Towards a Framework for Enterprise Architecture Analytics

Rainer Schmidt,
Michael Möhring

Munich University
Munich, Germany
Rainer.Schmidt@hm.edu,
Michael.Moehring@htw-aalen.de

Matthias Wißotzki,
Kurt Sandkuhl

University of Rostock
Rostock, Germany

Dirk Jugel,
Alfred Zimmermann

Reutlingen University
Reutlingen, Germany

Abstract—

Current approaches for enterprise architecture lack analytical instruments for cyclic evaluations of business and system architectures in real business enterprise system environments. This impedes the broad use of enterprise architecture methodologies. Furthermore, the permanent evolution of systems desynchronizes quickly model representation and reality. Therefore we are introducing an approach for complementing the existing top-down approach for the creation of enterprise architecture with a bottom approach. Enterprise Architecture Analytics uses the architectural information contained in many infrastructures to provide architectural information. By applying Big Data technologies it is possible to exploit this information and to create architectural information. That means, Enterprise Architectures may be discovered, analyzed and optimized using analytics. The increased availability of architectural data also improves the possibilities to verify the compliance of Enterprise Architectures. Architectural decisions are linked to clustered architecture artifacts and categories according to a holistic EAM Reference Architecture with specific architecture metamodels. A special suited EAM Maturity Framework provides the base for systematic and analytics supported assessments of architecture capabilities.

Keywords—Enterprise Analytics, Enterprise Architecture

I. Introduction

Enterprise Architecture represents the overall structure of the enterprise, which is composed of its business and IT structures, like stakeholders, strategy, business capabilities, domains and functions, business and IT processes, business products, business services, IT services, IT applications, and technologies. The quality and completeness of information, however, decreases when going from top to bottom. The top layers of architecture models contain complete and up-to-date information. This changes for lower-level information such as concrete IT services and applications. This information is hitherto difficult to collect and especially difficult to keep up-to-date.

Therefore it does not surprise that current approaches for enterprise architecture lack analytical instruments for cyclic evaluations of business and system architectures in real business enterprise system environments. The aim of our research is to close this gap and enhance analytical instruments for cyclic evaluations of information systems and enterprise architectures in real business environments.

On the other hand, modern infrastructure systems contain a lot of information that describes architectures on a low abstraction-level. This is due to the broad use of virtualization [1]. All items of a virtualized environment are completely described in associated management systems like configuration management systems[2]. However, the use of these data for enterprise architecture analytics had been hampered by shortcomings of information technology, limiting the volume, variety and velocity of data collection and analysis.

This has changed since the advancement of Big Data. Big Data can be best understood as a transition integrating several technological developments [3]. First, the management of data has become much more powerful. Today, it is possible to distribute even relational databases world-wide [4]. Second, new approaches for data processing based on commodity hardware allow to handle the processing of very large volumes of data such as Hadoop [5]. Third, data are processed increasingly as stream enabling decision making in real-time or near real-time.

Using technologies from the Big Data context it is possible to complement the architectural information by a bottom-up perspective. Data collected from infrastructure systems, processed and analyzed using Big Data technologies is possible to provide valuable additional architectural information. Enterprise Architectures may be discovered, analyzed and optimized. Furthermore the compliance of Enterprise Architectures may be verified.

In this paper we introduce a framework for enterprise architecture analytics, which we are integrating from an extended service-oriented enterprise architecture reference model in the context of Big Data analytics for architecture, new decision support methods for architecture alignment, and an original architecture maturity approach.

The paper is organized into the following sections. First, suitable abstracted Enterprise Architecture models and views are

clustered using the Enterprise Software Architecture Reference Cube to provide a normative classification base for EA Analytics considering architectural artifacts, which are prepared for associated decisions and architecture lifecycle information. Then the key concepts Enterprise Architecture Analytics are defined. To do the new potential for Enterprise Architecture created by Big Data and Business Analytics are discussed. An original EA Maturity Framework provides an example for using analyses created by our Enterprise Architecture Analytics approach. After discussing related work we conclude with intermediate results from our research showing also the ideas for future work.

II.   ENTERPRISE SERVICES REFERENCE ARCHITECTURE

In areas where flexibility or agility in business is important, services computing is the approach of choice to organize and utilize distributed capabilities for Cloud and Big Data applications. Innovation oriented companies have introduced in recent years service-oriented architectures to assist in closing the business - IT gap and making it cloud-ready. The benefits of SOA are recognized for systems on the way to cloud computing and being ready for extended service models. They comprise flexibility, process orientation, time-to-market, and innovation.

The OASIS Reference Model for Service Oriented Architecture [6] is the basic service-oriented abstract architecture framework, which guides the construction of Reference Architectures, like [7] [8]. We are using the fundamental concepts and definitions for Software Architecture, Architecture Reference Model, and Reference Architecture from [9], and expand these for our research on Service-oriented Enterprise Architectures, as in [10], [11], [12].
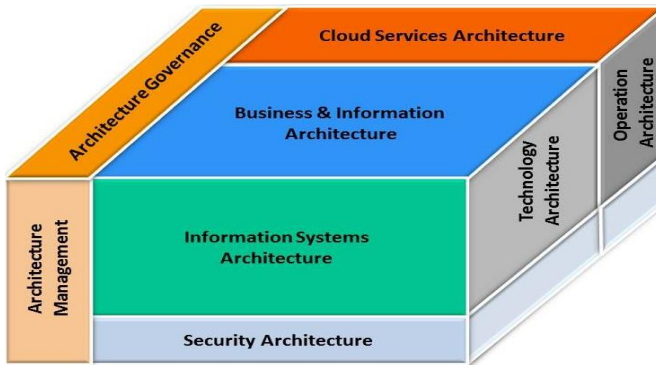

Fig. 1.   ESARC - Enterprise Software Architecture Reference Cube.

The ESARC – Enterprise Services Architecture Reference Cube [13] (Fig. 1) is an enterprise reference architecture model, which completes architectural standards from [14] and [15] for Service-oriented EAM – Enterprise Architecture Management. We have integrated additionally fundamental concepts and architectures for Services Computing, like in [16] [17] [18], as well as architecture references of Cloud Computing from [19] and [20] , as well as current Cloud Reference Architectures [21] [22] [23] [24]. ESARC is our still growing original Service-oriented Enterprise Architecture Reference Model, which provides an integral EAM model for main interweaved architectural viewpoints. ESARC abstracts from a concrete business scenario or from specific technologies.

The Open Group Architecture Framework [14] provides together with the current standard of ArchiMate [15] the basic blueprint and structure for our extended service-oriented enterprise architecture domains like: Architecture Governance, Architecture Management, Business and Information Architecture, Information Systems Architecture, Technology Architecture, Operation Architecture, and Cloud Services Architecture. ESARC provides a coherent aid for clustering, classification, examination, comparison, quality evaluation and optimization of enterprise architectures.

To be able to integrate architectural resources from the state of art and practice we have developed the Enterprise Services Architecture Metamodel Integration – ESAMI [25], as a correlation-based integration method for architecture viewpoints, views and models. The following few examples of interrelated reference architectures of ESARC are the result from correlation-based mappings of architectural models and their elements. The Business and Information Reference Architecture – BIRA provides a single source and comprehensive repository of business-related knowledge from which concrete corporate initiatives will evolve and link. This knowledge is model-based and defines an integrated enterprise business model, which includes organization models and business processes. The BIRA opens a connection to IT infrastructures, IT systems, and software as well as security architectures.

We are using metamodels to define architecture model elements and their relationships within ESARC [13] and [25]. We use metamodels as an abstraction for architectural elements and relate them to architecture ontologies [26]. Architecture ontologies represent a common vocabulary for enterprise architects who need to share their information based on explicitly defined concepts. Ontologies include the ability to automatically infer transitive knowledge. The Metamodel of the Business & Information Reference Architecture – BIRA consists of ESARC-specific concepts, which are derived as specializations from generic concepts such as Element and Composition from the Open Group's SOA Ontology [26].

The ESARC Information Systems Reference Architecture – ISRA is the application reference architecture and contains the main application-specific service types, defining their relationship by a layer model of building services. The core functionality of domain services is linked with the application interaction capabilities and with the business processes of the customer organization. In our research we are integrating reference architecture models for services computing [6] [7] [8], and extend them for cloud computing.

In the ESARC – Information Systems Reference Architecture we have differentiated layered service types. The information services for enterprise data can be thought of as data centric components, providing access to the persistent entities of the business process. The capabilities of information services combine both elementary access to CRUD (create, read, update, delete) operations and complex functionality for finding/searching of data or complex data structures, like data composites or other complex-typed information. Close to the access of enterprise data are context management capabilities, provided by the technology architecture: error compensation or

exception handling, seeking for alternative information, transaction processing of both atomic and long running and prevalent distributed transactions.

Cloud architectures are still under development and have not reached today their full potential of integrating EAM with Services Computing and Cloud Computing. The ESARC – Cloud Services Architecture provides a reference-model-based synthesis of current standards and reference architectures, like [21] [22] [23] [24]. The NIST Cloud Computing Reference Architecture [21] defines the Conceptual Reference Model from the perspectives of the following Actors in Cloud Computing: Cloud Consumer, Cloud Provider, Cloud Auditor, and the Cloud Broker. The NIST standard defines following deployment models: private cloud, community cloud, public cloud, and hybrid cloud. Cloud Computing offers essential characteristics like: on-demand self-services, broad network access, resource pooling, rapid elasticity, and measured services. The fundamental part of the NIST Reference Architecture is defined by following Cloud Service Models: IaaS – Infrastructure as a Service, PaaS – Platform as a Service, and SaaS – Software as a Service. Some Standard extensions like [22] provide practical additions for supporting more directly modern business architectures by BPaaS – Business Process as a Service and by giving a direct link by integrating with business services to a service-oriented Enterprise Architectures. Security additions from the CSA Security Guidelines for Critical Areas of Focus in Cloud Computing [23] defines a Jericho-Security-focused Service-oriented Reference Architecture for Cloud Computing and integrates the management perspectives from standards like ITIL and TOGAF. The Service-Oriented Cloud Computing (SOCCI) Framework [24] is an enabling framework for an integrated set of cloud infrastructure components. Basically it is the synergy of service-oriented and cloud architectures by means of a consistent As-a-Service-Mechanism for basically infrastructure cloud services. The fundamental characteristics of a Service-oriented Infrastructure (SOI) are: business-driven infrastructure on-demand, operational transparency, service measurement, and consumer provider model. The SOCCI-Service-Oriented Cloud Computing Framework is the extension of the Service-oriented Infrastructure (SOI) mapped to the SOA Reference Architecture [16]. The SOI-Framework is the layer on top of the basic infrastructure and provides the elements of SOCCI: Compute, Network, Storage, and Facilities. SOCCI extends these basic elements of SOCCI by Business and Operational SOCCI Management Building Blocks.

III. Enterprise Architecture Analytics

The beneficial effects of data-driven decision making on the performance of firms is well known [27]. So far, a data driven approach has not been applied to enterprise analytics due to a lack of information and limited computation capabilities. However, nowadays the situation has changed due to the advancement of virtualized infrastructure and big data. Virtualized infrastructures such as cloud-environments [28] are theoretically capable to describe every of the entities contained in them. Furthermore, such environments not only have a static

perspective for architectural elements, they also contain a dynamic perspective by logging all relevant elements in virtualized environments. In the past, the data collected in virtualized environments could not be exploited appropriately. The huge sets of data overburdened the existing computation capabilities. Furthermore, the data from the dynamic perspective contains a lot of semi-structured data that was difficult to process with existing approaches.

Through the advancement of Big Data, the situation has changed significantly. Big Data [29] is one of the most disruptive information technological developments [30] [31]. Big Data enables handling and analyzing more types of unstructured (e.g. user statements in social media) and semi-structured data as before [31] with higher velocity and volume [32]. The importance of this disruptive development is shown by an empirical study (worldwide online survey with over 1300 IT managers) from ZDNet showing that "70% will use data analytics by 2013" [33] (ZDnet 2012). Big Data is not identical with a certain technology. However the use of the highly distributed Hadoop [32] architecture is often associated with Big Data. Big Data enables the extension of analytics in the three dimensions volume, variety and velocity [34] as shown in Fig. 2. In comparison to business analytics it is now possible to analyze nearly in real-time large quantities of data from data sources with varying structure processed.
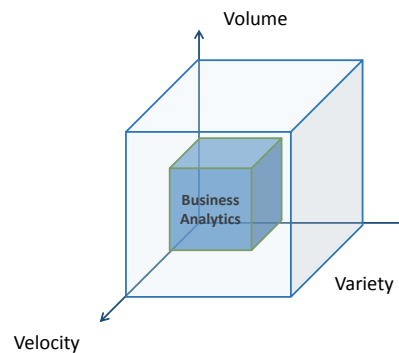

Fig. 2. Big Data scales out analytics based on [35]

Big Data applications are data-intensive applications, with a large volume of data, a high velocity of processing and a data variability of the existing IT solutions [31]. An example scenario for Big Data is the provisioning of real-time information to mobile users. Based on a stream of position information created by the mobile user's smartphone, the mobile receives information selected from a variety of sources and provided nearly in real-time. Various examples of real business cases show business impact of Big Data such as [36]. In this case study it is shown, that significant cost cuts could be achieved by decreasing the estimated and actual arrival time of aircrafts. Furthermore it is possible to increase sales through faster data analysis and thereby better personalized promotions. Business processes at the link to the customers and suppliers can be improved by using Big Data [37]. This is possible because of the increase of process quality, enabled by a better data quality for decision making [37]. Furthermore, IT applications can benefit

from the use of Big Data, because current empirical studies show that the perceived advantages of Big Data are very high in the field of technology development and IT [38]
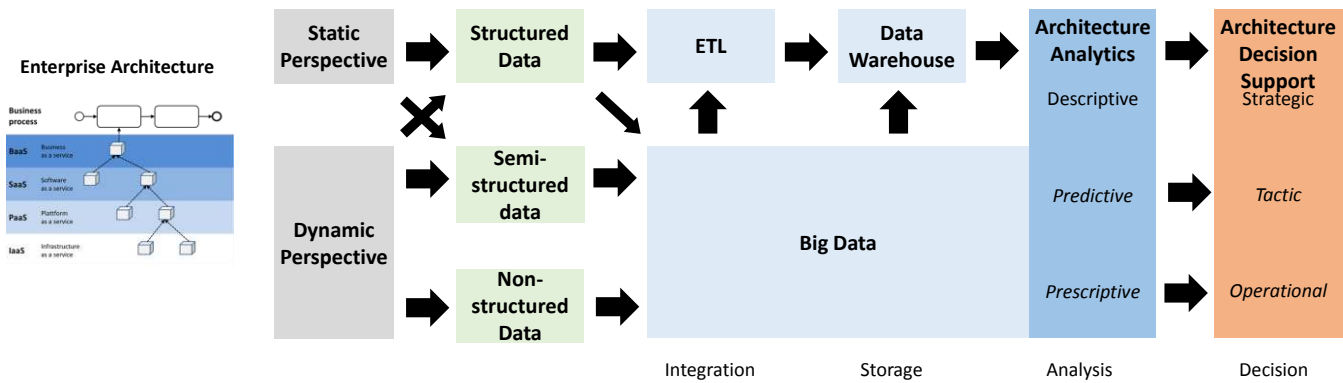

Fig. 3.   Enterprise Architecture Analytics

By combining conventional means of analytics such as data warehouses [39] with Big Data it is possible to use both data from the static and dynamic perspective of enterprise architecture, as shown in Fig. 3. Especially structural information from the static perspective may be processed using a conventional extract, transform and load (ETL) [39] approach in combination with a data-warehouse. The primary goal is to gain descriptive information about the architecture. There is, however, the possibility to use Big Data for tasks that are to computation intensive.

The dynamics perspective of enterprise architecture provides huge amounts of non- and semi-structured data from log files etc. Using them predictive and even prescriptive analyses are possible exploiting the capability of Big Data to process large volumes of data. Furthermore, these analyses are available in a shorter and shorter time scale. This and the decreased cost of creating analyses allow the use Enterprise Architecture Analytics not only in strategic decisions, but also tactical or even operational decisions. Now, the analysis of Enterprise Architectures using Business Analytics and Big Data shall be described. The core idea is depicted in Fig. 3. Enterprise Architecture is analysed using concepts and technologies provided by Business Analytics and Big Data. By this mean decision support is provided for designing and optimizing Enterprise architecture.

The use of enterprise architecture analytics can generate positive business value and business impacts. Important architecture decisions can now be done based on a better data quality and not only on a gut instinct. Therefore the following metrics [40] [41] can be better acquired:

- Cost metrics
- Scalability metrics
- Portability metrics
- Security metrics
- Etc.

A comparison of different architecture variants based on metrics show CIO's a better view of the possibilities and constraints of each. As a result better decisions can be applied.

This comparison can be applied for example by using a utility analysis [42]. An example of a group utility analysis based on the metrics above are shown in the following figure:

Fig. 4.   : Example of utility analysis of architecture variants

Furthermore a better understanding of processes in the enterprise architecture based on a larger base of data can improve IT services. If the demand of each IT service at each time is known or better predicted, preproduction costs can be decreased. Prediction algorithms and methods for the calculation of the demand of IT services (e.g. ARIMA [43], linear regression [44], neuronal networks [45] models) and their forecasts of demand be improved by a broader database. Generally the IT-Demand (e.g. normal use and peaks) are difficult to predict [46]. The prediction can be improved, because of using a broader database and using complex as well as robust (e.g. non-linear analysis) prediction algorithms (e.g. linear regression [44]) [47]. Therefore forecast errors can be decreased. A Forecast error is defined as the absolute value of the difference between the real and predicted demand [48] [49]:

$$forecast\ error = |\ demand_{real} - demand_{predicted}|$$

Equation 1: forecast error

The quality of prediction the forecast error can be improved by using other more detailed metrics (like RMSE [48] [49]).

A reduced forecast error outcome a cost saving for idle time costs:

$$idle\_time\_cost = \sum (forecast\ error * capacity_{costs}) + \sum preproduction\_costs$$

Equation 2: forecast error

Idle time is defined as the time where not all available IT resources are used. Furthermore some effort for preproduction is needed, because many IT services need some time or capacity to be established.

Therefore new price differentiation (internal or external) can be implemented. Times of less demand can decrease with a cost reduction of the lower price limit for the internal or external customers. In contrast times of a huge amount of demand can be managed be a increased price for customers. As a result idle time costs can be decreased and earnings increased. These facts are shown in the following formula.

$$income = \sum budget - \sum idle\_time\_costs - \sum used\_capacity\_costs - \sum other\_costs$$

Equation 3: cost aspects of IT services

The budget variable in equation 3 describes the budget of the IT department for enterprise architecture solutions. The IT budget includes the costs of salaries, maintenance, technology, R&D etc. [50]. Mostly this budget is defined by the CFO of each enterprise or organization. Each used capacity cost (e.g. cost for database access, business process use, esb-use, etc) are defined in the used_capacity_cost variable. Furthermore other costs (e.g. for administration, special indirect costs) defined by other_costs variable in equation 3.

An increase of incomes will implement trough a reduction of idle time costs through a better knowledge about the demand. If the IT department is implemented as a profit center, earnings will increase and improve the standing of the department in the enterprise or organization. In this case the variable budget must be renamed as revenue in the formula.

In the past IT departments were distinguished as a cost driver and collection of risks [51]. An applied enterprise architecture analytics approach can improve the standing of the IT department in the enterprise and maybe move the IT department from a cost driver to a profit center. This is possible because of more accurate enterprise architecture decisions and focusing on cost aspects, which are better supported by this approach than in the past and increase the outcomes of the enterprise (e.g., productivity [50].

IV. EA CAPABILITY ASSESSMENT

From an EA capability perspective, the core question in context of analytics and big data is which capabilities significantly would benefit from big data analytics. This section will investigate this question and derive requirements to big data analytics (BDA). Starting point for the discussion will be the capability catalog as part of an EAM maturity model developed by Wißotzki, which is published in [52] [53] .

Maturity models are specific management instruments, which define various degrees of maturities in order to evaluate to what extent a particular competency fulfils the qualitative requirements that are defined for a set of competency objects [54] and the development processes in organisations [55]. This abstract set of competency objects, which are representations of concrete objects from the real world. Beginning with very early stages of these entities, maturity models define anticipated, logical and consecutive development paths until observed objects reach an absolute maturity [56]. Having their origins in the software industry, maturity models are designed to measure the current state - the achieved level of competence - by means of assessment methods [57] [58].

The idea of EAM paradigm is modelling of the important enterprise elements and their relationships, which allows the analysis of as-is and target state dependencies [59]. In this context, EA models serve as maps with information of the current situation and strategies for future directions of the company [60]. Making the organizations more sensitive towards the interaction of business strategies, customers, application systems and organizational units, companies need to control enterprise-wide processes and adopt matching actions [61]. For this purposes, the concept of maturity was employed for EA which assigns different levels of achievement by means of a maturity assessment to processes, sub-processes, capabilities and characteristics [57].

Organizations will increasingly adopt maturity models to guide the development and implementation of their strategies. Yet it is a challenge for an organization to efficiently put the right capabilities into practice. In order to do so, organizations have to carry appropriate actions into execution, which later on should be turned into so-called "initiatives". For these actions to be taken there is a need for an integrated approach, which could be gained by implementing EAM. This is a prerequisite for an enhanced holistic enterprise view that reduces the management complexity of business objects, processes, strategies, information infrastructure and the relations between them. Nevertheless the successful adoption of EAM is accompanied by challenges that an

enterprise has to face and to overcome. In order to be able to implement the operationalized initiatives efficiently and achieve a specific outcome the enterprises require EAM capabilities.

In order to support this process, a capability model is to be developed and examined, which enables as-is performance assessment of EAM capabilities and suggests methods for their further development. The idea of constructing a EAM capability maturity model, called Enterprise Architecture Capability Navigator (EACN), was triggered by a research project from University of Rostock and alfabet AG (now Software AG) Berlin. Central purpose of the research is the identification of EAM capabilities and their transfer to a flexible, feature-related measurement system, which contains both the methodology for the determination and concepts for the further development of the relevant EAM capabilities of an enterprise. A capability is defined as the organization's capacity to successfully perform a unique business activity to create a specific outcome [62] and the ability to continuously deliver a certain business value in dynamically changing business environments [60]. An EAM capability describes a specific combination of know-how in terms of organizational knowledge, procedures and resources able to make this knowledge applicable in a specific process with appropriate resources to achieve a specific outcome for a defined enterprise initiative [53].

In this context, EACN is an elementary approach that identifies the EAM capabilities which are derived through structured processes and then gathered in an enterprise specific repository for an efficient operationalization of enterprise initiatives.
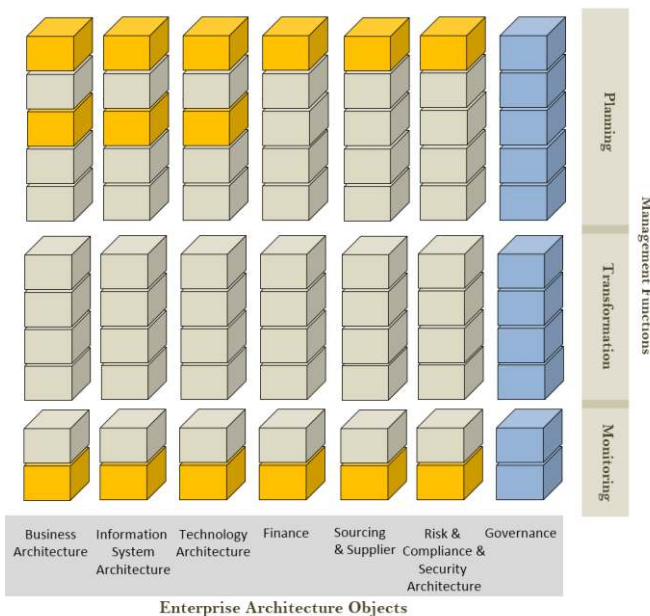

Fig. 5.   EAM Capability Solution Matrix

The construction of EACN elements is an ongoing process but still in this section we report the state of art. The EAM Capability Catalog represents one of the EACN core concepts. The EAM Capability Catalog involves a set of capabilities that are derived from the EAM Capability Solution Matrix. The solution matrix has two dimensions namely management functions (planning, transforming, monitoring) and EA objects (business architecture, information system architecture, technology architecture, finance, sourcing & supplier, Risk & compliance & security architecture, governance) and shows how the capabilities relate to each other (Figure 1).

The management functions are derived from [63]. The EA objects and its contents are constructed on the basis of The Open Group Architecture Framework [64]. An EAM Capability is performed on management functions and/ or different architecture objects. The capabilities can have i) information dependencies (information used to fulfill one capability that is produced by another), ii) support relationships (resources for use in other capabilities) and iii) close functional relationships (representing different aspects of the same area).

Fig. 6.   An exemplary EACN Capability (1st Layer)

In order to identify the most promising application areas for BDA within the capability catalog, we will use the dimensions management function and EA objects on the one hand side and the characteristics of big data on the other side. From the perspective of management functions, we have to consider the purpose of analytics, which basically is exploration and investigation of past business performance to gain insight and continue business planning. Since "planning" obviously is the premium function supported, all planning capabilities will be further investigated. Transformation is not considered a promising application field since it is primarily concerned with stepwise or iterative implementation of plans. Operation might also have benefits from analytics when using huge data volumes for detecting change needs in configurations and initiate future planning.

From the perspective of EA objects, we are particularly interested in objects exposed to the 3V challenges of big data presented in section III. Some of the EA objects obviously are neither dependent on high volume data nor on processing them fast: the core of "business architecture" planning is evaluating,

revising and adjusting strategic objectives, processes, services and their relationships to continuously changing demands. For enterprise level, this is not performed in short intervals requiring real-time data but on mid or long-term perspective. The same applies for "governance" of EA as such, but here the focus is on the governance process.

V. RELATED WORK

Up to now, the architecture of Big Data has been considered primarily from a technical point of view such as [65] and [66][67]. However, these approaches do not discuss the relationship of Big Data to enterprise architecture and strategy. A further example are the technical and performance considerations for moving the Big Data approach Hadoop into the cloud in [68]. New technologies such as the globally distributed spanner database attract a large amount of attention [69]. Approaches to develop an architecture for Big Data can also be found in [70]. Basic considerations addressing Big Data architecture are made in [71]. A service-oriented architecture for business intelligence systems is proposed in [72]. In [73] a framework using an Service-Oriented-Architectures is proposed to for real-time environments. An architecture adapted to requirements of the telecommunications industry is introduced in [74]. A delphi-study on business intelligence and service-oriented architecture is presented in [75]. A further service-oriented architecture for business intelligence is presented in [76]. Different ways to create multitenant architectures to support Big Data in the cloud are investigated in [77].

Several research approaches related to the framework created in this paper can be found in a number of papers. First approaches for moving Big Data into the cloud have been analyzed in [78]. Cloud-specific problems of data management are discussed in [79]. General technical considerations are found in [66] and [80]. The Big Data management system ASTERIX is introduced in [81]. A value chain for Big Data is conceptualized in [70]. First basic considerations on system architecture for Big Data are made in [82]. The relationship of cloud computing and Big Data in general is discussed in [83].

In [84] the term Decision as a Service is introduced for the separation of decisions logic from the application logic. However, no considerations about moving the decision service into the cloud are made. The decision process is completely put into one service. Our approach however differentiates into two dimensions. It allows to use separate services for every step of the process and to decide for each service either to use a cloud-service or to provide it locally.

One of the earliest examples for decision as a Service is Microsoft Tag [85]. Individual Quick Response Codes [86] (QR-Code) can be placed on products, newspapers etc. Using an app for mobile phones is provided. By following the link encoded in the QR-code, the uses may obtain information. At the same time, data and location of the activation is registered and can be analyzed. All services are provided as a cloud-service. For example, Whole Foods Market Inc. uses Microsoft-Tag to give their customers information about their products and track at the same time, where their customers are [84]. By this means, the effectiveness of marketing campaigns can be measured with much more detail.

Another area of related research are considerations how to deploy Big Data within Enterprise Architecture. The NIST cloud definition framework [87] differentiates four deployment models: private, public, hybrid and community. Private clouds are provided on premise only by the enterprise itself. Independent service provider delivers public cloud-services.. A hybrid cloud is a cloud composed of services provided both by private and public cloud-services. A community cloud is created provided by a community of cloud-users sharing their resources.

Big Data can be implemented in a public, private, hybrid or community cloud. A private Big Data pipeline is provided completely on premise. It provides a significantly better protection of intellectual property because the process definitions are no longer stored in the cloud-environment. A public Big Data pipeline uses only cloud-services by an independent cloud-service vendor. Another alternative is, that a community cloud provides the cloud-services used for implementing the Big Data pipeline. Splitting the Big Data pipeline between the cloud and on premise implementations defines a hybrid.

A purely managerial view without relation to enterprise architecture is used in research about information logistics such as [88]. A business oriented approach for structuring the business analytics lifecycle [89]. It consists of capturing information, analyzing information, aggregation and integration of information the use of gained inside to guide further strategies and the dissemination of information and insights. Only little research is done on the organizational impact of Big Data [90] and especially the influence on the role of IT departments in organizations.

VI. CONCLUSION

Many Enterprise Architecture Management projects still struggle to establish a sufficiently complete and up-to-date information basis for analysis, planning, transformation and operations. Especially information on lower levels of enterprise architecture are difficult to obtain and especially to keep up-to-date.

We introduced a new approach, Enterprise Architecture Analytics to provide sufficient and up-to-date information especially on the lower abstraction levels of enterprise architecture. Enterprise Architecture Analytics leverages the huge amounts of data available in today's virtualized infrastructures and exploits those using technologies from the context of Big Data. By this means it is possible to use both semi- and un-structured data for infrastructure systems. Enterprise Architecture Analytics embraces both a static and a dynamic perspective. The static perspective describes the aggregation and association of architectural elements, the dynamic perspective depicts the usage of architectural elements. In this way Enterprise Architectures may be discovered, analyzed and optimized. By combining the static and dynamic perspective it is possible to check the compliance of

Enterprise Architectures in operation. By applying technologies from the Big Data context, it is possible to decrease the latency between bottom-up initiated changes of architectural elements and their representation in the Enterprise Architecture model. Analysis results are done using fully distributed algorithms and use less intermediate data products.

By enlarging the set of available data and speeding up analysis, Enterprise Architecture Analytics strengthens the role of Enterprise Architecture as trigger for change processes in enterprises. Due to the reduced latency of analyses it possible to use increasingly agile methods for Enterprise Architecture Management. In this way Enterprise Architectures Analytics paves a way for a "real-time" Enterprise Architecture Management. This vision could materialize in an Enterprise Architecture Management Cockpit (EAMP) that visualizes the present status of Enterprise Architecture.

In parallel to this vision, established applications of Enterprise Architecture Management such as compliance checks and measurements of maturity profit from the reduced time needed for analysis. Furthermore, the reliability of compliance checks and maturity measurements is improved by the increased amount of data taken into account for investigation.

Based on knowledge gained by Enterprise Architecture Analytics, the past, current and future use of each element of the enterprise architecture becomes visible. New pricing models and cost savings can be implemented. In this way standing of the IT-departments can may become stronger.

Part of the future work is to evaluate the use of Enterprise Architecture Analytics in different contexts, and to investigate how to establish a suitable information and decision support basis. Industry and other context specific criteria can be proved by implementation of an empirical study (e.g. expert interviews or empirical experiments in different enterprises of different industry sectors). Furthermore risks of a data driven analysis (e.g. forecast error, metric problems) must be observed.

VII. REFERENCES

[1] Kai Hwang, Geoffrey C. Fox, and J. J. Dongarra, *Distributed and Cloud Computing: From Parallel Processing to the Internet of Things*. Morgan Kaufmann Pub, 2012.

[2] N. Ayachitula, M. Buco, Y. Diao, B. Fisher, D. Loewenstern, and C. Ward, "IT Service Management Automation–An Automation Centric Approach Leveraging Configuration Control, Audit Verification and Process Analytics," *Manag. Virtualization Netw. Serv.*, pp. 195–198, 2007.

[3] C. White, "Converting Big Data Hype into Big Value with Analytics," BI Research, Oct. 2012.

[4] J. C. Corbett, J. Dean, M. Epstein, A. Fikes, C. Frost, J. J. Furman, S. Ghemawat, A. Gubarev, C. Heiser, and P. Hochschild, "Spanner: Google's Globally-Distributed Database," *Appear Proc. OSDI*, p. 1, 2012.

[5] T. White, *Hadoop: The definitive guide*. O'Reilly Media, 2012.

[6] OASIS, "Reference Model for Service Oriented Architecture 1.0," 08-Feb-2006. [Online]. Available: http://www.oasis-open.org/committees/download.php/19679/soa-rm-cs.pdf. [Accessed: 08-Apr-2009].

[7] J. A. Estefan, K. Laskey, F. McCabe, and P. Thornton, "Reference architecture foundation for service oriented architecture version 1.0," *OASIS Comm. Draft*, vol. 2, no. 14, p. 2009, 2009.

[8] A. Arsanjani, N. Kumar, C. Harding, M. Gejnevall, T. Carrato, H. Kreger, and J. Diaz, "Soa reference architecture," *Open Group Httpwww Theopengroup Orgprojectssoa-Refarch Tech Rep*, 2009.

[9] L. Bass, P. Clements, and R. Kazman, *Software architecture in practice*. Upper Saddle River, NJ: Addison-Wesley, 2013.

[10] R. Sweeney, *Achieving service-oriented architecture: applying an enterprise architecture approach*. Wiley, 2010.

[11] M. O. 't Land, E. Proper, M. Waage, J. Cloo, and C. Steghuis, "The Enterprise Architect," in *Enterprise Architecture*, Springer Berlin Heidelberg, 2009, pp. 113–125.

[12] J. Schekkerman, *Enterprise architecture good practices guide: how to manage the enterprise architecture practice*. Victoria, BC: Trafford Pub., 2008.

[13] A. Zimmermann, M. Pretz, G. Zimmermann, D. G. Firesmith, and E. El-Sheikh, "Towards Service-oriented Enterprise Architecture for Big data Applications in the Cloud," in *IEEE-EDOCW*, Vancouver, Canada, 2013, pp. 130–135.

[14] V. Haren, "TOGAF Version 9.1," 2011.

[15] V. Haren, *Archimate 2.0 Specification*. Van Haren Publishing Series. Van Haren Publishing, 2012.

[16] T. Erl, *Service-oriented architecture*, vol. 8. Prentice Hall New York, 2005.

[17] L.-J. Zhang, J. Zhang, and H. Cai, *Services computing*. Springer, 2007.

[18] D. K. Barry, *Web Services, Service-Oriented Architectures, and Cloud Computing: The Savvy Manager's Guide (The Savvy Manager's Guides)*, 2 edition. Morgan Kaufmann, 2012.

[19] R. Buyya, J. Broberg, and A. Gościński, *Cloud computing: principles and paradigms*. Hoboken, N.J.: Wiley, 2011.

[20] N. Antonopoulos and L. Gillam, *Cloud Computing: Principles, Systems and Applications (Computer Communications and Networks)*, 2010 edition. Springer, 2010.

[21] F. Liu, J. Tong, J. Mao, R. Bohn, J. Messina, L. Badger, and D. Leaf, "NIST cloud computing reference architecture," *NIST Spec. Publ.*, vol. 500, p. 292, 2011.

[22] M. Behrendt, B. Glasner, P. Kopp, R. Dieckmann, G. Breiter, S. Pappe, H. Kreger, and A. Arsanjani, "Introduction and Architecture Overview IBM Cloud Computing Reference Architecture 2.0," *Draft Version V*, vol. 1, no. 0, 2011.

[23] G. Brunette and R. Mogull, "Security guidance for critical areas of focus in cloud computing v2. 1," *Cloud Secur. Alliance*, pp. 1–76, 2009.

[24] "SOCCI Framework Technical Standard." [Online]. Available: http://www.opengroup.org/soa/source-book/socci/index.htm. [Accessed: 16-Oct-2012].

[25] A. Zimmermann, K. Sandkuhl, M. Pretz, M. Falkenthal, D. Jugel, and M. Wissotzki, "Towards an integrated service-oriented reference enterprise architecture," in *Proceedings of the 2013 International Workshop on Ecosystem Architectures*, 2013, pp. 26–30.

[26] Open Group, "Service-Oriented Architecture Ontology," Open Group.

[27] E. Brynjolfsson, L. Hitt, and H. Kim, "Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?," 2011.

[28] R. Schmidt, "A framework for comparing cloud-environments," presented at the Computer Science and Information Systems (FedCSIS), 2011 Federated Conference on, 2011, pp. 553–556.

[29] A. McAfee and E. Brynjolfsson, "Big data: the management revolution.," *Harv. Bus. Rev.*, vol. 90, no. 10, p. 60, 2012.

[30] J. Bughin, M. Chui, and J. Manyika, "Clouds, big data, and smart assets: Ten tech-enabled business trends to watch," *McKinsey Q.*, vol. 56, 2010.

[31] S. LaValle, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, "Big data, analytics and the path from insights to value," *MIT Sloan Manag. Rev.*, vol. 52, no. 2, pp. 21–32, 2011.

[32] P. C. Zikopoulos, C. Eaton, and P. Zikopoulos, *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. Mcgraw-Hill Professional, 2012.

[33] "Bursting the Big Data Bubble | ZDNet." [Online]. Available: http://www.zdnet.com/bursting-the-big-data-bubble-7000002352/. [Accessed: 13-Aug-2012].

[34] H. Chen, R. H. L. Chiang, and V. C. Storey, "Business intelligence and analytics: From big data to big impact," *MIS Q.*, vol. 36, no. 4, pp. 1165–1188, 2012.

[35] R. Schmidt and M. Möhring, "Strategic alignment of Cloud-based Architectures for Big Data," in *Proceedings of the 17th IEEE International Enterprise Distributed Object Computing Conference Workshops (EDOCW)*, Vancouver, Canada, 2013.

[36] An interview with Erik Brynjolfsson and Andrew McAfee, "Beyond Enterprise 2.0," 01-Apr-2007. [Online]. Available: http://sloanreview.mit.edu/smr/issue/2007/spring/16/. [Accessed: 27-Aug-2008].

[37] M. Möhring, R. Schmidt, N. Wolfrum, M. Kammerer, S. Maier, and S. Höritz, "How Big Data Transforms the IT Department to a Strategic Weapon," in *Proceedings of the IADIS International Conference Information Systems 2013*, Lisbon, 2013, pp. 323–326.

[38] R. Schmidt, M. Möhring, S. Maier, J. Pietsch, and R.-C. Härting, "Big Data as Strategic Enabler - Insights from Central European Enterprises," in *In: Abramowicz, W., Kokkinaki, A. (Hrsg.). 17th International Conference on Business Information Systems (BIS) (in press). Springer LNBIP*, Larnaca, Cyprus, 2014.

[39] S. Chaudhuri and U. Dayal, "An overview of data warehousing and OLAP technology," *ACM Sigmod Rec.*, vol. 26, no. 1, pp. 65–74, 1997.

[40] S. Aier, M. Ahrens, M. Stutz, and U. Bub, "Deriving SOA evaluation metrics in an enterprise architecture context," in *Service-Oriented Computing-ICSOC 2007 Workshops*, 2009, pp. 224–233.

[41] A. Vasconcelos, P. Sousa, and J. Tribolet, "Information system architecture metrics: an enterprise engineering evaluation approach," *Electron. J. Inf. Syst. Eval.*, vol. 10, no. 1, pp. 91–122, 2007.

[42] P. Weirich, *Decision space: Multidimensional utility analysis*. Cambridge University Press, 2001.

[43] G. E. Box, G. M. Jenkins, and G. C. Reinsel, *Time series analysis: forecasting and control*. Wiley. com, 2013.

[44] N. R. Draper, H. Smith, and E. Pownell, *Applied regression analysis*, vol. 3. Wiley New York, 1966.

[45] D. E. Rumelhart, G. E. Hintont, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[46] D. Gmach, J. Rolia, L. Cherkasova, and A. Kemper, "Capacity management and demand prediction for next generation data centers," in *Web Services, 2007. ICWS 2007. IEEE International Conference on*, 2007, pp. 43–50.

[47] J. T. Connor, R. D. Martin, and L. E. Atlas, "Recurrent neural networks and robust time series prediction," *Neural Netw. IEEE Trans. On*, vol. 5, no. 2, pp. 240–254, 1994.

[48] G. Leitch and J. E. Tanner, "Economic forecast evaluation: profits versus the conventional error measures," *Am. Econ. Rev.*, pp. 580–590, 1991.

[49] L. Cao and A. S. Soofi, "Nonlinear deterministic forecasting of daily dollar exchange rates," *Int. J. Forecast.*, vol. 15, no. 4, pp. 421–430, 1999.

[50] A. Ross and K. Ernstberger, "Benchmarking the IT productivity paradox: Recent evidence from the manufacturing sector," *Math. Comput. Model.*, vol. 44, no. 1, pp. 30–42, 2006.

[51] N. G. Carr, "IT Doesn't Matter," *IEEE Eng. Manag. Rev.*, vol. 32, no. 1, pp. 24–32, 2004.

[52] M. Wißotzki and H. Koc, "A Project Driven Approach for Enhanced Maturity Model Development for EAM Capability Evaluation," presented at the IEEE-EDOCW, Vancouver, Canada, 2013, pp. 296–305.

[53] M. Wissotzki, H. Koc, T. Weichert, and K. Sandkuhl, "Development of an Enterprise Architecture Management Capability Catalog," in *Proceedings BIR 2013*, 2013, vol. 158, pp. 112–126.

[54] R. Wendler, "Reifegradmodelle für das IT-Projektmanagement," Techn. Univ. Fak. Wirtschaftswiss, Dresden, 2009.

[55] A. Back, "Reifegradmodelle im Management von Enterprise 2.0." BITKOM, 2010.

[56] J. Becker, R. Knackstedt, and J. Pöppelbuß, "Dokumentationsqualität von Reifegradmodellentwicklungen," 2009.

[57] M. Meyer, M. Helfert, and C. O'Brien, "An Analysis of Enterprise Architecture Maturity Frameworks," in *Perspectives in Business Informatics Research*, Springer, 2011, pp. 167–177.

[58] T. De Bruin, R. Freeze, U. Kaulkarni, and M. Rosemann, "Understanding the main phases of developing a maturity assessment model," 2005.

[59] I. S. Aier, D.-W.-I. C. Riege, and R. Winter, "Unternehmensarchitektur–Literaturüberblick und Stand der Praxis," *Wirtschaftsinformatik*, vol. 50, no. 4, pp. 292–304, 2008.

[60] J. Stirna, J. Grabis, M. Henkel, and J. Zdravkovic, "Capability Driven Development–An Approach to Support Evolving Organizations," in *The Practice of Enterprise Modeling*, Springer, 2012, pp. 117–131.

[61] M. Wißotzki and A. Sonnenberger, "Enterprise architecture management - state of research analysis & a comparison of selected approaches," in *Proceedings of the 5th IFIP WG 8.1 Working Conference on the Practice of Enterprise Modeling*, Rostock, 2012.

[62] J. Scott and A. Cullen, "Business capabilities provide the rosetta stone for Business-IT Alignment: Capability Maps Are A Foundation For Business Architecture."

[63] F. Ahlemann, E. Stettiner, M. Messerschmidt, and C. Legner, *Strategic Enterprise Architecture Management: Challenges, Best Practices, and Future Developments*. 2012.

[64] T. O. Group, *TOGAF 9*. 2011.

[65] V. R. Borkar, M. J. Carey, and C. Li, "Inside 'Big Data management': Ogres, onions, or parfaits?'," in *Proc. EDBT*, 2012.

[66] S. Sakr, A. Liu, and A. G. Fayoumi, "The Family of MapReduce and Large Scale Data Processing Systems," *ArXiv Prepr. ArXiv13022966*, 2013.

[67] E. E. Schadt, M. D. Linderman, J. Sorenson, L. Lee, and G. P. Nolan, "Computational solutions to large-scale data management and analysis," *Nat. Rev. Genet.*, vol. 11, no. 9, pp. 647–657, Sep. 2010.

[68] Y. Pingle, V. Kohli, S. Kamat, and N. Poladia, "Big Data Processing using Apache Hadoop in Cloud System," in *National Conference on Emerging Trends in Engineering & Technology*, 2012.

[69] J. C. Corbett, J. Dean, M. Epstein, A. Fikes, C. Frost, J. J. Furman, S. Ghemawat, A. Gubarev, C. Heiser, and P. Hochschild, "Spanner: Google's Globally-Distributed Database," *Appear Proc. OSDI*, p. 1, 2012.

[70] H. G. Miller and P. Mork, "From Data to Decisions: A Value Chain for Big Data," *IT Prof.*, vol. 15, no. 1, pp. 57–59, 2013.

[71] J. Ranjan, "Service-oriented architecture for business intelligence: a research agenda," *Int. J. Agile Syst. Manag.*, vol. 4, no. 4, pp. 301–318, 2011.

[72] M. Marefati and S. M. Hashemi, "Service Oriented Architecture for Business Intelligence Systems."

[73] J. Javanmard, S. Moaven, and J. Habibi, "Introducing a framework to use SOA in business intelligence for real-time environments," in *Software Engineering and Service Science (ICSESS), 2011 IEEE 2nd International Conference on*, 2011, pp. 94–99.

[74] T. Ishaya and M. Folarin, "A service oriented approach to Business Intelligence in Telecoms industry," *Telemat. Inform.*, vol. 29, no. 3, pp. 273–285, 2012.

[75] R. M. Müller, S. Linders, and L. F. Pires, "Business intelligence and service-oriented architecture: a Delphi study," *Inf. Syst. Manag.*, vol. 27, no. 2, pp. 168–187, 2010.

[76]   L. Wu, G. Barash, and C. Bartolini, "A service-oriented architecture for business intelligence," in *Service-Oriented Computing and Applications, 2007. SOCA'07. IEEE International Conference on*, 2007, pp. 279–285.

[77]   D. Agrawal, S. Das, and A. El Abbadi, "Big data and cloud computing: current state and future opportunities," in *Proceedings of the 14th International Conference on Extending Database Technology*, 2011, pp. 530–533.

[78]   S. Francia, "Myths of Cloud-Computing," in *Bad Data Handbook: Cleaning Up The Data So You Can Get Back To Work*, 1st ed., O'Reilly Media, 2012, pp. 175–185.

[79]   D. J. Abadi, "Data management in the cloud: Limitations and opportunities," *IEEE Data Eng Bull*, vol. 32, no. 1, pp. 3–12, 2009.

[80]   V. Borkar, M. J. Carey, and C. Li, "Inside Big Data management: ogres, onions, or parfaits?," in *Proceedings of the 15th International Conference on Extending Database Technology*, 2012, pp. 3–14.

[81]   V. R. Borkar, M. J. Carey, and C. Li, "Big data platforms: what's next?," *XRDS Crossroads ACM Mag. Stud.*, vol. 19, no. 1, pp. 44–49, 2012.

[82]   A. E. Gattiker, F. H. Gebara, A. Gheith, H. P. Hofstee, D. A. Jamsek, J. Li, E. Speight, J. W. Shi, G. C. Chen, and P. W. Wong, "Understanding System and Architecture for Big Data," 2012.

[83]   L.-J. L. Zhang, "Editorial: Big Services Era: Global Trends of Cloud Computing and Big Data," *IEEE Trans. Serv. Comput.*, pp. 467–468, 2012.

[84]   A. Zarghami, B. Sapkota, M. Z. Eslami, and M. van Sinderen, "Decision as a service: Separating decision-making from application process logic," in *Enterprise Distributed Object Computing Conference (EDOC), 2012 IEEE 16th International*, 2012, pp. 103–112.

[85]   "2D Barcodes and Mobile Tagging | Microsoft Tag." [Online]. Available: http://tag.microsoft.com/home.aspx. [Accessed: 21-Apr-2013].

[86]   M. Weiser, "Ubiquitous computing," *Computer*, vol. 26, no. 10, pp. 71–72, 1993.

[87]   P. Mell and T. Grance, "The NIST Definition of Cloud Computing," 07-Oct-2009. [Online]. Available: http://csrc.nist.gov/groups/SNS/cloud-computing/. [Accessed: 06-Jan-2011].

[88]   R. Winter, "Enterprise-wide information logistics: Conceptual foundations, technology enablers, and management challenges," in *Information Technology Interfaces, 2008. ITI 2008. 30th International Conference on*, 2008, pp. 41–50.

[89]   D. Kiron, P. K. Prentice, and B. Ferguson, "Innovating with Analytics," *MIT Sloan Manag. Rev.*, vol. 54, no. 1, pp. 47–52, Fall 2012.

[90]   R. Schmidt, "Data Science & Big Data." 05-Nov-2012.