IJIREEICE

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

Vol. 8, Issue 12, December 2020

DOI 10.17148/IJIREEICE.2020.81208

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

Designing Secure Data Pipelines for Regulatory Compliance in Cross-Border Tax Consulting

Pallav Kumar Kaulwar

IT Director, ORCID ID: 0009-0002-1142-0329

Abstract: Besides traditional on-premise or virtual data centers, a growing number of companies process their data in cross-border cloud environments. The taxation of such companies is a highly regulated task. As a result, protecting sensitive data is of utmost importance, while also keeping records available on their whereabouts due to potential requests from financial authorities. The same challenges can be observed in other business areas, such as the health sector, with sensitive patient data. This paper presents early design considerations on how to orchestrate cloud services while being able to comply with existing laws and regulations. To this end, a coherent architecture is outlined and key building blocks are explained in detail. In order to validate compliance with the orchestration-specific designs, the architecture must be set up to create appropriate compliance enforcement strategies.

To comply with the laws and regulations explicitly, the authors introduce initial design considerations on the high-level architecture to orchestrate a cross-border data pipeline in a compliant manner. Further, the key components of the architecture are discussed. In general, compliance verification measures need to be reusable on different data flows without additional manual effort. Such measures can be captured in a compliance enforcement strategy, which must be preconfigured for every set of compliance requirements. Authors envision that in order for information that flows between blocks of a data flow to be compliant, it has to conform to associated strategies, just like how a business process is expected to follow a business process model. Cloud computing has raised new opportunities and challenges for information and knowledge management. Due to the availability of ever greater computational resources and data, the development of ever more complex algorithms and statistical data analysis methods, data-driven decision making is on the rise. Compliance with existing regulation frameworks is a key challenge when building and operating cross-border data pipelines.

Keywords: Data Security, Regulatory Compliance, Cross-Border Data Transfer, Secure Data Pipelines, Data Encryption, Tax Consulting, GDPR Compliance, Data Sovereignty, Audit Trails, Data Governance, PII Protection, Real-Time Monitoring, Compliance Automation, Cloud Security, International Tax Regulation.

I. INTRODUCTION

In today's world, where billions of dollars worth of information are exchanged online, it has become imperative to ensure that these resources are processed in compliance with complex data protection regulations. These regulations are maintained in several jurisdictions across the globe and are being revised constantly, which leads to the emergence of various new compliance mandates almost daily. In an effort to facilitate the localization of compliance requirements across jurisdictions, different efforts have been taken up by organizations.

As cloud computing allows the deployment of systems in a frictionless global computing environment, it presents challenges for regulators and data processors in addressing compliance with this extensive and complex array of legal and regulatory obligations, particularly with regard to cross-border data flows. In this context, compliance with local and cross-jurisdictional legal and regulatory obligations has to be addressed in potentially untrusted environments, such as in the public cloud. In the public cloud environment, the processing may traverse multiple jurisdictions with different data protection laws. Each of these laws may require compliance on the part of several actors in the processing.

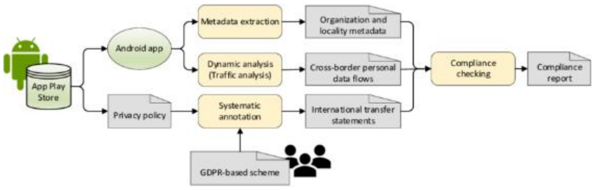
App Play Store
Android app
Privacy policy
Metadata extraction
Dynamic analysis (Traffic analysis)
Systematic annotation
GDPR-based scheme
Organization and locality metadata
Cross-border personal data flows
International transfer statements
Compliance checking
Compliance report

Fig 1: Cross-Border Tax Consulting

Most proposed frameworks focus on the post-mortem detection of non-compliance. This approach assumes that the requirements for compliance are well understood and the control measures have been specified. This has not been the case for the cloud, particularly with regard to the complex and extensive legal and regulatory obligations applicable in most jurisdictions. Absent a clear understanding and articulation of the relevant compliance obligations, adopting control measures would not assure compliance, especially in the cloud where systems are often used in a different manner or for a different purpose than intended. Hence, while advanced detection and validation techniques are invaluable for the investigation of suspected non-compliance, there is an urgent need for frameworks that assist organizations with the specification of their legal and regulatory obligations relevant to compliance.

In this paper, a framework addressing this need is introduced. The preference language is based on temporal logic, which is able to express relatively simple legal requirements such as those typically mandated by data protection laws. The types of temporal logic specifications at the heart of the framework comprise events, states or situations, and transitions between events, as well as mandatory, sleeping, and forbidden events or transitions. Each of these types of specifications relates to underlying control measures ensuring compliance with a specific requirement.

1.1. Background And Significance

Due to rapid automation and globalization, cross-border data flows (CBDF) or transborder data flows have become common for business operations and services. However, CBDF has made cross-border information control an urgent time-honored challenge for many governments. As modern digital services and applications become more advanced and widely adopted, an increasing number of countries are implementing stricter regulations on the data flow to avoid potential risks. One of the challenges is how to prevent the information leakage risks in cases of regulatory compliance violations during CBDF. A group of regulators that jointly monitors a CBDF provides an elegant solution to this challenge. Since they can share the privacy-preserving monitoring results with each other, such a group of regulators can identify a regulatory compliance violation even when the detailed information of sensitive features is not revealed to the public.

An important aspect of this compliance monitoring scenario is the privacy-preservation of CBDF auditors, where multiple regulators independently analyze the sensitivity of the payment data without leaking any sensitive information to each other. However, it is challenging to support the privacy-preserving compliance monitoring of CBDF with multiple regulatory frameworks. A pan-regional regulatory group must be formed first, which is usually impractical. Therefore, it is desirable to achieve the privacy-preserving compliance monitoring of CBDF across diverse regulatory mandates and frameworks.

When two regulatory agencies design competitive compliance monitoring for CBDF against cross-border e-commerce data by independently analyzing the data without revealing sensitive information to each other, two challenges arise: (1) Hospital A gives its collected data to Insurer A, but the insurance claim of a patient is a fraud. How do the regulators find the fraud without revealing the data itself? (2) A privacy-preserving auditing service is designed such that if either of the regulations fails, the auditors can prove it without revealing any sensitive information.

Equ : 1 Access Control Risk Score

$$R = \sum_{i=1}^{n} A_i \cdot P_i$$

- R: Total access control risk score
- A_i: Access frequency by user i
- P_i: Privilege level risk weight for user i

II. UNDERSTANDING REGULATORY COMPLIANCE

Regulatory compliance is an industry term that refers to how well organizations are able to comply guided by rules and/or regulations that govern their organization. Compliance is the process of ensuring adherence to rules, regulations and standards within a given domain. In the context of business, regulations can be internal, e.g. corporate governance and professional standards, or external, such as laws and statutes. Business processes can be defined as an interrelated series of tasks designed to produce a specific outcome.

Compliance constraints processes to adhere to rules, standards, laws and regulations. Non-compliance subjects enterprises to litigation, financial fines, loss of corporate image and restraining orders. Compliance relates to processes meeting a standard or required outcome. Organizations aspire to conduct business within the bounds of compliance regulations. However, compliance verification is a crucial requirement to deploy and implement processes in compliance management systems. Collaborative business processes imply that each of the collaborating organizations comply with a combination of internal and cross regional regulations. Organizations in the European Union must comply with the EU data privacy regulation, but also with each member state's data protection laws. Such a multitude of compliance regulations complicates verification of compliance of business processes and compliance with bigger, joined, or cross-organizational regulatory compliance. Typically, regulations can be categorized in different domains, also referred to as perspectives or layers. Organizations often have been verified for adherence to a company's policy, but a change in the governing regulation or in the business processes can render the adherence check no longer valid. It has been shown that compliance relates to conformance to different process perspectives/categories/layers, namely control flow, resources, data, and time.

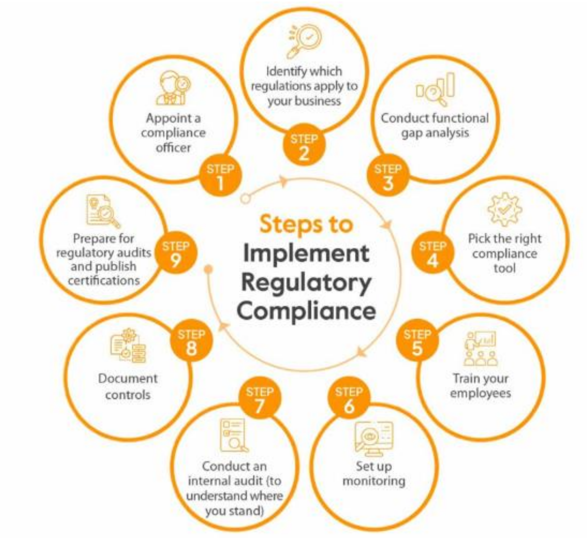
Steps to Implement Regulatory Compliance
STEP 1
Appoint a compliance officer
STEP 2
Identify which regulations apply to your business
STEP 3
Conduct functional gap analysis
STEP 4
Pick the right compliance tool
STEP 5
Train your employees
STEP 6
Set up monitoring
STEP 7
Conduct an internal audit (to understand where you stand)
STEP 8
Document controls
STEP 9
Prepare for regulatory audits and publish certifications

Fig 2: Steps to implement regulatory compliance

2.1. Overview of Regulatory Frameworks

There is a need for compliance verification within collaborative business processes to support end users in extricating individual business processes from the entire global one while guaranteeing compliance with internal policies and external regulations during the execution of those processes. The information security requirements of the generic business processes of the fiscal advisors that can be met are established and mapped to the relevant legislation and regulations. The Global Leader Tax Specialists network is presented, the selection of the process to map to legislation and regulations is explained, and a description of the compliance requirements to be met is given.

Each business process has its own process model, implementation, actors (process participants), data, and IT system that can potentially operate in industrial and public clouds. This information is stored in decentralized databases of the business process participants. For example, a required question is: What is the best way to securely transmit personal data of a customer from a Dutch tax advisor to a German tax advisor via a process with a fiscal attorney in the US? There are strict legislation and regulations, as this data contains tax return-related, sensitive personal information. The Financial Supervision Act prohibits moving this data to a country that is not included in the EU or for which there is no conditional suitability decision by the European Commission. Security mechanisms need to be in place around the business process to comply with the regulations. Some of these mechanisms are technological solutions or organizational measures. Not all mechanisms can be fulfilled in all collaborations as this is dependent on the individual business processes, agreements made between the parties, and the degree of trust between the parties.

2.2. Key Compliance Challenges

Compliance and identity verification. Organizations must comply with laws and regulations to limit the harm to themselves, their customers, and the society as a whole. Compliance with these laws must be verified. Efforts are made to define an organizations' claim regarding compliance to rules.

One method is to formally define and prove that if a system is in a certain state, an event in the future will lead to a new state, in which the organization will be compliant to certain regulations. In practice, the verification of compliance can be and is performed out of the box and before system implementation. These results of this verification must be converted to executable business rules. These rules can notify the absence of a certain obligatory event or warn of a potential loss of compliance.

Compliance in collaborative business processes. Business process models specify a specification in (natural) language. The process modelers use informal annotations, user-defined protocols, and simulated scenario runs to stipulate the compliance requirements of a business process model. Each enterprise utilizes their own compliance checkers in their own systems, and each partner will only focus on its own restrictions. However, while their own compliance requirements may be sufficiently verified, possible non-compliance of the overall cross-regional business process model remains unheard of. Non-compliance of the overall cross-regional business process model may obscure legal violations and expose extra sanctions for enterprises involved.

III. DATA PIPELINE ARCHITECTURE

A data pipeline is a directed acyclic graph (DAG) composed of a sequence of nodes. At least one source node produces the data at the beginning of the DAG, and at least one sink node finally receives the processed data. In terms of data pipeline output at runtime, at least one source node generates and outputs the data items, which are then sent to adjacent nodes, where they are processed. The internal execution and data flow on the graph can be executed concurrently. Once the source node sends out an output, the consumption of the output at the sink node triggers the termination of the data pipeline. From a practical point of view, data pipelines are a piece of software that automates the manipulation of data and moves them from diverse source systems to defined destinations. The primary purpose of this set of complex and interrelated software bundles is to enable efficient data processing, transfer, and storage, control all data operations, and orchestrate the entire data flow from source to destination. The main components of a data pipeline typically include data sources, processing component, data storage component, and destinations. The starting point of every data pipeline is its data sources, which can include databases, sensors, or text files. The sources define the origin of where the pipeline pulls the data from and publishes new data. A data source typically has a source type describing how and from where the data can be ingested. The second component is its processing component, which includes data ingestion, data preprocessing, and data loading. Data ingestion is responsible for retrieving the latest data from the data sources, while data preprocessing is responsible for cleaning the raw data to transform them into a standardized format, which is further transferred to the data loading component. The data loading component describes the transformation of data to structured storage systems, such as SQL and NoSQL databases. Finally, the data storage component represents the internal storage of the pipeline, which stores raw, ingested, and processed data. The fourth component describes the destinations where the data are finally provided.

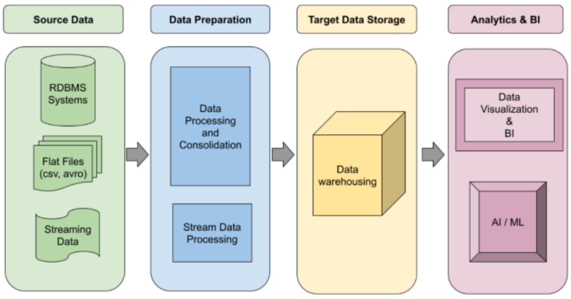

Fig 3: Data Pipeline Architecture

3.1. Components of Data Pipelines

The creation of a data pipeline requires many dealing components working together. The exact set of components varies across different data pipelines, but almost all data pipelines consist of the same main components: data sources, processing component, internal data storage component, and destinations. The starting point of every data pipeline is its data sources. Data sources are the origins of the data that are ingested into the pipeline. The second and central component of a data pipeline is its processing component. The primary focus of this paper is to develop, implement, and improve the processing component. A data pipeline connects one or more data sources to one or more destinations. The portion of the data pipeline that receives data from the source systems and stores it is called the data ingestion component.

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526
Vol. 8, Issue 12, December 2020
DOI 10.17148/IJIREEICE.2020.81208

The data pipelines process the ingested data or parts thereof and load the result into the respective sink systems, also called the destinations. The data storage component represents the internal storage of the pipeline. The modified data is stored and held within the pipeline until it is transferred to the destinations.

All data pipelines are designed to process data in a certain way based on the regulatory requirements to which the pipeline is subject. A data pipeline is always built with a processing component that takes care of the controlled process of one or multiple datasets at a time according to the needs of the organization and/or the regulations. Understanding the architecture and purpose of the processing component is crucial to develop an overview of its parts that contribute to security and regulatory compliance. Knowledge of legal requirements is essential as part of accomplishing the related tasks. However, this knowledge is only one part of the requirements for compliance that can be achieved when dealing with and processing data. In addition, the data processing component itself should be considered. Importantly, the implementation and software development definitions depend highly on the data technologies and environments. Though, many data processing features are almost the same across technologies.

3.2. Data Flow Design

The pipeline flows data from data owners to data consumers over three parties: the first party (i.e., data owners) is a data provider that generates data, the second party (i.e., third-party operator) is a single cloud service provider that acts as a data analyser, and the third party (i.e., data consumers) is an output dealer that conveys the output to data consumers. The first party encrypted sensitive data using the selected encryption scheme before sending it to the second party. This scheme performs secret sharing on encrypted data, regenerates the secret shares according to power consumptions, and securely executes analytical queries on secret shares while modulating the cost of data transferring and computing. The output is sent back to the third party over a secret-sharing channel for decryption. It also includes the security test to verify whether the scheme is correct. When building data pipelines with different parties, it requires a comprehensive discussion of the pipeline's design, setup, implementation, and security analysis through architectures, higher-level components, and runtime protocols. In this solution, four key offerings, including design tools, pipeline interoperability, pipeline schedulers, and clone debugging, are first framed by a cross-border tax consulting pipeline view. Generic component interfaces and runtime protocols are then provided to other providers. Different schedulers are discussed to compare performance for both non-residential and application-resident approaches. As data pipelines acquire increasing importance for contemporary businesses, more intricate pipelines with many components are being built and directed by practitioners. Ensuring these pipelines are executed correctly according to expectations is a considerable challenge, which is critical in applications with monetary costs or intricate regulatory compliance concerns.

An end-to-end debugging approach for such sophisticated systems is discussed. First, the impact of each component is simulated, enabling users to successively narrow down components responsible for the miscalibration. A semi-automated consideration is then supported, in which configurations that are most likely responsible for errors can be represented to the user. Implementing this solution as part of the offering toolbox is not a numerical challenge. When considering the practical need for highly interactive debugging, new and robust algorithms need to be designed that can track down errors using proxy models without confirming each calibration target every debugging iteration.

IV. SECURITY CONSIDERATIONS

Tax consultants utilize sophisticated modeling to calculate the tax-efficient structure of a cross-border transaction. The resulting calculations demand high computing capacity and therefore a number of participant's machines can be involved through the cloud. This poses a risk of information leakage because the agreed upon methods of workflow coordination have not been rigorously verified. The computations may be intercepted, corrupted or faked before they are discovered and acted upon. The concepts introduced and demonstrated in this work apply equally well to data analysis methods outside of the tax processing field. The choice of language used is not integral to the system – it could easily be implemented in Microsoft Excel, C/C++, Python or some other programming language. The participants would simply instantiate their own machines and tell the central machine how to connect with them.

The implementation focus is using R F3ildCrypt to write the cloud machine in F3ildCrypt and the tools needed by the participants in the R programming language. A major goal is to provide a highly reproducible working model. All aspects of the system from the data schema to the building of machines to the coordination language (R) are publicly available. It is hoped that this work can serve as a template for how Roda's F3ildCrypt, Roda's F3OMG, the set of open source exchange documents, and the digital notebook feature of F3frac can work together seamlessly. The format of the submission is two parts: a demonstration of the possible workflow with explanatory videos and articles.

One article details the documented construction of the implementation of the cloud side of a complete secure tax consulting workflow. The other article details the instantiation of participant tools and what an uncredentialed agent can reconstruct from an execution that they did not take part in.

Equ : 2 Compliance Audit Log Volume

$$V_{\text{log}} = R_{\text{event}} \times S_{\text{entry}} \times t$$

- V_{log}: Total log volume
- R_{event}: Rate of security/compliance events per second
- S_{entry}: Average size of one log entry
- t: Time in seconds

4.1. Data Encryption Techniques

To segregate data from the Other Party's infrastructure (ensuring data remains in controlled, compliant environments), it is continuously staged in a controlled, downloadable format; access tooling is effectively frozen, eliminating all interactions with tax computation engines and output rendering pathways. A data export execution order is to be established and, if feasible, processed in parallel. Completeness should be assured by re-importing machine-readable information, confirming tax data integrity, and assessing that no PW has inadvertently been excluded from export. Data consumed - The provided processing/audit software should compress and encrypt the on-the-fly export clones of code, executables, logs, and lastly the data files, creating a secured data process and audit trail. For verification, the on-the-fly export of "non-digital" assets is suggested to be done in plaintext. Post-execution data security - Data exported while code, executables and logs are deactivated should be scrubbed. Data exported from better-updated sources could be scrubbed with adequate encryption/tokenization in place, while unused archived improved or raw data (now not required for processing, including old tokenization settings) can remain secure in Secondary Storage unless otherwise stated in the agreement. Data structure serialization, re-enablement execution, and thus verifiability should be within the Other Party's controlled environment. For post-processed data security, contingency mechanisms covering expired licenses/settings, and overall settings reablement after export are suggested. Delivered as tailored softs - Horizontally encompassed softs shall be produced with appropriately modular interfaces. Encrypted softs shall subsist on easily updatable hardware without accretion. To complete delivery of hard copies, whatever protected by secrecy agreements should remain digital in sealed digital storage. Upon request for more thorough testing, migration laborers will quantify further considerations. Intensive performance testing will support precise time assessment and effective resource utilization.

4.2. Access Control Mechanisms

Safeguarding confidential data necessitates prudent data handling regardless of its storage method or format. Data security concerns are pivotal when data is transferred from one environment to another. Insufficient or improper handling, such as using insecure technologies for data import or export, can expose sensitive data. Generic software systems are often used for an initial import, though expanded usage raises regulatory compliance and confidentiality concerns.

Data transfer settings that collect, transfer, and process sensitive data or otherwise confidential datasets raise data protection questions that software vendors must answer. An outline of expectations based on regulatory and legal requirements followed by a processing firm, as well as data transfer specifics related to software systems, must be provided. A legal contract regarding the processing firm with specific compliance obligations is typically required, such as restrictions on technical access, mandatory records of processing and regulatory conversations, separate ISO13485 certification, and dedicated infrastructure updates.

The solution must specify how the questions posed above are balanced and mitigated, elaborating on necessary security mechanisms based on the individual circumstances of the software vendor. Security policies and protocols must be aligned with security mechanisms, specifying technologies, procedures, and practices. There must be a description of how each security control contributes to preserving the information assurance properties of the data transfer. Unlike security policies, which describe what must be implemented in an environment, security controls detail how this is achieved. There must be a discussion of practical aspects of vulnerability testing and maintenance for software security mechanisms. The aim, subject, and criteria of validation tests used in a data transfer setting must be specified.

DOI 10.17148/IJIREEICE.2020.81208

The security environment in closest vicinity to the software system that can be influenced must be discussed. Security-related shared keys or certificates, such as public encryption keys, TLS certificates, or shared AES keys, must be analyzed. There must be a list of checks that the construction and existence of these shared secrets need to conform to. All third-party services or components – systems that could access confidential data directly or indirectly – that are used when performing data transfers must be specified. The security status of all these additional services and components must also be discussed.

4.3. Audit and Monitoring Practices

Data processing and analysis of sensitive information comes with the need for an audit and monitoring process. Appropriate monitoring systems should be put in place in order to avert breaches. Auditing is the process of ensuring compliance with regards to rules and regulations, whereas monitoring is the act of watching over activities for possible violations. Consider a compliance task that verifies the flow of information between two parties in a simple model. For example, the necessary parties wish to ensure that data sent across and beyond a boundary is properly protected. In the simple model, every relation R says who sent information to whom. There is an additional relation T, which says that a means of communication M is prevented from exposing S to T, for example by encryption.

When checking compliance, the specific parties would like to know whether there is a relation R that describes a leakage of sensitive information S to undesired recipient T while preventing that leak with means of communication M. Summarising, there are rules, but it is ensured that parties are complying with them. However, there is a cost incurred for the data storage that records captured audit data which, depending on the amount of sensitive data flow that needs to be tracked, may prove significant. Fortunately, it is realistic to expect the adoption of techniques similar to those described in this paper.

To assess compliance with the purpose of its use, a detailed representation of relevant data handling must be formalised. In order to enable the information flow analysis with a set of declarative graph query languages, an approach is based on extending existing assurance certification representations with information flow. The query works by the intuition behind, the base of how querying an Information Flow Audit graph may be used retrospectively to verify compliance with regulation is introduced, along with a use case derived from a CNIL report describing best practice for smart metering services for electricity supply, including those mediated by the cloud.

V. CROSS-BORDER DATA TRANSFER

In a modern society increasingly dependent on the collection and cross-border transfer of data, concerns are raised regarding the possible harms caused by the misuse or negligent handling of such data, resulting in significant interest in the imposition of regulatory measures on data controllers and processors. As privacy and data protection regulations are drafted and implemented by jurisdictions around the world, they often contain provisions specifically regulating the transfer of personal data from one jurisdiction to another, more generally referred to as cross-border data transfer. These provisions, however, are young and largely untested in practice, resulting in uncertainty over both their interpretation and application. While important, this uncertainty is particularly pronounced in jurisdictions where data protection and transfer regulations are still in their infancy, or for business and industry sectors that are not ordinarily subject to data protection regulation, such as those engaged in legal or tax consulting.

Transfer pricing regulations are now being introduced worldwide to implement the two Pillars of the BEPS 2.0, a project to reform international taxation. In order to comply with these tax regulations, cross-border transfer pricing documentation and data regarding related party transactions are often required for tax audits or for filing in a tax return. However, these cross-border data are often subject to local data protection and privacy regulations, or data transfer laws, which restrict sending this data outside the jurisdiction. As related party transaction data are personal data containing sensitive information about taxpayers, data protection issues in cross-border transfer may arise. The risk of non-compliance is a heavy burden for tax consultants and multinational enterprises.

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526
Vol. 8, Issue 12, December 2020
DOI 10.17148/IJIREEICE.2020.81208

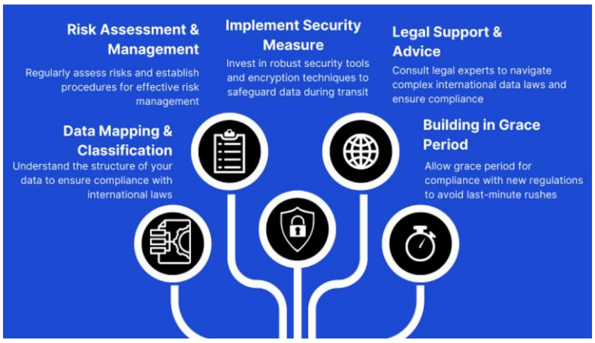
Risk Assessment & Management
Regularly assess risks and establish procedures for effective risk management
Implement Security Measure
Invest in robust security tools and encryption techniques to safeguard data during transit
Legal Support & Advice
Consult legal experts to navigate complex international data laws and ensure compliance
Data Mapping & Classification
Understand the structure of your data to ensure compliance with international laws
Building in Grace Period
Allow grace period for compliance with new regulations to avoid last-minute rushes

Fig 4: Cross-Border Data Transfer

5.1. Legal Implications

In recent years, compliance has become a critical issue and a hot topic of research. Compliance refers to the conformity, adherence, or obsequiousness of an enterprise's adherent evaluation, procedures, regulations, or laws. Compliance ensures adherence to different rules and regulations, guidelines, or predefined legal prescriptions like laws, norms, standards, and guidelines. Compliance constraints on business processes and activities ensure those processes or activities are in conformance with the laws and regulations. Non-compliance subjects the business organization or enterprises to different impacts like financial fines, litigation, and loss of corporate image. With the development of e-business and tight cooperation between virtual organizations, collaborative business processes, in which several enterprises work together towards a common goal, have become essential for enterprises to enhance their abilities and competitiveness. However, collaborative business processes are also a difficult issue for the enterprises due to their complexity and difficulty in compliance. Cross organizational collaborations imply the compliance with the internal regulations of every enterprise involved. Cross region collaborations also imply that the implemented processes comply with the regulations of all the regions involved in the collaboration.

In a tight cooperation between several suppliers working together to produce a product for the manufacturer, it is critical for the manufacturers to know the suppliers' process and ensure the correctness of the input data supplied by them. When bad input data is supplied, it can be very difficult to track errors back and hold someone responsible. In this cross-border collaboration, the enterprise has to ensure that all its controlling processes are working correctly to prevent and trace non-conformance of the input data from the supplier enterprises. In addition, this cross-border collaboration is also involved in extensive data sharing and data mining, for example, the cross-border sharing in collaboration with customs. European enterprises have to comply with the EU data privacy regulation and the implementation laws and regulations of each member state. Hence, when deploying and implementing a collaborative business process system shared over different enterprises and regions, business processes have to be checked to comply with the internal process regulations and the regulations of every region involved.

5.2. Data Localization Requirements

A significant obstacle to cross-border tax consulting are the obligations on organizations to store and process some or all data within the country. Generally speaking, the legal restrictions on data locations are known as data localization requirements. The accurate definition of the data localization requirements supports the selection between provided solutions most effectively.

A robust definition of the term “data localization requirements” is that the governmental restrictions are imposed on storing and processing some or all data in the country. Four primary variants of data localization requirements are the following: 1. Local Storage and/or Processing: The data must remain in the national territory throughout its whole lifecycle. This version of the data localization requirements is the purest one. It completely separates two countries on the level of data types, as one kind of the data would leave the domain of another country permanently. 2. Local Control: Although some data may be temporarily moved abroad, the surveillance agencies must be able to access the data located at foreign servers physically or through technical access to monitor it. 3. Local Knowledge: The organizations intending to compute data abroad must own local knowledge on these data. Although data still may be physically stored abroad, the companies are prohibited to send the breed of data that requires the state license abroad. Therefore, the previous two variants require the mandatory presence of a local cloud in this country where surveillance could be conducted. 4. Local Activities: The separate data use processes may not be conducted beyond the national border.

IJIREEICE
International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering
Vol. 8, Issue 12, December 2020
DOI 10.17148/IJIREEICE.2020.81208
ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

VI. TECHNOLOGY SOLUTIONS

Technology has a significant role to play in the regulation of data processing, whether through established rules or emerging regulation through local supervisory authorities. The core question is how to align technology solutions with the legal requirements governing data protection and privacy to prevent, monitor, and detect processes that breach the law. It is expected that a significant acceleration of regulation will occur in various jurisdictions over the next few years towards privacy and protection. However, tech solutions are emerging alongside such regulation, especially in the wake of the ongoing AI wave. The question is how to mitigate overspending, especially in the case of duplicative compliance solutions that only address parts of regulatory stipulations.discusses the evolution of technology and the consumerization of information in the new data processing environment, introducing numerous challenges that can affect companies' market share and the confidence that investors have in that company and the nation as a whole. To better understand the problem situation, the resultant queries, and some of the possible solutions, a framework containing seven main components is considered: society, laws, regulations, the industry, companies, people, and technology.

6.1. Cloud-Based Solutions

As a result of data protection laws, cross-border internal and external data flows are restricted. In order to guarantee compliance with regulations, cross-border data flows analyzed, potential risk mitigation actions proposed, and the processes necessary for their implementation outlined. Security and regulatory risk analysis activities have been taken into account to make sure that the findings and recommendations support the data protection law compliance requirements as needed. This purpose is to contribute to the on-going discussion of incorporating privacy by default into the entire lifecycle of the solution design and implementation. Specifically, building secure data pipelines using cloud services and data storage for international organizations. The resulting design and components enhancing European regulation GDPR compliance are shown. Methods on how to demonstrate the compliance of data handling on public cloud services, and data storage, Information security and data handling process, will become secure by design and by default.

Cloud-based applications and IT services running in the cloud can significantly reduce infrastructure costs while providing similar or even better performance and security. They can also dramatically decrease the time before implementation of a generic solution. Compared to on-premise systems, however, cloud services offer only high-level control and management interfaces. To keep sensitive data safe while enabling the analysis of such data, privacy and security relevant controls are vital. Regulations and laws require strict handling of sensitive personal and organizational data. Violation of such regulations and laws can have significant financial and reputational impact. Adversaries obtaining sensitive data can severely harm the business of organizations.

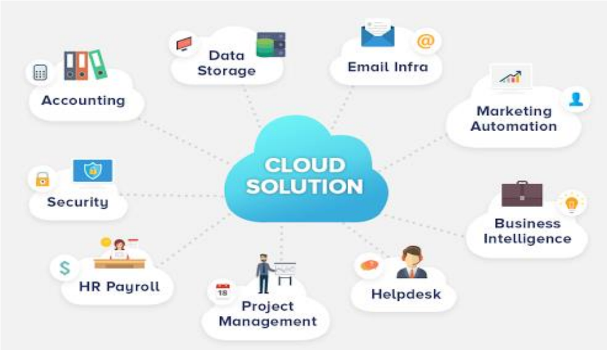

Fig 5: Cloud Solutions

Equ : 3 Data Encryption Overhead

$$T_{\text{total}} = T_{\text{base}} + T_{\text{enc}} + T_{\text{dec}}$$

- T_{total}: Total time for data processing
- T_{base}: Time without encryption
- T_{enc}: Time to encrypt the data
- T_{dec}: Time to decrypt the data

DOI 10.17148/IJIREEICE.2020.81208

6.2. On-Premises Solutions

While cloud solutions have their benefits, some organizations struggle to trust cloud service providers in terms of data protection and are thus bound to consider on-premises solutions featuring a greater guarantee of control over vital data. Such solutions do exist, although it still remains necessary to provide ways to demonstrate compliance similar to those already mentioned for cloud solutions. Such on-premises establishments would consist of a local cluster of storage devices, sometimes managed as storage area networks, to run the compliant binary of the storage system. In such a solution, compliance demonstrability goes through auditability facts relying on trusted hardware craftsmanship of the designated nodes implementing GDPS functions and an evaluation of the chain of trust of their endpoints. Compliance verification would proceed through a challenge-dependent digests of outer-tree and header integrity preserved by hash-locking mechanisms. Direct audibility to data plane procedures as per cloud solutions means the entire proof of eligibility process is essentially similar, with the notable exception that external attestations would not need to be managed since they concern inner nodes operating on untrusted storage clusters. There is also no need for any claims to be checked or processors to furnish digests from inner nodes in this case since compliance with EUD regulations is inherently built into such nodes. This alternative design facet also faces its own share of challenges relating to privacy and accountability. The massive overhead involved in data verifications would limit the number of verifiers to a handful of authorized requesters, disallowing support for situations with many transient verifiers. An alternative solution must be sought, as it will be too costly and cumbersome to physically audit storage systems containing data of millions of individuals similarly managed by the cloud. In such scenarios, compliance is generally addressed through auditability mechanisms designed to furnish external proof of compliance via a third party.

6.3. Hybrid Approaches

In tax consulting at MNCs, one of the main challenges is how to deal with the highly sensitive data which needs to be transferred across borders. With the change of regulations in data protection, there are many approaches and technologies available to address this problem. Many of them are already applied by banks, providing cross-border payment solutions that protect the privacy of their customers. However, these solutions are too expensive for tax consulting, whose margin is already thin compared to banks. Guidelines and recommendations are standard approaches but often of limited effect, especially if no official judiciary regulation is in place. This part elaborates on hybrid solution approaches that combine new technological approaches with compliance and consulting steps to achieve an acceptable immediate solution after which the business can adapt sufficiently to the regulations. As in every complex situation, first a solid understanding of the basics needs to be provided. This will include regulations that are really relevant, systems in place to handle the data and pipeline options that are already available or could be used in the future. The presentation of the findings will have to be adjusted to the audience. If the interviewees are heavily involved in a particular system, the interviewees need to focus on other systems to avoid overestimating the current knowledge. Based on the understanding of the context in which the challenge exists, security and regulatory perspectives need to be identified. Which data is a potential target for attacks? What are the regulations governing the data and threats posed to access the data? A high-level overview will be sufficient here, as the details will be elaborated in the next step. The general context may influence the relevant regulatory and security perspectives significantly. Additionally to the general context and potential threat related to the data in question the specific context in which the data will be applied needs to be understood. Here, the business and regulatory consequences of improper data usage will be elaborated on. This ensures a comprehensive understanding of how to shape the compliance strategy and go through the compliance processes, before the strategy is aligned with the enablement of tools and systems to ensure compliance. After these steps, knowledge gaps need to be identified. What information on regulations affects the pipelines? What systems in place are capable of safeguarding compliance? Which tools and technologies are in place and how capable are they of safeguarding safety? This step is essential as knowledge gaps first of all need to be filled before a compliance strategy can be drafted.

VII. BEST PRACTICES FOR DATA PIPELINES

Data pipelines are increasingly used to transfer, aggregate, and analyze data, especially since the emergence of data-intensive domains like Internet of Things and Big Data. Pipes assemble processes responsible for collecting data from sources, transforming it, and exporting it to sinks in a streamlined manner. The design process involves tool selection and configuration, pipe design, and maintenance. Quality assurance is challenging, with no established practices widely adopted in the industry despite quality-related research.

Pipes, used in computer programming and Unix sheets, work in tandem as a chain, feeding the data output from one pipe to the input of the next. The lower a pipe’s layer, the more primitive the operations that happen there. Pipes read and wiretap data frequency bands before describing frequency limits to their effective band-passed versions for batch and

DOI 10.17148/IJIREEICE.2020.81208

stream signals. Robustness to pipe design issues like coder format mismatches and data types is essential. In computer science, code stability concerns how semantics and language changes affect systems in graph-like form.

Most advertising nowadays is targeted, exploiting user tracking, databases, algorithms, and visualization tools. However, privacy intrusion issues have arisen. Generally, designing a data-intensive system is a complex task requiring a myriad of design decisions and activity choices. Its success relies on semantic knowledge about the processes and their interplay. Learning this knowledge involves training in formally defined abstractions with inherent limitations.

Multiple industries manage pipelines of processes producing and using data. In addition to problem-aware agents, for economically viable solutions, automated tools must be created. DSLs are program languages providing abstractions for tasks of a specific domain that are not necessarily natural languages or textual. Considerably more general design forms bring more productivity. Generalization techniques include equivalency-holding reasonings and planning methods.

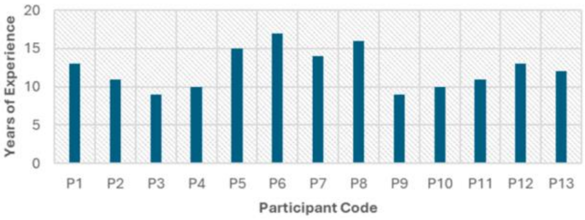

Fig 6: Regulatory compliance in cross-border tax consulting

7.1. Data Governance Framework

In the Accelerated Data Pipeline Development phase, a Data Governance Framework and related Governance and Monitoring tools have been designed and prototyped. The Design Matrix has created a multi-level framework defining how and where data governance can be dictated, controlled and executed, with the latter including parts on Data Domain Owners and related Monitor and Accountability routine. It is based on a model successfully piloted at the bank. Governance needs good tooling, as losing track of responsibilities, or failing to automate will hurt analysis and monitoring effort. Governance tools need to offer a complete picture of the data governance in place including accountability, transparency, discoverability, lineage, and specification. There is a conscious focus on on-premises tooling, given security and information risk concerns.

Monitoring is needed to oversee compliance with the governance and its desired outcome, e.g., quality, operational performance. Usually, documentation fails to achieve this due to missing updates or data usage changes not being tracked. Compliance checks can be built into the code at both the pipeline and governance tool level, offering the potential for real-time compliance checking and reporting. Monitoring tools should offer a comprehensive picture of compliance across all data assets against all checks.

Data Governance is a critical source of risk in data-centric industries, particularly given internal and external regulatory pressure to control the data used. A multi-level framework for governance covering definition, design and deployment of governance and related monitoring is presented and demonstrated. Given competitive investments here, tools are being kept confidential. The Framework has been successfully used in the Banking and Asset Management sectors. Future work relates to stress-testing the Governance Framework and related tools and new areas of application.

7.2. Continuous Compliance Monitoring

Compliance constraints are rules, standards, laws, and regulations that constrain processes. Processes not compliant to the constraints must be rectified. Regulatory non-compliance subjects enterprises to heavy litigation and monetary fines. Business processes are subject to compliance constraints. Collaborative business processes must comply with internal organizational policies, industry standards, and cross regional regulations. European enterprises must comply with two regulations: the EU data privacy regulation and each individual member state’s data protection laws. If non-compliance is found, compliance violations must be either prevented upfront or detected and rectified. Preemptive compliance verification is essential to the successful deployment and implementation of collaborative business processes. Compliance has two perspectives: the “universal” perspective with compliance constraints as a first requirement, and the “specific” perspective that confines a domain/region and thus its regulation.

Continuous compliance monitoring assesses process executions and the respective data artifacts with respect to compliance constraints. Compliance violation cases are produced together with evidence. The needs of corrective actions are prioritized. Monitoring rules are used to define compliance checking queries. These monitoring rules are extensions of the CEP language Correlation Query Language. Executions of specified processes are persistently analyzed. Aggregate information is continuously computed in a manner similar to materialized view maintenance and is employed to validate compliance constraints. Noncompliance is assessed using temporal logic rather than producing low level traces. Instead of hardcoded queries, a temporal query language is defined and can be used in the analysis. In cross-border situations with high compliance scarcity, collaborative business processes are adapted to federated systems embracing compliance verification services. Federated formalism for semistructured documents based on a general type hierarchy and shared concepts is defined, which facilitate the exchange of compliance constraint languages. A new formalism is developed to capture a broader categorization of compliance constraints from legal requirements, and a framework is proposed for a language consistency checking that can be used to detect bugs in compliance constraint languages before deploying the languages.

VIII. CASE STUDIES

In this section, two case studies illustrating tax consulting-related GDPR challenges from a regulatory compliance standpoint are provided. Both cases also describe how the design science approach can be utilized to produce potential solutions. The first illustration focuses on a situation involving an information datamart housing audit-related personal data. The second example discusses a project at a multinational consulting firm involving the design of a custom platform enabling compliance with cross-border country-by-country reporting regulations. These cases will be presented in a narrative format and are fictitious yet illustrative. However, they are sufficiently grounded in reality based on the author's experience in the field.

GDPR-related questions have arisen regarding the data mart and how personal data can be classified as “data protection by design and by default”. As a consequence, it was hindered from being taken out of the data mart by the regulator. A discussion occurred regarding whether the data could become “fair” if the purpose for processing it were changed from an audit to another, perhaps commercial purpose.

In the early days of regulation, a temporary solution was proposed by a design team. A nightly extraction process would downgrade the information from data mart A to B with pseudonyms persisted in an anchor table. Analysis rules resulting in reporting sets were to be moved to the new data mart as well. The old data mart process configuration needed to be copied in the new setting and adjusted to the new structure. A new design-conform reporting front was to be created. In this framework, very few grounds for re-identification remained. For authorities, continued access to the audited data mart would facilitate understanding of how data protection rights were respected, but they would need to be scrapped from mass analysis to solely follow up on individual audit trails. This would certainly complicate matters and err on the side of caution, but it would at least provide time to work on a sustainable solution. A chronology of descriptive events related to the recent past was crafted — this being akin to an enterprise architecture view type — so that the older and more explanatory analyses could reside in the old primary repositories.

Regulators later suggested that unless tangible forward-looking endorsements were in place to assure that the downgrading scheme would remain compliant, the whole downgrading logic would need to be quarantined, together with the data mart and the coding and all possible unwanted artifacts, until submission of “compliance as of” regression tests, or total ban of both machines and data mart contents.

8.1. Successful Implementations

In the past three years, a number of organizations have engaged in successful implementation of the design time approach. Based on strong academic foundations, this research has matured via several research projects. The development methodology, key algorithms, a process mining method for searching event logs for violations, and a variety of application domains have been published. As a consequence, compliance verification has matured from a pure proof of concept stage towards maturity in both comprehensiveness and usability. However, the need for adjustments and further enhancement is still present, similar to several other areas from other research. Nevertheless, the first implementations have been put into operational use.

An implementation in one of the largest Dutch NGOs is used for ensuring compliance to and monitoring of rules regarding trust, safety, and non-violation of human rights and privacy of minors in youth care services provision to minors.

DOI 10.17148/IJIREEICE.2020.81208

Another Dutch customer is a larger and rapidly expanding international consultancy organization which is using compliance verification in the area of compliance to and monitoring of routines regarding the cutoff of non-conformations regarding thermal energy provision. An implementation at the Dutch Ministry of Finance is being finalized. This implementation is focused on compliance verification of the data model of the national tax information system regarding adherence to rules regarding data protection status, security levels, and data sharing possibilities. A later implementation is focusing on compliance to data protection rules of tax assessing processes regarding involvement of domains, jobs, and roles.

A fully operational implementation is in use for the academic researchers of the Data Protection Research Group from the Department of Information and Computing Sciences at Utrecht University. For these researchers, the software prevents much time- and resource-consuming violations of the security of several hundreds of thousands of observations in extensive tax data sets through their intention elicitation task by means of being applied on the resulting annotated process model before demonstrating these process models in a widely used processware program.

8.2. Lessons Learned from Failures

In times of crisis, chaos is usually the main culprit. Reforms intended to maintain order only give ground to new power structures. After the Second World War and the Great Depression, the United Nations, International Monetary Fund, and World Bank all aimed to build an institutional order to ensure against any similar crises. Economies saw the emergence of institutions that others had to obey, just as handshakes had obliged the kings and emperors of old. The United Nations, the United Nations Security Council, and the World Trade Organization primarily keep peace. The International Monetary Fund and World Bank promote growth and development. Financial orders emerged with the Basel Committee on Banking Supervision, Financial Stability Board, Financial Action Task Force, and International Accounting Standards Board. The agencies intended to maintain lawfulness in these differing fields strive against each other for influence.

Each order attempted to ensure compliance by denying access to the benefits. However, comparable to the ‘rogue states’ disregarding formal authority, other institutions found a way to continue, sometimes at greater risk of hurting the system than before. This created a game of cat and mouse where those trusted to develop the rules misunderstood and ignored them until tragedy struck again. Thus, the agencies’ legitimacy gave rise to a string of compliance programs initially meant to prevent failures, regulatory acceptance of accountability without oversight, and a lack of promotion or reward of trustworthiness beyond hope. When agencies depend on self-regulation, legal authority disappears, although necessity needless repetitions beget acceptance. By then it may be impossible to modify expectations, while systems become increasingly opaque.

IX. FUTURE TRENDS IN DATA COMPLIANCE

The continuous rise of data security breaches means that data compliance will become paramount over the coming years. The recent rise of large Language Models (LLMs) is further accelerating data compliance checks since companies will increasingly have to validate the security of sending their data to cloud implementations of LLMs. Unfortunately, several countries have law rules concerning the type of data that can leave the national borders and the permission required for non-nationals to gain access to business-critical or sensitive data. Consequently, building a global tax consulting infrastructure based on these AI technology trends is still ahead for the regulatory challenges that come along with it. Large global firms have long been incapable of delivering compliance-controlled solutions guiding the stakeholders of the data pipelines designing processes that will deliver regulatory-compliant data movement. Current implementations are inconsistent, lack maintainability, and are extremely difficult to use. A new data compliance framework that defines a taxonomy of compliance regulations, constraints modeling a compliance verifiability approach, and finally compliance visualizations was proposed.

Nevertheless, with changing regulations and business models, it is required that compliance models and visualizations can advance as well. The current approach precomputes compliance checks for a static set of analysis scheduling. However, it is unclear how a compliance taxonomist can effectively maintain this compositional approach. Practical feedback from cross-border tax compliance consultants was incorporated to iterate towards more user-friendly views that remove redundancy. Novel compliance challenges due to AI technologies were documented and a regulatory taxonomy originated evolving the existing constraints and verification approach. Implementing a prototype framework supported both the technical extensions and the non-technical visualizations as a proof-of-concept application interface demonstrating the operational effectiveness of the derived use-cases.

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

DOI 10.17148/IJIREEICE.2020.81208

Finally, evaluation of the versatility of the framework across a range of different use-cases and the ability to maintain compliance constraints over time was sought.

It is striking that compliance has historically not been an agenda point of software systems while its pervasive inconvenience to organizations, regulators, and society was established. The data-driven transformation of institutes and businesses has amplified this. However, with the dawn of new technologies automating the entire data-compliance analysis pipeline from regulatory text to data visualization qualified standards are required. A formal regulatory modeling approach provides organizations the tools for domain-driven creation of compliance checks. Formal data-mapping capabilities can independently transfer and process data-pipeline designs implementing compliance checks. Considerably reducing the burden of compliance analysis while vastly improving the coverage and processing of analysis is possible.

9.1. Emerging Technologies

Emerging technologies, such as machine learning (ML) and AI, are increasingly being integrated into the existing work practices and software of professional services firms. In the tax sector, emerging technologies may be subjected to scrutiny under the BCRs proposed in the prior sections, especially section 5.2. In addition to these sections, model risk assessments should also be considered. Designing and implementing an AI/ML system is a multi-step process that includes – but is not limited to – data preparation, model training and evaluation, model deployment, and model monitoring. Data preparation is most often the most time-consuming part of a data science project and is critical in determining a model's performance. Relevant regulatory requirements for tax consulting firms in the EU and beyond can also impact the nature of the model risk assessment. While ML models are often referred to as black-box models, the realities surrounding their necessary production and ongoing monitoring practices indicate that they may not be fully opaque. With relevant consideration, an understanding of ML systems can be achieved. A model risk assessment examines the potential harm that the use of a deployed model may cause, and often involves a more technical examination of the inputs and outputs regarding the processing of data such as gradients, and feature importance. Due to the different foundations of some emerging technologies, they introduce broader systemic considerations beside those that are particularly pertinent to legal compliance in cross-border tax consulting. Regardless, consideration of the unique risk scenarios they introduce can help frame the discussion of firm responses, service shortcomings, and other developments pertinent to the ongoing and future implementations of those technologies. Methodologies used to anticipate looming threats posed by streaming financial information, news articles, and the evolution of geopolitical situations help frame the emergence of many technologies that have already transformed the description of regulatory events in compliance technologies.

9.2. Evolving Regulatory Landscape

The increasing digitization and automation of business processes has led to a proliferating use and flow of data assets between organizations. Sharing sensitive data is, however, often regulated. For instance, in processing business data, European enterprises must comply with the EU data privacy regulation and each EU member state's data protection laws. These regulations prohibit the admission of data from non-compliant organizations and countries. To deploy a data pipeline in accordance with such regulations, a multi-actor compliance specification must be designed that understands complex compliance needs covering the technicalities of querying and accessing data across organizational and legal borders. In order to learn from the sharing of sensitive data, compliance verification is essential on a proactive design-time level. The subject of this article is the design of client-side abstractions for verification methods tackling compliance specification dynamics. At the core of this abstract data type is a globally convergent high-level method that obtains and refines needed compliance specifications, subsequently enriching verification methods with that compliance knowledge base.

The terms regulatory landscape, compliance checking, and legal compliance cover a multitude of actors, possibilities, and interpretations. It has become clear that regulatory compliance is a central responsibility of organizations. A proactive compliance strategy refers to compliance processes that occur at design time prior to changes within the regulated process or organization, therefore dealing with design-time compliance. Taking a qualitative approach means to focus on ideal world models, emphasizing what ought to be (rather than what is or can be). A framework is a high-level description of methods, capturing important aspects of reasoning about compliance processes. As the study of regulatory compliance is broad, the framework and proof principles only focus on a restricted model. The primitives involved are legal documents, regulations, and compliance models but this can be expanded with unratified regulations and regulated processes. It's important to observe that for every organization compliance with legal obligations (completeness) is a requirement. Compliance with best practices is of higher risk and depends on the organization.

X. CONCLUSION

This article highlighted a risk-based data pipeline design framework to manage data subject rights in a cross-border compliance demanded work environment. This comprehensive framework was aligned with various compliance scopes and adapted to privacy experts' objectives. Combining the principles of cross-border data transfer and data protection by design, the framework emphasized governed data processing pipeline composition and user rights-focused processing plans. Beyond compliance analysis, the framework also facilitated systematic processing plan adjustment. The proposed framework was further demonstrated via a realistic sub-architectural messaging queue use case.

Privacy law touches multiple aspects of contemporary life and business. Internationally, the related compliance ends are far from agreement, given more than 140 data privacy and protection (DPP) regulations. Data transfers create an explosion of cross-border processing scopes. It is evident that no single compliance framework will fit different regulatory regions. Wider compliance examination of cross-processing pipelines with heterogeneous scopes is a common pain point for privacy experts in multi-context privacy assurance. To this end, attention is needed to enable privacy experts in compliance assessment across different scopes, collect metrics spanning heterogeneous scopes, and use federated checking tools.

Privacy automation and compliance-aware architecture design were fruitful, but the effort largely narrows down to a specific compliance framework. Wider compliance examination still requires formal understanding on broader scopes as well as tools tailored to compliance experts. With a more prospects' driven design story, connecting the top-down scope-centric formalism and automated model-checking tools would be valuable to enable tiered analysis on wider aspects of privacy engineering. It would be more general to base the framework on ensemble theory rather than on top of a model-checking tool, i.e., coupling the precision of the operator formalism and the generality of reasoning on unfinished/simplified design units with gap-filling and forwarding with Discovery or will-be-designed tools.

Shifting the scope from automated checking to general compliance assurance across heterogeneous systems will also be meaningful to integrate with wider automation efforts, such as compliance network design, quantitative reasoning, and DPP law text understanding.

10.1. Future Trends

Regulation of the environment and way of life of people and organizations has become increasingly difficult with rapid developments in multimedia, wireless networks, mobile communications, and cloud computing. Financial data travel across borders and cross multiple jurisdictions. Technologies reduce the cost of switching, provide incentives for this behaviour, and make it difficult to monitor. An important consequence of this is regulatory competition, referred to as the race to the bottom and the evolution of 'legal vacuums'. A number of ways have been proposed and commenced to mitigate these problems, including the development of enhanced coordination between different regulations, establishing global standards for the regulation of financial services, and attributing secondary jurisdiction through developed enforcement mechanisms. However, these systems may conflict with a number of domestic values and are still developing, and provide limited guidance to the challenges in complex settings where more than one regulation governing records and related data apply. Data processing applications act on records and associated data and, by virtue of their actions, are governed by related regulations. Cross-border data traffic and the complexities of the application context produce situations in which different laws and regulations govern audit stewardship and related data actions. Explicitly representing the qualitative semantics of compliance makes it possible to reason about and assess compliance with regard to the associated regulations. Interoperable, extensible and scalable compliance reasoning mechanisms make the solution applicable in practice.

Compliance across borders often requires models of the regulations and an ability to compare the models. This presents a computational model of regulatory compliance for data processing applications. The model represents, formally and explicitly, the compliance conditions imposed by a regulation, in terms of requirements on the application, data, its usage, associated responsibilities and liabilities, and the actions of a regulator. Any data processing application can be represented in terms of states, actions and input/output observables. Implicit observables and behaviours though sufficient to specify a compliant or non-compliant behaviour may be unrepresentable. Adapted to differing, potentially conflicting regulations, the reasoning mechanism may give conflicting views of the application behaviour.

REFERENCES

[1] Karthik Chava, "Machine Learning in Modern Healthcare: Leveraging Big Data for Early Disease Detection and Patient Monitoring", International Journal of Science and Research (IJSR), Volume 9 Issue 12, December 2020, pp. 1899-1910, https://www.ijsr.net/getabstract.php?paperid=SR201212164722, DOI: https://www.doi.org/10.21275/SR201212164722

[2] Data Engineering Architectures for Real-Time Quality Monitoring in Paint Production Lines. (2020). International Journal of Engineering and Computer Science, 9(12), 25289-25303. https://doi.org/10.18535/ijecs.v9i12.4587

[3] Vamsee Pamisetty. (2020). Optimizing Tax Compliance and Fraud Prevention through Intelligent Systems: The Role of Technology in Public Finance Innovation. International Journal on Recent and Innovation Trends in Computing and Communication, 8(12), 111–127. Retrieved from https://ijritcc.org/index.php/ijritcc/article/view/11582

[4] Mandala, V. (2018). From Reactive to Proactive: Employing AI and ML in Automotive Brakes and Parking Systems to Enhance Road Safety. International Journal of Science and Research (IJSR), 7(11), 1992-1996.

[5] Ghahramani, M., Qiao, Y., Zhou, M., O'Hagan, A., & Sweeney, J. (2020). AI-based modeling and data-driven evaluation for smart manufacturing processes. IEEE/CAA Journal of Automatica Sinica, 7(4), 1026–1037. https://doi.org/10.1109/JAS.2020.1003114