

Towards Building a Next-Generation Data Analytics Toolbox: Application of the Axiomatic Theories Fusion Methodology

Abou Eddahab-Burke, F.

Publication date

2024

Document Version

Final published version

Published in

2024 63rd Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)

Citation (APA)

Abou Eddahab-Burke, F. (2024). Towards Building a Next-Generation Data Analytics Toolbox: Application of the Axiomatic Theories Fusion Methodology. In *2024 63rd Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)* (pp. 432). SICE.

https://www.researchgate.net/publication/384256446_Towards_building_a_next-generation_data_analytics_toolbox_Application_of_the_axiomatic_theories_fusion_methodology

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Towards Building a Next-Generation Data Analytics Toolbox: Application of the Axiomatic Theories Fusion Methodology

Fatima-Zahra Abou Eddahab-Burke^{1†}

¹Department of Multi actor systems, Delft University of Technology, Delft, The Netherlands
(E-mail: f.aboueddahab-1@tudelft.nl)

Abstract: Reliable development of next-generation data analytics toolboxes (N-GDATs) requires robust underpinning theories, which cannot necessarily be inductively generated. Based on the axiomatic theories fusion (ATF) methodology we deductively developed a comprehensive theory supporting the development of a N-GDAT for white goods design based on middle-of-life data. Accordingly, theories about designer's needs, advanced technologies, data analytics, creative problem-solving, decision-making, and interoperability were fused following the ATF steps: (i) selection of component theories, (ii) axiomatic discretization of foundational theories, (iii) establishing relationships among axioms and postulates, (iv) transcription of system of axiomatic propositions into a textual format, and (v) validation of explanatory theory. The obtained new theory provides a robust basis for the targeted knowledge platform. It provides (i) decision-making, (ii) algorithmic concepts, (iii) learning, (iv) data management, (v) interfacing, (vi) reasoning, (vii) data types and characteristics, (viii) design issues, (ix) analytics techniques and methods, and (x) outputs requirements to develop N-GDATs.

Keywords: Data analytics tools, axiomatizations, theories fusion, middle-of-life data

1. INTRODUCTION

Data analytics tools have reached a level of sophistication that challenges product designers. Other challenges emerge from real-time generation of big data, multiplicity of data sources, as well as from the nature of products [1]. In the process of developing a next generation data analytics toolbox (N-GDAT), there is no specific theory in the literature that would be able to cover all important aspects. On the other hand, we make the hypothesis that *the conceptualization, implementation, and validation of N-GDATs may be efficiently supported by synthesized theories that include technological, human, social, computational, organizational, etc. chunks of knowledge*. In principle, the underpinning theory could be generated by exploiting the different epistemological relationships between various theories. Rigorously following these principles, a practical approach to theorizing in a deductive manner called axiomatic theories fusion (ATF) was introduced. The underpinning theory is based on the assumption that a number of properly selected theories can provide sufficient explanations for phenomena that have not been studied experimentally. The multiple theories must be considered simultaneously and must be interwoven in the reasoning. However, these theories are conceptually different and cannot be combined straightforwardly. They also may not be completely coherent and consistent. Therefore, a proper new theory should be synthesized by blending the existing in-part- sufficient theories (called source theories) into a new explanatory theory (called the target theory) with sufficient clarifying power and consistency. Following the ATF transforms a set of source theories into one target theory.

For our research we decide to use the principles of the ATF to develop a theory supporting the conceptualization and implementation of a N-GDAT for white goods design based on Middle-of-life-data (MoLD). The choice of the ATF is justified by the following hypothesis: *a robust and*

comprehensive conceptual basis for a knowledge platform in the context of multidisciplinary research needs to combine many composite theories. This was reflected in the conducted literature study. Integrating theories from different disciplines is becoming a need to develop new relevant theories [2]. The existing theories focused on one aspect separately without integrating all aspects in one theory. Grouping all aspects of research into a new theory was a good strategy for system development. As do other professionals, developers of white goods need data analytics tools that are tailored to their problems, needs, knowledge, and expertise. Theories that support the development of traditional user software and data analytics tools have proven to be insufficient in this context. Consequently, there was a need to provide proper theories that describe the new tools and explain what functions and computation are necessary. Theories of this kind are scarce, but the ATF is promising in deriving comprehensive supporting theories by semantic fusion of relevant component theories.

2. AXIOMATIC THEORIES FUSION

Axiomatic theorizing is a deductive approach that draws conclusions based on fundamental principles and accepted truths. Axiomatization specifies the content of a theory in which a set of axioms and postulates are given, and from which a set of propositions is derived [3]. The self-evident nature of axioms and postulates makes them trustable. The ATF forms a base for merging theories, since it (i) is logical, (ii) does not need to be proven, (iii) contains general statements as “axioms”, (iv) contains specific statements within specific contexts as “postulates”, and (v) contains propositions derived based on axioms and postulates. This approach is particularly appropriate in the case of design domains, where solving a problem, innovating, or improving requires different domains of expertise. The ATF methodology covers the lack encountered in some research domains (such as design theories, data analytics theories, and so on), where

theories are not up to date regarding processes and new techniques and technologies. Another advantage of this approach is that it allows the development of one's own sufficient theory dedicated to one's particular case and domain of interest. The procedural framework of the ATF methodology is shown in Fig. 1.

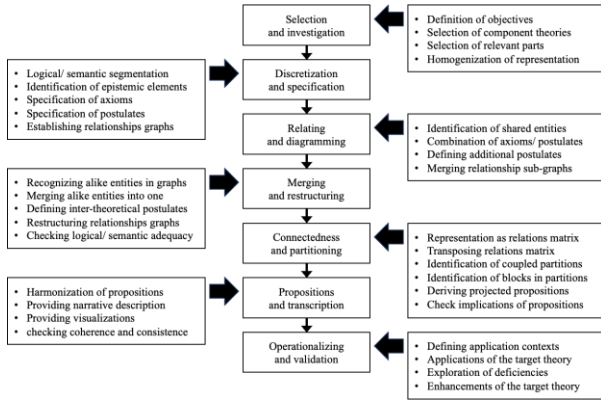


Fig.1 Procedural framework of the ATF methodology.

The process of ATF reflects the logic of deductive reasoning in that it is driven by a proper research hypothesis, and it tries to develop a descriptive, explanatory, or predictive theory by considering relevant existing scientific theories. However, this reflecting is not strong because two issues must be considered. First, a well-founded theory alone might be insufficient to explain the “truth” of the stated hypothesis due to its possible limited coverage. Second, since a source theory in deductive reasoning is typically broader than the target one, the specialization challenge should also be considered. Therefore, systematic combination of component theories was regarded as a way out of the trap of deduction. Nevertheless, a consistent fusing of multiple component theories needs further considerations. Building theory for next-generation data analytics toolbox by axiomatic theories fusion

Our overall objective is to develop a comprehensive theory supporting the conceptualization and implementation of a N-GDAT for functional and embodiment design of white goods based on aggregation and exploration of MoLD. The desired theory should provide an ontological description on what exists or should exist in the development of N-GDAT. It is supposed to specify crucial aspects of the N-GDAT manifestation. This can include behavioral and functioning expectations, and opportunities of the N-GDAT. To achieve this objective, we decided to follow the steps of the ATF methodology. This choice is underpinned by the fact that an axiomatic theory can provide theoretical basis for design processes and help designers in decision-making, reducing complexity and unnecessary repetitions of design procedures and imparting systematization and rationalization to design activities [5].

2.1 Selection of relevant component theories

Our reasoning for this deductive study is represented

in Fig. 2. It illustrates an overall process flow of enhancement of a particular family of products (white goods) by product developers (product designers) who are not data analysts. According to our conceptualization, these product developers used various data analytics tools to generate data about the middle-of-life (MoL) of products and converted these data into knowledge that served as the basis of idea generation for product enhancement. The most favorable enhancement options should be chosen by decision-making. Several important methodological connections among the elements of the overall process flow could be identified.

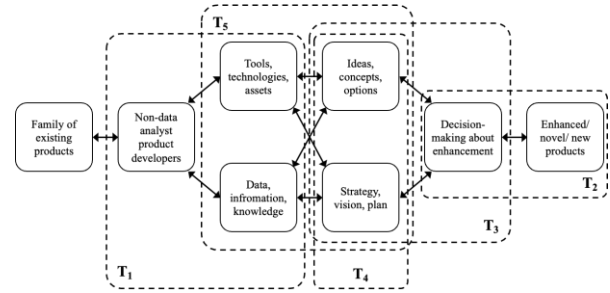


Fig.2 Reasoning model of our deductive study.

Based on previous knowledge, the theories to be fused were chosen based on the following assumption: “a robust and comprehensive conceptual basis for a knowledge platform for a next generation data analytics toolbox for white goods designers needs a combination of many composite theories about (i) explicit needs of designers, (ii) issues of interoperability, (iii) principles of decision-making, (iv) evolution of data analytics, and (v) enabling technologies”. After investigating existing theories, we selected the following theories: (1) T_1 [Theory concerning the needs of designers] that captures the needs of white goods designers related to data analytics tools through a set of inductively generated requirements that they should fulfill when used in this context [4]; (2) T_2 [Theory describing advanced technological enablers] that provides knowledge about how software and cyber-physical-system tools could be exploited as enablers [6]; (3) T_3 [Theory explaining the evolution of data analytics] that provides knowledge on the methods to be used by designers and the means for patterns extraction and for handling MoLD [7]; T_4 [Theory of combined creative problem-solving and decision-making] combines two relevant theories, that describe the methodological and epistemological relationships between preliminary knowledge generation, courses of action, effective solution development, and context-driven robust decision-making, one theory of creative problem solving, which relies on the concept of proactive decision support [8], and the other is the theory of creative decision-making [9]; T_5 [Theory of functional and structural interoperability] that helps determining the architecture of the system as well as the functional design and the structural arrangement of components [10].

2.2 Axiomatic discretization of component theories

- Semantic discretization of theories

Once all relevant theories were selected, we filtered their paragraphs. They were judged relevant based on their semantic meaning and their concrete relationship to the application case. Paragraphs with similar meanings were not considered. The chosen ones were objectively decomposed into a list of short meaningful statements without modifying their original meaning. The statements were written as sentences composed of entities and relationships between them. For example, if we consider the statement “data analytics generate knowledge”, the entities are “data analytics” and “knowledge”, and “generate” represents the relationship between them. They are represented as $E_{x,i}$, where “x” is the identifying number of the containing theory, and “i” is the order of appearance of the entity within the theory “T_x”. From all statements, we defined a set of entities for every component theory. The total number of entities extracted from T₁ was 129. From T₂, 218 entities were extracted. From T₃, 174 entities were extracted. From T₄, 78 entities were extracted. Finally, from T₅, 79 entities were extracted.

- Arrangement and composition of axioms and postulates structures

The statements were used to derive axioms and postulates for all concerned theories. An axiom is represented as $A_{x,j}$ (where “x” is the identifying number of the theory, and “j” is the order of formulation of the axiom). Postulates were distinguished in two forms: (i) Postulates derived directly from the theories, represented as $P_{x,k}^D$ (where “x” is the identifying number of the theory, and “k” is the order of formulation of the derived postulate), (ii) Auxiliary postulates added based on additional domain or problem knowledge, represented as $P_{x,l}^A$ (where “x” is the identifying number of the theory, and “l” is the order of formulation of the auxiliary postulate).

Table 1 Sample of axioms and postulates derived from some of the chosen component theories.

Theory	Axiom/postulate code	Involved entities	Denomination
T ₂	A _{2,97}	E _{2,84} ; E _{2,92}	(Uncertain data) ⁸⁴ [is used for] (inference) ⁹²
T ₅	P _{5,7} ^D	E _{5,13} ; E _{5,17}	(System-system interaction) ¹³ [is to be considered in designing] (cyber-physical systems) ¹⁷

In this step we generated for (i) T₁: 78 axioms, 10 derived postulates, and 89 auxiliary postulates, (ii) T₂: 153 axioms, 7 derived postulates, and 134 auxiliary postulates, (iii) T₃: 141 axioms, 3 derived postulates, and 113 auxiliary postulates, (iv) T₄: 67 axioms, 17 derived postulates, and 36 auxiliary postulates, and for (v) T₅: 57 axioms, 7 derived postulates, and 29 auxiliary postulates. Table 1 groups samples of axioms and postulates derived from the some of the component theories.

2.3 Semantic and visual capturing of relationships

- Creating relationships network

Every axiom and postulate captured an elementary statement of the theory in the form of semantic relationships. In this step, so-called relationships networks were built. We first graphically represented the entities and then linked them to each other considering the nature of the relationships between them (axiom(s) and/or postulate(s)). This graphical representation was needed not only to capture the connectivity between entities but also to visualize the unconnected entities and determine the possibility of connecting them through new logical auxiliary postulates to complement the original list of postulates. In this way, an augmentation of the relationships between entities was created. Connectivity between the disconnected parts of the graph was established only if it had a meaningful content and served the purpose. The added set of postulates should not, in any case, conflict with the original statements and logic of the source theory.

We graphically represented the relationships network in Microsoft Visio. Fig. 3 shows an example of the visualization obtained. The circles represent the various entities, the number in the circles is the order of the entity in the textual formulation of the theory, the black arrows represent relationships contained in an axiom, the red arrow represents relationships contained in a postulate derived directly from the theory, and the blue arrow represents relationships contained in an auxiliary postulate. The numbers within arrows indicate the numbers of relationships between entities. It can be seen that auxiliary postulates connect numerous disconnected and distant entities. The relationships network was needed to enrich individual theories, partially contextualize them, and capture the connectivity between entities, but the actual connection could not be detected from it. Thus, another type of representation was needed for this purpose.

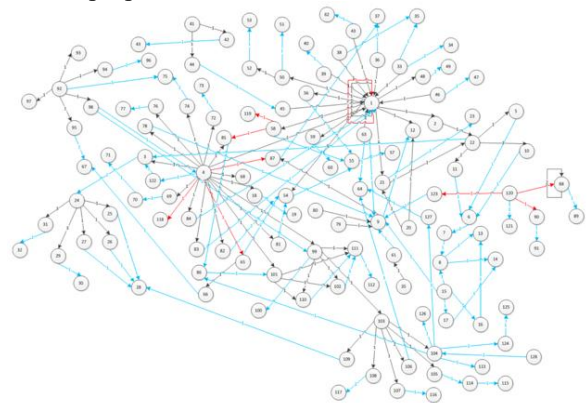


Fig. 3 Relationship network of T1.

- Matrix representation and rearrangement

For another dimension of visualization, we carried out the matrix decomposition of the five theories to facilitate movement toward a semantic capturing of relationships.

In this representation, entities were rows and columns of the matrix, and relationships connected the entities. The matrixes generated for all theories were large, which

made it difficult to capture and include as illustration of this step. The matrix $M(n \times n)$ is a symmetric construct of size “n”. If there is no relationship between entities (empty slots of the matrix), the value “0” was included for these entities. This step was completed using Microsoft Excel. The matrix decomposition does not represent the semantic distance between entities. For this reason, it is an intermediate step between relationships networks and matrix rearrangement. The rearrangement algorithm in Matlab reorders the initial matrix into a matrix that brings the entities into a closer (and more expressive) relationship with each other, instead of the initial, largely arbitrary order. In other words, it is a matrix of blocks called Block_M($n \times n$). We call these blocks of the matrix clusters. This method was used to group the syntactically linked entities. In the results of this procedure, 30 clusters were distinguished for T₁, 54 clusters for T₂, 40 for T₃, 2 large clusters for T₄, and 6 clusters for T₅. The matrixes obtained in the matrix decomposition process were huge and therefore could not be inserted in this dissertation. Therefore, we only present parts of the matrixes for illustration, see Fig. 4.

- Deriving proposition in the context of building a N-GDAT

Propositions represent a logical combination of axioms and postulates contained in the same cluster (they can be many) with regards to the application context. The propositions were categorized as (i) relevant, (ii) partially relevant, or (iii) irrelevant based on their implications and their importance in building a new theory. Irrelevant propositions (vague or out of context) were directly deleted and were not considered in further steps of the methodology. Relevant and partially relevant (those that showed a certain level of implication but were not concrete enough) propositions were kept so we could study their further implications in the theory fusion step.

	E _{5,39}	E _{5,38}	E _{5,35}	E _{5,33}	E _{5,36}	E _{5,34}	E _{5,23}	E _{5,37}	E _{5,19}	E _{5,40}	E _{5,32}	E _{5,22}	E _{5,21}
E _{5,39}	0	0	0	0	0	0	0	0	0	0	0	0	0
E _{5,38}													
E _{5,35}													
E _{5,33}													
E _{5,36}	0	0	0	0	0	0	0	0	0	0	0	0	0
E _{5,34}													
E _{5,23}	0	0	0	0	0	0	0	0	0	0	0	0	0
E _{5,37}	0	0	0	0	0	0	0	0	0	0	0	0	0
E _{5,19}	0	0	0	0	0	0	0	0	0	0	0	0	0
E _{5,40}	0	0	0	0	0	0	0	0	0	0	0	0	0
E _{5,32}	0	0	0	0	0	0	0	0	0	0	0	0	0
E _{5,22}													
E _{5,21}	0	0	0	0	0	0	0	0	0	0	0	0	0

Fig. 4. Part of T₅'s rearranged matrix.

By analyzing the clusters of each theory separately, we identified 15 relevant, 1 partially relevant, and 14 irrelevant clusters in T₁; 41 relevant, 4 partially relevant, and 9 irrelevant clusters in T₂; and 26 relevant, 7 partially relevant, and 7 irrelevant clusters in T₃. The two clusters of T₄ were both relevant, and T₅ contained 5 relevant clusters and 1 irrelevant cluster. By analyzing relevant and partially relevant clusters of each theory, we derived 34 propositions from T₁, 76 propositions from T₂, 45 propositions from T₃, 27 propositions from T₄, and 12 propositions from T₅. Examples of these propositions are

given in Table 2. Analyzing the implications of the propositions formulated in the application context revealed that they were all directly or indirectly linked to the research phenomenon. For this reason, no filtering of the propositions was done in this step.

Table 2 Sample of propositions derived from the component theories.

Theory	Denomination
T ₁	Designers need procedural reasoning and case-based reasoning
T ₂	Traditional and mathematical modelling do not solve complex real-world data driven problems
T ₅	Intelligent-based system-human interaction, proactivity and awareness are to be considered in the case of intelligent systems

2.4 Fusion of the component theories

- Syntactic processing and merging of component theories

Based on the ATF methodology, our first step in merging theories was to combine the lists of entities of component theories and then merge them into one list with no redundancies. Our final list contained 574 entities, including 81 common entities represented as $E_{c,x}$ (where “x” is the order of appearance of the common entity). The second step was the axiomatization of the combined theories, taking into consideration the final list of entities. This step required less effort, since axioms and postulates were already established and written in their final format, only duplications needed to be removed. We ended up with a list of 924 relationships, including 492 axioms, 44 derived postulates, and 388 auxiliary postulates.

Our third step was to establish a relationships network connecting all theories. To facilitate capturing the semantics of the relationships. This visual representation of the relationships network was converted into a matrix representation. From this representation, the statements concerning the particular entities could be grouped and further analyzed. The size of the matrix was 574×574 cells. It contained 878 defined relationships. Our fifth step was to convert the matrix in Matlab into a block-matrix. The later contains 95 clusters. The remaining two steps of the theory fusion are detailed below.

- Deriving propositions based on units of resultant theory

The total set of 95 clusters contained 40 relevant, 13 partially relevant, and 42 irrelevant clusters. The irrelevant clusters did not affect the relevance of the methodology, since clusters were not identical in terms of the numbers of included relationships. For example, one relevant cluster can contain more relationships than a group of 10 irrelevant clusters. From the relevant and partially relevant clusters, 82 propositions were constructed, of which 77 were relevant and 5 only partially relevant. We decided to keep only relevant propositions. By combining the relationships in each cluster semantically and putting them into the context of

the study, a set of propositions was derived. The natures of these propositions fell into four categories: (i) requirement, (ii) descriptive, (iii) explanatory, and (iv) control content.

Below is a sample of obtained propositions of different natures (the number represents the order in which the propositions were formed):

Proposition₁₁: “The N-GDAT incorporates self-learning, self-management, and self-adaptation capabilities”. *This proposition is of a descriptive nature.*

Proposition₁₃: “The intellectualization of the N-GDAT is provided by system learning mechanisms, situation awareness, strategy development, and system adaptation capabilities”. *This proposition is of a control content nature.*

Proposition₁₄: “The N-GDAT allows sentiment analysis in order to identify user’s opinion”. *This proposition is of an explanatory nature.*

Proposition₆₄: “The N-GDAT should include case-based reasoning and should learn from its applications”. *This proposition is of a requirement nature.*

By analyzing all obtained propositions, we observed that they are not at the same level. This checking distinguished two levels of propositions from the general to the particular (see Table 3), that could be put under a general category. This result was unexpected.

- Transfer of the propositions into a narrative description

The final set of propositions was dedicated to the context of developing a N-GDAT that goes beyond individual tools and covers white goods designer’s expectations. It described the expectations and the basis for the toolbox implementation requirements. Some of the requirements (algorithms and methods) were clearly formulated in the clusters and were evident without further investigation. We had to further investigate the remaining requirements to determine their implications for the construction of the toolbox. Examples of the obtained requirements are presented below:

- Large-scale data within the N-GDAT could be modeled using support vector machines, naive Bayes, or logistic regression.
- The N-GDAT could learn from experiential data using a back propagation algorithm.
- Validation of the ATF methodology

The ATF was used in the particular case of developing a theory to build a N-GDAT. To this end, we posed the following question: was our objective achieved?

By evaluating the propositions derived from the set of relationships supported by axioms and postulates, the result was satisfying, since it did not present contradictions and formed a knowledge platform about the ideation of a N-GDAT. This meant also that the set of axioms and postulates was relevant. The obtained requirements connected the specification of the new-generation toolbox and its ideation and conceptualization. Considering all findings, it is possible to conclude that chosen theories were relevant and allowed the

establishment of the needed knowledge platform for the application case. The objective of applying ATF in a particular data analytics design context was to develop a knowledge platform for a next generation N-GDAT for white goods designers. The outcomes exceeded the objective by providing a multilevel set of functionalities and requirements to be implemented for a next generation N-GDAT. In our study, the ATF methodology was relevant.

Table 3 Examples of propositions of different levels

General category	1 st level of propositions	2 nd level of propositions
Data type and characteristics	The N-GDAT identifies the user’s opinion	The SDATB allows sentiment analysis to identify user’s opinion

DISCUSSION

The new theory obtained based on the ATF application was able to deliver more knowledge than individual theories. The outcomes of ATF in this context were formulated as a list of 77 relevant propositions converted into requirements for the N-GDAT and categorized into ten clusters: (i) decision-making, (ii) algorithmic concepts, (iii) learning, (iv) data management, (v) interfacing, (vi) reasoning, (vii) data types and characteristics, (viii) design issues, (ix) analytics techniques and methods, and (x) outputs. The propositions obtained reflected the functionalities that need to be included in a N-GDAT. Some of these functionalities are already provided by existing tools, such as analyzing labeled data and modeling large-scale data. They also reflected what should not be part of the toolbox, such as (i) traditional analytics techniques inadequate for handling big data from smart products and (ii) deep neural networks that are computationally expensive and require long training times for pattern recognition. The major findings of the ATF application were the novel functionalities not yet covered by existing data analytics tools. Table 4 presents a sample of these functionalities.

In a previous study it was emphasized on the need for (i) step-by-step assistance, (ii) advice in selecting means, (iii) multifold data visualization, (iv) multichannel data management, (v) blending datasets, (vi) combining qualitative and quantitative data, (v) permanent accessibility, (vi) adaptation to users, (vii) case-based reasoning, and (viii) learning from applications in building N-GDAT [4]. This was completed by the new synthetic theory obtained from the ATF that is more concrete about the needs and the requirements for the N-GDAT. This theory suggested that the toolbox should include (i) context-driven decision-making, (ii) proactive decision-making, and (iii) algorithms able to process complex data. In addition, it should (iv) allow semantic interpretation, (v) blend data and datasets, (vi) merge

multiple data streams, (vii) allow high-speed and high-volume storage, (viii) provide permanent accessibility, (ix) deliver advice to designers based on their work context, (x) allow case-based reasoning, (xi) process structured, semi-structured, and multi-structured data, (xii) propose solutions to solve difficult design problems, (xiii) predict future outcomes, and (xiv) derive actionable insights.

Table 4 Sample of the major findings of the ATF application.

1 st level of requirements	2 nd level of requirements
Interfacing: N-GDAT permanently advises designers in their choices	N-GDAT is permanently accessible
	N-GDAT recognizes its user
	N-GDAT helps its user in his choices (step by step)
Reasoning: N-GDAT allows case-based reasoning	N-GDAT detects the context of the analysis
	N-GDAT reasons with cases
	N-GDAT offers solutions based on saved manipulations

By reflecting on the ATF methodology, it was indeed very useful but presents several limitations. The combination of its both manual and computer-aided methods and techniques made its manipulation challenging, and time-consuming. It can become very tiring in certain complex application cases that involve fusing a large number of theories. Nonetheless, these manual manipulations could not be fully automated given the need for human comprehension, semantic interpretation, logical reasoning, reductionist decomposition, consistency checking, and compliance testing.

CONCLUSIONS

Our ultimate goal is to develop a N-GDAT to support white goods designers in enhancing/designing products based on MoLD. To do so we needed to have a theory explaining how to do so. In the literature we encountered a lack of theories reflecting on our interest. Then a question arose: are there methodologies to follow to overcome this lack? The methodologies for theorizing proposed in the literature cannot be used directly in some research areas in which more than one aspect needs to be considered. The complications of using existing approaches became more persistent in the presence of semantically distant aspects. The methodology of ATF makes theories that are insufficient individually (in some contexts or when used alone) more valuable and insightful when combined. It was challenging to apply the ATF in a design context. The challenge was not only in the mathematical nature of the components (axioms and postulates) but also in the manual work done in major parts of the approach. This was a beneficial but time-consuming procedure because of the high level of precision and focus required in all steps to avoid mistakes that risked becoming apparent in later stages.

The ATF provided us with a new theory for building N-GDAT. We were able to merge originally distant theories with each other and convert the results of the merging into requirements for building a N-GDAT. We have detected no contradictions in the process of building the new theory. Also, we were able to put together theoretical and practical needs of white goods designers in a theory including a set of needs, expectations, techniques, and technologies needed for building N-GDAT. These findings were contained into clusters as presented in Table 4. allowing us to form a skeleton of the N-GDAT.

REFERENCES

- [1] A. Oussous, F. Z. Benjelloun, A. A. Lahcen, and S. Belfkih, "Big data technologies: A survey", *Journal of King Saud University-Computer and Information Sciences*, Vol. 30, No. 4, pp. 431-448, 2018.
- [2] T. R. Morgan, M. Tokman, R. G. Richey, and C. Defee, "Resource commitment and sustainability: A reverse logistics performance process model", *International Journal of Physical Distribution & Logistics Management*, Vol. 48, No. 2, pp. 164-182, 2018.
- [3] F. Suppe, "Axiomatization", *A Companion to the Philosophy of Science*, pp. 9-11, 2017.
- [4] F.-Z. Abou Eddahab-Burke and I. Horváth, "What do designers miss regarding the outputs of data analytics tools in the context of possible product improvements?", in *Proceedings of the 12th Symposium on Tools and methods of Competitive Engineering (TMCE)*, Las Palmas of Gran Canaria, Spain, 2018.
- [5] J. Qui, J. Zhang, Y. Ma, J. Feng and P. Wang, "Reliability-based robust design for kinematic accuracy of a mechanism with axiomatic theory", *Journal of Aerospace Engineering*, Vol. 232, No. 13, pp. 2412-2423, 2018.
- [6] R. Iqbal, F. Doctor, B. More, S. Mahmud, and U. Yousuf, "Big data analytics: Computational intelligence techniques and application areas", *Technological Forecasting and Social Change*, Vol. 153, No. 2020, 119253, 2018.
- [7] S. Sruthikha and N. Tajunisha, "A study on evolution of data analytics to big data analytics and its research scope", in *Proceedings of the International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, Coimbatore, India, 2015.
- [8] M. Mishra, D. Sidoti, G. V. Avvari, P. Mannaru, D. F. Ayala, K. R. Pattipati, and D. L. Kleinman, "A context-driven framework for proactive decision support with applications", *IEEE Access*, Vol. 5, pp. 12475-12495, 2017.
- [9] S. H. Kim, "A formal model of creative decision making", *Robotics and Computer-Integrated Manufacturing*, Vol. 8, No. 1, pp. 53-65, 1991.
- [10] I. Horváth and J. Wang J, "Towards a comprehensive theory of multi-aspect interaction with cyber-physical systems", in *Proceedings of the ASME Design Engineering Technology Conference*, Boston, USA, 2015.