sustainability

MDPI

Article

Valuable Business Knowledge Asset Discovery by Processing Unstructured Data

Maria-Isabel Sanchez-Segura [1,*], Roxana González-Cruz [2], Fuensanta Medina-Dominguez [1] and German-Lenin Dugarte-Peña [3]

[1] Computer Science and Engineering Department, Universidad Carlos III de Madrid, Avenida de la Universidad 30, 28011 Leganés, Spain
[2] Project Management Department, Telefónica España, Ronda de la Comunicación s/n, 28050 Madrid, Spain
[3] Higher Polytechnic School, Universidad Francisco de Vitoria, Carretera Pozuelo a Majadahonda, Km 1.800, 28223 Madrid, Spain
* Correspondence: misanche@inf.uc3m.es

Abstract: Modern organizations are challenged to enact a digital transformation and improve their competitiveness while contributing to the ninth Sustainable Development Goal (SGD), "Build resilient infrastructure, promote sustainable industrialization and foster innovation". The discovery of hidden process data's knowledge assets may help to digitalize processes. Working on a valuable knowledge asset discovery process, we found a major challenge in that organizational data and knowledge are likely to be unstructured and undigitized, constraining the power of today's process mining methodologies (PMM). Whereas it has been proved in digitally mature companies, the scope of PMM becomes wider with the complement proposed in this paper, embracing organizations in the process of improving their digital maturity based on available data. We propose the C4PM method, which integrates agile principles, systems thinking and natural language processing techniques to analyze the behavioral patterns of organizational semi-structured or unstructured data from a holistic perspective to discover valuable hidden information and uncover the related knowledge assets aligned with the organization strategic or business goals. Those assets are the key to pointing out potential processes susceptible to be handled using PMM, empowering a sustainable organizational digital transformation. A case study analysis from a dataset containing information on employees' emails in a multinational company was conducted.

Keywords: intangible assets; process mining; natural language processing; knowledge management; digital transformation; sustainability; design science

Citation: Sanchez-Segura, M.-I.; González-Cruz, R.; Medina-Dominguez, F.; Dugarte-Peña, G.-L. Valuable Business Knowledge Asset Discovery by Processing Unstructured Data. *Sustainability* **2022**, *14*, 12971. https://doi.org/10.3390/su142012971

Academic Editors: Hyunchul Ahn and Luigi Aldieri

Received: 14 August 2022
Accepted: 1 October 2022
Published: 11 October 2022

1. Introduction

Since the late 20th century, we have been evolving from a traditional to a digital economy. This is a new kind of economy based mainly on new information communications and technologies (ICTs) and the value hidden in huge amounts of digital data generated by companies. As a result, every company must have access to the information that it generates, processes and stores in its physical and/or virtual devices to help evolve regulations on the protection of personal and business data, as well as control data monetization [1,2]. Accordingly, market forces within the digital economy have forced companies to establish or redefine their sustainability strategy.

Today, more than 90% of CEOs state that sustainability is important to their company's success and companies develop sustainability strategies, market sustainable products and services, create positions, such as "chief sustainability officer", and publish sustainability reports for consumers, investors, activists and the public at large [3]. If enterprises aspire to excellent and efficient sustainable growth, they must utilize their own intangible resources (intellectual capital) to maintain a competitive advantage and adjust to dynamic changes

in the internal and external operating environment through investment in intellectual capital [4].

In the digital era, data are one type of intangible resource that enterprises can use to stay competitive and sustainable and there are several ways in which a company can achieve sustainability based on the data it owns.

Recent research on how to value, manage and protect intangible assets and their underlying information and knowledge to empower companies has revealed the need for improvements with respect to the step concerning the identification of the information at the disposal of organizations [5]. Information identification is, in fact, the first critical step for companies embarking on any project related to data recycling for their own corporate benefit in order to achieve the sustainability and digital transformation demanded by the market.

This is challenging across all company types or sectors because new work methods—either corporate devices or personal media—make data management processes more complex, requiring companies to acquire or develop new tools and guarantee technological integration, often at an unaffordable investment cost for many companies.

We focus on process mining as a means to analyze, understand and improve business processes based on the analysis of event data, i.e., data known as event logs that are stored in a software tool during process execution [6]. Process mining has proven to be a powerful tool for digitally mature companies, but it assumes that:

- Every company has logs to be used as an entry to the process mining tools, which is not true, bearing in mind that from the 333.34 million companies existing around the world, 332.99 million are Small and Medium Enterprises (Source: Eurostat; ILO; OECD; Statista; World Bank), with few resources to invest on process mining tools.
- For companies that already have logs, the identification of data that will lead to intangible asset identification among existing logs is a critical, albeit difficult, task for the business.

This is why we propose a set of steps to complement process mining methodologies and broaden their scope to include organizations in the process of improving their digital maturity. This should help more businesses approach the goal of being more sustainable and competitive, allowing them to use their own data to identify organizational intangible knowledge assets and control their status (which, by definition, improves organizational sustainability and competitiveness).

Following design science research (DSR) principles, we designed C4PM (Complement for process mining), a method that provides, as an output, the identification of potential intangible assets, providing CIOs, IT leaders, software developers and company information security managers with the opportunity to agilely and systematically identify enterprise information and intangible knowledge assets that are critical for their businesses. C4PM output points to potential sources of information or organizational processes that are worthy to be investigated to determine whether or not they are susceptible to be an input for process mining methodologies.

Apart from identifying and evaluating business-critical information and intangible knowledge assets, the C4PM method includes:

1. A systemic approach, based on the existing relationship between employee and stakeholder knowledge of the organizational business goals, which can, throughout the execution of the methodological process, operate within complex real-world environments where process mining software tools are used to more accurately represent the underlying activities in each of the respective processes.
2. An agile business intelligence and natural language processing approach, based on the performed data processing, which can, throughout the process, operate on and manage lots of datasets from heterogeneous sources and with different formats.

The remainder of this paper is structured as follows. Section 2 presents a state of the art examining the main contributions in areas related to this research. Section 3 describes

the materials and methods that were examined and used. Section 4 outlines the proposed solution to the problem of identifying a company's intangible assets based on an artificial intelligence and natural language processing approach. Section 5 reports an analysis of the application of this proposal to a real case with accessible information and future work related to this research.

2. State of the Art

This section contextualizes this proposal and analyzes research already developed within the respective areas.

Today, one of the greatest obstacles in the path of most digital transformation processes at companies is gaining absolute control over data. Data are usually represented in different formats, possibly from more than one source. They frequently occur as semi-structured or unstructured data and often represent people's knowledge about diverse organizational topics, i.e., they represent organizational know-how. However, this knowledge is not always formalized within the company's organizational processes and may or may not be digitized; that is, a software tool may or may not be used to enact this process. In any event, this poses a great challenge for companies seeking to identify their intangible knowledge assets, without which they cannot discover their hidden value. Companies that identify their intangible assets for transformation, updating and protection have an edge in this respect.

However, how can companies generate such business value? One intellectual (human, relational, structural) capital concept states that an organization's ability to take advantage of the dynamics of knowledge assets is at the core of its capacity to create value [7]. For example, the nexus between intellectual capital and corporate sustainable growth in three specific sectors was analyzed in [4] and they found that enterprises that want to excel at efficient sustainable growth must utilize intangible resources (intellectual capital) to maintain competitive advantage and adjust their investment in intellectual capital to dynamic changes in the internal and external operating environment. A review of the existing literature in the administrative management field [8] found that several of the management models and tools developed to evaluate intellectual capital spotlight the definition of metrics and indicators. Indeed, research in this area related to intellectual capital focuses primarily on intangible asset evaluation, paying little heed to the first and critical step: asset identification. Identification is the basis of any analysis and later valuation of intangible assets. This suggests that current analyses and business objective alignment with the valued intangible assets are essentially flawed, which is an obstacle to sound decision making with respect to the process of business digitalization and sustainability.

Since business digitalization processes have so far turned out to be a definite failure [9] and the change drivers needed to support digital transformation are IT managers, IT experts should propose and use solutions that combine the identification and valuation of intangible assets with the construction of software solutions to improve digital businesses. On this basis, we propose, from the IT sector, a solution that will enable the success of business digitalization processes by identifying the hidden knowledge assets behind processes that the company has in place.

One would expect the business intelligence (BI) field to put forward applicable formulas to identify hidden knowledge assets. Several studies published in the BI literature focus on the successful implementation of BI projects using agile methodologies in the areas of computer science, engineering, economics and information systems management. However, none define the information and knowledge assets related to business objectives. For example,

- Reference [10] gathers the most recent works around BI, capable of handling all data types. Existing works promote the rapid adaptation of the organization's BI systems to the changes required by the business, thus, promoting system flexibility, but it does not map data with the strategic and business objectives of the company.

- Reference [11] presents an extensive literature review on process mining, highlighting that process mining is a new kind of Business Analytics and has emerged as a powerful family of Process Science techniques for analyzing and improving business. Existing works fail to address unstructured data or define business-critical information and knowledge assets. It is stated that "The current scientific literature does not present an up-to-date research agenda specifying the directions the application of PM can take in Business Management".
- Reference [12] introduces a practical guide on how to implement an agile BI project, again improving the flexibility demanded by BI systems. However, it does not mention which data types can be processed, nor does it provide any steps to define business-critical information and knowledge assets.

Focusing on process mining software tools, it can be affirmed that they are useful for discovering and modeling business processes from the event logs stored in the software tools used by companies to execute digitalized processes. Thus, they can minimize the noise in the implementation of this type of project, resulting in a more accurate representation of the activity model of the analyzed process. RDM also does not consider business objectives. Therefore, any solution identified by a process mining methodology would, in any case, be unrelated to the value that it could potentially provide for the digital business.

Process mining represents a set of data techniques that supports the analysis, understanding and improvement in business processes based on the analysis of event data, known as event logs, that are stored in a software tool during process execution. Here, three main branches may be identified: process discovery, compliance checking and process improvement. In process discovery, the goal is to discover a process model that accurately describes the behavior recorded in an event log, i.e., a model that describes the actual process followed during process execution. In conformance testing, a process model is compared with the recorded process behavior to check for deviations between the model and the observed behavior. In process improvement, a process model is dynamically enriched with new information about the process based on new analyses of the process model and/or event log (e.g., detecting critical paths, predicting process performance indicators, repairing/simplifying process models, etc.) [6].

So far, the field of process mining, and process discovery in particular, has focused on the exploration and description of event data using models. Since modeling is usually based directly on a sample of event data, the question of whether they also apply to the actual process is often left unanswered. Since the underlying process is unknown in real life, unbiased estimators are needed to assess the system quality of a discovered model and to make subsequent assertions about the process. Jassenswillen and Depaire [13] describe and discuss an experiment to analyze whether existing fitness, accuracy and generalization metrics can be used as unbiased estimators of system fitness and accuracy. The results show that there are significant biases, making it currently almost impossible to objectively measure the ability of a model to represent the system.

Therefore, process mining requires event logs stored in company information systems to automatically identify new business process models by applying sophisticated artificial intelligence-based algorithms. Event logs must be prepared and/or adapted to the format required by the corresponding software tools. Some research found that process mining is not usually very accurate because the event logs may contain noise that affects the analysis being performed. Further, the company must be staffed by professionals with advanced IT knowledge to ensure that the software tools used in process mining are correctly deployed in order to effectively map the digitalized processes that the company needs to analyze.

A comparison of the 16 software tools most commonly used in process mining was conducted by [14]. The software tools were selected based on the most recent reports from three related consulting and research firms (Gartner, Everest Group and Forrester). To this end, vendors that did not provide a testing environment or were not exclusively dedicated to process mining were preliminarily excluded. Further, three open-source tools were discarded: ProM, which targets academic research [15], PM4Py [16], which does not

currently provide a graphical user interface and Apromore, which had a commercial license in a second test cycle.

Table 1 shows a list of 16 process mining software tools from the above comparison. Column 2 reports one of the variables analyzed in previous research and the other columns report new variables that we consider important and that are essential for a company to assess the feasibility of implementing a process mining project.

Table 1. Comparative table of process mining software tools based on new variables affecting the business.

Tool Name (Manufacturer)	Supported Log File Extensions	Compatibility with Most IT Tools	Definition of Strategic and Business Objectives Required	Level of Detail of Organizational Processes (Low, Medium, High)	Knowledge of Business and IT Systems Required	Known and Proven Methodologies Used
ABBYY Timeline (ABBYY)	csv	No	No	M, H	Yes	No
ARIS Process Mining (Software AG)	csv, xls	Yes	No	M, H	Yes	No
BusinessOptix (BusinessOptix)	csv, xes, xml	Yes	No	M, H	Yes	No
Celonis Process Mining (Celonis SE)	csv, xlsx, xes	Yes	No	M, H	Yes	No
Disco (Fluxicon BV)	csv, txt, xls(x), mxml, xes	Yes (Via API)	No	M, H	Yes	No
EverFlow (EverFlow)	csv, xls, xml, xes, json	No	No	M, H	Yes	No
LANA Process Mining (Lana Labs GmbH)	csv	Yes	No	M, H	Yes	No
Logpickr Process Explorer 360 (Logpickr)	csv, txt	Yes (Vía API)	No	M, H	Yes	No
MEHRWERK ProcessMining (Mehrwerk GmbH)	csv, excel files, xml, dif, json	Yes	No	M, H	Yes	No
Minit (Minit j.s.a.)	csv, xes, mxml, xls(x) (m)(b), mdb, accdb	No	No	M, H	Yes	No
MyInvenio (myInvenio Srl.)	csv, xes	Yes	No	M, H	Yes	No

Table 1. *Cont.*

Tool Name (Manufacturer)	Supported Log File Extensions	Compatibility with Most IT Tools	Definition of Strategic and Business Objectives Required	Level of Detail of Organizational Processes (Low, Medium, High)	Knowledge of Business and IT Systems Required	Known and Proven Methodologies Used
PAFnow (Process Analytics Factory GmbH)	csv, txt, excel formats	Yes	No	M, H	Yes	No
ProDiscovery (Puzzle Data Co., Ltd.)	csv, txt	No	No	M, H	Yes	No
QPR ProcessAnalyzer (QPR Software Plc)	csv, xes	Yes	No	M, H	Yes	No
Signavio Process Intelligence (Signavio GmbH)	csv, xes	Yes	No	M, H	Yes	No
UiPath Process Mining (UiPath Inc.)	csv, tsv, xls(x), txt, xes	No	No	M, H	Yes	No

These new variables are:

- Compatibility with most IT tools: the criterion applied for this variable to be rated as "Yes" was for the software tool to be compatible with at least six or more tools known on the market or in the corporate environment.
- Definition of strategic and business objectives required: this variable is based on the fact that process mining does not require the company to define its strategic and business objectives to implement the project.
- Level of detail of organizational processes (Low, Medium, High): this variable is assessed according to three levels: 1. Low (L), when company organizational processes have not been digitalized or formalized, 2. Medium (M), when some company organizational processes have been digitalized, although others may or may not be formalized and 3. High (H), when most company organizational processes have been digitalized and all have been formalized.
- Knowledge of business and IT systems required: the criterion applied for this variable to be rated as "Yes" was based on the professional effort required by a company to implement a process mining project.
- Known and proven methodologies used: this variable is based on the fact that the process mining definition and structure do not use any known or proven methodology in the IT and business environment since it is based exclusively on tool deployment and use in order to later analyze the automatically generated process models.

After analyzing Table 1, we can conclude that although process mining supports various file extensions, the event logs must be properly parameterized within a defined data format for correct software tool use. Therefore, we find that most of the data used are structured and the use of semi-structured or unstructured data is much more limited, so just the companies with existing logs (with enough digital maturity) can use those tools. On the one hand, we observe that more than 65% of the analyzed tools are compatible with most organizational IT tools, although, for the use of any process mining software tool,

the level of detail of organizational processes is necessarily M (medium) or H (high), since the company will have to rely on the digitalized processes that are part of the project. On the other hand, the company will require IT professionals that are very knowledgeable about business and IT systems, because they will be in charge of correctly deploying and operating the software tool. As they do not apply a methodology that is adapted to process mining projects (see Table 1 columns Known and proven methodologies used), there is a huge gap between business and IT. This can lead to analytical errors. However, a known, tried and tested methodology adapted to this type of projects would enhance its execution in the business environment; such a methodology would explicitly contribute to better strategic decision making on the digital transformation that the company is carrying out.

In view of the existing deficiencies in the field of BI and in current process mining tools to identify intangible assets hidden on structured of semi-structured data and aligned with the companies' business goals, the method that we propose in this paper can cover both deficiencies and effectively account for these new variables identified in Table 1, creating a source of data that can be useful as an entry to process mining methodologies without the use of existing process mining tools and is affordable for companies with less digital maturity.

3. Materials and Methods

We used design science research (DSR) as the basis to develop the C4PM method proposed in this paper. DSR aims to create novel artifacts in the form of models, methods and systems that help people develop, use and maintain informatics solutions. An artifact is a human-made object designed to solve a practical problem, possibly concerning many stakeholders. In the field of IT and information systems, there is a large number of artifacts that deliver business value to their owner, ranging from algorithms, logic programs and formal systems, through software architectures, information models and design guidelines, to demonstrators, prototypes and production systems [17]. Thus, DSR provides a solid foundation for creating new methods in the field of information technology, such as the one presented here. Its main objective is to improve support for strategic decision making based on the value of corporate intangible assets to facilitate the path towards digital transformation considering the organization's intellectual capital.

DSR canvas is a rectangle divided into several fields, providing a concise, simple, understandable and visually appealing overview of the components in the designed method [17]. The top (blue) part of the canvas defines the artifact in question, the problem addressed and the knowledge base used in the research. The middle (yellow) part describes the DSR framework activities, such as problem explanation, requirement definition, artifact development, artifact demonstration and artifact evaluation. The bottom (brown) part shows the results of a DSR project in terms of artifact structure, function and effects. The proposed C4PM method was defined according to the DSR canvas illustrated in Figure 1. Its components emerged after a conscientious process of reflection during the method design with respect to the following issues:

- Main requirement: create an agile, systemic and iterative approach using natural language processing techniques to discover behavioral patterns in the analyzed data. These steps should be applied in a balanced way according to the business objectives and should not be limited to a specific project; that is, experts in software and IT solution development and IT and security managers should be able follow these steps to adopt this method in their projects as a building block to implement any type of action on the identified information and knowledge assets.
- Find answers to the next four key questions:
    - Question 1: What is the flow of non-physical assets within the organizational processes?
    - Question 2: How can we discover the use to which these assets are put?
    - Question 3: How locatable are these assets?
    - Question 4: Which assets can be used to improve company sustainability?

Problem A lot of data in modern organizations tend to be unstructured, while they are very important from a strategic perspective since they often represent people's knowledge. This knowledge may not be digitized. However, it is a great challenge for companies to identify their intangible assets from the knowledge hidden in process data, since these assets are essential for the company to discover what they know and have before starting any digital transformation project. CIOs and IT leaders must apply new approaches to capitalize on the value delivered to companies through digital business improvement solutions. It is necessary to take advantage of data, available in different formats from various sources. In this way, they also promote enterprise innovation and sustainability.		**Artifact** This is a method designed for companies, whereby they can identify hidden information and knowledge assets of whose existence they are often unaware. This method entails applying a series of steps and requires information ranging from the knowledge of leaders, stakeholders, employees through to data from information systems that are used to perform their daily activities, such as email. This is an effective method, which integrates four information systems disciplines and can be simply, agilely and intelligently applied by experts in software and IT solutions development, as well as IT and security managers to implement any process mining project or any actions applied to information and knowledge assets in order to improve knowledge digitalization, management and control processes, contributing to business sustainability based on the business objectives.		**Knowledge Base** This method integrates four disciplines —1) Agile principles, 2) systems thinking, 3) business intelligence techniques and the use of process mining software tools, and 4) natural language processing techniques— to holistically and agilely analyze the behavioral patterns of semi-structured or unstructured data, on the basis of which to identify the valuable hidden information and output the related knowledge asset. By integrating these disciplines, it is possible to discover what is going on at the company within the real world and its environment, analyzing different data types that often contain implicit knowledge of which the company is unaware and that ultimately represent its information and knowledge assets.
Practice In the data era, most of these assets are unstructured. This makes it difficult for the company to recognize and extract value from the assets for the purpose of its business. Therefore, this method provides companies with an agile and systemic tool, which consists of applying a series of steps to identify the information and knowledge assets that are critical for their business. This method is mainly based on the information about company organizational processes, regardless of their level of maturity, as well as their strategic and business objectives, whereby analysts can identify the assets. This method is developed by the CEO and the company leaders who develop any of the activities being analyzed. After identifying these assets, companies will be able to implement and ensure the success of their process mining projects, thus promoting business sustainability.		**Requirements** The method requirements are as follows: 1. The CEO or company representative in charge of the project has to indicate the company's strategic objective, and select a single type of generic intangible asset (GIA) model for each iteration until all the GIAs that are likely to provide information and knowledge assets and may be causing bottlenecks within some company processes have been reviewed 2. The systemic analyst/consultant will be in charge of applying the method at the company 3. Process leaders/heads of areas will be involved in the analysis process.		**Constructs** The constructs required to apply the method and its solution are: soft systems methodology, BPMN language notation, agile BI methodology, intangible assets and natural language processing strategic management methodology, process mining. Based on this knowledge, it will be possible to correctly develop the method that will solve the company's problem and identify knowledge and information assets.
State Problem Companies today do not have control over their data, since an increasing volume of data is represented of which a growing percentage is unstructured. On this ground, CEOs, IT and information security leaders, as well as software developers have need of an effective method that facilitates decision making to undertake digital transformation projects and add value to the company, favoring their sustainability. For this purpose, we provide this method that helps to identify company information and knowledge assets as a starting point for any project implementing actions on such assets.	**Define Requirements** This method can address the stated problem, which poses a challenge to companies that need to stay in the market and increase their value in the digital economy. Companies interested in this method merely have need of a consultant with knowledge of soft systems methodologies, agile BI, strategic management of intangible assets and natural language processing, because all they need to develop this method is information about their business processes, which may or may not be digitalized and formalized, irrespective of their maturity level.	**Develop Artifact** This method integrates four disciplines: 1) Agile principles, 2) systems thinking, 3) business intelligence techniques and the use of process mining software tools, and 4) natural language processing techniques. In so doing, it provides companies with an agile and systemic method whereby they can efficiently identify all their information and knowledge assets. This is indispensable before starting up any process mining project, as well as any other project that requires the use of these assets.	**Demonstrate Artifact** Companies can find it difficult to control a lot of unstructured data, like the emails that all employees use daily to communicate and share information primarily related to the development of their activities. However, companies do not know if there are information and knowledge assets that they are unaware of and critical for their business hidden among all this information. After applying this method, it is possible to analyze employee emails and discover intangible assets that are critical for the company and on which it would be possible to implement a process mining project.	**Evaluate Artifact** Regardless of their maturity level, companies will be able to apply this method, whereby they will be able to identify the information and knowledge assets that are critical for their business that increase their business value and help improve their decision making in order to succeed in their digital transformation projects, thus favoring sustainability. Additionally, once these assets have been identified, it will facilitate the implementation of process mining projects or any other project that requires action to be taken on these assets.
Structure This method integrates four disciplines: 1) Agile principles, 2) systems thinking, 3) business intelligence techniques and the use of process mining software tools, and 4) natural language processing techniques. This synergetic method is composed of a series of steps, including the definition of business objectives, system conceptualization and identification of information and knowledge assets, which are iterated as many times as necessary until most of the information and knowledge assets that are hidden among the company's structured, semi-structured and unstructured data have been identified.		**Function** This method offers companies the possibility to analyze the mainly unstructured data that it has and discover a number of hidden information and knowledge assets of whose existence they are unaware. Therefore, they are not included in the company accounts, which detracts from the company business value and sustainability. In addition, the development of this method contributes to improving company strategic decision making to undertake any digital transformation project and succeed in its implementation and thus achieve its sustainability. In this way, it is also possible to launch new process mining projects and ensure their proper development aligned with the strategic objectives defined by the company.		**Effects** The application of this method has direct effects on the improvement of digital transformation project development and the promotion of company sustainability, since success depends on the discovery of the business critical information and knowledge assets One of the possible indirect effects is the possibility of launching any type of project requiring some kind of action on these assets, including process mining projects.

Figure 1. Design science research canvas for the discovery of proposed C4PM method.

Questions 1 and 2 respond to the need for companies to discover what data and information their employees have and how they use these data and information to perform their daily activities. In response to the first question, this research sets out to conduct a preliminary analysis of the information that the company has about its organizational processes based on its strategic objective and generic intangible knowledge assets model (see [18] for more information about generic intangible assets). This will serve to define the business objectives with which the identified information and intangible knowledge assets will be aligned. Thus, it is possible to discover asset flows and which actors are involved in each of the organizational processes in the analyzed system.

Question 3 is motivated primarily by the intensification of new ways of working (flexible working) that require the use of collaborative tools and any number of cloud services. They make it difficult to control information shared among internal employees and with partners, especially if there are no data traceability software tools.

In response to Questions 2 and 3, agile principles (interviews, workshops and/or dynamic meetings) or process mining software tools are used to identify the system actors and the activities that they perform in every process under analysis.

Finally, Question 4 responds to the fact that all companies need to use their own data and intangible knowledge assets to improve their market and customer value and, thus, attain a more advanced level of sustainability and competitiveness. This way, employees can enjoy new ways of working that improve their productivity and help deliver improved products without putting the business at risk. In response to Question 4, semi-structured or unstructured data sources are analyzed, applying natural language processing techniques in order to discover new information and intangible knowledge assets. They can be used to identify which organizational processes are digitalized or duly formalized, as well as facilitate the digitalization of non-digitalized organizational processes that are part of the analyzed system. This should open up the possibility of implementing process mining projects that contribute to the digital transformation of the company. Currently, process mining is only accessible to a few companies that are digitally mature enough as to have detailed organizational processes that can be used by software tools to generate and store event logs. This research embraces cases where digital maturity is limited and no such logs yet exist. Thus, it paves the way for the use of process mining in this context. Figure 1 shows the developed DSR canvas.

C4PM is a new method that can be enacted by companies of any size or sector and useful for pre-processing unstructured data that are available to most companies. Properly processed, these data are extremely valuable for discovering an organization's hidden intangible assets that can lead to the discovery of company processes that can be a potential input for process mining methodologies. In addition, the identification and appropriate reuse of these intangible assets, that is, pieces of organizational knowledge, discovered by applying C4PM, leads to a more productive and sustainable organization. Indeed, it has been proven that knowledge reuse improves the productivity and sustainability of businesses on the basis of "do it taking advantage of what you have". In sum, information is power; an organization that uses mechanisms to reuse its own information and discover its intangible knowledge assets can materialize the power of information.

C4PM can output a list of all the critical intangible business assets discovered based on the existing (automated or non-automated) organizational processes. These organizational processes will be susceptible to be properly detailed according to the knowledge of the people performing the activities conforming to the processes. The output of C4PM provides potential input for process mining and reduces noisy data used in the processes analyzed by process mining discovery software tools. By linking C4PM with process mining, we ensure that the implementation of the process mining project is properly aligned with company strategic and business objectives from the very beginning. As a result, the process mining project can be implemented more efficiently, because it will be driven by the intangible assets identified with C4PM.

C4PM method is composed of three-step business intelligence and natural language processing techniques and applied to identify information and knowledge assets. The three steps are illustrated in Figure 2:

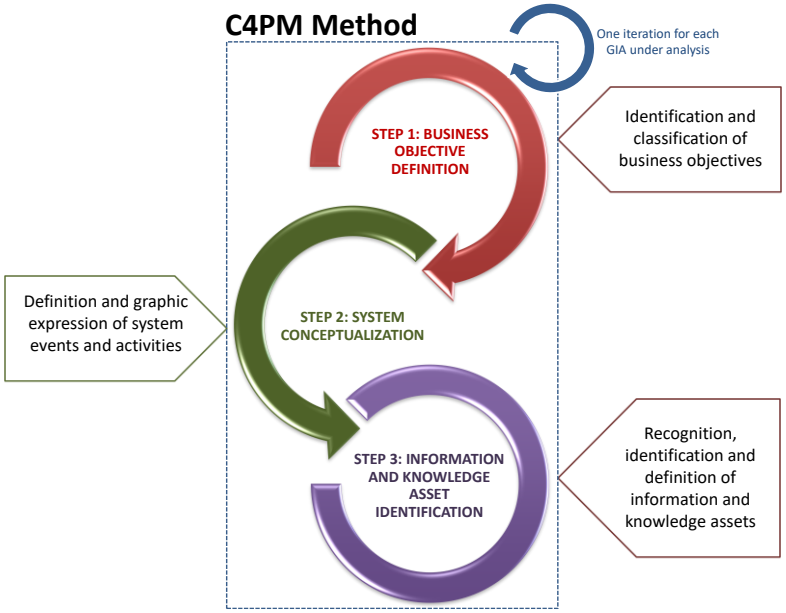

Figure 2. Steps of the proposed C4PM method.

STEP 1: Business objective definition. This step consists of identifying and classifying business objectives based on strategic objectives, according to the organizational processes provided by the company. It is composed of two activities: identification of business objectives and classification of business objectives.

For the company to enact the proposed method, it requires some initial information about its organizational processes. To this end, the CEO or company representative in charge of the project must specify the company's strategic objective and choose one type of generic intangible asset (GIA) for analysis in each iteration. Therefore, the C4PM method has to be iterated as many times as there are GIAs likely to provide information and specific knowledge assets that are possibly causing bottlenecks in company processes. Such bottlenecks may constitute a significant weakness on the path toward economic and/or technological progress and business sustainability, which is a key feature required within the new economy of connections. GIAs can be viewed as the anatomy of any company and provide the full picture, whereas each iteration selects the part of the company that needs to be analyzed. Therefore, this methodology is driven by GIAs. The following list of GIAs was sourced from other research within the intellectual capital (IC) discipline and was validated for its use as a tool to visualize the status of the business model canvas of industries in any business domain [18]:

- GIA 1. Production model/model of service execution: operational vision of production (goods and services) and related knowledge application and transfer practices.
- GIA 2. Commercial or customer model: company's commercial plan, including client management directives.
- GIA 3. Supply and diversification of services model/model of innovation: approach and process for any business development initiative related to the delivery of goods or services and their modification. It also includes management initiatives and policies for business improvement.
- GIA 4. Model of international geographic expansion: international business expansion guidance, procedures, policies and schemes.
- GIA 5. Models of human resources/professional development/principles and values: policies for the supervision, organization and administration of the company work-

force, including professional improvement and modernization and moral and ethical (professional and personal) basics for the business.
- GIA 6. Remuneration and property model: salary distribution policy, financial incentives and company ownership.
- GIA 7. Model of brand development: business vision and rules for marketing campaigns and improvements in public impact.
- GIA 8. Model of institutional relations and high-level networking/stakeholders: business regulations for professional alliances and industry and social involvement.
- GIA 9. Model of organization and processes: procedures and rules for the satisfactory operation of any business processes.
- GIA 10. Model of organizational strategy, mission and vision: high-level company strategies and directions and business philosophy, organizational and behavioral guidelines and relation with stakeholders and market institutions.
- GIA 11. Model of organizational knowledge management: company mechanisms, tools and models that enable the elicitation, gathering, recovery, use, evolution and valuation of company know-how.

For each C4PM iteration, one GIA from the above list is chosen. The precision of the information and knowledge assets that will be identified in C4PM Step 3, based on the chosen GIA will depend on how detailed the company's organizational processes are. To classify the information held by the company, we have three levels of detail:

- High: detailed description, formally established by the company, of its strategic objectives, organizational processes and mechanisms and qualitative and quantitative business objective performance indicators.
- Medium: less-detailed description of the strategic objectives and organizational processes that the company has formally established.
- Low: a generic description of the strategic objectives and organizational processes have not yet been formally established by the company.

STEP 2: System conceptualization. This step consists of identifying the areas working towards the business objectives of each organizational process, including the actors involved in the achievement of the business objectives of each of these areas, determining the type of relationship among actors and the type of knowledge that they share in the current system. It is composed of three activities: identification of areas involved, identification of system stakeholders and representation of system activities.

STEP 3: Information and knowledge asset identification. This step consists of identifying and classifying the information and knowledge transferred among the system stakeholders in each of the processes required to enact the organizational processes. It involves analyzing semi-structured and unstructured data to discover behavioral patterns that lead to the discovery of new information and knowledge assets for the company. It is composed of two activities: systemic conceptualization and identification of hidden information and knowledge assets.

4. Case Study Results

To test the C4PM method, we put the proposed method into practice on Enron, a multinational company that used to have a large share of the world energy market and, at the same time, developed several products in other industries, such as paper, water, etc. Enron went bankrupt in 2001 due to accounting fraud, leaving almost 21,000 people unemployed. Since this case was very well publicized around the world, there is a lot of information available on the Internet [17,19].

As C4PM is intended to complement process mining in cases where organizations' 'data are unstructured or semi structured', we decided to use data sources, such as email. Email is a key tool available at all companies and used by employees to transfer large volumes of data, usually containing important corporate information. Additionally, process mining is not yet capable of handling this type of unstructured data. Therefore, the application of C4PM helps to identify knowledge assets that were previously overlooked by

the company. These newly identified, potentially business-critical knowledge assets, such as organizational processes, can then be analyzed applying the principles of process mining.

In this respect, we discovered that several groups of researchers had conducted research with Enron employee email data [20,21]. These contained unstructured and very important data that were suitable for the purposes of applying the natural language processing techniques used in this research. In addition, it was possible to find a lot of relevant public information about Enron's operation.

The dataset used to test the current C4PM proposal was collected and prepared by the CALO (Cognitive Assistant that Learns and Organizes) project. CALO was an outcome of the PAL (Personal Assistant that Learns) framework developed by DARPA, a massive collaboration project led by SRI International (https://pal.sri.com/ last access on 10 October 2022). It contains data from about 150 users, mostly Enron senior managers, organized into folders. The corpus contains a total of about 0.5 M messages. These data were originally made public and posted on the web by the Federal Energy Regulatory Commission during its investigation.

The email dataset was later purchased by Leslie Kaelbling at MIT and turned out to have a number of integrity problems. A number of SRI members, notably Melinda Gervasio, worked hard to correct these problems and it is thanks to them that the dataset is available. The dataset does not include attachments and some messages were deleted "as part of a redaction effort due to requests from affected employees". Invalid email addresses were converted to something like user@enron.com whenever possible (i.e., recipient was specified in some parsable format, such as "Doe, John" or "Mary K. Smith") and to no_address@enron.com when no recipient was specified.

This dataset has been distributed as a resource for researchers who are interested in improving current email tools or understanding how email is currently used. These data are valuable because they are the only substantial public collection of "real" email. They are particularly suited for the purpose of the C4PM validation, because C4PM requires unstructured data. In this paper, we used the 7 May 2015 dataset version (about 1.7Gb, tarred and gzipped) [20].

Note that C4PM was validated using the Enron dataset because it is the only dataset that contains public unstructured data registering conversations in a real organizational environment. We do not set out to portray Enron as a good example of a digital company or to provide companies with a guide to digital transformation per se. Rather, we describe a model, used within the digital transformation process, to analyze unstructured data within emails to which every company has access but of whose value they are unaware. After applying C4PM to analyze these emails, these unstructured data can be turned into valuable knowledge assets.

As an example of the applicability of our proposal, each step of the proposed C4PM method was applied to the Enron dataset, selecting GIA 1. Production model/model of services execution in one iteration of the C4PM model execution. This method provided a holistic view of the system, leading to the expected results, even though it was not possible to interact with any actor in the (extinct) system. In fact, it was possible to apply natural language processing techniques to analyze the semi-structured data contained in Enron's emails. We discovered behavioral patterns in the email data, which we used to define more information and knowledge assets that were critical for the business. These assets are not usually considered by IT professionals in digital transformation processes or in software solutions that improve support for strategic business decision making.

In this case, based on the public information available on the Internet about Enron and the causes that led it to its bankruptcy, it was possible to gather some preliminary information regarding compliance with the requirements of the proposed method. Based on this information, the notable processes are related to Enron's operational vision of the production of goods and services, as well as the application and transfer of knowledge within the energy market. Therefore, as inputs, we were able to infer that the level of detail available for this analysis of the strategic objective and GIA 1 at Enron was low.

Accordingly, there were no defined and publicly available organizational processes to help analyze the production model or the service execution model, nor was it possible to ascertain which management tools could be used to automatically extract any process. Therefore, this information fails to detail or generically describe any of the organizational processes that the company would have established to achieve its strategic objective [21].

Table 2 shows the results of an analysis of the information required to launch the C4PM method available at the organization. Because the level of detail is low, the organizational processes necessary to develop **Step 1** to identify and classify business objectives are not defined.

Table 2. Level of detail of the information held by the company.

Inputs	Level of Detail
Strategic objective: To increase the value of the company's shares	Low
Generic intangible asset model: GIA 1. Production model/model of service execution	Low

During the development of **Step 2**, we identified two areas involved within the system: Accounting Department and Investment Relations Department, as well as information about the organizational processes.

Step 3 consists of the identification of intangible knowledge assets. The first set of intangible assets was identified by manually analyzing all the documentation available for each area of the company (accounting and investment departments, respectively) and those assets are shown in Figures 3 and 4. These information and knowledge assets are classified according to the type (structural, human or relational) of knowledge asset and its source.

ACCOUNTING DEPARTMENT										
ORGANISATIONAL PROCESS	SOURCE TYPE	STRUCTURAL KNOWLEDGE ASSETS			HUMAN KNOWLEDGE ASSETS			RELATIONAL KNOWLEDGE ASSETS		
		Structured Data	Semi-Structured Data	Unstructured Data	Structured Data	Semi-Structured Data	Unstructured Data	Structured Data	Semi-Structured Data	Unstructured Data
Not defined	INTERNAL	Product price lists	Email address	Corporate applications, file storage applications		Projects for the creation of financial products				
	EXTERNAL	Product price lists	Financial reporting Email address							

Figure 3. Information and knowledge assets of the accounting department.

INVESTMENT RELATIONS DEPARTMENT										
ORGANISATIONAL PROCESS	SOURCE TYPE	STRUCTURAL KNOWLEDGE ASSETS			HUMAN KNOWLEDGE ASSETS			RELATIONAL KNOWLEDGE ASSETS		
		Structured Data	Semi-Structured Data	Unstructured Data	Structured Data	Semi-Structured Data	Unstructured Data	Structured Data	Semi-Structured Data	Unstructured Data
Not defined	INTERNAL		Market research document, financial reports, business models, trade agreements, new business creation projects, financial product creation projects, E-mail	Marketplace, corporate applications, file storage applications		Gas transmission and production processes, electricity transmission and generation processes, projects for the creation of new companies, projects for the creation of financial products			Emails, contracts with suppliers, contracts with customers, outsourcing contracts	Minutes and audios of meetings with stakeholders, suppliers, managers to close business deals
	EXTERNAL		Projects for the creation of new companies, commercial agreements			Projects for the creation of new companies			Emails, outsourcing contracts, contracts with suppliers	

Figure 4. Information and knowledge assets in the investment relations department.

Figures 3 and 4 illustrate that email was identified as one of the information and knowledge assets, since it was found to be used transversally by the different areas of the company.

As a second part of **Step 3,** email was analyzed, which is a semi-structured data type from internal and external sources used in both the accounting and investment relations departments, to discover a set of new knowledge assets. It is expected to find much more

knowledge assets from the email analysis. To do this, natural language processing (NLP) techniques were applied to the entire available dataset, using Jupyter Notebook (a Python programming language interpreter for data analysis). As Jupyter Notebook uses different libraries for NLP, algorithms and data exploration techniques were applied to identify new information and knowledge assets by analyzing the text in the email dataset.

First, we converted the dataset that contains the Enron company emails ("enron_mail_20150507") to a format that is easily manipulable using Python (enron.mbox) and analyzed all the emails that were located in the inbox directories for 16 employees (see Figure 5).

```
df_original.head(2)
```

	Message-ID	Body	Date	From	To	
0	<14955894.1075855377681.JavaMail.evans@thyme>	_____DAVID COURSEY_____HOPE AHEAD: WHAT I LEARNED FROM 2001'S TRAGEDIES As years go, 2001 sucked. But adversity teaches us ...	Sun, 30 Dec 2001 22:49:42 -0800 (PST)	anchordesk_daily@anchordesk.zdlists.com	pallen@enron.com	ANCH Hc Wha fr
1	<7462038.1075855377703.JavaMail.evans@thyme>	Dear phillip,This e-mail is automated notification of the availability of yourcurrent Natural Gas Intelligence Newsletter(s). Please use yourusername of "pallen" and ...	Sun, 30 Dec 2001 23:42:30 -0800 (PST)	subscriptions@intelligencepress.com	pallen@enron.com	NGI P Dece

2 rows × 24 columns

Figure 5. Example of the content of the Enron company's email dataset.

After reading the records in a table format, we removed the columns that were of no use in the analysis, leaving only the columns that refer to the email recipient and the body of the message (see Figure 6).

	EMPLOYEE1	BODY_MAIL
EMPLOYEE		
Allen-P	allen-p	_____DAVID COURSEY_____HOPE AHEAD: WHAT I LEARNED FROM 2001'S TRAGEDIES As years go, 2001 sucked. But adversity teaches us ...
Allen-P	allen-p	Dear phillip,This e-mail is automated notification of the availability of yourcurrent Natural Gas Intelligence Newsletter(s). Please use yourusername of "pallen" and ...
Allen-P	allen-p	[IMAGE] [IMAGE] [IMAGE] [IMAGE] $ 2500 [IMAGE] [IMAGE] [IMAGE] Dear Phillip, You've got to spin to win! Play now! Spin the iWon Prize Machine 2 for...

Figure 6. Example of a simplified table showing the records of the columns to be analyzed.

As Figure 6 shows, the simplified table contains duplicate records. Therefore, we had to remove duplicates, integrate records and convert all text to lower case to facilitate analyzed text processing. Thus, Figure 7 illustrates that there are no more duplicate records (since they were merged by email recipient).

```
df_data = df_data.groupby(['EMPLOYEE1']).agg('sum')
df_data.pivot_table(index=['EMPLOYEE1'], aggfunc='size')
```

```
EMPLOYEE1
allen-p        1
arnold-j       1
arora-h        1
badeer-r       1
bailey-s       1
              ..
williams-w3    1
wolfe-j        1
ybarbo-p       1
zipper-a       1
zufferli-j     1
Length: 142, dtype: int64
```

Figure 7. Example of record integration.

For the data analysis process, the body of the message was merged into one long text (see Figure 8), which could be processed using data cleaning and NLP techniques.

```
df_data.BODY_MAIL.loc['dean-c']
```

```
'As our last day is Friday, November 30th, we would love to toast the good times and special memories that we have shared wit
h you over the past five years.  Please join us at Teala\'s (W. Dallas) on Thursday, November 29th, beginning at 5pm.  Looki
ng forward to being with you, Lara and Janel  Lara Leibman713.851.7770 (cellular)713.528.5281 (home)lleibman@houston.rr.com
<mailto:lleibman@houston.rr.com> Janel Guerrero713.851.3778 (cellular)713.524.1534 (home)travelgirl_janel@hotmail.com <mailt
o:travelgirl_janel@hotmail.com>Jerry Scarbrough\'s True OrangeThe Newsletter for the True Texas Longhorn FaithfulVolume 12, N
o. 6, November 26, 2001(Editor\'s Note - I\'m sending this in four pieces today because it is a biggerissue than usual due to
the spring recruiting in basketball, track andsoftball.)Volume 12, No. 6, November 26, 2001 Benson, Defense Stifle Ags; Huge
Upsets PropelHorns Into Big 12 Title Game Against ColoradoAs Longhorn defensive coordinator Carl Reese noted, it was an old-f
ashioned,rock-em, sock-em game in College Station before a state record football crowdof 87,555 Friday, and Texas won it with
a gritty defense and two latetouchdowns by outstanding freshman RB Cedric Benson.After taking the 21-7 victory, head coach Ma
ck Brown congratulated the teamon its fine 10-1 season and said he wasn\'t worried about the BowlChampionship Series (BCS) "b
ecause we\'ll be in it."I don\'t know if had a crystal ball, but immediately after the Longhorns\'expected victory, top-rated
Nebraska was humiliated, 62-36, by Colorado, andthe next day No. 3 Oklahoma, a 27-point favorite lost at home to OklahomaStat
e, 16-13. It was only the fifth time OSU has won at OU.The good news is those games moved Texas up to the No. 3 spot in both
majorpolls and put the Longhorns back in control of their own destiny for thefirst time since that 14-3 loss to OU back in Oc
tober.Texas plays Colorado Saturday at 7 p.m. in Irving\'s Texas Stadium for the Big12 Championship and an automatic berth in
one of the four BCS bowls, possiblyeven the Rose Bowl in the national title game.But the bad news is that a loss to the stamp
eding Buffaloes would knock theLonghorns out of the BCS mix.Brown and his players are tickled orange, however, at the chance
```

Figure 8. Example of merged text from the body of messages from employee "Dean-C".

After applying cleaning techniques to remove characters that are of no use in the analysis, remove blank spaces and remove URLs and email addresses that are of no use in the analysis either, we output a clean unified text for each of the email recipients, as shown in the example in Figure 9.

```
df_data.BODY_MAIL.loc['ermis-f']
```

```
'offer to fantasy members take off colosseum ncaa block and tackle jerseys and get into the game it s an easy play simply ent
er coupon code vryehygp into step of the check out process to receive your discount but don t fumble offer ends want to win y
our fantasy league our fantasy football guides are the source for strategy player ratings scouting reports team reports proje
ctions and more a must have for beginners and fantasy veterans alike special in season price going fast click here fantasy fa
ns subscribe to the sporting news now and get free issues hurry brought to you by sponsorship bar are receiving these e repor
ts because you have signed up for cbs sportsline com fantasy football to customize reschedule or turn off these reports pleas
e click here reports player updates nfl player newsterry glenn wr ne matty updated glenn did not report to practice on wednes
day it is uclear if and when he will play for the patriots again eddie kennison wr kc free agent updated kennison might see s
ome limited action against the raiders this week even if he plays extensively at some point kennison will have minimal fantas
y value doug brien random key k free agent updated brien was signed by the colts with mike vanderjagt nursing a sore back bri
en will be used on kickoffs and could see some occasional field goal work he will be an accurate reliable fantasy choice if v
anderjagt misses any playing time dave moore te tb free agent updated an mri on moore s hip was negative but he may still mis
s sunday s game against detroit the bucs activated mike roberg from the practice squad as a precaution don t use moore this w
eek anthony thomas rb chi martin updated thomas will test his strained hamstring in full speed practice wednesday we want to
try to get him healthy coach dick jauron told the chicago tribune hopefully by resting him for two weeks we ll have him back
healthy and james allen has done an outstanding job in the meantime if thomas is not ready to go we re comfortable with it mo
nitor thomas progress before you make a final decision on starting a chicago rb this week ricky watters rb sea free agent upd
ated watters is close to returning from a shoulder injury according to the seattle post intelligencer the plan is for ricky t
o practice th is week whether he can play this sunday that s another question coach and general manager mike holmgren said it
```

Figure 9. Example of clean merged text from the body of messages of employee ermis-f.

To analyze the text contained in the messages, the corpus (long text) was then converted into a matrix of TF-IDF characteristics, excluding all the English words that had little meaning (stop words). Then, CountVectorizer was used. CountVectorizer tokens the records and counts the token occurrences, which it then returns as a scattered matrix.

The text was analyzed based on this matrix and word clouds were used to represent the frequency of the words that make up the body of the emails from each Enron employee. Word clouds visualize the most frequently used words, which have a larger font and stand out from the others (see Figure 10).

Figure 10. Example of graphical representation of the frequency of words used in the body of the messages from employees Allen-p and Arnorld-j.

The intangible assets were identified after analyzing the emails of 16 out of the 150 employees contained in the dataset using Jupyter Notebook and extracting the most frequent words. The 16 selected employees were "allen-p", "arnold-j", "arora-h", "causholli-m", "corman-s", "crandell-s", "cuilla-m", "dasovich-j", "davis-d", "dean-c", "delainey-d", "derrick-j", "donoho-l", "donohoe-t", "dorland-c" and "ermis-f". A total of 962,385 words was analyzed for each of the 16 employees. The analysis identified the most frequent words used by each employee in their emails, which were then used to discover words that infer some type of knowledge asset. Finally, the knowledge assets were classified by type within each area.

Table 3 summarizes the words most frequently used by each employee. The 60 words most frequently used by all 16 employees whose emails were analyzed are highlighted in bold. Column 3 in Table 3 summarizes the identified assets, described in Sections 4.1 and 4.2.

Table 3. Most frequently used words and identified assets by employee.

Employee	Most Frequently Used Words	Identified Assets
allen-p	buy, mail, information, message, phillip, downgraded, **border**, account, email, use, know, enron, request, new, need, see, password, strong, **shares**, sent, time, recipient, change, like, **price**, review, receive, original, subject, free, **stock**, coverage, first, initiated, company, type, date, corp, visit, read, back, allen, send, **plan**, questions, **distribution**, make, year, look, section, upgraded, **cash**, call, days, available, home, meeting, last, want, today	- Order Request Procedures - Stock list - Border protocols
arnold-j	enron, company, said, new, **credit**, financial, **trading**, **energy**, billion, **stock**, week, **shares**, john, **dynegy**, business, jones, york, last, **investors**, year, message, dow, corp, mr, **deal**, news, houston, gas, sent, price, power, market, **service**, day, **companies**, street, rating, could, time, original, wall, million, rights, copyright, like, **cash**, fastow, **transactions**, know, email, reserved, monday, chief, reuters, friday, **debt**, **markets**, exchange, **partnerships**, offer	- Register of shares - Procedures for share purchases - List of services - Energy trading procedure - Agreements with other companies - Investor agreements

Table 3. *Cont.*

Employee	Most Frequently Used Words	Identified Assets
arora-h	size, align, right, nbsp, new, left, enron, request, day, **rates**, year, message, call, free, see, continental, miles, time, **report**, week, date, email, specials, **information**, available, **research**, **rate**, harry, sent, power, make, **analyst**, houston, options, subject, review, **price**, offer, daily, november, businesses, **market**, mail, online, questions, **purchase**, **investor**, original, table, dynegy, page, monday, want, gas, name, financial, **business**, hr, need, **data**	- List of customers
causholli-m	top, align, gt, right, news, story, nov, valign, cn, enron, **pulp**, **paper**, section, message, detail, control, function, nbsp, digest, table, november, page, view, **border**, mail, **information**, contact, **market**, email, sent, original, press, mailto, edition, time, left, li, release, ou, october, dna, year, name, recipient, today, new, corp, use, **service**, **products**, like, **cellpadding**, online, **finance**, lumber, company, forest, intended, **prices**	- Border protocols - Pulp extraction process for paper production - Forest exploitation reports
corman-s	**gas**, day, october, sent, **order**, original, enron, know, need, call, **information**, **related**, message, shelley, pipelines, **cash**, security, **market**, **report**, new, imbalance, time, **business**, net, tw, email, questions, month, attached, area, mail, line, march, **meeting**, number, control, like, monday, issues, imbalances, **interstate**, quantities, pmto, **standards**, **capacity**, dell, european, july, gisb, contact, revised, following, comments, current, nexis, gary, contract, pipeline, send, kim	- Cash flow reports - List of customers - Gas standards guidelines - Customer contracts in the gas sector - Partner agreements
crandell-s	message, day, original, sent, time, enron, power, **energy**, know, sean, new, **deal**, **product**, pmto, real, ahead, call, october, wednesday, **price**, like, news, amto, **meeting**, thursday, friday, need, west, **information**, **market**, rto, monday, tuesday, questions, mw, change, california, said, **deals**, commission, week, mail, november, center, today, **delivery**, **business**, fran, see, parking, financial, group, portland, could, **contract**, ferc, bpa, steve, month, **scheduling**	- Minutes of meetings agreements - Product delivery procedure in the energy sector
cuilla-m	class, align, game, allowed, right, size, day, yards, face, last, points, helvetica, fantasy, height, src, week, new, year, miles, nbsp, **rates**, rank, passing, weeks, november, target, bgcolor, continental, sunday, vs, color, self, **border**, houston, specials, message, enron, hilton, gas, free, left, city, cera, martin, time, **information**, **offers**, available, use, book, table, middle, **offer**, car, updated, hotels, **listed**, **energy**, rushing, airport	- List of rates - List of orders - List of customers
dasovich-j	enron, said, company, power, **energy**, new, **market**, **gas**, state, business, year, size, mail, message, million, billion, jones, nbsp, table, face, corp, california, time, **border**, dow, sent, **information**, financial, jeff, last, **stock**, trading, october, original, could, verdana, **service**, helvetica, week, companies, news, **cellpadding**, **credit**, commission, **electricity**, call, mailto, houston, align, **price**, right, know, cellspacing, subject, use, september, years, center, **shares**, like	- Sales reports - Border protocols
davis-d	enron, amp, new, mail, **market**, mitigation, **dynegy**, november, message, sent, **information**, **company**, power, day, mailto, original, fw, october, god, subject, commission, know, businesses, york, continue, make, time, see, **work**, size, amto, face, home, need, free, pray, **employees**, review, **people**, believe, **financial**, like, call, sec, year, every, including, proposed, today, **measures**, available, ferc, million, chocolate, bids, houston, help, questions, color, without	- Market reports on employee management in the energy sector - On-boarding and employee monitoring procedures - Agreements with other companies

Table 3. *Cont.*

Employee	Most Frequently Used Words	Identified Assets
dean-c	type, **database**, final, trans, unknown, date, **schedule**, epmi, sc, **hour**, **schedules**, mkt, **details**, preferred, found, **california**, point, data, iso, tie, portland, variances, detected, **price**, **scheduling**, **file**, log, messages, table, term, **process**, ancillary, continuing, **awarded**, cannot, enron, **perform**, closed, **operation**, wheel, start, variance, **energy**, short, occurred, attempting, engine, total, user, **time**, **transaction**, disk, **deal**, **progress**, **import**, **export**, sp, long, **trading**, **sale**	- Database maintenance procedure - Continuity planning procedures - Information systems audit reports - Information retrieval processes - Security standards - California power system operation scheduling document
delainey-d	**contracts**, december, message, david, enron, forward, issues, moving, **employee**, sent, original, call, unavailable, paid, **meeting**, notice, new, mailbox, brown, delainey, **contract**, friday, **expense**, **report**, time, following, know, bankruptcy, outlook, daniel, regarding, assets, **expenses**, terminate, pmto, need, retail, **goods**, contact, household, number, going, **payment**, think, status, amount, gas, yet, server, november, bill, worthy, agenda, received, luce, alan, **scheduled**, work, continue, today	- Payment reports - List of customers - Customer contracts in gas sector - Contract with customers and suppliers - Expenditure reports
derrick-j	page, enron, mail, message, **information**, sent, email, know, new, original, jim, state, epa, **business**, court, **meeting**, subject, **law**, june, power, time, corp, group, could, jr, mailto, company, **federal**, bna, **committee**, **report**, **legal**, review, year, week, board, think, need, gas, pmto, use, number, mark, **agreement**, intended, attached, work, national, today, contact, derrick, **pay**, **energy**, last, like, call, **risk**, **office**, **management**, **project**	- Risk Reports - Minutes of meetings agreements - Court reports
donoho-l	tw, enron, message, sent, original, **gas**, **capacity**, november, **report**, **market**, **information**, october, new, time, know, pipeline, socal, **dynegy**, pmto, like, **contract**, attached, company, year, friday, rate, **california**, **business**, power, comments, need, use, review, available, tuesday, **meeting**, amto, questions, **file**, said, changes, **project**, transwestern, **system**, mmbtu, next, email, last, wednesday, week, **order**, pd, make, **energy**, **service**, **volumes**, watson, core, mail, houston,	- Minutes of meeting agreements - Project initiation minutes - Agreements with other companies
donohoe-t	enron, **scheduled**, thru, **gas**, **outages**, **energy**, new, **information**, **power**, sat, contact, center, call, **report**, **deal**, pager, time, october, fri, original, **impact**, sun, london, ct, **market**, **risk**, **services**, **data**, natural, pt, available, email, impacted, company, **system**, help, houston, mail, server, questions, year, week, **business**, sent, **londonoutage**, message, november, **employees**, today, trading, **industry**, day, see, **management**, **operations**, north, free, know, **support**, cms	- Procedure for the operation of IT systems - List of gas company employees - Time scheduling documents for the operation of gas systems - Gas system operation procedure - Agreements with other companies
dorland-c	enron, message, **energy**, **information**, use, cera, **report**, mail, new, **research**, **market**, online, january, today, february, attachments, **distribution**, chris, email, contact, sent, **associates**, prohibited, access, cambridge, original, **area**, year, questions, password, **company**, day, intended, pira, strictly, contain, know, free, **gas**, **list**, **simulation**, **companies**, privileged, power, available, **oil**, executive, part, make, **conference**, follow, **western**, recipient, **electronic**, **disclosure**, changes, **events**, **data**, attachment, follows	- Research work in the energy, oil and gas industry - List of partners
ermis-f	updated, fantasy, week, free, **game**, agent, wr, sunday, rb, start, play, **season**, yards, injury, **reports**, **gas**, **player**, smith, expected, **league**, **practice**, day, new, back, starting, time, still, enron, year, te, **team**, **energy**, **football**, good, **signed**, available, **scheduled**, reserve, thru, **company**, weeks, exhibit, martin, last, green, monday, pts, **listed**, **information**, ankle, injured, passes, receiver, james, contact, **sportsline**, **numbers**, however, nfl, frank	- List of sports teams - Sports practice guide - Championship results report

"Most frequent words" provide a source of very useful information. Once they have been identified, the software engineer can be more efficient following the leads provided by those words and can, in a more efficient way, dig into the organization to identify procedures and processes and then check their digital maturity to realize whether or not the identified process can be an entry to the process mining methodology. The Identified Assets column in Table 3 was obtained digging into the most frequent words by analyzing the public documentation of the Enron company.

This result shows that C4PM is an effective and efficient technique for the first and one of the most difficult digital transformation tasks: discovery of company intangible assets. The audit of the company used to identify intangible assets uncovers company knowledge to underpin the digital transformation project.

Table 4 summarizes the number of valuable knowledge assets discovered by the C4PM method and how they are catalogued in terms of their impact on the intellectual capital of the organization.

Table 4. Summary of discovered valuable assets.

ENRON	Type of Intellectual Capital	Number of Knowledge Assets Discovered
Accounting Department (20 valuable assets discovered)	Structural Capital	16
	Human Capital	4
	Relational Capital	0
Investment Relations Department (26 valuable assets discovered)	Structural Capital	6
	Human Capital	13
	Relational Capital	7

The information and knowledge assets that were discovered for each particular department after scanning the body of the emails in the Enron employee inboxes are described below.

4.1. Valuable Assets Discovered for Accounting Department

The information and knowledge assets that were discovered are classified according to the type of knowledge they represent.

4.1.1. Structural Knowledge Type (16 Identified Knowledge Assets)

1. Order request procedure: The company can provide a clear definition of this asset to the respective area to improve the management of customer order requests. This asset can also be adapted and/or improved within the company's organizational processes.
2. Stock list: With this asset, the company will be able to analyze the available lists and define a template that facilitates stock management for its employees.
3. Register of shares: With this asset, the company will be able to keep control of all the shares it acquires, helping it to make decisions on its investments.
4. Procedure for share purchases: The company can provide a clear definition of this asset to the respective area to improve the management of share purchases. This asset can also be adapted and/or improved within the company's organizational processes.
5. List of services: With this asset, the company will be able to monitor the services offered and promote their internal and external dissemination, whereby it will be able to consider/improve new services.
6. Cash flow reports: With this asset, the company will be able to adapt/improve the cash flow reporting templates, providing the respective area with better financial control, which contributes to making better strategic decisions.
7. List of rates: The company will be able to provide a clear definition of this asset to the respective areas to control the tariffs applied to the services. Likewise, this asset can

be analyzed to define a single format that improves the time taken to define the tariffs applied to the services.

8. List of orders: With this asset, the company will be able to analyze the available lists and define a template that makes it easier for its employees to attend to orders, improving response times to customers.
9. Sales reports: With this asset, the company will be able to adapt/improve the sales report templates, providing the respective area with a better financial control of sales and, thus, helping to make better, more timely decisions.
10. Expenditure reports: With this asset, the company will be able to adapt/improve the expense report templates, providing the respective area with better financial control of expenses, thus, contributing to better strategic decision making.
11. Payment reports: With this asset, the company will be able to adapt/improve payment report templates, providing the respective area with better financial control of payments, thus, contributing to better strategic decision making.
12. Risk reports: With this asset, the company will be able to analyze and adapt/improve the available reports by defining a template, which facilitates better risk control over investments.
13. Database maintenance procedure: The company can provide a clear definition of this asset to the respective area to improve the management of the databases. This asset can also be adapted and/or improved within the company's organizational processes.
14. Procedure for the operation of IT systems: The company can provide a clear definition of this asset to the corresponding area to improve the management of the databases. This asset can also be adapted and/or improved within the company's organizational processes.
15. Continuity planning procedures: The company can provide a clear definition of this asset to the respective area to ensure the integrity and continuity of the systems. This asset may also be adapted and/or enhanced within the organizational processes of the company.
16. Information retrieval processes: The company can provide a clear definition of this asset to the respective area to ensure contingency plans for the recovery of system information. Likewise, this asset may be adapted and/or improved within the company's organizational processes.

4.1.2. Human Knowledge Type (Four Identified Knowledge Assets)

1. Security standards compliance criteria documents (ISO): With this asset, the company will be able to analyze how the process is carried out for improvement purposes and to establish formats to help the respective area understand, update and identify the company's shortcomings in terms of maintaining ISO standards.
2. Information systems audit reports: The company will be able to analyze this asset to improve the security, availability and integrity of its information systems. In addition, it will be able to adapt/create a template to improve decision making in the respective area.
3. Minutes of meeting agreements: The company will be able to analyze this asset to improve decision making, keeping the focus on meeting agreements. In this way, the company will be able to make meetings more and more productive.
4. Project initiation minutes: The company will be able to analyze this asset to improve decision making on projects in order to produce the desired deliverables within the established timeframe.

4.2. Valuable Assets Discovered for Investment Relations Department

The discovered information and knowledge assets are classified according to the type of knowledge they represent.

4.2.1. Structural Knowledge Type (Six Identified Knowledge Assets)

1. Border protocols: The company will be able to provide a clear definition of this asset to the respective area for adaptation and/or improvement within the company's organizational processes.
2. Energy trading procedure: The company can provide a clear definition of this asset to the respective area to improve energy service commercialization. This asset can also be adapted and/or improved within the company's organizational processes.
3. List of customers: With this asset, the company will be able to analyze the available lists and define a template that will enable its employees to improve customer service and promote customer loyalty.
4. Market reports on employee management in the energy sector: With this asset, the company will be able to adapt/improve the templates of market reports, helping the respective area to carry out better talent management in energy services.
5. On-boarding and employee monitoring procedures: The company will be able to provide a clear definition of this asset to the respective area to improve talent management in the organization. Likewise, this asset can be adapted and/or improved within the company's organizational processes.
6. List of gas company employees: With this asset, the company will be able to analyze the available rosters and define a template that will help the organization to understand the existing talent in the gas sector.

4.2.2. Human Knowledge Type (13 Identified Knowledge Assets)

1. Pulp extraction process for paper production: The company will be able to provide a clear definition of this asset to the respective area to improve the pulp extraction process and facilitate paper manufacturing.
2. Forest exploitation reports: With this asset, the company will be able to adapt/improve the templates of forest exploitation reports, providing the respective area with better control over this activity, thus, contributing to better strategic decision making.
3. Gas standards guidelines: With this asset, the company will be able to analyze how standards are evolving in the gas sector to enable the respective area to understand, update and identify gaps in the company's ability to maintain these standards.
4. Product delivery procedure in the energy sector: The company can provide a clear definition of this asset to the respective area to improve the delivery of energy service products to its customers. This asset can also be adapted and/or improved within the company's organizational processes.
5. Minutes of meeting agreements: The company will be able to analyze this asset to improve decision making, keeping the focus on meeting agreements. In this way, the company will be able to make meetings more and more productive.
6. Project initiation minutes: The company will be able to analyze this asset to improve decision making on projects and output the desired deliverables within the established timeframe.
7. Time scheduling documents for the operation of gas systems: With this asset, the company will be able to analyze how this process is carried out for improvement purposes and establish formats to help the respective area to understand, update and schedule gas system operation that guarantees service availability for customers.
8. California Power System Operation Scheduling Documents: With this asset, the company will be able to analyze how this process is carried out for improvement purposes and establish formats to help the respective area to understand, update and schedule energy system operation in a state (California) that guarantees service availability for customers.
9. Gas system operating procedure: The company will be able to provide a clear definition of this asset to the respective area to improve the gas system operation process.

10. Research work in the energy, oil and gas industry: With this asset, the company will be able to analyze scientific advances in the energy, oil and gas industry to improve its production processes and the strategy to be applied in each sector.
11. List of sports teams: With this asset, the company will be able to analyze the available rosters and define a template that will enable the respective area to better manage talent for sporting activities.
12. Sports practice guide: With this asset, the company will be able to analyze how employee sports activities are carried out and improve their management.
13. Championship results reports: With this asset, the company will be able to analyze the development of sporting activities in order to make better strategic decisions about the organization's talent.

4.2.3. Relational Knowledge Type (Seven Identified Knowledge Assets)

1. Agreements with other companies (Dynergy): The company will be able to analyze this asset to improve strategic decision making on agreements with other organizations.
2. Investor agreements: The company will be able to analyze this asset to improve strategic decision making on agreements with various investors.
3. Customer contracts in the gas sector: With this asset, the company will be able to analyze how contracts with customers in the gas sector are drawn up in order to improve their format and ensure compliance.
4. Partner agreements: The company will be able to analyze this asset to improve strategic decision making on the services provided, based on the agreements established with various partners.
5. Contracts with customers and suppliers: With this asset, the company will be able to analyze how contracts with customers and service providers are drawn up in order to improve their format and ensure compliance.
6. Court reports: With this asset, the company will be able to analyze and report the cases settled by a court to the respective area in order to establish the necessary corrective measures that lead to better strategic decision making.
7. List of partners: With this asset, the company will be able to analyze the available lists and define a template to help the respective area better manage activities with the organization's partners.

5. Discussion: Theoretical and Practical Implications

With C4PM, IT professionals have an effective, holistic, systemic, iterative business-goal-oriented and GIA-driven method to identify the information and knowledge assets that will ultimately support digital processes.

A noteworthy important aspect of this proposal is its power to identify an organization's intangible knowledge assets from information contained in the body of the text of real emails. Additionally, their classification by asset type (structural, human or relational) provides insights into the effect of these assets and the waterfall effect of their state of intellectual capital health. This issue is related to the first research question: What is the flow of non-physical assets within the organizational processes?

The use of NLP and BI techniques for the analysis of corporate unstructured data helps enterprises use existing information to discover other business-critical information and knowledge assets with which companies are unfamiliar. In addition, it is in the company's interest to expand the application of process mining with C4PM prior to the process mining discovery stage to help the company on its way to digital transformation and sustainability. This should ensure that organizational processes that have not yet been digitalized or formalized and of which the company is unaware are accounted for. C4PM also reduces the noise surrounding the implementation of process mining projects because it allows companies without event logs exported from the software tools to use process mining.

The use to which assets are put may be related to frequency of use and asset importance. These effects, related to the second research question (How can we discover the use to which these assets are put?), are the roots for future research, focused on creating and monitoring indicators based on the above asset identification.

Regarding the third question (How locatable are these assets?), we integrated an external source of knowledge in order to locate the assets from an intellectual capital architecture. The classification of knowledge assets by type makes it possible to directly connect specific company assets to a company's specific organizational structure. In the case of Enron, the intellectual capital architecture cannot be matched to the organizational structure, since Enron is no longer in operation, but future research will focus on identifying such connections.

With a focus on which assets may be considered to improve enterprise sustainability and competitiveness (Question 4), this proposal includes a process of refinement; as part of this process, the list of assets suggests which points should be looked at and measured to control and improve the business from the intangible knowledge assets perspective.

The application of C4PM, as described in the reported case study, identified a huge number of assets using just the text of employee emails. Therefore, any organization can use their email dataset to identify assets that are potentially invisible or hidden for the organization. Knowledge asset discovery is essential for any organization with a view to successfully applying process mining on the road to the development of a sustainable, productive and competitive digital transformation.

According to the assets identified in each area, Enron would have had the capacity to agilely evaluate, value and formalize its assets within its organizational processes. Having identified and classified these assets, Enron would have found it easier to seek new sustainable technological solutions (such as, for example, the use of the cloud) to adapt or automate assets, thus, improving the management of the organizational processes underpinning the company structure (with structural assets), talent management without being affected by talent turnover (with human assets) and the management of strategic alliances that maintain the satisfaction of partners and investors (with relational assets). Furthermore, having identified these assets, the company's level of competitiveness in the market would have improved, since it would be constantly, agilely and sustainably taking advantage of the knowledge contained in each of these assets, contributing to making better strategic decisions. For example, by identifying the gas system operating procedure human asset within the investment relations area, it would have been possible to value, safeguard, automate and update this asset, since the knowledge that it contains is the result of having invested resources in customizing the gas systems that kept this service running.

Based on this proposal, a company will be able to consider all the information and intangible knowledge assets that it owns in order to establish quantitative and qualitative business plans. This will generate more business value and drive a sustainable and successful digital transformation, when needed.

In future research work, we intend to apply the proposed method to other domains with a semi-structured or unstructured dataset other than email. We also plan to apply the proposed method together with the process mining discovery stage on the output information and knowledge assets.

Author Contributions: The authors M.-I.S.-S., R.G.-C., F.M.-D. and G.-L.D.-P. have participated actively and in a collative way on conceptualization, method definition, validation, coding, writing and reviewing. All authors have read and agreed to the published version of the manuscript.

Funding: R&D for a Smart Digital Transformation of Occupational Health And Safety (OHS) Research Chair (https://catedrairsst.uc3m.es accessed on 10 October 2022).

Institutional Review Board Statement: No applicable.

Informed Consent Statement: No applicable.

Data Availability Statement: No applicable.

Acknowledgments: This work was supported by the RESTART project "Continuous Reverse Engineering for Software Product Lines/Ingeniería Inversa Continua para Líneas de Productos de Software" (ref. RTI2018-099915-B-I00), 2018 National Societal Challenge-Oriented R&D&I Research Program Call for R&D Projects and, also, by the R&D+I for a Smart Digital Transformation of Occupational Health And Safety (OHS) Research Chair (https://catedrairsst.uc3m.es, accessed on 4 October 2022.

Conflicts of Interest: Authors declare no potential conflict of interest in the research.

References

1. Andriole, S. The Hard Truth About Soft Digital Transformation. *IT Prof.* **2020**, *22*, 13–16. [CrossRef]
2. U.N. Digital Economy Report 2019. In *Value Creation and Capture: Implications for Developing Countries*; United Nations: New York, NY, USA, 2019; ISBN1 9210042166, ISBN2 9789210042161.
3. Andrew, J.H. The next phase of business sustainability. In *Stanford Social Innovation Review*; Stanford University: Stanford, CA, USA, 2018; Volume 16, pp. 34–39. [CrossRef]
4. Xu, X.L.; Li, J.; Wu, D.; Zhang, X. The intellectual capital efficiency and corporate sustainable growth nexus: Comparison from agriculture, tourism and renewable energy sector. *Environ. Dev. Sustain.* **2021**, *23*, 16038–16056. [CrossRef]
5. van Zelst, S.J.; Buijs, J.C.A.M.; Vázquez-Barreiros, B.; Lama, M.; Mucientes, M. Repairing Alignments of Process Models. *Bus. Inf. Syst. Eng.* **2019**, *62*, 289–304. [CrossRef]
6. Sanchez-Segura, M.-I.; Ruiz-Robles, A.; Medina-Dominguez, F.; Dugarte-Peña, G.-L. Strategic characterization of process assets based on asset quality and business impact. *Ind. Manag. Data Syst.* **2017**, *117*, 1720–1737. [CrossRef]
7. Burton, B.; Scheibenreif, D.; Barnes, H.; Smith, M.; Buytendijk, F.; Bradley, A. Digital Business Gives Rise to the New Economics of Connections. *Gartner* 2015, 1–11, *Retrieved 2020*. Available online: https://www.gartner.com/smarterwithgartner/the-economics-of-connections (accessed on 5 October 2022).
8. Schiuma, G. The managerial foundations of knowledge assets dynamics. *Knowl. Manag. Res. Pract.* **2009**, *7*, 290–299. [CrossRef]
9. Otero, E.; Schwarz, M. Review of the literature on the techniques and methods for measuring Intellectual Capital. *Rev. Científica De La UCSA* **2018**, *5*, 41–60. Available online: http://ucsa.edu.py/yeah/wp-content/uploads/2018/05/7_AR2_Otero-E_Revisi{\protect\edefT1{T5}\let\enc@update\relaxó}n-de-la-literatura-de-las-t{\protect\edefT1{T5}\let\enc@update\relaxé}cnicas-y-m{\protect\edefT1{T5}\let\enc@update\relaxé}todos-de-medici{\protect\edefT1{T5}\let\enc@update\relaxó}n_41-60.pdf (accessed on 4 October 2022). [CrossRef]
10. Tabrizi, B.; Lam, E.; Girard, K.; Irvin, V. Digital Transformation is Not About Technology. *Harv. Bus. Rev.* **2019**, 2–7. Available online: https://hbr.org/2019/03/digital-transformation-is-not-about-technology (accessed on 4 October 2022).
11. Fakir, M.; Baslam, M.; El Ayachi, R. *Business Intelligence: 6th International Conference, CBI 2021Beni Mellal, Morocco, May 27–29, 2021, Proceedings*; Springer International Publishing: Berlin/Heidelberg, Germany, 2021. [CrossRef]
12. Zerbino, P.; Stefanini, A.; Aloini, D. Process Science in Action: A Literature Review on Process Mining in Business Management. *Technol. Forecast. Soc. Chang.* **2021**, *172*, 121021. [CrossRef]
13. Williams, M.; Ariyachandra, T.; Frolick, M. Business Intelligence- Success Through Agile Implementation. *J. Manag. Eng. Integr.* **2017**, *10*, 14–21. Available online: https://www.journalmei.com/_files/ugd/f76c8e_e5ea95afc7e9468cb279f9b89a825fb1.pdf (accessed on 4 October 2022).
14. Janssenswillen, G.; Depaire, B. Towards Confirmatory Process Discovery: Making Assertions About the Underlying System. *Bus. Inf. Syst. Eng.* **2018**, *61*, 713–728. [CrossRef]
15. Viner, D.; Stierle, M.; Matzner, M. A Process Mining Software Comparison. In Proceedings of the Conference: 2nd International Conference on Process Mining (ICPM), Padova, Italy, 4–9 October 2020; pp. 19–222. Available online: https://www.researchgate.net/publication/344657667_A_Process_Mining_Software_Comparison (accessed on 4 October 2022).
16. De Cnudde, S.; Claes, J.; Poels, G. Improving the quality of the Heuristics Miner in ProM 6.2. *Expert Syst. Appl.* **2014**, *41*, 7678–7690. [CrossRef]
17. Johannesson, P.; Perjons, E. *An Introduction to Design Science*; Springer International Publishing Switzerland: Stockholm, Sweden, 2014. [CrossRef]
18. Sanchez-Segura, M.I.; Dugarte-Peña, G.; Amescua, A.; Medina-Dominguez, F. Exploring how the intangible side of an organization impacts its business model. *Kybernetes* **2020**, *50*, 2790–2822. [CrossRef]
19. Chen, J. Enron (Company Profile). 2019. Available online: https://www.investopedia.com/terms/e/enron.asp (accessed on 4 October 2022).
20. Cohen, W.W. Enron Email Dataset. 2015. Available online: http://www.cs.cmu.edu/~{}enron/ (accessed on 4 October 2022).
21. Chabrak, N.; Daidj, N. Enron: Widespread myopia. *Crit. Perspect. Account.* **2007**, *18*, 539–557. [CrossRef]