ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 11, Issue, 11, pp. 51760-51764, November, 2021
https://doi.org/10.37118/ijdr.23273.11.2021

RESEARCH ARTICLE

OPEN ACCESS

INTEGRATED RISK MANAGEMENT USING ARTIFICIAL INTELLIGENCE IN AN ELECTRIC POWER TRANSMISSION ENTERPRISE

Roger Sousa Fiusa[1,*], Starch Melo de Souza[2], Hemir da Cunha Santiago[2], Michel Mozinho dos Santos[2], Pedro Henrique Silva Prado[3], Nathalia Melo Lima[2], Kelly Pereira de Lima[2] and Lorrany Fernanda Lopes da Silva[1]

[1]Integração Transmissora de Energia S.A. (INTESA)
[2]In Forma Software
[3]AVSystemGEO

ARTICLE INFO

Article History:
Received 25th August, 2021
Received in revised form
24th September, 2021
Accepted 19th October, 2021
Published online 28th November, 2021

Key Words:
Transmission Lines, Risk Management, Machine Learning, Artificial Intelligence.

Corresponding author:
Roger Sousa Fiusa

ABSTRACT

This work presents a computational tool for Integrated Risk Management in a transmission enterprise, through data integration and analysis using Artificial Intelligence (AI), statistical and Business Intelligence (BI) methods. The life cycle phases of transmission projects defined for this work are: (1) pre-auction, preparation for concession dispute; (2) feasibility, engineering and environmental licensing studies; (3) implementation, of construction of the transmission line and substations, with execution of environmental programs; 4) operation and maintenance, for the entry of the enterprise into commercial operation. The areas of interest defined for Risk Management are: (a) Environment; (b) Land ownership; (c) Implementation Engineering (Construction); (d) O&M Engineering (Operation and Maintenance); (e) Regulatory; (f) Relationship with Investors/New Businesses. The objectives of this work are: to identify possible data sources that can be associated with transmission project risks; and analyze, through the application of several methods, how these data are correlated with each other. For this, an optimized analysis of the main causes of delays in the delivery of projects, cost increases and interruption of energy transmission is carried out and, through surveys carried out by professionals from each of the defined areas of interest, previous causes that may be mapped occur throughout the life cycle of projects. Analyzes are presented in the form of dashboards, with the aim of helping managers to make more assertive decisions. Another objective is to identify unresolved issues during the implementation phase, preventing them from having any impact on the operation phase.

Citation: ***Roger Sousa Fiusa, Starch Melo de Souza, Hemir da Cunha Santiago, Michel Mozinho dos Santos et al.*** "Integrated Risk Management using Artificial Intelligence in an Electric Power Transmission Enterprise*", International Journal of Development Research*, 11, (11), 51760-51764.

NTRODUCTION

The term project risk is defined as an "uncertain event or condition which, if it occurs, will have a positive or negative effect on one or more project objectives". Thus, risks are composed of probability (uncertain event or condition) and degree of impact (effect). According to (Zampolli, 2012), risk can be defined in the broadest sense as the "uncertainty of the future". Thus, risk has two basic components which are frequency and severity. It is noteworthy that the author also reports the importance of risk management for proactive asset management and that it is an integral part of the entire process.

According to (Zampolli, 2015), there is a specific need to have knowledge of the risk identification and monitoring process, not only taking into account the actions of the current legislation, but in a practical way that enables the optimization and prioritization of actions based on costs, risks and performance. Furthermore, in an organizational context, integrated project risk management provides for the identification and control of these risks in various areas and levels of the organization. Some risks, identified at lower levels, will be delegated to the project team for management and others can be transferred to higher levels if they can be better managed externally. A coordinated approach to enterprise-wide risk management ensures alignment and consistency.

This approach creates risk efficiencies in the structure of programs and portfolios, providing the greatest overall value for a given level of risk exposure for a project. In this context, it is important that the risk analysis is done individually by calculating the probability and impact for each risk, in addition to the integration of the results of all individual risks seeking aggregated project results or segmented by areas, project phases, region etc. Thus, the work has as general objective to present a dashboard for the integrated risk management in a transmission enterprise, using Artificial Intelligence (AI), Statistics and Rule Induction algorithms techniques applied to a Data Lake obtained from an integration middleware of data sources (structured and/or unstructured). This paper is organized as follows: Section II presents the materials and methods from the development methodology. Section III presents the results obtained using the dashboards for risk management. Finally, Section V presents this work's conclusions.

MATERIALS AND METHODS

In this section, we present the guidelines for the development methodology: 1. Understanding the business through flowcharts and identification of risks; 2. Definition and treatment of data sources; 3. Creation of the system architecture; 4. Data modeling to calculate risk probability and impact; 5. Aggregation of the results of the risk probability and impact calculations; and 6. Creation of query dashboards.

Elaboration of Flowcharts and Risk Identification: Process flowcharts for all areas of interest were designed to facilitate understanding of the context, relating the causes and impacts of risks. Thus, the objectives for the elaboration of these flowcharts were: 1) To support the identification of risks (causes and impacts) in each phase and area of interest; 2) Map data sources for each of the identified risks; 3) Facilitate the understanding of how risks from different areas of interest and phases relate to each other. The risks and causes were registered in a specific database being organized and coded. An example flowchart for the areas of Deployment Engineering and O&M Engineering is shown in Section 3.

Definition and Treatment of Data Sources: The great diversity of data sources surveyed, for example, data referring to soil types (National Spatial Data Infrastructure), history of occurrence of fires (National Institute for Space Research) etc., are related to different areas of interest for the project. Figure 1 below illustrates some examples of the fonts used.

Figure 1. Examples of data sources used (Authors).

Once the data sources were defined, the study was carried out to find out if the data were structured, semi-structured and unstructured. Structured data is defined as arranged in a fixed tabular structure, that is, in the form of rows and columns so that the data is predictable. The unstructured ones, on the other hand, have greater flexibility and dynamics, such as free text, without imposing a form of organization, which generates unpredictability and greater difficulty for processing. And finally, semi-structured data generally has a fixed overall organization, but that allows flexibility in data fields, such as a tabular database that allows the column to contain free text.

System Architecture Construction: The software architecture describes at a high level of abstraction the components and data flow, serving for communication and planning, aiming at the development and implementation of the system. The system architecture can be seen in Figure 2. Among the components of the proposed architecture, the following stand out: 1) From databases collected from internal and external sources, the Middleware is the module responsible for performing the extraction, transformation and loading of data into a common and standardized repository called Data Lake; 2) Data Preparation is the component responsible for performing various types of processing such as data integration, statistical analysis and artificial intelligence, taking the results to a data repository called Data Warehouse from where they will be consumed by the Business Intelligence (BI) software for later presentation in dashboards; 3)

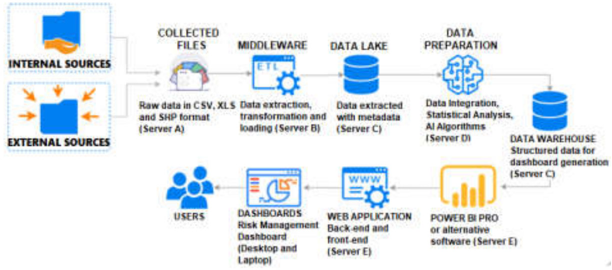

Figure 2. System Architecture (Authors)

On the other hand, the Web Application module represents the interface through which the user has access to functionalities such as registration of transmission projects, consultation of risk flowcharts, data reports from systems involved in the transition between the implementation phase and the operation phase, in addition to enabling the monitoring of risks through alerts, checklists and dashboards with risk indicators.

Data Modeling for Probability and Impact Calculation: Due to the number of identified risks and the diversity of types of data sources surveyed, and also, aiming to define a methodological structure that could standardize the way of calculating the probability and individual impacts of each risk, but already foreseeing the form of aggregation and segmentation Based on the results, a general methodology was created, using the calculation methods considered more viable for each type of data source related to the identified risks, namely: Clustering, Classification, Variable Correlation, Natural Language Processing (NLP), Multicriteria Heat Maps, Statistics and, finally, the Expert Manual Input.

Clustering: The method consists of grouping data that have characteristics in common, in such a way that the appropriate probabilities can be attributed to these formed groups. The clustering algorithm used in this work was the K-means algorithm (Likas, 2003), this algorithm applies an unsupervised machine learning method that searches for patterns among the data without the need to provide a reference for such, only the parameter K is provided to define the amount of clusterings.

Classification: Classification is one of the Machine Learning problem categories. The purpose of these algorithms is to learn a general rule of thumb with data inputs for mapping to other outputs correctly. Thus, the data is analyzed and categorized so that the mapping points to defined outputs according to each category. The classification technique used to define the probabilities of some risks was the Decision Tree (Brijain, 2014), this tree is the result of a data mapping among a series of related choices. Based on this mapping sequence and a data entry, it was possible to define the probability of occurrence of a certain risk.

Variable Correlation: The correlation of variables (Gogtay, 2017) was used to analyze data sources related to ANEEL, where BI (Business Intelligence) tools were used. Data were grouped by technical group (Environment, Engineering and Regulatory) and by region of Brazil. Through the situation of all events of all concession contracts, it was possible to define the historical probability of a certain event being delayed or not.

Natural Language Processing (NLP): NLP (Natural Language Processing) (Kang, 2020) was used for data sources whose information is semi-structured or unstructured. In this method, several terms formed by grouping words were extracted and, using textual semantic similarity, the terms generated were crossed with another group of terms related to the causes of risks that were defined by the specialists. The result originated a new set of terms, which were classified as relevant terms and then associations were made between the data source and the risk(s). From this association, the probability of the risk occurring was calculated by mathematical equations.

Multicriteria Heat Maps: The Multicriteria Heat Map was the method used to analyze risks related to the definition of the route with the least possible impact for a Transmission Line. The analysis was possible with the use of programs with Geographic Information System (GIS). Each of the data sources were grouped into areas, and weights were assigned to them. In addition, weights were assigned to each of the parameters (information) found within each of the data sources. The definition of the best spatial analysis technique was defined as a function of the parameters found in each source: Manual Weighting, Euclidean Distance, Slope, Kernel Density, Hot Spot Analysis, Topo to Raster, IDW (inverse distance weighting), Weighted Sum and Weighted Average.

Statics: Some statistical methods were used for data analysis, including: the use of separator measures, distribution modeling, and survival analysis. When considering the separator measures, there is the conception that there is a separation of the data ordered in some ways, such as, for example, the quartiles that divide the data into four equal parts. Thus, data can be grouped (characterized) according to these measures. The distribution modeling technique consists in the distribution adjustment to characterize the data from a probability density distribution (Gamma, Exponential, Weibull and Normal), with this it is possible to observe the data in terms of probabilities.
Survival analysis is a branch of statistics, generally used when time is the object of interest, as in the case of equipment aging or failure. The technique studies the time until the occurrence of an event or the risk of an event occurring per unit of time. In the case of this work, the exponential distribution was used to characterize the survival time.

Expert Manual Input: For those risks where no databases for modeling the calculations were found or even where they were found, but no feasibility for use was identified, the solution to maintain these risks in the analysis was the manual insertion of probability and impact on a scale predefined by the risk expert user.

Aggregation of the Results of the Probability and Impact of Risk Calculations: In real-world applications, there are situations where it is unreasonable to assume that a good alternative is one that simultaneously meets all the criteria. The decision maker must consider how restrictive each of the objective functions is relevant, and evaluate the consequences of their choices. The OWA operator - Ordered Weighted Averaging - was developed by Yager (Yager, 1988) in order to enable intermediate solutions, which are not completely restrictive or additive. The nature of the OWA procedure depends on some parameters, which can be specified by fuzzy quantifiers (fuzzy logic). The operator is composed of several mathematical equations that aim at segmenting risks by areas of interest, project phases and geographic region. It is also possible to have aggregation at the project and portfolio level.

Creating query dashboards: To develop the query dashboards, two information management tools were used: the Management Information System (SIG) (Marques, 2007) and the BI (Rud, 2009). The dashboards were all concentrated in a website application, being formed by graphs, maps and tables. The type of dashboard prepared for each risk was defined by the information present in the databases as well as the results of the modeling obtained through this information. The website also included the inclusion of all elaborated flowcharts, as well as specific dashboards created to identify unresolved issues during the implementation phase.

RESULTS

Risk Identification and Flowcharts: Extensive documentation on transmission business risks and processes was produced throughout the project, containing for each risk a description, cause, primary and secondary impacts, among others. Subsequently, a total of 31 flowcharts were prepared, organized by themes and areas of interest: 21 for the Environment, 3 for the Deployment Engineering and O&M Engineering areas, 2 for O&M Engineering, 1 for Land, 1 for Regulatory and 1 for Investor Relations/New Business. It is noteworthy that the Environment had a greater number of flowcharts prepared because in this case a flowchart was chosen for each environmental program. Figure 3 illustrates an example of a flowchart for the areas of Implementation Engineering and O&M Engineering, with the theme “Implementation Completion”. The number of identified risks can be seen in Table 1.

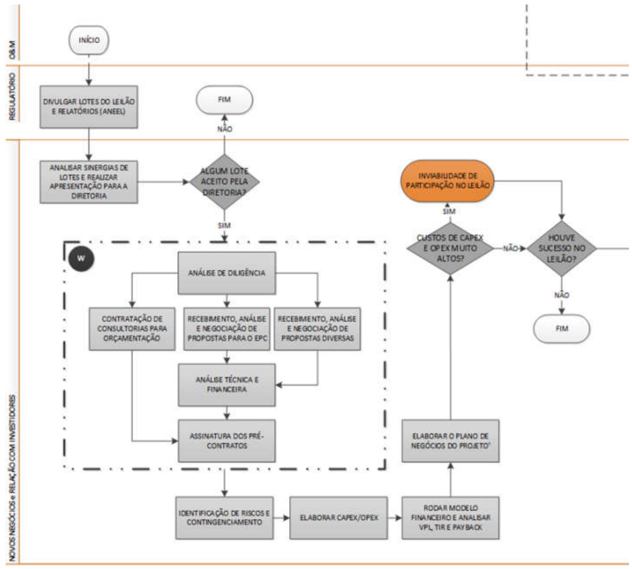

Figure 3. Flowchart involving the areas of Eng. Implementation and Eng. O&M. (Authors)

Table 1. Number of risks identified in the project (Authors)

Area / Phase	Pre-auction	Viability (Auction-LP)	Viability (LP-LI)
Environment	10	14	15
Landlord	5	3	0
Deployment Eng.	26	1	6
O&M Eng.	2	0	0
Regulatory	4	0	0
New Business	1	5	5
Total	48	23	26

Area / Phase	Implantation (LI-LO)	Operation	Total
Environment	36	2	77
Landlord	12	0	20
Deployment Eng.	31	0	64
O&M Eng.	1	28	31
Regulatory	5	11	20
New Business	6	4	21
Total	91	45	233

Query Dashboards: The website developed has a side menu for the user to navigate through the entire content of the risk management system. The main dashboard shows the periodic evolution of the Project/Portfolio Global Risk Degree, as well as the Risk Degree segmented by areas of interest and project phases, as can be seen in Figure 4.

Figure 4. Main dashboard (Authors)

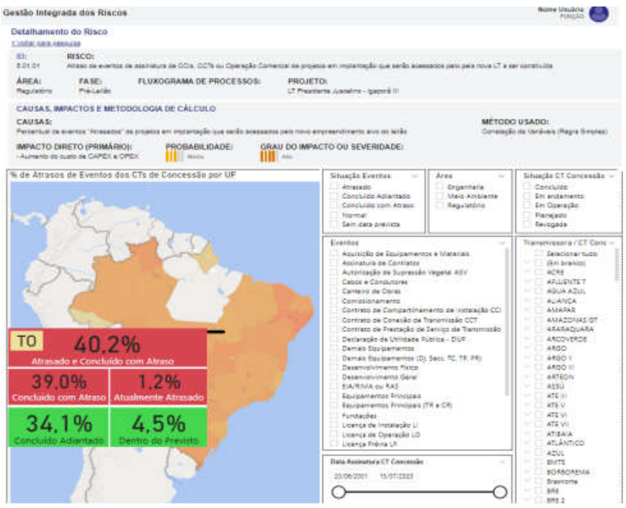

Figure 5. Example of delay in concession contract events (Authors)

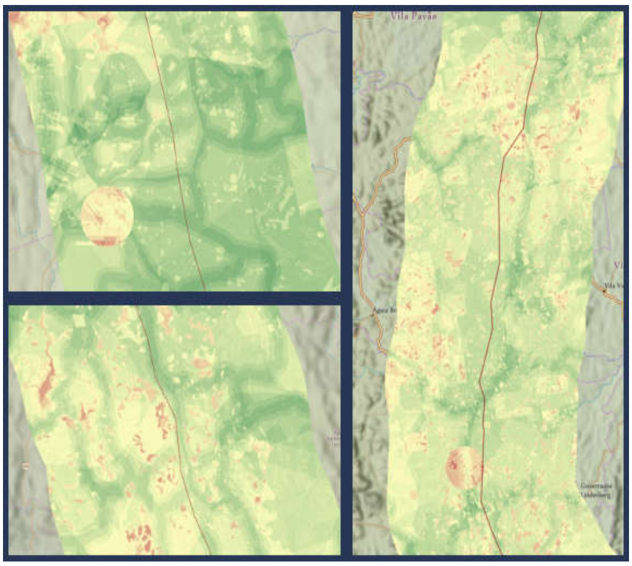

Figure 6. Result of the integrated analysis of the Multicriteria Heat Maps (Authors)

By accessing the Risks menu, the user has access to a list of all the risks that have been identified, the probabilities and impacts calculated by the modeling, the probabilities and impacts that can be edited by the risk specialist user, the result of the Degree of Risk, in addition to the status, phase, project and also a link to access the individual risk analysis dashboard. Accessing the links for each risk, the system opens a new window with various information from the risk record, such as causes, impacts, modeling method used, flowchart, among others. Figure 5 illustrates an example of the risk related to delays in events of Concession Contracts of projects that will be accessed by the TL (Transmission Line) project that is being studied/implemented by the transmission company.

The dashboards related to the Multicriteria Heat Maps (see Figure 6) allow a study of an ideal layout of a TL, as it overlays various spatial information, processes the data and provides an answer on higher risk regions considering attributes of the physical, biotic environment and anthropogenic, such as greater declivity, susceptibility to erosion, native vegetation, presence of traditional communities and other relevant factors for the route to avoid regions where there are risks of extra costs with obtaining licenses and possible delays in the work schedule. Another gain is to provide greater support for the determination of technicians and engineers to go into the field to evaluate the outlined alternatives. Unresolved issues in the Land and Environment areas of interest in the implementation phase were also worked on in specific dashboards. In addition to the managerial nature of these dashboards, an informative character was also sought, since the operation teams generally do not have knowledge of the land and environmental process during the implementation phase. In the Land area, we sought to present the location of properties and various information on the owners depending on the status of the negotiation. For the Environment area, the information is related to archeology, installation of anti-collision flags for avifauna, speleology, fauna, germplasm, paleontology, quilombola communities, social communication, erosion processes, Projects for the Recovery of Degraded Areas (PRDAs), and conditions of environmental licenses.

CONCLUSION

This work presents a software system for Integrated Risk Management in transmission line projects, as well as the mapping of risks and processes that were necessary for the development of the system. Starting from activities that sought to understand the transmission business for 6 areas of interest and over 4 phases, it was obtained: 233 identified risks including details such as cause and impact for each risk; 31 flowcharts of mapped risk processes; a computer system for data processing containing a Web Application to monitor risks through dashboards. The software system developed and the information mapped on risks must contribute to the management of transmission businesses by making available multidisciplinary analytical tools.

ACKNOWLEDGMENT

The authors acknowledge the financial support of Integração Transmissora de Energia S.A. (INTESA) through the ANEEL R&D Program for the development of the research project entitled: "PD-05456-0001/2019: Integrated risk management applied to the value chain of transmission projects using a data interconnection bus and artificial intelligence".

REFERENCES

Brijain, M., Patel, R., Kushik, M. R., & Rana, K. (2014). A survey on decision tree algorithm for classification.

Gogtay, N. J., & Thatte, U. M. (2017). Principles of correlation analysis. Journal of the Association of Physicians of India, 65(3), 78-81.

Kang, Y., Cai, Z., Tan, C. W., Huang, Q., & Liu, H. (2020). Natural language processing (NLP) in management research: A literature review. Journal of Management Analytics, 7(2), 139-172.

Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global k-means clustering algorithm. Pattern recognition, 36(2), 451-461.

Marques, W. L. (2007). Sistema de informações gerenciais. Clube de Autores (managed).

Rud, O. P. (2009). Business intelligence success factors: tools for aligning your business in the global economy (Vol. 18). John Wiley & Sons.

Yager, R. R. (1988). On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Transactions on systems, Man, and Cybernetics, 18(1), 183-190.

Zampolli, M. Guia básico para implantação da gestão de ativos em empresas de energia. International Copper Association Latin America. 2012

Zampolli, M. Gestão de Ativos - Guia para aplicação da Norma ABNT ISO 55001. International Copper Association Latin America. 2015.

****