Integrating XML Technology with Object-Relational Databases into Decision Support Systems

[1]Iuliana BOTHA, [2]Adela BÂRA, [3]Simona-Vasilica OPREA, [4]Ticiano COSTA JORDÃO
[1,2]Academy of Economic Studies, Bucharest
[3]Transelectrica Company, Bucharest
[4]University of Pardubice, Czech Republic
iuliana.botha@ie.ase.ro, bara.adela@ie.ase.ro, simona.oprea@transelectrica.ro, ticiano.costajordao@upce.cz

This paper presents some informatics technologies, like XML and object-relational databases, and the main motivations for implementing them in Decision Support Systems (DSS). Also, it is proposed and analyzed a conceptual model of a DSS prototype, which can be applied in the uncertain and unpredictable environments, like the production and the prediction of the wind energy.
Keywords: *Object-Relational Databases, XML, Decision Support Systems, Wind Energy, Power Plants*

1 Introduction

Early this century was marked by a series of major changes in the field of information and communication technology (ICT), but also in the field of economics, leading to the creation and development of a new type of economy, generally named most often, the new economy.

The role of informatics in the new economy, as producer and consumer of information, is crucial. All economic activity is strongly influenced by the evolution of the informatics context and vice versa.

Under these facts, it is appropriate the research of decision-making solutions development, built with advanced technologies, such as relational, object oriented and XML.

Decision support systems (DSS) form a distinct class of computer systems. It integrates specific IT tools to assist decision and general purpose tools to form a constituent part of the organization's overall information system. Any DSS is designed to provide IT support required to minimize the effects of restrictions (cognitive, communication, time, etc.) with which is facing the human decision maker during his activities. Decision support systems have become a real presence in the IT world. Web technologies, advanced databases and methods for accessing them, flexible user interfaces, are all pressing needs of IT technology. Therefore, today's decision support systems, in order to be efficient, integrate many of these technologies. The DSS proposed in this paper will use spatial, multimedia and XML data stored in object-relational databases.

2. Considerations about object-relational databases

The attempts to model the complex elements of the economic activities have caused the reorientation to the object-oriented technology. Thus, in the late 80s has appeared the third generation of databases, namely the object-oriented ones.

As stated in [2], the development of object-oriented data model was due to limitations of the relational model to handle efficiently massive amounts of data, with high complexity, found in new types of applications (multimedia, Internet, Geographic Information Systems, etc.).

Although object-oriented databases appear to meet the needs of applications required by the new economy, the market they use remains relatively low, often the invoked reason being the difficulty to query data and the high consumption of computing resources [10].

Currently, the solution that combines the advantages of relational and object-oriented data models, comes with a hybrid data model, the object-relational one, which involves object-oriented facilities (especially the fundamental characteristics of objects: encapsulation, inheritance, polymorphism) as extensions of relational model.

The object-relational data model was designed as result of the research in the 90s, by extending the concepts of relational databases with object properties. In the center of the new model architecture, the researchers decided to keep the declarative query language based on predicate calculation.

Thereby, the object-relational databases (ORDB) are the result of applying features of object-oriented technology in data storage and retrieval. Unlike object-oriented databases, the object-relational ones provide the ability to represent complex data structures through objects, while retaining the advantages of relational databases.

Broadly, we can define an object-relational database as a set of tables of objects, linked logically, organized in the external memory and accessible to multiple users in a timely manner.

The main features of object-relational databases are given by the flexibility and the structure of the used data model and refer mainly to the following:

- Treating consistently the entities, by modeling both the properties of objects and their behavior;
- Providing communication between data and among programs;
- Simplifying the data structure that provides ease of use and high portability of the resulting systems;
- Modeling and flexible storing the real world objects, which helps to address the complex areas and to use different types of data;
- Integrating the structural and behavioral description of the entities, which gives a dynamic aspect to the object;
- Inheriting the properties of object types, which provide the opportunity to realize complex operations and processing over the data.

Concepts used in object-relational databases are borrowed from the object-oriented terminology, but also from the relational one, and they were presented at defining the structure of the object-relational data model. Figure 1 illustrates the combination of these concepts in order to develop the object-relational databases.

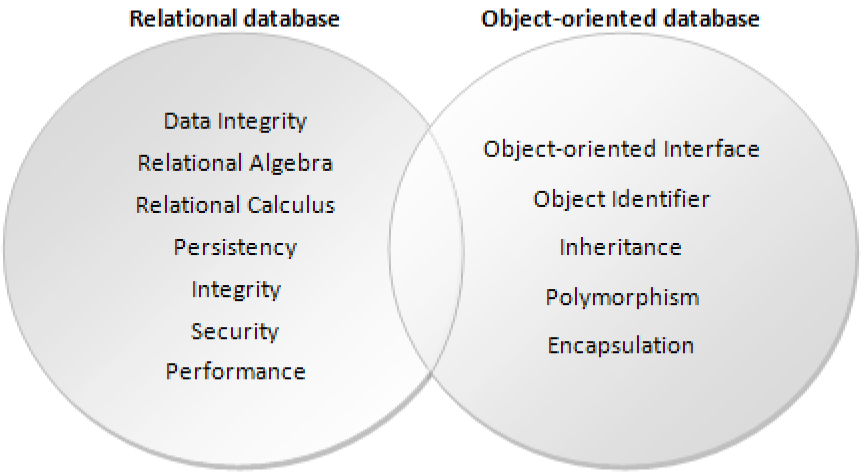

Fig. 1. The main concepts of the object-relational databases
(Source: Authors)

Defining the type of objects represents the mechanism for specifying the database schema. Type definitions include inheritance relationships (which generate supertypes and subtypes) and structural relationships between them.

A complete schema of the object-relational database consists of relational tables and object tables. The object tables can present inheritance relationships between types of objects that were the basis for creating these tables.

Can be identified two types of changes in an object-relational database schema. Thus, we can discuss about changes on how to define an object type, which can lead to changes in the specification of attributes and methods defined (e.g. changing the domain or the name of an attribute; adding or deleting an attribute or a method). Updates in the schema can also appear in the case of some changes in the hierarchy of types or in the supertype/subtype relationships. However, this version of modifying the schema is limited, depending on manufacturer.

According to [11], a fully object-relational database must necessarily meet the following four characteristics:

- Allowing the creation of extensions from basic types. Refers to the ability to create extensions of data types by defining abstract, simple or complex, data types;
- Enabling management of complex objects. There is a major difference between the ability to manipulate complex data using an object-oriented database and using relational database. The set of basic types of relational databases is very poor compared to the existing set of object-oriented databases. In this respect, [11] states that a database, to be considered object-relational should support a rich collection of complex data types, including mandatory: complex types, SET constructors, RECORD constructors, REF constructors;
- Accepting the inheritance property. In an object-relational database, a table is seen as a container used to hold instances of a type that can use the inheritance feature. This can refer both to the data and methods inheritance. Data inheritance applies only to the data types, but the property also can be propagated upon the tables built over these data types. Instead, methods inheritance apply to the user-defined functions and methods included in the types of objects. Also, according to [11], a fully object-relational database must support multiple inheritance, required in many applications;
- Using a system of rules. In a fully object-relational database should not miss a system of rules and triggers, necessary especially to ensure consistency of the database.

In conclusion, based on the object-relational data model, we can define an object-relational database as a collection of tables, relational or object, persistent, stored in the external memory through a system of rules, consistently organized, ordered in hierarchies and providing shared access to competing users. As a hybrid, the object-relational databases combine flexibility, scalability and security in using existing relational systems with object-oriented properties such as: abstraction, encapsulation, inheritance and polymorphism.

3. Integration of XML technology into object-relational databases

eXtensible Markup Language (XML) was designed as a standard for information exchange over the Internet. In [13], W3C offers an abbreviated definition for XML: „XML is a markup language for structuring arbitrary data. XML documents are made up of storage units called entities, which contain either parsed or unparsed data. Parsed data is made up of characters, some of which form character data, and some of which form markup. Markup encodes a description of the document's storage layout and logical structure. XML provides a mechanism to impose constraints on the storage layout and logical structure."

Databases and XML offer complementary functionality for storing data. While databases store data, ensuring their rapid retrieval, XML enables easy exchange of information, which allows interoperability between applications due to data encapsulation with metadata.

XML allows designers to create their own tags, allowing the definition, transmission, validation, and interpretation of data between applications and between organizations. In Romania there are several systems based on XML standard (e.g. Labour Inspection, Bucharest Stock Exchange) [9].

Can be identified two approaches in order to store XML data: we can use native XML databases, or we can map XML data and queries into a relational or object-relational database [6]. The advantage of using object-relational databases is that we can get the benefits of both relational and object-oriented technologies, while the disadvantage translates into lower performance due to XML data mapping to the relational data, which can produce a database schema with many relations.

When used in object-relational databases, XML data must be mapped into relations. In order to transfer the data between XML documents and object-relational structures are used specific mapping methods. The study [8] makes a basic classification of these mapping methods, as follows:

- *Generic methods* – that are not using any schema of stored XML documents;
- *Schema-driven methods* - that are based on existing schema of stored XML documents;
- *User-defined methods* - that are based on user-defined mapping.

The mapping process is represented in Figure 2 and detailed in [6]:

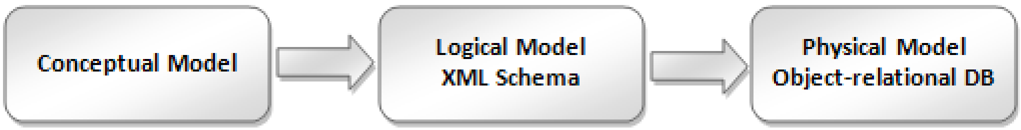

Fig. 2. The mapping process
Source: [6]

An algorithm for mapping XML schema to object-relational database schema is proposed in [7]. Also, in [5] are presented mapping algorithms and correspondences between object-relational databases and XML schema for the description and validation of the data. The correspondence is observed in reverse and includes in addition to abstract data types (those defined by the user, which corresponds to XML schema elements), persistent objects, specific to databases, in general.

Starting from [5], in table 1 we have completed the set of mapping rules, as follows:

Table 1. Mapping rules: ORDB – XML Schema
Source: adapted from [5] and completed

Object-relational database element	XML Schema element
Table	Complex type whose sequence contains elements corresponding to the table columns, including the constraints

Simple attribute, without constraints	Element of the complex type sequence, which is represented through a simple type
Attribute with primary key constraint	Attribute of the complex type having the property *use="required"*
Attribute with not null constraint	Element of the complex type sequence having the property *nullable="false"*
Attribute with uniqueness constraint	Element of the complex type sequence, with the attribute *xs:unique*
Attribute with referential constraint	Element of the complex type sequence, with the attribute *xs:keyRef*
User defined object type	Complex type having the elements corresponding to the attributes of the object type
REF type attribute	Reference specified directly as a property of the XML schema element *ref="pointer"*
ROW type attribute	Complex type whose elements represent attributes of type ROW
Collection type attribute	Complex type which contains a sequence of the element associated with the collection type, that has the property *maxOccurs* equal with the collection dimension
Inheritance	Complex type having complex content and the type *extension*

The DSS that we propose will use XML documents containing data which may influence wind park performances, like meteorological data. These XML data need to be stored in object-relational databases, in order to realize further analysis, necessary for decision making.

4. Conceptual model for a decision support system with object-relational databases implementation

Organization management involves different types of activities and therefore requires different types of information. For better manage, transform, process and analyze this information, it is needed a decision support system (DSS), seen as "an information system that draws on transaction processing systems and interacts with the other parts of the overall information system to support the decision-making activities of managers and other knowledge workers in organizations" [4].

For an efficient management of resources it is necessary to be able to achieve a prediction of energy produced with minimal errors.

Currently, as stated in [10], both nationally and internationally, there are various types of frameworks, architectures, solutions and systems to provide economic assistance for decision making processes and production environments with a relatively high degree of certainty.

However, as shown in [1], the existence of such a system to assist decision making at national level is necessary because currently there are no solutions to offer efficient predictions and with minimal errors for the environments with low predictability, such as wind energy, which depends exclusively by natural factors.

As shown in [12], in Europe there are some countries that have been developed decision support systems in the energy system for efficient management of wind resources (e.g. Germany, Spain), but the costs of building these systems is very high and the particularities of national energy potential makes difficult the application of these methods in Romania.

These issues concerning the impossibility of prediction accuracy, the precarious data integration from various local equipment and systems and the inefficient analysis of energy resources, lead to the need to develop solutions for a better prediction of energy produced, but also an integrated system, with a component for assisting decision making in this area.

The data sources provided for the decision-making process must be as varied and require integration. Some of the data collected within the operational activity will be stored in the object-relational database as: spatial data, multimedia data, Large Objects, user-defined object types. The spatial data are used in order to representing them on interactive maps monitoring the production of energy, multimedia data and LOB for using and manipulating large objects.

Also, since data integration from various sources is mandatory, we will use XML as a standard for integration and data warehouses for centralize data from heterogeneous data sources.

The decision support system will use data that can influence wind park performances, like meteorological data or data resulted from national energy production regular monitoring. The data will be received like XML documents and after validating them with XML Schemas, they will be stored in object-relational databases, in order to realize further analysis, necessary for decision making.

In the following paragraphs, we propose an architectural model for such a decision support system based on the typical architecture of a DSS, but with specific components of information requirements identified in the National Energy System (NES).

In [3] are identified four components of a decision support system: 1) interface; 2) database system; 3) system of analytical, mathematical and statistical models; 4) component to ensure communication. This architecture is implemented in most Business Intelligence solutions made by major manufacturers.

Starting from these standard DSS architecture and also based on the information requirements identified by the National Energy System, we have proposed the conceptual model of the informatics system. Thus, as stated in [10], the conceptual model contains the four levels presented above (Figure 3).

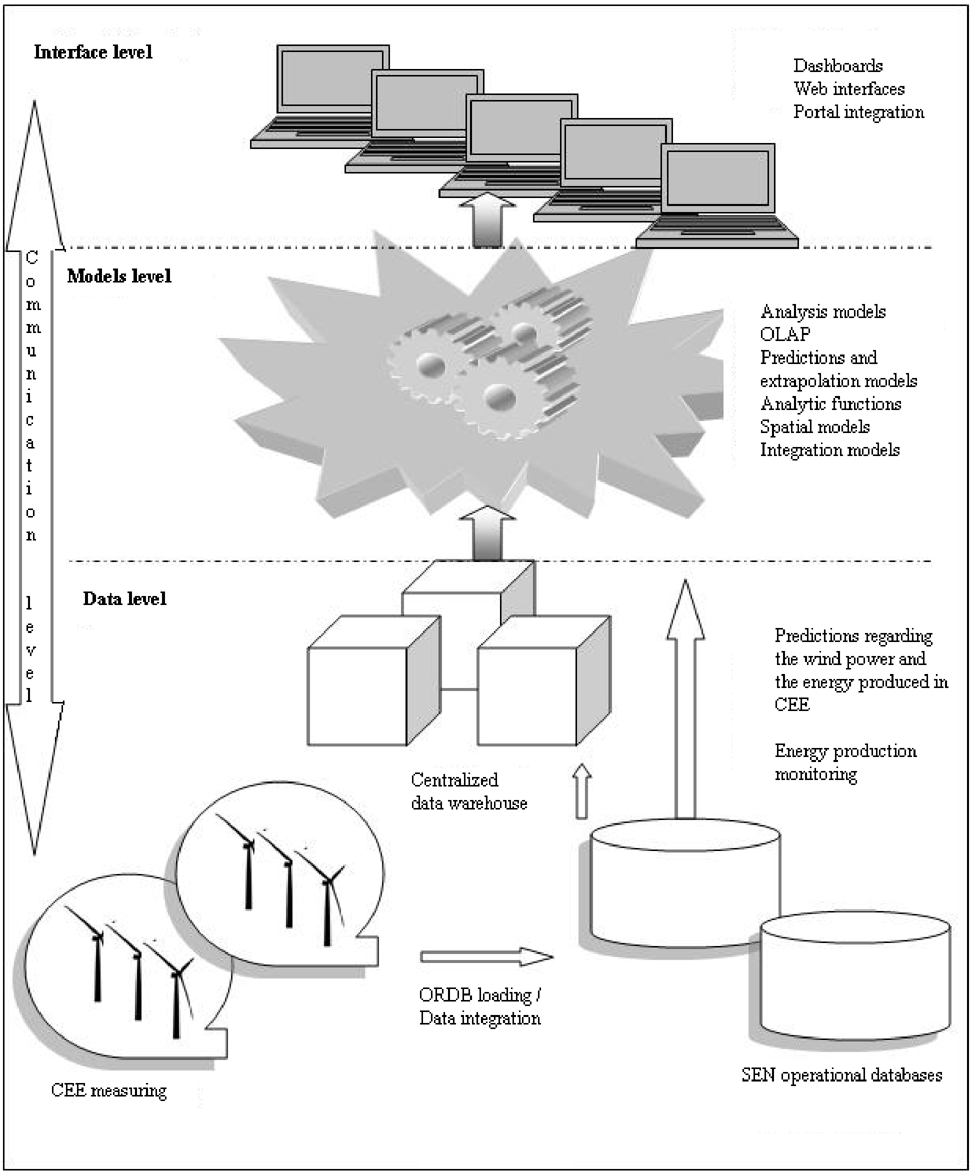

Fig. 3. The architecture of the proposed DSS
Source: [10]

The proposed DSS architecture can be seen in terms of achievement levels, from the bottom to the top of the pyramid, on three levels: the bottom-tier, middle tier and top-tier, being interconnected at the level of telecommunications.

1) The data level will consist of operational data sources, as well as spatial data and the data warehouse to be built to facilitate multi-dimensional analysis.

Some of the data collected on operational activity will be stored as data in a spatial DBMS (Database Management System Data), in order to represent them on maps for interactive monitoring of energy and resources.

The data warehouse will be a centralized repository, organizational, composed of data marts for each type of activity which is analyzed: operational (energy production) and financial. To load data warehouse from data sources is necessary to undertake a process of Extract, Transform and Load (ETL). This process will automatically run at regular intervals depending on the technical requirements of implementation.

2) The models level will include models for the current financial analysis and

forecasting of these activities, models for spatial data representation, and simulation models, extrapolation, prediction and analysis of energy production from wind sources. To achieve models will use new technologies for Business Intelligence (BI), such as OLAP, data mining, predictive analytic functions and algorithms.

3) The interface level will contain the data presentation elements and specific analysis and reporting for decision support systems. The system will be accessible via a Web interface so users can access reports without having to install client applications. It will choose to integrate all elements into a portal interface that allows a single, uniform type of authentication: Single Sign - On. The portal will integrate elements of the subsystems and previous levels, such as OLAP analysis, prediction and monitoring modules, reporting.

4) The level of telecommunications will provide support for access the decision support system both within and outside NES, using mobile devices. For this level will use the existing network components in the NES, where the access is authorized according to the role and access rights of each user.

Following analysis of the data sources provided by the wind power plant units, we found data heterogeneity, mainly due to the diversity of wind turbines and measuring instruments, and computer applications to process data. For this reason it is essential to use at data level techniques of data migration and data integration. In this manner, data can be loaded in a consistent manner in a centralized operational database which can be later used for prediction, simulation and analysis.

Conclusions and future work

The integration of the special data (like multimedia, LOB, XML, spatial data) into object-relational databases is a necessary characteristic for today's enterprises that maintain informatics systems in the unpredictable environments.

The paper provides a synthetic look at object-relational database features and possibilities for integration with XML technology. Also we have presented an implementation perspective of a DSS, which requires decision making in the wind energy prediction and production.

Our research will continue with the development and implementation of the conceptual model proposed in this paper. For this purpose, will be used specific technologies for each level of the model architecture. Thus, the data level will use solutions for organizing and integrating data, the models level will use solutions of multidimensional analysis, forecasting models, simulation and extrapolation, and for the interface level will be used solutions for data analysis and dynamic presentation of data.

Acknowledgment

This paper presents some results of the research project PN II, TE Program, Code 332: "Informatics Solutions for decision making support in the uncertain and unpredictable environments in order to integrate them within a Grid network", financed within the framework of People research program.

References

[1] Adela Bâra, Ion Lungu, Manole Velicanu, Simona Vasilica Oprea, "Intelligent Systems for Predicting and Analyzing Data in Power Grid Companies", *TheProceedings of the IEEE International Conf. on Information Society* (i-Society 2010) London.

[2] Iuliana Botha, AndaVelicanu, Adela Bâra, "Integrating Spatial Data with Object Relational-Databases", *Journal of Database Systems*, nr.1/2011, pp. 33-42, ISSN: 2069 – 3230

[3] Daniel Power, "Decision Support Systems: Concepts and Resources", Cedar Falls, 2000: DSSresources.com, http://dssresources.com/dssbook

[4] Ralph H. Sprague, Eric Carlson, “Building Effective Decision Support Systems”, Prentice-Hall, 1982, ISBN: 978-013-086-215-0
[5] Laila Alami Kasri, Noureddine Chenfour, “Model of Storage XML Database based on the Relational-Object Model”, *International Journal of Engineering Science and Technology*, Vol. 2(11), 2010, ISSN 0975–5462
[6] Eric Pardede, J.Wenny Rahayu, David Taniar, “On using Collection for Aggregation and Association Relationships in XML Object-Relational Storage”, *ACM Symposium on Applied Computing*, Cyprus, 2004, ISBN 1-58113-812-1
[7] Irena Mlynkova, Jaroslav Pokorny, “From XML Schema to Object-Relational Database – an XML Schema-driven mapping Algorithm”, *Proceedings of the 3rd IADIS International Conference WWW/Internet*, Madrid, Spain, 2004, pp 115 - 122, ISBN 972-99353-0-0.
[8] Irena Mlynkova, Jaroslav Pokorny, “XML in the World of (Object-) Relational Database Systems”, *Information Systems Development: Advances in Theory, Practice, and Education*, Vilnius, Lithuania, 2004, pp. 63 - 76, ISBN 978-0-387-25026-7.
[9] Vlad Diaconiţa, Iuliana Botha, “The Level of System Integration in Romania”, *Informatica Economica Journal*, Vol. XI, no. 3/2007, pp. 96-99, ISSN 1453-1305
[10] Iuliana Botha, Adela Bâra, Simona Vasilica Oprea, “An implementation perspective on object-relational databases in decision support systems”, *Proceedings of the Tenth International Conference on Informatics in Economy*, 2011, ISSN: 2247-1480
[11] Michael Stonebraker, Dorothy Moore, “Object-Relational DBMS - The Next Great Wave”, Morgan-Kaufmann, 1996, ISBN: 155-860-397-2
[12] Adela Bâra, Simona Oprea, Anda Velicanu, Vlad Diaconiţa, “Decision support systems - solutions for predicting, monitoring and integrating wind energy resources into the national grid companies”, *Development, Energy, Environment, Economics*, pp. 55-60, 2010, ISBN: 978-960-474-253-0
[13] http://www.w3.org/TR/xml/

Iuliana BOTHA is an Assistant Lecturer at the Economic Informatics Department at the Faculty of Cybernetics, Statistics and Economic Informatics from the Academy of Economic Studies of Bucharest. She has graduated the Faculty of Cybernetics, Statistics and Economic Informatics in 2006 and the Databases for Business Support master program organized by the Academy of Economic Studies of Bucharest in 2008. Currently, she is a PhD student in the field of Economic Informatics at the Academy of Economic Studies. She is co-author of 5 books, 12 published articles (2 articles ISI indexed and the other 10 included in international databases), 16 scientific papers published in conferences proceedings (among which 6 paper ISI indexed). She participated as team member in 4 research projects that have been financed from national research programs. From 2007 she is the scientific secretary of the master program *Databases for Business Support* and she is also a member of INFOREC professional association. Her scientific fields of interest include: Databases, Database Management Systems, Design of Economic Information Systems, Grid Computing, e-Learning Technologies.

Adela BÂRA is a Lecturer at the Economic Informatics Department at the Faculty of Cybernetics, Statistics and Economic Informatics from the Academy of Economic Studies of Bucharest. She has graduated the Faculty of Economic Cybernetics in 2002, holds a PhD diploma in Economics from 2007. She is the author of 7 books in the domain of economic informatics, over 40 published scientific papers and articles (among which over 20 articles are indexed in international databases, ISI proceedings, SCOPUS and 10 of them are ISI indexed). She participated as team member in 3 research projects and has gained as project manager one research contract, financed from national research programs. She is a member of INFOREC professional association. From May 2009, she is the director of the Oracle Excellence Centre in the university, responsible for the implementation of the Oracle Academy Initiative program. Domains of competence: Database systems, Data warehouses, OLAP and Business Intelligence, Executive Information Systems, Decision Support Systems, Data Mining.

Simona Vasilica OPREA is a Senior Engineer and Project Manager at Transelectrica National Power Grid Company. She has graduated the Polytechnic University in 2001, holds a Master Diploma in Infrastructure Management Program, Yokohama National University, Japan in 2007 and a PhD diploma from 2009. She is the author of over 20 articles, from which 3 ISI Web of Science indexed and 2 included in SCOPUS international database. Domains of competence: wind farm, investment opportunity analysis, studies of prognosis, stationary and dynamic regimes, short circuit calculations.

Ticiano COSTA JORDÃO is a Full Lecturer at the Faculty of Economics and Administration of the University of Pardubice (Czech Republic), director and co-founder of the Sustainability Crusade Observatory (CRUSUS), vice-president of the Czech-Brazilian Chamber of Commerce, and a member of the editorial board of the International Journal *Sustainable Development and Corporate Social Responsibility Review* (*SD&CSR*).
In order to know more about his professional work, please consult: www.crusus.org.