making as agent-based simulation, socio-ecological agent-based models, empirically grounded agent-based models). To identify relevant papers, keyword, backward and forward search was applied. The selected time span of the investigation is 2002-2017.

To provide insights, the retrieved papers were scanned, evaluated and synthesized. In an iterative bottom-up process, units of text were extracted, discussed and classified. With the help of a summary database, various constructs were identified. Major structuring concepts were model procedure, model application, model components as well as theoretical and empirical grounding was supplemented with subclasses. However, since the documentation of methodological details among the papers ranged from highly ambiguous to very specific, the individual paper could be only hardly classified. The main concepts represent key elements of the business specification of an agent-based diffusion model and were derived from [5] and [8]. Technical specifications like programming languages and development frameworks are not considered in this review. For more information, see [32] or [33]. Finally, the raised
issues of different authors of the papers as well as the discussion of the findings result in a research agenda, comprised of sharper and more insightful questions for future research.

4. Research findings

Empirically grounded agent-based models attract more and more the interest of both researchers and managers. At the same time, there seems to be a lack of documenting critical methodological details [9]. Following the objectives of this paper, the findings of the literature analysis are provided in this section. The modeling procedure is drafted in section 4.1, the application domain in section 4.2, the components modeling in section 4.3, the theoretical grounding in section 4.4 and the empirical grounding in section 4.5. While the latter sections are only based on retrieved papers using case-based applications, the first section is supplemented by papers describing developmental instructions.

4.1. Modeling procedure

Developing an empirically grounded agent-based diffusion model is a complex procedure. “Besides technical obstacles, there are methodical and fundamental problems such as a lack of confidence in the results, missing methodology for the development of agent-based simulations that would be easy and comprehensible, missing leading development frameworks for agent-based simulations, computational performance limitations and a lack of information about the method along the public.” [34]. Drawing on applied procedures seems promising for comparison and replication.

Even though most of the retrieved research papers do not explicitly describe the wide range of necessary steps is recognizable. According to this analysis, the ODD (Overview, Design, and Details) protocol and the UML (Unified Modeling Language) description are the only proven frameworks. Both help internals and externals to consider, to understand and to discuss model constructs. While the ODD protocol has been applied by [35, 36, 37, 38, 39, 40, 26], the UML description is considered by [41, 42, 37, 39].

The ODD protocol [14, 43] standardizes the model formulation and promotes the theoretical description. It aims to improve consistency and understanding. The harmonious use of the protocol is supported by the provision of questions and checklists as well as explanations, and notes. Overview as the first category comprises three guiding aspects: purpose; entities, state variables, and scales; process overview and scheduling. Design as the second category deals with design concept aspects in general and various sub-aspects in particular: basic principles, emergence, adaption, objectives, learning, prediction, sensing, interaction, stochasticity, collectives, observation, and explanation. Details as the third category conclude the protocol. Major aspects are initialization, input data, and sub-models.

The UML is a standard for specifying object-oriented software systems. The major aim is to visualize the design of a system. The different types of diagrams assist the modeler. Assistance regarding designing and modeling of agent systems with the UML standard is given in [44] and [45]. In the same way, other known business process modeling languages such as BPMN (Business Process Model and Notation) are applicable [46, 47].

Research papers with an overall procedural description like [48] present the stages of an ideal-typical research study centered on agent-based modeling. These are as follows: 1) identifying the “regularities at the societal or macro level” that one wants to explain and formulating general explanatory hypotheses; 2) specifying the agents, environment and model outcomes one is planning to collect; 3) writing a computer program that will make the simulation possible; 4) verifying the model; 5) validating the model; 6) comparing model “output” with the relevant empirical data; 7) assessing analysis results; 8) experimenting with the model. Additionally, a methodology called Agentology [34] guides the reader through the entire development process, from the formal definition of the problem through conceptual modeling and the selection of the development platform to the programming and debugging of the code itself and the final assessment of the model. Different diagrams like global agent diagrams, goal diagrams, activity diagrams, class diagrams and communication diagrams are suggested to support different development stages. An alternative logic with a stronger focus on data is proposed by [49]. With respect to data treatment, artificial intelligence tools are suggested. For robust modeling in marketing [17] proposes and describes four steps in terms of model development and two steps in terms of model accuracy. The four steps in the first case are (1) decide if the model is appropriate, (2) design the model, (3) construct the model and (4) analyze the model. Before selecting an approach, the appropriateness needs to be assessed first. For this, various indicators are presented in the paper regarding the consideration of an empirically grounded agent-based model. A research process with different supporting questions is also presented by [16].
process of [50], special attention is given to attributes and behavioral rules as well as the social and environmental system.

[51] introduce a descriptive grammar as well as an evaluation framework for empirically grounded agent-based models in the public health domain. It exhibits one of the most extensive frameworks in the literature to assist in designing and modeling of such systems. In this context, a grammar for a systematic and consistent description is proposed. This can guide the modelers early in the model development process just as the ODD protocol [43]. The use of the grammar intends to improve communication between the model development team and the policy-makers and will help to ensure that everyone is in agreement about the goals and intended uses of the model. The grammar is divided into seven broad categories: (1) basic model description; (2) model agents; (3) use of data and theories; (4) model context, (5) model outcomes; (6) policy aspects; (7) communication aspects. Each category is further divided into various descriptors with guiding questions for assistance. Furthermore, based on the literature review and expert knowledge, an evaluation framework for developers, funders, policy-makers, modelers, and scientists is presented. The evaluation framework is designed to cover the important aspects of designing, implementing, testing, and disseminating policy-relevant agent-based models, especially for tobacco control regulatory and policy efforts. It can be used to assess the model development processes as well as its outcomes. The framework is structured into five major sections: (1) resources; (2) activities; (3) outputs; (4) outcomes; (5) environment. Each of the sections can be selected for evaluation. The presented case of the framework is based on [52].

To represent human behavior and decision making in system models, [53] suggest a stepwise modeling procedure. The cyclic procedure allows a gradual increase in model complexity. Further guidance for characterizing and parameterizing socio-technological systems is provided by the CAP (Characterization and Parametrization) framework according to [54] and [8]. On the one hand, the framework demonstrates relevant design steps like model characterization, attribute parametrization, behavior parametrization, attribute-based typologization, and behavior-based typologization as well as model scaling. On the other hand, the framework assigns a set of methods to single design steps. [55] demonstrate a structured methodology called ICTAM (Interviews, Cognitive mapping, Time sequence UML, All-encompassing framework, and numerical agent-based Models) for integrating perceptions of stakeholders (qualitative) into the formal simulation models (quantitative). Thereby, five steps are necessary to transform the elicited mental model or personal construct of decision makers through semi-structured interviews (step 1) into the empirically grounded agent-based model (step 5). In between, the individual cognitive mapping is used to depict decision making (step 2). The different cognitive maps are merged together in the next step until a collective map arises (step 3). Special attention is given to the elaboration of differences and similarities among interviewees. Subsequently, the cognitive map is used to derive a sequence of conceptual models (step 4). For support of the first steps, the methodology of [56] can be taken considered. [28] presents verification and validation strategies at different design and model stages.

4.2. Application domain

Empirically grounded agent-based modeling is increasingly considered as an analysis tool for diffusion processes. This can be seen by the rising number of papers in the last years. Case-based applications “have an empirical space-time circumscribed target domain.” [28]. They are usually built “to provide forecasts, decision support, and policy analysis […]” [5]. The application domain of empirically grounded agent-based models of innovation diffusion is very versatile. With respect to the reviewed research papers, the approach has been applied to the following major substantive domains: mobility and logistics, consumption and retail, energy and utilities, nature, and the environment as well as public and education. An overview of the classification of the papers is given in Table 5.
Table 5: Identified application domains

<table>
<thead>
<tr>
<th>Domain</th>
<th>Example</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility and logistics</td>
<td>introduce an empirically grounded, spatially explicit, agent-based model, InnoMind (Innovation diffusion driven by changing MINDs), to simulate the effects of policy interventions and social influence on consumers' transport mode preferences.</td>
<td>[57], [52], [59], [60], [61], [62], [63], [64], [55], [65], [66], [57]</td>
</tr>
<tr>
<td>Consumption and retail</td>
<td>aim &quot;to gain insights on how social influences can affect the market inequalities in the motion picture market.&quot;</td>
<td>[68], [69], [70], [67], [71], [72], [73], [74]</td>
</tr>
<tr>
<td>Energy and utilities</td>
<td>propose an agent-based model to simulate how changes to the Italian support scheme will affect the diffusion of PV [photovoltaic] systems among single- or two-family homes.</td>
<td>[42], [59], [76], [77], [78], [79], [80], [75]</td>
</tr>
<tr>
<td>Nature and environment</td>
<td>develop an agent-based simulation model linked to Geographic Information System (GIS) data in order to investigate the spatial–temporal diffusion of agricultural biogas plants, given constraints on the local availability of feedstock resources.</td>
<td>[41], [81], [82], [83], [36], [84], [40]</td>
</tr>
<tr>
<td>Public and education</td>
<td>apply “a data-driven case study […] of residential mobility [in a medium-sized town in Germany] to systematically explore the role of model detail on model performance.”</td>
<td>[85], [86], [87], [88], [89], [26]</td>
</tr>
</tbody>
</table>

It is obvious that a large share of the retrieved research papers deals with environmental and energy-related innovations. Product innovations analyzed range from hybrid or electric cars [57, 52, 62, 60], through biofuels [66] to photovoltaic panels [80]. Moreover, smart meter diffusions have been analyzed [59]. One reason might be the high societal relevance of these innovations. Promoting consumer choices in respect to environmental technologies is crucial to meet the challenge of climate change and its associated impacts since the adoption of such environmental-friendly products generally only happens slowly [52]. This also falls in line with the statement of [57] that environmental-friendly technologies require influencing the demand side to diffuse on a large-scale. Another reason is the need for an individual-based modeling approach. They claim that “the key strength of agent-based-models is that they overcome the homogeneity assumption of traditional aggregate diffusion models.” [5]. Environmental innovations oftentimes polarize and divide consumers between proponents and opponents. To overcome this homogeneity assumption, agent-based-models seem to be appropriate for environmental and energy-related innovations.

4.3. Component modeling

An overview of the model entities and dynamics of reviewed empirically grounded agent based-models is given in Table 6. In addition to some considerations for consistency and comprehensiveness, they form the basis of the agent-based simulation model. A representation of a comprehensive agent based-model component configuration is given in Figure 2.
Table 6: Identified model components

<table>
<thead>
<tr>
<th>Component</th>
<th>Example</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent attributes</td>
<td>&quot;[... ] each household is assigned four additional attributes: its average income, electricity consumption level, type of housing, and its Sinus-Milieu®.&quot; [75].</td>
<td>[82], [70], [59], [90], [71], [52], [38], [39], [91], [40], [77], [75], [80], [26]</td>
</tr>
<tr>
<td>Consumer heterogeneity</td>
<td>&quot;[... ] was varied during simulations to test the importance of heterogeneity for the model output.&quot; [26].</td>
<td>[82], [83], [70], [92], [52], [59], [71], [37], [38], [87], [77], [40], [75], [66], [57], [26]</td>
</tr>
<tr>
<td>Consumer preferences</td>
<td>&quot;The population of agents is heterogeneous concerning social susceptibility and individual preference [...]&quot; [92].</td>
<td>[83], [70], [59], [90], [37], [37], [38], [39], [75], [26]</td>
</tr>
<tr>
<td>Corporation entity</td>
<td>&quot;This market game represents the interaction between electricity suppliers and the residential electricity consumers.&quot; [90].</td>
<td>[83], [70], [59], [90], [37], [37]</td>
</tr>
<tr>
<td>Policy actor</td>
<td>&quot;The government agent establishes policies in order to influence the vehicle production of manufacturers and the purchasing behavior of consumers.&quot; [59].</td>
<td>[59], [90], [87], [84]</td>
</tr>
<tr>
<td>Product attributes</td>
<td>&quot;First, all living units that lie within acceptable ranges with respect to costs and space provided (see below) and that release the trigger are stored in a list.&quot; [26].</td>
<td>[83], [82], [70], [59], [52], [38], [87], [77], [37], [39], [91], [80], [75], [26]</td>
</tr>
<tr>
<td>Product perception</td>
<td>&quot;We therefore introduce an attribute-specific parameter [...] that determines the 'observability' of an attribute.&quot; [66].</td>
<td>[83], [70], [52], [67], [86], [38], [37], [39], [87], [66], [57]</td>
</tr>
<tr>
<td>Social network</td>
<td>&quot;Each agent also has a social network comprising other agents of similar age, salary, and residential location (within a given agent-specific distance).&quot; [52].</td>
<td>[82], [92], [59], [90], [52], [87], [38], [39], [39], [72], [66], [80], [75], [57], [40], [57]</td>
</tr>
<tr>
<td>Communication channels</td>
<td>&quot;[... ] the information about smart metering and electricity suppliers [...] travels through the consumers' social network via word of mouth; [...]&quot; [90].</td>
<td>[82], [92], [67], [90], [52], [71], [38], [87], [38], [36], [66], [80], [75], [57], [57]</td>
</tr>
<tr>
<td>Marketing model</td>
<td>&quot;The promotion parameter is an indicator of promotional intensity and higher values indicate greater intensity. The brand density parameter is an indicator of distribution channel intensity.&quot; [70]</td>
<td>[70], [52], [71], [90], [67], [79], [57], [66]</td>
</tr>
<tr>
<td>Decision process model</td>
<td>&quot;Agents use two different kinds of decision rules to decide upon adoption or rejection of the modeled innovations: A cognitively demanding deliberate decision rule and a very simple decision heuristic.&quot; [82].</td>
<td>[83], [82], [70], [92], [59], [90], [52], [71], [38], [39], [36], [38], [77], [87], [79], [39], [84], [40], [75], [57]</td>
</tr>
<tr>
<td>Spatial model</td>
<td>&quot;[... ] we explicitly model the geographical distribution of the agents in order to account for the regional differences that have strongly influenced the PV diffusion in Italy.&quot; [75].</td>
<td>[82], [52], [90], [38], [87], [37], [39], [72], [75], [66], [80], [40], [57], [26]</td>
</tr>
</tbody>
</table>

Agent attributes are qualities of agents that characterize (individual) agents representing potential adopters. Most
agent attributes as the personal attributes of consumer agents are incorporated directly. Examples for this are the subjective norms [91], characteristics within the socio-economic coordinate system [57] (age, salary, car ownership [52]), or innovativeness [70]. Agent attributes can also describe certain characteristics of a group of agents such as opinion leaders [71], although in this case they often incorporate heterogeneity as well.

Consumer heterogeneity has been explicitly mentioned by various models based on heterogeneous social structure [82, 75, 52, 64, 38]. Other aspects not explicitly related to heterogeneity falling under this category were social demographics and some consumer attributes that were classified as (more broad-scoped) personal attributes. They can be understood as being attributes of consumer groups, such as demographic household group [52], the number of peers to communicate [52], and the decision strategy employed [38].

Modeling consumer preferences of agents absorbed the categories of ecological aspects of products, such as pollution [83] or environmental concerns [75] and certain aspects of social perception [66, 38, 70]. It further incorporates attitudes relating to preferences and might be used by decision processes [90] or in other words to weigh product aspects [52].

The corporate entity has been mentioned in terms of several aspects related to the representation of corporations in innovation diffusion models. These include strategic alignment [37], suppliers influence [90] [59] and competitors influence. In [70], corporate entities are modeled directly (and endogenously) as brand agents. In this role, corporation entities also interact with the consumer agents to prompt a decision process. This is even more so the case for models where the point-of-sale is a (physically distinct) entity with a range of products, product availabilities, and prices, and a purchase process as implemented by [66].

The policy agent is based on two components, namely the policy measures and policy-makers. The difference between these two is that policy measures are generally rather passive, exogenous components of the models, whereas the policymaker is an active agent that can bring about policy measures. Because of this, this activity can be subsumed, as a component endogenous to the model. This is done by [90], where economic regulations are set by authorities, and [84], where the administration is aggregated by a single policy agent.

Arguably the most important aspect of innovation diffusion modeling is the decision process. Since other actors influence the decision (or at least an actors perception), the influence model is situated primarily in this component. This is also particularly the case with social influences and various evaluation aspects. Due to the multitude of aspects playing into the actor’s decision, a number of other components interplay with it. This incorporates, for example, the channel described in [70], bringing together mass media and advertisement, communication, product attributes and
decision processes. Other aspects touching this category were social as a “person’s perception that most people who are important to him think he should or should not perform the behavior in question” [91]. The mental representation incorporated in [57] also touches the decision model. Additionally, the goals of different actors are influenced by economic circumstances, influencing the decision processes [39].

Obviously, for innovation diffusion models, the product innovations of interest are of fundamental importance. Due to the heterogeneity of the models analyzed, the technologies come in a number of flavors. With an abstract representation of products, however, this diversity can be captured, in particular, since product innovations come to life predominantly in relation to other model aspects. Modeling products also cover some aspects of markets [39] if market mechanisms are only modeled rudimentary.

Product attributes come in many varieties, ranging from technical parameters [59, 82, 87] over ecological characteristics [83, 75, 38] to cost-related parameters [52, 75, 59]. Since these are all characteristics potentially influencing the evaluation of a product, the various aspects can be abstracted into distinct product attributes.

To enable employing models sensitive to cognitive modeling, modeling not only the true qualities but also the perception of qualities is relevant. Perceptual aspects identified are ecological aspects (e.g. fuel efficiency [52]), social perception (e.g. observability [66]) or substitute availability, as mentioned by [83], which intends to model acquiring knowledge about qualities and existence of products. Another aspect of models falling under perception is learning. In different models, learning is modeled through memory and behavior adoption based on the knowledge derived from this memory. Another aspect (used by [57]) that is modeled through the interplay of several model aspects is a mental representation, which is subsumed by the interplay of communication, the social network, media, perception, and the decision model.

4.4. Theoretical grounding

The model entities and dynamics need to be based on different theories. For this, agent-based models represent a flexible tool to implement different relevant theories at various stages. An overview of the retrieved theoretical groundings is given in Table 7.

In accordance with the observation of [4] “[…] that the topology of the social network involved in consumers’ decision making is consistently found to have a large impact on innovation diffusion”, many papers analyzed statements about the importance of the network models. Graph topologies mentioned include small-world networks [59, 75, 60, 80, 82, 64, 76, 36], scale-free networks [67, 39, 87, 71, 92], and social circle [35]. A network topology that is rather used for spatial models are lattice-based topologies [26, 64, 78, 37, 84], or more rarely unspecified neighbor-based interactions [72]. Random graphs [59] were usually used only in contrast to small-world networks.

Modeling the decision process is most commonly grounded in decision theories or in utilitarian approaches [71, 75, 66, 70, 57, 80]. Utility functions calculate a utility value for each option which serves to decide for the adoption or rejection of an innovation. This is done by testing whether a threshold is exceeded, or by rating options against one another based on the calculated utility. Thus, the approach by [39], whose decision process is based on the AHP framework, falls within this category. A rather different approach to utilitarian decision making is taken by [52]. The adoption of an agent will take place when its G-value, which “[…] indicates how much weight the agent places on heuristically perceived benefits related to saving gasoline that is independent of rationally estimated financial benefits […]”, has a certain position relative to a reference agents’ G-value. The adoption decision is implemented as a state transition process [64, 78, 79]. Additionally, [59] uses game theoretic approaches in their decision processes. On the contrary [57] “[…] model agent decision-making with artificial neural networks that account for the role of emotions in information processing”. Simultaneously, an utilitarian approach is also employed.

The other major approach in modeling the decisions of agents is based on decision theories. The frameworks used in the models analyzed are based on the Theory of Planned Behavior [40, 82, 80, 59], the technology acceptance model [91], and the meta-theory of consumer behavior [76, 38]. Furthermore, the consumer framework [35] has been applied. The model used in [57] is based on the HOTCO (hot coherence) model “[…] according to which agents make decisions by maximizing the coherence of their current beliefs and emotions.” [82] use different decision procedures, depending on the typological class of the agent.
<table>
<thead>
<tr>
<th>Domain</th>
<th>Example</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision theory</td>
<td>“The formulation of our behavioral model is motivated by the Theory of Planned Behavior (TPB) a widely applied behavioral model in psychology.” [80].</td>
<td>[82], [59], [76], [93], [38], [87], [79], [35], [40], [57], [80]</td>
</tr>
<tr>
<td>Network models</td>
<td>“Agents are positioned in a social network. The social network is a connected graph [...] which means that a path between any couple of agents always exists.” [92].</td>
<td>[82], [92], [59], [76], [52], [71], [93], [67], [37], [86], [87], [38], [35], [64], [79], [78], [84], [72], [39], [40], [75], [66], [80], [57], [26]</td>
</tr>
<tr>
<td>Opinion dynamics approaches</td>
<td>“During each conversation, the communication partners assume the roles of speaker and listener simultaneously. Thus, belief adjustment is carried out for both agents [...].” [57].</td>
<td>[82], [71], [87], [78], [79], [80], [57]</td>
</tr>
<tr>
<td>Social influence approaches</td>
<td>“An agent’s threshold [...] is the proportion of PHEVs [plug-in hybrid electric vehicles] the agent must perceive in its combined geographic neighborhood and social network [...] to be willing to consider purchasing a PHEV.” [52].</td>
<td>[71], [59], [93], [67], [59], [52], [55], [37], [38], [72], [79], [64], [40], [57], [15], [80]</td>
</tr>
<tr>
<td>Social psychology approaches</td>
<td>“To operationalize adopter categories, the concept of Sinus-Milieus® [...] was chosen.” [82].</td>
<td>[82], [83], [70], [59], [53], [67], [66], [38], [78], [91], [64], [72], [80], [40], [57], [75], [26]</td>
</tr>
<tr>
<td>Stochastic models</td>
<td>“This parameter is a decaying stochastic cost term that accounts for poor products or poor information about the product when in its infancy. This is to account for the inertia that can exist in the early stages of technology adoption.” [64].</td>
<td>[83], [52], [67], [71], [59], [64], [72], [80], [38], [64], [80], [66]</td>
</tr>
<tr>
<td>Uncertainty factors</td>
<td>“[Agents] maximise their personal and social satisfaction taking their uncertainty into account.” [35].</td>
<td>[92], [36], [35], [66], [80]</td>
</tr>
<tr>
<td>Utilitarian approaches</td>
<td>“Using the agent’s utility function and current estimates of attribute values, the utility of each product in the evoked set is calculated.” [68].</td>
<td>[83], [92], [70], [53], [59], [71], [67], [38], [77], [37], [39], [64], [73], [64], [80], [75], [66], [57]</td>
</tr>
</tbody>
</table>

*Social influence* means that the attitudes and decisions of one agent connected to another agent influence the agents’ attitudes, or decision. Quantitative modeling approaches explicitly calculate the degree to which an agent’s influence fares against another agent. This indicator is, among others, used by [75] as the influence of communication. [80] implemented the social component to be the average of the number of adoptions in the neighborhood as well as contacts with adopters outside of the neighborhood. [52] models the fraction of adopters as an exclusion factor, using a social threshold. Another approach to modeling social influence was taken by [35, 36, 38] using the meta-theory of consumer behavior. Social influence is exerted when the agent cannot reach his/her own decision and thus imitates the behavior of the majority. A rather restrictive approach is put forward by [79], who assume that social influence takes place in the case of unanimity. In [57] social influence is exerted through communication processes.

Social influence is closely linked to the concept of *opinion dynamics*. “An opinion leader may influence the diffusion process by increasing the speed of diffusion and/or increasing the maximum adoption percentage.” [71]. [87] “[…] simulates the mechanism that consumers will use word of mouth to communicate (unexpected) events that altered their
perceptions.” [79] use the opinion formation model proposed and adapted it to the case of electricity tariff. [71] put the biggest emphasis on this aspect and gives a good overview of opinion dynamics. In regard to other approaches, [80] notes that “[many approaches] oversimplify the reality of opinion dynamics as a complex, multi-dimensional process”. Modeling emotional stands and belief change interaction, [37] implement belief adjustments in response to external input by changing connection weights between elements of the network.

“Social psychology approaches, arguably the most sophisticated and least parsimonious, are based on psychological theories of behavior.” [4]. Common theoretical bases for this are the “Theory of Planned Behaviour”, used in [59, 80, 38] and the concept of homophily [52, 36, 57, 40]. Furthermore, latter publications base behavior diffusion on Social Learning Theory. Another social theory employed is the theory of Reasoned Action, which was used by [91] and [78].

An additional aspect of social modeling of agents that is often closely linked to agent heterogeneity is modeling the socio-economic characteristics of agents. Different models group their agents with the help of the Sinus milieus® (Sinus Sociovision) as introduced by [82]. This classification differentiates between the social status and the basic values. [75] “[…] incorporate these socio-economic groups and their attitudes toward innovative technologies in the model […]”, whereas [82, 26] utilize these to “[…] operationalize adopter categories […]”.

Additionally, [70] base their model on the ideal adopter categories by [1]. Similarly, [64] work with distributions of characteristics for their agents. [38] use cluster analysis in a space spanned by “[…] income level and basic values to approximate the influence of lifestyle on attitudes toward a technology.” [83] utilize consumer classes. Each class can be thought of as a distinct consumer type, each with a lifestyle that is facilitated by a particular type of technology product. The aim is to create a more fluid classification. The approach of [80] views geography as a good proxy for additional socio-economic demographic variables not captured directly in our model, but which do impact attitudes. [72] use “[…] social, economic and environmental measures” that “[…] capture the individuality of nation agents within the social system.” To initialize their network, [57] used, in addition to cognitive–emotional parameters, individual socio-demographic properties and residential location of our survey respondents. All in all, in most models the approaches are used to construct the social networks of the agent, based on the assumption of the homophily of agents.

Aspects for which stochasticity modeling is used are related to the environment, agent composition and demographic factors, intra-agent aspects, and the social network. The use of environmental stochastic elements includes, among others, the estimation of snow cover and avalanche risk [37], electricity prices [52], the introduction of heterogeneity for the firms modeled [83], stochastic cost terms for non-mature products [64] and lamp lifetimes [87]. Aspects of the agent composition based on stochastic approaches include agent demographics and socio-economic factors [52]. Stochastic intra-agent aspects comprise agent profile preference thresholds for amenities [54], agent thresholds [52], choice of decision strategy [38], utility weights [67], utility and quality thresholds [71], population-wide estimates for agent attitudes at initialization [80], agents utility to model errors [66], and price sensitivities [59], as well as weights for choice criteria [87] and weights for social and personal satisfaction [54]. Characteristics of social networks that were stochastically established include the selection of friends to simulate social influence [52], the construction of an agents’ social network [38], as well as agents location within the network [35].

Uncertainty is generally understood as the lack of information (of the true value) of a variable, although it is interpreted in different ways in terms of the research papers reviewed. [36] and [55] uses uncertainty in order to choose a decision strategy when an agent shows a high variability of satisfaction. In this case, an agent employs strategies that consider decisions that neighboring agents took. [72] views uncertainty as high variance in adoption, when agents adopt in some simulation and not in others. For [80], uncertainty about attributes is a driver for the attitudinal module. [66] models the perception of the quality of a product, based on the assumption that the true quality of the product is unknown to the agents. However, in most of the models this is rather neglected.

4.5. Empirical grounding

Empirical grounding is essential in case-based models. However, as [26] notes, “[…] the ability of a model to explain real world patterns and dynamics is not straightforward […].” A major problem is the multitude of parameters for these models, and the related problem of how knowledge about the target system can be transformed into usable information. An overview of the applied data analysis methods is given in Table 8. Despite numerous papers stressing the importance of data analysis methods, few papers report in detail on their data analysis techniques. An introduction to several methods to characterize and parameterize behavioral responses of human empirically is given in [8].

15
Table 8: Identified empirical grounding

<table>
<thead>
<tr>
<th>Empiricism</th>
<th>Example</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geographic information</td>
<td>„The spatial units are georeferenced, based on GIS (geographical information system) layers concerning elevation, slope, aspect, land use and a thematic differentiation in areas of the destination.“ [37].</td>
<td>[64], [76], [77], [38], [84], [39], [66], [80]</td>
</tr>
<tr>
<td>Expert knowledge</td>
<td>“An energy market expert asserted that diffusion rates and market shares obtained from our model are reasonable.”</td>
<td>[59], [37], [65], [26]</td>
</tr>
<tr>
<td>Field or lab experiments</td>
<td>„The parameterization of the connection weights adjustment was based on data from a separate experimental study, aimed at quantifying how people change their beliefs about EVs [electric vehicles] in response to influence of others […].” [57].</td>
<td>[66], [57]</td>
</tr>
<tr>
<td>Focus groups</td>
<td>„The parameterization of the connection weights adjustment was based on data from a separate experimental study, aimed at quantifying how people change their beliefs about EVs in response to influence of others […].” [66].</td>
<td>[66], [57]</td>
</tr>
<tr>
<td>Interview study</td>
<td>“Classification and typification of foresters, private forest owners, and certain wood fuel consumers, based on qualitative interviews conducted with market participants, scientific studies of nonindustrial private forest owners […].” [39].</td>
<td>[67], [91], [39], [84], [40], [66]</td>
</tr>
<tr>
<td>Literature analysis</td>
<td>“We have identified these characteristics as being prominent amongst early stage adopters of EVs from the literature.”</td>
<td>[70], [59], [87], [37], [64], [40]</td>
</tr>
<tr>
<td>Market research</td>
<td>“The spatial distribution of lifestyle groups in percent per populated cell was derived from data provided by Microm.”</td>
<td>[82], [70], [67], [75]</td>
</tr>
<tr>
<td>Public (census) data</td>
<td>“[…] Demographic rates from 2001 to 2012 are used to calculate the number of births, deaths, emigrations, immigrations and separations. The probability to receive the respective event varies between different household groups.”</td>
<td>[82], [35], [37], [69], [39], [64], [75], [80], [26]</td>
</tr>
<tr>
<td>Social survey</td>
<td>“[…], consumer preferences were obtained from an online survey of a representative sample of 1000 subjects from the Austrian population carried out by a professional market research firm.”</td>
<td>[82], [59], [67], [71], [76], [38], [37], [87], [64], [39], [91], [57], [66], [80], [26]</td>
</tr>
</tbody>
</table>

Various research paper incorporates public data/census data, of which the most commonly employed is census data. Other data sources are authorities, but also institutional data sources like, for example, the Eurobarometer Survey, ISTAT 2009 data, data from the Italian energy authority [36], data from Holzenergie Schweiz, the third Swiss National Forest Inventory and the US Energy Information Administration [39], as well as national development indicators drawn from the World Bank’s databank [72].

Rather a category of data than a data source, although often based on existing data, is data based on geographic...
information systems (GIS). These data are especially important for spatial models intending to capture the geographical characteristics of spatial entities. [60] base their data on gas stations in Austria on OpenStreetMap data. [80] used data “[…] collected by the electric utility (Austin Energy) as a part of the implementation of Austin’s solar rebate program”, which they “[…] geocoded to street locations in a GIS […]” and which they “[…] overlaid with […] socio-economic demographics and environmental layers”. [39] derive “[t]he number, size, and location of wood fuel heating systems in Switzerland […]” from data by Holzenergie Schweiz. [77] uses GIS data “[…] to determine the local availability of substrate resources for biogas production given the underlying agricultural production of the local communities.”

The third category of data sources used by the analyzed models is based on existing literature. [64] base their characteristics of early adopters of electric vehicles on existing literature, assuming them to be like the characteristics of early adopters of hybrid electric vehicles since they report that they “[…] did not have access to econometric estimates of the parameters”. Additionally, [70] base “[t]he number of consumers within each […] on proportions defined by Rogers”, which they also use for their categories of ideal adopters (see the section about theoretical grounding above). [37] employs literature in order to draw “[t]ourist profiles […] from the literature […]” that they further develop “[…] through contacts with local tourists and economic actors.” [87] base some core parameters (especially to base weight factors) on literature, where [40] draw on literature for network design. [59] uses GIS data “[…] to determine the local availability of substrate resources for biogas production given the underlying agricultural production of the local communities.”

Another common approach is social surveys. The use of survey data often falls into the qualitative or the quantitative camp. Not necessarily qualitative in nature of the surveys, but rather in how the results were used. For example [67] used surveys to derive information on cross-cultural differences in socially influencing situations. Model mechanisms that were informed by surveys include pull-factors [26], the influence of word-of-mouth [59], behavioral, financial and social components of the adoption decision [80], and the decision-making process for adoption [76]. Approaches directed towards using survey data in order to inform entities or aspects of these include deriving consumer preferences [59, 64], the identification of opinion leaders [71], domain-specific knowledge about alternative fuel vehicles of consumers [59], customer expectations and concerns, perceived value proposition, willingness to pay and potential drivers for the intention of use [91] and establishing profiles or categorization of agents types [82, 37].

Survey data has just been utilized for parameterization or configuration. This includes, among others, the parameterization of the social model and configuration of model agents [76], parameterization of connection weights adjustment regarding belief strength and sender–receiver belief congruence [57], the distribution and correlation of household characteristics [26], and the choice probability of decision strategy [76]. Informing the model via social aspects of agents was done by [26] who use demographic aspects such as net equivalence income or [91], who used sociodemographic data derived through a survey. Survey results were also used for validation and elaboration of simulation findings by [87] and model validation by comparing model predictions with data on actual transport choices [57].

A more detailed approach than the use of surveys is possible in terms of interviews, focus groups and drawing on expert knowledge. [37] used interviews with tourists and economic actors to derive profiles for their agents. [39] based their “[c]lassification and typification of foresters, private forest owners, and certain wood fuel consumers […] on qualitative interviews conducted with market participants […].” [84] combined interviews and ethnographic analyses. Drawing a clear line between expert knowledge and the use of interviews is of course not always possible.

[66] used focus groups of consumers to identify product attributes. [57] utilized focus groups to maximize the empirical plausibility of a network representing individual agents. “They provided [them] with a detailed, in-depth picture of people’s needs regarding transport as well as their current cognitive and emotional representations […]”. [57]. Expert knowledge has been considered by [37] and [66] to complement previously acquired information. [26] employed expert knowledge from the city administration and assessing the number of uninhabitable flats.

Less common data sources and methods were field or lab experiments [57], marketing data [70, 75], micro-data [64] and proxy data [37]. For example [66] included an adaptive conjoint analysis, whereas [57] used a “[…] vignette experiment by asking participants to rate their agreement and their perceived belief change on a series of unrelated statements about the use of electric vehicles and combustion engine cars.”
5. Research agenda

The reviewed empirically grounded agent-based models already cover a wide range of required system characteristics. Different methods, components, and approaches define an excellent foundation for further model development. According to the findings of this review, the future agenda of modeling such simulation systems should be based on the following aspects: a generic framework 5.1, model coupling 5.2, research consistency 5.3, modular testing 5.4, actor involvement 5.5, behavior modeling 5.6, network foundation 5.7, data transparency 5.8.

5.1. Generic framework

Providing a generic framework for innovation diffusion modeling and integrating modeling approaches with a different focus, allows modelers to concentrate on system modeling as well as simulation analysis instead of model implementation. This would further allow them to test and compare specific model mechanisms with little effort and identify what kind of model would be most appropriate for the (sub-)system of their choice. In practice, generic models can help companies to evaluate business model innovations in a fast-changing environment and thus develop a sustainable business strategy.

5.2. Model coupling

Existing empirically grounded agent-based models should be coupled with different models to support decision makers in identifying the optimal strategy. In this context, among others [66] see the embedding of an agent-based model in a simulation optimization framework as a challenging but rewarding research endeavor. Additionally, [26] state that coupling of the agent-based model also enables the variety of different model details. In contrast, [64] plan to link the implemented model with engineering network-cost models.

5.3. Research consistency

Individual modeling decisions are stated without any defendable rationale according to the findings of this review. In the optimal case, modeling decisions should be derived from literature, based on (the comparison of) established models, empirical grounding of parameters, and analyses that yield which configuration seems to work best for the approach. Different protocols and processes are already available to support the researcher and the documentation process. First, existing process frameworks and documentation protocols need to be integrated and standardized. Second, future developments should refer to these guiding principles and add the completed protocols.

5.4. Modular testing

To make valid predictions and recommendations, models have to ensure they can effectively capture the essence of what they are trying to model. Thus, future works need to concentrate on exploring the effects of the interplay of the factors analyzed with other policy mechanisms like, for instance, the effect of a reduction in uncertainty or of campaigns specifically targeted to exploit social influences in fostering positive behaviors [52]. Thereby, a consequent modularization of single model elements that can be activated or disabled is necessary to evaluate the interplay and the influence of each module on the results [56].

5.5. Actor involvement

Analyzed models scarcely consider other actors than end consumers. Specifically, the model analysis is lacking in regard to the role organizational actors play in existing system architectures and the resulting impact they might have. However, they are rarely endogenously modeled. The explicitness of their decision-making processes would make them excellent subjects of investigation. In this sense, [70] that interactions among brands would further advance the existing model by incorporating the effect of brand agents on one another in terms of strategy adaptation and entry/exit behavior. Similarly, this applies to policy agents. [59] plan for the government agent to be endogenous to the model with the goal of maximizing the social good.
5.6. Behavior modeling

Arguably the most important aspect of agent-based research is decision modeling. The theoretical approaches are essential since they strongly influence the results of the simulation. At the same time, considered decision-making processes hardly vary between the heterogeneous agent groups. In future decision-making processes, there is a need for the integration and analysis of various influencing factors. In accordance with this, \[38\] want to improve the cognitive model of the decision process of agents, introducing a variable frequency of decision making and make this frequency dependent on external variables such as price developments and media attention. Similar to this, \[57\] intend to extend the model to account for technological innovations and marketing campaigns of manufacturers and other suppliers. At the same time, there is also a lack of a theoretical framework that explains the different influencing variables. Among others things, there is no communication model in existence that explains what information consumers exchange and how this affects their decision \[66\].

5.7. Network foundation

The interaction of agents depends on the relationship individual agents have with one another. The experience of one agent should not be valued the same as the experience of another agent who is spatially distributed and thus more far away. Additionally, trust and perceived competence, as well as the emotional stance towards the other agent might be very important and is rarely modeled in terms of network structures. Thus, future work needs to focus on networks which combine geographical and social influences. For this, \[71\] want to explore the relationship between the innovativeness of a consumer and the importance of the normative influence for this consumer, as well as to test the model with different network structures, for example, more complex network structures, such as dynamic networks.

5.8. Data transparency

The results of even the best theoretical models will highly depend on the underlying input data. Various future approaches for empirical grounding are aimed at improving model parameterization once more data becomes available \[59\] \[52\]. Simultaneously, further work needs to involve systematic data collection to better inform the choice of parameters \[64\]. Additionally, integrated data need to be validated \[26\]. This is also required since different researchers are willing to collect more data based on different approaches.

6. Concluding remarks

Empirically grounded agent-based models for innovation diffusion are of growing interest. Existing models differ strongly in their design and their grounding and therefore in their predictive power. At the same time, there are only partial reviews that help to compare different concepts to be used as a fundamental basis for model design. In this context, this review paper outlines and analyses various applied research papers in order to guide new research as well as to propose future research.

First, relevant terms have been defined for a systematic analysis. Subsequently, a systematic review has been conducted regarding the model development procedure, the model entities, and dynamics as well as the grounded theory of the models. The analysis covered various characteristics of the individual models to give an overview of existing research. Future empirically grounded agent-based models need to cover additional issues, such as generic modeling, model coupling, research consistency, modular testing, actor involvement, behavior modeling, network foundation and data transparency. In a further step, a novel model approach should integrate the findings into one holistic and generic design framework and a corresponding formal model.

Bibliography


