Optimization of purchasing business process in Moroccan public universities based on COBIT and Artificial intelligence techniques

Sakyoud Zakaria, Aaroud Abdessadek

Abstract

The Moroccan Court of Accounts pointes out often the Purchasing Business Process in public sector in general and public universities in particular as a budgetary loophole. In this research work we propose a new approach of organization and execution of the Purchasing Business Process PBP in Moroccan public universities. The approach we propose aims to set up a system of transparency and governance's best practices in order to minimize the financial and temporal wastes. The proposed approach enhances the actual purchasing business process with Control Objectives for Information and Related Technology COBIT 2019 guidelines and artificial intelligence techniques. The COBIT 2019 is intended to align the Purchasing Business Process with the conventional regulations and best practices in term of IT governance. Based on this guidlines an intelligent recommendation system is proposed to generate the optimal purchasing order and overcome this business process's limitations. The intelligent recommendation system is experimented on the purchasing of IT products (hardware, gear ...) in our faculty. Experimental results demonstrated that our approach contributes to limit the financial and time wastes, and promotes transparency in this business process.

I Introduction

As impact of the coronavirus 2019 pandemic, the world bank estimated that most major economies will lose at least 2.5 percent of the value their gross domestic product (GDP) [**world2020global**]. In developing countries like Morocco, the impact may be even more prominent and the government will inevitably look for loans to overcome the hard-hit on the economy. this means necessarily an increase of the external debt which puts the country in a vicious circle delaying the economic development. To limit the external, borrow, the government has to optimize the internal expenses and limit as much as possible the waste linked to bad governance. Our current research work fits in this endeavor. We propose a new approach of organization and execution of the Purchasing Business Process in Moroccan public universities. In fact, The Moroccan Court of Accounts pointes out often the purchasing business process in public sector in general and public universities in particular as a budgetary loophole. So, the proposed approach aims to set up a system of transparency and governance best practices in this business process.

The proposed approach enhances the actual purchasing business process with Control Objectives for Information and Related Technology COBIT 2019 guidelines and artificial intelligence techniques. The COBIT 2019 is intended to align the PBP with the conventional regulations and best practices in term of IT governance. Artificial intelligence techniques are in turn used to establish an intelligent recommendation system. This recommendation system takes as input a set of preferences and technical features, then executes a series of filters on local and crawled data to propose the suitable products in term of quality and cost. The final output of the recommendation system is hence an optimal Purchasing Order. In fact the actual state regulations on Purchasing Business Process stipulate that the optimal Purchasing Order is the one submitted by the lowest bidder vendor. The problem with this approach is the lake of guarantee

regarding the complaince with the market average costs. And the Purchasing Order can be excessive despite proposing the lowest price. Hence, the proposed recommendation system propose an alternate approache that define the optimal Purchasing Order based on the Utility Theory [**fishburn1968utility**].

The intelligent recommendation system is experimented on the purchasing of IT products and gears through. A purchasing scenario is set up in computer science department of the Faculty of Sciences El Jadida - Morocco. Experimental results demonstrated that our approach contributes to limit the financial and time wastes, and promote the transparency.

The paper is structured as follows:

- Section 2: presents the research question, and research method.
- Section 3: discusses related works and contributions, and positions the current paper.
- Section 4: presentation and modeling of the Purchasing Business Process in Moroccan public universities.
- Section 5: overviews the motivation scenario and the theoretical background of our proposition
- Section 6: presentation of the Intelligent recommendation system it's components and workflow
- Section 7: presents the implementation of the Intelligent Recommendation System. A prototype is developed and tested on a real purchasing use case in computer science department.
- Section 8: discusses the results
- Section 9: anticipates the perspectives of our future works

II Research question, research methodology

1 Research question

The key research question we aim to answer is: How can we promote transparency, limit financial wastes and reduces execution time in the Purchasing Business Process of Moroccan public universities, via information system governance and artificial intelligence.

The starting point of our approach, is to study the current Purchasing Business Process then formalizing this process in order to analyze and point out its weaknesses. This analysis enables us to approach the "to-be" Purchasing Business Process. Therefore, we designed an approach with two dimensions: the first dimension is the governance of information systems which brings the added value in the organizational component, the second-dimension component is more technical and consists of an artificial intelligence-based recommendation system.

2 Research methodology

The methodology we followed in this paper is Design Science Research (DSR) methodology for Information Systems [**peffers2012design**]. Design science research is a research paradigm in which a designer answers questions relevant to human problems via the creation of innovative artifacts, thereby contributing new knowledge to the body of scientific evidence. The designed artifacts are both useful and fundamental in understanding that problem [**hevner2010design**]. Correspondingly, this research work aims to provide an artificial intelligent based system as well as a set of COBIT's guide line to optimize the governance of purchasing process in Moroccan public universities. Concretely, the DSR implementation of this paper was inspired by [**peffers2007design**] [**hevner2010design**] [**hevner2004design**] [**hevner2004design**]. This implementation consists of six steps:

- Section 2: Problem identification and motivation: The problem identification is principally the analysis of the current Purchasing Business Process to point out the zones of inefficiency. A motivation scenario is also presented to express the motivations behind our proposition.
- Definition of the objectives for a solution: The objectives is the optimization of the Business Process from an information system governance point of view and also to build an AI based software platform to support an efficient and transparent execution of this Business Process.
- Design and development: The analysis, design and technological choices for the development of the platform are presented in section 4 and section 5. Fundamentally, we have adopted an Agile methodology [**highsmith2002agile**] all over the process of design and development. Agile methodology promotes and continuous iterative layers in development as well as testing throughout lifecycle of the project. To represent and formalize the Purchasing Business Process BPMN Business Process Diagrams [**white2004introduction**] have been used. This BPMN diagram is a conceptual contribution of this paper and fits in a larger BPMN modeling of all business process of the university.
- Demonstration The demonstration phase is represented by section 7, where we explore the developed plate-form and. This phase explains how the developer plate-form support Purchasing Business Process and simulate its execution in real-life use case.
- Evaluation: The developed platform has been evaluated and discussed regarding the initial objectifies.
- Communication: The research work has been communicated in the current paper.

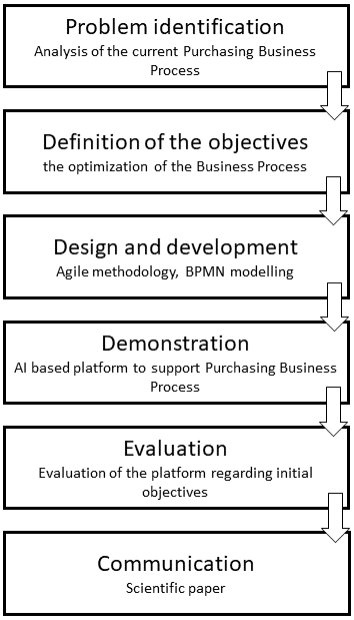

Figure 1: Implementation of Design Science Research Methodology for Information Systems

III Related works

In this section we will discusses related works and contributions. Then we will position our current research work according to this related work through a comparative table:

In [**hwangbo2018recommendation**], the authors proposed a system that extends the item-based collaborative filtering algorithm. In term of data this system combines online product click data (metadata) and offline product sale data. This combination represents the online / offline preferences of the customers. And these preferences are then traced over time. The recommendation system aims to offer substitute and complementary products by exploring product category information and also based on a scoring system. This field of this study is fashion products and the proposed system was adapted to this domain's characteristics. First, preferences for fashion products appear by using online click data and purchase data to generate recommendations. Second, the preference for fashion products tends to decrease over time. Finally, the product which the customer wishes to purchase is a product that replaces or supplements the product that the customer preferred before.

The authors in [**bang2020product**] also propose a product recommendation system, Unstructured data have the potential to be transformed into information that companies and users can employ, using appropriate processing and analyses. However, existing systems do not reflect the detailed information they collect, such as user characteristics, purchase preference, or purchase priority while analyzing review data. Thus, it is challenging to provide customized recommendations for various users. Therefore, in this study, we have developed a product recommendation system that takes into account the user's priority, which they select, when searching for and purchasing a product. The recommendation system then displays the results to the user by processing and analyzing their preferences. Since the user's preference is considered, the user can obtain results that are more relevant.

The research work proposed in [**martins2017intelligent**] fits into the emergence of data marketplaces as an alternative to traditional data commerce, as they provide appropriate online environments for data offering and purchasing. Nevertheless, as the number of available datasets to purchase increases, the task of buying appropriate offers is, very often, challenging. In this sense, we propose an intelligent decision support system to help buyers in purchasing data offers based on a multiple-criteria decision analysis. Experimental results show that our approach provides an interactive way that addresses buyers' needs, allowing them to state and easily refine their preferences, without any specific order, via a series of dataset recommendations.

The research work in [**yoshikawa2019product**] propose a recommendation system that suggests product alternatives. This research work points out the limitations of traditional recommendation systems that cannot recommend alternative products when existing products' reviews are negative about a specific product (e.g., price, battery). Hence, this paper, propose a purchasing recommendation system that focus on reviews data. In facts two kinds of reviews are analyzed from comments on e-commerce website: Complaint information and satisfaction information. The alternative products recommended are those who solving the complaint information on the product. The complaint informations are extracted based on product review analysis. This separation is performed by extracting negative information and positive information from product reviews. Then an reviews analysis outputs the higher-rated products which will be used to propose alternative products by solving complaints.

Research work	GUI	Dataset	Product information	Product reviews	Criteria ponderation	Similar products	Sentiment analysis	classification algorithms
1	console	Offline Online	Yes	No	No	Yes	No	-
2	console	Online	No	Yes	Yes	No	Yes	Morphological Analysis
3	console	Offline	Yes	No	Yes	Yes	No	SOM
4	console	Online	No	Yes	No	Yes	Yes	-
Current work	Web	Offline Online	Yes	Yes	Yes	Yes	Yes	SOM SVM

Figure 2: Comparative table of the related work

IV Purchasing business process in Moroccan public universities

1 Presentation

Purchasing in public sector is completely different and more challenging compering to private sector. In the public sector many procedures and guidelines need to be respected before making the final purchasing order. And this purchasing order need to satisfy specific state regulations. On another side, money has to be spent prudently and cleverly when purchasing in the public sector. Therefore, purchasing services and staffs are pressured into making efficient procurement decisions, and delivery time more efficient.

The purpose of the purchasing business process is to describe all acts that create or execute an obligation that entails an expense to be paid by the authorizing officer to meet the needs of the University. This process applies to all acts of engagement (Markets, Purchase Order, Contract, Agreement), normative references for this process are:

- Decree n ° 02-06 -388 of the 16 Moharrem 1428 (05-02-2007) bearing the regulation of the public markets
- Order No. 2-2471 of May 17, 2005 concerning the financial and accounting organization of universities
- Law 69.00 on state financial control of public enterprises and other sectors
- Dahir n ° 01-02-25 of the 03 April 2002 promulgating the law n ° 61-99 relating to the responsibility of the authorizing officers, the controllers and the public accountants
- Note 2-2471 DE / SPC of the DEPP concerning the financial and accounting organization of universities.
- Law 21-00 on the opening date and the closing date of the accounting years of universities.
- Decree 2-89-61 of 10 RABIA II 1410 (November 10, 1989) setting the rules applicable to the accounting of public institutions. (OO No 4023 of 6 December 1989).
- Royal Decree No. 330-66 of 10 March 1387 (April 21, 1967) on the General Regulation of Public Accounting. (Official Bulletin No. 2843 of April 26, 1967)
- Code of Obligations and Contracts (DOC)

- Paying Treasurer's Guide
- C.C.A.G.T
- General Accounting Standards Code (CGNC)
- C.C.A.G. MO

2 BPMN MODELLING Purchasing business process

based on the previous normative references we have established the PBMN [**white2004introduction**] model of the purchasing business process, BPMN stands for Business Process Model and Notation. BPMN is an ISO certified standard (ISO/IEC 19510:2013) for describing business process semantics, since its notation is generally easy to comprehend and is highly understandable for business and technical personnel [**ritter2011building**] BPMN provide businesses and organizations with the capability of understanding their internal business processes in a graphical notation in order to analyze and communicate these business processes in a standard manner.

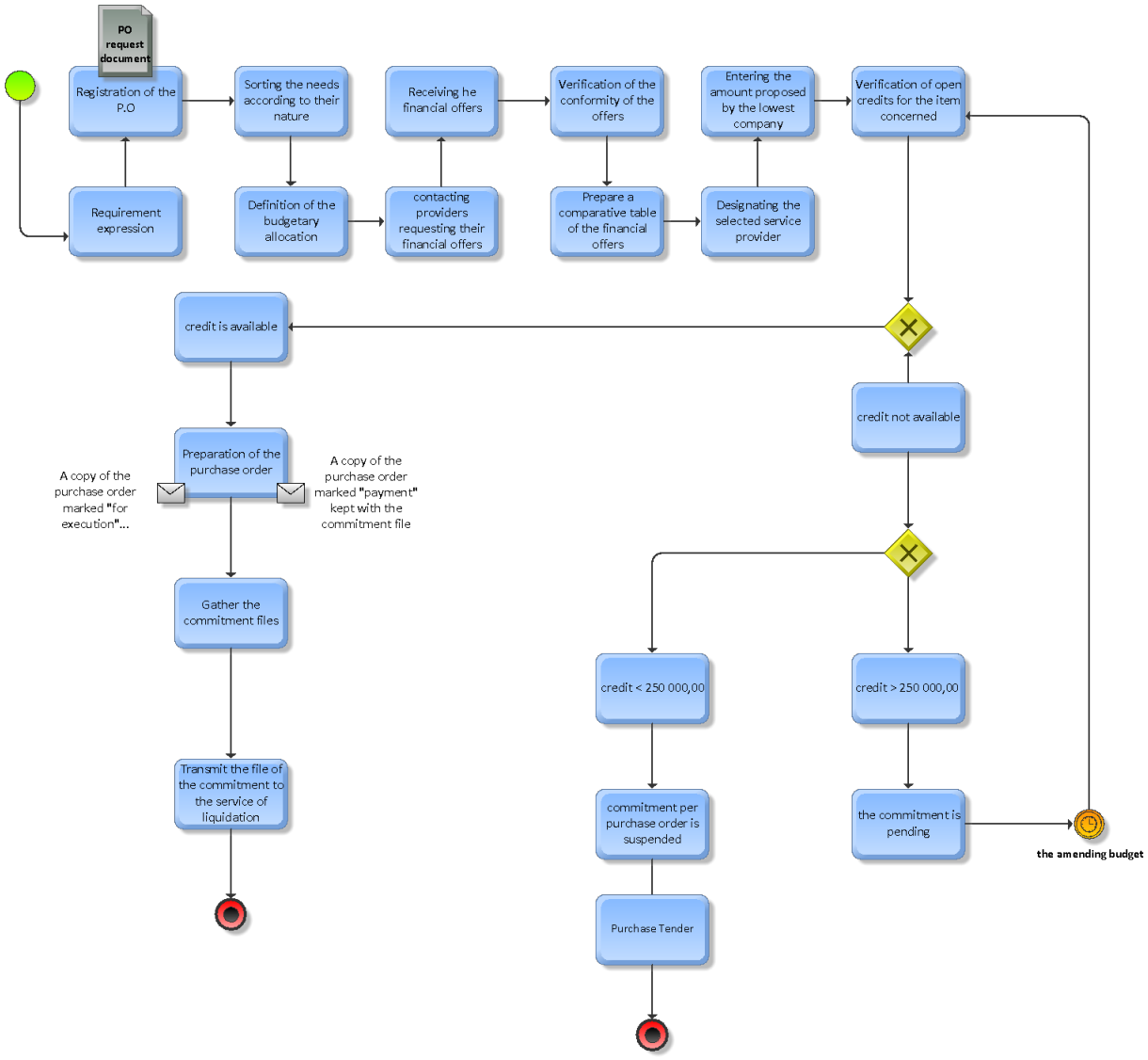

Figure 3: BPMN representation of the purchasing business process

V Limitations of the current PBP and challenging motivations

1 Motivation scenario

let us consider a scenario where the computer science department has, he identified a need for computer equipment, in particular powerful computer to run simulations we will not describe the execution of entire purchasing business process, we will be focusing on the establishment of the purchasing order:

Professor X is starting a project on the implementation machine learning technique on bigdata. He identifies the need of new laptop workstation with the following technical features:

- Graphics (GPU): NVIDIA GTX 1050 GPU with 4GB RAM
- Processing (CPU): 2.8GHz Intel Core i7-7700HQ
- RAM: 32GB 2400MHz DDR4 RAM
- Storage: 1TB SSD

A purchase order of tow Dell XPS 15 9560 laptops is established and communicated to the head of department. The purchase order is then sent to three vendors, vendor A, vendor B and vendor B, vendor to receive their financial offers. The Vendors A, B and C responds respectively with 1750$, 1.800$ and 1910$. the purchasing order find his way through the PBP and vendor A will be selected as he is the lowest bidder.

In fact, even if Vendor A has the lowest financial proposition, the proposed price exceeds the average price of the same product in Amazon by 250$. The actual business process doesn't cover this use case for tow raisons: first lowest bidder rule, second the absence of formal mechanism to assess how reasonable the prices are.

In addition to the financial side, the quality of the purchased reference is not assessed. And the possible technical problems and failures that may occur after sometime of use are not verified. In this case alternative product may be proposed even with minor differences with the initial product. And from an optimization point of view, we can check in the IT equipment stock to look for computers that meets approximately the initial criterion and requires only a maintenance effort.

2 Diagnosis of the current challenges

The diagnosis the current problem and position the proposed solution we will take a fragment from the purchasing business process BPMN model:

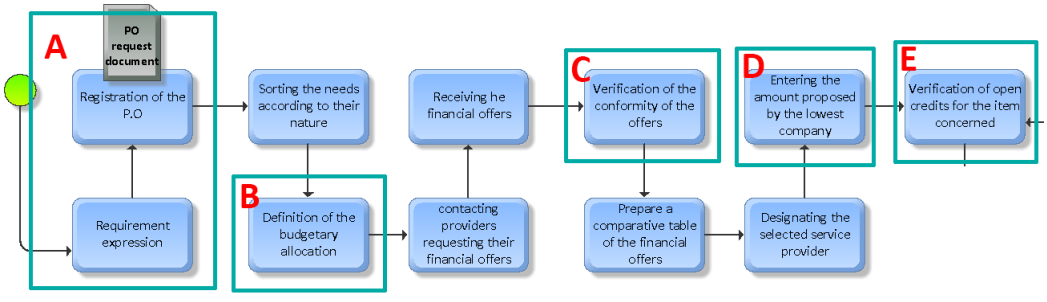

Figure 4: Diagnosis of the current challenges in the current the purchasing business process

A. The user expresses the needs through technical specification without any visibility about the future material and the expected budget in comparison with the market. On another hand the user can't consult the available products in local storage of the university.

B. The budgetary allocation is defined without a clear visibility about the actual prices in the market.

C. The verification of conformity is done comparing to initial requirements without consulting a formal technical repository for validation, and the preparation

D. The lowest bidder is selected. And as we stated before there is no assurance that it's offers is the good deal technically and financially.

E. the verification of the open credit is done after the consultation of the vendors. If the open credit is not sufficient, the purchasing order is blocked o reported while we have gone through several step and spends considerable time.

3 Positioning of the proposed solution

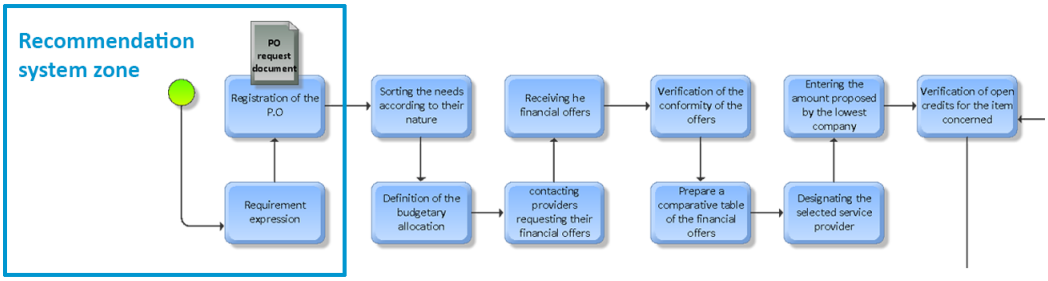

Figure 5: Recommendation system zone

As show in figure the recommendation system operates at the beginning of the process when the user express it's need. The recommendation system role is hence the establishment an optimal purchasing order. The recommendation system allows also to overcome the challenges A, B, C, D and E presented before:

A. The recommendation system provides the user with accurate and UpToDate informations about the future product. These informations concerns the technical characteristics but also the reviews of previous users. In addition to that the recommendation system propose also to the user the available products on local storage to check if a matching product exists.

B. The recommendation system provides the user with the actual market prices. At the end the defined budgetary allocation will be the closest possible to market costs.

C. The recommendation system verifies the conformity according to initial requirements and also propose alternative products by consulting online e-commerce data for technical characteristics as well as users' reviews.

D. The recommendation system replaces the "lowest bidder" new rule based not only on the lowest offer but the most reasonable offer comparing to the market in term of technical specification and cost.

E. Since the recommendation system provides the user with the actual market prices, the user can verification the budgetary allocation according to the open credit without waiting for vendors offers. This approach adds another layer of transparency because the vendors must align to the user budgetary allocation and not the opposite.

VI Theoretical background

1 Process optimization

Business process optimization is the automated improvement of business processes using pre–specified measures of performance representation the optimization' objectives [**tsakalidis2017towards**]. Generally, business process optimization aims to redesign a business process and generates various instances based on the same initial process specification. In term of implementation, business process optimization can involve techniques from different domains, not only managerial given the organizational aspect of business process, but also technological and engineering. Because in business process context, there is a need for a wider use of decision support systems based on artificial intelligence and expert systems [**gunasekaran2002modelling**]. Hence ITC can bring a great value to business process optimization in term of execution and also by offering tools to represent and simulate models of business processes in order to select the optimal design.

Our current research work fits in this perspective by using artificial intelligence as a decision support and PBMN models to design and simulate the purchasing business process.

2 Utility Theory in artificial intelligence

Utility theory study an theorize the choices of individuals and their decisions making process based on initial preferences The utility here reflects a subjective conception on an opinion quality and goodness compared to other options. Standard neo-classical economic theory describes utility as a set of internally-consistent assumptions about options in the wish to maximize utility [**fishburn1968utility**]. Utility theory has leveraged as one of the most dominant theories in economics as an underpinning of rational choice and game theory.

Utility functions are one of the elements of artificial intelligence solutions that are frequently mentioned but seldom discussed in details in AI articles. That basic AI theory has become an essential element of modern AI solutions. In some context, we could generalize the complete spectrum of AI applications as scenarios that involve a utility function that needs to be maximized by a rational agent.

In our recommendation system we are trying to mimic behavior of humans when maximizing their utility function. So, the role of this recommendation system is helping users to make decisions about final purchasing order, by proposing the most suitable IT equipment candidate. The recommendation system as utility function is

3 COBIT 2019

Information technology governance ITG is defined by [**calder2008iso**] as the system by which the current and future use of IT is directed and controlled. It involves evaluating and directing the use of IT to support the organization and monitoring this use to achieve plans and includes the strategy and policies for using IT within an organization. ITG and its frameworks provide managers with the structures considered necessary to facilitate IT services for academic and business process [**tawafak2020governance**].

In fact, the overall aim of ITG is to purify business processes and provide justifiable road map to organization strategy. Control Objectives for Information and related Technology 2019 (COBIT 2019) [**information2018cobit**] is an ITG framework intended for this purpose. The first principle of COBIT is that all IT related activities should support generating value for the enterprise. COBIT 2019 core model figure X illustrate the 5 domains covered by COBIT 2019 code model is composed of [**cobitdesignguide2019**]:

- Direct and Monitor (EDM): the governing body evaluates strategic options, directs senior management on the chosen strategic options and monitors the achievement of the strategy.

- Align, Plan and Organize (APO): addresses the overall organization, strategy and supporting activities for IT
- Build, Acquire and Implement (BAI): treats the definition, acquisition and implementation of IT solutions and their integration in business processes.
- Deliver, Service and Support (DSS): addresses the operational delivery and support of IT services, including security.
- Monitor, Evaluate and Assess (MEA): addresses performance monitoring and conformance of IT with internal performance targets, internal control objectives and external requirements.

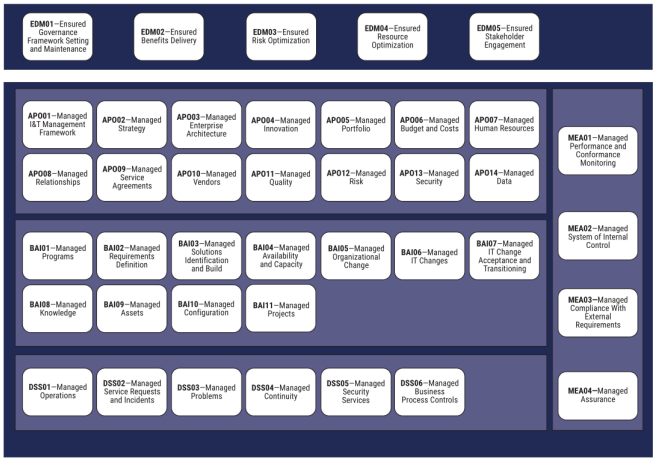

Figure 6: COBIT 2019 code model

Our research work fits in the first domain Align, Plan and Organize (APO), and uses more tow Governance and Management Objectives GMO which are in direct relation of this paper purpose these tow GMO are:

- APO06 (Budget and Cost Management): Foster a partnership between IT and enterprise stakeholders to enable the effective and efficient use of IT-related resources and provide transparency and accountability of the cost and business value of solutions and services. Enable the enterprise to make informed decisions regarding the use of IT solutions and services.
- APO10 (Managed Vendors): Optimize available IT capabilities to support the IT strategy and road map, minimize the risk associated with nonperforming or noncompliant vendors, and ensure competitive pricing.

VII Design of the Intelligent recommendation system

1 BPMN workflow of the recommendation system

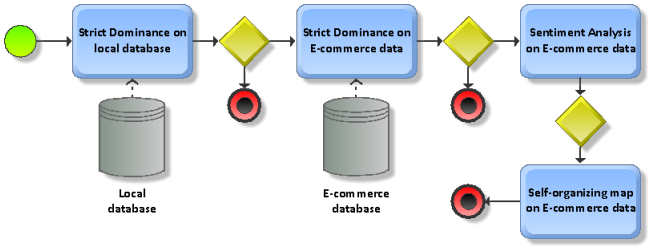

Figure 7: BPMN workflow of the recommendation system

The workflow illustrated in this BPMN model present the filtering process executed by the used when interacting with the recommendation system. For every interaction the workflow is executed as follow:

- Workflow is trigged when the user provides the initial requirements
- Based on the initial requirements a Strict dominance on a local database and the resulted products are presented to the user, if the user is satisfied the workflow is aborded else the used pass to step 3.
- Based on the initial requirement E-commerce data are scarped and a strict dominance is performed on the scraped data, the resulted products are presented to the user, if the user is satisfied the workflow is aborded else the used pass to step 4.
- The E-commerce data are subject to a sentiment analysis to reclassify them based on users' reviews, the resulted products are presented to the user, if the user is satisfied the workflow is aborded else the used pass to step 5.
- The resulted products from step 4 are processed by Self-organizing map classifier and provide the user with a matrix to explore the similarities between products and finally the workflow is aborded.

2 UML modeling of the recommendation system

2.1 UML sequence diagram

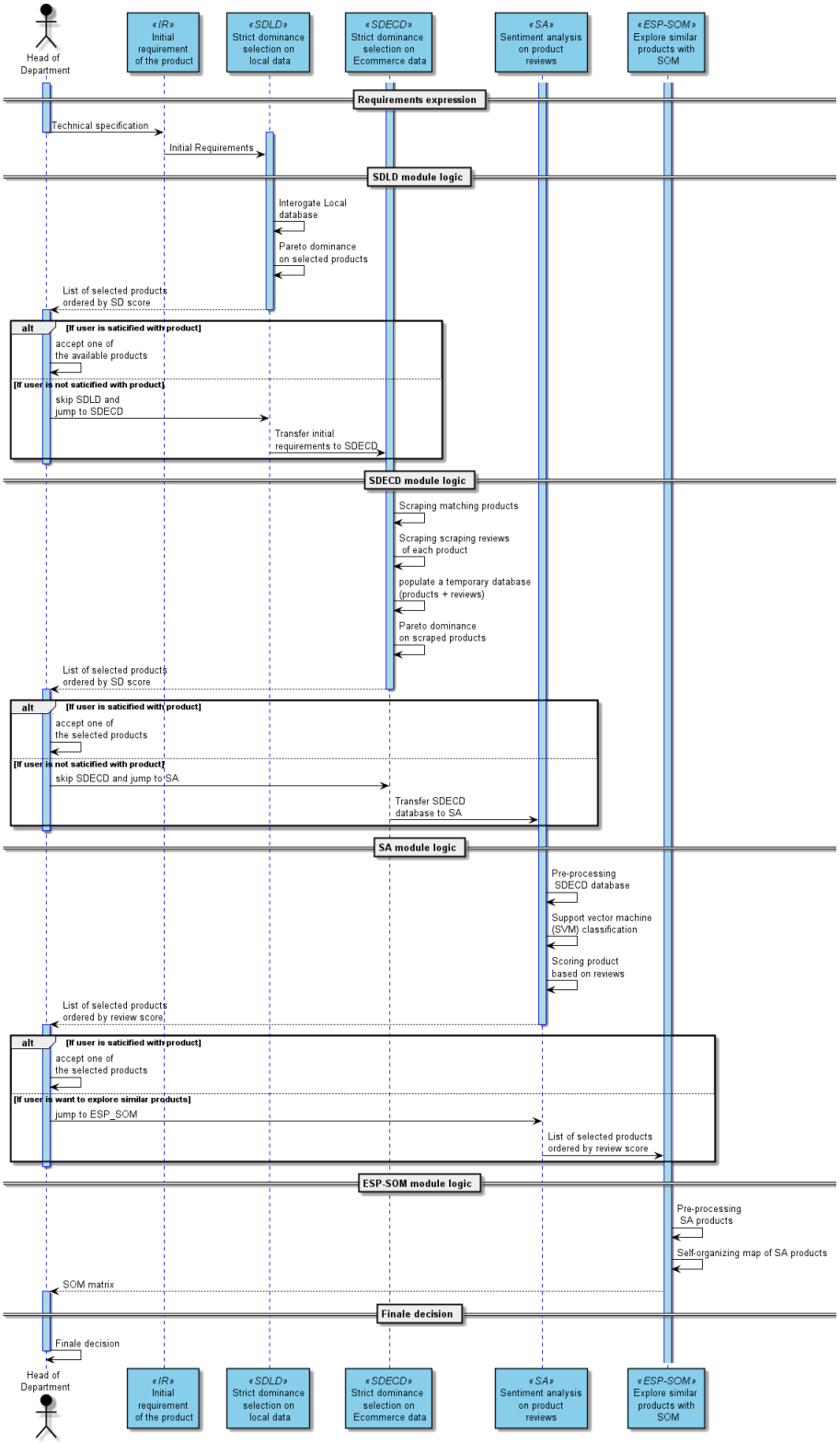

Figure 8: Sequence diagram of the Intelligent recommendation system

2.2 UML Component diagram

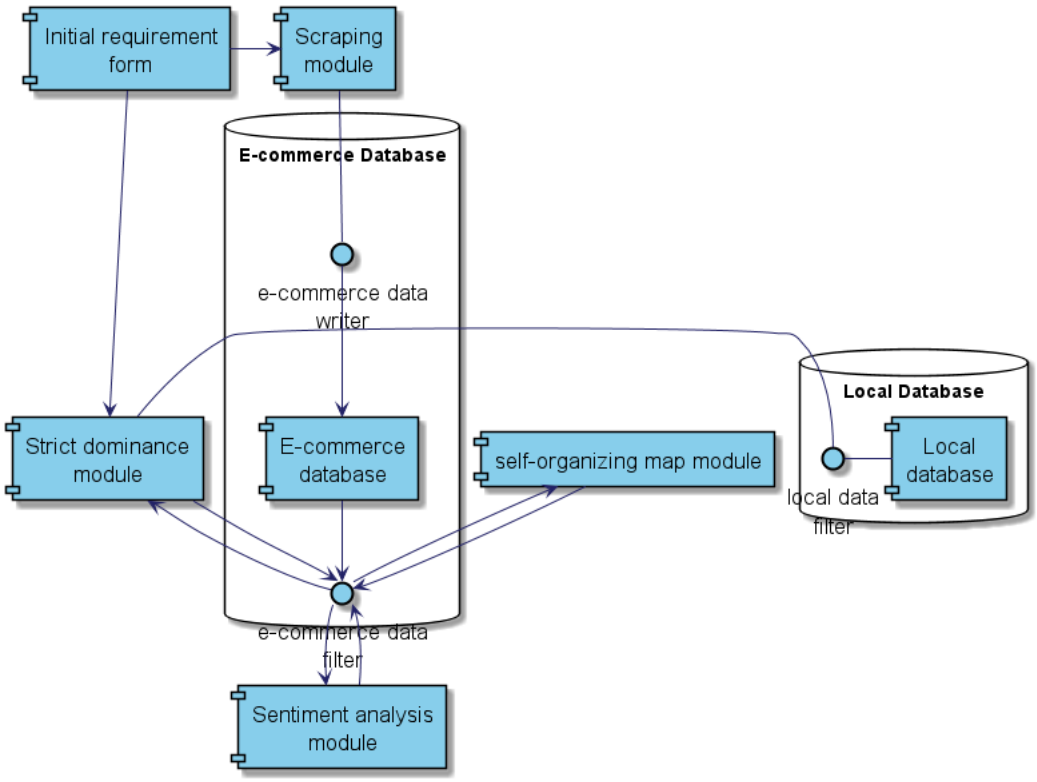

Figure 9: Component diagram of the Intelligent recommendation system

3 System components presentation

3.1 Initial Requirements

The initial requirements represent user technical preferences expressed as technical features (Eg. Processor generation, RAM capacity ...). These purchasing criteria are considered as a top priority of the final product. For more flexibility and in order to explore similar products at the end of the process, the entered values of the criteria are allowed a margin of maneuver. For example, the user expresses that the minimum RAM capacity is 16 Gb, which allow recommendation system to propose product with RAM higher then 16Gb and balancing with other technical features.

3.2 Strict dominance module

This module performs a data refinement and selection using Game theory 's strict dominance concept. The mean purpose of this this module is the calculate a strict dominance score for each one of the products given as input then classify them based on this score. In the recommendation process, we callout this module twice, first to classify products issued from the local database, second to classify the scraped data from e-commerce website. In both case this module performs a selection of products based on initial requirements. Then the products are refined using the strict dominance logic.

In game theory, a strategy if called strictly dominant when it is better than all another strategy. In the opposite side a strategy if called strictly dominated when all other strategies are better [**straffin1993game**]. In our proposition we are using the notion of dominance to calculate a dominance score to classify and the eliminate the dominated products. Which mean products with lower value on all attributes than all the others will be given a low score and vice versa. In other words, the strict dominance is used to narrow down the field of choices to the real candidates.

In order to implement the concept of strict dominance, we relied on pareto dominance algorithm. Pareto dominance enable the comparison of candidates based on two or more criteria and provide proper candidate optimal scheme for decision-makers decision [**zhang2018pareto**]

Let Up the Utility functions of the product. The Utility function Up represent the preferences between products based on criteria p appertaining to preferences set P:

A product X is dominated by product Y if the tow following conditions are satisfied:

- Y is better than X in every criterion, meaning $\forall p \in P, \quad Up(X) \geq Up(Y)$
- a product is considered pareto-optimal if it is not dominated by any other product in the selected list.

3.3 Local database

This database constitutes an inventory of the "out of service" IT equipment. This database centralizes maintenance information and facilitates the follow-up of the of maintenance processes.

3.4 E-commerce database

This database is filled using web scraping technic [Slamet, Cepy, et al. "Web scraping and Naïve Bayes classification for job search engine." IOP Conference Series: Materials Science and Engineering. Vol. 288. No. 1. IOP Publishing, 2018] on predefined e-commerce websites. Based on the initial requirements, a URL representing search pattern is constructed for each e-commerce website.

Upstream we have developed a middleware for each website that enable the browsing of its DOM and the extraction of product's information from its html tree. The crawled information is stored in two different table, first the technical features of each product is stored in the corresponding table. Then for every product a set of rows representing its reviews is created and stored in the corresponding table with reference to its parent product. For sentiment analysis efficiency only products with more than 10 reviews are selected

3.5 Sentiment analysis module

In the natural language processing (NLP), Sentiment Analysis (SA) has become one of the many fields of computational studies [**bautin2008international**]. We used the sentiment analysis to calculate an appreciation score for each product. The appreciation score is calculated trough the categorization of positive and negative customer reviews about this product. The polarization of positive and negative reviews is done by the mean of supervised learning model. The model is trained on one of the most reliable and large datasets used in this field which is amazon reviews dataset.

In fact, Amazon is one of the largest E-commerce sites as for that there are innumerous number of reviews that can be seen. We used data named Amazon product data which was provided by [**ni2019justifying**]. Amazon product data is a subset of a large 233.1 million Amazon review dataset. The current data includes reviews in the range May 1996 - Oct 2018. The dataset reviews include ratings, text, helpful votes, product description, category information, price, brand, and image features. The dataset is unlabeled and is widely used as reference to analyze user's positive and negative sentiments on products.

In our analysis we have used tow subset of the global data: "Electronics" and "Office Products". The total reviews in the principal dataset for this tow subsets is 26,575,666 reviews. In our prototype we have used portions of the "Small subsets" Offred by [**ni2019justifying**] for experimentation purpose. Hence the total review number used to train the sentiment analysis module is 1,539,947 reviews. A typical json object of a review in Amazon product data is:

```
{
  "reviewerID": "ID of the reviewer",
  "asin": "ID of the product,",
  "reviewerName": "name of the reviewer",
  "vote": 5,
  "style": {
    //a disctionary of the product metadata
    "key1": "values1",
    "key2": "value2"
  },
  "reviewText": "text of the review",
  "overall": "rating of the product from 0 to 5",
  "summary": "summary of the review",
  "reviewTime": "time of the review"
}
```


Figure 10: json structure of review object

To stablish to sentiment analysis workflow on our subset we have adopted the methodologies used in similar works [**elli2016amazon**], [**xu2015sentiment**], [**rain2013sentiment**], [**bhatt2015amazon**]. The workflow contains 4 steps:

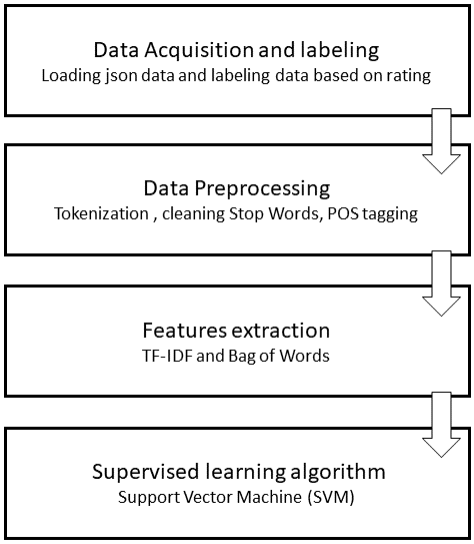

Figure 11: Sentiment analysis workflow

- Data Acquisition: As stated before, we have acquired our data form 2 subsets, each subset came in json format. Regarding the important number of reviews, it is very hard to label data manually. Hence, the data were labeled programmatically based on amazon reviews rating. This rating is represented by a 1 to 5 rating. A rating of 1 or 2 starts is considered positive while a rating of 4 or 5 is considered positive, 3 starts rating are considered neutral hence not processed in the labeling.
- Data Pre-Processing: This step performed trough three stages. First the tokenization which consists on untying a string sequence into separates entities called Tokens. Tokens can be individual words, phrases sentences. Second the Stop Words elimination which consists on removing punctuation and all unnecessary strings in text mining. This will enhance the accuracy and the performance of the analysis. Finally, the POS tagging which helps the model to understand the grammatical function of words [**pasupa2019thai**].

- Feature Extraction: This step performed trough Bag of Words and TF-IDF technics. The Bag of word is created from nouns and adjectives selected based on the previous the POS tagging. The bag of word is then used to generates term frequency as text characterization feature. This feature that we will be used as input in TF-IDF stage. TF-IDF in turn is will weighs the extracted term frequency and also inverse document frequency. a TF and IDF scores are then calculated for each word as well as the TF*IDF weight for each word. The value TF*IDF weight determine the rarity of words, lower TF*IDF value means o heigh frequency of the word.

- Support vector machine SVM : Support Vector Machine SVM is a machine learning method that has become exceedingly popular for neuroimaging analysis in recent years. Because of their relative simplicity and flexibility for addressing a range of classification problems, SVMs distinctively afford balanced predictive performance, even in studies where sample sizes may be limited [**pisner2020support**]. SVM has also proven it accuracy in sentiment analysis classification especially when trained over the amazon reviews dataset. In [**nasr2017building**] and [**haque2018sentiment**] the accuracy of SVM has being demonstrated comparing to other supervised learning algorithm.

3.6 SOM analysis module

The Self-Organizing Map (SOM), with its variants, is the most popular artificial neural network algorithm in the unsupervised learning category [**kohonen2012self**]. It provides a topology preserving mapping from the high dimensional space to map units. Map units, or neurons, usually form a two-dimensional lattice and thus the mapping is a mapping from high dimensional space onto a plane. [**kohara2018clustering**]. Each unit has a set of weights which is compared to input data in order to find the nearest unit (commonly measured with Euclidean distance) to an input data, this unit is called the Best Matching Unit BMU. The weights of the BMU and its neighbors are then adjusted. At the end of the training phase. A topological map is established (commonly 2D map). The topological neighborhood of the units reflects the similarities in the input data that activated them. Figure X illustrate the concept of the competitive learning algorithm [**kohonen2013essentials**]

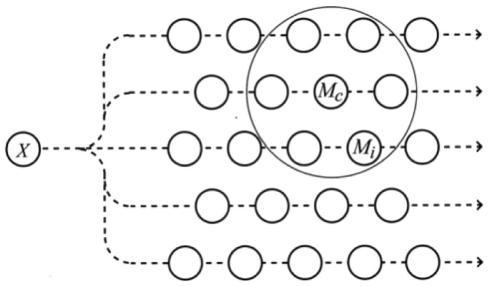

Figure 12: Illustration of a self-organizing map. An input data item X is broadcast to a set of models M i , of which M c matches best with X. All models that lie in the neighborhood (larger circle) of M c in the grid match better with X than with the rest [**wandeto2017detection**]

SOM was used as purchase decision making support in many research works [**kohara2006purchase**] [**martins2017intelligent**] [**kohara2013selecting**] [**cho2014clustering**]. in this perspective we use SOM in our recommendation system to give the user more purchasing options by exploring new opportunities. In fact, at the end of Sentiment analysis on primary data, the user has a first classification of the selected products based on technical features as well as review score, the purchasing order can be established based on this classification. Otherwise, the user can have a deeper discrimination of the classification by discovering similarities among the selected product using SOM analysis. for example, the product X rank is 5/10 but in SOM 2D grid this product is located near product A which rank is 1/10 and they are located in the same cluster. Hence the choice of product X may present a potential benefit.

VIII Experimental implementation

1 Technical specification

The implementation prototype was developed with Django framework. Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design [**forcier2008python**]. Django is based in the MVC design patten [**leff2001web**] and enable a fluent integration with a large number of python libraries.The following tables resume the technical specification and the used libraries.

Computer	• 15” Laptop • CPU-Intel i7-8650U vPro 8th Generation • 16G RAM based
Operating System	Windows 10 pro 64bit
Programming language	Python 3.6

Crawling	Beautiful Soup 4.9.0
Natural Language Processing	Spacy 2.3.5
Self-organizing Map	MiniSom
Feature Scaling	sklearn
Others	numpy, matplotlib, pandas

2 Data preparation

2.1 Local database

To prepare our demonstration we did an inventory in our computer science department to list and identify the existing desktop computers and laptops. Through this inventory we have built a local database that references the technical specifications of each product as well as its operating state. we mean here by operating state wither the product is functional or not. And if the product is out of service, identify the required level of maintenance.

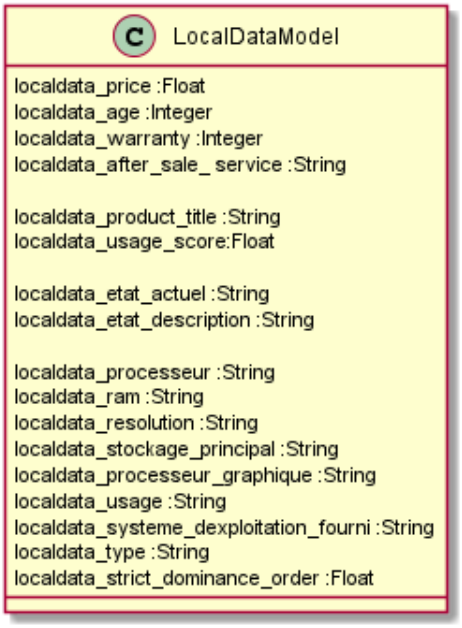

Figure 13: Local database structure

2.2 E-commerce scraped database

Leveraging Data from the Web presents both researchers and practitioners with big challenges as well. Apart from the need to learn and deploy new tools and technologies capable of accommodating Big Data, researchers and practitioners intending to use Web Scraping in their research projects need to comply with a number of legal and ethical requirements[**krotov2018legality**]. The development of our scraping module is done in respect to the following ethical requirements:

- Terms of Use: Respect to the website we used
- Purpose of Web Scraping: none profit usage
- Damage to the Website: respect the time out and amount of data
- Individual Privacy: we don't scrap any personal information
- Organizational Privacy and Trade Secrets: we don't scrape o have access to any privacy or thread secrets

2.3 Data enhancement

Due to a technical constraint, the reviews of each data are enhanced by external reviews from the amazon review database. In fact, when accessing the detail page of the product only first ten products are displayed, and to access other reviews a manual navigation must be done

3 Experimental use case

3.1 Initial Requirements

Int this from the user fill the minimum required preferences as well as the approximative budget allocated. further informations are also given about the purchasing order namely the subject and a short description of the purpose of the future purchasing order.

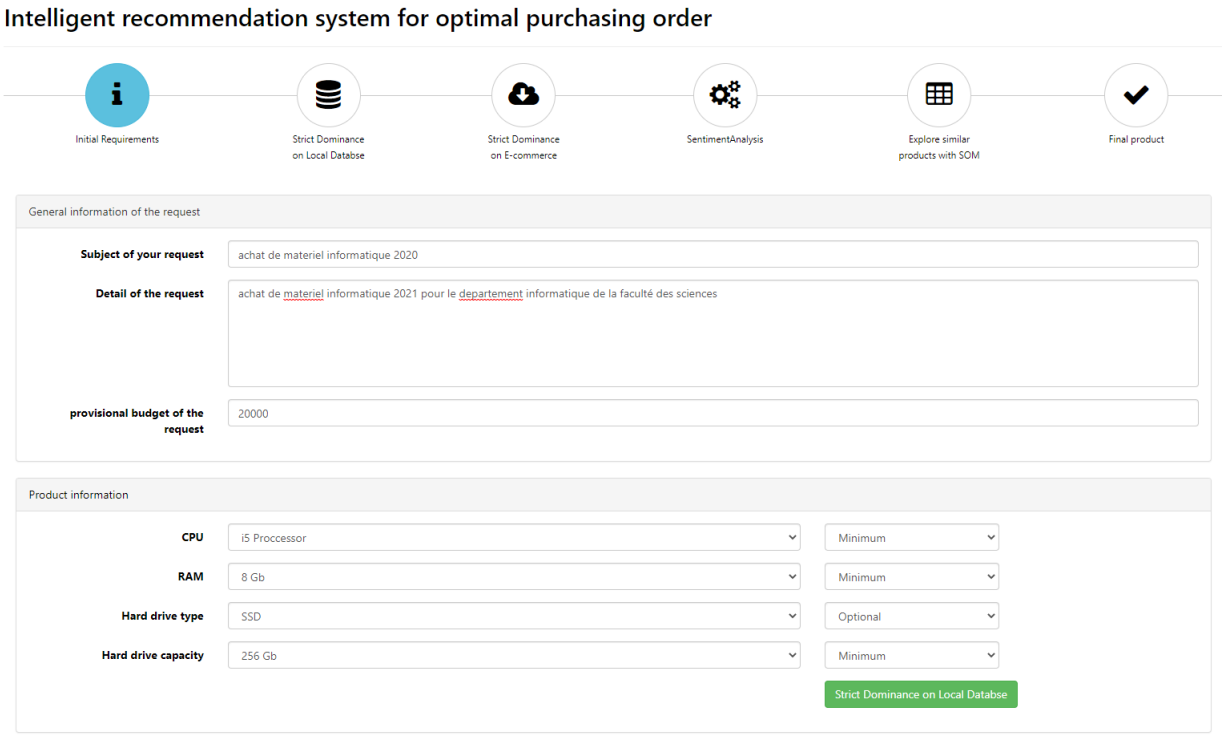

Figure 14: Initial Requirements

3.2 Strict Dominance on local database

In this step a selection query is performed based on the initial requirement data entered by the user. The filter is hence composed of the desired technical specification and the ponderation of each specification. Then a strict dominance score is calculated and the selected products are classified base on this score. Top 10 products are then presented to the user.

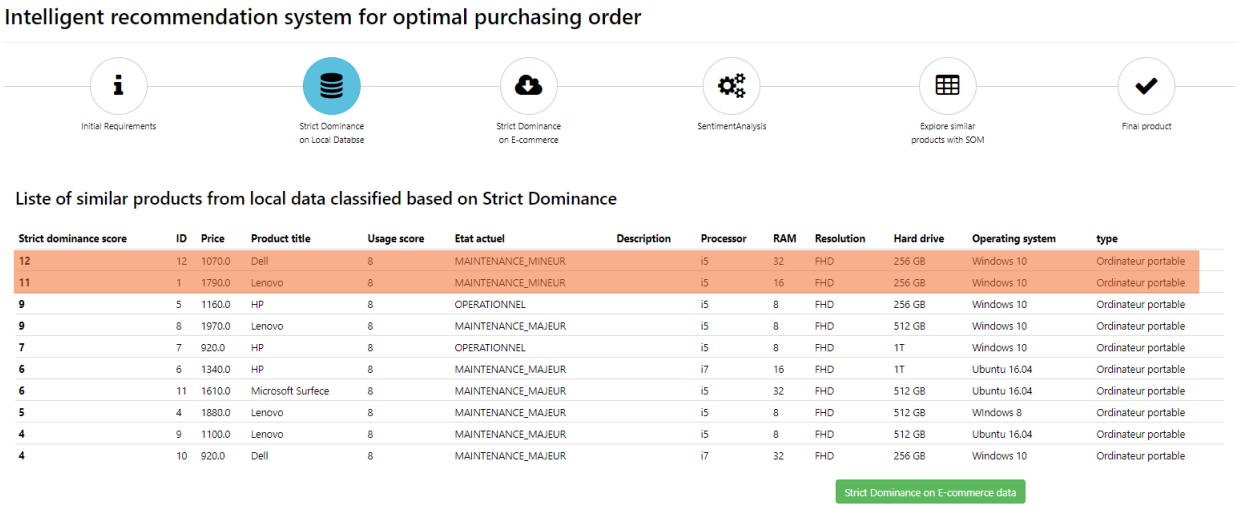

Figure 15: Strict Dominance on local database results

If the user is satisfied with the result, he will communicate the product ID to the administration to get the product through a recovery procedure. To browse the other step of the demonstration we considered that the user will go through all steps. The next one is the strict dominance on e-commerce data.

3.3 Strict Dominance on e-commerce data

In this step the initial requirement data entered by the user are used as entre for a scarping. This scraping module use the technical specifications and their ponderations to scrape data from e-commerce website. The strict dominance score is calculated this time for the selected products from the scraping module.

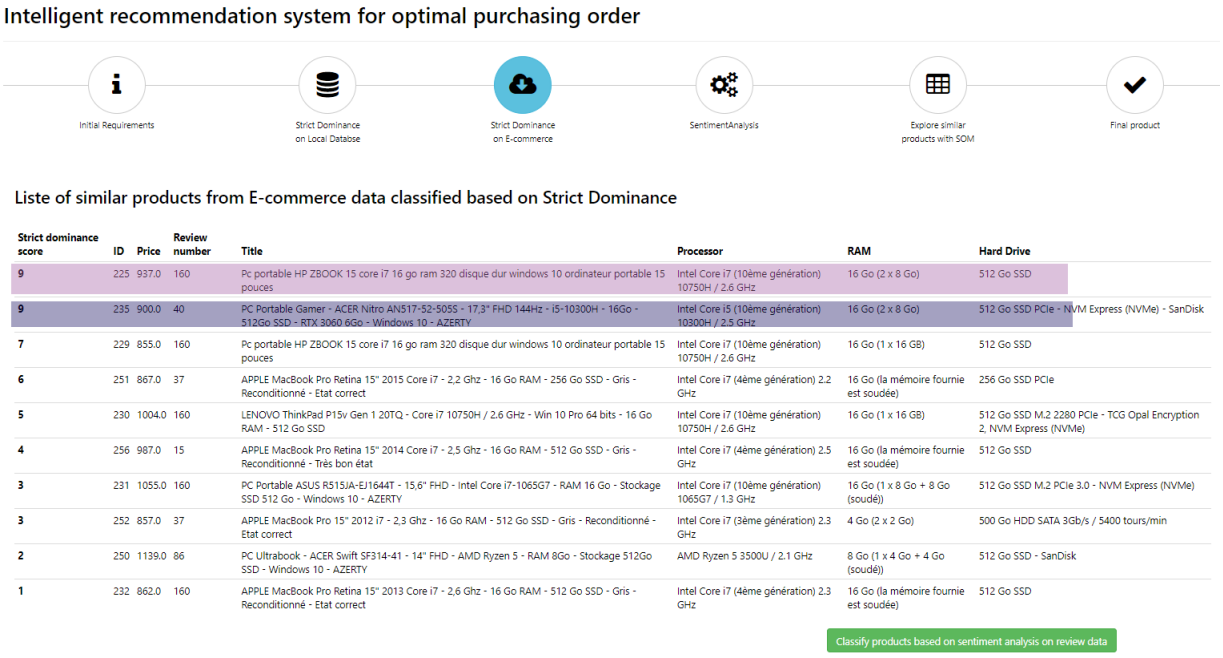

Figure 16: Strict Dominance on e-commerce data results

For each product scraping module import available product reviews and store them to be in the next step.

3.4 Sentiment analysis

For each product the Sentiment Analysis module perform a natural language processing on each one of the previously selected products. This analysis calculates the sentiment analysis score and order the product based on this score.

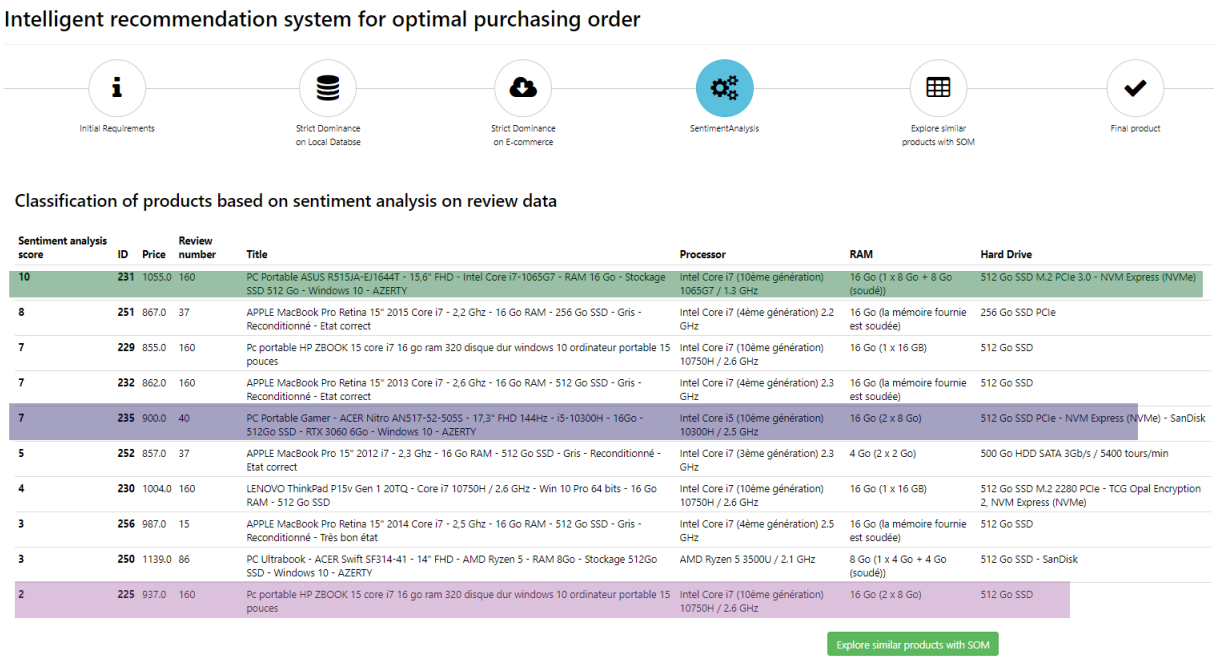

Figure 17: Sentiment analysis results

3.5 Explore similar products using SOM analysis

In this step a SOM map is presented to the user with nods representing products IDs. For each product nearby counterparts more similar than those who are further away. This enables the user to explore similar products and cheque each time more details by entering the ID of the product in the appropriate form.

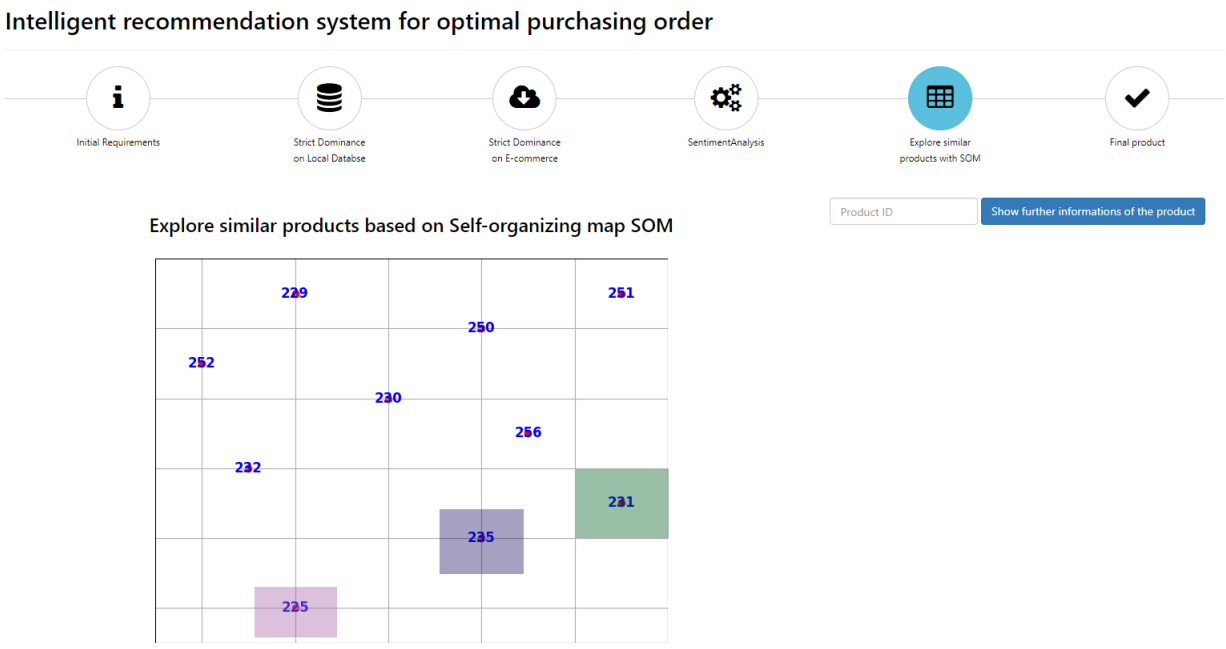

Figure 18: SOM map of the selected products from sentiment analysis step

3.6 Final product

the final product selected by the user. In the next section the process who lead to this choice will be explored and discussed.

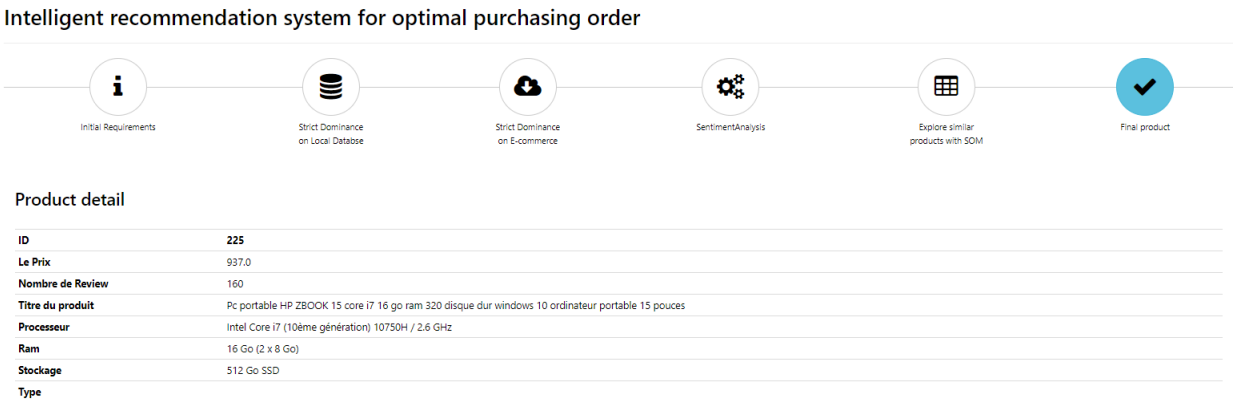

Figure 19: Final product informations

IX Results and discussion

To analyze the efficacy of our intelligent recommendation system we have, we simulated above a recommendation request of the computer science head department. The recommendation system concerns efficient laptops for developing new internal application. We will discuss here the execution and the results of each step of the recommendation process:

1 Initial requirements

The initial requirements entered by the user are:

Criteria	Value	Ponderation	Description
CPU	i5	Minimum	At least an i5 processor can be i7
RAM	8 Gb	Minimum	At least 8 Gb of RAM
Hard drive	SSD	Optional	Can be DDR
Capacity	256 Gb	Minimum	At least 256 Gb

2 Strict Dominance on local database

When performing the strict dominance on local data, the algorithm favorize products with miner maintenance with respect the i5 minimum processor value and maximizing the value of RAM.in this step the ponderation of the price is neutralized because the products are already purchased and are available internally. Then to top scored products are Product 12 and product 1 of the database.

3 Strict Dominance on E-commerce data

The scraped data are then object to a strict dominance processing, as demonstrated before the algorithm give a score of 9 to products 225 and 235.Product 225 has an i7 processor and 16 Gb in RAM with a price of 937.0, this product seems to be a good deal because it has a good configuration and reasonable price. Product 235 is relatively cheaper then 225 but for an i5 processor.

4 Sentiment analysis

The sentiment analysis on the scraped product gives a quite different result from the previous one. It's remarkable that the product 225 descended in the ranking with a sentiment analysis score of 2. While the product 235 kept a good ranking with a score of 7. It's if clear that from a user's reviews point of view the product 231 is a good deal with a score of 10 extracted from 160 review.

5 Explore similar products using SOM analysis

When exploring the SOM map, we can see that the product 225 is relatively far from other products, if we combine this information with the results of sentiment analysis, we can conclude that product 225 represent a good deal in term of technical specification and user's reviews. We can also notice that product 235 and 232 are the closest product to the 225, which mean they car be a protentional choice especially 235 which was recommended by strict dominance.

X Conclusion and perspectives

As initially stated in this paper, the main goal of the current research work relies in the optimization of the purchasing business process in Moroccan public university in term of transparency and budgetary optimization. In order to achieve this goal, we have used a functional-technical approach leading to the development of an intelligent recommendation system that assists concerned users in the choice of optimal equipment. All this, in total alignment with Moroccan normative references, and also with COBIT's guidelines in term of information system governance.

Hence the major contribution proposed int the current paper can be summarized as follow:

- The modeling of business processes in public university is established using BPMN in accordance with official regulations. The set of BPMN models constitute a powerful repository for business process execution but also for further optimization
- Governance aims in general to reduce budgetary wastes, our recommendation system demonstrates a technical and methodological approach enabling this feature.
- The implementation of artificial intelligence techniques can bring a great value in term of transparency and fluidity in purchasing business process execution