

**Universidad Católica de Santa María**  
**Facultad de Ciencias e Ingenierías Físicas y Formales**  
**Escuela Profesional de Ingeniería Electrónica**



**DISEÑO E IMPLEMENTACIÓN DE UN PROTOTIPO DE “DIGITAL TWIN” APLICADO A UNA PLANTA PSA GENERADORA DE OXÍGENO, CON EL USO DE INTERNET DE LAS COSAS.**

Tesis presentada por el bachiller:  
**Núñez Melgar Obregón, Jorge Luis**

Para optar por el título profesional de:  
**Ingeniero Electrónico**  
Con especialidad en:  
**Automatización y Control**

Asesor:  
**Ing. Rodríguez González, Pedro Alex**

**Arequipa – Perú**  
**2022**

UCSM-ERP

**UNIVERSIDAD CATÓLICA DE SANTA MARÍA**  
**INGENIERIA ELECTRONICA**  
**CON ESPECIALIDAD EN AUTOMATIZACIÓN Y CONTROL**  
**TITULACIÓN CON TESIS**  
**DICTAMEN APROBACIÓN DE BORRADOR**

Arequipa, 14 de Julio del 2022

**Dictamen: 005023-C-EPIE-2022**

Visto el borrador del expediente 005023, presentado por:

**2015222181 - NUÑEZ MELGAR OBREGON JORGE LUIS**

Titulado:

**DISEÑO E IMPLEMENTACIÓN DE UN PROTOTIPO DE "DIGITAL TWIN" APLICADO A UNA  
PLANTA PSA GENERADORA DE OXÍGENO, CON EL USO DE INTERNET DE LAS COSAS.**

Nuestro dictamen es:

**APROBADO**

**1546 - DELGADO BARRA LUCY ANGELA  
DICTAMINADOR**



**1843 - URRUTIA ESPINOZA MARIO WILLIAM  
DICTAMINADOR**



**2465 - ZEGARRA GAGO HENRY CHRISTIAN  
DICTAMINADOR**

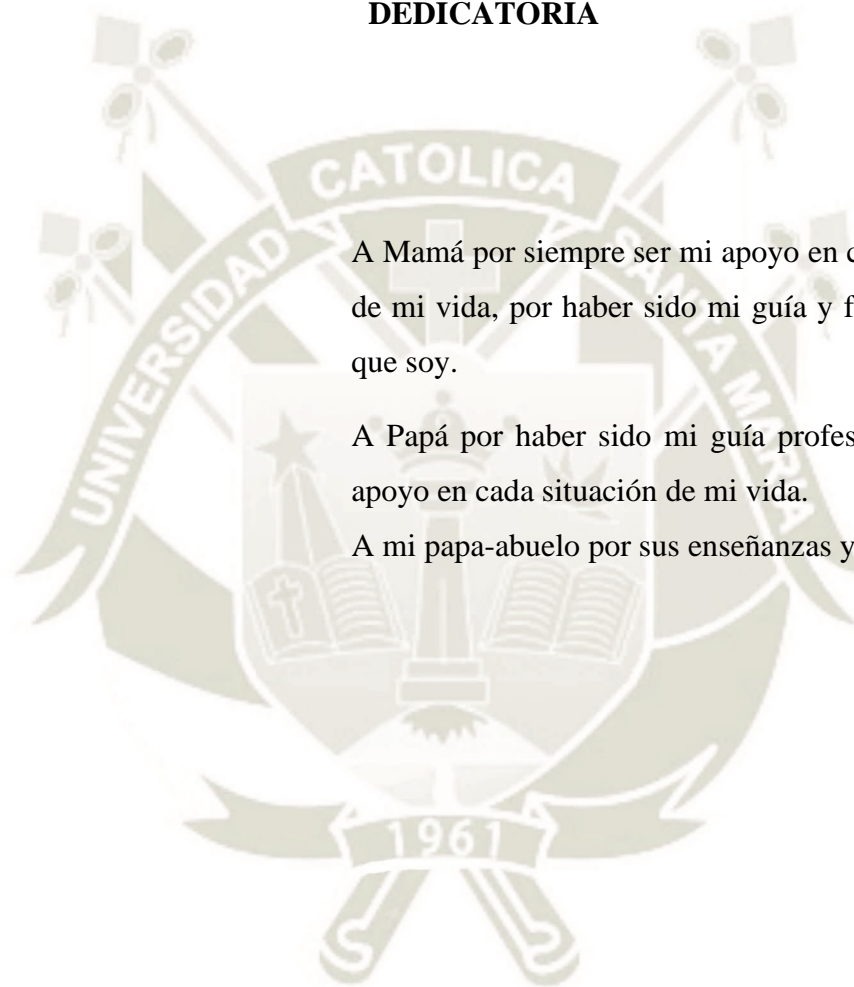


## DEDICATORIA

A Mamá por siempre ser mi apoyo en cada circunstancia de mi vida, por haber sido mi guía y forjar a la persona que soy.

A Papá por haber sido mi guía profesional y darme su apoyo en cada situación de mi vida.

A mi papa-abuelo por sus enseñanzas y ejemplo de vida.



## RESUMEN

El presente trabajo de investigación trata sobre el diseño e implementación de un “Digital Twin”. Empezando con el diseño del activo físico, desde la integración de sensores previamente instalados en el lugar, pasando por la programación usada en todos los controladores.

Posteriormente se diseña el activo digital, escogiendo el procesador teniendo en cuenta su mínimo costo y consumo de energía, pasando por la selección de servicios y la arquitectura que estos usaron, para finalmente interconectarlos.

Luego se generaron dashboards para poder observar con claridad la información recogida por los sensores y enviadas previamente, como muestra de esto también se implementa un sistema SCADA, reforzando la facilidad de integración que trae el “Digital Twin”.

También como parte de la investigación se desarrolla un modelo de Machine Learning como retroalimentación del “Digital Twin” hacia el activo físico, mejorando con esto la automatización del proceso de generación de oxígeno, regulando el flujo para poder mantener la pureza de oxígeno

**PALABRAS CLAVES:** Digital Twin, Industria 4.0, Machine Learning, automatización, Planta de oxígeno PSA.

## ABSTRACT

This research work deals with the design and implementation of a "Digital Twin". Starting with the design of the physical asset, from the integration of sensors previously installed on site, through the programming used in all the controllers.

Subsequently, the digital asset is designed, choosing the processor considering its minimum cost and energy consumption, going through the selection of services and the architecture they used, to finally interconnect them.

Then boards were generated to be able to clearly observe the information collected by the sensors and sent previously, as an example of this a SCADA system is also implemented, reinforcing the ease of integration that the "Digital Twin" brings.

Also, as part of the research, a Machine Learning model is developed as comments from the "Digital Twin" towards the physical asset, thereby improving the automation of the oxygen generation process, regulating the flow to maintain oxygen purity.

**KEY WORDS:** Digital Twin, Industry 4.0, Machine Learning, Automation, PSA oxygen plant.

## INTRODUCCIÓN

La nueva tecnología de “Digital Twins” nos permite comprender lo que está sucediendo en la línea de producción y predecir lo que sucederá en el futuro, lo que es esencial para maximizar los indicadores de rendimiento, optimizar la productividad y mejorar la rentabilidad comercial. Todas estas cosas son posibles gracias a las tecnologías de “Digital Twins”. Es por eso por lo que esta investigación se centra en el desarrollo de un prototipo para poder comprobar la veracidad de la información recogida.

El concepto de “Digital Twin” se introdujo hace más de una década, como una herramienta innovadora que lo abarca todo, con beneficios percibidos que incluyen monitoreo, simulación y pronóstico en tiempo real. Sin embargo, el marco teórico y las implementaciones prácticas de los Digital Twin (DT) aún están lejos de esta visión. Aunque existen implementaciones exitosas, los detalles de implementación no están disponibles públicamente, por lo tanto, es difícil evaluar su efectividad, establecer comparaciones y promover conjuntamente la metodología DT.

La siguiente investigación tiene por propósito revisar los múltiples conceptos de “Digital Twin” encontrados en la literatura, posteriormente realizar el diseño teórico de este para luego implementarlo físicamente, poniendo a prueba su rendimiento mediante la obtención masiva de datos y la automatización del proceso en una planta generadora de oxígeno PSA (Adsorción por Oscilación de Presión).

En el capítulo I se define el planteamiento del problema, discutiendo los objetivos principales y específicos de la investigación, así como distintas variables propias de esta.

En el capítulo II se revisa los conceptos y aplicaciones encontrados en la literatura acerca del “Digital Twin”, se realizan las revisiones categóricas de papers, además se da una revisión de los retos que presenta la implementación de un “Digital Twin”.

En el capítulo III se desarrolla el modelo teórico de la implementación del “Digital Twin”, repasando sus componentes elementales e imperativos, haciendo una revisión de sus propiedades. En el capítulo IV se revisa la implementación tanto física como digital del “Digital Twin”, mediante la revisión de los diagramas de conexión, programación en el activo físico e integración de servicios en el activo digital, posteriormente se muestran los resultados obtenidos.

En el capítulo V se repasan las conclusiones obtenidas de esta investigación y se dan las recomendaciones a ser usadas por los siguientes tesisistas en este tema.

## GLOSARIO DE TÉRMINOS

API: Application programming interface (Interfaz de programación de aplicaciones) CPU: Central Process Unit (Unidad central de procesos)

DDoS: Distributed Denial of Service (Denegación de servicio distribuida) DT: Digital Twin (Gemelo Digital)

GPU: Graphic Process Unit (Unidad gráfica de procesos) GUI: Graphic Interface User (Interfaz gráfica de usuario) IA: Inteligencia Artificial (Inteligencia Artificial)

IEEE: Institute of Electric and Electronics Engineers (Instituto de Ingenieros Eléctricos y Electrónicos)

IIoT: Industrial Internet of Things (Internet de las cosas industrial) IoT: Internet of things (Internet de las cosas)

IP: Internet Protocol (Protocolo de internet) M2M: Machine to machines (Máquina a máquina) MCU: Micro controller unit (Microcontrolador)

MQTT: Message Queue Telemetry Transport (Transporte de telemetría de cola de mensajes)

NASA: National Air Space Administration (Administración Nacional del Espacio y Aire)

OGP: Oxygen Generator Plant (Planta generadora de oxígeno)

PLC: Programmable Logic Controller (Controlador Lógico Programable) PoE: Power over Ethernet (Alimentación por internet)

PSA: Pressure Swing Adsorption (Presión Oscilación Adsorción) RAM: Random Access Memory (Memoria de acceso aleatorio) RF: Radiofrequency (Radiofrecuencia)

RGPD: Reglamento General de Protección de Datos

RTOS: Real Time Operating System (Sistema Operativo de Tiempo Real)

SCADA: Supervisory Control and Data Acquisition (Control de supervisión y Adquisición de Datos)

SDK: Software development kit (Kit de desarrollo de software)

SQL: Sequence Query Language (Lenguaje de secuencia de consultas)

SSD: Solid State Drive (Disco de estado sólido) SSH: Secure Shell (Escudo seguro)

SSL: Secure Sockets Layer (Capa de sockets seguros)

TCP: Transmission Control Protocol (Protocolo de Control de Transmisión) TI: Tecnologías de la información

TLS: Transport Layer Security (Capa de transporte de seguridad) TSDB: Time Series Data Base (Base de tiempo de series temporales)

## INDICE

CAPÍTULO I PLANTEAMIENTO DE LA INVESTIGACIÓN .....	20
1.1 Identificación y descripción del problema .....	20
1.2 Objetivos .....	20
1.2.1 Objetivo general.....	20
1.2.2 Objetivos específicos .....	20
1.3 Variables.....	2
1.3.1 Cuantitativas.....	2
1.3.2 Cualitativas.....	2
1.4 Alcances .....	2
1.5 Aportaciones.....	3
CAPÍTULO II PLANTEAMIENTO TEÓRICO .....	4
2.1 Digital Twin .....	4
2.1.1 Definición de “Digital Twin” .....	4
2.1.2 Conceptos Erróneos de Digital Twin.....	5
2.1.2.1 Modelo digital.....	5
2.1.2.2 Sombra digital.....	5
2.1.2.3 Digital Twin.....	5
2.1.3 Aplicaciones de los Digital Twin .....	6
2.1.3.1 Ciudades inteligentes .....	7
2.1.3.2 Fabricación .....	7
2.1.3.3 Sanidad .....	9
2.1.3.4 Digital Twin en la industria .....	10
2.2 Retos.....	11
2.2.1 Retos de la analítica de datos.....	11
2.2.1.1 Infraestructura de TI .....	11

2.2.1.2	Datos .....	11
2.2.1.3	Privacidad y seguridad.....	12
2.2.1.4	Confianza.....	12
2.2.1.5	Expectativas .....	13
2.2.2	Retos de la IoT/IIoT.....	13
2.2.2.1	Datos, privacidad, seguridad y confianza .....	13
2.2.2.2	Infraestructura.....	14
2.2.2.3	Conectividad.....	14
2.2.2.4	Expectativas.....	15
2.2.3	Retos del Digital Twin.....	15
2.2.3.1	Infraestructura de TI .....	15
2.2.3.2	Datos útiles .....	16
2.2.3.3	Privacidad y seguridad.....	16
2.2.3.4	Confianza.....	16
2.2.3.5	Expectativas.....	17
2.2.3.6	Modelización estandarizada .....	17
2.2.3.7	Modelización del dominio .....	17
2.3	Tecnologías facilitadoras.....	18
2.3.1	Breve historia del internet de las cosas .....	18
2.3.2	Breve historia del internet industrial de las cosas.....	20
2.3.3	Tecnologías facilitadoras de IIoT/IoT y bloques funcionales .....	21
2.3.4	Breve historia de la analítica de datos .....	23
2.3.4.1	Datos .....	23
2.3.4.2	Estadística.....	24
2.3.4.3	Inteligencia artificial.....	24
2.3.4.3.1	Aprendizaje automático .....	24
2.3.4.3.2	Aprendizaje supervisado.....	24

2.3.4.3.3	Aprendizaje no supervisado.....	25
2.3.4.3.4	Aprendizaje profundo .....	25
2.3.4.4	Visualización de datos .....	25
2.3.5	Tecnologías facilitadoras y bloques funcionales para el análisis de datos 26	
2.4	Tecnologías habilitadoras para el Digital Twin .....	27
2.5	Investigación actual.....	29
2.5.1	Metodología de revisión categórica.....	29
2.5.2	Organización de la revisión categórica.....	30
2.5.2.1	Paper .....	30
2.5.2.2	Tipo.....	30
2.5.2.3	Definición de gemelo.....	30
2.5.2.4	Gemelo real.....	30
2.5.2.5	Ámbito general .....	30
2.5.2.6	Área específica.....	30
2.5.2.7	Tecnología .....	31
2.5.3	Análisis de los documentos revisados. ....	31
2.5.3.1	Salud .....	32
2.5.3.2	Ciudades inteligentes.....	33
2.5.3.3	Fabricación .....	35
2.5.3.3.1	Reseñas sobre fabricación inteligente.....	35
2.5.3.3.2	Simulación e inteligencia artificial .....	36
2.5.3.3.3	Diseño y desarrollo de sistemas.....	37
2.5.3.3.4	Fabricación eficiente.....	39
2.6	Investigación abierta.....	40
2.6.1	Digital Twin en la fabricación .....	40
2.6.2	Digital Twin en salud .....	41
2.6.3	Digital Twin en ciudades inteligentes.....	43

2.6.4	Necesidades comunes y hallazgos específicos. ....	43
2.6.4.1	Modelos de datos .....	43
2.6.4.2	Sistemas heterogéneos .....	44
2.6.4.3	Inteligencia artificial .....	44
2.6.4.4	Seguridad .....	44
2.6.4.5	Intercambio de datos .....	44
2.6.4.6	Internet de las cosas .....	45
2.6.5	Desafíos de la investigación abierta .....	45
2.6.5.1	Multidisciplinariedad .....	45
2.6.5.2	Estandarización .....	45
2.6.5.3	Avances globales .....	45
<b>CAPÍTULO III DISEÑO TEÓRICO DEL DIGITAL TWIN .....</b>		<b>46</b>
3.1	Componentes elementales .....	46
3.1.1	Activo físico .....	46
3.1.1.1	Breve descripción del funcionamiento del activo físico .....	46
3.1.1.2	Lista detallada de parámetros a extraer .....	47
3.1.1.3	Arquitectura usada para la obtención de oxígeno .....	50
3.1.1.3	Arquitectura usada para la extracción de datos ESP8266 .....	54
3.1.1.4	Arquitectura usada para la extracción de datos Zigbee .....	55
3.1.1.5	Selección de componentes .....	56
3.1.1.5.1	PLC Siemens S7-1200 .....	56
3.1.1.6.2.	Expansor de entradas analógicas con comunicación Modbus TCP.....	57
3.1.1.6.3.	ESP8266 .....	57
3.1.1.6.4.	Zigbee Bridge .....	59
3.1.1.6	Comunicación MQTT .....	59
3.1.1.6.1	Funcionamiento de MQTT .....	60
3.1.1.6.2	Broker MQTT .....	60

3.1.2	Activo digital .....	60
3.1.2.1	Hardware principal ¿Por qué Raspberry? .....	60
3.1.2.1.1	Consumo .....	62
3.1.2.1.2	Velocidad .....	62
3.1.2.2.	Docker y los contenedores .....	62
3.1.2.2.1.	Contenedores Docker y la necesidad de usarlos .....	63
3.1.2.2.2.	Resolución de problemática .....	63
3.1.2.3.	Contenedores a usar .....	64
3.1.2.3.1.	Portainer .....	64
3.1.2.3.2.	EMQX .....	65
3.1.2.3.3.	Influxdb .....	65
3.1.2.3.4.	Zigbee2MQTT .....	66
3.1.2.3.5.	Telegraf .....	66
3.1.2.3.6.	Traefik .....	66
3.1.2.3.7.	Nodered .....	66
3.1.3	Relación sincronizada bidireccional entre ambos .....	67
3.2.	Componentes imperativos .....	67
3.2.1.	Dispositivos IoT .....	67
3.2.1.1.	Librería MQTT S7-1200 .....	67
3.2.1.2.	Ventajas del ESP8266 .....	67
3.2.1.3.	Ventajas de Zigbee .....	68
3.2.1.4.	Wi-Fi vs Zigbee .....	69
3.2.1.4.1.	Alcance inalámbrico .....	69
3.2.1.4.2.	Tipo de red .....	70
3.2.1.4.3.	Diferencias entre redes .....	70
3.2.1.4.4.	Velocidad y ancho de banda .....	71
3.2.1.4.5.	Interferencias .....	71

3.2.1.4.6.	Pruebas de velocidad en Zigbee .....	72
3.2.1.4.7.	Selección de dispositivos .....	72
3.2.2.	Herramientas de almacenamiento .....	73
3.2.2.1.	Bases de datos de series temporales vs relacionales .....	73
3.2.3.	Aprendizaje automático .....	75
3.2.4.	Seguridad .....	78
3.2.4.1.	SSH .....	78
3.2.4.2.	Claves públicas/privadas SSH .....	79
3.2.3.3.	Autenticación MQTT .....	79
3.2.3.4.	SSL en proxy inverso .....	81
3.2.5.	Propiedades de un DT .....	82
3.2.5.1.	Propiedades necesarias .....	82
3.2.5.2.	Propiedades dinámicas .....	82
3.3.	Relación entre el flujo de oxígeno y la pureza .....	83
<b>CAPÍTULO IV IMPLEMENTACIÓN DEL DIGITAL TWIN .....</b>		<b>85</b>
4.1.	Activo físico .....	85
4.1.1.	Plano de conexiónado de los transmisores .....	85
4.1.2.	Breve explicación de la programación .....	86
4.1.3.	Programación LADDER para la recolección de data mediante cliente Modbus	87
4.1.4.	Programación LADDER para publicación MQTT .....	89
4.1.5.	Programación MQTT en ESP8266 .....	92
4.2.	Activo digital .....	93
4.2.1.	Monitoreo por influxdata DB .....	94
4.2.2.	Componentes de la plataforma influxDB .....	96
4.2.3.	Raspberry Pi 4B+ y componentes adicionales .....	97
4.2.4.	Instalación docker .....	98
4.2.5.	Instalación de contenedores .....	99

4.2.5.1.	Portainer.....	99
4.2.5.2.	EMQX .....	101
4.2.5.3.	Influxdb .....	103
4.2.5.3.1.	Dashboards .....	105
4.2.5.4.	Telegraf.....	110
4.2.5.4.1.	Configuración .....	111
4.2.5.5.	Node-Red.....	112
4.2.5.6.	Configuración Node-Red.....	112
4.2.5.7.	Zigbee2MQTT.....	115
4.2.5.7.1.	Configuración MQTT.....	117
4.2.5.8.	Traefik. ....	118
4.2.5.8.1.	Necesidad de uso del proxy inverso. ....	120
4.2.5.9.	Loud ML.....	121
4.2.5.10.	Rapid SCADA.....	124
4.2.5.10.1.	Dashboards. ....	124
4.2.5.10.2.	Reportes. ....	127
4.2.5.10.3.	Alarmas.....	128
4.2.6.	Costos y presupuesto.....	128
4.2.7	Indicadores clave de desempeño y producción diaria.....	129
CONCLUSIONES .....		130
RECOMENDACIONES .....		132
REFERENCIAS .....		133
ANEXOS.....		142
1.	Diagramas de flujo .....	142
1.1.	Diagrama de flujo principal usado en la programación de la Raspberry. 142	
1.1.1.	Subrutina configuración bróker MQTT.....	143
1.1.2.	Subrutina configuración red Zigbee .....	143

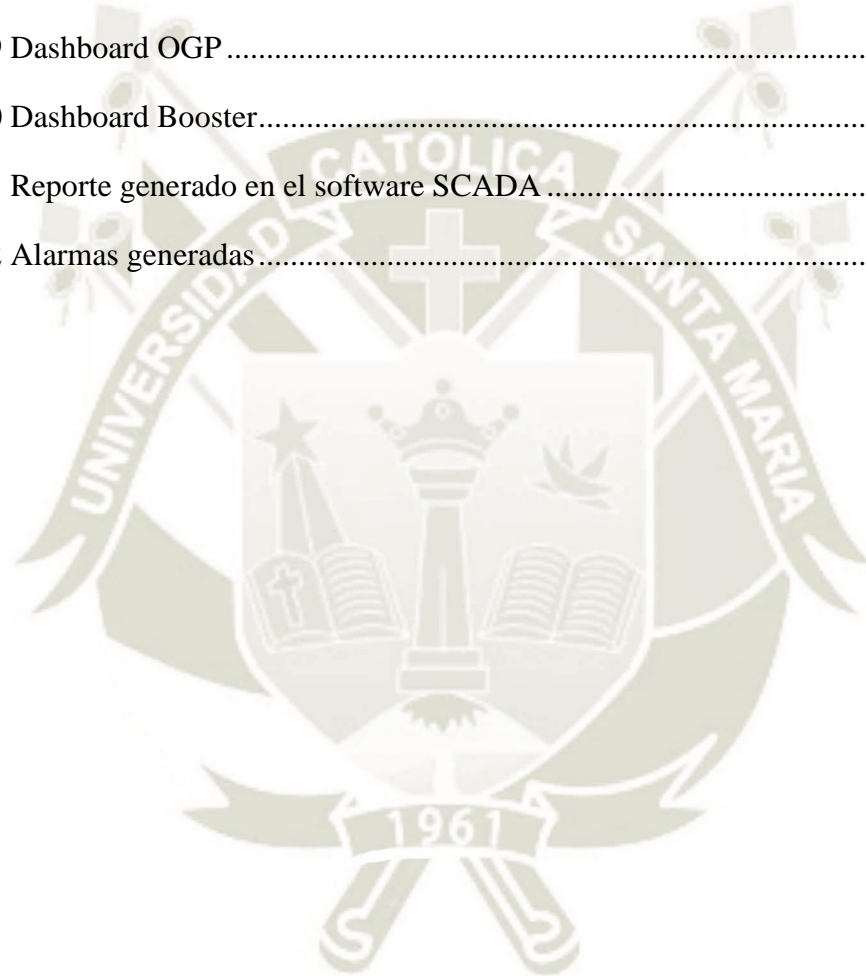
1.1.3.	Subrutina lectura de datos.....	144
1.1.4.	Subrutina almacenamiento de datos. ....	144
1.1.5.	Subrutina procesamiento por machine learning.....	145
1.1.6.	Subrutina envío de datos.....	145
1.1.7.	Diagrama de flujo principal usado en PLC S7-1200.....	146
1.1.8.	Subrutina adquisición de datos Modbus. ....	147
1.1.9.	Subrutina cambio de registros.....	148
1.1.10.	Subrutina envío de datos MQTT. ....	149
1.1.11.	Subrutina recepción de datos MQTT.....	149
1.1.12.	Subrutina regulación de flujo.....	150
1.1.13.	Diagrama de flujo principal usado en ESP8266.....	150
1.1.14.	Subrutina configurar I2C. ....	150
1.1.15.	Subrutina configurar SPI. ....	151
1.1.16.	Subrutina configurar Onewire. ....	152
1.1.17.	Subrutina enviar datos. ....	153

## INDICE DE FIGURAS

Figura 1 Modelo, sombra y gemelo Digital .....	6
Figura 2 Diagrama internet de las cosas.....	19
Figura 3 Crecimiento de dispositivos IoT .....	19
Figura 4 Revoluciones industriales .....	21
Figura 5 Porcentaje de áreas de investigación en el análisis de papers.....	32
Figura 6 Diagrama de bloques de una planta de oxígeno PSA .....	50
Figura 7 Arquitectura usada en la extracción de datos Siemens S7-1200 .....	51
Figura 8 Modulo expansión entradas analógicas .....	53
Figura 9 Siemens S7-1200 .....	53
Figura 10 Arquitectura usada en la extracción de datos ESP8266.....	54
Figura 11 Arquitectura usada en la extracción de datos Zigbee .....	55
Figura 12 Tiempo de respuesta sensores Zigbee.....	72
Figura 13 Diagrama Esquemático de adquisición de datos.....	85
Figura 14 Cliente Modbus.....	88
Figura 15 Bloque de reasignación de variable .....	88
Figura 16 Publicador MQTT.....	89
Figura 17 Funcion reasignacion de variable .....	90
Figura 18 Base de datos para comunicación MQTT.....	91
Figura 19 Variables para MQTT.....	92
Figura 20 Funciones para comunicación MQTT .....	92
Figura 21 Funciones de inicialización MQTT .....	93
Figura 22 Función para reconectarse al bróker .....	93
Figura 23 Arquitectura de la plataforma Influxdata a implementar.....	95
Figura 24 Raspberry Pi 4B+.....	97
Figura 25 Fuente de alimentación Raspberry Pi .....	97

Figura 26 Disco SSD Portable .....	98
Figura 27 Instalación de Docker mediante comandos .....	98
Figura 28 Sección del código docker compose correspondiente a Portainer .....	99
Figura 29 GUI Portainer.....	100
Figura 30 Sección del código Docker compose correspondiente a EMQX.....	101
Figura 31 GUI EMQX.....	102
Figura 32 Sección del código Docker compose correspondiente a Influxdb.....	103
Figura 33 GUI Influxdb.....	104
Figura 34 Dashboard compresor de tornillo.....	105
Figura 35 Dashboard OGP .....	106
Figura 36 Dashboard Booster.....	107
Figura 37 Dashboard estado del sistema .....	108
Figura 38 Revisión de alarmas programadas .....	109
Figura 39 Sección del código docker compose correspondiente a Telegraf .....	110
Figura 40 Configuración Telegraf plugin Influxdb.....	111
Figura 41 Configuración Telegraf plugin MQTT .....	111
Figura 42 Sección del código docker compose correspondiente a Nodered.....	112
Figura 43 Sección del código docker compose correspondiente a Nodered.....	112
Figura 44 GUI Nodered.....	113
Figura 45 Debugging de datos MQTT .....	114
Figura 46 Sección del código docker compose correspondiente a Zigbee2MQTT .....	115
Figura 47 GUI Zigbee2MQTT Mapa de conexión .....	116
Figura 48 Configuración MQTT en Zigbee2MQTT.....	117
Figura 49 Asignación de Pan ID para la seguridad en la red Zigbee .....	117
Figura 50 Clave de red Zigbee .....	118
Figura 51 Sección del código docker compose correspondiente a Traefik.....	118
Figura 52 GUI Traefik.....	119

Figura 53 Certificados SSL generados automáticamente para cada servicio.....	120
Figura 54 Menú entrenamiento Loud ML.....	121
Figura 55 Entrenamiento en proceso.....	122
Figura 56 Entrenamiento completado .....	123
Figura 57 Funcion predictiva Loud ML lista para utilizar .....	123
Figura 58 Dashboard Compresor .....	124
Figura 59 Dashboard OGP .....	125
Figura 60 Dashboard Booster.....	126
Figura 61 Reporte generado en el software SCADA .....	127
Figura 62 Alarmas generadas.....	128



## INDICE DE TABLAS

Tabla 1 Retos compartidos.....	18
Tabla 2 Tecnologías de habilitación y bloques funcionales: I/IoT. ....	21
Tabla 3 Tecnologías de habilitación y bloques funcionales: Análisis de datos .....	26
Tabla 4 Tecnologías de habilitación y bloques funcionales: Digital Twin .....	28
Tabla 5 Revisión categórica .....	31
Tabla 6 Parámetros a extraer.....	47
Tabla 7 Lista de dispositivos usados en la extracción de datos Siemens S7-1200 .....	52
Tabla 8 Lista de dispositivos usados en la extracción de datos ESP8266.....	55
Tabla 9 Lista de dispositivos usados en la extracción de datos Zigbee .....	56
Tabla 10 Costos y presupuesto.....	128
Tabla 11 KPIs de operaciones.....	129

## CAPÍTULO I PLANTEAMIENTO DE LA INVESTIGACIÓN

### 1.1 Identificación y descripción del problema

Ante la entrada de la Industria 4.0 a nuestra actualidad se requiere un modelo capaz de reunir la cantidad inmensa de datos por recolectar de miles de sensores de Internet de las Cosas (IoT) que recolectan datos de una fábrica industrial, en el caso de esta investigación una planta generadora de oxígeno de tipo PSA.

Las tecnologías actuales, tanto de información como de comunicación no están diseñadas para poder administrar todo el tráfico entrante de la data sensada, sin embargo, mediante nuevas tecnologías, estas se pueden adaptar para el desarrollo del internet de las cosas e Industria 4.0.

El modelo para la adquisición de datos, procesamiento, almacenamiento y toma decisiones es conocido como Digital Twin, el cual es una representación digital de un ambiente o proceso; el cual tiene una retroalimentación bidireccional para estar totalmente conectado con su gemelo físico cual permite tener un conocimiento constante de la situación, evaluar posibles escenarios futuros mediante una extrapolación de datos anteriores.

Sin embargo, el concepto de Digital Twin tampoco se encuentra del todo claro en la literatura estudiada, se tienen distintos conceptos e ideas erróneas que se repasarán a lo largo de esta investigación.

La presente investigación busca esclarecer el significado del concepto de Digital Twin y también diseñar e implementar uno para mostrar los beneficios que traerá al avance de la industria 4.0.

### 1.2 Objetivos

#### 1.2.1 Objetivo general

Diseñar e implementar un “Digital Twin” de una planta PSA generadora de oxígeno, que pueda mejorar la automatización del proceso, reflejando está en el incremento de producción de balones de oxígeno.

#### 1.2.2 Objetivos específicos

- Integrar al “Digital Twin” una red de sensores de tecnología Zigbee, WiFi y Modbus para la adquisición de datos y accionamientos necesarios.
- Diseñar e implementar un servidor de automatización en la plataforma Raspberry Pi4B.

- Programar e implementar una base de datos en la plataforma Influx db capaz de almacenar la data entrante.
- Implementar una interfaz gráfica de fácil entendimiento para mostrar los datos.
- Integrar todos los sistemas mencionados anteriormente.

### 1.3 Variables

#### 1.3.1 Cuantitativas

- Consumo energético de sensores WiFi y Zigbee.
- Velocidad en la red en malla Zigbee y en red estrella WiFi.
- Consumo eléctrico de un servidor basado en la plataforma Raspberry.

#### 1.3.2 Cualitativas

- Ventajas de la implementación de servicios mediante contenedores ante la forma tradicional.
- Ventajas de la implementación de bases de datos de series temporales frente a las bases de datos relacionales, en el uso de internet de las cosas.
- Mejoras en la historización de datos y la automatización del edificio mediante el “Digital Twin”.

### 1.4 Alcances

Mediante este proyecto se aspira desarrollar una representación virtual de un sistema denominado “Digital Twin” capaz de entregarnos información del entorno y analizar esta para tener respuestas a los posibles escenarios que se desarrollen en el futuro. Los alcances que se buscan tener con este proyecto son los siguientes:

- Diseñar e implementar una red de sensores inteligentes de tecnología Zigbee y WiFi capaces de comunicarse con el servidor de automatización para la adquisición de datos.
- Diseñar e implementar un servidor de automatización en la plataforma Raspberry Pi, debido a su bajo costo de adquisición y consumo eléctrico reducido.
- Diseñar e implementar una plataforma de visualización (Dashboard) en donde se puedan almacenar los datos recibidos, previamente almacenado en una base de datos, por los sensores y al mismo tiempo visualizar estos.
- Implementar algoritmos de control que se ejecuten con los datos adquiridos con los sensores y puedan a su vez accionar distintos actuadores para lograr las respuestas necesarias a posibles escenarios.

## 1.5 Aportaciones

Las aportaciones al desarrollo de la ingeniería brindados por esta investigación son las siguientes:

- Análisis de los distintos conceptos de Digital Twin y obtención de los componentes teóricos para la implementación de este.
- Implementación de un Digital Twin de un sistema industrial en el mundo real que contiene lo siguiente:
  - Diseño en instrumentación para la recolección de datos producidos por diferentes equipos propios de la generación de oxígeno.
  - Integración de protocolos industriales para la recolección de datos obtenidos por un autómata programable y sistemas embebidos.
  - Implementación de una base de datos de series temporales, debido a su adaptabilidad para la recolección de datos masivos y el Internet de las Cosas.
  - Ventajas informáticas a la hora de implementar servicios mediante contenedores Docker.
  - Implementación de software de Control Supervisor y Adquisición de Datos (SCADA) para el control y supervisión de distintos parámetros de la planta generadora de oxígeno.
  - Comparación entre protocolos de obtención de datos, evaluando su consumo energético y cantidad de transmisión de datos.

## CAPÍTULO II PLANTEAMIENTO TEÓRICO

### 2.1 Digital Twin

#### 2.1.1 Definición de “Digital Twin”

En esta sección se expone los orígenes del Digital Twin. En ella se exponen definiciones claras, al tiempo que se examinan algunos de los conceptos erróneos encontrados en los Digital Twin erróneamente identificados. Las ideas formales en torno a los Digital Twin existen desde principios de la década de 2000 [4]. No obstante, es posible que la definición de Digital Twin se haya hecho antes debido a la constante evolución de las definiciones. La primera terminología la dio Grieves en una presentación de 2003 y posteriormente se documentó en un libro blanco que establecía una base para el desarrollo de los Digital Twin [4]. La Administración Nacional del Espacio Aeronáutico (NASA) publicó en 2012 un documento titulado "The Digital Twin Paradigm para los futuros vehículos de la NASA y la Fuerza Aérea de los Estados Unidos", estableciendo un hito para la definición de los Digital Twin.

Nasa 2012 [5] "Un Digital Twin es una simulación integrada multifísica, multiescala y probabilística de un vehículo o sistema tal y como está construido que utiliza los mejores modelos físicos disponibles, actualizaciones de sensores, historial de la flota, etc., para reflejar la vida de su correspondiente gemelo volador."

Chen 2017 [6] "Un Digital Twin es un modelo computarizado de un dispositivo o sistema físico que representa todas las características funcionales y los vínculos con los elementos de trabajo."

Liu et al. 2018 [7] "El Digital Twin es en realidad un modelo vivo del activo o sistema, que se adapta continuamente a los cambios operativos cambios basados en los datos e información en línea recopilados, y puede predecir el futuro de la correspondiente contraparte física".

Zheng et al. 2018 [8] "Un Digital Twin es un conjunto de información virtual que describe completamente una producción física potencial o real desde el nivel micro atómico hasta el nivel macro geométrico."

Vrubic et al. 2018 [9] "Un Digital Twin es una representación digital de un artículo o conjunto físico que utiliza simulaciones integradas y datos de servicio. La representación digital contiene información de múltiples fuentes a lo largo del ciclo de vida del producto. Esta información se actualiza continuamente y se visualiza en una

variedad de formas para predecir las condiciones actuales y futuras, tanto en entornos de diseño como operativos, para mejorar la toma de decisiones."

Mandi 2019 [10] "Un Digital Twin es una instancia virtual de un sistema físico (gemelo) que se actualiza continuamente con los datos de rendimiento, mantenimiento y estado de salud de este último a lo largo del ciclo de vida del sistema físico." [10] La definición a) es una definición ambigua específica para el desarrollo de vehículos interplanetarios de la NASA [5] y es uno de los primeros artículos que definen los Digital Twin. A pesar de que han pasado más de seis años entre las publicaciones a) y f), el consenso sigue siendo que no hay un cambio fundamental o significativo. Tanto el mundo académico como la industria no han ayudado a distinguir los DT de los modelos informáticos generales y las simulaciones. El trabajo futuro requiere una definición más definitiva de Digital Twin. Esta investigación pretende contribuir al desarrollo de una definición actualizada, al tiempo que ayuda a analizar los trabajos relacionados y a señalar los Digital Twin erróneamente identificados.

## **2.1.2 Conceptos Erróneos de Digital Twin**

### **2.1.2.1 Modelo digital**

Un modelo digital se describe como una versión digital de un objeto físico preexistente o previsto, para definir correctamente un modelo digital no debe haber un intercambio automático de datos entre el modelo físico y el modelo digital. Ejemplos de un modelo digital podrían ser, entre otros, los planos de edificios, los diseños de productos y el desarrollo.

La característica más importante es que no hay ninguna forma de intercambio automático de datos entre el sistema físico y el modelo digital. Esto significa que, una vez creado el modelo digital, un cambio realizado en el objeto físico no tiene ningún impacto en el modelo digital. La figura 1 ilustra un modelo digital.

### **2.1.2.2 Sombra digital**

Una sombra digital es una representación digital de un objeto que tiene un flujo unidireccional entre el objeto físico y el digital. [5] Un cambio en el estado del objeto físico provoca un cambio en el objeto digital y no al revés. La figura 1 ilustra una sombra digital.

### **2.1.2.3 Digital Twin**

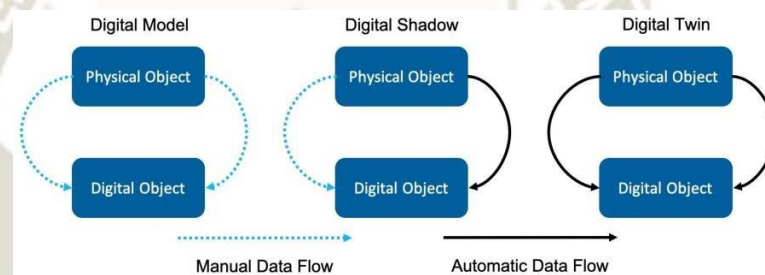
Si los datos fluyen entre un objeto físico existente y un objeto digital, y están totalmente integrados en ambas direcciones, esto constituye la referencia "Digital

Twin". Un cambio realizado en el objeto físico conduce automáticamente a un cambio en el objeto digital y viceversa. La figura 1 ilustra un Digital Twin.

Estas tres definiciones ayudan a identificar los conceptos erróneos más comunes que se observan en la literatura. Sin embargo, hay varios conceptos erróneos que no se limitan a estos ejemplos específicos. Entre los conceptos erróneos está la idea de que los Digital Twin tienen que ser un modelo 3D exacto de una cosa física. Por otro lado, se suele pensar que un Digital Twin es sólo un modelo 3D. [11]

La figura 1 y sus definiciones presentan los diferentes niveles de integración de un Digital Twin.

La tabla 5 de la sección 2.4 de esta revisión presenta una serie de publicaciones, en las que se destaca el nivel de integración declarado frente a la integración real basada en la definición anterior. Las definiciones y las cifras deberían ayudar al desarrollo y la identificación de futuros Digital Twin.



*Figura 1 Modelo, sombra y gemelo Digital*

*Nota: Fuente:*

*[https://www.researchgate.net/publication/337019778\\_Digital\\_Twin\\_Enabling\\_Technologies\\_Challenges\\_and\\_Open\\_Research](https://www.researchgate.net/publication/337019778_Digital_Twin_Enabling_Technologies_Challenges_and_Open_Research)*

### 2.1.3 Aplicaciones de los Digital Twin

La siguiente parte de esta investigación se centra en las aplicaciones de los Digital Twin. En primer lugar, se examinarán las posibles aplicaciones de los Digital Twin, analizando el ámbito, los sectores y los problemas específicos de la tecnología de Digital Twin.

Por el momento, el término y el concepto de Digital Twin están creciendo en el mundo académico, y los avances en IoT e inteligencia artificial (IA) están permitiendo que este crecimiento aumente [11] [12] [13] [14] [15] [16]. En este momento, las principales áreas de interés son las ciudades inteligentes y la fabricación, con algunas aplicaciones de la tecnología Digital Twin relacionadas con la sanidad.

### 2.1.3.1 Ciudades inteligentes

El uso y el potencial de los Digital Twin para ser dramáticamente efectivos dentro de una ciudad inteligente está aumentando año tras año debido a los rápidos desarrollos en la conectividad a través del IoT.

Con un número cada vez mayor de ciudades inteligentes desarrolladas, las comunidades más conectadas son, con esto viene más uso de Digital Twins.

No sólo esto, cuantos más datos recopilemos de los sensores IoT integrados en los servicios básicos de una ciudad, sino que también allanará el camino para la investigación destinada a la creación de algoritmos avanzados de IA [3] [17] [18]. La posibilidad de que los servicios e infraestructuras de una ciudad inteligente dispongan de sensores y sean supervisados con dispositivos de IoT es de gran valor para todo tipo de preparación para el futuro.

Puede utilizarse para ayudar en la planificación y el desarrollo de las ciudades inteligentes actuales y para ayudar a los desarrollos en curso de otras ciudades inteligentes. Además de los beneficios de la planificación, también hay beneficios dentro del mundo del ahorro energético. Estos datos ofrecen una excelente visión de cómo se distribuyen y utilizan nuestros servicios públicos. El avance para la ciudad inteligente es el potencial de utilizar la tecnología Digital Twin. Puede facilitar el crecimiento al ser capaz de crear un banco de pruebas vivo dentro de un gemelo virtual que puede lograr dos cosas; una, probar escenarios, y, dos, permitir que los Digital Twin aprendan del entorno analizando los cambios en los datos recogidos. Los datos recogidos pueden utilizarse para el análisis de datos y la supervisión. El alcance de los Digital Twin es cada vez más viable a medida que el desarrollo de las ciudades inteligentes aumenta la conectividad y la cantidad de datos utilizables [19] [20] [21] [22].

### 2.1.3.2 Fabricación

La siguiente aplicación identificada para los Digital Twin es dentro de un entorno de fabricación. La razón principal es que los fabricantes siempre buscan una forma de rastrear y supervisar los productos en un intento de ahorrar tiempo y dinero, un motor clave y una motivación para cualquier fabricante. De ahí que los Digital Twin parezcan tener el mayor impacto en este ámbito. Asimismo, con el desarrollo de una ciudad inteligente, la conectividad es uno de los mayores impulsores para que la fabricación utilice Digital Twins. El crecimiento actual está en consonancia con el concepto de Industria 4.0, acuñado como la cuarta revolución industrial, que aprovecha la

conectividad de los dispositivos para hacer realidad el concepto de Digital Twin en los procesos de fabricación [23] [1] [24] [25][26].

El Digital Twin tiene el potencial de ofrecer un estado en tiempo real sobre el rendimiento de las máquinas, así como información sobre la línea de producción.

Esto da al fabricante la capacidad de predecir los problemas antes. El uso de Digital Twin aumenta la conectividad y la retroalimentación entre los dispositivos, mejorando a su vez la fiabilidad y el rendimiento. Los algoritmos de IA acoplados a los Digital Twins tienen el potencial de una mayor precisión, ya que la máquina puede contener grandes cantidades de datos, necesarios para el análisis del rendimiento y la predicción. Los Digital Twin están creando un entorno para probar productos, así como un sistema que actúa sobre los datos en tiempo real, dentro de un entorno de fabricación esto tiene el potencial de ser un activo enormemente valioso [27] [28] [2] [29].

Otra aplicación de los Digital Twin es en la industria del automóvil, sobre todo la demostrada por Tesla. La capacidad de tener un Digital Twin de un motor o pieza de automóvil puede ser valiosa en términos de uso del gemelo para la simulación y el análisis de datos [30] [31]. La IA mejora la precisión de las pruebas, ya que puede realizar análisis de datos sobre los datos del vehículo en vivo para predecir el rendimiento actual y futuro de los componentes.

La industria de la construcción es otro sector que alberga una gama de aplicaciones para el uso de Digital Twin. La fase de desarrollo de un edificio o estructura es una aplicación potencial para un Digital Twin. La tecnología no sólo puede aplicarse en el desarrollo de edificios o estructuras de ciudades inteligentes, sino también como una herramienta de predicción y monitorización en tiempo real. El uso del Digital Twin y la analítica de datos proporcionará potencialmente mayor precisión a la hora de predecir y mantener los edificios y estructuras, ya que cualquier cambio realizado virtualmente se aplicará físicamente. El Digital Twin ofrece a los equipos de construcción una mayor precisión al realizar simulaciones, ya que los algoritmos pueden en tiempo real en el Digital Twin antes de que se construya el edificio físico.

Un objetivo común visto hasta ahora en el campo de los Digital Twin es la idea de la simulación en tiempo real, en contraposición a los estáticos de bajo nivel de detalle. El uso de estos modelos sirve para algo, pero no utilizan parámetros en tiempo real que limitan la previsibilidad y la capacidad de aprendizaje. El Digital Twin puede aprender y monitorizar simultáneamente, así como aplicar algoritmos de aprendizaje automático y profundo [32] [33] [34] [35].

### 2.1.3.3 Sanidad

El sector sanitario es otro ámbito de aplicación de la tecnología Digital Twin. El crecimiento y los desarrollos que la tecnología está teniendo en la asistencia sanitaria no tienen precedentes ya que lo que antes era imposible se está convirtiendo en posible. En términos de IoT los dispositivos son más baratos y fáciles de implementar, de ahí el aumento de la conectividad [36] [37]. El aumento de la conectividad no hace más que aumentar la aplicación potencial de en el sector sanitario. Una aplicación futura es un Digital Twin de un ser humano, que ofrezca un análisis en tiempo real del del cuerpo. Una aplicación actual más realista es un Gemelo Twin para simular los efectos de ciertos medicamentos. Otra aplicación es el uso de un Digital Twin para planificar y realizar procedimientos quirúrgicos [38].

De la misma manera, con otras aplicaciones dentro de un entorno sanitario, el uso de un Digital Twin ofrece a los investigadores, médicos hospitalares y proveedores de atención sanitaria la capacidad de simular entornos específicos para sus necesidades ya sea en tiempo real o con vistas a futuros desarrollos y usos. Además de esto, el Digital Twin puede utilizarse simultáneamente con algoritmos de IA para hacer predicciones y tomar decisiones más inteligentes. Muchas aplicaciones dentro de la atención sanitaria no incluyen directamente al paciente, pero son beneficiosas para el cuidado y el tratamiento en curso, de ahí el papel clave que tienen estos sistemas en la atención al paciente. DigitalTwin para la sanidad está en pañales, pero el potencial es enorme desde la gestión de camas hasta la gestión de salas y hospitales a gran escala.

Tener la capacidad de simular y actuar en tiempo real es aún más importante en la asistencia sanitaria, ya que puede suponer la diferencia entre la vida y la muerte. El Digital Twin también podría ayudar al mantenimiento predictivo y a la reparación continua de los equipos médicos. El Digital Twin en el entorno médico tiene el potencial, junto con la IA, de tomar decisiones que salven vidas basadas en datos históricos y en tiempo real [39] [40]. Aquí se identifican las aplicaciones de un Digital Twin, mostrando algunas de las aplicaciones cruzadas en el uso previsto, demostrando de mantenimiento predictivo es adaptable desde las máquinas de fabricación a la atención de pacientes.

También se muestran algunas de las aplicaciones en las que no se cruzan, y el uso de DigitalTwin es específico para su uso previsto. Los avances en IA, IoT e Industria 4.0 han facilitado el crecimiento de las aplicaciones de Digital Twin.

#### 2.1.3.4 Digital Twin en la industria

General Electric (GE) documentó por primera vez su uso de un Digital Twin en una solicitud de patente en 2016. A partir del concepto expuesto en la patente, desarrollaron una aplicación llamada plataforma "Predix" [41] que es una herramienta para crear DigitalTwin. Predix [41] se utiliza para ejecutar la analítica de datos y la monitorización. En los últimos años, GE ha reducido sus planes de crear un Digital Twin y ha decidido centrarse en su herencia como multinacional industrial en lugar de una empresa de software. Siemens, sin embargo, ha desarrollado una plataforma llamada "MindSphere" [42] que ha adoptado el concepto Industria 4.0 con un sistema basado en la nube que conecta las máquinas y la infraestructura física a un Digital Twin. Utiliza todos los dispositivos conectados y miles de millones de flujos de datos con la esperanza de transformar las empresas y proporcionar soluciones de Digital Twin [42]. Una plataforma alternativa para desarrollar tecnología de Digital Twin e IA es "ThingWorx" [22]. Esta plataforma, creada por PTC, es una plataforma de innovación industrial cuyo objetivo principal es recopilar datos de Internet Industrial de las Cosas (IIoT)/ Internet de las cosas (IoT) y presentarlos a través de una interfaz de usuario intuitiva y basada en roles que ofrece información valiosa a los usuarios. La plataforma facilita el desarrollo fluido de la analítica de datos, a la vez que desarrolla un entorno para una solución Digital Twin [22].

IBM ha desarrollado una plataforma llamada "Watson IoT Platform" [43] que se comercializa como una herramienta de datos de IoT completa que puede utilizarse para gestionar sistemas a gran escala, en tiempo real, a través de los datos recogidos de millones de dispositivos de IoT. La plataforma tiene varias características adicionales: servicios basados en la nube, análisis de datos, capacidades de borde y servicios de blockchain. Todo ello la convierte en una posible plataforma para un sistema Digital Twin [43].

Desde el punto de vista del código abierto, hay dos grandes proyectos a destacar. El primero es el proyecto "Ditto" de Eclipse [44], una plataforma lista para usar que puede gestionar los estados de un Digital Twin, dando acceso y control a los Gemelos físicos y digitales. La plataforma se encuentra en un papel de back-end que proporciona soporte a los dispositivos ya conectados y simplifica la conexión y gestión de los Digital Twin [44]. Otro proyecto de código abierto llamado "imodel.js" desarrollado por Bentley Systems [45] es una plataforma para crear, acceder y construir Digital Twin.

## 2.2 Retos

Cada vez es más evidente que los Digital Twin corren en paralelo con la tecnología de IA e IoT, lo que da lugar a retos compartidos. El primer paso para abordar los retos es identificarlos. Algunos de los retos comunes se encuentran tanto con la analítica de datos como con el Internet de las Cosas, y el objetivo final es identificar los retos compartidos para los Digital Twin.

### 2.2.1 Retos de la analítica de datos

A continuación, se enumeran algunos de los retos en el ámbito del aprendizaje automático y profundo.

#### 2.2.1.1 Infraestructura de TI

El primer gran reto es la infraestructura general de las Tecnologías de la Información (TI). El rápido crecimiento de la IA necesita ser atendido con una infraestructura de alto rendimiento en forma de hardware y software actualizado, para ayudar a ejecutar los algoritmos. Actualmente, el reto de la infraestructura se reduce al coste de instalación y funcionamiento de estos sistemas. Por ejemplo, el coste de las GPU de alto rendimiento que pueden ejecutar los algoritmos de aprendizaje automático y profundo es de miles de dólares, entre 1.000 y 10.000 dólares. Además de esto, la infraestructura necesita software y hardware actualizados para ejecutar dichos sistemas con éxito. La superación de este reto pasa por el uso de las GPU "como servicio", que proporcionan GPU a la carta y a precio de coste a través de la nube.

Amazon, Google, Microsoft y NVIDIA, por nombrar algunas, están ofreciendo servicios únicos bajo demanda similares a las aplicaciones tradicionales basadas en la nube, lo que rompe la barrera de la demanda, pero la escasa infraestructura y el elevado coste siguen siendo un reto para el análisis de datos. [55]

El uso de la nube para la analítica de datos y los Digital Twin sigue planteando desafíos a la hora de garantizar que la infraestructura de la nube ofrezca seguridad robusta.

#### 2.2.1.2 Datos

Desde el punto de vista de los datos, es importante asegurarse de que no son de baja calidad. Los datos deben clasificarse y limpiarse para garantizar que los algoritmos de IA reciban datos de la máxima calidad.

### 2.2.1.3 Privacidad y seguridad

La privacidad y la seguridad es un tema importante para cualquier persona relacionada con la industria de la informática y esto no es diferente cuando se realiza el análisis de datos. Las leyes y la normativa aún no se han establecido del todo debido a la infancia de la IA. El reto es un mayor escrutinio, regulación y medidas relativas a la IA en el futuro a medida que la tecnología crece. La futura regulación garantiza el desarrollo de algoritmos que tomen medidas para proteger los datos de los usuarios. El Reglamento General de Protección de Datos (RGPD) es una nueva normativa que garantiza la privacidad y la seguridad de los datos personales en el Reino Unido y en toda Europa. A pesar de ser un reglamento general relativo a los datos y la seguridad, esto pone de relieve las preocupaciones con el manejo de datos cuando se desarrollan algoritmos de IA. La regulación es un paso para garantizar la protección de los datos personales, mientras que otro método es el aprendizaje federado, un marco descentralizado para el entrenamiento de modelos. Permite que los datos de los usuarios en un modelo de aprendizaje permanezcan localizados sin que se compartan los datos, de datos, lo que permite abordar los problemas de privacidad y seguridad de datos en un Digital Twin. [60].

### 2.2.1.4 Confianza

La confianza es otro reto que preocupa a gran parte del campo de la IA. En primer lugar, porque es relativamente nueva y en segundo lugar porque a menos que el desarrollador esté familiarizado con la complejidad el uso de la IA puede ser desalentador. La ansiedad de que los robots y la IA se conviertan en una fuerza dominante en la Tierra, arrebatando a los humanos el control de las infraestructuras clave, es una barrera para la confianza. La cuestión de la confianza puede ser una barrera porque la representación de la IA se centra sobre todo en los efectos negativos que podrían producirse. Las historias positivas de los medios de comunicación en el campo de la inteligencia artificial son cada vez más comunes, pero el reto es evidente, y la necesidad de una mayor exposición de la IA y los usos positivos ayudaría a superar los problemas de confianza. Los problemas de privacidad y de seguridad contribuyen a estos problemas de confianza, pero una regulación más de la privacidad y la seguridad en la IA contribuye a generar confianza. [62]

### 2.2.1.5 Expectativas

El último reto de la analítica de datos es la expectativa de que pueda utilizarse para resolver todos nuestros problemas. Es vital considerar cuidadosamente el uso de la IA e invertir tiempo en ello para identificar la aplicación correcta, asegurando que los modelos estándar no puedan producir los mismos resultados. Al igual que otras nuevas tecnologías, tienen el potencial de trabajar mano a mano con el fortalecimiento de cosas como la fabricación y los desarrollos de las ciudades inteligentes. Los usuarios potenciales sólo ven los beneficios y creen que se ahorrará instantáneamente tiempo y dinero, de ahí las altas expectativas. El campo está todavía en sus inicios, y hay que tener en cuenta el reto que supone la aplicación de la analítica de datos.

Es evidente a través del número de escenarios que utilizan la "IA" para procesos que no los necesitan, en contraste con otras situaciones en las que debería utilizarse. Se necesita una mayor exposición y comprensión de la IA para que las personas adquieran los conocimientos básicos correctos del área, y así aprender cómo se puede aplicar.

### 2.2.2 Retos de la IoT/IIoT

A continuación, se enumeran los retos encontrados en el ámbito del Internet de las cosas y del Internet industrial de las cosas:

#### 2.2.2.1 Datos, privacidad, seguridad y confianza

Con el enorme crecimiento de los dispositivos IoT, tanto en el ámbito doméstico como en el industrial, surge el reto de recopilar cantidades considerables de datos. El reto consiste en tratar de controlar el flujo de datos, garantizando que puedan organizarse y utilizarse eficazmente.

El reto se convierte en un problema mayor con la llegada del big data. El uso de IoT aumenta los grandes volúmenes de datos no estructurados. Para que la IoT pueda gestionarla cantidad de datos, la clasificación y organización de estos es una necesidad y hará que más datos sean utilizables y aporten valor. De lo contrario, los datos recogidos a través de la IoT se perderán o será demasiado costoso extraer el valor de los enormes volúmenes acumulados. Como los datos podrían ser sensibles, podrían ser de valor para un delincuente, aumentando así la amenaza. La amenaza aumenta significativamente para las empresas cuando pueden estar tratando con datos sensibles de los clientes. Los ciberataques plantean más desafíos, ya que los delincuentes atacan los sistemas y los desconectan para paralizar la infraestructura

de una organización. Algunas organizaciones tienen miles de dispositivos IoT conectados, lo que supone un riesgo de que los ciberdelincuentes puedan atacarlos para tomar el control y utilizar los dispositivos para sus servicios. Un ejemplo de ello es el escándalo de la red de bots Mirai, en el que casi 15 millones de dispositivos IoT de todo el mundo fueron comprometidos y utilizados para lanzar un ataque de denegación de servicio distribuido (DDoS) [46].

El riesgo de ataques DDoS aumenta debido al rápido crecimiento del IoT. Además, el retraso en las prioridades en torno a las soluciones de privacidad y seguridad supone un riesgo adicional de ataque. A la hora de instalar los dispositivos, es necesario contar con las características de seguridad y protección más actualizadas; de lo contrario, se trata de una vulnerabilidad que ofrece una puerta trasera para que los delincuentes se infiltren en un entorno de IoT conectado más amplio.

#### **2.2.2.2 Infraestructura**

La infraestructura de TI en vigor está atrasada, debido al rápido crecimiento observado en la tecnología de la IoT en comparación con los sistemas existentes en la actualidad. La actualización de la infraestructura antigua y la integración de la nueva tecnología ayudan a facilitar el crecimiento de IoT. La infraestructura de IoT actualizada ofrece la oportunidad de beneficiarse de la última tecnología y aprovechar las aplicaciones y servicios disponibles en la nube sin una costosa actualización de los sistemas y la tecnología existentes.

Otro reto para los sistemas de IoT es la conexión de máquinas antiguas al entorno de IoT. Una de las formas de combatirlo es adaptar los sensores de IoT a las máquinas antiguas, asegurando que los datos no se desperdicien y que las máquinas antiguas no se pierdan que los datos no se desperdicien y que las máquinas antiguas puedan tener alguna forma de análisis. [30]

#### **2.2.2.3 Conectividad**

A pesar de este crecimiento en el uso de la IoT, los retos de la conectividad siguen existiendo. Estos son especialmente frecuentes cuando se trata de lograr el objetivo de la supervisión en tiempo real. Un gran número de sensores dentro de un proceso de fabricación supone un reto importante cuando se trata de conectarlos todos simultáneamente. Los problemas relacionados con atributos como los cortes de energía, los errores de software o los errores de despliegue en curso afectan a este objetivo general de conectividad. El mero hecho de que un sensor no esté totalmente

conectado podría afectar de forma drástica al objetivo general de un proceso determinado.

Por ejemplo, los dispositivos IoT son una fuente de alimentación de datos para los algoritmos de IA; esto puede convertirse en un reto importante, ya que se necesitan todos los datos para que funcionen con precisión y la falta de datos IoT podría afectar negativamente al funcionamiento del sistema. El reequipamiento de las máquinas y la recolección de los datos ya servidos por la máquina es un método para garantizar la recopilación de todos los datos. Los métodos de imputación son un proceso de búsqueda de valores de sustitución para los datos de los sensores de IoT que faltan, un concepto utilizado para garantizar la conectividad completa y facilitar la ejecución de modelos de IA con una gran precisión y sin apenas datos perdidos. [32]

#### **2.2.2.4 Expectativas**

Del mismo modo, con la IA, las expectativas asociadas a la IoT son un reto, debido a que las organizaciones y los usuarios finales no comprenden del todo qué esperar de las soluciones de IoT o cómo utilizarlas mejor. Un aspecto prometedor es que el rápido crecimiento de la IoT indica que los usuarios finales y las organizaciones reconocen el valor de la IoT y cómo un mundo conectado más inteligente puede beneficiarnos a todos. La expectativa de que la IoT puede utilizarse infinitamente sin conocimiento previo puede ser perjudicial, con el consiguiente efecto de aumentar la presión sobre la privacidad y la seguridad, lo que aumenta la carga de los problemas de confianza. Al igual que en el caso de la IA, es necesario un conocimiento previo de la IoT para que se aproveche todo su potencial. [35]

### **2.2.3 Retos del Digital Twin**

Esta sección se centra en los retos asociados a los Digital Twin. Sin embargo, a medida que avanza la investigación, es evidente que los desafíos encontrados en el análisis de datos, la IoT y la IIoT son similares a los encontrados en los desafíos de los Digital Twin, algunos de los cuales se discuten a continuación:

#### **2.2.3.1 Infraestructura de TI**

Al igual que en el caso de la analítica y el IoT, el reto está en la infraestructura informática actual. El Digital Twin necesita una infraestructura que permita el éxito de la IoT y la analítica de datos; esto facilitará el funcionamiento eficaz de un Digital Twin.

Sin una infraestructura de TI conectada y bien pensada, el Digital Twin no podrá ser eficaz en la consecución de sus objetivos.[52]

### **2.2.3.2 Datos útiles**

El siguiente reto se refiere a los datos necesarios para un Digital Twin. Deben ser datos de calidad, sin ruido y con un flujo de datos constante e ininterrumpido. Si los datos son deficientes e incoherentes, se corre el riesgo de que el Digital Twin no funcione bien, ya que actúa sobre datos deficientes y ausentes. La calidad y el número de señales de IoT es un factor esencial para los datos de Digital Twin. La planificación y el análisis del uso de los dispositivos son necesarios para identificar los datos correctos que se recogen y utilizan para un uso eficiente de un Digital Twin.

### **2.2.3.3 Privacidad y seguridad**

En el ámbito de la industria, está claro que la privacidad y la seguridad asociadas a los Digital Twin son un reto. En primer lugar, por la gran cantidad de datos que utilizan y, en segundo lugar, por el riesgo que esto supone para los datos sensibles del sistema. Para superar este reto, las tecnologías clave que permiten los Digital Twin -analítica de datos eIoT- deben seguir las prácticas y actualizaciones actuales en materia de seguridad y privacidad. La consideración de la seguridad y la privacidad de los datos de los Digital Twin contribuye a abordar los problemas de confianza con los Digital Twin. [40][42][44]

### **2.2.3.4 Confianza**

Los retos asociados a la confianza se plantean tanto desde el punto de vista de la organización como del usuario. La tecnología de los Digital Twin debe debatirse y explicarse más a fondo para garantizar que los usuarios finales y las organizaciones conozcan las ventajas de los Digital Twin, lo que permitirá superar el problema de la confianza.

La validación del modelo es otra forma de superar los retos de la confianza. Verificar que los Digital Twin funcionan como se espera es fundamental para garantizar la confianza de los usuarios.

Con una mayor comprensión, la confianza en los Digital Twin prevalece. La tecnología de habilitación proporcionará más información sobre los pasos que garantizan que se sigan las prácticas de privacidad y seguridad a través del desarrollo, superando a su vez los desafíos con la confianza. [49]

### **2.2.3.5 Expectativas**

A pesar de que los líderes de la industria, Siemens y GE están acelerando la adopción de Digital Twin, es necesario ser cauteloso para destacar los desafíos que existen para las expectativas de Digital Twins y la necesidad de una mayor comprensión. La necesidad de contar con bases sólidas para la infraestructura de IoT y una mayor comprensión de los datos necesarios para realizar análisis garantizará que las organizaciones hagan uso de la tecnología Digital Twin. También es un reto combatir el pensamiento de que el Digital Twin debe utilizarse únicamente por las tendencias actuales. Es necesario debatir los aspectos positivos y negativos de las expectativas de los Digital Twin para garantizar una acción adecuada cuando se desarrollen los sistemas de Digital Twin.

Está claro que los retos tanto para el IoT/IIoT industrial como para la analítica de datos son también retos compartidos para la aplicación de un Digital Twin.

A pesar de los retos que Digital Twin comparte con el IoT y la analítica de datos desde la perspectiva del usuario hasta los retos de privacidad e infraestructura de Digital Twin, también hay retos específicos relacionados con el modelado y la construcción del Digital Twin. [47]

### **2.2.3.6 Modelización estandarizada**

El siguiente reto en todas las formas de desarrollo de un Digital Twin está relacionado con la modelización de estos sistemas, ya que no existe un enfoque estandarizado para la modelización. Desde el diseño inicial hasta la simulación de un Digital Twin es necesario un enfoque estándar, ya sea basado en la física o en el diseño. Los enfoques estandarizados garantizan la comprensión del dominio y del usuario, al tiempo que aseguran el flujo de información entre cada etapa del desarrollo y la implementación de un Digital Twin.

### **2.2.3.7 Modelización del dominio**

Otro reto derivado de la necesidad de un uso estandarizado está relacionado con garantizar que la información relativa al uso del dominio se transfiera a cada una de las etapas de desarrollo y funcionalidad del modelado de un Digital Twin. Esto asegura la compatibilidad con dominios como el IoT y la analítica de datos, permitiendo el éxito de los usos del Digital Twin en el futuro.

Esto es importante para el futuro, ya que garantiza que se tengan en cuenta en el desarrollo futuro de los Digital Twin, así como en el uso de la IIoT/IoT y la analítica de datos.

La tabla 1 muestra un resumen de los retos para la analítica de datos y la I/IoT, al tiempo que muestra los retos generales combinados para un Digital Twin, con los retos 2.2.3.6 y de 2.2.3.7 específicamente para la implementación del Digital Twin. [45]

*Tabla 1*  
*Retos compartidos*

Digital Twin	
Análisis de datos	IIoT / IoT
Infraestructura De TI	Infraestructura de TI
Datos	Datos
Privacidad	Privacidad
Seguridad	Seguridad
Veracidad	Veracidad
Expectativas	Expectativas
	Conectividad

*Nota:* Fuente: Elaboración propia.

## 2.3 Tecnologías facilitadoras

En esta sección se analizan las tecnologías habilitadoras de los Digital Twin.

### 2.3.1 Breve historia del internet de las cosas

El Internet de las cosas es el término que se da a los dispositivos conectados a Internet. Se trata de dar a las llamadas "cosas" un sentido de inteligencia y la capacidad de recoger información sobre su entorno. El término se publicó por primera vez a finales de los años 90, cuando Kevin Ashton expuso su visión de la IoT [47]. La idea de que todos los dispositivos que están interconectados dan al desarrollador la capacidad de rastrear y supervisar todo lo que hacemos, lo que conduce a un mundo más inteligente. Un ejemplo de esto lo encontramos muchos años antes en la Universidad Carnegie Mellon de Pittsburgh. Aquí un programa conectaba una máquina de Coca-Cola a través de Internet para ver si la bebida estaba lista y lo suficientemente fría como para que un usuario la comprara y consumiera [48]: un caso de uso simple pero eficaz para la visión de Ashton.

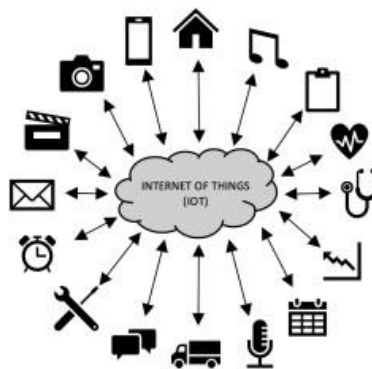


Figura 2 Diagrama internet de las cosas

Nota: Fuente:

[https://www.researchgate.net/publication/337019778\\_Digital\\_Twin\\_Enabling\\_Technologies\\_Challenges\\_and\\_Open\\_Research](https://www.researchgate.net/publication/337019778_Digital_Twin_Enabling_Technologies_Challenges_and_Open_Research)

El número de dispositivos IoT que se registra año tras año muestra el considerable crecimiento de esta tecnología. En 2018 la cifra superaba los 17.000 millones [49]. Para el año 2025, [50] predice que habrá más de 75.000 millones de dispositivos y se prevé que la industria tenga un valor de más de 5 billones de dólares [51]. La figura 3 muestra el crecimiento de los dispositivos IoT desde 2016. Estas cifras muestran el enorme impacto que están teniendo estos dispositivos y se suman a la visión expuesta por Ashton. El considerable número de dispositivos conectados ayuda a la visión de un mundo totalmente conectado, la figura 2 ilustra esta idea de servicios conectados a través de la IoT. La proliferación de dispositivos de la IoT es universalmente beneficiosa, ya que afecta al núcleo de la vida cotidiana, al sector de la comunicación, a la sanidad, a la construcción y al transporte, a las ciudades inteligentes y a la fabricación [51] [52].

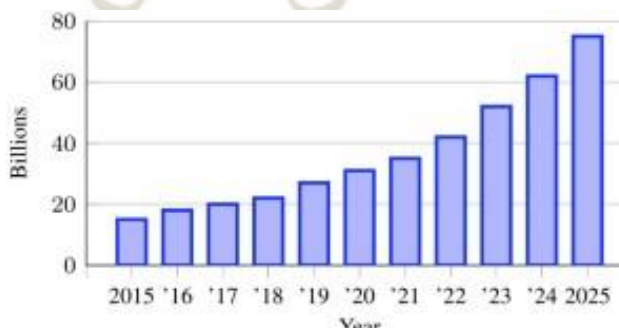


Figura 3 Crecimiento de dispositivos IoT

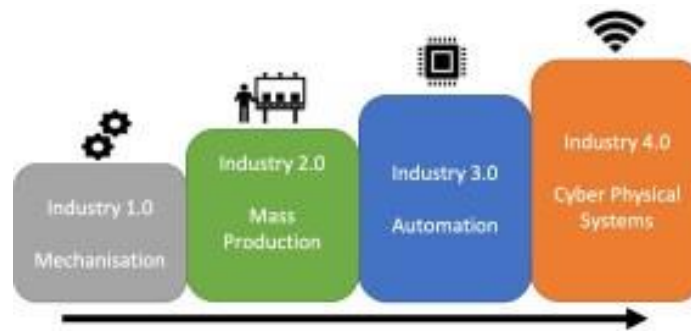
Nota: Fuente:

[https://www.researchgate.net/publication/337019778\\_Digital\\_Twin\\_Enabling\\_Technologies\\_Challenges\\_and\\_Open\\_Research](https://www.researchgate.net/publication/337019778_Digital_Twin_Enabling_Technologies_Challenges_and_Open_Research)

### 2.3.2 Breve historia del internet industrial de las cosas

El concepto de Internet Industrial de las Cosas (IIoT) procede del término IoT, redactado por Ashton [47]. La definición de IoT varía en el ámbito académico, y lo mismo ocurre con la definición de IIoT. El término tiene características similares a las de IoT, pero con un énfasis añadido en los procesos industriales. Boyes et al. presentan una serie de definiciones de IIoT, pero el enfoque principal es la mejora de la productividad para la industria [53]. En los entornos industriales y de fabricación, los sistemas originales son los sistemas de control industrial (ICS). Estos están bien documentados y se utilizan, pero los beneficios de que estos sistemas se vuelvan autónomos e inteligentes se ven potencialmente a través de la IIoT. Otra tecnología intrínsecamente vinculada a la IoT y a la IIT son los sistemas ciberfísicos (CPS) [54] [55]. Tanto los ICS como los CPS son como el IIoT, pero no son lo mismo. La principal diferencia es que los dispositivos de la IIoT requieren una conexión a Internet en lugar de estar encerrados en una arquitectura ICS [56]. Al igual que el IoT, el IIoT puede tener un gran impacto en la mejora de los procesos de fabricación, permitiendo que las tareas se evalúen con mayor conocimiento y respuestas en tiempo real a través de los dispositivos conectados, mejorando así el rendimiento, la tasa de producción, los costes, los residuos y muchos otros resultados críticos dentro del entorno industrial [57]. El IIoT no solo afecta a la fabricación, sino también a la agricultura, el petróleo, el gas y otros procesos a gran escala.

Del mismo modo, con el IoT, el IoT industrial está teniendo un impacto significativo dentro de la industria. Esto se ve especialmente con la predicción de Morgan Stanley de que el tamaño del mercado alcanzará los 110.000 millones de dólares en 2021 [58]. [59] informada que la IIoT podría añadir 14,5 billones de dólares a la economía mundial para 2030 [59]. La figura 4 muestra el desarrollo de la Industria 4.0, que es la introducción de la IIoT dentro de la línea de tiempo de la revolución industrial [60].



*Figura 4 Revoluciones industriales*

*Nota:* Fuente:

[https://www.researchgate.net/publication/337019778\\_Digital\\_Twin\\_Enabling\\_Technologies\\_Challenges\\_and\\_Open\\_Research](https://www.researchgate.net/publication/337019778_Digital_Twin_Enabling_Technologies_Challenges_and_Open_Research)

### 2.3.3 Tecnologías facilitadoras de IIoT/IoT y bloques funcionales

Tanto IoT como IIoT tienen una amplia gama de áreas esenciales que aseguran el funcionamiento de los sistemas conectados. Estas tecnologías habilitadoras se clasifican en cuatro dominios funcionales principales, como se describe en [61]. Estos dominios cubren las tecnologías habilitadoras individuales desde la comunicación de red, el hardware y el software hasta el procesamiento de datos, la alimentación y el almacenamiento de energía - todas ellas con objetivos específicos para permitir el desarrollo completo de un sistema IoT que facilite una arquitectura de Industria 4.0.

Los cuatro dominios tecnológicos habilitantes para la I/IoT comprenden de D1 el dominio de la Aplicación, D2 el dominio del Middleware, D3 el dominio de la red y D4 el dominio de los objetos, como se ve en la tabla 2.

*Tabla 2*

*Tecnologías de habilitación y bloques funcionales: I/IoT.*

<b>Dominio</b>	<b>Tecnología habilitadora</b>
<b>D1 Dominio de aplicación</b>	Aplicación I/IoT
	Arquitectura
	Software e Interfaz de programación de aplicaciones (API)
<b>D2 Dominio de middleware</b>	Plataformas en la nube
	Mecanismo de procesamiento de datos
<b>D3 Dominio de networking</b>	Almacenamiento de datos
	Protocolos de comunicación
	Interfaz de red
	Mecanismos de adopción

**D4 Dominio de objeto**

Plataforma de hardware

Objetos embebidos

Partes mecánicas y eléctricas

*Nota:* Fuente: Elaboración propia.

D1 se compone de tres capas. La primera es la capa de aplicación, que son las aplicaciones I/IoT; desde el hogar y las ciudades inteligentes hasta las granjas inteligentes. La siguiente es la capa de arquitectura, que puede consistir en arquitecturas de software; SOA (arquitectura orientada a servicios) o REST (transferencia de estado representativa), ambos ejemplos de lo que constituye la capa de arquitectura. La tercera capa, el software y las APIs, une el dominio de la aplicación con el dominio del middleware. el dominio de la aplicación con el dominio del middleware. Mantiene los sistemas operativos y el software. Por ejemplo, Android y los sistemas operativos personalizados que se utilizan para operar un sistema IoT. Este también podría estar formado por APIs hechas a medida para el despliegue de un sistema IoT, ambas son una tecnología clave para unir D1 con D2.

El dominio del middleware se compone de otras tres capas. La primera es la plataforma en la nube, que se compone de servicios que proporcionan recursos informáticos bajo demanda a través de la nube. Microsoft, Amazon y Google son los principales proveedores de servicios en la nube.

La segunda capa es el procesamiento de datos, habilitado mediante la minería de datos y los servicios de ejemplo proporcionados por BigQuery, Apache y Storm. La tercera capa de habilitación en D2 es el almacenamiento de datos. Esto es esencial en una infraestructura de I/IoT, y un ejemplo es MongoDB, que ofrece grandes motores de almacenamiento. La tercera parte del sistema IoT es D3, el dominio de la red, que está compuesto por tres capas habilitadoras. La primera es la capa de protocolo de comunicación; comprende los protocolos de aplicación, transporte y red de un sistema determinado, lo que permite una comunicación sin fisuras. La segunda capa de habilitación es la interfaz de red. Aquí se encuentran los estándares tecnológicos esenciales (por ejemplo, Radiofrecuencia (RF)) utilizados en todo el sistema de IoT, que también permiten la integración sin fisuras de IoT. La última capa de D3, el dominio de red, son los mecanismos de adopción. Consiste en la capa de adopción, que incluye estándares como 6TiSCH e (Instituto de Ingenieros Eléctricos y Electrónicos) IEEE 1095, que permiten una comunicación inalámbrica más fiable; asimismo, con la interfaz de conectividad y la capa de pasarela, que son estándares tecnológicos clave para el

desarrollo de un sistema de un sistema I/IoT. D4 es el último bloque del sistemaIoT, como se ilustra en tabla 2. El dominio del objeto está formado por tres capas habilitadoras. La primera son las plataformas de hardware, que consisten en las soluciones de hardware, con ejemplos como Raspberry Pi o Arduino. Este dominio reúne las tres últimas capas; ejemplos son los sensores, las etiquetas de radio, las pantallas y el firmware, son vitales para conectar el sistema. La última capa es la mecánica y eléctrica, formada por las baterías y las unidades de procesamiento necesarias para el funcionamiento del dispositivo. La división del dominio en cuatro bloques funcionales es más fácil para entender las doce secciones habilitadoras de este sistema de Internet de las Cosas; esto proporciona un marco integrado e interconectado.

### **2.3.4 Breve historia de la analítica de datos**

El término data analytics es un término paraguas que agrupa conceptos analíticos, como se ha visto a lo largo del documento y del mundo académico. Por lo tanto, es necesario comprender y analizar otros documentos. El término data analytics proviene del campo de la "Ciencia de los Datos", una materia multidisciplinaria que abarca una serie de conceptos, con énfasis en la recogida y presentación de datos para su análisis con el fin de obtener una mayor comprensión. La subsección siguiente presenta un análisis en profundidad del campo de la analítica de datos. Identificar y destacar estos temas ayudará a analizar otros trabajos y a ver dónde encaja esta investigación [62].

#### **2.3.4.1 Datos**

Para llevar a cabo el análisis de datos, la necesidad de datos brutos es primordial. Hay varias acciones necesarias para convertir estos datos en información utilizable, lista para su uso en algoritmos y análisis estadísticos. Se trata de los requisitos, la recogida, el tratamiento y la limpieza. Los requisitos establecen las necesidades necesarias de los datos y su uso, asegurando que se esbozan los requisitos específicos, teniendo en cuenta el uso previsto de los datos. La segunda etapa actúa sobre el requisito de recoger los datos pertinentes, identificando físicamente dónde y cómo se recogerán los datos. A continuación, los datos recogidos pasarán por una fase de procesamiento en la que se clasificarán según los requisitos específicos. La última fase, y posiblemente la más importante, es la limpieza de los datos. A pesar de que los datos se hayan recogido y clasificado, pueden tener importantes lagunas o datos erróneos. En esta fase de limpieza se utilizan los métodos de imputación, identificados anteriormente como retos para el análisis de datos. Estos métodos garantizan que no existan datos perdidos [63] [24].

### 2.3.4.2 Estadística

Estadística es el término que engloba la recopilación, la clasificación, el análisis y la interpretación de los datos. En este caso es relevante para el análisis de datos, ya que los modelos estadísticos sustentan los algoritmos de aprendizaje automático. La inferencia estadística y la estadística descriptiva son otra forma de utilizar el análisis de datos se utilizan para describir las observaciones de los datos recogidos. La IA y los siguientes temas muestran el crecimiento de la analítica de datos avanzada de datos [64] [63].

### 2.3.4.3 Inteligencia artificial

La (IA) es el primer tema de interés en la analítica de datos. La definición general de IA seremonta a finales de los años 50 con el concepto de crear "sistemas inteligentes" [65]. A continuación, se clasifican en temas de posible importancia para este proyecto [66] [67].

#### 2.3.4.3.1 Aprendizaje automático

Una subsección de la IA, el aprendizaje automático es la creación de algoritmos que pueden dar al ordenador la capacidad de aprender y actuar por el usuario sin estar directamente programado para hacerlo. El aprendizaje automático se utiliza para crear programas que utilizan sofisticados algoritmos para recoger y analizar datos de forma autónoma. Para un análisis más general, el aprendizaje automático puede encuadrarse endos tipos de aprendizaje: [54] [55].

#### 2.3.4.3.2 Aprendizaje supervisado

Es la forma más popular de aprendizaje automático. Los algoritmos utilizan grandes cantidades de datos etiquetados para analizar y aprender. El algoritmo se encarga de aprender y analizar los datos etiquetados para identificar correctamente una tareadeterminada; la clasificación de imágenes es un ejemplo [68].

Los algoritmos aprenden a partir de los datos de entrenamiento y, a continuación, se lesproporcionan datos de prueba para comprobar su grado de precisión en la predicción de lo que muestra una imagen, que se presenta mediante un porcentaje de exactitud. A continuación, el usuario analiza estas respuestas y cualquier error se corrige y se vuelva aprender, lo que ayuda a entrenar el modelo y a aumentar la precisión de un determinado algoritmo [69].

#### 2.3.4.3.3 Aprendizaje no supervisado

El aprendizaje no supervisado es otra forma de aprendizaje automático, no requiere datos costosamente marcados donde para cada patrón de entrada se ha determinado previamente la salida deseada: como se requiere para el aprendizaje supervisado [68]. Los algoritmos de aprendizaje no supervisado aprenden utilizando sus propios métodos para categorizar y resaltar patrones dentro de los datos en lugar de depender de la información del usuario. El clustering es un método de categorización de datos. Los algoritmos aprenden a agrupar conjuntos de datos sin etiquetar, mostrando potencialmente patrones ocultos que no eran identificables explícitamente [68].

#### 2.3.4.3.4 Aprendizaje profundo

El aprendizaje profundo es otra parte del campo de la analítica de datos y una subsección del aprendizaje automático. Los algoritmos de aprendizaje profundo aprenden datos no estructurados y no etiquetados utilizando redes neuronales complejas con extracción autónoma de características de entrada en contraposición a la extracción manual [70]. Estas redes utilizan el aprendizaje automático para crear modelos de aprendizaje profundo que pueden tardar más en entrenarse debido a que las redes neuronales son mucho más grandes, pero esto permite una mayor precisión. Otro tipo de aprendizaje es el semi-supervisado, que se define como tener algunos datos etiquetados, pero más datos sin etiquetar para ver cómo los algoritmos pueden aprender a ser más precisos [70]. En el campo de la ciencia de datos aparecen muchos más algoritmos, pero estos son los más comunes.

#### 2.3.4.4 Visualización de datos

El último subtema dentro de la analítica de datos es la visualización, definida como una representación gráfica o visualización de datos o resultados. El tipo de datos afecta a la forma de visualizarlos. Los más comunes son los datos multidimensionales, que pueden presentarse mediante gráficos y diagramas, tomando múltiples variables, por ejemplo, gráficos de barras o de tarta. Otro tipo de datos es el geoespacial; se trata de datos recogidos de la tierra a través de datos de localización, visualizados mediante mapas de distribución, mapas de conglomerados y, más comúnmente, mapas de contorno [71].

### 2.3.5 Tecnologías facilitadoras y bloques funcionales para el análisis de datos

La siguiente sección se refiere a la analítica de datos dentro del campo de la Inteligencia Artificial, el aprendizaje automático y el aprendizaje profundo. Las descripciones de las tecnologías habilitadoras para la analítica de datos [72] y el esquema de clasificacionesbozado por [61] se resumen en la abla 3. Las tecnologías habilitadoras se parecen a la IoT en muchos aspectos, pero tienen capas ligeramente diferentes en torno a la visualización y el aspecto algorítmico de la analítica. A continuación, se muestra una visión general, con los dominios etiquetados como D5, D6, D7 y D8. La tabla 3 es el resultado del análisis dela tabla 2 de Bibri y Krogstie [72]; sin embargo, su presentación es ligeramente diferente.La tabla comienza con D5, el dominio del objeto, seguido de D6, el dominio del middleware,D7, el dominio de la red, y por último D8, el dominio de la aplicación. En cada uno de los dominios hay una lista de tecnologías habilitadoras asociadas a la analítica de datos.

*Tabla 3*

*Tecnologías de habilitación y bloques funcionales: Análisis de datos*

<b>Dominio</b>	<b>Tecnología habilitadora</b>
<b>D5 Dominio de objeto</b>	Recolección de datos y preprocesamiento Repositorios de datos
<b>D6 Dominio de middleware</b>	Facilidades de almacenamiento Procesamiento de datos Técnicas de análisis y algoritmos
<b>D7 Dominio de networking</b>	Tecnología de red inalámbrica
<b>D8 Dominio de aplicación</b>	Hardware y Visualización de datos Aplicaciones de análisis de datos

*Nota:* Fuente: Elaboración propia.

D5 es el dominio de los objetos, que tiene al menos tres capas, lo que refleja la doble condición de las instalaciones de almacenamiento. La primera capa habilitadora es la recopilación de datos, que se ocupa del preprocesamiento de datos para las soluciones analíticas. El uso de herramientas y métodos de detección de datos permite la recogida de estos. Las unidades de procesamiento de señales digitales también garantizan la recogida de datos. La segunda capa es el repositorio de datos que facilita el almacenamiento y el uso de las bases de datos. La última capa de la D5, que enlaza con

la D6, es la de las instalaciones de almacenamiento, que permiten guardar grandes cantidades de datos mediante el uso de almacenamiento en servidores que permiten obtener datos a la carta. Esta capa es también la conexión con el procesamiento de los datos de almacenamiento al dominio del middleware. D6 es el dominio del middleware, que consta de tres capas de habilitación. La primera enlaza con la D5, que se refiere al procesamiento del almacenamiento. La segunda capa en D6 es el procesamiento de datos, que es la capa principal para habilitar la analítica de datos, los servicios en la nube y las principales arquitecturas de middleware, incluidos los sistemas de software y de bases de datos. La tercera capa en D6, que se ve en la tabla 2, es el análisis y los algoritmos. Esta capa facilita la tarea de la minería de datos, el aprendizaje automático, la estadística y la consulta de los datos recogidos. Así como los modelos de habilitación dentro de la analítica de datos; aprendizaje supervisado y no supervisado.

D7 es el dominio de las redes, que muestra la tecnología de habilitación de los protocolos de conectividad, considerando la tecnología inalámbrica y la comunicación, y cómo permiten la recopilación y el procesamiento eficientes de los datos de las capas y dominios anteriores. D7 también se refiere a las normas de habilitación relacionadas con los mecanismos de privacidad y seguridad. El último ámbito tratado es el D8, titulado aplicación. El D8 tiene dos capas tecnológicas habilitadoras. La primera es la capa de hardware y visualización. Esta capa permite que la tecnología tangible registre los datos y lleve a cabo el aprendizaje automático y profundo o el análisis estadístico. La capa de visualización permite mostrar fácilmente información útil sobre las tareas del usuario. Por último, la capa de aplicación destaca las aplicaciones relacionadas con la analítica de datos, como los coches auto conducidos, el reconocimiento de imágenes o los asistentes personales virtuales como Alexa de Amazon. En resumen, la tabla 2 presenta los dominios funcionales para las tecnologías habilitadoras asociadas a IoT/IIoT, mientras que la tabla 3 presenta los dominios funcionales para las tecnologías habilitadoras de la analítica de datos. La sección anterior proporciona un marco para crear una síntesis de los dominios funcionales y las tecnologías habilitadoras en el contexto de un Digital Twin, como puede verse en la tabla 4.

#### **2.4 Tecnologías habilitadoras para el Digital Twin**

De forma similar a la tabla 2 de Bibri y Krogtie [72], la tabla 4 proporciona una síntesis adicional de ideas para los dominios funcionales y las tecnologías habilitadoras de un Digital Twin. Empezando por el D9 para el dominio de la aplicación, pasando por

el D10 para el dominio del middleware, seguido por el D11 para el dominio de la red y finalmente el D12 para el dominio del objeto.

*Tabla 4*

*Tecnologías de habilitación y bloques funcionales: Digital Twin*

<b>Dominio</b>	<b>Tecnología habilitadora</b>
<b>D9 Dominio de aplicación</b>	Modelo de arquitectura y visualización
	Software y APIs
	Recolección de datos y preprocesamiento
<b>D6 Dominio de middleware</b>	Tecnologías de almacenamiento
	Procesamiento de datos
<b>D7 Dominio de networking</b>	Tecnología de comunicación
	Comunicación inalámbrica
<b>D8 Dominio de objeto</b>	Plataforma de hardware
	Tecnologías de sensores

*Nota:* Fuente: Elaboración propia.

El primer dominio es el D9, que es el dominio de la aplicación y se compone de tres capas importantes, siendo la primera la capa de arquitectura y visualización del modelo, esencial para crear modelos de alta fidelidad de la entidad física. Esta capa permite la visualización y el modelado de la arquitectura de los Digital Twin. Esto garantiza que los Digital Twin se modelen utilizando algo más que los comportamientos de las entidades físicas. Esto es posible gracias a herramientas como Simulink y Twin Builder. La segunda capa es el software y la API, que se utilizan específicamente para ayudar a modelar la arquitectura de los Digital Twin y facilitar la tercera capa: el preprocesamiento y la recopilación. Esta última capa del dominio de la aplicación es necesaria para asegurar que los datos se recogen correctamente, por ejemplo, con Predix, Mindsphere y Storm, por nombrar sólo algunas aplicaciones para la recogida de datos. Esta capa asegura que los datos se recogen correctamente para facilitar el uso de IoT y la analítica para un Digital Twin, a la vez que sirve de puente entre el dominio D9 y el D10. D10, el dominio del middleware consta de dos capas de habilitación. La primera es la tecnología de almacenamiento. Que facilita el almacenamiento de datos a través de Mongo DB, servicios MySQL y bases de datos bajo demanda que son necesarios para el uso de Digital Twin. La segunda capa, relacionada con el procesamiento de datos, es esencial para transferir los datos almacenados entre D10 y D11. D11 consiste en el

dominio de red con dos capas de habilitación, la primera es la capa de tecnología de la comunicación, esencial para asegurar que los datos recogidos se comunican entre los dominios. La segunda capa del bloque funcional del dominio de red para los Digital Twin es la capa de comunicación inalámbrica, necesaria para garantizar que la transmisión de datos de forma inalámbrica sigue el protocolo correcto dentro de una arquitectura de Digital Twin, así como para llevar los datos al siguiente dominio, D12.

D12, el dominio del objeto consta de dos capas de habilitación. La primera es la plataforma de hardware y la segunda es la tecnología de sensores. Ambas son necesarias para garantizar que se dispone del hardware adecuado para llevar a cabo el análisis de Digital Twin, así como para facilitar la recogida de datos mediante la tecnología de sensores.

## **2.5 Investigación actual**

En la siguiente sección se identifican los trabajos relacionados con el IoT/IIoT y el análisis de datos, centrándose en las publicaciones sobre Digital Twins, analizando una serie de publicaciones e identificando las lagunas en este campo. Esta visión del mundo académico ayudará a otros investigadores, permitiéndoles encontrar lagunas en la investigación y permitirá avanzar hacia una definición más completa de un Digital Twin.

### **2.5.1 Metodología de revisión categórica**

La primera parte de esta sección realiza una revisión categórica de la literatura relevante, siguiendo la metodología utilizada por Kritzinger et al, [73] para producir una tabla de revisión categórica de las publicaciones seleccionadas. Los principales elementos de la revisión se basan en los tres niveles de integración de un Digital Twin, como se describe en la sección 2.2, subsección B de este documento y por Kritzinger et al. [73]. La metodología utilizada en este documento se basa en el trabajo realizado por Kritzinger et al. [73], que implicó la categorización de cuarenta y tres artículos relacionados con el tema de los Digital Twin, que fueron publicados entre 2001 y 2017. Los artículos se obtuvieron a través de motores de búsqueda de literatura, como Google Scholar.

Para la búsqueda realizada en este trabajo, los autores utilizaron Google Scholar con búsquedas específicas dirigidas a los repositorios ACM, IEEE y Science Direct. De las investigaciones encontradas hubo 177 artículos para analizar desde 2015 hasta la actualidad (31 de diciembre de 2021), con solo 42, anteriores a 2017. Los términos de búsqueda incluían variaciones de Digital Twin (Digital-Twin, Digital Twins). Además del término Digital Twin, la búsqueda incluyó la adición de términos relacionados con las

amplias áreas de investigación (Industrial Digital Twin, Salud Digital Twin, Smart Cities Digital Twin). Se encontró veintiséis fuentes clave de tres áreas, fabricación, sanidad y ciudades inteligentes, para sintetizar la tabla 5.

### **2.5.2 Organización de la revisión categórica**

Los artículos están ordenados alfabéticamente en tres grandes áreas, fabricación, asistencia sanitaria y ciudades inteligentes, que se muestran en la tabla 5. En las siguientes secciones se analizan las columnas utilizadas en tabla 5:

#### **2.5.2.1 Paper**

La primera columna indica los autores y el año de cada publicación.

#### **2.5.2.2 Tipo**

La segunda columna identifica cada artículo por el tipo de investigación llevada a cabo: artículo de revisión o estudio de caso. También puede ser un nuevo concepto relacionado con los Digital Twin o una definición, cada uno se clasifica en consecuencia.

#### **2.5.2.3 Definición de gemelo**

Las dos columnas siguientes son la definición relativa a los niveles de integración. La primera identifica a qué se refieren los autores originales modelo digital, sombra o gemelo.

#### **2.5.2.4 Gemelo real**

La segunda columna identifica si el documento describe con precisión un Modelo Digital, una Sombra o un Gemelo. Las definiciones de la sección 2.1 se utilizan para clasificar la publicación, lo que permite comprender los posibles errores de definición.

#### **2.5.2.5 Ámbito general**

Una quinta columna identifica las áreas generales de investigación, a saber, fabricación, la asistencia sanitaria y las ciudades inteligentes.

#### **2.5.2.6 Área específica**

La sexta columna profundiza en el área amplia, reduciéndola a un área específica. Por ejemplo, en la industria manufacturera, esto se reduciría a una fábrica inteligente o a un área de interés de diagnóstico de fallos. En el caso de las ciudades inteligentes, se reduce a el tráfico o las infraestructuras.

### 2.5.2.7 Tecnología

La última columna identifica la tecnología utilizada, por ejemplo, simulación, análisis de datos e IoT. Los documentos están ordenados alfabéticamente en tres grandes áreas, fabricación, atención sanitaria y ciudades inteligentes, que se muestran en la tabla 5.

*Tabla 5 Revisión categórica*

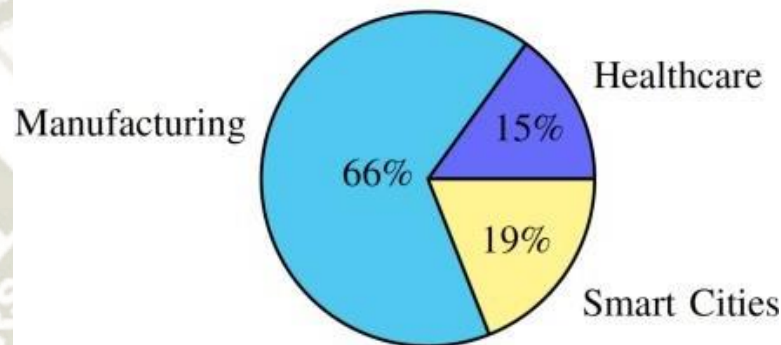
Paper	Tipo	Gemelo definido	Gemelo real	Ámbito general	Área específica	Tecnología
<b>Bilberg, Malik (2019) [1]</b>	Caso de estudio	DT	DS	Fabricación	Fábrica inteligente	<b>Simulación</b>
<b>Chhetri et al. (2019) [74]</b>	Estudio de caso	DT	DT	Fabricación	Línea de montaje	<b>IA, Sensores, Simulación</b>
<b>Él et al. (2018) [75]</b>	Revisión	DT	DS	Fabricación	Sistema de alimentación	<b>Simulación, IA, Analítica</b>
<b>Howard (2019) [11]</b>	Concepto	DT	DM	Fabricación	Desarrollo de productos	<b>EDA, Visualización</b>
<b>Jain et al. (2019) [76]</b>	Concepto	DT	DT	Fabricación	Diagnóstico de fallas	<b>Industria 4.0</b>
<b>Karadeniz et al. (2019) [77]</b>	Estudio de caso	DT	DS	Fabricación	Máquinas de helado	<b>AR, VR, Industria 4.0, IA, CPS</b>
<b>Kuehn (2019) [78]</b>	Concepto	DT	DS	Fabricación	Fábrica inteligente	<b>Simulación</b>
<b>Lu (2019) [79]</b>	Revisión	DT	Sin ejemplo	Fabricación	Fábrica inteligente	<b>Nube, CPS, Industria 4.0</b>
<b>Mandola et al. (2019) [2]</b>	Estudio de caso	DT	Sin ejemplo	Fabricación	Avión	<b>Blockchain, Visualización</b>
<b>Mawson, Hughes (2019) [27]</b>	Estudio de caso	DT	DT	Fabricación	Modelado energético	<b>Industria 4.0</b>
<b>Min et al. (2019) [35]</b>	Estudio de caso	DT	DS	Fabricación	Fábrica Petroquímica	<b>IA, Optimización</b>
<b>Qi, Tao (2018) [24]</b>	Revisión	DT	DT	Fabricación	Fábrica inteligente	<b>Industria 4.0, IA, Nube, Big Data</b>
<b>Shangguan et al. (2019) [80]</b>	Estudio de caso	DT	DM	Fabricación	Turbina eólica	<b>CPS</b>
<b>Sivalingam et al. (2018) [17]</b>	Revisión	DT	DS	Fabricación	Turbina eólica	<b>CPS, Simulación</b>
<b>Tao et al. (2019) [81]</b>	Revisión	DT	DT	Fabricación	Fábrica inteligente	<b>CPS, Industria 4.0, IA</b>

*Nota:* Fuente: Elaboración propia.

### 2.5.3 Análisis de los documentos revisados.

Las siguientes secciones se basan en los documentos revisados en la revisión categórica, discutiendo el documento con más detalle, destacando los conceptos y estudios de caso realizados. Las secciones no se limitan a, pero incluirán, las principales áreas de los documentos recogidos relacionados con la asistencia sanitaria, las ciudades inteligentes y la fabricación. El orden de las áreas presentadas refleja el nivel de investigación actual en cuanto al número de artículos encontrados. En el caso de la

sanidad, el número de artículos encontrados es limitado, pero prevalecen los beneficios que los Digital Twin pueden tener en la industria sanitaria [84] [32] [39] [40]. Le sigue de cerca el área de las ciudades inteligentes, con un número reducido de artículos. La mayoría de las investigaciones se enmarcan en el ámbito de la fabricación. La figura 5 muestra la proporción de artículos en función del área de investigación en la que se encuadran: sanidad, fabricación o ciudades inteligentes.



*Figura 5 Porcentaje de áreas de investigación en el análisis de papers*

*Nota: Fuente:*

[https://www.researchgate.net/publication/337019778\\_Digital\\_Twin\\_Enabling\\_Technologies\\_Challenges\\_and\\_Open\\_Research](https://www.researchgate.net/publication/337019778_Digital_Twin_Enabling_Technologies_Challenges_and_Open_Research)

### **2.5.3.1 Salud**

Un elemento para tomar de las diversas definiciones de Digital Twin es el concepto descrito por He [75], las "réplicas digitales" de una cosa física. El Saddik lo redefinió con la inclusión de réplicas de Digital Twin de cosas vivas, así como de entidades no vivas [39]. La presentación del uso potencial del Digital Twin para el sector sanitario demuestra que no se limita a la fabricación. Ross [40] presenta un trabajo con Hewlett-Packard en el que utiliza la IA y el IoT para crear avatares de Digital Twin de personas de diversas maneras. Desde el punto de vista de la salud, la tecnología de Digital Twin, combinada con algoritmos de IA, puede utilizarse para ver los efectos que determinados cambios de estilo de vida podrían tener en la salud de una persona, recomendando cambios específicos a partir del análisis de IA y Digital Twin. Este uso pone de relieve la plena integración de los datos del gemelo físico (el humano) y del Digital Twin (la réplica). Se ofrece al humano la posibilidad de ver el impacto que sus acciones están teniendo en el gemelo físico, a la vez que se le muestra el efecto que podrían tener algunos cambios en su estilo de vida [40].

En otro escenario más reciente, Laaki et al. [32] presentan un prototipo funcional de cirugía autónoma aprovechando la conectividad del IoT y la Industria 4.0 para crear un Digital Twin de un paciente. Los autores proponen una aplicación de cirugía remota a través de una red móvil. El prototipo utiliza un brazo robótico, realidad virtual (RV) con un entorno 4G, para ofrecer una cirugía de precisión. El artículo presenta la complejidad que entraña la investigación multidisciplinar, y cita este aspecto como una de las razones por las que el trabajo utilizó la simulación en lugar de un prototipo físico. Laaki et al. también discuten los problemas de integración del prototipo con un Digital Twin. Los autores exploran algunos de los avances de la IA y la Industria 4.0 y cómo facilitan los retos de conectividad, integración e investigación multidisciplinar [32]. Liu et al. [37] presentan un enfoque novedoso para la futura prestación de asistencia sanitaria que combina la tecnología de la nube con los Digital Twin para crear un marco que ayude a monitorizar, diagnosticar y predecir la salud de un paciente. Liu et al. lo consiguen mediante el avance y el uso de la tecnología vestible del IoT y de sensores en el hogar, haciendo hincapié en el uso para las personas mayores. Además de un marco de trabajo, los autores también presentan varias aplicaciones citando la viabilidad de cada una. La principal contribución del artículo de Liu et al. es la capacidad de predecir un problema con los pacientes con mayor precisión, a través del uso combinado de la tecnología IoT, Cloud y Digital Twin [37].

### **2.5.3.2 Ciudades inteligentes**

Esta sección se centra en la investigación actual sobre las ciudades inteligentes en relación con los Digital Twin. La investigación en los últimos años ha visto un crecimiento sustancial de la urbanización combinado con el aumento del IoT y la analítica de datos [3].

Mohammadi et al. [3] citan esto como una de las motivaciones de su trabajo e identifican los diferentes estados en el flujo espaciotemporal, haciendo hincapié en que es necesario comprenderlos para mantener el crecimiento. El concepto que presentan lo hace mediante el uso de Digital Twin y auriculares de realidad virtual, lo que les permite supervisar las fluctuaciones al tiempo que realizan predicciones mediante análisis en tiempo real [3]. Ruohomäki et al. [83] también presentan un marco, "mySMARTlife", que hace uso de los avances en IoT en las ciudades para crear un Digital Twin de ciudad inteligente. El documento presenta un caso de estudio para ayudar en la planificación urbana y los entornos construidos, pero tiene

usos particulares en el campo del consumo de energía con la capacidad de utilizar el Digital Twin para el seguimiento y la comparación del consumo de energía basado en el medio ambiente y el impacto humano. Ambos se utilizan en tiempo real y para futuros desarrollos [83]. Tanto Mohammadi et al. [3] como Ruohomäki et al. citan la necesidad de adoptar los conceptos de la Industria 4.0 para garantizar que el nivel de intercambio de datos sea lo suficientemente alto como para que el gemelo para que el gemelo funcione con precisión.

La alimentación de la red energética, junto con la implementación de la integración de métodos de energía renovable, es un reto. Con esto viene la necesidad de un suministro preciso dentro de una ciudad inteligente. La energía eólica es un ejemplo de fuente de energía renovable que debe ser suministrada, controlada y analizada. Pargmann et al. [18] presentan un sistema de monitorización Digital Twin basado en la nube que se utiliza para el desarrollo y la monitorización de parques eólicos. Los autores presentan un prototipo de trabajo que utiliza fuentes de datos y parámetros establecidos tanto desde un contexto técnico como empresarial, lo que permite crear un gemelo de trabajo del desarrollo de un parque eólico [18]. Sivalingam et al. [17] revisan y elaboran un estudio de caso que investiga el uso de los parques eólicos y el consumo de energía para una red inteligente.

El documento cita algunos de los retos que plantea la fiabilidad del consumo de energía y el mantenimiento general de los aerogeneradores. Los autores proponen una metodología de trabajo que hace uso de los sensores del IoT combinados con la analítica de datos dentro de un entorno Digital Twin para realizar y predecir con precisión el mantenimiento de las turbinas eólicas [17]. En cuanto a una ciudad inteligente, Jo et al. [15] y Chen et al. [22] presentan trabajos relacionados con las ciudades inteligentes y ambos utilizan Digital Twins. El primero, de Jo et al. [15], presenta un trabajo que produce un estudio de viabilidad sobre los usos potenciales de un Digital Twin para una granja inteligente. Los autores señalan la importancia de la Industria 4.0 para la realización de este proyecto, ya que hace que el despliegue de Digital Twin sea un reto en entornos complejos. Los autores destacan tres aplicaciones, Predix de GE, Ditto de Eclipse y Watson de IBM como contendientes para el despliegue de la tecnología Digital Twin, junto con algunas directrices sobre cómo la supervisión del ganado puede ser más inteligente mediante el uso de un Digital Twin [15]. Chen et al. [22] presentan Digital Twin para los coches y la gestión del tráfico. El documento explora los retos de la conducción, mostrando que es

necesario un mayor flujo de datos dentro del vehículo utilizado y una conexión con otros vehículos en las proximidades. Chen et al. [22] presentan un marco que utiliza un Digital Twin combinado con algoritmos de aprendizaje que supervisan y analizan la información basada en el comportamiento del usuario. Los algoritmos facilitan un Digital Twin del comportamiento de un conductor en tiempo real, proporcionando advertencias e instrucciones sobre cómo conducir de forma más segura para minimizar el riesgo [22]. Tanto Jo et al. como Chen et al. comparten la necesidad y el reto del intercambio de datos, y ambos citan la necesidad de una mayor conectividad, que es necesaria para lograr el óptimo resultado y la precisión de cada uno de los Digital Twin.

### 2.5.3.3 Fabricación

Por último, esta sección se refiere a un área que incluye la mayoría de las investigaciones relacionadas con los Digital Twin. También es un área que tiene muchos subcampos, desde las grandes fábricas inteligentes hasta las máquinas y herramientas más pequeñas. Por este motivo, la sección se divide en las siguientes subsecciones:

a) Reseñas sobre fabricación inteligente, b) Simulación e inteligencia artificial, c) Diseño y desarrollo de sistemas y d) Fabricación energéticamente eficiente.

#### 2.5.3.3.1 Reseñas sobre fabricación inteligente

Las revisiones en este campo son limitadas, pero la primera que hay que destacar es la presentada por Qi y Tao [24], que ofrece una visión completa de "360" sobre los Digital Twin para los grandes datos en un entorno industrial y de fabricación. Ofrece una comparación de las tecnologías que permiten los Digital Twin, además de argumentar la importancia de las tecnologías emergentes para el desarrollo de la fabricación inteligente [24].

Tao presenta dos artículos de revisión, [82] y [81]; el primero [82] compara los Digital Twins y los sistemas ciber físicos dentro de un entorno de fabricación inteligente. El segundo, de Tao [81], es un artículo sobre el estado del arte de los Digital Twin combinados con la industria. Ambos trabajos comparten tecnologías habilitadoras, IoT, nube, big data e inteligencia artificial, mostrando cómo estas tecnologías utilizan Digital Twins. Tao [85] también profundiza en algunos de los principales usos de las aplicaciones; simulación de Digital Twin, Digital Twin "como servicio", fusión de datos e interacción y

colaboración. Ambos citan el creciente desarrollo de las tecnologías de la Industria 4.0, con énfasis en la analítica de datos y el IoT, como factor de crecimiento y uso de los Digital Twins. [82] [81] [85].

#### 2.5.3.3.2 Simulación e inteligencia artificial

Los artículos anteriores han tocado algunas áreas clave relacionadas con el uso de DigitalTwin combinados con enfoques de simulación e inteligencia artificial para la fabricación. Kuehn [78] describe un concepto que cita "clones virtuales" de un sistema combinados con algoritmos de aprendizaje automático para mejorar el proceso de fabricación.

El autor categoriza las áreas clave de un proceso de fabricación para destacar sus objetivos y conceptos específicos para aplicar un Digital Twin al proceso de fabricación, dando a las empresas la capacidad de probar, simular y optimizar los procesos de fabricación en un entorno virtual que garantice una mayor calidad y eficiencia. De forma similar a Kuehn [78], Min et al. [35] presentan un trabajo que enumera de forma exhaustiva las principales tecnologías habilitadoras, haciendo hincapié en las soluciones de Digital Twin que utilizan IA, concretamente el aprendizaje automático, ofreciendo evaluaciones exhaustivas. El trabajo utiliza un caso de estudio para evaluar los pros y los contras del uso del aprendizaje automático y los Digital Twin para la industria petroquímica, aplicando cada uno de ellos a una fábrica petroquímica de IoT industrial [35]. Una limitación encontrada al realizar el caso de estudio es que el DigitalTwin y los algoritmos son únicos para una planta petroquímica. El aprendizaje de transferencia podría utilizarse para encontrar puntos comunes en los algoritmos para ayudar a crear soluciones transferibles a otros procesos de fabricación. Una subsección de la fabricación es el diagnóstico de fallos con dos artículos de interés encontrados. El primero, de Jain et al. [76], presenta un estudio de simulación que analiza un enfoque de Digital Twin para el diagnóstico de fallos en sistemas fotovoltaicos distribuidos. Los avances en la tecnología Digital Twin permiten al equipo desarrollar un Digital Twin que puede estimar con precisión los fallos relacionados con las unidades de energía fotovoltaica en tiempo real [76].

Un trabajo similar realizado por Xu et al. [23] propone una solución centrada en la IA para el diagnóstico de fallos en un entorno de fabricación inteligente. Xu et al. destacan los retos que plantea la cantidad de datos de

entrenamiento disponibles para crear algoritmos de IA precisos para los nuevos procesos de fabricación, algo que también es necesario para crear Digital Twin precisos. Para garantizar la mitigación de cualquier desafío, la primera fase del modelo utiliza un Digital Twin que aprende y diagnostica los fallos mientras produce datos de entrenamiento. En la segunda fase se puede utilizar el aprendizaje por transferencia con los datos de entrenamiento recogidos en la primera fase del algoritmo. Con la ayuda de un Digital Twin y el aprendizaje por transferencia se consigue un sistema de diagnóstico de fallos más preciso.

Los autores presentan un caso práctico con un entorno de fabricación de automóviles que prueba y evalúa la eficacia del concepto [23].

#### 2.5.3.3.3 Diseño y desarrollo de sistemas

Otra área clave para adoptar el uso de Digital Twin es a través de la etapa de diseño y desarrollo de los procesos y sistemas de fabricación. Shangguan et al. [80] analizan un enfoque que se basa en el uso de Digital Twin pero que también introduce la tendencia de los sistemas ciber físicos (CPS). Los autores presentan un Marco Jerárquico de Digital Twin (HDTM) para el diseño y desarrollo de CPS dinámicos en un entorno de fabricación inteligente. El Digital Twin se utiliza en todos los niveles del diseño de CPS, aprovechando la capacidad de reutilizar y probar con datos físicos en un gemelo virtual. Los autores realizan un estudio de caso que muestra los beneficios de un Digital Twin, que se ven mejor aquí, ya que da a los autores la capacidad de utilizar el gemelo en un entorno de diseño predictivo en tiempo real, así como el aprendizaje para los cambios del sistema a gran escala. El marco es una aplicación para el diseño de robots industriales [80]. Howard [11] presenta un trabajo que supone una visión de la tendencia del Digital Twin, evaluando sus usos y su idoneidad para el mundo de la fabricación inteligente en constante evolución. El principal objetivo alcanzado es el diseño y desarrollo de hardware electrónico de fabricación a través de la "validación virtual" utilizando un Digital Twin [11]. Ambos artículos citan los retos identificados por los líderes y visionarios del mundo de los Digital Twin, la analítica y la fabricación inteligente. Aparte del diseño de sistemas de alto nivel, Karadeniz et al. [77] y Chhetri et al. [74] elaboran documentos contextuales y estudios de casos sobre el desarrollo de sistemas Digital Twin para procesos de fabricación. El primero [77] analiza los avances de la Industria 4.0 y el IoT utilizando un caso de estudio para explorar cómo este crecimiento

está facilitando las tendencias de la RA (Realidad Aumentada) y la RV (Realidad Virtual). Los autores exponen un concepto que crea un Digital Twin de las cosas gastronómicas (dispositivos y procesos relacionados con la comida y la cocina) y lo presentan como "cosas eGastronómicas", similar al Internet de las Cosas.

Los procesos gastronómicos tienen sensores físicos del IoT que recogen datos para crear un Digital Twin. Los autores utilizan una máquina de helados como caso de estudio para mostrar cómo un Digital Twin puede ayudar a monitorizar y mantener el rendimiento de los procesos "eGastronómicos", concretamente la máquina de helados [77].

El último, de Chhetri et al. [74], también cita los avances en la digitalización para el crecimiento del diseño y los desarrollos de la fabricación inteligente. Los autores proponen una metodología y realizan un estudio de caso que aprovecha el crecimiento del IoT para construir un Digital Twin de un proceso de fabricación. En esta metodología, los sensores del IoT ayudan a recoger y almacenar flujos de datos que se utilizan para resaltar indirectamente los estados del canal lateral; la acústica, la potencia y la salida magnética de un proceso. Estos datos pueden utilizarse para localizar un fallo e identificar problemas en los procesos de fabricación. El equipo valida el trabajo con un estudio de caso de un sistema de modelado por deposición fundida (FDM) que produce una detección de anomalías de gran precisión, la primera de este tipo en este formato [74]. Mandolla et al. [2] analizan otra tendencia adaptada a la industria de la fabricación inteligente, Blockchain. Una lista creciente de bloques dentro de un libro de contabilidad descentralizado puede utilizarse para registrar datos en muchos ordenadores. Los bloques anteriores se vinculan mediante un hash criptográfico, un método de protección de datos que utiliza códigos cifrados. Estos hashes son únicos por cada bloque y contienen atributos como la marca de tiempo y los datos de la transacción. El crecimiento de Blockchain se atribuye a su creciente uso en procesos distintos a su uso original; por ejemplo, la criptomoneda [86] [87] [88]. Mandolla et al. [2] presentan un ejemplo de uso de blockchain, en un estudio de caso que muestra cómo lo han combinado con un Digital Twin. Los autores se centran en la fabricación aeroespacial en relación con el proceso de adición de metales. Mandolla et al. crean un Digital Twin de este proceso, a la vez que proporcionan una respuesta conceptual para asegurar el proceso a través de blockchain, y la monitorización

a través de un Digital Twin [2]. Del mismo modo, Bilberg y Malik [1] utilizan la robótica, para explorar cómo se puede adaptar para un proceso de fabricación, presentan un Digital Twin dirigido por eventos que trabaja en paralelo con un robot para realizar tareas en una línea de montaje.

El otro combina la tarea física y la tarea virtual para crear una línea de producción robótica en tiempo real, basada en las habilidades, que asigna con precisión una tarea al humano o al robot en función de la producción óptima que podrían alcanzar respectivamente [1]. Ambos documentos, mediante el uso de estudios de casos, identifican los beneficios de la digitalización, pero también citan los retos de la integración perfecta, algo necesario para el funcionamiento eficaz de una línea de montaje y de los Digital Twin [2] [1] [89].

#### 2.5.3.3.4 Fabricación eficiente

Desde el punto de vista energético en consonancia con los objetivos nacionales e internacionales, es necesario controlar y reducir la energía utilizada en la fabricación; de ahí la necesidad de encontrar soluciones potenciales para las líneas de producción eficientes desde el punto de vista energético. Tanto Lu et al. [79] como Mawson [27] presentan sistemas y arquitecturas para la fabricación energéticamente eficiente. Una de las razones para ello son los beneficios medioambientales. Sin embargo, una fabricación más eficiente también reducirá los costes, aumentará los beneficios y las inversiones futuras. Lu et al. presentan un artículo que se centra en el desarrollo de una arquitectura que implementa el modelo Digital Twin consciente de la energía con una plataforma llamada MCloud. Ambas facilitan un entorno de Industria 4.0 en el que los procesos de fabricación se monitorizan de forma continua y se autoconfiguran a través de CPSs y Digital Twins, siendo la eficiencia energética el objetivo general. Mawson et al. [27] también se basan en el avance de la Industria 4.0, citando una revisión sobre su eficacia para aumentar la automatización, la conectividad y la flexibilidad de los procesos de fabricación. Las principales aportaciones del artículo son una revisión holística y un estudio de casos que utilizan metodologías y marcos derivados del análisis del consumo de energía a nivel del proceso de la máquina. Mawson et al. citan varias herramientas de simulación, pero carecen de integración multinivel.

Para producir un modelo de alta precisión, se necesitan todos los aspectos del proceso de fabricación, desde los materiales hasta los flujos de recursos, para

analizar el consumo de energía con precisión. Estos son retos que demuestran cómo los Digital Twin, la RV y la RA facilitarán la investigación futura [27].

## 2.6 Investigación abierta

La penúltima sección de esta revisión analiza brevemente las cuestiones de investigación abiertas para los Digital Twin, así como la exploración de la literatura revisada para los desafíos que enfrenta la investigación futura.

### 2.6.1 Digital Twin en la fabricación

En la literatura se proponen Digital Twin de fabricación y totalmente integrados, pero actualmente no se realizan en la industria. En la sección 2.4 se han explorado una serie de estudios de casos, revisiones y conceptos relacionados con las publicaciones en la industria manufacturera. Hay una serie de publicaciones sobre desarrollos más pequeños de un Digital Twin. Los ejemplos de estudios de casos de la industria se enumeran en la tabla 5 y se analizan con más detalle en la sección 2.4, subsección 2.4.3, y se resumen a continuación. En la literatura no todas las publicaciones cubren todos los aspectos de un Digital Twin, desde el modelado físico y virtual hasta las partes de datos, conexión y servicio de modelado de un Digital Twin. El modelado y el escalado son necesarios para crear Digital Twin genéricos. Zheng et al. [8] y Schleich et al. [34] presentan una modelización basada en sensores de diferentes etapas de un Digital Twin, mientras que Shanguam et al. [80] sólo se centran en la modelización jerárquica de la modelización virtual y física de un Digital Twin; se necesita un modelo genérico para un Digital Twin completo. La literatura que cubre la fusión de datos es otro tema muy investigado en todas las áreas de la ciencia, pero está menos investigado cuando se aplica la fusión de datos y los Digital Twin en un entorno industrial.

Del mismo modo, con el modelado de Digital Twin, otra vía es incorporar la fusión de datos al desarrollar modelos genéricos. La literatura explora cómo el mantenimiento predictivo puede incorporar Digital Twins con la fusión de datos, lo que es un área de investigación prometedora [90] [7].

Lo que falta en la investigación es el modelado de Digital Twin que utilizan la fusión de datos virtuales y físicos, lo que también puede ser hablado de la fusión "ciber física". Es necesario de enfoques estandarizados a la hora de modelar la fusión de datos con Digital Twin. [90] [7] y [91], exploran algunas de las formas en que la fusión de datos puede utilizarse para los Digital Twin en la industria, al tiempo que destacan algunos de los retos potenciales de la fusión de datos, desde los problemas de conexión hasta las amenazas a la seguridad.

Años de investigación en el mantenimiento de máquinas han dado lugar al término gestión de la salud y el pronóstico (PHM) utilizado principalmente en un entorno industrial, ya que puede aplicarse a la salud y los procesos de fabricación, desde las plantas pequeñas hasta las grandes plantas a gran escala [92] [93] [23]. La tecnología Digital Twin promueve PHM como potencial para la investigación en áreas de diagnóstico de fallos y el mantenimiento predictivo de los procesos industriales, que son tangibles con el desarrollo de los Digital Twin [10] [94].

El Digital Twin también permite la potencial interacción y colaboración de las máquinas, dando la posibilidad de simulación de procesos, facilitando el objetivo de una fabricación más precisa [95] [96] [82].

En la bibliografía revisada anteriormente, siete autores presentan los resultados de varios estudios de casos sobre implementaciones de Digital Twin. En resumen, [1] [35] y [2], los tres trabajos y estudios de casos discuten cómo se pueden utilizar los Digital Twin para optimizar los procesos de fabricación. En [1] se exploran los beneficios de la optimización al aplicar un Digital Twin a una cadena de montaje. Min et al. [35] investigan cómo el uso de Digital Twin puede optimizar la producción en la industria petroquímica. [2] presenta un trabajo sobre cómo los Digital Twin y el blockchain pueden para apoyar y optimizar la industria de fabricación de aeronaves.

Las tres primeras publicaciones mencionadas se refieren a estudios de casos que exploran la optimización, mientras que las cuatro siguientes, [74] [77] [27] y [80], se refieren a la modelización de sistemas Digital Twin. [74] investiga un caso práctico sobre cómo modelar y mantener un Digital Twin con sensores de IoT; [27] utiliza un estudio de caso para [27] utiliza un estudio de caso para explorar cómo desarrollar modelos de Digital Twin utilizando un marco multinivel y [80] explora el uso de un marco de modelado jerárquico para el desarrollo de Digital Twin. jerárquico para el desarrollo de Digital Twin. El último Karandiz et al. [77] presenta un estudio de caso sobre el modelado de una máquina de helados que explora cómo se puede utilizar un Digital Twin.

Los resultados específicos de cada uno de estos estudios de casos contribuyen a la conclusión general de la bibliografía, que se presenta en la sección 2.6.4.

### **2.6.2 Digital Twin en salud**

La siguiente sección analiza las preguntas de investigación abiertas para Digital Twins en un entorno sanitario. Algunas de las investigaciones citan el potencial para

adaptar DigitalTwin tecnología para humanos. Un ejemplo es un Digital Twin de una persona para monitorear día a día salud y bienestar dando el potencial para un gemelo humano para simular lo que cambios de estilo de vida positivos y negativos podrían tener en el ser humano físico. La importante investigación abierta viene en forma de modelado y desglose de las barreras para modelar un cuerpo humano a un Digital Twin. Una vez más, hasta el tema de no tener métodos estandarizados de modelado de Digital Twin [91]. Similar a PHM en fabricación, este concepto es un área emocionante de investigación con el DigitalTwin siendo para monitorear y mantener la salud de las personas. Desde el día a día de la asistencia sanitaria hasta condiciones de salud en curso, el Digital Twin se puede usar de manera similar a PHM, combinar el análisis de datos para garantizar que los pacientes estén sanos. [32] presenta investigación abierta relacionados con la cirugía; uso de datos históricos con datos actuales en tiempo real para Digital Twin simulaciones de cirugía y atención médica en general. El objetivo es detectar los riesgos antes de que surjan utilizando el gemelo virtual del paciente. Un área de investigación es en el campo de los datos.

Citando la necesidad de investigación para tratar con precisión los datos recopilados y procesados para un Digital Twin, principalmente porque se trata de pacientes sensibles datos sobre el Digital Twin virtual y físico, adecuada interacción y convergencia son necesarios para una mayor confianza y uso; de ahí la necesidad de más investigación.

La cirugía y la atención médica remotas son otra área emocionante de investigación. La capacidad de un médico de realizar controles previos a la cirugía de forma remota a través de un Digital Twin es una forma prometedora de minimizar el riesgo para la vida. Otro concepto presentado por Lakki et al. [32] es la investigación abierta para cirugía remota apoyada en red; esto está en línea con los desarrollos de 5G para cirugía en red móvil, otra área de investigación futura para Digital Twin en el cuidado de la salud [32] [40].

La fusión de datos, el modelado, la cirugía remota y la implementación de Digital Twins para el cuidado de la salud facilitan áreas de investigación más específicas. [40] [39] [32] y [37] todos citan la preocupación constante por la seguridad de la recopilación y el procesamiento de datos para un Digital Twins, más importante cuando se trata de datos confidenciales de un entorno de atención médica. El objetivo del trabajo futuro es garantizar la privacidad de los datos utilizados para un Digital Twin.

### **2.6.3 Digital Twin en ciudades inteligentes**

La última área de interés son las preguntas de investigación abiertas para Digital Twins asociadas con ciudades inteligentes. La revisión informada en la Sección 2.4 anterior muestra que esta área es similar a la atención médica en términos de investigación académica limitada. Actualmente, los documentos cubren gran parte de esta visión más amplia de una ciudad inteligente; cualquier cosa, desde Digital Twin para toda la ciudad para ciudades inteligentes hasta áreas más específicas como Digital Twin para la gestión del tráfico [22] [97] y gestión del ganado [15] hasta energías renovables [18]. Los temas abordaron cómo se puede usar el Digital Twin a gran escala para un gran pueblo o ciudad totalmente conectado.

Una oportunidad emocionante es la combinación de Digital Twin y la infraestructura local, como se explora en [3] y [83], haciendo uso del modelado 3D para el desarrollo y mantenimiento de ciudades inteligentes.

Las áreas abiertas para investigar vienen en la forma de aplicar análisis de datos, como el análisis predictivo aplicado a un Digital Twin para desarrollar una ciudad inteligente. Un Digital Twin estandarizado para una ciudad inteligente es imprescindible. Sin embargo, esto requiere más investigación en áreas para las interacciones de fusión de datos para el intercambio físico y virtual de datos para un Digital Twin. [22] y [97] enfatizan la necesidad de modelar en una ciudad inteligente con la demanda de modelos más genéricos que tengan en cuenta todos los componentes de una ciudad inteligente. Chen [22] presenta un concepto de gestión de tráfico integrado sin necesidad de desarrollos adicionales o modelado más genérico, estableciendo un estándar para la compatibilidad entre edificios y tráfico inteligentes. Con el auge y el desarrollo de los Digital Twin para la fabricación, es evidente que aumentan las oportunidades para las ciudades inteligentes [3] [83] [22].

### **2.6.4 Necesidades comunes y hallazgos específicos.**

#### **2.6.4.1 Modelos de datos**

El primer conjunto de hallazgos se relaciona con el modelo Digital Twin y su arquitectura. Hay una falta de modelos unificados o una arquitectura genérica de Digital Twin en la literatura, sin consenso sobre cómo construir un sistema de Digital Twin. Desarrollar un paradigma de diseño para construir un Digital Twin podría ser una forma genérica de implementar un sistema básico de Digital Twin.

#### **2.6.4.2 Sistemas heterogéneos**

Dado que el entorno del Digital Twin es heterogéneo, aunque también está conectado a grandes redes distribuidas, es un objetivo continuo que se observa en todo el campo para que los investigadores obtengan una comprensión más profunda de cómo lidiar con dichos sistemas. Esto es algo que se logrará evaluando y comparando los sistemas Digital Twin actuales entre sí, explorar cómo se diferencian en el manejo de entornos de Digital Twin heterogéneos más pequeños y grandes.

#### **2.6.4.3 Inteligencia artificial**

Con el avance de la digitalización, la IA es uno de los líderes y facilitadores del crecimiento y la adaptabilidad para permitir Digital Twin y se está convirtiendo en el componente principal de dichos sistemas. Las áreas abiertas de investigación que evalúan el impacto de la IA, la máquina y los algoritmos de aprendizaje profundo podrían tener potencialmente en los avances y usos de la tecnología Digital Twin. Estos avances en los conceptos de Industria 4.0 han enriquecido la forma en que vivimos, trabajamos y nos comunicamos, a su vez, abriendo más oportunidades de investigación a medida que aumenta la demanda de tecnología Digital Twin. AI allana el camino para más estudios de casos en un entorno industrial para evaluar el mantenimiento predictivo y la vida útil restante de los sistemas como un caso de estudio.

#### **2.6.4.4 Seguridad**

Con los avances en tecnología como blockchain, surgen varias oportunidades para garantizar la seguridad con el enfoque principal en explorar soluciones que ayuden a proteger a los Digital Twin.

#### **2.6.4.5 Intercambio de datos**

Al igual que la seguridad y la IA, con más conectividad viene el aumento de los sistemas Digital Twin y sus tecnologías habilitadoras. Más investigación garantizará que el Digital Twin pueda prosperar al compartir datos entre dispositivos. Un área abierta está buscando soluciones para lograr la integración perfecta de datos para pequeños sistemas de IoT, así como para grandes sistemas heterogéneos. El desarrollo de pequeños Digital Twin debe ampliarse, teniendo en cuenta el débil intercambio de datos.

#### **2.6.4.6 Internet de las cosas**

La mayor parte de la investigación abierta es desde una perspectiva de IoT cuando se piensa en Digital Twin. Debe haber una manera de adaptar los sensores para garantizar que el intercambio de datos sea preciso y funcione de la mejor manera posible. Edge Computing es otra área abierta de investigación para la tecnología IoT y Digital Twin.

#### **2.6.5 Desafíos de la investigación abierta**

Hay una serie de desafíos que se encuentran al tratar de abordar preguntas de investigación abiertas para Digital Twins, como se analiza a continuación.

##### **2.6.5.1 Multidisciplinariedad**

Un desafío importante proviene de los entornos multidisciplinarios utilizados para el diseño y el desarrollo. Estos surgen debido a los diferentes campos de investigación involucrados en la colaboración, lo que puede ayudar a los avances, pero también un obstáculo, ya que los objetivos en varios campos conducen a diferentes direcciones para la investigación, lo que en última instancia conduce a resultados más lentos.

##### **2.6.5.2 Estandarización**

Como se destaca en este documento y se ve en muchas tecnologías nuevas y emergentes, otro desafío que enfrenta Digital Twin es la falta de estandarización. Esto da como resultado discrepancias entre proyectos en Digital Twin. Un factor que contribuye es la variedad de definiciones vistas; esto, junto con la falta de estandarización, es un desafío que ralentiza el progreso de la tecnología Digital Twin.

##### **2.6.5.3 Avances globales**

Con los grandes avances en Digital Twins y sus tecnologías habilitadoras, vienen muchos beneficios, pero también desafíos. Los objetivos futuristas y poco realistas conducen a una adopción y desarrollo más lentos de tecnologías y opiniones recientemente adaptadas. Con el avance de la Industria 4.0 y el rápido crecimiento de la digitalización viene el crecimiento de las tecnologías a diferentes ritmos. Sin embargo, el crecimiento acelerado genera una variedad de desafíos con la conectividad y el intercambio de datos. Por ejemplo, es posible que los sistemas de soporte o los dispositivos IoT no sean compatibles con Digital Twin. Este crecimiento a diferentes ritmos también puede plantear nuevos desafíos para la seguridad en términos de exposición de vulnerabilidades.

## CAPÍTULO III DISEÑO TEÓRICO DEL DIGITAL TWIN

### 3.1 Componentes elementales

#### 3.1.1 Activo físico

El activo físico en evaluación será una planta de oxígeno medicinal de tipo PSA, las cuales serán monitorizadas mediante la extracción de datos enviados por un Gateway de Transporte de Telemetría de Cola de Mensajes (MQTT) en el Controlador Lógico Programable (PLC) a un servidor local (activo digital). El diagrama de la planta de oxígeno se muestra en la figura 6.

##### 3.1.1.1 Breve descripción del funcionamiento del activo físico

Las plantas de oxígeno se basan en la conocida tecnología PSA. Dos recipientes presurizados con zeolita aseguran una producción continua para la separación de oxígeno de otros gases. Un generador O<sub>2</sub> PSA consta de dos recipientes a presión que almacenan zeolita. El aire comprimido seco se sopla a través de una válvula al primer recipiente donde se construye la presión para alcanzar 4 a 8 bar(g). El gas no deseado (N<sub>2</sub>) es adsorbido por la zeolita durante la formación de presión, y el oxígeno pasará al tanque de acumulación. Mientras se acumula presión en un recipiente, el otro recipiente se regenera a través de una disminución de la presión. Este proceso garantiza un flujo continuo. Antes de la siguiente etapa de adsorción o desorción, se igualará la presión entre los recipientes de adsorción. Cuando la presión disminuye en el primer recipiente y la zeolita se satura, el producto deseado se expulsa a través del sistema de escape. Ahora la presión en el segundo recipiente aumentará hasta que la presión entre ambos recipientes se iguale nuevamente.

- PSA es la abreviatura de adsorción por cambio de presión.
- Presión: Se necesitan presiones elevadas para el proceso, normalmente de 4 a 8 bar(g). En caso de que se requiera una presión más alta para su proceso, también se puede aplicar.
- Balancear: Mientras se presuriza un recipiente, se despresuriza el otro recipiente.
- Adsorción: La separación del aire atmosférico se realiza con el principio PSA; El oxígeno se separa del aire a través del proceso de adsorción.

### 3.1.1.2 Lista detallada de parámetros a extraer

*Tabla 6*  
*Parámetros a extraer*

<b>Componente</b>	<b>Parámetro</b>	<b>Sensor</b>	<b>Controlador</b>	<b>Comunicación</b>
<b>Compresor de tornillo</b>	Voltaje	Transformador de voltaje	Inovance	Modbus TCP
	Corriente	Transformador de corriente	Inovance	Modbus TCP
	Presión	Transmisor de presión	Inovance	Modbus TCP
	Temperatura	Transmisor de temperatura	Inovance	Modbus TCP
<b>Secador de aire</b>	Punto de rocío	Transmisor de temperatura	Siemens S7-1200	Modbus TCP
	Presión	Transmisor de presión	Siemens S7-1200	Modbus TCP
<b>Torre de carbón</b>	Presión	Transmisor de presión	Siemens S7-1200	Modbus TCP
<b>Tanque de aire</b>	Presión	Transmisor de presión	Siemens S7-1200	Modbus TCP
<b>Generador de oxígeno (OGP)</b>	Pureza O2	Pureza	Siemens S7-1200	Modbus TCP
	CO O2	PPM	Siemens S7-1200	Modbus TCP
	CO2 O2	PPM	Siemens S7-1200	Modbus TCP
	Punto de rocío	Transmisor de temperatura	Siemens S7-1200	Modbus TCP
	Flujo O2	Transmisor de flujo	Siemens S7-1200	Modbus TCP

<b>Tanque O2</b>	Presión	Transmisor de presión	Siemens S7-1200	Modbus TCP
<b>Booster</b>	Presión 1 Etapa	Transmisor de presión	Siemens S7-1200	Modbus TCP
	Presión 2 Etapa	Transmisor de presión	Siemens S7-1200	Modbus TCP
	Presión 3 Etapa	Transmisor de presión	Siemens S7-1200	Modbus TCP
	Presión 4 Etapa	Transmisor de presión	Siemens S7-1200	Modbus TCP
	Temperatura 1 Etapa	Transmisor de temperatura	Siemens S7-1200	Modbus TCP
	Temperatura 2 Etapa	Transmisor de temperatura	Siemens S7-1200	Modbus TCP
	Temperatura 3 Etapa	Transmisor de temperatura	Siemens S7-1200	Modbus TCP
	Temperatura 4 Etapa	Transmisor de temperatura	Siemens S7-1200	Modbus TCP
	Voltaje	Transformador de voltaje	Siemens S7-1200	Modbus TCP
	Corriente	Transformador de corriente	Siemens S7-1200	Modbus TCP
<b>Manifold</b>	Presión	Transmisor de presión	Siemens S7-1200	Modbus TCP
<b>Ambiente</b>	Temperatura	Transmisor de temperatura	ESP8266	MQTT

Presión	Transmisor de presión	ESP8266	MQTT
Humedad	DHT22	ESP8266	MQTT
Potencia eléctrica	POW R2	-	MQTT
Puerta abierta	Sonoff ZB2	ZB Bridge	MQTT
Temperatura	Sonoff ZB3	ZB Bridge	MQTT
Movimiento	Sonoff ZB4	ZB Bridge	MQTT
Interruptor	Sonoff ZB5	ZB Bridge	MQTT
Switch	Sonoff ZB6	ZB Bridge	MQTT

*Nota:* Fuente: Elaboración propia.

### 3.1.1.3 Arquitectura usada para la obtención de oxígeno

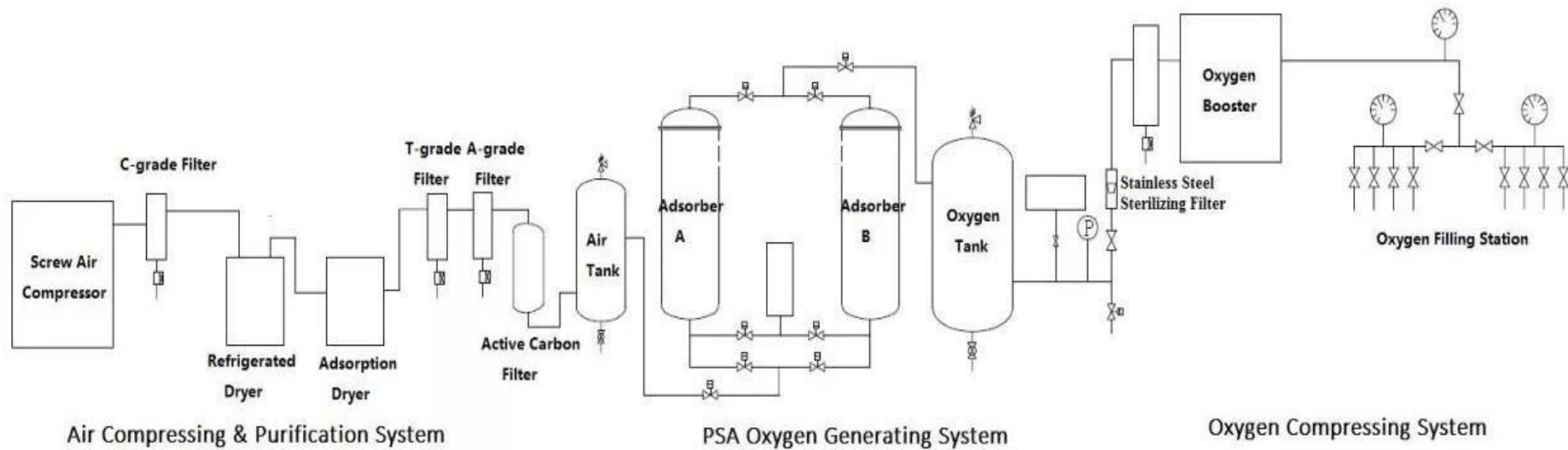


Figura 6 Diagrama de bloques de una planta de oxígeno PSA

Nota: Fuente: <https://m.n2o2gas.com/factory-directly-supply-nitrogen-generator-for-lcms-medical-o2-machine-ldh.html>

Según la figura 7 los sensores usados para la medición de parámetros en las plantas de oxígeno poseen una salida analógica de 4-20 mA, conociendo la limitada capacidad en cuestión de entradas analógicas del PLC S7-1200, se decidió por usar un módulo de expansión de entradas analógicas con comunicación Modbus TCP, esta información será enviada al PLC Siemens para su posterior publicación mediante el protocolo MQTT.

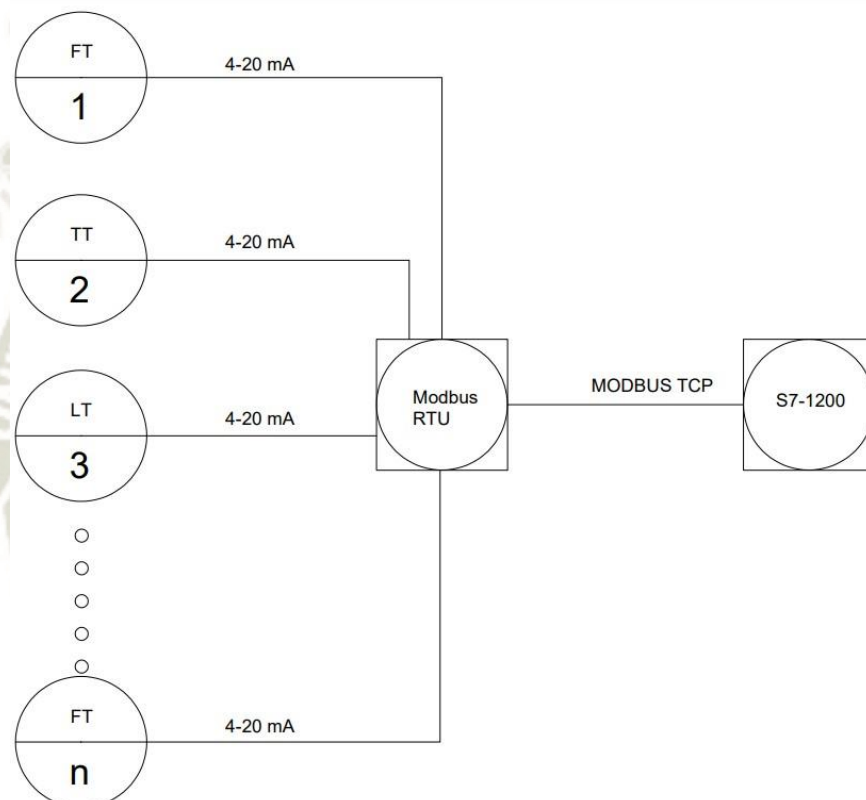


Figura 7 Arquitectura usada en la extracción de datos Siemens S7-1200

Nota: Fuente: Elaboración propia.

Tabla 7

*Lista de dispositivos usados en la extracción de datos Siemens S7-1200*

# de transmisor	Tipo	Señal	Marca	Modelo
1	Temperatura	4-20 mA	Yokogawa	YTA50
2	Presión	4-20 mA	Keller	PA-21Y
3	Presión	4-20 mA	Keller	PA-21Y
4	Presión	4-20 mA	Keller	PA-21Y
5	Pureza	4-20 mA	SOUTHLAND	OMD-755
6	Contaminación	4-20 mA	TESTO	360
7	Contaminación	4-20 mA	TESTO	370
8	Temperatura	4-20 mA	Yokogawa	YTA50
9	Flujo	4-20 mA	DWYER	HFT
10	Presión	4-20 mA	Keller	PA-21Y
11	Presión	4-20 mA	Keller	PA-21Y
12	Presión	4-20 mA	Keller	PA-21Y
13	Presión	4-20 mA	Keller	PA-21Y
14	Presión	4-20 mA	Keller	PA-21Y
15	Temperatura	4-20 mA	Yokogawa	YTA40
16	Temperatura	4-20 mA	Yokogawa	YTA50
17	Temperatura	4-20 mA	Yokogawa	YTA100
18	Temperatura	4-20 mA	Yokogawa	YTA200
19	Presión	4-20 mA	Keller	PA-21Y

*Nota:* Fuente: Elaboración propia.



*Figura 8 Modulo expansión entradas analógicas*

*Nota:* Fuente: Elaboración propia.



*Figura 9 Siemens S7-1200*

*Nota:* Fuente: <https://new.siemens.com/mx/es/productos/automatizacio n/systems/industrial/plc/s7-1200.html>

### 3.1.1.3 Arquitectura usada para la extracción de datos ESP8266

Según la figura 10 los sensores usados para la medición de parámetros en las plantas de oxígeno poseen una salida analógica de tipo protocolo de comunicaciones, por ejemplo: One Wire o I2C, esto debido a que conociendo la limitada capacidad en cuestión de entradas analógicas del ESP8266, solamente 1, se decidió por usar sensores por protocolo, netamente se usó una entrada de transmisor de temperatura de 0-5V para la entrada analógica.

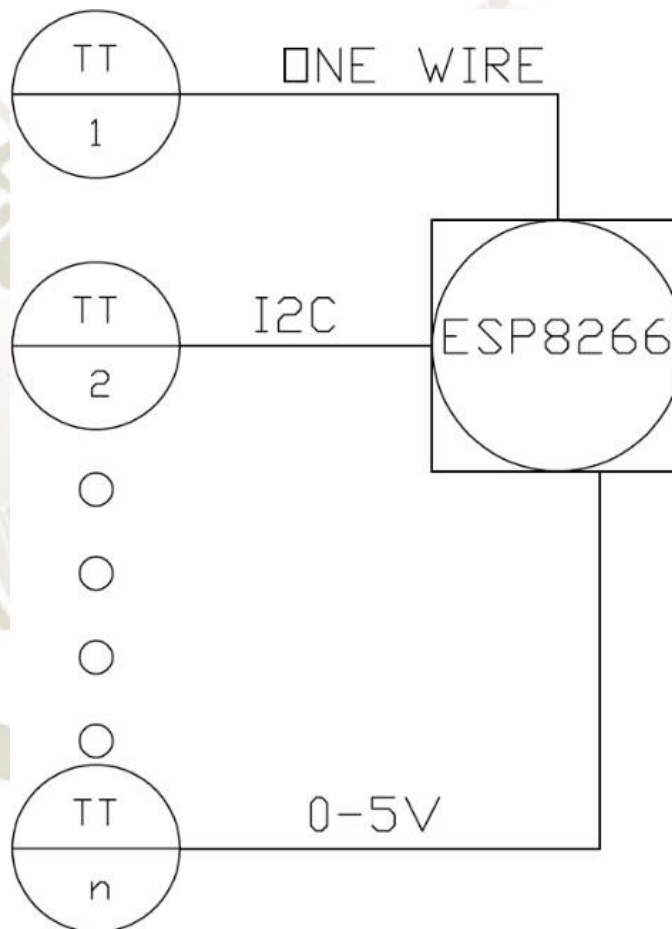


Figura 10 Arquitectura usada en la extracción de datos ESP8266

Nota: Fuente: Elaboración propia.

Tabla 8

Lista de dispositivos usados en la extracción de datos ESP8266

# de transmisor	Tipo	Señal	Marca	Modelo
1	Temperatura	One wire	DS1820	Dallas instruments
2	Temperatura	I2C	MCP5822	Microchip
3	Temperatura	0-5V	LM35	Texas instruments
4	Temperatura	I2C	MCP5822	Microchip

Nota: Fuente: Elaboración propia.

### 3.1.1.4 Arquitectura usada para la extracción de datos Zigbee

Según la figura 11 los sensores usados para la medición de parámetros en las plantas de oxígeno poseen son netamente de conexión inalámbrica mediante protocolo Zigbee, formando una red en malla teniendo al CC2652 como coordinador y principal responsable de las comunicaciones.

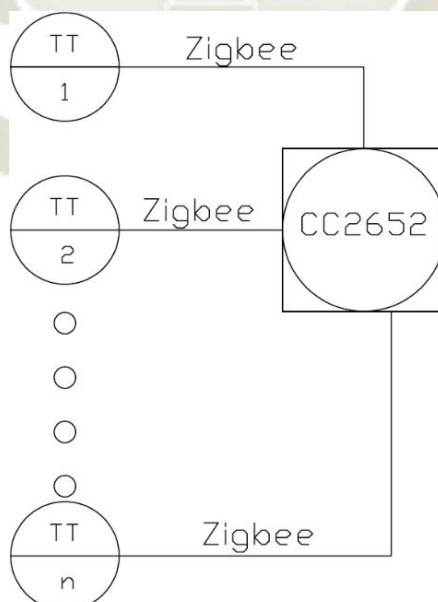


Figura 11 Arquitectura usada en la extracción de datos Zigbee

Nota: Fuente: Elaboración propia.

Tabla 9

*Lista de dispositivos usados en la extracción de datos Zigbee*

# de transmisor	Tipo	Señal	Marca	Modelo
1	Temperatura	Zigbee	ZB 01	Sonoff
2	Temperatura	Zigbee	ZB 01	Sonoff
3	Temperatura	Zigbee	ZB 01	Sonoff
4	Temperatura	Zigbee	ZB 01	Sonoff

*Nota:* Fuente: Elaboración propia.

### 3.1.1.5 Selección de componentes.

#### 3.1.1.5.1 PLC Siemens S7-1200

Los controladores SIMATIC S7-1200 son la opción ideal cuando se trata de realizar tareas de automatización de forma flexible y eficiente en el rango de rendimiento bajo a medio. Presentan una amplia gama de funciones tecnológicas y comunicación integrada, así como un diseño especialmente compacto y que ahorra espacio.

- Dispositivo previamente instalado en planta.
- La carga del proceso de la planta, accionamientos del proceso. No se debe sobrecargar al procesador.
- Limitación en cuanto a entradas analógicas.
- Compatibilidad Modbus TCP
- Posibilidad de enviar mensajes mediante protocolo MQTT.

Parámetros del sistema:

- Tensión de alimentación: 24VDC
- Consumo (valor nominal): 500mA, solo Unidad Central de Procesos (CPU)
- Con memoria integrada: 100 Kbyte
- Número de entradas digitales: 14
- Número de salidas digitales: 10
- Número de entradas analógicas: 2
- Interfaz PROFINET (estándar abierto de Ethernet)
- Soporta servidores web

- Dimensiones: 110x100x75mm

#### 3.1.1.6.2. Expansor de entradas analógicas con comunicación Modbus TCP

Dispositivo previamente instalado en planta, bajo coste de adquisición y buenas prestaciones.

Parámetros del sistema:

- CPU: chip importado estadounidense, procesador de alta velocidad de 32 bits para ATMEGA ARM, 72MHZ
- Sistema operativo: para GCOS, mecanismo de programación de 10 ms
- Voltaje de la fuente de alimentación: 7-30 V CC a 3 W, protección de conexión inversa de la fuente de alimentación, diseño de aislamiento
- Interfaz de comunicación: diseño de aislamiento DCDC, protección contra rayos de 2500 V, ESD, sobretensión, protección contra sobre corriente
- Dimensiones: Volumen: 155 x 95 x 39 mm (LxAnxAI)
- Método de instalación: fijación por tornillo, instalación en riel
- Indicadores ambientales: -40 a 85 °C, 5%~95%HR sin condensación, protección IP20
- Vigilancia de hardware: sistema de protección de vigilancia de hardware de 1segundo
- Aislamiento de canal: aislamiento de 2500 V CC, diseño de protección anti-interferencias.

#### 3.1.1.6.3. ESP8266

Se realizará una comparativa con el microcontrolador ESP32 de una gama superior. El ESP32 es el sucesor del ESP8266. Agrega un núcleo de CPU adicional, Wi-Fi más rápido, más GPIO y es compatible con Bluetooth 4.2 y Bluetooth de baja energía. Además, el ESP32 viene con pines sensibles al tacto que se pueden usar para despertar al ESP32 del sueño profundo, un sensor de efecto hall incorporado y un sensor de temperatura incorporado (las versiones recientes del ESP32 ya no vienen con un sensor de temperatura incorporado). Ambas placas son baratas, pero la ESP32 cuesta un poco más. Mientras que el ESP32 puede costar alrededor de 25 a 48 soles, el ESP8266 puede costar 45 a 60 soles.

Tabla 10 Comparativa entre ESP8266 y ESP32

	ESP8266	ESP32
<b>MCU</b>	Xtensa Single-core 32-bit L106	Xtensa Dual-Core 32- bit LX6 with 600 DMIPS
<b>802.11 b/g/n Wi-Fi</b>	HT20	HT40
<b>Bluetooth</b>	X	Bluetooth 4.2 y BLE
<b>Frecuencia</b>	80 MHz	160 MHz
<b>SRAM</b>	X	✓
<b>Flash</b>	X	✓
<b>GPIO</b>	17	34
<b>Hardware /Software PWM</b>	0 / 16 canales	0 / 16 canales
<b>SPI/I2C/I2S/UART</b>	2/1/2/2	4/2/2/2
<b>ADC</b>	10-bit	12-bit
<b>CAN</b>	X	✓
<b>Ethernet MAC Interfaz</b>	X	✓
<b>Sensor Touch</b>	X	✓
<b>Sensor de temperatura</b>	X	✓
<b>Sensor de efecto hall</b>	X	✓
<b>Temperatura de funcionamiento</b>	-40°C a 125°C	-40°C a 125°C

Nota: Fuente: Elaboración propia.

Parámetros del sistema (ESP8266):

- Procesador de 32 bit de bajo consumo
- Velocidad de 80MHz (máximo de 160MHz)
- 32 KiB Memoria de Acceso Aleatorio (RAM), instrucciones, 32 KiB RAMcache
- 80 KiB RAM para datos de usuario
- Memoria flash externa hasta 16MiB

- Pila de Protocolo de Control de Transmisión (TCP)/ Protocolo de internet (IP) integrada
  - Wifi 802.11 b/g/n 2.4GHz (soporta WPA/WPA2)
  - Certificado por FCC, CE, TELEC, WiFi Alliance y SRRC
  - 16 pins GPIO
  - PWM en todos los pines (10 bits)
  - Conversor analógico digital de 10 bits
  - UART (2x TX y 1x RX)
  - SPI, I2C, I2S
  - Voltaje de operación 3.0 a 3.6V
  - Consumo medio 80mA
  - Modo consumo stand-by (1mW) y deep sleep (1uA).
- 3.1.1.6.4. Zigbee Bridge

Parámetros del sistema:

- Voltaje de funcionamiento: 1,8 ~ 3,8 V, recomendado como 3,3 V
- Frecuencia de trabajo: 2402 MHz ~ 2480 MHz
- Potencia máxima de transmisión: + 20 dBm
- Sensibilidad de Recepción: -100 dBm @ 802.15.4 (2,4 GHz), -105 dBm @ Bluetooth 125-kbps (Código LE PHY)
- SRAM: 80 KB
- FLASH: 352 KB
- Cantidad de GPIO: 23
- Frecuencia de cristal: 48 MHz, 32.768KHz
- Interfaz de comunicación: UART, I2S, I2C, SPI, ADC
- Tamaño del módulo: 25x16,4mm
- Temperatura de funcionamiento: -40 °C ~ + 85 °C
- Temperatura de almacenamiento: -40 °C ~ + 125 °C

### 3.1.1.6 Comunicación MQTT

MQTT es el protocolo de mensajería más utilizado para Internet de las cosas (IoT). MQTT significa transporte de telemetría MQ. El protocolo es un conjunto de reglas que define cómo los dispositivos IoT pueden publicar y suscribirse a datos a través de Internet. MQTT se utiliza para mensajería e intercambio de datos entre

dispositivos IoT e IoT industrial (IIoT), como dispositivos integrados, sensores, PLC industriales, etc.

#### 3.1.1.6.1 Funcionamiento de MQTT

Con el objetivo de maximizar el ancho de banda disponible, el modelo de comunicación de publicación/suscripción (pub/sub) de MQTT es una alternativa a la arquitectura

cliente-servidor tradicional que se comunica directamente con un punto final. Por el contrario, en el modelo pub/sub, el cliente que envía un mensaje está desacoplado del cliente o clientes que reciben los mensajes (o los suscriptores). Debido a que ni los editores ni los suscriptores se contactan entre sí directamente, los terceros, los intermediarios, se encargan de las conexiones entre ellos.

Los clientes MQTT incluyen editores y suscriptores, términos que se refieren a si el cliente está publicando mensajes o suscrito para recibir mensajes. Estas dos funciones se pueden implementar en el mismo cliente MQTT. Cuando un dispositivo (o cliente) desea enviar datos a un servidor (o intermediario), se denomina *publicación*. Cuando la operación se invierte, se llama *suscripción*. Bajo el modelo pub/sub, varios clientes pueden conectarse a un broker y suscribirse a los temas que les interesan.

#### 3.1.1.6.2 Broker MQTT

Un broker MQTT actúa como intermediario entre los clientes que envían mensajes y los suscriptores que reciben esos mensajes. En una analogía con la oficina de correos, el intermediario es la propia oficina de correos. Todos los mensajes tienen que pasar por el broker antes de que puedan ser entregados al suscriptor.

### 3.1.2 Activo digital

El activo digital en esta investigación será implementado en un ordenador de bajo consumo, en esta ocasión se usará la plataforma Raspberry Pi 4B.

#### 3.1.2.1 Hardware principal ¿Por qué Raspberry?

Raspberry Pi elimina el costo del hardware de la computadora como una barrera para el descubrimiento y la experimentación. Basar el hardware en la tecnología ARM proporcionó la mejor base para ofrecer una computadora de bajo costo y alto rendimiento que pudiera aprovechar la innovación y la dedicación de la comunidad de código abierto.

Raspberry Pi ofrece alto rendimiento y confiabilidad a bajo costo; es portátil y usa muy poca energía, lo que lo convierte en un dispositivo perfecto para el monitoreo

y control industrial. Para ser utilizado en un entorno de TI empresarial, es útil tener un entorno operativo que cumpla con los requisitos empresariales de seguridad y soporte. Muchos clientes empresariales prefieren usar distribuciones empresariales de Linux porque requieren un sistema operativo seguro y compatible que ya saben cómo administrar.

#### Parámetros del sistema

- SoC Broadcom BCM2711, Cortex-A72 de cuatro núcleos (ARM v8) de 64 bits a 1,5GHz
- SDRAM LPDDR4-3200 de 1 GB, 2 GB, 4 GB u 8 GB (según el modelo)
- 2,4 GHz y 5,0 GHz IEEE 802.11ac inalámbrico, Bluetooth 5.0, BLE
- Gigabit Ethernet
- 2 puertos USB 3.0; 2 puertos USB 2.0.
- Encabezado GPIO de 40 pines estándar de Raspberry Pi (totalmente compatible conversiones anteriores de placas anteriores)
- 2 × puertos micro-HDMI (compatible con hasta 4kp60)
- Puerto de pantalla MIPI DSI de 2 carriles
- Puerto de cámara MIPI CSI de 2 carriles
- Puerto de audio estéreo y video compuesto de 4 polos
- H.265 (descodificación 4kp60), H264 (descodificación 1080p60, codificación 1080p30)
- OpenGL ES 3.1, Vulkan 1.0
- Ranura para tarjeta micro-SD para cargar sistema operativo y almacenamiento de datos
- 5V CC mediante conector USB-C (mínimo 3A\*)
- 5 V CC a través del cabezal GPIO (mínimo 3 A\*)
- Alimentación a través de Ethernet (PoE) habilitada (requiere PoE HAT separado)
- Temperatura de funcionamiento: 0 – 50 grados C ambiente

#### 3.1.2.1.1 Consumo

Según mediciones eléctricas hechas a la Raspberry Pi 4B se obtuvieron los siguientes datos:

- 3 W en reposo.
- 5.85 W reproduciendo vídeo.
- 6.25 W en estrés máximo.

Cálculos de consumo eléctrico:

- $6.25 \text{ W} \times 24 \text{ horas} = 150 \text{ W}$  al día.
- $150 \text{ W} \times 365 \text{ días} = 54750 \text{ W}$  al año = 54.75kWh al año.

Considerando el precio del consumo de electricidad con 0.76 soles por KWh

- $54.75 \times 0.76 = 41.61$  soles.

Lo que sumando al precio de adquisición del equipo con sus complementos (800 soles aproximadamente) resulta en un precio promedio de 2.3 soles por día en el uso del servidor de automatización, superando por mucho la oferta en cuanto a precio que se tiene con servicios en la nube.

#### 3.1.2.1.2 Velocidad.

- Procesador Broadcom BCM2711 1.5 GHz
- Memoria RAM LPDDR4, velocidad de 2.400MHz

Con estas características sumado a la utilización de contenedores en la implementación de los distintos servicios, tanto la memoria RAM como el uso del procesador no suele sobrepasar el 50% de su capacidad, nunca mermando su rendimiento, sumado a esto también se habilitó la inicialización del sistema desde un disco externo Disco de Estado Sólido (SSD) para hacerlo aún más rápido.

#### 3.1.2.2. Docker y los contenedores

Los contenedores son una solución al problema de cómo hacer que el software se ejecute de manera confiable cuando se mueve de un entorno informático a otro. Esto es lo que necesita saber sobre esta popular tecnología.

Docker irrumpió en escena en 2013 y desde entonces ha causado entusiasmo en los círculos de TI.

La tecnología de contenedor de aplicaciones proporcionada por Docker promete cambiar la forma en que se llevan a cabo las operaciones de TI tal como lo hizo la tecnología de virtualización unos años antes.

#### 3.1.2.2.1. Contenedores Docker y la necesidad de usarlos

Los contenedores son una solución al problema de cómo hacer que el software se ejecute de manera confiable cuando se mueve de un entorno informático a otro. Esto podría ser desde la computadora portátil de un desarrollador hasta un entorno de prueba, desde un entorno de prueba hasta la producción, y quizás desde una máquina física en un centro de datos hasta una máquina virtual en una nube privada o pública.

Los problemas surgen cuando el entorno de software de soporte no es idéntico, dice el creador de Docker, Solomon Hykes.

#### 3.1.2.2.2. Resolución de problemática

En pocas palabras, un contenedor consta de un entorno de tiempo de ejecución completo: una aplicación, más todas sus dependencias, bibliotecas y otros binarios, y los archivos de configuración necesarios para ejecutarlo, agrupados en un solo paquete. Al contener la plataforma de la aplicación y sus dependencias, se eliminan las diferencias en las distribuciones del sistema operativo y la infraestructura subyacente.

##### 3.1.2.1.1. *Diferencias entre contenedores y virtualización*

Con la tecnología de virtualización, el paquete que se puede pasar es una máquina virtual e incluye un sistema operativo completo además de la aplicación. Un servidor físico que ejecuta tres máquinas virtuales tendría un hipervisor y tres sistemas operativos separados ejecutándose sobre él.

Por el contrario, un servidor que ejecuta tres aplicaciones en contenedores con Docker ejecuta un solo sistema operativo y cada contenedor comparte el kernel del sistema operativo con los otros contenedores. Las partes compartidas del sistema operativo son desolo lectura, mientras que cada contenedor tiene su propio montaje (es decir, una forma de acceder al contenedor) para escritura. Eso significa que los contenedores son mucho más livianos y utilizan muchos menos recursos que las máquinas virtuales.

#### 3.1.2.2.3. *Otros beneficios.*

Un contenedor puede tener solo decenas de megabytes de tamaño, mientras que una máquina virtual con su propio sistema operativo completo puede tener varios

gigabytes de tamaño. Debido a esto, un solo servidor puede albergar muchos más contenedores que máquinas virtuales.

Otro beneficio importante es que las máquinas virtuales pueden tardar varios minutos en iniciar sus sistemas operativos y comenzar a ejecutar las aplicaciones que alojan, mientras que las aplicaciones en contenedores se pueden iniciar casi instantáneamente. Eso significa que los contenedores pueden crearse instancias "justo a tiempo" cuando se necesitan y pueden desaparecer cuando ya no se necesitan, liberando recursos en sus hosts.

Un tercer beneficio es que la contenedorización permite un mayor modularidad. En lugar de ejecutar una aplicación compleja completa dentro de un solo contenedor, la aplicación se puede dividir en módulos (como la base de datos, el front-end de la aplicación, etc.). Este es el llamado enfoque de microservicios. Las aplicaciones construidas de esta manera son más fáciles de administrar porque cada módulo es relativamente simple y se pueden realizar cambios en los módulos sin tener que reconstruir toda la aplicación. Debido a que los contenedores son tan livianos, los módulos individuales (o microservicios) se pueden instanciar solo cuando se necesitan y están disponibles casi de inmediato.

#### *3.1.2.2.4. Docker Compose.*

Docker Compose es una herramienta que se desarrolló para ayudar a definir y compartir aplicaciones de varios contenedores. Con Compose, podemos crear un archivo YAML para definir los servicios y, con un solo comando, podemos activar o desactivar todo.

La gran ventaja de usar Compose es que puede definir la pila de su aplicación en un archivo, mantenerla en la raíz del repositorio de su proyecto y permitir fácilmente que otra persona contribuya a su proyecto. Solo se necesitaría clonar el repositorio e iniciar la aplicación de redacción.

### **3.1.2.3. Contenedores a usar**

#### **3.1.2.3.1. Portainer**

Portainer está diseñado para organizaciones que no tienen acceso a la experiencia en administración de contenedores, pero que desean explotar el poder de los contenedores en la actualidad.

Con una Interfaz gráfica de usuario (GUI) intuitiva y un conjunto de valores predeterminados sensatos que hacen que los usuarios se pongan en marcha rápidamente, Portainer reduce drásticamente la necesidad de que las organizaciones

aprendan Docker y Kubernetes, lo que lleva a una adopción más rápida y a un ahorro de tiempo en toda la organización.

Reduce la complejidad operativa asociada a los contenedores, lo que acelera la adopción y reduce los errores. Aborda la escasez crítica de habilidades al hacer que la tecnología sea accesible de manera segura para todos dentro de la organización.

Simplifica la creación de configuraciones de seguridad 'seguras' dentro de Docker y Kubernetes a través de IAM centralizado.

#### 3.1.2.3.2. EMQX

EMQX (Erlang/Enterprise/Elastic MQTT Broker) es un intermediario de mensajes IoT/MQTT de código abierto basado en la plataforma Erlang/OTP.

Erlang/OTP es una excelente plataforma de desarrollo distribuida, de baja latencia y en tiempo real suave.

MQTT es un protocolo ligero de intercambio de mensajes que utiliza un patrón de publicación-suscripción.

EMQX está diseñado para el acceso masivo de clientes y realiza un enrutamiento de mensajes rápido y de baja latencia entre dispositivos de red físicos masivos.

Es estable para alojar conexiones de clientes MQTT a gran escala, y un solo nodo de servidor admite 2 millones de conexiones.

#### 3.1.2.3.3. Influxdb

InfluxDB es una base de datos de series temporales (TSDB) de código abierto desarrollada por la empresa InfluxData. Está escrito en el lenguaje de programación Go para el almacenamiento y la recuperación de datos de series de tiempo en campos como el monitoreo de operaciones, métricas de aplicaciones, datos de sensores de Internet de las cosas y análisis en tiempo real. También tiene soporte para procesar datos de Graphite.

InfluxDB no tiene dependencias externas y proporciona un lenguaje similar al Lenguaje de secuencia de Consultas (SQL), escuchando en el puerto 8086, con funciones integradas centradas en el tiempo para consultar una estructura de datos compuesta por medidas, series y puntos. Cada punto consta de varios pares clave-valor denominados conjunto de campos y una marca de tiempo. Cuando se agrupan por un conjunto de pares clave-valor llamado conjunto de etiquetas, estos definen una serie. Finalmente, las series se agrupan por un identificador de cadena para formar una medida.

Los valores pueden ser enteros de 64 bits, puntos flotantes de 64 bits, cadenas y valores booleanos. Los puntos están indexados por su tiempo y conjunto de etiquetas. Las políticas de retención se definen en una medida y controlan cómo se reducen y eliminan los datos. Las consultas continuas se ejecutan periódicamente y almacenan los resultados en una medición objetivo.

#### 3.1.2.3.4. Zigbee2MQTT

Zigbee2MQTT es una aplicación de puerta de enlace nodejs que conecta redes Zigbee a redes MQTT. Se ejecuta desde la línea de comandos, pero también se puede iniciar como un servicio en el inicio. Debido a que está escrito en nodejs, se ejecutará en Windows y Linux (Raspberry Pi).

#### 3.1.2.3.5. Telegraf

Telegraf es un agente basado en servidor para recopilar y enviar todas las métricas y eventos de bases de datos, sistemas y sensores IoT.

Telegraf está escrito en Go y se compila en un solo binario sin dependencias externas, y requiere una huella de memoria mínima.

#### 3.1.2.3.6. Traefik

Traefik es un proxy inverso y equilibrador de carga de código abierto. Se trata de un proxy perimetral que proporcione un enrutamiento simple sin todas las funciones de NGINX y HAProxy, Traefik es una buena opción. El enrutador proporciona una reconfiguración sin recarga, métricas, monitoreo y disyuntores que son esenciales cuando se ejecutan microservicios. También se integra muy bien con Let's Encrypt para proporcionar terminación de capa de sockets seguros (SSL), así como componentes de infraestructura como Kubernetes, Docker Swarm o Amazon ECS para seleccionar automáticamente nuevos servicios o instancias para incluir en su equilibrio de carga.

#### 3.1.2.3.7. Nodered

Node-RED es una herramienta de programación para conectar dispositivos de hardware, API y servicios en línea de formas nuevas e interesantes.

Proporciona un editor basado en navegador que facilita la conexión de flujos mediante la amplia gama de nodos de la paleta que se pueden implementar en su tiempo de ejecución con un solo clic.

### 3.1.3 Relación sincronizada bidireccional entre ambos.

La relación sincronizada entre el activo físico y digital se hará mediante el protocolo MQTT, con el uso de librerías en el caso del PLC S7-1200 y con el uso del gestor de datos Telegraf en el lado del activo digital.

## 3.2. Componentes imperativos

### 3.2.1. Dispositivos IoT

#### 3.2.1.1. Librería MQTT S7-1200

La biblioteca "LMQTT\_Client" ofrece una solución adecuada para implementar el protocolo MQTT en un controlador SIMATIC S7. La biblioteca "LMQTT\_Client" ofrece un bloque de función para SIMATIC S7-1500 y SIMATIC S7-1200. El bloque de funciones "LMQTT\_Client" integra la función de cliente MQTT y permite transferir mensajes MQTT a un intermediario (función de editor) y crear suscripciones (función de suscriptor). La comunicación se puede asegurar a través de una conexión de Capa de Transporte de Seguridad (TLS).

#### 3.2.1.2. Ventajas del ESP8266

El módulo WiFi ESP8266 es un módulo WiFi de montaje en superficie potente, pero rentable, con un sistema en chip (SoC) ESP8266 integrado. Es uno de los módulos WiFi más populares de bajo costo (15 a 40 soles) para controlar dispositivos a través de Internet.

El módulo tiene un transceptor WiFi inalámbrico que opera en un rango de frecuencia sin licencia de 2400-2484 MHz en el estándar IEEE 802.11 b/g/n, con soporte para la pila de protocolos de comunicación TCP/IP y seguridad WiFi que incluye WPA3. Puede funcionar con firmware para proporcionar conectividad Wi-Fi a Microcontroladores (MCU) host externas como Arduino o como MCU programables autosuficientes con un Kit de Desarrollo de Software (SDK) basado en Sistemas Operativos de Tiempo Real (RTOS) que puede ejecutar aplicaciones.

Además, es capaz de procesar datos a bordo, integrándose con sensores y otras aplicaciones a través de GPIO. Estos son una serie de módulos producidos por Espressif, Ai-Thinker y algunos otros fabricantes con modelos específicos como el transceptor inalámbrico ESP- WROOM-02, ESP-01 y ESP8266.

Características del módulo WiFi ESP8266

- Un rendimiento de RF superior con certificaciones de RF, ecológicas y de fiabilidad.

- Alta integración con conmutadores de antena, amplificadores, filtros, balun de RF, PLL, reguladores, DCXO y módulos de administración de energía integrados en el nivel SoC.
- Algunos módulos como el ESP8266EX tienen integrado un procesador RISC de 32 bits de bajo consumo que incluso puede alcanzar una velocidad de reloj máxima de 160MHz.
- Se admiten 96 KiB internos de RAM de datos, 64 KiB de RAM de instrucciones y memoria flash QSPI externa de 512 KiB hasta 16 MiB.
- Excelente rendimiento con hasta el 80 % de la potencia de procesamiento disponible para la programación y el desarrollo de aplicaciones de usuario.
- Conectividad en la nube lista para usar con Wi-Fi Direct (P2P) y soft-AP.
- Activa y transmite paquetes en menos de 2 ms con un consumo de energía en espera de menos de 1 mW (DTIM3) y una corriente de fuga de apagado de no más de 10  $\mu$ A.
- Soporta APSD para aplicaciones de VoIP; códigos de bloque de espacio-tiempo (STBC); Configuraciones 1x1 MIMO, 2x1 MIMO; Interfaces SDIO 1.1 / 2.0, SPI, I2C, UART, I<sup>2</sup>S con DMA; Agregación de A-MPDU y A-MSDU.

### 3.2.1.3. Ventajas de Zigbee

ZigBee es una variante de baja potencia y velocidad de datos de protocolos inalámbricos como WiFi, que utiliza muchas de las mismas tecnologías y protocolos de redes, seguridad y prevención de interferencias. Zigbee opera con la especificación de radio física IEEE

802.15.4 en sus dispositivos.

La especificación 802.15.4 en la que opera ZigBee obtuvo la aprobación oficial del Instituto de Ingenieros Eléctricos y Electrónicos (IEEE) en 2003. La especificación IEEE 802.15.4 es un protocolo de radio basado en paquetes diseñado para dispositivos de bajo costo que funcionan con baterías. para mejorar la eficiencia de la interoperabilidad.

Zigbee está diseñado para transportar pequeñas cantidades de datos a corta distancia y consumir muy poca energía. A diferencia de WiFi, que utiliza una red en estrella, Zigbee utiliza un estándar de red de malla, lo que significa que cada nodo de la red está conectado entre sí.

Hay cuatro tipos de dispositivos ZigBee:

- Enrutador: un enrutador ZigBee es un ciudadano típico de una red de malla. Enruta

mensajes hacia y desde otros dispositivos, incluidos otros enrutadores. En una situación ideal, cada enrutador es redundante.

En consecuencia, si un enrutador pierde la conexión a la red, el tráfico hacia los nodos restantes de la red permanece intacto gracias al reenrutamiento y la redundancia.

- **Dispositivo final:** un dispositivo final ZigBee no enruta el tráfico a otros nodos. Suele estar dormido la mayor parte del tiempo. Los dispositivos finales típicos de ZigBee son periféricos simples como interruptores, atenuadores, etc.
- **Coordinador:** Un coordinador ZigBee es un enrutador ZigBee con todas las características de un enrutador. El coordinador también controla la configuración de la red, es decir, confirma nuevos nodos para unirse a la red.
- **Centro de confianza:** el centro de confianza de ZigBee administra las claves de cifrado y descifrado y distribuye las claves a los nuevos nodos. El centro de confianza es casi siempre el mismo dispositivo que el coordinador.

#### **3.2.1.4. Wi-Fi vs Zigbee**

Tanto Wi-Fi como ZigBee tienen sus cualidades positivas, pero obviamente tienen aspectos negativos. Lo que gana en ancho de banda con WiFi se pierde en potencia de batería y operatividad, y lo que gana con la eficiencia energética de ZigBee lo pierde en alcance y ancho de banda. Entonces, como cualquier decisión basada en presupuestos de enlaces, comprender las compensaciones es crucial.

##### **3.2.1.4.1. Alcance inalámbrico**

Tanto los dispositivos inteligentes WiFi como Zigbee funcionan de forma inalámbrica, lo que les brinda mayor comodidad y simplifica la instalación. Los dispositivos WiFi y Zigbee funcionan con frecuencias de 2,4 GHz, que ofrecen un mayor alcance y permeabilidad a través de paredes y puertas. Sin embargo, debido a que funcionan en la misma frecuencia, pueden interferir entre sí si se colocan varios dispositivos uno al lado del otro.

Por lo tanto, los mejores sistemas domésticos inteligentes que incorporan dispositivos de los dos protocolos organizan los componentes de manera lineal y sin interrupciones.

Los dispositivos que funcionan con WiFi tienen un alcance considerablemente mayor que Zigbee, pudiendo comunicarse y operar desde 30

metros y hasta 100 metros. Por otro lado, los dispositivos Zigbee tienen un alcance de solo unos 20 metros.

#### 3.2.1.4.2. Tipo de red

El tipo de red que utiliza un sistema determina cómo se conecta cada dispositivo al concentrador central, que sirve como unidad de control para todos los componentes conectados. Los sistemas de adquisición de datos inteligentes utilizan principalmente dos tipos de redes: redes en estrella y en malla.

En el desarrollo de este prototipo se usarán ambos, debido a la necesidad de un ancho de banda grande para ciertos sensores, se usará la red WiFi (estrella) y para dispositivos de información discreta o de bajo ancho de banda se usará la red Zigbee (malla).

#### 3.2.1.4.3. Diferencias entre redes

En una red en estrella, cada dispositivo se comunica directamente con (y solo con) el concentrador central. Si un dispositivo está fuera del alcance del concentrador central, ese dispositivo no formará parte de la red.

Una red de malla permite que cada dispositivo de la red funcione como un repetidor y pase la señal a otro dispositivo. Esto permite que las redes de malla sean más versátiles. Pueden cubrir mayores distancias e incluso sortear obstáculos.

WiFi utiliza una red en estrella y todos sus dispositivos se conectan a un enrutador WiFi. Alternativamente, los dispositivos Zigbee usan una red de malla y se conectan entre sí para llegar a un concentrador Zigbee. Con la red de malla, un sistema doméstico inteligente con dispositivos Zigbee puede sincronizarse de manera más coherente y disfrutar de una gama más amplia de conectividad en general.

Es importante tener en cuenta que los dispositivos Zigbee que funcionan con baterías no se comportan como repetidores y no pueden pasar señales a otro dispositivo.

#### 3.2.1.4.1. El consumo de energía

Hablando específicamente sobre el consumo de energía, los sistemas WiFi tienden a consumir mucha más energía que sus contrapartes Zigbee.

Por otro lado, las redes basadas en ZigBee generalmente consumen solo alrededor del 25% de la potencia de las redes WiFi. Los dispositivos Zigbee funcionan con una potencia ultra baja, en comparación con WiFi, y consumen energía a niveles de microamperios. La duración de la batería de ZigBee es una gran ventaja sobre WiFi, ya

que dura meses y mucho más que los dispositivos WiFi, y debe tenerse muy en cuenta si sus terminales funcionarán con baterías.

#### 3.2.1.4.4. Velocidad y ancho de banda

Los dispositivos que funcionan con WiFi, aunque consumen más energía, también pueden transmitir más datos más rápido que Zigbee. La mayoría de las redes tienen una velocidad máxima de transferencia de datos de 11 Mbps, lo que las hace perfectas para transmisión de video en vivo en Full HD, como la de nuestro Buzz Smart Wireless Doorbell o Trace Video Tracking Camera.

El protocolo Zigbee tiene velocidades mucho más lentas que WiFi, con velocidades de transferencia de datos de solo 250 kbps. Debido a esta menor velocidad de transferencia, los dispositivos Zigbee pueden consumir cantidades minúsculas de energía. El dispositivo conectado Variantz Zigbee, como el contacto magnético inteligente Zig Snap, es uno de esos dispositivos que no sufre debido a las velocidades de datos más bajas, pero se beneficia enormemente del menor consumo de energía.

#### 3.2.1.4.5. Interferencias

Los dispositivos WiFi son más propensos a sufrir interferencias y es más probable que se vuelvan inoperables. Por un lado, WiFi que se conecta directamente a su enrutador genera una carga de datos significativa en su conexión WiFi e Internet. Es probable que esto genere velocidades de Internet más bajas o una conexión WiFi más deficiente si hay demasiados dispositivos conectados a través de WiFi y la carga es demasiado severa. Además, los dispositivos WiFi dependen de múltiples nubes de Internet para su funcionamiento. Si el fabricante decide dejar de funcionar, esto puede hacer que sus dispositivos WiFi queden obsoletos.

Alternativamente, es menos probable que los dispositivos Zigbee experimenten tales dificultades. Con un Smart Gateway, un concentrador Zigbee central procesa todos los datos localmente y no depende de la nube. Esto significa que no hay tensión en la red WiFi existente y que los componentes de la edificación inteligente siguen siendo completamente funcionales, incluso si el fabricante decide detenerse. Además, con Zigbee, incluso cuando Internet no funciona, aún se puede controlar el edificio inteligente. Zigbee no depende de Internet y los dispositivos funcionan normalmente siempre que el concentrador esté conectado

#### 3.2.1.4.6. Pruebas de velocidad en Zigbee

Mediante la realización de un flujo de funciones en Node-Red, se pudo medir la velocidad de respuesta entre el coordinador Zigbee y el dispositivo más lejano de la red, obteniéndolos siguientes resultados:

Siendo 58 ms la lectura de más tiempo, verificamos la rapidez de respuesta en Zigbee.

#### 3.2.1.4.7. Selección de dispositivos

Para la selección de los dispositivos que se usaran para la obtención de

```

25/10/2020 12:16:02 node: 9767aaaf.c59478
msg.payload : string[53]
"Total: 11 Fallidas: 0 Tiempo medio:
58.96366080605952"

25/10/2020 12:16:02 node: 9767aaaf.c59478
msg.payload : string[54]
"Total: 11 Fallidas: 0 Tiempo medio:
38.255502993080825"

25/10/2020 12:16:02 node: 9767aaaf.c59478
msg.payload : string[53]
"Total: 11 Fallidas: 0 Tiempo medio:
47.93608244987995"

25/10/2020 12:16:02 node: 9767aaaf.c59478
msg.payload : string[53]
"Total: 11 Fallidas: 0 Tiempo medio:
33.22485126063206"
    
```

*Figura 12 Tiempo de respuesta sensores Zigbee*

*Nota:* Fuente: Elaboración propia.

variables, se debetener en cuenta:

- Cantidad de datos.
- Fuente de alimentación.
- Lejanía.

En caso de tener una cantidad de datos de tipo discreta o baja, de trabajar con una fuente de alimentación a base de baterías y estar lejos del concentrador, se usará un dispositivo Zigbee, en caso de que los datos tengan un gran volumen, se

tenga una fuente de alimentación ininterrumpida y no sea tan lejano, se usará un dispositivo WiFi.

### **3.2.2. Herramientas de almacenamiento**

#### **3.2.2.1. Bases de datos de series temporales vs relacionales**

Una base de datos de series temporales es un tipo de base de datos que está optimizada para datos de series temporales y marcas de tiempo. Las series temporales incluyen mediciones o eventos que se rastrean, monitorean, muestrean y agregan a lo largo del tiempo. Pueden estar relacionados con las métricas del servidor, el rendimiento de las aplicaciones, los datos de la red, los datos de los sensores y otros tipos de datos analíticos.

Una base de datos de series temporales está optimizada para medir el cambio a lo largo del tiempo. Está diseñada para manejar métricas, eventos y medidas con marca de tiempo y permite a los usuarios crear, actualizar, destruir y organizar varias series temporales de manera más eficiente. La diferencia clave entre la base de datos de series temporales y una base de datos relacional está en la forma en que se gestiona el orden de los datos. Mientras que las bases de datos relacionales no implican un orden particular de los datos insertados (se debe usar la cláusula `order by`), las bases de datos de series temporales indexan y procesan los datos con el objetivo de optimizar las consultas que implican un orden en el tiempo. Por lo tanto, una de las principales diferencias entre los dos es la forma en que se implementan los índices. De lo contrario, las bases de datos relacionales y de series temporales tienen una arquitectura similar. Otras diferencias notables incluyen las ventajas y desventajas cuando se trata de adherirse a las propiedades ACID frente al rendimiento. Las bases de datos de series temporales no tienen por qué garantizar necesariamente la durabilidad y la consistencia sólida, ya que los datos de las series temporales suelen ser inmutables y se generan en un punto único en el tiempo.

Además, no es necesario recopilar todos los datos posibles, cuando se recopilan continuamente datos de un sensor, puede que no sea necesario almacenar todos los puntos de datos. Por ejemplo, InfluxDB proporciona consistencia eventual.

Hoy en día, un número cada vez mayor de empresas genera grandes flujos de métricas y eventos, lo que genera la necesidad de bases de datos de series temporales para facilitar el acceso eficiente a los datos. Sin embargo, las bases de datos de series temporales, debido a sus optimizaciones, no siempre muestran propiedades favorables a la hora de abordar tareas de propósito más general. Desde el punto de vista de la

ingeniería, es importante comprender en qué casos de uso las bases de datos de series temporales funcionan mejor en lo que respecta al tiempo de ejecución de consultas, así como al consumo general de recursos. Se ha dedicado cierto esfuerzo de investigación a abordar estos problemas. Por ejemplo, los autores prueban el rendimiento de InfluxDB en un conjunto de datos del mundo real muy grande que contiene 450 millones de lecturas de consumo de energía. Los autores centran las pruebas en la velocidad de inserción de datos y los parámetros de consumo de almacenamiento en disco. En nuestro artículo, nuestro objetivo es fortalecer aún más los resultados encontrados en la literatura mediante la realización de experimentos adicionales en un conjunto de datos del mundo real. Además de los experimentos de inserción de datos, también realizamos varios experimentos de agregación de datos.

Actualmente, una de las bases de datos de series de tiempo más populares es InfluxDB y, como tal, se ha decidido utilizarla con fines de comparación de rendimiento con una base de datos relacional. InfluxDB es una base de datos de series temporales sin esquema de código abierto con un conjunto de componentes opcionales. Está escrito en el lenguaje de programación Go y está optimizado para manejar datos de series temporales en particular. El entorno de la base de datos proporciona InfluxQL, un lenguaje de consulta similar a SQL. La versión de código abierto TICK Stack proporciona una plataforma de base de datos de serie temporal completa, incluido el núcleo InfluxDB.

Todos los datos en la base de datos InfluxDB tienen una columna llamada hora. La hora de la columna almacena marcas de tiempo en formato RFC3339 UTC y está asociada con datos particulares. Otra parte requerida de la estructura de datos de InfluxDB son los campos que constan de una clave de campo y un valor de campo. Las claves de campo son cadenas y almacenan metadatos. Los valores de campo pueden ser de tipo cadena, flotante, entero o booleano y representan los datos. La colección de pares de campo-clave y campo-valor forman un conjunto de campos. Los campos no están indexados y son una pieza requerida de la estructura de datos de InfluxDB. Las etiquetas son opcionales y consisten en una clave de etiqueta y un valor de etiqueta que son cadenas que almacenan metadatos. Diferentes combinaciones de todos los pares clave de etiqueta y valor de etiqueta forman un conjunto de etiquetas. A diferencia de los campos, las etiquetas están indexadas y son opcionales. La medición actúa como contenedor de etiquetas, campos y la columna de tiempo y el nombre de la medición representa la descripción de los datos que se almacenan en los campos asociados. Una

misma medida puede pertenecer a diferentes políticas de retención. La política de retención describe cuánto tiempo InfluxDB conserva los datos. En una base de datos InfluxDB, una serie es una colección de datos que comparten una política de retención, una medida y un conjunto de etiquetas. Finalmente, un punto es el campo establecido en la misma serie con la misma marca de tiempo.

### 3.2.3. Aprendizaje automático

Loud ML es la primera API de aprendizaje profundo de código abierto que simplifica la preparación, el entrenamiento y la implementación de modelos de aprendizaje automático y analiza los datos almacenados en sus bases de datos favoritas sin mover los datos. El usuario selecciona la serie temporal que desea modelar y establece los rangos de fechas del modelo, luego Loud ML construirá los modelos y los guardará para la inferencia en producción. Loud ML hace todo el trabajo y elimina la complejidad del aprendizaje automático con Tensorflow.

A continuación, algunos ejemplos de casos de uso para los que se usa Loud ML:

- Detectar caídas anormales en el tráfico de usuarios y responder a incidentes antes de que afecten la satisfacción de los clientes.
- Detección de valores atípicos en las fluctuaciones estacionales del comercio electrónico.
- Detección de carga anormal en una base de datos distribuida.
- Detectar dinámicamente patrones de tráfico de red y anticipar la congestión antes de que afecte la experiencia del cliente.
- Pronóstico de capacidad, uso y desequilibrio de carga para productores y proveedores de energía.
- Pronóstico de la demanda de inventario y optimización de la cadena de suministro.
- Detección de patrones de fraude anormales para operadores de redes móviles.
- Predecir fallas en equipos de red para la planificación de operaciones de mantenimiento.
- Anticipo de la capacidad del disco y descubra los problemas de capacidad antes de que genere problemas.
- Conociendo la carga futura de antemano y escalando automáticamente las instancias virtuales en la nube.

- Conocer de antemano la carga futura y ahorrar energía en los centros de datos.

El aprendizaje supervisado es un enfoque para crear inteligencia artificial (IA), donde un algoritmo de computadora se entrena con datos de entrada que han sido etiquetados para una salida en particular. El modelo se entrena hasta que puede detectar los patrones subyacentes y las relaciones entre los datos de entrada y las etiquetas de salida, lo que le permite generar resultados de etiquetado precisos cuando se le presentan datos nunca vistos.

El aprendizaje supervisado es bueno en problemas de clasificación y regresión, como determinar a qué categoría pertenece un artículo de noticias o predecir el volumen de ventas para una fecha futura determinada. En el aprendizaje supervisado, el objetivo es dar sentido a los datos dentro del contexto de una pregunta específica.

A diferencia del aprendizaje supervisado, se encuentra el aprendizaje no supervisado. En este enfoque, el algoritmo se presenta con datos sin etiquetar y está diseñado para detectar patrones o similitudes por sí solo, un proceso que se describe con más detalle a continuación.

¿Cómo funciona el aprendizaje supervisado?: Como todos los algoritmos de aprendizaje automático, el aprendizaje supervisado se basa en el entrenamiento. Durante su fase de entrenamiento, el sistema se alimenta con conjuntos de datos etiquetados, que le indican al sistema qué salida está relacionada con cada valor de entrada específico. Luego, al modelo entrenado se le presentan datos de prueba: estos son datos que se han etiquetado, pero las etiquetas no se han revelado al algoritmo. El objetivo de los datos de prueba es medir con qué precisión funcionará el algoritmo en datos no etiquetados.

En los algoritmos de redes neuronales, el proceso de aprendizaje supervisado se mejora midiendo constantemente las salidas resultantes del modelo y ajustando el sistema para acercarse a la precisión de su objetivo. El nivel de precisión que se puede obtener depende de dos cosas: los datos etiquetados disponibles y el algoritmo que se utiliza. Además, los datos de entrenamiento deben equilibrarse y limpiarse. Los datos basura o duplicados distorsionarán la comprensión de la IA; por lo tanto, los científicos de datos deben tener cuidado con los datos en los que se entrena el modelo.

La diversidad de los datos determina qué tan bien funcionará la IA cuando se le presenten nuevos casos; si no hay suficientes muestras en el conjunto de datos de entrenamiento, el modelo fallará y no brindará respuestas confiables.

Una alta precisión, paradójicamente, no es necesariamente una buena indicación; también podría significar que el modelo sufre de sobreajuste, es decir, está sobre ajustado a su conjunto de datos de entrenamiento particular. Dicho conjunto de datos podría funcionar bien en escenarios de prueba, pero fallar miserablemente cuando se le presenten desafíos del mundo real. Para evitar el sobreajuste, es importante que los datos de prueba sean diferentes de los datos de entrenamiento para garantizar que el modelo no obtenga respuestas de su experiencia previa, sino que la inferencia del modelo sea generalizada.

El algoritmo, por otro lado, determina cómo se pueden utilizar esos datos. Por ejemplo, los algoritmos de aprendizaje profundo se pueden entrenar para extraer miles de millones de parámetros de sus datos y alcanzar niveles de precisión sin precedentes, como lo demuestra GPT-3 de OpenAI.

Además de las redes neuronales, existen muchos otros algoritmos de aprendizaje supervisado (ver más abajo). Los algoritmos de aprendizaje supervisado generan principalmente dos tipos de resultados: clasificación y regresión.

**Algoritmos de clasificación:** Un algoritmo de clasificación tiene como objetivo ordenar las entradas en un número determinado de categorías o clases, en función de los datos etiquetados en los que se entrenó. Los algoritmos de clasificación se pueden usar para clasificaciones binarias, como filtrar el correo electrónico en spam o no spam y categorizar los comentarios de los clientes como positivos o negativos. El reconocimiento de características, como el reconocimiento de letras y números escritos a mano o la clasificación de medicamentos en muchas categorías diferentes, es otro problema de clasificación que se resuelve mediante el aprendizaje supervisado.

**Modelos de regresión:** Las tareas de regresión son diferentes, ya que esperan que el modelo produzca una relación numérica entre los datos de entrada y salida. Los ejemplos de modelos de regresión incluyen la predicción de precios inmobiliarios según el código postal, o la predicción de tasas de clics en anuncios en línea en relación con la hora del día, o la determinación de cuánto estarían dispuestos a pagar los clientes por un determinado producto en función de su edad.

Los algoritmos comúnmente utilizados en los programas de aprendizaje supervisado incluyen los siguientes:

- Regresión lineal
- Regresión logística
- Redes neuronales
- Análisis discriminante lineal
- Árboles de decisión
- Aprendizaje de similitud
- Lógica bayseiana
- Máquinas de vectores de soporte (SVM)
- Bosques aleatorios

Al elegir un algoritmo de aprendizaje supervisado, hay algunas cosas que deben tenerse en cuenta. El primero es el sesgo y la varianza que existen dentro del algoritmo, ya que existe una delgada línea entre ser lo suficientemente flexible y demasiado flexible. Otro es la complejidad del modelo o función que el sistema está tratando de aprender. Como se señaló, la heterogeneidad, la precisión, la redundancia y la linealidad de los datos también deben analizarse antes de elegir un algoritmo.

### **3.2.4. Seguridad**

#### **3.2.4.1. SSH**

Secure Shell (SSH) es un protocolo criptográfico para conectarse a servicios de red a través de una red no segura. Las aplicaciones comunes para SSH son el inicio de sesión remoto y la ejecución remota de comandos en los hosts, pero eso solo rasca la superficie de lo que puede hacer con SSH.

Un beneficio clave de SSH es su casi ubicuidad. Dado que OpenSSH se lanzó en 1999, SSH se puede encontrar en casi todos los sistemas operativos. Las aplicaciones SSH se basan en una arquitectura cliente-servidor, con un cliente SSH conectado a un servidor SSH.

Se reparará un ejemplo simple: se ha creado un nuevo servidor en un proveedor de nube. Durante la configuración, tuve que proporcionar mis claves SSH, o más específicamente, la clave pública SSH. Cuando se inicia el servidor, esta clave se agrega al servidor, el servidor se inicia y el proceso del servidor SSH espera una conexión. Ahora se puede conectar al host mediante SSH, `ssh ubuntu@host`, casi de forma

transparente. El terminal local se siente como si estuviera conectado al host con un monitor.

SSH es un protocolo, no un producto. El protocolo cubre 3 áreas principales.

- **Autenticación:** La autenticación implica probar la identidad de un usuario o de una máquina; es decir, quienes dicen ser. Para SSH, esto significa confirmar que el usuario tiene las credenciales correctas para acceder al servidor SSH.
- **Cifrado:** SSH es un protocolo cifrado, por lo que los datos son ininteligibles excepto para los destinatarios previstos.
- **Integridad:** SSH garantiza que los datos que viajan por la red lleguen inalterados. Si un tercero modificara el tráfico durante el tránsito, SSH lo detectaría.

#### **3.2.4.2. Claves públicas/privadas SSH**

De forma predeterminada, SSH se ejecuta en el puerto 22 y es común abrir el puerto a todo Internet. Esto significa que es un objetivo común para que los piratas informáticos prueben SSH. Si está utilizando una contraseña SSH, habrá muchos ataques de fuerza bruta. Es por eso por lo que se recomienda usar `ssh-keygen` para crear claves de autenticación sólidas para SSH. Obtenga más información sobre cómo configurar claves SSH, y es posible que desee considerar agregar un segundo factor a la autenticación SSH, en caso de que alguien obtenga sus claves. Si se implementa SSH sin contraseña, es importante comprender los pros y los contras de las opciones disponibles.

#### **3.2.3.3. Autenticación MQTT**

La autenticación es una parte importante de la mayoría de las aplicaciones. El protocolo MQTT admite la autenticación de nombre de usuario/contraseña. Habilitar la autenticación puede prevenir de manera efectiva las conexiones ilegales de clientes.

La autenticación en EMQX Broker significa que cuando un cliente se conecta a EMQX Broker, la configuración del servidor se usa para controlar el permiso del cliente para conectarse al servidor.

El soporte de autenticación de EMQX Broker incluye dos niveles:

El protocolo MQTT especifica el nombre de usuario y la contraseña en el paquete `CONNECT` por sí mismo. EMQX Broker admite múltiples formas de autenticación basadas en nombre de usuario, ID de cliente, HTTP, JWT, LDAP y varias bases de datos como MongoDB, MySQL, PostgreSQL, Redis a través de complementos.

En la capa de transporte, TLS garantiza la autenticación de cliente a servidor mediante certificados de cliente y garantiza que el servidor verifique el certificado de servidor para el cliente. También se admite la autenticación TLS/DTLS basada en PSK.

Estos métodos de autenticación compatibles con EMQX y los métodos de configuración de los complementos correspondientes se presentan en este artículo.

EMQX admite el uso de fuentes de datos integradas (archivos, bases de datos integradas),JWT, bases de datos convencionales externas y API HTTP personalizadas como fuentes de datos de autenticación.

La conexión de la fuente de datos y la lógica de autenticación se implementan a través de complementos. Cada complemento corresponde a un método de autenticación, y el complemento correspondiente debe habilitarse antes de su uso.

Cuando el cliente se conecta, el complemento implementa la autenticación de identidad del cliente al verificar si su nombre de usuario/cliente y contraseña son consistentes con la información de la fuente de datos especificada.

- Fuente de datos integrada

Autenticación de Mnesia (nombre de usuario/id de cliente). El archivo de configuración y la base de datos integrada de EMQX se utilizan para proporcionar una fuente de datos autenticada, que se administra a través de la API HTTP y es simple y liviana.

- Base de datos externa

- Autenticación LDAP.
- Autenticación MySQL.
- Autenticación PostgreSQL.
- Autenticación Redis.
- Autenticación MongoDB.
- La base de datos externa puede almacenar una gran cantidad de datos, al tiempo que facilita la integración con sistemas de gestión de dispositivos externos.

- Otros

- Autenticación HTTP.
- Autenticación JWT.

- La autenticación JWT puede emitir información de autenticación en lotes y la autenticación HTTP puede implementar una lógica de autenticación compleja.

*Autenticación Mnesia.:* La autenticación Mnesia utiliza la base de datos Mnesia integrada de EMQX para almacenar el ID/nombre de usuario y la contraseña del cliente, y admite la administración de datos de autenticación a través de la API HTTP. La autenticación de Mnesia no depende de fuentes de datos externas y es simple y liviana de usar.

#### **3.2.3.4.SSL en proxy inverso**

SSL funciona al garantizar que los datos transferidos entre usuarios y sitios web, o entre dos sistemas, sigan siendo imposibles de leer. Utiliza algoritmos de cifrado para codificarlos datos en tránsito, lo que evita que los piratas informáticos los lean cuando se envían a través de la conexión. Estos datos incluyen información potencialmente confidencial, como nombres, direcciones, números de tarjetas de crédito u otros detalles financieros. El proceso funciona así:

- Un navegador o servidor intenta conectarse a un sitio web (es decir, un servidor web) protegido con SSL.
- El navegador o servidor solicita que el servidor web se identifique.
- El servidor web envía al navegador o servidor una copia de su certificado SSL en respuesta.
- El navegador o el servidor comprueba si confía en el certificado SSL. Si lo hace, se lo indica al servidor web.
- Luego, el servidor web devuelve un reconocimiento firmado digitalmente para iniciar una sesión encriptada SSL.
- Los datos cifrados se comparten entre el navegador o el servidor y el servidor web.
- Este proceso a veces se denomina "apretón de manos SSL". Si bien parece un proceso largo, se lleva a cabo en milisegundos.

Cuando un sitio web está protegido por un certificado SSL, el acrónimo HTTPS (que significa HyperText Transfer Protocol Secure) aparece en la URL. Sin un certificado SSL, solo aparecerán las letras HTTP, es decir, sin la S de Seguro. También se mostrará un icono de candado en la barra de direcciones URL. Esto indica confianza y brinda tranquilidad a quienes visitan el sitio web.

### 3.2.5. Propiedades de un DT

Las propiedades del DT son lo que lo convierte en algo más que una representación digital.

#### 3.2.5.1. Propiedades necesarias

Estas son las propiedades inherentes a cualquier gemelo digital:

- a) Conexión en tiempo real con la entidad física.
- b) Autoevolución: una propiedad muy importante que no se ha explorado mucho. Con esto, DT puede aprender y adaptarse en tiempo real, proporcionando retroalimentación tanto al activo físico (a través del activo humano) como al DT. Esto se puede aprovechar fácilmente ahora debido al surgimiento de las herramientas de aprendizaje automático: para remodelar y rediseñar (como el aprendizaje por refuerzo). La frecuencia de esta sincronización depende de los escenarios de actualización, como, por ejemplo, basado en eventos (cadena de suministro), intervalos periódicos (aviones), basado en condiciones (logística), etc.
- c) Análisis continuo de aprendizaje automático (dependiendo de la frecuencia de la sincronización).
- d) Disponibilidad de datos de series de tiempo (o tiempo continuo) para monitoreo y como entrada para el sistema de aprendizaje automático.
- e) Dependencia de dominio (o servicios específicos de dominio): De acuerdo con el dominio, un DT puede brindar o priorizar servicios específicos de la industria. Estos son los mismos servicios 'específicos del dominio' que existen en el activo físico (por ejemplo, los problemas de optimización para un avión y una unidad de fabricación priorizarán o agregarán más peso a diferentes parámetros).

#### 3.2.5.2. Propiedades dinámicas

Sobre la base de estas propiedades dinámicas, se puede crear una jerarquía de gemelos digitales:

- a) Autonomía: un DT (o, para el caso, cualquier proveedor de información) podría realizar cambios en el activo físico en sí mismo, o un humano en control podría realizar cambios en el DT. Esto se aplica de manera diferente a las diferentes jerarquías de componentes presentes en el gemelo, como

algunas partes del sistema de aprendizaje automático o alguna parte del sistema de toma de decisiones. Por lo tanto, la propiedad de un DT de ser autónomo, no autónomo o parcialmente autónomo depende del caso. Esta clasificación también incluye el mecanismo de auto evolución del DT (qué cambios debe hacer en sí mismo y qué cambios debe aprobar un DT humano).

- b) Sincronización: La sincronización de datos puede ser continua o en ciertos intervalos de tiempo. Estos dependen de un número de factores como la tecnología, los recursos disponibles, la necesidad de los datos y el tipo de algoritmo de aprendizaje automático que se utiliza. Un DT podría tener subcomponentes que podrían sincronizarse en parte continuamente y en parte sincronizarse en función de eventos.

Esta sincronización puede resultar en diferentes jerarquías basadas en lo siguiente:

- ¿Con qué frecuencia se recopilan los datos?
- ¿Con qué frecuencia se almacenan los datos?
- ¿Con qué frecuencia se actualiza el DT?

Respondiendo al debate sobre si un DT debe ser un producto o un ciclo de vida del producto, después de analizar la literatura, nuestra opinión es que un DT puede ser, por ejemplo, tanto el DT de un automóvil como el DT del ciclo de vida de producción del automóvil. Esto se deriva de los componentes de DT definidos en la Sección 3.2; siempre que todos los componentes estén presentes, cualquier activo físico podría tener su DT.

### **3.3. Relación entre el flujo de oxígeno y la pureza**

Según [98] llegaron a la conclusión de que para afectar el aumento de la presión en el paso de adsorción como resultado de los aumentos de la tasa de flujo de alimentación (y la tasa de flujo de producto constante), el aumento de la presión de producción en más de 3 bar provoca una disminución tanto de la pureza como de la recuperación de oxígeno. En otro estudio de caso, Mendes et al. [26], el rendimiento de la unidad de PSA se estudió mediante experimentos y simulaciones. Informaron que, para afectar la tasa de flujo de purga en la pureza y recuperación de oxígeno, el aumento de la tasa de flujo de purga hará que disminuya la recuperación de oxígeno mientras que aumentará su pureza. La mejora en la automatización del llenado de balones de oxígeno viene puesta dado que, al tener una pureza correcta más estable a lo largo

del proceso, se evitarán tiempos muertos en la producción de oxígeno, he aquí la importancia de mantener este parámetro lo más estable posible.



## CAPÍTULO IV IMPLEMENTACIÓN DEL DIGITAL TWIN

### 4.1. Activo físico

#### 4.1.1. Plano de conexionado de los transmisores

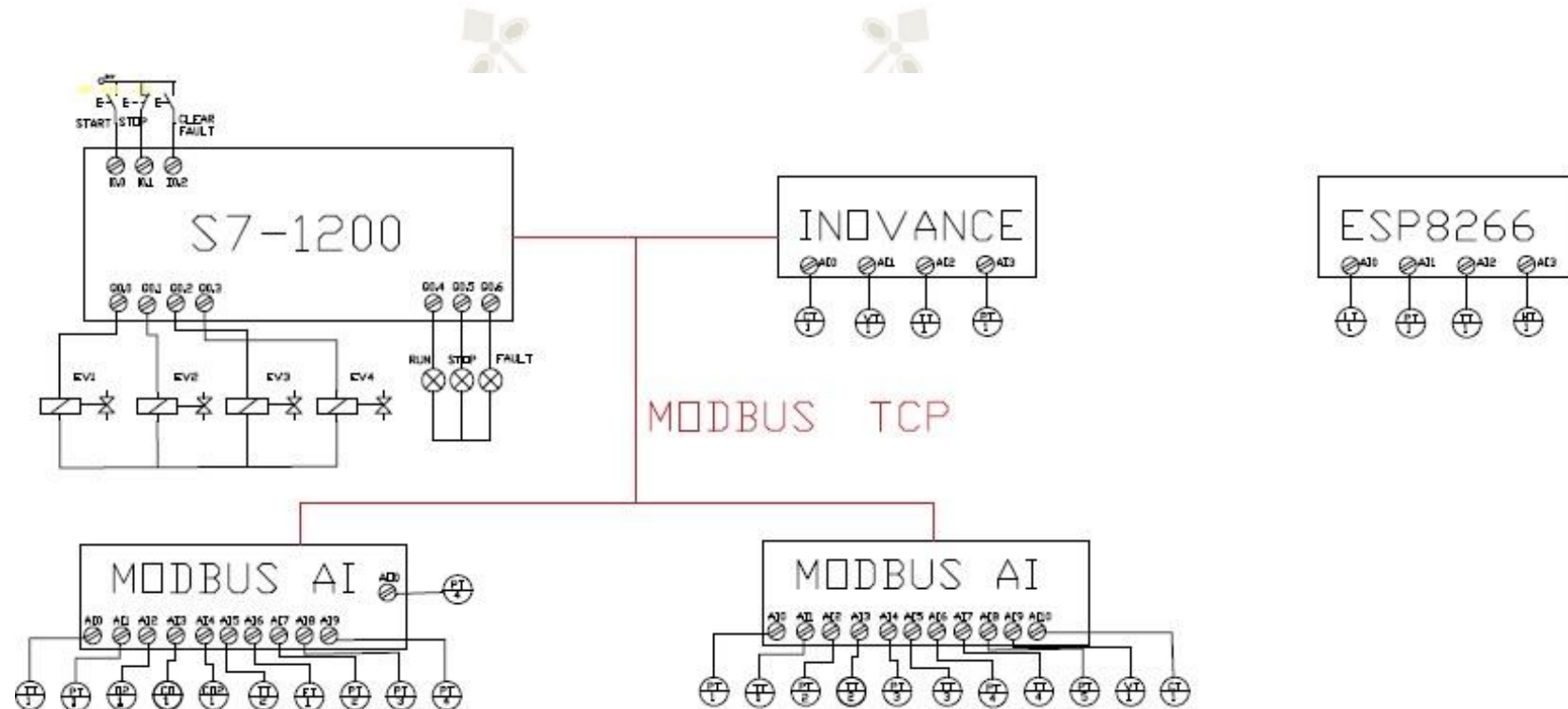


Figura 13 Diagrama Esquemático de adquisición de datos

Nota: Fuente: Elaboración propia.

#### 4.1.2. Breve explicación de la programación

Para la extracción de datos mediante el módulo analógico al PLC y su posterior publicacional broker MQTT se hará en el siguiente orden:

1. Configuración del cliente Modbus TCP: Esto se hará mediante el bloque Modbus TCP Client (figura 13) al cual se le debe programar los siguientes parámetros:
  - Connect: 1, debido a que la conexión siempre estará en marcha mientras el PLC este en estado RUN.
  - ConnectID: 1, identificación del dispositivo.
  - IP Octect 1-4: dirección IP del cliente Modbus para la extracción de datos.
  - MB\_MODE: 1 o 0 dependiendo si queremos leer data 1 o escribir 0.
  - ADDR: Dirección de memoria a la cual apuntar.
  - MB\_DATA\_LEN: Cantidad de bytes a leer desde la dirección de lectura.
  - MB\_DATA\_PTR: Registro en donde se guardarán los datos obtenidos. Posteriormente se tendrán las siguientes salidas:
    - DONE: Extracción de datos terminada.
    - BUSY: Operación en proceso.
    - ERROR: Error en la operación.
    - STATUS: Estado actual del bloque.
2. Extracción de datos obtenidos mediante cliente MQTT: Posteriormente a la extracción de datos del cliente Modbus se debe asignar estas a una cierta variable, por ejemplo, en la figura 14 se reasignan ciertas variables (temperatura, corriente y voltaje) a otros espacios de memoria, además de acomodar el formato, estando estas listas para convertirlas en datos de tipo STRING.
  - Para el envío de mensajes MQTT como se repasó previamente, estos deben ser de tipo STRING, por lo que necesitamos convertir los valores previamente obtenidos mediante el protocolo Modbus a una cadena de

caracteres, el encargado de realizar esta labor es la función VAL\_STRG, figura 15, cuyas entradas son:

- IN: Es aquí donde se coloca el valor a convertir.
- SIZE: El tamaño de la cadena de caracteres, es decir el espacio en memoria que se necesitará.
- PREC: La precisión decimal con la cual se hará la conversión.
- FORMAT: El formato con el que se hará la conversión.
- P: El bit de inicio para la conversión. Y salida:
- OUT: El valor ingresado en IN pero en formato de cadena de caracteres.

3. Finalmente se hace el envío de la cadena de caracteres mediante la función PubMQTT, figura 15, cuyas entradas son:

- Topic: El tópico en el cual se publicará el mensaje.
- Publish: El comando para realizar la publicación.
- Temperature/current: Estos parámetros son los datos para enviar, varían según la necesidad.

También para la publicación de este resultado es necesario datos almacenados en un DB con datos del broker MQTT (figura 17).

#### **4.1.3. Programación LADDER para la recolección de data mediante cliente Modbus**

La siguiente figura nos muestra la conexión al cliente Modbus mediante librerías del software TIA Portal v17.

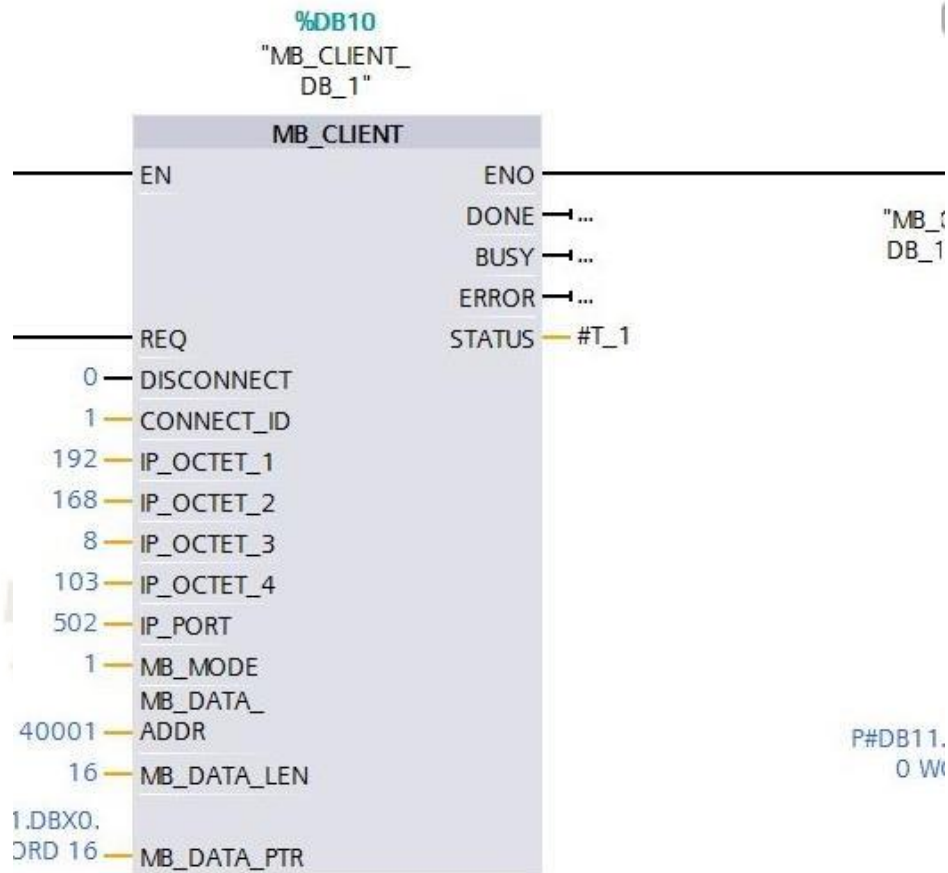


Figura 14 Cliente Modbus

Nota: Fuente: Elaboración propia.

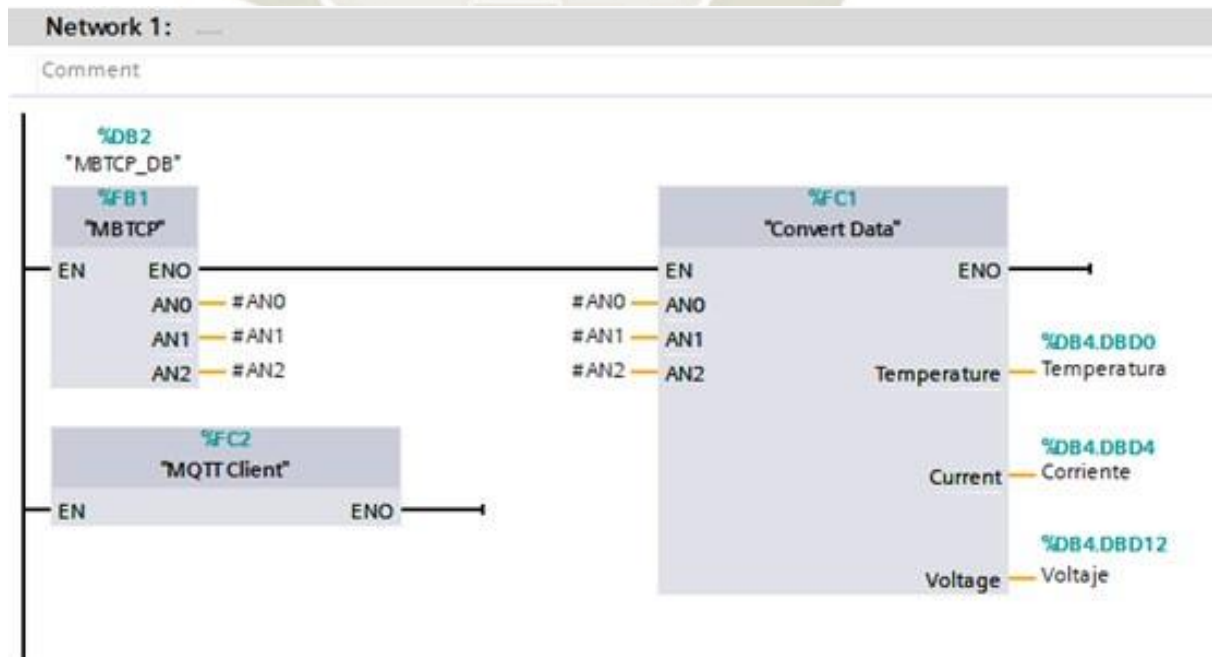
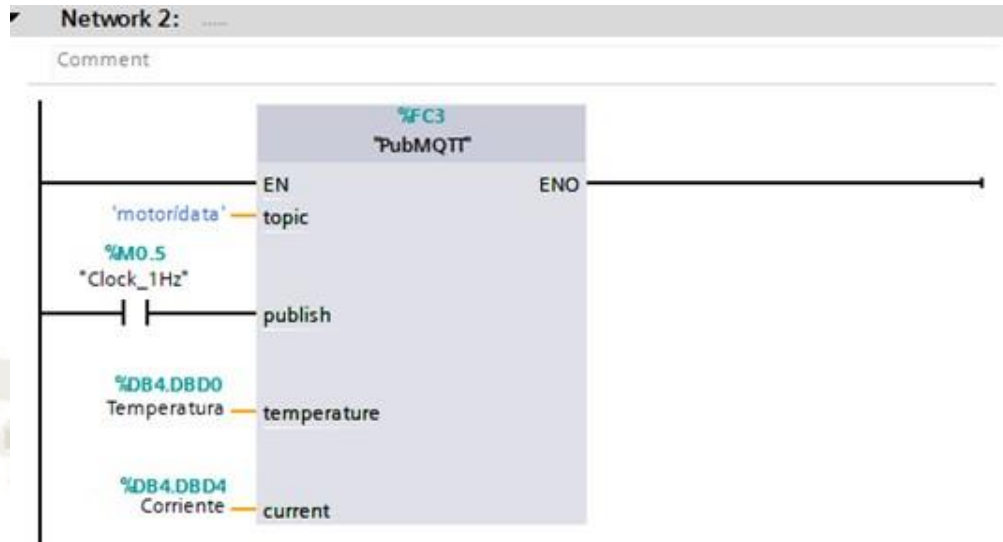


Figura 15 Bloque de reasignación de variable

Nota: Fuente: Elaboración propia.

#### 4.1.4. Programación LADDER para publicación MQTT

La siguiente figura nos muestra la conexión al cliente MQTT mediante librerías del software TIA Portal v17.



*Figura 16 Publicador MQTT*

*Nota:* Fuente: Elaboración propia.

typeUserData					
	Name	Data type	Default value	Accessible from HMI	Visible in HMI
1	realValue	Real	0.0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
2	size	USInt	0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
3	prec	USInt	0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
4	formWord	Word	16#0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
5	startP	UInt	0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
6	stringValue	String	"	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

DB					
	Name	Data type	Start value	Retain	Accessible from HMI
1	Static			<input type="checkbox"/>	<input type="checkbox"/>
2	convert	"typeUserData"		<input type="checkbox"/>	<input checked="" type="checkbox"/>
3	realValue	Real	0.0	<input type="checkbox"/>	<input checked="" type="checkbox"/>
4	size	USInt	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>
5	prec	USInt	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>
6	formWord	Word	16#0	<input type="checkbox"/>	<input checked="" type="checkbox"/>
7	startP	UInt	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>
8	stringValue	String	"	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Figura 17 Funcion reasignacion de variable

Nota: Fuente: Elaboración propia

	Name	Data type	Start value	Retain	Accessible f...	Writa...	Visible in ...	Setpoint	Comment
1	Static			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2	control	Struct		<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
3	output	Struct		<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
4	connparams	"LMQTT_typeConn...		<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Specifies connection parameters for the mqtt c
5	clientid	WString[20]	WSTRING#'Siem...	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
6	username	WString[20]	WSTRING#'emqx'	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
7	password	WString[20]	WSTRING#'publ...	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
8	willtopic	WString[20]	WSTRING#'willt...	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
9	willMessage	Array[0..99] of Byte		<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
10	willMessageCnt	UDInt	10	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
11	qos	USInt	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
12	topic	WString[100]	WSTRING#'a/b'	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
13	message	Array[0..999] of Byte		<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
14	pubMessageCnt	UDInt	11	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
15	retain	Bool	false	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
16	receivedTopic	WString[200]	WSTRING#''	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
17	receivedmessage	Array[0..999] of Byte		<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
18	receivedMesLen	UDInt	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
19	receivedMsgStatus	USInt	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	

Figura 18 Base de datos para comunicación MQTT

Nota: Fuente: Elaboración propia.

#### 4.1.5. Programación MQTT en ESP8266

Las siguientes ilustraciones muestran las variables y funciones necesarias para la creación de un cliente MQTT y la publicación de los datos obtenidos.

```
43 //MQTT
44 const char* mqtt_server = "10.0.0.60";
45 const int mqttPort = 1893;
46 const char* mqttUser = "emqx";
47 const char* mqttPassword = "public";
```

Figura 19 Variables para MQTT

Nota: Fuente: Elaboración propia.

```
void callback(String topic, byte* message, unsigned int length) {
  Serial.print("Message arrived on topic: ");
  Serial.print(topic);
  Serial.print(". Message: ");
  String messageTemp;

  for (int i = 0; i < length; i++) {
    Serial.print((char)message[i]);
    messageTemp += (char)message[i];
  }
  Serial.println();

  if(topic=="ESP8266-4/Bomba"){
    Serial.print("Bomba ");
    if(messageTemp == "on"){
      digitalWrite(RELEPIN, HIGH);
      Serial.print("On");
    }
    else if(messageTemp == "off"){
      digitalWrite(RELEPIN, LOW);
      Serial.print("Off");
    }
  }
  Serial.println();
}
```

Figura 20 Funciones para comunicación MQTT

Nota: Fuente: Elaboración propia.

```
//client.publish("ESP8266-4/sensores", msg);  
if (client.publish("ESP8266-4/sensores", msg) == true) {  
  Serial.println("Mensaje MQTT enviado");  
  digitalWrite(LED_BUILTIN, LOW);  
  delay(100);  
  digitalWrite(LED_BUILTIN, HIGH);  
} else {  
  Serial.println("Mensaje MQTT NO enviado");  
}
```

Figura 21 Funciones de inicialización MQTT

Nota: Fuente: Elaboración propia.

```
void reconnect() {  
  while (!client.connected()) {  
    Serial.print("Attempting MQTT connection..");  
    if (client.connect("ESP8266-4", mqttUser, mqttPassword )) {  
      Serial.println("connected");  
      client.subscribe("ESP8266-4/Bomba");  
      client.subscribe("room/lamp");  
    } else {  
      Serial.print("failed, rc=");  
      Serial.print(client.state());  
      Serial.println(" try again in 5 seconds");  
      delay(5000);  
    }  
  }  
}
```

Figura 22 Función para reconectarse al bróker

Nota: Fuente: Elaboración propia.

#### 4.2. Activo digital

Para la implementación del activo digital se eligió una Raspberry Pi de 8 GB de RAM como se mencionó previamente, se eligió la mayor capacidad de RAM para futuras implementaciones y asíno quedar sin recursos.

Adicionalmente se escogió la fuente de alimentación propia del fabricante para que cumpla los requerimientos de voltaje y corriente para el correcto funcionamiento del sistema.

Por último, para el almacenamiento se eligió un SSD portable, esto debido a la mayor de velocidad tanto de escritura como de lectura que tiene en comparación con el almacenamiento clásico en base a una tarjeta microSD.

#### 4.2.1. Monitoreo por influxdata DB

Todos los datos de los sensores tienen una marca de tiempo para ayudarlo a comprender cómo funcionan sus procesos y equipos a lo largo del tiempo para ayudar a una mayor innovación y mejora.

El mundo industrial tiene una larga historia de procesos de modernización para mantener la producción en funcionamiento de manera eficiente y segura mientras se minimiza el tiempo de inactividad. Sin embargo, muchos están encerrados en soluciones de historial de datos establecidas que son costosas y carecen de los métodos necesarios para proporcionar innovación e interoperabilidad.

Por el contrario, InfluxDB, la base de datos de series temporales de código abierto proporciona de forma inherente diversas perspectivas de diseño que no están disponibles en un único proveedor de software. Brinda la libertad de integrarse con otras soluciones y le permite adaptar el código para que se ajuste a los requisitos de su sistema en constante cambio.

##### Escalabilidad

La recopilación de datos de eventos de su equipo es solo el comienzo. La verdadera transformación digital requiere más fuentes de datos y más análisis de los datos combinados para obtener una mejor comprensión de sus sistemas. InfluxDB es un almacén de datos de alto rendimiento escrito específicamente para datos de series temporales. Permite ingesta de alto rendimiento, compresión y consultas en tiempo real.

##### Abierto

Con la disponibilidad de tantas herramientas de código abierto, ya no es necesario que los operadores compren soluciones misteriosas de código cerrado. En su lugar, pueden buscar la creación de sus propios reemplazos de historiadores de datos o comprar una solución lista para usar basada en código abierto. Esto le brinda al operador la libertad de innovar rápidamente y nunca estar encerrado en una sola solución que podría volverse obsoleta fácil y rápidamente.

##### Construido para desarrolladores

InfluxDB proporciona capacidades de escritura y consulta con una interfaz de línea de comandos, una API HTTP integrada, un conjunto de bibliotecas de clientes (por ejemplo, Go, Python y JavaScript) y complementos de Telegraf para formatos de datos comunes como OPC-UA, ModBus, MQTT y más.

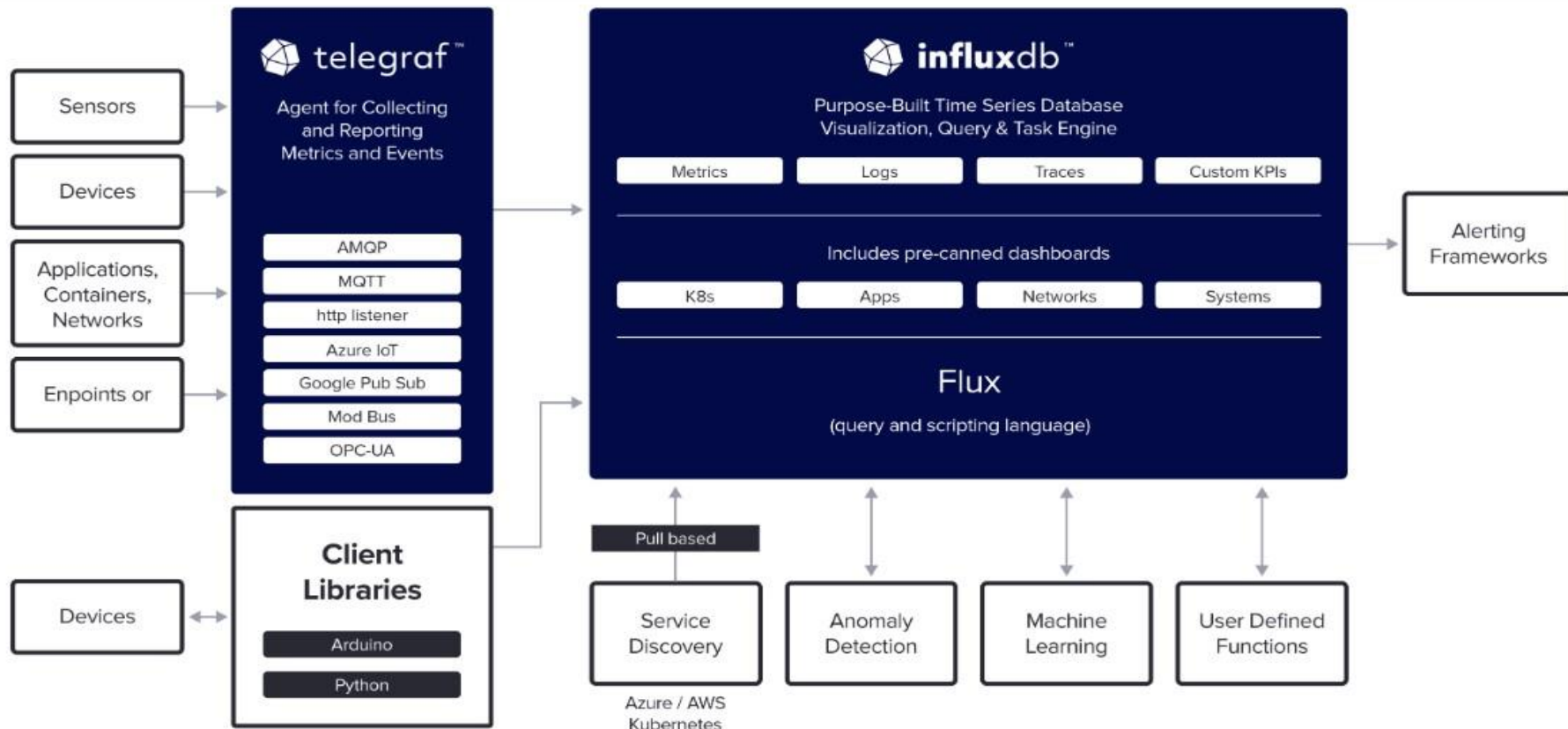


Figura 23 Arquitectura de la plataforma Influxdata a implementar

Nota: Fuente: Elaboración propia.

#### 4.2.2. Componentes de la plataforma influxDB

InfluxDB es fundamental para muchas soluciones de historial de datos que brindan ingestade alto rendimiento, compresión y consulta en tiempo real de esos mismos datos. La eficiencia y la eficacia deben comenzar en la estructura de datos, asegurando que los valores con marca de tiempo se recopilen con la precisión y los metadatos necesarios paraproporcionar flexibilidad y velocidad en la creación de gráficos, consultas y alertas.

El nombre de la medida es una cadena, el conjunto de etiquetas es una colección de pares clave/valor donde todos los valores son cadenas y el conjunto de campos es una colecciónde pares clave/valor donde los valores pueden ser int64, float64, bool o cadena. La compatibilidad con la codificación de datos más allá de los valores float64 significa que los metadatos se pueden recopilar junto con la serie temporal y no se limitan solo a valoresnuméricos (por ejemplo, la válvula está abierta o cerrada). Además, no hay límites estrictosen el número de etiquetas y campos.

Tener varios campos y etiquetas bajo la misma medida optimiza la transmisión de datos, lo cual es importante para los dispositivos remotos que envían métricas.

El nombre de la medición y los conjuntos de etiquetas se mantienen en un índice invertidoque hace que las búsquedas de series específicas sean muy rápidas.

Telegraf es el agente recopilador que cuenta con más de 200 complementos, la mayoría proporcionados por la comunidad de código abierto. Está escrito en Go, se compila en un solo binario sin dependencias externas y requiere una huella de memoria mínima. La variedad de complementos creados para Telegraf incluye la recopilación de datos de:

Sensores: recopile datos de estado críticos (niveles de presión, niveles de temperatura,etc.) de sensores y dispositivos IoT.

Bases de datos: conéctese a fuentes de datos como MongoDB, MySQL, Redis y otraspara recopilar y enviar métricas.

Sistemas: Recopile métricas de su pila moderna de plataformas en la nube, contenedoresy orquestadores.

InfluxDB también admite la capacidad de unir fuentes de datos temporales y no temporales

### 4.2.3. Raspberry Pi 4B+ y componentes adicionales



*Figura 24 Raspberry Pi 4B+*

*Nota: Fuente:* <https://www.raspberrypi.com/products/raspberry-pi-4-model-b/>



*Figura 25 Fuente de alimentación Raspberry Pi*

*Nota: Fuente:* <https://www.raspberrypi.com/products/raspberry-pi-4-model-b/>



*Figura 26 Disco SSD Portable*

*Nota: Fuente: <https://www.kingston.com/es/ssd/xs2000-portable-usb-c-solid-state-drive>*

#### **4.2.4. Instalación docker**

```
pi@raspberrypi: ~  
pi@raspberrypi:~ $ sudo bash get-docker.sh  
# Executing docker install script, commit: 26ff363bcf3b3f5a00498ac43694bf1c7d9ce16c  
+ sh -c 'apt-get update -qq >/dev/null'  
+ sh -c 'DEBIAN_FRONTEND=noninteractive apt-get install -y -qq apt-transport-https ca-certificates curl >/dev/null'  
+ sh -c 'curl -fsSL "https://download.docker.com/linux/raspbian/gpg" | apt-key add -qq - >/dev/null'  
Warning: apt-key output should not be parsed (stdout is not a terminal)  
+ sh -c 'echo "deb [arch=armhf] https://download.docker.com/linux/raspbian buster stable" > /etc/apt/sources.list.d/docker.list'  
+ sh -c 'apt-get update -qq >/dev/null'
```

*Figura 27 Instalación de Docker mediante comandos*

*Nota: Fuente: Elaboración propia.*

#### 4.2.5. Instalación de contenedores

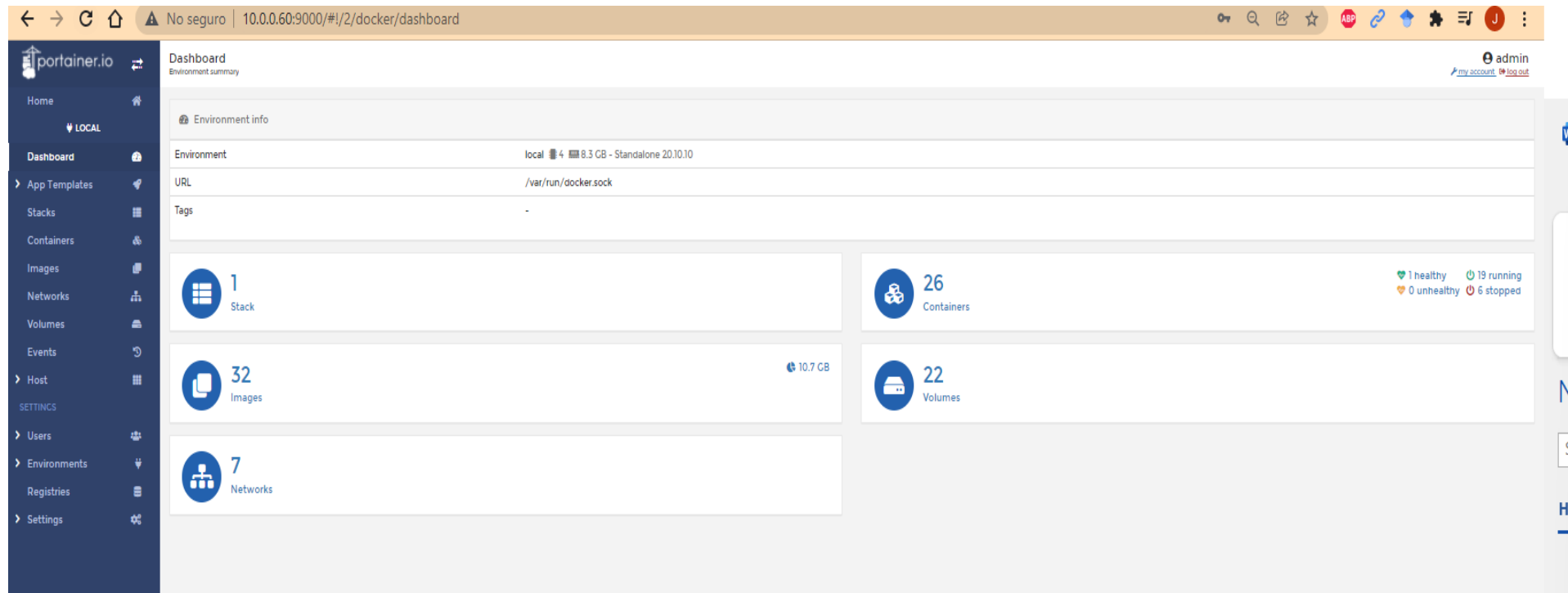
##### 4.2.5.1. Portainer

El siguiente código es parte del docker-compose para utilizar el contenedor Portainer, encargado de dar una interfaz más amigable al uso de contenedores Docker.

```
portainer:  
  container_name: portainer  
  image: portainer/portainer-ce:latest  
  restart: always  
  ports:  
  | - "9000:9000"  
  command: -H unix:///var/run/docker.sock  
  volumes:  
  | - /var/run/docker/portainer/var/run/docker.sock:/var/run/docker.sock  
  | - /portainer_data:/data  
  networks:  
  | - zigbee
```

*Figura 28 Sección del código docker compose correspondiente a Portainer*

*Nota:* Fuente: Elaboración propia.



*Figura 29 GUI Portainer*

*Nota:* Fuente: Elaboración propia

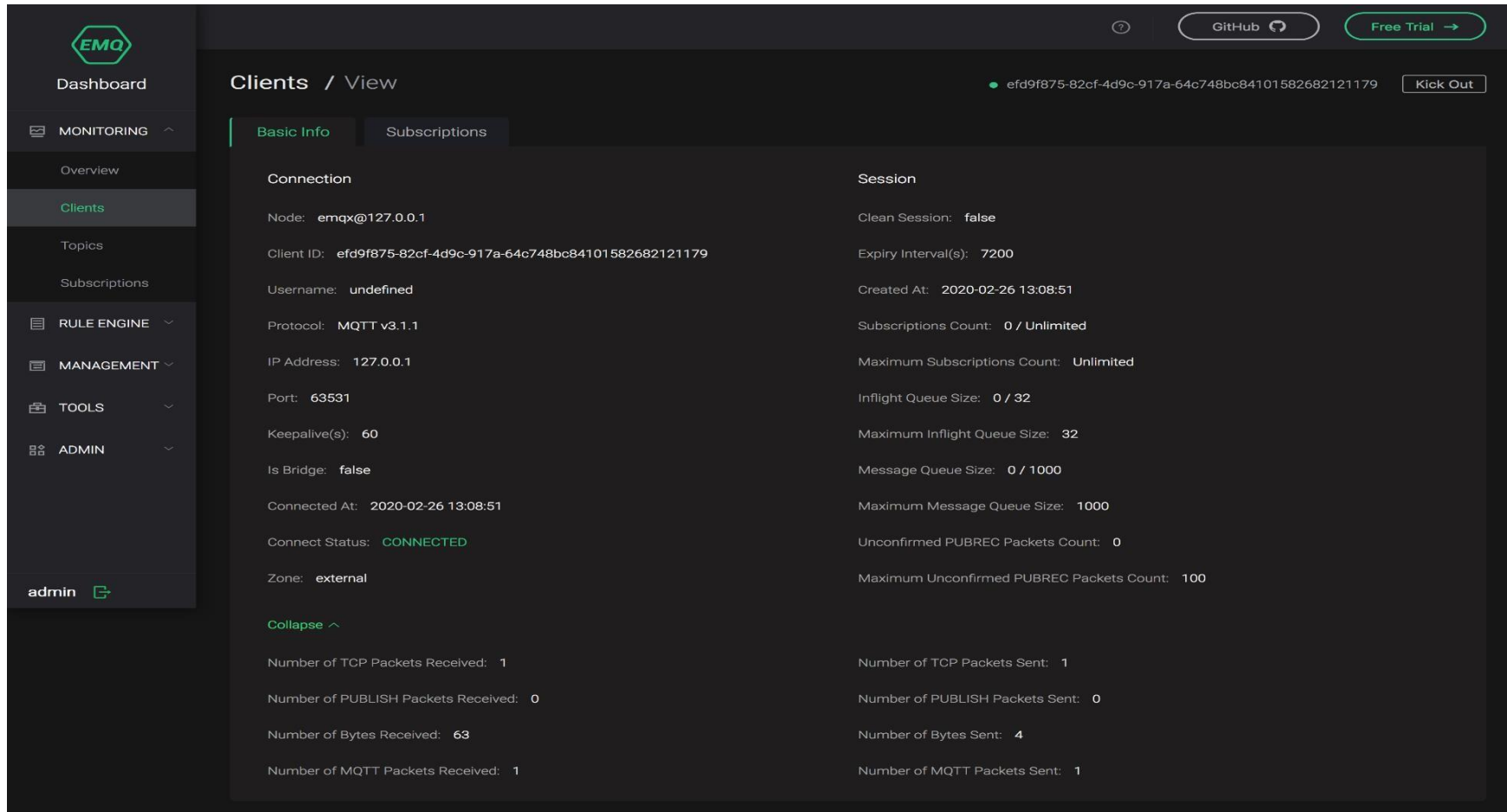
#### 4.2.5.2.EMQX

El siguiente código es parte del docker-compose para utilizar el contenedor EMQX, el broker MQTT en cargado de escuchar y reenviar mensajes a los clientes que escojan publicar o escuchar distintos tópicos.

```
emqx:
  container_name: emqx
  image: emqx/emqx:4.3.8
  restart: always
  environment:
    # - EMQX_NAME=emqx
    # - EMQX_LOG_LEVEL=debug
    - EMQX_ALLOW_ANONYMOUS=false
    # - EMQX_LISTENER_SSLEXTERNAL_KEYFILE=etc/certs/server.key #emqx.key
    # - EMQX_LISTENER_SSLEXTERNAL_CERTFILE=etc/certs/server.crt #emqx.crt
    # - EMQX_LISTENER_WSSEXTERNAL_KEYFILE=etc/certs/emqx.key
    # - EMQX_LISTENER_WSSEXTERNAL_CERTFILE=etc/certs/emqx.crt
    # - EMQX_LISTENER_SSL_EXTERNAL_CACERTFILE=etc/certs/ca.crt
    #- EMQX_LOADED_PLUGINS=emqx_auth_username,emqx_auth_http,emqx_web_hook,emqx_management,emqx
  ports:
    - 18083:18083 #Dashboard
    - 8093:8083 #MQTT/WS
    - 8094:8084 #MQTT/WSS
    - 1893:1883 #MQTT
    - 8893:8883 #MQTT/SSL
  #volumes:
    # - /home/pi/docker/emqx/data:/opt/emqx/data
    # - /home/pi/docker/emqx/etc:/opt/emqx/etc
    # - /home/pi/docker/emqx/log:/opt/emqx/log
```

Figura 30 Sección del código Docker compose correspondiente a EMQX

Nota: Fuente: Elaboración propia.



The screenshot displays the EMQX web interface. On the left is a navigation sidebar with sections: Dashboard, MONITORING (with sub-items Overview, Clients, Topics, Subscriptions), RULE ENGINE, MANAGEMENT, TOOLS, and ADMIN. The main content area is titled 'Clients / View' and shows details for a specific client with ID 'efd9f875-82cf-4d9c-917a-64c748bc84101582682121179'. The client is in a 'CONNECTED' state. The interface is divided into two columns: 'Connection' and 'Session'.

Connection	Session
Node: emqx@127.0.0.1	Clean Session: false
Client ID: efd9f875-82cf-4d9c-917a-64c748bc84101582682121179	Expiry Interval(s): 7200
Username: undefined	Created At: 2020-02-26 13:08:51
Protocol: MQTT v3.1.1	Subscriptions Count: 0 / Unlimited
IP Address: 127.0.0.1	Maximum Subscriptions Count: Unlimited
Port: 63531	Inflight Queue Size: 0 / 32
Keepalive(s): 60	Maximum Inflight Queue Size: 32
Is Bridge: false	Message Queue Size: 0 / 1000
Connected At: 2020-02-26 13:08:51	Maximum Message Queue Size: 1000
Connect Status: <b>CONNECTED</b>	Unconfirmed PUBREC Packets Count: 0
Zone: external	Maximum Unconfirmed PUBREC Packets Count: 100
<b>Collapse ^</b>	
Number of TCP Packets Received: 1	Number of TCP Packets Sent: 1
Number of PUBLISH Packets Received: 0	Number of PUBLISH Packets Sent: 0
Number of Bytes Received: 63	Number of Bytes Sent: 4
Number of MQTT Packets Received: 1	Number of MQTT Packets Sent: 1

*Figura 31 GUI EMQX*

*Nota:* Fuente: Elaboración propia

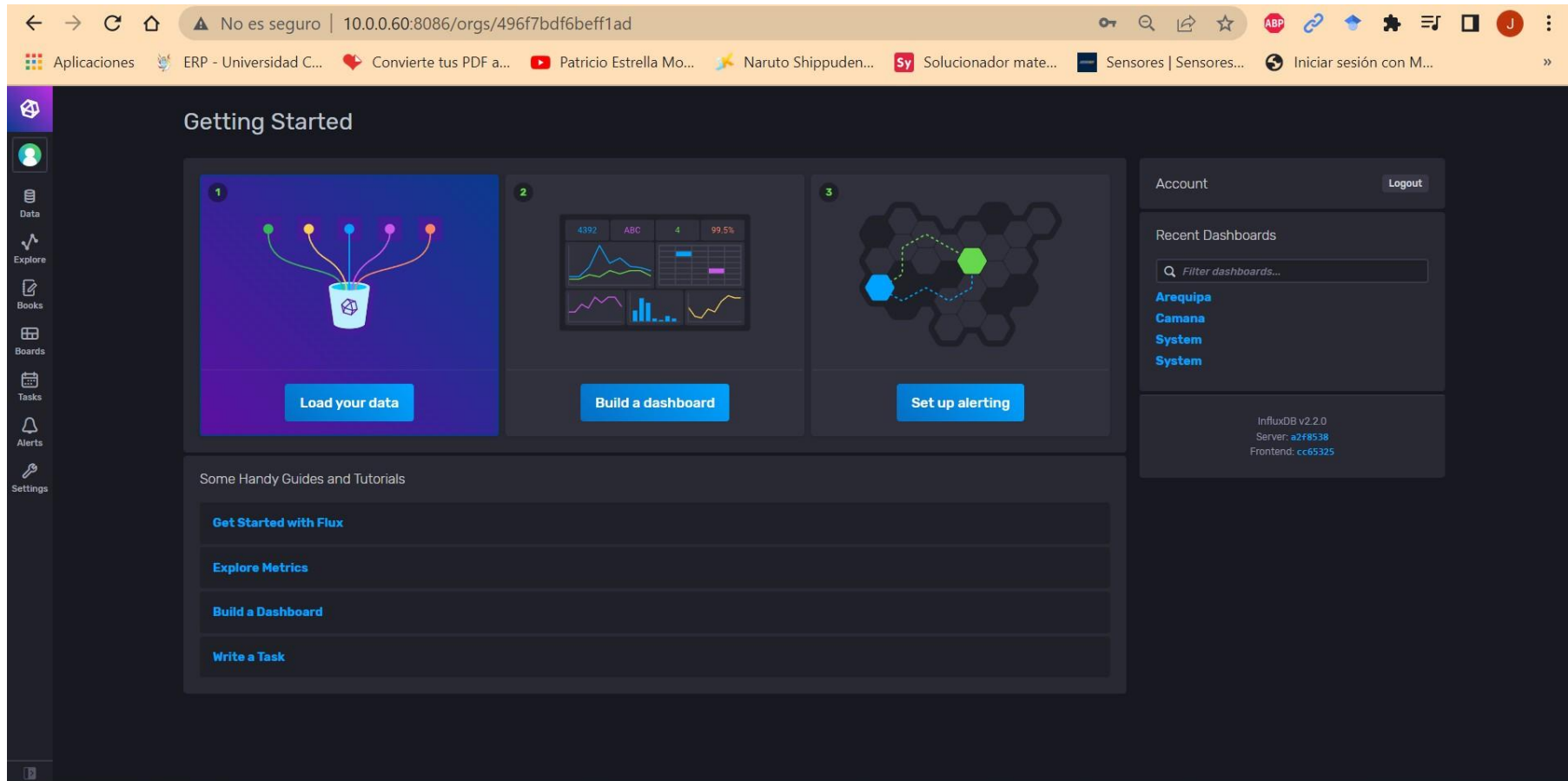
#### 4.2.5.3. Influxdb

El siguiente código es parte del docker-compose para utilizar el contenedor Influxdb, encargado de almacenar la información recibida mediante el servicio Telegraf, así como de elaborar los dashboards y generar alarmas para ciertas condiciones.

```
influxdb:
  container_name: influxdb
  image: influxdb:1.8 #quay.io/influxdb/influxdb:v2.0.4
  restart: always
  environment:
    - DOCKER_INFLUXDB_INIT_USERNAME=user
    - DOCKER_INFLUXDB_INIT_PASSWORD=user
  ports:
    - 8086:8086
    - 8083:8083
    - 2003:2003
  volumes:
    - /home/pi/docker/influxdb/var/lib/influxdb:/var/lib/influxdb
    - /home/pi/docker/influxdb/etc/influxdb:/etc/influxdb
  networks:
    - zigbee
```

*Figura 32 Sección del código Docker compose correspondiente a Influxdb*

*Nota: Fuente: Elaboración propia.*



*Figura 33 GUI Influxdb*

*Nota:* Fuente: Elaboración propia.

#### 4.2.5.3.1. Dashboards



*Figura 34 Dashboard compresor de tornillo*

*Nota:* Fuente: Elaboración propia.



Figura 35 Dashboard OGP

Nota: Fuente: Elaboración propia.



Figura 36 Dashboard Booster

Nota: Fuente: Elaboración propia.



Figura 37 Dashboard estado del sistema

Nota: Fuente: Elaboración propia.

4.2.5.1.1. Configuración de alarmas.



Figura 38 Revisión de alarmas programadas

Nota: Fuente: Elaboración propia.

En la figura 34, se puede apreciar el dashboard correspondiente al compresor de tornillo, En este dashboard correspondiente al compresor de tornillo se puede observar lecturas actuales (lado izquierdo) así como datos históricos en una gráfica de tipo XY (lado derecho).

En la figura 35, se puede apreciar el dashboard correspondiente al OGP en estas gráficas podemos observar el punto de rocío del secador de aire, presión del tanque de aire; así como el punto de rocío del generador de oxígeno, el flujo de consumo de este y finalmente la presión en el tanque de almacenamiento, se tienen lecturas en tiempo real acompañadas de un histórico en la parte trasera del plugin.

Finalmente, para concluir con la etapa de la planta, tenemos el dashboard del booster de oxígeno, en la figura 36; el cual nos muestra un registro histórico de las temperaturas de este, en distintos colores, y al lado derecho nos muestra la temperatura actual que se tiene en tiempo real.

Como adicional, en la figura 37 se tiene un dashboard creado para monitorear las variables informáticas del servidor, tenemos por ejemplo el tiempo que este encendido, la carga del sistema, el uso de memoria y la carga del sistema.

Finalmente en la figura 38 se pueden verificar las alarmas programadas para ciertos niveles de cualquier parámetro, dándonos aviso de la hora en que se dio la incidencia.

#### 4.2.5.4. Telegraf

El siguiente código es parte del docker-compose para utilizar el contenedor Telegraf, encargado de recibir la información mediante plugins para distintos protocolos de comunicación.

```
version: '3'
services:
  # Define a Telegraf service
  telegraf:
    container_name: telegraf
    image: telegraf
    volumes:
      - ./telegraf/telegraf.conf:/etc/telegraf/telegraf.conf:ro
    links:
      - influxdb
    ports:
      - "8092:8092/udp"
      - "8094:8094"
      - "8125:8125/udp"
```

Figura 39 Sección del código docker compose correspondiente a Telegraf

Nota: Fuente: Elaboración propia.

#### 4.2.5.4.1. Configuración

En las siguientes imágenes se puede observar la configuración de los plugins encargados de la comunicación con Influxdb y con el bróker MQTT.

```
[[outputs.influxdb_v2]]
## The URLs of the InfluxDB cluster nodes.
##
## Multiple URLs can be specified for a single cluster, only ONE of the
## urls will be written to each interval.
## ex: urls = ["https://us-west-2-1.aws.cloud2.influxdata.com"]
urls = ["http://10.0.0.60:8086"]

## API token for authentication.
token = "ZmP-tvs4kH2AqnrwGClxZFF0wPOn6jnPq8Fe1UtKjwBo0SBBfRiyaIaNLiej3comzqCL1EREDJcmXceZTlitwrQ=="

## Organization is the name of the organization you wish to write to; must exist.
organization = "Iotverse"

## Destination bucket to write into.
bucket = "telegraf"
```

Figura 40 Configuración Telegraf plugin Influxdb

Nota: Fuente: Elaboración propia.

```
docker-compose > ticktest > telegraf > telegraf.conf
486
487 [[inputs.mqtt_consumer]]
488 ## Broker URLs for the MQTT server or cluster. To connect to multiple
489 ## clusters or standalone servers, use a separate plugin instance.
490 ## example: servers = ["tcp://localhost:1883"]
491 ##          servers = ["ssl://localhost:1883"]
492 ##          servers = ["ws://localhost:1883"]
493 servers = ["tcp://10.0.0.60:1893"]
494
495 ## Topics that will be subscribed to.
496 topics = [
497     "telegraf/host01/cpu",
498     "telegraf+/mem",
499     "test/#",
500     "Arequipa/Compresor",
501     "Arequipa/OGP",
502     "Arequipa/Booster",
503     "Camana/Compresor",
504     "Camana/OGP",
505     "Camana/Booster",
506     "Mollendo/Compresor",
507     "Mollendo/OGP",
```

Figura 41 Configuración Telegraf plugin MQTT

Nota: Fuente: Elaboración propia.

#### 4.2.5.5. Node-Red

```
nodered:  
  container_name: nodered  
  image: nodered/node-red:latest  
  restart: always  
  ports:  
  | - 1881:1880  
  volumes:  
  | - /home/pi/docker/nodered/data:/data  
  | - /opt/vc:/opt/vc:ro  
# - nodered/node-red  
  devices:  
  | - /dev/vchiq:/dev/vchiq  
  environment:  
  | - LD_LIBRARY_PATH=/opt/vc/lib  
  networks:  
  | - zigbee
```

*Figura 42 Sección del código docker compose correspondiente a Nodered*

*Nota:* Fuente: Elaboración propia.

#### 4.2.5.6. Configuración Node-Red

```
nodered:  
  container_name: nodered  
  image: nodered/node-red:latest  
  restart: always  
  ports:  
  | - 1881:1880  
  volumes:  
  | - /home/pi/docker/nodered/data:/data  
  | - /opt/vc:/opt/vc:ro  
# - nodered/node-red  
  devices:  
  | - /dev/vchiq:/dev/vchiq  
  environment:  
  | - LD_LIBRARY_PATH=/opt/vc/lib  
  networks:  
  | - zigbee
```

*Figura 43 Sección del código docker compose correspondiente a Nodered*

*Nota:* Fuente: Elaboración propia.

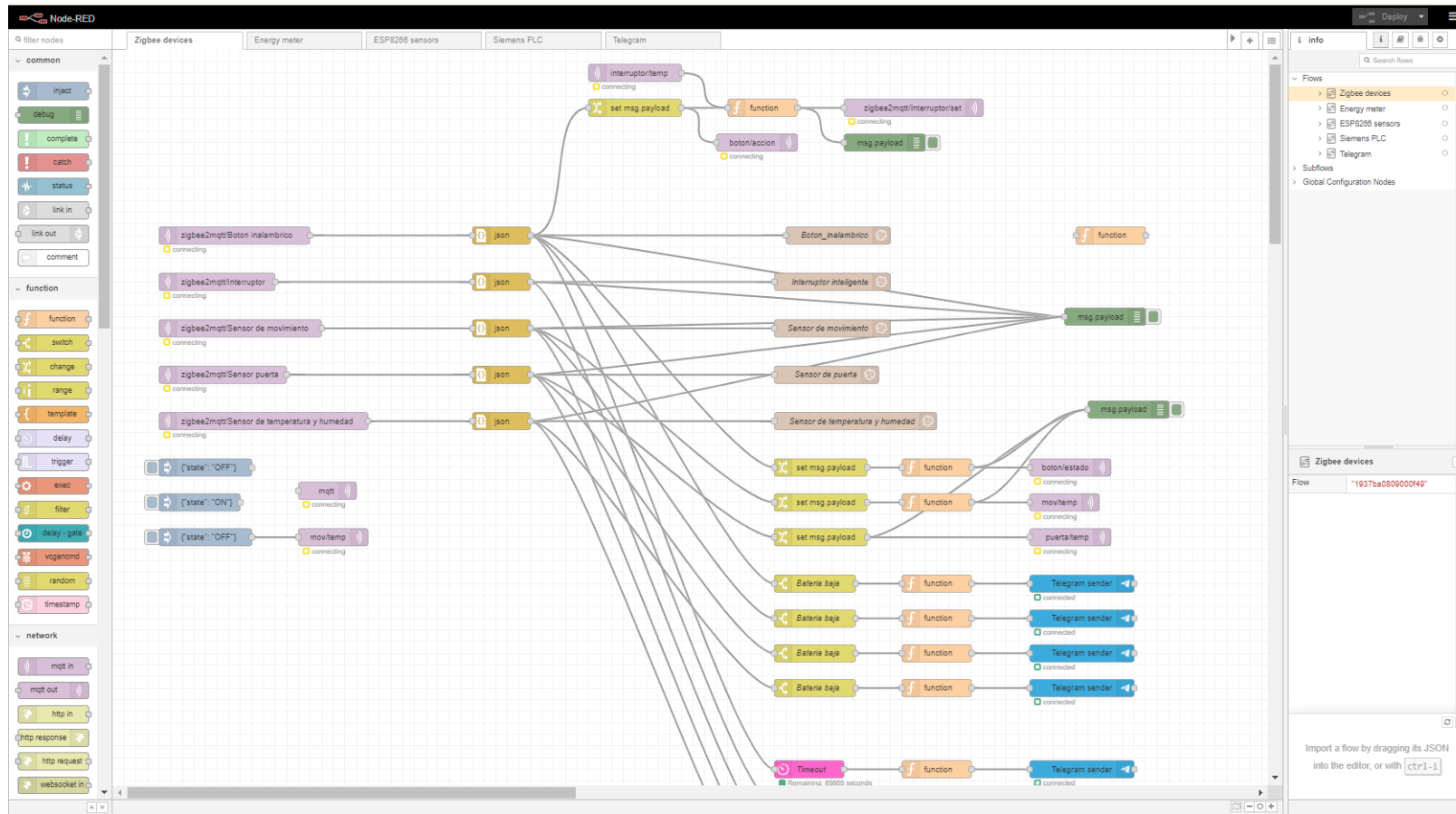


Figura 44 GUI Nodered

Nota: Fuente: Elaboración propia.

The screenshot shows the Node-RED web interface. On the left, a sidebar contains node categories: 'common' (inject, debug, complete, catch, status, link in, link call, link out, comment) and 'function' (function, switch, change, range, template, delay, trigger). The main workspace displays a flow named 'Flow 1' with 12 nodes: 8 'json' nodes and 1 'msg.payload' node. The 'json' nodes are connected to the 'msg.payload' node. The nodes are labeled as follows: Arequipa/Compresor, Arequipa/OGP, Arequipa/Booster, Camana/Compresor, Camana/OGP, Camana/Booster, Mollendo/Compresor, Mollendo/OGP, and Mollendo/Booster. The right-hand side features a 'debug' console showing a log of messages. The messages are JSON objects representing MQTT payloads, such as:
 

```

    {
      "C-Temperature": 38.41,
      "C-Pressure": 95.4,
      "C-Voltage": 223.32,
      "C-Current": 63.18,
      "S-Dewpoint": 1.72
    }
    
```

 and
 

```

    {
      "O2-Purity": 93.18,
      "O2-CO": 1,
      "O2-CO2": 123,
      "O2-Dewpoint": 4.1,
      "O2-Flow": 38.78
    }
    
```

Figura 45 Debugging de datos MQTT

Nota: Fuente: Elaboración propia.

En la figura 45, podemos observar que El contenedor Node-red tiene entre su uso el debugging de los mensajes MQTT que llegan a los tópicos que se subscriba, como se puede ver en el lado derecho de la imagen tenemos los mensajes que han llegado para así poder ser evaluados para la siguiente aplicación que estos tengan.

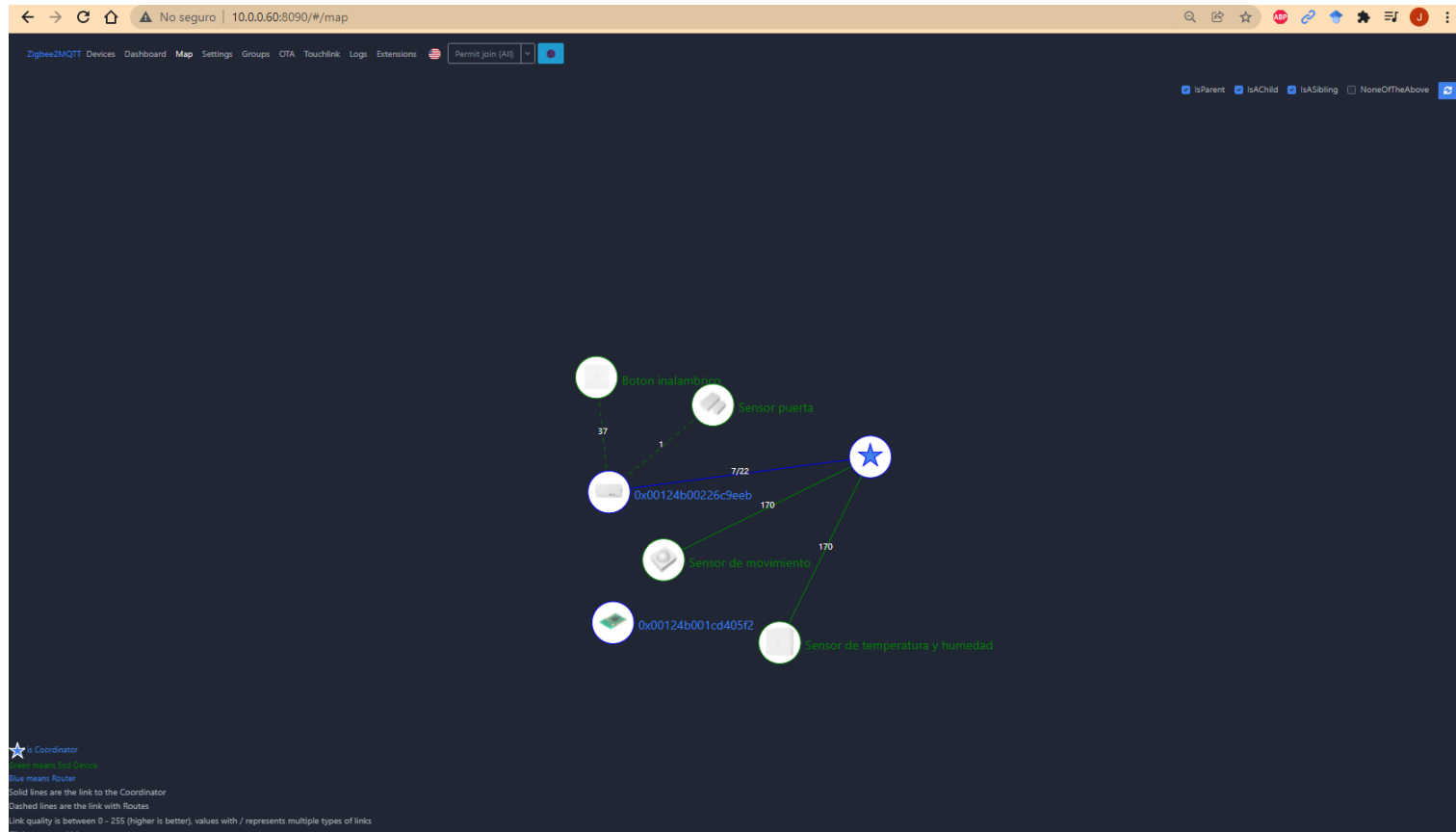
#### 4.2.5.7. Zigbee2MQTT.

El siguiente código es parte del docker-compose para utilizar el contenedor Zigbee2MQTT, encargado de recibir los datos del protocolo Zigbee y convertirlos en mensajes de tipo MQTT.

```
#zigbee2mqtt:
# image: koenkk/zigbee2mqtt:latest
# container_name: zigbee2mqtt
# user: 1000:1000
# group_add:
#   - dialout
# restart: always
# privileged: true # only required for autodetecting the port
# environment:
#   - TZ=America/Lima
# volumes:
#   - /home/pi/docker/zigbee2mqtt3:/app/data
#   - /run/udev:/run/udev:ro # only required for autodetecting the port
#
# devices:
#   - /dev/ttyUSB0:/dev/ttyACM0
# ports:
#   - 8090:8090
# networks:
#   - zigbee
```

Figura 46 Sección del código docker compose correspondiente a Zigbee2MQTT

Nota: Fuente: Elaboración propia.



*Figura 47 GUI Zigbee2MQTT Mapa de conexión*

*Nota:* Fuente: Elaboración propia.

#### 4.2.5.7.1. Configuración MQTT.

Para la conexión MQTT se debe cambiar la configuración e indicarle a que broker apuntar, adicionalmente se le debe asignar un canal Zigbee para la comunicación.

```
docker-compose > mqtt > zigbee2mqtt > data > ! configuration.yaml
1  homeassistant: false
2  permit_join: false
3  mqtt:
4    base_topic: zigbee2mqtt
5    server: mqtt://emqx
6    client_id: zigbee2mqtt
7    keepalive: 60
8    reject_unauthorized: true
9    version: 4
10   password: public
11   user: emqx
12  serial:
13    port: /dev/ttyUSB0
14  frontend:
15    port: 8080
16    auth_token: '!secret auth_token'
17  advanced:
18    pan_id: 52009
```

*Figura 48 Configuración MQTT en Zigbee2MQTT*

*Nota:* Fuente: Elaboración propia.

#### 4.2.5.7.2 Seguridad en la red.

```
docker-compose > mqtt > zigbee2mqtt > data > ! configuration.yaml
13  port: /dev/ttyUSB0
14  frontend:
15    port: 8080
16    auth_token: '!secret auth_token'
17  advanced:
18    pan_id: 52009
19    ext_pan_id:
20      - 221
21      - 221
22      - 221
23      - 221
24      - 221
25      - 221
26      - 221
27      - 221
28    channel: 25
```

*Figura 49 Asignación de Pan ID para la seguridad en la red Zigbee*

*Nota:* Fuente: Elaboración propia.

```
docker-compose > mqtt > zigbee2mqtt > data > ! secret.yaml
1  user: admin
2  password: adminroot
3  network_key: [130, 46, 49, 251, 168, 63, 192, 75, 159, 26, 252, 41, 17, 233, 220, 95]
4  auth_token: adminroot
5
```

*Figura 50 Clave de red Zigbee*

*Nota:* Fuente: Elaboración propia

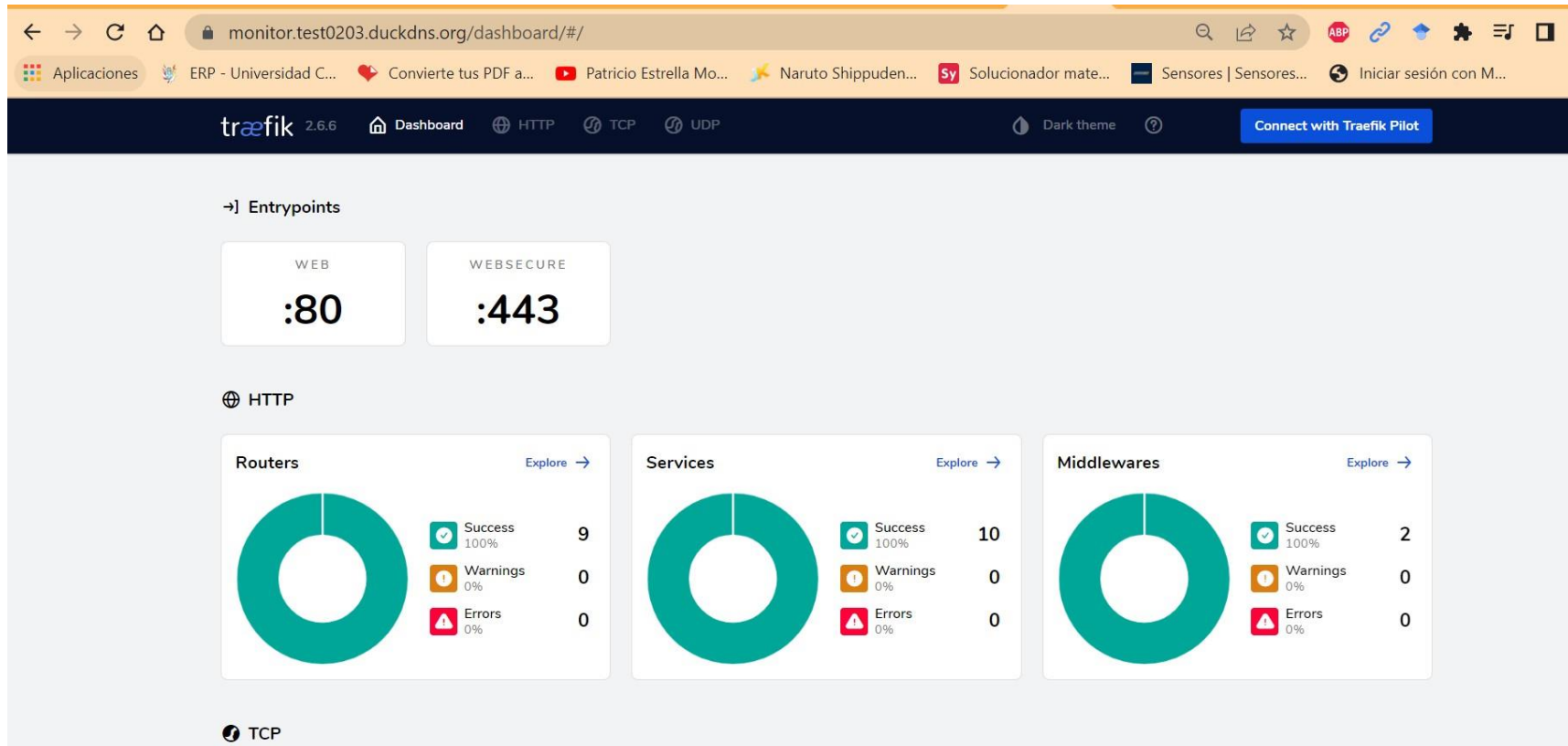
#### 4.2.5.8. Traefik.

El siguiente código es parte del docker-compose para utilizar el contenedor Traefik, encargado de actuar como balanceador de carga y proxy inverso.

```
96  traefik:
97    image: traefik:latest
98    container_name: traefik
99    #restart: always
100   volumes:
101     - ./traefik/traefik.toml:/traefik.toml
102     - ./traefik/traefik_dynamic.toml:/traefik_dynamic.toml
103     - ./traefik/acme.json:/acme.json
104     - /var/run/docker.sock:/var/run/docker.sock
105   ports:
106     - 81:80
107     - 443:443
108   networks:
109     - web
110
```

*Figura 51 Sección del código docker compose correspondiente a Traefik*

*Nota:* Fuente: Elaboración propia.



*Figura 52 GUI Traefik*

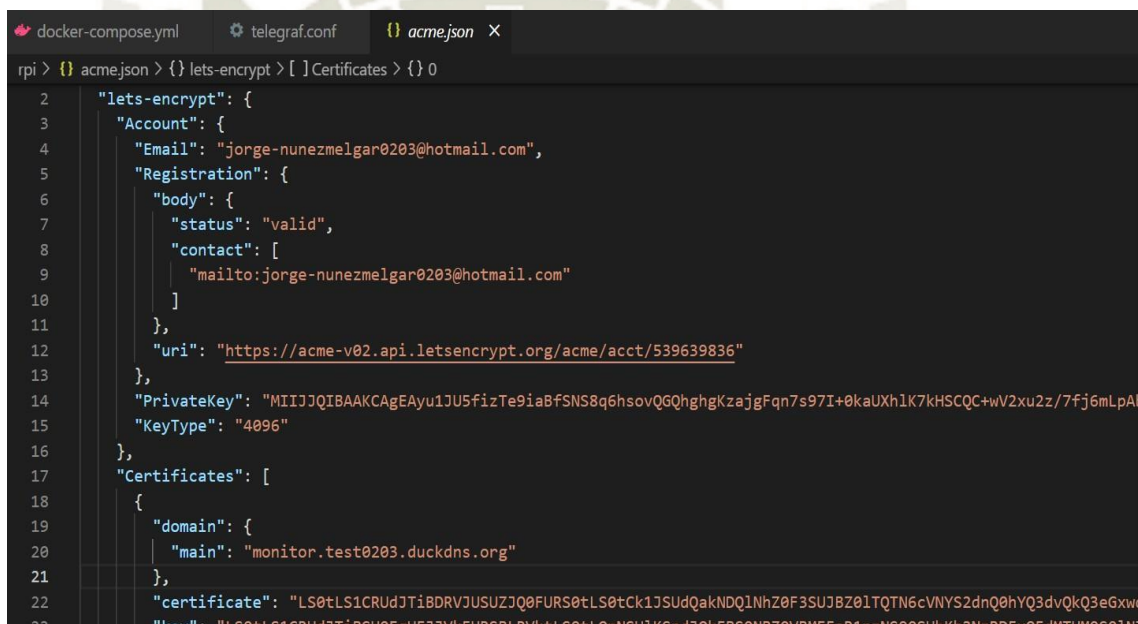
*Nota:* Fuente: Elaboración propia.

#### 4.2.5.8.1. Necesidad de uso del proxy inverso.

Traefik es un equilibrador de carga y proxy inverso moderno que facilita la implementación de microservicios. Traefik se integra con los componentes de su infraestructura existente y se configura de forma automática y dinámica.

Traefik está diseñado para ser lo más simple posible de operar, pero capaz de manejar implementaciones grandes y altamente complejas en una amplia gama de entornos y protocolos en nubes públicas, privadas e híbridas. También viene con un poderoso conjunto de middlewares que mejoran sus capacidades para incluir equilibrio de carga, puerta de enlace API, entrada de orquestador, así como comunicación de servicio este-oeste y más.

#### 4.2.5.8.2. Certificados SSL.



```

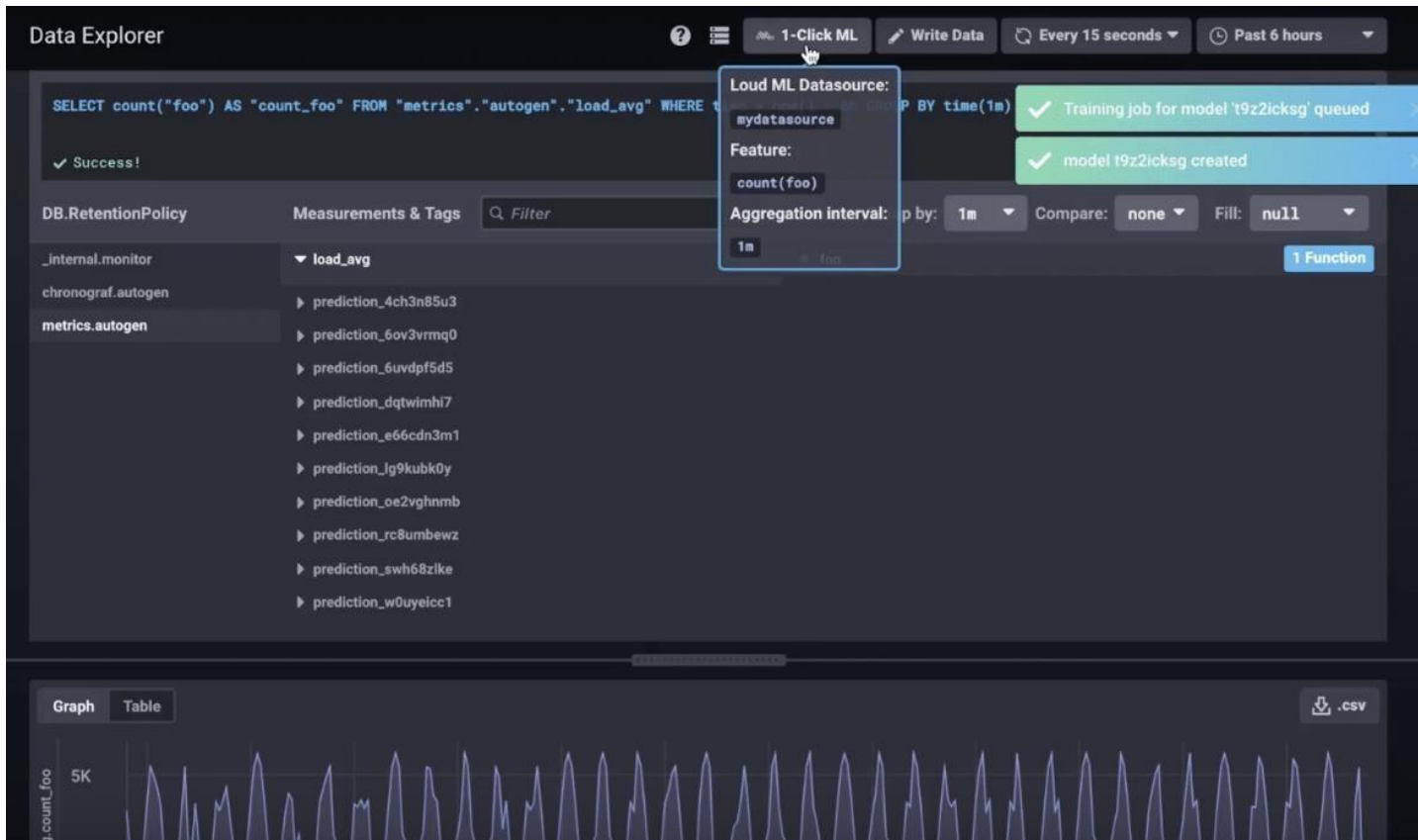
rpi > {} acme.json > {} lets-encrypt > [ ] Certificates > {} 0
2   "lets-encrypt": {
3     "Account": {
4       "Email": "jorge-nunezmelgar0203@hotmail.com",
5       "Registration": {
6         "body": {
7           "status": "valid",
8           "contact": [
9             "mailto:jorge-nunezmelgar0203@hotmail.com"
10          ]
11        },
12        "uri": "https://acme-v02.api.letsencrypt.org/acme/acct/539639836"
13      },
14      "PrivateKey": "MIIJJQIBAAKCAgEAYu1JU5fizTe9iaBfSNS8q6hs0vQGQhghgKzajgFqn7s97I+0kaUXh1K7kHSCQC+wwV2xu2z/7fj6mLpAb
15      "KeyType": "4096"
16    },
17    "Certificates": [
18      {
19        "domain": {
20          "main": "monitor.test0203.duckdns.org"
21        },
22        "certificate": "LS0tLS1CRUdJTiBDRVJUSUZJQ0FURS0tLS0tCk1JSUdQakNDQ1NhZ0F3SUJBZ01TQTN6cVNYNS2dnQ0hYQ3dvQkQ3eGxwd
23        "key": "LS0tLS1CRUdJTiBDRVJUSUZJQ0FURS0tLS0tCk1JSUdQakNDQ1NhZ0F3SUJBZ01TQTN6cVNYNS2dnQ0hYQ3dvQkQ3eGxwd

```

Figura 53 Certificados SSL generados automáticamente para cada servicio

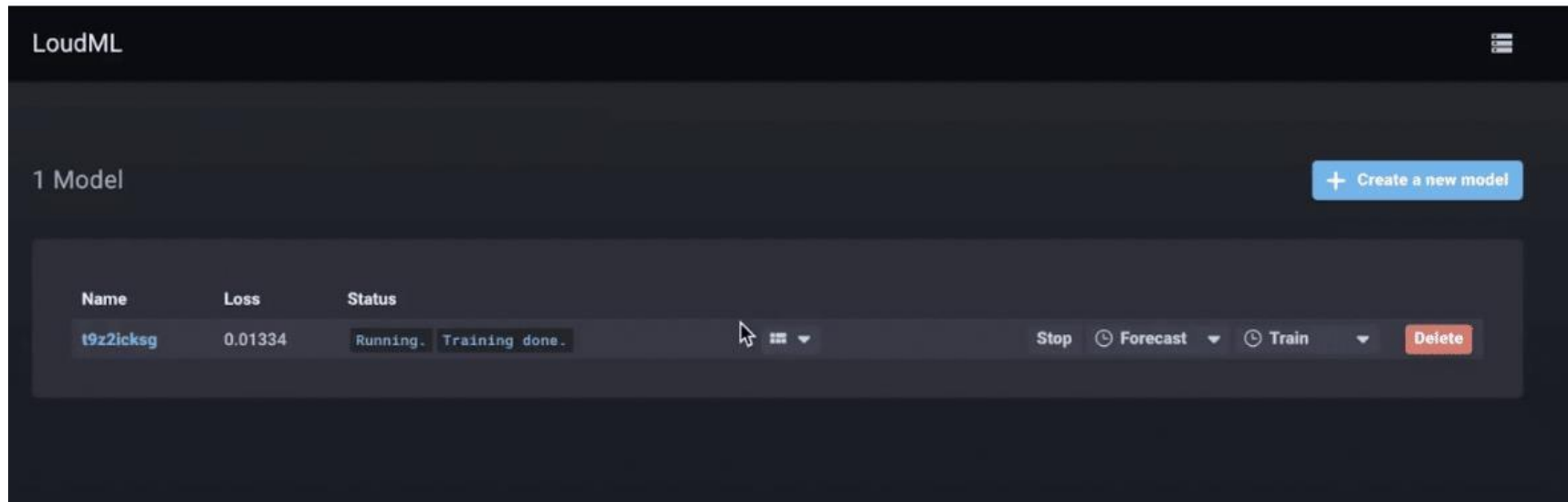
Nota: Fuente: Elaboración propia.

#### 4.2.5.9. Loud ML.



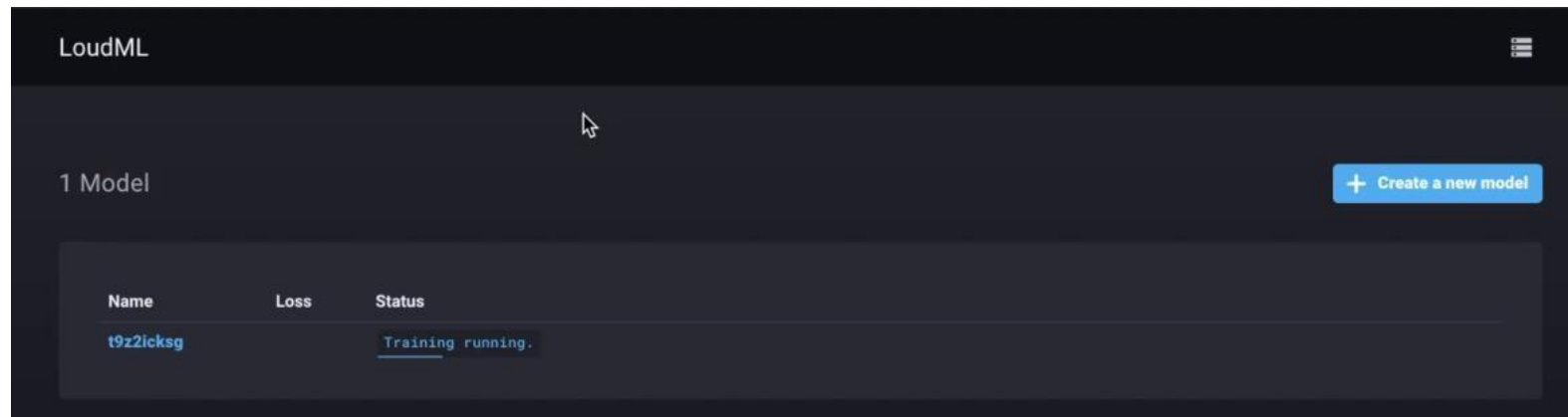
*Figura 54 Menú entrenamiento Loud ML*

*Nota:* Fuente: Elaboración propia.



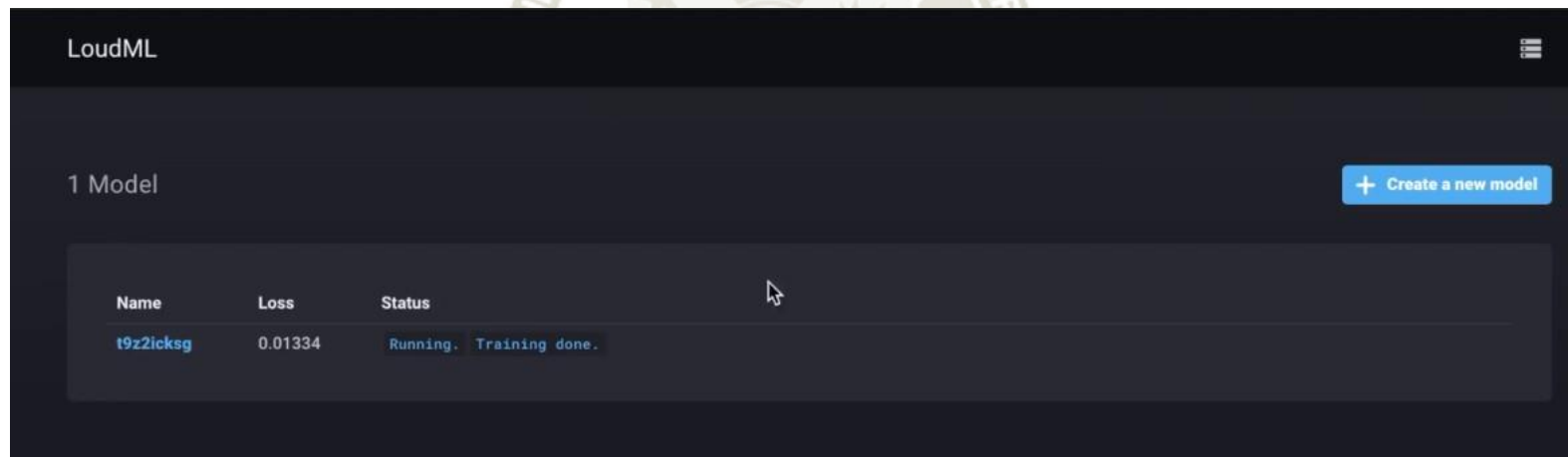
*Figura 55 Entrenamiento en proceso*

*Nota:* Fuente: Elaboración propia.



*Figura 56 Entrenamiento completado*

*Nota:* Fuente: Elaboración propia.



*Figura 57 Funcion predictiva Loud ML lista para utilizar*

*Nota:* Fuente: Elaboración propia.

#### 4.2.5.10. Rapid SCADA.

##### 4.2.5.10.1. Dashboards.

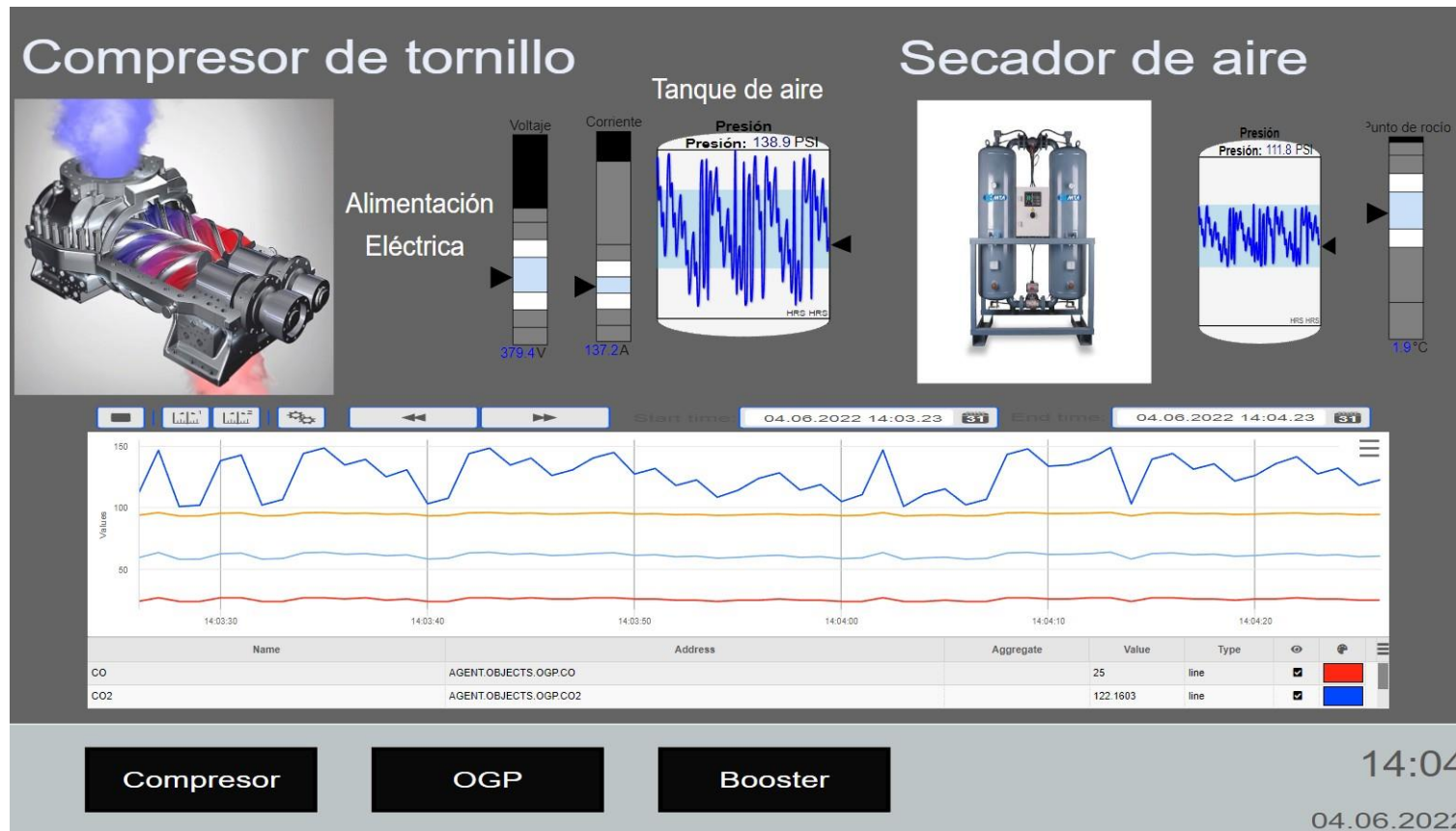


Figura 58 Dashboard Compresor

Nota: Fuente: Elaboración propia

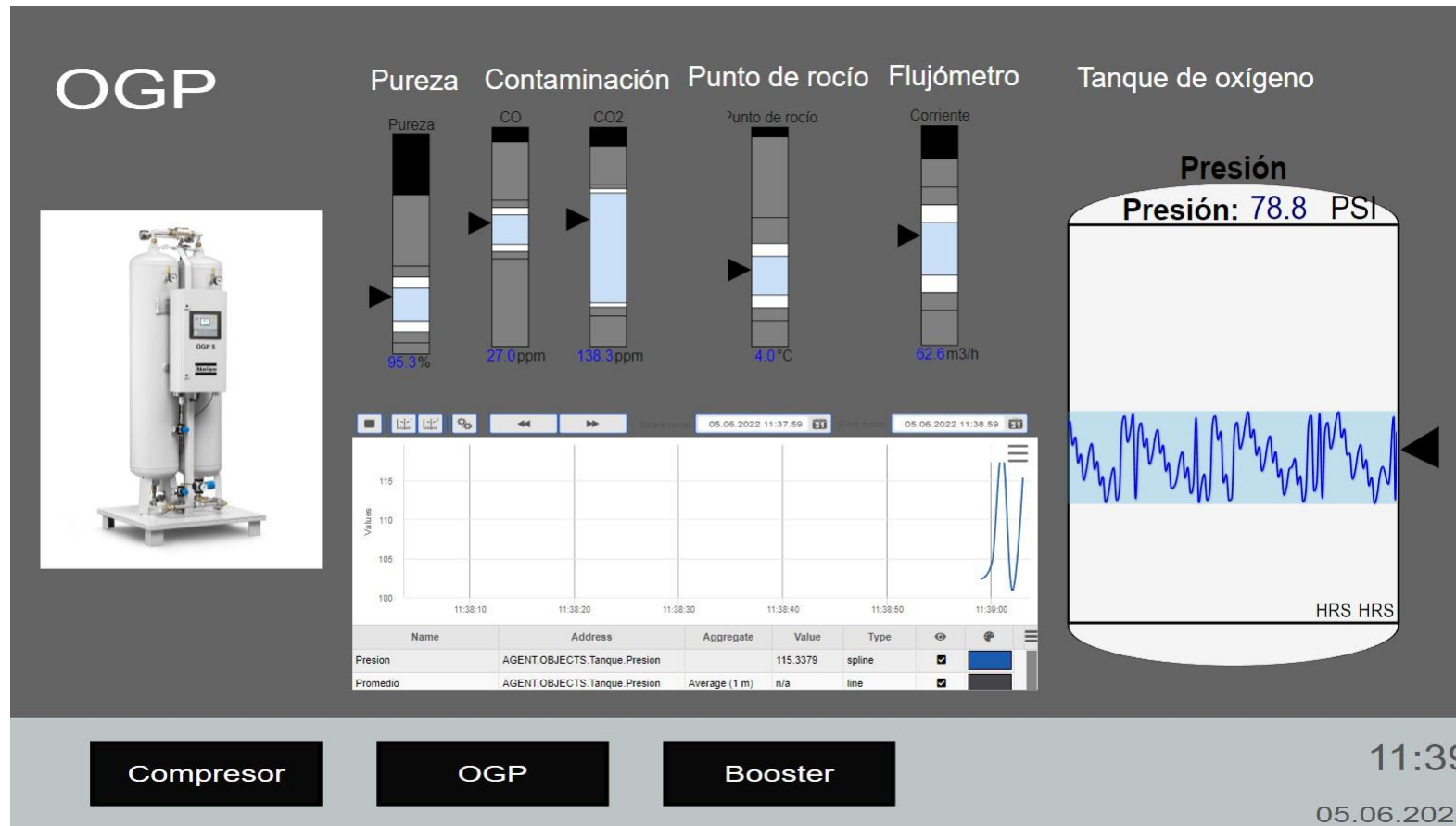


Figura 59 Dashboard OGP

Nota: Fuente: Elaboración propia.

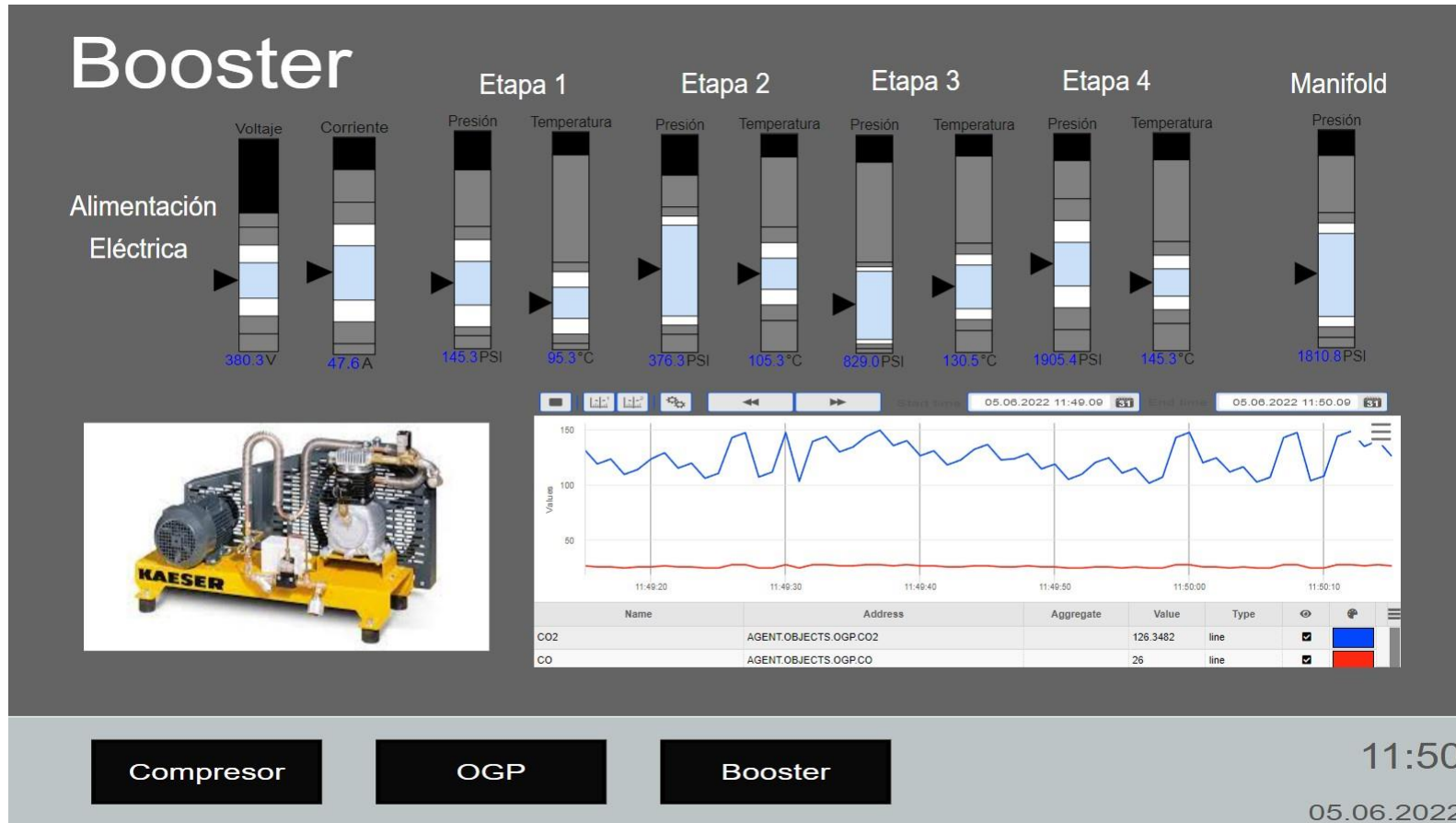


Figura 60 Dashboard Booster

Nota: Fuente: Elaboración propia.

En la realización de los Dashboard se tuvo en cuenta el un fácil entendimiento para el usuario, agregando ítems de poco cansancio para la visión, se tienen voltajes, corrientes eléctricas,presiones y datos históricos en el caso del compresor, acompañado de una pequeña ilustración parasu entendimiento. En el caso del OGP, se tiene los parámetros importantes como el de la pureza yla contaminación, así como la presión en el tanque de almacenamiento finalmente se tiene el flujode consumo. Para terminar con el Dashboard del booster se tiene voltaje y corriente eléctrica de consumo, acompañados por la presión y temperatura medidas en cada etapa de este, para finalmente acabar con registros históricos durante un tiempo predeterminado por el usuario.

4.2.5.10.2. Reportes.

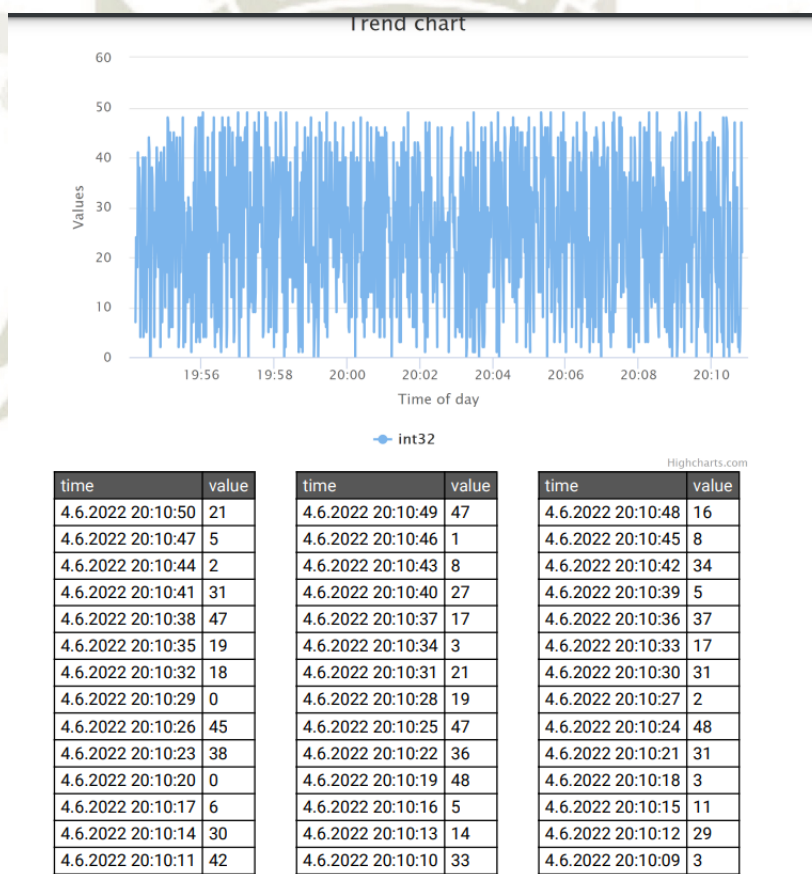


Figura 61 Reporte generado en el software SCADA

Nota: Fuente: Elaboración propia.

4.2.5.10.3. Alarmas.

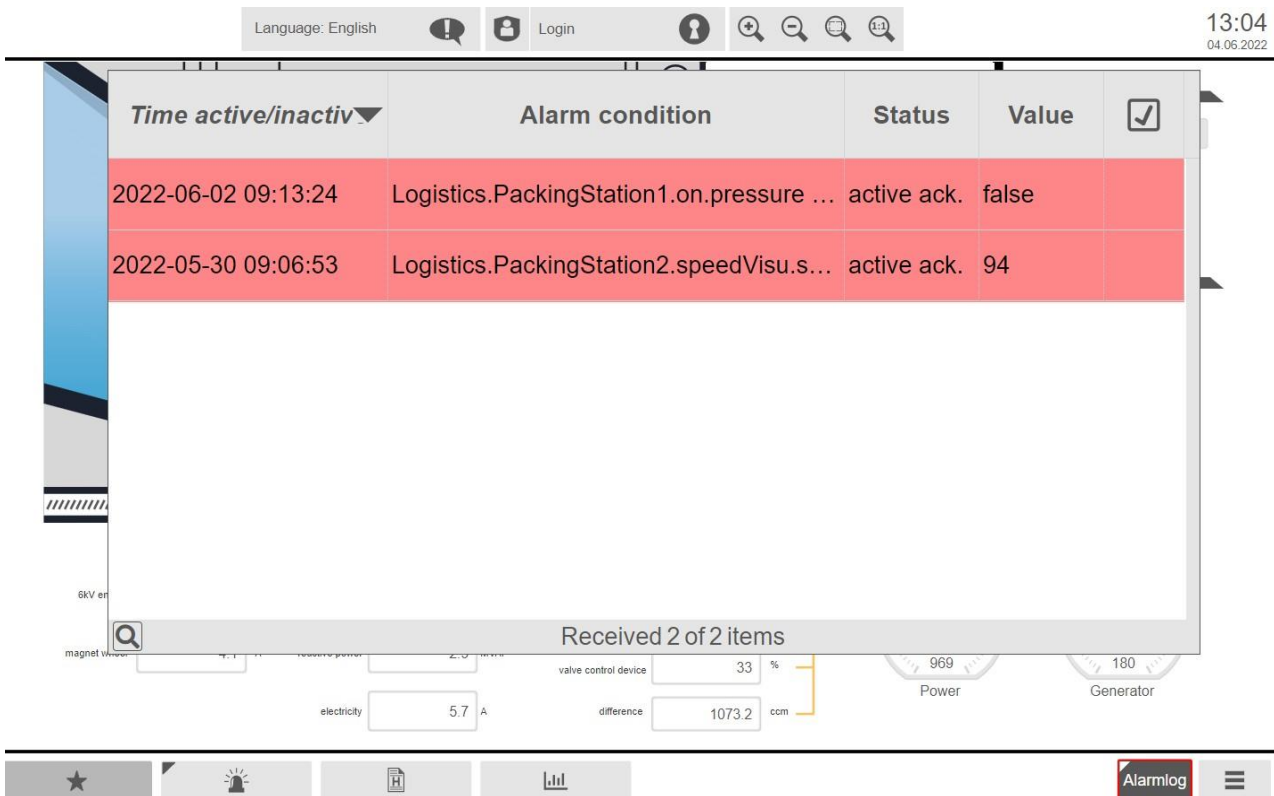


Figura 62 Alarmas generadas  
Nota: Fuente: Elaboración propia.

4.2.6. Costos y presupuesto.

Tabla 11 Costos y presupuesto.

ítem	Cantidad	Precio unitario	Precio total
Raspberry Pi	1	800	800
Fuente de alimentación	1	100	100
Costo anual electricidad	1	41.61	41.61
Disco SSD	1	200	200
Accesorios	1	100	100
Horas hombre	50	80	4000
<b>Total</b>			<b>5241.61</b>

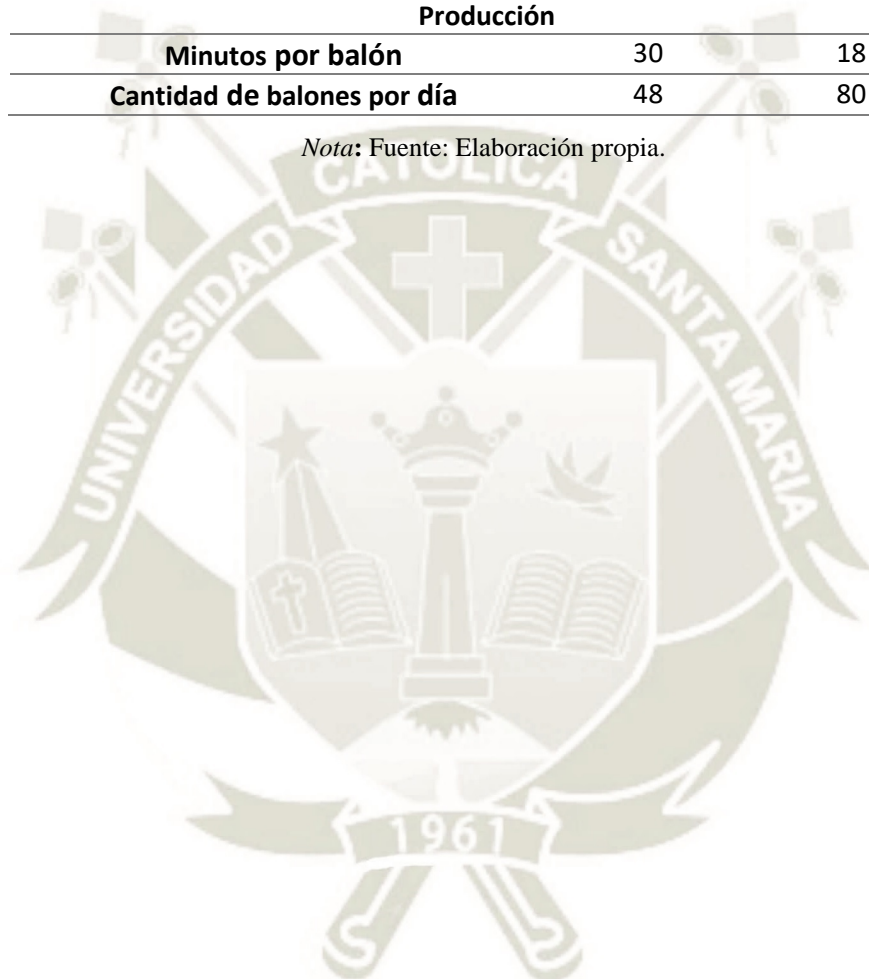
Nota: Fuente: Elaboración propia.

#### 4.2.7 Indicadores clave de desempeño y producción diaria

Tabla 12 KPIs de operaciones

KPI de operaciones		
	Antes (min)	Después (min)
Tiempo en alcanzar pureza definida	45	24
Tiempo de parada entre cargas	25	5
Tiempo de llenado (8 balones)	150	100
Tiempo de reposo	20	15
Tiempo total	240	144
Producción		
Minutos por balón	30	18
Cantidad de balones por día	48	80

Nota: Fuente: Elaboración propia.



## CONCLUSIONES

- Se pudo diseñar e implementar con éxito el Digital Twin de las plantas generadoras de oxígeno, mejorando la automatización mediante la adquisición de datos de los instrumentos y la retroalimentación del control de flujo para una pureza de oxígeno estable. La mejora de esta automatización se ve reflejada en lo que respecta al tiempo de llenado de los balones de oxígeno, anteriormente a la automatización se tenía un tiempo de llenado de 20-25 min por balón lo que significa 48 balones con un tiempo de operación de 22 horas al día, posterior a la automatización se tiene un tiempo de llenado de 12-15 min por balón lo que significa 80 balones en un tiempo de operación de 22 horas diarias, prácticamente duplicando la cantidad de balones que la planta puede producir.
- Se integraron correctamente al “Digital Twin” los sensores de tecnología Zigbee y Wifi para la obtención de datos, en la implementación se pudo demostrar la superioridad en cuanto a velocidad de los dispositivos WiFi, teniendo estos una velocidad máxima de 11 Mbps en contra de los 250 kbps de máxima transferencia en el protocolo Zigbee; la superioridad de los dispositivos Zigbee en cuanto a consumo eléctrico, siendo un 10% el consumo de los dispositivos WiFi dependiendo de la aplicación, además de esto los dispositivos Zigbee pueden mantenerse conectados a una batería mientras que los WiFi es preferible usar una fuente de alimentación interrumpida. No pudiendo decantarse netamente por un protocolo, ambos protocolos son utilizados dependiendo de la cantidad de datos a transportar, considerando su bajo precio, de 10-30 dólares, dependiendo de la aplicación, son una alternativa bastante útil.
- Se diseñó e implementó con éxito un servidor de automatización en el hardware Raspberry Pi, demostrando su bajo consumo eléctrico, 54.75 kWh al año y su capacidad de computación (1.5 – 6 GHz) para ejecutar los servicios asociados al Digital Twin, esto sumado a su bajo precio, 150 – 200 dólares, sumado a la facilidad de implementación demuestra ser una buena opción para la implementación del Digital Twin.
- Se implementó con éxito una base de tiempo de series temporales, demostrando la alta eficiencia que tiene para con los datos masivos, esto puede explicarse debido a que se puede insertar 100K a 1M datos cada segundo, tener funciones integradas de

transformación y agregación de series temporales y manejar un esquema optimizado para matrices de series temporales con fragmentación e indexación integradas; este tipo de bases de datos de series temporales son las preferidas sobre las relacionales para el desarrollo del IIOT.

- Se logró implementar con éxito interfaces gráficas amigables al uso del usuario, mediante el uso del contenedor influxdb, aprovechando la potencia de esta herramienta y además se implementó un sistema SCADA como se mostró en figuras anteriores aprovechando la conectividad MQTT.
- Se integraron todos los servicios para la correcta ejecución del Digital Twin, estos se crearon mediante contenedores Docker, que nos demostró una facilidad mayor a la hora de iniciarlos, así como de interconectarlos para que puedan compartir información y acciones, sin olvidar que no están instalados del todo en el servidor y que su reproducibilidad es la más rápida posible para futuros proyectos.
- Se identificó con claridad el concepto de Digital Twin mediante la revisión de varios trabajos de investigación; también se logró identificar los componentes teóricos adecuados para posteriormente diseñarlos e implementarlos.

## RECOMENDACIONES

- Se recomienda para implementaciones posteriores, el uso de un computador de más recursos (Intel Nuk) que pueda contar con procesadores de arquitectura x86, esto debido a que usualmente las imágenes de los servicios suelen tener sus primeras versiones y actualizaciones para esta arquitectura; además que en este tipo de computadores pueden cambiarse partes del hardware (Memoria RAM, Almacenamiento SSD).
- Debido a el análisis de datos que el Digital Twin es capaz de realizar se recomienda obtener más dispositivos actuadores para hacer la acción de realimentación en las plantas
- Se recomienda hacer la búsqueda de un sistema SCADA con un base de tiempo se series temporales, que sea compatible con los softwares de Influx data.
- Se recomienda implementar algún servicio de generación de reportes PDF en base a datos obtenidos, análisis y alarmas activadas durante la operación del Digital Twin.
- Se recomienda implementar un simulador que se alimente de la data obtenida con el DigitalTwin, para realizar pruebas sin entrar en contacto con el activo físico.
- Se recomienda integrar al sistema de comunicaciones, un Gateway IioT encargado netamente de extraer los datos del PLC u otros remotos y convertir estos a un formato MQTT, esto hará que el PLC libere la carga de comunicaciones y se limite netamente a acciones de control e instrumentación.
- Debido a la gran cantidad de data recolectada, esta puede ser usada en distintos planes de mantenimiento predictivo, relacionando ciertos eventos a la data generada para encontrar ciertos patrones y luego soluciones, posteriormente esto debe automatizarse mediante aprendizaje automático para tener un mejor performance.

## REFERENCIAS

- [1] Y. Zheng, S. Yang y H. Cheng, An application framework of digital twin and its case study, vol. 10, 2018: *Journal of Ambient Intelligence and Humanized Computing*, p. 1141–1153.
- [2] H. Zhang, Q. Liu, X. Chen, D. Zhang y J. Leng, A Digital Twin-Based Approach for Designing and Multi-Objective Optimization of Hollow Glass Production Line, vol. 5, *IEEE Access*, 2017, p. 26901–26911.
- [3] Y. Xu, Y. Sun, L. Liu y Y. Zheng, A Digital-Twin-Assisted Fault Diagnosis Using Deep Transfer Learning, vol. 7, *IEEE Access*, 2019, pp. 19990–19999.
- [4] D. Wu, C. Jennings, J. Terpenney y S. Kumara, «Cloud-based machine learning for predictive analytics: Tool wear prediction in milling,» de *2016 IEEE International Conference on Big Data (Big Data)*, 2016.
- [5] R. Vrabic, A. Erkoyuncu, P. Butala y R. Roy, Digital twins: Understanding the added value of integrated models for through-life engineering services, vol. 16, *Procedia Manufacturing*, 2018, p. 139–146.
- [6] R. Vishwakarma y A. K. Jain, «A survey of DDoS attacking techniques and defence mechanisms in the IoT network,» *Telecommunication Systems*, 2019.
- [7] Y. Umeda, J. Ota, F. Kojima, M. Saito, H. Matsuzawa, T. Sukekawa, A. Takeuchi, K. Makida y S. Shirafuji, Development of an education program for digital manufacturing system engineers based on ‘Digital Triplet’ concept, vol. 31, *Procedia Manufacturing*, 2019, p. 363–369.
- [8] F. Tao, J. Cheng y Q. Qi, IIHub: An Industrial Internet-of-Things Hub Toward Smart Manufacturing Based on Cyber-Physical System, vol. 14, *IEEE Transactions on Industrial Informatics*, 2018, p. 2271–2280.
- [9] F. Tao, Q. Qi, L. Wang y A. Nee, Digital Twins and Cyber-Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison, vol. 5, *Engineering*, 2019, p. 653–661.
- [10] F. Tao y M. Zhang, Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing, vol. 5, *IEEE Access*, 2017, p. 20418–20427.

- [11] F. Tao, H. Zhang, A. Liu y A. Nee, Digital Twin in Industry: State-of-the-Art, vol. 15, IEEE Transactions on Industrial Informatics, 2019, p. 2405–2415.
- [12] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang y F. Sui, Digital twin- driven product design, manufacturing and service with big data, vol. 94, The International Journal of Advanced Manufacturing Technology, 2018, p. 3563–3576.
- [13] I. Taleb, R. Dssouli y M. Serhani, «Big Data Pre-processing: A Quality Framework,» p. 191–198, 2015.
- [14] Statista, «IoT: number of connected devices worldwide 2012-2025».
- [15] R. Stark, C. Fresemann y K. Lindow, «Development and operation of Digital Twins for technical systems and services,» *CIRP Annals*, 2019.
- [16] M. Stanley, «IIoT & the New Industrial Revolution,» 2016.
- [17] S. Srivastava, A. Bisht y N. Narayan, «Safety and security in smart cities using artificial intelligence — A review,» de *2017 7th International Conference on Cloud Computing, Data Science Engineering - Confluence*, 2017.
- [18] Sonal, S. Reddy y D. Kumar, «Review of Smart Health Monitoring Approaches with Survey Analysis and Proposed Framework,» *IEEE Internet of Things Journal*, pp. 1-1, 2018.
- [19] R. Soe, «Smart Twin Cities via Urban Operating System,» de *Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance - ICEGOV '17*, New Delhi AA, 2017.
- [20] R. Soe, «FINEST Twins: platform for cross-border smart city solutions,» de *Proceedings of the 18th Annual International Conference on Digital Government Research - dg.o '17*, Staten Island, 2017.
- [21] K. Sivalingam, M. Sepulveda, M. Spring y P. Davies, «A Review and Methodology Development for Remaining Useful Life Prediction of Off-shore Fixed and Floating Wind turbine Power Converter with Digital Twin Technology Perspective,» de *2018 2nd International Conference on Green Energy and Applications (ICGEA)*, Singapore, 2018.
- [22] P. P. Shinde y S. Shah, «A Review of Machine Learning and Deep Learning Applications,» de *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2018.

- [23] D. Shangguan, L. Chen y J. Ding, «A Hierarchical Digital Twin Model Framework for Dynamic Cyber-Physical System Design,» de *Proceedings of the 5th International Conference on Mechatronics and Robotics Engineering - ICMRE'19*, Roma, 2019.
- [24] O. Sezer, E. Dogdu y A. Ozbayoglu, Context-Aware Computing, Learning, and Big Data in Internet of Things: A Survey, vol. 5, *IEEE Internet of Things Journal*, 2018, pp. 1-27.
- [25] B. Schleich, N. Anwer, L. Mathieu y S. Wartzack, Shaping the digital twin for design and production engineering, vol. 66, *CIRP Annals*, 2017, p. 141–144.
- [26] T. I. a. R. R. B. S V Engineering College, A Comprehensive Literature Review on Data Analytics in IIoT (Industrial Internet of Things), vol. 8, *HELIX*, 2018, p. 2757–2764.
- [27] T. Ruohomäki, E. Airaksinen, P. Huuska, O. Kesäniemi, M. Martikka y J. Suomisto, «Smart City Platform Enabling Digital Twin,» de *2018 International Conference on Intelligent Systems (IS)*, 2018.
- [28] D. Ross, Digital twinning [Information Technology Virtual Reality, vol. 11, *Engineering Technology*, 2016, pp. 44-45.
- [29] A. Reyna, C. Martín, J. Chen, E. Soler y M. Díaz, On blockchain and its integration with IoT. Challenges and opportunities, vol. 88, *Future Generation Computer Systems*, 2018, p. 173–190.
- [30] Q. Qi, F. Tao, Y. Zuo y D. Zhao, Digital Twin Service towards Smart Manufacturing, vol. 72, *Procedia CIRP*, 2018, p. 237–242.
- [31] Q. Qi y F. Tao, Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison, vol. 6, *IEEE Access*, 2018, p. 3585–3593.
- [32] D. Petrik y G. Herzwurm, «IIoT ecosystem development through boundary resources: a Siemens MindSphere case study,» de *Proceedings of the 2nd ACM SIGSOFT International Workshop on Software-Intensive Business: Start-ups, Platforms, and Ecosystems - IWSiB 2019*, Tallinn, 2019.
- [33] H. Pargmann, D. Euhausen y R. Faber, «Intelligent big data processing for wind farm monitoring and analysis based on cloud-technologies and digital twins: A quantitative approach,» de *2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, Chengdu, 2018.
- [34] S. Ornes, Core Concept: The Internet of Things and the explosion of interconnectivity, vol. 113, *Proceedings of the National Academy of Sciences*, 2016, p. 11059–11060.

- [35] O. Novo, Blockchain Meets IoT: An Architecture for Scalable Access Management in IoT, vol. 5, IEEE Internet of Things Journal, 2018, p. 1184–1195.
- [36] NetObjex, «Industrial Revolution,» *NetObjex*, 2018.
- [37] E. Negri, L. Fumagalli y M. Macchi, A Review of the Roles of Digital Twin in CPS-based Production Systems, vol. 11, Procedia Manufacturing, 2017, p. 939–948.
- [38] M. Najafabadi, F. Villanustre, T. Khoshgoftaar, N. Seliya, R. Wald y E. Muharemagic, Deep learning applications and challenges in big data analytics, vol. 2, Journal of Big Data, 2015.
- [39] A. Moujahid, M. Tantaoui, M. Hina, A. Soukane, A. Ortalda, A. ElKhadimi y A. Ramdane-Cherif, «Machine Learning Techniques in ADAS: A Review,» de *2018 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, 2018.
- [40] N. Mohammadi y J. E. Taylor, «Smart city digital twins,» de *IEEE Symposium Series on Computational Intelligence (SSCI)*, Honolulu, 2017.
- [41] A. Modi, «Review Article on Deep Learning Approaches,» de *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2018.
- [42] Q. Min, Y. Lu, Z. Liu, C. Su y B. Wang, «Machine Learning based Digital Twin Framework for Production Optimization in Petrochemical Industry,» *International Journal of Information Management*, 2019.
- [43] V. J. Mawson y B. R. Hughes, The development of modelling tools to improve energy efficiency in manufacturing processes and systems, vol. 51, Journal of Manufacturing, 2019, p. 95–105.
- [44] C. Mandolla, A. M. Petruzzelli, G. Percoco y A. Urbinati, «Computers in Industry. Building a digital twin for additive manufacturing through the exploitation of blockchain: A case analysis of the aircraft industry,» vol. 109, Science Direct, 2019, pp. 134-152.
- [45] M. Malik, A. Ahmad y A. Bilberg, «Virtual reality in manufacturing: immersive and collaborative artificial-reality in design of human-robot workspace,» *International Journal of Computer Integrated Manufacturing*, 2020.
- [46] R. Magargle, L. Johnson, P. Mandloi, P. Davoudabadi, O. Kesarkar, S. Krishnaswamy, J. Batteh y A. Pitchaikani, «A Simulation-Based Digital Twin for Model-Driven Health Monitoring and Predictive Maintenance of an Automotive Braking System,» 2017.

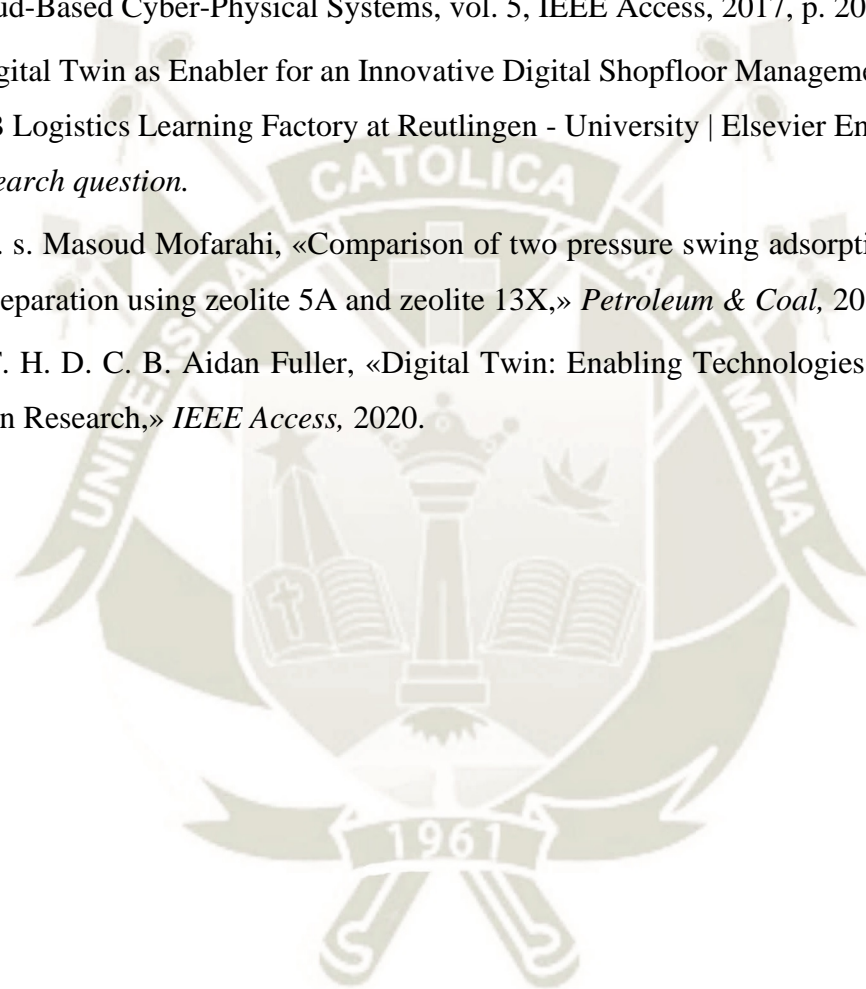
- [47] A. Madni, C. Madni y S. Lucero, Leveraging Digital Twin Technology in Model-Based Systems Engineering, vol. 7, Systems, 2019, p. 7.
- [48] Y. Lu, T. Peng y X. Xu, Energy-efficient cyber-physical production network: Architecture and technologies, vol. 129, Computers & Industrial Engineering, 2019, p. 56–66.
- [49] F. Longo, L. Nicoletti y A. Padovano, Ubiquitous knowledge empowers the Smart Factory: The impacts of a Service-oriented Digital Twin on enterprises' performance, Annual Reviews in Control, 2019, p. 221– 236.
- [50] M. Lohtander, N. Ahonen, M. Lanz, J. Ratava y J. Kaakkunen, Micro Manufacturing Unit and the Corresponding 3d-Model for the Digital Twin, vol. 25, Procedia Manufacturing, 2018, pp. 55-61.
- [51] Z. Liu, N. Meyendorf y N. Mrad, «The role of data fusion in predictive maintenance using digital twin,» de *Annual Review Of Progress In Quantitative Nondestructive Evaluation*, Provo, 2018.
- [52] Y. Liu, L. Zhang, Y. Yang, L. Zhou, L. Ren, F. Wang, R. Liu, Z. Pang y M. J. Deen, A Novel Cloud-Based Framework for the Elderly Healthcare Services Using Digital Twin, vol. 7, IEEE Access, p. 49088–49101.
- [53] J. Liu, X. Du, H. Zhou, X. Liu, L. Ei Li y F. Feng, A digital twin-based approach for dynamic clamping and positioning of the flexible tooling system, vol. 80, Procedia CIRP, 2019, p. 746–749.
- [54] K. Lasse Lueth, «State of the IoT 2018: Number of IoT devices now at 7b – Market accelerating».
- [55] H. Laaki, Y. Miche y K. Tammi, Prototyping a Digital Twin for Real Time Remote Control Over Mobile Networks: Application of Remote Surgery, vol. 7, IEEE Access, 2019, p. 20325–20336.
- [56] S. Kumar y A. Jasuja, «Air quality monitoring system based on IoT using Raspberry Pi,» de *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 2017.
- [57] S. Kumar y A. Jasuja, «Air quality monitoring system based on IoT using Raspberry Pi,» de *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 2017.

- [58] W. Kuehn, «Simulation in Digital Enterprises,» de *Proceedings of the 11th International Conference on Computer Modeling and Simulation – ICCMS 2019*, North Rockhampton, 2019.
- [59] M. Kritzler, J. Hodges, D. Yu, K. Garcia, H. Shukla y F. Michahelles, «Digital Companion for Industry,» de *Companion Proceedings of The 2019 World Wide Web Conference on - WWW '19*, San Francisco, 2019.
- [60] W. Kritzinger, M. Karner, G. Traar, J. Henjes y W. Sihn, Digital Twin in manufacturing: A categorical literature review and classification, vol. 51, *IFAC- PapersOnLine*, 2018, p. 1016–1022.
- [61] K. Kobara, Cyber Physical Security for Industrial Control Systems and IoT, vol. E99, *IEICE Transactions on Information and Systems*, 2016, p. 787–795.
- [62] G. Knapp, T. Mukherjee, J. Zuback, H. Wei, T. Palmer, A. De y T. DebRoy, Building blocks for a digital twin of additive manufacturing, vol. 135, *Acta Materialia*, 2017, p. 390–399.
- [63] A. Karadeniz, I. Arif, A. Kanak y S. Ergun, «Digital Twin of eGas- tronomic Things: A Case Study for Ice Cream Machines,» de *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, Sapporo, 2019.
- [64] M. Joordens y M. Jamshidi, «On The Development of Robot Fish Swarms in Virtual Reality with Digital Twins,» de *2018 13th Annual Conference on System of Systems Engineering (SoSE)*, Paris, 2018.
- [65] S. Jo, D. Park, H. Park y S. Kim, «Smart Livestock Farms Using Digital Twin: Feasibility Study,» de *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, 2018.
- [66] P. Jain, J. Poon, J. Singh, C. Spanos, S. Sanders y S. Panda, «A Digital Twin Approach for Fault Diagnosis in Distributed Photovoltaic System,» *IEEE Transactions on Power Electronics*, pp. 1-1, 2019.
- [67] S. Huh, S. Cho y S. Kim, «Managing IoT devices using blockchain platform,» 2017.
- [68] D. Howard, «The Digital Twin: Virtual Validation In Electronics Develop- ment And Design,» de *Pan Pacific Microelectronics Symposium (Pan Pacific)*, Kauai, 2019.
- [69] Y. He, J. Guo y X. Zheng, «From Surveillance to Digital Twin: Chal- lenges and Recent Advances of Signal Processing for Industrial Internet of Things,» *IEEE Signal Processing Magazine*, vol. 35, p. 120–129, 2018.

- [70] R. He, G. Chen, C. Dong, S. Sun y X. Shen, «Data-driven digital twin technology for optimized control in process systems,» *ISA Transactions*, 2019.
- [71] M. Grieves, «Digital Twin: Manufacturing Excellence through Virtual Factory Replication,» NASA, 2014.
- [72] E. Gleassgen y D. Stargel, «The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles,» de *53rd AIAA Structures, Structural Dynamics and Materials Conference*, Honolulu, 2012.
- [73] X. Fei, N. Shah, N. Verba, K. Chao, V. Sanchez-Anguix, J. Lewandowski, A. James y Z. Usman, CPS data streams analytics based on machine learning for Cloud and Fog Computing: A survey, *IEEE Internet of Things Journal*, CORREGIR.
- [74] O. Elijah, T. Rahman, I. Orikumhi, C. Leow y M. Hindia, An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges, vol. 5, *IEEE Internet of Things Journal*, 2018, p. 3758–3773.
- [75] A. El Saddik, Digital Twins: The Convergence of Multimedia Technologies, vol. 25, *IEEE MultiMedia*, 2018, pp. 87-92.
- [76] I. U. Din, M. Guizani, J. J. Rodrigues, S. Hassan y V. Korotaev, Machine learning in the Internet of Things: Designed techniques for smart cities, vol. 100, *Future Generation Computer Systems*, 2019, p. 826–843.
- [77] T. DebRoy, W. Zhang, J. Turner y S. Babu, Building digital twins of 3d printing machines, vol. 135, *Scripta Materialia*, 2017, p. 119–124.
- [78] J. David, A. Lobov y M. Lanz, «Leveraging Digital Twins for Assisted Learning of Flexible Manufacturing Systems,» de *2018 IEEE 16th International Conference on Industrial Informatics (INDIN)*, Porto, 2018.
- [79] V. Damjanovic-Behrendt, «A Digital Twin-based Privacy Enhancement Mechanism for the Automotive Industry,» de *2018 International Conference on Intelligent Systems (IS)*, 2018.
- [80] D. Curry, «ARM: One trillion IoT devices by 2035, \$5 trillion in market value,» 2017.
- [81] A. Coraddu, L. Oneto, F. Baldi, F. Cipollini, M. Atlar y S. Savio, Data-driven ship digital twin for estimating the speed loss caused by the marine fouling, vol. 186, *Ocean Engineering*, 2019, p. 106063.

- [82] S. Chhetri, S. Faezi, A. Canedo y M. Faruque, «QUILT: quality inference from living digital twins in IoT-enabled manufacturing systems,» de *Proceedings of the International Conference on Internet of Things Design and Implementation - IoTDI '19*, 2019.
- [83] Y. Cheng, «Integrated and Intelligent Manufacturing: Perspectives and Enablers,» de *Engineering*, vol. 3, 2017, pp. 588-595.
- [84] X. Chen, E. Kang, S. Shiraiishi, V. Preciado y Z. Jiang, «Digital Behavioral Twins for Safe Connected Cars,» de *Proceedings of the 21th ACM/IEEE International Conference on Model Driven Engineering Languages and Systems - MODELS '18*, Copenhagen, 2018.
- [85] Y. Cai, B. Starly, P. Cohen y Y. Lee, Sensor Data and Information Fusion to Construct Digital-twins Virtual Machine Tools for Cyber-physical Manufacturing, vol. 10, *Procedia Manufacturing*, 2017, p. 1031– 1042.
- [86] A. Čolaković y M. Hadžialić, Internet of Things (IoT): A review of enabling technologies, challenges, and open research issues, vol. 144, *Computer Networks*, pp. 17-39.
- [87] D. Burnett, J. Thorp, D. Richards, K. Gorkovenko y D. Murray-Rust, «Digital twins as a resource for design research,» de *Proceedings of the 8th ACM International Symposium on Pervasive Displays - PerDis '19*, Palermo, 2019.
- [88] C. Brosinsky, D. Westermann y R. Krebs, «Recent and prospective developments in power system control centers: Adapting the digital twin technology for application in power system control centers,» de *2018 IEEE International Energy Conference (ENERGYCON)*, Limassol, 2018.
- [89] H. Brandtstaedter, C. Ludwig, L. Hubner, E. Tsouchnika, A. Jungiewicz y U. Wever, «DIGITAL TWINS FOR LARGE ELECTRIC DRIVE TRAINS,» de *2018 Petroleum and Chemical Industry Conference Europe (PCIC Europe)*, Antwerp, 2018.
- [90] H. Boyes, B. Hallaq, J. Cunningham y T. Watson, The industrial internet of things (IIoT): An analysis framework, vol. 101, *Computers in Industry*, 2018, pp. 1-12.
- [91] A. Bilberg y A. Malik, «Digital twin driven human–robot collaborative assembly,» *CIRP Annals*, 2019.
- [92] S. E. Bibri y J. Krogstie, «The core enabling technologies of big data analytics and context-aware computing for smart sustainable cities: a review and synthesis,» *Journal of Big Data*, vol. 4, 2017.

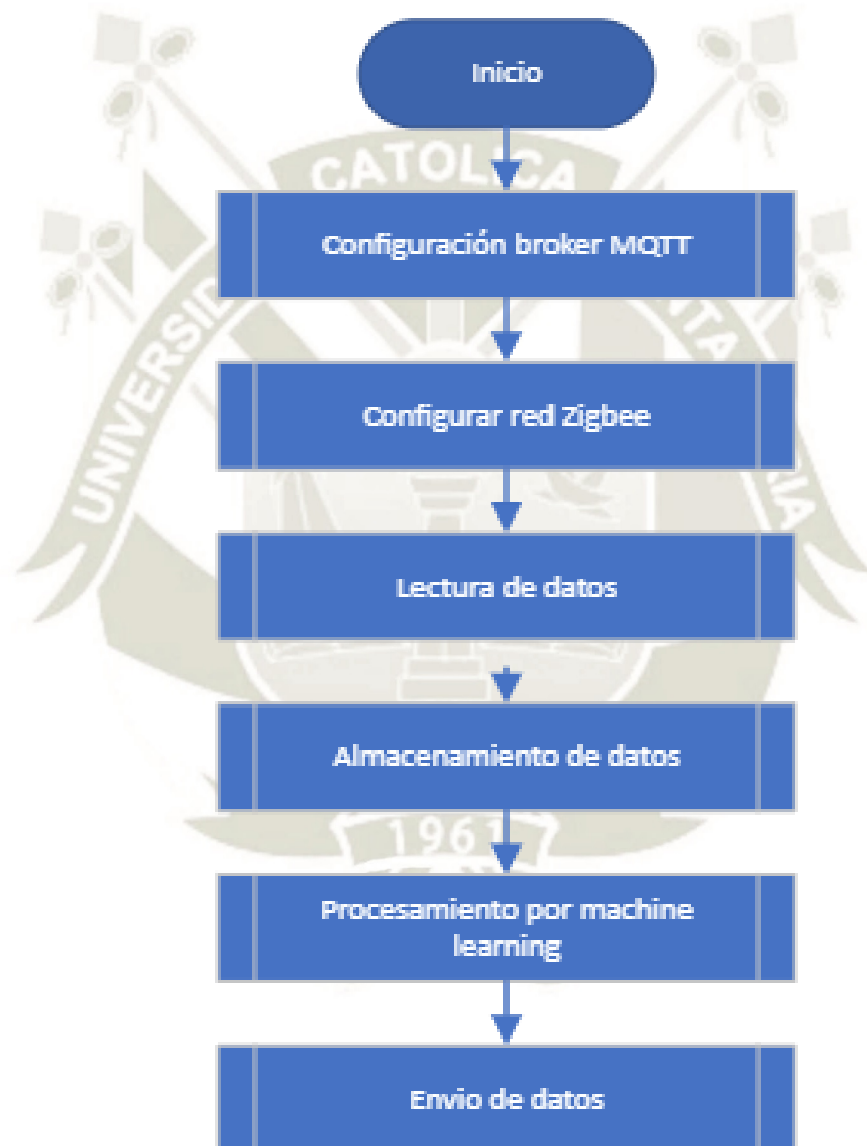
- [93] K. Barkwell, A. Cuzzocrea, C. Leung, A. Ocran, J. Sander- son, J. Stewart y B. Wodi, «Big Data Visualisation and Visual Analytics for Music Data Mining,» de *2018 22nd International Conference Information Visualisation (IV)*, 2018.
- [94] Automation, «Bentley Systems releases iModel.js open-source library,» 2018.
- [95] K. Ashton, «That 'Internet of Things' Thing,» *RFID Journal*, p. 1, 2009.
- [96] K. Alam y A. El Saddik, C2ps: A Digital Twin Architecture Reference Model for the Cloud-Based Cyber-Physical Systems, vol. 5, *IEEE Access*, 2017, p. 2050–2062.
- [97] «Digital Twin as Enabler for an Innovative Digital Shopfloor Management System in the ESB Logistics Learning Factory at Reutlingen - University | Elsevier Enhanced Reader,» *Research question*.
- [98] E. J. s. Masoud Mofarahi, «Comparison of two pressure swing adsorption processes for air separation using zeolite 5A and zeolite 13X,» *Petroleum & Coal*, 2013.
- [99] Z. F. H. D. C. B. Aidan Fuller, «Digital Twin: Enabling Technologies, Challenges and Open Research,» *IEEE Access*, 2020.



## ANEXOS

### 1. Diagramas de flujo

#### 1.1. Diagrama de flujo principal usado en la programación de la Raspberry.



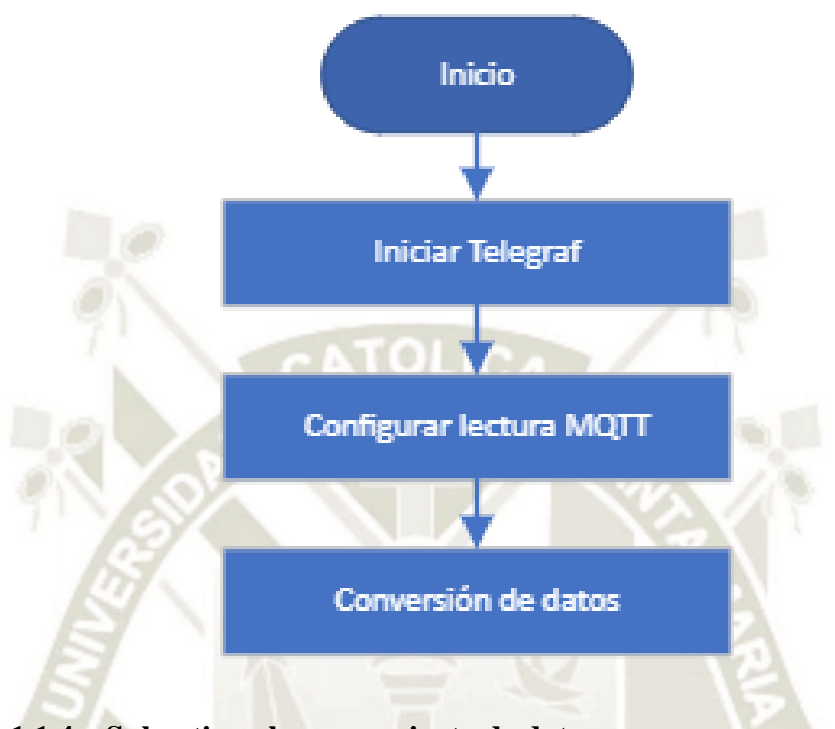
### 1.1.1. Subrutina configuración bróker MQTT.



### 1.1.2. Subrutina configuración red Zigbee



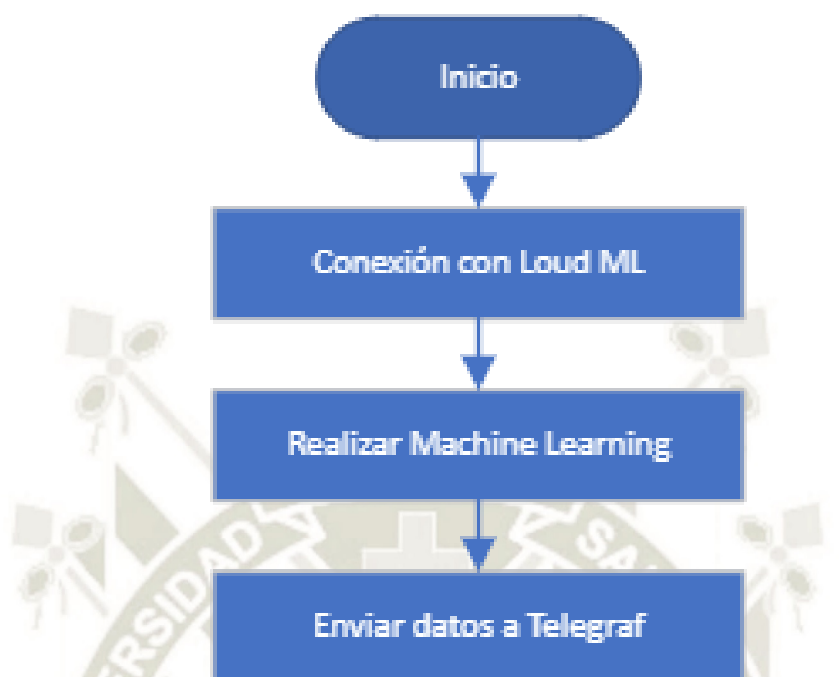
### 1.1.3. Subrutina lectura de datos.



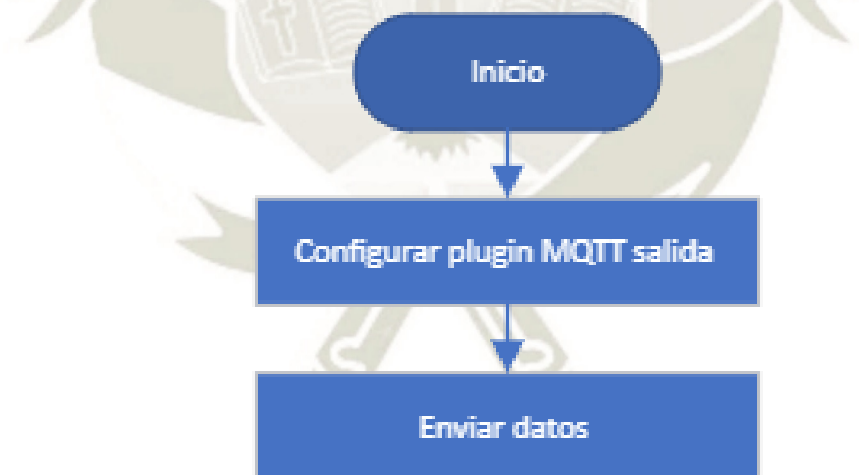
### 1.1.4. Subrutina almacenamiento de datos.



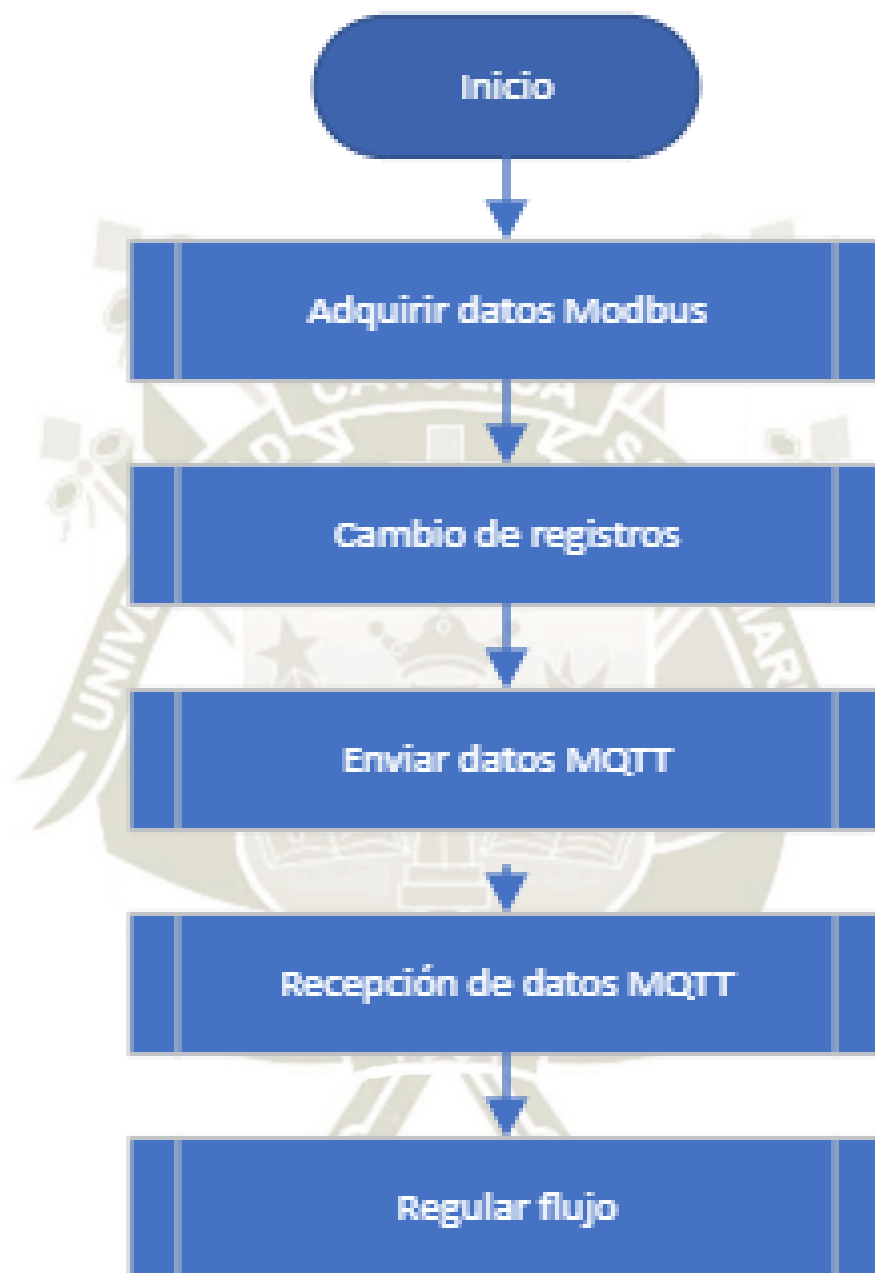
### 1.1.5. Subrutina procesamiento por machine learning.



### 1.1.6. Subrutina envío de datos.



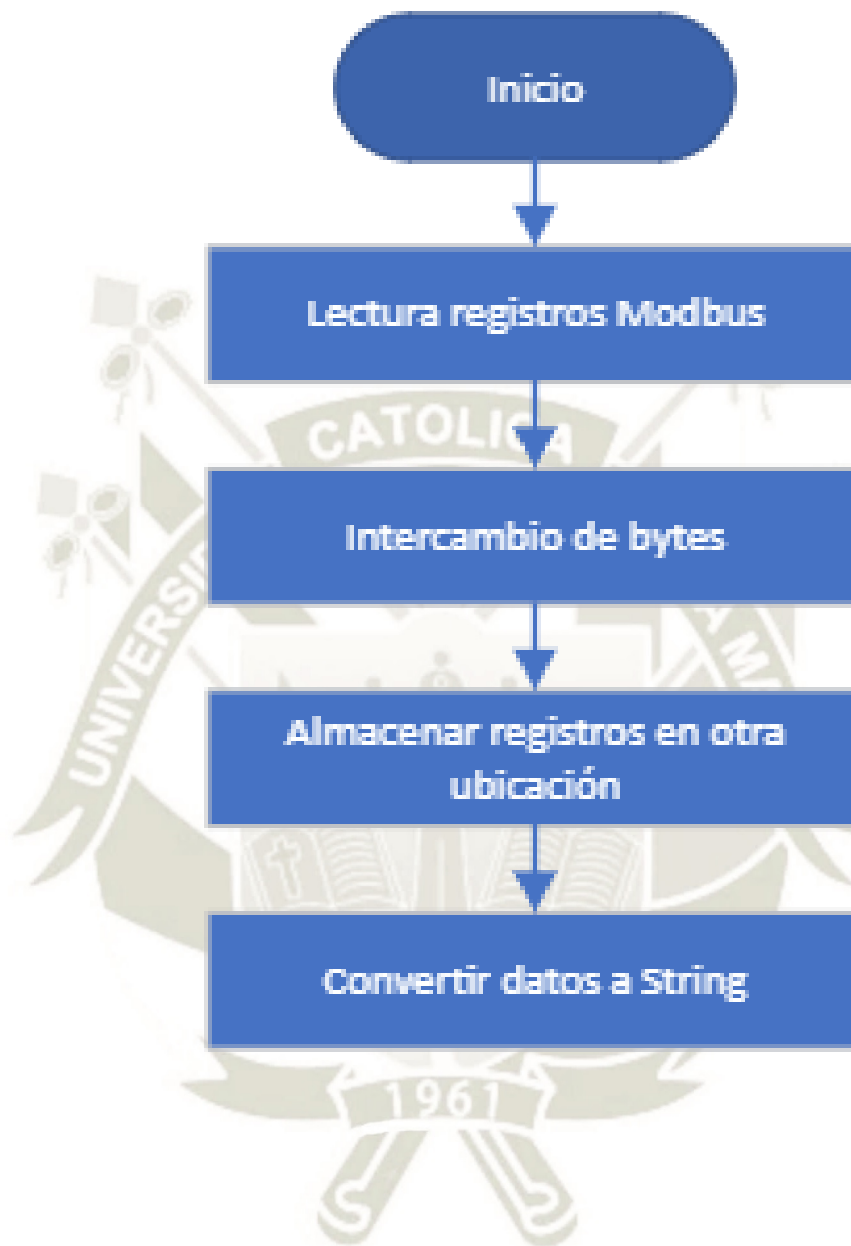
### 1.1.7. Diagrama de flujo principal usado en PLC S7-1200



### 1.1.8. Subrutina adquisición de datos Modbus.



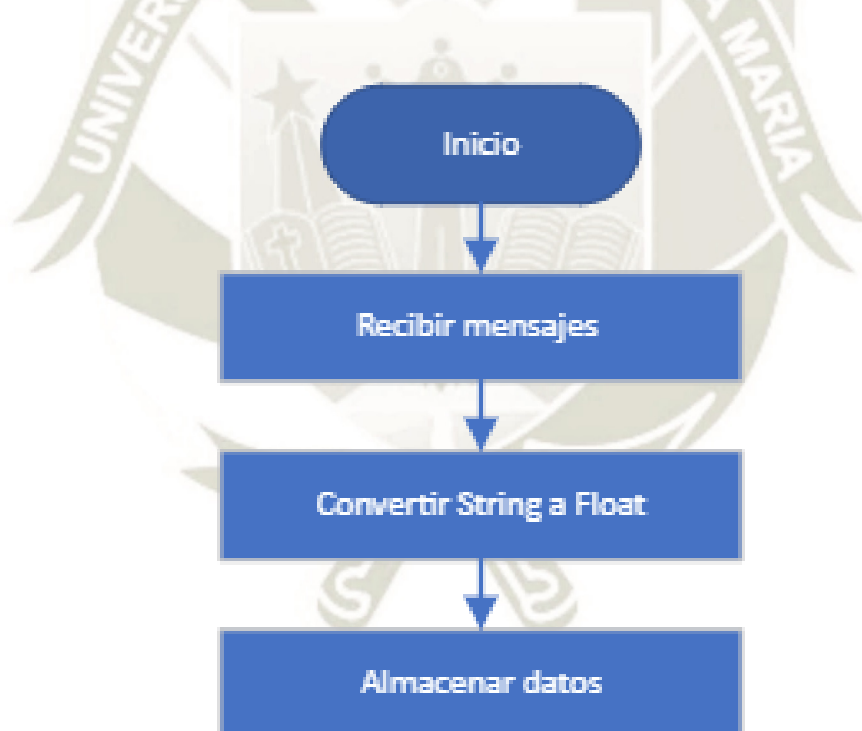
### 1.1.9. Subrutina cambio de registros.



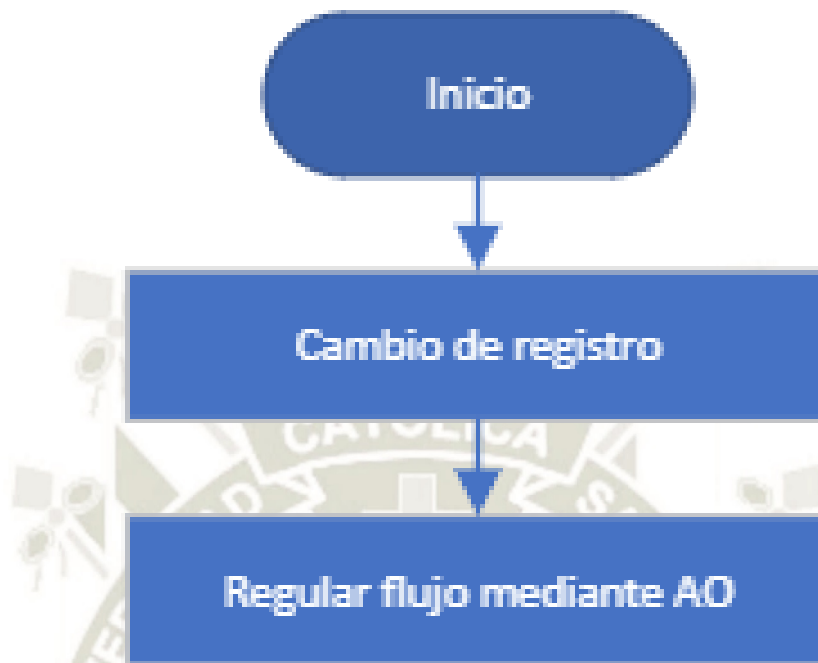
### 1.1.10. Subrutina envío de datos MQTT.



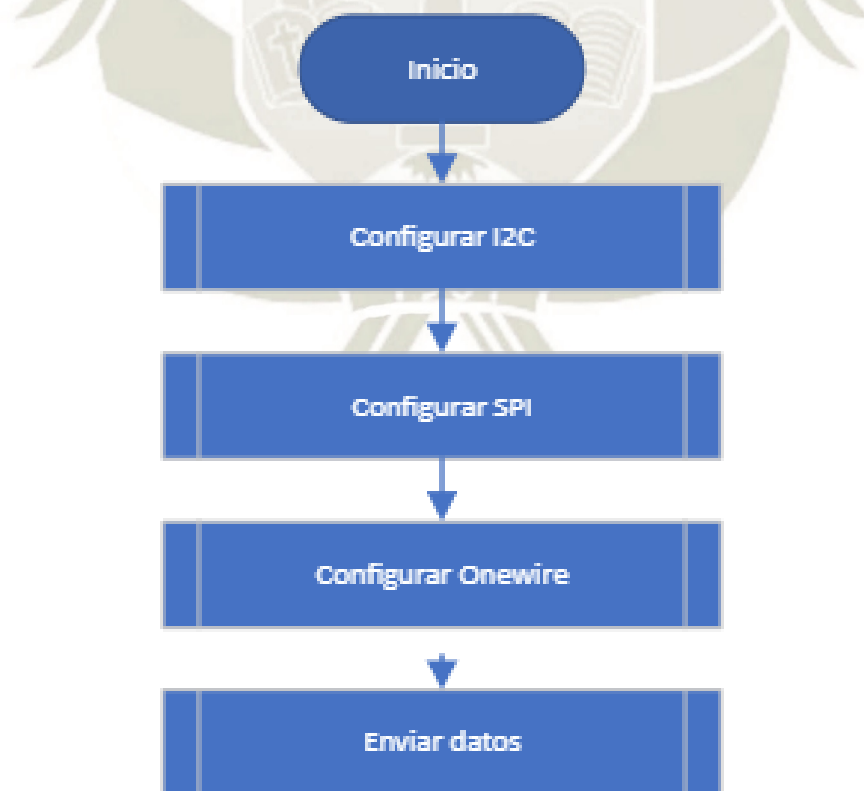
### 1.1.11. Subrutina recepción de datos MQTT.



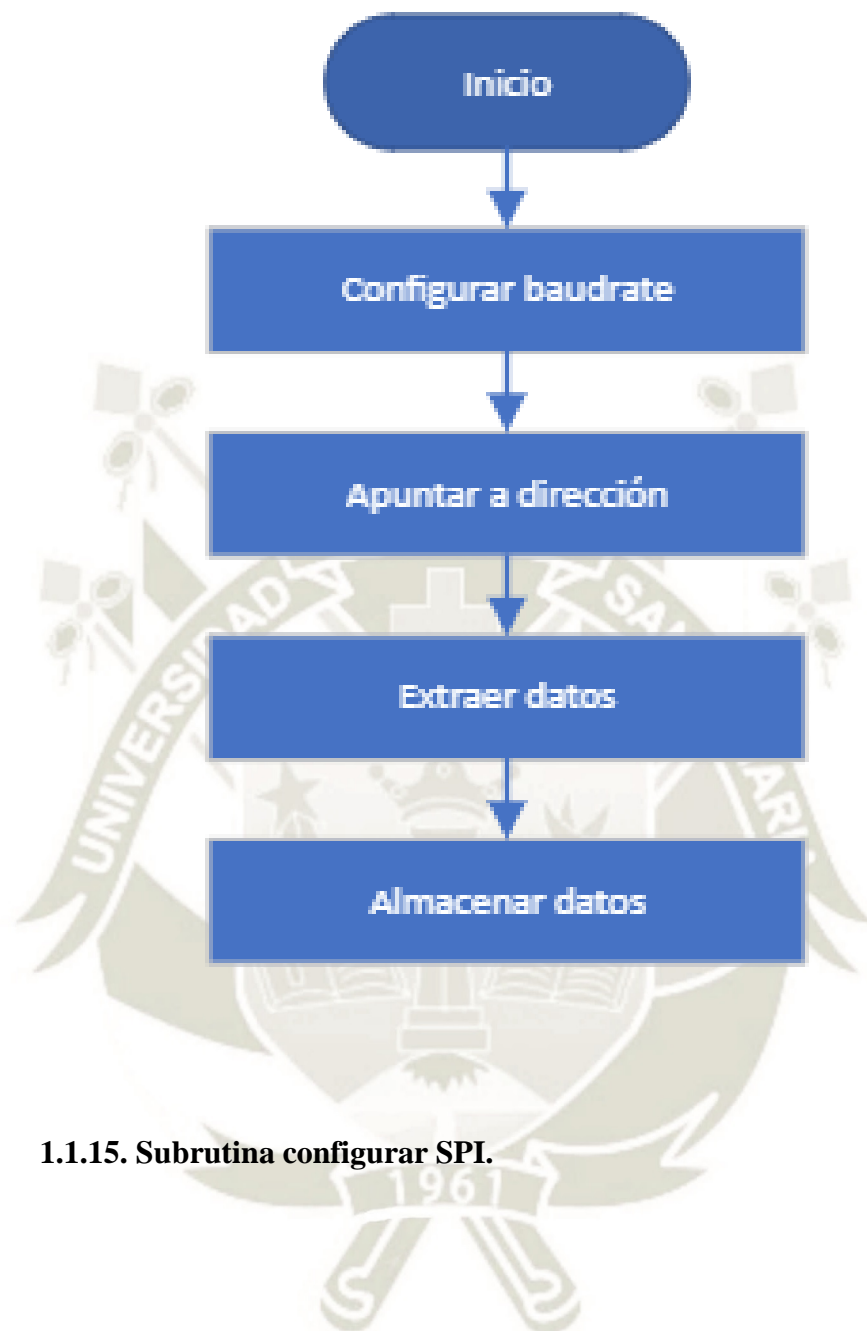
### 1.1.12. Subrutina regulación de flujo.



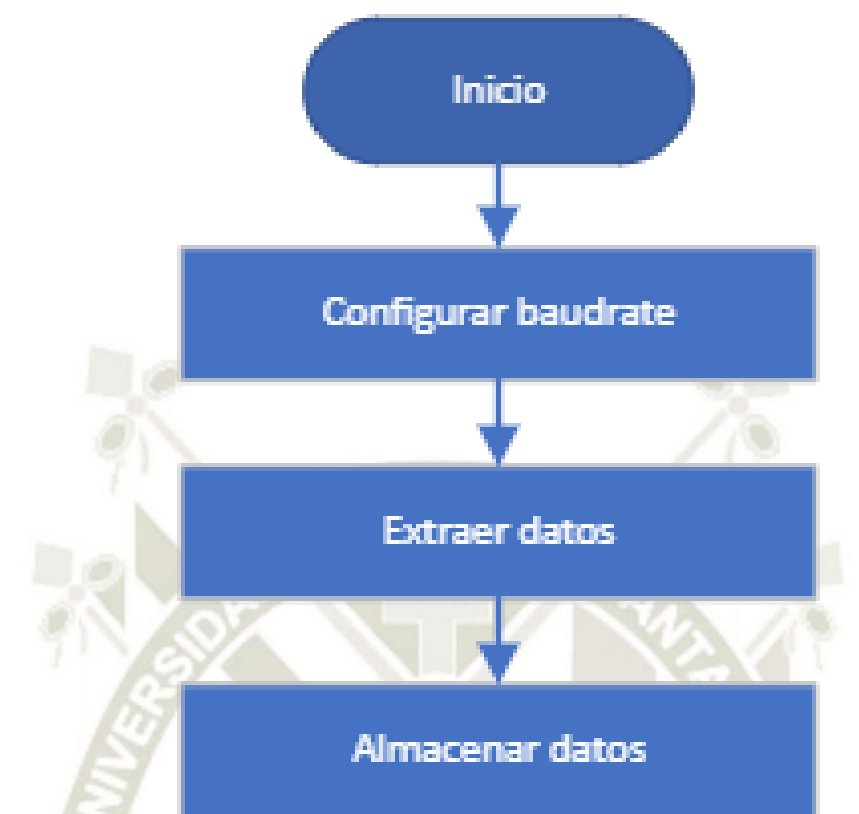
### 1.1.13. Diagrama de flujo principal usado en ESP8266



### 1.1.14. Subrutina configurar I2C.



**1.1.15. Subrutina configurar SPI.**



**1.1.16. Subrutina configurar Onewire.**



### 1.1.17. Subrutina enviar datos.

