A method for Ecological Process Optimization based on Compliance Checking

Patrick Lübbecke
Institute for Information Systems
German Research Center for Artificial Intelligence and Saarland University
Saarbrücken, Germany
Patrick.Luebbecke@dfki.de

Abhishek Goswami
Institute for Information Systems
German Research Center for Artificial Intelligence and Saarland University
Saarbrücken, Germany
Abhishek.Goswami@dfki.de

Peter Fettke
Institute for Information Systems
German Research Center for Artificial Intelligence and Saarland University
Saarbrücken, Germany
Peter.Fettke@dfki.de

***Abstract*— In this paper, we introduce a novel approach for optimizing the environmental footprint of Business Processes by using techniques from the area of Compliance Checking along with a set of ecological weakness patterns that were identified during prior research. Through our approach, we are able to identify ecological weaknesses in process models that can be revised to come up with a modified, more eco-efficient process. In our work we employ a set of existing weakness patterns and divide them into clusters or genres based on their semantical meaning. Consequently, we implement unique Compliance Checking methods for each different pattern cluster. The result of this paper represents a Compliance Checking Method Library which offers an exhaustive set of methods to identify many forms of ecological weaknesses in process models.**

Keywords- Sustainability, Ecological Weakness Patterns, Green BPM

I. Introduction

Business Process Compliance Checking (BPCC) is a management principle that helps companies to meet the requirements imposed on them through legislation, internal standards (e.g. ISO 9001), or direct agreements with their business partners [1]. Compliance management is predominantly applied in sectors where disobedience of regulations can cause physical or financial harm to people, such as in healthcare, finance, or insurance. For that reason, BPCC provides means to identify instances of non-compliant behavior in day-to-day business routines and allows for re-design of these activities.

BPCC is part of the Business Process Management (BPM) landscape. BPM implements the idea of improving corporate performance by modifying and optimizing a company's business processes [2], [3]. A challenging field of BPM that researchers have targeted in recent years is Green BPM which aims to improve the ecological performance of a process by effectively and efficiently designing resources-usage in IT-related processes with regard to ecological goals [4]. Sustainability refers to the property of a system to remain balanced, diverse and indefinitely productive [5]. From an ecological standpoint that refers to the idea of using resources wisely in the short term to maintain availability in the long run [6]. To achieve this goal, day-to-day processes must also be aligned with environmental laws, compete with best practice in the field and incorporate ecology-related agreements with business partners. BPCC can be applied in this area to foster process compliance with ecological goals.

Many BPCC approaches use patterns of non-compliant behavior to identify these patterns as sub-graphs in graph-based process models [7]–[9]. While the benefit of BPCC towards process improvement has already been empirically validated [10], BPCC methods have yet to be applied to the ecological area to foster ecological process improvement. Patterns of ecological good practice were proposed in the past, both, domain neutral ([11], [12]) as well as for IT-related processes in the area of public administration (PA) [13]. However, these patterns are currently not used in any BPM tool to support ecological process improvement through BPCC measures.

With this research paper we address this gap. Our goal is to design and implement an extension for the BPM platform ARIS from Software AG that can be used to identify ecological shortcomings and foster improvement of the ecological footprint of a business process. By ecological improvement, we mean reduction of resource consumption or emission of pollutants. We rely on existing ecological weakness patterns from public administration to demonstrate our approach. The advantage of IT-related processes in the PA is that the concepts can easily be transferred to most administrative processes in the private sector as well. However, we will draw some analogies to manufacturing or logistical processes throughout the paper as well.

The research method applied in this paper shall be characterized as design science research (DSR) [14], [15]. By following the process of [15] for creating design science artifacts, we presented a problem identification in this section. In section 2, we analyze related work and point out the shortcomings of existing approaches that makes necessary our artifact. We also outline the objectives of the solution. Then we provide the framework of knowledge preparation in section 3 and implement the solution in section 4. An exemplary application of the artifact will be conducted in section 5, before we conclude our research with a conclusion and discussion on future research opportunities in section 6.

II. Related Work

A. Patterns in (Green) Business Process Management

The concept of patterns was initially introduced by Alexander in the domain of architecture. Alexander described *"problems which occur[s] over and over again in our environment"* along with the core of the solution to that problem as patterns [16]. The concept found its way to software engineering through the work of Gamma et al. [17], who proposed *design patterns* to provide reliable solutions for recurring problems when designing software. Until this day the Design Patterns have made a tremendous impact.

Consequently, patterns were also applied in several areas of Business Process Management (BPM). Initially they were used to describe the logic behind Workflow-Management-Systems as *workflow patterns by* van der Aalst et al. [18], [19]. A workflow represents a systematic chain of cause and effect where any process runs from its start to the end with a certain degree of parallelization, completing small tasks at every step, to complete the final goal of the process. The authors identified patterns for basic control, branching, instantiation, cancelation, and state representation in workflows and eventually expanded their work to identify *data interaction*- and *resource handling patterns* in workflows [20], [21].

More recently, researchers worked on casting knowledge about good practice or compulsory rules to *process patterns* to be used for assessment of compliance of existing business processes with either internal or legal standards [22]. This kind of compliance check is especially important in domains that are liable to stringent regulation, for example in the financial sector [23] or healthcare [24]. These approaches either provide reference process constructs that can be used as a blueprint for reengineering existing processes or they synthesize weaknesses that should be avoided. These anti-patterns can be used to identify instances of poor process constructs in existing models [25].

Green BPM means defining, implementing, executing and improving business processes in corporations with the aim to support environmental objectives [26]. In the environmental domain, patterns are used to reduce the ecological footprint of business operations by optimizing, for example, the process flow. One of the earliest instances where patterns were adopted to that domain was when Nowak and Leymann, borrowing from the workflow patterns of van der Aalst et al., identified 9 generic patterns that lead to ecological benefits in processes, such as *process automation*, *resource change*, or *green compensation* [11]. In later work they expanded their collection of ecological patterns by including application architecture principles and cloud patterns [12], [27]. Lübbecke et al. advanced the work of Nowak et al. by identifying recurring process structures on the workflow-, operational-, and data-layer in processes of the PA domain [28]. The authors proposed a systematization to describe process constructs on these three levels as patterns. In their following work [13], they identified 26 weakness patterns on a very detailed level through a review of process models from the PA that can be used to identify ecologically adverse constructs in processes from that area.

The usefulness of weakness patterns for business process redesign has already been discussed by Becker et al. [10]. They conducted an empirical study where they automatically searched for the occurrences of exemplary weakness patterns in process models from the banking sector. From a corpus of 32 process models the authors searched for 10 different weaknesses through SQL statements on the database of the modeling tool. The weaknesses were identified during expert interviews prior to the study. From 30 expected matches the study yielded in 22 matches or 73 percent to be found.

B. Compliance Checking in Process Models

The detection of recurring patterns in process models provides the methodical foundations to find and remedy non-compliant behavior as sub-graphs in graph-based process models [9]. Regulative rules that companies must comply with can be deducted from different sources (Fig. 1). The most important source are governmental laws and policies, since non-compliance can lead to financial penalties. A prominent example is the European directive for *Registration, Evaluation, Authorisation and Restriction of Chemicals* (REACh), which restricts certain substances from being used during manufacturing. In this case, the manufacturing process must comply by not including one of the regulated substances listed in any of the tasks. In addition, ecological process improvement can be achieved if processes comply with good- or best-practice in the field, which makes these best-practice a second source for compliance rules. For example, a best practice for sending a telefax message in administrative processes would be by using a software-fax instead of hardware-fax due to the excess energy consumption of the latter. Strategic realignment means the pro-active adoption of measures that are neither compulsory by law, nor considered a best practice, but that might be implemented for strategic reasons. The last obvious source for compliance rules is explicit knowledge held by process experts that can be used to make the process comply with. Methods for compliance checking are implemented through graph-based query languages, including *Compliance Rule Graphs* [7], *BPMN-Q* [8], and *GMQL* [9].

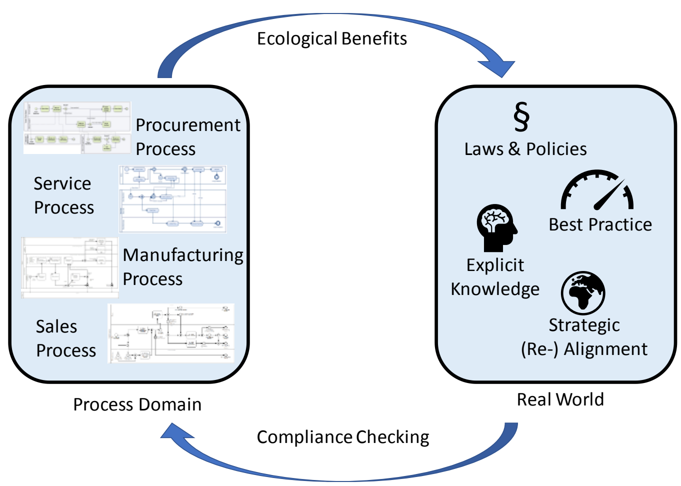

Figure 1. The Compliance Checking Optimization Cycle

Compliance Rule Graphs are a visualization of domain-specific compliance rules in the form of *antecedence-* and *consequence patterns*. The formal semantics of the compliance rules require a transformation into first-order predicate logic formulas [7].

BPMN-Q allows the query of process models with visual patterns that are based on the BPMN language and are translated to Petri nets for processing. The patterns are transformed into temporal logic formulas and further processed by the model checker on the Petri nets [8]. Although the approach is described as not being limited to BPMN models, it requires the pre-processing of models to the required representation as Petri net before.

The *Generic Model Query Language* (GMQL) is a graph pattern-based business process compliance checking approach. Just like BPMN-Q, GMQL is supposed to be applicable to conceptual models defined in any modeling language. Until so far, it only has been implemented proprietarily for PICTURE, a non-standard and domain-specific semantic modeling language for the domain of public administration [30]. The language uses so-called process building blocks to describe process activities along with a comprehensive annotation of process relevant information to the building blocks [9].

The limitations of the current approaches can be summarized as follows. First, none of the approaches mentioned above are implemented in mainstream BPM tools and are currently only available as a proprietary development from research facilities. That comes with all the implications in terms of software quality and support, which are important factors for corporate users. Second, the BPMN-Q approach only works with process models in the BPMN notation, which makes the application impossible for companies that use, for example, EPC or UML as their preferred modeling language. But the most important reason why existing approaches need to be reconsidered are the complexities that come with the application of the domain for ecological process improvement. BPMN-Q is only able to identify similarities based on graph structure. Behavioral semantics and operation semantics must be considered as well to identify ecological shortcomings in process models. Although GMQL has proven to be able to meet these requirements when applied on the strongly annotated modeling language PICTURE, most process models both, BPMN and EPC turn out to be poorly annotated. Hence, the approach that we propose will consider this issue and address it accordingly.

III. RESOURCE CONSUMPTION DURING PROCESS EXECUTION

A. *Environmental effects of various business sectors*

The analysis and optimization of Business Processes towards ecological goals is subject of the research discipline Green BPM [26]. The discipline is focusing on corporate processes to foster environmental optimization, e. g. the reduction of greenhouse gas. A process model describes the tasks, mostly performed by a human, that are necessary to achieve a certain process goal. Although many tasks are accomplished by humans, the environmentally-adverse effects that come with these tasks normally crop up due to the use of various technical devices. Depending on the type of process and the area where the process is being carried out, the type of technical device or transformation process can be very diverse, much like the consumption of natural resources and the resulting impact on environmental objectives that go along with it (Table 1).

Most processes use any kind of resources, energy, and ancillaries to achieve the specific objective set to the process. Industrial activities, and manufacturing processes in particular, have a large impact on the environment and are responsible for 33% of the primary energy use and CO_2 emission globally [31]. Manufacturing is supported by machines to perform mostly technical tasks such as lathing, milling, or many other robotic manufacturing activities. These kinds of machines are powered directly by electric energy or air pressure, which needs electric energy to be provided. Most importantly, manufacturing processes transform raw, auxiliary or operational materials into original products and result in waste as by-product. Generating electric energy causes emission of various gases whose role in causing global warming is widely accepted [32].

In the logistics sector, the engaged machinery might be different (e. g. trucks, fork lifts, automated high rack storage systems). However, the resulting emission remains very similar compared to manufacturing, as the primary energy carriers used are electric energy for warehousing and gasoline for transport of goods.

Product development is conducted computer-aided (CAx), sometimes supported by simulation models (e.g. Finite Element Analysis) or rapid prototyping techniques such as 3D-printing. Primary resource consumption in most of these cases is electric energy.

The resource consumption in administrative processes is largely homogenous. Although many different ICT devices are used in administrative processes, the resource consumption is largely limited to electric energy, resulting in the emission of different greenhouse gases, depending on the energy-mix. Instead of raw, auxiliary or operation materials, the primary resource that is being consumed in this area with the largest impact on environment is paper, which leads to deforestation. Notably, administrative processes are strongly repetitive and are conducted very similar in the public- and private sector. If we compare the administrative tasks necessary to process an order in a private company, we can find many similarities to the public administration for, let's say, requesting a birth certificate, as the example shows:

Private Company: Create new customer account in CRM, send correspondence (fax, (e-)mail), print delivery notes and invoices, arrange shipment, monitor payment.

Public Administration: Fill out request, send correspondence (fax, (e-)mail), print invoices/certificates, arrange shipment, monitor payment.

Hence, our ideas can be easily transferred from public to private sector as well as between companies with only minor adjustments. For that reason, we chose office processes in the public administration to implement our approach.

TABLE I. MOST IMPORTANT RESOURCE CONSUMPTIONS ACROSS PROCESS TYPES

Process type	Tools	Resource type → Environmental Impact
Manufacturing	– Lathe – Milling machine – Kiln – Robots	– Electric energy → Emission of CO_2, CH_4 (methane), SO_2 (sulfur dioxide) – Raw-, auxiliary-, operational materials → Depletion of ltd. resources, waste – Gasoline → Emission of CO_2, NO_x (nitrogen oxide) – Natural gas → Emission of CO_2
Logistics	– Truck – Fork lift – High rack storage – Air Cond.	– Electric energy → CO_2, CH_4, SO_2 – Gasoline → CO_2, NO_x
Product Development	– 3D printing – Computing, Simulation – ICT	– Filament → Polyactic Acid (renewable) – Electric energy → CO_2, CH_4, SO_2
Office Administration	– ICT	– Electric energy → CO_2, CH_4, SO_2 – Paper → Deforestation – „Storage room" → Gasoline + Elec. Energy

B. Ecological Weakness Patterns in Public Administration

Process models provide much information about the set of activities a company pursues to accomplish an objective. This information is available explicit through label names of nodes (events or activities), implicit through their structure (edges and conjunctions), or a combination of both. In many cases, models are enriched by context information like organizational units, data objects, and – most importantly – technical devices that are used for the task. That kind of information can be used to evaluate a process model in terms of its ecological impact. For example, a document that is printed but not further used anymore (waste of energy) can be identified by the activities' label name ("print") plus the relationship to an Information object that represents the printer in the case of the event-driven process chain (EPC).

Ecological weaknesses exist when a given as-is process deviates from a hypothetical best-practice or even good-practice, causing an additional resource consumption or emission of pollutants compared to the best-practice. Ecological weaknesses mostly depend on a process type and a sector where they are applied. In logistics, for example, such a weakness could exist when canned food is unnecessarily stored in frozen cargo areas or when frozen cargo is stored at the wrong temperature, causing it to perish. In administrative office processes, on the other hand, a weakness could exist in sending invoices as physical document instead of an e-mail, resulting in a vast amount of energy for the printing process. The two examples show that various attributes must be considered, depending on process type and business sector. In addition, different sources for data analysis must be included. For the food scenario, real-time data is necessary to evaluate the occurrence of ecological weaknesses. This real-time data can be provided, for example, by monitoring the product through Digital Product Memories [33] or by monitoring the storage facilities [34].

For administrative processes, on the other hand, most information or identifying ecological weaknesses can be extracted directly from static process models. These models provide information on the process flow (e.g. traversed paths, cycles, sequence of tasks), on the tools that are used to achieve the process goal and allow the identification of relevant drivers of resource consumption. Lübbecke et al. proposed a framework where they point out the importance of the *control-flow*, *operational*-, and *data-perspective* of process models for identifying ecological weaknesses [28].

Control-flow: This perspective addresses errors in the control-flow during process execution that result in ecological shortcomings. Control-flow errors might originate from a wrong order of tasks or quality problems of any kind which are responsible for recurring iterations of cyclic process constructs. Both error types cause unnecessary resource consumption. Many processes in PA are driven by legislation and constrained by law. Therefore, the processes executed in the administration of municipalities are extensively regulated and a require stringent documentations.

Operational perspective: In the operational perspective, an assessment of the tools involved to accomplish the process goal is conducted. In EPC (*information object*) or BPMN (*annotation*), such tools are normally modeled explicitly. Table 1 gives an impression on how broad the term "tool" can be understood for certain process types. An ecological weakness in the PA could be the use of a standalone fax instead of a software fax, which consumes an enormous amount of power even while in stand-by.

Data perspective: The data perspective holds two different kinds of information that can be used for identifying ecological weaknesses. If resource-related information about individual aspects of processes or tools is available, this information can be used to balance one tool or one technique against another. For example, if a certain transporter is used to carry out a logistical task, the fuel consumption of the transporter can be compared to other transporters in the inventory that might cause fewer fuel consumption. The same principles can be applied to ICT devices like printers. However, to be able to capitalize on this information, relevant aspects have to be included in the process model from early beginning to allow automatic information extraction from the models [35], [36]. In addition to that, information is an important asset in administrative processes and particularly in PA. Most of the resource consumption in administrative processes evolves around creating, storing, processing, and transferring information. With the importance of the information flow, the information handling can be analyzed in terms of ecological shortcomings. Examples would be: information is created, but never used thereafter or Information is sent multiple times.

Lübbecke et al. have identified a set of 26 ecological weaknesses during manual process analysis that relate to these three perspectives for the domain of PA (Table 2) [13]. To demonstrate our BPCC approach, we will rely on these weakness patterns.

TABLE II. ECOLOGICAL PROCESS WEAKNESSES IN PUBLIC ADMINISTRATION [13]

Nr., Description of the pattern	Improvement Potential
1: A form is printed and afterwards the application is checked for completeness.	The check for completeness should be performed prior to the print of documents to avoid unnecessary prints in case the application is missing some other documents.
2: Documents are printed and sent via mail, which results in consumption of a considerable amount of energy for the print process.	Documents should be sent via e-mail or provided through download instead to avoid unnecessary energy consumption.
3: Payment slip for registration fee is handed to the customer personally and mailed to his address as well.	The mailing of the slip should be skipped if the customer already received it personally.
4: Application received via third-party Information System (IS) is printed and then entered to the in-house IS.	Integration of the two Information Systems could spare the printing of the application form.
5: Physical Document is printed but never used afterwards.	The form should not be printed if not used afterwards.
6: Document data is entered into IT and subsequently examined or formally reviewed.	The sequence of both processes should be altered to make sure the data input happens only once, thus avoiding unnecessary energy consumption for the input process.
7: Information is sent via email in multiple steps instead of using multiple recipients, CC or BCC.	Every process of sending only a single e-mail causes unnecessary energy consumption. E-Mails should be pooled.
8: A document is forwarded twice. Once as e-mail and once as print without the print version being further processed in the future.	The creation and forwarding of the printed document is unnecessary and can be avoided to save energy.
9: A text document is scanned. Later, the information of the document is entered manually to an Information System (e.g. word processor).	The scan process is either not necessary or the information should be transferred to the IS after the scan automatically to save energy.
10: Redundant document storage in different media.	If a document is stored in different media (paper, HDD, SSD), the creation (print) and possibly the storage consumes energy (server). Redundant document storage should be avoided if possible.
11: A document is forwarded via mail. After the receipt, the receiver scans the document.	Document should be forwarded by e-mail to avoid energy consumption during transport, print and scan.
12: The same document is sent through different communication channels to various recipients at various times instead of collectively sending the document by e-mail.	The forwarding through different communication channels (mail, in-house mail) consumes energy for transport and prints of the document and should be replaced by a single e-mail process.
13: A created document is never used again after its creation in any process.	The process of document creation could be avoided. Leads to saving of electric energy.
14: The process contains numerous print tasks that exceed a given threshold.	Document printing with laser or photocopier consumes a high amount of energy due to the thermal processing.
15: Media disruption takes place from a physical to a digital representation of information or vice versa.	Energy for the transformation of the information's representation could be saved by keeping a single representation (e.g. print).
16: Document is being archived as a physical and digital representation.	Energy either for the print or the storage devices could be saved.
17: Document is printed and scanned later.	Energy for the transformation of the information's representation could be saved.
18: A clerk prints a document to review it. The changes are then made in the digital representation in the Information System.	Energy for the unnecessary print.
19: Data was burned on a CD. The CD is forwarded to a receiver who prints the documents on the CD.	The energy intensive process of burning a CD is superfluous and can be replaced by e-mail.
20: Digitally available information is exchanged personally by using a print process.	Digitally available information should be exchanged in a digital representation to avoid energy consumption for the print of a document.
21: Information is exchanged via FAX instead of using e-mail.	As with a photocopier, a FAX consumes a huge amount of energy even during standby. E-mail should be used to send documents instead.
22: Information between software applications is exchanged via e-mail.	The software applications should comprise proper interfaces to support a seamless data exchange to avoid the energy consumption resulting from e-mail processes.
23: Data is requested and exchanged via in-house mail instead of e-mail.	Document transport via in-house mail includes a thermal print of the document and consumes high amounts of electrical energy.
24: Data manually entered into an Information System.	The manual input of information to an Information System causes unnecessary energy consumption for the time the input process takes. Again, proper interfaces between the applications could solve this issue.
25: Data manually entered into multiple Information Systems.	Same as 23 but multiple energy consumption.
26: The exchange of information or documents exceeds a reasonable threshold.	Excessive information exchange always comes along with a huge consumption for the resource that carries the information from Point A to B (energy, gas, etc). Therefore, the number of information exchanges should be limited to a reasonable number.

IV. CONCEPTUAL CONSIDERATIONS FOR THE BPCC

The identified Ecological Weakness Patterns by nature represent only a subset of the entire number of weaknesses that might be latently present in process models. New patterns might occur if new technologies emerge and become best practice, if new policies are established, or if other strategic directions are taken. For example, while natural gas might be the resource of choice today for kilning in the ceramic industry, new taxes on that resource or even the rise of electric energy generated from renewable sources might change the rules in the future, making natural gas an effective weakness compared to electricity. In these instances, emerging weakness patterns would require the implementation of a new algorithm in the BPM software to be able to discover instances of these weaknesses in models, making the process of maintaining the pattern database very time-consuming. For the existing weakness patterns from Table 2, 26 different algorithms would have to be implemented. For that reason, it is necessary to abstract from the level of *individual weakness patterns* and form *clusters of weakness patterns with similar characteristics* instead that can be recognized using one common algorithm, with different parameters. This will strongly reduce the effort necessary for including future patterns. Figure 2 shows a graphical representation for two sets of two patterns that can be recognized focusing on the same characteristics.

The common characteristic of the first two patterns of Figure 2 is that a task creates an object (Form), which is later used in the process in an undesired way. In pattern #1, a task called 'Print Form' creates a physical printed copy of the 'Form' which in later steps is checked by the event 'Check Form for Completeness'. The media transformation of the form (from digital to physical paper) before any final checking procedure is carried out creates steps resulting in excessive use of energy and paper, as the form must be printed again in case of errors discovered during the check. The same principle applies in pattern #17. A digital representation of information is printed, and later scanned. Patterns #7 and #12 follow similar characteristics as well. In

#7, the same e-mail content is sent to three different recipients. In #12, the same content is sent to three recipients, but this time through three different communication channels. Both are ecological weaknesses.

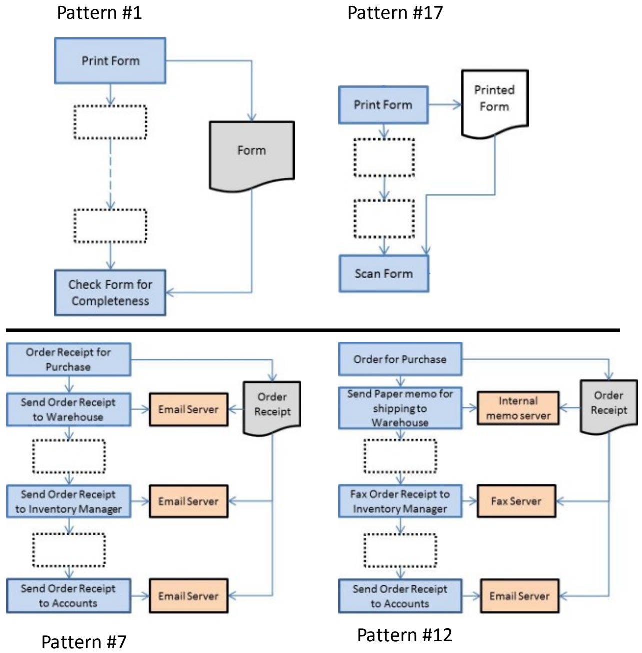

Figure 2. Weakness Patterns with similar characteristics

With that in mind, prior to the actual implementation, the set of weakness patterns from Table 2 can be further analyzed to identify similar pattern types that can be implemented using the same algorithms. In many cases, this proceeding allows to include some variability for the user. For example, the patterns #1 and #17 from Figure 2 can be identified, among other characteristics, by detecting tasks that contain the terms "Print" and is followed by a task that contains "Check" respectively "Print" and "Scan". In this or similar cases, the user of the BPM tool can be provided with a drop-down list of the possible combinations or with two text fields to enter own combinations.

Therefore, after analyzing the existing weakness patterns, we organized patterns that can be discovered using the same or similar techniques for pattern recognition in own clusters. The techniques that need to be applied vary from label analysis with NLP-methods to structural analysis of model elements (sequence analysis, duplicate detection) as well as semantical analysis. The latter is relevant to find out about the meaning of a process construct when no appropriate information such as distinct label names is provided. For the moment we could identify 5 weakness clusters that will be described in the following (Figure 3). Just like the weakness patterns, it must be noted that the weakness clusters are not exhaustive as well. With future research leading us to new kinds of weakness patterns, it is possible that new cluster(s) must be defined as well when a new pattern cannot fit into any of the existing clusters. New clusters could also be necessary when the concept is transferred to a different domain or process type (Table 1). In this paper, we focus on administrative processes. Logistical or manufacturing processes come with other premises in terms of what defines ecological sustainability, thus further weakness cluster(s) might be necessary.

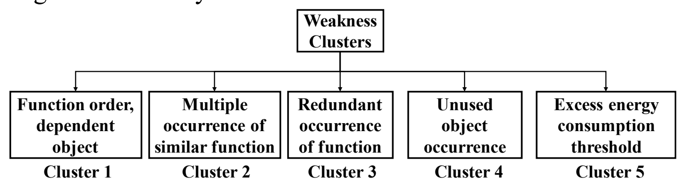

Figure 3. Weakness Clusters

Cluster 1: Function order, dependent object

A weakness pattern is found when a Function-Object-Function dependency loop is created, where the first function creates/modifies/accesses an Object which in turn is modified/accessed by a second function. The actual weakness lies in the sequence of the two functions namely when latter function should precede the first function to save energy or redundancy. Example of weaknesses: Checking for errors/completion after printing or permanently storing documents. Media exchange before final modifications, etc.

Cluster 2: Multiple occurrence of similar function

Refers to weakness patterns where the same or similar function is triggered at various time intervals on the same un-modified object thereby consuming more energy as well as resources. Instead of multiple occurrences of the function, single occurrence with multiple parameters can solve the problem. Example of weaknesses: Multiple use of the email server for sending unmodified data at various time intervals instead of using a CC can create unnecessary energy consumptions.

Cluster 3: Redundant occurrence of function

Refers to weakness patterns where a "two function loop" creates duplicate instances of information in same or different media. The second function accesses the object created by the first function and creates a second object which is a duplicate of the first object in the same or a different media. Output of a function being limited to a single existence saves storage, energy and physical resources. Example of weaknesses: Printing paper copy of document for physical file storage after having an electronic copy already, mailing paper copy of document after sending it by email before, etc.

Cluster 4: Unused object occurrence

Refers to weakness patterns where a Function-Object tree-structure is formed. This signifies that the object created or modified by the Function is never used or accessed in the workflow of the same or any other process after creation/modification. So, the non-existence of the function or the object itself saves energy and resources. Example of weaknesses: Document is created but not used. Termination or Deregistration form is printed but not used, Information is stored in various media (paper, HDD etc.) but used from only one keeping the others unused, etc.

Cluster 5: Excess energy consumption threshold

This weakness pattern is slightly different from the other patterns in other clusters. A weakness here is defined based on comparison of energy consumptions if a function can be carried out in multiple ways or if the function uses energy beyond a certain threshold. A weakness pattern is found when a function is spotted using more energy whereas it can be replaced by a function consuming less energy that carries out the same objective. Example of weaknesses: Using Fax for information exchange uses more energy than E-mail etc. Every information exchange takes up energy, which is why the number of exchanges that take the energy consumption beyond a threshold can also be a weakness.

The final organization of the weakness patterns from Table 2 into clusters looks as follows:

TABLE III. PATTERN ALLOCATION

Cluster ID	Weakness Pattern IDs
Cluster 1: Function order, dependent object	1,6,11,17,19
Cluster 2: Multiple occurrence of similar function	7,12,14,26
Cluster 3: Redundant occurrence of function	2,3,8,9,18,20
Cluster 4: Unused object occurrence	5,10,13,15,16
Cluster 5: Excess energy consumption threshold	21,22,23,24,25

V. IMPLEMENTATION OF THE BPCC METHODS

A. The development environment

To implement the BPCC methods for finding weakness patterns we first need a BPM tool to define the process workflow.

In our case we used ARIS 10[1] from Software AG with real world process models from the domain of public administration. ARIS is a mainstream BPM platform for process modeling, analysis, and execution that offers an API called ARIS Script. With ARIS Script, individual program code can be implemented to modify or, in our case, analyze the structure of a BPMN or EPC process model. We used the API to generate the necessary algorithms for the pattern detection along with the process models in EPC notation.

One of the main advantages of modeling a process-flow in ARIS is the ease of access. Every node in ARIS (both Objects and Events have a set of attributes like the name, type of object and its own GUID (Global Unique Identifier) which can be used as parameters throughout our BPCC logic.

ARIS has built-in classes to manipulate and modify user-designed process models. The classes consist of several methods that can be called just following the concepts of object-oriented programming. The methods can be parameter-free or can be called with parameter based on the desired result from the method. Two very important classes are:

- ObjDef: Class and methods defined based on definition of the object.
- ObjOcc: Class and methods defined based on occurrence of the object.

[1] See www. https://www.softwareag.com/corporate/ products/aris_alfabet/bpa/aris_connect/default

The logic behind every BPCC method revolves broadly around these two classes. With the methods in these two classes we can retrieve a list of all objects occurring, their labelled names, their types, their GUIDs. We can also retrieve information about how two consecutive nodes are connected to each other, with information of edges going out or coming in to a particular node. This serves our purpose to set up a basis to track a particular node as well as traverse through all nodes in a process model.

B. Implementing BPCC for Cluster1: Function Order, Dependent Object

Recapping the definition of Cluster 1, a weakness pattern is recognized if we have a function-object-function loop where the order of the functions creates weakness failing ecological standards.

We design a pseudocode for this implementation and then have a walkthrough of the logic.

- Start.
- Get the first and second function.
- Get the object type (mostly ‘information carrier’ as default object parameter).
- Check whether first function is followed by second function in direct or indirect succession.
- Check whether there is a common node accessed by both functions sequentially.
- Check whether the common node is an object of searched type (Default information carrier).
- Report weakness pattern.
- End

Walkthrough of the Pseudo-code

Identify the function nodes: Getting the first and second function refers to identifying the label names of functions in form of substrings. The modelling of a process flow may vary as per the designer and that includes the naming styles of function labels. To make the method more robust we use lexical analysis to identify the string labels. For instance, a function node that deals with any sort of printing can be named in several ways like ‘Print form’, ’Printing form’, ‘Send Printed document’ etc. but the substring ‘print’ remains a constant. Hence, we match substrings and filter out nodes that can be semantic equivalence of functions we are searching for. For instance, to find a weakness pattern where something is printed and later on scanned, we would search for the occurrence of a process with the term print, followed by a process with the term scan in their label names. The implementation allows us to search for several term-sets at the same time, thus identify several weakness patterns (e.g. print – scan, print – check, mail – mail).

Check for connection between function nodes: Once the potential list of first and second function is recovered we check for a direct or indirect connection from the former to the latter. For this we treat the process flow as a graph and traverse every node by outgoing edges. The list of nodes in the process model as output by ARIS is not ordered by the basis of workflow, so we create an adjacency matrix of our finite graph. The adjacency matrix indicates whether two nodes are directly connected to each other or not. Looping over our starting node (first function), we can traverse to the next directly connected node and so on until either the second function node is reached (connection between nodes found to exist) or all nodes have been traversed and end of graph has been reached but second function node has not been reached (connection between nodes do not exist).

Checking for common node of dependency: With the connection between the first and a second function established, the existence of a weakness pattern can be assumed if an object of common access is noticed between the functions, e. g. through printing, a new information carrier (physical document) was created. If a node is created/modified by the first function and later accessed/modified by a second function, the pattern can be found based on the type of edge connecting the three objects. As for the pair of functions, if there is an outgoing edge from the first function to a node and an incoming edge to the second function from the same node, it fits the profile of a weakness. The weakness pattern is confirmed if the common node is of default type 'Information Carrier' or the type specified initially by user.

Reporting Weakness: At the end of the detection process, the user is provided with a detailed compliance report. The detected weakness patterns are listed so that a process expert can begin to revise the process.

Important to note: Input of function labels and type of common object remain at the sole discretion of the user. Any combination that fits the logic of a weakness as defined in this cluster is reported, irrespective of the fact whether it is a weakness or not. To refrain from getting false positive results, function labels and object types should be carefully passed as parameters.

VI. Examplary Application of the BPCC Method

We did not conduct an empirical evaluation to test for accuracy and practical applicability of our implementation so far. In the following we demonstrate the application of our BPCC method on a process that is based on a real-world proceeding in the German public administration for request of residence permit by foreign students. We conclude this section by laying out our formal evaluation scenario that we plan as the next steps of our research.

The process from Figure 4 shows the proceeding necessary for foreign students to request a residence permission at the residents' registration office (RRO). The process features an exception from how resident permits are normally processed. When students from abroad request such a permission, an exchange of various kinds of information takes place with the Visa department at the Aliens Registration Authority (ARA) and the Tax Authority (TA). We decided to use this process because intensive information exchange is normally a source for ecological flaws in the PA. The process was modeled in EPC notation using the ARIS Designer. After the modeling of the process, the process expert can run one of the four scripts to find patterns of cluster 1-4. The scripts are parametrized with the proper parameters to discover the individual weakness patterns. Currently, each script needs to be run one after another with the respective parameters. In the next iteration of the implementation, the presence of each weakness pattern will be checked in one consecutive run.

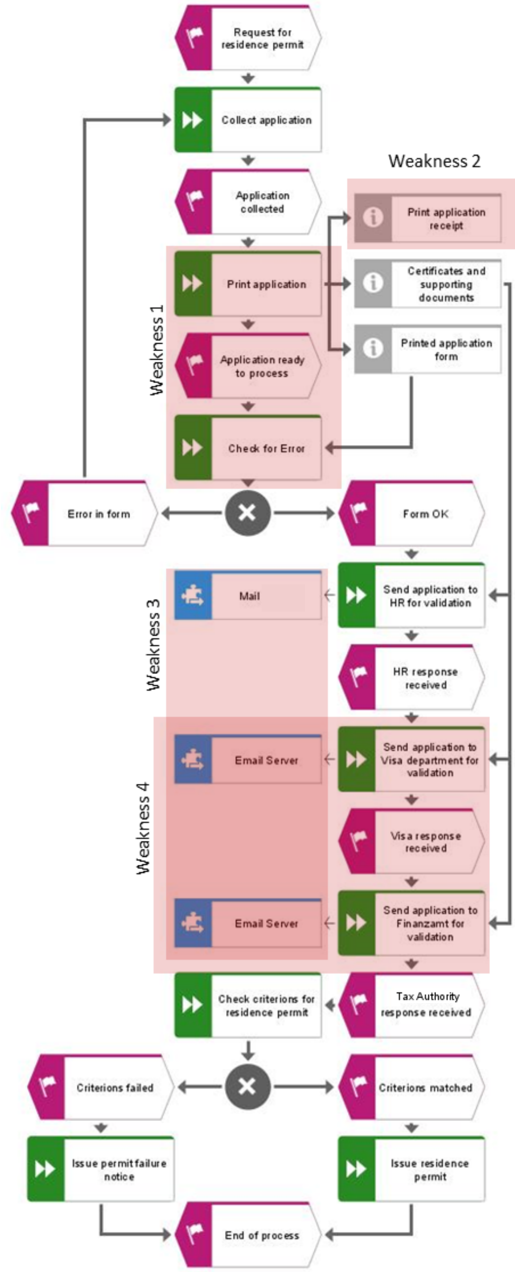

Figure 4. Demo process for requesting a residence permission

Weakness 1 (Pattern #1 from Table 2): After a foreign university student applies for a residence permit at the RRO by submitting an online registration form, the case manager at the RRO prints the form along with certificates and other documents to further process the application. In case of missing or wrong information, a request for providing the correct information is sent to the applicant. **Detection approach:** This flaw was detected by 1) a keyword search for the combination of a process containing a variation of the word “print[ed/ing]” followed by “check[ed/ing]” and 2) a unique document (*Printed application form*) that was created during the task with the word “print” in its label and which serves as an ingoing object to the task with the word “check” in its label. With that detection strategy, similar patterns can be recognized with the same algorithm: *a document is printed and then scanned, a document is printed and then sent by mail etc.*

Weakness 2 (#5): At the beginning of the process, three different documents were printed of which two are further used in the process. However, the *application receipt* is never used during the remainder of the process. This could be a hint for an ecological issue in the form of unnecessary resources consumed. This receipt may be handed directly to the applicant without explicitly stating it in the process. Nevertheless, such a process construct should be subject to further evaluation by a process expert, thus the detection makes sense. **Detection approach:** Whenever a task creates an object (outgoing edges as indicator), a check is performed if these newly created objects have outgoing edges on their own. In that case, they are input to another task. With that detection strategy, similar patterns can be recognized with the same algorithm: *document is being archived physically and digitally, while only one variant is further used.*

Weakness 3 (#12): If the information is complete, copies of the digital documents are forwarded to the Human Resources department of the University by mail to validate whether the student is enrolled for a course or not. As soon as the enrollment is confirmed, another copy is being sent to the Visa department of the ARA by e-mail. The Visa department semiannually evaluates, if the student is still active in his/her mayor. A third copy is sent to the tax authority by e-mail for creating a tax ID for the student in case an employment is pursued. In this pattern, the same documents are sent to different recipients using multiple communication channels instead of collectively using the most resource efficient way (e-mail, software fax) **Detection Approach:** While parsing the process model, the verb-noun combination “*send application*” is recognized along with an information carrier being generated “application”. In the remainder of the process, “send application” is discovered two more times, this time with two different objects being used “Email Server”. Through the identical verb-noun combination it can be recognized, that the supposedly same document is processed here, but through different communication channels.

Weakness 4 (#7): The documents to the ARA and tax authority are sent as multiple e-mails instead of using cc in a single e-mail. This causes an additional energy consumption which is admittedly low at first glance, but that adds up over the year for highly repetitive tasks. **Detection Approach:** As the patterns #7 and #12 which were recognized as weaknesses 3 and 4 belong to the same cluster, the same algorithm could be used to recognize this pattern. Again, a verb-noun combination was recognized that was used to identify similar occurrences of that combination in the remainder of the process. This time, two similar application systems were identified that support the two tasks (“Email Server”).

After presenting the application of the BPCC in this example, we plan to conduct an empirical evaluation during future research. However, this evaluation will come with some limitations. Not every real-world process includes one or more of the ecological weakness patterns. The number of available processes from the public administration would have to be accordingly higher to be able to apply statistical accuracy measures. For that reason, we decided to evaluate in a more controllable scenario. To have available a reasonable number of process models to test our BPCC against, we will have 3-4 researchers that are familiar with process modeling in the domain of PA compile up to 10 process models. Each model is supposed to contain up to 5 ecological weaknesses from Table 2. By having the processes designed by neutral experts, we accomplish two benefits. First, we avoid confirmatory bias of the software-designer, that would design the process to fit to his/her own BPCC script. Second, by having more than one process designer compile the models we induce variety in the process models’ terminology and structure which reflects reality.

VII. Discussion and Conclusion

We hence conclude our work with a few points of recap and a scope for further extension. We discussed the necessity of having a robust approach for Compliance Checking to find weaknesses in business processes which can be revised to attain better performance towards ecological goals. We analyzed existing weakness patterns and divided them into clusters based on their semantic similarities. We defined the clusters in a manner that helps to identify the unique parameters in a weakness pattern and thus allocate patterns to respective clusters. We devised a Compliance Checking Method for each cluster that fits all the weakness patterns in a particular cluster with the only variables being the parameters inherent to a weakness pattern. Finally, we presented a logical walkthrough of the methods taking the BPCC for Cluster 1 to be an example in our case.

What we have implemented is an extension for a BPM tool that takes in a process model and pre-defined or user-supplied parameters to define a pattern and that searches for weakness pattern matches. The tool takes into account several parameter-sets, thus, the BPCC can detect all weakness patterns in the same cluster.

However, our approach comes with some limitations as well. We learned during our activities, that process model quality much too often turns out to be very poor in terms of annotation with context information. To be able to evaluate certain aspects (e.g. which tools were used?), a minimum of information must be provided either through explicit information objects (unique objects for the tool) or in the label name. If that kind of information is not available, the weakness simply cannot be found. However, this aspect of model quality is subject to a plethora of research, including in the context of sustainability [36].

For the time being we are reliant on ARIS to model our processes and optimize them. Hence, a general prerequisite of familiarity with ARIS must be available to apply the approach. Furthermore, we bind ourselves to established weakness patterns and synthetic process models. An evaluation of the performance of our approach on real-life processes is not available yet and shall be subject to future work. However, judging by the successful results of the implementation on existing process models, a positive correlation can be drawn.

As a final word for future scope, we discussed that weakness patterns and the idea of clusters is not an exhaustive one. Further research might bring up more weakness patterns as more process types and business sector outside the PA will be analyzed. With newer weakness patterns we can either allocate them to existing clusters or define new clusters to fit the level of abstraction. The clusters that we defined in our work are guidelines and hence there lies chances of overlaps in the types of weakness. We keep it open to the discretion of the tester as weakness patterns apart from being technical are highly contextual and can have various representations under circumstances of a particular process model, though the inherent logic remains the same.

VIII. Acknowledgment

The research described in this paper was supported by a grant from the German Ministry of Education and Research (BMBF), project name GreenFlow, support code 01IS12050.

IX. References

[1] M. El Kharbili, A. K. A. de Medeiros, S. Stein, and W. M. P. van der Aalst, "Business Process Compliance Checking: Current State and Future Challenges," *MobIS*, vol. 141, pp. 107–113, 2008.

[2] W. M. P. van der Aalst, "Business Process Management: A Comprehensive Survey," *ISRN Softw. Eng.*, vol. 2013, pp. 1–37, 2013.

[3] M. Dumas, M. La Rosa, J. Mendling, and H. A. Reijers, *Fundamentals of Business Process Management*. Heidelberg: Springer, 2013.

[4] J. vom Brocke, S. Seidel, and J. Recker, "Green Business Process Management," Berlin: Springer, 2012, pp. 3–13.

[5] World Commission on Environment and Development, "Report of the World Commission on Environment and Development: Our Common Future (The Brundtland Report)," *Med. Confl. Surviv.*, vol. 4, no. 1, p. 300, 1987.

[6] H. C. von Carlowitz, *Sylvicultura Oeconomica*. 1713.

[7] L. T. Ly, S. Rinderle-Ma, and P. Dadam, "Design and Verification of Instantiable Compliance Rule Graphs in Process-Aware Information Systems," 2010.

[8] A. Awad, "BPMN-Q: A Language to Query Business Processes," *Proc. Conf. Enterp. Model. Inf. Syst. Archit. 2007*, pp. 115–128, 2007.

[9] P. Delfmann, M. Steinhorst, H. A. Dietrich, and J. Becker, "The generic model query language GMQL - Conceptual specification, implementation, and runtime evaluation," *Inf. Syst.*, vol. 47, pp. 129–177, 2015.

[10] J. Becker, P. Bergener, D. Breuker, and M. Räckers, "An empirical assessment of the usefulness of weakness patterns in business process redesign," *ECIS 2012 Proceedings*.

[11] A. Nowak, F. Leymann, D. Schleicher, D. Schumm, and S. Wagner, "Green business process patterns," in *Proceedings of the 18th Conference on Pattern Languages of Programs - PLoP '11*, 2011, pp. 1–10.

[12] A. Nowak and F. Leymann, "Green business process patterns - Part II," *6th International Conference on Service-Oriented Computing and Applications, SOCA 2013*, pp. 168–173.

[13] P. Lübbecke, P. Fettke, and P. Loos, "Sustainability patterns for the improvement of IT-related business processes with regard to ecological goals," in *LNBIP* vol. 281, M. Dumas and M. Fantinato, Eds. Cham: Springer International Publishing, 2017, pp. 428–439.

[14] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design Science - Hevner," *Des. Sci. Inf. Syst. Res. Author MIS Q.*, vol. 28, no. 1, pp. 75–105, 2004.

[15] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A Design Science Research Methodology for Information Systems Research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, pp. 45–77, 2007.

[16] C. Alexander, "The Timeless Way of Building," *Oxford University Press*, 1979.

[17] E. Gamma, R. Helm, R. Johnson, and J. Vlissides, "Design Patterns: Elements of Reusable Software," *Addison-Wesley Prof. Comput. Ser.*, p. 395, 1995.

[18] W. M. P. van der Aalst, A. P. Barros, A. H. M. ter Hofstede, and B. Kiepuszewski, "Advanced Workflow Patterns," *CoopIS 2000 Proceedings, Eilat, Isr. Sept. 6-8, 2000.* vol. 1901, pp. 18–29, 2000.

[19] W. M. P. Van der Aalst, A. H. M. Ter Hofstede, B. Kiepuszewski, and A. P. Barros, "Workflow patterns," *Distributed and Parallel Databases*, vol. 14, no. 1. pp. 5–51, 2003.

[20] N. Russell, A. H. M. Hofstede, D. Edmond, and W. M. P. Van Der Aalst, "Workflow data patterns," *Business*, vol. 66, no. FIT--TR--2004--01, p. 2004-01, 2004.

[21] N. Russell, A. H. M. Hofstede, and D. Edmond, "Workflow resource patterns," *Business*, vol. 3520, no. 5446, pp. 13–17, 2004.

[22] A. Ghose and G. Koliadis, "Auditing Business Process Compliance," *Proc. Int. Conf. Serv. Comput.*, pp. 169–180, 2007.

[23] J. Becker, M. Räckers, and A. Winkelmann, "Pattern-Based Semi-Automatic Analysis of Weaknesses in Semantic Business Process Models in the Banking Sector," *ECIS Proc.* 2010, p. 156, 2010.

[24] C. Stephenson and W. Bandara, "Enhancing Best Practice in Public Health: Using Process Patterns for Business Process Management," *ECIS Proceedings* 2007, pp. 2123–2134, 2007.

[25] R. Fourati, N. Bouassida, and H. Ben Abdallah, "A Metric-Based Approach for Anti-pattern Detection in UML Designs," in *Computer and Information Science 2011*, 2011, pp. 17–33.

[26] S. Seidel and J. Recker, "Implementing green business processes: the importance of functional affordances of information systems," *23rd Australas. Conf. Inf. Syst. 2012*, pp. 1–10, 2012.

[27] A. Nowak and F. Leymann, "An Overview on Implicit Green Business Process Patterns," Stuttgart, 2013.

[28] P. Lübbecke, P. Fettke, and P. Loos, "Towards ecological workflow patterns as an instrument to optimize business processes with respect to ecological goals," in *Hawaii International Conference on System Sciences*, 2016, pp. 1049–1058.

[29] I. Nonaka and G. von Krogh, "Tacit Knowledge and Knowledge Conversion: Controversy and Advancement in Organizational Knowledge Creation Theory," *Organ. Sci.*, vol. 20, no. 3, pp. 635–652, 2009.

[30] P. Delfmann and M. Hübers, "Towards Supporting Business Process Compliance Checking with Compliance Pattern Catalogues," *Enterp. Model. Inf. Syst. Archit.*, vol. 10, no. 1, pp. 67–88, 2015.

[31] "Energy Perspectives: Industrial and transportation sectors lead energy use by sector," *US Energy Information Administration*, 2012. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=9250. [Accessed: 22-Feb-2018].

[32] D. T. Shindell, G. Faluvegi, D. M. Koch, G. A. Schmidt, N. Linger, and S. E. Bauer, "Improved attribution of climate forcing to emissions," *Science (80-.).*, vol. 326, no. 5953, pp. 716–718, 2009.

[33] B. Brandherm and A. Kroner, "Digital Product Memories and Product Life Cycle," in *2011 Seventh International Conference on Intelligent Environments*, 2011, pp. 374–377.

[34] F. Kamoun, O. Alfandi, and S. Miniaoui, "An RFID solution for the monitoring of storage time and localization of perishable food in a distribution center," in *2015 Global Summit on Computer Information Technology (GSCIT)*, 2015, pp. 1–6.

[35] J. Mendling, H. A. Reijers, and W. M. P. van der Aalst, "Seven process modeling guidelines (7PMG)," *Inf. Softw. Technol.*, vol. 52, no. 2, pp. 127–136, 2010.

[36] P. Lübbecke, P. Fettke, and P. Loos, "Towards Guidelines of Modeling for Ecology-aware Process Design," in *BPM Workshops*, 2017, p. 10.