

# An Empirical Investigation of Enterprise Architecture Analysis Based on Network Measures and Expert Knowledge: A Case from the Automotive Industry

Alixandre Santana,<sup>1</sup> Matthias Kreimeyer,<sup>2</sup> Pascal Clo<sup>2</sup>,  
Kai Fischbach,<sup>3</sup> Hermano de Moura<sup>1</sup>

<sup>1</sup> Informatics center, UFPE, Brazil

<sup>2</sup> MAN Truck and Bus AG, Germany

<sup>3</sup> University of Bamberg, Germany

**Abstract:** Enterprise architecture (EA) may be considered an organizational blueprint that helps experts manage organizational complexity. In this regard, EA analysis is an emerging field gaining greater attention, and considering EA as an intertwined system of components and relationships and performing EA analysis from a structural perspective are promising areas of research. This paper analyzes EA data from a German commercial vehicle manufacturer, modeling a subset of its EA with the help of design structural and domain mapping matrices. Thus, we propose an analysis approach based on network measures that uses structural knowledge generated by the network analysis to validate or refine experts' tacit knowledge about EA key components from different layers. We refer to this approach as the diagnosis analysis method. Based on our results, we successfully combine the structural knowledge with expert knowledge and provide useful validations for experts.

*Keywords:* Enterprise architecture, network analysis, DSM modeling

## 1 Introduction

Enterprise architecture (EA) can be considered a blueprint that describes a general organization in terms of its components to help experts manage organization complexity. Management of all these components (e.g., goals, business process, applications, and IT infrastructure), their interdependencies, and their evolution allows organizational changes to be coordinated and aligned with mid- to long-term company objectives (Ahlemann, 2012). In companies that deal with a vast set of applications supporting several business processes, the task of gleaning additional value from current EA models (EA analysis) is made even more complex by the lack of suitable analysis tools (Santana et al., 2016).

EA analysis has attracted researcher attention in the last decade (Santana et al., 2016). The literature includes several different paradigms for EA analysis, such as probabilistic relation models (Buschle et al., 2010), EA intelligence (Veneberg et al., 2014), complexity management (Schneider et al., 2014), ontology-based analysis (Antunes et al., 2013) and network-based analysis (Dreyfus and Wyner, 2011).

Considering EA as an intertwined system of layers, components, and relationships and performing EA analysis from a structural perspective holds promise. However, the application of network measures in EA analysis still has considerable room to grow, according to Santana et al. (2016) and Simon and Fischbach (2013). A similar challenge

of analyzing interdependent components has also been faced in other fields such as system engineering and product engineering. In these areas, the literature has emphasized the importance of the design structure matrix (DSM). The principles on which DSM and other similar methods are based have proven to be valuable and viable in numerous applications (Kreimeyer et al., 2009). Related research also reports the application of the DSM through the use of strength-based graphs and algorithms from network theory (Kreimeyer et al., 2009).

In this paper, we take a similar approach to analyze EA data from a large commercial vehicle manufacturer located in Germany, modeling a subset of its EA with the help of DSMs and applying network measures. In the end, we offer two contributions: First, we frame an empirical subset of EA data as DSMs and apply matrix transformations to this subset, deriving a co-affiliation network. We then apply network measures to identify key components in the primary and derived networks. With that, we expect to foster the discussion for applying DSM to EA analysis with primary and derived data. Second, we propose an analysis approach based on network measures to be applied to our data, modeled as primary and derived DSMs. We take the expert knowledge and compare it with the structural information generated by the network measures in order to validate and/or refine experts' perceptions about key EA components. We refer to this approach as the *diagnosis analysis method*. We demonstrate the use of this approach, aiming to answer the following research question: *How can expert and structural knowledge about EA components be combined to help an expert perform EA analysis?*

This paper is organized as follows: Section 2 presents key concepts discussed. Our research design is described in Section 3. Section 4 presents results and our analysis. Section 5 presents our conclusions.

## 2 Key concepts

This section is a short introduction to the topics covered in the paper: enterprise architecture, EA analysis, the DSM and multiple-domain matrix (MDM), and EA network analysis. Related works are also detailed in this section.

### 2.1 Enterprise architecture

EA is defined in a variety of ways. We adopt the definition proposed in the literature review of Schütz et al. (2013), which is also supported by Open Group (2011). According to that definition, EA is a system formed by four subsystems: business (or business layer), data (or information layer), application (or application layer), and infrastructure (or technology layer). In this paper, we model these EA subsystems (which might also be called EA layers or architectures) as networks/graphs.

### 2.2 Enterprise architecture analysis

EA can be considered an organizational blueprint composed of the four layers above. Creating such a blueprint is worthwhile only if resultant models can add value to the architectural decision-making process, help in managing organizational complexity, and lower risks (Naranjo et al., 2014). As part of the broader EA management lifecycle, EA analysis initiatives might target different concerns such as EA domain redesign, application support to business processes support, identification of misalignment of resources, EA decision making, and so on. These analysis initiatives may have different

degrees of abstraction and coverage, ranging from a full-edged impact analysis over the entire model to an in-depth analysis of a specific domain (Naranjo et al., 2014).

### **2.3 Modeling EA as a complex network**

In the context of network theory, according to Scott (1992), a graph or network is a set of components (nodes or vertices) and links (edges or relations). The use of graph theory and network measures to analyze single software systems extends back to the 1980s (Hall and Preiser, 1984). Clustering algorithms, a common approach found in software modularity studies (group or cluster analysis), was introduced in the EA context by Aier and Winter (2009) to identify EA virtual domains and thus aid in EA redesign. However, analysis measures at the individual level—such as eigenvector and degree centrality—appear more frequently in EA research than does clustering, as shown in Santana et al. (2016).

When modeling EA as a set of networks or layers, nodes may represent different components. In the application layer, for instance, nodes might represent applications that support business functions that integrate the business layer (Simon and Fischbach, 2013). Nodes in the technology layer can represent IT infrastructure components such as application servers. Edges represent relationships and interdependencies between applications. In general, these relationships can take different forms (Simon and Fischbach, 2013). For example, links between application components in the application layer may indicate the same vendor or a data flow. These modeling choices for nodes and relationships are closely related to the EA concerns one may wish to analyze (Wasserman and Faust, 1994).

### **2.4 DSM, MDM**

If the goal of architecture is to provide flexibility, robustness, and adaptability, then any change to the architecture must be managed to minimize its complexity (Schmidt, 2013). Schütz et al. (2015) define EA complexity as having two dimensions: structural (interdependence of the elements) and heterogeneity. In this paper, we focus on the first of these two.

The DSM-based methods have been well established for many years, and are typically applied to systems engineering, product architecture, and so on, to analyze aspects such as dependency and modularity (Kreimeyer et al., 2009). We find, however, only a few applications of DSM in the context of enterprise architecture (Lagerström et al., 2013).

According to the literature review of Santana et al. (2016), research that considers DSM modeling in EA analysis is scarce; the exceptions are the works of Lagerström et al. (2013) and Lagerström et al. (2014). In the view of Lagerström et al. (2013), “interestingly, many of the problems encountered by software architects dealing with a single software system are similar to those that occur for enterprise architects on a system-of-systems level.” The authors took the DSM expertise from the analysis of single software architectures in Baldwin et al. (2013) and applied it to EA. In the end, Lagerström et al. (2013) proposed the “hidden structure method” to classify EA components into four categories according to their position in the network. Later, Lagerström et al. (2014) used their previous method and adding a correlation analysis

between the position occupied by the component in the network and the cost of change propagation and architecture flow.

The multiple-domain matrix (MDM) is an extension of the DSM to model entire systems consisting of multiple domains, each having multiple elements and connected by various relationship types. Bartolomei et al. (2012) present an MDM representing several domains of systems engineering for modeling large-scale, complex systems projects. They consider six component classes belonging to five domains. Hollauer et al. (2015) also propose an MDM that supports dynamic modeling of sociotechnical systems consisting of seven domains.

In this paper, we advocate and reinforce the use of EA as an organizational blueprint, together with DSM and MDM modeling and constructs taken from network theory, to build a toolbox to perform EA structural analysis. One possible application that arises immediately is to use an MDM to model EA and derive single DSMs to explore new analysis perspectives.

### 3 Methodological aspects

This work can be classified as exploratory and applied research. We have adopted a design science research approach (Hevner et al., 2004).

In practice, EA analysis depends critically on human cognitive abilities (e.g., expertise) (Hevner et al., 2004) to produce effective results (Simon and Fischbach, 2013). We believe that, particularly in organizations with dozens or even hundreds of business processes supported by a similar number of applications, one might want to use additional knowledge sources to add confidence to the analysis. As our design artifact, we develop a method to combine expert knowledge (subjective by nature) about EA components (critical business units (BUs), business process (BPs), and business objects (BOs)) with network measures outputs (structural criteria). Figure 1 depicts this approach.

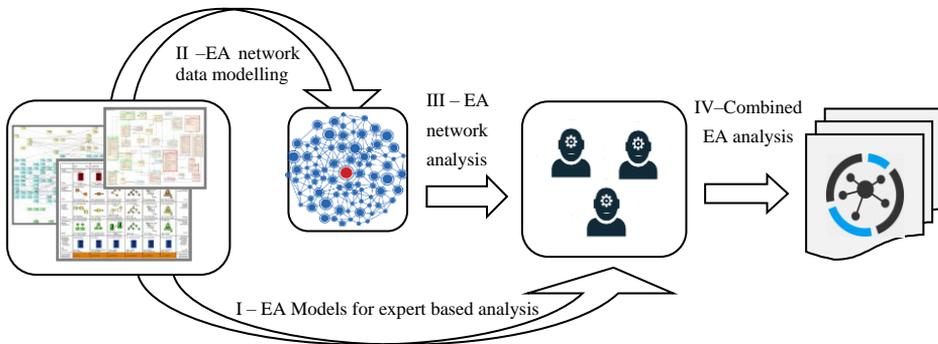


Figure 1: Combined EA analysis. Cycle I: EA models are analyzed with expert-based techniques; Cycle II: EA models are converted into EA network data; Cycle III: EA network analysis is performed using network measures; Cycle IV: Combined EA analysis

We believe this approach allows for designing a more robust method for EA analysis. We validated the artifact with a real-world example, using data from a business unit of a German automotive company with global operations. The company (henceforth referred

to as “Autocompany”) has multibillion-dollar revenues and more than 40,000 employees around the world.

To model our networks, we preprocessed documents made available by Autocompany that were complemented with a round of two meetings to model the data properly. Table 1 describes the subset of EA layers analyzed. Components such as application (from application layer) and technology (from technology layer) are not discussed in this paper due to space limitations.

We used these data to build the primary or original networks (i.e., BOxBO, BPxBP). The primary network was created based on the data provided by Autocompany. We also worked with a second category of data, so-called derived network data. For instance, the BUxBU is a network (or DSM) derived from the BUxBP network (or a MDM) by the co-affiliation mechanism described in Borgatti and Halgin(2016).

Table 1. Dataset description of our case

EA component	EA layer	Amount of components in dataset	Network model
Business unit	Business	15	BUxBU
Business process step	Business	101	BPxBP
Business object	Information	70	BOxBO

According to Scott (1992), different network measures can be used as proxies for various structural concepts. We use the set of measures described in Table 2.

Table 2. Network measures used, and their contextualization

Network measures	Meaning in the context of BPxBP network
Betweenness centrality	BPs that are important intermediary channels of information
In-Closeness centrality	BPs can be reached easily from other nodes (in our case, BPs that are common destinations of information flow)
Eigenvector	BPs connected with other significant (well-connected) BPs; these are structural nodes in the network of BPs
In-degree centrality	BPs that receive several inputs from other BPs
Out-degree	BPs that provide several inputs for other BPs
Total degree	BPs with high total degree centrality represent BPs that interact directly with several other BPs

To capture this diversity of concepts, we define a majority voting strategy to select the “Y” most recurrent outliers from our voting committee based on the “X” outliers identified by each network measure. With this voting strategy, we consider the TOP 15 (X=15) outliers of each of the six measures in Table 2. We then compute the most “cited” outliers among them to build a ranking containing the “Y” most recurrent components, as depicted in Figure 2. The “Y” and “X” parameters are adjusted *ad hoc* by the experts (one might want to analyze the 10, 20, or 30 most recurrent outliers among the TOP X outliers of each measure). This ranking represents a synthesis of the most significant outlier components in terms of structure.

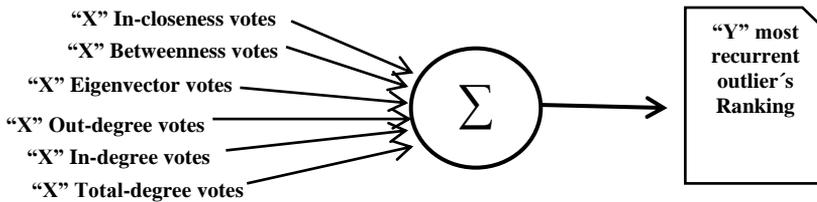


Figure 2. Voting strategy to build the outliers' ranking

We can use this information in different ways. The most obvious one is a context-independent analysis to identify the components that play important roles in terms of structure. We propose, thus, a second possibility, which is to combine this structural knowledge with expert's tacit knowledge (*a priori* knowledge). We call this *diagnosis analysis*. We are interested in validating the expert knowledge with the components that are true positives (i.e., components identified as critical both by network measures and the experts) and potentially providing new information for the experts with the false positives (i.e., components identified as critical only by the network measures).

We performed this analysis for the BUxBU, BPxBP, and BOxBO networks, aiming to answer the following research question: *How can expert and structural knowledge about EA components be combined to help an expert perform EA analysis?* Section 4 presents our results.

## 4 Results and discussion

In this section, present the three aspects of our network analysis: BPxBP, BOxBO, and BUxBU.

### 4.1 BPxBP network analysis

With this analysis, we aimed to answer the following question: *Can the business process outliers (central points of the network BPxBP) also be identified as elements of the critical path defined by experts?* Our EA analysis concern, then, was to identify key structural components of an EA layer (in this instance, the business process layer). We used as input for this analysis the BPxBP network (primary data) extracted from Autocompany's documents.

Our hypothesis H1 is that network measures will be able to identify the main components belonging to the critical path already defined by Autocompany's experts. Additional components will also be identified and may have their importance validated.

The experts identified seven BP components in the critical path (the data had to be anonymized). Following the algorithm of the diagnosis analysis method defined in Section 3, we selected the TOP 15 outliers generated for each of six network measures. We then took the most recurrent components among all measures, using the voting strategy. This resulted in the selection of 21 distinct components (BPs), among which it was possible to identify successfully, from a universe of 102 BPs, the seven BPs that constitute the critical path defined in Autocompany's documents. This selection also included components considered for further analysis by Autocompany's experts, who classified all of them as important BPs. As one expert remarked about these



experts but is identified as important due to local importance of the components (high number of in-connections or out-connections to other local BOs) and may be worth of the experts' attention. We claim that these results support H2.

### 4.3 BUxBU network analysis

We consider organization business units (BUs) as stakeholders that execute different workloads depending on each process phase. We consider two phases for analysis purposes: PROD (phase I) and KONZ (phase II), and we pose this question: *Can we identify key stakeholders in different process phases?* Thus, our analysis concern here is stakeholder management, which might be important when it comes to involving the right people (BUs) in the EA decision-making process. As input, we took the BUxBP and applied a matrix transformation to generate a derived BUxBU network. With this derivation method, if two processes BP1 and BP2 are connected an artificial connection is created between their respective BUs in the BUxBU network (a co-affiliation network). The hypothesis H3 formulated by the experts is then broken down as follows. H3.1: *In Phase I, the focus of the project management unit (BU1) should be fairly continuous as they manage all activities.* H3.2: *In Phase II, the focus will be more on the technology people, with a ramp up to production and logistics and possibly to purchasing.* H3.3: *Overall, in Phase II, design engineers will be fairly central, as they function as a sort of information hub around which all technical concept design focus.*

We combined two types of network analysis outputs: the network measures rankings described in Figure 2 and the heat maps depicted in Figure 4. With the heat map, it was possible to check the high intensity of the information flow from B1, BU5, and BU7, which is spread out in Phase I. There was intense activity inside BU1, as can be seen in the dark blue cell, confirming the importance of BU1 for Phase I (thus supporting H3.1). Figure 4 also suggests a strong interaction from BU5 to BU1. This might confirm that both BUs together are the most active BUs in Phase I in terms of process interactions. From Table 3, we notice that BU1, BU5, BU2, and BU7 also appear in different rankings of network measures, reinforcing our visual analysis results from Figure 3.

Table 3. Network analysis at the component level for BUxBU Prod (Phase I)

TOP Out-degree BUs		TOP Eigenvector BUs		Most recurrent BUs	
BU5	product management	BU1	project management	BU1	Project management
BU1	project management	BU5	product management	BU13	Prod. Preparation
BU7	total vehicle integration	BU2	controlling	BU2	Controlling
BU3	quality	BU7	total vehicle integration	BU5	Prod. management
BU2	control	BU13	production preparation	BU7	Total vehicle integration

For H3.2, we obtained the following results: BU13, responsible for production aspects, became imperative in Phase II (detected by high in-degree centrality components and eigenvector centrality); Purchasing (B15) had importance detected by high in-degree centrality and also was among the most recurrent outliers; Validation and integration (BU7, BU10, BU3) aspects received focus in Phase II, detected by eigenvector centrality, most recurrent outliers, and out and in-degree centralities. Although identified

by the experts *a priori*, logistics did not appear as a focus in Phase II. In conclusion, we found H2.3 to be partially supported.

For hypothesis H3.3, integration, validation, and preparation for production activities were the main ones in Phase II. So, we can conclude that H3.3 was also supported.

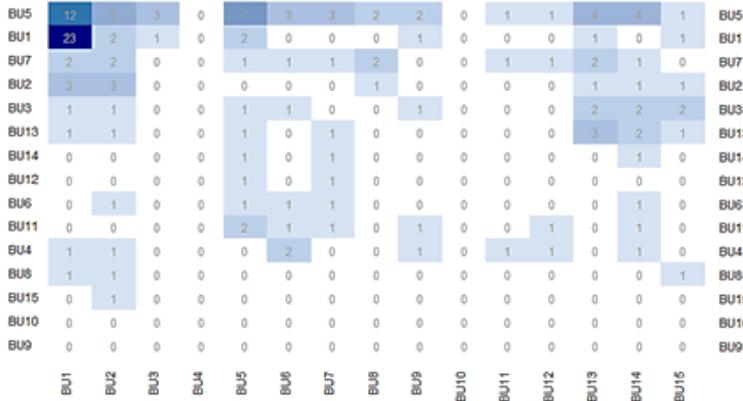


Figure 4. Heat map for BUxBU Prod phase

## 5 Conclusions

In this research, we propose an analysis approach we call *diagnosis analysis* that we believe provides two key information gains: 1) confirmation of components' importance, including from the structural perspective (confirming what experts identify prior to analysis as the critical BPs and BOs), and 2) suggesting for further analysis other components with similar network values and labeled as important by network measures, at first neglected by experts (based only on their own opinions), and ultimately validated as important by the experts. We also detected the expected shift of BUs' focus along the two process phases. In the end, we showed that combining expert and structural knowledge (the latter provided by primary and derived data) is a useful tool to assist experts in EA analysis.

There are some limitations to our research. First, the data collection and modeling processes were manual; future research might benefit considerably from use of a software plugin that can convert data from architecture models to network representations. Second, we analyzed only two EA layers. Other EA MDMs and DSMs (primary and derived) must be explored. Finally, we need to apply the proposed approach in other organizations and to other EA concerns to test whether it can be generalized.

We agree with Lagerström et al. (2013; 2014) that DSM and network analysis-based methods should be explored further in the context of enterprise architecture. Thus, we are working on the development of a matrix-based framework to support EA modeling with DSMs, including their possible co-affiliation networks.

## References

- Ahlemann, F. Strategic Enterprise Architecture Management: Challenges, Best Practices, and Future Developments. DOI: 10.1007/978-3-642-24223-6.
- Aier, S. and Winter, R. "Virtual Decoupling for IT/Business Alignment – Conceptual Foundations, Architecture Design and Implementation Example", *Business & Information Systems Engineering*, vol. 1, no. 2, pp. 150-163, 2009.
- Antunes, G., Bakhshandeh, M., Mayer, R., Borbinha, J. and Caetano, A. Using ontologies for enterprise architecture analysis, in *Proc. EDOCW*, 2013, pp. 361–368.
- Bartolomei, J. E., Hastings, D. E., de Neufville, R. / Rhodes, D. H. in *Systems engineering -new York-*; 15, 1; 41-61; *systems engineering -New York-*by John Wiley & Sons, Ltd; 2012
- Borgatti, S. and Halgin, S. Analyzing Affiliation Networks. Available: <http://www.behr.ufl.edu/sites/default/files/Analyzing%20affiliation%20networks.pdf>
- Buschle, M., Ullberg, J., Franke, U., Lagerström, R. and Sommestad, T., A tool for enterprise architecture analysis using the PRM formalism. *LNBIP*, v. 72 pp. 108–121.
- Dreyfus, D. and Wyner, G.M. "Digital Cement: Software Portfolio Architecture, Complexity, and Flexibility". *Proc. 17th AMCIS*, Detroit, Michigan, USA, 2011
- Hall, N. R. and Preiser, S. Combined Network Complexity Measures. *Ibm Journal of Research and Development (Impact Factor: 0.69)*. 02/1984; 28(1):15 - 27. DOI: 10.1147/rd.281.0015
- Hevner, A., March, S., Park, J., Ram, S.: *Design Science in Information Systems Research*. *MIS Quarterly*, v. 28, p. 75–105, 2004
- Hollauer, C., Wilberg, J., Mayada, O. A matrix-based framework to support dynamic Modeling of sociotechnical systems. In: *Proc. of the 17 DSM Conference* 2015.
- Kreimeyer, M., Maier, J., Fadel, G. and Lindemann, U, editor, *Proceedings of the 11th International DSM Conference*, Greenville, SC, 12 - 13 October 2009, page 369-381.
- Lagerström, R., Baldwin, C.Y. , MacCormack, A. and Aier, S, *Visualizing and Measuring Enterprise Application Architecture: An Exploratory Telecom Case*". Harvard Business School Working Paper, no. 13-103, 2013.
- Lagerström, R., Baldwin, C., MacCormack, A., and Aier, S. 2014. *Visualizing and Measuring Enterprise Application Architecture: An Exploratory Telecom Case*. In *Proc. of the HICSS*
- Naranjo, D., Sánchez, M.E and Villalobos, J: *PRIMROSE: A Graph-Based Approach for Enterprise Architecture Analysis*. *ICEIS (Revised Selected Papers) 2014*, p. 434-452
- Santana, A, Fischbach, K and De Moura, H, "Enterprise Architecture Analysis and Network Thinking : a Literature Review." In *Proc. HICSS 2016*.
- Schmidt, C "Measuring and Managing Enterprise Architecture Complexity".(online). Available: [http://scape-consulting.de/tl\\_files/scape/publications/CALM3\\_Scape.pdf](http://scape-consulting.de/tl_files/scape/publications/CALM3_Scape.pdf)
- Schneider, A. W. , Zec, M. and Matthes, F. "Adopting Notions of Complexity for Enterprise Architecture Management". *Proc. AMCIS*, 2014.
- Schütz, A., Widjaja, T. and Kaiser, J., Complexity in enterprise architectures: conceptualization and introduction of a measure from a system theoretic perspective, In *Proc. of 21st ECIS*
- Scott, J. *Social Network Analysis*. Newbury Park, CA: Sage, 1992.
- Simon, D. and Fischbach, K. "IT Landscape Management Using Network Analysis." In G. Poels (ed.), *LNBIP (Vol. 139)*, Springer Berlin Heidelberg, pp. 18–34, 2013
- The Open Group (2011). *TOGAF Version 9.1*. Van Haren Publishing.
- Veneberg, R. K. M., Iacob M., van Sinderen, M. J. and Bodestaff, L. "Enterprise Architecture Intelligence: Combining Enterprise Architecture and Operational Data". *Proc. EDOC*, pp. 22–31.
- Wasserman, S. and Faust, K.. *Social Network Analysis*. Cambridge: Cambridge University Press.1994.

**Contact:** Aloxandre Santana, UFPE, Informatics Center, Recife, Brazil, [atfs2@cin.ufpe.br](mailto:atfs2@cin.ufpe.br)

**About the Authors:**



**Alixandre Santana, UFPE, Brazil** (atfs2@cin.ufpe.br) – Alixandre is concluding his PhD degree in computer science at the informatics center, federal university of Pernambuco -UFPE, Brazil. He also works as assistant professor at UFRPE, Brazil, since 2010.



**Matthias Kreimeyer, MAN Truck & Bus AG** (matthias.kreimeyer@man.eu) – Dr. Matthias is head of product strategy & product management truck at MAN Truck & Bus AG.sfdfsdfs



**Pascal Clo, MAN Truck & Bus AG** (pascal.clo@man.eu) - Pascal Clo works as product architect at MAN Truck & Bus AG



**Kai Fischbach, University of Bamberg** (kai.fischbach@uni-bamberg.de) – Dr. Kai is chair of business informatics and social network analysis at University of Bamberg.



**Hermano de Moura, UFPE** (hermano@cin.ufpe.br) – Dr. Hermano is senior lecturer at the informatics center, UFPE, Brazil. He also worked as planning and finance vice-president at UFPE, Brazil, from 2012 until 2015.