# A Case Study of Research Improvements in an Service Industry upgrading the Knowledge Base of the Information System and the Process Management: Data Flow Automation, Association Rules and Data Mining

Alessandro Massaro[1,*], Palo Lisco[1], Alessandro Lombardi[1], Angelo Galiano[1], Nicola Savino[1]

[1]Dyrecta Lab, IT Research Laboratory, Via Vescovo Simplicio, 45, 70014 Conversano (BA), Italy.
(in collaboration with ACI Global S.p.A., Viale Sarca, 336 - 20126 Milano, Via Stanislao Cannizzaro, 83/a - 00156 Roma, Italy)

## ABSTRACT

*In this paper is analyzed a case study of an upgrade of an industry communication system developed by following 'Frascati' research guidelines. The goal of the proposed model is to enhance the industry knowledge Base –KB- by acting directly on information communication system improvements and data system integration, enabling automated process and data processing. The paper follow all the steps performed during the project development: the preliminary data infrastructure design, the information infrastructure improvements, and data processing. Data processing is performed by a calculus engine embedding data mining association rules and Artificial Neural Network –ANN- predictive algorithms thus improving the research. The calculus engine has been implemented by a multiple variables model where the contract data are preliminary processed in order to define functions classifying the operation processes and activating automatically the service process management. The business intelligence –BI- operations are performed mainly by the calculus engine optimizing industry performances. The goal of the paper is to show how research and development –R&D- can be applied by gaining and optimizing the knowledge and processes of an Italian industry working in car services. The project has been developed with the collaboration of the industry ACI Global working in roadside assistance services. By means of a research project resources, the information technology –IT- infrastructure has been improved by new solutions of the communication system and of the data transfer. The proposed case of study provides a model and a guideline to follow in order to apply research in industry acting directly on data and information network.*

## KEYWORDS

*Frascati Guideline, Information System Architecture, Data and Information Networks, Knowledge and information processing, Calculus Engine, Association Rules.*

## 1. INTRODUCTION: BASIC STATE OF THE ART DEFINING INITIAL PROJECT SPECIFICATIONS UPGRADING THE INFORMATION SYSTEM

In this section is discussed the state of the art of basic concepts adopted for the formulation of the preliminary information system architecture of the proposed case of study. A new information workflow suitable for industries and improving information quality, takes into account the following implementation phases [1]:

1. knowledge flow design and development;

2. process steps implementation (steps regarding data analysis and data treatment);

3. logical formalism (logic for association rules formulation oriented on the optimization of the industry performance). An efficient information system –IS- is mainly based on the data quality and data integrity [2]. For this purpose it is important to innovate the whole IS in order to generate many digital data sets (financial data, customer data, etc.) useful for carrying out business intelligence –BI- operations, thus improving business processes. Anyway, an important aspect for business process agility is the automation of management processes [3], which establishes a standardization of some processes by accelerating the activities. This topic is in line with the industry need of the case of study, to allocate efficiently resources by simplifying activities of the operational centers. According to BI strategies, different models have been studied in literature for the customer behavior based on data mining, specifically on "Decision Tree" algorithms [4], where the tree structure suggests a model suitable for the call center structure in order to reduce talk times increasing the accuracy of human resources assignments. The same tree structure can be adopted also to design other algorithms processing association rules enabling data mining processes. Data mining is generic term indicating tools that can be used in general for the activities of business intelligence, and for forecasting. Data mining processing and predictions can be performed by applying different algorithms such Artificial Neural Network –ANN-, Z-score classification, logistic regression, K-Means, random forests algorithm [5], and in general association rules. In data mining, association rules are one of the methods for extracting hidden relationships between data thus providing an important tool adding knowledge. In order to implement association rules, it is important to first pre-process information thus defining the main logic of association rules, and then refine the structure algorithm including different cases of the pre-classification. The formulation of the rules supports the BI for cross selling and upselling activities thus suggesting additional promotions and services based on correlations and combinations of products or services. An intelligent management of consumer data can contribute to define a specific BI tool [8], increasing profits and customer lifecycle [9]. In this context the Customer Relationship Management -CRM- was proposed in [10] as an important tool for the marketing performance. Another useful topic matching with car service call center activities is the Computer Telephony Integration –CTI- [11]-[13]. In this context, an example of integration of a CTI system with the Local Area Network –LAN- has been discussed in [14]. Concerning technologies supporting customer care services, the Interactive Voice Response –IVR- navigation tree can recognize the customer's voice and offer a precise service [15]. Another system which can be integrated with CTI and IVR is the Workforce Management –WM- able to optimize the management of human resources [16]. Other applications potentially suitable for a car service industry is the Optical Character Recognition –OCR- [17] enabling, by a mobile application, the car plate recognition function by automating assistance process of the operation centers. Also E-commerce platform can be useful for the company, mainly in digital marketing orientation [18]. Furthermore the Enterprise Resource Planning –ERP- platforms [19]-[21] is fundamental for process management, allowing the projects control (project management), the asset management, the maintenance, the human resources management, the accounting, the production control, the supplies management, the traceability of the storage, the distribution process, the marketing sales and financial management. Important productivity indices can be estimated by Plan Do Check Act –PDCA- model (Deming cycle) [22] which is a scientific and systemic method processing sequential phases, potentially useful for both executives and operators to address the business problems introduced by the new technology. These tools will also be of considerable importance for the redefinition of quality processes according to ISO 9001: 2015 [23].

From the conceptual and state of the art analysis outlined above, have been defined the preliminary idea of the discussed industry project sketched in Fig. 1, where is illustrated the initial architecture having the goal to upgrade the front office and the back office systems by the integration of different hardware and software tools. Below are listed the preliminary specifications:

- Enhancement of As400 database system: redefinition of the hardware layout of the current As400 infrastructure to adapt to the implementation of new technologies.
- Upgrade of the LAN and Wide Area Network –WAN-: integration of  new tools such as Private Automatic Branch eXchange –PABX- (call center server), and of different databases connected to different application server applications (system integration).
- Billing automation and reengineering of the invoice system: billing platform oriented on the automation of active and passive invoices thus reducing the time of the bill cycle; this tool will define indicators monitoring the effective financial management.
- Payroll (processing of the salaries of a company's employees and of all the related activities): this tool is able to detect attendance, shifts, transfers and budgets, providing useful elements for the costs rationalization and for indicators about the optimal allocation of resources.
- Decision tree call center: cloud system for the creation of decision-making system implementing a tree logic useful for call center operators (this tool could optimize the phone service and help to define the necessary Full-Time Equivalent –FTE- parameter); the system includes the integration of the IVR tool.
- Centralization of Delegate Center data: system for the creation of Customer Management Data Base –CMDB- to adopt for the creation of a detailed registry of the delegated centers concerning motor vehicles and of their disposal, and cars replacement.
- CTI: management of the E-mail/Fax and virtual channels in order to reduce  the time spent in conversation between the customers and the operators.
- Operative dashboards: implementation of graphical dashboards for business intelligence (histograms and plots of activities processes an data mining outputs).
- Invoice system: billing platform linked to dedicated web portal where customer can download invoices legally.
- Document management: document management platform enabling automatically business processes.
- Design and implementation of a Business to Business –B2B- E-commerce platform for sales and services of delegated centers and of the whole sale network;
- Improvement of the web portal (E-commerce access, interactions between operative centers, access to the electronic billing, customer management, etc.).
- Improvement of the mobile app by OCR tool enabling automation assistance processes (customer services);
- E- learning platform (SCORM protocol) for delegated centers and for internal human resources of operational centers (the whole car service network is made by delegated centers, and operative centers displaced in the national territory). The e-learning platform will support the technology transfer.
- Association rules and new algorithms based on new techniques for advanced BI (scientific research improvement).
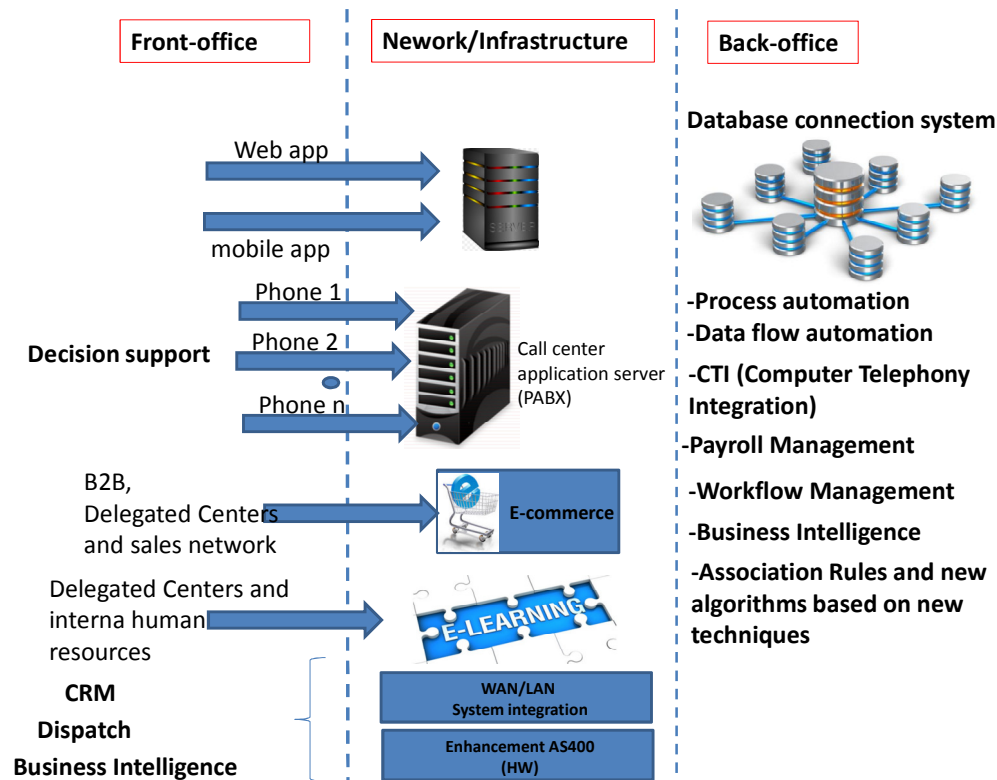
Figure 1. Preliminary information system architecture improving information digitalization, process automation and intelligent data processing.

## 2. UPGRADE OF THE INFORMATION SYSTEM AND OF KB FOLLOWING A RESEARCH APPROACH: ITALIAN REGULATORY FRAMEWORK AND FRASCATI GUIDELINES

### 2.1 Italian Regulatory Framework

The aforementioned project mainly refers to the innovation of information management and data flow processes, oriented on strategic BI keys. In this direction, the development of the project activities has been oriented on the resolution of technological problems on systemic basis regarding to: (i) the management of information that is structured differently for the different implemented technologies (technological and scientific problems mainly related to Industrial Research), (ii) the use of the "enhancement" of the research and of the development of a project able to enrich the knowledge stock promoted by the use of innovative technologies, applying them to a basic knowledge (Knowledge Base -KB-) managed by an upgraded information system. The research activities of the case of study therefore refer to the decree of the Minister of Economy and Finance in concert with the Minister of Economic Development of May 27, 2015 (published in the Official Gazette no. 174 of 29 July 2015), where, following the "Definitions" contained in paragraph 1.3, point 15, of the "Guidelines on State aid for research, development and innovation" in the Commission Communication (2014 / C 198/01) of the 27 June 2014 (published in the GUUE C / 198 of 27 June 2014) 1, are described all the activities related to "fundamental research", "industrial research" and "experimental development".

In particular, the proposed project is inscribed in the last two main activities. In fact, according to letter b) of paragraph 4 of Article 3, taken from letter b) of paragraph 1 of article 2 of the implementing decree, are classified in "industrial research" the following activities: planned research or critical investigations aimed at acquiring new knowledge to develop new products, processes and services and, at the same time, to allow an improvement of existing ones. Furthermore, based on letters c) and d) of paragraph 4 of article 3, taken from letters c) and d) of paragraph 2 of article 2 of the implementing decree, can be classified as "experimental development" the following activities: acquisition, combination, structuring and use of existing scientific, technological and commercial knowledge and skills, aimed at the production of plans, projects or designs for new, modified or improved products, processes or services, may also be other activities intended for conceptual definition, planning and documentation concerning new products, processes and services. These activities may include the development of projects, drawings, plans and other documentation not intended for commercial use (the BI and the outputs of the project is not to be considered a sales object, but helps only to increase the stock of knowledge apply them to advanced business strategies). It is therefore considered that the object of the research originates from the need to review processes and services closely related to the technological upgrade based on scientific methodologies and approaches. This goal is therefore a part able to facilitate activities process and service innovation, and to address "the creation of components of complex systems, necessary for industrial research ..." (see Agenzia delle Entrate Circular no. 5 / E of 16.3.2016 concerning: Article 3 of the Decree-Law of 23 December 2013, n. 145, converted with amendments by the law 21 February 2014, n. 9, as amended by paragraph 35 of article 1 of the Law of 23 December 2014, n. 190; see also Stability Law - Tax credit for research and development activities of 2015).

## 2.2 Development of the Project Specifications by Following Frascati Guidelines

In this paragraph are provided more useful details to better focus on research issues of the project following the 'Frascati' manual [24]. The project contributes, following different logics and specifications, to increase the systemic knowledge stock of a car service company working mainly on services. The knowledge gain becomes the primary goal of the research project. In this direction, in Fig. 2 is illustrated a functional summary diagram useful to understand the key points of the research. From this model it is possible to see how the software tools and the data flow systems (before the activation of the project activities) represent the initial knowledge base –KB- of the industry. The initial KB is represented by level 1 of Fig. 2 as a fragmentary knowledge base, which is not properly structured and not suitably interconnected with the information infrastructure according to a defined logic process. Starting from level 2 (beginning of the project development), and through new and evolved software architectures and databases, it is possible and to increase the knowledge base [24] (the gain of knowledge can be performed by the processing of different types of data coming from different data sources, data that are appropriately synchronized and managed by operational processes), by levels 3 and 4 of Fig. 2. The increase in the knowledge stock is achieved mainly through the following points:

- (a) the implementation of various functional modules with the aim of enriching the new knowledge base through new data, by matching and correlating them with other informatic modules to be processed in an innovative information system: data flow management, data analytics (big data analytics, ie. the analysis of massive amounts of structured and unstructured information, analytics of processed data) and new data processing approaches (by new artificial intelligence and self-learning algorithms or by data mining techniques), potentially carry out an advanced BI which is not feasible by means of traditional software and hardware technologies ('Frascati' manual: "…*the creation of new or more efficiency-based algorithms on new techniques*" page 66 [24]).
- b) the carried out activities are considered as a part of applied research, as subjects of original multidisciplinary investigations (BI investigations, research about information

sources, technological and scientific investigations, research about efficient system architectures, investigations on innovative algorithms capable to provide new outputs, etc.) targeted to acquire further new knowledge related to prediction, to BI and to the data mining for the achievement of a specific and practical application objective ('Frascati' manual: "*Applied research is original investigation undertaken in order to acquire new knowledge. It is, however, directed towards a specific, practical aim or objective.* "; point 1.35 [24]);

- c) the carried out activities are considered as a part of experimental development, because they refer to developments on a systemic basis (as represented by the architectural systems shown in the Fig.1), capable of producing an 'additional knowledge' for the optimization of the management processes of the business activities ('Frascati' manual: *Experimental development is systematic work, drawing on knowledge gained from research and practical experience and producing additional knowledge, which is directed to producing new products or processes or to improving existing products or processes.*", page 29, point 1.35 [24]).

- d) The activities were also aimed at resolving conflicts within the hardware and software (integration of structured and non-structured data and information), conflicts resulting from the reengineering process of the information infrastructure ('Frascati' manual: "*The effort to resolve conflicts within hardware or software based on the process of re-engineering a system or a network*", page 66 [24]). In this context, the reengineering of the data flow inherent the software modules and databases, has generated problems due to conflicts highlighting some limits such as: (i) management of structured data originated by different technologies (in order to overcome this limit has been implemented an Enterprise Service Bus –ESB- [25]); (ii) the migration of data coming from different database technologies towards big data technologies (for example using NoSQL technology [26]); (iii) the need to pre-process data to perform advanced analyses using innovative data mining algorithms (cleaning, filtering, etc.); (vi) the need to combine multiple data sources to optimize the activity management processes; (v) the possibility to consider in the post-analysis phase the uncertainty of the project outputs (uncertainty due to incorrect data extraction or to incorrect information); (vi) the need to integrate different technologies with a view to system integration and to embedded systems ('Frascati' manual: "*In the systems software area, individual projects may not be considered as R & D, the resolution of which will need R & D. Alternatively, a large project can R & D in its planning, but there may be elements in the project that would need additional R & D activity to the smooth integration of different technologies.* "page 66 [24]).

- e) The orientation to identify Research and Development towards new car services, and the orientation of the population to use the new services ('Frascati' manual page 70 [24]: "*Examples of R & D in some other service activities:*

  ● *the development of new methods for measuring consumer expectations and preferences;*

  ● *the development of new methods to deliver and measure social service outcomes that can be adapted into a variety of different socioeconomic or cultural settings*").

It is noted that the increase in the KB is marginal in the case where an implementation is limited to the first two levels of Fig.2 (further fragmented information added to the basic KB): in this case the outputs are distinct, and the increase of the KB will be limited because is not considered in an "embedded" system capable to integrate all the information. By adopting the architecture of Fig. 2 and the technological and scientific enhancements, into an embedded systems will be possible to produce new outputs by solving systematically technological and scientific uncertainties ('Frascati' manual page 65 [24]: "*Software development is an innovation related activity that is sometimes connected with R&D and incorporates, under specific conditions, some R&D. For a software development project to be*

*classified as R&D, its completion must be based on a scientific and / or technological advance, and the aim of the project must be the systematic resolution of a scientific and / or technological uncertainty*".

Furthermore 'Frascati' considers that the algorithms applicable to software outputs (or databases), even if not innovative but necessary, are able to produce scientific / technological upgrades ('Frascati' manual page 65/66 [24]: "*The software development component of such projects, however, may be classified as R&D if it leads to an advance in the area of computer software. Such advances are generally incremental rather than revolutionary. Therefore, an upgrade, addition or change to an existing program or system may be classified as R&D if it embodies scientific and/or technological advances that result in an increase in the stock of knowledge."*). So, the software and licenses used in research projects (indicated in Fig. 2 by SW) therefore become essential elements for structuring a knowledge base from which to implement an advanced BI necessary for the revisitation and the innovation of processes and services [27].
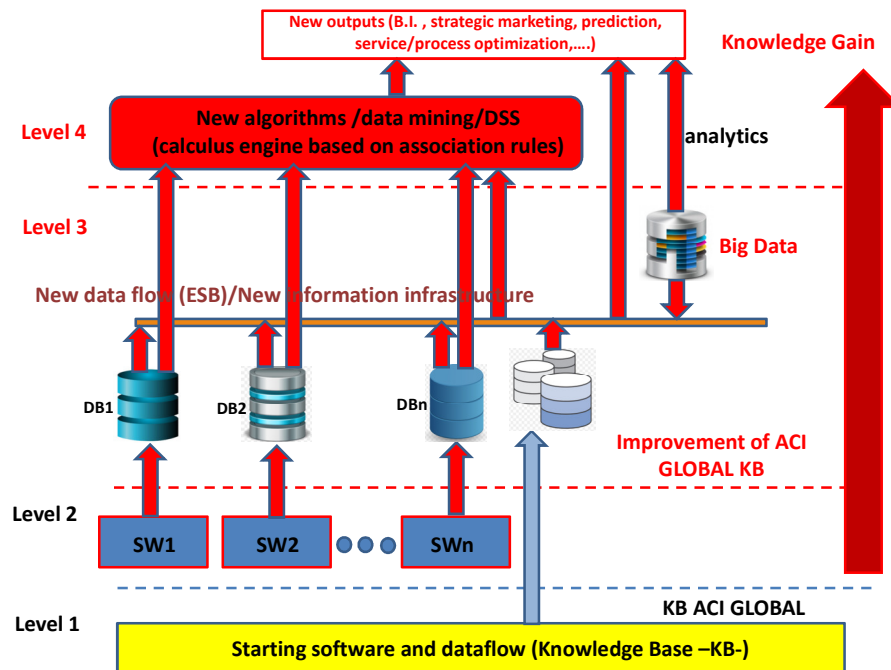


Figure 2.  Knowledge Base Gain concept in ACI Global S.p.A.: four levels model upgrading KB.

## 3. GAIN OF KB AND NEW PROTOTYPE INFORMATION SYSTEM NETWORK: IMPROVEMENTS OF THE PRELIMINARY ARCHITECTURE

In this section are discussed some main tools improving the preliminary architecture of Fig. 1 and the KB gain by discussing data flow of the final information system automation processes and BI improvements.

### 3.1  Upgrade of the CTI Tool: Integration of a Multi-Channel Information System Based on Priority Rules

In the first part of the project has been implemented the basic CTI tool. Starting from this tool the gain of the KB has been achieved by integrating new channels for user communication, by

defining new priority conditions and association rules. In the updated communication system contacts have been forwarded into the following heterogeneous channels:

- Voice;
- Fax;
- E-mail;
- Webpush (callback);
- No-voice lists (As400).

The contacts are characterized by different Service Level Agreements (SLA) depending on the nature of the contact. For example, a voice contact must be processed quickly in real time, besides asynchronous contacts (Fax, E-mail, webpush, non-voice information) are processed by a time table defined according to the order of call arrivals within defined time limits. All contacts are addressed to the new CTI platform. The contacts are distributed to the operators in charge according to the operators skills. Agents can have one or more skills and then can manage different types of contacts. The new CTI platform allows to optimize the queuing strategies towards operators. In Fig. 3 is illustrated the adopted technological infrastructure enabling the following priority levels (priority rules):

• High : voice or callback;
• Medium and medium low: fax, or e-mail or no-voice information.
And the following operators skills (no-voice contacts):
• Customer fax 1: skill profile X;
• Customer fax 2: skill profile Y;
• Customer e-mail 1: skill profile Z;
• Customer e-mail 2: skill profile K.
,
 where 1 indicates the major priority (urgent cases).
The way of forwarding contacts follows the one to one logic: each operator can only manage one interaction at a time independently to the contact type. The improvement of the CTI has been performed after an accurate mapping of the process related the service call system. The architecture of Fig. 3 is collocated into level 2 of the theoretical architecture of Fig. 2 where all tools (software -SW-) provide information by means of different server database (DB) system. The data stored into each DB can be processed in order to estimate Key Performance Indicators – KPI- of operators.
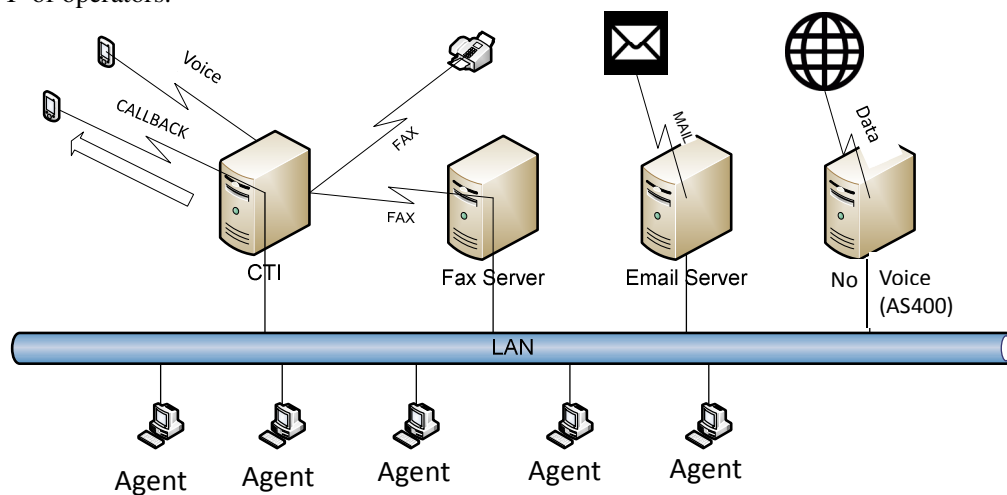


Figure 3.  Upgraded information infrastructure concerning CTI.

## 3.2 Improvement of the Information Infrastructure by ESB and by Graphical Dashboards

The actors mainly involved in the data source extraction (level 3 of Fig. 2) process are:
- Head of contractors –COT-;
- Consultants –COS-;
- BackOffice system;
- Team Leader –TL-;
- Agents.

The new technological infrastructure is made by (see Fig. 4):

• An Extract, Transform, Load –ETL-  platform that reads, with a predefined frequency depending on the availability of data by each individual platform, the data sources and aggregates them into a relational database;
• An ESB (the open source WSO2 [28]) able to connect different data formats coming to different tools and technologies by solving conflicts between hardware and software (the ESB will be useful also for the data migration from different database –DB- to big data, moreover the graphical interface will facilitate the management of data coming from the different data sources by allowing  an easy formulation of structured queries);
• Relational databases collecting the information to be analysed.
• A data visualization system that displays information in the form of real-time charts (dashboards updated every 3-5 minutes).
The information to be collected for each area by the upgraded system, are summarized in the following table:

Table 1.  Information and data collected by each management area.

| Area | Information/data type |
|---|---|
| Contact Area | - Queue status: call waiting, mail waiting, fax waiting, callback;<br>- State of the operators: moment by moment how many are ready, how many in busy, how many are responding to an email or fax, how many are in pause and for which reason, etc.<br>- Assignments Situation: step by step how many assignments have been assigned, how many are in progress, how many have been closed |
| Dossier  Area | - Practices (dossier ) opened;<br>- Practices in progress;<br>- Practices closed;<br>- Backlog practices; |
| Staffing Area | - Comparison between staffing (Work Force Management –WFM-) and actual presence (entry and exit stamps). |

The data source and related data that are transferred into the ESB infrastructure are mainly the following ones (Table 2):

Table 2.  Information and data collected by each management area.

| Tool | Information/data type |
|---|---|
| CTI | CRD (Call Record Detail): log in and out calls, FAX and Email, non-voice lists (currently managed via AS400 views) |
| IVR | IVR navigation string: calling log of selected options |
| TS ACG (Stamping) | Data extracted from garrison device (entry and exit data) |
| Business Process Modelling- BPM- | Practices opened, practices in progress, practices closed |
| Callback | Rescue requests coming from the mobile APP i803116 (user caller of Fig. 6) |

In Fig. 5 is shown an example of graphical dashboard for operative centers agents. The dashboard is oriented to give few data in a graphical format that give evidence of the room's performance supporting the agents to understand if their work is efficiently managed. For each actor is activated a proper graphical dashboard useful for decision making operations or for corrective management actions of different industry areas. The dashboards provide the information that was previously hidden by the fragmentary KB of level 1 of Fig. 2. The simultaneous analysis of data of different dashboards could provide new outputs for the BI and for the strategic marketing (level 4 of Fig. 2).
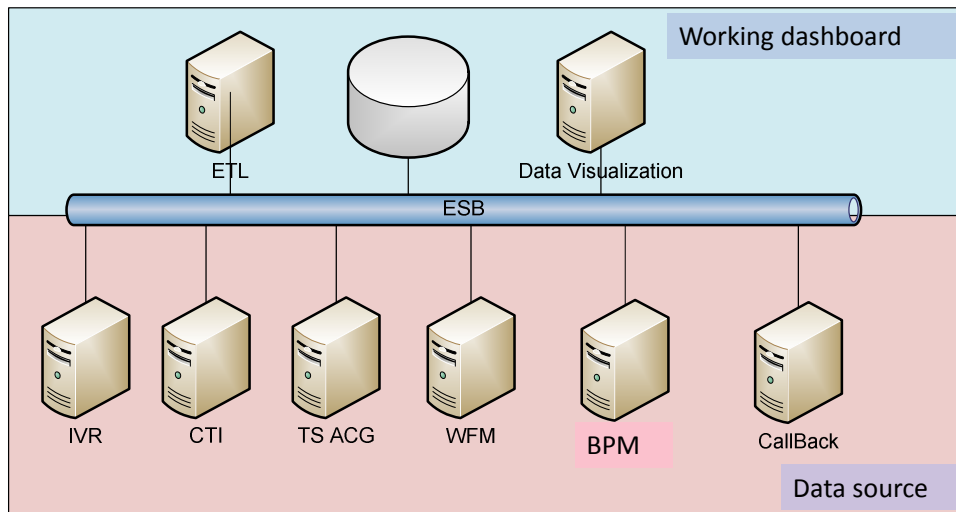


Figure 4.  Information infrastructure integrating different technologies and tools.
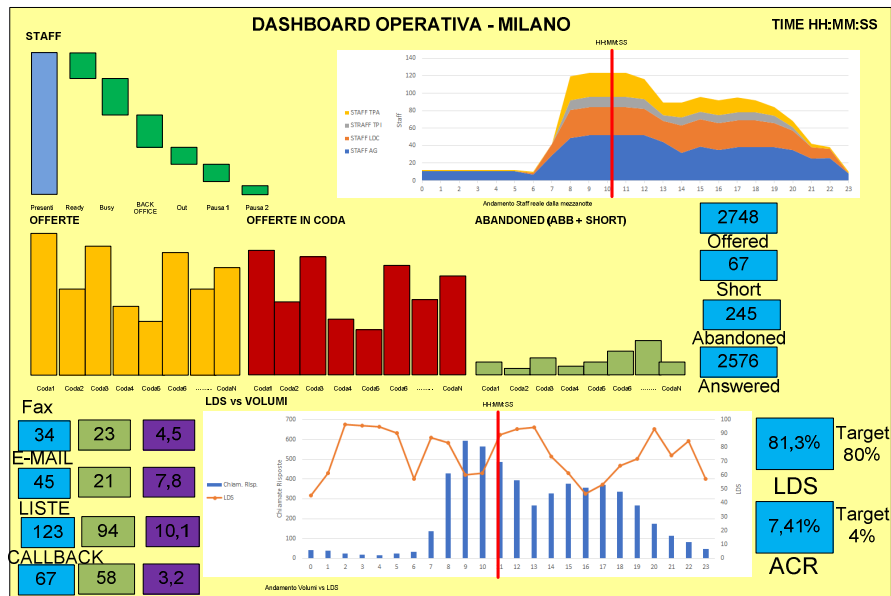
**Agent dashboard**



Figure 5.  Example of graphical working dashboard.

## 3.3  Upgrade of the Information System and of Data Flow Management

The preliminary architecture of Fig. 1 has been improved by the architecture illustrated in Fig. 6 where it is possible to distinguish better the frontend from the backend data flow. By means of this architecture are activated mainly the following services:

- Automatic rescue missions;
- Interventions of missions piloted by the operation centers;
- Interventions assignment totally controlled by the control panel operator.

The integration of the communication system with the external system that delivers the missions and automatically dispatch the requests for roadside assistance to the delegated centers, will optimize the whole operational performance. Thanks to the improvement of the user interfaces and to the simplification of the mission assignment process, is achieved the user experience of the system by refining  the activation logics of the missions. The performed dataflow integrations are related to:

- Integration between Automation and Tool Service Management –TSM-  system dossier through ESB protocol;
- Invoice automation (see data flux of Fig. 7);
- Server communication (communication of different hardware devices);
- Geolocation by GeoEngine and Map viewer (information useful to guarantee fast services);
- ERP and CRM providing data to BPM;
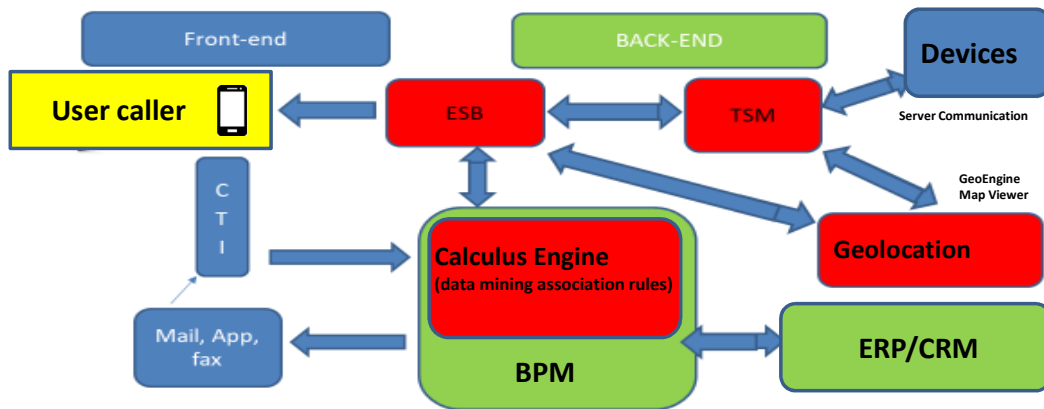- Linking of the BPM to a calculus engine improving association rules for business processes.

Figure 6.  Upgraded Information System.

Observing for example data flow of invoice automation process of Fig. 7, it is clear how invoice data can support the–BI- of the industry by predicting costs of each practice: the historical data can be adopted to create a training model able to predict by ANN algorithms the cost of a classified practice (see section 3.3.1).
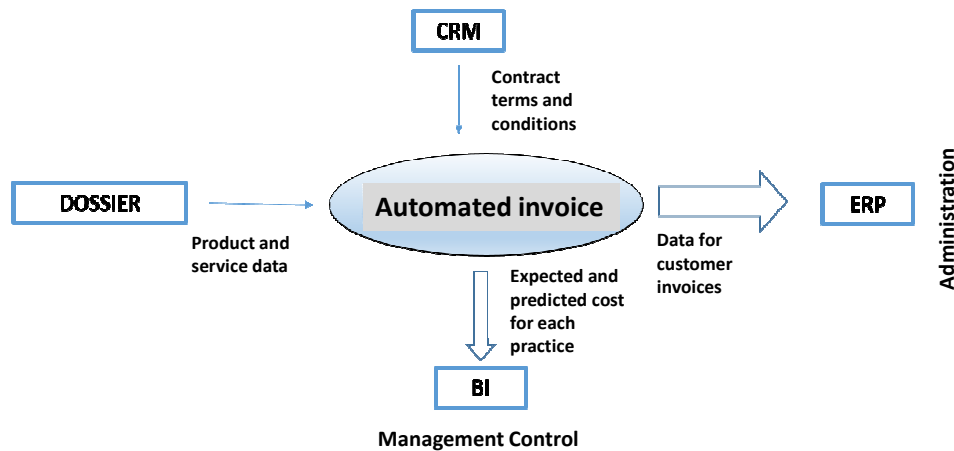


Figure 7.  Example of data flow involving automated invoice process.

The invoice automation process is framed into level 3 of Fig. 2.

## 3.3.1 Example of  Gain of Knowledge by ANN Algorithm (level 4 of Fig. 2): BI Improvement by Data Mining

An example of application of "gain of knowledge" is illustrated in Fig. 8 where is plotted the prediction of a practice cost (normalized values based on simulation results). The prediction results provide a decision support system -DSS- based on a multilayer perceptron –MLP-artificial neural network –ANN- [29]. The open source Konstanz Information Miner –KNIME- [30] workflow (see Fig. 8) has been adopted in this example for the data processing setting the following parameters: 210 samples of training dataset  (historical cost data), 12 testing dataset, 100 iterations, 2 hidden layers,  10 hidden neurons per layer. The simulation of Fig. 9 provides the prediction cost of the future 12 days where the cost of a practice can oscillates around a defined minimum and maximum values defined by the pre-classification of the practice.
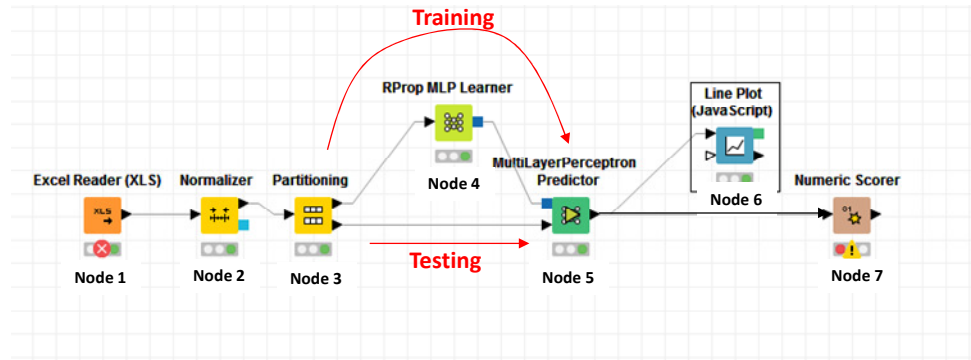
Figure 8.  KNIME workflow implementing ANN algorithm.

The workflow is structured in the following nodes:
- Node1 (node reading input data): data can be loaded from an excel file containing data of specific DB systems, or can be extracted through the ESB , or can be loaded from a big data by a python script);
- Node 2 (Normalizer): data are normalized in a scale from 0 to 1 in order to process different attributes having different scale values;
- Node 3 (Partitioning): this block establish the data percentage of the input dataset to be processed by the training operator (Node 4)  and by the MLP algorithm (Node 5);
- Node 4 (RProp MLP Learner): operator enabling the training of the model implementing RProp algorithm [31];
- Node 5 (MultiLayerPerceptron Predictor): node embedding MLP algorithm;
- Node 6 (Line Plot): block embedding ANN MLP algorithm;
- Node 7 (Numeric Scorer): block providing calculation performance.

The proposed workflow is a graphical user interface –GUI- which can be automated by connecting at the input a node containing commands (scripts) to extract periodically data from a data source. Data can be extracted from different data sources and merged into an unique local repository used for ANN processing.
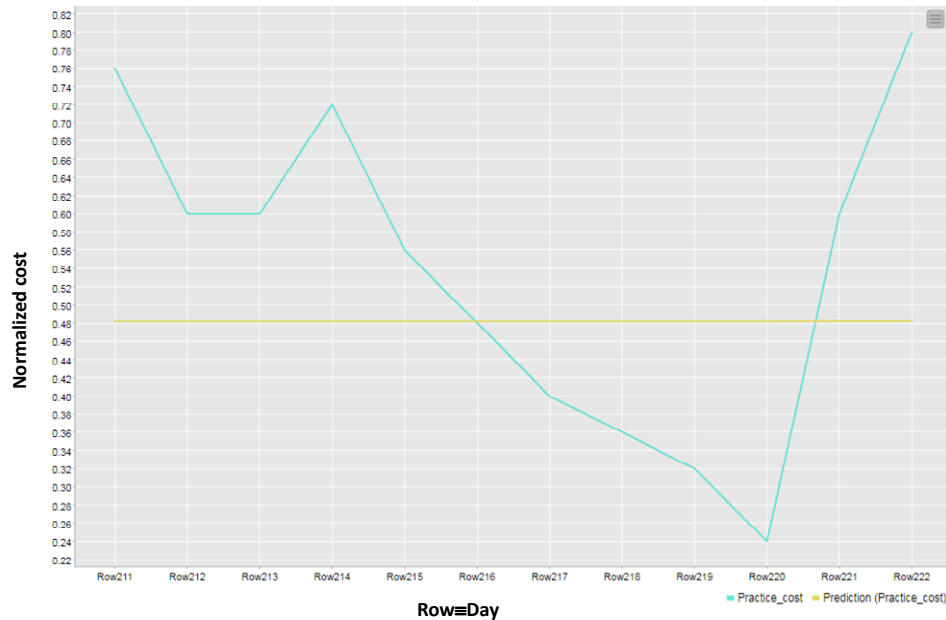


Figure 9.  Prediction results of practice cost.

The ANN algorithm structured for industry data process is framed into level 4 of Fig. 2. The MLP is adopted for the initial lower data flow. Other data mining algorithms will be considered for  larger dataset to process.

## 3.4  Automation Processes of Practices

In this section is described  the upgrade of the automation process managing practices (dossier automation system). From the application point of view the automated system takes into account the following functional aspects:

• Management of the "single-channel incoming", allowing the integration at the telephone level via CTI;

• Management of the "entitlement" system access;

• Process management (workflow) of different car service "*Help on Phone*", "*Depannage*", "*Towing*", "*Sub Traine Performance*", "*Substitutive Car*", and "*Accessory Performance*";

• Loading of the contract types.

• Management of contracts for rescue services and insurance policies linked to them.

• Management of tariffs with respect to services, suppliers and typologies of suppliers.

• Service delivery management: geo-localization via integration of the Google Maps service to assign the Supplier assignment via TSM and through the BPM application platform.

• Integration to the ERP system.

• Database integration.

From an architectural point of view, the data system takes into account a model of composite architectures based on technologies allowing the decoupling between the architecture levels, the internal services design, and the integration of external systems.

The architectural and functional choices are fundamental in order to provide a certain flexibility when the organizational and business processes vary. An approach based on loose coupling integrations has been considered for the system development. Below are indicated some aspects concerning technological choices:

   a. Data exchange via standard script (JavaScript Object Notation –JSON-, eXtensible Markup Language –XML-) following the suggestions of Table 3:

Table 3.  Case Uses of JSON and XML script.

| Script | Case |
|--------|------|
| **JSON** | - Data exchange occurs between internal applications or in any case does not need a public bees; <br> - Data exchange occurs between a server and a mobile device; <br> - Saving bandwidth; <br> - Data exchange for very large data (tens or hundreds of MB). |
| **XML** | - Data exchange which take place between software developed by many different companies including the possibility to have diagrams greatly helping the development thus facilitating the creation of web services; <br> - For data management originally in XML format. |

   b. Encapsulation of Legacy applications and / or pre-existing services.
   c. Breakdown of applications into services.
   d. Implementation of processes based on services.
   e.  Orchestration of services.

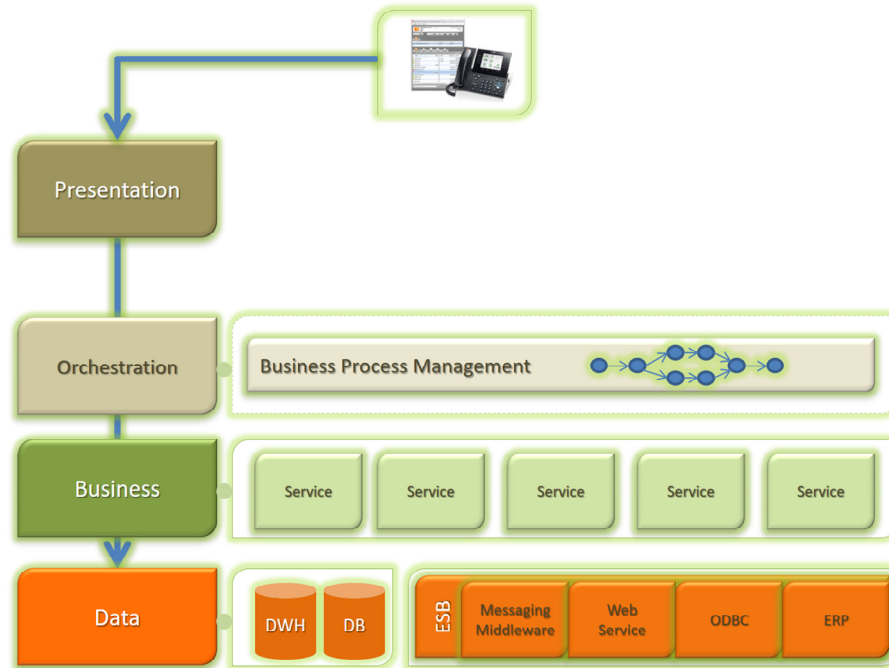The following Fig. 10 schematizes the  architecture of the Automation Dossier tool.

Figure 10.  Architecture of Automation Dossier management tool.

The following description describes   the reference technological stack of the whole implementation solution. The implemented technologic stack of Fig. 11 is  summarized by the following table 4:

Table 4.  Technologies adopted for the Automation Dossier tool.

| Level | Information/data type |
|---|---|
| **Web/presentation layer** | • Laravel Model-view-controller –MVC-: open source framework written in PHP for the development of web applications;<br>• Dingo API package to use in the Laravel framework (speed of development warranty);<br>• Laravel / Blade simple but powerful model engine supplied with Laravel. Unlike other popular template templates for PHP, Blade allows you to use PHP code in your visualizations. Blade views are compiled into a simple PHP code and cached until they are changed, which means that Blade essentially does not  add overheads to the application.<br>• Vue.js JQUERY: Vue.js  is a JavaScript framework, dedicated to the realization of reactive web interfaces that exploit the dual-binding between data and view model; jQuery is JavaScript library for web applications; it was created with the aim of simplifying the selection, the manipulation, and the management of events and the animation of DOM elements (Document Object Model) in HTML pages, implementing also AJAX functionality.<br>• $HTML_5$ :markup language for structuring web pages, published as W3C Recommendation.<br>• $CSS_3$ : it is  a language used to define the formatting of HTML, XHTML and XML documents. |
| **Business layer** | • BPM enterprise-class platform for the realization of Business Process Management. |

| | |
|---|---|
| | • Laravel / MVC software architecture framework that separates business logic from the rest of the user interface (by separating the application into three parts: the model, the view and the controller).<br>• WSO2 ESB is a complete integration solution that allows communication between various and disparate applications. Instead of having applications communicate directly with each other in all their various formats, each application simply communicates with WSO2 EI, which handles the transformation and routing of messages to their appropriate destinations. |
| **Data/MetaData layer** | • Laravel / Eloquent: the ORM Eloquent included in Laravel provides a simple, ActiveRecord implementation to work with the database. Each database table has a corresponding "template" that is used to interact with that table.<br>• Mysql 5.7 RDBMS, relational database manager.<br>• Guzzle / http client: PHP HTTP client that facilitates sending HTTP requests and it is easy to integrate with web services. |

During the project development has been evaluated the possibility of abandoning a too structured and rigid standard experimental technology, moving to a more flexible BPM.
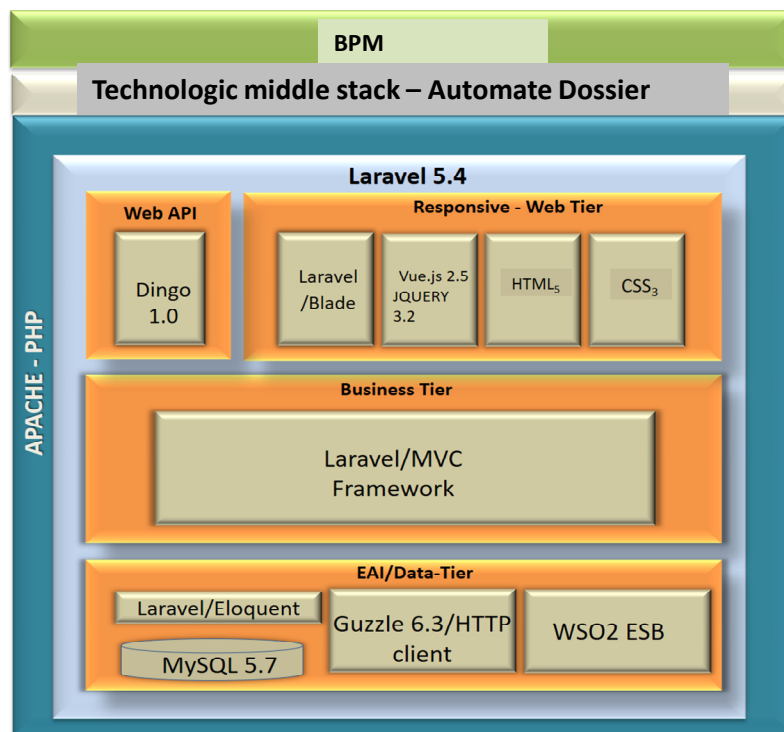The Automation process of practices is framed into level 3 of Fig. 2.



Figure 11. Technological stack used for the Automation Dossier solution (Automation Dossier framework).

## 3.5 Association Rules implemented into the Calculus Engine: BI as KB Gain

The calculus engine of Fig. 6 represent an tool improving further "gain of knowledge". The architecture of the calculus engine expressing it functionalities is "exploded" in Fig. 9.
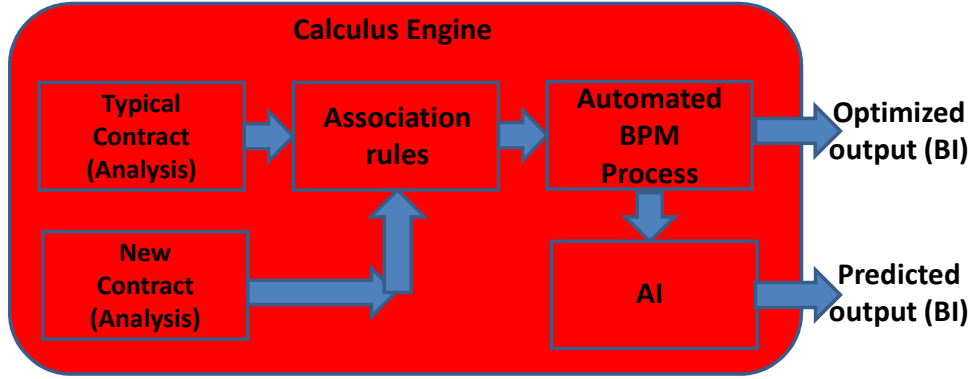
Figure 11.  Calculus engine architecture: association rules embedded into BPM tool.

The scheme indicates a pre-processing of the existing contracts useful for the definition of the main variables able to classify by association rules the main classes. New contracts are analysed in order to define further variables able to refine the classification. The association rules model will provide automatically to the BPM information about the management of the classified activities thus improving BI facilities. The use of the artificial intelligence –AI- algorithms will provide also information about predictions (prediction of services, sales, cost evolution of practices, etc.) thus providing a Decision Support Systems –DSS- based on predicted outputs (advanced BI).

A preliminary contracts analysis showed that, in some cases, the components of the association rules function are dependent on several variables. For this reason has been formulated a preliminary basic model representing a generic function on a single real-x variable and defined as:

$$f(x) = \begin{cases} c_{11}, & x < f_1, \\ c_2, & x \in [f_1, f_2) \\ c_3, & x \in [f_2, f_3) \\ c_3 + c_0 x, & x \in [f_3, \omega) \end{cases} \tag{1}$$

Equation (1) defines some important  intervals (or ranges) of a size (km, days, weight, time ...) and the constants $c_0$ define the classified costs per unit (for example, the distance component has a fixed price within mileage ranges, and a price per km over a certain limit). The asocial rules have been extended for function which depends on two or more variables, expressed by:

$$f(x,y) = \begin{cases} c_{11}, & x < f_1, \\ c_{12}, & x \in [f_1, f_2) \\ c_{13}, & x \in [f_2, f_3) \qquad y < p_1 \\ c_{13} + c_{10} x, & x \in [f_3, \infty) \\ c_{21}, & x < f_1, \\ c_{22}, & x \in [f_1, f_2) \\ c_{23}, & x \in [f_2, f_3) \qquad y \in [p_{1,\infty}) \\ c_{23} + c_{20} x, & x \in [f_3, \infty) \end{cases} \tag{2}$$

In this case, the distance component would have a price still dependent on the mileage, but this price would also depend, for example, on the weight range $p_i$ defined by a decision tree, where each level is defined by one of the following main quantities:

- distance in km;
- duration in days;
- type of road;
- weight range;
- time slot / day of the week.

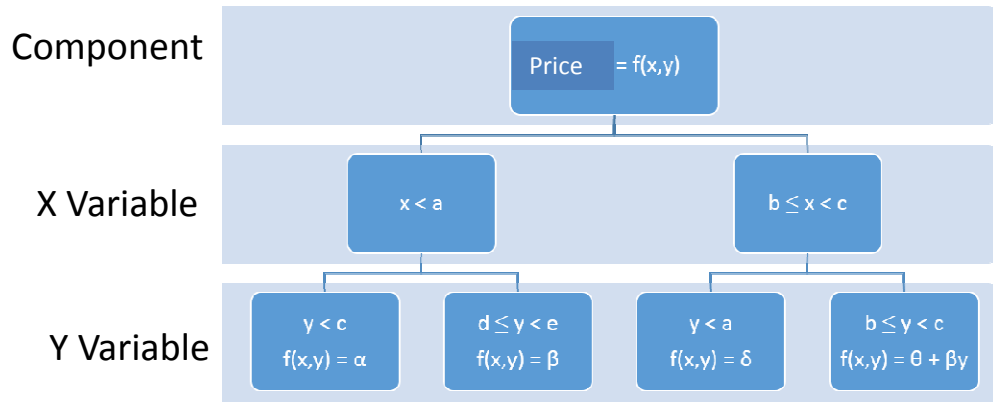In Fig. 12 is illustrated the association rules model implementing the function with two variables.



Figure 12. Calculus engine architecture.

In the tree structure each level represents a variable and each branch depends on the value of this variables, besides the leaves contain the parameter value.

The contracts have been integrated into the multi-variable component, placing themselves at the level defined by the function classifier. This last model evaluates which branch to follow, if a lower level is present, or allows to calculate the price of the component (leaf of the tree). In Fig. 13 is illustrated an alternative representation of the multi-variable model.
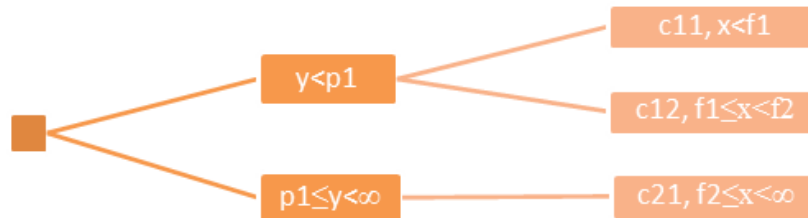


Figure 13. Alternative representation of the decision tree model implementing a multi-variable function.

The weight $p_i$ can be correlated to the type of service

The multi-variable model has been applied by providing the Unified Modeling Language –UML-design of Fig. 14 integrating the decision tree cost diagram of Fig. 15.
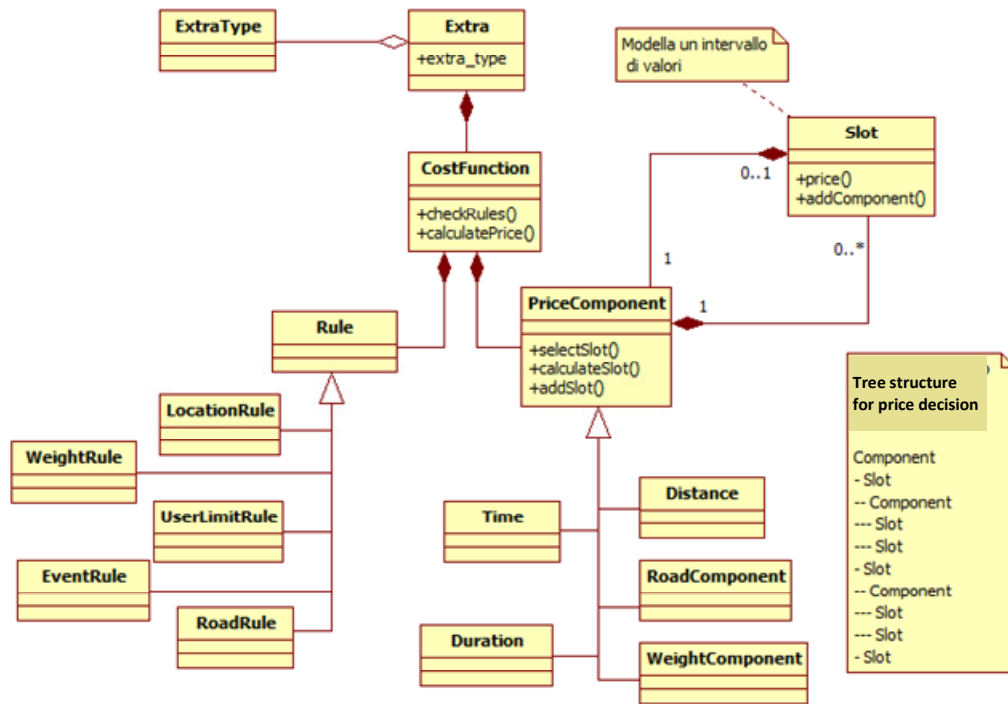
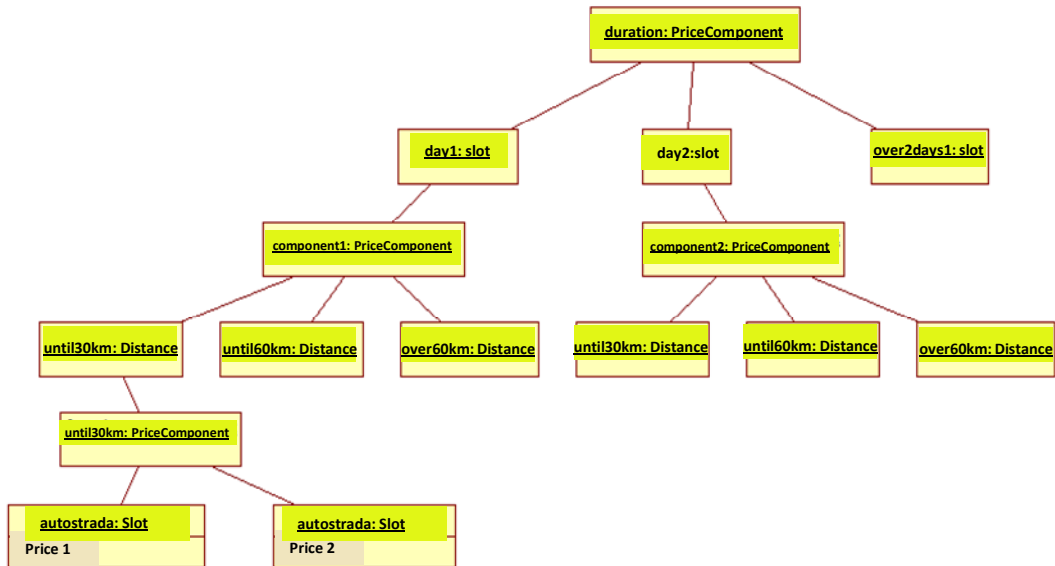Figure 14.  UML diagram integrating cost function and a cost component.



Figure 15.  Object diagram for a cost component.

In table 5 are reported all the executed propedeutic steps regarding the implementation of the multi-varable decision tree model:

Table 5.  Propedeutic Steps of decision tree model.

| Step | Description |
|---|---|
| 1. **Contract pre-analysis** | Pre-processing of classified contracts. The classification has been performed for customers and for sector (automotive, health, etc.). |
| 2. **Rules Extraction** | Rules extraction from the analysed contracts (discursive contracts) and their generalization. During this steps are the defined the logics which can be generalized and combined with other ones. |
| 3. **Logics combination** | Modelling of a contract performed by combining different logics. |
| 4. **Implementation in the calculus engine** | Implementation of models into the calculus engine and linking with the BPM. |

The proposed model has been tailored for the following processes improving automation:
- Active billing;
- Passive billing (costs);
- Choice of network to be used for the destination of the vehicle (logistic).

The calculus engine is framed into level 4 of Fig. 2.

For future software development will be considered also the implementation of fuzzy neural networks [32] suitable for an efficient knowledge extraction providing simultaneously proper association rules.

## 4. CONCLUSION

The goal of the paper is to show results of a case of a study regarding research and development - R&D- applied for a car service industry. Following "Frascati" R&D theory, the research has been focused on the upgrade of the industry information system by describing some improvements performed in strategic management areas such as contact area, dossier area and staffing area. Starting from a preliminary architecture design have been discussed all the tools improving the KB of the whole information systems. The hardware infrastructure and the adopted software tools are necessary to optimize the service processes and to construct a further upgrade enabling an advanced BI. The studied architecture is fundamental to perform the gain of the KB making available the whole digitalised information. The KB is available in digital format by means of the implementation of the ESB protocol allowing data automation transfer and data extraction from different database technology. For this purpose ESB is a key element for system integration. Some association rules defined the level priority for call center system and for the practice management oriented on BI improvement. The new process of practice management has been implemented by the "Automation Dossier" management tool by describing the adopted framework and related technologies including open source tools. The KB gain is mainly improved by the calculus engine integrating a multi-variable association rule model classifying automatically the contracts and the car service practices. An example of a cost decision tree classification provided by the implantation of the multi variable model has been proposed. The calculus engine is embedded into a BPM tool thus enabling business process automation. This calculus engine shows how it is possible to match industry research with scientific methodologies focusing on logics formulation and on data processing. The paper show also how it is possible further add KB gain by data mining algorithms such as ANN: a KNIME ANN workflow has been applied, as example, for the prediction simulation of practice cost. The results are synergistic with those of other research activities of the same industry concerning the development of data mining

algorithms and big data systems. The proposed work can be a reference for the researchers working in industry projects involving information systems and automated processes.

## ACKNOWLEDGEMENTS

## REFERENCES

[1]    Kuechler, B. & Vaishnavi, V. (2011) "Promoting Relevance in IS Research: An Informing System for Design Science Research", Informing Science: the International Journal of an Emerging Transdiscipline, Vol. 14, pp 125-138.

[2]    Arachchi, S. M., Chong , S. C. & Madhushani, A. G. I. (2015) "Quality Assurance and Quality Control in ERP Systems Implementation", American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), Vol. 11, No. 1, pp 70-83.

[3]    Seethamraju, R. & Sundar, D. K. (2013) "Influence of ERP Systems on Business Process Agility", IIMB Management Review, Vol. 25, No. 3, pp 137-149.

[4]    Eichinger, F., Nauck, D. D. & Klawonn, F. (2006) "Sequence Mining for Customer Behaviour Predictions in Telecommunications",  Proceedings of the Workshop on Practical Data Mining: Applications, Experiences and Challenges, Berlin, Germany, Vol. 1, 2006, pp 3-10.

[5]    Tsai, C.-F. & Lu, Y.-H. (2010) "Data Mining Techniques in Customer Churn Prediction", Recent Patents on Computer Science, Vol. 3, No. 1, pp 28-32.

[6]    Pinedo, M., Seshadri, S. & Shanthikumar, J. G. (2000) "Call Centers in Financial Services: Strategies, Technologies, and Operations", Springer book: Creating Value in Financial Services, chapter 18, pp 357-388.

[7]    Sehat, M. & Flores, R. P. (2012) "Customer Data Management", Master of Science Thesis INDEK 2012:89 KTH Industrial Engineering and Management, Stockholm.

[8]    Lia, M. (2015) "Customer Data Analysis Model using Business Intelligence Tools in Telecommunication Companies", Database Systems Journal, Vol. 6, No. 2, pp 39-43.

[9]    Kamakura, W., Mela, C. F., Ansari, A., Bodapati, A., Fader, P., Iyengar, R., Naik, P., Neslin, S., Sun, B., Verhoef, P. C., Wedel, M. & Wilcox, R. (2005) "Choice Models and Customer Relationship Management",  Marketing Letters, Vol. 16, No. 3-4, pp 279–291.

[10]   Soliman, H. S. (2011) "Customer Relationship Management and Its Relationship to the Marketing Performance", International Journal of Business and Social Science, Vol. 2, No. 10, pp 166-182.

[11]   Messerschmitt, D. G. (1996) "The Future of Computer-Telecommunications Integration", IEEE Communications Magazine, Vol. 34, No. 4, pp 66-69.

[12]   Katz, R., Joseph, A., Czerwinski, S., Hodes, T., Hohlt, B., Kiciman, E., Ludwig, R., Mukkamalla, S., Oden, K., Ordonez, A., Raman, B., Shih, J., Wang, H. & Zhao, B. (1999) "A Scalable Service Architecture for Computer-Telephony Integration", Proceedings: In IEEE Communication Magazine, pp 1-21.

[13]  Parihar, A., Chorage, D., Bichkule D. & Kumbhar, P. (2016) "CTI Integration Using Salesforce.com CRM", An International Peer Reviewed & Referred Scholarly Research Journal for Interdisciplinary Study, Vol. 3/23, pp. 1671-1676.

[14]   Credle, R., Adukkadukkath, R., Rajendiran, R., Singh, R. P. (2007) "IBM Sales Center with Computer Telephony Integration", IBM Redpaper.

[15]   Benzi, V. (2010) "Un IVR Intelligente per Supportare le Necessità del Servizio Clienti", VoiceCom News 2010.

[16] "Industry Insider Guide- Workforce Management" 2018. [Online]. Available: http://www.afrc.org/iso_album/wfm_insider_guide.pdf

[17] Patel, C., Shah, D. & Patel, A. (2013) "Automatic Number Plate Recognition System (ANPR): A Survey", International Journal of Computer Applications, Vol. 69, No.9, pp 21-33.

[18] Kingsnorth, S. (2016) "Digital Marketing Strategy: An Integrated Approach to Online Marketing", Paperback, Kogan Page book, pp 1-338.

[19] Garefalakis, A., Mantalis, G., Vourgourakis, E., Spinthiropoulos K. & Lemonakis, C. (2016) "Healthcare Firms and the ERP Systems", Journal of Engineering Science and Technology Review, Vol.9, No. 1, pp 139 – 144.

[20] Motwani, B. & Sharma, R. K. (2016) "A Study on the Effect of Enterprise Resource Planning (ERP) on People of an Organization", Journal of Technology Management for Growing Economies, Vol. 7, No. 1, pp 73-84.

[21] Mitra P. & Mishra, S. (2016) "Behavioral Aspects of ERP Implementation: A Conceptual Review", Interdisciplinary Journal of Information, Knowledge, and Management, Vol. 11, pp 17-30.

[22] Bertocco M., Callegaro P. & De Antoni Migliorati D. (2006) "Ingegneria della qualità", CittàStudi, Torino.

[23] Cibien, M. (2016) "La Nuova ISO 9001:2015 e l' Evoluzione della Gestione per la Qualità", U&C , No. 1, pp. 17-36.

[24] Frascati Manual 2015: The Measurement of Scientific, Technological and Innovation Activities Guidelines for Collecting and Reporting Data on Research and Experimental Development. OECD (2015), ISBN 978-926423901-2 (PDF).

[25] Massaro, A. Maritati, V., Galiano, A., Birardi, V. & Pellicani, L. (2018) "ESB Platform Integrating KNIME Data Mining Tool oriented on Industry 4.0 Based on Artificial Neural Network Predictive Maintenance", International Journal of Artificial Intelligence and Applications (IJAIA), Vol.9, No.3, pp 1-17.

[26] Galiano, A., Massaro, A., Barbuzzi, D., Pellicani, L., Birardi, G., Boussahel, B., De Carlo, F., Calati, V., Lofano, G., Maffei, L., Solazzo, M., Custodero, V., Frulli, G., Frulli, E., Mancini, F., D'Alessandro, L. & Crudele, F. (2016) "Machine to Machine (M2M) Open Data System for Business Intelligence in Products Massive Distribution oriented on Big Data", International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 7, No. 3, pp 1332-1336.

[27] Massaro, A., Calicchio, A., Maritati, V., Galiano, A., Birardi, V., Pellicani, L., Gutierrez Millan, M., Dalla Tezza, B., Bianchi, M., Vertua, G., Puggioni, A. (2018) "A Case Study of Innovation of an Information Communication system and Upgrade of the Knowledge Base in Industry by ESB, Artificial Intelligence, and Big Data System Integration", International Journal of Artificial Intelligence and Applications (IJAIA), Vol. 9, No.5, pp. 27-43.

[28] "WSO2" [Online]. Available: https://wso2.com/products/enterprise-service-bus/

[29] Massaro, A., Maritati, V., Savino, N., Galiano, A., Convertini, D., De Fonte, E. & Di Muro, M. (2018) "A Study of a Health Resources Management Platform Integrating Neural Networks and DSS Telemedicine for Homecare Assistance", Information, Vol. 9, No. 176, pp 1-20.

[30] "KNIME" [Online]. Available: https://www.knime.com/

[31] Igel, C., Toussaint, M., Weishui, W. (2005) "RProp Using the Natural Gradient", In Trends and Applications in Constructive Approximation; Mache, D.H., Szabados, J., de Bruin, M.G., Eds.; Birkhäuser: Basel, Switzerland, Vol. 151, pp 259–272.

[32] De Campos Souza, P. V., Guimaraes, A. J., Araujo, V. S., Rezende, T. S., Araujo, V. J. S. (2018) "Regularized Fuzzy Neural Network to Aid Effort Forecasting in the Construction and Software Development", International Journal of Artificial Intelligence and Applications (IJAIA), Vol.9, No.6, pp 13-26.

**Corresponding Author**

**Alessandro Massaro**: Research & Development Chief of Dyrecta Lab s.r.l.