

## Exploration of data analytics for ground segment in space systems

Business Modeling and Software Design - 10th International Symposium, BMSD 2020, Proceedings

Tekinerdogan, Bedir; Acar, Bedia; Cabioğlu, Çağrı; Savaş, Damla; Vuran, Nebi et al

[https://doi.org/10.1007/978-3-030-52306-0\\_25](https://doi.org/10.1007/978-3-030-52306-0_25)

This article is made publicly available in the institutional repository of Wageningen University and Research, under the terms of article 25fa of the Dutch Copyright Act, also known as the Amendment Taverne. This has been done with explicit consent by the author.

Article 25fa states that the author of a short scientific work funded either wholly or partially by Dutch public funds is entitled to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed under The Association of Universities in the Netherlands (VSNU) 'Article 25fa implementation' project. In this project research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and / or copyright owner(s) of this work. Any use of the publication or parts of it other than authorised under article 25fa of the Dutch Copyright act is prohibited. Wageningen University & Research and the author(s) of this publication shall not be held responsible or liable for any damages resulting from your (re)use of this publication.

For questions regarding the public availability of this article please contact [openscience.library@wur.nl](mailto:openscience.library@wur.nl)



# Exploration of Data Analytics for Ground Segment in Space Systems

Bedir Tekinerdogan<sup>1</sup>(✉), Bedia Acar<sup>2</sup>, Çağrı Cabioglu<sup>2</sup>, Damla Savaş<sup>2</sup>, Nebi Vuran<sup>2</sup>, Şenol Tekdal<sup>2</sup>, and Ümit Gürsoy<sup>2</sup>

<sup>1</sup> Information Technology, Wageningen University & Research, Wageningen, The Netherlands  
bedir.tekinerdogan@wur.nl

<sup>2</sup> TUBITAK Space Technologies Research Institute, Ankara, Turkey  
{bedia.acar, cagri.cabioglu, damla.savas, nebi.vuran, senol.tekdal, umit.gursoy}@tubitak.gov.tr

**Abstract.** Space systems have to deal with massive spatio-temporal Earth and Space observation data collected by space-borne and ground-based sensors. Despite the data latency in communications, data is collected at enormous rates, and a sophisticated network of ground stations is set up to collect and archive telemetry data. The data that is received at the ground segment can be made available to the end-users. Beyond archiving data, the available data provides opportunities for data analytics that can support the decision-making process or provide new insight for the target requirements. Unfortunately, for practitioners, it is not easy to identify the potential and challenges for data analytics in the space domain. In this paper, we reflect on and synthesize the findings of existing literature and provide an integrated overview for setting up and applying data analytics in the space systems context. To this end, we first present the process as adopted in space systems, and describe the data science and machine learning processes. Finally, we identify the key questions that can be mapped to data analytics problems.

**Keywords:** Space systems · Ground segment · Data analytics

## 1 Introduction

In the context of spaceflight, a satellite is an artificial object which has been intentionally placed into orbit. Such objects are sometimes called artificial satellites to distinguish them from natural satellites such as the Moon. A space system consists basically of three segments, including Ground Segment, Launch Segment, and Space Segment [5–7, 9]. Each segment by itself is a complex system integrating hardware and software. Space systems have to deal with massive spatio-temporal Earth and Space observation data collected by space-borne and ground-based sensors. Despite the data latency in communications, data is collected at enormous rates, and a sophisticated network of ground stations is set up to collect and archive telemetry data [13]. The data that is received at the ground node can be made available to the end-users. Beyond archiving data, the available data provides opportunities for data analytics that can support the decision-making process or provide new insight for the target requirements.

In this context, the term Big Data usually refers to data sets with sizes beyond the ability of commonly used software tools to capture, curate, manage, and process data within a tolerable elapsed time [1, 2, 12, 14, 16]. The extraction of useful information out of the big data is done using data analytics. Data analytics includes inspection, pre-processing, and modeling of data to discover useful information. The data analytics processes have been typically automated using machine learning and/or deep learning algorithms. In principle, four different types of data analytics can be distinguished. *Descriptive analytics* describes what has happened over a given period. *Diagnostic analytics* focuses on why something has happened. *Predictive analytics* aims to find out what is likely going to happen in the near term. Finally, *prescriptive analytics* suggests a course of action.

Like in many industrial domains, we can observe that data science in general and (big) data analytics, in particular, has triggered opportunities for creating smart space systems based on information derived from raw data. Unfortunately, for practitioners, it is not easy to identify the potential and challenges for data analytics in the space domain. Several studies discuss the application of data analytics to specific space domain problems but a general overview is largely missing. As such, we reflect on and synthesize the findings of existing literature and provide an integrated overview for setting up and applying data analytics for the ground segment in the space systems context. In particular, we aim to answer the following research questions:

- RQ1. What are the important questions for the ground segment of space systems?*
- RQ2. Which type of data analytics will be required to answer these questions?*
- RQ3. What are the required machine learning approaches for these data analytics problems?*

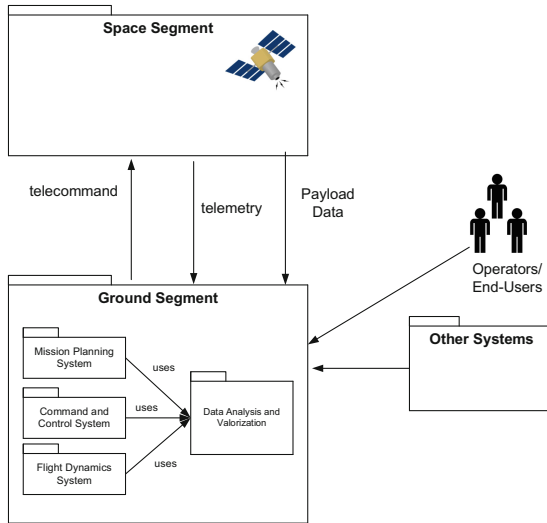
Our aim is not to provide a detailed analysis to derive the analysis for these questions, but rather to provide an overview of the current literature that will help practitioners identify the key solution directions. This objective was directly derived from our own industrial context of space domain projects. To this end, we first present the background and the process as adopted in space systems, together with an overview of data science and machine learning processes.

The remainder of the paper is organized as follows. In Sect. 2, we shortly describe the overview of space systems, thereby focusing on the ground segment. In Sect. 3, we discuss the problem statement. Section 4 presents the approach, and finally, we conclude the paper in Sect. 5.

## 2 Space System

Figure 1 shows a conceptual model of a typical space system consisting of the ground segment and space segment. The space segment has onboard computers, data-handling systems, attitude and orbit control systems, all of which contain software. The ground segment has command and control systems, simulators, flight-dynamics systems, mission-analysis tools, communications networks, and ground-station data systems such as telemetry and telecommand processors, as well as ‘downstream processing’ systems

for payload data. These all contain software, often of considerable complexity. The results of the ground segment are typically used by external systems, which include application-specific software.



**Fig. 1.** Conceptual architecture of space system

The ground segment can send so-called *telecommands* to the satellite (through radiofrequency links), to initiate behaviour of the satellite, and sometimes to override internal decisions taken by the onboard software [10]. Telecommands are normally sent as asynchronous data packets, and are always received by the satellite central computer (the OBDH computer). Depending on the type of mission, the link with the ground station can be continuous or extend only over a section of the orbit.

The satellite on its turn can send *data* to the ground, which can be divided into two broad categories [10]:

- *The payload data* (mission data) represents the data collected by the satellite. In the case of an astronomical telescope, for instance, the telemetry contains the pictures taken by the telescope.
- *Housekeeping data* (telemetry) gives information about the general status of the satellite. Telemetry data are sent in packets.

Data analytics is typically part of the ground segment, which is the system set up on Earth to manage and control a space mission, and to receive and process the data produced by a spacecraft’s instruments, and if necessary, to send out and archive any generated products. A principal telecommunications device of the ground station is the parabolic antenna. When a satellite is within a ground station’s line of sight, the station is said to have a view of the satellite. It is possible for a satellite to communicate with

more than one ground station at a time. The ground segment consists of the following systems:

### **Mission Planning System**

Mission planning is defined as the task of preparing, organizing, and planning all relevant activities that happen during the mission, onboard as well as on-ground. It is important that the plan is correct and conflict-free and, as such, needs to be thoroughly verified.

### **Command and Control System**

Command and control system manages satellites orbit, usually from the point of lift-off up to the orbit, and the end of the mission. From the ground system, a staff of flight controllers monitor all aspects of the mission using telemetry, and send commands to the satellite.

### **Flight Dynamics System**

Spacecraft flight dynamics includes functionality for the management of the performance, stability, and control of satellites. In general, spacecraft flight dynamics involve three forces: propulsive force, gravitational force exerted by the Earth or other celestial bodies; and aerodynamic lift and drag.

### **Data Analytics and Valorization**

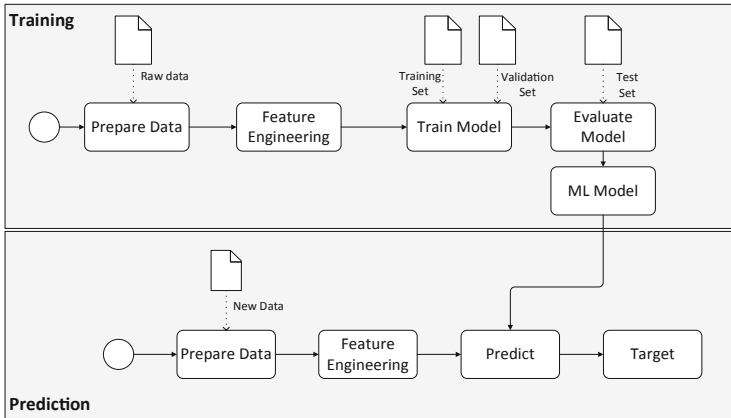
As shown in Fig. 1 the above three modules use the Data Analytics and Valorization module that focuses on the clearing and pre-preprocessing of the received raw data to derive information and likewise to support the decision making. The decision making process on its turn can support the other three functionalities of mission planning, command and control and flight dynamics.

## **3 Data Science and Machine Learning**

To identify the needs for data science for the ground segment modules we will first focus on the conventional data science process. Data science includes a process for obtaining raw data and converting it into information useful for decision-making by users. Data science consists of the following steps [9, 15, 18, 19]:

*Data collection* is a systematic approach for gathering and measuring information from a variety of sources. *Data pre-processing* defines the step in which the data is transformed and prepared to bring it to such a state that it can be further processed by an algorithm. This is often an essential and time-consuming step since data is often taken from different sources using different formats. Data pre-processing includes sub-steps such as data cleaning (e.g. missing, duplicate, noisy data), data transformation (e.g. normalization), and data reduction (e.g. dimensionality reduction). *Data processing* includes the analysis and analytics of the data to derive useful information from the pre-processed data in order to support further understanding and/or decision-making process. *Data visualization* is the graphic representation of the acquired information in order to communicate the relationships among the represented data.

For processing the data and as such realizing data analytics machine learning (ML) can be applied. ML enables computer programs to perform complex tasks such as prediction, diagnosis, planning, and recognition by learning from historical data. Figure 2 shows the traditional process for machine learning. The process consists of two steps, training and prediction. In the training process a ML model is developed that can be used to predict the outcome for new data. For training the model, the initial raw data must be first pre-processed after which the necessary features are extracted for which we wish to analyze the correlations. In training activity the data is split up in a training set and validation set (for tuning so-called hyperparameters). In the evaluation step the model is tested using a test data set and a final ML model is provided. In the prediction activity, the new data can be prepared, and the features extracted, and an outcome can be predicted using the provided ML model. As stated above we distinguish between payload data and housekeeping data. The payload data depends on the mission of the satellite and as such a broad set of ML applications can be identified here. The housekeeping data is often based on structured data that needs to conform to the standards as defined in the space domain.



**Fig. 2.** Machine learning process

In the literature, different ML types can be distinguished [19]: In *supervised learning* a function or predictor is derived from a set of labeled training data. *Unsupervised learning* looks for previously undetected patterns in a data set with no pre-existing labels and with a minimum of human supervision. Unsupervised learning is usually used when relationships among input variables are not known. In several cases, only a small portion of the data is labeled while the majority of the data points are unlabeled. In that case, *semi-supervised learning* algorithms can be used to train a model with both labeled and unlabeled data, which can provide better accuracy compared to the supervised model that uses very limited labeled data. Finally, *reinforcement learning* adopts agents that observe the environment, perform some actions, and get some rewards (negative/positive) based on the selected action, upon which the model is updated accordingly. These machine learning types are agnostic to the application domain data, and as such, can be applied to

support the questions of the mission planning, command and control and flight dynamics of the ground segment in space systems.

Besides of ML types, a number of ML task are distinguished in the literature. *Classification* maps input features to one of the discrete output variables. The output variable represents a class for the underlying problem. For binary classification, the output variable can only be one or zero. For *multi-class classification*, the output variable can consist of several classes. *Regression* maps the input features to a numerical continuous variable. The output variable can be either an integer or a floating-point number. *Clustering* divides data points into relevant groups. This grouping is based on the similarity pattern between data points. Similar points are grouped together and provide valuable information to data scientists. *Data reduction* tasks can reduce the number of features for building models faster. This task is mainly used as an auxiliary method for other machine learning tasks

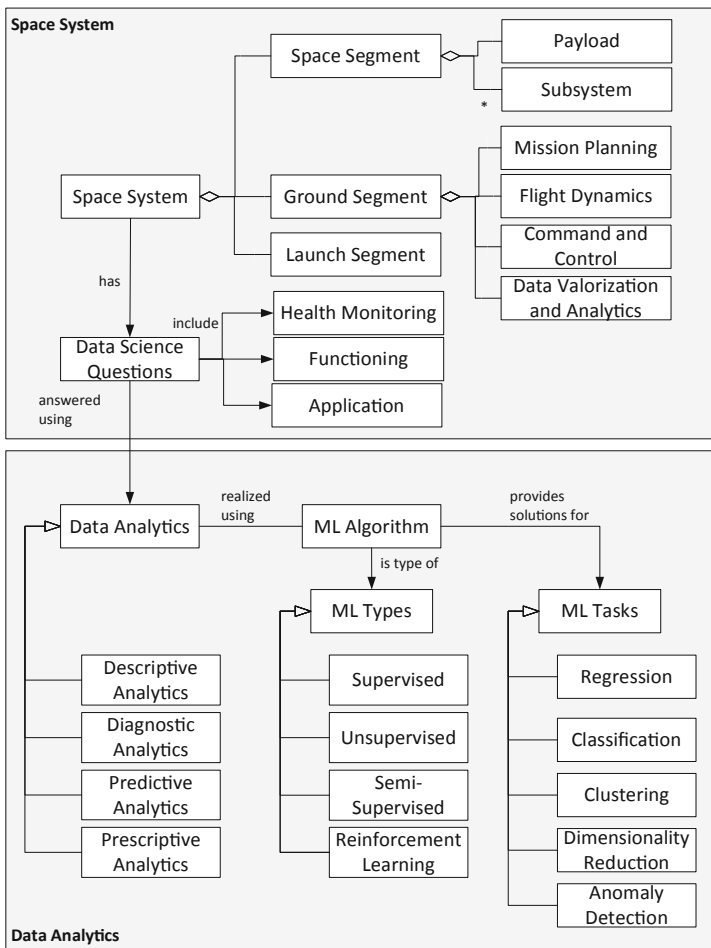
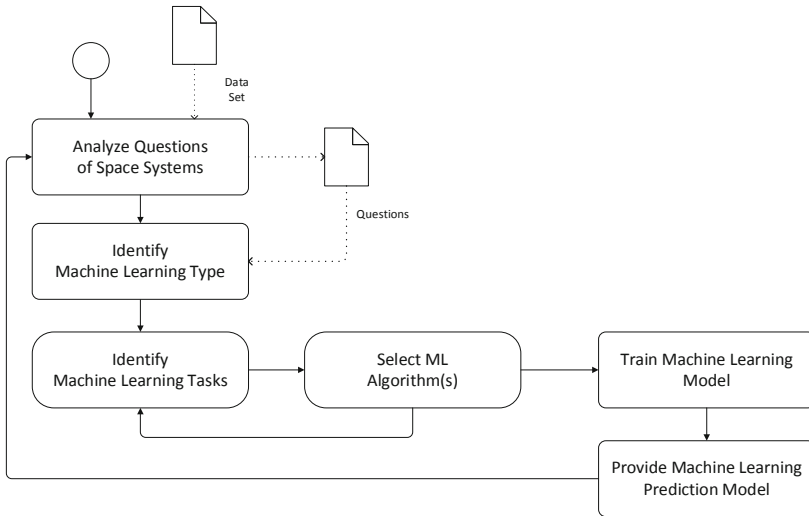


Fig. 3. Metamodel for big data analytics for space systems

such as regression and classification. *Anomaly detection* is usually handled with unsupervised learning methods. Similar to clustering, anomaly detection algorithms group the samples and aims to determine the outliers. In Fig. 3, we summarize the concepts in a metamodel that we can use for our further analysis of data science for the space domain.

### 4 Data Analytics for Space Domain

As stated before, space systems concern the collection of a large amount of data and as such we can identify an important opportunity to process this raw data using data analytics. In Fig. 4, we show the process of guiding this activity.



**Fig. 4.** Workflow process for adopting machine learning process for space systems

Typically, we will start with analysing the key questions that need to be answered. These questions could relate to health monitoring or be more specific to the payload data. Subsequently, the corresponding ML type and ML task is identified. This is followed by the step in which an ML model is trained using selected ML algorithms. Finally, the trained and validated ML model is deployed for usage in the prediction process. In this paper we focus on the steps for identifying ML types and ML tasks. In Table 1 we present the identified space domain questions that require a data science and in particular big data analytics approach.

It should be noted that the questions in Table 1 is not an exhaustive list. The main objective of this study was to synthesize and provide an overview of the concepts. Detailed analysis and application of these questions will be performed in our future work. The required data analytics will be implemented using the corresponding ML tasks. Table 2 shows the identified ML tasks that were derived from the literature [3, 5, 8, 11, 13, 15, 17, 21, 22].



**Table 1.** The identified space domain questions and the relation to data analytics type

Data analytics	Identified space domain questions
Descriptive analytics	Satellite orbit monitoring; Satellite maneuver monitoring; Orbit determination; Non-reliable data source identification
Diagnostics analytics	Health monitoring (Failure Diagnosis)
Predictive analytics	Health monitoring; Fuel consumption prediction; Lifetime-prediction of satellite subsystem; Maneuver prediction; Conjunction analysis; Orbit determination; Orbit propagation; Initial orbit determination; Orbit drift prediction; Satellite simulator of ground segment
Prescriptive analytics	Health monitoring; Radio resource allocation; Maneuver mission planning; Mission planning scheduling Payload re-configuration

**Table 2.** The identified ML tasks

ML task	Identified space domain problems/tasks
Regression	Health monitoring; Telemetry prediction Fault diagnosis; Orbit determination Precise orbit prediction; Initial orbit determination Orbit propagation; Lifetime-prediction of satellite subsystem; Conjunction analysis
Classification	Satellite stability prediction; Maneuver classification; Non-reliable data source identification (Security); Maneuver prediction
Clustering	Health monitoring; Telemetry prediction; Anomalous orbital behavior prediction
Dimensionality reduction	Pre-processing step for telemetry mining applications to reduce high-dimensional telemetry data
Anomaly detection	Anomaly detection of power unit; Anomaly detection of the subsystems; Satellite telemetry data anomaly detection

## 5 Related Work

In this paper we have provided a general and conceptual overview of the space domain and the potential for exploiting the available data analytics techniques in the literature. This idea is not new and a number of studies have been reported about the application of specific techniques to the identified questions in this paper.

Hassanien et al. [13] explore in their edited book the concepts, algorithms, and techniques of data mining in analyzing telemetry data of satellites for health monitoring. Three key domains are distinguished including health monitoring, telemetry data analytics and applications, and security issues in telemetry data.

We have provided a general reference model for the data analytics. Complementary to this is the need to develop a proper architecture for data-intensive space systems. In this context, the Consultative Committee for Space Data Systems (CCSDS) [4] has proposed the Reference Architecture for Space Data Systems (RASDS) [14]. This reference architecture can be used to design concrete space data systems, support reuse, and likewise reduce the cost and time of development of space data systems.

In our earlier work, we have focused on model management and analytics [19] which can be considered as a special case of data analytics. Model Management and Analytics for Large Scale Systems covers the use of models and related artefacts (such as metamodels and model transformations) as central elements for tackling the complexity of building systems and managing data. Also in the space domain a large number of various models are generated which are amenable for further processing to derive useful insight. The integration of model analytics and the selection and development of the corresponding model platform [20] is part of future research.

## 6 Conclusion

A massive amount of data is collected at the ground segment of space systems, which can be made available to the end-users for further processing. In this position paper we have indicated the potential for data analytics that can support the decision-making process or provide new insight for the target requirements. We have reflected on the current space system domain from a data science perspective and developed the metamodel that integrates the important concepts. Three key modules of the ground segment have been identified, including mission planning, command and control, and flight dynamics. We have stated that each of these modules can benefit from data analytics. Further, we have identified the key data analytics approaches, the machine learning techniques and machine learning tasks. We can conclude that there is indeed huge potential for applying data science techniques in the space domain, which have not yet been fully explored. This paper complements on the initiated studies in this domain and aims to further pave the way for future work. In our own future work, we aim to address the key research questions using various machine learning and deep learning approaches.

## References

1. Avci Salma, C., Tekinerdogan, B., Athanasiadis, I.: Feature driven survey of big data systems. In: Proceedings of the International Conference on Internet of Things and Big Data, Rome, pp. 348–355 (2016)
2. Avci Salma, C., Tekinerdogan, B., Athanasiadis, I.: Domain-driven design of big data systems based on a reference architecture. In: Mistrik, I., Bahsoon, R., Ali, N., Heisel, M., Maxim, B. (eds.) Software Architecture for Big Data and the Cloud, pp. 49–68 (2017)

3. Bradley Knox, W., Mengshoel, O.: Diagnosis and reconfiguration using Bayesian networks: an electrical power system case study. In: Twenty-First International Joint Conference on Artificial Intelligence (2009)
4. Consultative Committee for Space Data Systems (CCSDS). <http://www.ccsds.org>. Accessed May 2020
5. Dibona, P., Foster, J., Falcone, A., Czajkowski, M.: Machine learning for RSO maneuver classification and orbital pattern prediction (2019)
6. Eickhof, J.: On-Board Computer, On-Board Software, and Satellite Operations. Springer, Heidelberg (2012). <https://doi.org/10.1007/978-3-642-25170-2>
7. European Cooperation for Space Standardization (ECSS): Space Engineering – Ground systems and operations - Telemetry and telecommand packet utilization (2003). (ECSS-E-70-41A)
8. Ibrahim, S.K., Ahmed, A., Eldin Zeidan, M.A., Ziedan, I.: Machine learning methods for spacecraft telemetry mining. *IEEE Trans. Aerospace Electron. Syst.* **55**, 1816–1827 (2018)
9. Fortescue, P., Swinerd, G., Stark, J.: *Spacecraft Systems Engineering* (2011)
10. Garcia, P.A.: *Satellite Ground System Architecture* (2006)
11. Gremyachikh, L., et al.: *Space Navigator: A Tool for the Optimization of Collision Avoidance Maneuvers* (2019)
12. Han, J., Kamber, M., Pei, J.: *Data Mining: Concepts and Techniques*, 3rd edn. Morgan Kaufmann Publ. Inc., San Francisco (2011)
13. Hassanien, A.E., Darwish, A., El-Askary, H. (eds.): *Machine Learning and Data Mining in Aerospace Technology. Studies in Computational Intelligence*. Springer, Heidelberg (2019). <https://doi.org/10.1007/978-3-030-20212-5>
14. ISO 13537:2010: Space data and information transfer systems—reference architecture for space data systems (2015)
15. Lee, B.S., Hwang, Y., Kim, D.W., Kim, W.G., Lee, J.: Machine learning approach to initial orbit determination of unknown LEO satellites. In: 15th International Conference on Space Operations (2018)
16. Marz, N., Warren, J.: *Big Data: Principles and Best Practices of Scalable Realtime Data Systems*. Manning Publications Co. (2015)
17. Mital, R., Cates, K., Coughlin, J., Ganji, G.: A machine learning approach to modeling satellite behavior. In: *Proceedings - 2019 IEEE International Conference on Space Mission Challenges for Information Technology (SMC-IT 2019)* (2019)
18. Pääkkönen, P., Pakkala, D.: Reference architecture and classification of technologies, products and services for big data systems. *Big Data Res.* **2**(4), 166–186 (2015)
19. Tekinerdoğan, B., Babur, Ö., Cleophas, L., van den Brand, M., Akşit, M.: *Model Management and Analytics for Large Scale Systems*. Elsevier, Amsterdam (2020). <https://doi.org/10.1016/C2018-0-00106-1>
20. Tekinerdoğan, B., Bilir, S., Abatlevi, C.: Integrating platform selection rules in the model driven architecture approach. In: Aßmann, U., Aksit, M., Rensink, A. (eds.) *MDAFA 2003-2004*. LNCS, vol. 3599, pp. 159–173. Springer, Heidelberg (2005). [https://doi.org/10.1007/11538097\\_11](https://doi.org/10.1007/11538097_11)
21. Yairi, T., Kawahara, Y., Fujimaki, R., Sato, Y., Machida, K.: *Telemetry-mining: a machine learning approach to anomaly detection and fault diagnosis for space systems* (2006)
22. Wang, Y., Ma, Z., Yang, Y., Wang, Z., Tang, L.: A new spacecraft attitude stabilization mechanism using deep reinforcement learning method. In: 8th European Conference for Aeronautics and Space Sciences (EUCASS) (2019)