NOVA IMS Information Management School

MEGI

Mestrado em Estatística e Gestão de Informação

Master Program in Statistics and Information Management

Intelligent architecture to support second generation general accounting

Rita Moura Martins Galiére da Cunha

Dissertation presented as partial requirement for obtaining the Masters' degree in Statistics and Information Management, Specialization in Information Analysis and Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

INTELLIGENT ARCHITECTURE TO SUPPORT SECOND GENERATION GENERAL ACCOUNTING

By

Rita Moura Martins Galiére da Cunha

Master Thesis presented as partial requirement for obtaining the Master's degree in Statistics and Information Management, with a specialization in Information Analysis and Management.

Advisor: Professor Doutor Vitor Manuel Pereira Duarte dos Santos

November 2022

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledge the Rules of Conduct and Code of Honor from the NOVA Information Management School.

[student signature]

[place, date]

ACKNOWLEDGEMENTS

Firstly, I would like to thank Professor Vitor Duarte dos Santos for his dedication and interest, his commitment, his guidance and his competence during the preparation of this thesis. The professor accompanied me one hundred percent in this project that closes another chapter of my school life.

I would also express my gratitude to my family and my boyfriend who accompanied me on this journey and gave me the strength to continue. All this path that I have traced will take me far on a professional and personal level, a path that I have traced together with my family who have been present at all stages.

I am grateful to have had the opportunity to write a thesis, many people do not have the possibilities and opportunities to attend this course and become a master. Therefore, I am grateful for all the work I did because I know that later, I will be rewarded.

ABSTRACT

This study aimed to innovate the world of accounting software. After so many years, accountants are faced with an unbelievable amount of work, which is not always productive, effective and efficient for both the accountant and the company that provided him with the data required to carry out the accounting. There is already accounting software with various automation processes, from ornamentation to profitability analysis and management reporting. There is also software that is updated in accordance with the accounting laws, i.e., the platform changes its mechanisms according to the changes in the law.

Despite the existence of this software, manual work remains, and the amount of information accountants are faced with is still very large. It is difficult for accountants to do a 100% reliable job with so much information and data they have. One of the most common situations in the accounting world is undoubtedly the miscalculation or forgetting of some financial or non-financial data found in accounting operations (income statements, balance sheets, etc.). To render accounting operations efficient, effective and productive, error-free and 100% reliable, an intelligent architecture has been developed to support second generation general accounting. This architectural design was developed with a view to make the existing software smarter with the help of artificial intelligence.

A study was carried out on accounting keys and concepts, on AI and main process automation techniques to build the model. With these studies it was intended to acquire all possible requirements for the creation of the architecture. Towards the end of the thesis the model was validated.

KEYWORDS

Intelligent architecture; general accounting; artificial intelligence; RPA and OCR; Machine Learning

INDEX

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AA	Autonomous Agents
AI	Artificial Intelligence
AIPA	Augmented Intelligent Process Automation
AIS	Accounting Information System
CNC	Accounting Standards Committee
DSR	Design Science Research
ERP	Enterprise Resource Planning
IPA	Intelligent Process Automation
KDT	Knowledge Discovery Database
KNN	K-nearest neighbor
MIS	Management Information System
MRP II	Manufacturing Resource Planning
NCM	Accounting Normalisation for Micro-entities
OCR	Optical Character Recognition
PAA	Primavera Accountant Automation
PDF	Portable Document Format
PSDB	Public System of Digital Bookkeeping
RPA	Robotic Process Automation
SAGE	Software Aided Group Environment
SNC	Accounting Normalisation System
SNCP	Public Accounting Normalisation System
SVMs	Support Vector Machines
VAT	Value Added Tax

1. INTRODUCTION

Accounting is a management tool with four functions, the recording function, the control function, the evaluation function and the forecasting function. Its purpose is to provide meaningful data on company and business assets. In this way, its addressees (managers and decision-makers) get to know the economic and financial situation of the company, which enables them to analyse, create strategies and take management decisions, trying to minimise management failures.

As a support to accounting there are programmes that perform accounting operations and analysis of results (management analysis) already widely used by accountants and companies. Despite the innovation made through these software's, manual work remains, and mistakes continue to be made, often caused by the large amount of information available.

Thus, it is intended to create an architecture of an intelligent system to support general accounting.

1.1. Context

Accounting is divided between cost accounting, also known as management or internal accounting, and financial accounting (general or external accounting).

There are quite significant differences between them, being that cost accounting has as its main focus the evaluation of the company's performance for the creation of new strategies; it is intended for company managers or financial directors; it is an optional accounting; it does not have defined periodicity; its operations are: balance sheets and income statements and it has a future application. As for financial accounting, its objective is the registration and control of operations and the preparation of financial statements; it is intended for investors, suppliers, shareholders, customers and the State; it is a mandatory accounting, has defined periodicity (monthly, quarterly or annually); its operations are financial statements, and its application is immediate.

For this research the accounting to be studied will be financial accounting. Within this we can find transactions such as: the balance sheet, income statements, equity statements, cash flow statements and annex to the balance sheet and income statements.

Accounting proves to be a very complex world and includes a wide range of methods such as recording, classifying, summarizing, analysing and reporting financial transactions over an accounting period, financial position and cash flows to oversight agencies, tax collection entities and regulators. Accounting is complex in that the income, assets, liabilities, income statement, balance sheet and statement of cash flows are essential to corporate decision-making. Good decision-making can lead to better results and consequently increase the growth and success of the company as well as help to understanding what strengths and weaknesses are present in the company.

The platforms already in existence allow for the performance of various accounting activities and are currently a great help to the world of accounting. However, most of the data is still entered manually, which can lead to errors such as: missing values due to forgetfulness or another incident, entry of wrong values, double entry of values, error in the calculation of taxes.

Artificial intelligence is a major technological breakthrough, which allows autonomous decision making without human intervention, and may allow very relevant results through various technological solutions for accounting. To be precise, Birol, Kaya and Turkyilmaz (2019) define AI as *"the combination of cognitive automation, machine learning, reasoning, hypothesis generation and analysis, natural language processing, and intentional algorithm mutation producing insights and analytics at or above human capability".* AI and other technologies are going to revolutionize the way we all work (Govil, 2020).

1.2. Motivation

After so many years spent at a desk with countless financial and non-financial documents, with a high volume of data, AI and other technologies have revolutionized the way professionals work. Digital technology is changing workplaces in the broader accounting profession (Fawcett, 2015; Guthrie and Parker, 2016; Kruskopf et al.,2020). This change will create new opportunities and will consequently require more advanced technological skills from employees (Greenman, 2017; Lehner et al., 2019; Neely and Cook, 2011).

According to Jeb Su (2018), accounting has seen no innovation since the creation of double-entry bookkeeping, a system where each transaction is recorded in two accounts: debit and credit, and this is one of the great advances since many years ago.

With the aim of innovating the world of accounting, increasing productivity, efficiency, effectiveness and to improve output quality and decision-making, it would be interesting to adding Artificial Intelligence to accounting operations.

Mohammad S. J et al., (2020) said that *"AI has made great progress and it is set to shape the future. It has penetrated to all fields of life and its ever-increasing use in business decision making has made it vital tools for success for all major corporations around the world".*

AI and other technologies together will enable the elimination of errors commonly made by accountants, so there will be no more bias and self-interest on the part of accountants (Hamad et al., 2020).

Thus, the main research question of this thesis, is:

What will be the future architecture of accounting systems?

The question can be unfolded into several secondary questions, namely:

- How can we use Robotic Process Automation (RPA) in accounting / Optical Character Recognition (OCR) in general accounting (Scanning of documents and respective automatic document calculations)?
- How can we do automatic recognition of accounting documents?
- How can we use Machine Learning for massive data analysis and pattern identification through algorithms?
- How would be possible to detect anomalies in the various accounting operations through Artificial Intelligence and from that understand their origin and identify the main problems and solutions?

- How can we create an architecture to support second generation general accounting model using AI (using the above points in this model)?

1.3. Objectives

Aiming help answer the research questions, it was established the following research goal: create an Intelligent architecture to support second generation general accounting.

In order to achieve this goal, the following intermediate objectives were defined:

- Study the general accounting key areas and concepts
- Study the Artificial Intelligence and process automation main techniques
- Build an Intelligent architecture to support second generation general accounting
- Validate the model

1.4. Study relevance and importance

Nowadays, it is very common to store accounting information in digital form, through accounting software, but with the application of Artificial intelligence tasks can be carried.

In certain tasks is it possible not to have human intervention, therefore with all the research that will be done I intend to create an Intelligent architecture to support second generation general accounting. This architecture aims to turn existing programmes into more intelligent ones. This will allow data to be automatically stored on the platform.

The combination of AI and other technologies will enable increased productivity, efficiency and effectiveness of accounting processes and will play a key role in organizations as new talent strategies are implemented and insights for more competitive customers are gained (Wong, 2020). Thus, the elaboration of the model will speed up these processes and will help to increase their agility and security. This is achieved because AI can recognize patterns and learn contextually.

Despite the automation of processes, it is worth noting that the Artificial Intelligence will not eliminate accountants (Su, 2018). Accountant's participation is fundamental to the whole process, as they can focus on critical insights that machines may not be able to perform. For example, he/she can have more time to develop the reporting transparency which helps create trust between the accounting professional and his client, explain how they are creating long-term value (Wollmert, 2018) and create strategies. This is only possible, because with the use of AI and other technologies the accountant will not be burdened with tasks that can be done automatically. This more customer-focused involvement and creative insights will enable the accounting organizations to evolve.

Finally, according to the 2018 EY Global Financial Accounting and Advisory Services (FAAS) corporate reporting survey, 72% of finance leaders around the world believes that AI will have a significant impact on how finance crafts insights regarding data (Wong, 2020).

2. METHODOLOGY

An architecture that supports second generation general accounting, it can be considered as an artifact, for this reason, the most suitable methodology is Design Science Research (DSR).

The architecture of a system that will allow the automatic classification of documents (movements of accounting entries), analysis of balance sheets and analysis of results, will be the result that occurred from the preparatory or investigative procedure, DSR.

With this methodology, the final result of this artefact will be to propose changes to the traditional architecture of general accounting support programs, in order to make them more intelligent to pave the way for a second-generation accounting software.

2.1. Design Science Research

Design Science Research was influenced by predecessors of design research such as March and Smith (1995), Nunamaker (1990), and Walls (1992), authors who built research careers directed at building physical information systems (Peffers et al., 2018). According to Venable (2006), the Design Science Research *"has a broader objective: to generate knowledge that is applicable and useful for solving problems, improving existing systems and also creating new solutions and/or artifacts".*

Horváth (2007) and Baskerville et al. (2015) have the same idea regarding the methodology under analysis. The DSR is fundamental in two main ways: (1) utilise all the knowledge that has been acquired with the aim of solving problems, creating changes or improving existing solutions, and (2) to create new knowledge and theoretical explanations and insights.

Design Science Research methodology include six steps, according Peffers et al., (2007): (1) Problem Identification and motivation; (2) Objective Definition; (3) Design and Development; (4) Demonstration; (5) Evaluation; and (6) Communication. These six activities are represented in figure 1.

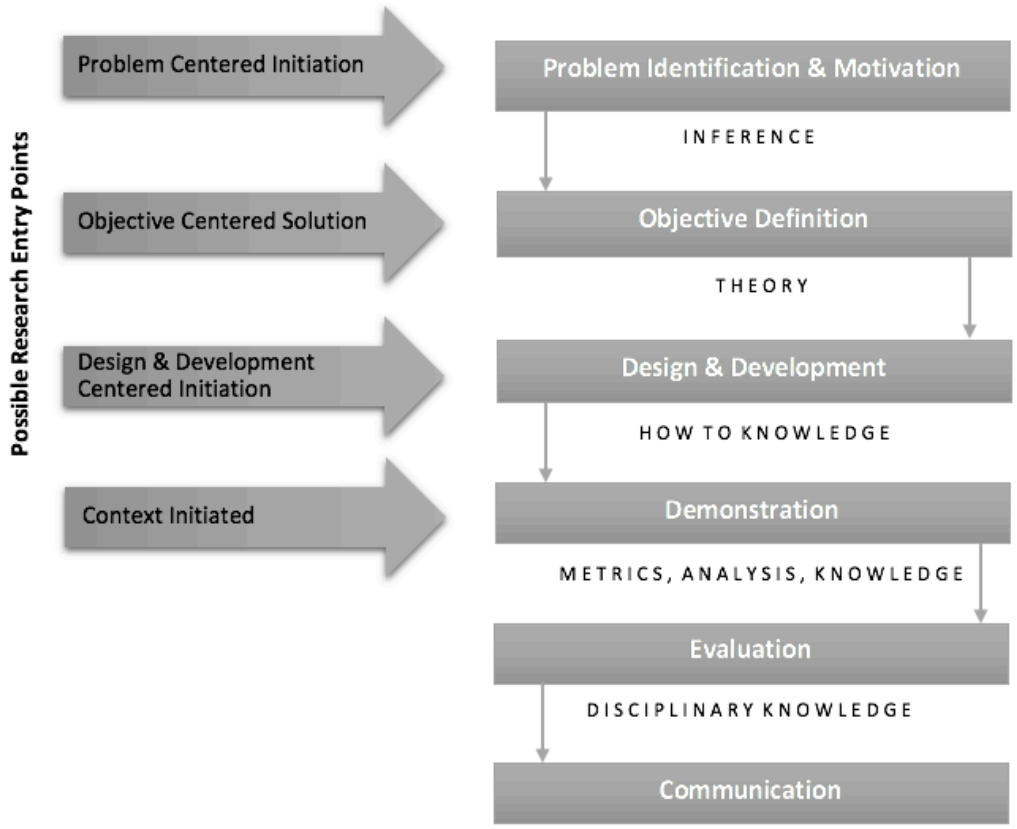


Figure 1.Design Science Research Stages - adapted from (Peffers et al. 2007)

These six steps will enable the construction of the intelligent architecture that will serve as support for second generation accounting, because, as has been mentioned, this methodology aims to solve problems by improving existing systems by creating new solutions that enable the improvement and growth of these same systems. Peffers et al. (2007) mentioned that the design science research methodology "emphasises the design and construction of applicable artefacts, such as systems, applications, methods, and others, that could potentially contribute to the efficacy of IS in organizations".

After the artefact has been constructed, it is evaluated, looking for both the environmental usefulness and the knowledge that has been implemented (new knowledge). For the evaluation both the quality of the artefact and the quality of the knowledge need to be addressed (Venable et al., 2016).

2.2. Research Strategy

One of the most common situations in the world of financial accounting is undoubtedly the calculation error or the forgetting of data. The motivation is to make accounting work more efficient and reliable, contributing to a greater and detailed assessment of the performance of companies and their activities. In the first step, literature review will be prepared in order to understand the area of accounting, more specifically general accounting. Within general accounting, we will study how it works, what accounting operations are carried out, what types

of companies use it, what standards are established in this accounting, what financial data are used, what types of analysis and reports are usually prepared in this type of accounting.

It will be needed a high level of understanding about the AI and other technologies, as Robot Process Automation, Optical Character Recognition and Machine Learning. This is to understand how the automatic recognition of documents can be used to make accounting entries and then a balance sheet analysis and a final analysis of the results obtained, so that later on beneficial business strategies can be created for the company or client in question.

The objective of this research will be the implementation of an architecture for a financial accounting system capable of turning already existing software into more intelligent programs, starting from the automatic recognition of documents for the preparation of accounting entries, analysis of results and management. In the second step, all the literature review will be concluded, and all the information analysed insofar as it can bring together all the requirements necessary for the elaboration of the intelligent architecture to support second generation general accounting, so that it is an architecture that can turn existing software into smarter general accounting software.

In the Design and Development phase, after all the requirements and all the research done, we will develop the artefact. I will combine AI with other technologies studied, so that the intelligent architecture to support general accounting will be formed with the aim of proposing changes to the architecture of traditional accounting software, more specifically general accounting. This architecture will consist more specifically of the technologies of OCR, Text Mining, Rule-based classification (automatic classification), Cloud Computing and Business Intelligence.

In phase four, a use case of the general accounting support system architecture will be built in order to understand if the conceptual artefact meets the pre-established objectives and if this system really has the ability to transform the existing software into smarter software, since the final result of this research will be to propose changes to the traditional architecture of general accounting support programs in order to pave the way for the existence of a second generation accounting software. This Demonstration will be executed in three stages: stage 1- presentation of an example invoice; stage 2- Text Mining, i.e., the information present in the invoice will go to the system and will be organized in a table with the same data present in it; stage 3- accounting of the invoice. The validation of the model will be essential to see if it is effective, efficient, fast and easy to use. This phase is quite relevant for research, since if something is not in line with the defined objectives, the project is still in time to be modified.

In the Evaluation phase, the results of the architecture will be evaluated and interpreted. To elaborate this phase, a brief presentation about the architecture project in question will be elaborated presentation to three interviewees who will answer three quality assessment questions.

The final phase corresponding to communication, the interpretation, description and communication of data relating to the architecture will be carried out taking into account that it has fulfilled all the pre-defined objectives and requirements.

3. LITERATURE REVIEW

A Literature Review is essential to any research project and discipline and can be, also referred to as a 'Theoretical Framework' or 'Research Background'. To be considered relevant and reliable, a literature review must follow certain steps, be clear to those who read it and convey confidence and veracity (Snyder, H., 2019).

This review consists of describing what has already been done, what has already been discovered, supported by reliable and clear reports (Moher et al., 2009; Snyder, 2019). According to Snyder, H., (2019) "a literature review provides the basis for the construction of a new conceptual model or theory", that is, by investigating what already exists and what has already been discovered, it is possible, most of the time, to find gaps or gaps in the research of a certain topic and from these gaps and gaps, various solutions can be created, allowing the development of a new model or new theory.

Given this information of what a literature review is, the available literature will be explored and the most pertinent information will be gathered for the topic being studied and worked on. This will require a literature review by reading articles associated with the topic in question and then gathering the most relevant information for the analysis of the research being undertaken.

3.1. General Accounting

3.1.1. Concepts

Financial Accounting summarises and reports the transactions of a company in order to provide an accurate and clear picture of its financial performance and positioning, what end results the company has achieved, whether positive or negative, what strategies or solutions for the company to pursue the next step.

"*Accounting is the process of systematically recording, reporting and analysing the financial transactions of a company. "*, according to the International Journal of Accounting Research. It is one of the most important tools for a company.

Financial and accounting data about the company is collected in order to guide company-related decisions. The accounting system is based on old technology dating back to Pacioli in 1494 and still used by the market. It is a double-entry system, which keeps track of the company's stocks and financial flows (Christensen, 2010).

The preparation of financial statements that correspond to the balance sheet, income statement and cash flow statement is the main objective of financial accounting and gives the financial manager insight into the company's past performance and its current position (Renzetti, 2015), making it possible to compare past and present data. This analysis of the company's performance through financial accounting allows the financial manager to make forecasts for the future (Renzetti, 2015) and create new strategies, always keeping the company's success in mind.

Accounting presents a set of standards, by which an accountant must follow and respect, as well as regulations called Generally Accepted Accounting Principles.

The Accounting Standards Committee (CNC) is an independent organization, in which public and private entities with an interest in accounting are represented at a national level (Comissão de Normalização Contabilística, 2013).

Its purpose is to issue standards, opinions and recommendations to the aforementioned entities within the corporate and public sector, in order to guarantee accounting procedures harmonised with European and international standards of the same nature (Comissão de Normalização Contabilística, 2013).

This entity issues the accounting standards that are applied in the Accounting Normalisation System (SNC), in the Accounting Normalisation for Micro-entities (NCM) and in the Public Accounting Normalisation System (SNCP) (Comissão de Normalização Contabilística, 2013).

The SNC is an accounting standardisation model based more on principles than on explicit rules and is composed of the following instruments: Bases for the Presentation of Financial Statements; Models of Financial Statements; Code of Accounts; Accounting and Financial Reporting Standards; Accounting and Financial Reporting Standards for Small Entities and Interpretative Standards as presented in annex version 070618 of the accounting standardisation system.

Although the SNC is the main accounting model, smaller entities may use, by option, the NCM which is an accounting model for micro-entities (Couto, M., 2010). The SNCP is applicable only to public accounting and seeks to bring national public accounting into harmony with global accounting practices (Ferreira, R., 2012).

In general, we can consider the Accounting Standardisation System (SNC). Accounting model for entities in general, whose decree-law in force is Decree-law nº 158/2009 (Comissão de Normalização Contabilística, 2013).

3.1.2. Areas & Tools

Accounting comprises various areas such financial accounting, management accounting, government accounting, public accounting, cost accounting, forensic accounting, tax accounting and auditing. The area on which this research will focus is area corresponding to financial accounting where their specialty is in external reporting involving accounting standards that have to be respected as they are mandatory and may change according to the country where the company is based (Bragg, S., 2021).

Today, many accountants already use innovative digital tools that optimize the entire accounting process, thus reducing paper consumption. Many wonders whether accountants will be replaced by robots, but quite the contrary, various technologies such as artificial intelligence will empower them by reducing manual data entry, will speed up processes and will improve the quality of data and final results (Chandi, N., 2018).

According to Chandi, N. (2018) publication present in Forbes, it is predicted that by 2020 labour-intensive and repetitive tasks such as preparing taxes, audits, payroll, banking, among others, will be fully automated – "*a trend that is considered the biggest transformation since the introduction of double-entry bookkeeping 500 years ago" and it is predicted also that by 2026 "the global market for accounting software will have a value of $11.8 billion".* Accounting in general is constantly changing for the better, and great opportunities and challenges are predicted for this world in the future.

3.1.3. Challenges and Opportunities

The initial thought that occurs in people's minds is that technology has come to devalue the accountant's work, however this type of thinking is completely wrong, since technology in this context is responsible for favoring the professional's performance, for better guiding the development in companies. The introduction of new technology has improved activities that used to be carried out in handwriting, which was a costly and exhausting job (Gera et al., 2013).

The accounting systems themselves present problems, since there is a discrepancy between the accounting model, which is linear and the economy, which is non-linear, thus reducing its usability for decision-making purposes for the company (Christensen, 2010). However, these accounting systems are critical in providing information to decision-makers. According to Horngren et al. 2009 "*It produces figures that have proved useful in pricing commodities and produces data that serve as inputs to the management-by-balance sheet assessment exercise".*

Financial accounting often faces challenges related to mandatory standards, which go according to the SNS, and which can be changed. Once these are changed, accountants need to be aware of these changes in order to do their job properly and now with COVID-19 some difficulties may arise. With the help of new technologies, some accounting software already keeps up with changes in norms, which allows the accountant not to have to worry about these changes, since the system takes care of them.

Around 2% of large companies have implemented machine learning or AI, and approximately one in five make plans to implement these new technologies as well. As more companies adopt new technologies, accountants have to develop other skills, mainly business competency strategies, such as data analytics and financial modelling and forecasting. In addition to these technical skills, so-called soft skills will be needed to obtain the ability to work independently and in virtual teams, have attention to detail and the accountant must be able to adapt to change and explore new subjects. (Beaver, S., 2021).

3.2. Processes Automation

3.2.1. Concepts

In the era of digital economy, automation and robotics technologies are gaining a large dimension in different sectors of the economy. These sectors are business services, such as financial and accounting processes, human resource management, sales and delivery and customer service. (Arkadiusz Januszewski et al., 2021). Process automation utilizes technologies with the aim of automating business processes that are time-consuming and quite complex and it is features three functions that concern process automation, centralization of information and reduction of the need for people input.

To implement process automation, it is necessary to identify if the tasks/processes are repetitive, time consuming, how many people are needed to perform the task, if they need auditing and if they are rules based or standardized. After all these aspects are identified we analyse if this task will have a high return on investment if automated. Finally, a project is created to change to new technologies such as Process Automation, and these tasks can be automated for better efficiency and speed (What is Process Automation, n.d.).

3.2.2. Approaches

Process Automation presents several types of approaches and one of them is Artificial Intelligence. “The capacity of artificial intelligence to solve any question by itself without the aid of any human involvement may be loosely described as machines”, according to Singh, J., et al. (2021). Artificial intelligence is composed of the following technologies, Robotic Process Automation (RPA), Intelligent Process Automation (IPA), Augmented Intelligent Process Automation (AIPA) and Autonomous Agents (AA). These technologies that constitute AI currently solve a number of business problems.

3.2.3. Robotic Process Automation

One of the best-known AI technologies is Robotic Process Automation and it enables the automation of workflows or business processes that are repetitive, with the input of structured data. The decision logic of RPA can be formulated into rule-based decisions. For the use of RPA, the business process or workflow must contain a low level of process complexity and zero or limited cognitive capabilities, this is because it has a low level of exceptional manipulation capabilities and intelligence. RPA presents a limited human supervision.

Robotic Process Automation is understood as the automatic execution of administrative, scientific or industrial tasks using robotics for the execution of multiple tasks that replaces the role of humans (Jorge Ribeiro et al., 2021).

RPA tools provide a set of techniques and processes that aim to improve work by decreasing the number of repetitive tasks.

RPA is defined as “a preconfigured software instance that uses business rules and predefined activity choreography to complete the autonomous execution of a combination of processes, activities, transactions, and tasks in one or more unrelated software systems to deliver result or service with human exception management” (IEEE Corporate Advisory Group, 2017).

Another definition of Robotic Process Automation RPA is that this type of automation is a form of technology that serves to speed up decision making efficiently through rules and with highly structured data. Human supervision is limited in business process management as it is enabled by robots, chatbots or software agents. The best known, so-called classic examples of RPA include help desk, scheduling systems, forms processing and call centre operations and sales process support. (Ng, K. K. et al., 2021).

According to Ribeiro, J. et al. (2021), *“In addition to the use of RPA, the complement with Artificial Intelligence (AI) - algorithms and techniques - allows to improve the precision of the execution of automated processes*.”

Although RPA is a powerful and quite effective technological tool that enables faster execution of routine and repetitive tasks, “its applications are limited only to highly rule-based, structured, mature, standardized, re-standardized and well-documented decision logic for easy tasks/processes with digitized structured data input” as mentioned by the Ng, K. K. et al. (2021).

3.2.4. Intelligence Process Automation

IPA consists of RPA and AI technologies and is able to perform prescriptive analysis and, to perform decision logic with input of unstructured data, such as images, text, videos and voices. IPA enables a certain level of cognitive decision making that combines Artificial Intelligence and soft computing techniques thus enabling the imitation of human decision making.

This technology enables business process improvement and holds tools that remove repetitive, routine and replicable tasks. This promotes the simplification of customer-business interactions, improving response time, speeds up processes and reduces operational risks.

IPA mimics tasks performed by humans and with time and experience learns to perform them better, with greater efficiency. (Berruti, F. et al., 2017)

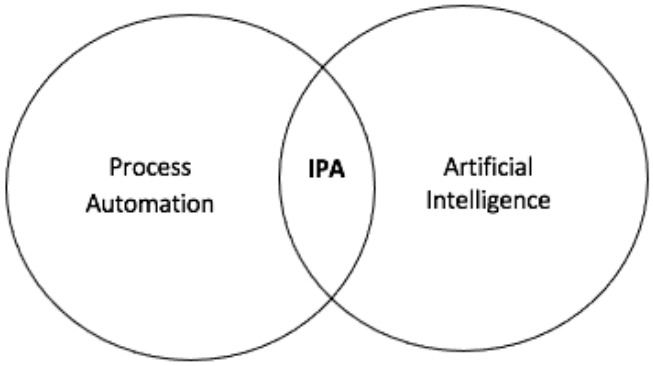


Figure 2.Components of the Intelligence Process Automation

3.2.5. Workflow Automation

The concept of workflow focuses on routines in manufacturing and office to create work efficiency. This routine work is done through tasks, roles, rules and certain procedures and was initially carried out solely by human beings. According to Georgakopoulos et al. (1995), "With *the introduction of information technology, workplace processes are partially or fully automated by information and information systems, i.e., computer programs that perform tasks and apply rules that were previously implemented by human beings".*

Workflows comprise activities that involve the coordinated execution of multiple tasks and, for this reason, information systems, i.e., the automatic processes that enable the execution of these tasks efficiently and quickly, are essential to business. The components of the tasks are relevant and crucial to the control and coordination of their performance (Baharudin, B., Lee, L. H., & Khan, K., 2010).

3.3. AI-based automation of financial accounting operations

3.3.1. Image recognition

Digital transformation across sectors is occurring at a great speed compared to the understanding and time needed for workers to adapt to these changes. This happens with accountants and the constant changes in digital. The transformation in the organisation of accounting could enable a number of services that increase the efficiency and effectiveness of processes and, consequently, a high success of the accounting product. The accounting systems that accompany this technological revolution are transformed into "information systems" that in real time provide information such as the financial and business performance of companies and bases for predictive models for better strategic management of companies (Boyanov, B.,2018).

Optical Character Recognition (OCR) is one of the most relevant and efficient applications in this digital age when it comes to accounting. This system corresponds to a technology that converts scanned images, Portable Document Format (PDF) files and photographs into readable and editable text (Boyanov, B.,2018). Phangtriastu, M. R. et al. (2017) refers that “OCR is one of popular research area in artificial intelligence and patter recognition since it makes possible to read the text and convert it into electronic files, which can be edited. Many applications have been developed using OCR”.

This type of digital technology is extremely relevant and useful for the work of accountants, as it will eliminate the difficulties of accounting processes. That will result in a significant decrease in human errors, limited accounting ransomware that is easy to manipulate, easier and more efficient storage and processing of accounting information accounting information and will provide an easy way to search the archive data. Incorporating OCR technology into accounting can help with fraud prevention and financial control, for example of income and expense accounts. One of the steps in the digital transformation of accounting processes will be to limit the constant use of paper documents through OCR (Boyanov, B.,2018).

The procedure for optical character recognition is composed of three sentences: preprocessing, text detection, text recognition. The phases mentioned can be named differently but they mean the same thing, for example the first phase preprocessing is also called Character Extraction, the second phase Text Detection means Character Classification and finally the last phase Text Recognition can also be called Character Recognition (Phangtriastu, M. R. et al., 2017). We then use the phases that make up the OCR Deep Learning Model where we start with an image (input image) and end with a text (output text).

The steps of an OCR Deep Learning Model:

Figure 3.Steps of an OCR Deep Learning Model - adapted from (Label Your Data., 2021, February 5)

3.3.2. Text Mining

Text Mining is one of the most relevant areas that stems from digital growth over the years and is applied in situations such as database applications that store information in the form of text. More precisely, text data mining is meant for the process of extracting interesting and non-trivial patterns or knowledge through text documents and is a variation on a field called data mining that attempts to find interesting and meaningful patterns from large databases. *"Text mining is the process of transforming unstructured text data into machine-processable structured form to discover hidden patterns, also known as a knowledge discovery database from the text (KDT), it deals with the machine learning supported analysis of the textual data*" mentions Thakur et al., 2021.

As for text mining we find several applications such as "entity" extraction (dates, names, abbreviations and acronyms), text summarization, text categorization, language identification, information extraction and information retrieval, information clustering, authorship relationship identification, sentence structure identification and key phrase detection (Thakur et al., 2021).

Thakur et al., 2021 refers "Textual data is extracted from semi-structured and unstructured datasets such as emails, full-text documents, HTML files ". The process of Text Mining starts with a collection of documents from various resources and the text mining tool retrieves a particular document which is pre-processed, where the document format and character sets are checked in detail. Then, the text present in the document goes through a semantic analysis phase in order to obtain high quality information through the text. Text analysis techniques can be repeated until the information is completely extracted. The resulting information can be placed into a management information system, producing an abundant amount of knowledge for the user of that system. (VijayGaikwad, S., Chaugule, A., & Patil, P.,2014).

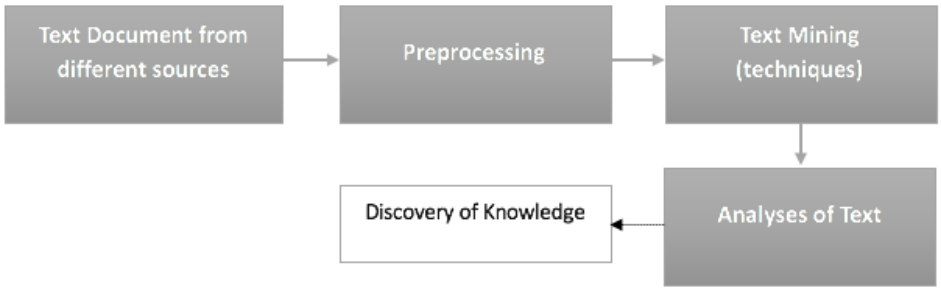


Figure 4.Text Mining Process - adapted from (VijayGaikwad, S., Chaugule, A., & Patil, P.,2014)

3.3.3 Rule based classification

Classification is a relevant issue in machine learning and data mining and has been used in many real-world applications. To build a classifier, a user needs to collect training examples that are named with predefined classes and thus a classification algorithm is applied to the training data in order to obtain a classifier that is used in assigning predefined classes to test instances or future instances (Li, X. et al., 2014).

Automatic classification refers to a process of managing text and unstructured information that is made from the clustering or categorization of text. This technology enables the organization of content or data in a fast and efficient way. The data present in the various accounting documents will be automatically classified and thus directed to the corresponding categories present in the accounting system.

Automatic classification of documents into different predefined categories is a subject that has been extensively studied over the years and is progressing immensely and there seems to be rapid progress in this area as well as the machine learning approaches used for this rule-based automation. The approaches correspond to Bayesian classifier, Decision Tree, Support Vector Machines (SVMs), Latent Semantic Analysis, Rocchio's Algo-Rhythm, K-nearest neighbor (KNN), Fuzzy Correlation, Neural Networks and Genetic Algorithms, etc. (Baharudin, B., Lee, L. H., & Khan, K., 2010).

The approaches mentioned here are all alternatives for the automatic classification of accounting documents so that the data is organized into different categories in the accounting software. Therefore, this would be the last step of the architecture after image recognition (OCR) and text mining (document uploading).

3.4. Existent solutions to support automatic accounting

The accounting area entails different changes with regard to the technological part and presents a term associated with the fourth industrial revolution which is Accounting 4.0. The so-called fourth industrial revolution aims at innovating and optimizing digital processes through various tools in the accounting area.

In the past, accounting was performed in a handwritten way and nowadays it is possible to enter data in proper systems prepared to receive such information. However, special care must be taken when entering data or information. Many errors can occur when entering data. What are accounting systems for? These systems make it possible to make decisions in an effective and efficient manner, optimize knowledge and better attend to clients.

According to Franco G. et al., 2020, "*The information technology is the most used tool by the managers of a company and to keep up with the competition the companies must follow this technological acceleration through the integrated systems that generate information*".

A current integrated system that generates information is the Enterprise Resource Planning (ERP) or integrated management system that aims at the automation of the procedures carried out by a company. This system carries out the planning, execution and control of data both at economic and financial level in an effective and efficient way, increasing the speed of any work method. ERP presents the integration of different sectors acting in different areas (commercial, industrial, etc.) by transmitting information in real time (Haberkorn, 2015).

The sector that most understands the importance of this tool is accountability, since any posting and movement has an impact on accounting reports whether a loss or gain will always reflect on the company's assets, according to the view of Haberkorn (2015).

MRP II, manufacturing resource planning, was a technology for large industries that stood out for its security, agility and greater management of work processes, guaranteeing business management control. Therefore, ERP emerged from MRP II (Franco G., 2020).

There is the integrated management system (ERP), as previously mentioned, as well as the management information system (MIS) and the accounting information system (AIS). The management information system was developed with the function of diminishing the potential risks that arise from business and with the function of making the analysis and interpretation of data more effective and efficient (Xavier & Rodrigues, 2019).

The accounting information system (AIS) has functions such as recording, processing and reporting of both non-financial and financial information/data to assist management decision making at the internal and external control levels of companies/organizations (Richardson et al., 2014).

The Public System of Digital Bookkeeping (PSDB) is a solution that was created in order to formalize the digital files generated by accounting and tax bookkeeping, that is, or the creation of PSDB, the Federal Government gained access to accounting records and thus apply a greater regime of inspection (Sasso et al., 2011). The PSDB files are subject to audits conducted by the PROSOFT software that reports inconsistencies in analytical reports, so *that "companies adopt the necessary measures to meet the obligations of current legislation",* stated Franco G., 2020.

Another well-known software in the accounting area is Software Aided Group Environment (SAGE), which is present in 23 countries with around 3 million clients and 13 thousand employees. SAGE combines Sage accounting, Sage fiscal, Sage payroll and Sage asset control. It has Sage solutions made for customer service based on reliable and updated information/data (Franco G., 2020).

The Sage Accounting software is a simple and easy to use software that allows you to automate several tasks, allowing you to be more customer-focused and provide a better service. It is possible to access the software from any connection, thus maintaining control of the company and ensuring proper compliance with tax obligations (Software de Contabilidade|Sage PT, n.d.).

SAGE software presents a tool that corresponds to a cloud digital archive for accountants that allows the organization of documents and ensures their security. The cloud digital archive refers to an add-on (plug-in, extension or complement used to add functionality to larger programmes) that has the function of archiving documents in the cloud and automating accounting entries.

How does this add-on work? Since documents are stored in the cloud, they can be accessed quickly whenever you want and will reduce the costs of office supplies (Arquivo Digital Cloud para documentos contabilísticos | Sage PT, n.d.).

One of the most important mechanisms for this thesis is the automatic classification of accounting documents. The traditional method of classification corresponds to a manual process where experts first know and understand the content of the document and then assign attributes of the taxonomy to the designated manuscript (Chakraborty, V. et al., 2014). According to Fisher et al., 2010, “While the extant research on accounting thought development and evolution applied the traditional manual classification method for long, concerns have been raised in the literature that it is fairly time consuming, costly, and could introduce inconsistent classifications by different researchers”.

Accounting for some companies is still done from slow, manual and outdated processes involving spreadsheets, paper export reports and immense data entry. Fortunately, automated classification technologies are now available to make accounting work faster, more productive and more accurate.

Automatic classification software has benefits that differ greatly from traditional, time-consuming automatic classification. The benefits are, saving time by reducing old, slow, inefficient and repetitive tasks; be more productive, as you can save time the accountant gains extra time for more important matters that add value to you and your client (e.g., decision-making, strategy development); get more accurate data, through software that is designed to perform endless repetitive tasks, always the same way, whereas a human being with a plethora of documents is more likely to make mistakes; retrieve information instantly if the accounting automation software is cloud-based, thus meaning that any documents, data or information will be at your disposal whenever you need it (you don't have to search and books or filing cabinets); use secure file storage (digital files are more secure than piles of paper moved from one place to another) and work from anywhere, from the time of COVID-19 many accountants were able to work from home as they were using one of the automated classification software with the cloud and therefore had access to all the files/documents needed to run the accounting (Whatman, P. 2022).

Four important features that should constitute an automated classification software are, digital solutions to boring, repetitive and stressful problems; easy to use, as accounting already presents complicated concepts and thus there will be no need for intensive trainings for users; cloud-based system, so that users of the platform can work wherever they want and finally, user-friendly, in the sense that good platforms allow sharing of reports to clients (Whatman, P. 2022).

Xero software is one of the most popular systems for small, medium and large companies and accountants. It enables the availability of financial data to all relevant parties, including accountants and external advisors.

Accountants do not have to recreate the reports for the rest of the company, clients can log into the platform and thus get the information they need. There is other software as an alternative to XERO with different components. **FreshBooks** offers unlimited screening time at the lowest level of its plan, while Xero users for a screening of information would have to opt for the highest level (most expensive level) to get that feature. **GoDaddy Bookkeeping** allows small businesses to accept payments directly from their phones. It's quick and easy to use and requires no prior bookkeeping knowledge.

QuickBooks Online is the ideal software if your business is growing with more specific accounting needs. Like Xero, it can easily scale with your business.

Going back to the Sage tool that includes **Sage 50cloud's Cash Flow Manager** is a tool to better understand future transactions and how they can impact the business. It allows you to analyze the business over the years.

The **Wave software** (free platform) is an excellent choice for small businesses looking for free software (Woock, K., 2022).

QuickBooks online is the best accounting software for medium-sized businesses whose advantages are advanced features, numerous integrations, ease of use and strong accounting. The disadvantages are poor customer support and support for only 25 users, which for a large company would not work.

For large companies **Xero** is the best cloud-based accounting software and its advantages are advanced features, best security when it comes to cloud-based, unlimited users (unlike QuickBooks), fairly strong accounting and numerous integrations. As for the cons these are, poor customer support and steep learning curve.

Another software for large businesses is Finsync with unlimited users, affordable, easy to use and strong cash flow and analytics. On the downside, there is no mobile app (so it can't be used anywhere), limited integrations and features. This software, also cloud-based, has unique aspects when it comes to its cash flow functionality, allowing you to analyze it and therefore plan the future of the business. This software, unlike its predecessors, already has indispensable customer support. For companies that are wary of switching to new technologies, with FINSYNC you can hire an expert to set up your account.

For large companies that want to grow in terms of the software they use, the idea would be to move from the current accounting software to full ERP software. ERP software consists of a complete package of functionality. These are: accounting, analytics, business planning, inventory, sales management, and human resources. NetSuite ERP is above XERO and QuickBooks Online, it has a lot more to offer as it has a lot of features like its strong security, advanced dashboards and user permissions, it suits any type of business and niche and, it is high level of analytics and reports available where the software displays the different KPI'S of the company. The cons of this platform correspond to the difficult learning and high price.

For companies that are not yet completely prepared to take on ERP software, they can alternatively opt for **QuickBooks Enterprise**. QuickBooks Enterprise is a version above the QuickBooks Online system that features a larger feature set and support for 40 users instead of 25. Advantages of this version, more than 240 integrations, numerous features, allows 24 users and good customer support. As for the disadvantages, it is expensive and has a steep learning curve. It should be mentioned that this version constitutes a unique feature not found in any accounting software (not ERP) which is, business plans and cash-flow projections (Seppala, E., 2022).

The last software is the **Primavera** software, which allows the automation of accounting, document archiving in a cloud, business reporting and fiscal reporting. This software is only available in one language, English, which prevents some accountants from using the software. The Primavera software commonly used in Portugal has the disadvantage of not having an OCR system, i.e., the documents that are in the digital archive of the system in question available for automatic launches are loaded via email or via QR CODE. Scanning is not included in this system, so Primavera users have to own a separate scanner.

4. AN ARCHITECTURE PROPOSAL

4.1. Assumptions

The purpose of this thesis came about while debating an issue about the hard work of accountants in general. The manual work is quite immense which can lead to accounting errors. One of the most common situations in the world of financial accounting is undoubtedly the calculation error or the forgetting of data.

The question was - how can I make the accounting work in financial accounting more efficient and productive by using artificial intelligence and its tools? Since in this new age of digital, robotics and automation are increasingly growing, it will be possible to build a suitable architecture for an accounting system that allows manual work to be reduced or even eliminated. This architecture will allow the automatic classification of documents (movements of accounting entries), analysis of balance sheets and analysis of results and this artefact will be to propose changes to the traditional architecture of general accounting support programs, in order to make them more intelligent to pave the way for a second-generation accounting software.

The advantages inherent in building this architecture are greater productivity, efficiency and effectiveness of accounting processes through the combination of AI with other technologies, performance a key role in organisations as new talent strategies is implemented and knowledge is gained for more competitive customers (Wong, 2020), acceleration of processes that will help increase their agility and security. This is achieved because AI can recognise patterns and learn contextually. The non-elimination of accountants is another advantage (Su, 2018). Accountants' participation is critical to the whole process as they can focus on critical insights that machines may not be able to perform. For example, they can have more time to develop the transparency of reporting that helps build trust between the accounting professional and their client, explain how they are creating long-term value (Wollmert, 2018) and create strategies. This is only possible, because with the use of AI and other technologies the accountant will not be burdened with tasks that can be done automatically. This more customer-centric engagement and creative insights will allow accounting organisations to evolve.

For this architecture it is necessary:

- Perform automatic recognition of accounting documents;
- Use Machine Learning for massive data analysis and identification of patterns through algorithms;
- Identify a specific document and extract all the information to the accounting system;
- Categorize the information, i.e., distribute the data into their different groups within the system
- Detect anomalies in the various accounting operations through Artificial Intelligence and from there understand their origin and identify the main problems and solutions.

Based on what has been studied in the literature review the exact information we have for this analysis is:

- "Accounting is the process of systematically recording, reporting and analysing the financial transactions of a company. ", according to the International Journal of Accounting Research. It is one of the most important tools for a company;

- The accounting system is based on old technology dating back to Pacioli in 1494 and still used by the market. It is a double-entry system, which keeps track of the company's stocks and financial flows (Christensen, 2010).
- The area on which this research will focus is area corresponding to financial accounting where their specialty is in external reporting involving accounting standards that have to be respected as they are mandatory and may change according to the country where the company is based (Bragg, S., 2021);
- Many accountants already use innovative digital tools that optimize the entire accounting process, thus reducing paper consumption. Many wonders whether accountants will be replaced by robots, but quite the contrary, various technologies such as artificial intelligence will empower them by reducing manual data entry, will speed up processes and will improve the quality of data and final results (Chandi, N., 2018);
- According to Chandi, N. (2018) publication present in Forbes, it is predicted that by 2020 labour-intensive and repetitive tasks such as preparing taxes, audits, payroll, banking, among others, will be fully automated – "*a trend that is considered the biggest transformation since the introduction of double-entry bookkeeping 500 years ago" and it is predicted also that by 2026* "*the global market for accounting software will have a value of $11.8 billion";*
- The accounting systems themselves present problems, since there is a discrepancy between the accounting model, which is linear and the economy, which is non-linear, thus reducing its usability for decision-making purposes for the company (Christensen, 2010);
- According to Horngren et al. 2009 "*It produces figures that have proved useful in pricing commodities and produces data that serve as inputs to the management-by-balance sheet assessment exercise".*
- About 2% of large companies have implemented machine learning or AI, and approximately one in five make plans to also implement these technologies. As more companies will embrace new technologies, accountants will need to develop other skills, primarily business competency strategies such as data analysis and financial modelling and forecasting. In addition to these technical skills the so-called soft skills will be needed to obtain the ability to work independently and in virtual teams, have an attention to detail and the accountant must be able to adapt to change and explore new subjects. (Beaver, S., 2021).
- In the age of the digital economy, robotization and automation is increasingly growing in different sectors of the economy. These sectors are business services, such as financial and accounting processes, human resource management, sales and delivery and customer service. (Arkadiusz Januszewski et al., 2021).
- To implement process automation, it is necessary to identify if the tasks/processes are repetitive, time consuming, how many people are needed to perform the task, if they need auditing and if they are rules based or standardized. A project is created to change to new technologies such as Process Automation, and these tasks can be automated for better efficiency and speed (What is Process Automation, n.d.).
- RPA is defined as "a preconfigured software instance that uses business rules and predefined activity choreography to complete the autonomous execution of a combination of processes, activities, transactions, and tasks in one or more unrelated software systems to deliver result or service with human exception management" (IEEE Corporate Advisory Group, 2017).
- According to Ribeiro, J. et al. (2021), *"In addition to the use of RPA, the complement with Artificial Intelligence (AI) - algorithms and techniques - allows to improve the precision of the execution of automated processes*."
- IPA mimics tasks performed by humans and with time and experience learns to perform them better, with greater efficiency. (Berruti, F. et al., 2017)
- Phangtriastu, M. R. et al. (2017) refers that "OCR is one of popular research area in artificial intelligence and patter recognition since it makes possible to read the text and convert it into electronic files, which can be edited. Many applications have been developed using OCR".

- “Text mining is the process of transforming unstructured text data into machine-processable structured form to discover hidden patterns, also known as a knowledge discovery database from the text (KDT), it deals with the machine learning supported analysis of the textual data” mentions Thakur et al., 2021.
- Text Mining applications: "entity" extraction (dates, names, abbreviations and acronyms), text summarization, text categorization, language identification, information extraction and information retrieval, information clustering, authorship relationship identification, sentence structure identification and key phrase detection (Thakur et al., 2021).

4.2. Architecture

The architecture is composed of two systems: the first corresponds to the Introduction and classification of documents that comprises the following technologies: Cloud Computing, OCR, Text Mining, Rule Based Classification, Optimization and adjustments and the second corresponds to the Business Intelligence where the data analysis takes place (indicator evaluation and data analysis over the years).

Accounting documents are stored in the cloud at the start of the process and during the process. Storing them ensures greater security and allows documents to be accessed anywhere, at any time. In this way, physical filing is eliminated, and the costs associated with office supplies are reduced (paper, physical space and printing). Accountants' time and resources become more efficient.

OCR converts scanned images, PDF files and photographs into readable and editable text (Boyanov, B.,2018). Incorporating OCR technology into accounting can help in fraud prevention and financial control, for example, of income and expense accounts. One of the steps in the digital transformation of accounting processes will be to limit the constant use of paper documents through OCR (Boyanov, B.,2018). The procedure of optical character recognition consists of three phrases: pre-processing, text detection, text recognition.

Text mining starts with a collection of documents from given resources. A specific document is retrieved, and the mining tool performs its pre-processing, checking the character set and its format. Then, the document goes through a text analysis phase (semantic analysis to obtain high quality information from the text in question).

Text analysis techniques can be repeated until the information is completely extracted (VijayGaikwad, S., Chaugule, A., & Patil, P.,2014).

Automatic classification refers to a process of managing text and unstructured information that is done from the grouping or categorization of text. This technology allows classification of documents (example: invoice) and enables the organization of the content or data in a fast and efficient way. The data present in the various accounting documents will be automatically classified and thus directed to the corresponding categories present in the accounting system. These accounting documents are stored in the cloud, which ensures greater security and allows access to the documents anywhere, anytime. This technology allows classification of documents (example: invoice) and allows the organization of the content present in the documents in a fast and efficient way. Accountants' time and resources become more efficient.

In stage 4, Accounting Records, the human will have to validate what has been done in the previous phases, if what has been done by the system is not correct or needs some adjustments, it will have to occur the phase 4.1, Optimization and adjustments, for optimization of the rules and the system, this way the system in a later

process already knows what it has to do. Returning to phase 4, Accounting Records, if what has been done by the system previously is correct, we move to phase 5, Data Analysis.

Data analytics brings together various processes such as: data collection, cleaning, inspection, transformation, storage, modelling and querying. The aim of this technology is to produce information and knowledge to enable decision-making.

The accounting system allows:

- Preparation of general accounting, which is the compulsory accounting, and which is outward-looking (clients, banks, partners, suppliers);

- Digital Cloud File, allows secure storage of documents and access to them anywhere;

- Access anywhere via smartphones and tablets;

- Data analysis, with tracking and graphing tools;

- Individual customer profiles with all their accounting data and analysis;

- Several languages in the program for international reach.

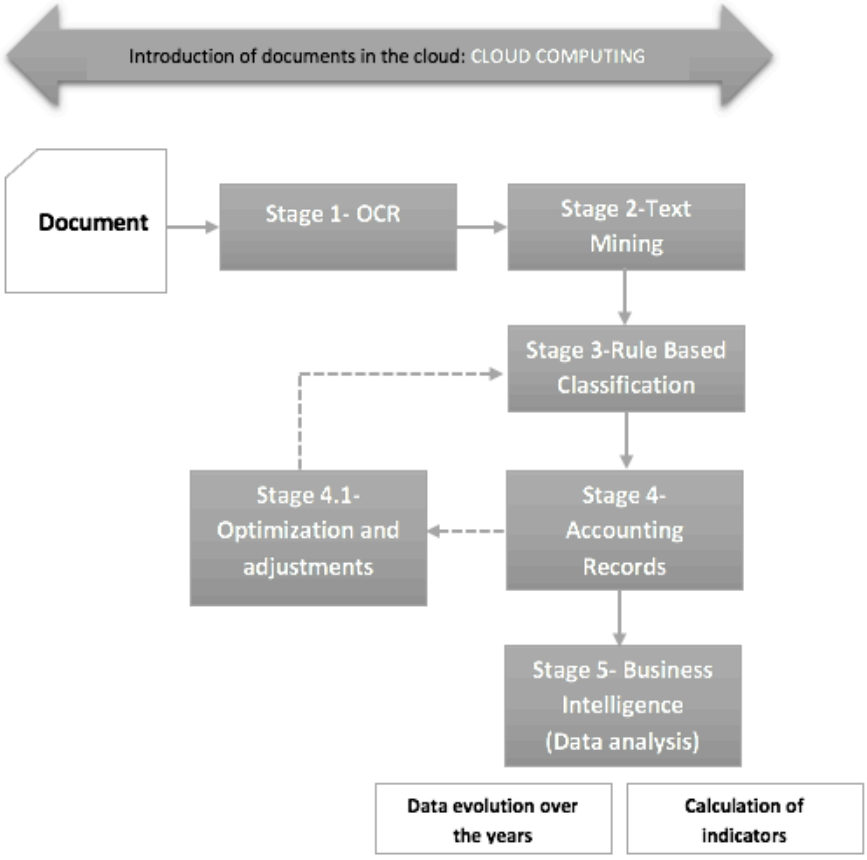


Figure 5.Architecture of the intelligent general accounting system

First system: **Introduction and classification of documents**	**Second System:** **Business Intelligence**
OCR- Optical Character Recognition	Accounting Records- Cloud computing
Text Mining	Business Intelligence- Data Analysis
Rule Based Classification	-
Optimization and adjustments	-

Table 1.Architectural composition

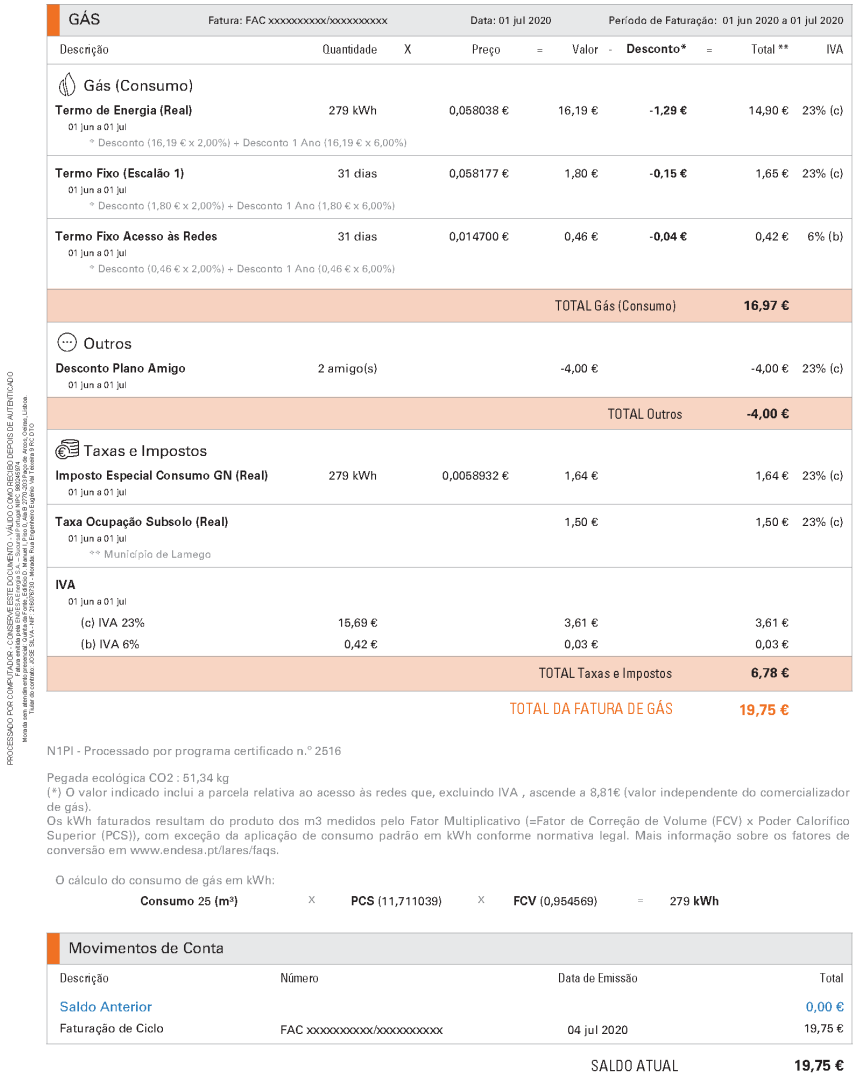
4.3. Use Case

At this stage of the thesis, I will give an example of what will happen in each phase of the architecture after loading the document.

For this example, I used a Gas invoice from July 2, 2020, with the billing period from June 1, 2020 to July 1, 2020.

In this invoice there are categorical elements such as: quantity, price, value, discount, total and value added tax (VAT). The elements to be categorized are: Gas (consumption)- energy term (real), fixed term (tier 1), fixed term access to networks; Others- discounts buddy plan; Rates and Taxes- Special GN Consumption Tax (real), Subsoil Occupancy Tax (real), VAT (23% (c), 6% (b)); total gas bill.

Stage 1: Document



GÁS — Fatura: FAC xxxxxxxxxx/xxxxxxxxxx — Data: 01 jul 2020 — Período de Faturação: 01 jun 2020 a 01 jul 2020

Descrição	Quantidade	X	Preço	=	Valor	- Desconto*	= Total **	IVA
Gás (Consumo)								
Termo de Energia (Real) 01 jun a 01 jul	279 kWh		0,058038 €		16,19 €	**-1,29 €**	14,90 €	23% (c)
* Desconto (16,19 € x 2,00%) + Desconto 1 Ano (16,19 € x 6,00%)								
Termo Fixo (Escalão 1) 01 jun a 01 jul	31 dias		0,058177 €		1,80 €	**-0,15 €**	1,65 €	23% (c)
* Desconto (1,80 € x 2,00%) + Desconto 1 Ano (1,80 € x 6,00%)								
Termo Fixo Acesso às Redes 01 jun a 01 jul	31 dias		0,014700 €		0,46 €	**-0,04 €**	0,42 €	6% (b)
* Desconto (0,46 € x 2,00%) + Desconto 1 Ano (0,46 € x 6,00%)								
					TOTAL Gás (Consumo)		**16,97 €**	
Outros								
Desconto Plano Amigo 01 jun a 01 jul	2 amigo(s)				-4,00 €		-4,00 €	23% (c)
						TOTAL Outros	**-4,00 €**	
Taxas e Impostos								
Imposto Especial Consumo GN (Real) 01 jun a 01 jul	279 kWh		0,0058932 €		1,64 €		1,64 €	23% (c)
Taxa Ocupação Subsolo (Real) 01 jun a 01 jul					1,50 €		1,50 €	23% (c)
** Município de Lamego								
IVA 01 jun a 01 jul								
(c) IVA 23%	15,69 €				3,61 €		3,61 €	
(b) IVA 6%	0,42 €				0,03 €		0,03 €	
					TOTAL Taxas e Impostos		**6,78 €**	
					TOTAL DA FATURA DE GÁS		**19,75 €**	

N1PI - Processado por programa certificado n.º 2516

Pegada ecológica CO2 : 51,34 kg
(*) O valor indicado inclui a parcela relativa ao acesso às redes que, excluindo IVA , ascende a 8,81€ (valor independente do comercializador de gás).
Os kWh faturados resultam do produto dos m3 medidos pelo Fator Multiplicativo (=Fator de Correção de Volume (FCV) x Poder Calorífico Superior (PCS)), com exceção da aplicação de consumo padrão em kWh conforme normativa legal. Mais informação sobre os fatores de conversão em www.endesa.pt/lares/faqs.

O cálculo do consumo de gás em kWh:
Consumo 25 (m³) x **PCS** (11,711039) x **FCV** (0,954569) = 279 **kWh**

Movimentos de Conta

Descrição	Número	Data de Emissão	Total
Saldo Anterior			0,00 €
Faturação de Ciclo	FAC xxxxxxxxxx/xxxxxxxxxx	04 jul 2020	19,75 €
		SALDO ATUAL	**19,75 €**

Figure 6.Gas invoice from July 2, 2020

Stage 2 - Text Mining Result

		Quantity	Price	Value	Discount	Total	VAT
ENERGY TERM (REAL)		279 kWh	0.058038€	(quantity *price) 16.19€	-1.29€	(value-discount) 14.90€	23%
FIXED TERM (STEP 1)		31 days	0.058177€	1.80€	-0.15€	1.65€	23%
FIXED TERM ACCESS TO NETWORKS		31 days	0.014700€	0.46€	-0.04€	0.42€	6%
				TOTAL GAS		16.97€	
FRIEND PLAN DISCOUNT		2 friends		-4.00€		-4.00€	23%
				TOTAL		-4.00€	
EXCISE DUTY GN (REAL)		279 kWh	0.0058932€	1.64€		1.64€	23%
SUBSOIL OCCUPANCY RATE (ACTUAL)				1.50€		1.50€	23%
VAT 23%		15.69€		(15.69*0.23) 3.61€		3.61€	
VAT 6%		0.42€		(0.42*0.06) 0.03€		0.03€	
				TOTAL TAXES AND FEES		6.78€	

Table 2.Gas Invoice

Stage 3 - After application of the rules

C #22		**19.75€ (total)**
D 6242 3	**13.14€**	
D 2432 (23%)	**3.02€**	
D 6813	**2.55€**	
D 2432 (23%)	**0.59€**	
D 6242 3	**0.42€**	
D 2432 (6%)	**0.03€**	

Table 3.Invoice Accounting

Account description:

C- Credit

D- Debit

Account 62423- Fuels, natural gas.

Account 2432 (23%)- (VAT 23%, Normal rate), Deductible VAT is a tax that can be reclaimed by taxable persons (businesses or self-employed people in the VAT regime) (IVA dedutível: Tudo o que precisa de saber sobre este imposto, 2020).

Account 6813- taxes.

Account 2432 (6%)- (VAT 6%, reduced rate), Deductible VAT.

Account 22- supplier's account. Account 22 is debited and account 12 is credited.

4.4. Evaluation

According to the Scientific Design Research (DSR) methodology, we proceed to the Validation phase, where the architecture and its features are validated. In this phase three quality assessment questions will be executed, as stated in the sixth guideline of Hevner et al. (2004), where the artefacts, which adhere to this methodology, were presented to three IA specialists who answered quality and utility assessment questions. For the presentation of the quality assessment questions, a presentation in PowerPoint (Annex 1) format was prepared

where the following points were mentioned: objective of the thesis, problem statement, motivation of the project, the required elements for the architecture of the intelligent accounting system, structure of the architecture and explanation and, finally, detailed explanation of each phase of the architecture.

Four people were selected within the accounting area who daily work with accounting processes and face the adversities of this area.

4.4.1. Interviews description

Presentation of the interviewees, education and current professional situation and presentation of the questions that were performed in the presentation and PowerPoint (Annex 1).

- **Ana Rita Almeida:** A graduate in Accounting and Finance who has worked in various companies in the accounting field, such as Azulconta Gabinete de Contabilidade, Lda; Fernando Borges Sousa & Filhos, Lda and currently the Grupo Luísa Todi, accounting for several companies.

- **Beatriz Gonçalves:** A graduate in Accounting and Administration in the area of Taxation, and almost a master's degree in Accounting and Management of Financial Institutions. Performs accounting at MTConsulting.

- **Joana Silva:** Degree in Finance, Master in Accounting and Management of Financial Institutions and Master student in Management Control and Performance Evaluation. Currently she is the tutor of the final year students of ISCAL.

Each interview was followed by the artifacts' presentation (Annex ...) and included tree questions:

- **Question 1 (Q1):** Do you consider the proposed architecture as useful and why? If not, why do you believe it is not?

- **Question 2 (Q2):** Do you have any criticism towards the proposed framework? Please explain.

- **Question 3 (Q3):** Do you have any recommendation or suggestions for further improvements pf the proposed framework?

4.4.2. Discussion

The interviews for the evaluation and validation of the architecture proposal of an intelligent accounting system allowed a discussion about its utility, possible criticisms to the structure of the system, technologies involved in the system, its functioning, among other important factors and, the interviews also permitted to perceive if there are recommendations that improve the architecture.

- **Question 1 (Q1): Do you consider the proposed architecture as useful and why? If not, why do you believe it is not?**

The interviewees agree with the proposal presented and say that it is a "little" different from what is available on the market. A "little", because there is already software that carries out some of the stages mentioned, such as the Primavera Accountant Automation (PAA) software, but the architecture presented is more complex than this existing software. The PAA software performs function 3 regarding the automatic launches according to the template that is chosen and then you have to confirm the launch. It also presents a digital archive but not as complex as the proposal mentions.

In general, they consider the proposal very good and useful for the current labor market, as it allows a more efficient accounting than the manual one, because the colossal number of documents that an accountant has to register, validate and reconcile every month, are hours and hours dedicated to a task that doesn't bring added value. These operations are repetitive, routine and very subject to human error. With this architecture, the execution of processes will be faster and there will be no margin for error and professionals will have more time to dedicate to more creative and strategic actions, such as the analysis and interpretation of information.

- **Question 2 (Q2): Do you have any criticism towards the proposed framework? Please explain.**

Two of the interviewees consider that they have no criticism regarding the structure of the architecture, as it is well designed, well-structured and all phases have meaning and sequence. However, they refer that, taking into account the present time, the processes are becoming more and more automated, which in a few years is expected to reduce the number of employees and with the implementation of the architecture they believe that there may be an abrupt reduction in the number of accountants, since the system will be able to control everything and only one accountant would be necessary, for example, to confirm and validate the operations. They mention that as businesswomen this would be ideal, but as employees the feeling would not be the same. Therefore, in terms of criticism, the general criticism is that it reduces employability.

The third interviewee considered the proposal to be very successful and stated that the automatic integration of data is the key to the great revolution that the service provision sector is currently undergoing. She also states that, the architecture allows transforming the daily routine of an accountant, increasing the efficiency of the processes, increasing the operationalization and accuracy of the data and through the system it will be possible to manage a greater number of clients, with much less effort.

- **Question 3 (Q3): Do you have any recommendation or suggestions for further improvements pf the proposed framework?**

The interviewees do not present recommendations and adjustments to be made, as they consider the architecture proposal to be well structured and well explained when compared to the PAA software. They also consider that the architecture of the intelligent accounting system has all the necessary phases for the

preparation of accounting records and that once it is implemented and explained to all entities that want to implement it, it will be easy to understand and apply.

The interviewees congratulate for the excellent work presented.

According to the interviewees' answers, the architecture meets the requirements and is an asset for the current market. The system is well structured and presents more complex phases than existing software, allowing for greater efficiency in the fulfilment of accounting processes and greater speed without margin for error.

As for the second question, two interviewees mentioned that if this architecture is implemented, there will be a decrease in the number of accountants due to the complexity of the system. However, the focus on the decision and strategy phase will be greater, since accountants have more time for data analysis.

5. CONCLUSION

5.1. Synthesis of the Developed Work

Accounting is a management tool with four functions, the recording function, the control function, the evaluation function and the forecasting function. Its purpose is to provide meaningful data on company and business assets. In this way, its addressees (managers and decision-makers) get to know the economic and financial situation of the company, which allows them to analyze, create strategies and take management decisions, trying to minimize management failures.

To facilitate the work of accountants, there are programs that perform accounting operations and analysis of results (management analysis) already widely used by accountants and companies. However, manual and exhaustive work remains so the existing software's are not so complex.

To make the accounting processes less complex, more efficient and faster an investigation was made, in order to elaborate an architecture of an intelligent system to support general/financial accounting. This architecture was built not only to make processes more efficient, faster and more productive, but also to be able to make good decisions, to drive better results and consequently increase the growth and success of the company, as well as to help understand what strengths and weaknesses are present in the company.

To build this architecture of an intelligent accounting system, the Scientific Research in Design (DSR) was used as the most suitable methodology, since it can be considered as an artefact. With this methodology, the end result of this artefact will be to propose changes to the traditional architecture in order to make them more intelligent.

The DSR methodology is used to act in two situations: in using all the acquired knowledge with the aim of solving problems or improving existing solutions and in creating new knowledge and theoretical explanations and insights as Horváth (2007) and Baskerville et al. (2015) refer to and are in harmony with.

The six steps of the DSR methodology (according to Peffers et al., (2007): (1) Problem Identification and motivation; (2) Objective Definition; (3) Design and Development; (4) Demonstration; (5) Evaluation; and (6) Communication) were fulfilled, which allowed the construction of the intelligent architecture that will support the second generation accounting.

The first phase required a high level of understanding about AI and other technologies such as Robotic Process Automation, Optical Character Recognition and Machine Learning.

In the second stage, all the literature review was completed, and all the information was analyzed so that all the requirements were gathered to elaborate the intelligent architecture.

In the third phase of Design and Development, after all the requirements and all the research carried out, the artefact was developed. The artefact corresponding to the architecture was composed of two systems: the first corresponding to the Introduction and classification of documents that comprised the following technologies: Cloud computing, OCR, Text Mining, Rule Based Classification, Optimization and adjustments and the second one corresponding to the Business Intelligence system where the data analysis takes place (indicator evaluation and data analysis over the years).

In phase four, an architecture use case was built where three phases (stage 1- presentation of the example invoice, stage 2- result of the text mining technology, stage 3- after application of the rules) were described after loading the accounting document. A Gas invoice from July 2, 2020 that constitutes the following elements was used: quantity, price, value, discount, total and VAT.

In the Evaluation phase, the results of the architecture were evaluated and interpreted from interviews with real people who work in this sector of accounting and are knowledgeable about the subject. The interviewees responded to three questions after a brief presentation about the architecture.

Finally, the research objective was achieved, and the research questions were partially answered. There was only one question that was not well examined and investigated regarding anomaly detection through Artificial Intelligence and identifying the source and what the solution was. This could have been better substantiated.

5.2. Research Limitations

The limitations of this thesis project are mainly due to time and scope, as this research was designed to complete a master's thesis that features deadlines and other strands.

The main limitation of this project was that the system was not built, only an architecture was performed. The fact that there is no system, it is not possible to have the full experience of this intelligent accounting system. The architecture is merely a demonstration of how the real system might work, meaning that it explains the various aspects and functions of a system. A system design was also not prepared, simply the structure that comprises the phases and technologies that are part of the system.

A literature review was done on Artificial Intelligence and other technologies that require high knowledge and understanding to build a true accounting system and certain technologies were mentioned in each phase of the architecture and were explained, however can be explained in a more specific way if there was more time to do the research. Each phase and technologies that composes the Artificial Intelligence will be specified, the structure of the system will be more complex with formulation of subdivisions so that there is a higher understanding of the process and greater detail.

In the conclusion it is stated that it was not investigated in a deeper way the theme related to the detection of anomalies, which their origins and solutions. The implementation of the system in the companies and the training of the users were another essential issue to be discussed and investigated.

5.3. Future Work

As a future project for this system proposal to be more complex, detailed and with substance, each layer of the architecture will be developed in a more complex way, more detailed and other studies will be developed regarding each technology used. In the current project and investigation each phase consists of a different technology whose explanation is present in the Architecture chapter, however in the future work a deeper investigation will be done and will be put into practice, i.e., the system will be developed. The architecture will transition to a real system with the different technologies in operation according to the structure presented here.

For the design of the intelligent accounting system I will need the assistance of computer production systems engineers.

Other recommendations for future work are: to investigate and study further how the system will be implemented in medium and large companies; will only an app be used to access the system or can users also access the system from a website (the app is fundamental to access the software anywhere); how will the graphics be designed, what will be the aesthetics to be used; what will be the cost of joining the programme (will it have a monthly cost, only a registration cost, elaboration of different budget packages); to investigate which languages to use for international reach and, how will the individual customer profiles with their data and analysis of them be executed.

In conclusion, it will be crucial to explore all mentioned factors and future features of the program and turn the proposed architecture into a real system with the support of computer engineers capable of producing a complex accounting system.

BIBLIOGRAPHICAL REFERENCES

Arquivo Digital Cloud para documentos contabilísticos | Sage PT. (n.d.). Retrieved from https://www.sage.com/pt-pt/contabilistas/arquivo-digital/

Baharudin, B., Lee, L. H., & Khan, K. (2010). A Review of Machine Learning Algorithms for Text-Documents Classification. *Journal of Advances in Information Technology*, *1*(1). https://doi.org/10.4304/jait.1.1.4-20

Baskerville, R., Kaul, M., & Storey, V. (2015). Genres of Inquiry in Design-Science Research: Justification and Evaluation of Knowledge Production. *MIS Quarterly, 39*(3), 541-564. Retrieved June 16, 2021, from https://www.jstor.org/stable/26629620

Beaver, S. (2021, April 22). *15 Biggest Accounting Challenges and Solutions in 2021*. Oracle NetSuite. https://www.netsuite.com/portal/resource/articles/accounting/accounting-challenges.shtml

Birol, B., Turkyilmaz, M., & Kaya, C. T. (2019). Impact of RPA Technologies on Accounting Systems. *The Journal of Accounting and Finance*, *82*, 235-250. Doi: 10.25095./mudad.536083

Bragg, S. (2021, April 10). *Types of accounting*. AccountingTools. https://www.accountingtools.com/articles/what-are-the-types-of-accounting.html

Chakraborty, V., Chiu, V., & Vasarhelyi, M. (2014). Automatic classification of accounting literature. *International Journal of Accounting Information Systems*, *15*(2), 122–148. https://doi.org/10.1016/j.accinf.2014.01.001

Chandi, N. (2018, September 13). *Accounting Trends Of Tomorrow: What You Need to Know*. Forbes. https://www.forbes.com/sites/forbestechcouncil/2018/09/13/accounting-trends-of-tomorrow-what-you-need-to-know/?sh=3ea35639283b

Christensen, J. (2010). Accounting Errors and Errors of Accounting. *The Accounting Review*, *85*(6), 1827–1838. https://doi.org/10.2308/accr.2010.85.6.1827

Li, X., & Liu, B. (2014). Rule-Based Classification.

Fawcett, T. (2015), "The digital disruption", *Academic Leadership Series*, *6*, 34-40

Fisher, I. E., Garnsey, M. R., Goel, S., & Tam, K. (2010). The Role of Text Analytics and Information Retrieval in the Accounting Domain. *Journal of Emerging Technologies in Accounting*, *7*(1), 1–24. https://doi.org/10.2308/jeta.2010.7.1.1

Franco, G., Faria, R. O. P., Maciel, A. L. M., & Duarte, S. (2020). Contabilidade 4.0: análise dos avanços dos sistemas de tecnologia da informação no ambiente contábil. *CAFI - Contabilidade, Atuária, Finanças & Informação*, *4*(1), 55–73. https://doi.org/10.23925/cafi.v4i1.51225

Georgakopoulos, D., Hornick, M., & Sheth, A. (1995). An overview of workflow management: From process modeling to workflow automation infrastructure. *Distributed and Parallel Databases*, *3*(2), 119–153. https://doi.org/10.1007/bf01277643

Govil, S. (2020). What AI does for accountants. https://www.accountingtoday.com/opinion/what-ai-does-foraccountants

Greenman, C. (2017), "Exploring the impact of artificial intelligence on the accounting profession", *Journal of Research in Business, Economics and Management*, *8*, p. 1452.

Guthrie, J. and Parker, L.D. (2016), "Whither the accounting profession, accountants and accounting researchers? Commentary and projections", *Accounting, Auditing and Accountability Journal*, *29*, 210, doi: 10.1108/aaaj-10-2015-2263

Haberkorn, E. (2015). Material didático do curso Teoria do ERP. 2. Ed.

Horngren, C., S. Datar, G. Foster, M. Rajan, and C. Ittner. 2009. *Cost Accounting: A Managerial Emphasis*.

IVA dedutível: Tudo o que precisa de saber sobre este imposto. (2020, June 18). Contabilidade. https://numerosecardinais.pt/como-funciona-o-iva-dedutivel/

New York, NY: Prentice Hall.

Horváth I. (2007). Comparison of Three Methodological Approaches of Design Research. *Proceedings of ICED 2007, the 16th International Conference on Engineering Design, Paris, France,* 28.-31. 07.2007 I

International Journal of Accounting Research. (n.d.). Retrieved from https://www.longdom.org/peer-reviewed-journals/financial-accounting-27883.html

Khusbu Thakur;Vinit Kumar; (2021). *Application of Text Mining Techniques on Scholarly Research Articles: Methods and Tools . New Review of Academic Librarianship, (), –*.doi:10.1080/13614533.2021.1918190

Kruskopf, S., Lobbas, C., Meinander, H., Soderling, K., Martikainen, M. and Lehner, O. (2020), "Digital accounting and the human factor: theory and practice", *ACRN Journal of Finance and Risk Perspectives*,*9,* 78-89, doi: 10.35944/jofrp.2020.9.1.006

Label Your Data. (2021, February 5). *OCR with Deep Learning: The Curious Machine Learning Case*. Label Your Data Team. https://labelyourdata.com/articles/ocr-with-deep-learning

Lehner, O., Leitner-Hanetseder, S. and Eisl, C. (2019), "The whatness of digital accounting: status quo and ways to move forward", *ACRN Journal of Finance and Risk Perspectives, 8*, I-X, doi:

10.35944/jofrp.2019.8.2.00

Mohammad, S. J., Hamad, A. K., Borgi, H., Thu, P. A., Sial, M. S. & Alhadidi, A. A. (2020). How Artificial Intelligence Changes the Future of Accounting Industry. *International Journal of Economics and Business Administration*, *5(*3), 478-488

Neely, M.P. and Cook, J.S. (2011), "Fifteen years of data and information quality literature: developing a research agenda for accounting", *Journal of Information Systems, 25*, 79-108

Peffers, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24* (3), 45-77.

Peffers, K., Tuunanen, T., & Niehaves, B. (2018). Design science research genres: introduction to the special issue on exemplars and criteria for applicable design science research. *European Journal of Information Systems*, *27*(2), 129–139. https://doi.org/10.1080/0960085x.2018.1458066

Sistema de Normalização Contabilística. (n.d.). Retrieved from http://www.cnc.min-financas.pt/_siteantigo/SNC_projecto/DL_SNC_anexo.pdf

Software de Contabilidade | Sage PT. (n.d.). Retrieved from https://www.sage.com/pt-pt/contabilistas/

Sasso, A., Rosa, I. C., & Barbosa, A. (2011). O SPED e seus reflexos na profissão contábil. VI EPCT – Encontro de produção Científica e Tecnológica, outubro. Recuperado de: http://www.fecilcam.br/nupem/anais_vi_epct/PDF/ciencias_sociais/03_Soc_Aplic_Co mpleto.pdf.

Seppala, E. (2022, August 1). *Best Accounting Software For Large Businesses & Enterprises*. Merchant Maverick. https://www.merchantmaverick.com/best-accounting-software-larger-businesses/

Su, J. (2018). Why Artificial Intelligence is the future of accounting: Study.

https://www.forbes.com/sites/jeanbaptiste/2018/01/22/why-artificial-intelligence-is-the-future-ofaccounting-study/?sh=773deda2337b

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: a Framework for Evaluation in Design Science Research. *European Journal of Information Systems*, *25*(1), 77–89. https://doi.org/10.1057/ejis.2014.36

VijayGaikwad, S., Chaugule, A., & Patil, P. (2014). Text Mining Methods and Techniques. *International Journal of Computer Applications*, *85*(17), 42–45. https://doi.org/10.5120/14937-3507

What is Process Automation?. (n.d.). Retrieved from https://www.tibco.com/reference-center/what-is-process-automation

Whatman, P. (2022a, February 11). *10 smart accounting automation software systems to try*. Spendesk. https://blog.spendesk.com/en/accounting-automation-software

Wollmert, P. (2018). Is everything that counts being counted? https://www.ey.com/en_sg/assurance/howdigital-transformation-of-reporting-connects-trust-and-long-term-value

Wong, R. (2020). How to harness artificial intelligence in accounting. https://www.ey.com/en_sg/ai/how-toharness-artificial-intelligence-in-accounting

Woock, K. (2022, May 2). *Xero Accounting Review 2022: Pricing, Features, Alternatives*. NerdWallet. https://www.nerdwallet.com/article/small-business/xero-review

Seppala, E. (2022, June 2). *Best Accounting Software For Large Businesses & Enterprises*. Merchant Maverick. https://www.merchantmaverick.com/best-accounting-software-larger-businesses/#QuickBooks-Enterprise:-Best-Enterprise-Accounting-Software-For-Large-Businesses-

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Renzetti, M. (2015). Corporate Finance: Financial Control. *International Encyclopedia of the Social & Behavioral Sciences*, 927–931. https://doi.org/10.1016/b978-0-08-097086-8.73030-3

Comissão de Normalização Contabilística:: (2013). Comissão de Normalização Contabilística. http://www.cnc.min-financas.pt/sobre.html

Normativo Contabilístico para Pequenas e Microentidades: implicações ao nível do Relato Financeiro. (2010). Maria Cristina Ribeiro da Silva Couto. https://www.occ.pt/news/comcontabaudit/pdf/119.pdf

SNCP - Sistema de Normalização Contabilística Público. (2012). Rui Miguel Lopes Ferreira. https://www.repository.utl.pt/bitstream/10400.5/4378/1/DM-RMLF-2012.pdf

IEEE Corporate Advisory Group, 2017.IEEE Guide for Terms and Concepts in Intelligent Process Automation. IEEE, New York City.

Ribeiro, J., Lima, R., Eckhardt, T., & Paiva, S. (2021). Robotic Process Automation and Artificial Intelligence in Industry 4.0 – A Literature review. *Procedia Computer Science*, *181*, 51–58. https://doi.org/10.1016/j.procs.2021.01.104

Januszewski, A., Kujawski, J., & Buchalska-Sugajska, N. (2021). Benefits of and Obstacles to RPA Implementation in Accounting Firms. *Procedia Computer Science*, *192*, 4672–4680. https://doi.org/10.1016/j.procs.2021.09.245

Ng, K. K., Chen, C. H., Lee, C., Jiao, J. R., & Yang, Z. X. (2021). A systematic literature review on intelligent automation: Aligning concepts from theory, practice, and future perspectives. *Advanced Engineering Informatics*, *47*, 101246. https://doi.org/10.1016/j.aei.2021.101246

Phangtriastu, M. R., Harefa, J., & Tanoto, D. F. (2017). Comparison Between Neural Network and Support Vector Machine in Optical Character Recognition. *Procedia Computer Science*, *116*, 351–357. https://doi.org/10.1016/j.procs.2017.10.061

Singh, J., & Dhiman, G. (2021). A survey on machine-learning approaches: Theory and their concepts. *Materials Today: Proceedings*. Published. https://doi.org/10.1016/j.matpr.2021.05.335

Xavier, L. M., & Rodrigues, A. T. L. (2019) Industria 4.0 e avanços tecnológicos da área contábil: perfil, percepções e expectativas dos profissionais. **Trabalho de Conclusão de Curso.** Departamento de Ciências Contábeis e Atuariais da Faculdade de Ciências Econômicas da Universidade Federal do Rio Grande do Sul. Recuperado de: https://lume.ufrgs.br/bitstream/handle/10183/197623/001098028.pdf?sequence=1&is Allowed=y.

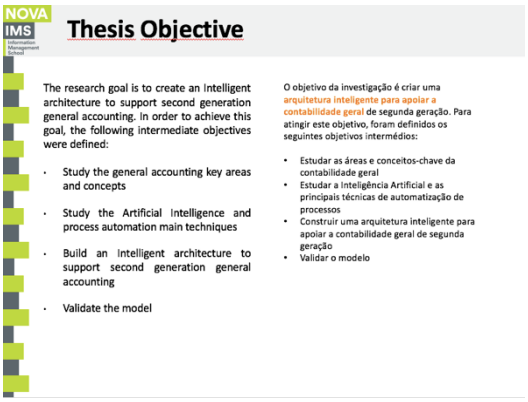
Mazzi, F. (2021). CMOs and AI: Can trained machine learning be justified with the concept of know–how? *World Patent Information*, *65*, 102036. https://doi.org/10.1016/j.wpi.2021.102036

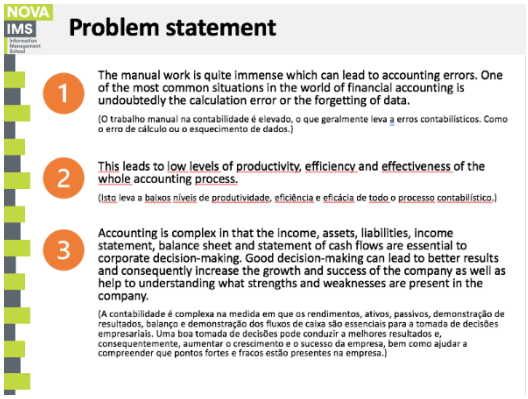
Maiorino, A., del Duca, M. G., & Aprea, C. (2022). ART.I.CO. (Artificial Intelligence for Cooling): An innovative method for optimizing the control of refrigeration systems based on Artificial Neural Networks. *Applied Energy*, *306*, 118072. https://doi.org/10.1016/j.apenergy.2021.118072

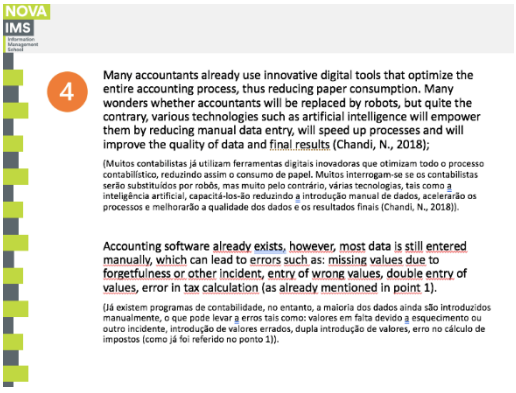
Berruti, F., Nixon, G., Taglloni, G., & Whiteman, R. (2017). *Intelligent process automation: The engine at the core of the next-generation operating model.*

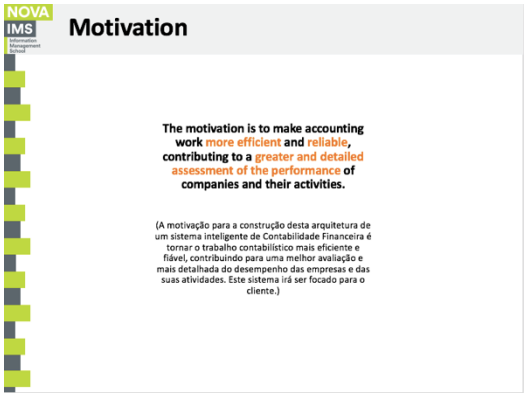
ANNEXES

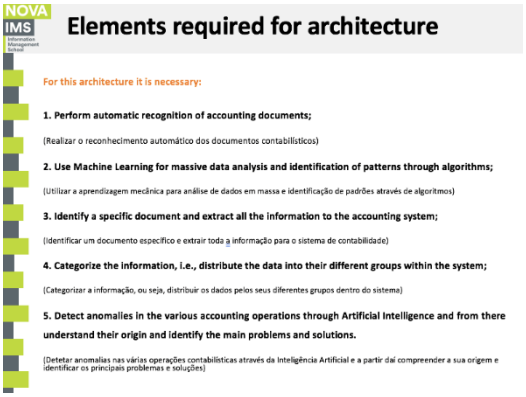
ANNEX 1- INTERVIEWS PRESENTATION



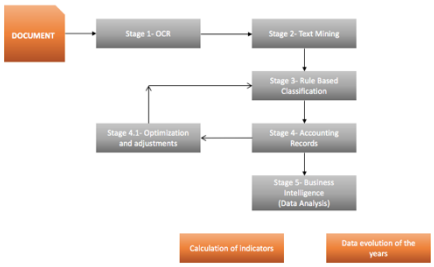








Architecture



Detailed explanation of architecture

The architecture is composed of two systems: the first corresponds to the introduction and classification of documents that comprises the following technologies: OCR, Text Mining, Rule Based Classification, Optimization and adjustments and the second corresponds to the Business Intelligence system with cloud computing and data analysis.

Accounting documents are stored in the cloud at the start of the process and during the process. Storing them ensures greater security and allows documents to be accessed anywhere, at any time. In this way, physical filing is eliminated, and the costs associated with office supplies are reduced (paper, physical space and printing). Accountants' time and resources become more efficient.

OCR converts scanned images, PDF files and photographs into readable and editable text (Boyanov, B.,2018). Incorporating OCR technology into accounting can help in fraud prevention and financial control, for example, of income and expense accounts. One of the steps in the digital transformation of accounting processes will be to limit the constant use of paper documents through OCR (Boyanov, B.,2018). The procedure of optical character recognition consists of three phrases: pre-processing, text detection, text recognition.

Text mining starts with a collection of documents from given resources. A specific document is retrieved, and the mining tool performs its pre-processing, checking the character set and its format. Then, the document goes through a text analysis phase (semantic analysis to obtain high quality information from the text in question).

Detailed explanation of architecture

Text analysis techniques can be repeated until the information is completely extracted (VijayGaikwad, S., Chaugule, A., & Patil, P.,2014).

Automatic classification refers to a process of managing text and unstructured information that is done from the grouping or categorization of text. This technology allows classification of documents (example: invoice) and enables the organization of the content or data in a fast and efficient way. The data present in the various accounting documents will be automatically classified and thus directed to the corresponding categories present in the accounting system. These accounting documents are stored in the cloud, which ensures greater security and allows access to the documents anywhere, anytime. This technology allows classification of documents (example: invoice) and allows the organization of the content present in the documents in a fast and efficient way. Accountants' time and resources become more efficient.

In stage 4, Accounting Records, the human will have to validate what has been done in the previous phases, if what has been done by the system is not correct or needs some adjustments, it will have to occur the phase 4.1, Optimization and adjustments, for optimization of the rules and the system, this way the system in a later process already knows what it has to do. Returning to phase 4, Accounting Records, if what has been done by the system previously is correct, we move to phase 5, Data Analysis.

Data analytics brings together various processes such as: data collection, cleaning, inspection, transformation, storage, modelling and querying. The aim of this technology is to produce information and knowledge to enable decision-making.

Interview Questions

1. Do you consider the proposed architecture as useful and why? If not, why do you believe it is not?
   (Considera que a arquitetura proposta é útil e porquê? Se não, por que razão acredita que não é?)
2. Do you have any criticism towards the proposed framework? Please explain.
   (Tem alguma crítica em relação à arquitetura proposta? Explique, por favor.)
3. Do you have any recommendation or suggestions for further improvements of the proposed framework?
   (Tem alguma recomendação ou sugestão para melhorar ainda mais a arquitetura proposta?)

Thank you | Obrigada!

Address: Campus de Campolide, 1070-312 Lisboa, Portugal
Phone: +351 213 828 610 Fax: +351 213 828 611

Acreditações e Certificações

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

ANNEX 2- INTERVIEWS TRANSCRIPTION

First interview with Ana Rita Almeida:

Q1: "Yes, the proposed Architecture is very useful, because the colossal number of documents that an Accountant has to register, validate and reconcile every month, are hours and hours dedicated to a task that does not bring any added value. These operations are repetitive, routine and subject to human error. Putting "systems" in place to perform these operations seems somewhat self-evident these days. The execution is immeasurably faster, there is no margin for error and professionals have more time to effectively dedicate to more creative and strategic actions, related to the analysis and interpretation of information."

Q2: "The proposed Architecture is very well done; the automatic integration of data is the key to the great revolution that the accounting services sector is currently undergoing. It is the one that allows transforming the daily routine of an accountant, bringing more efficiency to the processes, speed of operationalization and accuracy to the data. It is through it that it is possible to manage many more clients, with much less effort."

Q3: "As I mentioned in the previous question, the proposed Architecture was very well done, as I have no recommendations/suggestions. Simply congratulate for the excellent work presented. "

Second Interview with Beatriz Gonçalves:

Q1: "No doubt the proposal you presented is very good and a little different from what is available in the market. I say "a little" because there are already tools that do some of the phases you described, like, for example, the Primavera Accountant Automation (PAA), in the Primavera Software that already makes automatic launches (your phase 3) depending on the template that is chosen, and it is only necessary to confirm the launch (your phase 4). Within PAA there is also the so called "Digital Archive" that allows you to put the invoices in a cloud (more or less your phase 1), however I believe that it is not in such a complex way as you explain. Your proposal is very good and useful in the labour market because it allows the speed and efficiency of the accounting process because, those who are in the area, know that it is not always easy to launch invoices one by one either in paper format or in digital format; it also allows the reduction of human error (as you explained) because there will be no error with regard to values or duplication of documents launched and reduces the probability of error in the placement of the account to which that expense/income, supplier/customer relates."

Q2: "In terms of criticism of the architecture itself, despite the complexity of your architecture, I have none because I think everything is very well structured and all phases have sense and sequence, however, in terms of real life application of this architecture, I consider my opinion very similar to the general opinion: the more automatic the accounting the less accountants there will be. Explanation: Suppose I have a company with 100 clients under which I do the accounting; if I have 4 employees and each one of us has a portfolio of 20 clients, we (who don't have your architecture proposal) will take about 4 hours (just an example) in each client; however,

if I have your program, I might take half the time, I'm faster and consequently more productive and effective. But for this I won't need 4 collaborators and maybe I'll only need 2. As a businesswoman that would be ideal, but as a salaried employee the feeling would not be the same. So, in terms of criticism, my general criticism is that it reduces employability. Then, it's a very automated programme, which doesn't allow older people to access and understand it so easily. However, if we talk about future generations it will be much easier and useful."

Q3: "I have no recommendations or suggestions regarding the architecture itself because, as I said, I see that everything is very well structured, and all the phases make sense and have a sequence."

Third Interview with Joana Silva:

Q1: "The proposal presented is a little different from what exists in the market, taking for example the Primavera software, which in some stages ends up being similar, but in others is not as complex as the model you presented. It should be noted that nowadays we increasingly want information available at the moment and, as such, your model allows for this speed of information and when applied correctly, it also allows for much more efficient accounting than manual or existing accounting to date. Implicit in the mechanization of the process is the reduction of possible human errors, since the duplication of values (for example) no longer occurs, which simplifies the whole process, since there is no need to make corrections afterwards, either in amounts or in posting accounts."

Q2: "Nowadays, processes are becoming more and more mechanized, which in a few years' time will mean a large reduction in the number of employees. For example, with the implementation of your project I believe that it would lead to an abrupt reduction in the number of accountants, since there would be a system capable of controlling everything and only one person (for example) would be needed to confirm and validate such operations. Thus, my criticism, which I don't even know if I should call it that, would be the unexpected and significant reduction in the number of accountants."

Q3: "Taking into account the structure/architecture presented and well-grounded in all stages, I do not feel that there are any major recommendations to make, because I feel that once it is implemented and explained to all entities that want to implement it, that it is easy to understand and applicability."