eKNOW 2019

The Eleventh International Conference on Information, Process, and Knowledge Management

ISBN: 978-1-61208-690-3

February 24 – 28, 2019

Athens, Greece

eKNOW 2019 Editors

Jaime Lloret Mauri, Universitat Politecnica de Valencia, Spain

eKNOW 2019

Forward

The eleventh edition of the International Conference on Information, Process, and Knowledge Management (eKNOW 2019) was held in Athens, Greece, February 24 - 28, 2019. The event was driven by the complexity of the current systems, the diversity of the data, and the challenges for mental representation and understanding of environmental structure and behavior.

Capturing, representing, and manipulating knowledge was and still is a fascinating and extremely useful challenge from both theoretical and practical perspective. Using validated knowledge for information and process management and for decision support mechanisms raised a series of questions the eKNOW 2019 conference was aimed at.

eKNOW 2019 provided a forum where researchers were able to present recent research results and new research problems and directions related to them. The topics covered aspects from knowledge fundamentals to more specialized topics such as process analysis and modeling, management systems, semantics processing and ontology.

We take this opportunity to thank all the members of the eKNOW 2019 Technical Program Committee as well as the numerous reviewers. The creation of such a broad and high-quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to the eKNOW 2019. We truly believe that, thanks to all these efforts, the final conference program consists of top quality contributions.

This event could also not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the eKNOW 2019 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that eKNOW 2019 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in knowledge management research.

We also hope that Athens provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

eKNOW 2018 Chairs

eKNOW Steering Committee
Roy Oberhauser, Aalen University, Germany
Andrzej M.J. Skulimowski, AGH University of Science and Technology - Krakow, Poland
Dirk Malzahn, Dirk Malzahn Ltd. / HfH University, Germany
Peter Bellström, Karlstad University, Sweden
Susan Gauch, University of Arkansas, USA
Edy Portmann, Institute of Information Systems - University of Bern, Switzerland
Nitin Agarwal, University of Arkansas at Little Rock, USA

eKNOW Industry/Research Advisory Committee
Gil Ad Ariely, Interdisciplinary Center Herzliya (IDC), Israel
Daniel Kimmig, solute GmbH, Germany
Mauro Dragoni, Fondazione Bruno Kessler, Italy
Dayu Yuan, Google Inc., USA
Faisal Azhar, HP UK LTD, UK

eKNOW 2019

COMMITTEE

eKNOW Steering Committee
Roy Oberhauser, Aalen University, Germany
Andrzej M.J. Skulimowski, AGH University of Science and Technology - Krakow, Poland
Dirk Malzahn, Dirk Malzahn Ltd. / HfH University, Germany
Peter Bellström, Karlstad University, Sweden
Susan Gauch, University of Arkansas, USA
Edy Portmann, Institute of Information Systems - University of Bern, Switzerland
Nitin Agarwal, University of Arkansas at Little Rock, USA

eKNOW Industry/Research Advisory Committee
Gil Ad Ariely, Interdisciplinary Center Herzliya (IDC), Israel
Daniel Kimmig, solute GmbH, Germany
Mauro Dragoni, Fondazione Bruno Kessler, Italy
Dayu Yuan, Google Inc., USA
Faisal Azhar, HP UK LTD, UK

eKNOW 2019 Technical Program Committee

Gil Ad Ariely, Interdisciplinary Center Herzliya (IDC), Israel
Chris Adetunji, University of Southampton, UK
Nitin Agarwal, University of Arkansas at Little Rock, USA
Zeyar Aung, Khalifa University of Science and Technology, UAE
Faisal Azhar, Hewlett Packard (HP), UK
Michal Baczynski, University of Silesia in Katowice, Poland
Zbigniew Banaszak, Koszalin University of Technology, Poland
Gianni Barlacchi, University of Trento, Italy
Peter Bellström, Karlstad University, Sweden
Martine Cadot, LORIA - University Henri Poincaré Nancy I, France
Ali Çakmak, Istanbul Sehir University, Turkey
Enrico G. Caldarola, STIIMA - CNR, Bari, Italy
Ricardo Campos, Polytechnic Institute of Tomar / LIAAD INESC TEC, Portugal
Massimiliano Caramia, University of Rome "Tor Vergata", Italy
Shayok Chakraborty, Arizona State University, USA
Dickson Chiu, The University of Hong Kong, Hong Kong
Ritesh Chugh, Central Queensland University, Australia
Paolo Cintia, University of Pisa / KDDLab Isti Cnr, Italy
Marco Cococcioni, University of Pisa, Italy
Andrea De Salve, Institute of Informatics and Telematics (IIT) - National Research Council (CNR), Italy
Cláudio de Souza Baptista, Federal University of Campina Grande, Brazil
Chiara Di Francescomarino, Fondazione Bruno Kessler (FBK), Italy
Giuseppe A. Di Lucca, University of Sannio - RCOST (Research Center on Software Technology), Italy
Vasiliki Diamantopoulou, University of the Aegean, Greece
Mihaela Dinsoreanu, Technical University of Cluj-Napoca, Romania

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

An Exploratory Study on Notational Characteristics of Visual Notations Used in Decision Management

Sam Leewis
Optimizing Knowledge-Intensive Business Processes
Zuyd University of Applied Sciences
Sittard, The Netherlands
sam.leewis@zuyd.nl

Koen Smit, Matthijs Berkhout, Ruben Post
Digital Smart Services
HU University of Applied Sciences Utrecht
Utrecht, The Netherlands
koen.smit@hu.nl, matthijs.berkhout@hu.nl,
ruben.post@hu.nl

Martijn Zoet
Optimizing Knowledge-Intensive Business Processes
Zuyd University of Applied Sciences
Sittard, The Netherlands
martijn.zoet@zuyd.nl

***Abstract*— The visual representation of Information System (IS) artefacts is an important aspect in the practical application of visual representations. However, important and known visual representation principles are often undervalued, which could lead to decreased effectiveness in using a visual representation. Decision Management (DM) is one field of study in which stakeholders must be able to utilize visual notations to model business decisions and underlying business logic, which are executed by machines, thus are IS artefacts. Although many DM notations currently exist, little research actually evaluates visual representation principles to identify the visual notations most suitable for stakeholders. In this paper, the Physics of Notations framework of Moody is operationalized and utilized to evaluate five different DM visual notations. The results show several points of improvement with regards to these visual notations. Furthermore, the results could show the authors of DM visual notations that well-known visual representation principles need to be adequately taken into account when defining or modifying DM visual notations.**

Keywords-Decision Management; Visual Notations; Evaluation; Physics of Notations (PoN)

I. INTRODUCTION

Decisions are amongst the most important assets of an organization [1], and therefore should be managed adequately. A decision is defined as: "*A conclusion that a business arrives at through business logic and which the business is interested in managing*" [2]. Furthermore, business logic can be defined as "*a collection of business rules, business decision tables, or executable analytic models to make individual business decisions*" [3]. Examples of decisions are: 1) determine what illness a patient has, 2) determine the loan default risk factor for a specific customer, or 3) determine the maximum credit rating of an organization. If an organization can't consistently make and execute the right decision(s), large risks are taken that can eventually lead to high costs, reputation damage, or even bankruptcy. Following the previous example, imagine what will happen when a doctor makes the wrong decision continuously or a customer with a high-risk classification gets a low-risk classification.

One important aspect of Decision Management (DM) is modelling decisions and business logic using a visual representation. Such visual representations are often referred to as notations or modeling standards. An example of a decision modelling notation is the Decision Modeling and Notation (DMN) proposed by the Object Management Group [2] or The Decision Model, defined by von Halle and Goldberg [4].

While empowering the semantic modeling capabilities of notations is desirable, notations also need to be cognitively effective [5]. Cognitive effectiveness, in the context of visual notations, refers to ''*the speed, ease and accuracy with which a representation can be processed by the human mind*'' [6]. Generally speaking, important and known visual representation principles are often undervalued in the design of visual notations, which could lead to decreased cognitive effectiveness[7], [8]. Furthermore, these notations are usually not designed with all stakeholders in mind, from someone who never modelled on the one hand (Decision modelling novice) to a Decision modelling expert on the other hand [6], and have no design rationale nor a scientific basis for the choices in the structure of the visual representation [5]. Decision modelling novices do have different requirements in comparison to users who are considered a DM expert. An expert will need more advanced functionalities in comparison to a novice, however, a novice should be able to learn the notation quickly to get started.

This paper examines whether these problems exist in the notations specifically designed for the DM domain, as, to the knowledge of the authors, no earlier studies exist that focus on evaluating multiple DM notations. To do so, a

 ISBN: 978-1-61208-690-3

proper framework to evaluate known visual representation principles needs to be selected.

Several frameworks to evaluate visual notations exist, for example, the Cognitive Dimensions framework [9], the ontological analysis framework [10], and the Guidelines of Modelling (GoM) framework [11]. The most complete and referenced framework on the assessment of visual notations is the Physics of Notations (PoN) theory [6]. This theory is partly based on the Cognitive Dimensions framework, which was the predominant theoretical paradigm in visual notations research [12]. The framework is developed and devoted to design, evaluate, and compare visual notations and is based on theory and empirical evidence obtained from different disciplines, such as perceptual psychology, cognitive psychology, cartography, graphic design, human-computer interfacing, linguistics, and communication theory. Furthermore, one advantage of the PoN framework is that it also offers clear evaluation procedures and metrics so that researchers can easily operationalize them to be evaluated in practice. The PoN framework has been applied by many researchers to evaluate visual notations [13][14]. Since we selected a framework to evaluate DM visual notations with, the following research question is stated: "*How do the selected DM visual notations score with regards to the PoN framework?*"

The rest of the paper is structured as follows. First, the theory underlying visual notations and PoN are elaborated upon in the background and related work. This is followed by the research method utilized to conduct the research presented in this paper. Then, the data collection and analysis processes are explained. Next, in the results section, the PoN scores for the selected visual notations are presented. Lastly, the paper concludes with a discussion, conclusions, and directions for future research.

II. BACKGROUND AND RELATED WORK

DM notations can best be categorized by their complexity and linguistic power. Complexity refers to the ease of understanding the DM notation and linguistic power refers to the amount of results it can produce, indicating its richness. Five different types of DM notations have been defined: 1) labels, textual markers, 2) graphical aids, symbols representing semantic constructs, 3) structured languages, semantic representations of logic, 4) constrained natural languages, ontology defined by base terms and grammar, and 5) pure natural languages, unbound syntax, see Figure 1.

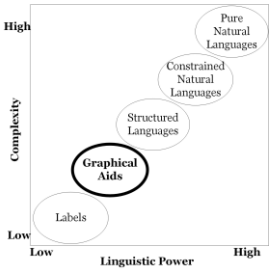


Figure 1. DM notation categorization [15]

The PoN framework is aimed towards visually represented DM notations. Therefore, this study evaluated DM visual notations of the graphical aids type.

The selected elements are drawn from the PoN framework [6]. This framework attempts to evaluate DM notations based on their visual representation, as these are often undervalued principles. It offers nine different principles by which the visual representation of a DM notation is measured against. The principles are as follows [6]:

Semiotic Clarity refers to every symbol having a one-to-one correspondence to their referent concept. If not, one or more of the following four anomalies can occur: 1) symbol redundancy occurs when multiple symbols can be used to represent the same concept, 2) symbol overload occurs when different concepts can be represented by the same symbol, 3) symbol excess occurs when symbols do not correspond to any concept, and 4) symbol deficit occurs when there are concepts that do not correspond with any symbols.

Perceptual Discriminability refers to the ability to differentiate symbols based on their graphical appearance. This can be improved by increasing the number of graphical attributes a symbol represents. For example, adding color, additional shapes, or text to a notation can improve the ability to differentiate between symbols.

Semantic transparency refers to the extent to which the meaning of a symbol can be inferred from its appearance. For example, a rectangle representing a decision has a scarce semantic transparency, while an icon of a calculator representing a formula has high semantic transparency.

Complexity Management refers to the ability of a visual notation to represent information without overloading the human brain. The complexity our brains can handle can be improved by the usage of different concepts. For example, modularization can be used to reduce the complexity of a large system by dividing it into smaller parts or making use of subsystems. Additionally, hierarchy can be incorporated into the notation by representing information on different levels of details.

Cognitive Integration refers to the extent to which a notation enables multiple diagrams to represent a system without overloading the human brain. This can be supported by two concepts, conceptual integration and perceptual integration. Conceptual integration can be achieved by providing a summary diagram as a whole or parts of the diagram or by contextualization, a technique where contextual information on each diagram is showing its relation to elements on other diagrams. Perceptual integration is achieved by providing navigational tools in the notation. Commonly used navigational tools are, for example, lines to provide direction of the flow or a map in which the entire diagram is shown if only a part of the diagram is to be shown on the screen.

 ISBN: 978-1-61208-690-3

Visual Expressiveness is defined as the number of visual variables used in a notation. If a notation has a high number of visual variables, the perceptual discriminability increases, making the notation easier to use. Visual variables are size, brightness, color, texture, shape, orientation, and text.

Dual Coding refers to the use of both visual and textual attributes in a notation. For example, the semantic transparency can be increased by adding a keyword of the semantic concept to the visual representation of the symbol, consequently achieving dual coding.

Graphic Economy refers to the number of graphical symbols used in a notation. The human brain can discriminate around six categories simultaneously, defining the limit of graphical symbols a notation should contain. There are three concepts by which excessive graphic complexity can be reduced: 1) reduce semantic complexity, 2) introduce symbol deficit, and 3) increase visual expressiveness.

Cognitive Fit refers to the Cognitive fit theory, which states that different methods of representation of information are suitable for different tasks and different audiences. This can be respected by creating multiple visual filters for, for example, expert-novice differences or representational mediums.

III. RESEARCH METHOD

The goal of this research is to evaluate several DM visual notations with regards to the PoN framework [6]. When selecting an appropriate research method, one should take into account the maturity of the research domain [16]. Research with regards to visual notations to express business decisions and business logic is scarce [11]. Therefore, a qualitative research approach is selected as our research method.

To evaluate DM visual notations, a structured technique must be selected. We utilize a technique to do so from the body of knowledge regarding visual notations, as it is rather mature, compared to the body of knowledge on DM. Based on the PoN framework, the researchers constructed a template which covers the nine principles of visual notations indicated in [6] (the template itself is available upon request and omitted due to space limitations). Each of the nine principles consists of specific elements characterizing each principle, e.g., the principle '*semiotic clarity*' has four elements of which one represents '*symbol overload*'. Every element is represented by a question whether the element is available in the visual notation, and if present, to what extent.

Instead of using a quantitative approach, it is more appropriate to use a mix of quantitative collection and analysis with qualitative thematic coding, as our template also aims to collect motivations of researchers evaluating the visual notations. The coding of the evaluations for the selected visual notations consists of three rounds of thematic coding according to the process of open coding, axial coding, and selective coding described in [17]. During the coding rounds, four researchers coded the five graphical aids-type notations separately from each other. The results of the coding rounds were compared and their meaning discussed among the four researchers. The process of data collection and analysis is described in more detail in the following section.

IV. DATA COLLECTION AND ANALYSIS

Before the data collection and analysis started, the research team needed to decide which visual notations to evaluate. For this study, the amount of visual notations to evaluate was five. The DM visual notations are selected based on the following criteria: 1) the notation should be applied in practice by multiple organizations, 2) the documentation for the notation should be accessible to be able to evaluate it in detail, and 3) the notation should be a DM graphical aid type. The selected visual notations for evaluation are: *Beinformed* [18], *Berkeley Bridge* [19], *Decision Model and Notation (DMN)* [2], *The Decision Model (TDM)* [4], and *Visual Rules* [20].

The data collection for this study occurred over a period of two months, between March 2018 and April 2018. The data collection is conducted by four researchers representing different levels of expertise on visual notations. Two researchers representing the expert group (researcher 1 and 2) and two researchers representing the novice group (researcher 3 and 4). Separating the coders increases the inter-reliability in the coding [21] and internal validity of the research [22]. Researcher 1 is a lecturer and postdoc researcher with seven years of practical and research experience in the field of DM; Researcher 2 is a PhD-candidate with five years of practical and research experience in the field of DM; Researcher 3 is a Master student with four years of practical and research experience in the field of DM; Researcher 4 is a Bachelor student with two years of research experience in the field of DM. All researchers have experience with visual notations, and completed at least two or more projects in which DM models had to be produced to be utilized in practice. It took the research team a week to gather all data required to evaluate the visual notations. The data consisted of webpages, client case documents, learning documents, meta-models, demo applications, and video repositories with tutorials.

A template is created and utilized by the researchers to cover the nine principles of Moody [6]. Every principle has it's own characteristics and thereby every principle in the template has different elements with each their related questions. For each element, a five-point Likert-scale ranging from 1) very poor; 2) poor; 3) neutral; 4) good; 5) very good. Additionally, the value 6) Not Applicable (NA) could be chosen. If NA was chosen it needed to be further specified why. Therefore, the dataset represents a total of four filled-in templates for each of the five visual notations selected.

 ISBN: 978-1-61208-690-3

The data analysis comprised three rounds of thematic coding based on the data analysis techniques described by Strauss & Corbin [17]. The first round of coding identifies the symbols and constructs of each notation, e.g., the different node-types as part of the BeInformed visual notation or the transition-types as part of DMN.

TABLE 1. EXAMPLE CODING NOTATION.

		Visual notation: BeInformed			
		Coders			
		Expert		Novice	
		R1	R2	R3	R4
Perceptual Discriminability	**Redundant coding**	4	4	4	4
	Perceptual popout	3	4	3	2
	Textual differentiation			2	3
	Iconic differentiation	2	1	3	4

The second round of coding refines and differentiates concepts that are already available and code them into categories [23]. The axial coding round consisted of the indication of the values (using a five-point Likert-scale) for each visual notation together with the principles of Moody [6], as shown in Table 1.

The first and second coding rounds were based on knowledge derived from sources described earlier, however, the coders did not follow courses or applied the visual notation in practice for this specific research.

The third and last round of coding represents the identification of functional categories [23]. The selective coding round included the identification of any consistencies or inconsistencies (using the color grey) within the notations or difference in expertise (Expert/Novice), as shown in Table 1.

The five-point Likert-scale is used to enable calculation of averages used for the comparison of notations, and to create a standard quantification mechanism for the coders to use during the coding of the notations. If doing any quantitative analysis, the Likert-scale is the most accepted and used scale for this purpose [24].

V. RESULTS

In this section, the results from the data collection and analysis phase are shown and further discussed. The results include the differences in values, based on percentages or a five-point Likert-scale, when the coder is of a different expert level (Expert/Novice). Table 2 shows the average of all the analysed visual notations against the nine principles mentioned by Moody [6]. Further on in this section, the results of each PoN principle are discussed in detail.

- **Semiotic clarity**

The ideal notation does not have any Excess, Deficit, Redundant, or Overload in symbols. Therefore, any occurrence in this is seen as a negative (as shown in Table 2). The Beinformed and Visual Rules notation have excess, and/or redundant symbols. The researchers identified 18,75% of the BeInformed symbols as Excess and Redundant. The Visual Rules notation was identified with a 7,69% Excess in symbols.

- **Perceptual discriminability**

A visually strong notation which discriminates itself by the use of text, icons, and visual spacing, in order to stimulate faster identification of the different symbols. Therefore, a higher value is an indication that the notation has a high perceptual discriminability. The BeInformed notation with a 3,06 has the highest perceptual discriminability of the analyzed notations, compared to the DMN notation with a 1,50 (lowest).

- **Semantic transparency**

Semantic transparency covers if the visual appearance of the symbols suggests their meaning. A higher value in this principle is an indication that the notation seems to have semantic transparent symbols. The Berkeley Bridge notation has the highest semantic transparency with a 4,4. This seems the result of the low number of symbols, which is two. The BeInformed and Visual Rules notation seems to have the same result but by their high number of symbols, these notations have the lowest semantic transparency (BeInformed 2,88, and Visual Rules 2,71).

- **Complexity management**

The complexity management principle covers the ability to scale the notation. A higher value in this principle is an indication that the notation is useful on a larger scale by utilizing modularization and hierarchical structuring. The BeInformed notation has the highest value (4,17) in complexity management and seems better when dealing with larger scale projects. The TDM notation seems to be impacted by the low number of symbols in their notation to score the lowest (2,17) in complexity management.

- **Visual expressiveness**

The visual expressiveness principle covers the use of visual variables (colour, 3d symbols, and textual encoding). A higher value indicates that the notations are visually expressive. The TDM notation has a total score of 4,5 and thereby seems to be the highest scoring notation in visual expressiveness, compared to the Berkeley Bridge notation with a 1,5 (lowest).

- **Graphical Economy**

The graphical economy principle covers the number of symbols a human brain is able to discriminate between, this number is estimated to be limited to six. A value above six would be a negative impact on the graphical economy of the notation, which is the case for BeInformed (16), Visual Rules (13), and DMN (9).

- **Dual Coding**

The dual coding principle covers the complement of graphics with text, which is more effective than using each of them on their own. A higher value in this principle indicates that the notation uses dual coding as the most optimal notation. The Beinformed (4.25) and Visual Rules

 ISBN: 978-1-61208-690-3

(4.00) notations are the only notations, out of the analyzed five notations, where dual coding was identified.

TABLE 2. CODING RESULTS

		Beinformed	*Visual Rules*	*DMN*	*TDM*	*Berkeley Bridge*
Average Total		**2,87**	**2,97**	**2,38**	**2,89**	**2,53**
Cognitive Integration		2,88	3,83	3,00	2,67	1,92
Cognitive Fit		2,75	2,25	4,13	4,50	3,88
Dual Coding		4,25	4,00	N.A.	N.A.	N.A.
Graphical Economy		*16	*13	*9	4	2
Visual Expressiveness		3,13	4,00	2,25	4,50	1,50
Complexity Management		4,17	3,33	3,83	2,17	3,50
Semantic Transparency		2,88	2,71	2,92	3,56	4,40
Perceptual Discriminability		3,06	2,69	1,50	2,83	2,53
Semiotic Clarity	**Excess**	18%	7%	N.A.	N.A.	N.A.
	Redundancy	18%	N.A.	N.A.	N.A.	N.A.

- **Cognitive fit**

The cognitive fit principle covers the theory that different representations of information are suitable for different audiences. The Visual Rules notation scored the lowest with a 2,25, compared to that of TDM, which scored the highest with a 4,50.

- **Cognitive integration**

The cognitive integration principle covers the range of mechanisms available for dealing with multiple diagrams thereby, helping the reader assemble information from separate diagrams. A higher value indicates that the notation has the mechanisms available to help the reader assemble information when multiple diagrams are shown. The Visual Rules notation has the highest value (3,83) in cognitive integration, compared to the Berkeley Bridge notation (1,92) which does not have the mechanisms to support the reader when dealing with separate diagrams (lowest).

- **Difference Expert/Novice**

Taking into account that having experience in the use of a visual notation, in this case, a modelling language, influences the attitude towards several of the Moody principles. For example, a notation could be more complex for a novice but not for an expert. Therefore, a difference is made between the results of the expert researchers and novice researchers.

TABLE 3. RESULTS DIFFERENCE EXPERT/NOVICE

	Expertise:	**Average Total Expert/Novice**	**Average Total**
Beinformed	**Expert**	2,79	2,87
	Novice	2,95	
Visual Rules	**Expert**	2,99	2,97
	Novice	2,96	
DMN	**Expert**	2,25	2,38
	Novice	2,51	
TDM	**Expert**	2,92	2,89
	Novice	2,86	
Berkeley Bridge	**Expert**	2,39	2,53
	Novice	2,67	

VI. CONCLUSION, DISCUSSION AND FUTURE RESEARCH

In this paper, a study is conducted in which five DM visual notations, namely: Visual Rules, Berkeley Bridge, Decision Model and Notation, The Decision Model, and BeInformed, were evaluated using the PoN framework [1]. From our analysis, Visual Rules scores best according to the average total of all PoN framework principles. From a theoretical perspective, our study and its results give meaning to the operationalization of the PoN framework. Furthermore, it will enable further exploration of the application of the PoN principles, as well as other DM visual notations not included in this study. Moody [1, p.772] describes the theoretical interactions between the described principles. Our results show that these interactions are, to a large extend, verified. From a practical perspective, the results presented in this paper contribute towards a better awareness for taking into account validated visual notation principles and guidelines. Our results could be utilized by organizations to either evaluate for themselves which visual notation is most adequate or to utilize a visual notation based on our results.

This study has multiple limitations. The first limitation concerns the research team that carried out the evaluation of the visual notations using the PoN framework. This study included evaluations of four researchers, two novice level researchers and two expert level researchers on the DM topic. Therefore, one could argue that the results and conclusions are potentially biased by a low amount of data points for the evaluation of the visual notations included. However, most studies conducted with a focus on evaluating one or multiple visual notations are often centered on the evaluation of the visual notation using one or two researchers. Future research should focus on evaluating visual notations utilizing larger sample sizes that will add to the generalizability of the results and conclusions about the evaluated visual notations. The second limitation concerns the method and framework utilized to evaluate the visual notations, the PoN framework and its operationalization by creating and utilizing a template with the goal to structure data collection and analysis. Utilizing the PoN framework is an explicit choice, however, limits the results because the PoN framework represents a specific lens. Future research could, therefore, focus on applying other frameworks and

 ISBN: 978-1-61208-690-3

theories that focus on uncovering and describing essential notational principles, e.g., Guidelines of Modeling (GoM) [11]. Furthermore, the operationalization of the PoN framework described in [6] is left open for interpretation and perception of the researchers applying it, a good example is the lack of weighting of the nine PoN principles. Therefore, our template is another limitation. This phenomenon becomes clear in the work of [5], which shows that the operationalization of the PoN framework by different research teams often do not always seem to take into account all principles described. To our knowledge, our operationalization included, one-on-one, all principles described in the work of [5]. Future research, however, should focus on how these principles are best measured in practice, i.e., whether Likert scales or other less quantitative measurements are adequate or not. The last limitation concerns the visual notations selected. Although we choose two well-known visual notations, as well as three visual notations applied in the DM practice a lot, the selection of visual notations could coincidentally have resulted in a bias and affect the generalizability of our results. We argue that this risk is more or less mitigated as most studies conducted that utilize the PoN framework focus on only one visual notation, see also [5], while this study reports upon the evaluation of five visual notations. Future research could also focus on evaluating additional DM visual notations.

REFERENCES

[1] M. W. Blenko, M. C. Mankins, and P. Rogers, "The Decision-Driven Organization," *Harv. Bus. Rev.*, no. june, p. 10, 2010.

[2] Object Management Group, "Decision Model and Notation," 2016.

[3] Object Management Group, "ArchiMate® 3.0 Specification," 2016.

[4] B. Von Halle and L. Goldberg, *The Decision Model: A Business Logic Framework Linking Business and Technology*. CRC Press, 2009.

[5] D. Van Der Linden and I. Hadar, "A Systematic Literature Review of Applications of the Physics of Notation," *IEEE Trans. Softw. Eng.*, pp. 1–1, 2018.

[6] D. L. Moody, "The 'Physics' of Notations: Towards a Scientific Basis for Constructing Visual Notations in Software Engineering," *IEEE Trans. Softw. Eng.*, vol. 35, no. 5, pp. 756–778, 2009.

[7] H. a. Reijers and J. Mendling, "A Study Into the Factors That Influence the Understandability of Business Process Models," *IEEE Trans. Syst. Man. Cybern.*, vol. 41, no. 3, pp. 449–462, 2011.

[8] D. L. Moody, P. Heymans, and R. Matulevičius, "Visual syntax does matter: Improving the cognitive effectiveness of the i* visual notation," *Requir. Eng.*, vol. 15, no. 2, pp. 141–175, 2010.

[9] A. Blackwell and T. Green, *Notational systems–the cognitive dimensions of notations framework*, HCI Models. Morgan Kaufmann, 2003.

[10] M. Rosemann, P. Green, and M. Indulska, "A reference methodology for conducting ontological analyses," in *Proceedings of the International Conference on Conceptual Modeling*, 2004, pp. 110–121.

[11] R. Schuette and T. Rotthowe, "The guidelines of modeling–an approach to enhance the quality in information models," in *Proceedings of the Conceptual Modeling–ER'98*, 1998, pp. 240–254.

[12] F. Saleh and M. El-Attar, "A scientific evaluation of the misuse case diagrams visual syntax," *Inf. Softw. Technol.*, vol. 66, pp. 73–96, 2015.

[13] N. Genon, P. Heymans, and D. Amyot, "Analysing the Cognitive Effectiveness of the BPMN 2.0 Visual Notation," 2011.

[14] D. Moody and J. van Hillegersberg, *Evaluating the visual syntax of UML: an analysis of the cognitive effectiveness of the UML family of diagrams*, LNCS vol. Springer, 2009.

[15] Gartner, "Taking the Mystery Out of Business Rule Representation," 2013. [Online]. Available: https://www.gartner.com/doc/2371915/taking-mystery-business-rule-representation. [Accessed: 01-Feb-2019].

[16] R. D. Galliers and F. F. Land, "Choosing appropriate information systems research methodologies," *Commun. ACM*, vol. 30, no. 11, pp. 901–902, 1987.

[17] A. Strauss and J. Corbin, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 3rd ed., vol. 3. Thousand Oaks, CA: SAGE Publications Ltd., 2015.

[18] BeInformed, "Be Informed Business Process Platform," 2017. [Online]. Available: https://www.beinformed.com/. [Accessed: 02-Jan-2019].

[19] Berkeley Bridge, "Berkeley Bridge Platform," 2018. [Online]. Available: https://www.berkeleybridge.com. [Accessed: 15-May-2018].

[20] Bosch, "Business Rules Management using Visual Rules," 2018. [Online]. Available: https://www.bosch-si.com/bpm-and-brm/visual-rules/business-rules-management.html. [Accessed: 02-Jan-2019].

[21] H. E. Tinsley and D. J. Weiss, "Interrater Reliability and Agreement," in *Handbook of Applied Multivariate Statistics and Mathematical Modeling*, San Diego, CA: Academic Press, 2000, pp. 95–124.

[22] H. T. Reis and C. M. Judd, Eds., *Handbook of Research Methods in Social and Personality Psychology*, 2nd ed. Cambridge: Cambridge University Press, 2014.

[23] A. Böhm, B. Glaser, and A. Strauss, "Theoretical Coding: Text Analysis in Grounded Theory," *A Companion to Qual. Res.*, pp. 270–275, 2004.

[24] I. E. Allen and C. A. Seaman, "Likert scales and data analyses," *Qual. Prog.*, vol. 40, pp. 64–65, 2007.

 ISBN: 978-1-61208-690-3

Comparison of Machine Learning Algorithms on Smartphone Energy Consumption Modeling Issue Based on Real User Context Data

Weizhu Qian
CIAD, Univ. Bourgogne Franche-Comt, UTBM, F-90010
Belfort, France
Email: weizhu.qian@utbm.fr

Franck Gechter
CIAD, Univ. Bourgogne Franche-Comt, UTBM, F-90010
Belfort, France
Email: franck.gechter@utbm.fr

Fabrice Lauri
CIAD, Univ. Bourgogne Franche-Comt, UTBM, F-90010
Belfort, France
Email: fabrice.lauri@utbm.fr

***Abstract*—Nowadays, billions of smartphones are used worldwide. The energy consumption is a critical issue when using such devices. In this context, smartphone power modeling is a mandatory step to better understand energy drain. On that way, the most widespread methods are based on specific hardware/software level analysis. As opposed to these classical approaches, we propose, in this paper, an alternative method aimed at constructing smartphones power models based on user context data provided by Device Analyzer, an Android application developed by the University of Cambridge. From a very large-scale smartphone usage data, we extract the energy-related events. Then, the energy-related context is formulated as the input features of the energy models. So as to predict the energy consumption of a smartphone, we compare four different machine learning models: a Linear Regression model, an AdaBoosted Decision Trees model, a Gradient Boosted Regression Tree model and a Random Forest model. The proposed energy models are then validated on a real user dataset in different usage scenarios.**

***Index Terms*—smartphone data; user behavior; energy modeling; regression model.**

I. Introduction

Since the embedded system related technology is progressing rapidly in recent years, mainly thanks to the support of numbers of applications and advanced hardware, smartphone devices have now become platforms integrated with various functions. Obviously, users not only can use their handsets to make phone calls and send SMS as before, but also can browse webs, play video games, listen to music and take photos, etc. Meanwhile, these various functions are tied to a huge growth in energy costs. Despite the progresses made in battery technology with an increase of energy storage, the smartphone usage is still limited by the energy drain of the components and of software needs. Hence, profiling, modeling and characterizing energy drain in the smartphone have become an important research topic in recent years.

In literature, many research work studied smartphone energy consumption focusing on either hardware or operating systems points of view [1]–[4]. Some other work studied how applications are consuming energy [5] or how the network environment changes affect smartphone energy drain [6]. Although this works proposed various accurate smartphone energy models, it should be aware of that the discharging process of a smartphone is a complex process, i.e., due to their unique hardware properties and usage context, different smartphone devices may have different discharging characteristics even when they are of the same type. Thus, one other reasonable alternative for modeling smartphone energy drain is to develop a personalized energy model for one specific smartphone device/user couple.

The purpose of our study is to find the relationship between user behavior and the energy drain in smartphone devices. This method can be a universal way to build energy models for smartphones because it does not need to measure the hardware components energy consuming properties or develop a specialized kernel tools as in precedent research [1], [3]. Our approach treats the energy consumption as the consequence of all the users' operation on their smartphone devices. Based on that, we develop a methodology to model the smartphone energy drain based on user context data. First, the energy-related events are chosen and extracted from the raw usage data. Then according to the selected events, the input features for energy models are generated. Afterwards, a series of machine learning models, the Linear Regression model, the AdaBoosted Decision Tree model, the Gradient Boosted Regression Tree model and the Random Forest model, are tested and compared in different usage scenarios.

The paper is organized as follows. Section II discusses some earlier related research work in this field and introduces the dataset and machine learning algorithms adopted. Section III presents the machine learning pipeline of our energy modeling process. Section IV describes how the features are generated and selected from the raw collected user data. Then, in Section V, the experiments for different scenarios and their associated results are presented. Finally, the conclusions and perspectives are given in Section VI.

II. Related Work and background

A. Power modeling issue

Smartphone power modeling is to use some energy-related factors, e.g., hardware properties (battery voltage, battery current, CPU utilization, etc.), system calls, user context(screen-

 ISBN: 978-1-61208-690-3

on time, phone-call time, WiFi status, etc.), as input variables of models to predict smartphone energy consumption [7]. A general smartphone energy model can be described as follows, where E is the energy consumption of the handset:

$$E = P_{base} \cdot D + \sum_{i} \beta_i \cdot x_i \cdot d_i \quad (1)$$

The first term of the sum represents the fixed amount of energy consumption for the total experiment time D, and P_{base} is the basic energy consumption per time unit. The second term of the sum represents the variable amount of energy consumption. x_i is a value which represents the property of energy-related factors, for instance, for hardware elements it can be the CPU utilization and for user context it can be the WiFi status. β_i is a scalar indicates the energy-drain effect of x_i and d_i is influencing duration of x_i.

According to (1), from the point of view of hardware, smartphones devices are made of a series of sub-hardware components, CPU, display, cell module, WiFi module, audio, camera, etc. If the energy consumption of each component can be known, then naturally if we adopt a reductionist point of view, the total energy consumption of whole smartphone handset can be regarded as the sum of energy drain of all hardware components. According to the results in [1], this method can estimate smartphone energy drain accurately, but it needs external equipment to measure the power of each hardware, which is not always practical to implement. Besides, this method did not consider the interaction between the components which can represents 10% of the total energy drain [1]. Another frequently used method of modeling smartphone energy consumption is to trace the system calls of the smartphone [3]. As the different system calls are related to different power states in the smartphone, once the benchmarks of different power states are determined, the relationships between power states and system calls can be represented by a Finite State Machine (FSM) energy model.

From the aspect of smartphone usage context, a several research works analyzed how the user behavior and the network environment affect the handset energy drain [4] [6] [8]. Moreover, with the fast-growing crowdsourcing technology, researchers can have the access to a very large amount of user context data from a large number of users, which provides the researchers a new approach to study the interaction between users and smartphone [9] [10]. For instance, [11] [12] presented user-data based methods to predict the smartphone energy consumption and [13] proposed a big data combined with machine learning method to predict handset's battery life. In a similar way, we planned to exploit user context data to develop the smartphone energy model and to profile how the energy consumption cause by user behavior in our research work.

B. Device Analyzer

So as to study how user behavior affects energy drain in the handset, naturally, the first step of the process is to collect user context data in a large-scale. The dataset presented in this paper is the Device Analyzer dataset [14], which contains comprehensive and detailed information about the Android smartphone usage, e.g., events types, event values and relevant time stamps. Considering the trade-off between overhead and performance, the Device Analyzer application records the sampled events, such as battery levels, network traffic and screen brightness levels, at a time interval of 5 minutes, while records the immediate events, such as phone-calls, screen-on/off and WiFi on/off when they occur. Table I presents a list of the collected events. The exhaustive usage information are selected and formulated as the input features of our proposed energy models and the change of battery level of smartphones over 5 minutes of sampling interval is treated as the modeling targets. By using the context data as inputs and energy consumption as outputs, our task is to find proper energy models to imply how user behavior leads to energy drain in smartphone devices.

TABLE I PART OF COLLECTED EVENTS

Event category	Event name
Setting	Airplane mode Ring mode Audio volume, etc.
CPU	Number of cores Maximum frequency Minimum frequency, etc.
Battery	Battery level Battery voltage Charging status Battery temperature, etc.
Screen	Brightness level Screen state Screen size
Application	APP installed APP updated APP service APP foreground APP background, etc.
Phone call	Roaming state Network location Signal type Calling state, etc.
SMS	SMS received SMS sent, etc
Data transfer	Received bytes Received packets Send bytes Send packets
WiFi	WiFi scan WiFi connected WiFi state, etc.
Sensor	Type Max range Values Delay, etc.

C. Machine Learning Algorithms

As explained before, our research purpose is to predict energy consumption based on user behavior. Mathematically, it can be regarded as a regression problem, which is to find a function that matches real outputs from inputs accurately. Due to the powerful modeling ability of machine learning techniques, we resort to different advanced machine learning

regression models. The basics of the models is described as below.

1) Linear Regression: Though the Linear Regression (LR) model is considered as the simplest regression model, it can obtain satisfying results in many research cases [13]. For this reason, the LR model is used in our research as the benchmark models to make comparisons to other tree-based models.

2) Decision Trees: The smartphone energy consumption is a continuous process, but we also assume that it consists of a series of power state change caused by different usage context. In this case, the regression decision tree models can be the appropriate models. Classification And Regression Decision Trees (CART) models [15] are tree-shaped models, it can be applied to either classification tasks or regression tasks. In our regression tasks, to approximate continuous target values, the CART algorithm uses the Least Square Difference (LSD) or the Least Absolute Difference (LAD) as criterion to split nodes. To improve the performance of tree models, the CART models are used as the basic estimators, combined with two different classes of ensemble methods: the boosting method and the bagging method.

- AdaBoosted Decision Tree (ADT): The boosting ensemble method aims to reduce the bias of each tree. AdaBoosted Decision Tree model [16] combines a series of weak learners with different weights which are calculated by their predicting accuracies to build a stronger learner.
- Gradient Boosting Decision Tree: The other Boosted Decision Tree model we utilize is Gradient Boosting Decision Tree (GBDT) [17]. GBDT uses the residuals between predictions and targets of former decision trees as new training targets to adapt its parameters.
- Random Forest (RF): The other class of ensemble method is the bagging method and a frequently-used model of bagging method is Random Forests. To overcome the over-fitting of each tree, Random Forests utilize the bootstrap aggregating technique [18] to decrease the variance of each tree.

In Section V, the performance of the four different models will be compared on the smartphone energy modeling issue.

III. Methodology

In order to construct the smartphone battery level predicting model from raw usage data, we propose a machine learning based modeling methodology. Though, the original data from the Device Analyzer dataset records extremely detailed information about users' usage, most of the information is not useful for our research purpose. Thus, the first step we take in the modeling process is to preprocess the raw data to extract the energy-related events. After the energy-related usage events are selected, according to their properties, the chosen contextual data are formulated as multidimensional matrix space consisting of a series of vectors. Each vector represents the occurrences and values of each energy-related events over a certain time range. We set the time range as 5 minutes because the battery level value is sampled every 5 minutes by the Device Analyzer application. Afterwards, the generated vectors are feed to the machine learning models as input variables. Finally, after the training process, the machine learning energy models are examined on the testing data and their results are analyzed. The pipeline of our methodology is shown in Figure 1. Four types of machine learning model are adopted in our research: Linear Regression, AdaBoosted Decision Trees, Gradient Boosted Regression Trees and Random Forests.

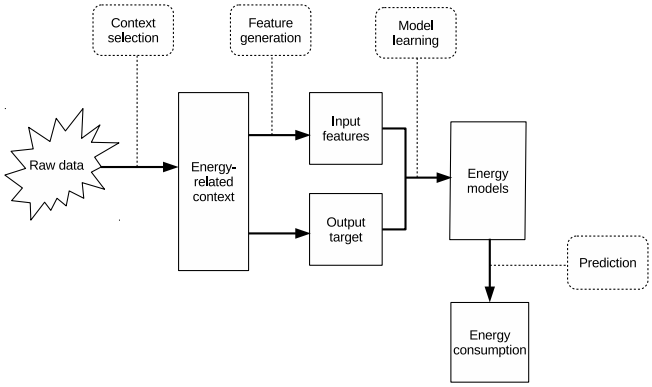


Figure. 1 Pipeline of smartphone energy prediction

IV. User Context

A. Event Selection

The dataset for the experiments is collected from one Android smartphone user. It records the user's information for 80 days, and contains 13716 instances. Although the dataset that we use is from one individual user, our methodology is not limited to one specific user or one specific smartphone device. Hence, it can be a re-usable paradigm for smartphone energy consumption modeling. To exploit the useful context for the energy models, it needs to extract energy-related events from the raw dataset at the first step. Although the raw data collected corresponds to very detailed usage information, not all the recorded events directly make contribution to energy consumption of the handset, for instance, the location information and mobility sensors information are not necessary useful at first steps. Meanwhile, it should be reminded that our energy modeling approach is based on user behavior data. Hence, some hardware information will not be used neither, even they can reflect the energy drain condition, e.g., battery voltage and CPU utilization.

The principle of selecting the features is that the chosen features not only imply the significant information of user context but also are essential to profile energy consumption in smartphone handsets. For instance, when the screen is turned on, and several minutes later, it is turned off. During this time period, the user might watch a video using the smartphone and as a result, the battery level decreased. Apparently, this screen-on event caused significant energy consumption, meanwhile, it also represents essential information of the user-smartphone interaction. Similarly, according to previous works on investigating the energy-consuming effects caused by energy-

 ISBN: 978-1-61208-690-3

TABLE II SELECTED ENERGY-RELATED EVENTS

Category	Event Type
Battery	Battery Level
Screen	Screen On/Off Brightness
User Context	APP Foreground APP Background Phone Call Plugged In/Out
Data Transfer	Received Bytes Received Packets Send Bytes Send Packets
WIFI	WIFI On/Off

TABLE III VARIABLES DESCRIPTION

Event	Value Type	Unit
Battery Level	Variation	Percentage
Screen-on Time	Duration	Second
Brightness Level	Average	Scale(0˜255)
APP Foreground	Aggregation	-
APP Background	Aggregation	-
Phone-call Time	Duration	Second
Plugged State	State	1/0
Received Bytes	Aggregation	Byte
Received Packets	Aggregation	-
Send Bytes	Aggregation	Byte
Send Packets	Aggregation	-
WiFi-on Time	Duration	Second

related events [4] [6] [13], other energy-related events, such as network data traffic, phone-call, application usage and WiFi status are also important to describe smartphone usage and reflect the energy drain rate change. And they function in the similar way as the screen-on events in our proposed energy models. Moreover, the charging state of the smartphone battery apparently affects the battery level, therefore it is utilized in our models as well. Eventually, the events that we choose to build our energy model are screen-on time, the level of screen brightness, received data bytes, received data packets, sent data bytes, sent data packets, phone-call time, application foreground, application background, WiFi state and charging state.

In our approach, the energy consumption is measured by the smartphone battery level change. As mentioned in Section II-B, the battery level is sampled every 5 minutes by Device Analyzer, thus our task is to predict the energy consumption over time period of 5 minutes. Since the battery level is obtained by the smartphone battery indicator, the unit of daily energy drain is percentage of whole battery capacity. For example, if the value is 7, it means that the energy consumed by the smartphone device over this time period is 7% of the battery capacity. All the collected events utilized in our approach are listed in Table II.

B. Feature Generation

The measurement of energy drain is the level of the smartphone battery which is sampled every 5 minutes by Device Analyzer. Therefore, the values of each event is quantized within the time window between two battery level sampled time points. Furthermore, to make the events' values feed the model appropriately, i.e., according to their usage properties, we classify the possible values of the energy-related events into different categories: the variation, the duration, the average, the aggregation and the state as it is demonstrated in Section III. The 5 categories are described as follow:

- **Variation**: It should be noticed that all the events' contribution to the smartphone energy consumption over a time period is in fact the variation of the battery capacity over this time period. Thus, instead of using the battery percentage at the sampling time point as modeling target directly, it is more reasonable to use the variation of battery level during the sampling interval.
- **Duration**: If an event is the human-smartphone interaction that lasts for a certain time period, then we use the range of this time period to measure its influence on smartphone energy drain. For instance, the screen of a smartphone at time point T_1 is turned on and the screen is turned off at time point T_2, naturally, the total screen-on time will be T_2-T_1. This category includes screen-on time, phone-call time and WiFi-on time.
- **Average**: For the brightness level of screen, we use the average value over the sample interval.
- **Aggregation**: Some of the event may occur more than once, such as application foreground and application background, so we use sum of the occurrences to represent its energy contribution. As for the network transfers, received Bytes, received packets, send Bytes and send packets are aggregated values as well.
- **State**: The charging state of the battery can be whether plugged-in or plugged-out, so it is a boolean value.

Based on the categories we introduced, the energy-related events is quantized to a series of input vectors with 11 items at different sampling time points. Each item of a vector represents its related event's quantity. The target value is the energy consumption at the same sampling time point. Moreover, due to the malfunction of the data-collecting process, some values are missing. In this case, the average values are used to fill in the blanks. Before the input vectors are fed to the model, they are normalized to eliminate the effect of the different events' scales.

Once the input variables and the output target are well defined, the following step is to find an appropriate model to describe the mathematical relationship between the inputs and the outputs.

V. Experiments and Results

A. Experiments

To investigate the smartphone energy consumption under different usage scenarios, we select different variables combinations as inputs of the energy models. The 6 different experimental scenarios is demonstrated in Table IV. In the first scenario, we put all the energy-related events together. Then, in scenario 2, scenario 3, scenario 4, scenario 5 and scenario 6, we emphasize each energy-related events individually to study

 ISBN: 978-1-61208-690-3

TABLE IV EXPERIMENTAL SCENARIOS

Scenario	Context
Scenario 1	All contexts
Scenario 2	Screen-on
Scenario 3	Phone-call
Scenario 4	WiFi-on
Scenario 5	Plugged-in
Scenario 6	Plugged-out

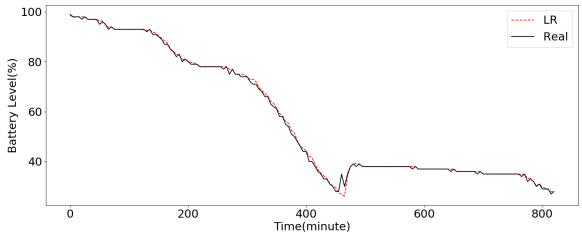
each event's contribution to energy consumption. It means that the data collected only in the scenario where screen-on events, phone-call events, WiFi-on events, plugged-in events or plugged-out events occur is utilized, respectively.

In order to evaluate the energy models, we resort to cross-validation method, the total dataset is split into two sub datasets, the training dataset and the testing data, with the Bootstrap method. The training subset data is to generate the energy models and the testing subset data is to examine generalization of the model to measure over-fitting. The criterion of a good energy mode is that it should not only fit the training dataset well but also have good performance on testing data. 70% of the whole dataset is used to train the energy models and the 30% rest is used to test the energy models. The proposed energy models are trained and tested on the dataset for 20 times for each scenario. The metrics to evaluate the energy models is the Root Mean Square Error (RMSE), and the mean value and Standard Deviation (STD) of RMSE are calculated.

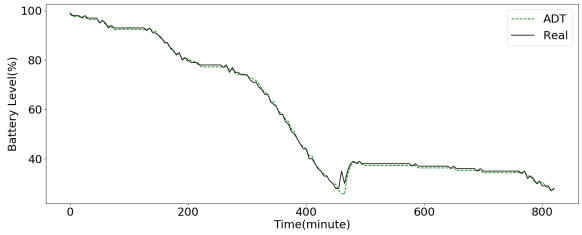
B. Results

In Figure 2, the battery level curves are reconstructed by predicted results of our devised models, and it shows that each energy model is capable of predicting battery level in the smartphone devices, either when the battery is discharging or charging. From the more detailed results shown in Table V and Table VI, in Scenario 1, when putting all contextual information together, the Random Forest regression model has the smallest training error and testing error. In Scenario 2, where the display is turned on, once again the RF energy model has the best performance both on training dataset and testing dataset among all the models. In Scenario 3, where calling events occur, the ADT model has the least training error and surprisingly, the Linear Regression model has the most accurate results on testing data. And in Scenario 4, where the WiFi module is turned on, the ADT model perform best on training dataset and the Random Forest model performs best on testing data. In Scenario 5, where the battery is plugged in, the GBDT model and the RF model has the smallest training error and has the smallest testing error, respectively. In Scenario 6, where the smartphone is not charging, the ADT model outperforms other models on training dataset and the RF model outperforms other models on testing dataset.

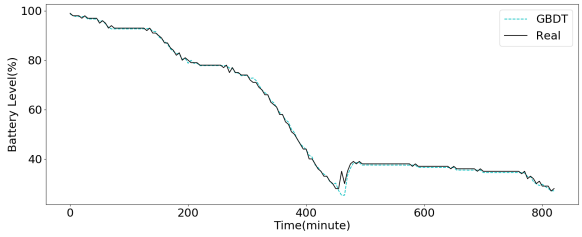
It also can be seen that, among all the scenarios, all the models' accuracy decreases in Scenario 5. As opposed to this, in Scenario 6, all the models have their own best performance on testing dataset. The reason could be that the



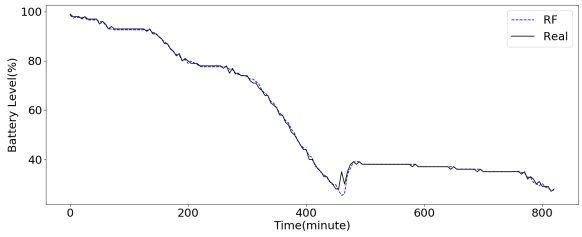
(a) LR Predicted Results



(b) ADT Predicted Results



(c) GBDT Predicted Results



(d) RF Predicted Results

Figure. 2 Predicting smartphone battery level over a time period

smartphone energy consumption is used to measured change of the battery level. However, it will not be accurate when the smartphone reaches its maximum energy volume (the battery level is 100%) and remains at the same battery level even energy-consuming events occur. Overall, in all scenario, except Scenario 3, all the tree-based models perform better than the Linear Regression model on the training dataset and testing dataset and the Random Forest energy model has the best performance on testing sub dataset. Thus, we can say that the RF is the most appropriate energy model in our research.

VI. CONCLUSION

In this paper, we presented a novel smartphone energy consumption modeling methodology based on user contextual data and machine learning algorithms. First, we extract the 11

TABLE V RMSE OF TRAINING RESULTS

Scenario	LR		ADT		GBDT		RF	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Scenario 1	0.8550	0.0063	0.4703	0.0107	0.5250	0.0072	**0.4181**	0.0058
Scenario 2	0.9251	0.0101	0.4197	0.0526	0.5115	0.0112	**0.3966**	0.0096
Scenario 3	0.6457	0.0372	**0.2227**	0.0248	0.3517	0.0450	0.3068	0.0142
Scenario 4	0.6999	0.0688	**0.2153**	0.0204	0.3453	0.0381	0.2963	0.0272
Scenario 5	1.5406	0.0110	0.8763	0.0242	**0.6525**	0.0173	0.8630	0.0181
Scenario 6	0.6179	0.0053	**0.3147**	0.0082	0.3605	0.0074	0.4600	0.0052

TABLE VI RMSE OF TESTING RESULTS

Scenario	LR		ADT		GBDT		RF	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Scenario 1	0.8612	0.0147	0.7230	0.0131	0.7052	0.0126	**0.6975**	0.0145
Scenario 2	0.9327	0.0235	0.8628	0.0326	0.8538	0.0306	**0.8411**	0.0325
Scenario 3	**0.8206**	0.0743	0.8554	0.1132	0.8944	0.1409	0.8421	0.1182
Scenario 4	0.9213	0.1772	0.8217	0.1732	0.8375	0.1205	**0.7814**	0.1430
Scenario 5	1.5584	0.0269	1.2881	0.0276	1.2352	0.0346	**1.2327**	0.0326
Scenario 6	0.6170	0.0123	0.5825	0.0154	0.5638	0.0146	**0.5624**	0.0137

most important energy-related contexts from the comprehensive and detailed raw user data. Then, so as to feed the machine learning model properly, a series of features are generated from the properties of the extracted events. Afterwards, the energy consumption is used as the output target and the generated features are used as input variable for the energy model. Four different machine learning regression models are trained, tested and compared on the dataset. The final results indicate the feasibility of our proposed methodology.

For future work, we plan to take into account more energy-related events, for instance, the GSM signal strength and its types. Besides, we also plan to improve the accuracy of the existing models and to make a comparison with an agent based method developed in our previous work [7].

ACKNOWLEDGEMENT

Authors would like to thank Alastair Beresford and Andrew Rice from the University of Cambridge for providing the Device Analyzer dataset.

REFERENCES

[1] L. Zhang *et al.*, "Accurate online power estimation and automatic battery behavior based power model generation for smartphones," in *Proceedings of the eighth IEEE/ACM/IFIP international conference on Hardware/software codesign and system synthesis*. ACM, 2010, pp. 105–114.

[2] A. Pathania, A. E. Irimiea, A. Prakash, and T. Mitra, "Power-performance modelling of mobile gaming workloads on heterogeneous mpsocs," in *Proceedings of the 52nd Annual Design Automation Conference*. ACM, 2015, p. 201.

[3] A. Pathak, Y. C. Hu, M. Zhang, P. Bahl, and Y.-M. Wang, "Fine-grained power modeling for smartphones using system call tracing," in *Proceedings of the sixth conference on Computer systems*. ACM, 2011, pp. 153–168.

[4] W. Jung, K. Kim, and H. Cha, "Userscope: A fine-grained framework for collecting energy-related smartphone user contexts," in *Parallel and Distributed Systems (ICPADS), 2013 International Conference on*. IEEE, 2013, pp. 158–165.

[5] A. Banerjee, L. K. Chong, S. Chattopadhyay, and A. Roychoudhury, "Detecting energy bugs and hotspots in mobile apps," in *Proceedings of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering*. ACM, 2014, pp. 588–598.

[6] L. Sun, H. Deng, R. K. Sheshadri, W. Zheng, and D. Koutsonikolas, "Experimental evaluation of wifi active power/energy consumption models for smartphones," *IEEE Transactions on Mobile Computing*, vol. 16, no. 1, pp. 115–129, 2017.

[7] F. Gechter, A. R. Beresford, and A. Rice, "Reconstruction of battery level curves based on user data collected from a smartphone," in *International Conference on Artificial Intelligence: Methodology, Systems, and Applications*. Springer, 2016, pp. 289–298.

[8] N. Vallina-Rodriguez, P. Hui, J. Crowcroft, and A. Rice, "Exhausting battery statistics: understanding the energy demands on mobile handsets," in *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds*. ACM, 2010, pp. 9–14.

[9] D. Ferreira, A. K. Dey, and V. Kostakos, "Understanding human-smartphone concerns: a study of battery life," in *International Conference on Pervasive Computing*. Springer, 2011, pp. 19–33.

[10] A. Rahmati, A. Qian, and L. Zhong, "Understanding human-battery interaction on mobile phones," in *Proceedings of the 9th international conference on Human computer interaction with mobile devices and services*. ACM, 2007, pp. 265–272.

[11] E. A. Oliver and S. Keshav, "An empirical approach to smartphone energy level prediction," in *Proceedings of the 13th international conference on Ubiquitous computing*. ACM, 2011, pp. 345–354.

[12] E. Oliver and S. Keshav, "Data driven smartphone energy level prediction," *University of Waterloo Technical Report*, 2010.

[13] H. Li, X. Liu, and Q. Mei, "Predicting smartphone battery life based on comprehensive and real-time usage data," *arXiv preprint arXiv:1801.04069*, 2018.

[14] D. T. Wagner, A. Rice, and A. R. Beresford, "Device analyzer: Understanding smartphone usage," in *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*. Springer, 2013, pp. 195–208.

[15] L. Breiman, "Bagging predictors," *Machine learning*, vol. 24, no. 2, pp. 123–140, 1996.

[16] Y. Freund *et al.*, "Experiments with a new boosting algorithm," in *Icml*, vol. 96. Citeseer, 1996, pp. 148–156.

[17] J. H. Friedman, "Stochastic gradient boosting," *Computational Statistics & Data Analysis*, vol. 38, no. 4, pp. 367–378, 2002.

[18] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

 ISBN: 978-1-61208-690-3

Adaptive Method for Trends in Ranking of Tourist Spots

Yusaku Takano
Graduate School of System Design
Tokyo Metropolitan University
Hino, Tokyo, Japan
Email: takano-yuusaku@ed.tmu.ac.jp

Masaharu Hirota
Department of Information Science
Okayama University of Science
Okayama, Japan
Email: hirota@mis.ous.ac.jp

Daiju Kato
WingArk First Inc.
Shibuya, Tokyo, Japan
Email: kato.d@wingarc.com

Tetsuya Araki
Graduate School of System Design
Tokyo Metropolitan University
Hino, Tokyo, Japan
Email: araki@tmu.ac.jp

Masaki Endo
Polytechnic University
Kodaira, Tokyo, Japan
Email: endou@uitec.ac.jp

Hiroshi Ishikawa
Graduate School of System Design
Tokyo Metropolitan University
Hino, Tokyo, Japan
Email: ishikawa-hiroshi@tmu.ac.jp

***Abstract*—In recent years, for tourists deciding on a destination for a trip, demand for websites, such as TripAdvisor, for obtaining tourist information is increasing. These websites usually rank tourist spots with user reviews. Regarding tourist spots, trends appear in their popularity. When ranking tourist spots solely by reviews posted on these websites for obtaining tourist information, the possibility exists that new tourist spots and tourist spots with actively changing popularity are not ranked higher in rankings. However, a user needs to find the most enjoyable tourist spots at the time of the visit in the ranking of tourist spots. Therefore, we use tweets in this study as comments for tourist spots with high recency to generate a ranking of tourist spots considering popularity variation. Then, we model the ranking of tourist spots on TripAdvisor using tweets. After analyzing the contribution rates of classification for linguistic features and statistical features on tweets, we performed ranking of learning with effective features and generated a ranking of tourist spots considering popularity variation. Finally, our experimental results showed that tourist spots holding exhibitions or other activities frequently are ranked higher.**

Keywords–tourism; ranking learning; social recommendation.

I. INTRODUCTION

In recent years, with the increasingly widespread use of smartphones and tablets, people have become able to acquire useful information easily from the web. Along with this progress, when tourists decide on a trip destination, demand for websites, such as TripAdvisor [1], for obtaining tourist information is increasing. As the importance of the internet for travel increases, the way of using the internet during trips has changed [2]. In addition, particularly addressing user reviews, the importance of reviews on websites for obtaining trip-related tourist information has increased [3]. These websites usually present tourist spots in a ranking format and rank tourist spots with user reviews. Roughly speaking, the ranking of tourist spots is a comprehensive evaluation of past impressions. Additionally, these websites fundamentally keep the ranking algorithm secret to avoid deliberate manipulation of rankings and to prevent their use by other web services.

Regarding tourist spots, trends exist in the popularity of certain tourist spots. In other words, the popularity of tourist spots varies for different reasons, such as events or seasons. When ranking tourist spots using only reviews posted in the past, the possibility exists that new tourist spots and tourist spots with actively changing popularity are not ranked higher in among all rankings. However, most people might browse only higher ranked tourist spots. Users need to find the most enjoyable tourist spot at the time of their visit. In short, the ranking of tourist spots using only reviews posted in the past might not satisfy a user's need. For example, as an accurate ranking for a user's need, famous tourist spots for the viewing of cherry blossoms, such as Shinjuku Gyoen, can be expected to be ranked highly in the spring. For this study, we rank tourist spots in consideration of their current popularity variation in addition to the comprehensive evaluation of them in the past.

Many comments related to tourist spots are also posted to Twitter [4] and Facebook [5] which is a typical Social Network Service (SNS). These posts are comments about tourist spots with higher recency than reviews on websites for obtaining tourist information. In addition, these posts include user impressions about tourist spots. In this study, we propose an adaptive method for popularity variation of tourist spots using tweet messages from Twitter. Although we propose a method to make a ranking considering the latest popularity of tourist spots, the reliability of information related to user posts is low because many posts are made by heterogeneous users and robots on Twitter. Castillo et al. [6] pointed out a difficulty for the reliability of information related to posts. They proposed a method to extract only trusted information from Twitter. As described above, the information of tweets is unreliable. It is difficult to rank tourist spots solely by impressions of tourist spots given in the contents of tweets.

Therefore, in this study, our main idea of ranking tourist spots is modeling the ranking on TripAdvisor using tweets. Reviews by users in each post on websites for obtaining tourist information are more informative than Twitter and are reflected in the details for tourist spots. Although many websites exist for obtaining tourist information, TripAdvisor is a popular website for obtaining tourist information. Filieri et al. [7] examined the reliability and its influence by its contents and revealed that many tourists are using TripAdvisor. We will generate a ranking of tourist spots incorporating changes in popularity by application of recent tweets to the model. Consequently, tourist spots related to events and seasons will be ranked at the top of the ranking, which makes it possible

 ISBN: 978-1-61208-690-3

to propose a ranking that satisfies the most enjoyable tourist spots at the time of the visit.

The remainder of the paper is organized as follows. Section II presents works related to feature selection for posts by users on the internet and tourism recommendations using posts on SNS. Section III presents a discussion of the results to analyze the contribution of classification with linguistic features and statistical features on tweets at tourist spots. Section IV presents results obtained using logistic regression and ranking learning with effective features for classification. Section V explains a discussion of the obtained results. Section VI concludes the paper with a discussion of results and avenues for future work.

II. Related Work

In this section, we present related works about analysis of user-posted reviews and recommendation of tourism spots.

A. Analysis of user-posted reviews

Currently, reviews on the internet are increasing. Many studies use these reviews. Mukherjee et al. [8] examined the spam filter system in Amazon Mechanical Turk with linguistic features and statistical features for reproducing the spam filter system in Yelp. In addition, Guy et al. [9] proposed a method of automatically collecting useful information for sightseeing with linguistic features on TripAdvisor. Therefore, in this study, we model the ranking of tourist spots on TripAdvisor by considering effective features with linguistic features and statistical features.

B. Recommendation of tourism spots

Impressions of tourist spots on SNS are increasing. These posts are comments by users with higher recency than reviews on websites for obtaining tourist information. Many studies applied these posts to recommendation systems for sightseeing. Borras et al. [10] classified these recommendation systems technically and considered their importance on trips. Ye et al. [11] [12] proposed a method to recommend geographical information, such as Points-Of-Interests (POIs), from check-in information and user attributes in Foursquare [13]. Choudhury et al. [14] and Lim et al. [15] proposed a method to recommend POI with user behavior on Flickr [16]. Ishihara et al. [17] reported a recommendation system for sightseeing while considering posts on Twitter from the perspective of sensitivity engineering. Mizutani et al. [18] proposed a recommendation system for sightseeing that is more suitable for users individually using posts on SNS, such as Twitter.

For these studies, they do not consider the effects of changes by enviromental factors and so on, but Missaoui et al. [19] pointed out the importance of taking into account changes of a user's preferences and environments for recommendation systems, and proposed a method to predict a user's preferences in the tourism domain. Additionally, these studies analyzed recommendation system for sightseeing only with posts and metadata on the SNS. However, many noise posts are sent to SNSs. The reliability of the information is questionable, as described in Section I. This study applies feature selection while considering the contribution rates of logistic regression. Also, this study proposes a system that generates a ranking of tourist spots incorporating a trend of tourist spots by modeling the ranking on the TripAdvisor where the reliability of information is higher than Twitter as training data.

In addition, various ranking methods use tweets. Duan et al. [20] proposed ranking recommendation to search tweets using ranking learning. Qupta et al. [21] proposed a recommendation method to rank important events with various linguistic features on tweets. Chang et al. [22] reported a method to remedy a shortage of recency on the ranking of web searches using tweets. This study models the ranking of tourist spots by application of ranking learning using tweets, which is one of the ranking methods.

III. Analysis of Features

As described herein, to examine the effective features for ranking learning, we divide the ranking of tourist spots into certain intervals and classified them. This section presents examination of the effective features for the classification of tourist spots and ranking learning from the distribution of each feature while considering linguistic features and statistical features of tweets.

A. Datasets

For this study, we use tourist spots among the top 100 rankings in Tokyo on TripAdvisor acquired in 2017. To analyze seasonal changes by month, we extracted tourist spots for which more than 100 tweets are posted each month. To obtain tweets at each tourist spot, we used GooglePlaceAPI [23] and NominatimAPI [24]. We were able to find tweets for a tourist spot if the latitude and longitude annotated to a tweet would be within the area of tourist spots obtained using these APIs. The number of tweets written in Japanese in 2017 was 1,691,521 at the 50 tourist spots.

B. Discussion of linguistic features

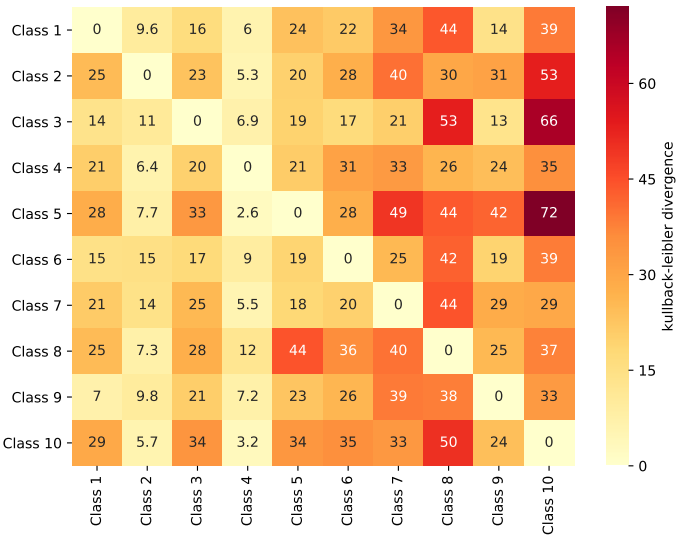


Figure 1. Distance found with KL-Divergence of Unigram for each class.

We first consider the distribution of linguistic features for each tourist spot. We compare Unigram, which is the probability distribution of words for each tourist spot. Tweets used for analyses were pre-processed using morphological analysis, for which we used Mecab [25]. Additionally, we used

TABLE I. TOP FIVE WORDS WITH HIGH BIAS IN UNIGRAM FOR EACH TOURIST SPOT.

Rank	Class 1	Class 2	Class 3	Class 4	Class 5
1	Taito Ward	Ota Ward	Shibuya Ward	Ota Ward	Ota Ward
2	Harajuku Station	Tokyo International Airport	Shibuya Station	Taito Ward	Tokyo International Airport
3	Tokyo International Airport	Passenger Terminal	Setagaya Ward	Tokyo International Airport	Passenger Terminal
4	Shibuya Ward	Chuo Ward	Tokyo International Airport	Passenger Terminal	Taito Ward
5	Tokyo Teleport Station	Haneda Airport	Tokyo Teleport Station	Tokyo Teleport Station	Haneda Airport
Rank	Class 6	Class 7	Class 8	Class 9	Class 10
1	Harajuku Station	Ota Ward	Tokyo Teleport Station	Ota Ward	Toyoshima Ward
2	Shibuya Ward	Taito Ward	Taito Ward	Nippon Budokan	Shibuya Ward
3	Jinbouchou Station	Minato Ward	Nippon Budokan	Minato Ward	Sunshine City
4	Taito Ward	Shinjuku Ward	Shibuya Ward	Sumida Ward	Nippori Station
5	Tokyo International Airport	Nippon Budokan	Tokyo International Airport	Tokyo Teleport Station	Arakawa Ward

the Good–Turing smoothed unigram language model in Kylm [26] for extracting Unigrams.

To examine the usefulness for classification with linguistic features, we divided tourist spots into 10 classes of $\{Class1, Class2, \cdots, Class10\}$ with class width of 10 based on the rank. In this study, our focus is not to reproduce the ranking of tourist spots on TripAdvisor completely, but to generate the ranking incorporating changes in popularity of tourist spots. Therefore, we set the number of classes to 10 for analyzing features that affect large ranking fluctuations rather than small fluctuations. Figure 1 is the result of comparing the difference of Kullback-Leibler Divergence (KL-Div) between Unigrams of linguistic features in each class. KL-Div has no symmetry which is an axiom of distance. Therefore, it can not be defined precisely as a distance. We confirmed the difference of KL-Div for each class. We assumed that linguistic features are useful for classification to a certain degree. Moreover, a large difference from $Class7$ to $Class10$ exists on average, which was regarded as caused by a special word distribution. Although we also compared KL-Div in each class by month, a similar tendency was apparent.

Table I presents results of comparing contribution rates of wordwise KL-Div in each class; then it displays the top five words of high bias. Many place names, station names, and facility names are apparent throughout all classes. Additionally, we confirmed a similar tendency even when verified by each month. As a cause for these results, we considered that many tourist spots are located in some areas belonging to a particular class. After examining the relation between the geographical factor belonging to a particular area and the ranking of tourist spots, we should consider whether a need exists to eliminate these geographical words during pre-processing. For the results described above, a certain difference is apparent between the distributions of words on linguistic features. In addition, results show that linguistic features contribute to the classification of ranking tourist spots. However, Twitter has many noise features, which implies that preprocessing and feature selection are extremely important.

C. Discussion of statistical features

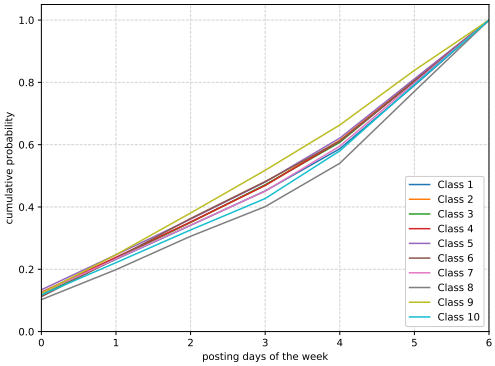
TABLE II. MEAN CORRELATION COEFFICIENT BETWEEN DIMENSIONS OF THE DISTRIBUTION OF STATISTICS FOR EACH TOURIST SPOT.

Date	Time	Character	Emoji	Emoticon	Post	Repeat	Revisit
0.52	**0.40**	0.22	0.29	**0.51**	0.20	0.22	**0.44**

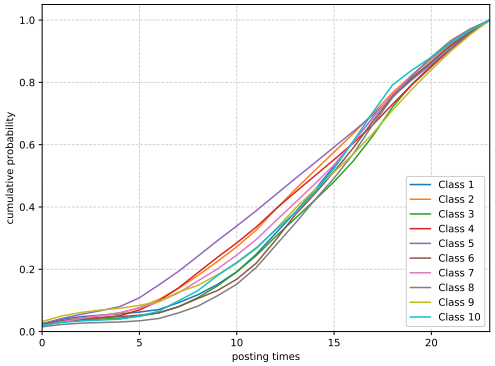
Next, we consider the distribution of statistical features for each tourist spot. We compare the cumulative probability distribution in the number of posting days of the week and times (hereinafter, Date and Time) and the number of posted characters, emoji, and emoticons (hereinafter Character, Emoji and Emoticon), which are statistics by tweets for each tourist spot. Similarly, we also compare the cumulative probability distribution in the number of posts, repeat posts and revisits (hereinafter Post, Repeat and Revisit), which are statistics by users on Twitter for each tourist spot.

To examine the usefulness for classification with statistical features, we divided tourist spots into 10 classes of $\{Class1, Class2, \cdots, Class10\}$ similarly to the process described in Section III-B. Figure 2 is the result of comparing the cumulative probability distribution of statistical features in each class. Figure 2a presents the cumulative probability distribution of each class. The horizontal axis shows days of the week, from Monday through Sunday. We confirmed distinct differences between the weekdays (0–4) and holidays (5–6) from data shown in Figure 2a. Figure 2b portrays the cumulative probability distribution of posting times of tweets posted in each class. The horizontal axis shows 24 hours in a day. A similar tendency is apparent from Figure 2a between the daytime (6–18) and night time (0–5, 18–23). Therefore, we consider that these features make a certain contribution to classification. Figure 2c displays the cumulative probability distribution of the number of posted characters in each class. The distribution in each class is complicated. Therefore, it remains unclear whether this feature contributes to classification. Figure 2d exhibits the cumulative probability distribution of the number of posted emoji in each class. Because of differences in each class, we can infer that this feature contributes to classification to a certain degree. However, a high bias is apparent by the number of occurrences of 0 throughout all classes. Figure 2e depicts the cumulative probability distribution of the number of posted emoticons in each class. We were unable to see differences in each class. Therefore, this feature is not regarded as contributing to classification.

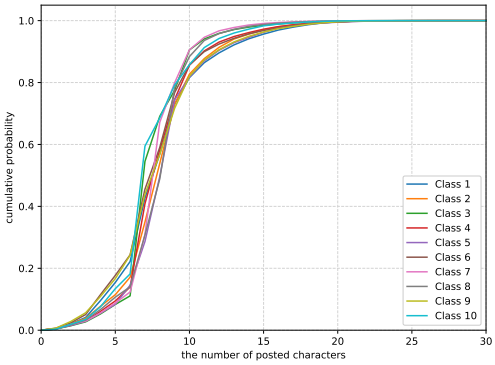
Figure 2f is the cumulative probability distribution of the number of posts by user on Twitter in each class. We were able to detect a certain difference in each class. However, a high bias exists by the number of occurrences of 1, similarly to Figure 2d. Figure 2g is the cumulative probability distribution of the number of repeat posts by users on Twitter in each class. There are differences in each class and a high bias to the number of occurrences of 0 similarly in Figure 2d and Figure 2f. Figure 2h is the cumulative probability distribution of the number of revisits by users on Twitter in each class, which closely resembles Figure 2g. From the above, we were able to confirm several differences in the distributions of



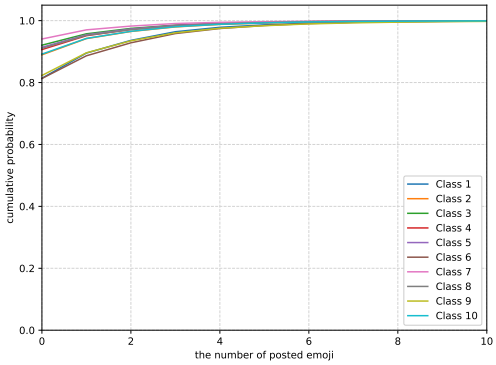
a. Cumulative distribution functions based on posting days of the week of tweets.



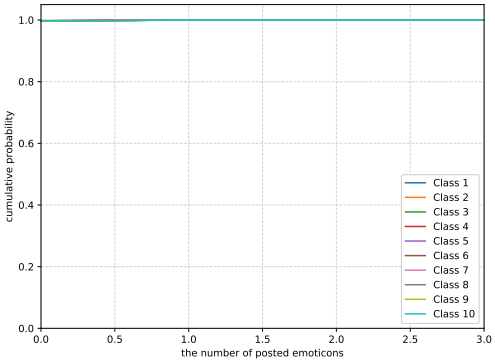
b. Cumulative distribution functions based on posting times of tweets.



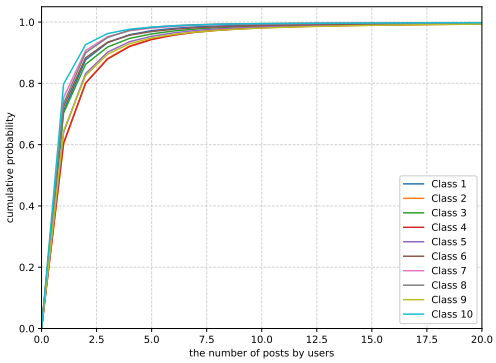
c. Cumulative distribution functions based on the number of posted characters on tweets.



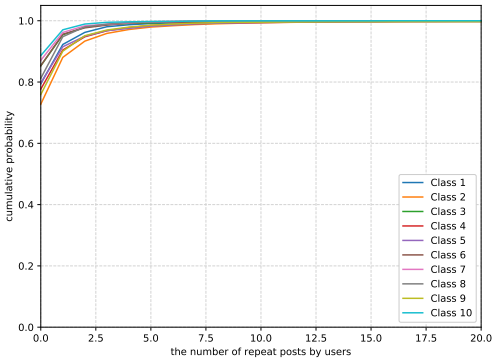
d. Cumulative distribution functions based on the number of posted emoji on tweets.



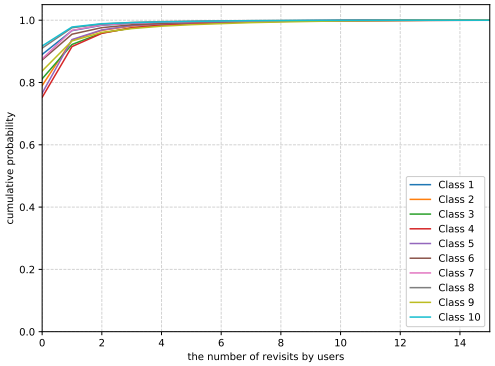
e. Cumulative distribution functions based on the number of posted emoticons on tweets.



f. Cumulative distribution functions based on the number of posts by users on Twitter.



g. Cumulative distribution functions based on the number of repeat posts by users on Twitter.



h. Cumulative distribution functions based on the number of revisits by users on Twitter.

Figure 2. Cumulative distribution functions of the statistical features for each tourist spot.

 ISBN: 978-1-61208-690-3

statistics by tweets and Twitter users. However, a high bias exists in the distribution of statistics by Twitter users. We must consider the influence of classification.

Table II shows mean correlation coefficients between dimensions of the features in each statistic. The criterion of correlation is 0.4, which is generally regarded as representing some degree of correlation. We should consider the possibility of negative influence on classification by multicollinearity because of a high mean correlation coefficient between dimensions in Date, Time, and Revisit. As described earlier, a certain difference exists between the distribution of the statistical features. We consider that several statistics, such as Date and Time, will make some contribution to classification. However, we must devote some attention to the influence of multicollinearity in classification.

IV. EXPERIMENT

In this section, to model the ranking of tourist spots on TripAdvisor for generating the ranking using tweets, we used 50 tourist spots in Section III-A. First, we analyze features by logistic regression and perform ranking learning using the result.

A. Classification

TABLE III. COMPARISON OF SUBSET ACCURACY BY LOGISTIC REGRESSION WITH LINGUISTIC FEATURES.

Features	Accuracy (+/- Error Rate)
Unigram	0.90 (+/- 0.02)
Unigram + IG	0.71 (+/- 0.02)
Unigram + TFIDF	**0.92 (+/- 0.01)**
Bigram	**0.94 (+/- 0.01)**
POS Unigram	0.26 (+/- 0.03)
Word2Vec	0.38 (+/- 0.03)

TABLE IV. TOP 20 WORDS WITH HIGH CONTRIBUTION RATES OF SPEARMAN'S RANK CORRELATION COEFFICIENT (SRCC) OF 0.4 OR MORE IN CONTRIBUTION RATES BY LOGISTIC REGRESSION.

Word	Coefficient
Collection	**0.6261**
Beautiful	**0.6071**
Walk	0.5860
Doing	0.5784
Fall	0.5599
Amazing	**0.5342**
Meal	0.5227
Buy	0.4644
Last night	0.4512
Meet	0.4497
Customer	0.4369
Good	**0.3802**
Many	0.3717
Flow	0.3438
Combination	0.3294
Ballet	0.3273
Old man	0.3082
Workplace	0.2948
Near	0.2648
Sit	0.2357

Here, we perform logistic regression and analyze effective features for ranking learning from contribution rates. We used the scikit-learn [27] algorithm for the implementation of logistic regression. To generate the highest performance model, we performed a grid search with five cross validation for hyperparameters.

TABLE V. COMPARISON OF SUBSET ACCURACY BY LOGISTIC REGRESSION WITH STATISTICAL FEATURES.

Features	Accuracy (+/- Error Rate)
All Statistics	0.39 (+/- 0.03)
All Statistics - Emoji - Emoticon	0.40 (+/- 0.03)
Tweet Statistics	0.58 (+/- 0.02)
User Statistics	0.25 (+/- 0.04)
Date + Time + Character	**0.64 (+/- 0.05)**

We used linguistic features and statistical features for logistic regression. For this experiment, we used four linguistic features of Unigram, Bigram, Part-Of-Speech (POS) Unigram, and Word2Vec. However, we performed feature selection with Information Gain (IG) and Term Frequency-Inverse Document Frequency (TFIDF) to Unigram by reducing dimensions of the features below each mean value. As statistical features, we used statistics by tweet (hereinafter, Tweet Statistics), statistics by users on Twitter (hereinafter, User Statistics) and All Statistics combining them in Section III-A.

For classification with logistic regression, we divided tourist spots into 10 classes of $\{Class1, Class2, \cdots, Class10\}$ with the class width of 10 based on the rank. Table III presents the subset accuracy of logistic regression with linguistic features. This subset accuracy takes 1 when the set of labels predicted by the test data exactly matches the set of labels on the answer data in the multi-class classification. As a result, the model generated by Unigram + TFIDF and Bigram has the highest performances in linguistic features. Because 3,521 dimensions of Unigram + TFIDF are far fewer than the 1,203,938 dimensions of Bigram, we infer that the effective features on Unigram + TFIDF were extracted by feature selection for classification.

Table IV shows the top 20 words of classification contribution rates for which Spearman's Rank Correlation Coefficient (hereinafter, SRCC) is 0.4 or more in the classification model in Unigram + TFIDF. As with Section III-B, we set the criterion of correlation to 0.4. We were able to confirm that "Collection", which seems to express a short-term event and "Beautiful" and "Good", which are favorable impressions of tourist spots, are placed higher in the ranking. In the case of comparing words solely by contribution rates, we confirmed that many geographical words appear in the top of the ranking. We consider that feature selection by rank correlation coefficient is useful to a certain degree.

Table V presents the accuracy of logistic regression with statistical features. As described in Section III-B, some features, such as Emoji and Emoticon, do not contribute to the classification of ranking of tourist spots. In classification with tweet statistics and user statistics, we confirm that the classification with tweet statistics shows high performance. Moreover, the higher accuracy of classification was recorded by eliminating Emoji and Emoticon based on Section III-B. As a result, Date + Time + Character, which differed in each class and which were free from bias in Section III-B exhibited the highest performance. As described previously, we consider that user statistics did not make a contribution to classification because of high bias. Next, we will model the ranking of tourist spots by ranking learning after selecting features with a high rank correlation coefficient of classification.

 ISBN: 978-1-61208-690-3

TABLE VI. COMPARISON OF PAIRWISE ACCURACY AND MEAN NDCG BY THE RANKING LEARNING MODEL WITH FEATURES: LF, UNIGRAM + TFIDF; SF, DATE + TIME + CHARACTER.

Features	Pairwise Accuracy	Mean NDCG
LF	0.8864	0.8918
SF	0.6666	0.1741
LF + SF	0.8329	0.7692
LF (SRCC $>$ 0.4)	0.7915	**0.8998**
SF (SRCC $>$ 0.4)	0.6440	**0.2038**
LF + SF (SRCC $>$ 0.4)	0.8528	**0.8075**

B. Ranking Learning

From the result presented above, we performed ranking learning with linguistic features and statistical features to model the ranking of tourist spots on TripAdvisor. We used RankSVM, which is a ranking learning algorithm incorporating Kendall's rank correlation coefficient to Support Vector Machine, for which we adopted the algorithm of LIBSVM [28] proposed by Lee et al. [29]. To generate the highest performance model, we also performed a grid search with five cross validation for hyperparameters.

As features for RankSVM, we used effective features for classification in Section IV-A, which is Unigram + TFIDF in Linguistic Features (hereinafter, LF) and Date + Time + Character in Statistical Features (hereinafter, SF). Additionally, we verify the usefulness of SRCC for ranking learning in each feature.

Table VI shows pairwise accuracy and mean Normalized Discounted Cumulative Gain (hereinafter, NDCG) of RankSVM in each features. NDCG takes a value from 0 to 1. It is closer to 1 when the result of ranking conforms to answer data, which is a widely used prediction result indicator of ranking method. Additionally, we apply selected features for which SRCC is higher than 0.4 in Table IV. Results show that the values of NDCG were improved by feature selection with SRCC in both linguistic features and statistical features. We also confirmed a high performance of ranking learning while combining both features. As a result, the model generated by ranking learning with LF exhibited the highest performance.

V. Discussion

We performed ranking learning by considering SRCC of contribution rates by logistic regression. In terms of linguistic features, from the accuracy of classification in Table III, large differences exist among tourist spots. Many geographical words were ranked higher when comparing the features solely by contribution rates according to Section III-A. However, considering feature selection by rank correlation coefficients, we were able to confirm that words which give positive impressions to tourist spots were ranked higher in Table IV.

We also discussed statistical features by tweets and Twitter users. Regarding contribution rates, it is apparent that Date, Time, and Character, which have statistics by tweets, have a high contribution rate. However, we confirmed certain variations in the distribution of Post, Repeat and Revisit in Section III-B. We consider that these statistics did not contribute to classification because the bias to the number of occurrences of 0 or 1 is too large. Additionally, we confirmed a certain correlation for dimensions of the distribution of Date and Time in Table II. Regarding multicollinearity, this is regarded as negatively influencing the accuracy of classification. Therefore, if we were to have sparse aggregation of statistics, then we might improve the accuracy of classification with statistical features.

In ranking learning, we performed feature selection based on SRCC of contribution rates by logistic regression. Consequently, NDCG recorded high performance for both linguistic features and statistical features. We infer that feature selection by a rank correlation coefficient of contribution rates is useful to a certain degree. Both pairwise accuracy and mean NDCG were also the highest scores with linguistic features. As described previously, the case of performing precise preprocessing demonstrates that linguistic features are extremely important for modeling the ranking of tourist spots.

Table VII presents the result of applying recent tweets to the model in ranking learning by LF and LF (SRCC $>$ 0.4), which are high performance in Section VI. We confirm that more tourist spots hold seasonal events at the top of the ranking by Twitter than by TripAdvisor. For instance, "The National Art Center, Tokyo", "Sunshine City" and "Tokyo International Forum" are in the top three of the ranking. These spots hold exhibitions, events or other activities frequently. Additionally, there are some differences in ranking from the viewpoint of rank correlation for feature selection. When the contents of these increase, we can not determine if it is profitable. Therefore, it is necessary to conduct an experiment by users.

As a result, we could confirm that linguistic features are more effective for ranking of tourist spots incorporating changes in popularity than statistic features. We speculate that this is because the number of dimensions of our statistic features is not enough to deal with this problem. Finally, the result of our experiment demonstrated that tourist spots which hold events frequently are ranked higher by applying recent tweets to the model in ranking learning. Although further experiments would be required, we could also confirm that the values of NDCG are improved by considering SRCC of contribution rates.

VI. Conclusion and Future Work

As described in this paper, we generated a ranking that incorporates popularity variation by modeling the ranking of tourist spots on TripAdvisor using tweets. Twitter has numerous noise posts. Therefore, we performed feature selection by a rank correlation coefficient of contribution rates by logistic regression. Results demonstrate that we were able to improve the performance of the model in ranking learning. Eventually, the model by RankSVM showed the highest NDCG of 0.89. However, pairwise accuracy decreased from the highest score of 0.88 to 0.79.

Future work must address the contribution rates themselves by logistic regression for feature selection. As described in this paper, we used only rank correlation coefficients for feature selection because the contribution rate of geographical words to classification is too large. Therefore, we expect improvement of classification accuracy in ranking learning by eliminating geographical words during pre-processing. Additionally, it is necessary to consider evaluation after proposing ranking by the proposed method to the user. From results of demonstration experimentation, we expect to verify whether tourist spots to enjoy most at the time of visit are suggested in the ranking, which is information that users actually need.

 ISBN: 978-1-61208-690-3

TABLE VII. TOP TEN TOURIST SPOTS BY THE RANKING OF TRIPADVISOR AND TWITTER.

Rank	TripAdvisor	Twitter	Twitter (SRCC $>$ 0.4)
1	Ryogoku Kokugikan	The National Art Center, Tokyo	Sunshine City
2	Asakusa	Sunshine City	The National Art Center, Tokyo
3	Roppongi Hills	Tokyo International Forum	Tokyo International Forum
4	Chidorigafuchi	Bunkamura	Bunkamura
5	Happoen Garden	Roppongi	Ameya Streets
6	Meiji Jingu Shrine	Yanaka Cemetery	Spa LaQua
7	Tokyo Camii & Turkish Culture Center	Odaiba Palette Town	Omoide Yokocho
8	Senso-ji Temple	Ginza Namiki Streets	Roppongi
9	Tokyo Metropolitan Government Buildings	Ryogoku	Tokyo Character Streets
10	Tokyo National Museum	Kitanomaru Park	Imperial Palace

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Numbers 16K00157 and 16K16158, and a Tokyo Metropolitan University Grant-in-Aid for Research on Priority Areas "Research on Social Big Data."

REFERENCES

[1] TripAdvisor, "TripAdvisor: Read Reviews, Compare Prices & Book." [Online]. Available: https://www.tripadvisor.com/ [Accessed: Jan., 2019].

[2] Z. Xiang, D. Wang, J. T. O' Leary, and D. R. Fesenmaier, "Adapting to the internet: trends in travelers' use of the web for trip planning," Journal of Travel Research, vol. 54, no. 4, 2015, pp. 511–527.

[3] U. Gretzel and K. H. Yoo, "Use and impact of online travel reviews," Information and communication technologies in tourism 2008, 2008, pp. 35–46.

[4] Twitter, "Twitter. It's what's happening." [Online]. Available: https://twitter.com/ [Accessed: Jan., 2019].

[5] Facebook, "Facebook." [Online]. Available: https://www.facebook.com/ [Accessed: Jan., 2019].

[6] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in Proceedings of the 20th international conference on World wide web. ACM, 2011, pp. 675–684.

[7] R. Filieri, S. Alguezaui, and F. McLeay, "Why do travelers trust tripadvisor? antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth," Tourism Management, vol. 51, 2015, pp. 174–185.

[8] A. Mukherjee, V. Venkataraman, B. Liu, and N. S. Glance, "What yelp fake review filter might be doing?" in ICWSM, 2013, pp. 409–418.

[9] I. Guy, A. Mejer, A. Nus, and F. Raiber, "Extracting and ranking travel tips from user-generated reviews," in Proceedings of the 26th international conference on world wide web. International World Wide Web Conferences Steering Committee, 2017, pp. 987–996.

[10] J. Borràs, A. Moreno, and A. Valls, "Intelligent tourism recommender systems: A survey," Expert Systems with Applications, vol. 41, no. 16, 2014, pp. 7370–7389.

[11] M. Ye, P. Yin, and W.-C. Lee, "Location recommendation for location-based social networks," in Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems. ACM, 2010, pp. 458–461.

[12] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. ACM, 2011, pp. 325–334.

[13] Foursquare, "Foursquare." [Online]. Available: https://foursquare.com/ [Accessed: Jan., 2019].

[14] D. Choudhury et al., "Automatic construction of travel itineraries using social breadcrumbs," in Proceedings of the 21st ACM conference on Hypertext and hypermedia. ACM, 2010, pp. 35–44.

[15] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency," Knowledge and Information Systems, vol. 54, no. 2, 2018, pp. 375–406.

[16] Flickr, "Find your inspiration. - Flickr." [Online]. Available: https://www.flickr.com/ [Accessed: Jan., 2019].

[17] S. Ishihara, M. Nagamachi, and T. Tsuchiya, "Development of a kansei engineering artificial intelligence sightseeing application," in International Conference on Applied Human Factors and Ergonomics. Springer, 2018, pp. 312–322.

[18] Y. Mizutani and K. Yamamoto, "A sightseeing spot recommendation system that takes into account the change in circumstances of users," ISPRS International Journal of Geo-Information, vol. 6, no. 10, 2017, p. 303.

[19] S. Missaoui, M. Viviani, R. Faiz, and G. Pasi, "A language modeling approach for the recommendation of tourism-related services," in Proceedings of the Symposium on Applied Computing. ACM, 2017, pp. 1697–1700.

[20] Y. Duan, L. Jiang, T. Qin, M. Zhou, and H.-Y. Shum, "An empirical study on learning to rank of tweets," in Proceedings of the 23rd International Conference on Computational Linguistics. Association for Computational Linguistics, 2010, pp. 295–303.

[21] A. Gupta and P. Kumaraguru, "Credibility ranking of tweets during high impact events," in Proceedings of the 1st workshop on privacy and security in online social media. ACM, 2012, p. 2.

[22] Y. Chang et al., "Improving recency ranking using twitter data," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 4, no. 1, 2013, p. 4.

[23] Google, "Places - Google Maps Platform - Google Cloud." [Online]. Available: https://cloud.google.com/maps-platform/places/ [Accessed: Jan., 2019].

[24] OpenStreetMap, "OpenStreetMap Nominatim: Search." [Online]. Available: https://nominatim.openstreetmap.org/ [Accessed: Jan., 2019].

[25] Kyoto University, "MeCab: Yet Another Part-of-Speech and Morphological Analyzer." [Online]. Available: http://taku910.github.io/mecab/ [Accessed: Jan., 2019].

[26] G. Neubig and X. Yao, "Kylm - Kyoto Language Modeling Toolkit." [Online]. Available: http://www.phontron.com/kylm/ [Accessed: Jan., 2019].

[27] scikit-learn, "scikit-learn: machine learning in Python." [Online]. Available: http://scikit-learn.org/stable/ [Accessed: Jan., 2019].

[28] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," ACM Transactions on Intelligent Systems and Technology, vol. 2, 2011, pp. 27:1–27:27.

[29] C.-P. Lee and C.-J. Lin, "Large-scale linear ranksvm," Neural computation, vol. 26, no. 4, 2014, pp. 781–817.

 ISBN: 978-1-61208-690-3

Defend the Template – Supporting the Standardisation and Harmonisation of Process Change Programmes with SAP templates

Markus Grube

VOQUZ IT Solutions GmbH
Hamburg, Germany
email: markus.grube@voquz.com

Abstract – **Many companies are aware of the benefits of using a template in SAP implementation. This research paper deals with the question of how a company uses a template in a business process management environment and, above all, what happens when there are requests for changes to the template. Based on a literature review and personal interviews, this article describes the promises that manufacturing companies make about the template implementation within SAP business process management and why an overarching goal could be: 'Defend the template'.**

Keywords – SAP; ERP; BPM; Business Process Management; Template; Process change.

I. INTRODUCTION

During this research on the relationships between Business Process Management (BPM), maturity models and the utilisation in the environment of SAP Enterprise Resource Planning (ERP) systems, one of the interviewees provided the following slogan regarding the introduction of SAP business processes and the usage of templates: "Defend the template". The application of this slogan was described by the interviewed expert as a fundamental principle for process standardisation and process implementation in his company. In general, templates are used when implementing SAP business processes. There are numerous examples in the literature, and the author is also aware of this from practical experience in SAP projects.

Foth [1] defines a template as a unified set-up containing the necessary documentation and programmes for construction of a new system. The use of a template offers numerous advantages, for example:

- reduction in costs of system development [2]
- re-use of processes among various enterprises
- reduction of required manpower [3]
- increased level of conformity and data consistency [4]
- greater stability of business processes
- quicker worldwide process improvements [1].

The given slogan of "Defend the template" raised the question of how the defence of a template can support the use of process templates in business processes. This paper considers the slogan and examines how templates are used by the interviewed experts in standardising and harmonising business processes.

Following this introductory section, Section 2 discusses the relevant literature and develops the research question for this paper. Section 3 describes the research methodology. Section 4 provides a summary of the findings. Section 5 contains an analysis for the research question, and finally, Section 6 presents the conclusion and suggestions for possible future research.

II. LITERATURE REVIEW

With the globalisation of business activities, there is a growing need to structure and shape business across an enterprise. For example, when setting up an overall process, binding corporate rules should be integrated [5].

In general, 'the main goal of process standardisation is the development of one standard or best-practice process to be used as a template for all instances of the process' [2]. Gavidia [4] illustrates that it is common practice to develop a template as the base configuration for all areas and locations of a company to increase compliance and data consistency. This new template will then have a consistent set of reports and will be adapted to the existing business processes within the company.

Systems like SAP ERP offer a variety of "best-practice" processes that exist for different industries and business processes. However, for various reasons, the suggested best practice does not work for most companies: thus many companies define their own BPM templates to meet their specific needs [6]. In addition, it may be necessary to distinguish between a global template and local requirements, and to define mechanisms that solve the problems resulting from these differences [7].

From an Information Technology (IT) perspective, it makes sense to standardise business processes and their configurations and transactions. Costs are reduced if IT processes do not have to be changed and if the business can adapt to the IT, rather than vice versa [8]. However, this approach is criticized in the literature. For example, Gavidia [4] reports that blanket implementation of a uniform template reflects at the failure of management and does not ensure that the process fits with existing customer relationships. A process template can only be effective if it takes into account the knowledge and requirements of all areas of a company. It is not sufficient for the template development to only consider a basic IT process or system from the parent company.

 ISBN: 978-1-61208-690-3

For successful template implementation, all employees and groupings must be included. Only then can productive implementation be carried out successfully, with employees understanding the benefits and accepting the changes. A company should also consider that a template rollout requires a long-term implementation phase, which can take 2–5 years, and may require a budget of $10 million or more [7].

The enforcement of a template is a practical difficulty in the process rollout. Often, this causes problems for a company gets stuck and business goals are affected. A global template must be approved and accepted by all companies and employees. This requires discipline from all employees involved in the implementation of the template, because things may be repeated during the rollout or there may be resistance against the implementation [5].

The literature describes what a template is and how it can be defined. Additionally, statements in the existing literature show that the implementation of a template can be quite difficult. We must consider: what instructions should be given to a team introducing a template? What different personal recommendations can respondents give in addition to the literature? The literature recommends the use of templates: however, can a slogan like "Defend the template" support the introduction of a process template? Based on these issue, the research question (RQ) of this paper can be defined as:

RQ: How and why can the slogan "Defend the template" support the use of process templates?

III. RESEARCH METHODOLOGY

The research philosophy for this paper adopts a post-positivist position and is based on the perspectives of Ryan [9] and Guba [10]. The goal of research from the post-positivist perspective is to generate new knowledge that other people can learn from and on which they can base their own decisions [11].

For this research, semi-structured and in-depth interviews are used to gather practical experience of companies that use SAP templates and BPM. These interviews are followed by an analysis of the best explanation regarding the collected facts, and this is done within an abductive approach [12].

As an explanatory study, described by Collis and Hussey [13], this research uses interviews with experts who have already gained practical experience in the areas of BPM in ERP projects and the use of SAP templates. It is quite possible that the inclusion of other experts would have led to a different result, but from the post-positivist position, this research understands that the world is much more complex than the opinions of some experts suggest.

To address the research question, this research combines two research methods. Semi-structured interviews are conducted with experts in their field. In addition, a brief literature search is carried out. The goal of this combination is to achieve greater depth within the research and the complex environment being studied [14]. The following Figure *1* illustrates the research strategy and the different sources used for this research:

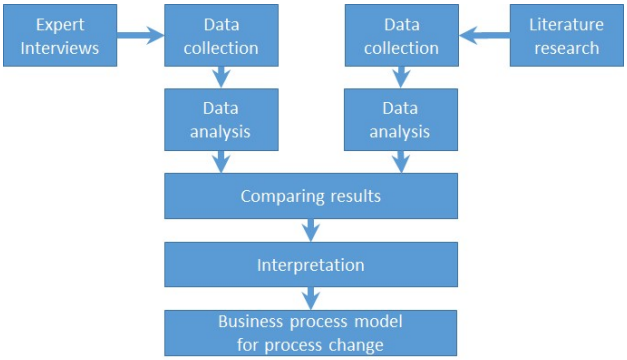


Figure 1. Method of data collection and analysis

The time horizon for this research is a cross-sectional snapshot study [15]. The research analyses the current practical situation and evaluates the state of affairs at the time of the study [16].

For the interviews, an expert is defined as a person who has experience in the areas of SAP and BPM. As a prerequisite to be an expert in this context, practical experience of approximately ten years or more in the mentioned areas is expected. Due to the small number of interviews, this research does not differentiate between which module of SAP these experts use or within which industrial sector they currently work. This research paper has the general objective of analysing the slogan "defend the template" in term of the general standardisation and harmonisation of process change programs within SAP environments.

Respondents were selected with the aim of gaining as much expert knowledge as possible from practice. Using semi-structured interviews provided flexibility to scrutinise the understanding and explanation of the experts' opinions [17].

The qualitative data analysis and the comparison of the transcribed interviews were supported by using the software tool MAXQDA. The coding functionality of the MAXQDA tool was used to analyse the experts' answers. The software was used to manually encode the qualitative data within the interviews and to obtain an overview of all the experts' answers based on the encodings.

Due to the small number of interviews, the aim is not to reach an overall and general conclusion for all industries, organisations or SAP modules. However, the results of this research paper could be used as part of the theoretical grounding for further research projects.

IV. FINDINGS

The search for potential interviewees was conducted within the author's personal network of business and personal contacts in Germany, Austria, and Switzerland. A total of 64 people were identified as potential experts in the business process and SAP environment. These people were then invited for an interview. From this initial pool, 11 people confirmed that they were willing to be interviewed. The personal, face-to-face interviews revealed the following findings, which the

 ISBN: 978-1-61208-690-3

experts have identified in their day-to-day business when using SAP and business process templates. All statements and insights described here are the personal opinions of the experts. Due to the total duration of the interviews being over 14 hours combined, only excerpts can be presented here.

The experts explained within the interviews that the SAP ERP system offers a variety of "out-of-the-box" business processes for the general SAP system, as well as for special sectors, such as real estate or the pharmaceutical industry. With these basic processes, a company can cover almost all of its standard process cases and can set up a solid basis for its daily business processes. Such SAP processes support a company in standardising its processes and implementing them relatively quickly. For example, sales processes are similar in different companies and across different industries. It is helpful to use a standard SAP template for the best-practice approach as starting point, and to show all the steps that belong to a process sequence. In this standard SAP process template, there is a lot of industry expertise that has been accumulated from various industries and many projects. However, this does not mean that this template has been optimised for a particular company, nor that the company's own process steps have that have been developed and implemented in-house, cannot deviate from the SAP process specifications. The experts pointed out that such an approach with given SAP processes is usually only used if a company has no existing processes or IT infrastructure.

Normally, a company has already gained its own experience and has established its own process flows. Therefore, it would not fully use the SAP "best-practice" standard template, instead developing its own process template. This template must then have a certain degree of flexibility. A tax-paying process in Germany is different from on in, say, Brazil or India. Therefore, process templates must be generated that can be applied globally across all countries. These processes must have a certain amount of local variation because of the different rules. Local adaptation must only be set up where a process, or a process step, is genuinely local and cannot be supported by a global SAP template. However, this means that the development of a company's own solution has nothing to do with the global template, and thus a shadow process is created, which is only used for local requirements. The disadvantage is that process adaptation then becomes more difficult because the adaptation is known only locally, on the level of one company or business unit.

A process template should be as flexible as possible from the beginning of the set-up. It should be scalable and applicable to specific local needs. In the interviews, the experts mentioned this as the biggest challenge for a process template. It must be considered whether new requirements that emerge can still be incorporated into the applied process template, and how quickly changes can be implemented company-wide. Therefore, a process template must allow for a certain amount of agility.

Regarding the use of standard processes, one expert mentioned the practical problem that business consultancies frequently use standard SAP templates without examining the individual processes within a company. This means that templates are established without optimising the existing processes to suit into the special behaviours of each company.

The experts also want to achieve harmonisation and standardisation of the process world through the introduction of a process template. This means that not every local manufacturing facility conduct its business according to a business process it wants and likes, but that a company-wide standard is set up and must be followed. The more complex and individualised the established processes are, the harder it is to run the business on one new company-wide process template. Uniform processes can create synergies and clarity, and a process template should help to minimise fragmentation and process variations in the company.

One expert explains that, when the processes are described in much more detail, it is no longer a template description, but more like a localisation. Localisation of processes is a challenge because it can only be developed in collaboration with the local business departments. One expert highlighted that the business must then maintain the processes, because only the local team knows how they really work. Another expert declared that the aim is to provide a standardised process template that is not a localisation and could be protected against further splitting and any diversification within the organisation.

Based on the previously discussed findings, one interviewee mentioned the slogan "Defend the template". The goal of the slogan is that a business should no longer be able to make decisions on process changes. Another expert also declared that a template should be a leader for all processes, with functional processes introduced via a company-wide, valid template. The question is then no longer how the individual divisions or areas carry out task; instead, one valid process is established throughout the whole company. A process in the enterprise can then only be arranged based on these template specifications; however, there exist a number of different variations, which are offered by a catalogue structure within the template. By minimising process diversity, companies expect enormous savings potential.

The goal of a process template must be to offer stable and mature process solutions for a company. These processes are then rolled out company-wide, and it is ideal to undertake a process without local on-site adaptation, because these local adaptations cost in term of time, money, and manpower.

One difficulty is to avoid being too detailed in process description for a template. The more detailed the description of a process, the more likely it is to be a description of localised or specific particularities. The goal must be to develop an enterprise-wide process template that works equally well in a five-person facility as for a business unit with 800 employees in another country.

It is important that a business should learn that not every process has to be individualised, and that specific and individualised processes are rarely the best solution for a company. Using fewer individualised solutions and more

 ISBN: 978-1-61208-690-3

standard process templates means that process changes can be implemented more quickly. Every process change, every system transport, and every release change could collapse and thus carries a significant risk, which is reduced if there are fewer variations in the template catalogue that need to be taken into account.

The slogan "Defend the template" aims to encourage the optimal use of process templates. With this harmonisation, businesses should and can be convinced that the use of standard processes is not a negative approach.

Even if a template is to be defended, it must be possible to change it; however, such a change should never be decided by a business entity itself. The sovereignty over processes must be located with a superordinate process organisation, which consists of, for example, process owners and process managers. Only this process organisation can decide whether and to what extent processes can be changed or whether a business unit may use a local adaptation.

V. ANALYSIS

The defined research question for this document is: How and why can the slogan "Defend the template" support the use of process templates? From both the literature review and the expert interviews, a generalised answer to this research question can be proposed as follows:

Both research methods have shown that there is a risk of a lack of standardisation and harmonisation of process templates leading to too much diversification. A company must be careful to avoid this. Applying the slogan "Defend the template" can encourage companies to realise that a new deviation in the standard process template can lead to difficulties with future system changes.

Companies need to be aware that harmonising and optimising global business processes in an SAP landscape is an extensive project that can last for years. If a template has more variations, it will be more difficult and time-consuming to maintain this template, and to manage every future changing process within it. Companies should search for the lowest common denominator and defend the existing template as far as they can.

If there is a great need within a company to capture and document processes, then the company should set up business process management. Within the change process, the company can visualise, sort and structure its BPM. During this reorganisation, it can often be identified that processes in different business units are similar. A process template must be built from these documented and structured processes, including all recognised process flows, and, if appropriate, excluding certain process features and declaring them invalid. In the practical implementation and introduction of the newly developed processes, the process team then has the task of establishing and defending these process templates and processes, even if they encounter resistance. Just because a process has been carried out in a company for years in a particular way does not mean that this approach was the optimal solution for the business. However, when such a traditional and established process is changed, it is important to clarify why the new solution is a more effective approach, and to present it benefits to all parties involved. Especially in such a case, the new process template must be defended, even if there may be strong resistance to it.

The slogan "Defend the template" is only one essential step in a procedure for investigating process change requests. The difficulty is estimating the extent to which a template should be defended, as there can be legitimate doubts that a particular process template is the optimal solution for a company. In this case, a process owner or a process organisation must decide either to change the process, or to apply a local adaptation. Decisions in such exceptional cases should not be made by a business unit itself, but by a superior process organisation, which is responsible for the process and thus has to approve a change or local adaptation.

After analysing the literature and the expert interviews, this research developed a procedure for changing a process template. The goal was to develop an easy-to-handle business process model to defend possible process changes. Such a procedure could consist of the steps shown in Figure 2. This figure contains two important gateways for a process change. First, the process organisation must be informed about a desired process change. If the process organisation already knows about the desired process change and has opted for another solution, then this alternative solution must be defended. The same applies to the second gateway. If the central process organisation does not accept a new change request, this decision must be defended by the process organisation against any resistance that may arise.

 ISBN: 978-1-61208-690-3

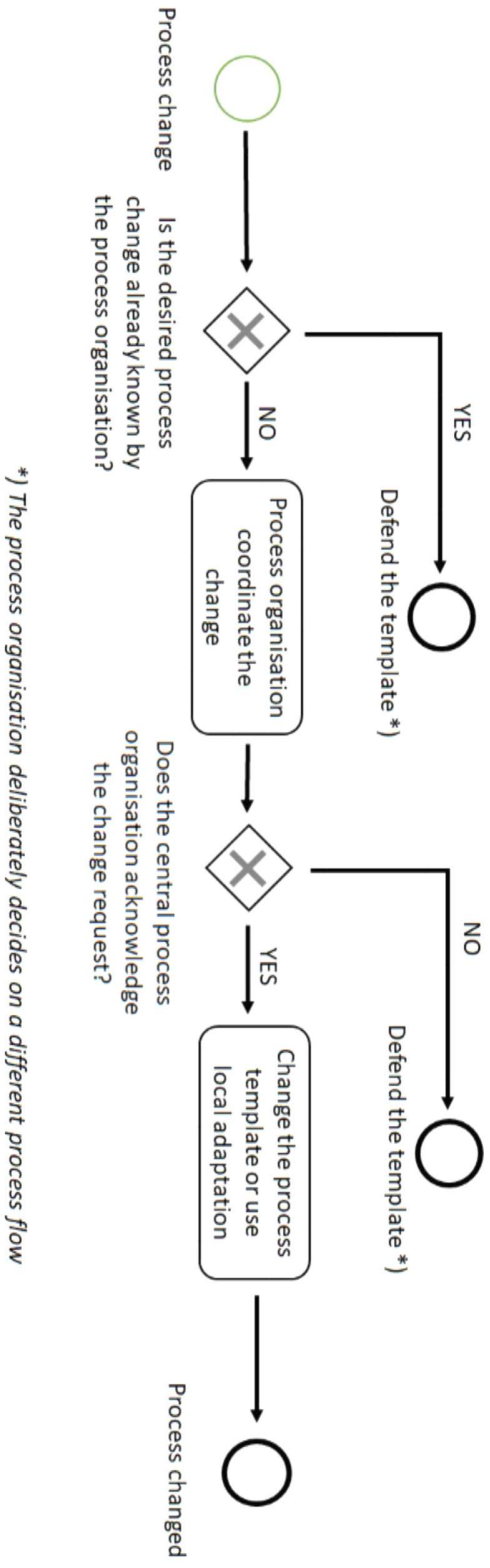


Figure 2. Business process model for process change

The process steps described in Figure 2 demonstrate that a template can be defended in several different locations within a process change. The advantage of this rigorous approach is that an escalation in local adaptions and shadow processes can be avoided.

VI. Conclusion and Future Work

The results of the literature review and the expert interviews suggest that a suitable compromise must be found between standardisation, (i.e., the usage of templates) and flexibility to adapt to local requirements (by coordinating with the organisation). The practical experience of the experts interviewed in this research demonstrates that the slogan "Defend the template" can also be useful. The recommendations of the literature and the respondents should be followed, and the processes of all units within a business should be considered before a process template is created. Despite this preliminary work, a process template may have to be adapted later, but this adaptation must be coordinated by the process organisation and must not be decided locally by a single business unit. As long as an adaptation is not supported by the process organisation, the slogan "Defend the template" applies, as exemplified in Figure *2*.

This investigation has only interviewed a small numbers of experts. It is quite possible that the inclusion of other experts or diversification into different industries, organisations, or SAP modules would have led to different or additional results. Future research should closely examine the argument of this paper and also consider diversification across different areas and industries. In addition, further research with more participants could significantly increase responses and lead to more diverse outcomes and analysis.

Future work could confirm the results of this research paper by conducting an online survey to check whether the illustrated process steps are acceptable to the participants in the business practice. An online questionnaire would, allow the collection of a larger amount of data from more participants in a shorter time and a more flexible manner than personal face to face interviews [15]. Participants should include a variety of people, such as users, process managers, researchers, and consultants for SAP and/or process management.

References

[1] E. Foth, "Excellent business processes with SAP: Practice of use in business groups", (Exzellente Geschäftsprozesse mit SAP: Praxis des Einsatzes in Unternehmensgruppen), Springer Verlag, Berlin, Germany, 2010.

[2] H. Romero, R. Dijkman, P. Grefen, and A. van Weele, "A literature review in process harmonization: a conceptual framework", Beta Working Paper series 379, Eindhoven, 2012.

[3] T. Kobayashi, M. Tamakia, and N. Komoda, "Business process integration as a solution to the implementation of supply chain management systems", Information & Management, vol. 40, pp. 769-780, 2003.

[4] J. V. Gavidia, "Impact of parent-subsidiary conflict on ERP implementation", Journal of Enterprise Information Management, vol. 29 no. 1, pp. 97-117, 2016

 ISBN: 978-1-61208-690-3

[5] E-3 Magazin, "Template-based SAP rollouts", Vol. September, Available from: https://www.cbs-consulting.com/files/pdf/information/Templatebasierte-SAP-Rollouts_E3_092011.pdf, September 2011, [retrieved: Jan. 2019].

[6] T. Kobayashi, S. Onoda, and N. Komoda, "Workflow Business Template for Application Processes in Administration Department", Information Technology and Management, vol. 3, no. 43, doi: 10.1023/A:1013160725421, 2002, [Retrieved: Jan. 2019].

[7] SAP Blogs, "Overview of Global SAP Template Rollout Programs", Available from: https://blogs.sap.com/2011/12/11/overview-of-global-sap-template-rollout-programs/, 2011, [Retrieved: Jan. 2019].

[8] A. Richen, and A. Steinhorst, "Standardization or Harmonization? You need Both", available from: https://www.bptrends.com/publicationfiles/11-05-ART-StandardizationorHarmonizationv-RickenSteinhorst.pdf, 2005. [Retrieved: Jan. 2019].

[9] A. B. Ryan, "Post-Positivist Approaches to Research Researching and Writing your thesis: a guide for postgraduate students"; MACE, Maynooth Adult and Community Education, pp. 12-26, 2006.

[10] E. G. Guba, "The Paradigm Dialog"; Newbury Park, CA, Sage Publications, 1990.

[11] Z. O'Leary, "The Social Science Jargon Buster", London, UK, SAGE Publications Ltd, 2007.

[12] G. Thomas, "How to Do Your Case Study: A Guide for Students and Researchers", London, UK SAGE Publications, 2011.

[13] J. Collis and R. Hussey, "Business research: A practical guide for undergraduate & postgraduate students"; Basingstoke, UK, Palgrave Macmillan, 2009.

[14] R. K. Yin, "Case Study Research: Design and Methods", London, UK, SAGE Publications, 2009.

[15] M. Saunders, P. Lewis, and A. Thornhill, "Research methods for business students", New York NY, Prentice Hall, 2009.

[16] R. Kumar, "Research Methodology: A Step-by-Step Guide for Beginners", London, UK, SAGE Publications, 2011.

[17] A. Bryman, and E. Bell, "Business research methods", Oxford, UK, Oxford University Press, 2007.

 ISBN: 978-1-61208-690-3

A Novel Ontology-Based Smart Service Architecture for Data-Driven Model Development

Gernot Steindl, Bernhard Heinzl, Wolfgang Kastner

Institute of Computer Engineering
Research Unit Automation Systems
TU Wien, Vienna, Austria
Email: gernot.steindl@tuwien.ac.at

***Abstract*—In the proposed concept, an ontology-based information model is designed to support the data-driven model development process of industrial energy systems. Smart Services for data-driven model development are conceptually deployed in the context of the Reference Architecture Model for Industry 4.0 (RAMI4.0) residing inside the information layer. The ontology-based Smart Service architecture provides interoperability by means of data integration and data exchange between for various applications, part of the functional layer. Due to its modular structure, these Smart Services and their functional enhancements can easily be extended and used by a variety of other applications at different layers of the RAMI4.0.**

Keywords–data-driven model development; ontology; Smart Service; industrial energy systems; Industry 4.0; RAMI4.0

I. INTRODUCTION

Currently, a number of research projects and development efforts address the further digitalization in the industrial sector [1], which is often called the fourth industrial revolution or Industry 4.0. Besides digitalization, the main goals of Industry 4.0 are optimization and customization of production as well as enhanced automation and adaption [2]. In this context, Industry 4.0 can also facilitate the sustainable energy transition [3]. Technology of Industry 4.0 is a forerunner to let the plant's energy demand become more controllable and providing means for flexibility, which helps to integrate a higher share of volatile renewable energy sources into the energy system. To utilize this flexibility of industry processes, model-based and predictive control strategies are suitable to handle the conflicts between different objectives and constraints [4]. This allows a factory to react, for example, to a demand response request, while meeting the production constraints. Accurate process models are obligatory for this kind of applications. Developing such process models, however, is the most challenging and time-consuming task in an implementation process [4] [5]. One reason for this is the cumbersome manual analysis of data from heterogeneous sources, which often have insufficient data quality, are only partly available [6] or stored in different formats, encodings and databases. Data-driven model development also relies on implicit domain knowledge [7], which is often scattered and not clearly communicated by the domain experts or hard-coded into various applications. In order to reduce the model development effort and facilitate knowledge reuse, systematic data and information management is a vital part in oder to increase data quality and consistency, as well as make it available over the whole plant life-cycle.

The goal of this work is to develop a framework for systematic support of data-driven model development in the context of industrial energy systems. Comprehensive information management is achieved explicitly by stating domain knowledge and making it available for different services in a formal ontology-based information model. This ontology will act as a semantic abstraction layer, to facilitate integration of different kinds of data.

In Section 2, the data-driven model development process will be discussed. It will be stated how information, which is stored in a formal ontology, could support the individual process steps. A service architecture will be introduced in Section 3, that facilitates the reuse of stored knowledge in the context of Industry 4.0.

II. DATA-DRIVEN MODEL DEVELOPMENT PROCESS

Data-driven models are usually identified during the operational phase of the plant using process data and are developed for a certain purpose or application, like plant control, monitoring, optimization, prediction or performance assessments. The main steps in the data-driven model development process for industrial processes are described in [8].

The first step is to decompose the process and its data based on process knowledge, like location of the components and the plant topology. If this knowledge is available in a machine-readable form or formal ontology, this step can be automated.

Afterwards, pre-processing of the data has to be performed, like removing outliers, normalizing the data, filling or estimating missing data and synchronizing the sampling rates of different process variables. Data analysis methods can be used in combination with domain knowledge from the proposed ontology to analyze and semantically enrich the time series data, e.g. to add reliability metrics to the data points. Also, inconsistencies like falsely labeled data or outliers can be identified automatically, if semantic information of a certain data point is available.

The subsequent step of sample and variable selection is very crucial as different operating conditions can occur in the monitoring data. The samples should be chosen with regard to the scope of application, in order to be able to train an accurate model. Therefore, domain knowledge has to be applied which should also be made explicitly available in the ontology to support this step.

 ISBN: 978-1-61208-690-3

The final model selection and identification can be performed in an iterative way by the control engineer. Design decisions, used training and evaluation data, as well as model performance are also important information, which should be stored in an ontology, to enable comparison and reuse of different component models.

III. Ontology-based Smart Service Architecture

For Industry 4.0, an architectural reference model (RAMI4.0) is available, which combines all elements and IT components in a layered model, taking into account life-cycle aspects [9]. In the context of RAMI4.0, a service-oriented architecture for "Smart Services" was proposed in [10]. We build on this idea and apply it to support the data-driven model development process by offering services, which can be located on different layers and hierarchies in RAMI4.0, in particular on the Information and Functional Layer.

Figure 1 shows the architecture of the proposed services at different layers of RAMI4.0. The data is available in various databases an formats. An ontology-based information model is used as a foundation for interoperability and data integration. Engineering data as well as operational data are integrated with the help of this ontology. The information model is one of the main components, which consists of smaller, interlinked ontologies. These are specialized for certain services or describe general concepts, like plant topology, plant components, etc. Existing information models, like the P&ID specification of the DEXPI initiative [11] should be reused and adapted. These ontologies manage data access and data exchange between different services. Process data, which typically occurs as time series, will stay in their databases, as these are optimized for this kind of purpose, but are accessible through the ontology to facilitate data integration. Thus, the ontologies can be connected to a "Linked Factory Data" repository and new information can easily be added into the modular ontology. The available services and their capabilities can also be stated in the ontology itself, to enhance interoperability.

The base service for data-driven model development is the "Smart Data" service, which is able to automatically analyze, pre-process and validate the available data, utilizing information from the ontology. For data-driven model development, a "Model Identification" service will be established, which uses domain knowledge from the ontology to find related input data to a certain model component and to support the model developer with information of the plant's operating conditions and perform proper sample selection. Information on the identified model will be stored in the ontology as well, to enable model comparison and reuse. The "Linked Data Management" service is the central point for entering and maintaining the ontology-based information model itself.

Based on the proposed architecture, additional services and applications on the different layers of RAMI4.0, like "Optimized Control" or "Fault Detection", can be implemented, which are also able to use and contribute to the information stored in the "Linked Factory Data".

IV. Conclusion

The goal of this work is to develop a systematic support for a data-driven model development process in the context of industrial energy systems. The proposed ontology-based architecture can help to reduce the manual and error-prone

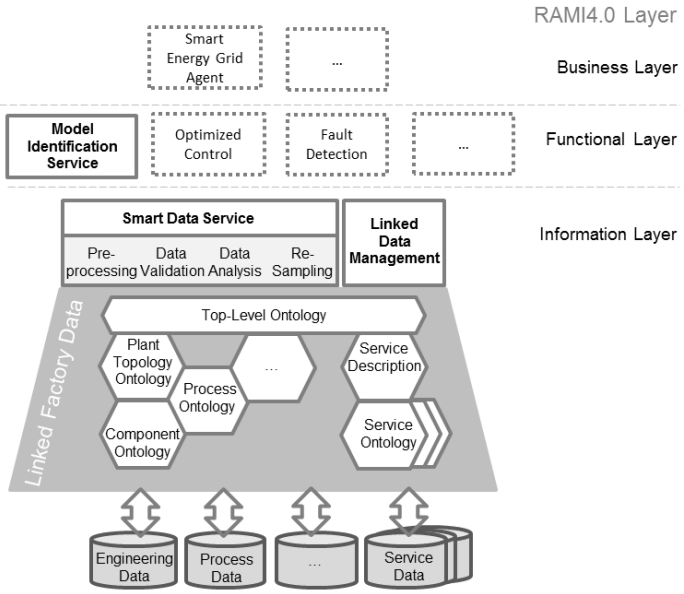


Figure 1. Ontology-based Smart Service Architecture

modeling effort and increase the quality, consistency and expressiveness of the available data. The next steps will be to specify the information which has to be included into the ontology and evaluate different methods for ontology-based data access. Afterwards, a real use case of an industrial production plant will be used to implement a "Smart Data" and "Model Identification" service as proof-of-concept.

References

[1] Y. Lu, "Industry 4.0: A survey on technologies, applications and open research issues," Journal of Industrial Information Integration, vol. 6, 2017, pp. 1–10.

[2] V. Roblek, M. Meško, and A. Krapež, "A Complex View of Industry 4.0," SAGE Open, vol. 6, apr 2016.

[3] UNIDO, "Accelerating clean energy through Industry 4.0: Manufacturing the next Revolution," United Nations Industrial Development Organization, Vienna, Austria, Tech. Rep., 2017.

[4] M. Killian and M. Kozek, "Ten questions concerning model predictive control for energy efficient buildings," Building and Environment, vol. 105, 2016, pp. 403–412.

[5] M. Morari and J. H. Lee, "Model predictive control: past, present and future," Computers & Chemical Engineering, vol. 23, may 1999, pp. 667–682.

[6] Y. Zong, S. You, J. Hu, X. Han, C. Jiang, Y. Zhang, and G. M. Böning, "Challenges of using model predictive control for active demand side management," 4th International Conference on Microgeneration and Related Technologies, 2015.

[7] G. A. Silver, O. Al-Haj Hassan, and J. A. Miller, "From domain ontologies to modeling ontologies to executable simulation models," Proceedings - Winter Simulation Conference, 2007, pp. 1108–1117.

[8] Z. Ge, "Review on data-driven modeling and monitoring for plant-wide industrial processes," Chemometrics and Intelligent Laboratory Systems, vol. 171, 2017, pp. 16–25.

[9] P. Adolphs, H. Bedenbender, D. Dirzus, and E. Martin, "Reference Architecture Model Industrie 4.0 (RAMI4.0)," VDI/VDE-Gesellschaft, Tech. Rep. July, 2015.

[10] G. Koschnick, "Industrie 4.0: Smart services," German Electrical and Electronic Manufacturers' Association, Frankfurt am Main, Germany, Tech. Rep. July, 2016.

[11] M. Theißen and M. Wiedau, "DEXPI Data Exchange in Process Industry P&D Specification," DECHEMA, Tech. Rep., 2016.

 ISBN: 978-1-61208-690-3

Information Model Integration for Service-oriented Manufacturing Operation Management Systems

Hartmut Schweizer
Institute for Applied Computer Science
TU Dresden
01062 Dresden
email: hartmut.schweizer@tu-dresden.de

Martin Wollschlaeger
Institute for Applied Computer Science
TU Dresden
01062 Dresden
email: martin.wollschlaeger@tu-dresden.de

***Abstract*—This paper presents a concept for the combined application of information models in automation as a basis for a service-oriented, component-based architecture of Manufacturing Operation Management Systems (MOM). It will be shown how the information models of the Automation Markup Language (AutomationML) and the OPC Unified Architecture (OPC UA) can be integrated into the Business To Manufacturing Markup Language (B2MML). This allows achieving an improved transparency and availability of process data and knowledge data within automated manufacturing processes. Here, the extension functionalities and the transaction specifications of B2MML are used. Finally, an application example is shown and is discussed.**

Keywords – Automation; Model integration; B2MML; AutomationML; OPC UA.

I. INTRODUCTION

In automated production, the organizational architecture, driven by increasingly powerful hardware, higher networking, and developments, such as the Industrial Internet of Things (IIoT) or Cyber-Physical Production Systems (CPPS), is changing from the classical pyramid structure to component-based architectures (Fig. 1). This applies both to the hardware, i.e., Personal Computers (PCs), Programmable Logic Controllers (PLCs), and field devices, and to the software used for tasks in the different company levels, e.g., Manufacturing execution System (MES) or Enterprise Resource Planning (ERP) [1].

This also results in new requirements for the form and availability of data and information within these distributed structures, e.g., addressing decentralization, interoperability and virtualization. The same applies to an information model, which needs to meet these requirements [2]. In the following, a concept for such a heterogeneous information system, consisting of OPC UA, B2MML, and AutomationML, is presented for a component-based organizational structure with a service-oriented architecture. The main advantage of this approach is given by the integration of existing information models into a harmonized information system.

II. INFORMATION MODELS IN AUTOMATION

In the automated manufacturing, numerous different information models are used. Due to their different focus and degree of detail as well as other characteristics, such as service suitability and legibility, however, they are not all equally suitable for all aspects at every (company) level [11]. Among other things OPC UA [3], AutomationML [3], and the IEC 62264 implementation B2MML are three models that cover the required information spectrum for a MOM [11]. These are briefly presented below. Especially the elements of B2MML, which are used in the application example in section V, will be explained detailed.

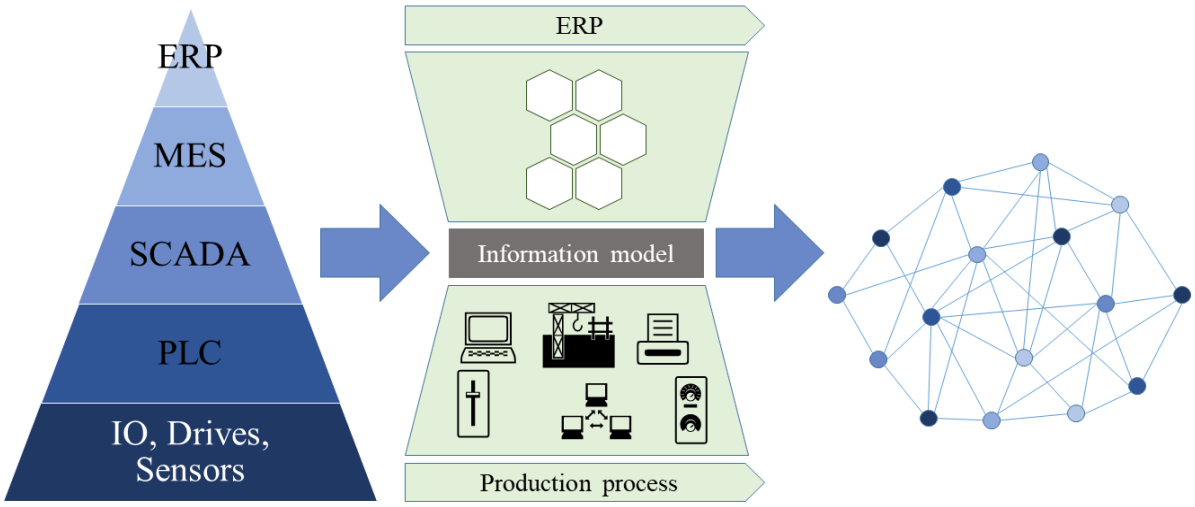


Figure 1. Development of the automation pyramid

 ISBN: 978-1-61208-690-3

A. OPC UA

The standard OPC UA (IEC 62541) defines an object-oriented, flexible industrial information model with a defined communication protocol. The OPC UA architecture follows a Service-Oriented Approach (SOA). Its structure consists of several layers. All basic services (base services) defined by OPC UA are abstract method descriptions. Thus, they are protocol-independent and provide the basis for the entire OPC UA functionality. The communication layer provides access to these methods using a protocol that serializes and deserializes the data and sends it over the network (in binary or XML format) [4].

The OPC information model is a so-called full mesh network of nodes. A node resembles an object from object-oriented programming and can have attributes. It is also possible to define and call methods [4]. Furthermore, events are supported that can be sent to exchange certain information between devices. Further information models, e.g., for specific application domains, can be based on this base layer [4].

B. CAEX and AutomationML

AutomationML is a neutral, XML-based data format for the storage and exchange of plant design and engineering data. It is based on the Computer Aided Engineering Exchange (CAEX) data format defined in IEC 624242. The primary goal of AutomationML is the exchange of engineering data in a heterogeneous tool landscape of modern engineering tools, covering various disciplines such as mechanical design, electrical design, HMI development, PLC programming or robot control. AutomationML is standardized in IEC 62714. It supports the integration of other file formats. In particular, the integration of PLCopen XML [5] is provided for embedding PLC programs and logic, while COLLADA format [6] is intended for geometry information [7].

C. IEC 62264 and B2MML

The international standard IEC 62264 defines models, activities, and data exchange for Manufacturing Operations Management systems. B2MML [8] is an XML implementation of the IEC 62264 specifications in the form of XML schemata, which is designed as a link between ERP and materials management (supply chain) and to the production level [9]. Part 2 of IEC 62264 defines objects and attributes for the integration of enterprise management and control systems. This includes the resource types *Personnel*, *Equipment*, *Material*, and *Physical Asset* as well as the production process type *ProcessSegment*. In the following, the equipment model (Fig. 2) and the process segment model for the resources are explained in more detail.

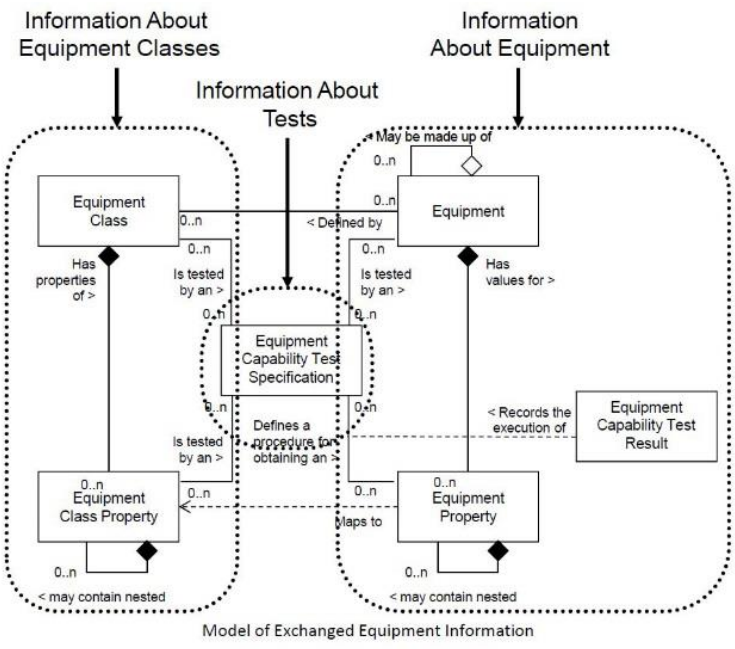
1) The role-based equipment model of IEC 62264

Each *Equipment* resource can be composed of other *Equipment* resources and aggregates *Equipment Properties*. It can also be defined using one or more *Equipment Classes*, which can have any number of their own class properties. Test specifications can also be defined for *Equipment* and *Equipment Classes* and their properties (Fig. 2).

Equivalent models are defined for the other resource classes *Personnel*, *PhysicalAsset*, and *Material* [10].

2) The process model of IEC 62264

Process segments in IEC 62224 are defined as the smallest elements of manufacturing activities that are visible to business processes. The process model is hierarchically




```
<xsd:complexType name = "EquipmentType">
  <xsd:sequence>
    <xsd:element name = "ID"
                 type = "IdentifierType"/>
    <xsd:element name = "Description"
                 type = "DescriptionType"
                 minOccurs = "0" maxOccurs = "unbounded"/>
    <xsd:element name = "Location"
                 type="LocationType"
                 minOccurs="0"/>
    <xsd:element name = "HierarchyScope"
                 type = "HierarchyScopeType"
                 minOccurs = "0" />
    <xsd:element name = "EquipmentLevel"
                 type = "HierarchyScopeType"
                 minOccurs = "0" />
    <xsd:element name = "EquipmentAssetMapping"
                 type = "EquipmentAssetMappingType"
                 minOccurs = "0" maxOccurs = "unbounded"/>
    <xsd:element name = "EquipmentProperty"
                 type = "EquipmentPropertyType"
                 minOccurs = "0" maxOccurs = "unbounded"/>
    <xsd:element name = "Equipment"
                 type = "EquipmentType"
                 minOccurs = "0" maxOccurs = "unbounded"/>
    <xsd:element name = "EquipmentClassID"
                 type = "EquipmentClassIDType"
                 minOccurs = "0" maxOccurs = "unbounded"/>
    <xsd:element name = "EquipmentCapabilityTestSpecificationID"
                 type = "EquipmentCapabilityTestSpecificationIDType"
                 minOccurs = "0" maxOccurs = "unbounded"/>
    <xsd:group   ref  = "Extended:Equipment"
                 minOccurs = "0" maxOccurs = "1"/>
  </xsd:sequence>
</xsd:complexType>
```

Figure 2. Equipment model IEC 62264 [10] and B2MML – Implementation of EquipmentType [8]

 ISBN: 978-1-61208-690-3

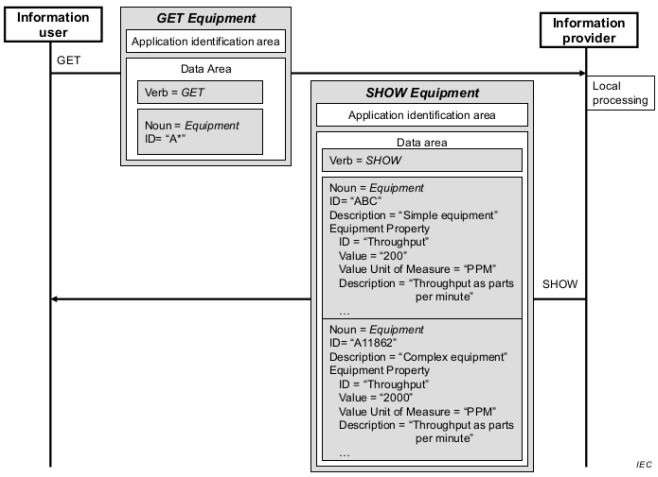


Figure 3. Transaction in IEC 62264 [15]

structured, i.e., each process segment can contain further subordinate process segments. The interdependencies between the process segments are realized by references in the form of *ProcessSegmentDependencies.* All resources can be aggregated from process segments as resource specifications via references. The processes in turn can correspond with *operation* or *product definition segments* [10].

3) The transaction model of IEC 62264

Part 5 of IEC 62264 defines transactions for the exchange of information between applications of business processes and applications of manufacturing processes. The extensive transaction specifications are one of the main reasons for using B2MML as a data model in this concept. The *GET/SHOW* transaction mechanism is briefly explained above (Fig. 3). The transaction model for GET and SHOW is the *PULL transaction model*. It is used when a user submits a data request to an information provider. A GET request can be issued for a specific object via its ID, or by means of a so-called "wildcard" (*) for several objects (Fig. 3) [15].

III. HETEROGENEOUS INFORMATION MODEL FOR MOM AS A DISTRIBUTED SERVICE-ORIENTED SYSTEM

As already explained above, there are mechanisms in both OPC UA and AutomationML to integrate and use other information models. In OPC UA and AutomationML, Companion specifications exist, for example for ISA-95 (B2MML) [12] [13]. In addition to COLLADA and PLCopen XML, other information models – such as B2MML – can also be integrated and referenced in AutomationML [14]. From the point of view of the enterprise level, these two variants are bottom-up approaches. Thus, information from or for the enterprise management level (ERP) is represented in information models of the plant and process control level (OPC UA, AutomationML). If this information was not already created in the modelling using companion specifications, it must be integrated into the respective model by transformation. This is

- costly
- error-prone
- complex and difficult to read

To solve this problem, this paper presents an approach in which the information models of OPC UA and AutomationML are integrated into B2MML (Fig. 4 and Fig. 5). This allows using information from all three models in combination without having to modify the information

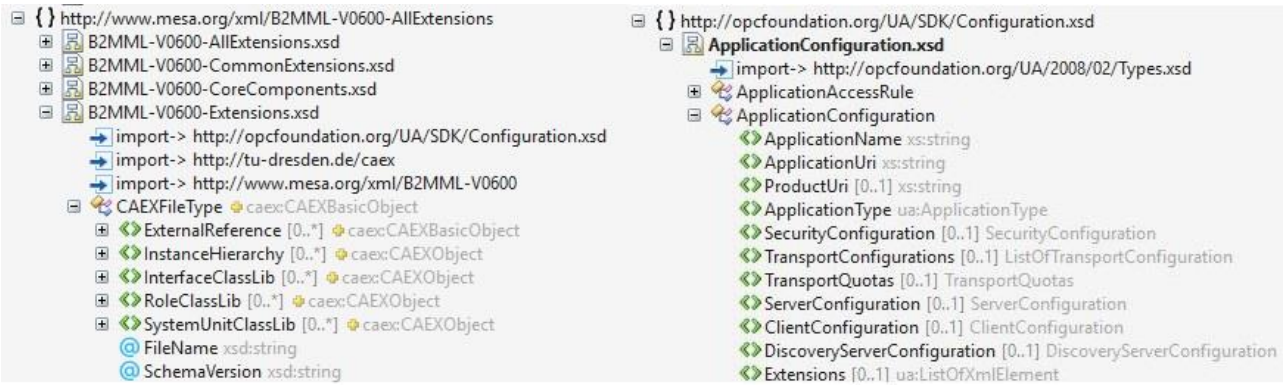


Figure 4. Integration of OPC UA and CAEX schema elements in the B2MML model schemata

```
<xsd:group name = "Equipment">
        <xsd:sequence>
                <!-- add extended elements here -->
                <xsd:element name="CAEXFile" type="CAEXFileType" minOccurs="0" maxOccurs="unbounded" />
                <xsd:element name="OPCUAServerConfig" type="opcua:ApplicationConfiguration" minOccurs="0" maxOccurs="1" />
                <xsd:element name="OPCUAClientConfig" type="opcua:ApplicationConfiguration" minOccurs="0" maxOccurs="1" />
        </xsd:sequence>
</xsd:group>
```

Figure 5. Enhancement of the Equipment Schema of B2MML [8] with OPC UA and AutomationML elements

 ISBN: 978-1-61208-690-3

structure of the individual information blocks, or to adapt them. This paper does not deal with the interoperability and transformation of the treated information models; these mechanisms remain unaffected by this concept. Nevertheless, their necessity can be reduced, because in the presented concept every stakeholder normally gets the needed data in the suitable format (Fig. 6).

IV. INTEGRATION OF INFORMATION MODELS FROM AUTOMATIONML AND OPC UA INTO B2MML

B2MML provides so-called extension points for the extension of all main classes such as process segments or resources (*Equipment*, *PhysicalAsset*, *Personnel*, and *Material*). These are used to include additional information structures in the model or in the corresponding submodels.

The following demonstrates how this mechanism is used to extend B2MML with the information models of OPC UA and AutomationML. The first step is to import the basic information models, i.e., the XML schemata of AutomationML (CAEX) and OPC UA, into the B2MML schema (Fig. 4). Then the partial models, into which corresponding data are to be inserted, can be extended accordingly. Figure 5 shows the extension of the B2MML equipment model by an AutomationML information and an OPC UA configuration.

It should be noted that, in the case of AutomationML, external references can be used to link the contents of other (XML) files. If this option is used, the schemas of these models and the referenced files must also be included. In the case of the PLCopen XML and COLLADA, these are the corresponding schemas found in [5] and [6].

There are two requirements for enabling communication for this extended B2MML schema: Both communication partners must have implemented (i) the transaction functionalities defined in [15] and (ii) the B2MML XML-schema extensions as described above. In case of AutomationML the *ChangeMode* attribute of AutomationML [16] and the transaction concept of B2MML have to be combined and synchronized. Table I shows the resulting affiliation.

For each transaction of a B2MML element including CAEX code, the appropriate AutomationML *ChangeMode* parameter has to be set. The following example illustrates how this extension can be used in practice, and which advantages result from it.

V. APPLICATION EXAMPLE: INFORMATION ON PLANNING, SIMULATING, AND EXECUTING A PRODUCTION PROCESS RESTRUCTURING WITH EQUIPMENT REPLACEMENT

The described application case deals with a process restructuring in a process chain. The process shall be changed from injection molding to 3D printing within an automated production process for the manufacturing of an element made from plastic. During the restructuring, as little interference as possible shall be made in the production flow of the other processes running. For this purpose, following questions must be answered and work steps carried out.

- What is the actual status of the plant (Monitoring: OPC UA)?
- Which operations and processes as well as interrelated processes are affected (ERP: B2MML)?
- Which resources are affected, deleted, or required (ERP: B2MML)?
- Estimation of the effort required for the necessary modifications (plant engineering data: AutomationML)
- Execution (plant data: AutomationML, personnel planning: B2MML)
- Commissioning, monitoring, quality control (Monitoring: OPC UA, Quality criteria: B2MML)

It can be recognized that production process restructuring affects all levels of automation. By means of the extensions of B2MML by OPC UA and AutomationML elements, the required data can be communicated directly in the required form. This can be performed between the necessary functional components at the various levels using a service-based approach on top of the transaction mechanisms of B2MML (Fig. 3) [15].

Figure 6 shows a simplified representation of the production process restructuring as described above. In particular, the organizational sequence of the action, in combination with the data required in the respective work steps is explained. On a customer request ("Customer Order"), a feasibility check is carried out with regard to the product and the production line. This leads to a review of the processes involved, which include references to the required equipment.

Figure 6, Listing 1 shows this *ShowProcessSegment* information. The equipment can now be called up via the references in the *EquipmentSegmentSpecifications*.

Figure 6, Listing 2 shows the content of the *ShowEquipment* transaction response for injection molding machine with its plant planning data in AutomationML, the OPC UA configuration data, and its test specification properties. A GET call [15] to the *EquipmentcapabilityTestSpecificationID*s allows the corresponding properties of the injection molding machine to be queried.

Figure 6, Listing 3 shows the content of the *ShowEquipment* transaction response for the 3D printer with its plant design data in AutomationML, the OPC UA configuration data, and its test specification properties. By a GET call [15] of the *EquipmentcapabilityTestSpecificationID*s, the corresponding properties of the 3D printer can be queried and compared with those of the injection molding machine.

TABLE I. AFFILIATION OF B2MML AND AUTOMATIONML COMMUNICATION CONCEPTS

B2MML	**AutomationML**	
Sync (Change)	Change	PUBLISH
Sync (Delete)	Delete	PUBLISH
Sync (Add)	Create	PUBLISH
Get	State	PULL
Show	State	PULL
Confirm, Cancel, Acknowledge, Process, Respond, Change		PUSH

 ISBN: 978-1-61208-690-3

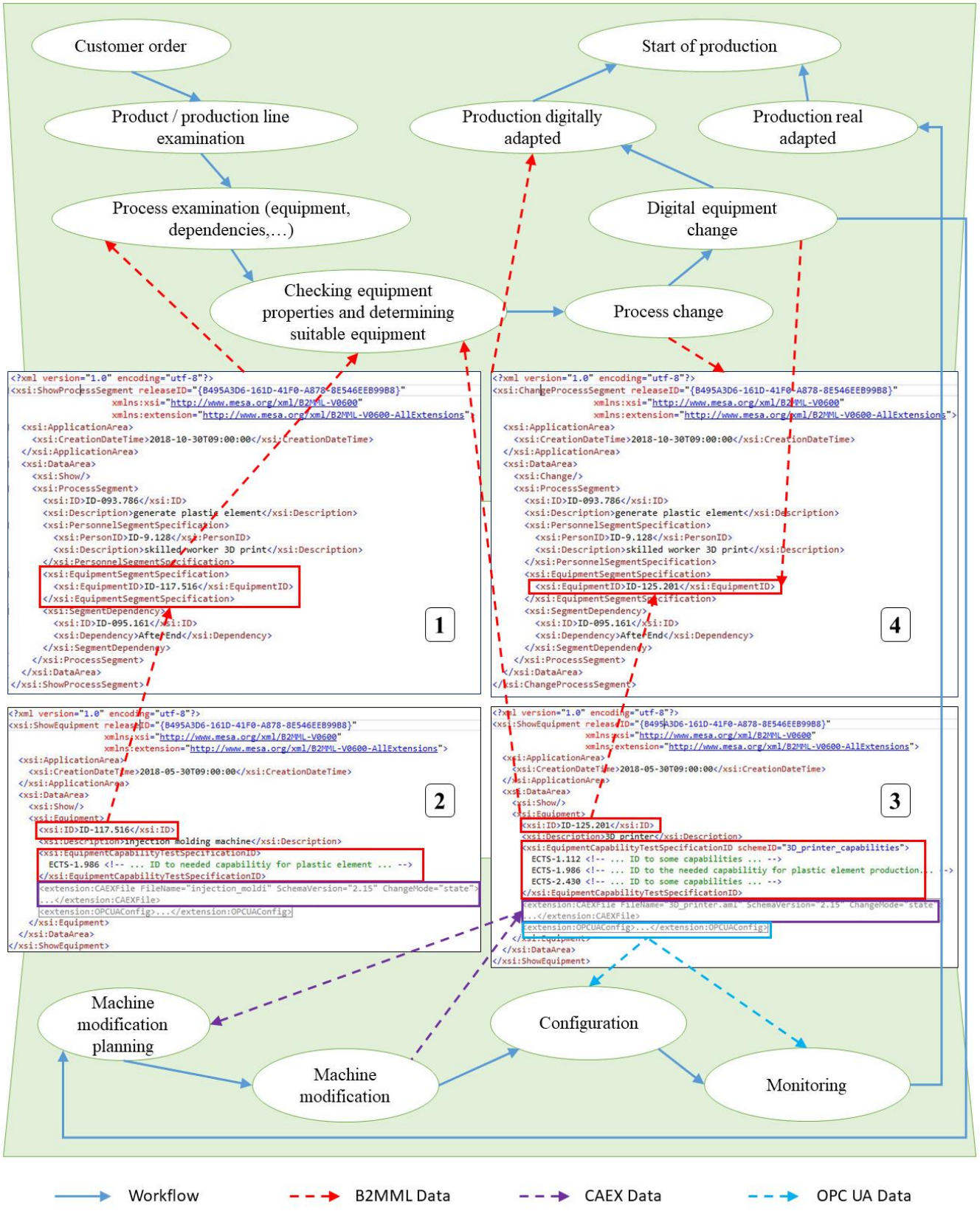


Figure 6. Application Example

Since the printer meets the required property, the process can be restructured. With this information, the process can now be updated with new *EquipmentSegmentSpecification* using the *ChangeProcessSegment* transaction [15] by entering the ID of the 3D printer for the required equipment. This means that the new equipment including communication functionalities (OPC UA configurations and hardware plans (AutomationML)) are technically updated for the restructured process (Fig. 6, Listing 4).

Furthermore, the required connection and installation information (AutomationML and OPC UA) are available simultaneously with the *EquipmentSegmentSpecifications* of the 3D printer (Fig. 6, Listing 3). In addition, the corresponding personnel with required capabilities was assigned to the process segment in the *PersonnelSegmentSpecification* at the same time.

When using the OPC UA-Companion specifications for ISA-95, the corresponding elements would have to be transformed to B2MML in order to make them readable for a B2MML-based ERP system. Equivalent to this, the B2MML information would have to be transformed into AutomationML for an AutomationML-based system or into B2MML for a B2MML-based system. Manual activities would be necessary to complete the allocation of the resources [14], if the Companion specifications were not used.

VI. DISCUSSION

As explained in the previous sections, AutomationML, OPC UA, and B2MML provide mechanisms for extending or integrating and using other information models. For the concept presented, the variant of using the extension of B2MML with AutomationML and OPC UA via the appropriate extension points was chosen. This has the advantage that, on the one hand (usually), no data transformation is necessary. All data is contained in the data package in the required information model. On the other hand, the transaction specifications of B2MML for AutomationML information [1] and OPC UA information can also be used for the enterprise level. Finally, the communication specifications of OPC UA for the MES and production level are also available. The Companion specifications of OPC UA and AutomationML remain unaffected and can be used without restrictions [17].

VII. CONCLUSION

It was shown how the information models OPC UA and AutomationML can be integrated by means of the extension mechanisms of B2MML and how they can be used together as data formats in a service-oriented MOM implementation. The transaction mechanisms of IEC 62264 are used here. At the same time, the scope of information provided by OPC UA and AutomationML is available without restriction. A limitation in the application of the concept results from the fact that B2MML transactions only allow queries via IDs. For an application between companies or independent company structures, these would have to be extended by mechanisms, which access other parameters by semantic means, e.g., the description elements (Fig. 2) [8].

REFERENCES

[1] D. Brandl, New integration architectures for federated systems, ISA Publications, InTech Magazine, April 2016, https://www.isa.org/intech/20160403/, retrieved: November 2018.

[2] B. Vogel-Heuser, Guest Editorial Industry 4.0 – Prerequisites and Visions, IEEE Transactions on Automation Science and Engineering, Vol. 13, No. 2, April 2016, pp. 411-413.

[3] M. Schleipen, Syed-Shiraz Gilani, Tino Bischoff and Julius Pfrommer, OPC UA & Industrie 4.0 - enabling technology with high diversity and variability, 49th CIRP Conference on Manufacturing Systems (CIRP-CMS 2016), December 2016, pp. 315-320.

[4] OPC Foundation: OPC-UA, https://opcfoundation.org/about/opc-technologies/opc-ua/, retrieved: November 2018.

[5] PLCopen, TC6–XML Schemes http://www.plcopen.org/pages/tc6_xml/, retrieved: November 2018.

[6] KHRONOS GROUP, COLLADA, https://www.khronos.org/collada/, retrieved: November 2018.

[7] IEC 62714 - Engineering data exchange format for use in industrial automation systems engineering - AutomationML, www.iec.ch, International Electrotechnical Commission, 2014.

[8] MESA International, Business To Manufacturing Markup Language (B2MML), http://www.mesa.org/en/B2MML.asp, retrieved: November 2018.

[9] IEC. IEC 62264-1:2013 Enterprise-control system integration – Part 1: Models and terminology. International Electrotechnical Commission, Geneva, 2013.

[10] International Electrotechnical Commission (IEC), Enterprise-control system integration-Part 2: Objects and attributes for enterprisecontrol system integration, International Standard, Rev. 2.0, IEC 62264-2:2013, June 2013.

[11] R. Silva Peres et al., “Selection of a data exchange format for industry 4.0 manufacturing systems”, IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, October 2016, pp. 5723–5728.

[12] OPC Foundation: ISA-95 for OPC UA. https://opcfoundation.org/markets-collaboration/isa-95/, retrieved: November 2018.

[13] B. Wally, Application Recommendation Provisioning for MES and ERP – Support for IEC 62264 and B2MML, November 2018.

[14] B. Wally, C. Huemer and A. Mazak, “Entwining Plant Engineering Data and ERP Information: Vertical Integration with AutomationML and ISA-95”, 2017 3rd International Conference on Control, Automation and Robotics, April 2017, pp. 356 - 364.

[15] IEC. IEC 62264-5:2016 Enterprise-control system integration – Part 5: Business to manufacturing transactions. International Electrotechnical Commission, Geneva, 2016.

[16] J. Prinz and L. Hundt, “AutomationML − Engineering Workflow”, SPS-Magazin, November 2013, pp. 60-62.

[17] B. Wally, C. Huemer and A. Mazak, “AutomationML, ISA-95 and Others: Rendezvous in the OPC UA Universe”, 14th International Conference on Automation Science and Engineering (CASE 2018) August 2018, pp. 1381-1387.

 ISBN: 978-1-61208-690-3