

Generic Solution Architecture Design of Regulatory Technology (RegTech)

Benny Firmansyah, Arry Akhmad Arman

School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia

Article Info

Article history:

Received Feb 20, 2023

Revised May 4, 2023

Accepted June 8, 2023

Keywords:

RegTech

Regulatory Technology

Regulatory Compliance

Solution Architecture

Generic Solution Architecture

ABSTRACT

Regulatory Technology, or RegTech, uses new technology that assists the financial industry, such as FinTech and banks, in meeting regulatory compliance. RegTech automates various regulatory compliance activities that were previously manual, such as regulatory interpretation and regulatory reporting, amidst the challenges of the increasing volume of regulations and operational data. Some cutting-edge technologies discovered at RegTech include big data analytics, artificial intelligence, machine learning, robotic process automation, and cloud computing. Although very dominant in the financial industry, RegTech solutions have the potential to be applied in other regulated industries besides finance. Several studies have explored the potential for applying RegTech in industries other than finance, such as charitable organizations, real estate marketplace, pharmaceuticals, and healthcare. Therefore, this study aims to design a generic RegTech solution architecture so that it can be adopted and applied in various regulated industries achieve regulatory compliance more efficiently. Based on the evaluation results, the proposed architecture can be applied in an industrial environment other than financial to be considered generic. Furthermore, an evaluation of the comparison of regulatory compliance business processes without and by implementing RegTech can produce a time efficiency of 95.16%. These results show that RegTech solutions can achieve regulatory compliance more efficiently.

Copyright © 2023 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Benny Firmansyah,
School of Electrical Engineering and Informatics,
Institut Teknologi Bandung,
Bandung, Indonesia.
Email: 23221114@std.stei.itb.ac.id

1. INTRODUCTION

1.1 Background

The banking world experienced significant changes after the financial crisis in 2008 [1], [2]. Almost all countries in the world strengthened their economic and regulatory policies significantly to prevent post-crisis risks from occurring [1]. Von solm [1] stated that the financial industry faces challenges due to the volume and complexity of regulations that must be followed. At the same time, digital transformation has also occurred in the financial industry, including banking, with the emergence of Financial Technology (FinTech) companies and applications [2]. The rapid evolution of FinTech is driving increased operational data and the emergence of new, more challenging risks such as fraud and cybersecurity [2]. Butler and Brien [3] stated that more than 50,000 regulations were issued between 2009 and 2012 at the G20. One of the major impacts of these challenges is the costs incurred to achieve regulatory compliance [1], [3]. Von Solms [1] also stated that based on research results from "The Trade", banks spent over 100 billion dollars on regulatory compliance in 2016. These costs will continue to increase because the process is still manual to comply with regulations, such

as employing hundreds or even thousands of professionals in compliance and risk including regulatory interpretation, compliance analysis, monitoring and reporting of compliance [3].

The procedure of reporting data, which must be done regularly and given to the supervisor, is an instance of this compliance requirement [1], [3]. The amount of data that must be reported and the number of regulations that must be followed continue to grow steadily, making this reporting procedure extremely difficult and time-consuming [4]. According to Von Solms [4], several obstacles must be faced, including using human resources to understand and interpret every existing regulation. This procedure could be more efficient because the number of regulations, both new and revised, continues to grow [4]. Furthermore, training human resources to understand and interpret regulations incurs additional costs [4]. This method also does not guarantee that the results of the understanding can be applied accurately and maintained over time [4].

A new solution developed to simplify the management of regulatory compliance in the financial industry is Regulatory Technology or RegTech [5]. Buckley et al. [6] stated that RegTech originally evolved to address regulatory challenges in the financial system through new technological innovations in FinTech. According to Arner et al. [7], RegTech uses new monitoring, reporting, and regulatory compliance technology. Automation processes help identify risks and comply with regulations better and more accessible. In addition, RegTech digitized the still manual reporting and regulatory compliance processes carried out by the financial industry [8]. RegTech implements several new technologies, such as big data analytics, natural language processing (NLP), robotic process automation (RPA), distributed ledger technology (DLT), artificial intelligence (AI), cloud computing, machine learning (ML), and application programming interfaces (API) [4]. These technologies can be applied separately or in combination according to the service needs to be met [4].

The application of RegTech is dominant in the financial industry because it was born in this environment [9]. However, it has the potential to be applied to other industries or entities whose operations are regulated by regulations and supervised by supervisors not only in finance [9]–[11]. It is also supported by Johansson et al. [12] by stating that one of the characteristics of RegTech is providing intensive services for various industries following the regulations that govern them, not only in the financial sector. Several studies have examined the potential for applying RegTech to industries or regulated entities other than the financial industry, such as real estate marketplace [13], charitable organizations [14], healthcare [15], and pharmaceuticals [16]. Therefore, a solution is needed that supports the application of RegTech in various industrial environments, not only in the financial sector.

1.2 Research Problem

RegTech solutions have the potential to be applied in a variety of industrial environments or regulated entities that require strict monitoring and reporting other than financial industry, such as health [15], pharmaceutical [16], real estate [13], and charitable organizations [14]. In addition, RegTech solutions can also be applied to non-profits or government agencies. It is also supported by Johansson et al. [12] by stating that one of the characteristics of RegTech is providing intensive services to various industries following the regulations that govern them, not only in the financial sector.

Currently, research related to RegTech is still very dominant in the financial sphere [9], such as the use of semantic technology in RegTech for the compliance reporting process in the banking industry [3], the application of RegTech to prevent money laundering [17], the application of RegTech in online financial services [2], and the application of RegTech for regulatory management so that support the management of treasury activities as well as demands for reporting as a form of compliance in the banking industry [1]. Therefore, to support and realize the application of RegTech in various industrial environments, it is necessary to analyze the specificity of RegTech in the financial industry, what the industry needs in general, and the more generic solution architecture.

Based on the problems, challenges, and several studies carried out by various parties and places, until now, there has yet to be an answer to the RegTech generic solution architecture [9], [10]. Therefore, this research was conducted to fill this gap. So, the problem discussed in this study is how to design a generic RegTech solution architecture. The architecture can be applied across various industries or regulated entities to achieve regulatory compliance more efficiently.

1.3 Objective

To answer the previously defined research problem, we determined that this study aimed to design a generic RegTech solution architecture. Solution architecture deals with the design and definition of IT solutions so that these solutions can be implemented, used, operated, and supported safely and efficiently [18]. The proposed architecture was designed by analyzing the specificity of RegTech in the financial industry, and the industry needs in general. In addition, the architecture is not specific to certain industries and designed at a high level. The results of the proposed architectural design are expected to contribute to the application of RegTech in various industrial fields or regulated entities in realizing regulatory compliance more efficiently.

2. DESIGN METHODOLOGY

2.1 Methodology

This research principally carries out the process of designing an artefact, namely the solution architecture. The research method used in this study refers to the Design Science Research Methodology (DSRM) proposed by Peffers et al. [19]. Figure 1 shows the stages of the activities carried out in this study.

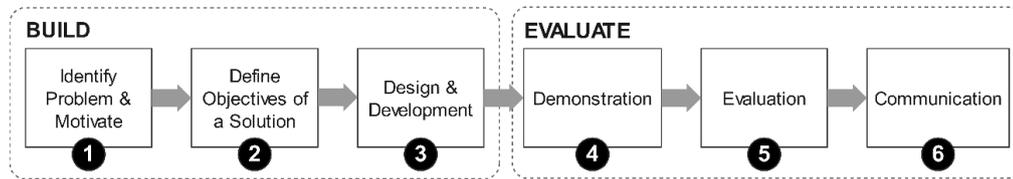


Figure 1. DSRM process model [19], [20]

In general, the DSRM stages consist of six parts, namely (1) problem identification and motivation; (2) objective definition of a solution; (3) design and development; (4) evaluation; (5) demonstrations; and (6) communication. The first section of this paper, the introduction, which comprises the background, research problem, and objective, explains the first and second stages. The third stage entails creating artefacts, including the RegTech generic solution architecture. The fourth stage evaluated the proposed artefacts' suitability for application in various case simulations [21]. Furthermore, in the fifth stage, artefact measurements were carried out regarding time efficiency [20]. The last step is to describe and present the research's findings by producing this article.

2.2 Design Method

This section describes in more detail the third stage of DSRM: creating and producing generic RegTech solution architecture artefacts. The architectural design adopts the enterprise architecture (EA) layer division model in the TOGAF framework. The division of these layers includes business, data, applications, technology or infrastructure and security, as shown in Figure 2. In general, there are two main stages in designing the architectural artefacts of this solution, namely requirements analysis and architectural modelling. In the first stage, a needs analysis is carried out for each layer, including solution components, including people, organization, process, information or data, and technology [22]. Next, modelling the solution architecture at each layer is carried out based on the results of identifying needs that have been analyzed previously. The modelling of the adopted EA artefacts can be in the form of catalogues (lists of things), matrices (showing relationships between things), and diagrams (pictures of things) [23].

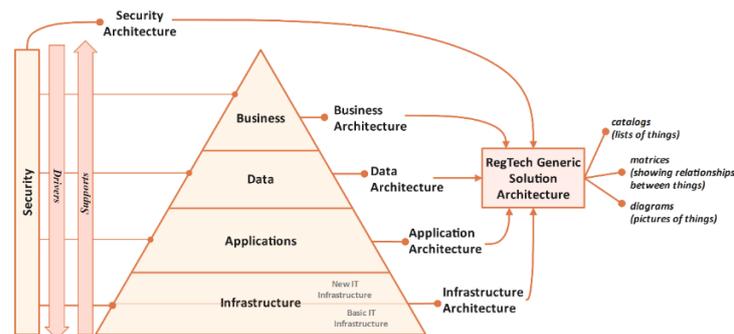


Figure 2. Solution architecture layer or domain categorization (adopted and modified from [22]–[24])

Lovatt [22] states that the solution components intersect with the four EA domains, as shown in Figure 3. Therefore, these components are also used to form the solution architecture in each corresponding domain. The difference is that the scope of the EA domain is for enterprises, while the solution architecture domain is only for certain specific solutions. In this study, it is the RegTech solution. Solution architecture is part of EA and must continue to coordinate with EA in its components to remain consistent. The details of the generic RegTech solution architecture design stage in each layer follow.

1. A generic regulatory compliance business process analysis is carried out in the business layer to obtain what is needed, including processes, activities, actors, and organizations. Based on the needs analysis results, business architecture modeling is carried out.

2. In the data layer, a needs analysis is carried out by identifying data entities and data storage needs in supporting the generic business processes of the RegTech solutions identified previously. Furthermore, data architecture modeling is carried out.
3. At the application layer, a needs analysis is carried out by identifying applications that support business processes and manage data previously identified at the business and data layers. Based on these needs, data architecture modeling is carried out.
4. Next, a needs analysis is conducted at the infrastructure layer by identifying a catalog of basic and new IT infrastructure needed to support applications and data management. Next, infrastructure architecture modeling is completed.
5. Finally, the security layer analyzes the security requirements based on the infrastructure architecture to support the security of applications and data and the running of business processes in the generic RegTech solution. A high-level security architecture modeling is performed based on the needs analysis results.

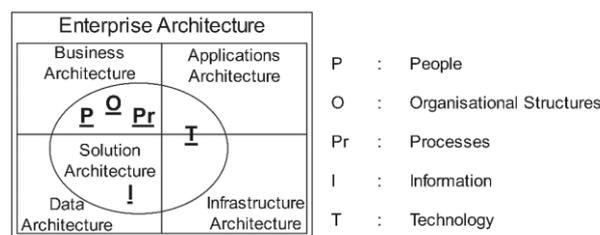


Figure 3. Relationship between solution architecture and EA [22]

2.3 Evaluation

This section describes stages four and five of the DSRM: demonstration and evaluation. The demonstration was carried out by using artifacts to illustrate logical scenarios in the non-financial industry. Peffers et al. [21] stated that architectural artifacts could be evaluated using scenario illustrations. An illustration of the scenario can be seen in Table 1. Furthermore, from the scenario illustration, a simulation and evaluation of business processes are carried out using the time analysis feature with the help of the Bizagi Modeler application. It aims to see the time efficiency of the regulatory compliance process before and after implementing the RegTech solution. Sonnenberg et al. [20] stated that the efficiency variable could be used to carry out a measurable evaluation of architectural artifacts.

Table 1. Illustration of demonstration scenario

Company Name	Scenario
ABC Ltd.	<p>ABC Ltd. is one of the cosmetic manufacturing companies in Indonesia and holds several leading brands. In carrying out its operations, ABC Ltd. is supervised by the Food and Drug Supervisory Agency (In Indonesia, it is known as BPOM). BPOM has a role in ensuring that cosmetic products sold in Indonesia are safe and comply with applicable regulations.</p> <p>ABC Ltd. wants to develop and implement RegTech solutions to ensure operations related to the manufacture and distribution of cosmetics comply with regulations issued by BPOM. With the RegTech solution, it is hoped that non-compliance with these operations can be detected earlier and in real-time so that preventive decisions can be taken immediately. In addition, the RegTech solution is also expected to help automate the creation of compliance reports.</p>

3. RESULTS AND DISCUSSION

The architecture of the generic RegTech solution is covered in this section, alongside the results of the proposed and evaluated artefacts in this study. In addition, it also discussed related future research perspectives.

3.1 RegTech Generic Solutions Architecture Results

3.1.1 Business Architecture

At this stage, the focus is on the three main components of the solution architecture design, namely, people, organization, and business processes. Business process analysis is carried out based on literature studies related to regulatory compliance business processes by applying RegTech to the financial industry as described by [1], [3], [25], [26]. Furthermore, the business process flow is adopted and adjusted to become a generic

regulatory compliance business process in regulated entities. Generic business processes are general and do not refer to a particular industry or entity. Based on the literature study, it was concluded that regulated entities' regulatory compliance business process consists of four processes: data collection, data processing, data analysis, and data presentation [10]. Furthermore, the needs of each process are identified to support generic regulatory compliance by implementing RegTech solutions. The results of identifying these needs can be seen in Table 2.

Table 2. The results of identification of needs in regulatory compliance generic business processes by implementing RegTech solutions

Process	Need
Data Collection	Automation of regulatory (real-time) and organizational operational (real-time and batch) data collection.
Data Processing	Automation of regulatory (real-time) and organizational operational (real-time and batch) data extraction and transformation.
Data Analytics	Automation of regulatory (real-time) and compliance (real-time and batch) data analytics using AI/ML
Data Presentation	Automate the presentation of regulatory data analytics results, compliance analytics results, and compliance reporting. This activity can be displayed in real-time and batch.

Based on the needs analysis in Table 2, a business architecture model is carried out in the form of a diagram covering the components of business processes, people, and organizations. The division of business process levels adopts the hierarchical process model made by Mahal [27]. The results of business architecture modelling can be seen in Figure 4.

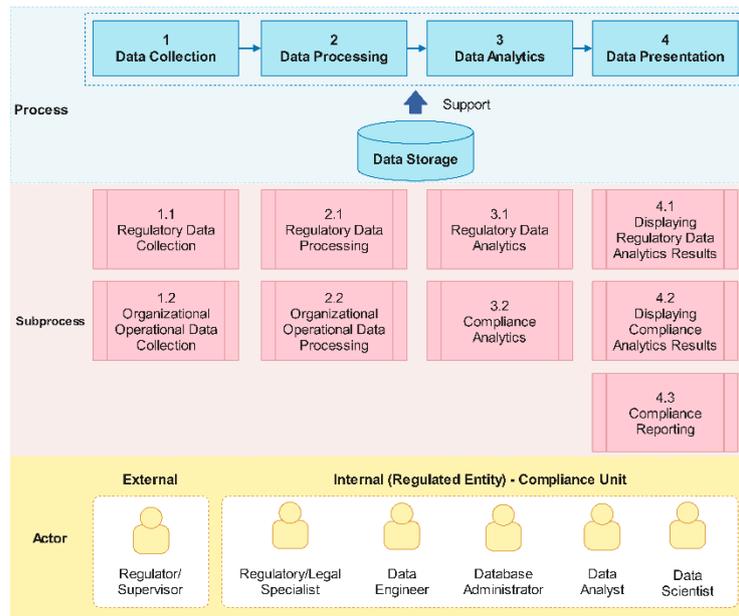


Figure 4. Generic business architecture of RegTech solution

There are two levels in the generic business architecture of RegTech solutions (Figure 4). First is level 1 business processes, including data collection, processing, analytics, and presentation. Each of these processes is supported by a data store. Second, level 2 business processes include all sub-processes in each level 1 business process. Altogether there are nine sub-processes. Each subprocess has several interrelated activities to fulfil regulatory compliance using RegTech.

Actors involved in the generic business processes of RegTech solutions are categorized into external and internal (Figure 4). External actors are regulators/supervisors who have a role as parties that issue various regulations and receive compliance reports submitted by regulated entities as a form of compliance. Internal actors are regulated entities or organizations whose business activities are regulated by regulation. Ideally, this entity has a compliance unit or compliance work unit that is focused on regulatory compliance issues. Some of the roles contained in this actor are regulatory/legal specialist, data engineer, database administrator, data analyst, and data scientist.

3.1.2 Data Architecture

At this stage, the focus is on analysing data components, including data entities and storage. Data entity identification was carried out for each generic business process RegTech solution in the early stages

based on the previous business architecture design results. Each business process identified data entities involved at a high level. Next, place every business process that requires data storage. Results identification of data entities and data storage requirements in generic business processes RegTech solutions can be seen in Table 2.

Table 2. Results identification of data entities and data storage requirements in generic business processes RegTech solutions

Process	Data Entity	Data Storage Needs
Data Collection	Input : Regulation (regulatory data published by regulator/supervisor), organizational operational (organizational operational data contained in various internal sources of the organization), and metadata . Output : Regulation (regulatory data that has been collected and stored in storage), organizational operational (organizational operational data that has been collected and stored in storage), and metadata .	<ul style="list-style-type: none"> Storage of regulatory data published by regulators/supervisors (raw data) and metadata. Storage of organizational operational data collected from various internal sources of the organization (raw data) and metadata.
Data Processing	Input : Regulation (regulation data already collected), organizational operational (operational data of the organization that has been collected), and metadata . Output : Regulation (regulation data that has been transformed), organizational operational (transformed organizational operational data), and metadata .	<ul style="list-style-type: none"> Storage of transformed regulatory data (valid or clean data) and metadata. Formatted and standardized store of organizational operational data for compliance analytics and metadata.
Data Analytics	Input : Regulation (transformed regulatory data), organizational operational (transformed organizational operational data that conforms to formats and standards for compliance analytics), model (AI/ML model for compliance analytics), and metadata . Output : Regulation (regulatory data analysis results), compliance analytics result , compliance report , and metadata .	<ul style="list-style-type: none"> Storage of AI/ML models for compliance analytics. Data storage of compliance analysis results. Metadata storage.
Data Presentation	Input : Regulation (regulatory data analysis results), compliance analytics result , compliance report , and metadata . Output : Metadata	Storage of metadata related to data presentation.

Based on the analysis of high-level identification of data entities and data storage needs, the following are the general requirements that must be met by the generic data architecture of the RegTech solution that will be designed.

1. Generally, the data entities must be available are regulation, organizational operational, model, compliance analytics result, compliance report, and metadata.
2. Have a data storage system that has general capabilities as follows: (i) supports storage and handling of data in structured, semi-structured, and unstructured formats in large volumes; (ii) supports data processing in real-time and batch; (iii) supports data/metadata catalog storage; (iv) supports on-demand scalability so that capacity is easily added according to various needs; and (v) support access by systems that require stored data, such as data science, visualization, and reporting tools.

Furthermore, high-level relationship modeling between entities and data architecture modeling of RegTech generic solutions are carried out based on the needs analysis results. Figure 5 shows the relationship between data entities that have been identified previously. Modeling the relationship is done conceptually using entity relationship (ER) diagrams. Meanwhile, RegTech generic data architecture modeling can be seen in Figure 6.

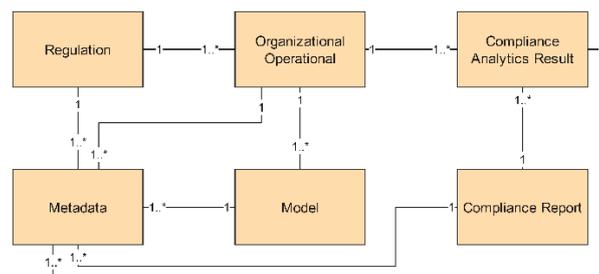


Figure 5. Conceptual data model

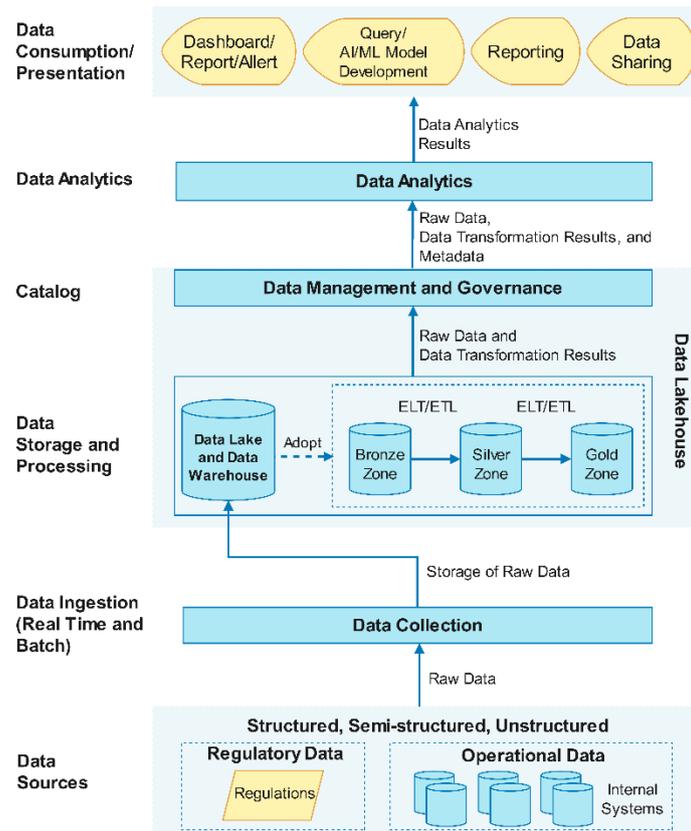


Figure 6. Generic data architecture of RegTech solution

The data storage system adopts the data lakehouse architecture. A data lakehouse can be defined as "a data management system based on low-cost, directly accessible storage that also provides database management system performance and management features for traditional analytics such as ACID (atomicity, consistency, isolation, durability) transactions, data versioning, auditing, indexing, caching, and query optimization" [28]. Lakehouse combines the main benefits of data lakes and data warehouses, namely low-cost storage in an open format that various systems from the data lake and the data warehouse's power management and optimization features can access [28]. Data lakehouses can store data in structured, semi-structured and unstructured formats [29]. The data sources that RegTech inputs are regulatory data published by regulators/supervisors and various organizational operational data. The process of ingestion or collecting data can be batch and real-time. Regulatory data is collected in real-time so that every time there is a change in regulations, the data will be processed immediately. Organizational operational data are identified according to the needs related to the collection in real-time or batch.

The data lakehouse divides the data storage zone into three, namely: bronze, silver, and gold zone [29]. All raw data collected is stored in the bronze zone. Data from the bronze zone transformed through the ELT (Extract, Load, Transform)/ETL (Extract, Transform, Load) process is stored in the silver zone. The gold zone stores the data transformation results from the silver zone through the ELT/ETL process. Data in the gold zone has been aggregated or has a certain scheme for business needs. The catalogue layer is in the form of data management and governance, such as metadata so that it can manage and organize data, such as a data warehouse. This is what distinguishes data lakehouses from data lakes. This layer supports storing all data formats and incorporates best practices from database management features to support data management, governance, and securing data. The data analytics layer performs analytical processing for regulation and compliance in real-time and batch. This process can use AI/ML models for real-time monitoring related to compliance or certain query schemes. Analytical result data is presented or displayed in the presentation or consumption layer using the application programming interface (API) as a data exchange.

3.1.3 Application Architecture

At this stage, an analysis of how software applications support business processes (business architecture) and help manage data (data architecture) in RegTech generic solutions is carried out. The first thing to do is to identify applications that support the generic business processes of RegTech solutions that

have been defined earlier: data collection, data processing, data analytics, and data presentation. The identification results and the data entities involved can be seen in Table 3 below.

Table 3. Identify applications based on the generic business processes of RegTech solutions

Process	Application Functional Requirements	Application	Engaging Data Entities
Data Collection	<ul style="list-style-type: none"> Extract data in structured, semi-structured and unstructured formats automatically from various sources, either in real-time or batch. Save extracted data in storage. 	Data Ingestion Tool, Data Management System	Regulation, Organizational, Operational, Metadata
Data Processing	<ul style="list-style-type: none"> Process automatically the results of data extraction either in real-time or in batch for the process of compliance analytics and regulatory analytics. Storing processed data in storage. 	ETL/ELT Tool, Data Management System	Regulation, Organizational, Operational, Metadata
Data Analytic	<ul style="list-style-type: none"> Perform regulatory data analytics automatically and in real-time using AI/ML, such as optical character recognition (OCR), natural language processing (NLP), and text mining. Storing data on the results of regulatory analysis in storage. Develop AI/ML models for regulatory and compliance analytics. Storing AI/ML models in storage. 	Analytic Engine, Data Management System, API Tool	Regulation, Metadata
	Perform further automated processing of pre-processed organizational operational data according to business requirements for compliance analytics in real-time or batch.	ETL/ELT Tool	Regulation, Organizational, Operational, Metadata
	<ul style="list-style-type: none"> Perform compliance analytics processes automatically, either in real-time or in batch. Store compliance analytics results in storage. 	Analytic Engine, Data Management System, API Tool	Regulation, Organizational, Operational, Model, Compliance Analytics Result, Metadata
	Automatically generate compliance reports	RPA Tool	Compliance Analytics Result, Metadata
	Automatically send compliance reports to regulators/supervisors	RPA Tool, API Tool	Compliance Report, Metadata
Data Presentation	<ul style="list-style-type: none"> Presents the results of regulatory analytics Presents the results of compliance analytics Provide compliance reports and data sharing related to compliance reports 	API Tool, Visualization Tool, Reporting Tool, Data Science Tool, Data Management System	Regulation, Compliance Analytics Result, Compliance Report, Metadata

Based on the results of the application identification described in Table 3, the following are the general requirements that must be met by the generic application architecture of the RegTech solution to be designed.

- Supports automatic data collection from certain sources in real-time and batch.
- Supports automated data processing for real-time and batch regulatory and compliance analytics.
- Supports automating regulatory and compliance analytics processes in real-time and batch.
- Provides data analytics functionality using AI/ML.
- Provides functionality for AI/ML model development.
- Provides functions to share or exchange data.
- Provides functions for data storage and management.
- Support the process of making compliance reports and submitting compliance reports automatically to regulators/supervisors.
- Provides functions to present data on regulatory and compliance analytics results such as alerts, real-time monitoring, dashboards, and reporting.

Furthermore, the generic application architecture modelling of the RegTech solution is carried out based on the results of the needs analysis identified in the previous stage. The architectural modelling is done by grouping applications based on groups according to the main function requirements (Figure 7) and diagrams that explain the interactions of these various applications (Figure 8).

Data Ingestion	Data Processing	Data Storage and Management
Data Ingestion Tool	ELT/ETL Tool Analytics Engine	Data Management System
Data Analytics	Data Presentation	Supporting System
Analytics Engine Data Science Tool	Visualization Tool	Scheduler Tool API Tool RPA Tool

Figure 7. RegTech solutions generic application group

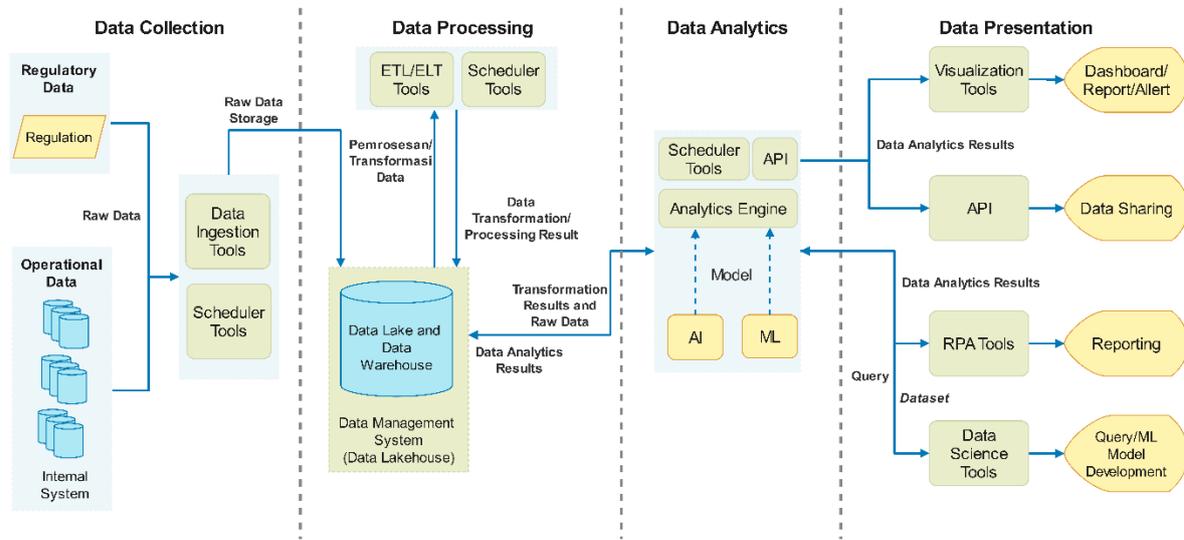


Figure 8. Generic application architecture of RegTech solution

3.1.4 Infrastructure Architecture

At this stage, the analysis and modelling of infrastructure architectures are grouped into two categories, namely Basic IT Infrastructure and new IT infrastructure or infrastructure 4.0. The basic IT infrastructure components support the implementation of new IT infrastructure components. Tables 4 and 5 show the results of identifying basic and new IT infrastructure catalogues needed to implement the business architecture, data architecture, and application architecture in RegTech generic solutions.

Table 4. Catalog list of basic IT infrastructure components

Basic IT Infrastructure Components	Description
Network	The network allows various application components at RegTech to communicate with each other, such as access to compliance analysis results in data stored in storage via a server for visualization or reporting. The network allows data exchange between devices, such as sending compliance reports to regulators/supervisors.
Compute and Storage	Compute devices are used to process regulatory and compliance analytics from data stored in storage, including regulatory and organizational operational data. In addition, computing devices can be used for other services supporting RegTech, such as web servers. This component must have good scalability to handle growing or increasing workloads.
Container	This virtualization technology can package RegTech applications and all the dependencies they need in one remote unit so that these applications can run in various environments without different configurations.
Data Center	Physical facilities are used to store applications and data. The main components of a data centre design include various networks, compute and storage devices. This facility must have a good and tested disaster recovery plan (DRP) and strategy before a disaster or disturbance occurs.

Based on the results of the analysis in Table 4, it is concluded that four main components are needed in the basic IT infrastructure that supports the implementation of data and application architectures in RegTech generic solutions, namely networks, compute and storage that has good scalability, container, and data center with disaster recovery plan (DRP).

Table 5. Catalog list of new IT infrastructure components

New IT Infrastructure Components	Description
Big Data Analytics	Big Data Analytics is the main infrastructure that RegTech must adopt. This technology can process and analyse very large and complex organisational operational data into information related to regulatory compliance, such as real-time compliance monitoring and reporting. This technology is used for every RegTech business process, namely data collection, data processing, data analytics, and data presentation. Various applications or big data platforms are components of RegTech applications in supporting business processes and data management.
AI/ML	AI/ML supports advanced analytics in big data such as diagnostic, predictive, and prescriptive related to regulatory compliance. This technology finds patterns related to data that do not comply with regulations, makes predictions, and makes decisions regarding the data. These technologies, such as NLP and text mining, can also extract information from regulatory documents into machine-readable rules.
RPA	RPA helps automate repetitive tasks such as generating and submitting compliance reports to regulators/supervisors.
Edge Computing	This technology can reduce the amount of data traffic on the network. Collecting and processing data on RegTech solutions can be done close to the data source or on edge computing devices if the data volume is still very large. Furthermore, data processing results are only forwarded to the main compute and storage devices for analysis so that data traffic to these main devices is manageable and bandwidth can be used optimally.
API	The API provides a communication channel allowing applications or systems to communicate and share data. In RegTech solutions, this technology provides data or information on the results of these analytics for visualization systems or applications such as real-time compliance monitoring and reporting.
Cloud Technology	Cloud technology can be used in various forms of services, such as infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS). RegTech can run on top of this technology according to service needs, including application components, data, and various other supporting technologies that support the business processes of RegTech solutions.

Based on the analysis results in Table 5, it is concluded that six main components are needed in the new IT infrastructure that supports data and application architectures, namely big data analytics, AI/ML, RPA, edge computing, API, and cloud technology. Furthermore, infrastructure architecture modelling is carried out that supports the implementation of the data architecture and generic application architecture RegTech solutions that have been designed previously. Figure 9 shows the generic infrastructure architecture model for RegTech solutions at a high level by adopting a layer three-tier architecture and adding a basic layer as the basic architecture. The following describes the layers in the RegTech generic infrastructure architecture solution (Figure 9).

1. The base layer provides the basic IT infrastructure, including software and hardware, that supports the data, application, and presentation layers. This layer includes network, compute and storage, and data center.
2. The data layer stores all the data required by a RegTech solution, such as regulatory data, organizational operations, and various analytical results, including regulatory and compliance analytics. The data will later be processed at the application layer and forwarded to the presentation layer to be displayed to the user. This layer uses data lakehouse technology to store various data for further use as needed, such as regulatory and compliance analytics.
3. All the RegTech logic functionality resides in the application layer. RegTech implements a big data analytics architecture consisting of four stages: data ingestion, data processing, data analytics, and data presentation. All these stages correspond to the generic business processes of the RegTech solution. Edge computing can be applied if the volume of data processed is very large by carrying out the collection and processing stages close to the data source so that the data sent for the analytical process in the main data center is not too large. RegTech uses an AI/ML model for compliance and regulatory analytics. In addition, RPA also supports the process of automatically creating and sending regulatory reporting to regulators/supervisors. RegTech uses API technology for data exchange, such as sending data from analytics results to the presentation layer.
4. The presentation layer provides the end-user interface for interacting with RegTech. The compliance unit can see various compliance analysis results and notifications of regulatory changes in real-time through the visualization tool on this layer. Data scientists can also interact with various data to create AI/ML models using the data scientist tools at this layer. The API provides data processed by various logical functions in the application layer and sends it to the presentation layer. Various devices can access the RegTech interface, such as desktops, laptops, and mobiles.

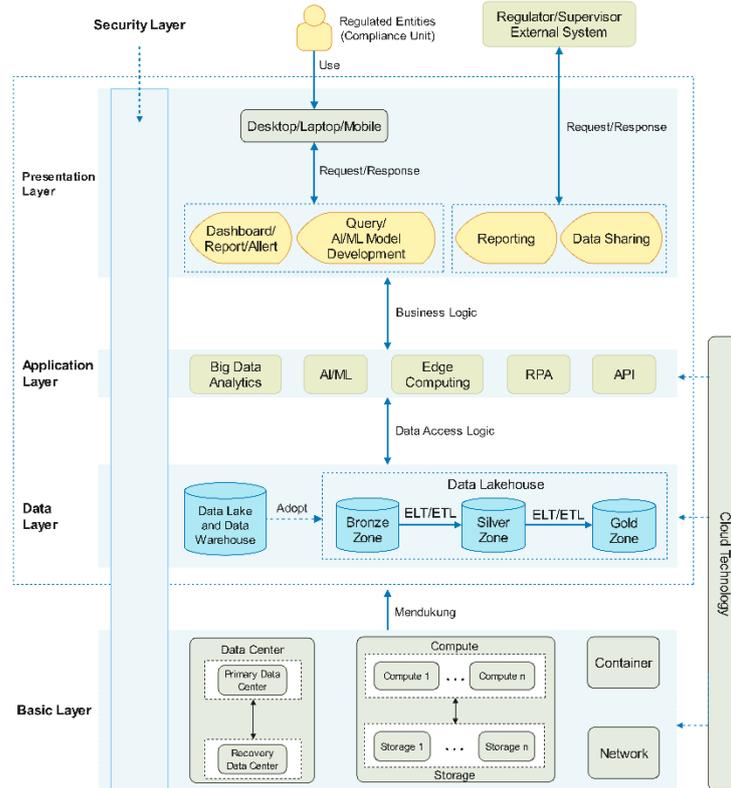


Figure 9. Generic infrastructure architecture of RegTech solution

Cloud technology also provides internet-based services, such as data storage, applications, and basic IT infrastructure, that can be adopted for implementing RegTech solutions. Cloud services can be implemented for the base, data, and application layers. Cloud services are managed by certain providers so that users only receive various services according to their needs.

3.1.5 Security Architecture

Based on the generic infrastructure architecture of the RegTech solution in Figure 9, security includes four layers: basic, data, application, and presentation. Security requirements at this layer can be seen in Table 6 below.

Table 6. Identify security needs

Layer	Description of Requirements
Basic	<p>Data Center Security If we maintain our own data center, the physical security of the data center is very important. Data center access from unauthorized parties must be blocked. Logical access to the server must be secured by network security which can be done by configuring the appropriate firewall. So, data center security depends on network security.</p>
	<p>Network Security Network security helps prevent overall IT resources supporting implementations of RegTech solutions from being open to external users. Network security helps prevent unauthorized system access, host vulnerabilities, and port scanning. Intrusion Detection System (IDS) and Intrusion Prevention System (IPS) can support network security using firewalls.</p>
Data	<p>Data Security Data is a core component of the RegTech solution because everything is sourced from data, processed, and analyzed to produce various information related to regulatory compliance. Here are some solutions for data security:</p> <ul style="list-style-type: none"> • Data Classification This solution is practiced by classifying data based on the sensitivity so that it is easy to plan data protection, data encryption, and data access requirements. • Data Encryption Practice this solution by converting data from plaintext to ciphertext format, which is encoded using the encryption key so that only authorized users will have access to the decryption key. • Data Encryption at rest and in transit Every layer where data is located or temporarily resides must be secured, such as data exchange and when data is in the database. When data is in transit (data-in-transit), it can be secured using Secure Socket Layer/Transport Layer Security (SSL/TLS) and (a security certificate). Meanwhile, when data is not active or is in the database, it can be secured by encryption.

Layer	Description of Requirements
Application	<p>Application Security</p> <p>RegTech is implemented on the network so that it is prone to cyber-attacks, especially for sensitive organizational data that will be very vulnerable and dangerous if it gets this attack. Web Application Firewall (WAF) can block malware, SQL injection attacks on applications, and malicious IP addresses. WAF can be combined with a Content Distribution Network (CDN) to help prevent and respond to distributed denial-of-service (DDoS) attacks.</p>
Presentation	<p>Endpoint Security</p> <p>Endpoint security protects individual devices in the network from attacks and security threats. The endpoint can be a computer, laptop, tablet, or mobile device connected to the network to access the RegTech front end. Endpoint security ensures that the devices are protected from malware, phishing, and other attacks that can cause damage or steal personal information. Technologies used in endpoint security include antivirus, anti-malware, firewalls, and other security software.</p> <p>User Access Security</p> <p>RegTech is accessible to users according to their respective roles in the regulated organization or entity. Therefore, it is very important to manage this access so that it does not interfere with RegTech's operations. Solutions that can be applied are authentication and authorization. Authentication means determining who can access the system, and authorization is applied to the user's activities after logging into the system or application. The solution that can be used for this security is user access via SSO (Single Sign-On) or active directory.</p>

Based on the results of the analysis of security needs that have been identified previously, then the generic security architecture modeling of RegTech solutions is carried out at a high level. The architectural modeling can be seen in Figure 10. The security architecture includes four layers, namely basic, application, data, and presentation. The security at the base layer consists of data center and network security. Data center security depends on network security. Security at the base layer supports data security. Data security can be addressed with data classification, data encryption, and data encryption solutions at rest and in transit. Next is security at the application layer using a Web Application Firewall (WAF) solution. Lastly is security at the presentation layer, covering user access security with authentication and authorization solutions and endpoint security with antivirus, anti-malware, and firewall solutions.

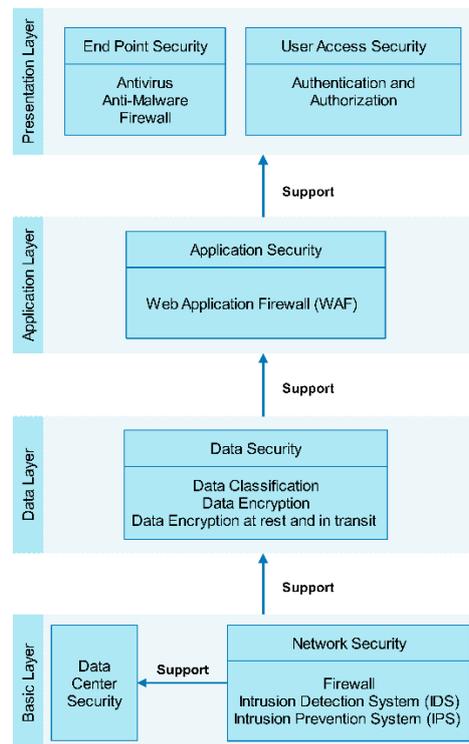


Figure 10. Generic security architecture of RegTech solution

3.2 Evaluation Results

At this stage, generic RegTech solution architecture artefacts generated in the scenario illustration described in section 2.3 will be implemented. The application of these artefacts includes business architecture, data, applications, infrastructure, and security. The modelling results can be seen in Figures 11 to 14.

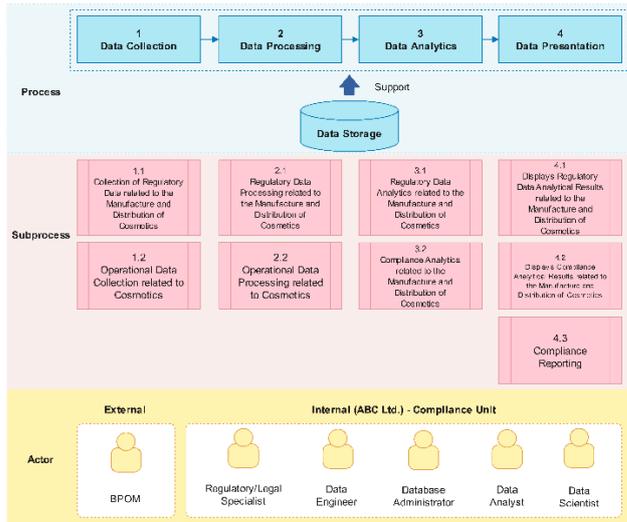


Figure 11. RegTech solution business architecture diagram at ABC Ltd.

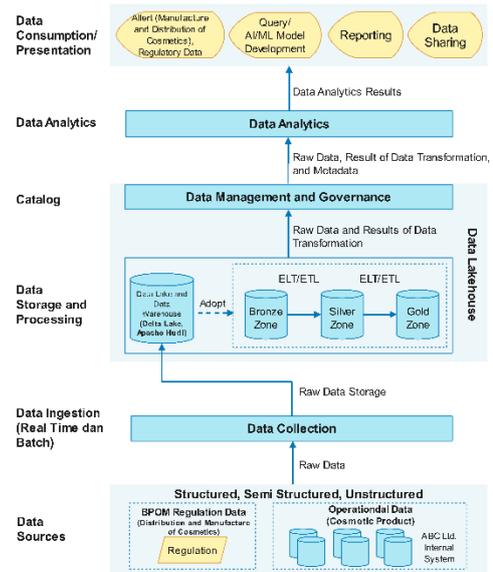


Figure 12. RegTech solution data architecture diagram at ABC Ltd.

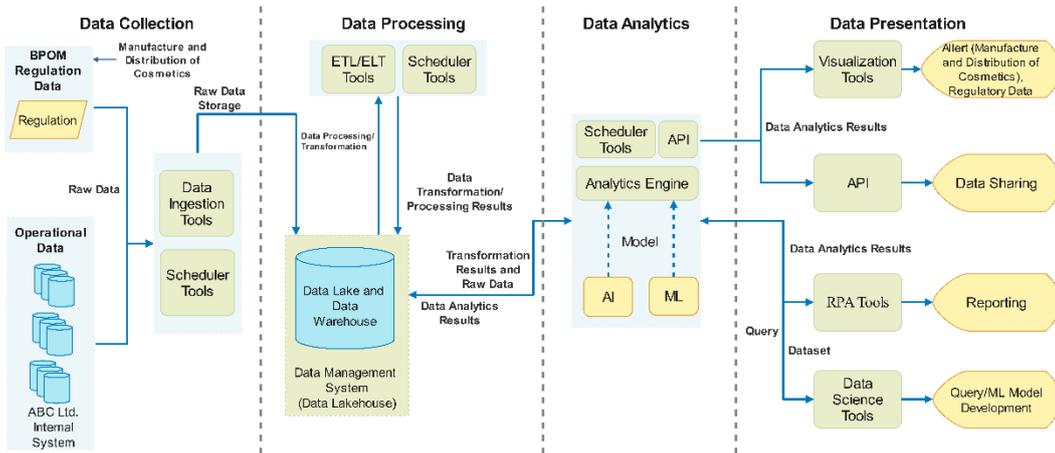


Figure 13. RegTech solution application architecture diagram at ABC Ltd.

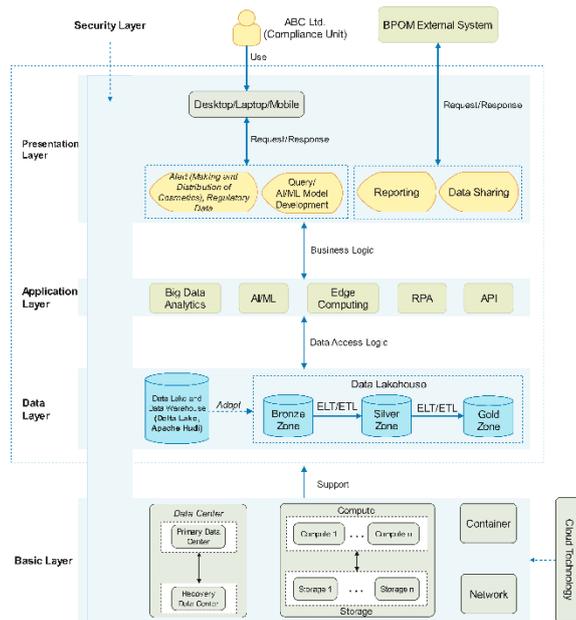


Figure 14. RegTech solution infrastructure architecture diagram at ABC Ltd.

The security architecture can adopt the generic security architecture of the RegTech solution by considering the security of the data center, network, data, applications, endpoints, and users. Various solutions for each security can be seen in Figure 10, previously related to the generic security architecture diagram of RegTech solutions. Based on Figures 11 to 14, the artefacts from this research can be applied to logical scenarios outside of finance. Next, a time analysis is performed using the Bizagi Modeler application. Business processes from scenarios before (As-Is) and after (To-Be) implementing RegTech were modelled with BPMN (Business Process Model and Notation). The business process simulation results can be seen in Figures 15 to 22 below. The results of evaluating business processes using time analysis with the Bizagi Modeler can be seen in Table 7.

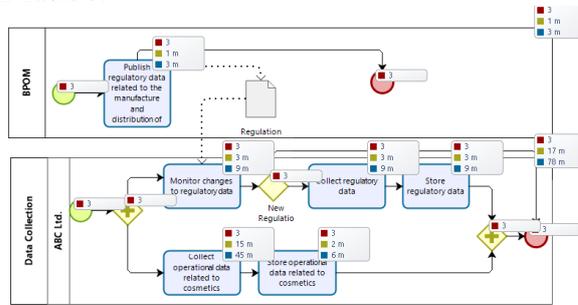


Figure 15. Data collection business processes at ABC Ltd. (As-Is)

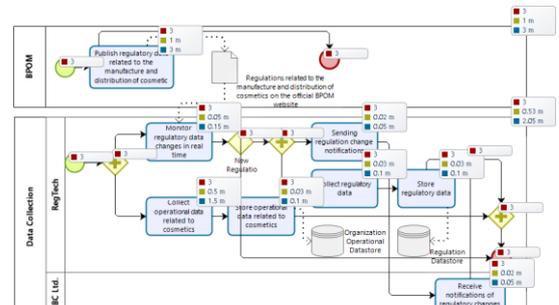


Figure 16. Data collection business processes at ABC Ltd. (To-Be)

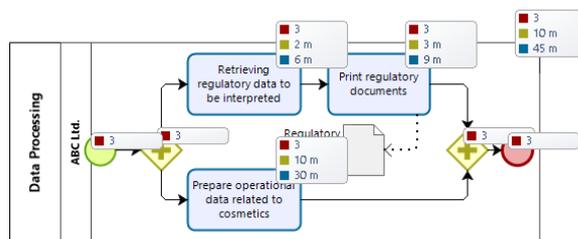


Figure 17. Data processing business processes at ABC Ltd. (As-Is)

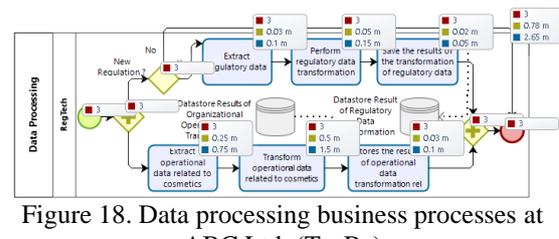


Figure 18. Data processing business processes at ABC Ltd. (To-Be)

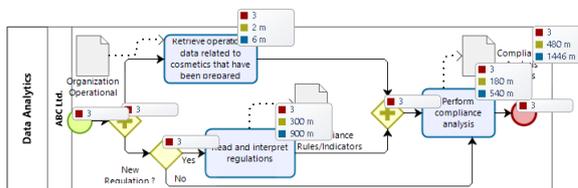


Figure 19. Data analytics business processes at ABC Ltd. (As-Is)

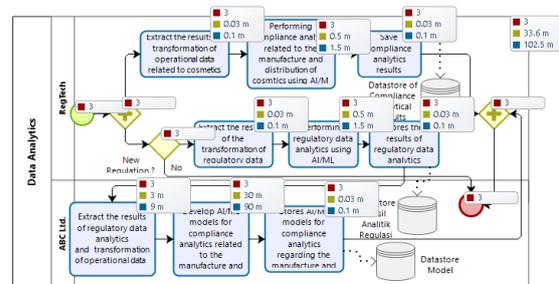


Figure 20. Data analytics business processes at ABC Ltd. (To-Be)

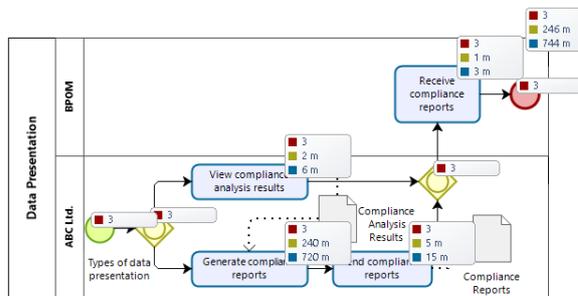


Figure 21. Data presentation business processes at ABC Ltd. (As-Is)

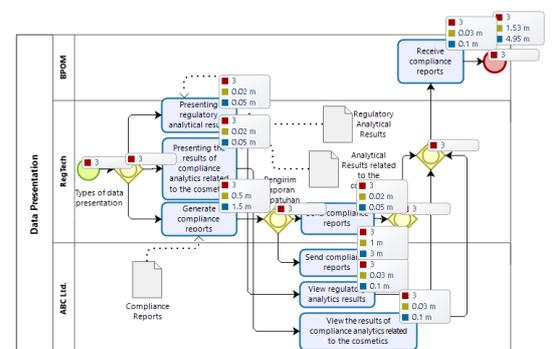


Figure 22. Data presentation business processes at ABC Ltd. (To-Be)

Table 7. The results of evaluating business processes using time analysis

Regulatory Compliance Business Processes	Average Time As-Is (minute)	Average Time To-Be (minute)	Efficiency (%)
Data Collection	17.00	0.53	96.88%
Data Processing	10.00	0.78	92.20%
Data Analytics	480.00	33.60	93.00%
Data Presentation	246.00	1.53	99.38%
Total	753.00	36.44	95.16%

Based on the table above, regulatory compliance business processes by implementing RegTech (To-Be) have an excellent time efficiency value compared to the old method (As-Is). The RegTech solution has the potential to provide a time efficiency of 95.16%. Regulatory interpretation and compliance analytics using AI/ML significantly impact RegTech implementation to achieve regulatory compliance more efficiently.

4. CONCLUSION

This research has produced a proposed RegTech Solution Architecture design by adopting the layer division model and EA artifacts referring to the TOGAF framework, including business, data, application, infrastructure, and information security architectures. The evaluation results by applying case studies to industries other than finance show that the proposed RegTech Solution Architecture is generic. Furthermore, an evaluation of the comparison of regulatory compliance business processes without and by implementing RegTech can produce a time efficiency of 95.16%. These results show that RegTech solutions can achieve regulatory compliance more efficiently.

The results of the RegTech generic solution architecture design can be applied as a case study in subsequent research for the development of RegTech in more specific regulated entities, especially other than financial entities or industries. This architectural proposal includes business, data, application, infrastructure, and security architectures that are still conceptual or high-level design, so when implemented in more specific industries or certain regulated entities, it is advisable to decompose business processes, data, applications, infrastructure, and security in more detail according to the needs of the industry or regulated entity. The various technological solutions defined in this architectural proposal can still change along with the increasingly rapid development of technological trends.

REFERENCES

- [1] J. von Solms, "Integrating Regulatory Technology (RegTech) into the digital transformation of a bank Treasury," *J. Bank. Regul.*, vol. 22, no. 2, hal. 152–168, 2020, doi: 10.1057/s41261-020-00134-0.
- [2] J. Du dan L. Wei, "An analysis of regulatory technology in the Internet financial sector in conjunction with the logit model," *Proc. - 2020 2nd Int. Conf. Econ. Manag. Model Eng. ICEMME 2020*, hal. 428–431, 2020, doi: 10.1109/ICEMME51517.2020.00091.
- [3] T. Butler dan L. O'Brien, "Understanding RegTech for Digital Regulatory Compliance," *Disrupting Financ.*, hal. 85–102, 2019, doi: 10.1007/978-3-030-02330-0_6.
- [4] J. von Solms, "Integrating Regulatory Technology (RegTech) into the digital transformation of a bank Treasury," *J. Bank. Regul.*, vol. 22, no. 3, hal. 191–207, 2021, doi: 10.1057/s41261-020-00138-w.
- [5] M. N. Kondratyeva, D. D. Svirina, dan A. I. Tsvetkov, "The role of information technologies in ensuring banking security," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1047, no. 1, 2021, doi: 10.1088/1757-899X/1047/1/012069.
- [6] R. P. Buckley, D. W. Arner, D. A. Zetzsche, dan R. H. Weber, "The road to RegTech: the (astonishing) example of the European Union," *J. Bank. Regul.*, vol. 21, no. 1, hal. 26–36, 2020, doi: 10.1057/s41261-019-00104-1.
- [7] D. W. Arner, J. Nathan Barberis, dan R. Buckley, "FinTech and RegTech in a Nutshell, and the Future in a Sandbox," no. October 2018, 2017, doi: 10.2139/ssrn.3088303.
- [8] D. W. Arner, "FinTech , regTech , and the reconceptualization of financial regulation," *Northwest. J. Int. Law Business, Forthcoming, Univ. Hong Kong Fac. Law*, no. January, 2017.
- [9] M. Becker, K. Merz, dan R. Buchkremer, "RegTech—the application of modern information technology in regulatory affairs: areas of interest in research and practice," *Intell. Syst. Accounting, Financ. Manag.*, vol. 27, no. 4, hal. 161–167, 2020, doi: 10.1002/isaf.1479.
- [10] B. Firmansyah dan A. A. Arman, "A Systematic Literature Review of RegTech: Technologies, Characteristics, and Architectures," *2022 Int. Conf. Inf. Technol. Syst. Innov. ICITSI 2022 - Proc.*, hal. 310–315, 2022, doi: 10.1109/ICITSI56531.2022.9970830.
- [11] A. D. Kristanto dan A. A. Arman, "Towards A Smart Regulatory Compliance, The Capabilities of RegTech and SupTech," in *2022 International Conference on Information Technology Systems and Innovation (ICITSI)*, 2022, hal. 300–309, doi: 10.1109/ICITSI56531.2022.9970801.
- [12] E. Johansson, K. Sutinen, J. Lassila, V. Lang, M. Martikainen, dan O. M. Lehner, "RegTech-A Necessary Tool to Keep Up with Compliance and Regulatory Changes?," *ACRN J. Financ. Risk Perspect.*, vol. 8, no. Digital Accounting, hal. 71–85, 2019.
- [13] P. Treleaven, J. Barnett, A. Knight, dan W. Serrano, "Real Estate Data Marketplace," *AI Ethics*, vol. 1, no. 4, hal. 445–462, 2021, doi: 10.1007/s43681-021-00053-4.

- [14] C. Singh, L. Zhao, W. Lin, dan Z. Ye, "Can machine learning, as a RegTech compliance tool, lighten the regulatory burden for charitable organisations in the United Kingdom?," *J. Financ. Crime*, vol. 29, no. 1, hal. 45–61, 2022, doi: 10.1108/JFC-06-2021-0131.
- [15] J. H. M. Bergmann, "The Emerging Field of Medical Regulatory Technology and Data Science," *Prosthesis*, vol. 4, no. 2, hal. 169–171, 2022, doi: 10.3390/prosthesis4020017.
- [16] E. N. Lallas, I. Santouridis, G. Mountzouris, V. C. Gerogiannis, dan A. Karageorgos, "An SQWRL-Based Method for Assessing Regulatory Compliance in the Pharmaceutical Industry," *Appl. Sci.*, vol. 12, no. 21, hal. 10923, 2022, doi: 10.3390/app122110923.
- [17] M. Turki, A. Hamdan, R. T. Cummings, A. Sarea, M. Karolak, dan M. Anasweh, "The regulatory technology 'RegTech' and money laundering prevention in Islamic and conventional banking industry," *Heliyon*, vol. 6, no. 10, hal. e04949, 2020, doi: 10.1016/j.heliyon.2020.e04949.
- [18] A. McSweeney, *Introduction to Solution Architecture*. Independently published (February 20, 2019), 2019.
- [19] K. Peffers, T. Tuunanen, M. A. Rothenberger, dan S. Chatterjee, "A design science research methodology for information systems research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, hal. 45–77, 2007, doi: 10.2753/MIS0742-1222240302.
- [20] C. Sonnenberg, "Design Science Research in Information Systems. Advances in Theory and Practice," vol. 7286, no. August 2014, 2012, doi: 10.1007/978-3-642-29863-9.
- [21] K. Peffers, L. Vegas, M. A. Rothenberger, L. Vegas, T. Tuunanen, dan R. Vaezi, "Design Science Research Evaluation," vol. 7286, no. July 2021, 2012, doi: 10.1007/978-3-642-29863-9.
- [22] M. Lovatt, *Solution Architecture Foundations*. Swindon, United Kingdom: BCS, The Chartered Institute for IT, 2021.
- [23] The Open Group, *The TOGAF® Standard, Version 9.2*. The Open Group, 2018.
- [24] C. Gebayew dan A. A. Arman, "Modify TOGAF ADM for Government Enterprise Architecture: Case Study in Ethiopia," *Proceeding 2019 5th Int. Conf. Wirel. Telemat. ICWT 2019*, hal. 1–6, 2019, doi: 10.1109/ICWT47785.2019.8978260.
- [25] J. Barberis, D. W. Arner, dan R. P. Buckley, *The RegTech Book, The Financial Technology Handbook for Investors, Entrepreneurs and Visionaries in Regulation*, vol. 4, no. 1. Wiley, 2019.
- [26] BIAN, "BIAN Service Landscape," *BIAN*, 2023. bian.org.
- [27] A. Mahal, *Business Process Management, Basics and Beyond. How works Get Done*. Technics Publications, LLC, 2010.
- [28] M. Armbrust, A. Ghodsi, R. Xin, dan M. Zaharia, "Lakehouse : A New Generation of Open Platforms that Unify Data Warehousing and Advanced Analytics," *Conf. Innov. Data Syst. Res.*, 2021.
- [29] N. E. Janssen, "The Evolution of Data Storage Architectures : Examining the Value of the Data Lakehouse," Master Thesis, University of Twente, 2022.

BIOGRAPHY OF AUTHORS



Benny Firmansyah is undergoing a master's program at the School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia. He received his bachelor of statistical computing (S.S.T) from STIS Statistics Polytechnic, Indonesia, in 2011. He can be contacted at email: 23221114@std.stei.itb.ac.id



Arry Akhmad Arman is a lecturer at the School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia. He received his bachelor's (Ir), master's (M.T.), and doctorate (Dr.) from Institut Teknologi Bandung, Indonesia, in 1990, 1995, and 2004, respectively. He can be contacted at email: arry.arman@yahoo.com