Empowering Business Users to Analyze Enterprise Architectures: Structural Model Matching to Configure Visualizations

Sascha Roth, Matheus Hauder, Marin Zec, Alexej Utz, and Florian Matthes
Technische Universität München
Boltzmannstr. 3
85748 Garching, Germany
{roth, matheus.hauder, marin.zec, alexej.utz, matthes}@tum.de

Abstract—Visualizations are common means to analyze Enterprise Architecture (EA) models and support decision makers with relevant information on organizational processes, information systems, infrastructure, and their interconnections. These EA visualizations are tailored typically to the specific information demand of stakeholders. Currently, creating such stakeholder-specific visualizations requires experts with a strong technical background due to complex configurations and inflexible tool solutions. In particular business users often lack technical expertise. At the same time, their concerns and questions often arise spontaneously. In this vein, visualizations that can be generated without expert knowledge would enable business users to perform ad-hoc analyses of an EA information model. Against this background, we propose a solution which facilitates analyses of arbitrary EA information models by non-technical stakeholders. Our approach is based on ad-hoc, configurable visualizations. We introduce an end-user friendly wizard that lowers the barrier for the creation of EA visualizations. The wizard is based on structural pattern matching of models.

Key to our approach are abstract viewpoints that model best-practice knowledge of EA visualizations and abstract view models which model the information demand of abstract viewpoints. Our contributions in this paper are 1) a meta-information model capable to capture both the technical information demand of an abstract viewpoint and the information offer of an EA information model, 2) a pattern matching algorithm calculating viable configurations for bindings of visualizations to information, and 3) a wizard to support non-technical stakeholders with the creation of these visualizations. We present an implementation of our approach and show user interface design of the wizard. The wizard recommends feasible configurations automatically in order to unburden the configuration process for non-technical stakeholders.

Keywords—Enterprise Architecture, visual analysis, structural model matching, EA visualizations

I. INTRODUCTION

Over the past decades, business applications and respective business application landscapes of organizations have grown to systems of systems intrinsically complex and hard to manage as a whole [1], [2]. At the same time, frequently changing market conditions force enterprises to transform their business to outperform competitors. In this vein, the mutual alignment of business and IT enables faster business transformations and ultimately contributes to increase shareholder return. Due to the intrinsic complexity of both business and IT, realizing effective and efficient IT support of constantly changing business processes is a major challenge of enterprises today. As a result, Enterprise Architecture (EA) management seeks to manage the arising complexity in order to increase business and IT alignment while reducing operational costs. Well executed EA management is commonly perceived as a strategic advantage since enterprises can respond faster and more reasonably to continuously changing market conditions.

Information about the EA may serve multiple stakeholders with different concerns and potentially unclear goals [3]. Common means for the analysis of this information are visualizations of these EA information models [4], [5], [6]. Due to the diverse nature of these concerns, decision makers require individual viewpoints that are tailored to their specific demand [7], [8]. Large enterprises commonly contain a large number of information systems, organizational processes, and infrastructure systems that are interconnected via interfaces. The information volume leads to large visualizations that are typically created with model-driven approaches by expert users [9], [10], [11]. Considering the background of EA visualizations\(^1\) the EA management community employs a finite number of views with variations identified as patterns (cf. [14]), such that they can be predefined in an abstract manner including some variability points [12], [11]. This means, the EA information model is unknown to the visual designer [11] and bound at runtime via model mapping techniques [10].

EA endeavors commonly start by documenting the current state of the organization in an EA repository (e.g. [15], [16], [17], [18]). This status quo of the EA is typically documented with respect to a formalized information model (cf. [19]) capturing organizational processes, information systems, infrastructure, and their interrelations. The resulting EA information model consists of highly interlinked entities. Although there are proposals for a standard EA information model [20], as of today, enterprises tend to develop their

\(^1\)We refer the interested reader to [12], [13], [3], for a predefined set of best-practice EA visualizations gathered from industry.
own organization-specific EA information model [21] to capture changing requirements and information demands from stakeholders. Thereto, EA management initiatives use highly specialized EA tools to manage EA information. With respect to the observations made above, some of these EA tools, in the following referred to as EA repositories, enable an adaptation of their underlying EA information model [22]. At the same time, EA information models may change frequently in case the data is automatically gathered from federated information sources (cf. e.g. [23], [24], [25], [26]).

Prevailing approaches for generating EA visualizations typically take into account multiple steps performed by a variety of different actors as shown in Figure 1. During the modeling step the required information is gathered from stakeholders or existing information sources in the organization (cf. e.g. [23], [24], [25], [26]). The result of the data binding step is a mapping between information model elements and an abstract viewpoint definition. Typically, this step requires specific knowledge from expert users. The definition of the abstract viewpoint is performed within a visualization design step. During configuration, the abstract viewpoint is detailed with e.g. color encodings, axis descriptions, and symbols. As a result, these subsequent steps lead to a communication effort between the various actors.

Answering above research question, the remainder of this paper is structured as follows. First, we briefly revisit related work relevant for the scope of this paper in Section II. Then, we present our approach in Section III followed by a description of its implementation in Section IV. The paper concludes with an outlook on further research in Section V.

II. RELATED WORK

Decoupling of visualizations and (model) data is discussed by research as well as industry (e.g. [27]). In particular in the field of software visualization, Panas et al. [28] describe an approach and a corresponding proof-of-concept for configurable software visualization. Instead of hand-coding mappings between model and visualization, they introduce a general and flexible framework for online-configuration of views. Their approach is based on graph structures of both the data model (i.e. software entities) as well as view model. These graph structures consist of nodes modeling entities or n-ary relations, edges modeling binary relations, and arbitrary properties of nodes and edges. Panas et al. provide a predefined set of mapping functions to map elements from the data graph to elements of the view graph. A visualization is configured by assigning a data graph property and the desired binding function to each view property. Then, the view graph is mapped to a scene graph which is specific to the rendering engine and serves as actual basis for the visualization. While Panas et al. distinguish between model, view, and scene and thus allow reuse of mappings, they do not offer suggestions for the end-user who might not have the required technical skills to specify a mapping from information offer to information demand.

Krule et al. [10] discuss the motivation and advantages of decoupling models and visualizations in the EA domain. They acknowledge that the specification of a model transformation of an information model to a view model is typically a difficult task for business users. Krule et al. suggest to use higher order transformations in order to avoid the need of specifying a new transformation for each viewpoint. They provide a prototypical implementation of their approach using the Eclipse Modeling Framework2. However, Krule et al. recognize that the visualization configuration process is still “knowledge-intensive and error prone” and call for an intuitive interface.

In their works, Goldschmitt et al. [29] coin the term view-based modeling, i.e. using a domain-specific modeling language to manipulate underlying information. The abstract concept of manipulating model information by stakeholder-specific views is similar to one of the solutions presented in this paper. This research group concentrates on text-based approaches to synchronize abstract and concrete syntax. However, they do neither apply their techniques to best-practice EA visualizations nor empower business users to configure these.

Figure 1. Overview of the current state and the scope of this paper.

We aim at supporting end-users to create visualizations on their own to reduce this effort and allow the creation of ad-hoc visualizations. With regard to the considerations made above, we conclude to the following research question:

How can business users be empowered to visually analyze data of arbitrary EA information models?

Our contributions in this paper are 1) a meta-information model capable to capture both the technical information demand (abstract view model) of an abstract viewpoint (cf.[11]) and the information offer of an EA information model, 2) an algorithm for structural pattern matching of models calculating viable configurations for bindings of visualizations to information, and 3) a wizard to support non-technical stakeholders with the ad-hoc creation of EA visualizations.

Visual notations, e.g. approaches like the Graphical Modeling Project\(^3\), are based on concrete models and thus can only be used for concrete EA models. Since there is no common standard for EA models [30], organizations tend to develop an organization-specific model that describes their EA with respect to the organization’s terminology and concepts.

The visualization approaches mentioned above do not address two major issues. First, ad-hoc information needs cannot be met spontaneously by business users because knowledge-intensive modeling is required for view configuration in one way or another. Second, enterprise models tend to be dynamic [31], [32]. Visualizations should be flexible and adjustable to changes in the information model without demanding from the business user to have programming or sophisticated modeling skills.

In [11], we present both requirements and a conceptual framework for interactive EA visualizations. In this vein, we show how to define domain-independent visualization types and distinguish between information offer and information demand of a visualization type. The former is captured in the information model, the latter is captured in a so-called abstract view model. The information demand needs to be met by an associated information offer in order to be able to generate a particular visualization in a model driven manner. While the information model and the visualization are decoupled in our approach, we did not detail how the information model is mapped to the abstract view model. However, we did not discuss how business users can perform the rather abstract task of mapping the information model to the abstract view model without prior knowledge of the information model’s details, modeling expertise, and/or programming skills. Especially in [22], [33] Matthes et al. analyze the state-of-the-art in EA management tools and conclude that current EA management tools require experts with programming skills to adapt visualizations.

Schweda [34] presents formal theories describing when two information models embed into each other. Thereby, Schweda distinguishes between strong subsumption (read/write) of models and weak subsumption of models (read-only). Based on his theories, we described the underlying idea of the solution presented in this paper in [35]. However, we neither presented a concrete implementation of the pattern matching algorithm nor did we provide an user interface design enabling business users to configure rather complex model mappings. The scope of this work is to show a prototypical implementation for visualization-driven ad-hoc analysis in the EA domain. We chose a structural over a semantic model matching since the latter requires semantic annotations and/or rule definitions. Those approaches often demand technical knowledge which we intend to avoid in the first place.

III. EMPOWERING BUSINESS USERS TO CONFIGURE EA VISUALIZATIONS

As we detailed in [11], model-driven approaches to generate EA visualizations, e.g. [10], require a mapping of an EA information model to a particular abstract viewpoint, i.e. a ‘visualization type’. Since the information model in particular may evolve over time [32], questions and concerns of stakeholders cannot be entirely anticipated. This also holds true for visualizations addressing them. At the same time, EA stakeholders often lack technical expertise and cannot be considered model experts nor experienced programmers.

As a reaction, we illustrate how business users (end-users) can be empowered to provide a model mapping of an EA information model to an abstract view model. As above-mentioned such a mapping is a basic prerequisite for generating an EA visualization in a model driven manner. First, we introduce a meta-information model used as a common denominator for both models. Subsequently, we introduce an algorithm that operates on two models complying to this meta-information model in order to compare two models structurally.

A. A Meta-information model for matching model structures

In line with model driven approaches to generate visualizations [10], [36], [11], we distinguish between technical information demand of a visualization, the so-called abstract view model, and the information demand given by an EA information model. Figure 2 illustrates a simplified meta-information model (m3) that is used to abstract from the concrete implementation of both the EA information model and the abstract view model, i.e. information offer and demand (models).

![Figure 2. Meta-information model](http://www.omg.org/mof/, last-accessed April 20, 2013)

The meta-information model covers basic concepts of the meta object facility (MOF)\(^4\), i.e. Classes, Attributes, and Re-
lationships. For the latter, we capture transitive relationships explicitly and a path that gives information how to traverse the model to navigate through this transitive relationship. Eventually, this information is used in the later model-driven generation of an EA visualization. UML distinguishes between navigation forward, backward, or both. In our model, we consider relationships to be bidirectionally traversable. Against this background, our approach focuses on read-only EA visualizations (cf. also Schweda [34]). For our considerations, start and end of a relationship are swapped if the underlying information model uses directed relationships that are technically traversable in both directions.

B. An algorithm to match model structures

After abstracting the EA information model to an information offer model $m_o$ and the abstract view model of an abstract viewpoint to an information demand model $m_d$ (cf. [11]) to above described meta-information model, we use these models to facilitate the configuration of the visualizations.

Algorithm 1 matches both models, i.e. retrieves $\text{matches} \subseteq m_o$ that structurally fit $m_d$. The result of this algorithm may serve end-users in a twofold manner. On the one hand, structural matches of both models deliver valid configuration variants to the end-user. In this vein, a valid configuration offer model to an abstract viewpoint to an information demand model is calculated which may serve as a suggestion for the latter, we capture transitive relationships explicitly and a path that gives information how to traverse the model to navigate through this transitive relationship. Eventually, this information is used in the later model-driven generation of an EA visualization. UML distinguishes between navigation forward, backward, or both. In our model, we consider relationships to be bidirectionally traversable. Against this background, our approach focuses on read-only EA visualizations (cf. also Schweda [34]). For our considerations, start and end of a relationship are swapped if the underlying information model uses directed relationships that are technically traversable in both directions.

Algorithm 1 Pattern matching

```
—— compareModel (md, mo)
rels ← getRelationships(md)
foreach rel ∈ rels,transitive do
  foreach co ∈ mo do
    if co classifier ≡ rel startclassifier ∧ rel startattributes ∩ co attributes then
      co attributes ← co attributes ∪ rel startattributes
    end
  end
  r ← 0
  matchCount ← 0
  foreach cd ∈ md do
    elementsToIgnore ← ∅
    r ← compareClass(cd, co, elementsToIgnore)
    if r ≡ 0 then
      matchCount ← matchCount + 1
      matches ← matches ∪ (cd, co)
    break
  end
  return md.classes.count – matchCount
—— compareClass (cd, co, elementsToIgnore)
foreach ad ∈ cd.attributes do
  if ad notElementOf(elementsToIgnore) then
    r ← compareAttribute(ad, ao, elementsToIgnore)
  if r ≡ 0 then
    matches ← matches ∪ (ad, ao)
  else
    closestMatches ← closestMatches ∪ (ad, ao, r)
  return r
—— compareAttribute (ad, ao, elementsToIgnore)
if ad.type ≡ STRING then
  return 0
else if ad.type ≡ RELATION ∧ ao.type ≡ RELATION then
  r ← compareRelationship(ad, ao, elementsToIgnore)
  if r ≡ 0 then
    matches ← matches ∪ (ad, ao)
  return r
else if ao.type ≡ ao.type then
  matches ← matches ∪ (ad, ao)
  return 0
else
  closestMatches ← closestMatches ∪ (ad, ao, 1)
  return 1
—— compareRelationship (rd, ro, elementsToIgnore)
relToIgnore ← relToIgnore ∪ (rd, ro)
r ← compareClass(rd.start, ro.start, relToIgnore)
  r ← r + compareClass(rd.end, ro.end, relToIgnore)
  if r ≡ 0 then
    matches ← matches ∪ (rd, ro)
  else
    closestMatches ← closestMatches ∪ (rd, ro, r)
return r
```
Comparing models is done by comparing respective classes with each other. In order to compare two classes the \texttt{compareClass(c_d, c_o, elementsToIgnore)} function (line 18) is used. The comparison is performed by analyzing and comparing the attributes of the classes. Every attribute of a class from the demand model is compared to each attribute of the offer model class (line 22). The result of the comparison is either 1 or 0, whereby 0 indicates a match of the attributes. This indicator is summed up over all attributes and returned as the result of \texttt{compareClass}. The matched pair of attributes is collected (line 24) for later use in the configuration of the visualization.

Attributes are compared calling \texttt{compareAttribute} (line 22). This function returns either 0 indicating a match or 1 for a mismatch. For a structural match, attributes are compared by their data types and cardinalities. First we check whether cardinalities of an attribute in \(m_o\) satisfy the multiplicity of the demand model attribute (line 29). In the next step, we check the data type of the demand model attribute. In line with Schweda [34], the algorithm takes into account that data types \{byte, integer, decimal, boolean, date, time, dateTime, duration, URI, ...\} (cf. also [37]) have a lexical representation\(^5\), i.e. they can be converted to a string. That is, when an attribute \(a_d\) is of type string, then \(a_o\) matches irrespective of its own data type. We assume that even relationships can be converted to string representations. Attribute \(a_d\) may also be a relationship (line 35) or of any other type (line 39).

Relationships between classes are compared in \texttt{compareRelationship}. The algorithm compares source (line 50) and target (line 51) classes of a relationship utilizing the \texttt{compareClass} (line 18). In order to avoid circularity the current relationships are ignored while comparing the classes (Line 49). If both start and end of \(r_d\) match \(r_o\), then \texttt{compareRelationship} returns 0 and respective information of the match is added to the result. To avoid infinite recursion through recursive relationships of a model, the algorithm takes into account collection \texttt{elementsToIgnore} representing elements that are not traversed any further.

During the design phase of the algorithm, we considered that an algorithm could extend \(m_o\) in a first step. Such an extension would basically transfer the offer model in such a way that attributes become relationships to classes whereas the type of these classes equals the respective primitive type of the attribute. This way, on the one hand, \(m_d\) is minimal in terms of required information to generate the visualization and on the other hand, chances that \(m_o\) fits \(m_d\) are increased. However, although this is technical feasible, end-user feedback led our design into a different direction. The major reason for our design decision was that such a solution would produce unnecessary choices and thus possibly confuse end-users.


IV. PROTOTYPICAL IMPLEMENTATION OF THE VISUALIZATION WIZARD

The pattern matching algorithm described above serves as a basis for our visualization wizard to recommend feasible configurations to stakeholders. This wizard guides the user through three essential steps. First, a predefined abstraction viewpoint, e.g. a binary matrix, a portfolio diagram, or a cluster map, has to be selected (cf. Figure 5 1). Our framework can be extended by new abstract viewpoint definitions.

Every abstract viewpoint definition contains both a sample preview of the viewpoint (cf. Figure 5 2) and a potential information model (Figure 5 3). The screenshot in Figure 5 shows a ternary matrix to realize a process support map (see [13]). Furthermore, an explanation of the ternary matrix concept (cf. 3 in Figure 5) and a typical usage scenario for EAM is given (4 in Figure 5).

\[\text{Figure 3. Abstract view model (}\text{m}_d,\text{) of a ternary matrix}\]

When an abstract viewpoint has been selected, the wizard guides the user through the configuration of the data binding between the EA information model and the abstract viewpoint. The information demand for the chosen abstract viewpoint is captured by the abstract view model, which describes the information required to generate the visualization. Figure 3 shows the respective abstract view model of a ternary matrix, i.e. \(x\)-axis, \(y\)-axis, and a cell element connected via a relationship such that each triple can be identified uniquely. Technically, elements of this abstract view model have to be bound to elements of an EA information model in order to generate a concrete visualization, i.e mappings to an EA information model has to be provided by the user.

Although the EA information model can be an arbitrary information model for our approach, we use a concrete EA information model to exemplify our approach. Thereto, Figure 4 illustrates a subset of the ArchiMate 2.0 [20] specification. Subsequently, it serves as an example for an
Figure 5. First step of the visualization wizard: diagram selection

Figure 6. Second step of the visualization wizard: configuration of the data binding

Figure 7. Third step of the visualization wizard: customize diagram
EA information model frequently applied in practice, which captures the following basic concepts and their interrelations:

- business layer, i.e. business processes and business objects,
- application layer, i.e. application service, application component, application interface, and data object, and
- infrastructure layer, i.e. node devices.

Algorithm 1 applied to \( m_{ds} \) and \( m_{os} \) provides valid, i.e. feasible, configurations of data bindings between the abstract viewpoint \( (m_{ds}) \) and the EA information model \( (m_{os}) \). The wizard assists stakeholders in the configuration of the desired viewpoint since only feasible configuration settings can be specified. For the selected ternary matrix (see Figure 6), the user can choose the EA information model (Figure 6 1) and the support relationship items (Figure 6 2). The number of respective instances for every type, attribute, and relationship is given in parentheses. The user can select elements of \( m_{os} \) to define a mapping to elements of \( m_{ds} \). For each of this elements the user can refine the selection of objects with the instance filter as shown in Figure 8. In the illustrated example, a filter already has been defined and is active for the type business object (Figure 6 3).

The filter dialog shows the context, i.e. selected model and type (Figure 8 1), and filters business objects by their attribute owner for a particular value, i.e. for the owner ‘John Smith’ (Figure 8 2). Additional attribute filters can be defined and applied in conjunction (Figure 8 3). The wizard further lists available relationships including transitive relationships (Figure 6 4). Since only feasible configurations are shown to the user, a generic preview of the actual visualization can be generated and is shown to the user (Figure 6 5).

In the third and final step, visual properties of the visualization may be customized (Figure 7). A variety of customization options are available for any given abstract viewpoint. For instance, the border color of the cell elements can be specified (Figure 7 1) or the symbol shapes can be customized, e.g. from rectangle to chevron shapes (Figure 7 2). Each visualization generated can be downloaded as PowerPoint (.pptx) slide (Figure 6 6) which can be easily modified and included in an executive presentation.

**V. CONCLUSION AND OUTLOOK**

Due to human cognition, visualizations are essential to support stakeholders with relevant information about an organization’s architecture. Analyzing EA models using EA visualizations is often difficult because of inflexible configurations in prevailing tools that can only be created by experts with strong technical background. Most likely, an expert with programming or modeling skills is required. As a result, ad-hoc analyses of the organization’s architecture often can not be performed in a timely fashion. However, stakeholders require instantaneous solutions in practice to satisfy ad-hoc information demands.

In this paper, we presented a lightweight meta-information model for matching structures of different models and an algorithm operating on respective models to calculate feasible configurations for bindings of visualizations to information. We further presented an implementation of our concepts in a user-friendly wizard to support stakeholders with the generation of EA visualizations and exemplified our approach based on an EA information model conforming to the ArchiMate 2.0 specification [20]. Thereby, we revealed user interface design and implementation details of the wizard. The wizard automatically recommends feasible configurations to the stakeholder in order to simplify the configuration process. First end-user feedback is positive. Currently, we are evaluating the presented approach with industry in real-world settings.
REFERENCES


