

Adaptación del estándar ISO/IEC 25012 en el diseño de una plataforma de gestión de los datos generados en una red AMI



Wilson Geovanny Carvajal Molina
Edwin Alexander Marulanda Gavilán

Director: Ing. Pablo Augusto Magé Imbachí

Universidad del Cauca
Facultad de Ingeniería Electrónica y Telecomunicaciones
Departamento de Sistemas
Popayán, Julio de 2021

Agradecimientos

Para el desarrollo del siguiente trabajo expresamos nuestro mayor agradecimiento en primer lugar a Dios todo poderoso por el gran don de la vida y con ella, sus retos cotidianos.

A nuestros padres: Hilda de Socorro Molina- Luis Alberto Carvajal y Luz Ena Gavilán Reina - Gabriel Marulanda Aguirre (+), quienes han colocado todo su esfuerzo para llegar y perseverar en el *alma mater*. A mi familia: cada una de nuestras esposas: Deisy Cristina Piñeros y M^a del Rosario Aguirre, -gracias por ser la luz de nuestro camino, jamás encontraré la forma de agradecer todo el apoyo constante y la compañía incondicional brindada a lo largo de este proceso de formación-. A nuestros hijos: Samuel José -M^a Juliana y Nicolás Marulanda; quienes han sido el regalo más grande en este camino de vivir. Haciéndose en cada uno de nosotros y a su tiempo, quienes motivan el alcance de este título profesional y nuestra constante superación personal.

A nuestros compañeros, nuestros amigos y compañeros de viaje, por tantas horas de trabajo que nos juntamos a lo largo de nuestra formación; no podemos dejar de agradecerles por su apoyo y constancia, al estar en las horas más difíciles, por compartir horas de estudio.

Con quienes además compartimos anécdotas memorables y con quienes fuimos forjando nuestra personalidad, carácter y disciplina que requiere esta vida y nuestra carrera académica. Con quienes se compartieron momentos que en definitiva marcaron nuestra historia de vida; y con quienes esperamos compartir más y valiosas experiencias de sus vidas personal o laboral, entre otras.

A nuestro director, Pablo Augusto Magé Imbachí. Ya que, sin usted y sus virtudes; paciencia y constancia en este trabajo, no lo hubiésemos logrado. Su orientación fue siempre útil para organizar nuestras ideas al escribir lo que hoy hemos logrado. Usted formó parte importante en esta investigación con sus aportes profesionales característicos; muchas gracias por todo su esfuerzo y dedicación.

A nuestros docentes, quienes contribuyeron semestre a semestre a nuestra formación profesional, Pues detrás de sus aportes enmascarados en notas calificativas se escondieron los procesos de mejora que hoy, nos permiten ser competitivos en un entorno laboral. Gracias por sus instrucciones rigurosos y precisas, les debemos nuestros conocimientos. Donde quiera que vayamos, los llevaremos en nuestro desarrollo profesional. Gracias por su paciencia, por compartir sus conocimientos de manera profesional e invaluable.

De igual manera agradecemos inmensamente, a todos los héroes que sin reconocimiento lograron sostenernos y que fuera posible nuestro día a día, los que no pusieron obstáculos cuando no había dinero para comer, para vivir, los que sin recelo confiaban en nuestra honradez, y los que sin palabras lograron identificar que los necesitábamos, aquellos les quedamos alta mente agradecidos toda la vida, y esperamos en cualquier momento poder compensar su humildad, bondad y misericordia para con nosotros.

Resumen

Este trabajo fue desarrollado con finalidad de generar un estándar de calidad de datos para un contexto eléctrico a partir de la adaptación de la norma ISO 25012, para esto, se generó una caracterización de las unidades de medida y los equipos empleados para su medición. Identificando así el funcionamiento operativo de las compañías eléctricas. Posteriormente, se registraron los datos relevantes en bases de datos por medio del reconocimiento de la información suministrada en visita de campo, entes de control y facturas de empresas prestadoras del servicio eléctrico, con estos datos se construyó una serie de matrices que permiten definir las características relevantes al momento de verificar la calidad de datos dentro del contexto eléctrico tomando como referencia el proceso de la norma ISO 25040. Posteriormente fueron evaluadas en un caso de estudio que permitió verificar la aplicación de los estándares mínimos planteados, concibiendo así una guía que permita simplificar la verificación de la calidad de datos en el contexto mencionado.

Palabras Clave

AMI; Base de datos; Calidad de datos; Medición eléctrica; PLC; Smart grids.

Abstract

This project was developed with the purpose of generating a data quality standard for an electrical context based on the adaptation of ISO 25012, for this purpose, a characterization of the measurement units and the equipment used for their measurement was generated. Thus, the operational functioning of the electric companies was identified. Subsequently, relevant data were recorded in databases by means of the recognition of the information provided in field visits, control entities and invoices of electric service providers. With these data, a series of matrices were built to define the relevant characteristics at the time of verifying the quality of data within the electric context, taking as a reference the ISO 25040 standard process. Afterwards, they were evaluated in a case study that allowed verifying the application of the minimum standards proposed, thus conceiving a guide that allows simplifying the verification of data quality in the mentioned context.

Keywords

AMI; Database; Data quality; electrical metering; PLC; Smart grids.

Tabla de contenido

Capítulo 1. Problema y objetivos.	1
1.1. Planteamiento del problema.....	1
1.2. Justificación.....	2
1.3. Pregunta de investigación.....	3
1.4. Objetivos	3
1.4.1. General	3
1.4.2. Específicos.....	3
1.5. Metodología	3
1.6. Estructura de la monografía	4
Capítulo 2. Contexto teórico.	6
2.1. Marco Teórico	6
2.1.1. AMI (Advanced Metering Infrastructure)	6
2.1.2. Amarre	7
2.1.3. SAA (Sistema de Amarre Automático).....	7
2.1.3.1. PLC_MMS	11
2.1.3.2. PLC_MC.....	11
2.1.3.3. PLC_TU.....	11
2.1.4. Balance de energía.....	11
2.1.5. Transformador de distribución	12
2.1.6. Modelo C4 de documentación para la Arquitectura de Software	13
2.1.6.1. Nivel 1: El diagrama de contexto del sistema	13
2.1.6.2. Nivel 2: El diagrama del contenedor.....	14
2.1.6.3. Nivel 3: El diagrama de componentes.....	14
2.1.6.4. Nivel 4: El código.....	15
2.1.7. Framework.....	17
2.1.8. Calidad de dato.....	17
2.1.9. AMI_SAA	17
2.1.10. NORMA ISO/IEC 25012	18
2.2. Estado del Arte.....	20
2.2.1. Revisión sistemática	20
2.2.1.1. Fuentes de búsqueda.....	21
2.2.1.1.1. Google Scholar	21
2.2.1.1.2. IEEE Xplore Digital Library	21

2.2.1.2. Estrategia de búsqueda.....	21
2.1.2. Documentos	22
2.1.2.1. Una Metodología Basada en ISO/IEC 15939 para la Elaboración de Planes de Medición de Calidad de Datos.	22
2.1.2.2. Framework de evaluación de calidad de la información: Ampliación del modelo de calidad de datos ISO 25012.	24
2.1.2.3. Un framework dirigido por modelos para la gestión de calidad de datos en el internet de las cosas.	24
2.1.2.4. Comparison of Data-Driven Models for Cleaning eHealth Sensor Data: Use Case on ECG Signal.	27
2.1.2.5. Hacia un framework de evaluación de calidad de información en foros de discusión técnicos.	28
Capítulo 3. Caracterización de los tipos de medida en un contexto eléctrico.	32
3.1. Características físicas de los medidores.....	32
3.1.1. Según su construcción.....	32
3.1.1.1. Mecánicos o de inducción	32
3.1.1.2. Estáticos o electrónicos.....	33
3.1.2. Según energía medida.....	34
3.1.2.1. Energía Activa	34
3.1.2.2. Energía Reactiva	34
3.1.3. Según su conexión en red	34
3.1.3.1. Monofásicos	34
3.1.3.1.1. Bifilares	34
3.1.3.1.2. Trifilares	34
3.1.3.2. Bifásicos.....	34
3.1.3.2.1. Trifilares	35
3.1.3.3. Trifásicos.....	35
3.1.3.3.1. Tetrafilares	35
3.1.3.3.2. Trifilares	35
3.1.4. Según la conexión de los filamentos.....	36
3.1.5. Según la complejidad de su construcción.....	36
3.1.6. Según la exactitud en su medida.....	37
3.2. Características técnicas de los Tipos de medida	37
3.2.1. Medición indirecta:	37
3.2.2. Medición semidirecta:	38
3.2.3. Medición directa.....	39

Capítulo 4. Calidad de datos en un contexto eléctrico.....	41
4.1. Visita de campo, Acompañamiento en el proceso de facturación de la compañía energética de occidente.	44
4.2. Datos en los entes reguladores del servicio público eléctrico.	46
4.3. Facturas de energía eléctrica.....	48
4.4. Selección de características de calidad de datos.....	57
4.4.1. Establecer los requisitos de la evaluación	60
4.5. Estrategia de evaluación.....	63
4.5.1. Especificar la evaluación	63
4.5.2. Diseño de la evaluación.....	65
Capítulo 5. Aplicación de la estrategia de evaluación de calidad de datos en un contexto eléctrico en la plataforma AMI-SAA.	66
5.1. Diseño de estudio de caso	66
5.2. Arquitectura plataforma AMI-SAA	67
5.3. Implementación de la estrategia de evaluación en la base de datos de la plataforma AMI-SAA.	71
5.3.1. Especificación de la evaluación	71
5.3.2. Diseño de la evaluación	73
5.3.3. Ejecución de la evaluación	75
5.3.4. Resultados de la evaluación	76
5.3.5. Conclusión de la evaluación	78
Capítulo 6. Conclusiones y recomendaciones.....	79
6.1. Conclusiones.....	79
6.2. Lecciones aprendidas	79
6.3. Trabajos futuros	80
Capítulo 7: Referencias.....	81

Índice de figuras

Figura 1 Cobertura de un sistema AMI Fuente [8].....	7
Figura 2 Vinculación de un usuario monofásico antes de un daño o incidente en la red Fuente [10].....	8
Figura 3 Vinculación de un usuario monofásico después de un daño o incidente en la red Fuente [10].....	9
Figura 4 Vinculación de un usuario bifásico antes de un daño o incidente en la red Fuente [10].....	9
Figura 5 Vinculación de un usuario bifásico después de un daño o incidente en la red Fuente [10].....	10
Figura 6 Transformador de distribución Fuente [19].....	12
Figura 7 Nivel C1 Fuente Elaboración propia.....	13
Figura 8 Nivel C2 Fuente Elaboración Propia.....	14
Figura 9 Nivel C3 Fuente Elaboración propia.....	15
Figura 10 Nivel C4 Fuente Elaboración Propia.....	16
Figura 11 Vista principal de la aplicación AMI SAA Fuente Elaboración propia....	18
Figura 12 Calidad de producto de datos Fuente [23].....	20
Figura 13 Actividad 1: Establecer y Mantener el Compromiso en la Medición de la CD Fuente [24].....	23
Figura 14 Actividad 2: Elaborar el Plan de Medición Fuente [24].....	23
Figura 15 Workflow de la propuesta. Fuente elaboración propia.....	25
Figura 16 Estructura del proyecto Fuente [26].....	26
Figura 17 Oracle Event Processing Fuente [26].....	26
Figura 18 Workflow para la limpieza de datos. Fuente elaboración propia.....	27
Figura 19 Diagrama de flujo búsqueda por técnico Fuente [28].....	29
Figura 20 Diagrama de la información contenida en un foro de discusión técnico Fuente [28].....	30
Figura 21 Relación características del foro con características de calidad Fuente [28].....	30
Figura 22 Medidor de inducción. Fuente [32].....	33
Figura 23 Medidor estático. Fuente (Blanco, 2011).....	33
Figura 24 Conexión en red. Fuente elaboración propia.....	35
Figura 25 Conexión de medidor monofásico bifilar simétrico. Fuente [36].....	36
Figura 26 Conexión de medidor monofásico bifilar asimétrico. Fuente [36].....	36
Figura 27 Medidor de energía indirecta en poste. Fuente elaboración propia.....	38
Figura 28 Medidor de energía Semi-directa. Fuente elaboración propia.....	39
Figura 29 Medidor de energía directa. Fuente elaboración propia.....	40
Figura 30 Interfaz de registro de lectura usado por la CEO. Fuente elaboración propia anexo B.....	45
Figura 31 Esquema factura ENEL pagina 1. Fuente elaboración propia.....	48
Figura 32 Esquema de factura ENEL pagina 2. Fuente elaboración propia.....	49
Figura 33 Esquema factura CEO. Fuente elaboración propia.....	50
Figura 34 Esquema factura Air-e Pag 1. Fuente elaboración propia.....	51

Figura 35 Esquema factura Air-e Pag 2. Fuente elaboración propia.....	52
Figura 36 Esquema factura EPM. Fuente elaboración propia.....	53
Figura 37 Esquema factura ElectroHuila Pagina 1. Fuente elaboración propia. ...	54
Figura 38 Esquema factura ElectroHuila Pagina 2. Fuente elaboración propia. ...	55
Figura 39 Esquema factura de Electro Caquetá. Fuente elaboración propia.	56
Figura 40 Nivel C1 Plataforma AMI-SAA. Fuente elaboración propia.....	67
Figura 41 Nivel C2 Plataforma AMI-SAA. Fuente elaboración propia.....	68
Figura 42 Nivel C3 Plataforma AMI-SAA. Fuente elaboración propia.....	68
Figura 43 Nivel C3 Plataforma AMI-SAA (Gestión consumo de energía). Fuente elaboración propia.....	69
Figura 44 Base de datos general AMI-SAA. Fuente elaboración propia.....	70
Figura 45 Datos relacionados a nivel 2 PLC-MMS a PLC-TU. Fuente elaboración propia.....	71

Índice de Tablas

Tabla 1 Clasificación de medidores según la complejidad de su construcción. Fuente (Aguirre & Wilkar, 2021).....	36
Tabla 2 Selección de los medidores de energía. Fuente (Aguirre & Wilkar, 2021) 40	40
Tabla 3 Características inherentes y dependientes (Caro, Fuentes, & Soto, 2013).....	43
Tabla 4 Datos visita de campo, fuente elaboración propia.....	45
Tabla 5 Formato ZNI C6 Fuente (SUPERSERVICIOS, 2014).....	46
Tabla 6 Tabla O3 Facturación de energía. Fuente (SUPERSERVICIOS, 2014)...	47
Tabla 7 Datos antes de control. Fuente Elaboración propia.....	47
Tabla 8 Datos de facturas. Fuente Elaboración propia.....	57
Tabla 9 Comparación final de atributos de datos por cada fuente. Fuente Elaboración propia.	58
Tabla 10 Datos críticos. Fuente elaboración propia.....	59
Tabla 11 Datos secundarios. Fuente elaboración propia.....	59
Tabla 12 Características VS Actores. Fuente elaboración propia.....	62
Tabla 13 Características por nivel de relevancia.....	63
Tabla 14 Tabla ejemplo paso 1. Fuente elaboración propia.....	64
Tabla 15 Tabla ejemplo paso 2. Fuente elaboración propia.....	64
Tabla 16 Tabla ejemplo paso 3. Fuente elaboración propia.....	65
Tabla 17 Tabla ejemplo paso 4. Fuente elaboración propia.....	65
Tabla 18 selección de características por atributo tabla cliente. Elaboración propia.....	71
Tabla 19 selección de características por atributo tabla producto. Elaboración propia.....	72
Tabla 20 selección de características por atributo tabla evento consumo. Elaboración propia.....	72
Tabla 21 Criterios de evaluación Exactitud semántica. Elaboración propia.....	72
Tabla 22 Criterios de evaluación Exactitud sintáctica. Elaboración propia.....	72
Tabla 23 Criterios de evaluación Completitud. Elaboración propia.....	73

Tabla 24 Criterios de evaluación consistencia. Elaboración propia.....	73
Tabla 25 Criterios de evaluación credibilidad. Elaboración propia	73
Tabla 26 Criterios de evaluación credibilidad. Elaboración propia	73
Tabla 27 Reglas Exactitud semántica. Elaboración propia	74
Tabla 28 Reglas exactitud sintáctica. Elaboración Propia.....	74
Tabla 29 Reglas Completitud. Elaboración propia	74
Tabla 30 Reglas Consistencia. Elaboración Propia.....	74
Tabla 31 Reglas Credibilidad. Elaboración propia.....	75
Tabla 32 Reglas Actualidad. Elaboración propia.....	75
Tabla 33 Evaluación exactitud semántica. elaboración propia.....	75
Tabla 34 Evaluación exactitud sintáctica. elaboración propia	75
Tabla 35 Evaluación completitud. elaboración propia	76
Tabla 36 Evaluación Consistencia. elaboración propia.....	76
Tabla 37 Resultado evaluación exactitud semántica. elaboración propia	76
Tabla 38 Resultado evaluación exactitud sintáctica. elaboración propia.....	77
Tabla 39 Resultado evaluación completitud. elaboración propia.....	77
Tabla 40 Resultado evaluación consistencia. elaboración propia	77

Capítulo 1. Problema y objetivos.

1.1. Planteamiento del problema.

A la fecha en nuestra región, la lectura del consumo de energía se realiza de forma manual por medio de visitas mes a mes a cada una de las viviendas. Estos datos, proporcionados por el visitador, además de generar costos adicionales tanto para las compañías de energía eléctrica, como para los usuarios. No representan una medición totalmente confiable debido a los múltiples errores y obstáculos que pueden surgir al momento de lectura y digitación del consumo. Adicionalmente, al no representar una medida en tiempo real, imposibilita el control y vigilancia permanente de fraudes y conexiones piratas. Dichos fraudes de fluidos eléctricos generan pérdidas por alrededor de 250 mil millones de pesos por año a las compañías de energía eléctrica [1].

Con el fin de tener un control más eficiente del servicio, las empresas se ven en la necesidad de implementar nuevas tecnologías que le permitan mejorar su servicio, mantener control y vigilancia constante sobre su red y posibilita la toma de datos de consumos eléctricos en tiempo real. Es allí donde las redes inteligentes hacen su contribución junto con la tecnología AMI (Advanced Metering Infrastructure) [2] estas Smart Grids¹ constituidas por medidores digitales avanzados con dos vías de comunicación [3] monitorea el consumo de energía, al mismo tiempo que se comunica entre sí y realiza control para optimizar el uso de energía.

AMI implementa sistemas de gestión para almacenar los datos de medición y control de procesos, generando beneficios financieros y permitiendo mejoras en los servicios. Dada la importancia que han tomado los datos en las empresas para realizar sus operaciones se ha hecho énfasis en su calidad, entendiendo esta como el grado en que los datos satisfacen los requisitos definidos por cada empresa, logrando instituir transparencia y confiabilidad.

Dada la criticidad de la información, se han creado normas que permiten asegurar la calidad y confiabilidad de los datos y del software en las empresas desarrolladoras de productos software, dentro de ellas se cuenta con el estándar ISO/IEC 25012:2008 [4], en el cual se presenta un modelo de calidad de datos que establece las características que se deben tener en cuenta al evaluar un producto de datos. Sin embargo, estos estándares constituyen una generalidad de calidad de datos y no existe ninguna adaptación u normativa específica que permita determinar los criterios necesarios para precisar la calidad de los datos tomados

¹ Las Smart Grids o redes inteligentes pretenden optimizar la red de distribución de energía eléctrica de forma que se haga un uso eficiente y sostenible de este recurso.

en una Smart Grid a través de tecnologías AMI y mucho menos, una que se refiera a la calidad de datos en un contexto eléctrico.

1.2. Justificación

El Grupo de Automática industrial (GAI), perteneciente a la Universidad del Cauca en convenio con la Compañía Energética de Occidente (CEO) realizó el proyecto denominado SAA el cual tenía como objetivo general diseñar, implementar y validar un prototipo de sistema de amarre automático, entre un transformador de distribución de energía y sus usuarios. Posteriormente GAI y CEO junto al Grupo de Investigación y Desarrollo en Ingeniería del Software (IDIS) desarrollan el proyecto AMI_SAA en el cual se pretende con ayuda de la tecnología AMI dar solución a muchas de las problemáticas que se presentan en las redes eléctricas relacionadas con la calidad de servicios. Esta tecnología permite obtener mediciones de energía² de una forma más precisa y eficiente, también obtiene la información del consumo de energía eléctrica en tiempo real y simultáneamente puede ser transmitido a través de un sistema de comunicación a las empresas prestadoras del servicio. También proporciona a las empresas la capacidad de detectar problemas en sus sistemas y así gestionar la demanda, con el fin de operar más eficientemente [5]. Hasta el momento se está implementando un prototipo software que permite monitorear y acceder a los datos que se generan en una red AMI_SAA que se basa en tecnologías de balances energéticos en transformadores de distribución, pero no cuenta con un proceso que permita asegurar que los datos que se reciben cumplan con la calidad necesaria para ser utilizados en la empresa.

La implementación de lectura remota del consumo de energía es un gran avance para la empresa prestadora del servicio, dados todos los beneficios ya mencionados lo que permite una mejor toma de decisiones y gestión en la prestación del servicio eléctrico, sin embargo, la información que brinda debe contar con los más altos estándares de calidad [3] y [6], puesto que una alteración en los datos puede representar pérdidas considerables para la empresa [7]. Es por esto y con base a la problemática anteriormente expuesta y partiendo de que en Colombia ya se comenzaron a desarrollar los esquemas necesarios para la incorporación de las Smart Grids en los sistemas de suministro de energía eléctrica, que surge la necesidad de adaptar los estándares generales de calidad de datos para establecer los criterios mínimos necesarios para determinar la calidad de los datos de las redes AMI en sistemas de suministro de energía eléctrica, lo que relacionado a los trabajos ya mencionados desarrollados por el GAI, la CEO y el IDIS dirigirá el desarrollo de esta investigación en la búsqueda de estándares que permitan la obtención de datos de calidad en las redes soportadas

² Mediciones de energía: Consumo de energía eléctrica.

con la tecnología AMI-SAA con la finalidad de contribuir a la mejora del servicio que brinda la compañía energética local CEO, y en general, a brindar un estándar para determinar la calidad de datos obtenidos de una red AMI.

1.3. Pregunta de investigación.

¿Cómo adaptar el estándar ISO/IEC 25012 con el fin de obtener datos de calidad que son medidos¹ en una red de energía eléctrica soportada con la tecnología AMI_SAA?

1.4. Objetivos

1.4.1. General

- Adaptar el estándar ISO/IEC 25012 para la evaluación de los datos medidos³ en una red AMI_SAA implementada en un contexto eléctrico, con el fin de determinar la calidad de estos.

1.4.2. Específicos

- Caracterizar los tipos de medidas⁴ que se generan en un contexto eléctrico⁷ con el fin de determinar la calidad de los datos que se manejan actualmente.
- Establecer las características relevantes de calidad de datos del estándar ISO/IEC 25012 con respecto al estado de la información⁵ para definir la estrategia de aplicación del estándar en un contexto eléctrico⁷.
- Evaluar mediante un estudio de caso la calidad de los datos medidos⁵ en un contexto eléctrico⁶ utilizando las estrategias definidas en el estándar utilizando un prototipo software que tiene como funcionalidad monitorear y acceder a los datos en una infraestructura de medición avanzada (AMI) basada en tecnología de balances energéticos en transformadores de distribución.

1.5. Metodología

El desarrollo de esta investigación surge como continuidad de los esfuerzos desarrollados por la universidad del Cauca, el grupo de investigación IDIS, la GAI

³ Datos medidos: Consumo de energía eléctrica.

⁴ Tipos de medidas: Consumo de energía eléctrica.

⁵ Estado de la información: Datos de medición a usar que de acuerdo con estos se seleccionaran las características de la norma ISO 25012 que pueden ser aplicadas.

⁶ Contexto eléctrico: Prototipo AMI_SAA convenio Compañía energética de occidente – CEO y Universidad del Cauca.

y en apoyo conjunto con la CEO. Esta continuidad consiste en la generación de una plataforma que pueda recopilar y almacenar los datos obtenidos del sistema de amarre SAA, y debido a esto surgió la necesidad de desarrollar una adaptación normativa que permitiera verificar la calidad de los datos que allí se almacenasen, y que en general, permitiera la construcción de un estándar en el contexto eléctrico. Para esto se realizaron las siguientes actividades:

- a. **Búsqueda bibliográfica:** Búsqueda de carácter normativo de fuentes secundarias, en la que se hallaron los estándares actuales en este tema. Tras analizar dicha información bibliográfica se pudo determinar a la ISO 25012.
- b. **Caracterización:** Caracterización de los tipos de energía y medidores asociados, que reforzarían teóricamente los conceptos del contexto eléctrico.
- c. **Recopilación de información:** Consistió en la visita de campo a la CEO, en la que se pudo observar el proceso de toma de muestras junto a un visitador y la recepción de esa información en la central, generando así, una fuente primaria de información junto a los datos contenidos en la paginas públicas de las entidad reguladora de servicios públicos y los datos contenidos en los recibos de energía de 6 compañías del país, con estos datos se pudieron generar los propósitos de evaluación desde la perspectiva de 3 actores, cliente, ente de control y entidad prestadora del servicio, que permita desarrollar una estrategia de evaluación de calidad de datos en el contexto eléctrico.
- d. **Ejecución de la evaluación:** permitió aplicar la estrategia de evaluación por medio de un estudio de caso.

1.6. Estructura de la monografía

El desarrollo de este documento se encuentra seccionado en 7 capítulos, descritos a continuación.

Capítulo 2. Contexto teórico: Presenta todos los conceptos de relevancia para este trabajo. En los que se incluyen datos relacionados con calidad de datos y el contexto eléctrico tales como tipos de medida, Infraestructura de Medición Avanzada entre otros. Igualmente se presenta investigaciones y artículos relacionados con la calidad de datos y el consumo eléctrico.

Capítulo 3. Caracterización de los tipos de medida en un contexto eléctrico: Presenta los atributos relacionados con los dispositivos empleados para las mediciones del servicio eléctrico, y la relación de estos dispositivos con los tipos de energía. Todo esto direccionado a generar un conocimiento general sobre el

funcionamiento de una red eléctrica, el cual concluye con la definición de los tipos de medida.

Capítulo 4. Calidad de datos en un contexto eléctrico: Expone un análisis de contenido presente en las facturas de 6 compañías prestadoras del servicio eléctrico y los resultados obtenidos en la visita de campo a la CEO, dicho análisis resulta en la identificación de 3 actores: cliente, entidad prestadora de servicio y entidad reguladora, a partir de estos actores y los datos que son de interés para cada actor se generan los propósitos de evaluación que permitieron la selección de las características relevantes de la norma ISO/IEC 25012, una vez definidas las características se define la estrategia de evaluación de calidad de datos de la plataforma AMI-SAA.

Capítulo 5. Aplicación de la estrategia de verificación de calidad de datos en un contexto eléctrico en la plataforma AMI-SAA: Exhibe la aplicación de la estrategia planteada en el capítulo anterior, para esto se toma como caso la plataforma AMI-SAA, inicialmente se muestra la arquitectura de dicha plataforma, posteriormente son relacionados los atributos de las tablas con los grupos de la estrategia, para finalmente evaluar las características y determinar la calidad de los datos confinados en la plataforma.

Capítulo 6. Conclusiones y Recomendaciones: En este capítulo se resume los resultados obtenidos en la construcción de este documento, y relacionados con los productos generados al evaluar la estrategia en la plataforma AMI-SAA, igualmente se generan las recomendaciones necesarias para generar una continuación a este trabajo.

Capítulo 7: Referencias: Incluye toda la bibliografía y cibergrafía empleada para la construcción de este documento.

Capítulo 2. Contexto teórico.

2.1. Marco Teórico

2.1.1. AMI (Advanced Metering Infrastructure)

La infraestructura de medición avanzada es un conjunto de tecnologías usadas por las empresas de servicios públicos que buscan tomar medidas de consumo de una forma avanzada que permita no sólo recolectar datos sino convertirlos en información que ayude al usuario a gestionar sus gastos través de un mayor conocimiento sino también a las empresas en la toma de decisiones. Para la aplicación de una AMI es necesario el uso de los **Medidores Inteligentes**. Según el Grupo de Industrias de Medidores Inteligentes Europeos (ESMIG) hay cuatro funcionalidades básicas que un medidor debe tener para ser considerado “inteligente” [8].

- Lectura remota de consumo.
- Comunicación en dos vías.
- Soporte para sistemas de pagos y tarifas avanzado.
- Capacidad para habilitar y deshabilitar el servicio de forma remota.

La lectura remota de consumo evita que los operarios vayan de forma física al sitio, evitando así que se produzcan errores humanos e inconvenientes climatológicos. Además, agiliza el proceso, ya que la información está disponible en todo momento para el prestador del servicio. La comunicación en dos vías va enlazada a las otras capacidades del medidor, ya que con esta se pueden tomar los datos de consumo y facilita que el servicio se habilite o inhabilite según sea la necesidad. El poder deshabilitar el servicio de forma remota apoya tanto a la empresa como al usuario económicamente debido a que los costos por desplazamiento se eliminarían y la operación se puede realizar rápidamente.

Adicional a los beneficios ya mencionados, esta tecnología permite (Según la implementación del AMI y los objetivos de la empresa) Adicionar, la detección de robos, medidores con funcionamientos errados, apagones y la detección de la ubicación exacta de un fallo [8].

Un sistema AMI opera a partir de la red de distribución y llega hasta el medidor del cliente final como se puede apreciar en la figura 1:

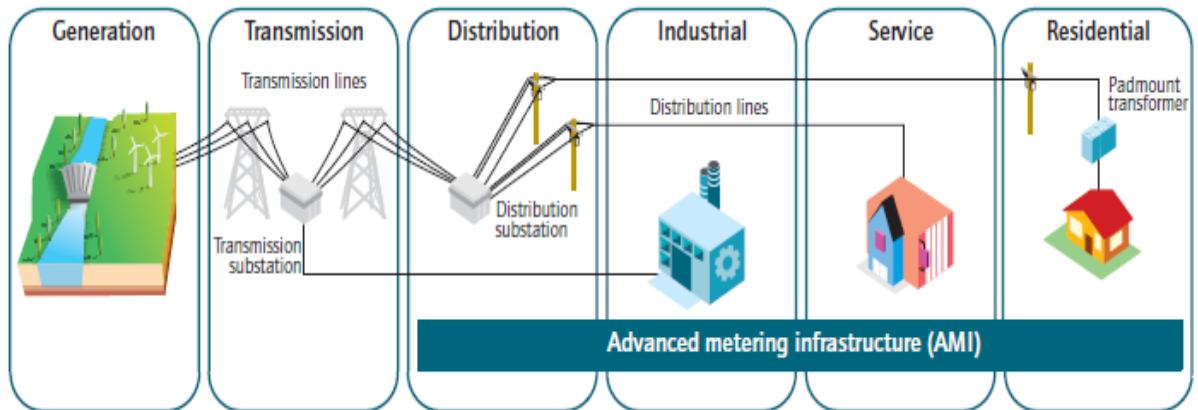


Figura 1 Cobertura de un sistema AMI Fuente [8]

2.1.2. Amarre

Son el resultado de unir 2 o más terminaciones de cable o alambre de manera que se conserve la continuidad eléctrica. Este amarre debe ser tan fuerte o más que el mismo cable o alambre utilizado pues deben mantenerse unidos por periodos muy largos de tiempo [9]. Para el contexto del proyecto se tiene en cuenta ya que representa la posibilidad de conectar o desconectar el servicio de energía eléctrica. También son conocidos como empalmes y existen de dos tipos principalmente:

- Amarres para unir dos conductores y así formar uno solo.
- Los que se usan para hacer derivaciones de y para otros conductores.

Al momento de elegir el tipo de amarre a implementar, se debe tener bajo consideración las condiciones climatológicas a las que van a estar expuestos los cables y la tensión eléctrica que van a transmitir. Entre otros aspectos para garantizar la durabilidad y calidad de los amarres.

2.1.3. SAA (Sistema de Amarre Automático)

Este sistema que permite la verificación y monitoreo de amarres entre un usuario y su respectivo transformador de forma que se potencializa la correcta realización de balances de energía en transformadores de distribución. SAA es el resultado de un proyecto conjunto de I+D de la Universidad del Cauca y la Compañía Energética de Occidente con patente No. 15-191386-0000-0000 denominado “Sistema de Amarre Automático”, el cual culminó en el desarrollo una tecnología compuesta por dos tipos de dispositivos electrónicos denominados respectivamente PLC_MMS (Power Line Communication - Main Monitoring

System) y PLC_TU (Power Line Communication - User Terminal). El sistema de amarre automático responde a la necesidad surgidas de los siguientes problemas:

- a) Crecimiento natural de la red por causas como el crecimiento urbanístico (Usuario monofásico): Este problema se refiere a la situación que se presenta cuando un cliente que se encontraba conectado a un transformador, como se aprecia en la figura 2, requiere ser conectado a otro transformador para aprovechar la disponibilidad o la cercanía de un nuevo transformador lo que provoca un error en el balance de energía realizado por la empresa prestadora del servicio de energía eléctrica. Esta misma situación ocurre cuando hay un daño en la red eléctrica y con el fin de brindarle una mayor disponibilidad al cliente se conectan los usuarios afectados a un transformador diferente durante el periodo que se realizan las reparaciones del transformador al que estaban vinculados como se aprecia en la figura 3 [10].

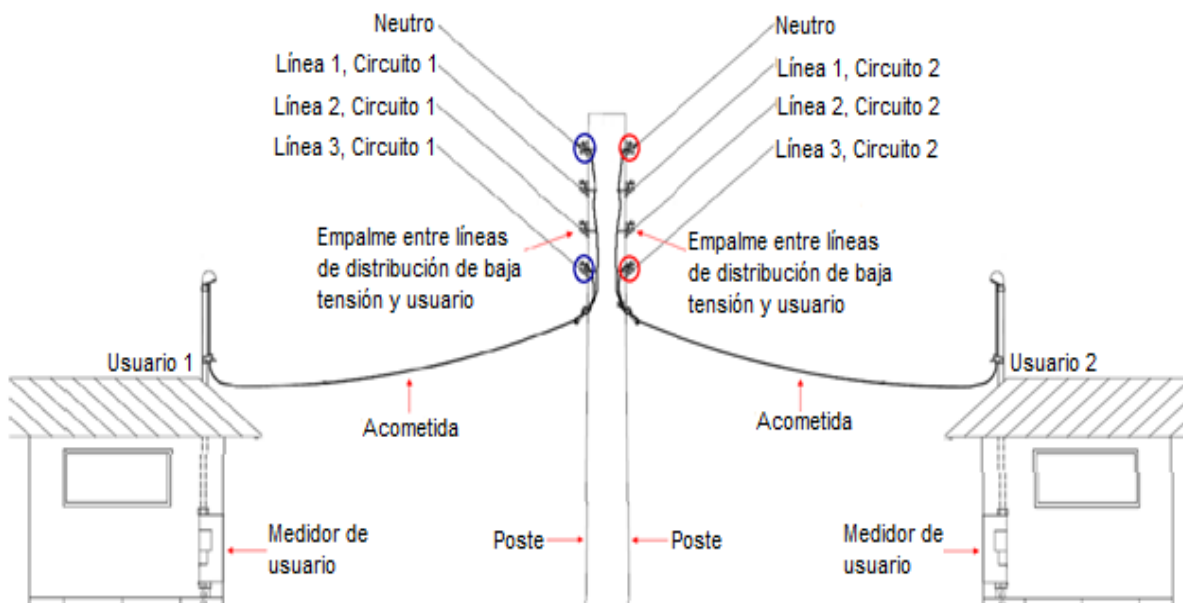


Figura 2 Vinculación de un usuario monofásico antes de un daño o incidente en la red Fuente [10]

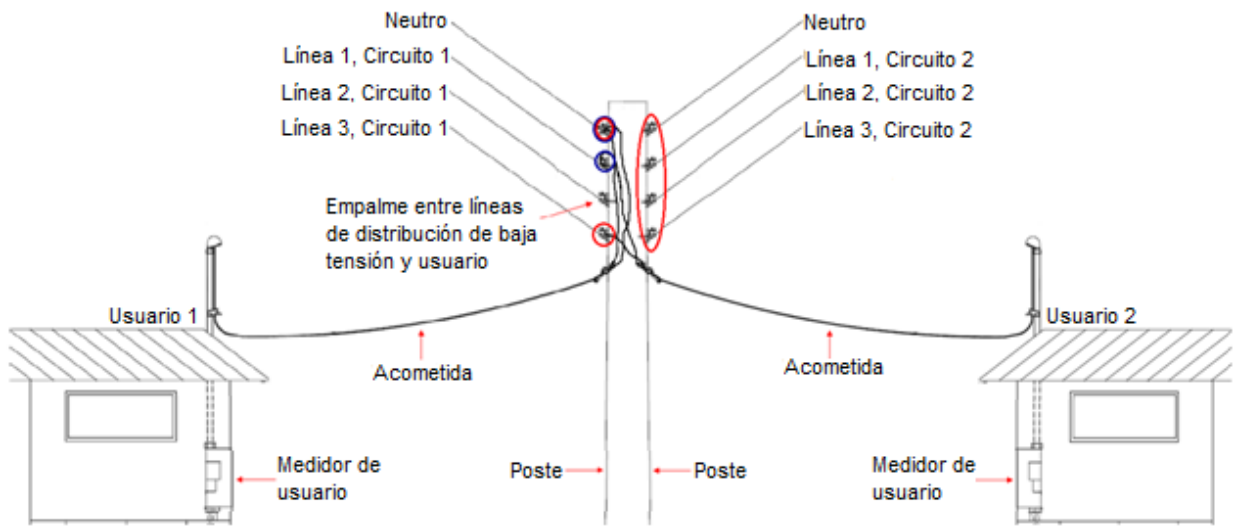


Figura 3 Vinculación de un usuario monofásico después de un daño o incidente en la red Fuente [10]

- b) cimiento natural de la red por causas como el crecimiento urbanístico (Usuario bifásico): Los problemas de amarre en los casos de usuarios bifásicos, se presentan de igual manera que los descritos en el apartado anterior usuario monofásico como se aprecia en la figura 4; con la variante que cuando las brigadas realizan revisiones o mantenimiento a los circuitos de baja tensión, por descuido pueden vincular una o las dos fases, a un circuito que no le corresponde como se aprecia en la figura 5, de modo que el usuario puede ser alimentado de dos circuitos diferentes [10].

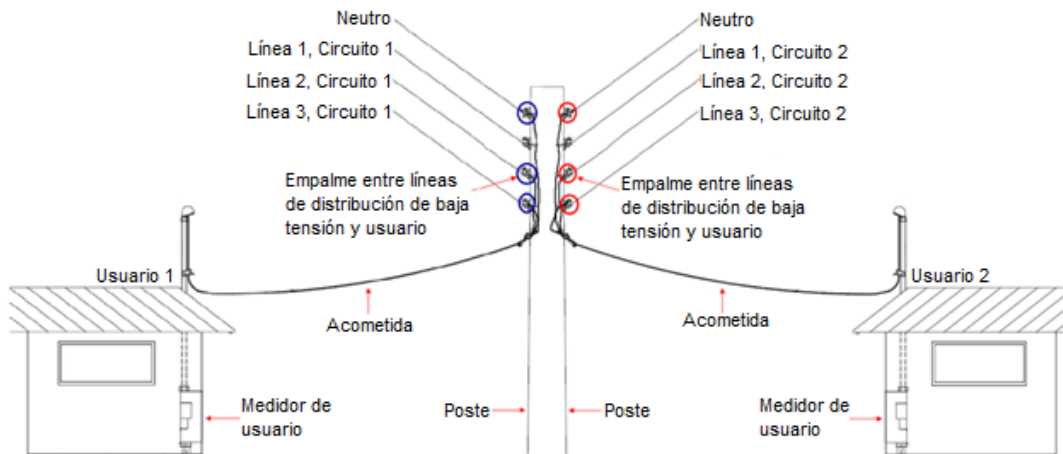


Figura 4 Vinculación de un usuario bifásico antes de un daño o incidente en la red Fuente [10]

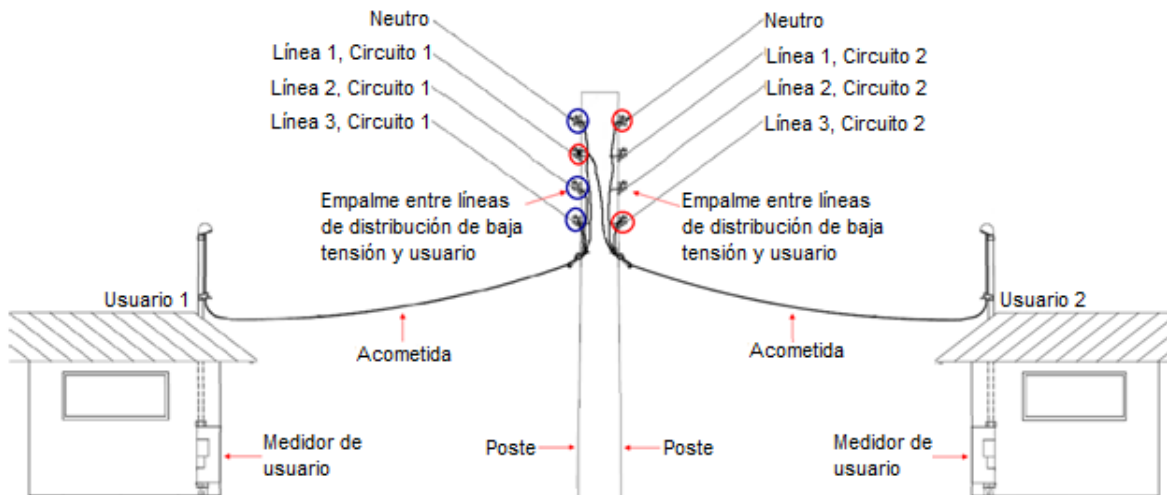


Figura 5 Vinculación de un usuario bifásico después de un daño o incidente en la red Fuente [10]

Estos problemas bajo circunstancias normales se detectan con algunos de los siguientes métodos:

- **Seguimiento visual:** Consiste en revisar la continuidad eléctrica entre dos puntos a través de la visualización de los cables en todo el tramo. Esto puede generar errores ya que la información levantada no es confiable y los balances obtenidos pueden arrojar datos inconsistentes [10].
- **Continuidad eléctrica:** se le asigna este nombre a la presencia de una ruta completa para el flujo de corriente. De esta manera, un interruptor cerrado que está en funcionamiento dentro de la ruta tiene continuidad. Una prueba de continuidad podría ser, un multímetro digital enviar una pequeña corriente por un circuito para medir la resistencia en el flujo. O a mayor escala, el uso de un cable de apoyo que puede ir desde el transformador hasta el nodo de un usuario, permitiendo de manera eficaz encontrar la ubicación a un problema de amarre eléctrico [10].
- **Envío de señales:** Estas hacen referencia a las diferencias o variaciones de tensión o corriente que hay entre dos puntos que poseen cargas eléctricas. Constan de valores codificados mediante los que se transmite información, la cual deberá servir en un circuito para realizar procesos electrónicos, de lo contrario no se le considerará señal [11]. De modo que el envío de una señal desde el transformador donde esté conectado un nodo usuario hasta el nodo usuario, deberá responder a un acuse de recibido para confirmar el estado del amarre eléctrico. El envío de esta señal se enfrenta a problemas como las distancias entre el usuario y el transformador, la posible contaminación de la señal con interferencias

externas, o cuestiones ambientales como por ejemplo los estados climatológicos, entre otros.

2.1.3.1. PLC_MMS

Es el Sistema de Monitoreo Principal, el cual está encargado del control, procesamiento y vigilancia de la información. Este se ocupa de entregar una vez por día [10]:

- a) La lista de medidores que enviaron su código de identificación.
- b) El valor del consumo de energía registrado por cada uno de los medidores.
- c) El valor del consumo de energía registrado por el macro medidor asociado a un transformador.

2.1.3.2. PLC_MC

Es el sistema de medida concentrada, es el único de los sistemas PLC que no está enlazado directamente a un medidor, y su única función es agrupar los datos de ID y Consumo de los PLC_TU asociados a él.

2.1.3.3. PLC_TU

Ubicados cerca o en el medidor de usuario son los responsables de responder a la encuesta de amarre del PLC_MMS [10] O al PLC_MC según corresponda.

Básicamente la tecnología PLC se refiere a la capacidad de transmitir datos a través de la red eléctrica. El PLC_MMS es un transformador con tecnología PLC que recibe la información de los PLC_TU que son como los medidores corrientes de consumo de energía, pero con tecnología PLC adicionada.

2.1.4. Balance de energía

En la transmisión de energía eléctrica no se logra que a los clientes finales llegue toda la energía producida ya que durante el proceso parte de ella se pierde a manera de calor que se disipa en los componentes de la red durante las etapas de generación, transmisión y distribución [12].

Es por esta razón que, para las empresas de servicio eléctrico se hace necesario emplear mecanismos que permitan reconocer y cuantificar pérdidas con el fin de elaborar estrategias para controlarlas y mitigarlas [13]. Para ello se hace uso de balances los cuales permiten a las empresas comercializadoras de energía eléctrica determinar las pérdidas totales a partir de una comparación entre la

cantidad de energía comprada a los diferentes agentes generadores y la energía facturada a sus clientes en un mismo periodo de tiempo [14].

Un balance de energía en un circuito de distribución de baja tensión se realiza a partir de la diferencia entre la energía entregada por un determinado transformador y la energía total que registran los medidores de los usuarios asociados al mismo. La diferencia entre estos valores representa las pérdidas totales de energía. Dichas pérdidas se clasifican como técnicas y no técnicas, siendo las primeras debidas a la disipación de energía por efecto Joule [15] y las segundas a errores de facturación, hurto de energía y asociaciones incorrectas entre el usuario y un transformador [16].

Cabe resaltar que la generación de balances de energía no permite diferenciar entre pérdidas técnicas y no técnicas, debido a que las primeras, en algunos países no se rigen bajo estándares internacionales como consecuencia de la arcaica infraestructura eléctrica en dichas regiones.

2.1.5. Transformador de distribución

Este dispositivo que se puede observar en la figura 6 son utilizados para distribuir energía eléctrica en media tensión. Se pueden aplicar en zonas industriales, urbanas; en la industria minera, petrolera o, incluso, en centros comerciales [17], las características de este tipo de transformadores son:

- Se fabrican en potencias normalizadas que van desde 25 hasta 1000 kVA.
- Se construyen en tensiones primarias de 13.2, 15, 25, 33 y 35 kV.
- Las variaciones de tensión se hacen con un conmutador exterior de accionamiento sin carga.
- Proveen en frecuencias de 50 a 60 Hz.



Figura 6 Transformador de distribución Fuente [19]

2.1.6. Modelo C4 de documentación para la Arquitectura de Software

Propuesto por Simon Brown como una solución para ayudar a los equipos de desarrollo a entender la arquitectura de un producto, se basa en el principio de ir de lo general a lo particular, haciendo zoom en cada elemento importante, el modelo propone 4 niveles de especificación [18]. En donde se busca describir la arquitectura de la aplicación de una forma mucho más sencilla. Este modelo se caracteriza por mostrar una serie de capas que, dependiendo del nivel de profundidad, definirán el nivel de detalle al que se lleve. El nivel que se muestra debe ser sólo el necesario, con el fin de no confundir a lectores que no tengan (dado el caso) el conocimiento técnico requerido para entender todo lo que aparece en el modelo. Sin embargo, si se necesita, se pueden mostrar detalles, en el caso de que el modelo vaya dirigido a personas con conocimientos e intereses en temas como bases de datos, código y pruebas de la aplicación. En general los niveles del C4 son los siguientes:

2.1.6.1. Nivel 1: El diagrama de contexto del sistema

Este nivel tiene como objetivo, visualizar el sistema desde un alto nivel (Generalidad), poniendo lo que se va a desarrollar en el centro y rodeado de todos los sistemas y usuarios que interactúan con él como se puede apreciar en la figura 7. En este nivel los detalles no son relevantes, al igual que los protocolos y los mecanismos de comunicación, pues busca centrarse netamente en las interacciones con el exterior [20].

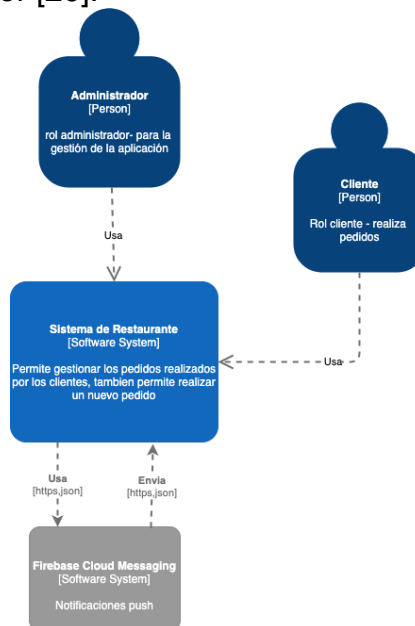


Figura 7 Nivel C1 Fuente Elaboración propia

2.1.6.2. Nivel 2: El diagrama del contenedor

Es una mirada introspectiva de lo que se va a desarrollar, siendo esto una ampliación del sistema a desarrollar, ilustrando los elementos que contienen y sus interacciones, pero de una manera general, por ejemplo, estos elementos podrían ser las bases de datos y sus interacciones, sistemas que exporten APIs y los elementos propios de soporte de la infraestructura como los gateways, esto se puede apreciar en la figura 8.

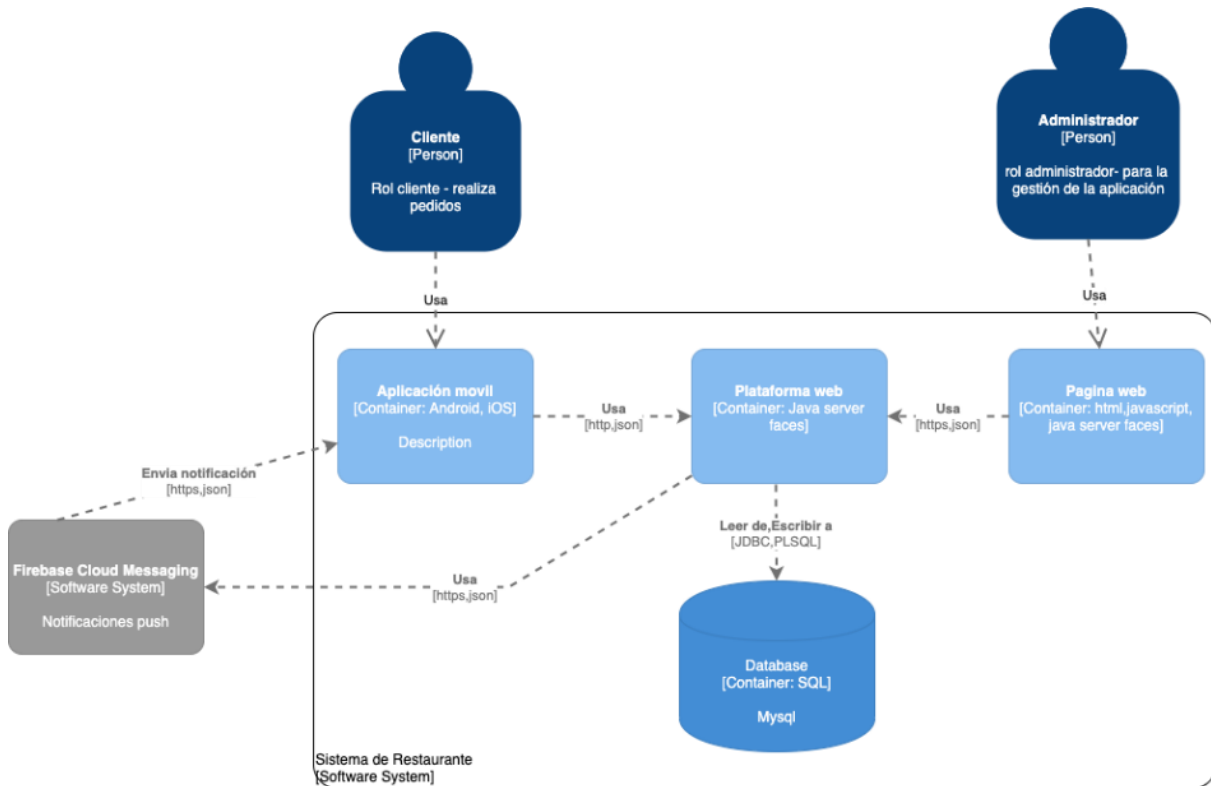


Figura 8 Nivel C2 Fuente Elaboración Propia

2.1.6.3. Nivel 3: El diagrama de componentes

En este nivel se describe en detalle un elemento en particular del sistema que se va a desarrollar. Sin incluir el detalle de componentes terciarios [20], como las bases de datos. Entonces, este nivel mostraría los paquetes, clases, vistas y demás que forman parte del elemento como se aprecia en la figura 9, pero sin entrar en detalles de código, sólo mostrando su existencia y la interacción que hay entre cada uno de ellos de forma simple.

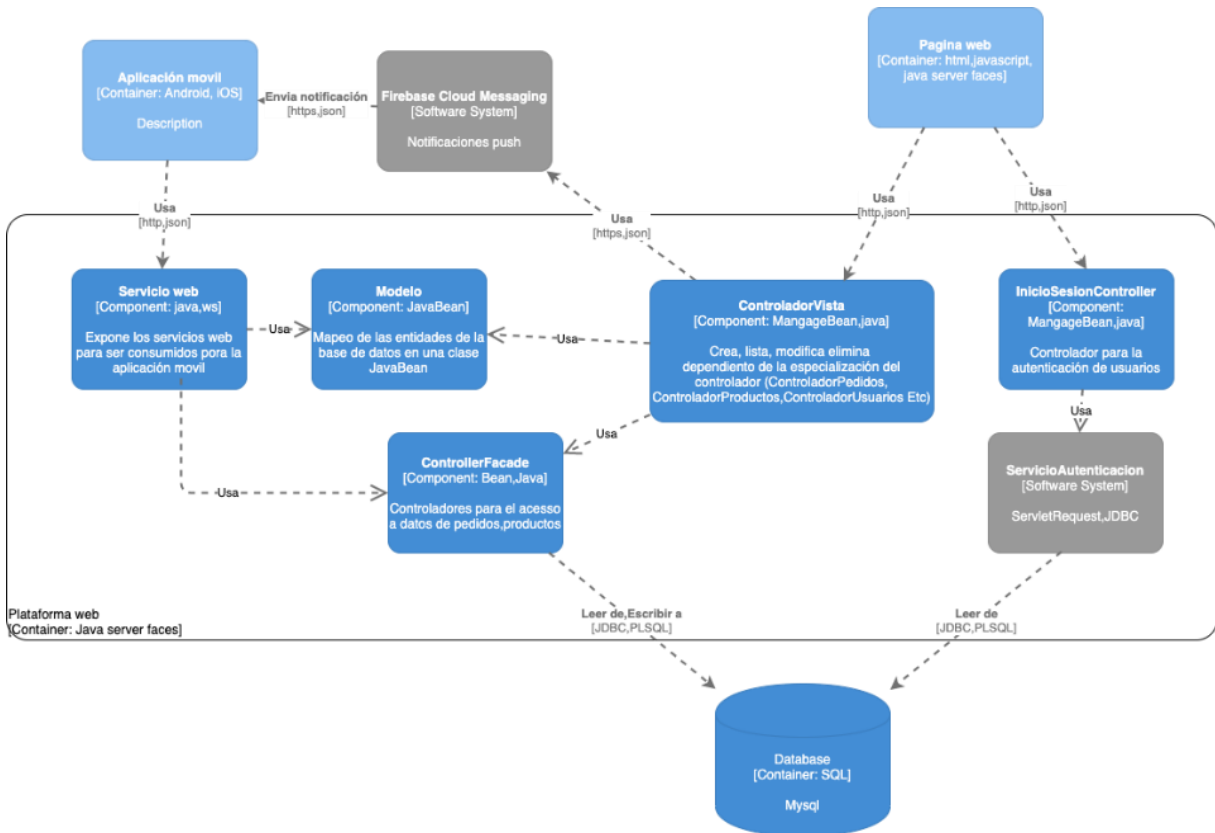


Figura 9 Nivel C3 Fuente Elaboración propia

2.1.6.4. Nivel 4: El código

Este nivel especifica cada uno de los componentes de manera específica, Explicando los atributos, funciones y demás de las clases, interfaces y otros componentes que se necesiten mostrar y sus interacciones como se aprecia en la figura 10. Este nivel sólo se muestra si el público objetivo es el personal técnico de desarrollo o similares, dada su complejidad. Y por esto mismo este nivel se emplea poco, ya que con los niveles anteriores es suficiente para transmitir el sistema a desarrollar [20].

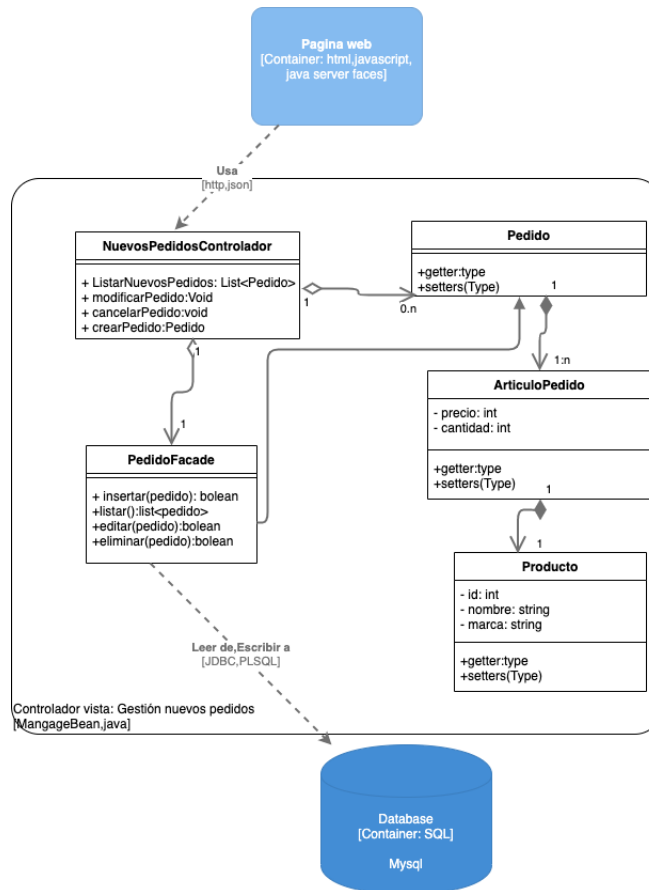


Figura 10 Nivel C4 Fuente Elaboración Propia

En el modelo C4 se puede llegar hasta el nivel que se necesite de profundidad, generando cada vez mayor nivel de detalle según sea requerido, sin perder el objetivo del modelo, que es, mostrar de forma clara y simple la arquitectura de una solución tecnológica. Es por esto que por lo general no se pasa del Nivel 4.

El modelo C4 no tiene una notación específica, se busca que las cosas sean tan simples como sea posible, pero teniendo en cuenta que a cada elemento es bueno asignarle un nombre, un tipo de elemento como “Persona” o “Base de datos” y un corto texto descriptivo que ayuda a evitar ambigüedades que a veces se presentan en los modelos. Si se desea se puede usar la notación de UML y mostrar todo de forma un poco más rigurosa, lo importante es que el modelo sirva para que el público objetivo entienda de qué está formada la solución. Como no hay una notación específica, si se necesita, se puede agregar una leyenda para describir los elementos gráficos como líneas, colores y bordes que se usan en el modelo y obviamente estos se deben usar de forma consecuente en este.

2.1.7. Framework

Qué es un framework en informática o programación. La palabra Framework es la combinación de dos palabras, es decir, Marco (frame) y Trabajo (work). Esto significa que previamente se ha diseñado un marco bajo patrones de diseño, segura y con buenas prácticas de desarrollo, por lo que el desarrollador puede funcionar en este marco para cumplir con los requisitos de su proyecto.

En general, un marco es una estructura real o conceptual destinada a servir como soporte o guía para la construcción de algo que expande la estructura en algo útil. Es solo una herramienta que ayuda al desarrollador a codificar mejor y más rápido [19].

2.1.8. Calidad de dato

Los datos son un activo valioso que debe ser gestionado a medida que se mueve a través de una organización. A medida que las fuentes de información son cada vez más numerosas y diversas, y las iniciativas de conformidad normativa se enfocan más, la necesidad de integrar, acceder y reutilizar información de estas fuentes dispares de forma consistente y confiable se está volviendo crítica [13].

2.1.9. AMI_SAA

Es una plataforma que implementa la tecnología AMI. SAA significa Sistema de Amarre Automático y permite que los sistemas de los PLC_TU se comuniquen automáticamente con los PLC_MMS. Esta plataforma es un proyecto de Colciencias desarrollada por la Universidad del Cauca que tiene patente No. NC2017/0002981 que se encuentra actualmente en trámite de aprobación. Esta plataforma permite que los datos capturados por el sistema SAA sean enviados y procesados para darles una utilidad lo que los convierte en información. Los datos son utilizados principalmente para la medición del consumo de los usuarios, consultas de conexión, corte y reconexión de un usuario y balances de energía. También puede hacer un sondeo de amarre por fase con el fin de detectar si todas las fases están debidamente amarradas y así encontrar posibles fallas o fraudes.



Figura 11 Vista principal de la aplicación AMI SAA Fuente Elaboración propia

Dentro de las opciones disponibles en la plataforma también se encuentran algunas que aportan al proceso, pero no se encontraban dentro del modelo inicial de la misma, un ejemplo de esto es la sección de “Gestión de Macro medidores”. Entendiendo estos como medidores asociados a los transformadores. Estos macromedidores se relacionan con los dispositivos PLC_MMS que son los encargados de recibir la información de consumo del macromedidor y los consumos particulares contenidos en los dispositivos PLC conectados a él.

2.1.10. NORMA ISO/IEC 25012

La norma ISO/IEC 25012 es uno de los principales referentes a la hora de identificar las características de calidad que permiten juzgar un conjunto de datos. Existen 3 bloques de características que se evalúan en el modelo como se puede apreciar en la figura 12, estos bloques son:

- Calidad de datos inherente.
- Calidad de datos dependiente del sistema.
- Calidad de datos inherente y dependiente del sistema.

La calidad de datos inherente hace referencia a aquellas características cuya medición depende de la naturaleza de los datos [20]. Desde el punto de vista dependiente del sistema, la Calidad de Datos depende del dominio tecnológico en el que los datos se utilizan, y se alcanza mediante las capacidades de los componentes del sistema informático tales como: dispositivos hardware (en general Respaldo Software para alcanzar la Recuperabilidad), y otro software (en general Herramientas de migración para alcanzar la Portabilidad) [4]. La calidad de datos inherente y dependiente del sistema se refiere a aquellas características de calidad que en encuentran aquellas cuya medición dependen tanto de la naturaleza de los datos como de la definición del sistema de información en el que los datos se alojan [22].

Dentro de las características que se evalúan al aplicar este modelo están:

- **Exactitud:** proximidad del valor de datos a un conjunto de valores válidos predefinidos en un dominio dado para un contexto específico.
- **Compleitud:** el grado en que los datos asociados a la entidad tienen valores para todos los atributos necesarios asociados para representar correctamente a la entidad.
- **Consistencia:** grado en el que los datos están libres de contradicción y son coherentes con otros datos en un contexto de uso específico. Puede ser analizada en datos que se refieran tanto a una como a varias entidades comparables.
- **Credibilidad:** grado en el que los datos tienen atributos que se consideran ciertos y creíbles en un contexto de uso específico. La credibilidad incluye el concepto de autenticidad (la veracidad de los orígenes de datos, atribuciones, compromisos).
- **Actualidad:** la medida en que los datos se actualizan correctamente y en el momento oportuno.
- **Accesibilidad:** grado en que los datos pueden ser accedidos en un contexto específico de uso, particularmente por personas que necesitan una tecnología de soporte o una configuración especial debido a alguna incapacidad.
- **Conformidad:** grado en que los datos cumplen estándares, convenciones o legislaciones o reglas similares relativas a la calidad en un contexto específico de uso.
- **Confidencialidad:** grado en que los datos tienen atributos que aseguran que sólo pueden ser accedidos por usuarios autorizados en un contexto específico de uso.
- **Eficiencia:** grado en el que los datos tienen atributos que pueden ser procesados y proporcionados con los niveles de rendimiento esperados mediante el uso de cantidades y tipos adecuados de recursos en un contexto de uso específico.
- **Precisión:** grado en el que los datos tienen atributos que son exactos o proporcionan discriminación en un contexto de uso específico.
- **Trazabilidad:** grado en el que los datos proporcionan un camino de auditoría de acceso a los datos y sobre los cambios realizados en los datos en un contexto específico de uso.
- **Comprensibilidad:** grado en el que los datos tienen atributos que permiten ser leídos e interpretados por los usuarios y son expresados utilizando lenguajes, símbolos y unidades apropiados en un contexto de uso específico.
- **Disponibilidad:** grado en el que los datos pueden ser recuperados por usuarios o por aplicaciones autorizadas en un contexto específico de uso.

- **Portabilidad:** grado en el que los datos tienen atributos que les permiten ser instalados, reemplazados o eliminados de un sistema a otro, preservando el nivel de calidad en un contexto de uso específico.
- **Recuperabilidad:** grado en el que los datos tienen atributos que permiten mantener y preservar un nivel específico de operaciones y calidad, incluso en caso de fallos, en un contexto de uso específico.

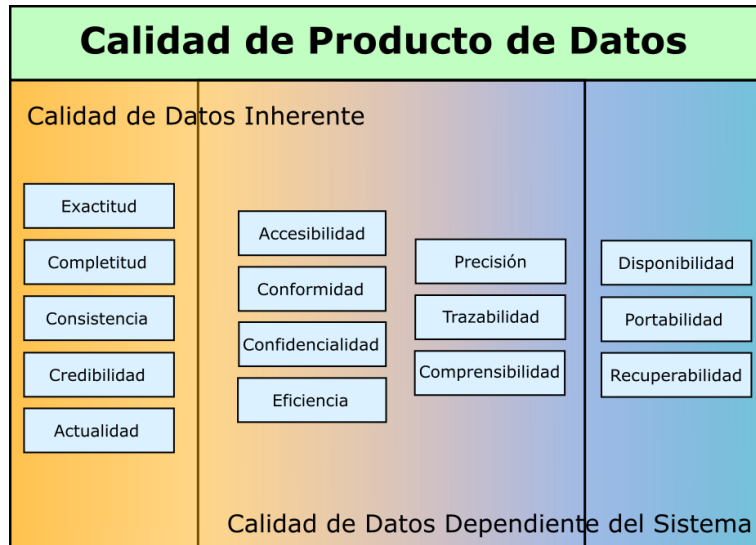


Figura 12 Calidad de producto de datos Fuente [23]

2.2. Estado del Arte.

Se realizó una revisión sistemática de la literatura con el fin de identificar documentos en los cuales aplicaran un modelo para la evaluación de calidad de datos. A continuación, se presenta la revisión sistemática y los documentos encontrados.

2.2.1. Revisión sistemática

Revisión sistemática es un estudio integrativo, observacional, retrospectivo, secundario, en el cual se combinan estudios que examinan la misma pregunta [21].

El proceso de búsqueda y selección de información, es un conjunto de técnicas que se adquieren de acuerdo a la información que se necesite y las herramientas que se tengan a disposición, como los motores de búsqueda, buscadores académicos y bibliotecas digitales que se encuentran en la red; con ello se

encontraran un sin número de artículos, publicaciones en revistas e información relacionada con el tema a buscar.

En el presente ítem se realiza la revisión sistemática con el fin de filtrar artículos relacionados con la tecnología AMI, Smart Grid, calidad de los datos, calidad de los datos en un contexto eléctrico y normas y modelos de calidad de datos.

2.2.1.1. Fuentes de búsqueda.

2.2.1.1.1. Google Scholar

Es un motor de búsqueda especializado en búsqueda de documentos científicos y académicos [22], facilita la búsqueda de trabajos de investigación en grandes áreas como: informática, física, derecho y medicina, debido a que emplea algoritmos diseñados para el entorno académico.

El rango de fecha utilizado para la búsqueda está comprendido entre el año 2010 y el 2019, se hizo uso de palabras claves donde se incluyeron artículos en inglés y español.

2.2.1.1.2. IEEE Xplore Digital Library

Es un recurso que proporciona información de contenido científico-tecnológico gestionado por IEEE. IEEE Xplore proporciona acceso a la Web a más de cuatro millones de documentos de texto completo de algunas de las publicaciones más citados del mundo en ingeniería eléctrica, informática y electrónica [23].

2.2.1.2. Estrategia de búsqueda

Para realizar la búsqueda se usaron diferentes estrategias las cuales permitían que la búsqueda fuese más precisa:

- **Búsqueda avanzada:** permite realizar búsqueda por frases, seleccionar un dominio, formato, limitar por idioma del documento, etc.
- **Búsqueda por palabras clave:** Permite realizar una búsqueda por una palabra específica Ej: Infraestructura de medición avanzada (AMI), Smart Grid, AMI, Data Quality, Calidad de datos, sector eléctrico, Data Quality standard, Data Quality model, ISO/IEC 25012.
- **Busqueda de frase exacta:** Permite que el resultado tenga específicamente la frase que se está buscando Ej. “Calidad de datos sector eléctrico”.

- **Búsqueda en título del documento:** Permite una búsqueda en la cual los resultados obtenidos incluyan la palabra o frase en el título del documento Ej.: intitle: advanced metering infrastructure.
- **Búsqueda con operadores AND, OR y NOT:** Esta búsqueda permite el uso de operadores lógico Ej.: Data quality AND (AMI OR electric power).

2.1.2. Documentos

2.1.2.1. Una Metodología Basada en ISO/IEC 15939 para la Elaboración de Planes de Medición de Calidad de Datos.

Debido a que poco a poco se ha incrementado la necesidad de tener mayor cantidad y mejor calidad de datos, se fue ampliando la necesidad de tener un marco de trabajo o una metodología con la cual medir la calidad de los datos con los cuales se está trabajando. La calidad de los datos está ligada a diferentes factores como su confiabilidad y disponibilidad o el contexto para el que se usan, pero hay un acuerdo común. La calidad de un dato viene dada principalmente por el nivel de cumplimiento que tiene este con el objetivo para el cual es usado. Este trabajo presenta una metodología para la medición de los datos llamada MEPLAMECAL [24]. Esta metodología, basada en la ISO/IEC 15939 se compone de dos actividades principales:

- Establecer y Mantener el Compromiso en la Medición de la CD, como se puede apreciar en la figura 13.
- Elaborar el Plan de Medición, como se puede apreciar en la figura 14.

Cada una de estas actividades tiene sus propias tareas y un conjunto de productos entregables que deben quedar como resultado de cada una. Estos entregables surgen en su mayoría de los entregables que normalmente se ven en la Ingeniería de Software.

Las tareas asociadas con **Establecer y mantener el compromiso de la organización** están orientadas a definir un equipo que se va a encargar de realizar el proceso de medición de la calidad de los datos, dentro de los cuales debe haber personas expertas en datos y calidad de datos y también personas de otras disciplinas que hagan uso de los mismos y formen parte de la definición del modelo de los datos usados.

Para la **Elaboración del plan de medición** se definen tareas adicionales que buscan conocer en profundidad los datos que se están trabajando: la cantidad, las fuentes, los conceptos medibles de calidad de datos y las necesidades de información. Una vez hecho esto se procede a planear cuándo y cómo se va a

realizar la medición. Como resultado se obtiene un plan de medición que tiene en cuenta variables como la disponibilidad de talento humano y de recursos económicos necesarios para ejecutar el plan, esto debido a que muy posiblemente las personas involucradas en la ejecución del plan ya tengan otras tareas asignadas, por lo que su disponibilidad estaría reducida.

Cabe resaltar que como resultado de la medición se obtienen datos que sirven para juzgar la calidad de los datos, pero no se decide que los datos sean buenos o malos ya que esto hace parte de un proceso de evaluación, en otras palabras, la medición entrega un conjunto de parámetros de entrada que pueden ser usados en un proceso de evaluación [24].

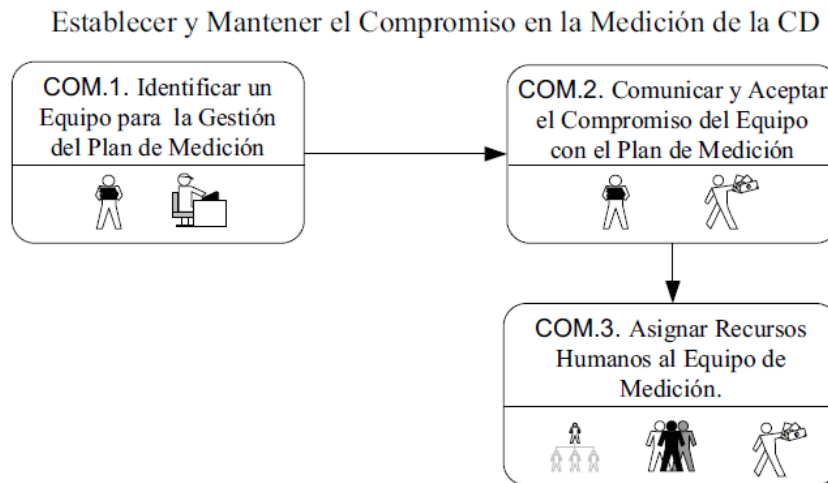


Figura 13 Actividad 1: Establecer y Mantener el Compromiso en la Medición de la CD Fuente [24]

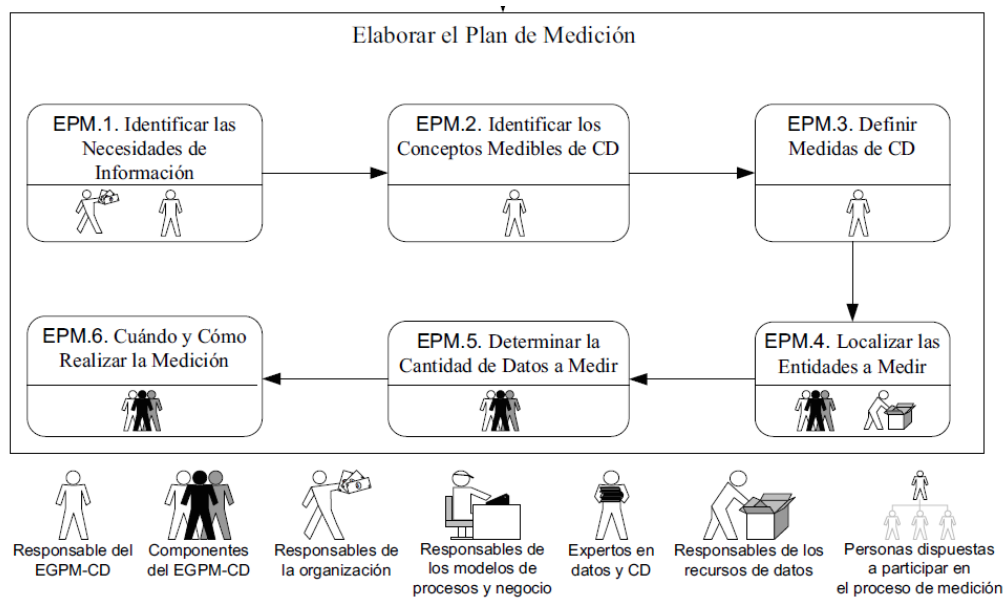


Figura 14 Actividad 2: Elaborar el Plan de Medición Fuente [24]

2.1.2.2. Framework de evaluación de calidad de la información: Ampliación del modelo de calidad de datos ISO 25012.

Este trabajo propone un acercamiento en la forma de especificar los requerimientos en calidad de la información para las aplicaciones web en particular al extender la ISO 25012 del 2008 del modelo de calidad de datos. Según este trabajo la información es presentar datos de una forma significativa semánticamente. Entregar información debe ser considerado un aspecto funcional de las aplicaciones software. La calidad de la información entregada por cualquier aplicación debe ser juzgada con base en calidad de los datos y la idoneidad de las funciones que transforman esos datos en información.

Dentro del framework propuesto se utilizan todas las características de calidad de la ISO 25012 a excepción de la característica de Conformidad. Se adicionaron dos características:

- **Valor agregado:** define qué tan ventajoso es usar un dato en particular respecto a otro
- **Exactitud representacional:** el nivel en que un dato o una información es presentado al usuario de una forma concisa, flexible y organizada con la debida relevancia para conseguir un fin especificado [25].

Cualquier información, aunque sea precisa y completa, si es presentada de una forma poco estructurada no ofrece las mismas ventajas que una que sí lo hace. Esta puede representar problemas para su entendimiento y aprendizaje. Condición que no es tenida en cuenta en la ISO 25012 por lo cual se adicionó.

2.1.2.3. Un framework dirigido por modelos para la gestión de calidad de datos en el internet de las cosas.

El Internet de las cosas es un entorno con un enorme flujo de datos enviados y recibidos por elementos inteligentes. La calidad de los datos es un criterio clave para los consumidores de esos datos, especialmente teniendo en cuenta la incertidumbre inherente en los sensores. La Calidad de los datos es un concepto altamente subjetivo y no hay un consenso general acerca de qué se considera “buenos” datos. Para lidiar con este problema se propone en este trabajo una arquitectura dirigida por modelos que permitan al consumidor de los datos reflejar sus necesidades con los datos y generar una infraestructura para la gestión de la calidad de los datos. Los requerimientos de los datos se expresan a través de una herramienta “fácil de usar” que provee recursos al usuario para mostrar su visión de los requerimientos en calidad de los datos. Esta arquitectura es probada en dos casos de flujos de datos en un entorno de la vida real.

Básicamente el objetivo de este trabajo consta de tomar de parte de los desarrolladores de IoT sus necesidades a través de una herramienta y luego convertir estas necesidades a un modelo generado automáticamente que normalice estos requerimientos para que lo que se expresa de una forma u otra entregue al final los mismos resultados [26].

Los pasos a seguir son los siguientes. Tal y como puede apreciar en la figura 15:

- El desarrollador de IoT define las especificaciones de Calidad de los datos usando un editor de modelos gráfico.
- Un conjunto de transformaciones automáticas es usado para bajar el nivel de abstracción del modelo de Calidad de los datos especificado.
- Una plataforma para la gestión de la Calidad de los datos es generada de tal forma que se adapte a los requerimientos del consumidor de los datos.
- Si los requerimientos en Calidad de datos del consumidor evolucionan, como por ejemplo al descubrir cuestiones de negocio, se puede volver al primer paso [26].

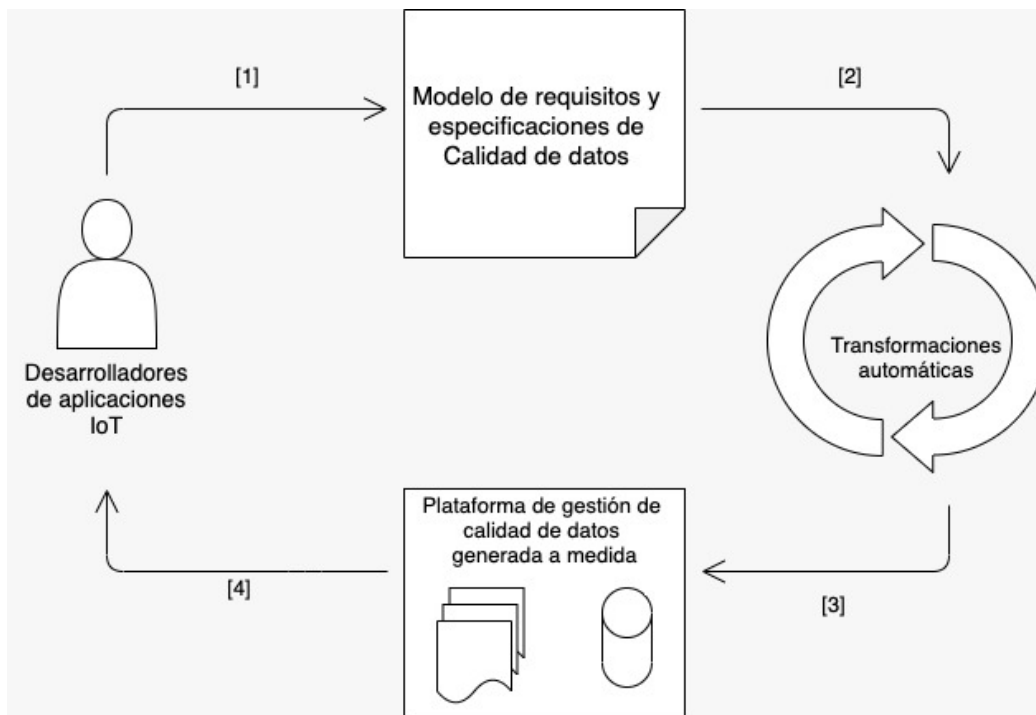


Figura 15 Workflow de la propuesta. Fuente elaboración propia.

Para el desarrollo de las herramientas se usó JavaSE. En particular la herramienta para que el desarrollador defina sus requerimientos en Calidad de los datos como se puede ver en la figura 16:

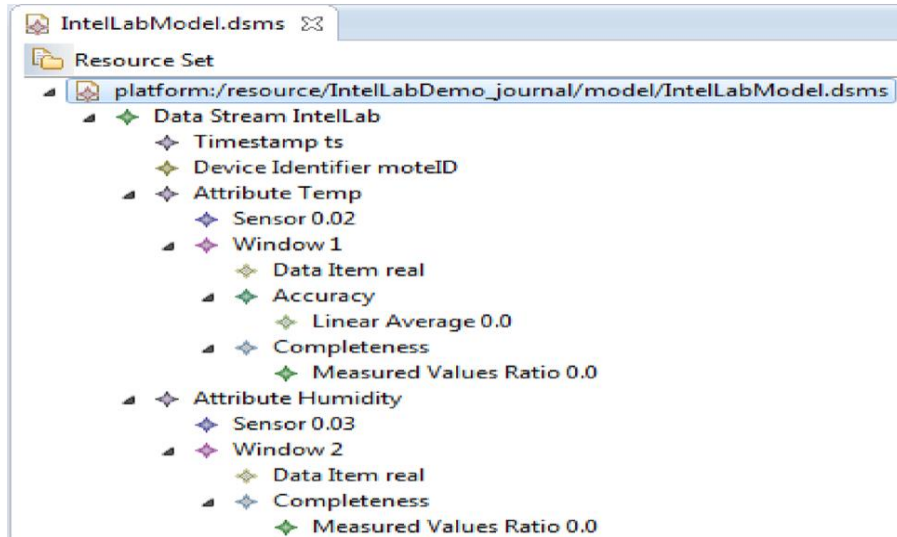


Figura 16 Estructura del proyecto Fuente [26]

Como se puede ver se están teniendo en cuenta dos datos tomados de unos sensores que se encontraban en un laboratorio de investigación (IntelLab), en este caso los valores medidos fueron la temperatura y la humedad del ambiente, de los cuales se buscan las características de calidad Completitud y Precisión.

Después de aplicar la lógica para cubrir esos requerimientos se generó un diagrama con una herramienta de Oracle llamada Oracle Event Processing como se ve en la figura 17:

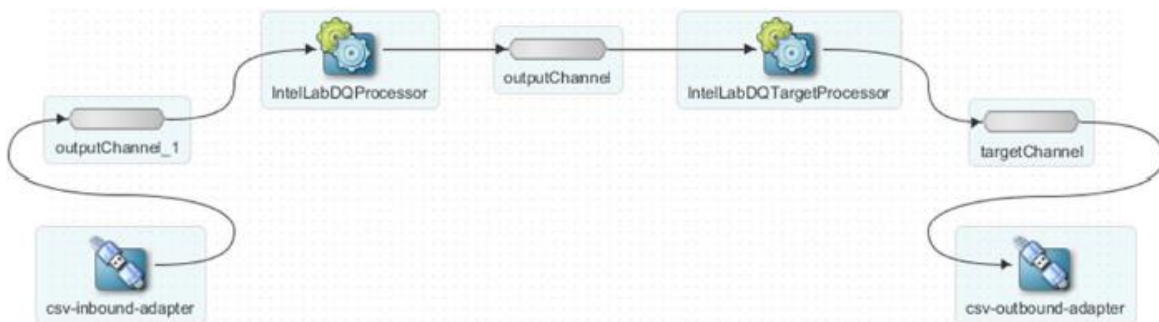


Figura 17 Oracle Event Processing Fuente [26]

Este diagrama muestra un proceso que es ejecutado por una aplicación generada que usa un documento CSV como entrada y genera una salida “limpia” que debería cumplir con los requerimientos.

De acuerdo a los resultados que se muestran en este trabajo se logró un alto nivel de flexibilidad para ajustarse a los requerimientos de Calidad de datos de los especialistas y los consumidores de estos datos, sin embargo, cabe resaltar que es necesario un aprendizaje de las herramientas proveídas para que se pueda hacer una efectiva gestión de la calidad de los datos [26].

2.1.2.4. Comparison of Data-Driven Models for Cleaning eHealth Sensor Data: Use Case on ECG Signal.

Los **Registros electrónicos de salud** permitieron almacenar y procesar datos entregados por sensores, lo que permite un nivel de atención mejorado respecto a tiempos anteriores, pero estos datos están expuestos a múltiples fallas tanto de los dispositivos como generadas por el ambiente y otros elementos, lo que lleva a posibles diagnósticos errados o inexistentes, siendo así, la calidad de los datos recolectados resultan ser un tema crítico para la vida de los pacientes. Este trabajo mostró una comparación entre dos modelos dirigidos por datos para la limpieza de los datos obtenidos por los sensores de salud, como dispositivos que se visten como un marcapaso, con el fin de asegurarse que los datos recolectados sean correctos y relevantes para ser añadidos a un Registro electrónico de salud [27].

Dentro de las técnicas que hay para la limpieza de los datos se encuentra la gestión de datos que superan la media, gestión de datos faltantes y la detección de valores repetitivos. El flujo que se sigue en este trabajo para la limpieza de los datos como se ve en la figura 18:

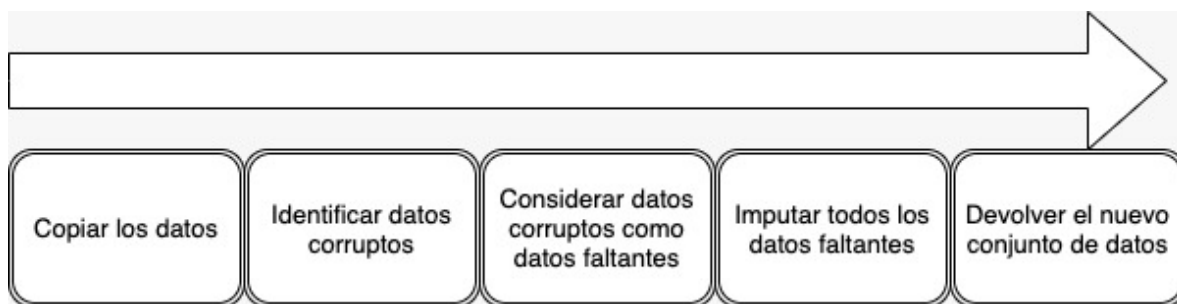


Figura 18 Workflow para la limpieza de datos. Fuente elaboración propia.

Primero se inicia por hacer una copia de los datos, esto con el fin de no modificar los datos originales y poder hacer posteriores procesos con ellos de ser necesario. Luego de esto se procede a identificar los datos corruptos, estos pueden

generarse por distintos factores y por lo general son valores que discrepan en mayor o menor medida la mayoría de los datos recolectados en tiempos cercanos. También se identifican situaciones en las que hay datos inexistentes bien sea de manera intermitente o por periodos de tiempo continuos. Posterior a eso se pasa a considerar los datos corruptos como datos inexistentes. Se eliminan esos datos considerados como inexistentes y por último se retorna un nuevo conjunto de datos “limpio” que debería servir de mejor insumo para la gestión y el diagnóstico de enfermedades.

Para poder mandar alertas y diagnosticar de forma temprana enfermedades, se aplican a los datos recolectados modelos predictivos como por ejemplo regresión lineal múltiple, arboles de decisión, redes neuronales, entre otros. En particular para este trabajo se usó un dataset que se encuentra disponible de forma pública llamado MHEALTH (Mobile Health), que reúne información de sensores y muestra datos como el movimiento de algunas partes del cuerpo, el ritmo cardiaco, señales usadas en electrocardiogramas y es usado para generalizar actividades rutinarias como correr o estar de pie.

En el trabajo se compararon resultados arrojados al aplicar limpieza de los datos usando vario modelos dirigidos por datos y se determinó que los que mejores resultados arrojados eran la regresión lineal múltiple y las redes neuronales. Estos se adaptaron para tener en cuenta las características de calidad: precisión, completitud, correctitud y consistencia [27].

2.1.2.5. Hacia un framework de evaluación de calidad de información en foros de discusión técnicos.

Teniendo en cuenta las preguntas y respuestas que se presentan en foros de discusión técnicos en los sitios web. Este artículo presenta un conjunto de características que permiten realizar una evaluación más precisa de la calidad de la información presentada en las respuestas de cada pregunta realizada, adicional a esto, se propone un modelo conceptual que permite la clasificación de esta información, basándose principalmente en el estándar ISO/IEC 25012:2008 [28].

Dentro del artículo se describe el proceso por el que pasa un técnico al momento de consultar sobre un tema que es de su interés como se ve en la figura 19:

- a. Digita el tema o pregunta pertinente en un navegador.
- b. Selecciona los enlaces que más se acercan a lo que desea consultar o ahondar. Estos resultados están delimitados por el buscador usado y sus algoritmos de ponderación.

- c. Evalúa la exactitud de lo que se encuentra en cada enlace respecto a su duda y también verifica (dado el caso) la confiabilidad de la fuente de los datos.

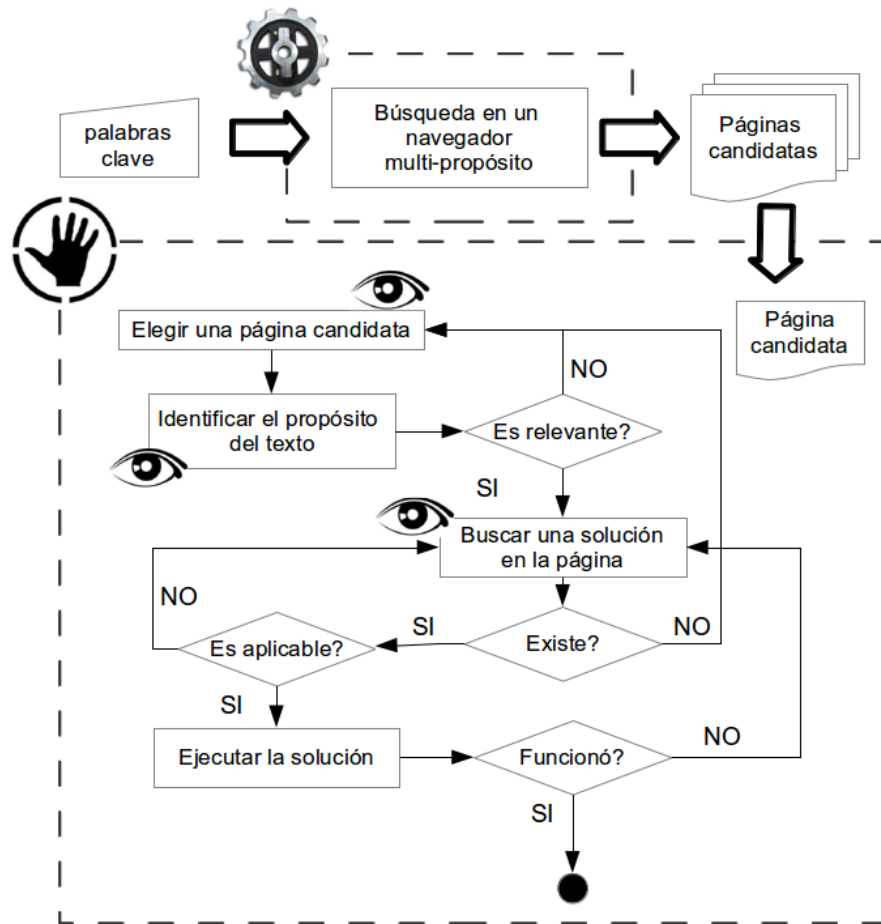


Figura 19 Diagrama de flujo búsqueda por técnico Fuente [28]

Para poder definir la calidad de la información que se encuentra en un foro se parte por definir los tipos de usuarios específicos para este tipo de aplicaciones web. Los tipos de usuarios definidos para este trabajo son:

- Administrador
- Moderador
- Usuario participante
- Usuario externo

Cada usuario tiene unas características, acciones permitidas y privilegios sobre la información en el foro, siendo el administrador el que más tiene y el usuario externo (en general) sólo un consumidor de la información que allí se encuentra. De acuerdo con la interacción y al interés mostrado por un usuario externo, este puede ir ascendiendo en el nivel de privilegios dentro de la aplicación.

El contenido que se encuentra en los foros de discusión técnico se encuentra dentro de un conjunto de hilos, cada uno con cero o más respuestas de los diferentes usuarios. En la figura 20, se presenta un modelo que describe el contenido de la información en un foro:

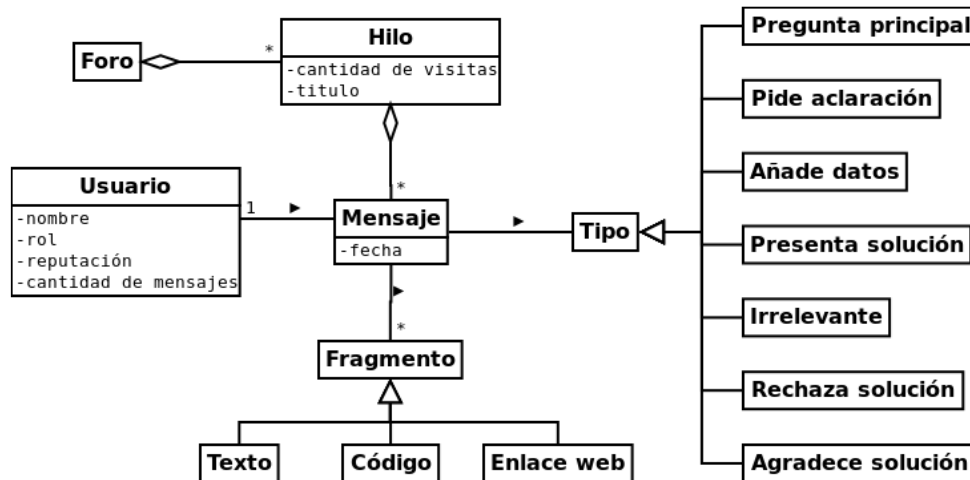


Figura 20 Diagrama de la información contenida en un foro de discusión técnico Fuente [28]

De acuerdo con las características inherentes al foro se definieron las relaciones con las características de calidad de datos de la norma ISO/IEC 25012 con las cuales se puede determinar la calidad de los datos existentes en el foro. En la figura 21 se ve el resultado de este estudio:

	Dimensiones de Calidad	Exactitud	Complejidad	Consistencia	Credibilidad	Actualidad	Accesibilidad	Conformidad	Confidencialidad	Eficiencia	Precisión	Trazabilidad	Entendibilidad	Disponibilidad	Portabilidad	Recuperabilidad	Valor Agregado	Adec. Represent.	
		Atributos																	
HILO	titulo	●	●																
	cantidad visitas	●	●							●									
USUARIO	nombre																		
	cantidad mensajes	●	●																
	reputación	●	●		●					●	●		●						
MENSAJE	fragmento	texto	●	●									●					●	
		código	●	●				●				●	●		●		●	●	
		enlace web	●	●									●	●				●	
	fecha					●													
	tipo	pregunta principal	●	●															
		añade datos	●	●							●	●	●	●	●			●	●
		irrelevante																	
		agradece solución	●	●		●					●	●	●	●		●		●	●
		presenta solución	●	●		●					●	●	●	●	●	●		●	●
		pide aclaración																	
rechaza solución																	●		

Figura 21 Relación características del foro con características de calidad Fuente [28]

De este modelo cabe destacar que hay relaciones que no parecen ser muy claras inicialmente, pero que revisando a fondo resultan necesarias para definir si se tiene información útil para el usuario. Por ejemplo, el título del hilo está relacionado con la característica de exactitud, esto, debido a que un título de un hilo puede contener todas las palabras en la búsqueda de un usuario, pero aun así no ser un resultado relevante para dar solución a la necesidad de información, por lo tanto, se debe evaluar si de verdad ofrece lo que se busca.

Capítulo 3. Caracterización de los tipos de medida en un contexto eléctrico.

En este capítulo se hace una clasificación de los tipos de medidores que se usan actualmente en el sector eléctrico, se exponen los 3 tipos de medición para finalmente presentar las características de los diferentes tipos de medida.

Actualmente, la energía eléctrica en Colombia se mide a través de dispositivos medidores mecánicos o de inducción y los estáticos Y/o electrónicos estos últimos compuestos por elementos electrónicos, capaces de registrar los consumos de energía activa o de energía reactiva, sea de manera independiente o simultánea y en un periodo de tiempo determinado, cuantificando los kilovatios – hora (kWh) o los kilovoltamperios reactivos – hora (kVAr-h) respectivamente. Existen tres (3) tipos de medición, que dependerán de los elementos necesarios para el registro de los consumos [29]. La selección entre un tipo de medición u otro, se hace dependiendo de la demanda de la instalación. Ellos son: medición directa, medición semidirecta y medición indirecta [30].

En relación con la medida. Los medidores se clasifican en función de sus características y la exactitud en su medida:

3.1. Características físicas de los medidores

3.1.1. Según su construcción

3.1.1.1. Mecánicos o de inducción

Compuesto por un conversor electromecánico que actúa sobre un disco, cuya velocidad de giro es proporcional a la potencia demandada, provisto de un dispositivo integrador. El principio de funcionamiento es muy similar al de los motores de inducción y se basa en La teoría de la relación de corriente eléctrica con los campos magnéticos [31].

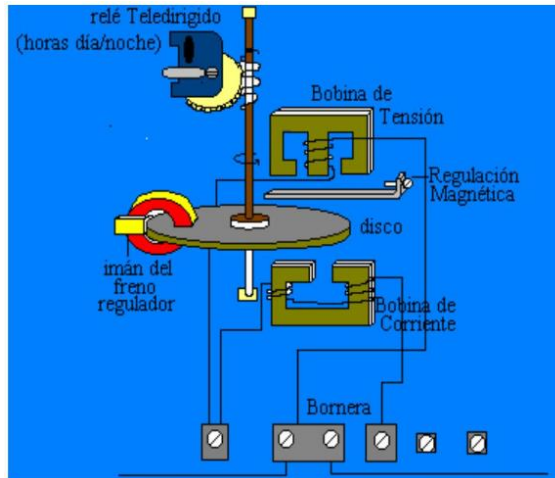


Figura 22 Medidor de inducción. Fuente [32]

3.1.1.2. Estáticos o electrónicos

Medidores en los cuales la corriente y la tensión actúan sobre elementos de estado sólido (electrónicos) para producir pulsos de salida y cuya frecuencia es proporcional a los Vatios-hora (KWh). Están contruidos con dispositivos electrónicos, generalmente son de mayor precisión que los electromagnéticos y por ello se utilizan para medir en centros de energía, donde se justifique su mayor costo [31].



Figura 23 Medidor estático. Fuente [32]

3.1.2. Según energía medida

3.1.2.1. Energía Activa

Mide el consumo de energía activa en kilovatios – hora. (KWh) [31], entendiendo la energía activa como aquella que genera trabajo útil al transformarse en luz, calor, movimiento, ETC... [33].

3.1.2.2. Energía Reactiva

Mide el consumo de energía reactiva en kilo vares – hora. (KVArh) [31], entendiendo la energía reactiva como aquella que no genera trabajo útil, y es generado por todos los dispositivos que utilicen campos eléctrico y electromagnéticos para su funcionamiento, como motores, refrigeradores, Etc... [33].

3.1.3. Según su conexión en red

3.1.3.1. Monofásicos

Este tipo de conexión se realiza generalmente para usuarios residenciales, donde la carga instalada no supera los 7 kW. También se puede observar en servicios comerciales donde la actividad puede ser: papelerías, pequeñas tiendas, modisterías, ETC... Para este tipo de conexión se instalan únicamente medidores monofásicos de clase 1 o clase 2 [34].

Estos a su vez se agrupan de acuerdo con la cantidad de sus filamentos transportadores ver figura 24 en:

3.1.3.1.1. Bifilares

Se utiliza para el registro de consumo en una acometida que tenga un solo conductor activo o fase y un conductor no activo o neutro [31].

3.1.3.1.2. Trifilares

Se utiliza para el registro del consumo de una acometida monofásica de fase partida (120/240 V) donde se tienen dos conductores activos y uno no activo o neutro [31].

3.1.3.2. Bifásicos

Este tipo de conexión se realiza generalmente en sectores residenciales, donde la carga instalada no supera los 15 kW. También se pueden observar en servicios comerciales donde la actividad requiera equipos que funcionan a 220 V. Para este tipo de conexión se pueden instalar medidores bifásicos clase 1 o clase 2 [34].

Estos a su vez se agrupan de acuerdo con la cantidad de sus filamentos transportadores ver figura 24 en:

3.1.3.2.1. Trifilares

Se utiliza para el registro del consumo de energía de una acometida en B.T de dos fases y tres hilos, alimentadas de la red de B.T de distribución trifásica [31].

3.1.3.3. Trifásicos

En pocas ocasiones se encuentra este medidor instalado en servicios residenciales, es más común en servicios comerciales e industriales donde la carga instalada supera los 15 kW, para una conexión directa. Para la medición semi-directa se utiliza medidores de energía activa y/o reactiva y un transformador de corriente (T.C) por cada fase que alimenta la carga. En este tipo de medición, la conexión de las señales de corriente provenientes de los devanados secundarios de los T.C. y de las señales de tensión provenientes de la acometida al medidor, deben realizarse mediante un bloque de prueba, para garantizar una correcta medida de la energía consumida [34].

Estos a su vez se agrupan de acuerdo con la cantidad de sus filamentos transportadores ver figura 24 en:

3.1.3.3.1. Tetrafilares

Se utiliza para el consumo de energía de una acometida trifásica en B.T de tres fases y cuatro hilos [31].

3.1.3.3.2. Trifilares

Se utiliza para el registro de consumo de energía de una acometida trifásica de tres fases sin neutro [31].



Figura 24 Conexión en red. Fuente elaboración propia.

3.1.4. Según la conexión de los filamentos.

Definida por la disposición de las bobinas y la distribución de las fases de entrada y salida [32].

- Simétricos.
- Asimétricos.

Estas conexiones están relacionadas con las clases de conexiones en red y sus filamentos, anteriormente mencionados [35].

Medidor monofásico bifilar: medidor conectado a una acometida de dos líneas una fase y un neutro, estos pueden ser simétrico o asimétrico como se observa en las figuras 25 y 26, estos son utilizado en usuarios con carga o consumo bajo. Se puede instalar en una red proveniente de un transformador trifásico o monofásico [35].

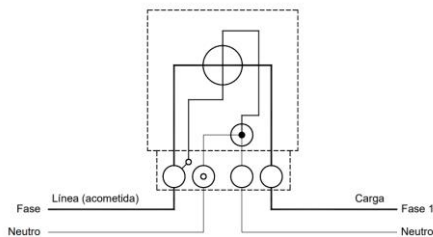


Figura 25 Conexión de medidor monofásico bifilar simétrico. Fuente [36]

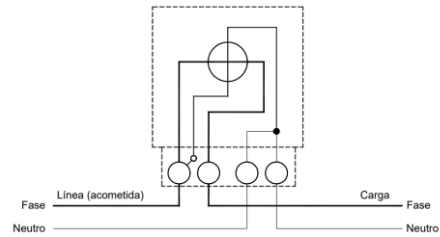


Figura 26 Conexión de medidor monofásico bifilar asimétrico. Fuente [36].

3.1.5. Según la complejidad de su construcción

Los medidores de energía se clasifican de acuerdo con su complejidad en: Básicos, Multienergía, Multifunción y Medidores de funciones adicionales. En la Tabla 1 se pueden observar los detalles de esta clasificación.

Clasificación	Descripción
Básicos	Medidores de energía sencillos, sin dispositivos internos de control de carga o tarifa; con o sin salida de impulso; con o sin puerto de comunicación óptico
Multienergía	Medidores que, en una única carcasa, miden más de un tipo de energía, con o sin salida de impulso; con o sin puerto de comunicación óptico
Multifunción	Medidores básicos o de multienergía, que incluyen funciones adicionales a las metrológicas básicas, tales como registro de demanda máxima, registro de tiempo de uso, dispositivo de control de tarifa y/o carga, como un interruptor horario o un receptor de telemando centralizado
Medidores con funciones adicionales	Medidores con otras unidades funcionales como PLC, comunicación telefónica o por radio, lectores de bonos de pago, etc...

Tabla 1 Clasificación de medidores según la complejidad de su construcción. Fuente [35]

3.1.6. Según la exactitud en su medida

Los medidores se clasifican de acuerdo al margen de error que presentan en la medición de consumo eléctrico [36]:

- **Clase 2:** garantiza que el error se encuentra entre más o menos el 2%.
- **Clase 1:** garantiza que el error se encuentra entre más o menos el 1%.
- **Clase 0.5:** garantiza que el error se encuentra entre más o menos el 0.5%.
- **Clase 0.2:** garantiza que el error se encuentra entre más o menos el 0.2%.

Estas clases simbolizan el margen máximo aceptado, por lo que puede poner medidores de clase inferior, por ejemplo, si se exige un medidor clase 1, se pueden usar medidores clase 0.5.

3.2. Características técnicas de los Tipos de medida

Dependiendo del nivel de tensión (NT) y de la magnitud de la carga a la cual se le va a medir el consumo de energía, la medición puede ser realizada en forma directa, semi-directa o indirecta [35].

3.2.1. Medición indirecta:

Tipo de conexión en el cual las señales de tensión y de corriente que recibe el medidor provienen de los respectivos devanados secundarios de los transformadores de tensión y de corriente utilizados para transformar las tensiones y corrientes que recibe la carga [37].

Para la medición indirecta de energía se utiliza generalmente un medidor estático multifuncional de energía y un juego de transformadores de medida compuesto por Transformadores de Corriente y Transformadores de Tensión. En este tipo de medición, la conexión de las señales de corriente proviene de los devanados secundarios de los TCs, y las señales de tensión provienen de los devanados secundarios de los TTs [29]. Así, la conexión al medidor debe realizarse mediante un bloque de pruebas. Las bobinas de corriente y de tensión se conectan a la red por intermedio de transformadores de corriente y transformadores de tensión. Se utilizan para medir la energía en circuitos de alta tensión, generalmente en subestaciones eléctricas [34].

Para la Compañía Energética del Occidente usuarios con cargas mayores o iguales a 112,5 kVA en 13,2 kV y 250 kVA en 34,5 kV la medición de energía eléctrica se realizará en media tensión, utilizando transformadores de tensión, transformadores de corriente y medidores electrónicos multifuncionales de energía

activa, reactiva y demanda máxima; conectados por medio del bloque de pruebas y conexión, y el sistema de comunicación o modem [29].



Figura 27 Medidor de energía indirecta en poste. Fuente elaboración propia.

3.2.2. Medición semidirecta:

Tipo de conexión en el cual las señales de tensión que recibe el medidor son las mismas que recibe la carga y las señales de corriente que recibe el medidor provienen de los respectivos devanados secundarios de los transformadores de corriente utilizados para transformar las corrientes que recibe la carga [35] y [37].

Las bobinas de corriente se conectan a la red por medio de transformadores de corriente. Se utilizan cuando la corriente de la instalación es superior a 100 A. Las bobinas de tensión se conectan directamente a la red. Normalmente son medidores trifásicos. Se conectan con tensiones hasta 440 V y están diseñados para una medida de corriente de 1 a 10 A, también llamados medidores de 5 A [34].

Los medidores serán electrónicos para energía activa y reactiva, conectados al secundario de los transformadores de corriente por intermedio del bloque de pruebas y conexión, para usuarios con demandas mayores a 36 kVA y menores o iguales a 112,5 kVA [35].



Figura 28 Medidor de energía Semi-directa. Fuente elaboración propia.

3.2.3. Medición directa

Tipo de conexión en el cual las señales de tensión y de corriente que recibe el medidor son las mismas que recibe la carga [29].

Las bobinas de tensión y de corriente se conectan directamente a la red. Se utilizan para corrientes máximas de 100 A. Pueden ser monofásicos, bifásicos o trifásicos [34].



Figura 29 Medidor de energía directa. Fuente elaboración propia.

De acuerdo con los lineamientos anteriormente conceptuados, se presenta a continuación un resumen en la tabla 2. relacionada a la norma técnica colombiana 5019 del 2007 referente a la selección de equipos de medición de energía eléctrica:

Tipo de medición	Tipo de servicio	Capacidad instalada (CI) en kVA	Descripción del medidor ¹⁾²⁾				
			Medidor	Energía ³⁾	Clasificación ⁴⁾	Clase ⁵⁾	
						Electromecánico	Estático ⁶⁾
Directa	Monofásico bifilar	≤ 12	Monofásico bifilar	Activa	Básico	2	1
	Monofásico trifilar	≤ 24	Monofásico trifilar ó Bifásico trifilar	Activa	Básico	2	1
				Activa y Reactiva	Multienergía	--	1 2
	Bifásico trifilar	≤ 24	Bifásico trifilar	Activa	Básico	2	1
				Activa y Reactiva	Multienergía	--	1 2
	Trifásico tetrafililar	≤ 36	Trifásico tetrafililar	Activa	Básico	2	1
Activa y Reactiva				Multienergía	--	1 2	
Semi-directa	Monofásico trifilar	> 24	Monofásico trifilar ó Trifásico trifilar	Activa y Reactiva	Multifunción	--	1 2
	Trifásico tetrafililar	> 36	Trifásico tetrafililar	Activa y Reactiva	Multifunción	--	1 ó 0,5S ¹⁰⁾ 2
Indirecta	Trifásico trifilar	>112.5	Trifásico trifilar ⁷⁾ ó Trifásico tetrafililar ⁸⁾	Activa y Reactiva	Multifunción	--	0,5S 2
			Trifásico tetrafililar ⁹⁾	Activa y Reactiva	Multifunción	--	0,2S 2

Tabla 2 Selección de los medidores de energía. Fuente [35]

En el marco y contexto del objetivo de estudio planteado para este documento, se seleccionará como ítem de interés el tipo de medidor de cargas directas y se tomarán en referencia únicamente los dos primeros tipos de servicio, es decir las asociadas a monofásicos y bifásicos dado que manejan un intervalo de cargas eléctricas instaladas iguales o inferiores a 24 kVa [35] o 15 Kw [34], rangos que involucran voltajes de 110v y 220v utilizados generalmente para uso y consumos residenciales y comercios establecimientos en los que se prestan diversos servicios o se realizan diferentes productos, que por su naturaleza y características no tienen un impacto fuerte en su entorno y no generarán una afectación a sus vecinos, como papelerías, tiendas, modisterías, floristerías, fotocopiadoras entre otros [38].

Capítulo 4. Calidad de datos en un contexto eléctrico.

Los dispositivos que se usan para la medición de energía eléctrica y los tipos de energía asociados a ellos determinan en gran medida la operación de las compañías prestadoras del fluido eléctrico, dicha operación está sufriendo una actualización, en la que los dispositivos medidores han mejorado la comunicación entre ellos y la compañía, permitiendo así una reducción de costos operativos, y un uso más inteligente de la energía eléctrica generada. La tecnología responsable de dicha comunicación es conocida como Power Line Communications o PLC, estos dispositivos PLC permiten usar las líneas convencionales de energía eléctrica con un propósito adicional al de transportar la corriente eléctrica, enviar datos entre los hogares y las empresas [38]. Esta tecnología genera la necesidad de entender como los dispositivos PLC adicionan características a los medidores, que permiten a las líneas constituir una red de comunicación bidireccional y conformar así una parte fundamental de las Smart grid, o redes inteligentes.

Para esto, se requiere generar una separación de la información digital de la señal eléctrica de forma similar a las líneas ADSL logra separar la voz de los datos por medio de un filtrado de frecuencias, la red eléctrica con grandes voltajes y frecuencias cercanas a los 50 MHz y los datos en altas frecuencias y bajos niveles de tensión. Estas frecuencias y voltajes se filtran posteriormente en los dispositivos PLC, logrando traducir estos impulsos en los datos [39].

Dado que esta modernización de las redes eléctricas ha generado una serie de beneficios financieros a quienes las han implementado, y adicionalmente permite tener un mejor control del consumo de energía y su regulación permitiendo así una mejor eficiencia energética. En Colombia, un grupo de ministerios en colaboración con el banco iberoamericano de desarrollo y la UPME, han trazado la ruta hacia la eficiencia energética del país a 2030. Este estudio muestra la ruta que debe tomar el país en la búsqueda de una red energética eficaz [40].

Por su parte, la universidad del cauca adiciona la solución AMI-SAA, directamente relacionada con el desarrollo de este trabajo, y que busca a través del sistema de amarre SAA y los dispositivos PLC relacionados con dicho sistema generar una Smart grid novedosa. En el funcionamiento de las redes AMI-SAA se emplean los dispositivos PLC-TU, PLC-MC y PLC-MMS, Siendo el segundo de estos, el único dispositivo PLC no asociado a un dispositivo medidor, por su parte los otros dos suministran datos sobre el consumo y estado de medidor asociado que posteriormente, serán ordenados en archivos y enviados por los dispositivos PLC a las compañías eléctricas.

Para la gestión de la administración y envío de datos, el sistema AMI-SAA se agrupa en 3 niveles de conexión, en los que se determina de manera jerárquica la

conexión de los dispositivos entre entidades PLC buscando optimizar la administración de los datos, y Siendo estos.

- Nivel 1: PLC-MMS: Los datos proceden únicamente de los consumos registrados en el transformador en su respectivo macromedidor (PLC-MMS).
- Nivel 2: PLC-MMS – PLC-TU: relación directa entre los medidores residenciales (PLC-TU) y los transformadores (PLC-MMS), este tipo de conexión se realiza generalmente en barrios.
- Nivel 3: PLC-MMS – PLC-MC – PLC-TU: Los medidores de cada vivienda (PLC-TU) se conectan a un dispositivo de medida concentrada (PLC-MC) que se encargara de generar un archivo con los datos recolectados de los medidores asociados y posteriormente enviara al transformador (PLC-MMS), este nivel está asociado a edificios de vivienda horizontal o edificaciones que contengan en si una gran cantidad de medidores asociados.

En cuanto a funcionalidad, estos niveles trabajan de la misma manera. El PLC maestro envía una señal solicitando la información a el PLC que continúe en jerarquía, y este, a su vez enviara los datos solicitados, permitiendo así una comunicación en doble vía entre dispositivos PLC. Siendo así. Estos niveles determinan los archivos o grupos de archivos que se procesaran posteriormente en la plataforma AMI-SAA, y que finalmente determinara los valores a cobrar correspondientes a los consumos de los usuarios.

Dada la criticidad que adquieren los datos obtenidos por los dispositivos PLC y procesados por AMI-SAA para mantener la estabilidad financiera de las compañías de energía eléctrica que decidan tomar esta tecnología para el control eficiente de sus medidores. Es necesario verificar que los datos aquí contenidos sean de calidad, asegurando tanto a usuarios como operadores la certeza de un cobro apropiado.

En búsqueda de esta certificación de calidad, se destaca la norma ISO 25012 que en un modelo de Calidad de Datos se establecen las características de Calidad de Datos que se deben tener en cuenta a la hora de evaluar las propiedades de un producto de datos determinado [4]. En esta normativa se presentan 15 atributos que, al ser evaluados según la necesidad particular del dato, puede determinar su nivel de calidad. Estas características se clasifican en 2 grandes categorías, Inherentes y dependientes como se observa en la tabla 3. Entendiendo las características inherentes como las características que dependen netamente del dato para su evaluación y las dependientes, cuando su evaluación está relacionada directamente con el software y el hardware empleado por la empresa.

Características	Puntos de vista de la DQ	
	Inherente	Dependiente del Sistema
Exactitud	X	
Completitud	X	
Consistencia	X	
Credibilidad	X	
Actualidad	X	
Accesibilidad	X	X
Conformidad	X	X
Confidencialidad	X	X
Eficiencia	X	X
Precisión	X	X
Trazabilidad	X	X
Comprensibilidad	X	X
Disponibilidad		X
Portabilidad		X
Recuperabilidad		X

Tabla 3 Características inherentes y dependientes [41]

Estos grupos nos permiten generar un primer filtro de selección de características, pues, como se definió anteriormente, las características dependientes son muy subjetivas y específicas de las condiciones de la empresa que hace la evaluación, y por tanto no es viable incluir las características de disponibilidad, portabilidad y recuperabilidad a una evaluación general de calidad de datos debido a su especificidad en el proceso de evaluación de calidad.

Una vez analizada en primera instancia las características de calidad, resulta necesario para el proceso de evaluación de calidad relacionar las características de calidad de datos con los datos a evaluar dentro del contexto eléctrico. Para tal fin se presentará un análisis de 3 fuentes o actores presentes y que generan los datos contenidos en las bases de datos de las compañías prestadoras del servicio de energía eléctrica. La primera de estas fuentes corresponde a una visita de campo realizada en la compañía energética de occidente, la segunda atañe a los datos registrados en los entes reguladores SUPERSERVICIOS y en el Servicio único de información de servicios públicos domiciliarios (SUI) y por último y dado que, como se mencionó en el anterior capítulo, este trabajo busca centrarse en redes de medidas directas de capacidades instaladas cercanas e inferiores a los 15 KWh generalmente relacionadas con redes residenciales se analizan los datos entregados directamente a los usuarios de estas compañías generando listas de los datos contenidos en las facturas de las compañías ENEL, CEO, AIRE, EPM,

ElectroHuila y Electrificadora del Caquetá. Finalmente, todos estos datos serán comparados y seleccionados

4.1. Visita de campo, Acompañamiento en el proceso de facturación de la compañía energética de occidente.

Como primer grupo de atributos de datos analizados están los atributos obtenidos de la visita de campo realizada en diciembre de 2019 a la compañía CEO y en la cual se pudo hacer el acompañamiento a un visitador en campo, aquí se pudo observar la interfaz utilizada por la CEO para la recolección de datos de consumo de los usuarios, como se puede apreciar en la figura 30, junto con estos también se encuentran datos del usuario básicos (Nombre, Ciudad, ubicación del medidor y número de contrato) que permiten identificar al usuario y generar de manera correcta el cobro. además de atributos del tipo de medida y medidor asociado. Finalmente se encuentra una casilla vacía en la que el visitador digita la lectura del mes, que posteriormente se albergará en la base de datos junto a las lecturas de meses anteriores, la fecha de lectura se genera automáticamente al anotar la lectura y se encuentra en la parte superior de la interfaz.

De igual manera se logró observar el proceso que se ejecuta en la central de información de la misma compañía, en esta se observó como los datos proporcionados por los visitadores eran albergados y examinados por los operadores de la plataforma, estos datos posteriormente eran usados para generar las facturas de cada uno de los usuarios (de este proceso no se tiene registro gráfico debido a que la interfaz usada por la compañía contiene gran cantidad de información de sus usuarios, y por tanto es de suma reserva por motivos de seguridad). En este proceso se observó las dificultades que se tiene al hacer el proceso de toma de lectura de la forma tradicional, pues los visitadores muchas veces tienen problemas para tomar algunas de las lecturas de medidores con difícil acceso, lo que llevaba a que la compañía generara un cobro por el promedio de viviendas de similares características, o por otro lado, generar un cobro elevado para que los usuarios tuviesen que acercarse a una oficina de la compañía, en la que se le pedía generar el cambio de localización del medidor, para asegurar futuras lecturas correctas.

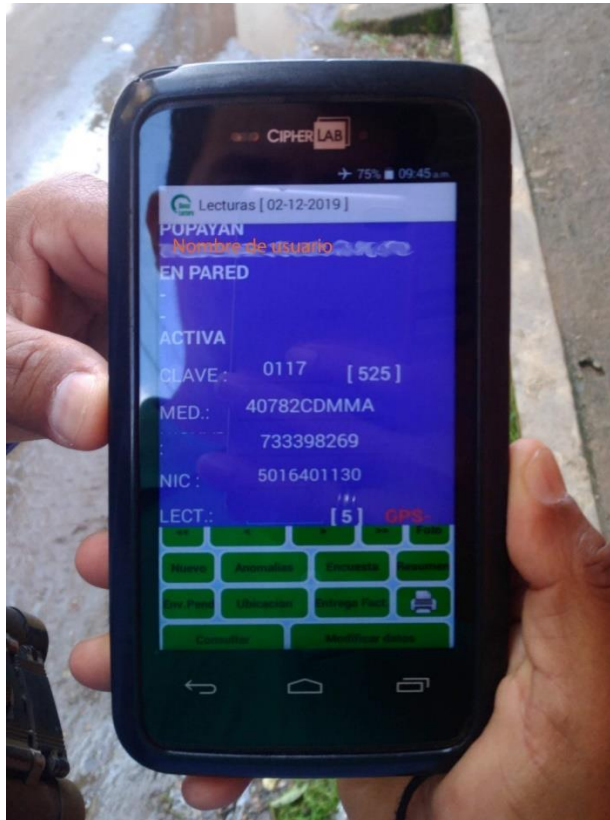


Figura 30 Interfaz de registro de lectura usado por la CEO. Fuente elaboración propia anexo B

Como resultado de esta visita de campo se construyó una lista en la que se pueden observar los atributos relacionados con el proceso de lectura y facturación de la compañía energética de occidente. Dichos atributos de datos se agruparon para facilitar su posterior relación con los atributos obtenidos tanto de los entes reguladores, como de las facturas de las distintas compañías prestadoras del servicio de energía eléctrica. Siendo estos grupos: Datos de usuario, datos de medidores, datos de Facturación y datos de ubicación como se observa en la tabla 4.

Datos de Usuario	Datos de medidores	Datos de facturación	Datos de ubicación
<ul style="list-style-type: none"> • Nombre de usuario • Número de identificación de contrato. 	<ul style="list-style-type: none"> • Clave • Id medidor • Id Transformador • Tipo de energía • posición del medidor 	<ul style="list-style-type: none"> • Lectura • Fecha lectura 	<ul style="list-style-type: none"> • Ciudad.

Tabla 4 Datos visita de campo, fuente elaboración propia.

4.2. Datos en los entes reguladores del servicio público eléctrico.

En segunda instancia se analizaron los atributos datos que los entes de control requieren de las compañías prestadoras del servicio de energía eléctrica, siendo el órgano encargado la superintendencia de servicios públicos domiciliarios o también llamada SUPERSERVICIOS. Dicho ente solicita periódicamente información a las distintas empresas de servicios públicos (E.S.P) para hacer un control constante de las operaciones y cobros de estas empresas, para tal fin se dispuso del Sistema Único de Información de servicios públicos domiciliarios (SUI) en el que los ciudadanos en general pueden consultar los datos que las distintas empresas de servicios públicos han presentado a la SUPERSERVICIOS.

Para la realización de este análisis, se generó una revisión general de los distintos formularios dispuestos por la SUPERSERVICIOS para la solicitud de información de este ente regulador a las empresas prestadoras del servicio eléctrico. Dicha revisión concluyo con el análisis a detalle del formato ZNI C6, como se aprecia en la tabla 5, dicho formulario fue seleccionado dado la gran cantidad de datos solicitados, y su clara descripción.

Superservicios Superintendencia de Servicios Públicos Domiciliarios		FICHA TÉCNICA FORMULARIOS / FORMATOS SUI			
SERVICIO	ENERGIA	CLASIFICACIÓN	N/A	FORMATO / FORMULARIO	FORMATO ZNI C6
NOMBRE FORMATO	INFORMACION SUSCRIPTORES COMUNITARIOS EN LOCALIDADES MENORES			PERIODICIDAD REPORTE	SEMESTRAL
TÓPICO	COMERCIAL	ACTO ADMINISTRATIVO	CIRCULAR SSPD - IPSE NO. 0001 DE 2004	No. CAMPOS	43
AÑOS DISPONIBLES	2014-2017				
DEFINICIÓN DE CAMPOS					
IDENTIFICADOR EMPRESA	Corresponde al código de identificación en base de datos, de la empresa que realiza el reporte de la información al SUI.				
EMPRESA	Nombre de la empresa que reporta la información, conforme a como esta registrado en el Registro Único de Prestadores de Servicios - RUPS.				
AÑO	Año de carga de la información.				
TRIMESTRE	Corresponde al trimestre de carga de la información.				
MES	Corresponde al periodo de reporte de la información solicitada.				
CODIGO_DANE	Corresponde a la codificación dada por el DANE a la división político-administrativa de Colombia. Con la siguiente estructura: DDMMMCC, donde DD es el código del departamento, MMM corresponde al código del municipio y CCC corresponde al código del centro poblado. Para los casos en que no aplique el centro poblacional se debe diligenciar 000. Si la localidad no se encuentra en la codificación DANE, se registrará el código de la cabecera municipal.				
DEPARTAMENTO	Corresponde al nombre del departamento donde se expidió la facturación.				
MUNICIPIO	Corresponde al nombre del municipio donde se expidió la facturación.				
CENTRO_POBLADO	Corresponde al nombre del Centro poblado donde se expidió la facturación.				
NOMBRE_LOCALIDAD	Es el nombre de la localidad menor en mayúsculas y como aparece en las resoluciones del IPSE, en donde se está aplicando el esquema de medición comunitaria tal y como se estableció en el Decreto 4974 de 2007 del Ministerio de Minas y Energía.				
NIU	Número de identificación del suscriptor comunitario asignado por el comercializador. - (Variable dissociada - Reserva por razones de habeeas data)				
ID_FACTURA	Número de la factura o identificación de la factura asignada por el prestador. - (Variable dissociada - Reserva por razones de habeeas data)				
FECHA_EXPEDICION	Se refiere a la fecha de expedición de la factura comunitaria, de acuerdo al formato establecido en el Anexo A.				
FECHA_INICIO	Se refiere a la fecha desde la cual comienza a registrarse el consumo a facturar, de acuerdo al formato establecido en el Anexo A.				
DIAS_FACTURADOS	Corresponde al número de días facturados en el periodo de facturación.				
NUMERO_HOGARES	Corresponde al número de hogares del sector residencial que están inmersos en la zona especial de prestación del servicio y los cuales hacen parte del acuerdo comunitario.				
USUARIOS ESTRATO1	Corresponde al número de usuarios de estrato 1 que están inmersos en la zona especial de prestación del servicio y los cuales hacen parte del acuerdo comunitario.				
USUARIOS ESTRATO2	Corresponde al número de usuarios de estrato 2 que están inmersos en la zona especial de prestación del servicio y los cuales hacen parte del acuerdo comunitario.				
USUARIOS ESTRATO3	Corresponde al número de usuarios de estrato 3 que están inmersos en la zona especial de prestación del servicio y los cuales hacen parte del acuerdo comunitario.				
USUARIOS ESTRATO4	Corresponde al número de usuarios de estrato 4 que están inmersos en la zona especial de prestación del servicio y los cuales hacen parte del acuerdo comunitario.				
USUARIOS ESTRATO5	Corresponde al número de usuarios de estrato 5 que están inmersos en la zona especial de prestación del servicio y los cuales hacen parte del acuerdo comunitario.				
USUARIOS ESTRATO6	Corresponde al número de usuarios de estrato 6 que están inmersos en la zona especial de prestación del servicio y los cuales hacen parte del acuerdo comunitario.				
USUARIOS OFICIAL	Corresponde al número de usuarios oficiales que están inmersos en la zona especial de prestación del servicio y los cuales hacen parte del acuerdo comunitario.				
USUARIOS COMERC INDUSTRIAL	Corresponde al número de usuarios comerciales e industriales que están inmersos en la zona especial de prestación del servicio y los cuales hacen parte del acuerdo comunitario.				
CONSUMO ESTRATO1	Consumo total de los usuarios de estrato 1 del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
CONSUMO ESTRATO2	Consumo total de los usuarios de estrato 2 del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
CONSUMO ESTRATO3	Consumo total de los usuarios de estrato 3 del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
CONSUMO ESTRATO4	Consumo total de los usuarios de estrato 4 del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
CONSUMO ESTRATO5	Consumo total de los usuarios de estrato 5 del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
CONSUMO ESTRATO6	Consumo total de los usuarios de estrato 6 del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
CONSUMO OFICIAL	Consumo total de los usuarios oficiales del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
CONSUMO COMERC INDUSTRIAL	Consumo total de los usuarios comerciales e industriales del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
CONSUMO TOTALIZADOR	Energía medida en el equipo de medición comunitaria.				
FACTURACION ESTRATO1	Facturación total de los usuarios de estrato 1 del acuerdo comunitario, utilizado para liquidar la factura comunitaria (incluye subsidios).				
FACTURACION ESTRATO2	Facturación total de los usuarios de estrato 2 del acuerdo comunitario, utilizado para liquidar la factura comunitaria (incluye subsidios).				
FACTURACION ESTRATO3	Facturación total de los usuarios de estrato 3 del acuerdo comunitario, utilizado para liquidar la factura comunitaria (incluye subsidios).				
FACTURACION ESTRATO4	Facturación total de los usuarios de estrato 4 del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
FACTURACION ESTRATO5	Facturación total de los usuarios de estrato 5 del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
FACTURACION ESTRATO6	Facturación total de los usuarios de estrato 6 del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
FACTURACION OFICIAL	Facturación total de los usuarios oficiales del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
FACTURACION COMERC INDUSTRIAL	Facturación total de los usuarios comerciales e industriales del acuerdo comunitario, utilizado para liquidar la factura comunitaria.				
FACTURACION COMUNITARIO	Valor de la facturación comunitaria en el periodo de facturación reportado (incluye subsidios).				
SUBSIDIO SUSC COMUNITARIO	Monto total de los subsidios asignados al suscriptor comunitario.				
				FECHA ACTUALIZACION	14/3/2018
				FICHA TÉCNICA	

Tabla 5 Formato ZNI C6 Fuente [42]

Al igual que con la categoría anterior, los atributos de datos obtenidos de este formato fueron agrupados por grupos, esta vez se adicióno la categoría de Datos

de tiempo. Igualmente, este formato es complementado con los datos hallados en la plataforma O3, usada como herramienta de análisis de información, y para facilitar la veeduría ciudadana de las empresas prestadoras de servicios públicos, como se puede apreciar en la tabla 6.

ESTRATO	MERCADO	USO	TIEMPO	Medidas		Facturación total (C)	Consumo total (kWh)	Número de Suscripciones	Valor Facturado por kW consumido (C/kWh)	Promedio de Consumo (kWh/Suscriptor)	Promedio de Facturación por consumo (C/Suscriptor)	Promedio de Facturación total (C/Suscriptor)	Contribuciones (C)	Subsidios (C)	Total
				Facturación por consumo (C)	Consumo (kWh)										
Estrato 1	MERCADO	Residencial	2020	84.276.703.241 (C)	1.711.079.756 (kWh)	21.376.176kwh	277.423	439.5 (C/kWh)	82.8 kWh/usuc	52.254 (C/usuc)	211.433 (C/usuc)	0 (C)	2.555.707.285 (C)	0 (C)	85.933.408.526 (C)
Estrato 2	MERCADO	Residencial	2020	18.177.878.853 (C)	4.110.055.583 (kWh)	3.502.177kwh	10.005	820.4 (C/kWh)	104.8 kWh/usuc	84.855 (C/usuc)	230.039 (C/usuc)	0 (C)	1.852.328.058 (C)	0 (C)	22.507.708.611 (C)
Estrato 3	MERCADO	Residencial	2020	3.872.862.215 (C)	1.824.088.149 (kWh)	3.094.294kwh	28.000	821.8 (C/kWh)	187.0 kWh/usuc	46.108 (C/usuc)	133.892 (C/usuc)	0 (C)	221.604.642 (C)	0 (C)	8.021.107.272 (C)
Estrato 4	MERCADO	Residencial	2020	2.010.000.078 (C)	1.080.169.191 (kWh)	1.787.204kwh	15.783	416.3 (C/kWh)	172.0 kWh/usuc	60.048 (C/usuc)	132.988 (C/usuc)	0 (C)	151.223.214 (C)	0 (C)	3.001.602.299 (C)
Estrato 5	MERCADO	Residencial	2020	454.358.558 (C)	398.019.738 (kWh)	495.395kwh	3.500	415.7 (C/kWh)	178.5 kWh/usuc	78.781 (C/usuc)	169.573 (C/usuc)	0 (C)	0.040.693.283 (C)	0 (C)	1.016.018.832 (C)
Estrato 6	MERCADO	Residencial	2020	315.728.300 (C)	68.909.777 (kWh)	111.888kwh	608	415.8 (C/kWh)	143.1 kWh/usuc	180.438 (C/usuc)	418.904 (C/usuc)	0 (C)	151.788.134 (C)	0 (C)	388.681.978 (C)
Industrial	MERCADO	No Residencial	2020	2.902.181.888 (C)	3.018.977.862 (kWh)	8.180.573kwh	110	380.0 (C/kWh)	4.326.8 kWh/usuc	1.671.435 (C/usuc)	3.408.958 (C/usuc)	0 (C)	181.871.321 (C)	0 (C)	61.188.803.643 (C)
Comercial	MERCADO	No Residencial	2020	23.584.005.249 (C)	3.311.758.995 (kWh)	5.623.781kwh	1.203	588.9 (C/kWh)	229.450 (C/usuc)	3.851.648 (C/usuc)	187.021.418 (C/usuc)	0 (C)	187.021.418 (C)	0 (C)	27.585.608.178 (C)
Oficial	MERCADO	No Residencial	2020	4.716.175.043 (C)	1.068.521.333 (kWh)	1.927.991kwh	5.643	551.6 (C/kWh)	332.0 kWh/usuc	293.418 (C/usuc)	1.294.695 (C/usuc)	0 (C)	0 (C)	0 (C)	1.789.454.483 (C)
Albergado Público	MERCADO	No Residencial	2020	1.218.586.211 (C)	415.766.214 (kWh)	1.637.236kwh	783	550.6 (C/kWh)	214.3 kWh/usuc	114.243 (C/usuc)	1.246.657 (C/usuc)	0 (C)	0 (C)	0 (C)	2.005.188.042 (C)
Total				120.012.662.971 (C)	29.807.518.748 (kWh)	51.516.294kwh	382.430	574.6 (C/kWh)	6.411.9 kWh/usuc	1.748.374 (C/usuc)	2.418.654 (C/usuc)	0 (C)	10.064.519.220 (C)	0 (C)	162.080.642.561 (C)

Tabla 6 Tabla O3 Facturación de energía. Fuente [42]

Esta tabla intercepta 2 tipos de dato entre sí, el primer grupo de estos datos, encontrado a la izquierda, ilustra el estrato y uso del bien en el que se presta el servicio además del año a que hace referencia. El segundo grupo de datos está ubicado en la parte superior, y evoca los atributos de consumo, facturación, cobros, subsidios y aportes realizados por los usuarios en función a el uso y estrato encontrados en el primer grupo de datos. Adicionalmente se muestra cuantos usuarios registrados hay en cada estrato y uso.

Este sistema de tablas funciona de igual manera para todas las compañías prestadoras del servicio de energía eléctrica. Pero, para ejemplificar este documento, se seleccionó solo los datos albergados en la plataforma O3, correspondientes a la compañía energética de occidente para el año efectivo 2020.

Finalmente, y al igual que el anterior grupo de datos, se generó una lista de datos como se observa en la tabla 7, que posteriormente se relacionara con los datos obtenidos de los otros 2 grupos. En cuanto a los Atributos encontrados en este grupo, se encuentra una gran cantidad de datos generados, por ejemplo, promedios de consumo y promedio de facturación. Que como se explicó al inicio de este capítulo serán descartados, debido que para obtener estos datos solo hay que hacer una operación matemática de otros datos ya almacenados.

Datos de Usuario	Datos de facturación	Datos de ubicación	Datos de tiempo
<ul style="list-style-type: none"> Número de identificación de contrato Número de hogares Estrato USO 	<ul style="list-style-type: none"> Id factura Fecha de expedición días Facturados Valor de facturación contribución Subsidios 	<ul style="list-style-type: none"> Código DANE Municipio Centro poblado Localidad Departamento 	<ul style="list-style-type: none"> Año Trimestre Mes

Tabla 7 Datos entes de control. Fuente Elaboración propia

4.3. Facturas de energía eléctrica.

Finalmente se realizó una recopilación de facturas de 6 compañías de energía eléctrica del país, con la intención de mostrar las variaciones y condiciones especiales de distintos departamentos del país en cuanto a facturación de energía eléctrica se trata, en estas facturas se generó una búsqueda de atributos los cuales determinan las características en cuanto a facturación y pretende ser una muestra general de este registro.

Inicialmente, se encontró en las facturas residenciales de electricidad entregados por la compañía ENEL a sus usuarios, varios campos con una clara definición del contenido encontrado en dichos espacios, estos, se encuentran en ambas caras de la hoja de la factura.

En la primera de las páginas, como se aprecia en la figura 31, se encuentra la mayoría de los datos visibles en esta factura, encontrando 2 cuadros de información en la parte superior, uno relacionado a los datos del cliente, y el otro a información del contrato. Adicionalmente se encuentra el número del contrato en la parte superior derecha.

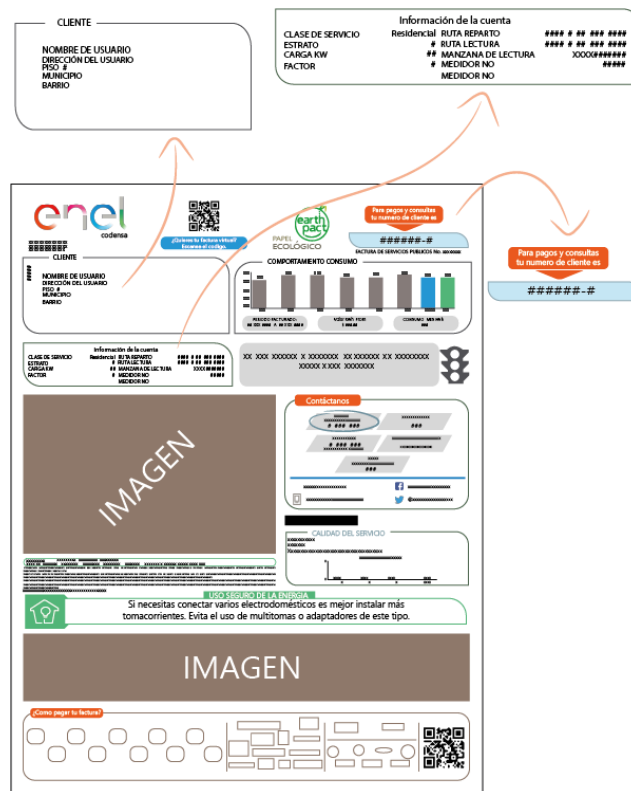


Figura 31 Esquema factura ENEL pagina 1. Fuente elaboración propia.

La segunda página, como se aprecia en la figura 32, la factura contiene lo relacionado a información de facturación y nuevamente el número de contrato.



Figura 32 Esquema de factura ENEL pagina 2. Fuente elaboración propia

Por su parte, la compañía CEO organiza en una única página todos los datos como se aprecia en la figura 33. Se puede observar cinco tipos de datos agrupados, dos de estos se repiten (el número de contrato y periodo de consumo) los grupos de datos restantes aportan información técnica sobre el contrato y cobro del servicio, datos de usuario e identificadores de medidores y transformadores.

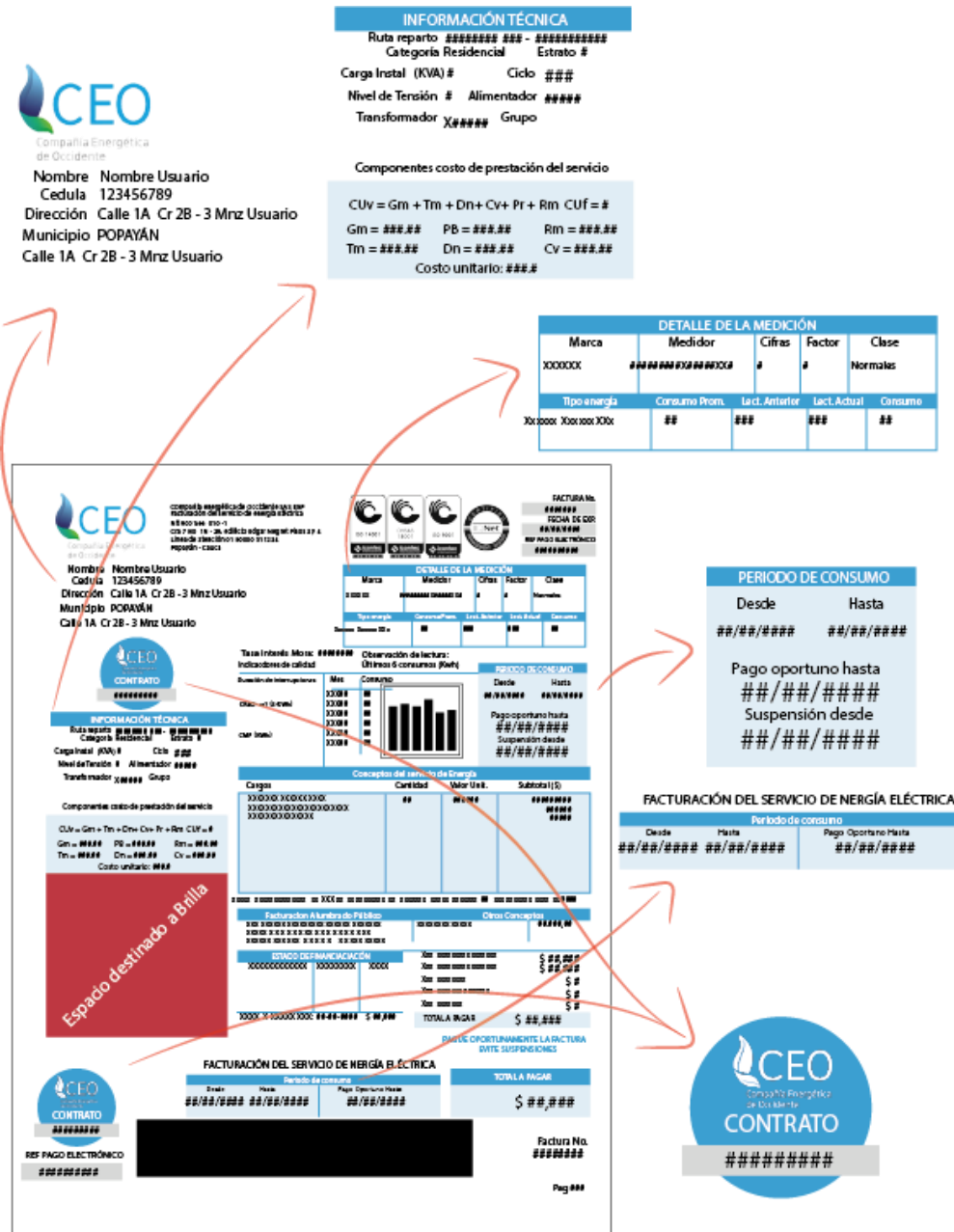


Figura 33 Esquema factura CEO. Fuente elaboración propia.

En cuanto a la factura perteneciente a AIR-E, distribuye los datos en las dos caras de la hoja, en la primera de estas, como se aprecia en la figura 34, se encuentra el logo de la compañía junto con un resumen de los cobros asociados al total final. Igualmente se encuentran los datos de usuario, información del consumo, información de lectura y datos relacionados al consumo.

DATOS DEL USUARIO Y/O SUSCRIPTOR

Titular de Pago XXXXXX XX	Dirección de suministro XXXXXXXX XXX	Dirección de suministro XXXXXXXX XXX
Usuario o suscriptor XXXXXX XX	Barrio Municipio	Barrio Municipio
Estrato/Clasificación: Resid. Estrato # E.Caribe	NIU: #####	

INFORMACIÓN DEL CONSUMO

Consumo	Tarifa en \$/KWh	Consumo kWh	Valores \$
Consumo	### x ##	##	###.##
Subsidio	### x ##	##	-###.##

INFORMACIÓN DE LECTURA

Fecha Lectura Anterior	Fecha Lectura Actual	Días Facturados
##/##/##	##/##/##	##
Medidor	Lectura Actual	Factor Multiplo
XXXXXX	XXXXXX	##
Consumo kWh	##	##

NIC: #####

Operador de Red: Air-e S.A.S. E.S.P.
 Dirección: Calle 27B # 59B-27 Biquita

NIC: #####

Valor a pagar

Anterior Consumo	Actual Consumo	Variación Anterior	\$ ###.###
###	###	###	\$ ###.###
			\$ ###.###
			\$ ###.###
			\$ ###.###
Pago oportuno	Suspensión a partir de	TOTAL MES	\$ ###.###
##/##/####	##/##/####	##	\$ ###.###

DATOS DEL USUARIO Y/O SUSCRIPTOR

Titular de Pago XXXXXX XX	Dirección de suministro XXXXXXXX XXX	Dirección de suministro XXXXXXXX XXX
Usuario o suscriptor XXXXXX XX	Barrio Municipio	Barrio Municipio
Estrato/Clasificación: Resid. Estrato # E.Caribe	NIU: #####	

INFORMACIÓN DEL CONSUMO

Consumo	Tarifa en \$/KWh	Consumo kWh	Valores \$
Consumo	### x ##	##	###.##
Subsidio	### x ##	##	-###.##

INFORMACIÓN DE LECTURA

Fecha Lectura Anterior	Fecha Lectura Actual	Días Facturados
##/##/##	##/##/##	##
Medidor	Lectura Actual	Factor Multiplo
XXXXXX	XXXXXX	##
Consumo kWh	##	##

Nivel actual en la lectura: # Propiedad del Activo: Empresa

Consumo mes \$ ###.## Consumo día \$ ##

Consumos últimos 6 meses kWh

Consumo actual

DE INTERES

Propiedad Día \$ ##

IMAGEN

Documentos Vencidos: # NIC: ##### Pago oportuno: ##/##/####

Total a pagar: \$###.### Total Mes: \$###.###

Representante legal

Figura 34 Esquema factura Air-e Pag 1. Fuente elaboración propia.

En la segunda página, como se aprecia en la figura 35, se encuentran los desgloses de los cobros asociados, y el cálculo de la fórmula de costo unitario. Adicionalmente se encuentra nuevamente el número de contrato.

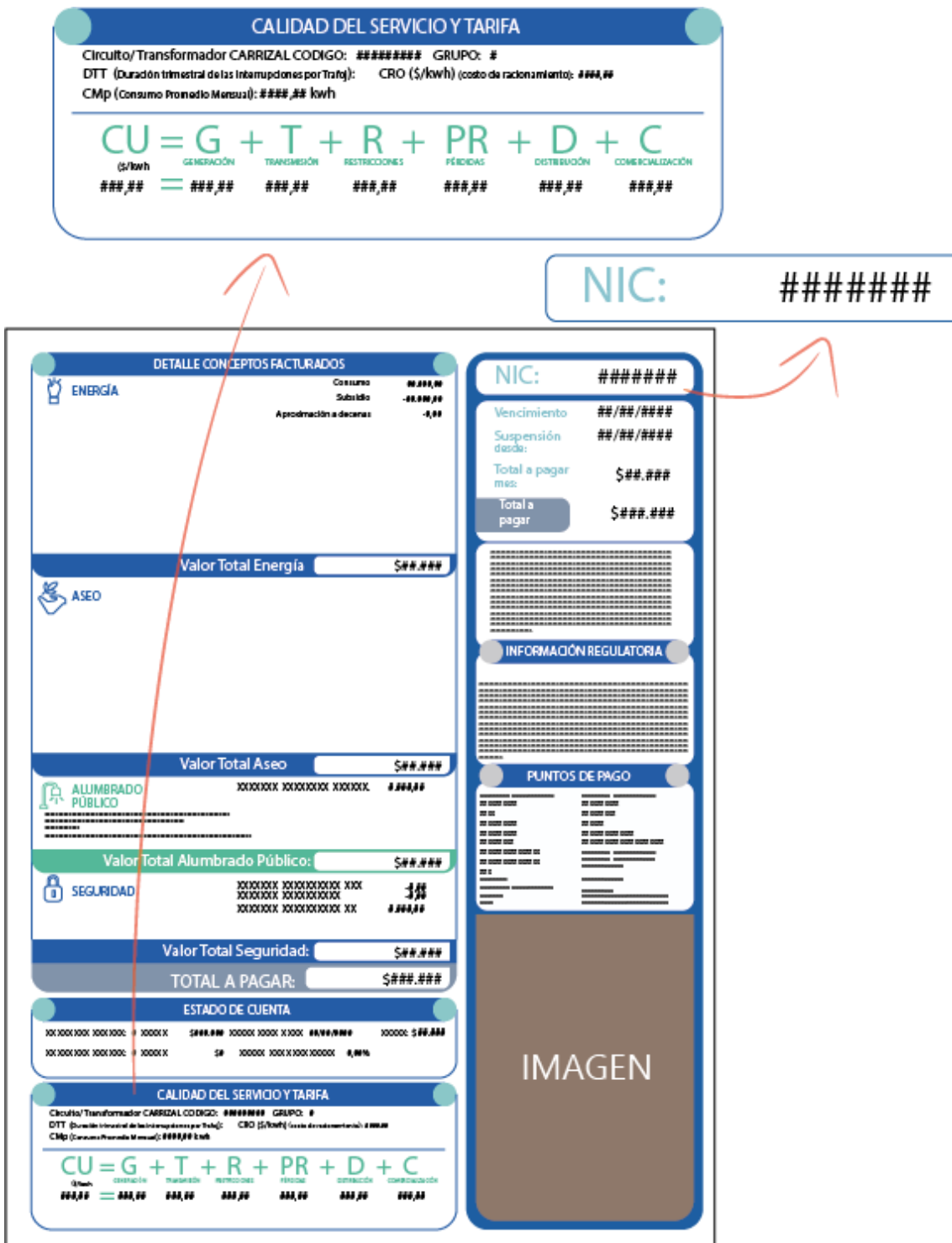


Figura 35 Esquema factura Air-e Pag 2. Fuente elaboración propia.

El caso de EPM es sumamente particular, como se aprecia en la figura 36, pues en la misma factura se encuentran los cobros de acueducto, gas, electricidad, aseo y alcantarillado, por lo que la información suministrada por esta compañía es muy reducida, solo se puede encontrar algunos datos personales y numero de contrato, identificador del medidor y datos del consumo y facturación.



Figura 36 Esquema factura EPM. Fuente elaboración propia.

Para el caso de ElectroHuila, La empresa divide los datos en 2 páginas, la primera, como se aprecia en la figura 37, Contiene información del usuario y sus datos personales, además de algunos datos del contador y ruta de reparto. Adicionalmente, se encuentra el número de contrato.

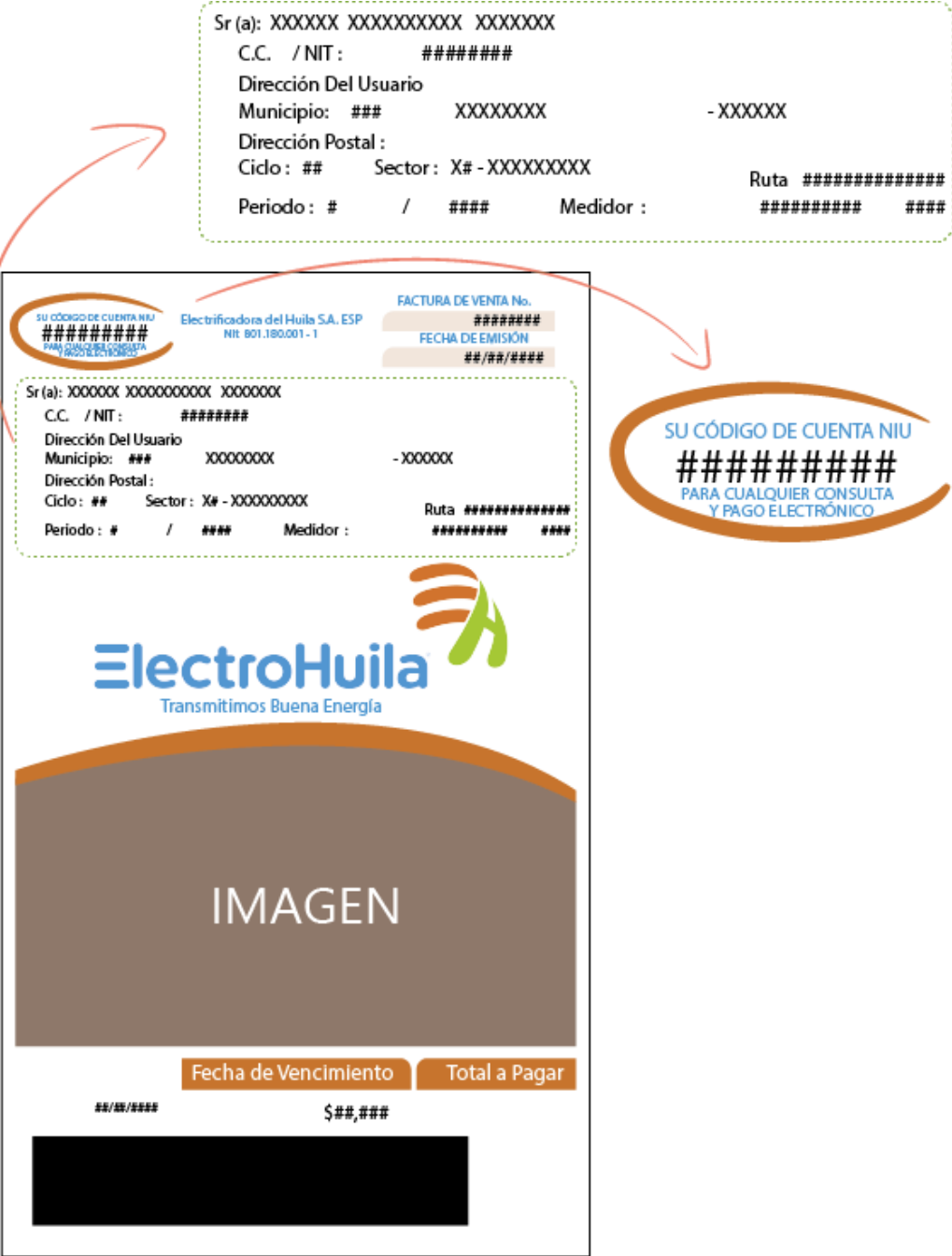


Figura 37 Esquema factura ElectroHuila Pagina 1. Fuente elaboración propia.

En la segunda página, como se aprecia en la figura 38, se generan distintos recuadros de información, entre estos se encuentran nuevamente los datos personales, datos necesarios para determinar el consumo y el periodo de dicho consumo, y el valor unitario considerado para el mes facturado.

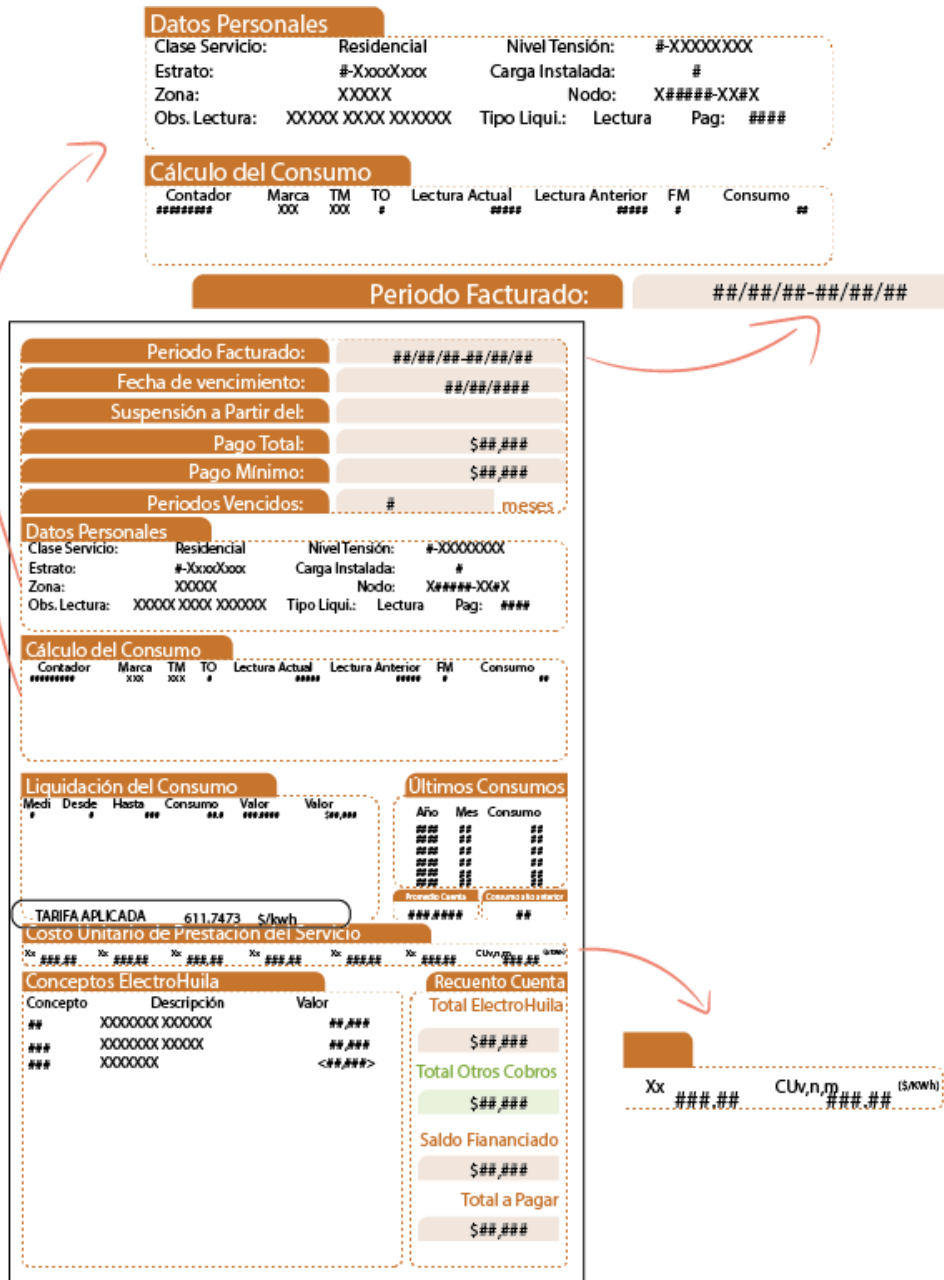


Figura 38 Esquema factura ElectroHuila Pagina 2. Fuente elaboración propia.

La última factura analizada corresponde a la de la electrificadora del Caquetá, como se aprecia en la figura 39. En esta se presenta la mayor cantidad de datos técnicos del suministro eléctrico, datos de contacto, y relacionados al contrato, identificadores de medidores y transformadores, cálculos de valores aplicados en la facturación al igual que los consumos del periodo facturado.

ELECTRIFICADORA DEL CAQUETA S.A ESP
 Carrera 1 N35- 99A, A. 404 Florencia - Caquetá NIT 891.190.127-3 NLR 2-1800 1000 pbs4355400
 Fax 4355414 www.electrocaqueta.com.co Linea gratuita 01 800 89 6812

INFORMACIÓN CLIENTE: FACTURA No: ##### - #####
 Nombre de usuario Tel.:
 Dirección de usuario **CLASE DE SERVIDO** Residencial.

INFORMACIÓN TÉCNICA

Matricula	Ruta	Municipio	Concentración	Estrato		
#####	#####	XXXXXXXXXX	###	# X,XXXX		
Nodo	Circuito	Grupo	Carga	Dtt	CRO m - I	Cmp
###	###	###	###	###	###	###
XX	XXXXXXXXXX					
#	#	#	#	#	###.##	###.##

PERIODO FACTURADO DIAS FACTURADOS **VENCIMIENTO** **SUSPENSIÓN**
 ##/##/## a ##/##/## ##/##/## ##/##/##

INFORMACIÓN CONSUMO

Serie Contador	Marca	TE	Lec. Ante.	Lec. Actual	Prom	Factor	Consumo
#####	XXX	X	#####	#####	##	#	##

COSTO UNITARIO DE PRESTACIÓN DEL SERVICIO

Gm	Tm	PRum:	Dm:cc	Rm:	Cm:cc	Cu (\$/kwh)
###.##	###.##	###.##	###.##	###.##	###.##	###.##

LIQUIDACIÓN DE ALUMBRADO PÚBLICO Y ASEO

Valor factura periodo: \$ 000.000
 Compensación DES: \$ 000.000
 Ajuste al Precio: \$ 000.000

VALOR A PAGAR: \$ 000.000

Nombre de usuario: Celé
 Fecha Vencimiento: ##/##/##
 Fecha Corte: ##/##/##

MEC: XXXXXXXX \$ 000.000
 DISTRIBUCIÓN: ### - XXXXXXXX \$ 000.000
 MATRÍCULA: ##### \$ 000.000
VALOR TOTAL A PAGAR: \$ 000.000

Figura 39 Esquema factura de Electro Caquetá. Fuente elaboración propia.

Al igual que los grupos anteriores, este grupo de atributos se consolida en la siguiente lista que se puede observar en la tabla 8, para posteriormente ser comparado con los atributos encontrados en los otros 2 grupos. Para el caso de este último grupo de atributos, se encontraron datos de todos los grupos de atributos ya mencionados, y se adiciona el grupo de Datos de contrato.

Datos de Usuario	Datos de facturación	Datos de ubicación	Datos de medidores	Datos de tiempo	Datos de contrato
<ul style="list-style-type: none"> • Nombre • Estrato • Teléfono • Numero de Documento 	<ul style="list-style-type: none"> • Lectura • Factor de multiplicación • Tarifa aplicada 	<ul style="list-style-type: none"> • Dirección • Municipio 	<ul style="list-style-type: none"> • Id medidor • Id transformador • Clase • Lugar de instalación • Marca • Cifras • Clave 	<ul style="list-style-type: none"> • Periodo facturado 	<ul style="list-style-type: none"> • Clase de servicio • Numero de contrato • Nivel de tensión • Carga instalada • Ruta de reparto • Tipo de energía

Tabla 8 Datos de facturas. Fuente Elaboración propia

4.4. Selección de características de calidad de datos

Una vez analizados las 3 fuentes de grupos de datos dentro del contexto eléctrico, siendo estas Empresa (visita de campo a la CEO) Entes de control (Superservicio y SUI) y Usuarios (Facturación) se realiza una comparación, como se puede observar en la tabla 9, con la finalidad de decantar todos los atributos hallados y generar un único paquete de atributos de datos para la futura generación de estrategias que permitan determinar la calidad de datos dentro del contexto eléctrico.

Fuente	Datos de Usuario	Datos de facturación	Datos de ubicación	Datos de medidores	Datos de tiempo	Datos de contrato
Facturación	<ul style="list-style-type: none"> Nombre Estrato Teléfono Numero de Documento 	<ul style="list-style-type: none"> Lectura Factor de multiplicación Tarifa aplicada 	<ul style="list-style-type: none"> Dirección Municipio 	<ul style="list-style-type: none"> Id medidor Id transformador Clase Lugar de instalación Marca Cifras Clave 	<ul style="list-style-type: none"> Periodo facturado 	<ul style="list-style-type: none"> Clase de servicio Numero de contrato Nivel de tensión Carga instalada Ruta de reparto Tipo de energía
Visita	<ul style="list-style-type: none"> Nombre de usuario Número de identificación de contrato. 	<ul style="list-style-type: none"> Lectura Fecha lectura 	<ul style="list-style-type: none"> Ciudad. 	<ul style="list-style-type: none"> Clave Id medidor Id Transformador Tipo de energía posición del medidor 		
Ente de control	<ul style="list-style-type: none"> Número de identificación de contrato Número de hogares Estrato USO 	<ul style="list-style-type: none"> Id factura Fecha de expedición Días facturados Valor de facturación contribución Subsidios 	<ul style="list-style-type: none"> Código DANE Municipio Centro poblado Localidad Departamento 		<ul style="list-style-type: none"> Año Trimestre Mes 	

Tabla 9 Comparación final de atributos de datos por cada fuente. Fuente Elaboración propia.

Como resultado del proceso de análisis de la visita de campo, los entes de control y las facturas de estas compañías, se encontraron un total de 23 atributos de datos que fueron clasificados en función a su relevancia en la operación de la empresa y su relevancia tanto para los usuarios como para los entes reguladores, reduciendo así a solo 2 categorías; críticos y secundarios, como se puede observar en las tablas 10 y 11. Lo que corresponde igualmente a un total de 17 datos críticos y 6 datos secundarios clasificando el total de atributos anteriormente registrados. Dichos atributos son.

Críticos	
Dato	Descripción
Nombre	Nombre completo del propietario del inmueble donde está instalado el servicio de energía eléctrica.
Dirección	Dirección del inmueble donde está instalado el servicio de energía eléctrica.
Municipio	Lugar donde está instalado el servicio de energía eléctrica.
Clase de servicio – Categoría	Clasificación del tipo de servicio instalado (residencial).
Estrato	Nivel de estratificación del inmueble en el cual se instala el servicio de energía eléctrica.
Factor de multiplicación	constante aplicada por la entidad prestadora de servicio de energía eléctrica.
ID de medidor	Numero identificador del medidor instalado en el inmueble
No Contrato	numero identificador asignado al servicio de energía eléctrica instalado.
Periodo facturado	Rango de fechas.
Lectura – Consumo	Total del consumo de energía eléctrica en un determinado rango de fechas.
No Documento	Número de identificación del propietario del del inmueble donde está instalado el servicio de energía eléctrica
ID Transformador	Número de identificador del transformador al cual está conectado el medidor instalado en el inmueble
Clase	Porcentaje de error de la medida
Nivel de tensión	Baja, media o alta
Teléfono	Número telefónico del propietario del inmueble donde está instalado el servicio de energía eléctrica
Carga instalada	Capacidad de energía eléctrica disponible instalada en el servicio
Tarifa aplicada Kw – KVA-r	Valor asignado para el cobro del consumo de energía eléctrica

Tabla 10 Datos críticos. Fuente elaboración propia.

Secundarios	
Dato	Descripción
Ruta de reparto	Número asociado a las líneas de distribución donde se encuentra instalado el servicio
Tipo de energía	Energía activa / energía reactiva
Lugar de instalación	Ubicación donde se encuentra instalado el servicio
Marca	Marca del medidor
Cifras	Número de dígitos del medidor
Clave	Número identificador asociado al inventario de los medidores por parte de la entidad prestadora del servicio

Tabla 11 Datos secundarios. Fuente elaboración propia.

Una vez determinados los atributos o lista de datos que se encuentran en el contexto eléctrico surgen la necesidad de encontrar un proceso para llevar a cabo la evaluación de calidad, para esto, se ha tomado como base la ISO/IEC 25040 que establece el proceso de evaluación para productos de software, esta ISO propone 5 pasos o actividades claramente definidas para la evaluación de productos de software. Dichas Actividades son.

Actividad 1: Establecer los requisitos de la evaluación. Esta actividad consiste en establecer el propósito de la evaluación, Identificar las partes o actores, riesgos (si los hubiera) y el modelo de calidad a usar. Para el caso de este trabajo las partes definen los propósitos de investigación, siendo estos actores los entes de control, las empresas prestadoras del servicio eléctrico y los usuarios tal y como se ha presentado en el actual capítulo, y el modelo de calidad a usar, será la ISO/IEC 25012 como ya se ha mencionado anteriormente.

Actividad 2: Especificar la evaluación. Dentro de esta actividad se especifican los módulos de evaluación (métricas, herramientas y técnicas) junto con los criterios de evaluación de los mismos. Esta actividad corresponde al proceso que se llevara a cabo en el próximo capítulo, y establece por medio de un formato la evaluación que relaciona lo ya dicho en esta actividad.

Actividad 3: Diseñar la evaluación. En esta actividad se define el proceso que permite realizar la evaluación de calidad de los datos, al diseñar la evaluación se especifica la documentación que permita identificar y asignar las características que serán evaluadas, así como las reglas de evaluación de cada una de ellas.

Actividad 4: Ejecutar la evaluación. En esta actividad se realiza el proceso de evaluación que consiste en ejecutar las consultas SQL mapeadas a partir de las reglas diseñadas en la actividad anterior.

Actividad 5: Concluir la evaluación. En esta actividad se presentan las tablas con el resultado de la ejecución de la evaluación y se culmina el proceso de evaluación con el informe de resultados finales.

4.4.1. Establecer los requisitos de la evaluación

Una vez definido el proceso a seguir para conseguir la evaluación de calidad de datos se determina los propósitos de la evaluación, dichos propósitos surgen del interés y la postura encontrada en los datos anteriormente presentados de cada uno de los actores del contexto eléctrico.

Ente de control

Los datos hallados en las bases de datos de los entes de control, y la manera en la que estos son presentados, permite suponer el interés principal de este actor. Siendo este interés el de analizar las relaciones de consumo, facturación, cobro, uso y nivel socioeconómico de los usuarios, todo esto enlazado directamente con un periodo temporal específico (mes, trimestre, semestre, año) finalmente estos datos y relaciones permiten generar índices estadísticos que facilitan su lectura al público general, y que a futuro permite generar comparaciones con periodos de tiempo similares.

Por esta razón es de suponer que el propósito de evaluación por parte de los entes de control es determinar si los datos considerados necesarios están presentes, si dichos datos cumplen con los formatos y parámetros esperados tanto por las convenciones fijadas por la empresa (caracteres, tipo de carácter, cantidad de carácter, ETC) como por las convenciones normativas actuales, adicionalmente determinar la facilidad de lectura de los datos y considerar que tan actualizados se encuentran. Con tal propósito, se pueden determinar las características de Exactitud (Semántica y Sintáctica), Completitud, Conformidad, Comprensibilidad y Actualidad como las características de interés para la evaluación por parte de este actor.

Usuario

Por otro lado, el usuario lejos de estar interesado en las especificidades técnicas del servicio, o de hacer un control que relacione grupos poblacionales con cobros y periodos de tiempo, solo requiere claridad en las facturas que permita hacer un análisis de su consumo, relacionar ese consumo con consumos anteriores y finalmente generar los pagos por los servicios tomados.

Por esto que el propósito de evaluación por parte del usuario es determinar que los datos considerados necesarios estén presentes, que dichos datos sean ciertos y creíbles, analizar si los datos son coherentes, determinar la facilidad de su lectura, y considerar que tan actualizados se encuentran, por esto, las características de interés de este actor son Exactitud, Credibilidad, Consistencia, Comprensibilidad y Actualidad.

Empresa

Finalmente la empresa al ser la portadora de todos los datos, y ser la encargada de realizar tanto los procesos de facturación para los usuarios, como de entregar los datos requeridos por los entes de control, requiere de una revisión más exhaustiva de sus datos, lo que supone que agruparía los propósitos de los otros

dos actores, sin embargo, es de gran interés para la operación de la empresa hacer un enfoque en los procesos de facturación y cobranza, pues de esto depende la estabilidad financiera de la empresa y por ende, que la empresa pueda seguir operando a futuro. Por esta razón, este actor recoge varios de los propósitos de los actores anteriores, pero adiciona algunos más. Siendo así, el propósito de evaluación de este actor sería, determinar que los datos considerados obligatorios estén presentes, analizar que los datos cumplan con los formatos esperados, que dichos datos sean coherentes y libres de contradicción, e igualmente que sean ciertos y creíbles, considerar que tan actualizados están los datos, que cumplan con las convenciones normativas actuales aplicables, determinar el nivel de seguridad de acceso a los datos confidenciales, considerar la facilidad de indagación del registro de acontecimientos en la base de datos, y determinar la facilidad de lectura de los datos. Consecuentemente las características de interés para la evaluación por parte de este actor son Exactitud (Semántica y Sintáctica), Completitud, consistencia, credibilidad, actualidad, conformidad, confidencialidad, trazabilidad y comprensibilidad.

Siendo así, se presentan las características de interés de cada una de las partes interesadas, como se puede observar en la tabla 12, dentro del contexto eléctrico, lo que permite tener un filtro de selección de características que, dependiendo de lo exhaustivo de la evaluación, determinara las características mínimas de revisión para considerar datos de calidad dentro del contexto eléctrico.

	Empresa	Usuario	Ente de control
Exactitud	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Completitud	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Consistencia	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Credibilidad	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Actualidad	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Accesibilidad	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Conformidad	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Confidencialidad	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eficiencia	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Precisión	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Trazabilidad	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Comprensibilidad	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Tabla 12 Características VS Actores. Fuente elaboración propia

Una vez determinadas las características de calidad de interés para cada uno de los actores, se puede generar un sistema de adición que permita definir cuales características son de mayor interés para el contexto eléctrico en general, como se puede observar en la tabla 13.

	Nivel de importancia
Exactitud	1
Compleitud	2
Consistencia	2
Credibilidad	2
Actualidad	1
Accesibilidad	4
Conformidad	2
Confidencialidad	3
Eficiencia	4
Precisión	4
Trazabilidad	3
Comprensibilidad	1

	Nivel de importancia
Exactitud	1
Actualidad	1
Comprensibilidad	1
Compleitud	2
Consistencia	2
Credibilidad	2
Conformidad	2
Confidencialidad	3
Trazabilidad	3

Tabla 13 Características por nivel de relevancia

Esta agrupación permite definir las características de mayor importancia para los distintos actores, por lo que es recomendable que las características de nivel 1 sean evaluadas prioritariamente, o como mínimo en una evaluación general.

4.5. Estrategia de evaluación.

Siguiendo las actividades de la norma ISO/IEC 25040 definidas anteriormente se plantea la estrategia de evaluación.

4.5.1. Especificar la evaluación

Para la ejecución del sistema de evaluación de calidad de datos se debe definir los atributos a evaluar estos a su vez están relacionados con una tabla correspondiente a la base de datos, y según los propósitos definidos anteriormente definir las características de calidad que serán evaluadas en cada atributo, como se observa en la tabla 14.

Tabla	Atributo	Exactitud	Complejidad	Consistencia	Credibilidad	Actualidad	Conformidad	Comprensibilidad
Cliente	Atributo 1	x			x			
	Atributo 2	x	x		x			
	Atributo 3	x	x		x			
	Atributo 4	x	x			x		
	Atributo 5	x						
	Atributo 6	x	x	x		x		
	Atributo 7		x	x		x		
	Atributo 8							

Tabla 14 Tabla ejemplo paso 1. Fuente elaboración propia.

En este ejemplo, solo se tuvieron en cuenta las características con nivel de relevancia 1 y 2, definiendo así las características de calidad de datos que serán aplicados a cada uno de los atributos de la tabla que serán evaluados.

Una vez se seleccionan las características que se aplicaran en la evaluación se especifican los criterios de evaluación para esto se genera una tabla en la que se desglosa las variables, formulas, escala de calidad por cada atributo a evaluar con cada característica, como se observa en la tabla 15.

Característica	Tabla	Atributo	Escala		Variables	Formula
Característica1	Tabla 1	Atributo1	Baja			
			Media			
		Atributo 2	Alta			
	Tabla 2	Atributo 1	Baja			
		Atributo 2	Media			
			Alta			
	Tabla N	Atributo1	Baja			
		Atributo2	Media			
			Alta			

Tabla 15 Tabla ejemplo paso 2. Fuente elaboración propia.

Los campos escala, variables, formula deben ser especificado en función a lo siguiente.

Variables: Son las convenciones o significados de las letras a usar en la formula, se pueden obtener de la norma ISO 25024 [42].

Formula: Es la formula responsable de determinar el rango resultante de calidad.

Escala: Determina los rangos aceptables, regulares, no aceptables, de calidad, o sin calidad con el que se clasificara posteriormente el resultado, Estos rangos son especificados por la empresa.

4.5.2. Diseño de la evaluación

Posteriormente, se genera una tabla por característica, como se observa en la tabla 16, en la que serán evaluados los atributos relacionados con dicha característica, inicialmente se deben describir las reglas que se tendrán en cuenta para generar la evaluación de esta forma.

Característica	Tabla	Atributo	Reglas
Exactitud (Semántica)	Tabla 1	Atributo 1	se esperan numeros mayores a cero
		Atributo 2	se esperan (1,2,3,4,5,6)
	Tabla 2	Atributo 3	debe ser un numero entre -90 y 90
		Atributo 4	debe ser un numero entre -180 y 180
	Tabla 3	Atributo 5	se esperan numero mayor o igual a 0
		Atributo 6	se esperan fechas entre 1 de enero 2020- 28 diciembre 2020

Tabla 16 Tabla ejemplo paso 3. Fuente elaboración propia.

Entendiendo estas reglas como las características o determinantes que serán evaluados en la base de datos, estas reglas son puestas por la empresa según se considere relevante evaluar una u otra condición.

Posteriormente se genera una nueva tabla en la que se relaciona la cantidad total de datos asociados a ese atributo y la cantidad de datos positivos, para determinar el porcentaje de datos de calidad por cada característica evaluada, como se observa en la tabla 17, el rango de calidad aceptable debe ser definido por cada empresa que genere la evaluación, determinando así la calidad de los datos, tablas o base de datos en general.

Característica	Tabla	Atributo	Regla	Numero de datos	Cumplen	Porcentaje
Exactitud (Semántica)	Tabla 1	Atributo 1	se esperan numeros mayores a cero	100	100	100
		Atributo 2	se esperan (1,2,3,4,5,6)		100	100
	Tabla 2	Atributo 3	debe ser un numero entre -90 y 90	100	100	100
		Atributo 4	debe ser un numero entre -180 y 180		90	90
	Tabla 3	Atributo 5	se esperan numero mayor o igual a 0	1000	1000	100
		Atributo 6	se esperan fechas entre 1 de enero 2020- 28 diciembre 2020		586	58.6

Tabla 17 Tabla ejemplo paso 4. Fuente elaboración propia.

Capítulo 5. Aplicación de la estrategia de evaluación de calidad de datos en un contexto eléctrico en la plataforma AMI-SAA.

Para llevar a cabo la aplicación de la estrategia anteriormente descrita, se desarrolló el análisis de un estudio de caso, el cual permitió no solo comprobar el funcionamiento de lo desarrollado en el capítulo anterior, sino, aportar una calificación de calidad de datos que permita mejoras en el desarrollo del trabajo desarrollado por la universidad del Cauca.

5.1. Diseño de estudio de caso

En este capítulo se presenta la aplicación y evaluación de la guía desarrollada a través de un estudio de caso, incluyendo el contexto de la investigación, resultados y análisis. Para la evaluación de la guía fue seleccionada la metodología de estudio de caso debido a que según Yin [43] es un método que estudia un fenómeno en su contexto real, buscando mantener la integridad y las características significativas de los eventos.

El estudio de caso ha sido desarrollado siguiendo el protocolo de Runeson y Höst descrito en [44] y es de tipo simple-holístico según el enfoque presentado en [58], debido a que la estrategia fue aplicada en un contexto organizacional, correspondiente a una compañía del sector eléctrico.

Objetivo: Evaluar la calidad de los datos haciendo uso de la selección de las características nivel 1 y 2 del modelo de calidad de datos propuestos: Exactitud, Completitud, Consistencia, Credibilidad, Actualidad, Conformidad.

Objeto: El objeto del estudio de caso son los datos almacenados en las tablas de la base de datos que involucran los eventos de consumo eléctrico y eventos de amarre del prototipo software, plataforma de gestión AMI-SAA.

Aspecto evaluado: el estudio de caso pretende evaluar la calidad de los datos almacenados en la base de datos de la plataforma AMI-SAA en términos del modelo de calidad adaptado el cual incluye 7 características de calidad de datos: Exactitud, Completitud, Consistencia, Credibilidad, Actualidad, Conformidad.

Pregunta de investigación: ¿los datos almacenados en la base de datos de la plataforma AMI-SAA cumplen con las características de calidad seleccionadas de la norma ISO/IEC 25012?

Contexto y unidades de análisis: la evaluación se realizó sobre los datos almacenados en la base de datos de la plataforma AMI-SAA que fue desarrollada

durante los proyectos: ID4278 y ID4249, en el cual participaron el Grupo de Automática Industrial (GAI), Grupo de Investigación y Desarrollo en Ingeniería de Software (IDIS) y la Compañía Energética de Occidente (CEO), estos datos fueron generados en el laboratorio de Automática Industrial durante las fases de prueba de la aplicación, las cargas instaladas en los medidores eran constantes.

Los datos evaluados corresponden a las entidades de la base de datos que involucran los eventos de consumo eléctrico y eventos de amarre las cuales son: trafo, plc_mms, plc_tu, eventos_amarre, eventos_consumo, producto, cliente, medidor, ciudad.

Criterio de selección: el criterio para seleccionar la organización fue que perteneciera al sector eléctrico en el ámbito de prestación de servicio de distribución de energía eléctrica y fuera una empresa que operara en la región.

Para la ejecución de la evaluación se siguen las actividades de la norma ISO/IEC 25040 descritas en el capítulo anterior.

5.2. Arquitectura plataforma AMI-SAA

Con tal finalidad, se elaboró una plataforma que permitiera recibir y almacenar los datos generados en el sistema AMI-SAA. Para el desarrollo de esta plataforma se generó inicialmente el diseño de la arquitectura con base en al modelo C4 para el diseño de arquitectura de aplicaciones software. Estos niveles permiten entender la relación entre la plataforma, la base de datos, los contenidos de esta y los dispositivos PLC involucrados en la toma y envío de datos.

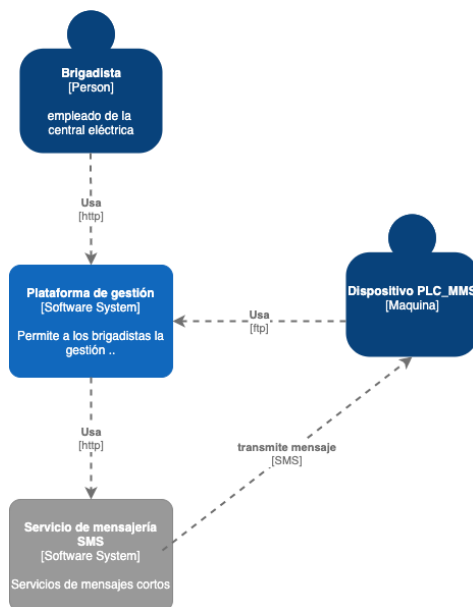


Figura 40 Nivel C1 Plataforma AMI-SAA. Fuente elaboración propia

Nivel C1. Este nivel muestra de forma macro la relación entre los dispositivos PLC, los servicios de mensajería SMS usados para el envío de señales en estos dispositivos, y la plataforma, además de la relación entre esta y los brigadistas. Únicos agentes humanos involucrados en este nivel.

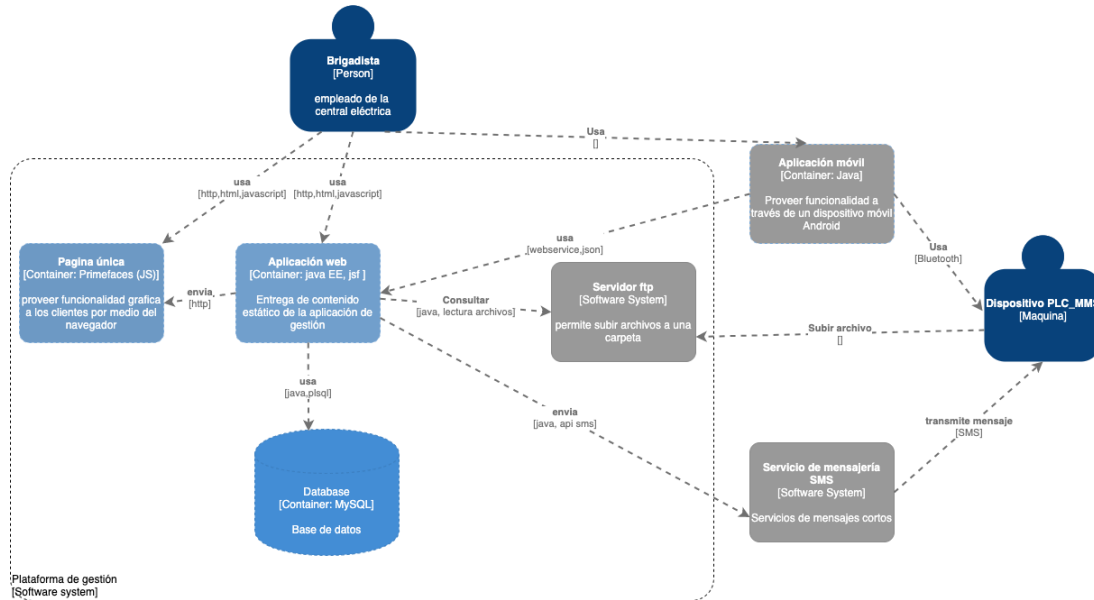


Figura 41 Nivel C2 Plataforma AMI-SAA. Fuente elaboración propia.

Nivel C2. Este nivel muestra de forma detallada la relación entre la plataforma, la base de datos (MySQL), la página única y los datos obtenidos del PLC-MMS, único dispositivo PLC con relación directa hacia la plataforma. Además de las relaciones de estos con el brigadista.

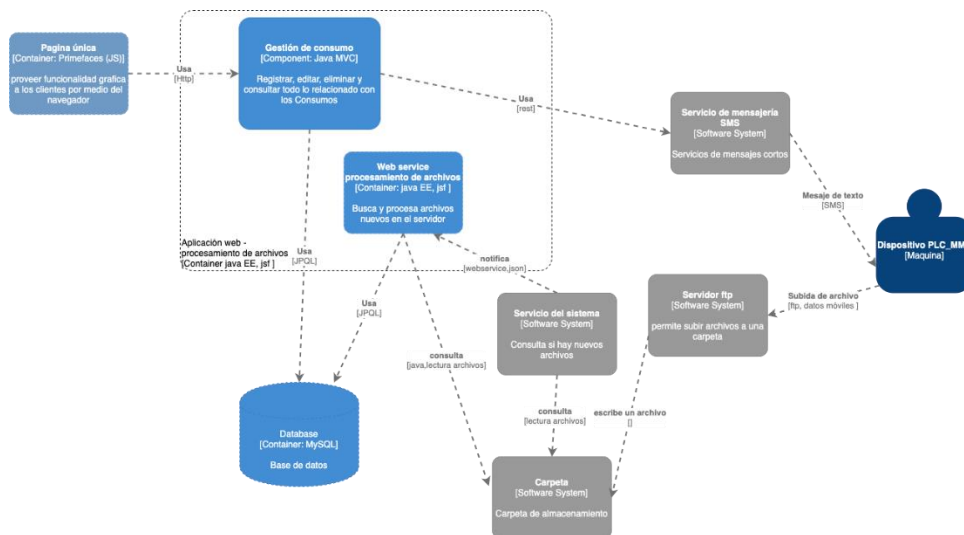


Figura 42 Nivel C3 Plataforma AMI-SAA. Fuente elaboración propia.

Nivel C3. Este nivel se va profundizar en un componente de la Plataforma de gestión el cual es Aplicación web, se exponen varios modelos dependiendo de la funcionalidad, gestión de consumos, gestión de usuarios, gestión de dispositivos (MMS, PLC-TU etc.), se expone la gestión de consumo como se puede ver en la figura 43.

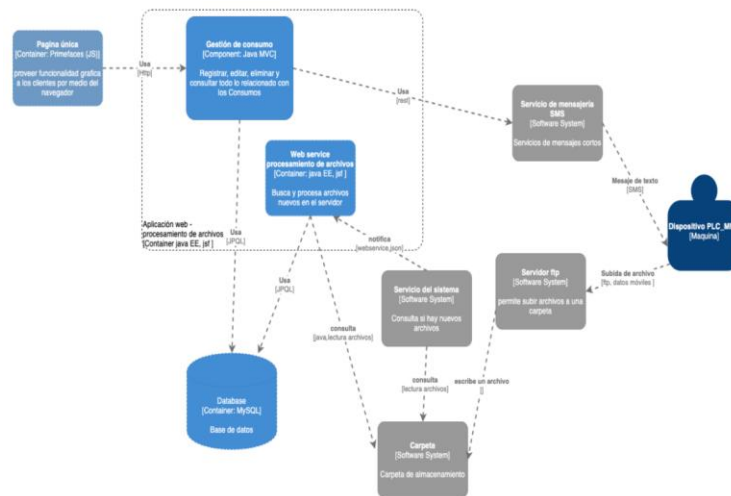


Figura 43 Nivel C3 Plataforma AMI-SAA (Gestión consumo de energía). Fuente elaboración propia.

Nivel C4. Este nivel se compone de gran diversidad esquemas que de forma muy detallada enseña el funcionamiento de la plataforma. Este nivel no es Expuesto en este documento debido a que para entenderlo se debe tener un nivel de conocimientos técnicos mínimo. Este nivel se encuentra en el anexo A Arquitectura de plataforma AMI-SAA.

Dando continuidad con el proceso, se logró hacer el llenado de datos en la plataforma AMI-SAA por medio del uso de una plataforma de generación de datos que simulan los datos encontrados en un contrato eléctrico, y las medidas asociadas durante un año. Estos grupos de información se agrupan en tablas, estas tablas no tienen una generalidad, y, por tanto, solo pueden ser definidas por la empresa a la que esta base de datos pertenezca, bajo los criterios que la misma empresa disponga.

En este caso en particular, la base de datos de AMI-SAA, dispone de 17 tablas en la que se encuentra gran variedad de atributos como se aprecia en la figura 44. Dichos atributos están relacionados con los cobros asociados a la facturación, el funcionamiento interno de la compañía eléctrica, datos de usuario y contrato con los que se podrán relacionar consumos, dispositivos y cobros al cliente final. Esta gran cantidad de información requiere de una evaluación de calidad de datos para

umentar el nivel de competitividad de la empresa, mejorar costos, y en general constituir una red eléctrica eficiente. Para esto, los atributos contenidos en cada una de las tablas se relacionan con los 3 niveles de conexión mencionados en el capítulo anterior. Y por tanto desglosa la información en grupos de tablas más pequeños que permite generar un sistema de evaluación de calidad de datos más eficiente.

Estos niveles determinan los tipos de amarre y conexión entre dispositivos PLC. encontrando usos desde sectores industriales de gran consumo energético hasta los usos comunes domésticos. Este último se relaciona por lo general con un amarre de nivel 2 en el que los dispositivos PLC-TU y PLC-MMS, son los únicos que actúan sobre la evaluación del consumo de las viviendas. Y dado que el interés de este proyecto ha sido enfocado al uso de mayor ocupación en los municipios, esta relación de nivel 2 será el único en ser evaluado las tablas relacionadas con este nivel se pueden apreciar en la figura 45.

Para el análisis de esta información, y dado que la plataforma anteriormente expuesta solo tiene la función de recolectar y almacenar la información del sistema AMI-SAA, es necesario implementar un servicio de gestión de datos como lo es MySQL.

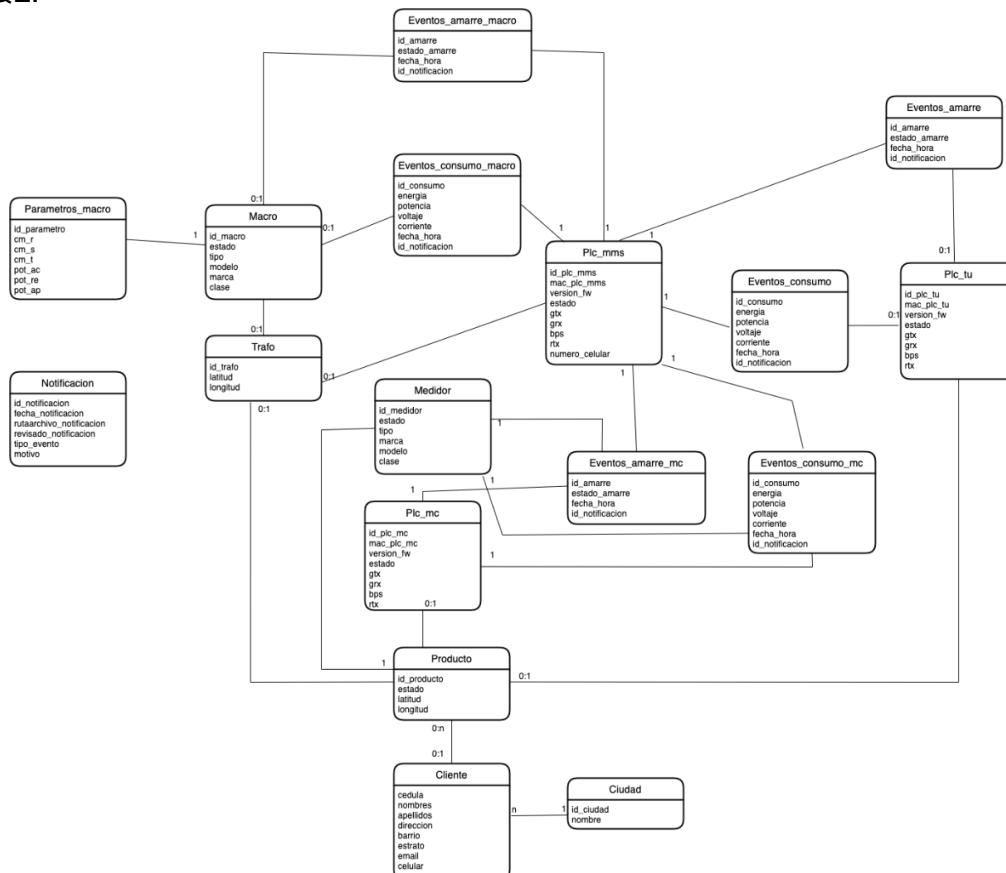


Figura 44 Base de datos general AMI-SAA. Fuente elaboración propia

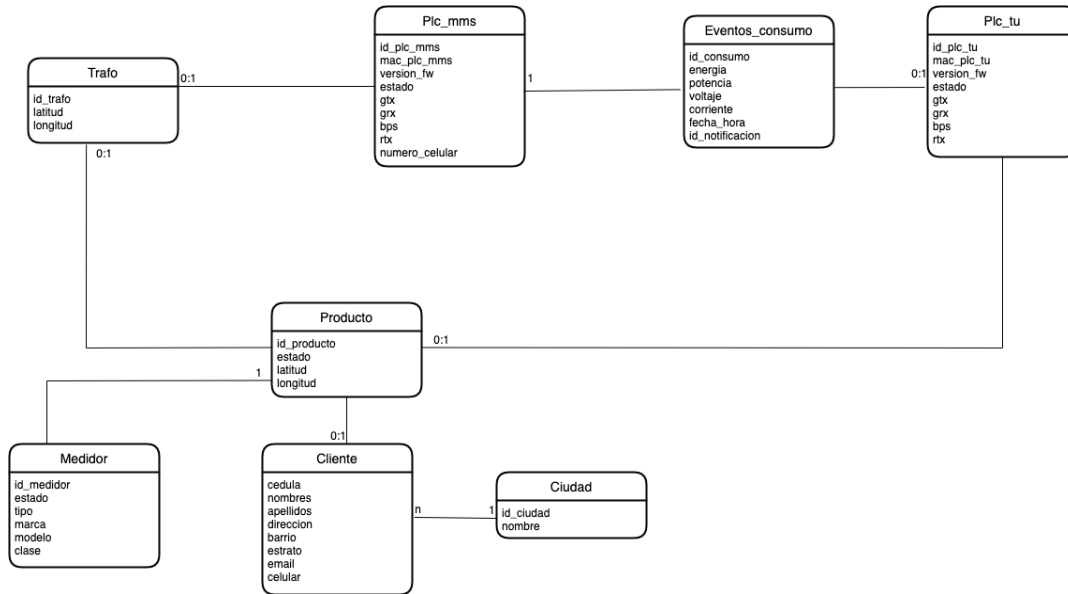


Figura 45 Datos relacionados a nivel 2 PLC-MMS a PLC-TU. Fuente elaboración propia

5.3. Implementación de la estrategia de evaluación en la base de datos de la plataforma AMI-SAA.

5.3.1. Especificación de la evaluación

Una vez enseñadas las tablas de la base de datos del proyecto AMI-SAA con sus respectivos atributos, se seleccionaron 3 tablas con la mayor variedad de datos posible, estas tablas corresponden a Cliente, Producto, Eventos de consumos, como se puede observar en las tablas 18, 19 y 20. Posteriormente se inició determinando las características que serían evaluadas en cada uno de los atributos presentes en las tablas.

Tabla	Atributo	Exactitud	Complejidad	Consistencia	Credibilidad	Actualidad	Conformidad	Comprensibilidad
Cliente	cedula	x			x			
	nombres	x	x		x			
	apellidos	x	x		x			
	direccion	x	x			x		
	estrato	x						
	email	x	x	x		x		
	celular		x	x		x		
	id_ciudad							

Tabla 18 selección de características por atributo tabla cliente. Elaboración propia

Tabla	Atributo	Exactitud	Complejidad	Consistencia	Credibilidad	Actualidad	Conformidad	Comprensibilidad
Producto	id_producto							
	cedula		x					
	id_trafo		x					
	latitud	x		x				
	longitud	x		x				

Tabla 19 selección de características por atributo tabla producto. Elaboración propia

Tabla	Atributo	Exactitud	Complejidad	Consistencia	Credibilidad	Actualidad	Conformidad	Comprensibilidad
Eventos Consumos	energia	x	x	x				
	fecha_hora	x		x				
	id_plc_mms							
	id_plc_tu		x	x				

Tabla 20 selección de características por atributo tabla evento consumo. Elaboración propia

Una vez definidas las características a evaluar en cada uno de los atributos, se especifican las fórmulas variables y escalas para cada característica de calidad, como se observa en las tablas 21 a 26.

Característica	Tabla	Atributo	Escala		Variables	Formula
			Baja	R <= 0.7		
Exactitud (Semántica)	Cliente	cedula estrato	Media	R > 0.7 y R <= 0.95	A = Número de datos que cumple con la regla B = Número de datos totales	R = A/B
			Alta	R > 0.95		
			Producto	latitud longitud		
	Media	R > 0.8 y R <= 0.95				
	Alta	R > 0.95				
	Evento consumos	energia fecha_hora	Baja	R <= 0.95		
			Media	R > 0.95 y R <= 0.98		
			Alta	R > 0.98		

Tabla 21 Criterios de evaluación Exactitud semántica. Elaboración propia

Característica	Tabla	Atributo	Escala		Variables	Formula
			Baja	R <= 0.70		
Exactitud (Sintáctica)	Cliente	cedula nombres direccion estrato email	Media	R > 0.70 y R <= 0.95	A = Número de datos que cumple con la regla B = Número de datos totales	R = A/B
			Alta	R > 0.95		
			Producto	latitud longitud		
	Media	R > 0.8 y R <= 0.95				
	Alta	R > 0.95				
	Evento consumos	energia fecha_hora	Baja	R <= 0.95		
			Media	R > 0.95 y R <= 0.98		
			Alta	R > 0.98		

Tabla 22 Criterios de evaluación Exactitud sintáctica. Elaboración propia

Característica	Tabla	Atributo	Escala		Variables	Formula
Compleitud	Cliente	nombres apellidos direccion email celular	Baja	R <= 0.90	A=Número de datos que no están vacíos o no son datos nulos B = Número de datos totales	R = A/B
			Media	R>0.90 y R<=0.98		
			Alta	R>0.98		
			Baja	R <= 0.60		
			Media	R>0.60 y R<=0.80		
			Alta	R>0.80		
	Producto	cedula id_trafo	Baja	R <= 0.70		
			Media	R>0.70 y R<=0.80		
			Alta	R>0.80		
Evento consumos	energia id_plc_tu	Baja	R <= 0.95			
		Media	R>0.95 y R<=0.99			
		Alta	R>0.99			

Tabla 23 Criterios de evaluación Compleitud. Elaboración propia

Característica	Tabla	Atributo	Escala		Variables	Formula
Consistencia	Cliente	Email celular	Baja	R <= 0.95	A =Número de datos que no presentan problemas de inconsistencia B = Número de datos totales	R = A/B
			Media	R>0.95 y R<=0.98		
			Alta	R>0.98		
	Producto	latitud longitud	Baja	R <= 0.90		
			Media	R>0.90 y R<=0.95		
			Alta	R>0.95		
	Evento consumos	energia fecha_hora id_plc_tu	Baja	R <= 0.95		
			Media	R>0.95 y R<=0.98		
			Alta	R>0.98		

Tabla 24 Criterios de evaluación consistencia. Elaboración propia

Característica	Tabla	Atributo	Escala		Variables	Formula
Credibilidad	Cliente	cedula nombres apellidos	Baja	R <= 0.95	A =Número de datos verificados como ciertos y creibles B = Número de datos totales	R = A/B
			Media	R>0.95 y R<=0.98		
			Alta	R>0.98		

Tabla 25 Criterios de evaluación credibilidad. Elaboración propia

Característica	Tabla	Atributo	Escala		Variables	Formula
Actualidad	Cliente	direccion email celular	Baja	se desconoce la fecha de actualizacion de los datos	-	-
			Media	fueron actualizados en el transcurso del año actual		
			Alta	fueron actualizados en el transcurso del mes actual		

Tabla 26 Criterios de evaluación credibilidad. Elaboración propia

5.3.2. Diseño de la evaluación

Siguiendo con el proceso una vez definidos los criterios de evaluación se determinan las reglas de evaluación que estas deben cumplir para determinar su calidad, como se observa en las tablas 26 a 32.

Característica	Tabla	Atributo	Reglas
Exactitud (Semántica)	Cliente	cedula	se esperan numeros mayores a cero
		estrato	se esperan (1,2,3,4,5,6)
	Producto	latitud	debe ser un numero entre -90 y 90
		longitud	debe ser un numero entre -180 y 180
	Evento consumos	energia	se esperan numero mayor o igual a 0
		fecha_hora	se esperan fechas entre 1 de enero 2020- 28 diciembre 2020

Tabla 27 Reglas Exactitud semántica. Elaboración propia

Característica	Tabla	Atributo	Reglas
Exactitud (Sintáctica)	Cliente	cedula	debe ser numerico maximo 10 digitos
		nombres	deben estar escritos en mayuscula
		apellidos	deben estar escritos en mayuscula
		direccion	formato de dirección de colombia
		estrato	debe ser un numero de un digito
		email	debe tener formato de correo electronico
	Producto	latitud	debe ser numerico
		longitud	debe ser numerico
	Evento consumos	energia	debe ser numerico
		fecha_hora	se espera un formato(YYYY-MM-DD HH:MM:SS)

Tabla 28 Reglas exactitud sintáctica. Elaboración Propia

Característica	Tabla	Atributo	Reglas
Complejidad	Cliente	nombres	no puede ser una cadena vacia
		apellidos	no puede ser una cadena vacia
		direccion	no puede ser una cadena vacia
		email	no debe ser un campo nulo ni ser una cadena vacia
		celular	no debe ser un campo nulo ni ser una cadena vacia
	Producto	cedula	debe haber un cliente asociado
		id_trafo	producto debe estar asociado a un dispositivo trafo
	Evento consumos	energia	no debe ser un campo nula
		id_plc_tu	no debe ser un campo nulo

Tabla 29 Reglas Complejidad. Elaboración propia

Característica	Tabla	Atributo	Reglas
Consistencia	Cliente	Email	el correo electronico debe ser unico en la tabla cliente
		Celular	el numero de celular debe ser unico en la tabla cliente
	Producto	latitud	el punto de ubicación debe ser unico en la tabla producto
		longitud	
	Evento consumos	energia	los consumos son incrementales por cada producto en el tiempo
		fecha_hora	las fechas son incrementales por cada dispositivo plc_tu
		id_plc_tu	debe existir en cada periodo de facturacion almenos un registro de consumo del dispositivo plc_tu

Tabla 30 Reglas Consistencia. Elaboración Propia

Característica	Tabla	Atributo	Reglas
Credibilidad	Cliente	cedula	La numero de cedula exista en colombia
		nombres	los nombres y apellidos correspondan con el
		apellidos	numero de cedula

Tabla 31 Reglas Credibilidad. Elaboración propia

Característica	Tabla	Atributo	Reglas
Actualidad	Cliente	direccion	fecha de la ultima actualizacion
		email	fecha de la ultima actualizacion
		celular	fecha de la ultima actualizacion

Tabla 32 Reglas Actualidad. Elaboración propia

5.3.3. Ejecución de la evaluación

Una vez definidas todas las reglas para el proceso de evaluación, se generan todas las consultas en la base de datos para determinar las condiciones de cada atributo, como se observa en las tablas 33 a 36, este proceso de consultas puede ser revisado en el anexo D reservado para las consultas y proceso de evaluación.

Característica	Tabla	Atributo	Reglas	B	A	R
Exactitud (Semántica)	Cliente	cedula	se esperan numeros mayores a cero	331	331	1
		estrato	se esperan (1,2,3,4,5,6)		331	1
	Producto	latitud	debe ser un numero entre -90 y 90	331	331	1
		longitud	debe ser un numero entre -180 y 180		327	0,99
	Evento consumos	energia	se esperan numero mayor o igual a 0	3972	3972	1
		fecha_hora	se esperan fechas entre 1 de enero 2020- 28 diciembre 2020		3972	1

Tabla 33 Evaluación exactitud semántica. elaboración propia

Característica	Tabla	Atributo	Reglas	B	A	R
Exactitud (Sintáctica)	Cliente	cedula	debe ser numerico maximo 10 digitos	331	331	1
		nombres	deben estar escritos en mayuscula		310	0,94
		apellidos	deben estar escritos en mayuscula		310	0,94
		direccion	formato de direccion de colombia		10	0,03
		estrato	debe ser un numero de un digito		331	1
		email	debe tener formato de correo electronico		318	318
	Producto	latitud	debe ser numerico	331	331	1
		longitud	debe ser numerico		331	1
	Evento consumos	energia	debe ser numerico	3972	3972	1
		fecha_hora	se espera un formato fecha y hora (YYYY-MM-DD HH:MM:SS)		3972	0

Tabla 34 Evaluación exactitud sintáctica. elaboración propia

Característica	Tabla	Atributo	Reglas	B	A	R
Compleitud	Cliente	nombres	no puede ser una cadena vacia	331	331	1
		apellidos	no puede ser una cadena vacia		331	1
		direccion	no puede ser una cadena vacia		331	1
		email	no debe ser un campo nulo ni ser una cadena vacia		318	0,96
		celular	no debe ser un campo nulo ni ser una cadena vacia		318	0,96
	Producto	cedula	debe haber un cliente asociado	331	331	1
		id_trafo	producto debe estar asociado a un dispositivo		331	1
	Evento consumos	energia	no debe ser un campo nulo	3972	3972	1
id_plc_tu		no debe ser un campo nulo	3972	3972	1	

Tabla 35 Evaluación completitud. elaboración propia

Característica	Tabla	Atributo	Reglas	B	A	R
Consistencia	Cliente	Email	el correo electronico debe ser unico en la tabla	318	309	0,97
		Celular	el numero de celular debe ser unico en la tabla clier	318	318	1
	Producto	latitud	el punto de ubicación debe ser unico en la tabla producto	331	149	0,45
		Longitud				
	Evento consumos	energia	los consumos son incrementales por cada producto en el tiempo	3972	3972	1
		feha_hora	las fechas son incrementales por cada dispositivo plc_tu	3972	3972	1
		id_plc_tu	debe existir en cada periodo de facturacion almenos un registro de consumo del dispositivo plc_tu	3972	3972	1

Tabla 36 Evaluación Consistencia. elaboración propia

5.3.4. Resultados de la evaluación

Una vez realizado el proceso de evaluación siguiendo las reglas establecidas anteriormente y como lo describe el proceso tomado de la ISO/IEC 25040, se presentan las tablas calificadas de acuerdo a la escala establecida, como se observa en las tablas 37 a 40.

Característica	Tabla	Atributo	Calidad
Exactitud (Semántica)	Cliente	cedula	Alta
		estrato	Alta
	Producto	latitud	Alta
		longitud	Alta
	Evento consumos	energia	Alta
		fecha_hora	Alta

Tabla 37 Resultado evaluación exactitud semántica. elaboración propia

Característica	Tabla	Atributo	Calidad
Exactitud (Sintáctica)	Cliente	cedula	Alta
		nombres	Media
		apellidos	Media
		direccion	Baja
		estrato	Alta
		email	Alta
	Producto	latitud	Alta
		longitud	Alta
	Evento consumos	energia	Alta
fecha_hora		Baja	

Tabla 38 Resultado evaluación exactitud sintáctica. elaboración propia

Característica	Tabla	Atributo	Calidad
Compleitud	Cliente	nombres	Alta
		apellidos	Alta
		direccion	Alta
		email	Alta
		celular	Alta
	Producto	cedula	Alta
		id_trafo	Alta
	Evento consumos	energia	Alta
		id_plc_tu	Alta

Tabla 39 Resultado evaluación completitud. elaboración propia

Característica	Tabla	Atributo	Calidad
Consistencia	Cliente	Email	Alta
		Celular	Alta
	Producto	latitud	Baja
		Longitud	
	Evento consumos	energia	Alta
		feha_hora	Alta
		id_plc_tu	Alta

Tabla 40 Resultado evaluación consistencia. elaboración propia

5.3.5. Conclusión de la evaluación

Se puede concluir de manera general que la base de datos diseñada para la plataforma de gestión AMI-SAA, según las reglas de evaluación que se usaron, tiene un grado de calidad alto, sin embargo, existen algunas falencias en algunos atributos en los que se encontraron calificación de media y baja incluso algunos con R=0 por lo que es importante realizar correcciones particulares en estos atributos con bajos niveles de calidad.

Adicionalmente se aclara que las características de Credibilidad y Actualidad no fueron evaluadas al encontrarse con obstáculos en la plataforma para su evaluación. Por el lado de la característica de Credibilidad, es necesario validar la información con una base de datos oficial como lo puede ser la registraduría o cualquier otra entidad en la que se pueda hacer cruce de información y validación de la misma, en el caso de esta evaluación no se logró hacer cruce con ninguna entidad debido a que los datos aquí expuestos fueron generados únicamente con el fin de probar la funcionalidad de la plataforma AMI-SAA y no representan datos reales de ninguna compañía eléctrica ni de sus usuarios.

Por otro lado, la plataforma AMI-SAA no cuenta con ningún sistema que permita visualizar el registro de actividad o actualización de datos, por lo que no es posible realizar una evaluación de Actualidad. Se recomienda generar una actualización al sistema que permita recaudar esta información, y posteriormente hacer evaluaciones periódicas de estas actualizaciones con el fin de determinar qué tan actualizada se encuentra la base de datos.

Capítulo 6. Conclusiones y recomendaciones.

6.1. Conclusiones.

- Los tipos de medida están directamente relacionados con los instrumentos para su medición. De estos dependerá igualmente, las características de tensión y corriente que se asocian a un producto o contrato.
- Este trabajo se fundamentó en la existencia de 3 actores principales convergentes, sin embargo, debe considerarse la posibilidad de otros actores que involucren nuevos propósitos de evaluación.
- Los datos contenidos, almacenados y de carácter públicos son equivalentes en la mayoría de empresas prestadoras del servicio eléctrico. Por lo que es coherente pensar que las bases de datos de estas, guardan una relación entorno a su funcionamiento. Por tanto, se puede generar una alternativa frecuente que permita determinar la calidad de los datos alojados en cada base de datos relacionado a este tipo de entidad.
- Se encontró que las características como: recuperabilidad, disponibilidad y portabilidad, dependen de manera directa de los procesos de sistemas de gestión internos. De esta manera, la evaluación del proceso de cada una de estas características, dependerá de la importancia acorde y requerida a la necesidad específica de cada compañía del sector eléctrico.
- Teniendo conocimiento del modelo de negocio de un contexto eléctrico, se encontraron problemas de diseño que son justificables tratándose de un prototipo.
- Para que los datos asociados al contexto eléctrico tengan un alto nivel de calidad es necesario que durante el proceso de desarrollo del software se integre un proceso de calidad de datos.
- La mejora en la calidad de datos no está lineada exclusivamente solo a través de procesos sistemáticos por tanto se debe prestar apoyo de mejoras al componente recurso humano.

6.2. Lecciones aprendidas

- Debe implementarse un formato, formulario o método por el cual el estudiante pueda solicitar información a otras instituciones, avalado por parte de profesor o dependencia que adelante la investigación. Esto optimizaría los tiempos en la diligencia, entrega de información y su confiabilidad.
- Es necesarios conocer los procesos de gestión de una institución, pues de ellos derivan la información.
- Integrar al proceso de investigación los cargos técnicos, pues estos son un punto clave dentro de la toma y la valoración de la calidad de un dato.
- Es bueno apoyarse en profesionales multidisciplinarios que retroalimente los conceptos y utilidades del proyecto de manera que pueda tenerse presente durante la ejecución del proyecto. y el servicio que así mismo pueda este

prestar. Esto permite que se amplie la visión de la necesidad del perfil del ingeniero de sistemas en diferentes sectores empresariales. Dando oportunidad así a la ampliación de nuevas plazas laborales.

- Debe saberse solicitarse y sobresaltar la utilidad de la información y dejar en claro el uso exclusivo de investigación, debido que puede ser negada por políticas internas (viéndose camuflada ante acciones como: cambios en roles y/o de manejo del personal, negligencia por parte del funcionario o, entre otras)
- Cuando se hace visita de campo el éxito de la actividad está dada por la disponibilidad y disposición de los actores involucrados
- Los recursos de tiempo, recurso humano y económico destinado para la ejecución de un proyecto investigativo deben ser fácilmente ajustables acorde a Extra planificaciones, en caso particular debió ser modificadas algunas tomas de información por motivos de emergencia sanitaria – Pandemia Sarc-CoviD 19.

6.3. Trabajos futuros

- Realizar un proceso de evaluación donde se involucren todos los niveles importancia obtenidos como resultados de la selección de características.
- Realizar nuevos procesos de evaluación de otras bases de datos que se encuentren en producción de compañías eléctricas que permita comprobar la veracidad de este trabajo de investigación.
- Se presenta la posibilidad de generar una guía que permite simplificar el proceso de evaluación para la calidad de datos en un contexto eléctrico, basada en la normativa estándar ISO 25012.
- De igual manera, dicha guía pretende iniciar procesos para la unificación de manera “estándar” en la evaluación de calidad de datos para un contexto eléctrico, pensando en brindar mejoras continuas a futuros lineamientos y nuevos proyectos del sector eléctrico; sea contribuyendo a las ramas administrativas como entes de regulación o como apoyo profesional y técnico a actividades procedimentales y/o de gestión institucionales.

Capítulo 7: Referencias.

- [1] J. Higuera. (2013, 01/15). *Hurto de energía se baja con más medidas sociales*. Available: <https://acortar.link/hCWAZo>
- [2] S. McLaughlin, D. Podkuiko, and P. McDaniel, "Energy theft in the advanced metering infrastructure," in *International Workshop on Critical Information Infrastructures Security*, 2009, pp. 176-187: Springer.
- [3] E. M. Igna. Ortega., "Redes de comunicación en smart grid," no. 7, pp. 36-55, 2012.
- [4] International. Organization. of. Standardization. (2008). *ISO/IEC 25012*. Available: <https://iso25000.com/index.php/normas-iso-25000/iso-25012>
- [5] O. E. Morán Mora and L. F. Ortiz Fernández, "Estudio de la infraestructura de medición avanzada (AMI), principales requerimientos y beneficios," 2012.
- [6] M. T. de Torres, A. G. Roza, A. Torres, M. Duque, and C. J. R. d. I. García, "La medición remota de energía eléctrica: un campo de oportunidades para la industria Colombia," no. 7, pp. 37-42, 1996.
- [7] M. Eppler and M. Helfert, "A classification and analysis of data quality costs," in *International Conference on Information Quality*, 2004, pp. 311-325.
- [8] IEA. (2011). *Technology Roadmap - Smart Grids*. Available: <https://acortar.link/sAyulm>
- [9] J. C. Reyes. (2014). *Amarres eléctricos*. Available: <https://acortar.link/xWBI2Y>
- [10] L. Pineda, "Diseño de protocolos de comunicación para el sistema SAA," Automática Industrial, Universidad del Cauca - Facultad Ingeniería Electrónica y Telecomunicaciones, 2013.
- [11] como-funciona.com. (2020). *¿Cómo funcionan las señales eléctricas?* Available: <https://como-funciona.com/senales-electricas/>
- [12] T. V. Babu, T. S. Murthy, and B. Sivaiah, "Detecting unusual customer consumption profiles in power distribution systems—APSPDCL," in *2013 IEEE International Conference on Computational Intelligence and Computing Research*, 2013, pp. 1-5: IEEE.
- [13] M. Madrigal, J. J. Rico, and L. J. I. L. A. T. Uzcategui, "Estimation of technical energy losses in electrical distribution systems," vol. 13, no. 10, pp. 3310-3316, 2015.
- [14] A. Meffe and C. C. B. de OLIVEIRA, "Technical loss calculation by distribution system segment with corrections from measurements," in *CIREN 2009-20th International Conference and Exhibition on Electricity Distribution-Part 1*, 2009, pp. 1-4: IET.
- [15] S. Kalambe, G. J. r. Agnihotri, and s. e. reviews, "Loss minimization techniques used in distribution network: bibliographical survey," vol. 29, pp. 184-200, 2014.

- [16] R. D. González Rodríguez, "Intelligent system for non-technical losses management in electricity users," Universidad del Norte, 2017.
- [17] F. Endesa. (2019). *El transformador eléctrico*. Available: <https://acortar.link/B4s3q>
- [18] Y. P. Villazón. (2019, 15 jun). *Diagramando la arquitectura de software en microservicios. Modelo C4*. Available: <https://acortar.link/ZBwHWd>
- [19] A. Ortiz, "Qué es un framework en informática o programación," vol. 6, 2018.
- [20] I. Caballero Muñoz Reja, A. I. Gómez Carretero, F. Gualo Cejudo, J. Merino García, B. Rivas García, and M. G. J. E. d. I. U. B. R.-M. E. Piattini Velthuis, "Calidad de datos," 2019.
- [21] O. A. Beltrán, "Revisiones sistemáticas de la literatura," vol. 20, no. 1, pp. 60-69, 2005.
- [22] D. Torres-Salinas and Á. C. J. A. T. Clavijo, "Herramientas para la evaluación de la ciencia en universidades y centros I+ D: descripción y usos," vol. 6, no. 1, pp. 142-146, 2012.
- [23] IEEE. (2017). *About IEEE Xplore® Digital Library*. Available: <https://acortar.link/oFd5VQ>
- [24] E. Verbo, I. Caballero, R. Pérez-Castillo, C. Calero, and M. Piattini, "Una Metodología Basada en ISO/IEC 15939 para la Elaboración de Planes de Medición de Calidad de Datos," in *JISBD*, 2008, pp. 253-264.
- [25] I. Rafique, P. Lew, M. Q. Abbasi, Z. J. I. J. o. C. Li, and I. Engineering, "Information quality evaluation framework: Extending ISO 25012 data quality model," vol. 6, no. 5, pp. 568-573, 2012.
- [26] A. Karkouch, H. Mousannif, H. Al Moatassime, T. J. J. o. A. I. Noel, and H. Computing, "A model-driven framework for data quality management in the Internet of Things," vol. 9, no. 4, pp. 977-998, 2018.
- [27] A. Koren, M. Jurčević, and R. J. W. P. C. Prasad, "Comparison of Data-Driven Models for Cleaning eHealth Sensor Data: Use Case on ECG Signal," vol. 114, no. 2, 2020.
- [28] G. N. Aranda, N. Martínez Carod, P. Faraci, and A. Cechich, "Hacia un framework de evaluación de calidad de información en foros de discusión técnicos," in *XIV Simposio Argentino de Ingeniería de Software (ASSE)-JAIIO 42 (2013)*, 2013.
- [29] Compañía. Energetica. de. Occidente. CEO, "Norma de Acometidas y Medicion Version 03," pp. 1-119, 2014.
- [30] ICONTEC, "NTC 2147 - Equipos de medición de energía eléctrica (C.A.). Requisitos particulares. Medidores estáticos de energía activa (clases 0,2 s y 0,5 s)," ed, 2019.
- [31] A. G. Acosta. (2019, 19 nov.). *8.7.4 Teoría y técnica*. Available: <https://acortar.link/ZZWGZz>
- [32] A. Blanco. (2011, Ene. 25). *Medidores de energía eléctrica*. Available: <https://acortar.link/B4G9wX>

- [33] E. P. d. M. EPM. (2013, 25 Enero). *¿Qué es la energía reactiva?* Available: <https://acortar.link/TWVVNs>
- [34] D. M. H. Londoño, "Desarrollo de una guía enfocada a medidores de energía y conexiones de medidores," Universidad Tecnológica de Pereira. Facultad de Tecnologías. Tecnología ..., 2013.
- [35] Z. Aguirre and J. Wilkar, "Modulo portátil de medida semidirecta para análisis de cargas de una instalación eléctrica," 2021.
- [36] S. F. chacón cajo, "Implementación de un equipo de contrastación para contadores de energía eléctrica monofásico, para el laboratorio de medidores de la empresa eléctrica quito," Quito, 2018.
- [37] EPM. (2021, 25 Enero). *Selección y conexión de equipos del sistema de medida de energía eléctrica.* Available: <https://acortar.link/7mvAyi>
- [38] T. M. Negocio. *Establecimientos de bajo impacto.* Available: <https://acortar.link/yupobC>
- [39] PACO RODRIGUEZ. (2013). *Redes PLC (I): Qué son y para qué sirven.* Available: <https://acortar.link/iZFiWD>
- [40] P. stone. (2020). *Smart grid colombia visión 2030: La ruta de transición hacia la eficiencia energética.* Available: <https://acortar.link/46ew3E>
- [41] A. Caro, A. Fuentes, and M. A. J. I. R. c. d. i. Soto, "Desarrollando sistemas de información centrados en la calidad de datos," vol. 21, no. 1, pp. 54-69, 2013.
- [42] SUPERSERVICIOS. (2014, Mar. 09). *Formato ZNI C6.* Available: <https://acortar.link/WqZli2>
- [43] K. Yin Robert, "Case study research and applications: design and methods," ed: Los Angeles, CA: Sage Publications, 2017.
- [44] P. Runeson and M. J. E. s. e. Höst, "Guidelines for conducting and reporting case study research in software engineering," vol. 14, no. 2, pp. 131-164, 2009.