Security and privacy concerns in cloud-based scientific and business workflows

Soveizi, Nafiseh; Turkmen, Fatih; Karastoyanova, Dimka

Published in:
Future generation computer systems

DOI:
10.1016/j.future.2023.05.015

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Review article

Security and privacy concerns in cloud-based scientific and business workflows: A systematic review

Nafiseh Soveizi [*], Fatih Turkmen, Dimka Karastoyanova

University of Groningen, Groningen, The Netherlands

ARTICLE INFO

Article history:
Received 3 August 2022
Received in revised form 10 May 2023
Accepted 20 May 2023
Available online 26 May 2023

Keywords:
Security
Scientific workflows
Business processes
Cloud computing
Workflow adaptation
Life cycle

ABSTRACT

Today, the number of data-intensive and compute-intensive applications like business and scientific workflows has dramatically increased, which made cloud computing more popular because of its ability to deliver a large number of computing resources on-demand. Security is a critical issue affecting the wide adoption of cloud technologies, especially for workflows that are mostly dealing with sensitive data and tasks. In this paper, we carry out a review of the state-of-the-art on how security and privacy concerns in scientific and business workflows in cloud environments are being addressed and identify the limitations and gaps in the current body of knowledge in this area. In this extensive literature review, we first present the state-of-the-art security solutions organized according to the phases of the workflow life cycle they target for both business and scientific workflows. The analysis shows that most of the existing literature focuses on the modeling and execution phases, while the monitoring and adaptation phases are not covered adequately by a scarce amount of publications thus leaving a huge gap in the body of knowledge relevant to detection, prevention of and reaction to security violations in cloud-based workflows.

Contents

[*] Corresponding author.
E-mail address: n.soveizi@rug.nl (N. Soveizi).

https://doi.org/10.1016/j.future.2023.05.015

1. Introduction

Workflows are commonly used application models that consist of a series of computational tasks logically connected by data- and control-flow dependencies [1]. There are two main types of workflows: The first type – *scientific workflows*, typically involve a large amount of data processing, analysis, and computing, requiring high computing and storage capacities. The second type of workflows is *business workflows*, applied predominantly in (business) information systems and enables process automation and improvement. Unlike scientific workflows, individual activities/tasks in business workflows usually require lower computing power and fewer storage resources, although the number of concurrently running workflow instances is typically large and the communication time between tasks needs to be kept as short as possible [2]. Both academia and industry have been active in setting the foundations of workflow management (as a subfield of Business Process Management) as a discipline in the last several decades by developing the workflow technology that allows for modeling, executing workflows, and thus automating enterprise processes.

Cloud computing [3] plays a key role in workflow management since it can deliver a large amount of computing resources on-demand [2] for running data-intensive and compute-intensive applications. It also promises to reduce running costs and maximize revenues while maintaining or even improving the Quality of Service (QoS). Making use of cloud computing by Workflow Management Systems (WfMSs) can further increase the productivity of the system. Seen from the point of view of cloud computing users, cloud workflows [4,5] provide an abstract definition of complex distributed applications, flexible configuration, and automated scalable operation, and also improve the QoS of the workflow execution environments (i.e. the WfMS) and hence the processes themselves.

From the perspective of the providers of cloud computing services, cloud workflows enable the automatic scheduling and monitoring of tasks (the process of mapping tasks to cloud resources within the required QoS) and management of resources [6].

Despite all the above-mentioned advantages of cloud-based workflows, cloud security is a major area of concern [7,8] that is restricting their use for certain applications, especially for the workflows dealing with sensitive data and tasks. In fact, when a workflow or part of it is outsourced to the cloud, the WfMS loses control over tasks that can lead to increased security risks and make them vulnerable to malicious attacks. The security challenges associated with cloud computing arise primarily from two factors. Firstly, the nature of cloud infrastructure means that computing and storage resources are shared with other users, and sensitive data is transferred among cloud components, such as Data Centers (DCs), over potentially untrusted network channels. Secondly, the distributed nature of workflows means that they can dynamically bind to cloud services during execution for various reasons, while these services may experience security issues that were not known at the time of binding. In addition, the cloud is honest-but-curious in the sense that the cloud service provider may faithfully follow the established protocols but at the same time, it may be curious to deduce valuable information about the users' data and the workflow logic. Since the deduced information may be leaked or even sold to third parties by malicious cloud providers [9], some users are reluctant to use the cloud (deployment model).

In search for solutions to these concerns, there have been a number of studies on the topic of security properties of processes and workflow management systems from different perspectives. In the scope of our own current research towards automated and proactive process adaptation while processes are being executed [10–12], we first need to understand the currently available approaches towards addressing workflow security concerns, in particular in Cloud-based environments, and identify the gaps in the body of available knowledge. Therefore, with this study we aim to achieve two main goals:

(1) To establish the current state of the art in addressing and maintaining security properties of cloud-based workflows throughout their complete life cycle, including a special focus on an additional life cycle phase accounting for runtime adaptation.

(2) To identify the gaps in the state of the art and subsequently the needs for future research.

The methodology selected to achieve these goals is presented in Section 4 in detail. We use a combination of a Systematic Mapping Review (SMR) [13] and a Systematic Literature Review [14] as they best fit the objectives of our study. The SMR step resulted in establishing the state of the art in the field and identifying the areas with missing research, whereas the SLR allowed for a focused assessment of the area with missing research.

Using this methodology, in our study we identified that most of the solutions for ensuring the security properties of cloud-based workflows focus on the modeling aspects of workflows, mainly providing modeling concepts to express the security properties of workflows. Solutions towards enforcing these properties are rare, narrow in scope, and in some cases implementation-specific. One significant finding is that there is only very scarce research reported on approaches for reacting to security violations during the execution/runtime phase of cloud-based workflows, which in the workflow lifecycle fits in the runtime adaptation phase.

The rest of the paper is organized as follows: Section 2 introduces the basic concepts. In Section 3, we discuss the existing literature about the security and privacy of cloud-based workflows. Section 4 presents the review process and data collection. In Section 5, the main results of our study are presented. Section 6 illustrates the open issues and the challenges that still need to be addressed in this topic. Finally, Section 7 concludes this paper.

2. Background: Similarities and differences between scientific and business workflows

The Business Process Management (BPM) field (that covers also workflow management technology) and scientific workflows are established research fields and have been regarded as separate fields of research. In the last decade though, there have been several attempts to apply approaches from the BPM field in the field of scientific workflows both in terms of modeling approaches and principles, as well as in terms of using workflow management environments to run scientific workflows for different application fields [15,16]. At the same time, these works are based on the observations that there are both similarities and differences between these seemingly disparate fields. In this section, we highlight these differences and similarities, and provide the necessary background information on the topic.

Based on the available literature, we can summarize the definitions of both terms as follows: A *scientific workflow* describes a series of computations that enable the analysis of data in a

Table 1
The similarities between scientific and business workflows.

Criterion	Business and scientific workflows
Security	Both types of workflows need to satisfy the fundamental security principles of CIANA (Confidentiality, Integrity, Availability, Non-Repudiation, and Authenticity) [17] during the whole workflow life cycle [18].
Robustness	Requirements like the ability to be error-resistant and recoverable are similar for both types of workflows [15].
Scalability	Both types of workflows require the ability to scale with the number of users, services, data resources, and involved participants [15].

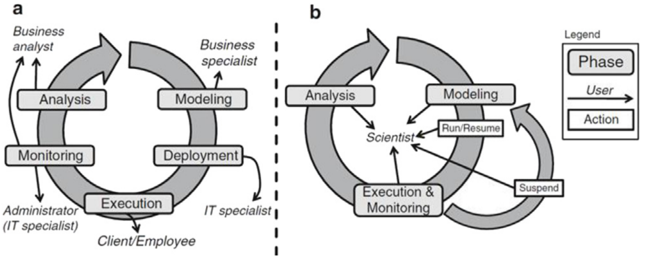

Fig. 1. (a) Business Workflow Life Cycle (b) Scientific Workflow Life Cycle [15].

structured and distributed manner. It orchestrates and automates scientific applications in a way that reduces the complexity of managing scientific experiments [19].

A business workflow is the automation of a business process, in whole or in part, during which documents, information, or tasks are passed from one participant to another for action, according to a set of procedural rules [20].

The life cycles of workflows that both fields follow have a different focus as depicted in Fig. 1. The two life cycles clearly show that the two fields view the workflows from different perspectives: business workflows are viewed as a software artifact that can be used by several user roles with a focus on different aspects of the management of the artifacts (in most cases modeling, execution, monitoring, and analysis), whereas the scientific workflows revolve around one user role, namely the scientist, who is dealing with both management of the software artifacts and their use. Note that we will mostly refer to the business process life cycle in our study, as it is the more detailed one and hence provides a more detailed basis for the comparison of the existing works.

Despite the seemingly different focus of the two fields, the literature shows [15,16,18,19,21,22] similarities as well, which were the reasons for the recent technological advances mentioned above. Based on these similarities especially in the security concerns, in this study, we look into both types for workflows together. Our goal of investigating these two types side by side is three-fold. First, the two areas exhibit many similar requirements, and there already exist lots of standards and tools in business areas that can be applied to scientific workflows too [15]. Second, despite the differences, it is possible to provide an integrated solution in order to support both types of workflows at the same time as supported by [11,15,16,23–30] and security related solutions could be reused across these two workflow types. Last but not least, on the one hand, the focus of the publications that we identified in our study showed significant differences in the scope of the solutions (modeling time focus in business workflows vs execution time approaches for scientific) whereas we want to obtain a complete overview of the state of the art and on the other hand, the number of works per workflow type is low. The details about the analysis of the available publications are presented in Section 5.1. We summarize the similarities and the differences between scientific and business workflows in Table 1 and Table 2 respectively [15,16,18,19,21,22].

3. Related surveys

Reviews are typically divided into two types: Systematic literature Review (SLR) [31] and Traditional Literature Review (TLR)[1] [32,33]. SLRs usually try to answer well-defined questions by following a specific search strategy. On the other hand, TLRs usually do not mention their search strategy for finding relevant publications, and therefore in TLRs, searches may be ad-hoc and are thus not fully comprehensive.

There are very few literature reviews that can be related to the security and privacy of cloud-based workflows. In Table 3, we provide an overview of the relevant literature reviews.

The role of trust in service workflows has been examined and explored in [23]. The authors have defined trust as a complement to conventional security services (e.g., authentication, and authorization). Therefore, the main focus of this paper is trust that can improve security in workflows where security requirements are locally defined, globally integrated, and distributedly enforced. Based on their findings, workflows need to be more flexible in terms of trust mechanisms to enable an increase in the degree of automation.

The survey [6] that is closest in scope to our survey, provides an initial overview of cloud workflow security. It has mapped the specification of QoS to the workflow life cycle phases as follows: The QoS specification is done in the workflow modeling stage; QoS aware service selection happens in the instantiation stage of the workflow where the appropriate software and hardware services are selected based on the requirements specified in the previous stage; and QoS consistency modeling and QoS violation handling happen in the workflow execution stage. However, based on the publication year (2014) of this paper and also the type of its review, it does not provide a comprehensive overview of the recent developments. The TLR presented in [24] has surveyed the existing works by defining the factors needed in securing scientific workflows during execution, identifying several domains in which security is essential and sources of security threats. The paper only focuses on the scheduling phase of the scientific workflow.

In [25], the security concerns in resource scheduling have been investigated. The authors identified the different types of security constraints and classified models into three categories: data security, data center security, and infrastructure security. The focus of this paper is only on the scheduling phase.

These literature reviews have different goals and/or do not cover all phases of the workflow life cycle. Hence, we can conclude that there is a lack of a comprehensive study of the security and privacy concerns of the cloud-based workflows during the whole workflow life cycle and their effect on the WfMS architecture.

[1] Narrative review.

Table 2
The differences between business and scientific workflows.

Criterion	Business workflows	Scientific workflows
Workflow definition and execution	Business processes typically define control and data flow in a process model using a generic (domain-independent) notation. They are executed multiple times on a generic process execution environment [15,21].	Scientific workflows are defined either using a programming language or a domain-specific scientific workflow language or notation. Execution is system-specific too [19,22].
Data flow vs. control flow	Control-flow oriented, focus on tasks/activities and their ordering [15,21].	Data-flow oriented, explicit focus on data and its processing [15,19,21].
Life Cycle	Fig. 1. a: [15,16,21]. – Explicitly defined life cycle where phases focus on managing processes/workflows – One model and many instances – Different groups of users	Fig. 1. b: [15,16,21]. – Life cycle phases focus on managing the scientific computation from point of view of the user – No explicit distinction between workflow models and their instances – Scientists are the only user group [15,19]
Duration	– Short and long running processes – Number of instances may be huge [15]	– Short and long running computations – Number of instances smaller [15]
Flexibility (a.k.a. dynamicity)	Usually, workflows are pre-defined during the modeling phase. Most academic research results are available in process evolution and adaptation [15].	Need a high degree of flexibility because they are carried out in a trial-and-error manner [15].
Reproducibility	Less need for reproducibility [15,21].	Need reproducibility [15,21].
Fault Handling	Processes must be guaranteed to be complete and if any fault occurs, it should be handled. Means are available for fault and exception handling in the existing technologies [15]..	Conduct experiments that may or may not succeed. However, technical faults (e.g. server unavailability, network connection error) that may occur during the execution should be handled. FH and EH on scientific workflow level specific for the domain language, if at all available [15].
Interaction with participants	Data can be processed by machines or humans. In most cases, several users are involved [21].	In most cases, data is processed only by machines, and the scientists just manage and monitor the workflow execution [21].

Table 3
The summary of the related literature reviews.

Paper	Type of review	Type of considered workflow	Covered phases of the life cycle	Focus	Main findings
Viriyasitavat and Martin [23], 2012	TLR	Both business and scientific workflows	Deployment and Execution	Trust exhibited in service workflows (trust is considered as a complement to the conventional security services)	Formal definitions of trust need more study to be used as means for decision making in dynamic distributed environments and as a result, increase the degree of automation.
Anupa and Sekaran [6], 2014	TLR	Both business and scientific workflows	All phases of the workflow lifecycle	An overview of cloud workflows and security	There must be cloud-specific standards for securing the workflows in the cloud.
Francis et al. [24], 2018	TLR	Scientific workflows	Deployment and Execution	Security of the scientific workflows during the execution	There is a need of developing more models which will consider different parameters such as (execution) environment, and CPU configuration settings, for more than one workflow.
Sheikh et al. [25], 2019, covers 2006–2015	SLR	Both business and scientific workflows	Deployment and Execution	Security concerns in resource scheduling	The main focus of their reviewed studies is limited to Integrity, Availability, and Security.

4. Research methodology

As mentioned in the previous section, up to now, there is no comprehensive review that establishes the existing evidence and evaluates the security and privacy concerns in cloud-based business or scientific workflows. To overcome this, we use a combination of an SLR [14] and a Systematic Mapping Review (SMR) [13] to identify the current research challenges and also existing gaps that can give an overview of research in the area. Since the articles are not evaluated in much detail in practice according to the SMR protocol and hence more articles can be considered, as a first step, we used SMR to portray the relationship between literature and categories and identify gaps, and show in which topic areas there is a shortage of publications [13]. Subsequently, we use the mapping as a road map for the next steps, namely an SLR, with which we show further details about existing works on the identified research question. Our Review Methodology Structure is presented in Fig. 2. As this figure shows, there are two steps in the *Conducting phase*: (1) The research conducted for the SMR in order to select and evaluate the papers, and (2) The research conducted for the SLR in order to select the papers to investigate in more detail. Also in the *Documentation phase*, we have two steps: (1) Analyzing the data and classifying the results of SMR, and (2) Reporting the results of SLR based on the detailed study. In Fig. 3, we show how exactly we carried out the research (as described in Section 4.3.2) including the libraries used and mapping to the research questions answered at each phase of the study.

4.1. Research questions

The research questions we define for our research are as follows:

RQ1: What is the state of the art in security and privacy in cloud-based business and scientific workflows in each stage of their life cycle?

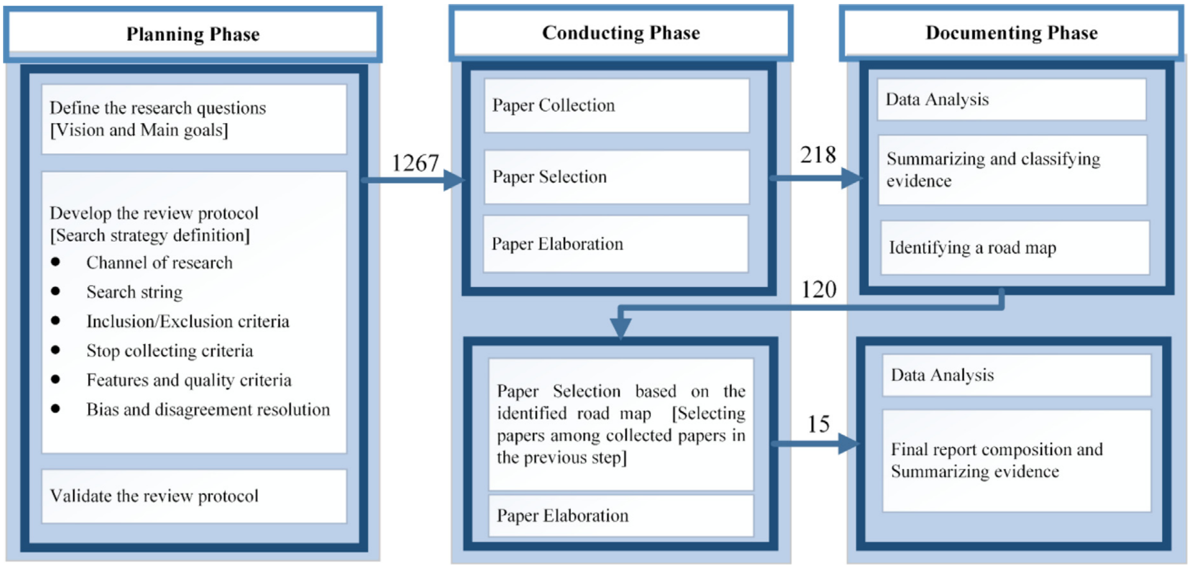
Planning Phase
Define the research questions [Vision and Main goals]
Develop the review protocol [Search strategy definition]
Channel of research
Search string
Inclusion/Exclusion criteria
Stop collecting criteria
Features and quality criteria
Bias and disagreement resolution
Validate the review protocol
1267
Conducting Phase
Paper Collection
Paper Selection
Paper Elaboration
218
Documenting Phase
Data Analysis
Summarizing and classifying evidence
Identifying a road map
120
Paper Selection based on the identified road map [Selecting papers among collected papers in the previous step]
Paper Elaboration
15
Data Analysis
Final report composition and Summarizing evidence

Fig. 2. Review Methodology Structure (based on [14]).

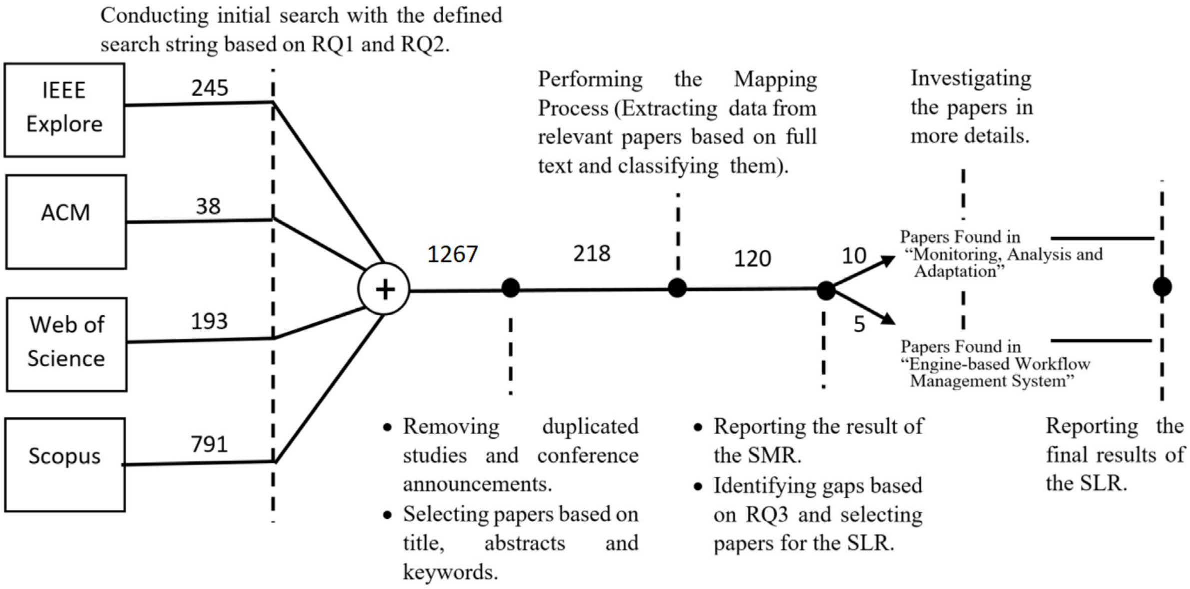
Conducting initial search with the defined search string based on RQ1 and RQ2.
IEEE Explore
245
ACM
38
Web of Science
193
Scopus
791
+
1267
218
120
10
5
Performing the Mapping Process (Extracting data from relevant papers based on full text and classifying them).
Investigating the papers in more details.
Papers Found in "Monitoring, Analysis and Adaptation"
Papers Found in "Engine-based Workflow Management System"
Removing duplicated studies and conference announcements.
Selecting papers based on title, abstracts and keywords.
Reporting the result of the SMR.
Identifying gaps based on RQ3 and selecting papers for the SLR.
Reporting the final results of the SLR.

Fig. 3. The study design process.

RQ2: Which security issues are addressed in which phase of the life cycle and what mechanisms are employed?

RQ3: Based on the research identified, what are the existing research gaps on which further research should focus?

4.2. Search strategy

The search was performed in four scientific databases, namely Scopus, Web of Science, ACM Digital Library, and IEEE Xplore. We also scanned the reference lists included in the papers in order to ensure that this review would be more comprehensive. The search was limited to papers in English published between January 2010 and December 2021.

The initial search string used to find the related papers is as follows:

((“security” OR “privacy”) AND (“scientific workflow” OR ""“business workflow”"" OR “business process” OR “service Composition” OR “orchestration”) AND (“cloud”))

In order to cover all papers that are related to the same or similar concepts in the literature, especially in the business context, we also used “business process“, ”service composition“, and ”orchestration” in the search string.

4.3. Study selection criteria and procedures

This section describes the inclusion/exclusion criteria that set the boundaries for the systematic review and also the procedures for performing the selection.

4.3.1. Inclusion/exclusion

The inclusion criteria for the selection of papers are:

- Published Conferences, Workshops, Journals, and peer-reviewed papers that address any aspect of security or privacy in cloud-based business or scientific workflows in one or more stages of their life cycle.
- Any previous literature reviews in this area.

The exclusion criteria are:

- Studies with inadequate information that are only available as presentations, abstracts, or otherwise incomplete.
- Duplicate reports of the same study. When several reports of a study exist in different journals, the most complete version of the study was included in the review.
- Forms of publication that have not been subjected to a formal review process (non-peer reviewed literature or gray literature), including journals such as ACM Software Engineering Notes (unless containing conference proceedings) and technical reports.
  - Opinion papers.

4.3.2. Procedures for selection

We found 1267 papers by the search performed in Section 4.2. After removing the duplicate publications and conference announcements, we reviewed the titles, abstracts, and keywords of these studies, and the approved studies (210 papers) were selected for further analysis based on the inclusion/exclusion criteria. Next, we extracted data from the relevant papers based on full text (we selected 120 relevant papers) and tried to classify the existing solutions that aim to answer the research questions RQ1 and RQ2. In the subsequent phase, we identified 15 papers that we investigated in order to answer the RQ3. The study design process is depicted in Fig. 3.

5. Results and discussion

In this section, we provide a summary of the results of the study and discuss our main findings. We consider the current body of knowledge from four perspectives: in Section 5.1, we give an overview of the available security solutions; in Section 5.2, we investigate the coverage of the workflow life cycle by the available literature; in Section 5.3, we focus on the life cycle phases monitoring, analysis and adaptation and in Section 5.4, we summarize the literature providing a concrete WfMS solution addressing security requirements. The following analysis is in the answer of the research question RQ1: "What is the state of the art in security and privacy in cloud-based business and scientific workflows in each stage of their life cycle?"

5.1. Security challenges and solutions

The papers found in the first phase of the review (120 papers) have tried to address the different security objectives of the workflows in the cloud environments. Fig. 4 shows the percentage of the security objectives considered by the papers. As the chart shows, Confidentiality (data and logic), Integrity (data and task), and Availability (CIA) are the most important security properties considered in the literature.

Some of these research works (109 papers) have tried to provide concrete solutions to achieve these objectives. Fig. 5 depicts the classification of the selected papers based on their proposed solutions in the context of cloud computing environments.

We have organized the provided solutions in three groups, namely: (1) virtualization and security services which focuses on solutions that reflect the perspective of cloud infrastructure providers, (2) administrative decision, which is a group of approaches reflecting the perspective of the workflow owners or WfMSs and the related decisions they need to make and (3) audit mechanism, which groups available mechanisms developed to audit workflow executions with the goal to capture and diagnose security violations and prevent them. In Fig. 6a we see that the bigger number of the publications in the first group comes from the scientific workflow field, whereas the larger number of publications dedicated to the second group focuses on business workflows. The third group is equally represented by both fields.

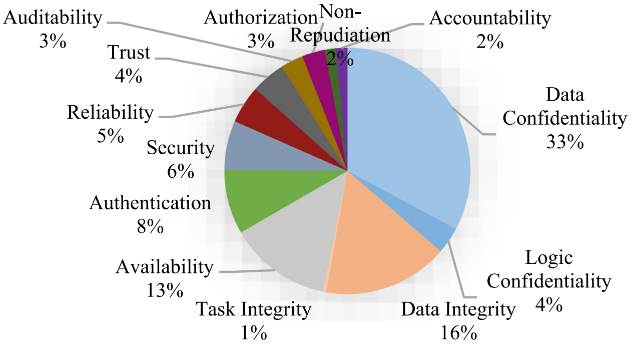

Fig. 4. Percentage of Security Objectives covered by the papers.

5.1.1. Virtualization and security services

Usually, cloud providers offer different security services and various levels of isolation guarantees for each service. For example, different availability methods (e.g. protective redundancy models and overload protection), encryption (SEAL, RC4, RC5, and IDEA), integrity services (e.g. hash functions like MD4, MD5, RIPEMD, RIPEMD-128, SHA-1, RIPEMD-160, and Tiger), and authentication mechanisms (e.g. HMAC-MD5, HMAC-SHA-1, and CBC-MAC-AES) are used to meet the availability, confidentiality, integrity, and authentication goals, respectively. These security services can be applied based on the level of security required by the user. Also, in order to provide confidentiality and user privacy in the workflow engine and/or on the client-side before deploying the process into an engine, different obfuscation and diversification techniques are proposed:

(a) *Data Obfuscation*: In this method, the confidentiality of the data is achieved by encrypting the data or obfuscating it on the client-side (e.g., splitting data, noise injection, and deleting sensitive data) before sending the data to the cloud;

(b) *Diversification*: This method tries to diversify the cloud execution environment continuously and lower the chance and the time for the attacker to discover the execution environment and its vulnerabilities. In other words, before the attacker acquires the knowledge about the execution environment, it would already be changed to a new one in order to render the acquired knowledge obsolete [34]. Diverse physical resources like servers, hypervisors, operating systems, or the workflow execution environments can be used to confuse the attackers [35];

(c) *Logic obfuscation* (BP obfuscation): In this method, the user or broker tries to split the process (splitting the BP model in a choreography of BP (fragments)) so that each cloud has only a partial view of the model.

(d) *Information Flow Checking*: This method tries to quantify the information flow to evaluate the intra-service leakage between different inputs and outputs of each service. It also aims to ensure inter-service flow security by evaluating the candidate services in the service chain [36].

5.1.2. Administrative decision

This group of approaches considers solutions enabling workflow owners directly or using an appropriate WfMS to make different security-related decisions at different workflow life cycle phases. These decisions have an impact on the security properties of the workflows throughout their life cycle. There is a variety of approaches that we grouped into the following three categories:

(a) *Extending Modeling and Execution Tools*: The papers in this group mostly focused on extending languages and/or modeling tools to specify user security and privacy requirements. For instance, some of them try to capture the access control requirements in business and scientific process specifications and then

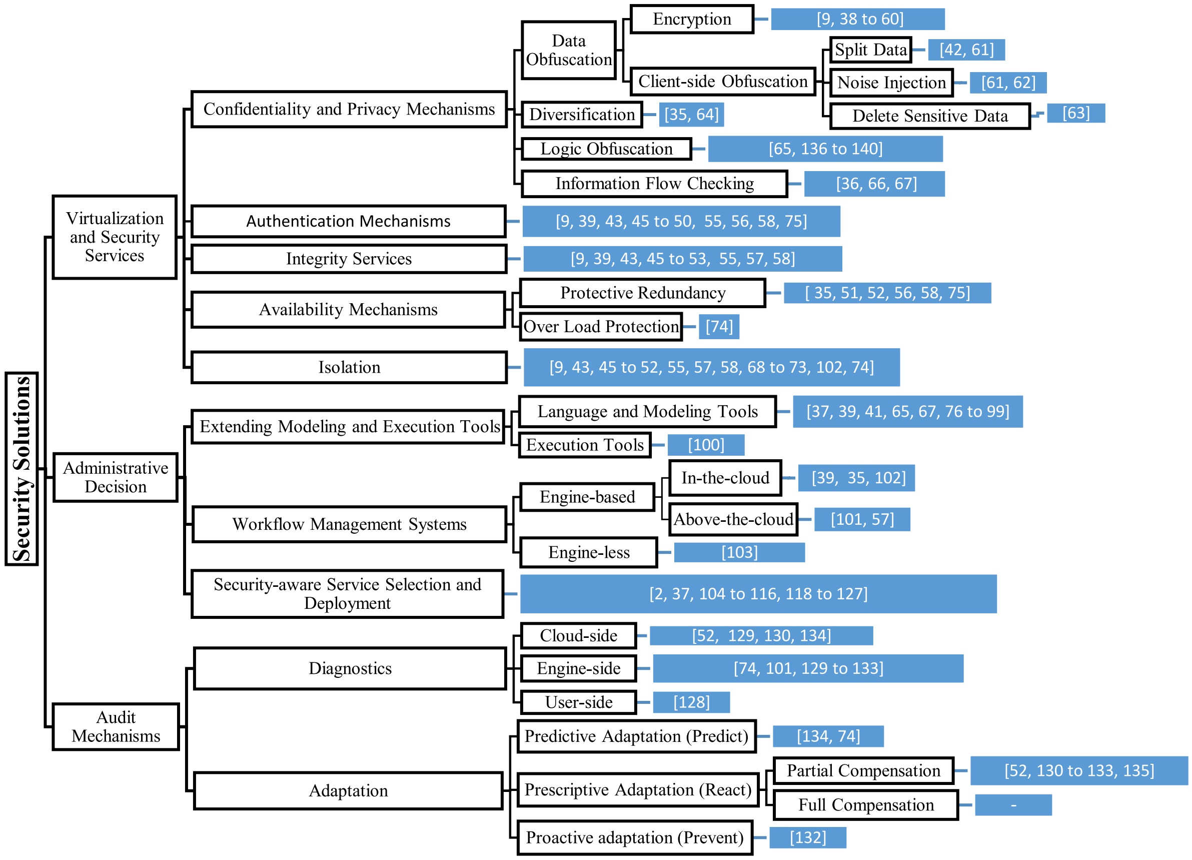

Fig. 5. Classification of the Security Solutions (see [2,9,35–140]).

tried to establish mechanisms to enforce these requirements, like supporting the principle of strict least privilege, the delegation of authority, the integrity principle, scalability, efficiency, and revocation [141–143]. Furthermore, a few papers try to verify the security of the workflows based on the predefined secure service composition (SCO) patterns [37].

(b) *Workflow Management System*: These papers addressed the architecture of the WfMS that can handle the security and privacy requirements. Most of the proposed WfMSs are engine-based. It means that they have an engine that is responsible for controlling the execution of the workflows. This engine can be placed in the private cloud or user-end system (i.e., Above-the-cloud) or it can be fully deployed in the public cloud (In-the-cloud) and thus the cloud has full control over the execution and monitoring of the workflows [6]. WfMS can also be centralized (controlling the workflow execution from one location in a single engine) or distributed (multiple workflow engines that provide the ability to cope with peaks in the system load and distributed environment) [103]. Note that in the literature identified in this study we found no evidence of the existence of a distributed cloud-based WfMS. More details about these engine-based security-aware WfMSs are discussed in Section 5.4. There is only one article [103] that introduces the notion of engine-less WfMS. The key idea of this kind of engine is to have the workflow process instance be self-protected and not need a workflow engine to secure the data therein. A detailed discussion on the engine-less WfMSs is out of the scope of this paper.

(c) *Security-aware Service Selection and Deployment*: These methods try to enable selection of services based on the user requirements. They must be able to make a balance between different user demands such as time, cost, and security during the scheduling of the processes. These papers addressed different aspects of security challenges during the scheduling of the workflows like confidentiality, integrity, authentication, availability, reliability, and trust. Most of them considered the level of employed security methods for each deployed VM that can fulfill the defined security requirements for each task. Also, for privacy, several papers tried to define some privacy protection constraints which help to select the best services for each task (e.g., data sensitivity constraints, data usage purpose constraints, data retention time constraints [107]) or restrict the sensitive tasks to be executed in the pre-defined locations (like a private cloud or specific hosts).

5.1.3. *Audit mechanisms*

These papers report on mechanisms developed to audit the workflow execution and prevent security violations as much as possible. They can be divided into two groups:

(a) The ones that only try to detect violations via monitoring. There are three different strategies in terms of the location of the monitoring functionality: (a1) Cloud-side Monitoring: These papers rely on a cloud monitoring module in order to detect security violations (e.g. [134,135], and [52]); (a2) Engine-side Monitoring: Violation detection is built into a new module/component of the workflow engine (e.g., [74,101,130], and [131]); (a3) User-side Monitoring: In this strategy, the user is responsible for monitoring and detecting violations in the workflows based on, for example, a log file (e.g., [128]). Each of these monitoring techniques has its limitations. For example, the first strategy is not fully resistant to some types of attacks. In other words, cloud insiders can misuse the access privileges to undermine the confidentiality, integrity, and availability of the systems [144]. This can be done by performing malicious activities in cloud logs with the aim of destroying attack traces, modifying and deleting log

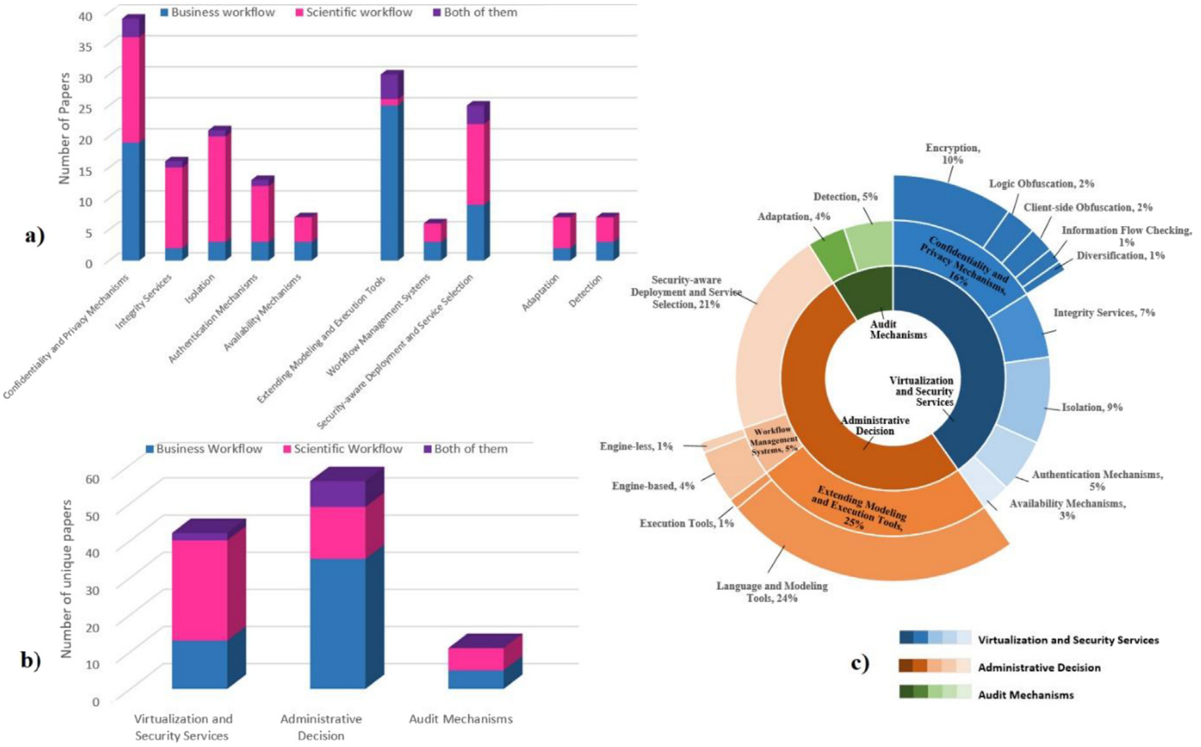

Fig. 6. Percentage of Security Solutions proposed by the papers disaggregated by the type of workflow, i.e. business, scientific workflows, or both types. (a) Percentage of each subcategory disaggregated by the type of workflow, (b) The number of unique papers in each category disaggregated by the type of workflow, (c) Percentage of each solution in detail. **(color should be used in print).**

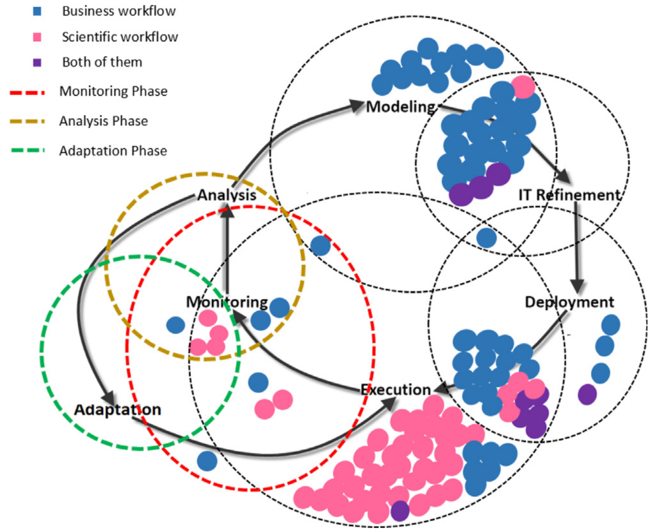

Fig. 7. Workflow life cycle coverage by the available security and privacy-related publications disaggregated by the type of workflow.

data, diverting the investigation process in other directions so as to hide them, extracting sensitive data, and others [145]. On the other hand, the second method can lead to time and cost overhead for the workflow engine, whereas the last one is often not scalable and can only be applied for small workflows by skilled users.

(b) Those that also try to predict, prevent, or react to violations after detection. In other words, these papers try to: (b1) *Predict* violations that may happen in the future by extracting knowledge from the security logs and past violations, (b2) *React* to them by using recommended actions to eliminate or reduce the effects of already occurred violations in real-time (Full or Partial Compensation), or (b3) *Prevent* them from happening by detecting any violation sign.

More details about these methods can be found in Section 5.3.

Fig. 6 also shows the percentages of each security solution proposed by the papers disaggregated by the type of workflow. It should be noted that because most of the papers have dealt with more than one objective with more than one solution, Fig. 6b shows the number of unique papers in each category.

5.2. Workflow life cycle and security

In Fig. 7, we show the workflow life cycle that also includes the phase of workflow adaptation. It also visualizes the coverage of the selected papers in the SLR regarding the security and privacy concerns per phase disaggregated by the type of workflow. The following analysis addresses our research question RQ2, namely, it focuses on the life cycle coverage of the publications we considered. As shown in Fig. 7, most of the papers on scientific workflows only focus on the execution phase. In the modeling and IT refinement phases, just a very few research works tried to address the security and privacy requirements of scientific workflows, whereas almost all papers considering business workflows focus on this phase. One of the reasons is that scientists do not distinguish between workflow modeling and executions, as explained in Section 2. Scientists develop their workflows in a trial-and-error manner and hence the modeling and execution phases are not arranged in a strict sequence [15]. As a result, there is a gap in the literature regarding the modeling or specification of security requirements in the scientific workflows in these phases of the life cycle.

Furthermore, as we can see in the same figure (Fig. 7), the majority of the papers for both scientific and business workflows focused on security aspects at either the modeling or execution phase thus leaving a huge open research space in the monitoring, analysis, and especially in the adaptation phase of workflows in cloud environments. More specifically, to the best of our knowledge, there is no paper proposing a solution to the adaptation of the currently running instances of business workflows in case of security violations. Most of the papers that address the modeling phase (which are mostly also including IT refinement), focused on extending language and modeling tools to specify/capture user security requirements at different levels of abstraction. Many of them define extensions to the widely known BPMN (Business Process Model and Notation) to support the specification of the non-functional requirements such as security. These works used graphical interfaces or textual notations to enrich BPMN diagrams with security properties [146]. Furthermore, some of these research works tried to define initial approaches in order to meet various security goals. For example in [39], firstly, the security goals are described in degrees (e.g., "high" and "low"), and also some actions are defined to meet these different degrees of security. Then, the predefined actions are enforced based on the security goal degree of each task. For example, for the tasks with a high level of confidentiality, the more secure cryptography algorithm, authentication, and access restriction must be applied, whereas for the low-level ones, a less secure cryptographic algorithm is sufficient. Considered from a different perspective, the work [77] models and analyzes the security threats for each security goal and proposes security requirements on the realization based on these threats. For instance, they defined certain threats such as the disclosure of information for confidentiality, and then proposed some requirements (like data obfuscation) as solutions to prevent these threats. Besides these extensions, some papers tried to consider the cloud features in the modeling phase and model the processes based on these features. These papers tried to use

BP obfuscation in order to preserve the privacy of processes in the cloud. Moreover, after the modeling phase, some papers check the models in terms of information flow or control flow via various model checking techniques (as mentioned in Section 5.1).

In the *deployment and execution phase*, the works propose the selection of the best fit services for each task and apply security services to them based on the user security requirements. At the same time, the aim is to find a balance between these requirements and other user preferences. The predominantly used decision-making criteria include trust, reliability, reputation, cost, or risk assessment in order to find the best fit services.

In the *monitoring, analysis, and adaptation phases*, the papers mostly used audit mechanisms to monitor the execution of processes, and then based on the obtained information, they aimed at preventing violations or reacting to them. These works are discussed in more detail in Section 5.3

In Fig. 8, we present the percentages of the selected papers per workflow life cycle phase.

5.3. Monitoring, analysis, and adaptation

In this section, we discuss the papers that addressed the security concerns in the monitoring, analysis, and adaptation phases. The presented analysis and findings complement the ones presented in Section 5.2 in addressing research question RQ2. The papers under consideration can be categorized based on the type of adaptation they focus on. We identify the following categories, as inspired by [147]: (1) Diagnostics; (2) Predictive Adaptation; (3) Prescriptive Adaptation; (4) Proactive Adaptation. Table 4 shows the summary of the 10 papers, which are the result of our last stage of selection in the research methodology (see Fig. 3). The works belonging to each of these groups are summarized in the following subsections.

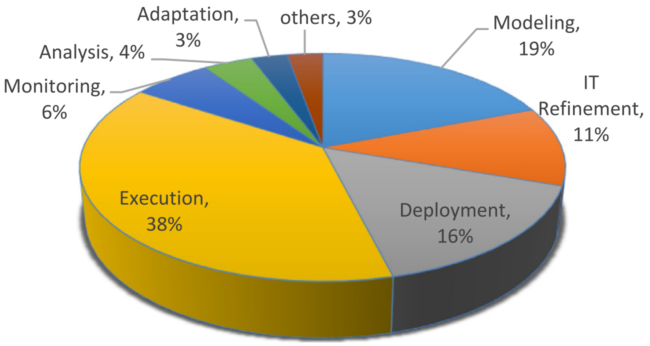

Fig. 8. Percentage of Security Objectives covered by the papers.

5.3.1. Diagnostic

The research works in this group focused only on the detection of the security violations during the monitoring phase of the workflows. The contributions per publication are summarized below.

The authors of [128] provided an architecture that allows users to verify the correctness of the business process executions remotely. For this purpose, they used a mechanism to log sensitive activities of business processes. After the process is completed, the client can request a signed version of the log and check it periodically to verify the correctness of the process execution. This paper did not consider real-time monitoring during execution and automatic adaptation.

In [101], the authors used provenance information for security purposes. They extended the Kepler provenance module and added the Security Analysis Package (SAP) to it in order to analyze provenance information in the security context. They focused on three data-flow oriented security properties: (1) input validation: using a whitelist of acceptable inputs to detect and filter unauthorized input, (2) remote access validation: implementing an internal firewall that contains the valid URLs and IP addresses, and (3) data integrity: comparing on-hash and post-hash of data to check data integrity. However, the provenance information was only used for detecting a few security violations and defining a follow-up action for the future scientific workflow executions in an attempt to prevent attacks.

In [129] the authors used cloud-wide auditing to uncover security issues. They defined Vulnerability Diagnostic Trees (VDTs) to formally manifest vulnerability patterns across several audit trails. This method can be used for implementing automated detection mechanisms that take various audit trails as input and pinpoints threats using their type and location. However, they did not consider the different requirements of users and services during their diagnostics. Furthermore, this method is not scalable as investigating all possible attacks using this method for all services can bring unnecessary time and computing overhead to the system reducing the scope of the work down to only a few possible attacks and violations in the service composition. Clearly, this paper only detected some violations based on the audit log and could not prevent or react to them.

5.3.2. Predictive adaptation

The research works in this category attempted to extract knowledge from the security logs to predict the future states, outcomes, or properties of a process instance or group of process instances and then prevent or reduce further violations based on these predictions and improve future decisions. In other words, these methods try to use dynamic scheduling based on real-time information to prevent further violations in the new

Table 4
The summary of the related papers in the monitoring, analysis, and adaptation phases in the cloud.

Paper	Adaptation type				Type of workflow	Trigger (adaptation reason)	Monitoring and detecting module
	Diagnostics (Detect)	Predictive adaptation (predict)	Prescriptive adaptation (react)	Proactive adaptation (prevent)			
[128], 2011	Verifying the Integrity of the computing platform and the correct execution of the outsourced processes by the user.	–	–		Business workflows	–	The user verifies the log file (User-side Monitoring)
[101], 2015	Input validation (detecting unauthorized input), remote access validation, and data integrity.	–	–		Scientific workflows	–	Security Analysis Package (SAP) embedded in the Kepler WfMS (Engine-side Monitoring)
[129], 2013	Detecting security issues (information leakage, Man-in-The-Middle (MiTM) attacks, DoS attacks, and resource misuse that can affect confidentiality, integrity, and availability of information)	–	–		Business workflows	–	Vulnerability Diagnostic Tree (VDT) model by using the audit log (it can be Engine-side or Cloud-side Monitoring)
[134], 2020	–	Minimize the security risk and improve the future composite decisions using this log information	–		Business workflows	Security issues (availability, integrity, and confidentiality)	Assumption that all the abnormal behaviors (security attacks) are logged in a broker module (Cloud-side Monitoring)
[74], 2016	Detecting any KPI violations because of the peak-loads period	Find a near-optimal solution based on this information to avoid further violations while considering security requirements.	–		Business workflows	KPI violations	Engine-side Monitoring
[130], 2020 [131], 2021	Detecting the sub-tasks with low confidence based on the lagged decision mechanism	–	Preserving the intermediate data for re-execution of sub-tasks with low confidence and avoiding the execution of the workflows from the beginning.		Scientific workflows	The sub-tasks with low confidence	Lagged Decision Mechanism embedded in the workflow monitor module (Engine-side Monitoring)
[52], 2018	–	–	Rescheduling the affected tasks		Scientific workflows	Any kind of malicious behavior in a VM	The Cloud Administrator module
[135], 2020	–	–	Rescheduling the uncompleted services.		Scientific workflows	Change of cloud resources availability (like clouds fail and availability of new clouds)	The Cloud Monitor module (Cloud-side Monitoring)

(continued on next page)

Table 4 (*continued*).

Paper	Adaptation type				Type of workflow	Trigger (adaptation reason)	Monitoring and detecting module
	Diagnostics (Detect)	Predictive adaptation (predict)	Prescriptive adaptation (react)	Proactive adaptation (prevent)			
[133], 2021	Detecting failure of tasks	–	Re-executing the failure tasks	–	Scientific workflows	Resource unavailability, an expiry of the deadline	Engine-side Monitoring
[132], 2021	Detecting a high or very high state of future resource load by predicting the resource load status in each VM	–	A live VM migration to transfer the current load of the damaged VM to another VM	Balancing the processing load in resources dynamically	Scientific workflows	Unavailability of VMs	Engine-side Monitoring

instances, and hence, they cannot prevent attacks in already running instances.

[134] used the providers' security log (which registers cloud anomalies such as cybersecurity attacks) to quantify the Cloud Security Risk (CSR) regarding the user's business process and then formulated the web service composition problem as a bi-objective optimization problem with service cost and multi-cloud risk viewpoints. We categorize this work in the Diagnostics group as it uses log information to improve the process design by minimizing the security risk and hence diagnoses potential issues with a cloud provider. However, this work may also be considered as a prescriptive approach with respect to the composition to carry pot, as it prescribes the design of a service composition, however, no adaptation is taking place in this work.

[74] presented an adaptive and context-aware decision system that can predict peak-load times of business processes where the KPI thresholds are mostly surpassed/violated. The authors use a decision tree technique which is applied to the business process execution log to extract this knowledge. Then, the penalty-based genetic algorithm is applied to find a near-optimal solution while considering the required level of security for each task. However, like the previous paper, this work also cannot adapt the currently running process instances.

5.3.3. Prescriptive adaptation

The papers discussed here aim at detecting violations and automatically recommending actions to prevent or reduce the effects of already occurred violations in real-time. In other words, these works focus on the adaptation of already running instances in order to fully or partially compensate for the effect of the occurred violations by taking proper actions.

The authors of [130] proposed an intrusion tolerant scientific workflow system. They used different techniques to secure the workflow execution: (1) Executing the same sub-task in parallel in multiple heterogeneous Virtual Machines (VMs) to enhance reliability; (2) Proposing the dynamic task scheduling strategy based on resource circulation to cut off the attack chain; (3) Designing the temporary workflow intermediate data backup strategy. This preserved the intermediate data which can be used for the re-execution of the workflow sub-tasks with low confidence; the confidence of tasks is assessed by the so-called lagged decision mechanism as the assessment is provided after the tasks have been executed. In [131], the same authors tried to solve a limitation of the previous work by scheduling the replicas of sub-tasks so that the attacker cannot destroy the whole workflow just by compromising one VM. This method can be considered as an adaptable technique that can avoid executing compromised workflows from the beginning. Furthermore, this paper only focused on detecting integrity violations during the execution phase of scientific workflows and could not discover the other types of security violations.

The research presented in [135] introduced a dynamic rescheduling method to handle the changes in cloud resources availability like the failure of the existing resources or the availability of new resources. In other words, if a changed status is detected, the cost model dynamically calculates the cost of deploying uncompleted tasks onto the currently available cloud resources and reschedules them to handle run-time failures. Note that resource availability during the execution of the scientific workflow is the only security factor considered by this work.

An attack response approach to reduce the security threat in scientific workflows is introduced in [52] in which the security threat is calculated by considering one task to be malicious and then estimating the number of affected tasks (using a simple variant of the decision tree). Then after detecting any kind of malicious behavior in a VM by the cloud administrator, the attack response module tries to reschedule the high-risk tasks (including the running tasks and future tasks). The paper only addressed the integrity and availability of the data during the execution of scientific workflows and no other security requirements are considered.

5.3.4. Proactive adaptation

These methods try to predict the violations before they occur and then adapt the already running instances in order to prevent these violations from happening.

The work presented in [132] proposed a combination of two adaptation approaches during the scheduling of the scientific workflows in order to tolerate faults (i.e., the unavailability of VMs): (1) Proactive Adaptation: Applying a prediction model to proactively control resource load fluctuation. In other words, this model can increase the failure prediction accuracy before fault/failure occurrence. (2) Prescriptive Adaptation: Applying a reactive fault tolerance technique for when a processor fails and the scheduler must allocate a new VM to execute the workflow tasks. However, this paper only focused on the VM faults and did not address other security issues.

5.4. Workflow management system

In this section, we provide details about the contributions that mention the development of new engine-workflow management systems or extend the already existing ones to handle the security requirements.

[35] proposed a framework of, what they call, a "mimic cloud workflow execution system" with three strategies: heterogeneity (diversification of physical servers, hypervisors, and operating systems), redundancy (Lagged Decision Mechanism), and dynamics (switching workflow execution environment). However, this system only covered the execution and monitoring phases of the workflow life cycle and cannot carry out adaptation of the process instances.

Table 5
Different WfMSs regarding security concerns in the cloud.

WfMS	Type of supported workflow	Supported representation model	Extension	Execution environment	Covered security objectives
[35], 2018	Scientific	DAG	–	Cloud-based	• Data Integrity • Data Confidentiality
Sec-DATAVIEW [57], 2019	Scientific	DAG	DATAVIEW	Execute the kernel of WfMS (the components that process confidential data) inside SGX enclaves (other components are executed on the trusted premises such as private cloud computing platforms or the user side premise).	• Data/Task Integrity • Data/Logic Confidentiality
[101], 2015	Scientific	DCG	Kepler	Flexibly arranged to run locally or on a cloud platform	• Data Integrity
BPA-Sec4Cloud [39], 2016	Business	BPMN	–	Cloud-based	• Data Confidentiality • Data Integrity • Authentication
[102], 2019	Business	BPMN	jBPM4	Cloud-based	• Data/Logic Confidentiality

[57] developed a secure big data workflow management which they called SecDATAVIEW, based on DATAVIEW [148]. This system leverages the hardware-assisted trusted execution environments (TEEs) such as Intel Software Guard eXtensions (SGX) and AMD Secure Encrypted Virtualization (SEV) to protect the execution of big data workflows and the data used by them. They also proposed a secure architecture and the WCPAC (Workflow Code Provisioning and Communication) protocol for securing the execution of workflow tasks in remote worker nodes. However, this system is still vulnerable to some attacks including network traffic-analysis, denial-of-service, side-channel attacks, and fault injections. Furthermore, it only protects workflows from possible attacks during execution, and if an attack occurs, it terminates the workflow execution. In other words, there is no adaptation module in this system.

[101] extended the Kepler provenance module and added the Security Analysis Package (SAP) to it in order to analyze provenance information in the security context using three security properties (i.e., input validation, remote access validation, and data integrity). As mentioned earlier, this module can only detect a few violations and cannot adapt running workflows to react to security violations.

[39] proposed an integrated environment named BPA-Sec4Cloud, which aims to provide a holistic and integrated cloud-based solution to address the automation of security-aware business processes from its modeling to its deployment. For instance, as already discussed in Section 5.2, they used a BPMN-based Editor with security abstractions and service modeling capabilities support. However, the monitoring, analysis, and adaptation phases are not supported by the system.

[102] presented a cloud workflow engine based on the extended jBPM4 (Java Business Process Management) that can support privacy protection between different tenant workflow instances in the cloud workflow systems. They defined three levels of isolation: (1) data isolation: protecting the private data produced in the execution process of tenants' workflow instances; (2) performance isolation: protecting the process instance information belonging to different tenants at run-time; (3) execution isolation to meet the different performance requirements for various tenants. However, they cannot detect the security violations in workflows and like others, did not consider the adaptation phase.

These WfMSs are summarized in Table 5. We observe that none of the existing WfMSs are able to handle all of the security concerns during the whole workflow life cycle. Furthermore, there is no WfMS that can adapt running workflows instances in order to prevent security violations or react to them.

6. Findings and future work directions

In this section, we summarize our main findings and discuss some of the challenges and open issues in different phases of the workflow life cycle for further investigation and hence answer research question RQ3. We present the findings and identified gaps organized per workflow lifecycle phase.

Our first finding related to the modeling phase is that there is a *need to automate the modeling of the security requirements and their subsequent translation* to execution related artefacts. For this purpose, we identify a need for a standard workflow modeling notation that can cover all security requirements and subsequently provide clear mapping rules in order to support specific solutions for the intended security goals. Such a support for automation can facilitate the understanding and use of the system even by non-security specialists for the purposes of conflict analysis, reuse, and validation of the model. In that context, we observed that each of the existing security-aware modeling languages, like the existing BPMN extensions, propose their own model and notation which creates a barrier to their wider adoption. Furthermore, despite existing formal specifications of security goals/properties and available security patterns, there is a *lack of automated model checking* for satisfying the workflow's security requirements and quantifying the information leakage and control flow risks between tasks.

Our study also shows that up till now, there are only a few research works that try to *consider the characteristics of Clouds in the modeling phase* before deploying workflows to the cloud. Such methods should model cloud-compatible workflows by striking a balance between functionality and different security principles. For instance, these methods should be aware of the conflicting effects of using the data-minimization mechanisms (like Client-side Obfuscation or BP Obfuscation) on other security requirements like accountability [149]. Therefore, we can conclude that there is *a lack of research work that accounts for the cloud features during workflow modeling* towards specifying all relevant information (using modeling tools and languages) for outsourcing (parts of) processes to the cloud.

In the deployment and execution phase, since most security challenges in cloud infrastructures are due to Virtualization Technology (VT), it is important to consider the relationship between virtualization technology and security properties and select the right VT. This can provide QoS to the end-users at lower cost and also a cost-effective solution and efficient resource utilization to the Cloud Service Providers (CSP). So, there is *a need of proposing a security-aware scheduling method* for selecting the proper VT (like Virtual machines (VM), Virtual Containers (VCs), Containers within VM, Lightweight VM, or Unikernel [150]) at the task-level and also at the workflow-level based on the workflow's characteristics and users' requirements.

This literature overview reveals a *significant gap in* the research under discussion regarding *the monitoring, analysis, and adaptation phases*. In the monitoring phase – which happens simultaneously with the execution phase –there is no reliable and scalable strategy that can detect all kinds of possible attacks during workflow execution. Besides, in the analysis and adaptation phases, there is a huge knowledge gap as there are no approaches for preventing and reacting to security and privacy violations. In particular, for the business workflows, there is almost no work dealing with adapting workflows to counteract such violations. In the scientific workflows context, only a few research works tried to adapt workflows considering only one or two types of violations. *Therefore, we conclude that there is no sufficient research for both scientific and business workflows that allows to detect, prevent, and react to security violations and compensate for part or all of the damage while the processes are running.*

Finally, as a result of this review of the literature, we can conclude that *there is no cloud-based WfMS that can handle the security and privacy concerns throughout the whole workflow life cycle* for both business and scientific workflows (or either type for that matter). Our study shows that there is no agreement neither within the two communities nor between them on which parts of the WfMS have to address security violations, nor there is a reference architecture as a guidance.

Based on the identified gaps we summarize next the potential *directions for future research.*

In the *modeling phase*, there is a need to provide an agreed-upon workflow model that can capture the fundamental security principles of CIANA. This requires establishing a set of adaptation strategies (e.g., skipping, re-working, re-sequencing) for each task in the workflow model and determining their potential impact on the overall value of the workflow and on the balance between value and security properties. As an example of such specification, consider the scenario of skipping certain tasks within the workflow, some tasks may be less critical and skipping them may have little impact on the overall workflow value. However, for other tasks, such as authentication tasks, skipping is not even an option in case of any violation, as they are essential to the security of the workflow. In such cases, the impact of skipping could be significant and compromise the overall security of the workflow. Therefore, by incorporating such specifications for each task in the modeling phase, the cost of each adaptation action could be estimated in execution time, thus allowing to select the most viable adaptation option for any violations that may arise. Furthermore, learning from the selected adaptation actions during runtime can provide insights for improving the workflow model. This means that the model can be continuously updated and refined based on real-world scenarios, leading to a more robust and effective system.

In the *execution phase*, the WfMS needs mechanisms to be able to make the right choices of the wide range of offers in the cloud environments based on the specified security related demands of the workflow models. Since the *deployment phase* is influenced by the execution environments, the mechanisms for transformation from security requirements on the model level to executable security mechanisms have to be considered in a coordinated manner.

The *monitoring phase* requires that the WfMS monitors the running workflow instances and the execution infrastructure, including the cloud providers/platforms to detect security violations (based on the Service-Level Agreement (SLA) or Key Performance Indicator (KPI)), unexpected behavior, failure of a task or instance, unavailability of a service or resource and so on. Detection of security violations can be considered as a trigger of activities to react to such violations appropriately and also prevent them from propagating in the whole workflow execution. Monitoring capabilities are in fact crucial, as there is no guarantee that services and cloud providers, which workflows may have selected for use, would not experience any security violations.

As to the adaptation phase, there is a significant lack of support in existing WfMSs. In other words, there is a need for a WfMS with detection and adaptation modules that can handle any possible violations. Furthermore, considering all potential malicious parties is also extremely important as they render the problem multifaceted, for which however there are no available solutions as of now and need to be developed for the different phases of the workflow lifecycle. For example, most of the existing WfMSs do not consider the cloud as a malicious party whose behavior needs to be monitored by the WfMS and therefore it is not possible to detect or prevent the providers' attacks. This is also true for the malicious behavior of different users during the execution of the workflows. In our view, such a cloud-based WfMS should provide an execution and monitoring environment that allows each user/participant to configure and run their workflows, meet their various functional and non-functional requirements, monitor the behavior of the workflows to detect any potential violations, and allow for reaction to security violations through adaptation while these workflows are being executed. To achieve this, it is necessary to incorporate multiple adaptation modules into different components of the WfMS architecture. One such module could be responsible for making informed decisions about which adaptation action is most appropriate for a given violation at the workflow level, taking into account the adaptation cost including the security requirements of the specific task, the type of attack, and the potential impact of attack on the overall workflow. Another module could focus on monitoring and adapting to violations occurring at the level of cloud services and users that are external to the system. This module would be responsible for detecting any anomalies or suspicious behavior and taking appropriate action to prevent any security breaches in the cloud services and users. A third module could be developed to predict potential violations and take proactive measures to prevent them from occurring in the first place. This research direction will involve designing machine learning algorithms and other predictive techniques to identify potential threats before they can cause any harm. By incorporating these various adaptation modules into the architecture, the system can maintain a high level of security and adaptability, effectively mitigating potential threats and ensuring the integrity of the system over time.

7. Conclusions

This paper presents a comprehensive overview of the security and privacy properties in cloud-based business and scientific workflows. We used a two-step methodology with which we aimed at establishing the state of the art on the topic and at identifying the main challenges and potential research directions in this area of research. The first step followed the SMR protocol in which we classified the available literature based on the proposed security solutions and their target phases in the workflow life cycle. To do so, we devised three categories covering various aspects of security and privacy concerns in workflows: Virtualization and Security Services, Administrative Decision, and Audit Mechanisms. Our findings show that most of the available relevant literature focuses predominantly on the modeling phase and less so on the execution phase. We established that there is a lack of sufficient attention to the monitoring, analysis, and adaptation phases for both business and scientific workflows, while the adaptation phase is in fact neglected up till now. As an additional result of this step, we identified a total of 120 publications that we investigated in the second step of our study.

In the second step, we followed the SLR protocol which we used on the set of publications identified by the SMR and the corresponding gaps in the state of the art related to security concerns in the monitoring, analysis, and adaptation phases of workflows. We compared the works using the criteria adaptation type, type of considered workflow, adaptation reason, and monitoring and detection module/mechanism. Based on that, we conclude that there is a gap in the state of current research on reliable and scalable approaches that can detect, prevent and react to security violations and compensate for part or all of the damage for both scientific and business workflows that are cloud-based. Furthermore, we investigated the existing WfMS implementations with respect to the type of supported workflow, the supported representation model, the covered phases of the workflow life cycle, and the covered security objectives. Based on our findings, we conclude that there is no comprehensive workflow management system in cloud environments that can handle the security and privacy concerns during the whole workflow life cycle neither in business nor in scientific workflow research. The conclusions of our study clearly identify a huge potential for future research in approaches and WfMSs that address the security concerns during the whole life cycle of cloud-based workflows and in particular in the monitoring, analysis, and adaptation phases.

CRediT authorship contribution statement

Nafiseh Soveizi: Conceptualization, Methodology, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization. **Fatih Turkmen:** Validation, Writing – review & editing. **Dimka Karastoyanova:** Supervision, Validation, Conceptualization, Methodology, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] M. Dumas, M. La Rosa, J. Mendling, H.A. Reijers, Introduction to business process management, Fundamentals Business Process Manage. (2018) 1–33.

[2] X. Ye, S. Liu, Y. Yin, Y. Jin, User-oriented many-objective cloud workflow scheduling based on an improved knee point driven evolutionary algorithm, Knowl.-Based Syst. 135 (2017) 113–124, http://dx.doi.org/10.1016/j.knosys.2017.08.006.

[3] P. Mell, T. Grance, et al., The NIST definition of cloud computing, 2011.

[4] H. Hua, Z. Yi-Lai, Z. Min, A survey of cloud workflow, Adv. Mater. Res. (1343) (2013) http://dx.doi.org/10.4028/www.scientific.net/AMR.765-767.1343, 765–767–1348.

[5] W. Li, Q. Zhang, J. Wu, J. Li, H. Zhao, Trust-based and qos demand clustering analysis customizable cloud workflow scheduling strategies, in: Proc. - 2012 IEEE Int. Conf. Clust. Comput. Work. Clust. Work. 2012, 2012, pp. 111–119, http://dx.doi.org/10.1109/ClusterW.2012.21.

[6] J. Anupa, K.C. Sekaran, Cloud workflow and security: A survey, in: Proc. 2014 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2014, 2014, pp. 1598–1607, http://dx.doi.org/10.1109/ICACCI.2014.6968496.

[7] S. Varshney, R. Sandhu, P.K. Gupta, Qos based resource provisioning in cloud computing environment: A technical survey, in: International Conference on Advances in Computing and Data Sciences, 2019, pp. 711–723.

[8] S.T. Maguluri, R. Srikant, L. Ying, Stochastic models of load balancing and scheduling in cloud computing clusters, in: 2012 Proceedings IEEE Infocom, 2012, pp. 702–710.

[9] W. Liu, S. Peng, W. Du, W. Wang, G.S. Zeng, Security-aware intermediate data placement strategy in scientific cloud workflows, Knowl. Inf. Syst. 41 (2) (2014) 423–447, http://dx.doi.org/10.1007/s10115-014-0755-x.

[10] A.Y. Ghahderijani, D. Karastoyanova, Autonomic process performance improvement, in: Proc. - IEEE Int. Enterp. Distrib. Object Comput. Work. EDOCW, 2021, pp. 299–307, http://dx.doi.org/10.1109/EDOCW52865.2021.00061.

[11] L. Pufahl, D. Karastoyanova, Enhancing business process flexibility by flexible batch processing, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), in: LNCS, vol. 11229, 2018, pp. 426–444, http://dx.doi.org/10.1007/978-3-030-02610-3_24.

[12] D. Karastoyanova, A. Buchmann, Extending web service flow models to provide for adaptability, in: Proc. OOPSLA, No. May, 2004, [Online]. Available: http://130.83.166.137/publications/pdf/WS-flow-Adaptability-OOPSLA04.pdf.

[13] K. Petersen, R. Feldt, S. Mujtaba, M. Mattsson, Systematic mapping studies in software engineering, in: 12th International Conference on Evaluation and Assessment in Software Engineering (EASE), vol. 12, 2008, pp. 1–10.

[14] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, S. Linkman, Systematic literature reviews in software engineering - a systematic literature review, Inf. Softw. Technol. 51 (1) (2009) 7–15, http://dx.doi.org/10.1016/j.infsof.2008.09.009.

[15] K. Görlach, M. Sonntag, D. Karastoyanova, F. Leymann, M. Reiter, Conventional workflow technology for scientific simulation, 2011.

[16] M. Sonntag, D. Karastoyanova, Next generation interactive scientific experimenting based on the workflow technology, 2010.

[17] T.T.W. Group, et al., The notorious nine: cloud computing top threats in 2013, Cloud Secur. Alliance (2013) 1–10.

[18] I.J. Taylor, E. Deelman, D. Gannon, M.S. Shields, Workflows for e-Science: Scientific workflows for grids, Work. E-Sci. Sci. Work. Grids (2007) 1–523, http://dx.doi.org/10.1007/978-1-84628-757-2.

[19] M.A. Rodriguez, R. Buyya, Scientific Workflow Management System for Clouds, first ed., Elsevier Inc., 2017.

[20] M. Weske, Business process management architectures, in: Business Process Management, Springer, 2012, pp. 333–371.

[21] J. Liu, et al., A survey of data-intensive scientific workflow management to cite this version : HAL Id : lirmm-01144760, J. Grid Comput. 13 (4) (2019) 457–493.

[22] M. Mattoso, et al., Towards supporting the life cycle of large scale scientific experiments, Int. J. Bus. Process Integr. Manag. 5 (1) (2010) 79–92, http://dx.doi.org/10.1504/IJBPIM.2010.033176.

[23] W. Viriyasitavat, A. Martin, A survey of trust in workflows and relevant contexts, IEEE Commun. Surv. Tutor. 14 (3) (2012) 911–940, http://dx.doi.org/10.1109/SURV.2011.072811.00081.

[24] A.O. Francis, B. Emmanuel, D.D. Zhang, W. Zheng, Y. Qin, D.D. Zhang, Exploration of secured workflow scheduling models in cloud environment: A survey, in: Proc. - 2018 6th Int. Conf. Adv. Cloud Big Data, CBD 2018, 2018, pp. 71–76, http://dx.doi.org/10.1109/CBD.2018.00022.

[25] A. Sheikh, M. Munro, D. Budgen, Systematic literature review (SLR) of resource scheduling and security in cloud computing, Int. J. Adv. Comput. Sci. Appl. 10 (4) (2019) 35–44, http://dx.doi.org/10.14569/ijacsa.2019.0100404.

[26] D. Karastoyanova, F. Leymann, Making scientific applications on the grid reliable through flexibility approaches borrowed from service compositions, Handb. Res. P2P Grid Syst. Serv. Comput. Model. Methodol. Appl. 2 (2010) 635–656, http://dx.doi.org/10.4018/978-1-61520-686-5.ch027.

[27] M. Reiter, U. Breitenbücher, O. Kopp, D. Karastoyanova, Quality of data driven simulation workflows, in: 2012 IEEE 8th Int. Conf. E-Science, E-Science 2012, 2012, http://dx.doi.org/10.1109/eScience.2012.6404417.

[28] M. Sonntag, S. Hotta, D. Karastoyanova, Using services and service compositions to enable the, 242–253.

[29] M. Sonntag, D. Karastoyanova, F. Leymann, The missing features of workflow systems for scientific computations, in: Lect. Notes Informatics (LNI), Proc. - Ser. Gesellschaft Fur Inform., vol. P-160, 2010, pp. 209–216.

[30] L. Stage, D. Karastoyanova, Provenance holder: Bringing provenance, reproducibility and trust to flexible scientific workflows and choreographies, in: Lect. Notes Bus. Inf. Process, in: LNBIP, vol. 362, 2019, pp. 664–675, http://dx.doi.org/10.1007/978-3-030-37453-2_53.

[31] B.A. Kitchenham, et al., Refining the systematic literature review process-two participant-observer case studies, Empir. Softw. Eng. 15 (6) (2010) 618–653, http://dx.doi.org/10.1007/s10664-010-9134-8.

[32] E.T. Rother, Systematic literature review x narrative review, Acta Paul. Enferm. 20 (2) (2007) v–vi.

[33] D.S. Cruzes, T. Dyb, Research synthesis in software engineering: A tertiary study, Inf. Softw. Technol. 53 (5) (2011) 440–455, http://dx.doi.org/10.1016/j.infsof.2011.01.004.

[34] S. Hosseinzadeh, S. Hyrynsalmi, M. Conti, V. Leppänen, Security and privacy in cloud computing via obfuscation and diversification: A survey, in: Proc. - IEEE 7th Int. Conf. Cloud Comput. Technol. Sci. CloudCom 2015, 2016, pp. 529–535, http://dx.doi.org/10.1109/CloudCom.2015.29.

[35] Y. wen Wang, J. xing Wu, Y. fei Guo, H. chao Hu, W. yan Liu, G. zhen Cheng, Scientific workflow execution system based on mimic defense in the cloud environment, Front. Inf. Technol. Electron. Eng. 19 (12) (2018) 1522–1536, http://dx.doi.org/10.1631/FITEE.1800621.

[36] N. Xi, C. Sun, J. Ma, J. Lv, Distributed quantitative information flow evaluation for service composition in clouds, in: Proc. - 2019 18th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. IEEE Int. Conf. Big Data Sci. Eng. Trust. 2019, 2019, pp. 200–207, http://dx.doi.org/10.1109/TrustCom/BigDataSE.2019.00035.

[37] L. Pino, G. Spanoudakis, M. Krotsiani, K. Mahbub, Pattern-based design and verification of secure service compositions, IEEE Trans. Serv. Comput. 13 (3) (2020) 515–528, http://dx.doi.org/10.1109/TSC.2017.2690430.

[38] I. Khabou, M. Rouached, A. Viejo, D. Sánchez, Privacy-preserving orchestrated web service composition with untrusted brokers, Int. J. Inf. Technol. Web Eng. 13 (4) (2018) 78–103, http://dx.doi.org/10.4018/IJITWE.2018100105.

[39] F. Lins, J. Damasceno, R. Medeiros, E. Sousa, N. Rosa, Automation of service-based security-aware business processes in the cloud, Computing 98 (9) (2016) 847–870, http://dx.doi.org/10.1007/s00607-015-0476-3.

[40] D. Derler, C. Hanser, H.C. Pöhls, D. Slamanig, Towards authenticity and privacy preserving accountable workflows, IFIP Adv. Inf. Commun. Technol. 476 (644962) (2016) 170–186, http://dx.doi.org/10.1007/978-3-319-41763-9_12.

[41] I. El Kassmi, Z. Jarir, Towards security and privacy in dynamic web service composition, in: Proc. 2015 IEEE World Conf. Complex Syst. WCCS 2015, 2016, pp. 1–6, http://dx.doi.org/10.1109/ICoCS.2015.7483260.

[42] S.A. Ghafour, P. Ghodous, C. Bonnet, Privacy-aware cloud services composition, in: Proc. - 2015 IEEE 24th Int. Conf. Enabling Technol. Infrastructures Collab. Enterp. WETICE 2015, 2015, pp. 140–142, http://dx.doi.org/10.1109/WETICE.2015.49.

[43] J. Angela Jennifa Sujana, T. Revathi, S. Joshua Rajanayagam, Fuzzy-based security-driven optimistic scheduling of scientific workflows in cloud computing, IETE J. Res. 66 (2) (2020) 224–241, http://dx.doi.org/10.1080/03772063.2018.1486740.

[44] Z. Li, et al., A security and cost aware scheduling algorithm for heterogeneous tasks of scientific workflow in clouds, Futur. Gener. Comput. Syst. 65 (2016) 140–152, http://dx.doi.org/10.1016/j.future.2015.12.014.

[45] A.R. Arunarani, D. Manjula, V. Sugumaran, FFBAT: A security and cost-aware workflow scheduling approach combining firefly and bat algorithms, Concurr. Comput. 29 (24) (2017) 1–13, http://dx.doi.org/10.1002/cpe.4295.

[46] L. Zeng, B. Veeravalli, X. Li, SABA: A security-aware and budget-aware workflow scheduling strategy in clouds, J. Parallel Distrib. Comput. 75 (2015) 141–151, http://dx.doi.org/10.1016/j.jpdc.2014.09.002.

[47] S. Hammouti, B. Yagoubi, S.A. Makhlouf, Workflow security scheduling strategy in cloud computing, in: International Symposium on Modelling and Implementation of Complex Systems, vol. 1, 2020, pp. 48–61, [Online]. Available: http://link.springer.com/10.1007/978-3-319-33410-3.

[48] H.Y. Shishido, J.C. Estrella, C.F.M. Toledo, Multi-objective optimization for workflow scheduling under task selection policies in clouds, in: 2018 IEEE Congr. Evol. Comput. CEC 2018 - Proc, 2018, http://dx.doi.org/10.1109/CEC.2018.8477799.

[49] X. Zhu, P. Jiao, Y. Zha, H. Chen, Security-aware workflow scheduling with selective task duplication in clouds, Simul. Ser. 48 (4) (2016) 114–121, http://dx.doi.org/10.22360/springsim.2016.hpc.048.

[50] H. Djigal, J. Feng, J. Lu, Performance evaluation of security-aware list scheduling algorithms in iaas cloud, in: Proc. - 20th IEEE/ACM Int. Symp. Clust. Cloud Internet Comput. CCGRID 2020, 2020, pp. 330–339, http://dx.doi.org/10.1109/CCGrid49817.2020.00-60.

[51] Y. Wang, Y. Guo, Z. Guo, W. Liu, C. Yang, Securing the intermediate data of scientific workflows in clouds with ACISO, IEEE Access 7 (2019) 126603–126617, http://dx.doi.org/10.1109/ACCESS.2019.2938823.

[52] F. Abazari, M. Analoui, H. Takabi, S. Fu, MOWS: multi-objective workflow scheduling in cloud computing based on heuristic algorithm, Simul. Model. Practice Theory 93 (2019) 119–132.

[53] D. Kim, M.A. Vouk, A formal model towards scientific workflow security in a cloud, Internat. J. Cloud Comput..

[54] Z. Wen, J. Cala, P. Watson, A. Romanovsky, Cost effective, reliable, and secure workflow deployment over federated clouds, in: Proc. - 2015 IEEE 8th Int. Conf. Cloud Comput. CLOUD 2015, 2015, pp. 604–612, http://dx.doi.org/10.1109/CLOUD.2015.86.

[55] M.R. Thanka, P. Uma Maheswari, E.B. Edwin, An improved efficient: Artificial bee colony algorithm for security and QoS aware scheduling in cloud computing environment, Cluster Comput. 22 (2019) 10905–10913, http://dx.doi.org/10.1007/s10586-017-1223-7.

[56] M. Anisetti, C.A. Ardagna, E. Damiani, F. Gaudenzi, G. Jeon, Cost-effective deployment of certified cloud composite services, J. Parallel Distrib. Comput. 135 (2020) 203–218, http://dx.doi.org/10.1016/j.jpdc.2019.09.003.

[57] S. Mofrad, I. Ahmed, S. Lu, P. Yang, H. Cui, F. Zhang, SecDataView: A secure big data workflow management system for heterogeneous computing environments, in: ACM Int. Conf. Proceeding Ser, 2019, pp. 390–403, http://dx.doi.org/10.1145/3359789.3359845.

[58] W.F. Ouedraogo, F. Biennier, P. Merle, Optimizing service protection with model driven security@run.time, in: Proc. - 9th IEEE Int. Symp. Serv. Syst. Eng. IEEE SOSE 2015, vol. 30, 2015, pp. 50–58, http://dx.doi.org/10.1109/SOSE.2015.50.

[59] J. Lei, Q. Wu, J. Xu, Privacy and security-aware workflow scheduling in a hybrid cloud, Futur. Gener. Comput. Syst. (2022) http://dx.doi.org/10.1016/j.future.2022.01.018.

[60] J. Kakkottakath Valappil Thekkepuryil, D.P. Suseelan, P.M. Keerikkattil, An effective meta-heuristic based multi-objective hybrid optimization method for workflow scheduling in cloud computing environment, Cluster Comput. 24 (3) (2021) 2367–2384, http://dx.doi.org/10.1007/s10586-021-03269-5.

[61] E. Goettelmann, N. Mayer, C. Godart, Integrating security risk management into business process management for the cloud, in: Proc. - 16th IEEE Conf. Bus. Informatics, CBI 2014, vol. 1, 2014, pp. 86–93, http://dx.doi.org/10.1109/CBI.2014.29.

[62] X. Liu, et al., Cloud workflow system quality of service, in: SpringerBriefs Comput. Sci, 2012, pp. 27–50, http://dx.doi.org/10.1007/978-1-4614-1933-4_4, (9781461419327).

[63] C. Ke, Z. Huang, F. Xiao, L. Liu, Privacy data decomposition and discretization method for SaaS services, Math. Probl. Eng. (2017) 2017, http://dx.doi.org/10.1155/2017/4785142.

[64] Y. Wang, Y. Guo, Z. Guo, T. Baker, W. Liu, CLOSURE: A cloud scientific workflow scheduling algorithm based on attack–defense game model, Futur. Gener. Comput. Syst. 111 (2020) 460–474, http://dx.doi.org/10.1016/j.future.2019.11.003.

[65] E. Goettelmann, A. Ahmed-Nacer, S. Youcef, C. Godart, Paving the way towards semi-automatic design-time business process model obfuscation, in: Proc. - 2015 IEEE Int. Conf. Web Serv. ICWS 2015, 2015, pp. 559–566, http://dx.doi.org/10.1109/ICWS.2015.80.

[66] T. Abdellatif, E2SM: a security tool for adaptive cloud-based service-oriented applications, IET Softw. 13 (1) (2019) 3–13, http://dx.doi.org/10.1049/iet-sen.2018.5016.

[67] P. Pullonen, R. Matulevičius, D. Bogdanov, PE-BPMN: privacy-enhanced business process model and notation, in: International Conference on Business Process Management, 2017, pp. 40–56.

[68] Y. Yang, X. Peng, J. Cao, Trust-based scheduling strategy for cloud workflow applications, Inform. 26 (1) (2015) 159–180, http://dx.doi.org/10.15388/Informatica.2015.43.

[69] M. Bidaki, S.R.K. Tabbakh, M. Yaghoobi, H. Shakeri, Secure and efficient SOS-based workflow scheduling in cloud computing, Int. J. Secur. Appl. 11 (2) (2017) 41–58, http://dx.doi.org/10.14257/ijsia.2017.11.2.05.

[70] S. Shahul Hammed, B. Arunkumar, Efficient workflow scheduling in cloud computing for security maintenance of sensitive data, Int. J. Commun. Syst. (July) (2019) 1–11, http://dx.doi.org/10.1002/dac.4240.

[71] M. Nguyen, S. Debroy, P. Calyam, Z. Lyu, T. Joshi, Security-aware resource brokering for bioinformatics workflows across federated multi-cloud infrastructures, 2020, pp. 1–10, http://dx.doi.org/10.1145/3369740.3369791.

[72] K. Boukadi, R. Grati, M. Rekik, H. Ben-Abdallah, Business process outsourcing to cloud containers: How to find the optimal deployment? Futur. Gener. Comput. Syst. 97 (2019) 397–408, http://dx.doi.org/10.1016/j.future.2019.02.069.

[73] Z. Wen, et al., GA-par: Dependable microservice orchestration framework for geo-distributed clouds, IEEE Trans. Parallel Distrib. Syst. 31 (1) (2020) 129–143, http://dx.doi.org/10.1109/TPDS.2019.2929389.

[74] M. Rekik, K. Boukadi, H. Ben-Abdallah, Towards an autonomic outsourcing to the cloud decision, in: Proc. - 25th IEEE Int. Conf. Enabling Technol. Infrastruct. Collab. Enterp. WETICE 2016, 2016, pp. 20–25, http://dx.doi.org/10.1109/WETICE.2016.15.

[75] M. Tao, K. Ota, M. Dong, Dependency-aware dependable scheduling workflow applications with active replica placement in the cloud, IEEE Trans. Cloud Comput. 7161 (c) (2016) http://dx.doi.org/10.1109/tcc.2016.2628374, 1–1.

[76] H. Chen, X. Zhu, D. Qiu, L. Liu, Z. Du, Scheduling for workflows with security-sensitive intermediate data by selective tasks duplication in clouds, IEEE Trans. Parallel Distrib. Syst. 28 (9) (2017) 2674–2688, http://dx.doi.org/10.1109/TPDS.2017.2678507.

[77] S. Zareen, A. Akram, S.A. Khan, Security requirements engineering framework with BPMN 2.0.2 extension model for development of information systems, Appl. Sci. 10 (14) (2020) http://dx.doi.org/10.3390/app10144981.

[78] O. Altuhhov, R. Matulevičius, N. Ahmed, An extension of business process model and notation for security risk management, Int. J. Inf. Syst. Model. Des. 4 (4) (2013) 93–113.

[79] M.E.A. Chergui, S.M. Benslimane, A valid BPMN extension for supporting security requirements based on cyber security ontology, in: International Conference on Model and Data Engineering, 2018, pp. 219–232.

[80] C.L. Maines, B. Zhou, S. Tang, Q. Shi, Adding a third dimension to BPMN as a means of representing cyber security requirements, in: 2016 9th International Conference on Developments in ESystems Engineering (DeSE), 2016, pp. 105–110.

[81] R. Matulevičius, Security risk-oriented BPMN, in: Fundamentals of Secure System Modelling, Springer, 2017, pp. 63–76.

[82] M. Salnitri, F. Dalpiaz, P. Giorgini, Designing secure business processes with SecBPMN, Softw. Syst. Model. 16 (3) (2017) 737–757.

[83] N. Argyropoulos, H. Mouratidis, A. Fish, Attribute-based security verification of business process models, in: Proc. - 2017 IEEE 19th Conf. Bus. Informatics, CBI 2017, vol. 1, 2017, pp. 43–52, http://dx.doi.org/10.1109/CBI.2017.37.

[84] K.S. Sang, B. Zhou, BPMN security extensions for healthcare process, in: 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, 2015, pp. 2340–2345.

[85] K. Zarour, et al., A BPMN extension for business process outsourcing to the cloud, in: World Conference on Information Systems and Technologies, vol. 2, 2019, pp. 833–843, http://dx.doi.org/10.1007/978-3-030-16181-1.

[86] N. Maroua, A. Adel, Z. Belhassen, Document business process model extension for modeling secure ubiquitous documents, in: Workshops of the International Conference on Advanced Information Networking and Applications, 2020, pp. 628–639.

[87] W. Viriyasitavat, A. Martin, Formal trust specification in service workflows, in: Proc. - IEEE/IFIP Int. Conf. Embed. Ubiquitous Comput. EUC 2010, 2010, pp. 703–710, http://dx.doi.org/10.1109/EUC.2010.111.

[88] M. Rekik, et al., Towards outsource-ability enabled BPMN, in: 2015 10th International Joint Conference on Software Technologies, vol. 1, 2015, pp. 1–10.

[89] X. Lin, X. Zhang, Workflow and role based access control model for cloud manufacturing, in: Proc. - 2013 IEEE 11th Int. Conf. Dependable, Auton. Secur. Comput. DASC 2013, 2013, pp. 65–71, http://dx.doi.org/10.1109/DASC.2013.39.

[90] J. Anupa, K.C. Sekaran, Securing cloud workflows using aggressive Chinese wall security policy, in: 1st Int. Conf. Networks Soft Comput. ICNSC 2014 - Proc. (2014), pp. 85–91, http://dx.doi.org/10.1109/CNSC.2014.6906714.

[91] N. Maroua, Z. Belhassen, A. Adel, Formal approach for authorization in distributed business process related task document role based access control, in: 2019 15th Int. Wirel. Commun. Mob. Comput. Conf. IWCMC 2019, 2019, pp. 1964–1970, http://dx.doi.org/10.1109/IWCMC.2019.8766520.

[92] J. Yanez-Sierra, A. Diaz-Perez, V. Sosa-Sosa, J.L. Gonzalez, A digital envelope scheme for document sharing in a private cloud storage, in: 2015 12th Int. Conf. Expo Emerg. Technol. A Smarter World, CEWIT, vol. 2015, 2015, pp. 1–6, http://dx.doi.org/10.1109/CEWIT.2015.7338158.

[93] B. Schwarzbach, B. Franczyk, L. Petrich, A. Schier, M. Ten Hompel, Cloud based privacy preserving collaborative business process management, in: Proc. - 2016 16th IEEE Int. Conf. Comput. Inf. Technol. CIT 2016, 2016 6th Int. Symp. Cloud Serv. Comput. IEEE SC2 2016 2016 Int. Symp. Secur. Priv. Soc. Netwo, 2017, pp. 716–723, http://dx.doi.org/10.1109/CIT.2016.59.

[94] L. Compagna, P. Guilleminot, A.D. Brucker, Business process compliance via security validation as a service, in: Proc. - IEEE 6th Int. Conf. Softw. Testing, Verif. Validation, ICST 2013, 2013, pp. 455–462, http://dx.doi.org/10.1109/ICST.2013.63.

[95] O. Rayis, A. Dogru, Authorization model definition for an adaptable workflow within cloud environment, ACM Int. Conf. Proc. Ser. (2019) 49–53, http://dx.doi.org/10.1145/3358505.3358526.

[96] M. Amini, F. Osanloo, Purpose-based privacy preserving access control for secure service provision and composition, IEEE Trans. Serv. Comput. 12 (4) (2016) 604–620, http://dx.doi.org/10.1109/tsc.2016.2616875.

[97] L. Lin, J. Hu, J. Zhang, Packet: a privacy-aware access control policy composition method for services composition in cloud environments, Front. Comput. Sci. 10 (6) (2016) 1142–1157, http://dx.doi.org/10.1007/s11704-016-5503-9.

[98] A.D. Brucker, I. Hang, Secure and compliant implementation of business process-driven systems, in: Lect. Notes Bus. Inf. Process, in: LNBIP, vol. 132, 2013, pp. 662–674, http://dx.doi.org/10.1007/978-3-642-36285-9-66.

[99] N. Maroua, A. Adel, Z. Belhassen, A new formal proxy-based approach for secure distributed business process on the cloud, in: Proc. - Int. Conf. Adv. Inf. Netw. Appl. AINA, 2018-May, 2018, pp. 973–980, http://dx.doi.org/10.1109/AINA.2018.00142.

[100] H.Y. Shishido, J.C. Estrella, C.F.M. Toledo, S. Reiff-Marganiec, A CloudSim extension for evaluating security overhead in workflow execution in clouds, in: Proc. - 2018 6th Int. Symp. Comput. Networking, CANDAR 2018, 2018, pp. 174–180, http://dx.doi.org/10.1109/CANDAR.2018.00031.

[101] D. Kim, M.A. Vouk, Securing scientific workflows, in: Proc. - 2015 IEEE Int. Conf. Softw. Qual. Reliab. Secur. QRS-C 2015, 2015, pp. 95–104, http://dx.doi.org/10.1109/QRS-C.2015.25.

[102] H. Huang, Y.L. Zhang, M. Zhang, Research on cloud workflow engine supporting three-level isolation and privacy protection, in: Proc. - 5th IEEE Int. Conf. Big Data Secur. Cloud, BigDataSecurity 2019, 5th IEEE Int. Conf. High Perform. Smart Comput. HPSC 2019 4th IEEE Int. Conf. Intell. Data Secur, 2019, pp. 160–165, http://dx.doi.org/10.1109/BigDataSecurity-HPSC-IDS.2019.00038.

[103] G.H. Hwang, Y.C. Kao, Y.C. Hsiao, Scalable and trustworthy cross-enterprise WfMSs by cloud collaboration, in: Proc. - 2013 IEEE Int. Congr. Big Data, BigData 2013, 2013, pp. 70–77, http://dx.doi.org/10.1109/BigData.Congress.2013.19.

[104] Y. Xiao, A.C. Zhou, X. Yang, B. He, Privacy-preserving workflow scheduling in geo-distributed data centers, Futur. Gener. Comput. Syst. 130 (2022) 46–58, http://dx.doi.org/10.1016/j.future.2021.12.004.

[105] M. Alam, M. Shahid, S. Mustajab, SAHEFT: Security aware heterogeneous earliest finish time workflow allocation strategy for IaaS cloud environment, in: 2021 IEEE Madras Sect. Conf, 2021, pp. 1–8, http://dx.doi.org/10.1109/mascon51689.2021.9563503.

[106] F. Lahmar, H. Mezni, Security-aware multi-cloud service composition by exploiting rough sets and fuzzy FCA, Soft Comput. 25 (7) (2021) 5173–5197, http://dx.doi.org/10.1007/s00500-020-05519-x.

[107] L. Liu, H. Zhu, S. Chen, Z. Huang, Privacy regulation aware service selection for multi-provision cloud service composition, Futur. Gener. Comput. Syst. 126 (2022) 263–278, http://dx.doi.org/10.1016/j.future.2021.08.010.

[108] R. Medara, R.S. Singh, Energy efficient and reliability aware workflow task scheduling in cloud environment, Wirel. Pers. Commun. 119 (2) (2021) 1301–1320, http://dx.doi.org/10.1007/s11277-021-08263-z.

[109] C. Ke, Z. Huang, W. Li, Y. Sun, F. Xiao, Service outsourcing character oriented privacy conflict detection method in cloud computing, J. Appl. Math. 2014 (2014) http://dx.doi.org/10.1155/2014/240425.

[110] H. Abrishami, A. Rezaeian, M. Naghibzadeh, A novel deadline-constrained scheduling to preserve data privacy in hybrid cloud, in: 2015 5th Int. Conf. Comput. Knowl. Eng. ICCKE 2015, No. October, 2015, pp. 234–239, http://dx.doi.org/10.1109/ICCKE.2015.7365833.

[111] W. Hu, X. Li, T. Ding, R. Ruiz, A trust constrained workflow scheduling method in cloud computing, ACM Int. Conf. Proc. Ser. F 1311 (2017) 197–200, http://dx.doi.org/10.1145/3127404.3127440.

[112] M. Farid, R. Latip, M. Hussin, N.A.W. Abdul Hamid, Scheduling scientific workflow using multi-objective algorithm with fuzzy resource utilization in multi-cloud environment, IEEE Access 8 (2020) 24309–24322, http://dx.doi.org/10.1109/ACCESS.2020.2970475.

[113] H. Hu, et al., Multi-objective scheduling for scientific workflow in multicloud environment, J. Netw. Comput. Appl. 114 (February) (2018) 108–122, http://dx.doi.org/10.1016/j.jnca.2018.03.028.

[114] X. Xu, S. Fu, W. Li, F. Dai, H. Gao, V. Chang, Multi-objective data placement for workflow management in cloud infrastructure using NSGA-II, IEEE Trans. Emerg. Top. Comput. Intell. 4 (5) (2020) 605–615, http://dx.doi.org/10.1109/TETCI.2019.2910242.

[115] Y. Wen, J. Liu, W. Dou, X. Xu, B. Cao, J. Chen, Scheduling workflows with privacy protection constraints for big data applications on cloud, Futur. Gener. Comput. Syst. 108 (2020) 1084–1091, http://dx.doi.org/10.1016/j.future.2018.03.028.

[116] S. Naser, S. Kamil, N. Thomas, A case study in inspecting the cost of security in cloud computing, Electron. Notes Theor. Comput. Sci. 318 (2015) 179–196, http://dx.doi.org/10.1016/j.entcs.2015.10.026.

[117] E. Goettelmann, K. Dahman, B. Gateau, C. Godart, A formal broker framework for secure and cost-effective business process deployment on multiple clouds, in: Lect. Notes Bus. Inf. Process. 204, 2015, pp. 3–19, http://dx.doi.org/10.1007/978-3-319-19270-3_1.

[118] E. Goettelmann, K. Dahman, B. Gateau, E. Dubois, C. Godart, A security risk assessment model for business process deployment in the cloud, in: Proc. - 2014 IEEE Int. Conf. Serv. Comput. SCC 2014, 2014, pp. 307–314, http://dx.doi.org/10.1109/SCC.2014.48.

[119] P. Xiao, D. Liu, User QoS enhanced web service composition framework in cloud platforms, Int. J. Netw. Virtual Organ. 13 (4) (2013) 351–364, http://dx.doi.org/10.1504/IJNVO.2013.064462.

[120] D.S. Marcon, et al., Workflow specification and scheduling with security constraints in hybrid clouds, in: 2nd IEEE Lat. Am. Conf. Cloud Comput. Commun. LatinCloud 2013, 2013, pp. 29–34, http://dx.doi.org/10.1109/LatinCloud.2013.6842219.

[121] Y. Wen, W. Dou, B. Cao, C. Chen, Towards scheduling data-intensive and privacy-aware workflows in clouds, in: Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng., in: LNICST, vol. 201, 2017, pp. 474–479, http://dx.doi.org/10.1007/978-3-319-59288-6_44.

[122] C. Hochreiner, Privacy-aware scheduling for inter-organizational processes, in: ZEUS, 2015, pp. 63–68.

[123] J. Shu, H. Jain, C. Liang, Business process driven trust-based task scheduling, Int. J. Web Serv. Res. 16 (3) (2019) 1–28, http://dx.doi.org/10.4018/IJWSR.2019070101.

[124] S. Sharif, P. Watson, J. Taheri, S. Nepal, A.Y. Zomaya, Privacy-aware scheduling SaaS in high performance computing environments, IEEE Trans. Parallel Distrib. Syst. 28 (4) (2017) 1176–1188, http://dx.doi.org/10.1109/TPDS.2016.2603153.

[125] A. Rezaeian, H. Abrishami, S. Abrishami, M. Naghibzadeh, A budget constrained scheduling algorithm for hybrid cloud computing systems under data privacy, in: Proc. - 2016 IEEE Int. Conf. Cloud Eng. IC2E 2016 Co-Located with 1st IEEE Int. Conf. Internet-of-Things Des. Implementation, IoTDI 2016, 2016, pp. 230–231, http://dx.doi.org/10.1109/IC2E.2016.42.

[126] H. Abrishami, A. Rezaeian, G.K. Tousi, M. Naghibzadeh, Scheduling in hybrid cloud to maintain data privacy, in: 5th Int. Conf. Innov. Comput. Technol. INTECH 2015, No. Intech, 2015, pp. 83–88, http://dx.doi.org/10.1109/INTECH.2015.7173369.

[127] W. Li, J. Wu, Q. Zhang, K. Hu, J. Li, Trust-driven and QoS demand clustering analysis based cloud workflow scheduling strategies, Cluster Comput. 17 (3) (2014) 1013–1030, http://dx.doi.org/10.1007/s10586-013-0340-1.

[128] S. Alsouri, S. Katzenbeisser, S. Biedermann, Trustable outsourcing of business processes to cloud computing environments, in: Proc. - 2011 5th Int. Conf. Netw. Syst. Secur. NSS 2011, 2011, pp. 280–284, http://dx.doi.org/10.1109/ICNSS.2011.6060015.

[129] K.J. Han, B.Y. Choi, S. Song, High performance cloud auditing and applications, in: High Perform. Cloud Audit. Appl, 2013, pp. 1–360, http://dx.doi.org/10.1007/978-1-4614-3296-8, 9781461432.

[130] Y. Wang, Y. Guo, Z. Guo, W. Liu, C. Yang, Protecting scientific workflows in clouds with an intrusion tolerant system, IET Inf. Secur. 14 (2) (2020) 157–165, http://dx.doi.org/10.1049/iet-ifs.2018.5279.

[131] Y. Wang, Y. Guo, W. Wang, H. Liang, S. Huo, INHIBITOR: An intrusion tolerant scheduling algorithm in cloud-based scientific workflow system, Futur. Gener. Comput. Syst. 114 (2021) 272–284, http://dx.doi.org/10.1016/j.future.2020.08.004.

[132] M. Alaei, R. Khorsand, M. Ramezanpour, An adaptive fault detector strategy for scientific workflow scheduling based on improved differential evolution algorithm in cloud, Appl. Soft Comput. 99 (2021) 106895, http://dx.doi.org/10.1016/j.asoc.2020.106895.

[133] Z. Ahmad, B. Nazir, A. Umer, A fault-tolerant workflow management system with quality-of-service-aware scheduling for scientific workflows in cloud computing, Int. J. Commun. Syst. 34 (1) (2021) http://dx.doi.org/10.1002/dac.4649.

[134] M. Hosseini Shirvani, Bi-objective web service composition problem in multi-cloud environment: a bi-objective time-varying particle swarm optimisation algorithm, J. Exp. Theor. Artif. Intell. 00 (00) (2020) 1–24, http://dx.doi.org/10.1080/0952813X.2020.1725652.

[135] Z. Wen, R. Qasha, Z. Li, R. Ranjan, P. Watson, A. Romanovsky, Dynamically partitioning workflow over federated clouds for optimising the monetary cost and handling run-time failures, IEEE Trans. Cloud Comput. 8 (4) (2020) 1093–1107, http://dx.doi.org/10.1109/TCC.2016.2603477.

[136] M. Skouradaki, V. Ferme, F. Leymann, C. Pautasso, D.H. Roller, BPELanon': Protect business processes on the cloud, in: CLOSER 2015-5th Int. Conf. Cloud Comput. Serv. Sci. Proc, 2015, pp. 241–250, http://dx.doi.org/10.5220/0005427502410250.

[137] A.A. Nacer, E. Goettelmann, S. Youcef, A. Tari, C. Godart, A design-time semi-automatic approach for obfuscating a business process model in a trusted multi-cloud deployment: A design-time approach for bp obfuscation, Int. J. Web Serv. Res. 15 (4) (2018) 61–81, http://dx.doi.org/10.4018/IJWSR.2018100104.

[138] M.N. Lacheheub, R. Maamri, A formal model for business process decomposition based on resources consumption with security requirement, in: ICAASE 2016 - Proc. Int. Conf. Adv. Asp. Softw. Eng, 2017, http://dx.doi.org/10.1109/ICAASE.2016.7843866.

[139] A. Ahmed Nacer, C. Godart, G. Rosinosky, A. Tari, S. Youcef, Business process outsourcing to the cloud: Balancing costs with security risks, Comput. Ind. 104 (2019) 59–74, http://dx.doi.org/10.1016/j.compind.2018.10.003.

[140] E. Goettelmann, W. Fdhila, C. Godart, Partitioning and cloud deployment of composite web services under security constraints, in: Proc. IEEE Int. Conf. Cloud Eng. IC2E 2013, 2013, pp. 193–200, http://dx.doi.org/10.1109/IC2E.2013.22.

[141] D.G. Cholewka, R.A. Botha, J.H.P. Eloff, A context-sensitive access control model and prototype implementation, IFIP Adv. Inf. Commun. Technol. 47 (2017) 341–350, http://dx.doi.org/10.1007/978-0-387-35515-3_35.

[142] J.D. Moffett, Control principles and role hierarchies, in: Proc. ACM Work. Role-Based Access Control, 1998, pp. 63–69, http://dx.doi.org/10.1145/286884.286900.

[143] S. Oh, S. Park, Task-role based access control (T-RBAC): An improved access control model for enterprise environment, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 1873, 2000, pp. 264–273, http://dx.doi.org/10.1007/3-540-44469-6_25.

[144] C.N. Modi, K. Acha, Virtualization layer security challenges and intrusion detection/prevention systems in cloud computing: a comprehensive review, J. Supercomput. 73 (3) (2017) 1192–1234, http://dx.doi.org/10.1007/s11227-016-1805-9.

[145] S. Khan, et al., Cloud log forensics: Foundations, state of the art, and future directions, ACM Comput. Surv. 49 (1) (2016) http://dx.doi.org/10.1145/2906149.

[146] F.A.A. Lins, E.T.G. Sousa, N.S. Rosa, A survey on automation of security requirements in service-based business processes, Int. J. Web Eng. Technol. 13 (1) (2018) 3–29, http://dx.doi.org/10.1504/IJWET.2018.092398.

[147] M. Dumas, Automated process improvement: Status, challenges, and perspectives, in: Adv. Inf. Syst. Eng. - 32nd Int. Conf. CAiSE 2020, Grenoble, Fr. June (2020) 8-12, Proceedings, in: Lect. Notes BPMDS EMMSAD pre-conference events, 2020, no. June.

[148] A. Kashlev, S. Lu, A system architecture for running big data workflows in the cloud, in: Proc. - 2014 IEEE Int. Conf. Serv. Comput. SCC 2014, 2014, pp. 51–58, http://dx.doi.org/10.1109/SCC.2014.16.

[149] Q. Ramadan, D. Strüber, M. Salnitri, V. Riediger, J. Jürjens, Detecting conflicts between data-minimization and security requirements in business process models, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), in: LNCS, vol. 10890, 2018, pp. 179–198, http://dx.doi.org/10.1007/978-3-319-92997-2_12.

[150] A.S. Gowri, et al., Impact of virtualization technologies in the development and management of cloud applications, Int. J. Intell. Syst. Appl. Eng. 7 (2) (2019) 104–110.

Nafiseh Soveizi received the B.S. and M.S. degrees from Ferdowsi University, Mashhad, Iran, in 2011 and 2015, respectively. She is currently pursuing a Ph.D. degree in computer science at the Information Systems Group, University of Groningen, The Netherlands. Her research interests include security and privacy in workflows, workflow adaptation, and cloud security.

Fatih Turkmen is a tenure track Assistant Professor at the University of Groningen working on security and privacy. His particular research interests include authentication and authorization infrastructures, privacy-preserving technologies (PET) and multi-party computing, and (semi)formal methods for security analysis. He has broad experience in the design and development of security infrastructures related to authentication and authorization when accessing data or services over the cloud.

Dimka Karastoyanova is a professor of Information Systems and a head of the Information Systems Group at the University of Groningen. Before that, she had a joint appointment as an associate professor of Data Science at The KLU in Hamburg and as a Senior Researcher at HPI, University of Potsdam, Germany and prior to that she was a junior professor in Simulation Workflows at the University of Stuttgart, Germany. She received her doctoral degree in Computer Science in 2006 from the Technische Universität Darmstadt, Germany. Her research interests are in the field of adaptive software systems with special focus on flexible workflows and corresponding middleware, process/workflow automation and autonomic process performance improvement, service-oriented systems and application, flexible data processing pipelines and enterprise information systems for various application domains.