

Ph.D. Research Project

Candidate Name	Muhammad Suffian
Project Title	FORMAL MODELS, DATA ANALYSIS AND SCIENTIFIC COMPUTING
Keywords	Data Analysis, Digital Transformation, Artificial Intelligence

2- Research Area

The aim of the research is to concentrate on developing a data-driven computational framework, based on AI approaches, which will be able to perform data analysis and prediction. The development of framework will be done in the scope of the digital transformation for different application domains, including Public Administration, Digital Finance, and Digital Industry. The validation of the framework will be performed addressing problems in the application domains, by exploiting the data sets and services integrated in the Digital Hub.

3- Introduction and State of the Art

What is the way to more intelligent for digital finance and administration that can settle on information driven choices? People around the globe are creating, sharing, and interacting with billions of new digital artifacts, generating an exponential explosion in the amounts of data. The most challenging task for the public administration is to deal with large amount of data. This big data on modern humanity accompanied by massive digitization of social media content, financial data, public administration relating data and economic data. This requires some general framework driven by such data to transform the public administration, digital hubs, finance and industry into digitalization. In a setting where information have been immensely promoted as an answer for some issues, the utilization of huge information investigation on a government level is a drifting point. As a theme and possible technique, information is on the plan of specialists and strategy creators around the world, in view of the expectation that the utilization of computational intelligence could empower better personal satisfaction in the above mentioned application fields. The general framework proposes an integrative system to more readily see all the segments of these ideas and their interrelationships.

4- Research Objective

The main objectives of the Ph.D. research are:

1. Understanding and insights data hubs and analysis of the data with an eye towards the future.
2. Understanding of the existing data-driven approaches for digitalization and Understanding the framework modelling in general and with artificial intelligence for the application tasks proposed in the project.
3. To design and develop the computational framework driven by big data for each digitalized application domain and an integrative approach for computational framework as all in one for the Digital Hub initiated by the project financier.
4. Use of data analysis, machine/deep learning, natural language processing, visualization, and data mining techniques for prediction.
5. Validation of the framework by using domain datasets from the Digital Hub.

Ph.D. Research Project

5- Literature Review

Researcher's community proposed various data-driven approaches for framework in specific domains. In this part background of relevant systems and related work to computational framework driven by data with artificial intelligence is given. At the beginning of 90's people started digitalization in public administration by collecting data manually but it was not accurate due to the traditional tools used for this purpose [1]. In 2005 Italian system has adopted the legislative reforms for the administrative procedures, in which the transparency was insured and the information available as democratically [2]. Increase in the data made an impact on manufacturing industry, in [3] they have proposed a conceptual framework for the smart manufacturing driven by big data with. Borodulin, Kirill, et al. introduced a Digital-Twins concepts, in which they have proposed a simulation model for each system as a computational framework and then for their integration they used the cloud services [4]. Lafferty [5] performed different experiments and analyses on the industrial manufacturing and its revenue, cost and efficiency by using the data collected from different places like British Science week etc. He used the structural equation modelling for designing the framework with data driven approach. Yassine et.al proposed a computational and theoretical framework [6] for product development systems in enterprises. Their framework integrate the dependencies on data, product and processes; also it provides insights about the influence of factors on the product and suggest the mechanism of acquiring of data, its interpretation and applying.

5- Methodology and Expected Results

Research methodology comprises the following steps:

1. Data collection

Get and acquired the data required for the project from the real world entity that can be organization, social network or third party etc. The Digital Hub would be the main source of data.

2. Data Processing

It involves data cleansing, data transformation, and data combining. This step cope with the missing values, outliers, physical impossible values, aggregating data, features dimensionality reduction and merging data sets.

3. Domain knowledge-guided partitioning

All the tasks related to application domains will be performed with domain knowledge expertise. Machine learning based model will be filtering and partitioning the tasks according to the knowledge rules for the better task performance Like in Finance framework model the policies and the actual processes in the real setting will be partitioned for the better data-driven operations.

4. Framework modelling

Computational framework will be modelled and developed for each domain application separately with artificial intelligence and will be tested on the datasets by exploiting the Digital Hub. After the correct operations of each specific framework, the integrative computational framework be designed and the data streams will be utilized from digital hub for the testing purposes.

5. Visualizations and interfacing

Ph.D. Research Project

In this part, data visualization will be done for the better task operations through the dialogue based interface. This interface will provide analysis and predictions in visual forms to domain enterprises for decision making.

6. Framework validation

The framework will be evaluated from the datasets present on the digital hub by exploiting the required data for the specific problem in application domain. Here, the feedback system will be incorporated by using deep learning RNN model which will feed the output back into the system as an input.

6- References

1. Mergel, Ines, R. Karl Rethemeyer, and Kimberley Isett. "Big data in public affairs." *Public Administration Review* 76.6 (2016): 928-937.
2. Carloni, Enrico. "New perspectives of administrative transparency: from access to documents to the availability of information." *Public Law* 11.2 (2005): 573-600.
3. Tao, Fei, et al. "Data-driven smart manufacturing." *Journal of Manufacturing Systems* 48 (2018): 157-169.
4. Borodulin, Kirill, et al. "Towards digital twins cloud platform: Microservices and computational workflows to rule a smart factory." *Proceedings of the 10th International Conference on Utility and Cloud Computing*. 2017.
5. Lafferty, Clive. "Sustainable Industry 4.0: Product decision-making information systems, data-driven innovation, and smart industrial value creation." *Journal of Self-Governance and Management Economics* 7.2 (2019): 19-24.
6. Yassine, A. A., and Bradley, J. A. (March 15, 2013). "A Knowledge-Driven, Network-Based Computational Framework for Product Development Systems." *ASME. J. Comput. Inf. Sci. Eng.* March 2013; 13(1): 011005. <https://doi.org/10.1115/1.4023166>

7- Tentative 3 Years Ph.D. Proposed Research Time Line and Targets

Time Line	Targets
November 2020 - April 2021	Literature review Learning tools and technologies
November 2021 - April 2022	Framework modelling Preparing two conference submissions
May 2021 – October 2021	Data collection, Data cleansing Data processing
May 2022 – October 2022	Interfacing and integration using AI Preparing a journal submission.
November 2022 - April 2023	Validation, prediction, analysis
May 2023 – October 2023	Results Comparison, Thesis Writing Complete data product