Measuring Ecosystem Complexity - Decision-Making Based on Complementarity Graphs

Rainer Schmidt
Munich University of Applied Sciences
rainer.schmidt@hm.edu

Rainer Alt
Leipzig University
rainer.alt@uni-leipzig.de

Alfred Zimmermann
Reutlingen University
alfred.zimmermann@reutlingen-university.de

Abstract

Platforms feature increasingly complex architectures with regard to interconnecting with other digital platforms as well as with a variety of devices and services. This development also impacts the structure of digital platform ecosystems and forces providers of these services, devices, and services to incorporate this complexity in their decision-making. To contribute to the existing body of knowledge on measuring ecosystem complexity, the present research proposes two key artefacts based on ecosystem intelligence: On the one hand, complementarity graphs represent ecosystems with an ecosystem's functional modules as vertices and complementarities as edges. The nodes carry information about the category membership of the module. On the other hand, a process is suggested that can collect important information for ecosystem intelligence using proxies and web scraping. Our approach allows replacing data, which today is largely unavailable due to competitive reasons. We demonstrated the use of the artefacts in category-oriented complementarity maps that aggregate the information from complementarity graphs and support decision-making. They show which combination of module categories creates strong and weak complementarities. The paper evaluates complementarity maps and the data collection process by creating category-oriented complementarity graphs on the Alexa skill ecosystem and concludes with a call to pursue more research based on functional ecosystem intelligence.

Keywords: Ecosystems, Digital Platforms, Assistant Platforms, Complementarities, Ecosystem Intelligence, Computationally Intensive Theory Building

1. Introduction

Ecosystems have emerged as a concept underlying most of today's successful businesses (Cusumano et al., 2020). It is a rather challenging form of organization since it foresees the interaction of multiple elements from independent users and providers. Ecosystems are closely related to digital platforms, which initially featured a clear purpose: while the early platforms were either transaction (Gawer & Cusumano, 2014) or innovation platforms (Gawer & Cusumano, 2014) many of today's platforms follow a hybrid approach (Cusumano et al., 2020) (that combines treats from innovation and transaction platforms. In parallel, more and more of these platforms are digital platforms that consist of a software core functionality accessible via an interface and enables the creation of modules (de Reuver et al., 2018). Among the digital platforms in this regard are the app stores with their underlying operating systems from Google (Google. Inc, 2011) and Apple (Roma & Ragaglia, 2016).

The variety of resources, which means products, services, etc., provided by platforms drives platform complexity (Alt, 2021). Initially, many platforms were limited to one resource. For example, Uber offered only transport from A to B. Today, more and more platforms provide different resources, such as services and products. For example, Airbnb offers apartments but also services connected with these apartments, such as cooking events. The platforms' variety of resources enabled by a composite architecture encapsulates resources as modules and provides mechanisms for their composition. Examples of digital platforms with a composite architecture are assistant platforms, such as Amazon Alexa or Google Assistant (Schmidt et al., 2021). They use a composite platform architecture and combine it with a declarative voice-based interface that eases access to a broad variety of devices and services.

Ecosystem intelligence has been applied to understand the dynamics and mechanisms in ecosystems. Today, ecosystem intelligence approaches depict the structure and relationships between the actor groups of the ecosystem as undirected, simple graphs (Basole et al., 2012). They are based on the theory of network effects (e.g. Katz & Shapiro, 1994) and postulate that an increase in users of a product or service increases its value. The present research aims to enhance existing ecosystem intelligence approaches for complexity measure-

URI: https://hdl.handle.net/10125/103057
978-0-9981331-6-4

ment to improve the understanding of these digital platform ecosystems for researchers and decision-makers alike. Research suggests that value creation in composite and assistant ecosystems should be measured from a module perspective to capture broader value creation details. We see our thesis supported, by the fact, that a more complete perspective on ecosystems has been introduced on a theoretical level by Jacobides et al. (2018), which current ecosystem intelligence approaches have not adopted. In particular, this endeavor promises to identify value-creating relationships within ecosystems from a module perspective.

Two artefacts are suggested to provide intelligence on complex ecosystems of composite and assistant platforms: A proper representation of the ecosystems and a procedure to collect the data, which leads to a first research question:

RQ1: How can we improve the representation of ecosystems?

Another challenge to ecosystem intelligence is the lack of publicly available data. Although digital platforms have abundant data, hardly any data on categories and the like are publicly available for competitive reasons. These data may be seen as a competitive advantage for platform providers, who use this data for decisions on products, categories, campaigns, product bundles, and the like. To allow non-platform providers access to such information, procedures are needed on how to collect the data for ecosystem intelligence in an indirect way, which motivates a second research question:

RQ2: How can we collect data on the ecosystem indirectly?

To answer both research questions, this paper proceeds as follows: We present the research background in section two and explain the methodological approach in section three, which is based on the design science research approach from Johannesson and Perjons (2021). Section four describes intelligence challenges in complex ecosystems and derives requirements from them. We develop two artefacts in section five to meet these requirements and demonstrate their use in section 6. In the seventh section, we demonstrate the creation of category-oriented complementarity maps on the Alexa assistant platform. The eighth section discusses our findings and concludes the paper.

2. Research Background

2.1. Ecosystems

Ecosystems represent a loose form of coordination among participants without hierarchical governance mechanisms such as those used within companies (Jacobides et al., 2020). They emerge around platforms driven by network effects (e.g. Katz & Shapiro, 1994) – that describe an increase in the value of membership for one participant if other participants join the ecosystem - and complementarities (Jacobides et al., 2018).

Non-generic complementarities (Jacobides et al., 2018) incentivize ecosystem participants to accept coordinating structures. They allow for stronger coordination than in purely market-based environments. Therefore, modularity is a key element of ecosystems, since the definition of interfaces fosters the coupling of modules and thus offsets complementarities without knowing details about the internal realization of the individual modules (Jacobides et al., 2018). Non-generic complementarities are divided into four categories (Jacobides et al., 2018): Unique and supermodular complementarities, which may both appear in production and in consumption.

Unique complementarities refer to one module. They emerge in production if products are produced more efficiently, better, faster, etc., through coordination within the ecosystem. In production, a unique complementarity improves a module's quality and efficiency by using the ecosystem's coordination mechanisms, particularly its interface definitions. In consumption, unique complementarities reveal if the collaborative consumption of products ("joint consumption") yields advantages to isolated consumption. They manifest through increasing the value of a module if it is used with other modules of the platform.

Supermodular complementarities describe the impact of one module on another module. Supermodular complementarities occur in production and consumption (Jacobides et al., 2018). Supermodular complementarities in production are present if the increase in the production of one module increases the quality of another module. However, other authors also identified negative effects (Srinivasan, 2021). Supermodular complementarities in consumption refer to the increased use of a product or service that leads to increases in the value of other products or services. The consumption of one module, together with other modules, positively impacts other modules.

2.2. Ecosystem intelligence

Ecosystem intelligence is a data-driven approach to improve the understanding of existing ecosystems. It denotes the structured analysis of ecosystem-related data to support decision-making (Basole, 2020). A review of ecosystem intelligence and modeling as approaches to empirical research on ecosystems are given in (Järvi & Kortelainen, 2017) and (Jussila et al., 2014). Ecosystem data contains many entities, relationships, activities, and issues (Basole, 2020) relevant to the present research. Nevertheless, only a few approaches use web mining as one technique within data science to scrape data from enterprise websites (e.g. (Kinne & Axenbeck, 2020).

An existing approach that applies network analysis for ecosystem intelligence is presented by (Basole et al., 2015), which is based on analyzing the network effects and relationships between actors. While this is helpful, more data beyond actors is available in ecosystems. In addition, ecosystems have differentiated network effects between product and service segments (Basole & Park, 2019). Scholten (2013) introduces a visual notation to depict and manage network effects. A graph-oriented approach is also suggested to depict complementarities in supply chain relationships (Benali & Burlat, 2012). Digital forensic approaches for investigating the Amazon Alexa ecosystem are presented (Chung et al., 2017).

A challenge of ecosystem intelligence is to make more comprehensive data available for analysis. Using this data, customer-oriented metrics to assess ecosystem health may be created (Pidun et al., 2021). Existing ecosystem intelligence approaches rely on third-party data sources if this data is unavailable. Basole (2020), for example, evaluated annual reports and other documents to create a data basis. This indirect approach has also been applied to a qualitative investigation of the Artificial Intelligence ecosystem (Jacobides et al., 2021).

3. Methodology

This paper applies a five-step research method based on the framework of Johannesson and Perjons (2021). It implements the activities of the framework, as shown in Table 1.

Activity	Implementation
Problem explication	Ecosystem representation, unavailability of data
Define requirements	Representation of complex ecosystems, data collection process
Design and develop artefact	Complementarity graph design, proxy-based data collection
Demonstrate artefacts	Category-oriented complementarity maps
Evaluate artefact	Alexa ecosystem

Table 1. Research Methodology

We start by explicating the problem that existing ecosystem intelligence approaches increasingly have difficulty capturing modern ecosystems on the level of detail and the speed of change. Following (Baldwin & Clark, 2000), we consider a system as complex if it cannot be designed or understood by a single person in all its details. The underlying platforms drive the complexity of ecosystems. Their complexity has significantly increased due to the rise of hybrid and composite platforms. Therefore, ecosystems have also increased in complexity and are difficult to capture with existing intelligent approaches.

To improve ecosystem intelligence, we need two new artefacts. First, an ecosystem representation should capture more details than the prevailing network effects-based approaches. The developed artefact shall capture the externalities that emerge in ecosystems based on advanced platforms, such as composite platforms. Second, a process artefact must be created to collect and process the data necessary to depict complex ecosystems. It must cope with the lack of publicly available data because, due to platform completion, not all data on platforms are publicly available.

We then create two artefacts. We conceptualize complementarity graphs by analyzing externalities and especially complementarities of ecosystems. The modules are mapped to the nodes of the graphs, and the complementarities are mapped to different edge types. The data collection process itself is based on proxies. Proxies substitute non-accessible variables that allow determining complementarities, although the original data are unavailable to the public.

We introduce category-oriented complementarity maps in section six to demonstrate the use of our artefacts. They are materialized aggregations of queries on the complementarity graphs. To evaluate our findings, we create category-oriented complementary maps for the ecosystem of the Alexa platform in section seven.

4. Problem Explication and Requirements

Based on the two research questions formulated in the first chapter, two problem areas, as well as important requirements, are described.

4.1. Ecosystem representation

Existing ecosystem approaches assume that a simple, undirected graph can describe ecosystems (Battistella et al., 2013) and Basole (2016). Actors are represented as vertices, and the relationships, such as network effects, are represented as an undirected edge between the actors, which fails to differentiate edges into types and the like. Ecosystem actors may learn about other actors they relate to, but they do lack further information on the type and strength of the relationships.

However, ecosystems' complexity and their description have grown in three dimensions. First, there are more differentiated theories on the externalities driving the emergence of ecosystems. For example, the complementarities framework of Jacobides et al. (2018) differentiates four types of externalities. Second, more and more composite platforms integrate multiple services and products from different categories and are no longer just one category, as with monolithic platforms.

Thus, the modules of composite platforms, such as assistant platforms, are differentiated by categories. Third, ecosystems are no longer based on just one type of platform, such as innovation or transaction platforms, but contain both traits and thus are called hybrid platforms (Cusumano et al., 2020).

4.2. Platform competition and data access

Data on the ecosystem of platforms are available to the platform operator and the module vendors. The platform operator knows the usage of the modules, usage patterns, etc. However, it is not in the interest of the platform operator to share information because this would give competitors important insights. The module vendors receive information on their modules' quality, efficiency, interaction, and usage patterns, which helps improve them. For example, the Alexa dashboard provides such information to module vendors (*Alexa Developer Documentation*, 2021). However, as same as the platform operators the module vendors do not disclose these data due to competitive reasons.

4.3. Requirements

The above presentation has identified two problems: The representation of ecosystems and data collection. For both issues, concrete requirements will now be developed. To reflect the increased complexity of ecosystems, a refined representation is to be created that covers as many complexity dimensions as possible. For a better representation along the first complexity dimension, the differentiation by types of externalities is required, which sheds light on different types of externalities. The second complexity dimension, the transition from monolithic to composite ecosystems, requires a differentiation of the categories according to the categories of modules. This paper will not address the third dimension of complexity and is the subject of future research.

To cope with the competition-induced data unavailability, it is necessary to replace the data needed for ecosystem intelligence. It is not to be expected that the operators of platforms and the providers of modules will move away from the secrecy of certain data sets.

5. Approach for Complex Ecosystems

Two artefacts are suggested to address the requirements above. First, complementarity graphs represent externalities of ecosystems and the categories of modules. Second, a data collection process aims to tap data unavailable to non-platform providers due to platform competition.

5.1. Complementarity graphs

Complementarity graphs use the modules as vertices and the complementarities as edges (see Figure 1) to address the first dimension of complexity as described in section 4.1. They complement the existing network graphs (Basole et al., 2016) that depict actors as nodes and network effects as edges. Complementarity graphs are multi-edged graphs because multiple complementarities may exist between the vertices.

A complementarity graph represents ecosystems as a collection of vertices that represent modules. Between them, one or multiple edges exist that represent complementarities. To differentiate the complementarities depicted as edges, the complementarity framework of Jacobides et al. (2018) posits that complementarities can be unique or supermodular and either emerge in production or consumption.

Complementarity graphs have four types of edges representing unique complementarities in production and consumption and supermodular complementarities in production and consumption. The edges are directed because complementarities are directed: The complementarities between module X and module Y can differ from those between module Y and module X.

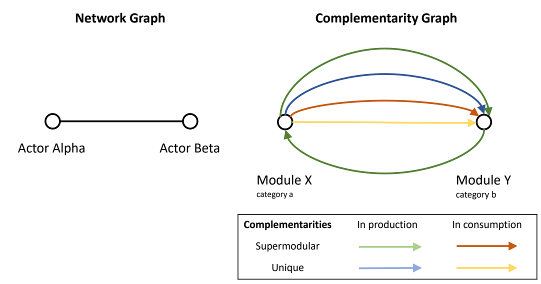

Figure 1. Network and Complementarity Graphs

To address the second dimension of complexity, heterogeneity, we augment the vertices by category information. The vertices then carry the information on the module's category. This category-oriented differentiation of modules enables the creation of category-specific subgraphs and their comparison. It allows the strength of complementarities between different categories to be determined.

5.2. Ecosystem intelligence data collection

We create a data collection process as an artefact by combining proxies with web scraping data collection, as shown in Figure 2. Proxies substitute the non-accessible data. The use of proxies is an approach that has been successfully used in several research areas. For example, the ranking of products is used as a proxy for market share (Brynjolfsson & Smith, 2000). Similarly, Garg and Telang (2013) use proxies to estimate app demand.

Web scraping yields a number of advantages (Landers et al., 2016) for data collection: The collected data are behavioral, and the collection of large datasets is possible with minimal effort. Furthermore, web scraping eliminates the risk of research contamination, the time to conduct a research study is dramatically reduced, and there is a low threshold.

The data collection process contains the following tasks. We start by the identification of variables needed for ecosystem intelligence. Then we check the accessibility of the variables in the ecosystem. To substitute non-ecosystem variables, we reengineer the available data sources of the ecosystem. Based on it, we define proxies for the non-accessible variables. Finally, we define web scraping procedures for collecting the accessible variables and the data for proxy creation.

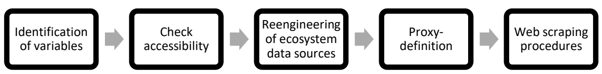

Figure 2, Ecosystem Data Collection Pipeline

First, we identify the variables for ecosystem intelligence by analyzing all four types of complementarities, as shown in Table 2. All these variables are not accessible and have to be substituted by proxies.

		Variables	Proxies	Web Scraping
Unique	Production	Specialization	Number of categories	Categories
Unique	Consumption	Value creation	Average module rating	Module rating
Supermodular	Production	Increase of production	Chronological progression of the number of modules	Module IDs
Supermodular	Production	Quality, efficiency, etc. of another product or service	Improvement of the rating and reviews of other modules	Module rating and reviews
Supermodular	Consumption	Increase of consumption	Increase in count of reviews of co-activated modules	Co-activations
Supermodular	Consumption	Value of another product or service	Improvement of the rating and reviews of other modules	Module rating and reviews

Table 2. Variables and Proxies

We use the re-engineered data model of assistant platforms (Schmidt, Alt, et al., 2022), which recognizes modules as a key element for assistant platform ecosystems. To organize the available modules, they are stored in a module registry (see Figure 3), which is an established solution in service- or component-oriented architectures.

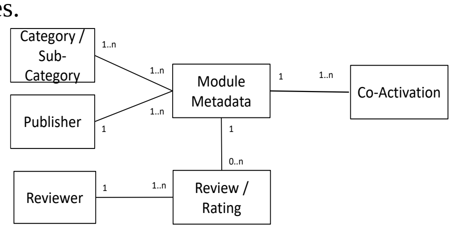

Figure 3. Meta-Data in the Module Registry

The metadata in the module registry supports the matching of modules (often referred to as orchestration) on the platform. It also contains information on the publisher and the reviews are identified by a unique ID and connected to the module description via the ID of the modules. In addition, the reviews contain a unique identification of the reviewer, and co-activations deliver information about modules that are activated together frequently. Metadata also allows users to retrieve suitable products and service providers more quickly. For example, co-activations offer users hints to other modules that are of possible interest to them. They intensify the motivation to use the platform and increase exit barriers and switching costs. Analyzing the metadata in the module registry serves to identify the data relevant to creating proxies. First, the publisher information allows identifying modules from the same publisher. This is interesting for proxies that describe complementarities in production. The publishers also describe the functionality of their modules with the help of categories that the platform operator has standardized. They are supposed to support the users in the search for modules by enabling a quick narrowing down of the functionality of the modules. The categorization in turn supports the identification of sets of functionally similar modules.

A second source of information for creating proxies are co-activations, which the platform operator for each module lists. They reflect the number of modules that are activated together and may point users to other modules that may be useful to them. The third group of information stems from users who provide ratings and reviews. Such evaluations based on social mechanisms are valuable since modules feature high specificity, low fungibility, and long-tail characteristics (Schmidt, Kirchner, et al., 2022). These reviews also provide insights into which modules are used together. To enable the definition of proxies, we create web scraping procedures for collecting categories, module ratings and reviews, module ids, and co-activations.

6. Category-Oriented Complementarity Maps

Complementarity maps are materialized aggregations of complementarity graphs that demonstrate the use of the artefacts. They are a visualization of complementarity relationships differentiated by the categories involved. Complementarity maps aggregate the information about the different complementarities in four fields, as shown in Figure 4 Two for the representation of unique complementarities in production (blue) and consumption (yellow) as well as another two for the representation of supermodular complementarities in production (green) and consumption (red). We will also use the color scheme in section 7.

Category-oriented complementarity maps are two-dimensional representations of complementarities differentiated by their category of functionality, as shown in Figure 4. They have two dimensions representing the categories of the ecosystem. Each intersection of two categories depicts the complementarities between two categories. We use the crossing of category "a" and category "b" as example. There, the means of the strength of the complementarities are depicted according to the legend. The two top cells give the means of the supermodular complementarities in production and consumption; the lower cells give the means of the unique complementarities.

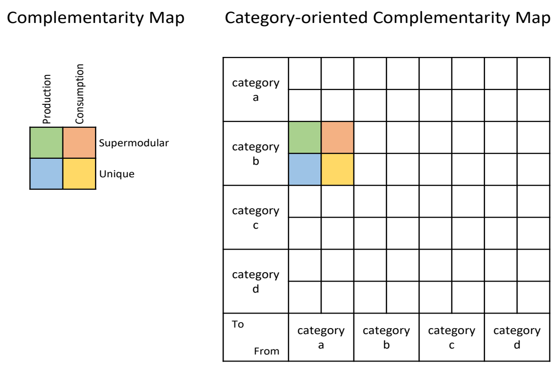

Figure 4. Complementarity Maps

The unique complementarities refer to the value created by the start category (here, category a). The supermodular complementarities refer to the value created by the target category (here, category b).

The unique complementarities in production (blue) represent the mean value created by a module of category "a" if it is co-deployed with a module of category "b". Using this information, the module provider can select the categories his module works best with.
The unique complementarities in consumption (yellow) depict the mean value created by a module of category "a" if it is used together with a module of category "b". This information enables the module provider to detect usage patterns of his module. In particular, he learns about beneficial combinations with other module categories. The supermodular complementarities in production (green) express the mean value created by a module of category "b" by an increase of production of a module of category "a". The information on supermodular complementarities is helpful to module providers that provide modules in both of categories "a" and "b". The supermodular complementarities in consumption (red) represent the value created by a module of category "b" because it is used together with a module of category "a". The knowledge of supermodular complementarities is helpful to module providers that provide or consider providing modules of both categories.

The platform provider can drive platform growth and value creation by incentivizing module categories with strong complementarities, ideally with a high turnout. Similarly, the platform provider can use the information on weak complementarities to start targeted initiatives for strengthening complementarities. Another important feature of category-oriented complementarity maps is their ability to predict the complementarities of a yet-to-be-designed module from its category.

7. Category-oriented Complementarity Maps on the Alexa Ecosystems

To evaluate the use of category-oriented complementarity maps, we use them to depict the complementarities of the Alexa ecosystem. This ecosystem is highly heterogeneous due to its 22 categories of functionality. Amazon Alexa is the assistant platform with the largest market share, closely followed by Google Assistant (Vailshery, 2021). Figures have shown a strong increase in Alexa skills in all important markets (*Amazon Alexa*, 2021) and a growth in the number of devices accessible via Alexa. In 2020, over 100,000 different types of devices were connected to the Alexa ecosystem worldwide (*Infographic*, 2020). The ecosystem notion also applies since assistant platforms associate devices outside their original platforms, such as TV sets and even microwave ovens, become access points (Chung et al., 2017).

We collected raw data from the skill section of the Amazon website (*Alexa Skills*, 2021) during April and May 2021 using several web scraping tools (Gunawan et al., 2019). The scraping mechanisms provided two CSV (comma-separated value) files. The first CSV file contained the description of the skills, and the second file contained the reviews. The files are linked via the ASIN, the Amazon Standard Identification Number used globally for Amazon products. We reverse-engineered the data model of the Alexa platform until we met data fields such as descriptions, details, and permissions that contain semi- and unstructured text entries. For their interpretation, the documentation of the Alexa skills was used (*Alexa Skills*, 2021).

For some data, we had to develop special procedures. For example, a list of up to ten co-activated skills is given for each skill. The skills on this list are activated together with the respective skill. To avoid bias by the limit to ten skills, we collected the co-activations of all skills into a table with the ASIN of the skill as the primary key. We created an entry for each co-activated with the ASIN of the co-activated skill as the foreign key. Then the ASIN of the co-activated skill became the primary key, and the former primary key became a foreign key. Finally, we created a conceptual data model using the scraped data by applying normalization procedures from relational database theory (Codd, 2001). The analyses have been done in a Python Jupyter notebook (*Project Jupyter*, 2019) containing Pandas (*Pandas - Python Data Analysis Library*, 2021).

We created category-oriented complementarity maps for both unique and supermodular complementarities in the production and consumption of the Alexa ecosystem. The following category-oriented complementarity maps use the same color scheme as in Figure 4. The strength of complementarities is depicted as a value between 1 and 5. A missing value indicates a lack of co-activations between the modules.

7.1. Unique complementarities

To measure unique complementarities in production, we used the number of co-activations as a proxy. We put it in relation to the average rating of the module as a proxy for its quality. Our findings show that due to the Alexa ecosystems' heterogeneity, the unique complementarities are distributed unevenly. We further investigated the unique complementarities in production by analyzing how the number of categories covered by the modules of a publisher influences the quality of the modules. We compared the co-reviews with the average rating depending on the category and got the following results (see Figure 5).

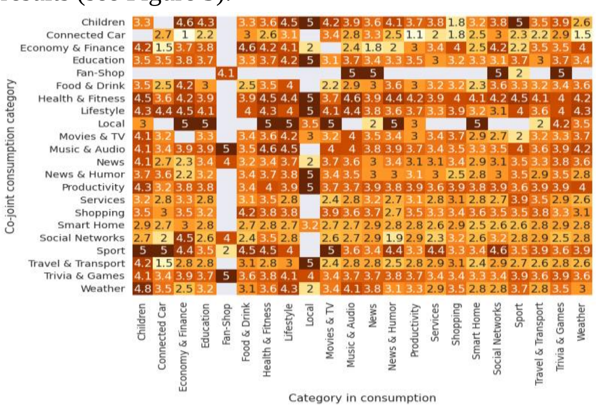

Figure 5. Unique Complementarities in Consumption

Unique complementarities in consumption describe assistant platforms that use them as a skill and increase their value to the users. We used co-reviews, reviews by the same user for different skills, as a proxy to measure the joint consumption of skills. The results also show that the unique complementarities in consumption are dependent on categories, as expected for a heterogeneous ecosystem.

7.2. Supermodular complementarities

Supermodular complementarities in production are present through the increase of skill quality, etc., with an increase in production of another skill. We used the number of co-activations as a proxy for production and investigated the impact of rating other skills as a proxy for quality.

Supermodular complementarities in consumption manifest as the increase of value creation of one module through the increase of consumption of another module. As the consumption of modules is not directly observable, we use the number of reviews as a proxy for the consumption of a module. Therefore, we used the number of co-reviews as a proxy for the joint consumption of skills and the average rating for the value created by the connected skills. The results in Figure 6 again confirm the heterogeneity of the Alexa ecosystem by the highly variant distribution of complementarities.

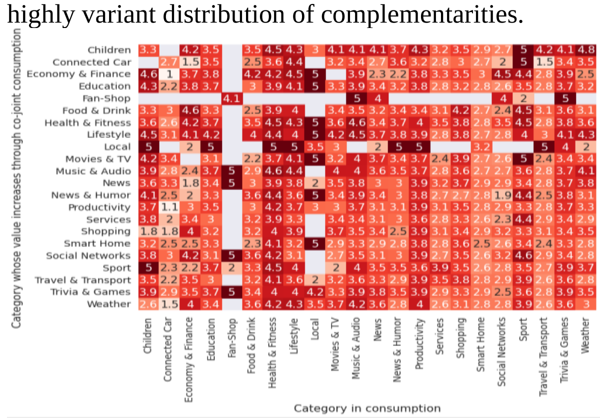

Figure 6. Supermodular Complementarities in Consumption

8. Discussion and Conclusion

Existing ecosystem intelligence approaches, such as those suggested by Basole (2020) and Battistella et al. (2013) have proven valuable in understanding the complexity of ecosystems, which has also been recognized as an important field in information systems research (see interview in (Alt, 2022)). However, they are limited to ecosystems where the modules' functionality is abstracted into one category of functionality. In ecosystems of composite platforms, they fail to differentiate relationships according to the category of the underlying functionality.

With the diffusion of the platform model, a growing number of composite platforms enable the creation of modules integrating different categories of functionality. For example, assistant platforms feature a composite architecture that enables the creation of modules that belong to different categories of functionality or may even mix them. The same applies for many other platforms such as the Google (Google. Inc, 2011) and the Apple

App Store (Roma & Ragaglia, 2016). We, therefore, expect that the findings from this research may be applied to a wide spectrum of ecosystems.

8.1. Contribution

This paper pursued two research questions to shed more light on highly complex ecosystems. The first addresses the fundamental representation of complementarities, which are recognized as key elements for ecosystems in the extant literature. These complementarities are particularly challenging in complex ecosystems, where multiple cross-effects among modules are possible. Complementarity and category-oriented augmented complementarity graphs depict ecosystems in greater detail and more precisely than existing ecosystem intelligence approaches. In particular, complementarity graphs consider that actors may provide multiple modules of different functionality. They employ modules as vertices instead of actors, and they classify the relationships between modules based on the complementarity framework of Jacobides et al. (2018). Multiple edges may represent different types of complementarities, and the edges are directed to show asymmetric relationships. By augmenting complementarity graphs with category information, it is possible to differentiate the depiction of an ecosystem in two dimensions. First, this allows the analysis of different types of complementarities. Second, subgraphs may be drawn to differentiate the complementarities according to the modules they are created from.

Another important challenge of ecosystem intelligence is obtaining the data for analysis. It has been argued that neither the platform operator nor the module vendors are often not interested in releasing data essential for the ecosystem. The suggested process for collecting ecosystem data is conceptually based on creating proxies that replace data not released by the platform operator and module vendors. The process obtains the proxies by analyzing the data released by the platform operators to support platform matching, which is the meta-information about the modules.

We have shown how complementarity maps can be formed to demonstrate the application of the concepts we have developed. Category-oriented complementarity maps provide important information to platform operators and module vendors to answer questions such as: How strongly does a module benefit from being consumed with others? Which categories of modules should we incentivize? In addition, category-oriented complementarity maps could also predict the strength of complementarities for a module to be developed and address questions such as: How strong are the complementarities to be expected by the yet to be designed module? The application of this approach in the Alexa ecosystem suggests that such insights may be obtained even in complex ecosystems. We expect that it also applies to other digital platform ecosystems.

8.2. Research implications

Information systems research has long relied on manual data collection when direct access to information system properties via interfaces is impossible. Extensive studies of information systems are based on interviews with experts and surveys. In addition to involving time and effort, this approach assumes that the experts are real experts and are free of bias. However, expert selection faces the paradox that one would have to be an expert to select experts. Large numbers of human participants could reduce the influence of individual bias, but this approach increases the risk that also non-experts are participating in the study. We intend to avoid the limitations of human intermediators and directly access data that reflects the properties and structures of information systems.

The suggested approach has connections to methodological research on using big data (e.g. (Rizk et al., 2020) (Kar & Dwivedi, 2020)) or using web scraping for information systems research (Boegershausen et al., 2021). In addition, it contributes to the further development of computationally intensive theory building (Miranda et al., 2022), which is part of the computational detection of scientific knowledge (Džeroski et al., 2007). Our proxy-oriented and web scraping data collection process provides the foundation for synchronic and diachronic studies, for example, the development of ratings. This includes the development of complexity metrics and the measurement of the complexity of ecosystems. Finally, the combination of web scraping and ecosystem intelligence supports faster decision-making by replacing the manual steps and including more up-to-date data. Such shorter turnaround times make the approach particularly interesting for dynamic information systems and may open new avenues for information systems research (Baskerville et al., 2020).

8.3. Management implications

Complementarity graphs and category-oriented complementarity maps offer more detailed insights for platform providers. They now see their ecosystem's complementarity hotspots and trouble spots in greater detail. Based on these insights, platform providers can create incentives to foster underperforming module categories and drive platform value by identifying the most value-creating complementarities.

Category-oriented complementarity graphs and maps provide key information to platform and module providers. Platform providers may identify categories of modules that provide strong complementarities and use this information to increase platform value by fostering those module categories. Module vendors can identify modules with strong complementarities and use this information to promote them. Module providers are enabled to map their strategic positioning. They obtain insight into competing modules and the potential for creating complementarities and thus support for their modules. Module providers use the information from complementarity maps to identify functionalities that create the strongest complementarities. In the same way, they learn which functionalities deem attention due to their weak complementarities.

8.4. Limitations and future research

Although the presented approach promises advantages over traditional approaches in general ecosystem intelligence and information systems research, we also see some limitations. For example, bias may occur in the ratings and reviews. Despite a sufficient number of ratings and reviews can reduce this risk, it cannot be excluded. Furthermore, the sample size may not reach this level for all modules. Another limitation is the representation of the data that may obstruct insights.

A key area of further research will be comparing and analyzing complementarity graphs and maps. First, comparing complementarity maps and graphs of different ecosystems may yield important insights into the strengths and weaknesses of ecosystems. This is of particular relevance for ecosystems that feature higher levels of complexity. Second, comparing the complementarities within an ecosystem may provide insights into correlations. Third, the analysis of category maps on correlations between categories of functionality may also provide critical insights. To support these tasks, we aim for a tool that automates the analyses and visualizes the data using a dashboard to identify relationships and depict all stakeholders' developments over time. Ultimately, we expect to see methods that describe heterogeneity and complexity in ecosystems and explain and calculate these important properties.

9. References

Alexa Developer Documentation. (2021). Amazon (Alexa). https://developer.amazon.com/en-US/alexa/alexa-developer-documentation-welcome.html

Alexa Skills. (2021). Alexa Skills. https://www.amazon.de/skills

Alt, R. (2021). Electronic Markets on platform complexity. *Electronic Markets*, *31*(4), 737–742. https://doi.org/10.1007/s12525-021-00517-9

Alt, R. (2022). Managing AI is managing complexity—An interview with Rahul C. Basole. *Electronic Markets*. https://doi.org/10.1007/s12525-022-00585-5

Amazon Alexa: Skill Count in Selected Countries 2021. (2021). Statista. https://www.statista.com/statistics/917900/selected-countries-amazon-alexa-skill-count/

Baldwin, C. Y., & Clark, K. B. (2000). *Design Rules: The Power of Modularity*. MIT Press.

Baskerville, R. L., Myers, M. D., & Yoo, Y. (2020). Digital First: The Ontological Reversal and New Challenges for Information Systems Research. *MIS Quarterly*, *44*(2), 509–523. https://doi.org/10.25300/MISQ/2020/14418

Basole, R. (2020). Understanding Ecosystem Data. *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 5718–5727. https://doi.org/10.24251/HICSS.2020.702

Basole, R., Huhtamäki, J., Still, K., & Russell, M. G. (2016). Visual Decision Support for Business Ecosystem Analysis. *Expert Systems with Applications*, *65*, 271–282. https://doi.org/10.1016/j.eswa.2016.08.041

Basole, R., Mengdie Hu, Patel, P., & Stasko, J. T. (2012). Visual Analytics for Converging-Business-Ecosystem Intelligence. *IEEE Computer Graphics and Applications*, *32*(1), 92–96. https://doi.org/10.1109/MCG.2012.27

Basole, R., & Park, H. (2019). Interfirm Collaboration and Firm Value in Software Ecosystems: Evidence From Cloud Computing. *IEEE Transactions on Engineering Management*, *66*(3), 368–380. https://doi.org/10.1109/TEM.2018.2855401

Basole, R., Russell, M. G., Huhtamäki, J., Rubens, N., Still, K., & Park, H. (2015). Understanding Business Ecosystem Dynamics: A Data-Driven Approach. *ACM Transactions on Management Information Systems*, *6*(2), 1–32. https://doi.org/10.1145/2724730

Battistella, C., Colucci, K., De Toni, A. F., & Nonino, F. (2013). Methodology of business ecosystems network analysis: A case study in Telecom Italia Future Centre. *Technological Forecasting and Social Change*, *80*(6), 1194–1210. https://doi.org/10.1016/j.techfore.2012.11.002

Benali, M., & Burlat, P. (2012). Modeling cooperation links within networks of firms: A case study. *Journal of Intelligent Manufacturing*, *23*(6), 2451–2457. https://doi.org/10.1007/s10845-011-0511-y

Boegershausen, J., Borah, A., Datta, H., & Stephen, A. T. (2021). Fields of Gold: Generating Relevant and Credible Insights Via Web Scraping and APIs. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3820666

Brynjolfsson, E., & Smith, M. D. (2000). Frictionless commerce? A comparison of Internet and conventional retailers. *Management Science*, *46*(4), 563–585.

Chung, H., Park, J., & Lee, S. (2017). Digital forensic approaches for Amazon Alexa ecosystem. *Digital Investigation*, *22*, 15–25. https://doi.org/10.1016/j.diin.2017.06.010

Codd, E. F. (2001). A Relational Model of Data for Large Shared Data Banks. In *Pioneers and Their Contributions to Software Engineering* (pp. 61–98). Springer.

Cusumano, M., Gawer, A., & Yoffie, D. B. (2020). The Future of Platforms. *MIT Sloan Management Review - Special Issue on Disruption*, *61*(3), 46–54.

de Reuver, M., Sørensen, C., & Basole, R. (2018). The Digital Platform: A Research Agenda. *Journal of Information Technology*, *33*(2), 124–135. https://doi.org/10.1057/s41265-016-0033-3

Džeroski, S., Langley, P., & Todorovski, L. (2007). Computational Discovery of Scientific Knowledge. In S. Džeroski & L. Todorovski (Eds.), *Computational Discovery of Scientific Knowledge* (Vol. 4660, pp. 1–14). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-73920-3_1

Garg, R., & Telang, R. (2013). Inferring App Demand from Publicly Available Data. *MIS Quarterly (MISQ)*, *37*(4), 1253–1264.

Gawer, A., & Cusumano, M. A. (2014). Industry Platforms and Ecosystem Innovation. *Journal of Product Innovation Management*, *31*(3), 417–433. https://doi.org/10.1111/jpim.12105

Google. Inc. (2011). *Google Apps Marketplace*. https://www.google.com/enterprise/marketplace/

Gunawan, R., Rahmatulloh, A., Darmawan, I., & Firdaus, F. (2019). Comparison of Web Scraping Techniques: Regular Expression, Html Dom and Xpath. *ICoIESE 2018*, 283–287.

Infographic: The Growing Footprint of Amazon's Alexa. (2020). Statista Infographics. https://www.statista.com/chart/22338/smart-home-devices-compatible-with-alexa/

Jacobides, M. G., Brusoni, S., & Candelon, F. (2021). The Evolutionary Dynamics of the Artificial Intelligence Ecosystem. *Strategy Science*, *6*(4), 412–435. https://doi.org/10.1287/stsc.2021.0148

Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a Theory of Ecosystems. *Strategic Management Journal*, *39*(8), 2255–2276. https://doi.org/10.1002/smj.2904

Jacobides, M. G., Cennamo, C., & Gawer, A. (2020). *Distinguishing between Platforms and Ecosystems: Complementarities, Value Creation, and Coordination Mechanisms*. https://www.jacobides.com/_files/ugd/0b15b1_7e0678d2815541bb920c77ec41e4d305.pdf

Järvi, K., & Kortelainen, S. (2017). Taking Stock of Empirical Research on Business Ecosystems: A Literature Review. *International Journal of Business and Systems Research*, *11*(3), 215–228. https://doi.org/10.1504/IJBSR.2017.085469

Johannesson, P., & Perjons, E. (2021). *An Introduction to Design Science*. Springer International Publishing. https://doi.org/10.1007/978-3-030-78132-3

Jussila, J., Kärkkäinen, H., Kortelainen, S., Huhtamäki, J., Aho, T., & Tebest, T. (2014). New Era of Business Analytics: Making Sense of Business Ecosystems. *Proceedings of the 18th International Academic MindTrek Conference*, 266–268. https://doi.org/10.1145/2676467.2676517

Kar, A. K., & Dwivedi, Y. K. (2020). Theory Building with Big Data-Driven Research – Moving Away from the "What" Towards the "Why". *International Journal of Information Management*, *54*, 102205. https://doi.org/10.1016/j.ijinfomgt.2020.102205

Katz, M. L., & Shapiro, C. (1994). Systems Competition and Network Effects. *Journal of Economic Perspectives*, *8*(2), 93–115. https://doi.org/10.1257/jep.8.2.93

Kinne, J., & Axenbeck, J. (2020). Web Mining for Innovation Ecosystem Mapping: A Framework and a Large-Scale Pilot Study. *Scientometrics*, *125*(3), 2011–2041. https://doi.org/10.1007/s11192-020-03726-9

Landers, R. N., Brusso, R. C., Cavanaugh, K. J., & Collmus, A. B. (2016). A Primer on Theory-Driven Web Scraping: Automatic Extraction of Big Data from the Internet for Use in Psychological Research. *Psychological Methods*, *21*(4), 475–492. https://doi.org/10.1037/met0000081

Miranda, S., Berente, N., Seidel, S., Safadi, H., & Burton-Jones, A. (2022). Computationally Intensive Theory Construction: A Primer for Authors and Reviewers. *MIS Quarterly*, *46*.

pandas—Python Data Analysis Library. (2021). https://pandas.pydata.org/

Pidun, U., Reeves, M., & Wesselink, E. (2021). How Healthy is Your Business Ecosystem? *MIT Sloan Management Review*, *62*(3), 31–38.

Project Jupyter. (2019). Project Jupyter. https://www.jupyter.org

Rizk, A., Ståhlbröst, A., & Elragal, A. (2020). Data-driven innovation processes within federated networks. *European Journal of Innovation Management*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/EJIM-05-2020-0190

Roma, P., & Ragaglia, D. (2016). Revenue models, in-app purchase, and the app performance: Evidence from Apple's App Store and Google Play. *Electronic Commerce Research and Applications*, *17*, 173–190. https://doi.org/10.1016/j.elerap.2016.04.007

Schmidt, R., Alt, R., & Zimmermann, A. (2021). A Conceptual Model for Assistant Platforms. *54th Hawaii International Conference on System Sciences (HICSS)*, 4024–4033. https://doi.org/10.24251/HICSS.2021.490

Schmidt, R., Alt, R., & Zimmermann, A. (2022). Ecosystem Intelligence for AI-based Assistant Platforms. *55th Hawaii International Conference on System Sciences (HICSS)*, 4316–4325. https://doi.org/10.24251/HICSS.2022.527

Schmidt, R., Kirchner, K., & Razmerita, L. (2022). *Leveraging Emergent Social Interactions for Value Co-Creation on Transaction Platforms*. 3317–3326. https://doi.org/10.24251/HICSS.2022.406

Scholten, U. (2013). *Dynamic network notation: A Graphical modeling language to support the visualization and management of network effects in service platforms*. KIT Scientific Publishing.

Srinivasan, R. (2021). Network Effects. In R. Srinivasan, *Platform Business Models* (pp. 35–44). Springer Singapore. https://doi.org/10.1007/978-981-16-2838-2_3

Vailshery, L. (2021). *Global Smart Speaker Market Share 2020*. Statista. https://www.statista.com/statistics/792604/worldwide-smart-speaker-market-share/