

MODELO ONTOLÓGICO PARA LA INTEGRACIÓN DE INFORMACIÓN ENTRE DOS APLICACIONES INDUSTRIALES



**Edgar Alexander Aguilar Bolaños
Juan Pablo Diago Rodríguez**

**Monografía para optar al título de
Ingeniero en Automática Industrial**

**Director
Msc. Miguel Ángel Niño Zambrano**

**Asesora
Ing. Elizabeth Granados Pemberty**

Universidad del Cauca
Facultad de Ingeniería Electrónica y Telecomunicaciones
Departamento de Electrónica, Instrumentación y Control
Ingeniería en Automática Industrial
Popayán, Mayo de 2009

AGRADECIMIENTOS

A la Universidad del Cauca y a su Programa de Ingeniería en Automática Industrial por forjarnos como personas y profesionales que buscan la excelencia con el propósito de mejorar la calidad de vida de la sociedad.

Al MSc. Miguel Ángel Niño Zambrano, director del trabajo de investigación, por su apoyo, instrucción, consejo, y paciencia sin lo cual no hubiese sido posible la realización de éste proyecto.

A la Ing. Elizabeth Granados Pemberty, co-directora del trabajo de investigación, por sus aportes, motivación y apoyo constante e incondicional durante el desarrollo del proyecto.

A nuestras familias por su infinito amor y respaldo. Por hacer de nosotros personas de bien, con profundos valores.

Y a todas aquellas personas que de una u otra forma, colaboraron en la realización de esta investigación, hacemos extensivos nuestros más sinceros agradecimientos.

TABLA DE CONTENIDO

TABLA DE CONTENIDO	3
LISTA DE FIGURAS	5
LISTA DE TABLAS	7
INTRODUCCIÓN	9
1 MARCO TEÓRICO	12
1.1 ONTOLOGÍAS	12
1.2 ELEMENTOS QUE CONFORMAN UNA ONTOLOGÍA	14
1.3 APLICACIONES Y VENTAJAS DEL USO DE LAS ONTOLOGÍAS.....	15
1.4 CLASIFICACIÓN DE LAS ONTOLOGÍAS	16
1.4.1 CLASIFICACIÓN CON BASE EN EL CAMPO Y EL USO DE LA CONCEPTUALIZACIÓN.....	17
1.4.2 CLASIFICACIÓN POR MOTIVACIÓN	18
1.4.3 CLASIFICACIÓN SEGÚN EL ÁMBITO DEL CONOCIMIENTO	19
1.4.4 CLASIFICACIÓN SEGÚN EL TIPO DE AGENTE.....	19
1.4.5 CLASIFICACIÓN SEGÚN EL NIVEL DE ABSTRACCIÓN Y RAZONAMIENTO LÓGICO	20
1.5 METODOLOGÍAS PARA LA CONSTRUCCIÓN DE ONTOLOGÍAS.....	20
1.5.1 METODOLOGÍA CYC.....	21
1.5.2 METODOLOGÍA DE USCHOLD Y KING.....	22
1.5.3 METODOLOGÍA DE GRÜNINGER Y FOX	23
1.5.4 METODOLOGÍA KACTUS	24
1.5.5 METHONTOLOGY.....	25
1.5.6 METODOLOGÍA SENSUS.....	26
1.5.7 METODOLOGÍA ON-TO-KNOWLEDGE	27
1.5.8 TERMINAE	28
2 METODOLOGÍA ONTOPRIN	30
2.1 DESCRIPCIÓN DE LA METODOLOGÍA ONTOPRIN.....	33
2.1.1 PASO 1: IDENTIFICACIÓN DEL PROPÓSITO.....	33
2.1.2 PASO 2: EXTRACCIÓN MANUAL DE TERMINOLOGÍA DE DIFERENTES FUENTES DE INFORMACIÓN.....	34
2.1.3 PASO 3: CONFRONTACIÓN DE LA TERMINOLOGÍA EXTRAÍDA CON UNA TERMINOLOGÍA TÉCNICA.....	37
2.1.4 PASO 4: PROCESO DE CONCEPTUALIZACIÓN	38
2.1.5 PASO 5: IMPLEMENTACIÓN.....	40
2.1.6 PASO 6: REFINAMIENTO Y ESTRUCTURACIÓN DE LA ONTOLOGÍA.....	41
3 APLICACIÓN DE LA METODOLOGÍA ONTOPRIN A UN CASO DE ESTUDIO	43
3.1 PASO 1: IDENTIFICACIÓN DEL PROPÓSITO.....	43
3.2 PASO 2: EXTRACCIÓN MANUAL DE TERMINOLOGÍA DE DIFERENTES FUENTES DE INFORMACIÓN	44

3.3 PASO 3: CONFRONTACIÓN DE LA TERMINOLOGÍA EXTRAÍDA CON UNA TERMINOLOGÍA TÉCNICA.....	48
3.4 PASO 4: PROCESO DE CONCEPTUALIZACIÓN	50
3.5 PASO 5: IMPLEMENTACIÓN	62
3.6 PASO 7: REFINAMIENTO Y ESTRUCTURACIÓN DE LA ONTOLOGÍA	66
<u>4 DESCRIPCIÓN DEL PROTOTIPO CREADO.....</u>	<u>67</u>
4.1 DESCRIPCIÓN DEL PROTOTIPO DESARROLLADO.....	67
4.1.1 ACTORES DEL SISTEMA	68
4.1.2 MARCO DE TRABAJO DE JENA.....	69
4.1.3 MODELO CONCEPTUAL DEL PROTOTIPO	69
4.1.4 MODELOS DE BASES DE DATOS	71
4.1.5 ARQUITECTURA DEL SISTEMA	79
4.2 CASOS DE USO DE ALTO NIVEL	84
4.3 CASOS DE USO REALES	87
4.3.1 CASO DE USO REAL SOLICITAR REPORTE DE DESEMPEÑO DE LA PRODUCCIÓN.....	87
4.3.2 CASO DE USO REAL SOLICITAR REPORTE INSUMOS Y/O PRODUCTO	89
<u>5 VALIDACIÓN DEL MODELO ONTOLÓGICO CONSTRUIDO.....</u>	<u>94</u>
5.1 PROCESO DE VALIDACIÓN	95
5.2 RESULTADOS.....	96
5.3 CONCLUSIONES DE LA VALIDACIÓN	98
<u>6 CUMPLIMIENTO DE OBJETIVOS</u>	<u>100</u>
6.1 CUMPLIMIENTO DE OBJETIVOS	100
6.1.1 INDICADORES USADOS.....	100
6.1.2 OBJETIVOS Y SUS INDICADORES DE CUMPLIMIENTO	101
<u>7 CONCLUSIONES Y TRABAJO FUTURO</u>	<u>108</u>
7.1 CONCLUSIONES.....	108
7.2 TRABAJO FUTURO	111
<u>REFERENCIAS BIBLIOGRÁFICAS</u>	<u>112</u>

LISTA DE FIGURAS

Figura 1. Metodología de Grüninger y Fox [21]	24
Figura 2. Diagrama del ciclo de vida en METHONTOLOGY [21]	26
Figura 3. Metodología SENSUS [21]	27
Figura 4. Encuesta para la extracción de terminología informal.....	35
Figura 5. Ejemplo de un sistema SCADA.....	35
Figura 6. Diagrama de torta para el análisis de información	37
Figura 7. Diagrama de barras para el análisis de información	37
Figura 8. Tabla de ejemplo para el análisis información.	37
Figura 9. Ejemplo de un árbol de conceptos (clases).....	39
Figura 10. Modelo de desempeño de producción (ANSI/ISA-95.00.02).....	52
Figura 11. Modelo de eventos (ANSI/ISA-88.00.04).	55
Figura 12. Modelo de conjunto de datos (ANSI/ISA-88.00.04).....	57
Figura 13. Diagrama del modelo de desempeño de producción ISA-95.00.02 formalizado en OWL.....	63
Figura 14. Diagrama del modelo del conjunto de datos ISA-88.00.04 formalizado en OWL	63
Figura 15: Diagrama del modelo eventos ISA-88.00.04 formalizado en OWL.	64
Figura 16. Ejemplo de definición de equivalencia entre clases en el lenguaje OWL	65
Figura 17. Ejemplo de definición equivalencia entre propiedades en el lenguaje OWL....	65
Figura 18. Modelo conceptual del prototipo construido	69
Figura 19: Modelo de bases de datos de históricos de Tablazo	73
Figura 20: Tablas de Eventos y Alarmas de Tablazo	74
Figura 21. Modelo de la base de datos desempeño producción Tablazo	75
Figura 22. Modelo de bases de datos de históricos de Tulcán	76
Figura 23. Tablas de Eventos y Alarmas de Tulcán	77
Figura 24. Modelo de la base de datos desempeño producción Tulcán	78
Figura 25. Arquitectura del prototipo desarrollado.....	79
Figura 26. Mapeo entre una clase y una tabla a través del lenguaje D2RQ	84
Figura 27. Mapeo entre un atributo y una columna a través del lenguaje D2RQ.....	84
Figura 28. Casos de uso comunes a todos los actores del sistema.....	85

Figura 29. Diagrama de casos de uso para el administrador del sistema85
Figura 30. Diagrama de casos de uso para el operador de planta.....86
Figura 31. Diagrama de casos de uso para el encargado de producción86

LISTA DE TABLAS

Tabla 1. Ventajas y desventajas de las metodologías existentes para crear ontologías ...	32
Tabla 2. Ejemplo de una tabla para especificar el propósito de la ontología	34
Tabla 3. Tabla de confrontación de terminología informal y terminología más general	38
Tabla 4. Descripción del propósito de la ontología.....	44
Tabla 5. Etapas dentro del proceso de potabilización de agua cruda	45
Tabla 6. Equipos dentro del proceso de potabilización de agua cruda	46
Tabla 7. Variables dentro del proceso de potabilización de agua cruda	47
Tabla 8. Insumos del proceso de potabilización de agua cruda.....	47
Tabla 9. Paralelo entre terminologías identificadas en relación al modelo conjunto de datos.....	49
Tabla 10. Paralelo entre terminologías identificadas en relación al modelo eventos	49
Tabla 11: Paralelo entre terminologías identificadas en relación al modelo desempeño producción	50
Tabla 12. Conceptos contenidos dentro del modelo de desempeño de producción	53
Tabla 13. Relaciones contenidas dentro del modelo de desempeño de producción	54
Tabla 14. Atributos del concepto desempeño de producción.....	55
Tabla 15. Conceptos contenidos dentro del modelo de eventos.....	56
Tabla 16: Relaciones contenidas dentro del modelo de eventos	56
Tabla 17. Atributos del concepto evento alarma	56
Tabla 18. Conceptos contenidos dentro del modelo de conjunto de datos	58
Tabla 19. Relaciones contenidas dentro del modelo de conjunto de datos	58
Tabla 20. Atributos de la clase conjunto de datos	59
Tabla 21. Instancias de propiedad de material producido, planta Tulcán	59
Tabla 22: Instancias de material consumido, planta Tulcán.....	60
Tabla 23: Instancias de la clase especificación de etiqueta, planta Tulcán	61
Tabla 24. Descripción de tablas del modelo de históricos de Tablazo	73
Tabla 25. Descripción de tablas del modelo de desempeño producción Tablazo.....	75
Tabla 26: Descripción de tablas del modelo de históricos de Tulcán.....	76
Tabla 27. Descripción de tablas del modelo de desempeño producción Tulcán.....	78
Tabla 28. Constructores soportados por las distintas configuraciones del razonador OWL [35].....	82

Tabla 29. Caso de uso real solicitar reporte de desempeño de la producción89
Tabla 30. Caso de uso real solicitar reporte insumos y/o producto93

INTRODUCCIÓN

Desde la aparición de los sistemas de automatización basados en computador a principios de 1960, los desarrollos en la automatización han ido de la mano de los avances de las tecnologías de la información (*Information Technology, IT*) en términos de hardware, software y aplicaciones [1]. Los sistemas de automatización hoy en día integran automatización básica, control de máquinas, control de calidad y monitoreo de condiciones, dando un gran énfasis a la administración de información para el manejo eficiente y seguro de sistemas grandes y complejos [1][2]. De esta forma, además de controlar el proceso, un sistema de automatización tiene la responsabilidad de recolectar información acerca del proceso productivo [2][3]. Los datos del proceso (señales de control, mediciones, estado del proceso) son almacenados en servidores con el propósito de brindar información actualizada acerca de las actividades del proceso que sirva de soporte a las tareas de toma de decisiones [2][3][4].

La toma de decisiones con base en la información presente en varias fuentes hace necesario su integración de tal forma que brinde a los usuarios la ilusión de interactuar con un único sistema que facilite la búsqueda y recuperación de información con la finalidad de establecer una mejor comprensión de cierta situación [5].

Históricamente, el problema de proporcionar un acceso uniforme a la información proveniente de diversas fuentes ha sido tratado ampliamente en el campo de las redes de computadores y de bases de datos distribuidas dando como resultado el desarrollo de estándares como ODBC (*Open Data Base Connectivity, Conectividad abierta de bases de datos*) y XML (*eXtensible Markup Language, Lenguaje extensible de marcas*) [1][6]. Sin embargo la integración de varias fuentes implica

un desafío adicional, ocasionado por las diferencias en la interpretación (semántica) que las personas asignan a la información [6][7].

Más específicamente, en el área de las bases de datos, la semántica representa la interpretación que las personas tienen acerca de los datos y elementos de un esquema de acuerdo con su comprensión del mundo dentro de contexto en particular [5]. Las formas más simples de discrepancia en la interpretación de la información son los homónimos (el empleo del mismo término con diferente significado) y sinónimos (el uso de diferentes términos con el mismo significado) [6].

Teniendo en cuenta lo anterior, la integración semántica consiste entonces en agrupar o combinar información proveniente de diversas fuentes empleando para ello una semántica explícita y formal con el objeto de evitar que información semánticamente incompatible sea mezclada estructuralmente y con ello proporcionar resultados más significativos [5][7].

Esta necesidad de definiciones formales y explícitas acerca de la semántica de la información ha llevado a la adopción y aplicación de ontologías formales como una solución potencial a los problemas de heterogeneidad semántica [7]. Esto se debe a que las ontologías están pensadas para la reducción de ambigüedades terminológicas y conceptuales que provee un marco de unificación (lenguaje común) con base en un conjunto de axiomas lógicos que especifican el significado preciso de los términos para una comunidad en particular [8][9].

Con base en los aspectos anteriormente presentados, este trabajo pretende utilizar las ontologías con el fin de obtener un dominio de conocimiento común dentro de un proceso productivo. Dicho dominio común se plantea como un medio para integrar los datos obtenidos de los productos empleados en el campo de la automatización de procesos, con el dominio de conocimiento que maneja el cliente

interno de dicho ámbito dominio con el fin de dar soporte al cliente, en sus tareas de toma de decisiones.

Para la realización del proyecto se utilizaron las herramientas SCADA propietarias iFIX™ de intellution y RSView®32™ de Rockwell Software®, paquetes que constituyen dos de las más importantes, potentes y comunes soluciones empleadas a nivel local e internacional en las industrias de procesos [10][11].

En resumen, las herramientas software se emplearon para simular el comportamiento de las etapas dentro de un proceso productivo manufacturero¹, a partir de las cuales se obtuvo una serie de información que fue descrita explícitamente a través del “modelo ontológico²”.

Para la realización del modelo se planteó una metodología ajustada al dominio de la automatización de procesos industriales debido al bajo nivel explicativo presentado en las encontradas en la literatura.

Finalmente, una vez construido el modelo ontológico siguiendo la metodología planteada, se procedió al desarrollo de un prototipo software que permitiera su procesamiento, con el fin de brindar una serie de funcionalidades a los usuarios del dominio seleccionado.

¹ El proceso caso de estudio seleccionado corresponde la de tratamiento de agua potable. Dentro del Anexo B se encuentra una descripción detallada de este proceso.

² El presente trabajo adopta dicha denominación, debido a que un modelo ontológico es aquel que toma como base de representación y conceptualización una ontología, para poder desarrollar un conjunto de aplicaciones [13][18][19].

1 MARCO TEÓRICO

1.1 ONTOLOGÍAS

Las ontologías fueron desarrolladas en el campo de la Inteligencia Artificial (AI) para facilitar que el conocimiento pudiera ser compartido y reutilizado. Desde comienzos de los noventa las ontologías se han convertido en un tema de investigación importante para diferentes comunidades de la Inteligencia Artificial, como lo son: la ingeniería de conocimiento, procesamiento del lenguaje natural y representación de conocimiento. Más recientemente, la noción de ontología se ha popularizado en campos como integración inteligente de información, sistemas cooperativos de información, recuperación de información, comercio electrónico y gestión de conocimiento. La razón por la cual las ontologías se han convertido en tan populares es, en gran medida, debido a lo que prometen una comprensión compartida y común de algún dominio que puede ser comunicado entre individuos y sistemas de aplicación [12][13][14].

Las ontologías se han desarrollado con el fin de proporcionar una semántica computacionalmente procesable de las fuentes de información que pretendan describir y, con esto, facilitar la comunicación entre diferentes agentes (software o humanos) [14].

Desde sus inicios (para ahondar en la historia de las ontologías, remitirse al Anexo A), muchas definiciones de ontologías han venido dándose por diferentes autores. Sin embargo, una de las definiciones más destacadas y que mejor caracteriza la esencia de una ontología es la siguiente [13][14]:

“Una ontología es una especificación formal y explícita de una conceptualización compartida”.

En este contexto, conceptualización se refiere a una visión simplificada de algún fenómeno en el mundo que se quiere representar con algún propósito particular, a través de la identificación y definición de los conceptos³, objetos y entidades relevantes a dicho fenómeno, así como de las relaciones existentes entre ellos [1][9][12].

Explícita significa que tanto los conceptos como las restricciones de su uso se definen sin ambigüedades [13].

Formal significa que una ontología debe ser expresada en un lenguaje de representación de conocimiento que proporcione una semántica formal. Este aspecto asegura que la conceptualización que contiene una ontología sea procesable a nivel computacional [15][16].

Compartida se refiere al hecho de que la ontología captura conocimiento consensual, esto es, que no se restringe a un individuo sino que es aceptado por un grupo [13].

En esencia, las ontologías son muy similares a las técnicas de modelado conceptual existentes, como el Modelo Entidad Relación o el Lenguaje Unificado de Modelado UML [16]. Sin embargo, las ontologías difieren de estos métodos y tecnologías existentes en tres aspectos: (i) el objetivo principal de las ontologías es permitir el consenso sobre el significado de un vocabulario específico de términos y, de esta forma, facilitar la integración de información entre aplicaciones; (ii) las ontologías son formalizadas en lenguajes de representación basados en lógica, aspecto que permite que su semántica sea especificada sin ambigüedades; (iii) los lenguajes de representación permiten realizar consultas y especialmente razonamientos a través de cálculos en tiempo de ejecución [16].

³ Un concepto constituye la representación del significado de una cosa [12].

1.2 ELEMENTOS QUE CONFORMAN UNA ONTOLOGÍA

La conceptualización que pretende representar una ontología se especifica generalmente usando cinco tipos de componentes: conceptos, relaciones, funciones, axiomas e instancias [13].

- **Conceptos o clases:** Son las ideas básicas que se intentan representar. Los conceptos pueden ser clases de objetos, métodos, planes, estrategias, procesos de razonamiento, etc. En general dentro del campo de la automatización industrial, los conceptos pueden ser vistos como las variables más importantes que representan de la mejor manera el funcionamiento de ambiente industrial específico.
- **Relaciones:** Representan la interacción y enlace entre los conceptos de un dominio. En algunos casos suelen formar la taxonomía del dominio. Por ejemplo: subclase-de, parte-de, parte-exhaustiva-de, conectado-a, etc.
- **Funciones:** Son un tipo concreto de relación donde se identifica un elemento mediante el cálculo de una función que considera varios elementos de la ontología. Por ejemplo, pueden aparecer funciones como: asignar-fecha, categorizar-clase, etc.
- **Reglas de restricción:** Los axiomas son expresiones que son siempre ciertas. Pueden incluirse en una ontología con muchos propósitos, tales como definir el significado de los componentes ontológicos, definir restricciones complejas sobre los valores de los atributos, argumentos de relaciones, etc. Las ontologías que poseen una gran cantidad de axiomas se denominan ontologías pesadas, en contraste con las ontologías ligeras que no incluyen axiomas.

- **Instancias:** Representan una manifestación concreta de una clase (un objeto con valores determinados).

1.3 APLICACIONES Y VENTAJAS DEL USO DE LAS ONTOLOGÍAS

A continuación se presentan algunas de las aplicaciones y usos de las ontologías más comunes [12] [15] [20] [22]:

- **Integración de Información:** Un campo de aplicación prometedor de las ontologías es su uso en la integración de fuentes de información heterogéneas a nivel de esquema. Esto debido a que en muchas ocasiones, bases de datos diferentes almacenan el mismo tipo de información empleando diferentes modelos de datos. En este escenario, una ontología puede emplearse para mediar entre los esquemas de las bases de datos, permitiendo integrar la información de diferentes fuentes e interpretar los datos de una fuente bajo el esquema de otra.
- **Recuperación de Información:** La recuperación de información desde documentos es el mayor campo de aplicación para las ontologías. La idea detrás de la recuperación de información basada en ontologías es incrementar la precisión de los resultados obtenidos teniendo en cuenta la semántica de la información contenida dentro de las consultas como en los documentos, pasando con esto, de palabras clave a conceptos y relaciones ontológicas.
- **Administración de conocimiento:** Al interior de compañías u otras asociaciones organizadas, el conocimiento individual puede verse como un recurso estratégico que es susceptible de compartirse y mantenerse

sistemáticamente dentro de la denominada gestión de conocimiento. En este aspecto, las ontologías proporcionan un modelo conceptual compartido de un dominio, que permite unificar la administración de conocimiento, conectando los sistemas técnicos para la navegación, almacenamiento búsqueda e intercambio del mismo.

- **Sistemas expertos:** Este tipo de aplicaciones permite formalizar el dominio de conocimiento de un experto, como por ejemplo el de diagnóstico médico o el de asesoramiento legal, mediante una ontología de dominio sobre la cual se presentan interrogantes sofisticados que son resueltos empleando mecanismos de razonamiento.

Los beneficios de utilizar ontologías se pueden resumir de la siguiente forma:

- Proporcionan una forma de representar y compartir el conocimiento utilizando un vocabulario común.
- Proporcionan un protocolo específico de comunicación.
- Permiten una reutilización del conocimiento.

1.4 CLASIFICACIÓN DE LAS ONTOLOGÍAS

Actualmente existen varios criterios que permiten clasificar los diferentes tipos ontologías con base en algunos parámetros [12]. Por esta razón a continuación se presentarán algunos de los más relevantes encontrados en la literatura.

1.4.1 Clasificación con base en el campo y el uso de la conceptualización

Esta clasificación está determinada por el nivel de complejidad empleado para describir los conceptos y las relaciones entre ellos [13][21].

A continuación se presenta las categorías de menor a mayor grado de complejidad que se distinguen dentro de esta clasificación [13][21]:

- **Ontologías terminológicas, lingüísticas:** Especifican los términos usados para representar el vocabulario de un dominio.
- **Ontologías de información:** Especifican la estructura de los registros de la base de datos. Los esquemas de bases de datos serían un ejemplo.
- **Ontologías para modelar conocimiento:** Especifican conceptualizaciones de conocimiento. Estas ontologías tienen una estructura interna mucho más rica que los anteriores tipos de ontologías y éstas son las ontologías que interesan a los desarrolladores de sistemas basados en conocimiento.

Una clasificación alternativa a la anteriormente presentada propone las siguientes tres categorías [13][21]:

- **Ontologías del dominio:** Corresponden a ontologías que contienen todos los conceptos asociados a un dominio particular.
- **Ontologías de tarea:** Establecen la forma en la cual se puede usar el conocimiento del dominio para realizar tareas específicas. De esta forma, una aplicación podría realizar búsquedas de información mientras otra podría gestionar la asignación de bloques de memoria libre.

- **Ontologías generales:** Contienen descripciones generales sobre objetos, eventos, relaciones temporales, relaciones causales, modelos de comportamiento y funcionalidades.

1.4.2 Clasificación por motivación

Dentro de esta clasificación se distinguen las siguientes categorías [12][13]:

- **Ontologías para la representación de conocimiento:** Permiten explicar las conceptualizaciones que subyacen de los formalismos de representación de conocimiento
- **Ontologías genéricas:** Definen conceptos considerados genéricos en diferentes áreas. Ejemplos de tales conceptos serían componente, subclase, proceso, estado, etc. Estas ontologías son reutilizables en diferentes dominios. Se llaman también ontologías abstractas o super teorías porque permiten definir conceptos abstractos.
- **Ontologías del dominio:** Definen conceptualizaciones específicas del dominio. Las metodologías actuales de adquisición de conocimiento distinguen entre ontologías y conocimiento del dominio; el último describe situaciones actuales del dominio, mientras que las ontologías imponen descripciones sobre la estructura y contenido del conocimiento del dominio.
- **Ontologías de aplicación:** Están ligadas al desarrollo de una aplicación concreta. Tales ontologías cubren los aspectos relacionados con aplicaciones particulares. Típicamente, estas ontologías toman conceptos

de ontologías del dominio y genéricas, así como métodos específicos para realizar la tarea, por lo que no son muy adecuadas para ser reutilizadas.

1.4.3 Clasificación según el ámbito del conocimiento

Clasifica las ontologías con base en el ámbito del conocimiento al cual se apliquen en las siguientes categorías [12][13]:

- **Ontologías generales:** Son las ontologías de nivel más alto ya que describen conceptos generales (espacio, tiempo, materia, objeto, etc.).
- **Ontologías de dominio:** Describen el vocabulario de un dominio concreto del conocimiento.
- **Ontologías específicas:** Son ontologías especializadas que describen los conceptos para un campo limitado del conocimiento o una aplicación concreta.

1.4.4 Clasificación según el tipo de agente

Permite la clasificación de ontologías teniendo en cuenta el agente (software o humano) al cual vayan destinadas [12][13].

- **Ontologías lingüísticas:** Se vinculan a aspectos lingüísticos, esto es, a aspectos gramáticos, semánticos y sintácticos destinados a su utilización por los seres humanos.
- **Ontologías no lingüísticas:** Destinadas a ser utilizadas por robots y agentes inteligentes.

- **Ontologías mixtas:** Combinan las características de las anteriores.

1.4.5 Clasificación según el nivel de abstracción y razonamiento lógico

Esta clasificación permite diferenciar las ontologías considerando el nivel de abstracción y el grado de razonamiento que permitan realizar a partir de la forma de descripción empleada [12][13].

- **Ontologías descriptivas:** Incluyen descripciones, taxonomías de conceptos, relaciones entre los conceptos y propiedades, pero no permiten inferencias lógicas.
- **Ontologías lógicas:** Permiten inferencias lógicas mediante la utilización de una serie de componentes como las axiomas formales.

1.5 METODOLOGÍAS PARA LA CONSTRUCCIÓN DE ONTOLOGÍAS

Similar a cualquier componente software, la construcción de una ontología puede ser mejorada mediante la aplicación de algún tipo de metodología. El propósito de emplear una metodología es procurar un óptimo resultado siguiendo un conjunto de pasos, los cuales usualmente están basados en procedimientos cuyo éxito ha sido comprobado en la práctica. Las ontologías no son la excepción a esta regla, y de esta manera la mayoría de metodologías destinadas a su construcción están basadas en la experiencia de personas o grupos involucradas(os) en su desarrollo [12]. En varios casos, estas metodologías son extraídas de la manera en la cual ha sido construida una ontología particular, de la cual han heredado su nombre.

Las metodologías propuestas en la literatura se pueden clasificar en dos grupos: (1) las destinadas a la creación de ontologías desde cero, y (2) las destinadas a la reutilización de ontologías disponibles [20].

En este sentido, debido que a que no se encontró una ontología que pudiese reutilizarse para satisfacer las necesidades de este proyecto, entonces se vio la necesidad de crear una a partir de cero. Por esta razón, a continuación se describirán brevemente algunas de las metodologías más importantes encontradas dentro de la categoría anteriormente mencionada.

1.5.1 Metodología CYC

La metodología Cyc surgió de la experiencia obtenida a partir del desarrollo de la base de conocimiento Cyc, la cual contiene una gran cantidad de conocimiento común o general [17] [21].

La metodología Cyc recomienda los siguientes pasos [22][23][24]:

1. **Codificación manual de conocimiento:** Consiste en la extracción de conocimiento a partir de documentos, sin la utilización de sistemas de procesamiento de lenguaje natural o de aprendizaje. Este paso se realiza de forma manual, debido a que dichos sistemas no cuentan con el suficiente vocabulario común que les permita efectuar una adecuada búsqueda de conocimiento.
2. **Codificación de conocimiento asistido por herramientas software:** Consiste en la utilización parcial de herramientas de procesamiento de lenguaje natural y de aprendizaje, con el propósito de descubrir nuevo conocimiento común a partir del conocimiento obtenido en el paso anterior.

3. **Codificación de conocimiento realizado por herramientas software:** En este paso el descubrimiento de nuevo conocimiento es realizado principalmente por las herramientas. La función del desarrollador consiste en recomendar las fuentes de conocimiento y explicar algunos fragmentos de texto.

1.5.2 Metodología de USCHOLD Y KING

Esta metodología nace de la experiencia obtenida en el desarrollo de “*Enterprise Ontology*”, una ontología para modelar los procesos empresariales [17][25]. Esta metodología proporciona las siguientes pautas para desarrollar ontologías [21][22][23]:

1. **Identificar el propósito de la ontología:** Consiste en determinar el porqué de la construcción de la ontología y cuáles serán los usos de la misma.
2. **Construir la ontología:** Comprende el desarrollo de la ontología a través de tres actividades. La primera consiste en capturar la ontología a través de la identificación de los conceptos y relaciones entre conceptos, así como los términos utilizados para referirse a dichos elementos. La segunda actividad consiste en codificar la ontología empleando un lenguaje formal. La tercera actividad consiste en integrar la ontología obtenida con alguna ontología existente.
3. **Evaluar la ontología:** Consiste en realizar un juicio técnico de la ontología, entorno software asociado y documentación con respecto a un marco de referencia (especificación de requerimientos, preguntas de competencia, etc.).

1.5.3 Metodología de GRÜNINGER Y FOX

Esta metodología (Figura 1) surgió del desarrollo del proyecto TOVE (Toronto Virtual Enterprise), dentro del cual se desarrollaron un conjunto de ontologías para el dominio de los procesos de negocios. Esta metodología propone los siguientes seis pasos [12][17][21][22][25]:

1. **Identificar los escenarios de motivación:** Aquí se reconocen las posibles aplicaciones en las cuales podrá emplearse la ontología.
2. **Elaborar preguntas informales de competencia:** Las cuales son realizadas en lenguaje natural y son empleadas con la finalidad de determinar el alcance de la ontología.
3. **Especificar la terminología utilizando lógica de primer orden:** Se realiza la formalización de la terminología extraída a partir de las preguntas de competencia realizadas en el paso anterior.
4. **Especificar preguntas formales de competencia:** Consiste en formalizar las preguntas de competencia, para lo cual las preguntas empleadas en el paso dos son re-escritas en lógica de primer orden.
5. **Especificar los axiomas empleando lógica de primer orden:** Esta metodología propone hacer uso de axiomas formales para definir los términos y restringir su interpretación dentro de la ontología.
6. **Establecer las condiciones para calificar la integridad de la ontología:** Consiste en definir las condiciones bajo las cuales la ontología puede dar respuesta a las preguntas de competencia que han sido formalizadas.

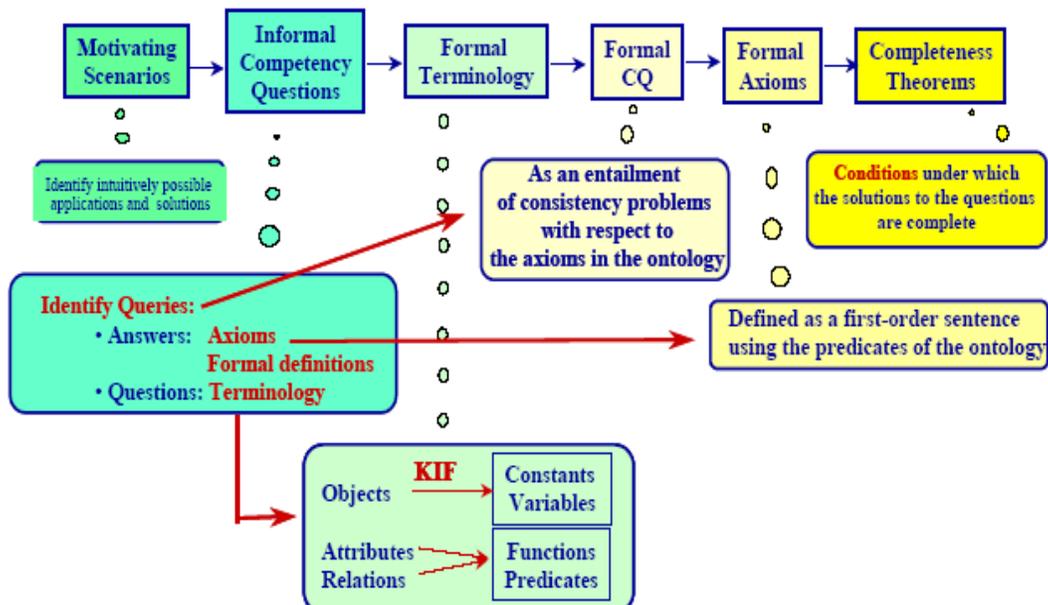


Figura 1. Metodología de Grüninger y Fox [21]

1.5.4 Metodología KACTUS

La metodología KACTUS basa la construcción de una ontología (que representa el conocimiento requerido por una aplicación) en la reutilización de conceptos definidos en otras ontologías [23]. De esta manera el proceso de construcción está condicionado al análisis y adaptación de conceptos de otras ontologías [23].

La metodología KACTUS comprende las siguientes etapas [17][21][25][26]:

1. **Especificación de la aplicación:** Se identifica el contexto de la aplicación a través de una lista de términos relacionados del dominio.
2. **Diseño preliminar de la ontología:** Comprende la evaluación de las ontologías superiores disponibles con el fin de seleccionar sus categorías relevantes, las cuales posteriormente son redefinidas y/o extendidas para finalmente incluirlas en la nueva ontología.

3. **Refinamiento y estructuración:** Se descompone la ontología para reorganizar su estructura con el fin de obtener un diseño modular.

1.5.5 Methontology

Es una metodología creada en el Laboratorio de Inteligencia Artificial de la Universidad Técnica de Madrid [21]. La Figura 2 muestra el conjunto de actividades comprendidas dentro de Methontology, las cuales son descritas a continuación [17][22][25][27]:

1. **Actividades de administración del proyecto:** Estas actividades involucran la planeación, seguimiento de tareas y control de la calidad con el fin de obtener una ontología de calidad.
2. **Actividades orientadas al desarrollo:** Dentro de las cuales se encuentran:
 - **Especificación:** La cual consiste en definir el alcance y la granularidad de la ontología.
 - **Conceptualización:** Aquí se realiza la organización y estructuración del conocimiento adquirido a través de un medio de representación (tablas, diagramas UML, etc.) que sea independiente del formalismo y lenguaje de implementación en que vaya a ser formalizada o implementada la ontología.
 - **Implementación:** Consiste en transformar el modelo conceptual obtenido en un modelo formalizado, el cual es implementado a través de un lenguaje de ontologías como OWL.
 - **Mantenimiento.**

3. **Actividades de soporte:** Incluyen la recopilación de conocimiento, evaluación de la ontología, reutilización de la ontología y documentación.

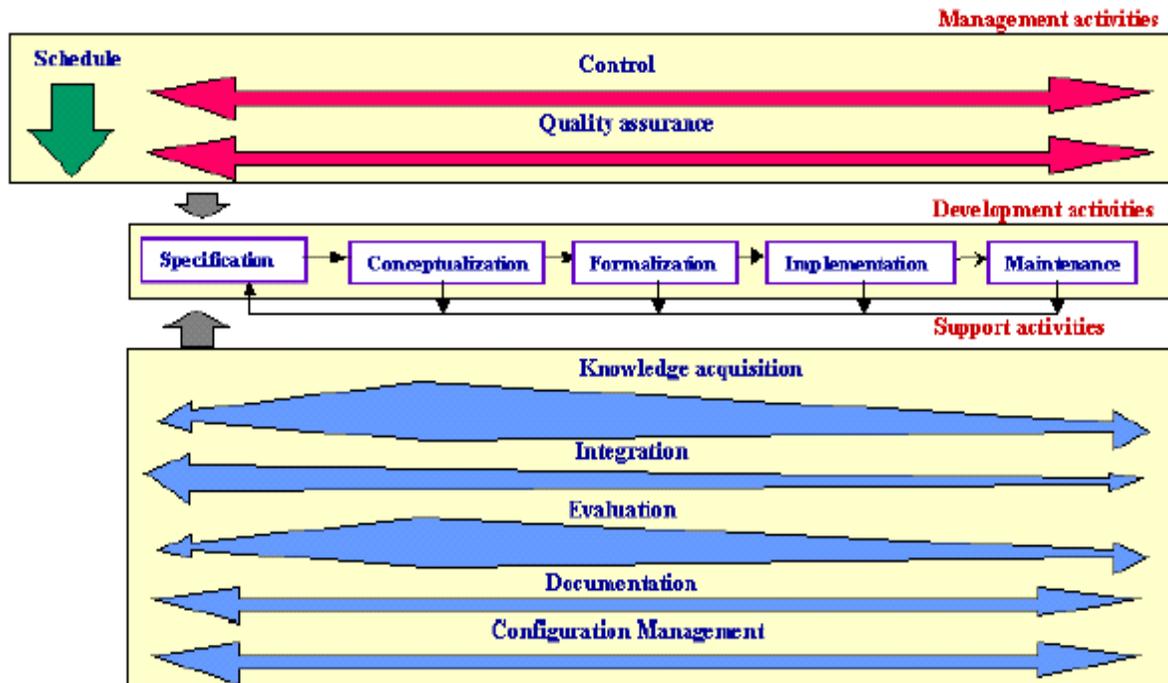


Figura 2. Diagrama del ciclo de vida en METHONTOLOGY [21]

1.5.6 Metodología Sensus

Esta metodología (Figura 3) permite derivar ontologías de dominio a partir de ontologías de grandes proporciones a través de un enfoque *top-down* (de lo general a lo particular) [21]. Para realizar esto, los autores proponen identificar un conjunto de términos semilla que son relevantes dentro de un dominio particular. Una vez obtenidos dichos términos, éstos se enlazan manualmente a una ontología de grandes proporciones, que en este caso corresponde a la ontología Sensus (contiene más de 70,000 conceptos). A continuación se ejecuta un algoritmo que retorna todos los términos que se encuentran entre la trayectoria de

los términos semilla y la raíz de Sensus. Si por algún motivo se considera que falta algún término que sea relevante dentro del dominio, entonces se procede a enlazar el término a la ontología y posteriormente ejecutar nuevamente el algoritmo de búsqueda de términos [21] [17] [25].

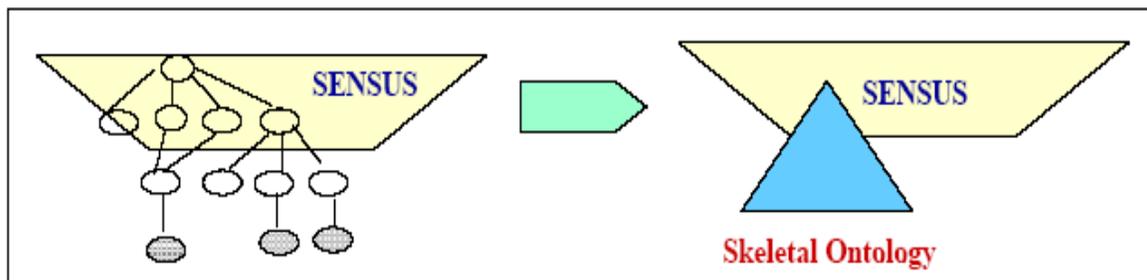


Figura 3. Metodología SENSUS [21]

1.5.7 Metodología On-To-Knowledge

La metodología On-To-Knowledge (OTK) está enfocada hacia la construcción de sistemas basados en conocimiento, en los que se empleen las ontologías como parte fundamental [17][25].

OTK define las siguientes etapas para la creación de una ontología [21] [22] [25]:

1. **Estudio de factibilidad:** OTK sigue la metodología CommonKADS con el propósito de establecer si se justifica la construcción de la ontología.
2. **Inicio:** Se determina el alcance y el objetivo de la ontología. Además de esto se identifican las entidades más importantes así como las fuentes de conocimiento y los expertos humanos que puedan ser consultadas(os).

3. **Refinamiento:** En esta etapa primero se extrae el conocimiento relevante desde las fuentes y expertos humanos identificados en la etapa anterior, para luego proceder a su formalización.
4. **Evaluación:** Permite verificar la utilidad de la ontología desarrollada y de su entorno software asociado.
5. **Mantenimiento:** Se establece quién es el responsable de la labor de mantenimiento y cómo ésta debería llevarse a cabo.

1.5.8 Terminae

Terminae aporta tanto una metodología como una herramienta para la construcción de ontologías a partir de textos [17][29]. Se basa en un análisis lingüístico de los textos, el cual se realiza mediante la aplicación de diferentes herramientas para el procesamiento del lenguaje natural. En particular se usan dos herramientas [29]: (1) Syntex para identificar términos y relaciones; y (2) Camaleón para identificar roles o relaciones. Estas herramientas se basan en la misma hipótesis lingüística: el significado de las frases y las palabras es específico para un dominio y puede ser inferido de la observación de regularidades en documentos. La metodología funciona como sigue [29]. Mediante la aplicación de Syntex obtenemos una lista de posibles palabras y frases del texto y algunas dependencias sintácticas y gramaticales entre ellas. Estos datos se usan como entrada para el proceso de modelado junto con el texto original. De esta forma, la identificación de conocimiento se basa en dos tareas que se realizan alternativamente [29]:

1. Explorar los resultados Syntex para identificar conocimiento importante o decidir cómo representar alguna información de acuerdo con el uso de las palabras en el texto.
2. Extraer sistemáticamente del texto tanto conocimiento como sea posible.

Cada pieza de conocimiento puede ser representada en el modelo de conocimiento de Terminae, cuyo lenguaje de representación de conocimiento posee las siguientes primitivas: fichero terminológico (términos), conceptos genéricos (clases), conceptos primitivos (instancias) y roles (relaciones). El siguiente paso es normalizar el conocimiento para obtener una ontología bien estructurada, donde cada concepto quede justificado por sus relaciones con otros conceptos. Esta metodología sugiere aplicar criterios diferenciadores para hacer explícitas las propiedades comunes y diferentes de un concepto con sus respectivos conceptos padre y hermanos debidas a sus roles. La última etapa es la formalización de la ontología en el lenguaje formal Terminae, que es un tipo de lógica descriptiva. Una función de clasificación sirve para comprobar la corrección de las definiciones de conceptos genéricos, ya que sólo pueden ser definidos si tienen roles diferenciados [29].

2 METODOLOGÍA ONTOPRIN

Como se presentó en el capítulo anterior, el proceso de desarrollo de una ontología no cuenta con un método o metodología estándar para su construcción. Por el contrario dentro de la literatura se pueden encontrar un gran número de metodologías, en su mayoría obtenidas a partir de los pasos empleados para la creación de una ontología en particular.

En la Tabla 1 se resumen las ventajas y desventajas que fueron encontradas a partir de la revisión de las metodologías presentadas anteriormente.

METODOLOGÍA	VENTAJAS	DESVENTAJAS
CYC	La extracción manual permite incluir una mayor cantidad de conocimiento común	Las actividades no son claras y no están explicadas en detalle lo que genera una dificultad para desarrollarlas totalmente. No se recomiendan técnicas o herramientas que se puedan emplear a través de cada uno de los pasos de esta metodología y que podrían colaborar en el desarrollo de la ontología.
USCHOLD Y KING	Hace referencia a la identificación del propósito de la creación de la ontología. Se plantea un proceso de	No están descritas ni explicadas las actividades que desarrollar de una forma clara y son muy generales. No se realizan recomendaciones para la

	<p>evaluación y documentación de la ontología.</p>	<p>formalización del conocimiento.</p> <p>No se recomiendan técnicas dentro de cada uno de los pasos de desarrollo de la ontología.</p> <p>Sólo se ha aplicado en el campo de los negocios.</p>
GRÜNINGER Y FOX	<p>El desarrollo de preguntas en caminadas a obtener la información necesaria para la creación de ontologías y el establecimiento de escenarios.</p>	<p>Las actividades no están explicadas en detalle.</p> <p>No hay recomendaciones de técnicas o herramientas para el desarrollo de las actividades.</p>
KACTUS	<p>Se define el contexto de la ontología través de la identificación de términos relevantes.</p>	<p>Metodología bastante pequeña y poco detallada.</p> <p>No se muestran recomendaciones para la formalización del conocimiento.</p> <p>No se recomiendan técnicas para el desarrollo de ontología con esta metodología.</p>
METHONTOLOGY	<p>Posee mucha documentación en cada paso. Es la excepción a las demás metodologías.</p>	<p>A pesar de que es una de las metodologías más detalladas, es muy compleja para ser aplicada en nuestro campo de acción.</p> <p>No hay recomendaciones para la formalización del conocimiento.</p> <p>Puede ser usada en diferentes campos especialmente en los químicos.</p>

SENSUS	La obtención de una estructura de términos a partir de una ontología de grandes proporciones.	En esta metodología no se tiene en cuenta el objetivo o propósito de la ontología. No se recomiendan técnicas que puedan ser utilizadas para el proceso de desarrollo y construcción de la ontología.
ON-TO-KNOWLEDGE	El estudio de la viabilidad previo a la creación de la ontología	Es una metodología muy corta y sobre la cual en la literatura encontrada no se explican en detalle la forma como se desarrollan los pasos que la conforman, tampoco se encuentran recomendaciones para la construcción de una ontología basada en esta metodología.
TERMINAE	El uso de herramientas que agilizan la construcción de la ontología.	No se explica en detalle cómo se realiza cada paso ni cada actividad dentro de ella. Está ligada a la formalización en el lenguaje Terminae.

Tabla 1. Ventajas y desventajas de las metodologías existentes para crear ontologías

En términos generales, a pesar de que se han propuesto un gran número de metodologías para la construcción de ontologías, el bajo nivel de detalle presentado por cada una de las metodologías estudiadas hace que se generen interrogantes de cómo ejecutar cada paso que las componen de una manera efectiva. De esta manera, el inconveniente anteriormente mencionado brinda la posibilidad de que se plantee una metodología que permita guiar el proceso de desarrollo de la ontología que se pretende realizar dentro de este trabajo.

2.1 DESCRIPCIÓN DE LA METODOLOGÍA ONTOPRIN

ONTOPRIN (Ontologías para Procesos Industriales) constituye la metodología planteada para guiar el proceso de desarrollo de la ontología que se aborda dentro de este trabajo y que ha sido concebida para su utilización en el dominio de la automatización de procesos industriales de manufactura, que es en el cual está enmarcado el presente proyecto.

La metodología propuesta ha surgido a partir de un estudio de las metodologías (CYC, USCHOLD Y KING, GRÜNINGER Y FOX, KACTUS, METHONTOLOGY, SENSUS, ON-TO-KNOWLEDGE y TERMINAE) descritas en el capítulo anterior. De dichas metodologías se han tomado las indicaciones o recomendaciones clave que permitieran guiar de una manera clara el proceso de construcción de una ontología destinada al dominio de la automatización de procesos industriales de manufactura.

La metodología ONTOPRIN consta de los pasos que serán descritos a continuación.

2.1.1 Paso 1: Identificación del propósito

Este paso consiste en definir clara y precisamente lo que se quiere lograr (los objetivos que se quieren alcanzar con la implementación de la ontología) con el fin de delimitar el alcance de la ontología que se pretende construir. Para esto es importante identificar el ámbito en el cual va emplearse la ontología, así como las aplicaciones que van emplear la ontología, el tipo de usuarios para los que está destinada la ontología y la funcionalidad que pueden brindar las aplicaciones a los

usuarios. La Tabla 2 muestra un ejemplo de una forma en la que se puede especificar el propósito de la ontología.

ÁMBITO DE TRABAJO DE LA ONTOLOGÍA	
OBJETIVO GENERAL	
OBJETIVOS ESPECÍFICOS	

Tabla 2. Ejemplo de una tabla para especificar el propósito de la ontología

2.1.2 Paso 2: Extracción manual de terminología de diferentes fuentes de información

Este paso consiste en obtener la terminología empleada para representar conceptos, atributos, relaciones e instancias dentro del dominio de interés de diversas fuentes (primarias, secundarias, etc.). Algunos ejemplos de este tipo de fuentes en un proceso industrial pueden ser los operarios, supervisores, libros de registro, sistemas de información, señalizaciones, etc. Este proceso se realiza mediante un estudio de campo, por ejemplo a través del desarrollo de encuestas (Figura 4), análisis de videos que se tengan sobre el proceso, visitas técnicas, análisis de sistemas de información (Figura 5), análisis de folletos, manuales del proceso, etc.

CUESTIONARIO SOBRE EL FUNCIONAMIENTO DE UNA PLANTA DE TRATAMIENTO DE AGUA POTABLE

El siguiente cuestionario se redactó con el fin de conocer sobre el proceso de producción de agua potable y obtener la información necesaria para un mejor conocimiento de su distribución que se presenta dentro del proceso de tratamiento de agua, por tal motivo se de vital importancia la información que se da como resultado de responder de la siguiente manera: cuando se pide el plantamiento de objetivos y desarrollo de metas a dicho proceso que no tienen mayor importancia, así sea necesario.

INTRODUCCION

1. Empresa: _____
2. Nombre: _____
3. Cargo: _____
4. Lugar de trabajo: _____
5. Horas de trabajo: _____

VARIABLES DEL PROCESO
NIVEL, TORREDAZ, CARGA, FLUJO, DOSIFICACION, VALVULAS Y CAPTACION

6. Cuales son las funciones principales que se realizan dentro de la empresa? _____
7. Cual es el nivel de la planta que se desarrolla el proceso de producción de agua en la planta?
 - a) nivel 1
 - b) nivel 2
 - c) nivel 3
 - d) nivel 4
8. De que parte o etapa del proceso de producción de agua tiene a cargo dentro de la planta? _____

9. De dicho proceso como mencionado, cual es el nivel que presenta la mayor importancia para el desarrollo? _____
10. Explique brevemente como es el desarrollo de dicho proceso? _____
11. Cual es el nivel que presenta mayor importancia? _____
12. Dónde se encuentran y como se desarrolla el desarrollo como que puede ser mejorado dicho desarrollo? _____
13. Dónde se encuentran y como se desarrolla el desarrollo como que puede ser mejorado dicho desarrollo? _____
14. Dónde se encuentran y como se desarrolla el desarrollo como que puede ser mejorado dicho desarrollo? _____
15. Dónde se encuentran y como se desarrolla el desarrollo como que puede ser mejorado dicho desarrollo? _____

Figura 4. Encuesta para la extracción de terminología informal

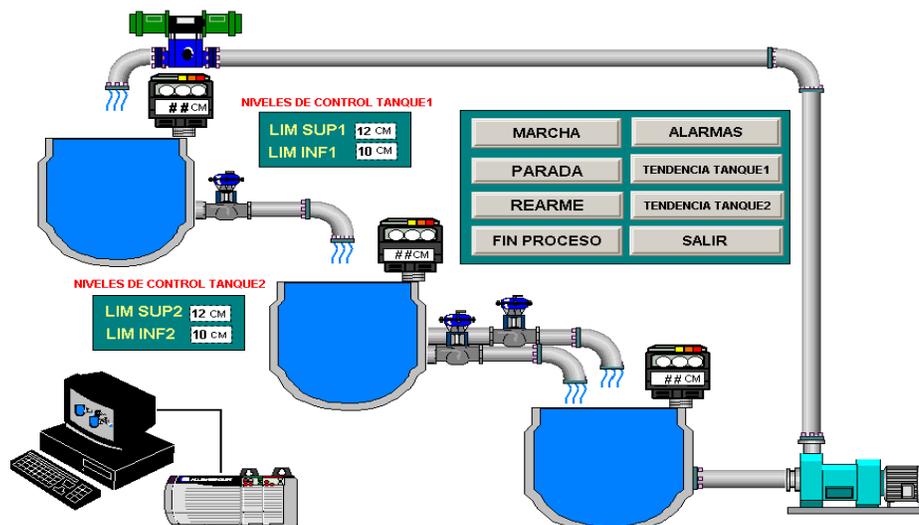


Figura 5. Ejemplo de un sistema SCADA

Para el desarrollo de este paso es necesario tener en cuenta algunos aspectos importantes que nos ayuden garantizar que la información obtenida sea la adecuada y la realmente necesaria; estas recomendaciones son:

1. En caso de realizar una encuesta o una entrevista, la información que se obtenga debe estar en bruto (términos o palabras utilizadas por operarios); esto quiere decir que no esté manipulada y que en el proceso de extracción no se debe influenciar de ninguna manera una respuesta por parte del encuestador. Las preguntas aquí pueden ser de tipo cerradas (delimitan la respuesta a algunas pocas alternativas, para lo cual se debe tener mucho tacto y cuidado en la selección de las alternativas) o abiertas (no delimitan la respuesta). Las preguntas deben ser claras y comprensibles. No deben incomodar al entrevistado. Una pregunta debe indagar un único aspecto, las preguntas no deben incluir respuestas, deben estar hechas pensando en el grado de intelectualidad de los encuestados, etc. El entrevistador debe procurar tener una expresión verbal muy buena y debe estar preparado para realizar preguntas que no estén preparadas para obtener alguna información importante.
2. Con respecto a la extracción de información que se realiza mediante videos, folletos, etc., la información sobre el proceso debe ser extraída tal cual como se expresa en cada una de estas fuentes (es decir, empleando los términos que en ellos se emplean).

Luego de realizada la extracción de información es necesario organizarla, sin que ello signifique modificarla, con el fin de analizar los datos y la información obtenida. Este proceso se realiza mediante herramientas como lo son cuadros sinópticos, diagramas de tortas (Figura 6), diagramas de barras (Figura 7), cuadros de descripción (Figura 8), etc.

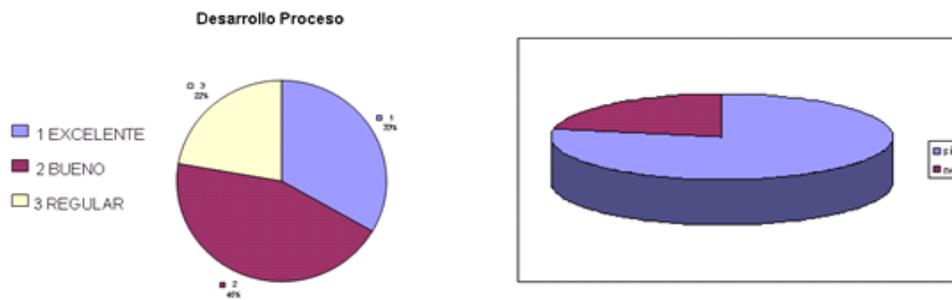


Figura 6. Diagrama de torta para el análisis de información

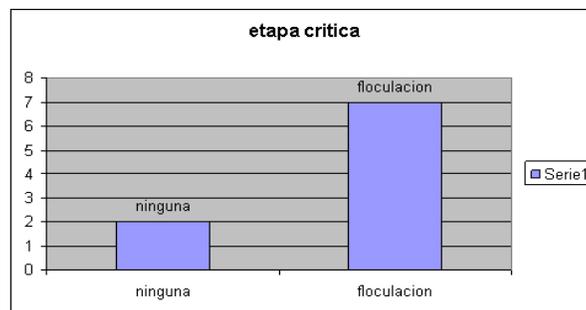


Figura 7. Diagrama de barras para el análisis de información

ETAPA	IMPORTANCIA	DATOS	UTILIDAD

TIPO DE EMERGENCIA	ALARMAS	PLAN DE CONTINGENCIA	EXISTENCIA DE MANUAL

Figura 8. Tabla de ejemplo para el análisis información.

2.1.3 Paso 3: Confrontación de la terminología extraída con una terminología técnica

Luego de extraer la terminología a partir del estudio de campo, se realiza una comparación con una terminología más técnica dependiendo del ambiente industrial. Por ejemplo, con la definida dentro de algún estándar o norma existente

(preferiblemente de carácter internacional) dentro del dominio de interés. La idea es confrontar la terminología informal extraída en el paso anterior con una terminología (igualmente informal) que represente los conceptos de una manera comúnmente aceptada por un grupo de usuarios y empresas del dominio de interés. De esta manera se pueden establecer relaciones entre términos similares (sinónimos) que puedan ser manejadas por la ontología con el fin de evitar ambigüedades.

Este proceso puede realizarse a través de una tabla comparativa como la que presenta en Tabla 3.

TERMINOLOGÍA ESTÁNDAR	TERMINOLOGÍA PARTICULAR	DESCRIPCIÓN

Tabla 3. Tabla de confrontación de terminología informal y terminología más general

2.1.4 Paso 4: Proceso de conceptualización

Consiste en identificar los conceptos, relaciones entre conceptos, atributos e instancias dentro del dominio de interés empleando para ello la terminología obtenida en los pasos anteriores. El propósito de este paso es obtener un modelo que sea fácilmente entendible, para lo cual se pueden emplear representaciones basadas en mapas conceptuales, diagramas UML, tablas o su combinación. De esta forma se obtiene un modelo independiente del lenguaje de implementación que sea seleccionado.

Para la identificación y organización de conceptos dentro de una taxonomía se pueden emplear las siguientes estrategias propuestas dentro de la metodología de Ushold and King:

1. *Top-Down*: Se identifican en primera instancia los conceptos más generales y luego éstos son especializados hacia conceptos más específicos.
2. *Bottom-Up*: Propone que los conceptos más específicos sean identificados en primera instancia para luego ser generalizados hacia conceptos más abstractos.
3. *Middle-Out*: Sugiere que los conceptos más importantes sean identificados en primera instancia para que posteriormente sean generalizados y especializados hacia otros conceptos.

La Figura 9 muestra un ejemplo en la cual se representa una taxonomía de conceptos encontrada dentro del dominio de la automatización de procesos industriales empleando un árbol de conceptos como medio de representación.

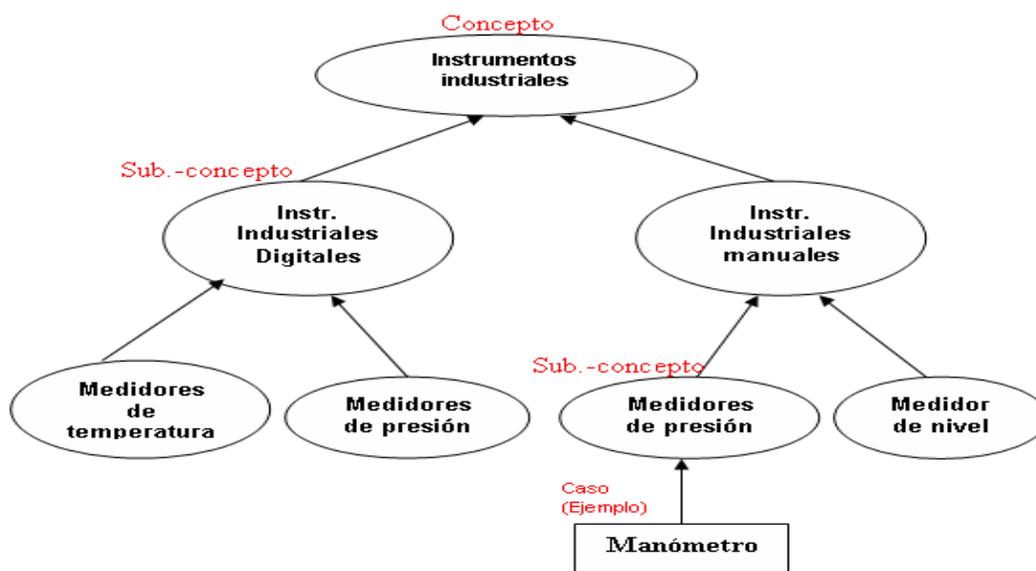


Figura 9. Ejemplo de un árbol de conceptos (clases)

Finalmente cabe anotar que la labor de conceptualización puede facilitarse a través de la utilización de modelos conceptuales previamente definidos, como son los encontrados en algunos estándares industriales (por ejemplo los estándares ISA-88, ISA-95, etc.), en los cuales se establecen las clases, relaciones y atributos dentro de un ámbito particular. Esta alternativa permitiría que la ontología construida tenga una mayor validez debido a la utilización de modelos que han sido el resultado de un consenso por parte de un grupo de expertos del dominio y cumpliría cabalmente una de las características contenida dentro de la definición de ontología que es la de capturar una “conceptualización compartida”.

2.1.5 Paso 5: Implementación

Representa la formalización de la conceptualización obtenida en el paso anterior a través de un lenguaje ontológico (OWL, DAML+OIL RDFS, etc.). Esta etapa permite la obtención de modelos procesables a nivel computacional, con el propósito de obtener información implícita a partir de las descripciones realizadas de manera explícita. Para su realización se debe seleccionar el lenguaje en que se quiera codificar la ontología (influenciado por algún formalismo de representación de conocimiento como la Lógica de Primer Orden o la Lógica Descriptiva) y la herramienta que facilite dicho proceso (editor de ontologías).

Para la selección de la herramienta de codificación se pueden considerar aspectos como la facilidad de aprendizaje, facilidad de desarrollo, funcionalidad, soporte y compatibilidad con otros sistemas o software con que se vaya a trabajar (razonadores, por ejemplo). Esto con el fin de escoger la herramienta que brinde las mejores prestaciones y se adapte a los requerimientos particulares por parte del desarrollador o grupo de desarrolladores.

Para la realización de la codificación es recomendable emplear convenciones de nombrado con el fin de obtener una ontología consistente. Algunas recomendaciones útiles se presentan a continuación [30]:

- Iniciar el nombre de los conceptos con letra mayúscula y los nombres de las relaciones e instancias con letra minúscula.
- En el caso de nombres compuestos se puede emplear el delimitador de guión bajo “_”. o emplear la notación CamelCase, la cual elimina la necesidad de delimitadores.
- Emplear nombres en singular para los conceptos.
- Emplear los prefijos “tiene” o “es” y el sufijo “de” cuando se emplee el prefijo “es” para nombrar las propiedades.
- Evitar abreviaciones, a menos que sean ampliamente reconocidas como URI para Identificador Uniforme de Recursos.

2.1.6 Paso 6: Refinamiento y estructuración de la ontología

Son los últimos ajustes y las últimas adecuaciones al modelo ontológico que se ha realizado en los pasos anteriores. No hay un método o unas pruebas estándar(es) para evaluar una ontología (de manera análoga a las metodologías para la creación de las mismas); sin embargo, existe un serie de recomendaciones para realizar su evaluación que se pueden seguir para comprobar que el modelo ontológico ha sido bien construido.

Una vez construido un modelo ontológico, éste se ha de evaluar exhaustivamente con los interesados para comprobar que refleja el fenómeno que se quiere representar. Para esto es importante tener en cuenta los siguientes aspectos [31]:

- **Lo que se evalúa:** Es el modelo creado y su funcionamiento.
- **Los que lo evalúan:** Corresponden a los diferentes tipos de clientes (usuarios que van a usar el modelo ontológico) que determinan dos aspectos: el primero es si el modelo ontológico representa adecuadamente el dominio para el cual ha sido desarrollado y el segundo es si el modelo cumple con los objetivos identificados al inicio de su construcción.

Para la realización de este proceso se pueden emplear encuestas, con el fin de indagar a los usuarios que van a emplear el modelo ontológico sobre los resultados obtenidos teniendo como referencia los objetivos planteados al inicio del proceso. El análisis de los resultados obtenidos es de gran importancia ya que a partir de ellos se puede conocer si la metodología o el modelo ontológico requieren de algún ajuste con el fin de lograr los resultados esperados.

3 APLICACIÓN DE LA METODOLOGÍA ONTOPRIN A UN CASO DE ESTUDIO

Luego de haber planteado la metodología ONTOPRIN en el Capítulo 2, a continuación se presentará una descripción de la forma como se abordó cada uno de los pasos que la componen, con el fin de llevar a cabo la construcción del modelo ontológico de este proyecto.

El caso de estudio seleccionado para la aplicación de la metodología ONTOPRIN corresponde al proceso de tratamiento de agua potable de la empresa de Acueducto y Alcantarillado de Popayán (ver Anexo B para una descripción completa de dicho proceso).

3.1 Paso 1: Identificación del propósito

ÁREA DE INTERES DE LA ONTOLOGÍA	El área de interés para el desarrollo de la ontología es la de procesos industriales. Más específicamente el proceso de tratamiento de agua potable de la empresa de Acueducto y Alcantarillado de Popayán.
OBJETIVO GENERAL	Realizar un modelo ontológico básico que establezca un dominio común y compartido entre la información generada por las aplicaciones SCADA (iFIX™ y RSView@32™) y el cliente interno, quien es el encargado de tomar las decisiones en el proceso de tratamiento de agua potable de la empresa de Acueducto y Alcantarillado de Popayán.
OBJETIVOS ESPECÍFICOS	Obtener un modelo ontológico que se adapte a las fuentes de información que se pretenden describir formalmente entre las aplicaciones iFIX™ y RSView@32™, con los requerimientos de información por parte del cliente interno.
	Construir un prototipo básico software que permita validar el modelo ontológico desarrollado, en materia de mejoramiento de procesamiento

	y refinamiento de la información que se le suministra al cliente interno para que realice sus diferentes tareas de toma de decisiones.
USUARIOS IDENTIFICADOS (PROTOTIPO)	Operador: Persona capacitada para el manejo de operaciones dentro del proceso de tratamiento de agua potable.
	Encargado de Producción: Persona responsable de la supervisión del proceso de tratamiento de agua potable.
FUNCIONALIDAD PROPORCIONADA (PROTOTIPO)	Consultar históricos. El operador y el encargado de producción a través del prototipo consultan los valores de las variables del proceso en un instante o periodo de tiempo.
	Consultar eventos. El operador y el encargado de producción a través del prototipo consultan los acontecimientos registrados de la ejecución del proceso.
	Solicitar reporte de desempeño de la producción: El encargado de producción a través del prototipo solicita un reporte de desempeño de la producción de las plantas de tratamiento de agua potable.
	Solicitar reporte de insumos y/o producto. El encargado de producción a través del prototipo solicita un reporte de los insumos y/o producto consumidos y/o obtenido en las plantas de tratamiento de agua potable.

Tabla 4. Descripción del propósito de la ontología

3.2 Paso 2: Extracción manual de terminología de diferentes fuentes de información

La información fue tomada mediante el desarrollo de encuestas a nivel operativo, visitas técnicas guiadas por el jefe de planta y por algunos operarios que se encontraban de turno, lectura de folletos de información de la empresa y análisis del sistema de supervisión y control encargado del proceso (Ver Anexo C).

A continuación se presenta la terminología obtenida junto con una breve descripción de la misma. Dicha terminología ha sido agrupada por categorías (etapas del proceso, equipos del proceso, variables del proceso e insumos del proceso) con el fin de organizarla y facilitar su comprensión.

- 1 **Etapas del Proceso:** Constituyen las etapas en las que se divide el proceso de tratamiento de agua potable que son identificadas por los encuestados (Tabla 5).

Término	Descripción
Bocatoma	Etapa donde se toma o capta el agua del río mediante un dique o a través de captación lateral.
Oxigenación	Etapa donde se quiere que el agua cruda gane moléculas de oxígeno, perdidas en el transporte del agua desde la bocatoma hasta la planta de tratamiento.
Dosificación Sulfato	Etapa en la que se adhiere sulfato de aluminio (agente decantador).
Mezclado rápido	Etapa donde se mezcla uniformemente el sulfato de aluminio agregado con el agua cruda.
Sedimentación	Etapa en la que por acción del sulfato de aluminio se forma el floc (agrupación de partículas en suspensión o mugre que se va al fondo).
Floculación	Etapa en la que se separa el floc del agua por desborde de la misma.
Filtración	Etapa donde se busca eliminar las partículas en suspensión restantes y que no se eliminaron en la sedimentación.
Desinfección	Etapa donde se adhiere cloro para eliminar las bacterias en el agua, además se le adhiere cal para mantener el pH óptimo y para proteger las tuberías
Bombeo ⁴	Etapa donde se realiza el envío de agua mediante bombeo a una parte de la ciudad que se encuentra a una mayor altura que la de la planta de tratamiento.
Almacenamiento	Etapa donde se guarda momentáneamente el agua tratada o potable antes de ser distribuida a la población.

Tabla 5. Etapas dentro del proceso de potabilización de agua cruda

⁴ Etapa encontrada únicamente en la planta de tratamiento del Tablazo

2 Equipos del proceso: Comprende los equipos y unidades relevantes dentro del proceso (Tabla 6).

Término	Descripción
Dosificador Sulfato	Equipo empleado para la dosificación de sulfato de aluminio, en la zona de vertedero.
Dosificador Cal	Equipo empleado para la dosificación de carbonato cal, en la zona de desinfección.
Dosificador Cloro	Equipo empleado para la dosificación de cloro, en la zona de desinfección.
Cisterna	Tanque en el cual se realiza el proceso de desinfección mediante la adición de cloro líquido.
Tanque Almacenamiento	Tanque en el cual se almacena el agua obtenida del proceso de potabilización, y partir del cual se distribuye hacia la comunidad.
Filtro	Unidad que realiza el proceso físico de separar las partículas coloidales y en suspensión del agua haciéndola circular a través de un material granular.
Bomba	Equipo que permite impulsar el agua hacia el sector que se encuentra a una mayor altura que la planta de tratamiento.
Medidor de pH	Dispositivo que permite medir el nivel de acidez del agua.
Medidor de turbiedad	Dispositivo que permite medir la apariencia del agua, debida a la presencia de material en suspensión que obstruye el paso de la luz.
Medidor de color	Dispositivo que permite medir el color en el agua debido a la presencia de metales, materia orgánica u otras sustancias disueltas o suspendidas en ella.
Computador	Permite visualizar niveles en los tanques, estado de algunas válvulas y bombas, caudales de entrada y salida, pH, turbidez y color en el tanque de almacenamiento.

Tabla 6. Equipos dentro del proceso de potabilización de agua cruda

- 3 **Variables del proceso:** Comprende las variables medidas en el agua procesada (Tabla 7).

Término	Descripción
Caudal	Volumen de agua que pasa en una unidad de tiempo determinada. Medida en Litros por segundo.
Nivel	Altura hasta donde se ha llenado un tanque con agua. Medida en metros.
pH	Nivel de acidez del agua.
Turbidez	Medida de la cantidad de partículas en suspensión en el agua cruda.
Color	Medida que representa la presencia de metales, materia orgánica u otras sustancias disueltas o suspendidas en el agua. Muestra de sustancias que están en el agua como minerales y bacterias. Medida en NTU (Unidades Nefelométricas de Turbidez).

Tabla 7. Variables dentro del proceso de potabilización de agua cruda

- 4 **Insumos del proceso:** Comprende las materias primas empleadas en el proceso de potabilización de agua (Tabla 8).

Término	Descripción
Sulfato de Aluminio	Coagulante empleado en el proceso de proceso potabilización de agua. Medido en Kilogramos.
Cloro	Desinfectante empleado en el proceso de proceso potabilización de agua. Medido en Litros.
Cal	Modificador de pH empleado en el proceso de proceso potabilización de agua. Medido en Kilogramos.

Tabla 8. Insumos del proceso de potabilización de agua cruda

3.3 Paso 3: Confrontación de la terminología extraída con una terminología técnica

La terminología identificada en este paso corresponde a la manejada dentro de los modelos de bases de datos y a la empleada por los operarios y el encargado de la producción para referirse a la información (terminología no incluida en el paso anterior). Para la realización de la confrontación se he empleado como referencia la terminología sugerida dentro de los estándares ISA-88.00.04 e ISA-95.00.02, los cuales han sido empleados para la creación de los modelos de bases de datos.

Las tablas presentadas a continuación muestran el paralelo entre los términos asociados al lenguaje de operador, encargado de producción, los modelos de bases de datos y el definido dentro de los modelos seleccionados de los estándares.

La Tabla 9 muestra un paralelo entre las diferentes terminologías establecidas y la terminología empleada en el modelo conjunto de datos propuesto dentro de ISA-88.00.04.

Término Estándar ISA 88	Término Modelo Base de datos Tablazo	Término Modelo Base de datos Tulcán	Término Operador	Término Encargado Producción
Conjunto Dato	Histórico	Histórico	Histórico	Histórico
Tiempo Inicio	Fecha Hora Inicio	Fecha Hora Inicio	Momento Inicio	Instante Inicio
Tiempo Finalización	Fecha Hora Finalización	Fecha Hora Finalización	Momento Finalización	Instante Finalización
ID Fuente Dato	Nombre Etiqueta	ID Tag	ID Origen Dato	ID Fuente Dato
Alias	Alias	ID Alternativo	Alias	Alias
Unidad Medición	Unidad Ingeniería	Unidad Medida	Unidad	Unidad

			Medida	Ingeniería
Fecha Registro	Fecha Hora	Fecha Hora	Momento Registro	Tiempo Registro

Tabla 9. Paralelo entre terminologías identificadas en relación al modelo conjunto de datos

La Tabla 10 muestra un paralelo entre las diferentes terminologías establecidas y la terminología empleada en el modelo de eventos propuesto dentro de ISA-88.00.04.

Término Estándar ISA 88	Término Modelo Base de datos Tablazo	Término Modelo Base de datos Tulcán	Término Operador	Término Encargado Producción
Valor Previo	Valor Anterior	Valor Pasado	Valor Pasado	Valor Previo
Mensaje Texto	Descripción	Descripción	Mensaje Informativo	Detalle Evento
ID Personal	ID Operador	ID Empleado	ID Operario	ID Empleado
Referencia Elemento de Procedimiento	Etapa	Procedimiento	Etapa	Fase
Categoría	Clasificación	Categoría	Clasificación	Clasificación
Tipo Alarma	Tipo Alarma	Clase Alarma	Clase de Alarma	Tipo de Alarma
Prioridad	Importancia	Relevancia	Importancia	Prioridad

Tabla 10. Paralelo entre terminologías identificadas en relación al modelo eventos

La Tabla 11 muestra un paralelo entre las diferentes terminologías establecidas y la terminología empleada en el modelo desempeño de la producción propuesto en ISA-95.00.02. En esta Tabla no se incluye la terminología del operador de planta debido a que dentro de sus funciones no está la revisión o el análisis de reportes de desempeño de la producción.

Término Estándar ISA 95	Término Modelo Base de datos Tablazo	Término Modelo Base de datos Tulcán	Término Encargado Producción
Localización	Sede	Sede	Sede
Definición Material	Definición Producto	Tipo Producto	Producto
	Químico	Insumo	Insumo químico
Unidad Medición Cantidad	Unidad Ingeniería	Unidad Medida	Unidad Ingeniería
Nombre propiedad	Propiedad	Característica	Característica
Unidad Medición Valor	Unidad Ingeniería	Unidad Medida	Unidad Ingeniería
Material Consumido	Insumo Consumido	Insumo químico	Insumo consumido

Tabla 11: Paralelo entre terminologías identificadas en relación al modelo desempeño producción

3.4 Paso 4: Proceso de conceptualización

Con el fin de obtener un modelo ontológico que no sea específico a un proceso de manufactura particular, como es en este caso el de tratamiento del agua potable, para el desarrollo de la etapa de conceptualización se emplearon los modelos conceptuales propuestos en los estándares ISA-88.00.04 (modelo de conjunto de datos, modelo de eventos) e ISA-95.00.02 (modelo de desempeño de la producción).

Estos modelos se han seleccionado con el fin de que, una vez concluida la creación del modelo ontológico, éste sea empleado para la descripción explícita de información asociada a los eventos, históricos del proceso (conjunto de datos) y desempeño de la producción contenida dentro de las bases de datos.

Cabe anotar que debido a que el modelo de desempeño de producción hace referencia a elementos de los modelos de personal, materiales, equipos y de segmento del proceso, propuestos dentro del estándar ISA-95.00.02, dichos modelos también fueron empleados en el paso de conceptualización del modelo ontológico y se encuentran consignados dentro del Anexo D de este trabajo.

Con base en lo anterior, y teniendo en cuenta la clasificación de ontologías de acuerdo con el ámbito de conocimiento al que se aplique, que es una de las más destacadas y que resulta ser la más general, el modelo ontológico que se pretende construir se puede definir como uno de dominio debido a que captura conceptualización aplicable al dominio de la automatización, común a varios procesos industriales, por lo que las descripciones que contiene pueden utilizarse en otros procesos diferentes al empleado en este trabajo (“tratamiento de agua portable”)

A continuación se presentan los modelos conceptuales de los modelos seleccionados (expresados en notación UML). A partir de los modelos se pueden extraer los conceptos (clases), relaciones (con su dominio, rango y cardinalidad) y atributos necesarios para llevar el paso de implementación de una manera mucho más sencilla. Finalmente se muestran algunos ejemplos de las instancias de las clases identificadas.

- **Modelo de desempeño de producción**

La Figura 10 muestra el modelo de desempeño de la producción propuesto dentro del estándar ANSI/ISA-95.00.01.

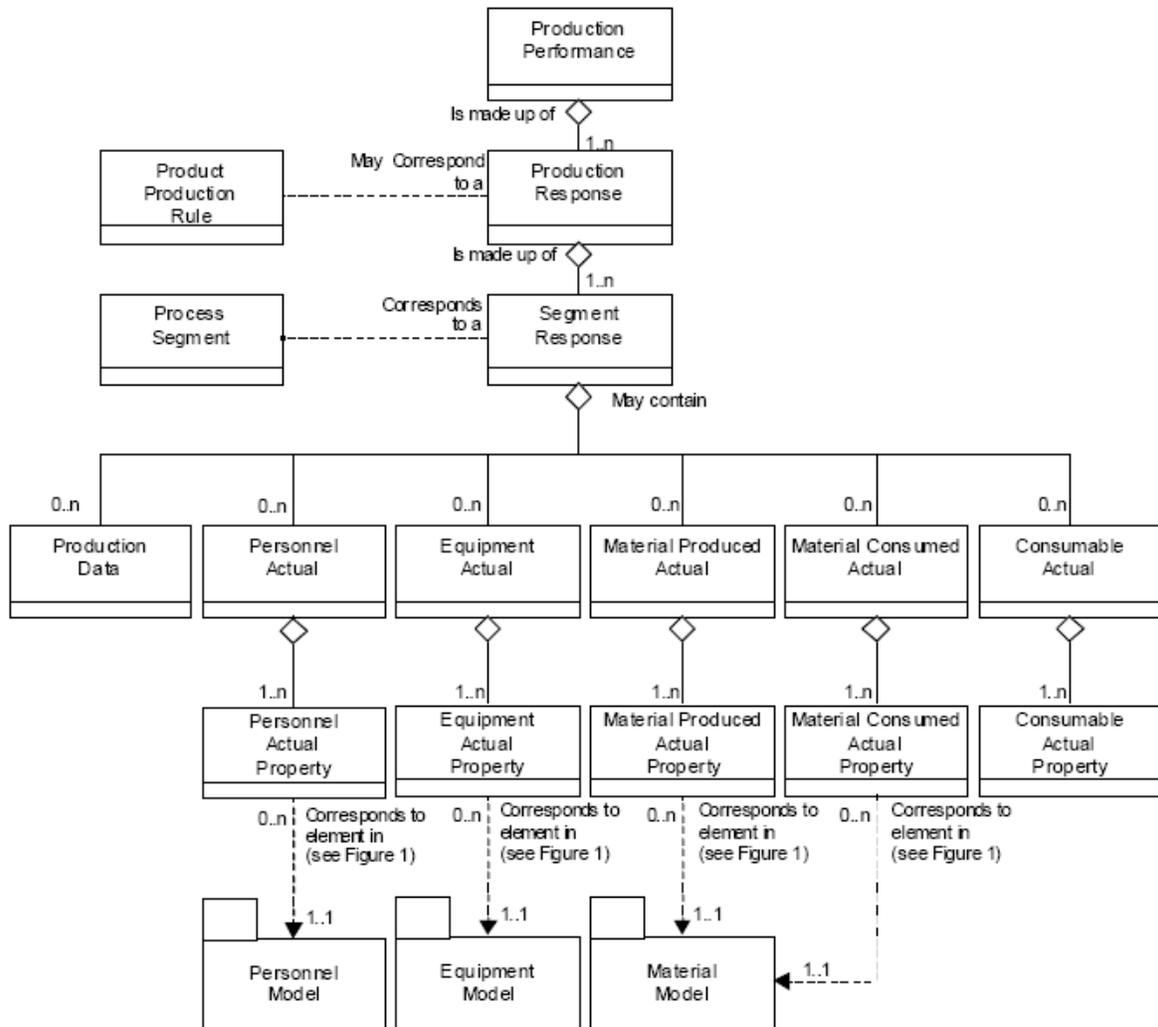


Figura 10. Modelo de desempeño de producción (ANSI/ISA-95.00.02)

- **Identificación de conceptos:** A continuación se procede a describir los conceptos (clases) contenidos dentro del modelo de desempeño de la producción (Tabla 12).

Concepto	Descripción
Desempeño de Producción	Permite determinar el nivel de rendimiento o productividad de una determinada solicitud de manufactura.
Respuesta de Producción	La respuesta de producción contiene información que está relacionada con una solicitud de producción.
Respuesta de segmento	Contiene información referente a la respuesta de producción para un segmento de proceso específico.
Datos de producción	Contiene información que está relacionada con los productos fabricados, tal como el comportamiento de las variables más relevantes del proceso, desviaciones en los procedimientos registrados a través de eventos y alarmas, y el comportamiento de los operarios determinado por sus intervenciones, acciones, y comentarios.
Equipo Real	El equipo real dentro de una respuesta de producción permite identificar la capacidad de equipo empleada durante un segmento de producto específico.
Propiedad de Equipo Real	Define las características para los objetos "equipo real".
Material real producido	Identifica el material producido durante un segmento de producto definido.
Propiedad de material real producido	Define las características para los objetos "material real producido".
Material Consumido	Identifica el material empleado durante un segmento del producto específico.
Propiedad de material consumido	Define las características para los objetos "material real consumido".
Personal Real	Identifica el personal empleado durante un segmento del producto específico
Propiedad de Personal Real	Define las características para los objetos "personal real".

Tabla 12. Conceptos contenidos dentro del modelo de desempeño de producción

- **Identificación de relaciones:** A continuación se procede a describir las relaciones contenidas dentro del modelo de desempeño de la producción (Tabla 13).

Nombre Relación	Dominio	Rango	Cardinalidad
compuesto de	Desempeño de Producción	Respuesta de Producción	1..n
compuesto de	Respuesta de Producción	Respuesta de segmento	1..n
puede contener	Respuesta de segmento	Datos de producción	0..n
puede contener	Respuesta de segmento	Equipo Real	0..n
puede contener	Respuesta de segmento	Material real producido	0..n
puede contener	Respuesta de segmento	Personal Real	0..n
puede contener	Respuesta de segmento	Material consumido	0..n
tiene propiedad	Equipo Real	Propiedad de Equipo Real	1..n
tiene propiedad	Material real producido	Propiedad de material real producido	1..n
tiene propiedad	Material Consumido	Propiedad de material consumido	1..n

Tabla 13. Relaciones contenidas dentro del modelo de desempeño de producción

- **Identificación de atributos:** A continuación se muestran los atributos del concepto “desempeño de la producción” (Tabla 14). Los atributos para los otros conceptos (clases) del modelo de desempeño de la producción se encuentran consignados dentro del Anexo D.

Nombre del Atributo	Descripción	Tipo
ID	Identificador único del reporte de desempeño de la producción, que podría incluir información de versión y revisión	Entero
Descripción	Contiene información adicional del reporte de desempeño de la producción	Cadena
Cronograma de producción	Identificador del cronograma de producción asociado, si aplica. El reporte de desempeño de la producción puede no estar relacionado con un cronograma de producción, puede ser un	Entero

	reporte de toda la producción durante un tiempo específico, o reportado con base en eventos del nivel de planta.	
Fecha de Inicio	La fecha de inicio que abarca el reporte de desempeño de la producción, si aplica.	Fecha
Fecha de Finalización	La fecha de finalización que abarca el reporte de desempeño de la producción, si aplica	Fecha
Fecha de Publicación	La fecha y hora en la cual el reporte de desempeño de la producción fue generado o publicado.	Fecha y Hora
Localización	Identificador del elemento dentro del modelo de jerarquía de equipos asociado al reporte.	Cadena
Tipo de Elemento	Definición del tipo de elemento dentro del modelo de jerarquía de equipos asociado. Por ejemplo: empresa, sede, área.	Cadena

Tabla 14. Atributos del concepto desempeño de producción

- **Modelo de Evento**

La Figura 11 presenta el modelo de eventos propuesto dentro del estándar ANSI/ISA-88.04.

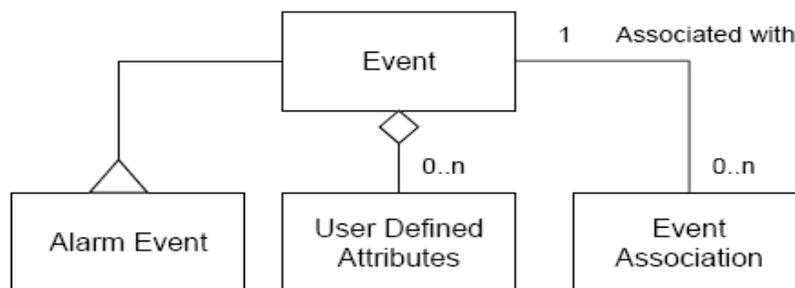


Figura 11. Modelo de eventos (ANSI/ISA-88.04).

- **Identificación de conceptos:** A continuación se procede a describir los conceptos (clases) contenidos dentro del modelo de eventos (Tabla 15).

Concepto	Descripción
Evento	Define un acontecimiento en el tiempo
Evento Alarma	Define un acontecimiento en el tiempo que tiene asociada una presencia real o inminente de una amenaza.
Asociación de Evento	Define una relación entre eventos.
Atributos definidos por el usuario	Elemento usado para incluir información adicional acerca de un evento

Tabla 15. Conceptos contenidos dentro del modelo de eventos

- **Identificación de relaciones:** A continuación se procede a describir las relaciones contenidas dentro del modelo de eventos (Tabla 16).

Nombre Relación	Dominio	Rango	Cardinalidad
tiene atributo usuario	Evento	Atributos definidos por el usuario	0..n
asociado con	Evento	Asociación de Evento	0..n

Tabla 16: Relaciones contenidas dentro del modelo de eventos

- **Identificación de atributos:** A continuación se muestran los atributos del concepto “evento alarma” (Tabla 17). Los atributos para los otros conceptos (clases) del modelo de eventos se encuentran consignados dentro del Anexo D.

Nombre del Atributo	Descripción	Tipo
Evento alarma	Registra los diferentes eventos durante el ciclo de vida de una alarma.	Cadena
Limite Alarma	Valor que al ser excedido por una variable origina la alarma	Real
Tipo de Alarma	El tipo de alarma desde la perspectiva del proceso.	Cadena
Prioridad	Indicación de la importancia del evento alarma.	Cadena

Tabla 17. Atributos del concepto evento alarma

- **Modelo de conjunto de datos**

La Figura 12 presenta el modelo de conjunto de datos propuesto dentro del estándar ANSI/ISA-88.04.

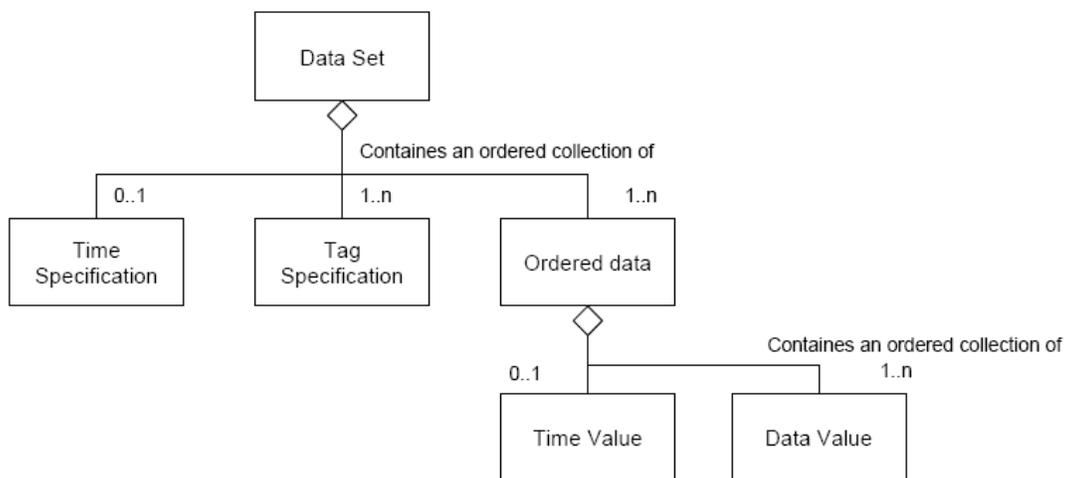


Figura 12. Modelo de conjunto de datos (ANSI/ISA-88.00.04)

- **Identificación de conceptos:** A continuación se procede a describir los conceptos contenidos dentro del modelo conjunto de datos (Tabla 18).

Concepto	Descripción
Conjunto de Dato	Representa un grupo de datos registrados en un intervalo dentro de una serie temporal ⁵ o una agrupación de datos correlacionados ⁶
Especificación de Tiempo	Define un conjunto de atributos que se aplican al elemento tiempo empleado en los objetos de la clase conjunto de datos
Especificación de Etiqueta	Permite identificar de manera única una fuente de datos.
Dato Ordenado	Contenedor de objetos de las clases Valor de Dato y Valor de Tiempo. Para los valores de los datos, por lo que dicha clase no posee atributos.

⁵ Una serie temporal corresponde a un conjunto de datos, donde uno de sus elementos representa el tiempo y los restantes hacen referencia a los valores específicos de las etiquetas en dicho momento.

⁶ Un conjunto de datos correlacionados se caracteriza por la carencia del elemento tiempo, como por ejemplo en una relación presión vs temperatura.

Valor de Dato	Define el valor de una etiqueta.
Valor de Tiempo	Define el tiempo asociado a los valores de los datos (instante de registro).

Tabla 18. Conceptos contenidos dentro del modelo de conjunto de datos

- **Identificación de relaciones:** A continuación se procede a describir las relaciones contenidas dentro del modelo de conjunto de datos (Tabla 19).

Nombre Relación	Dominio	Rango	Cardinalidad
contiene (una colección ordenada de)	Conjunto de Dato	Especificación de Tiempo	0..1
contiene (una colección ordenada de)	Conjunto de Dato	Especificación de Etiqueta	1..n
contiene (una colección ordenada de)	Conjunto de Dato	Datos Ordenado	1..n
contiene (una colección ordenada de)	Datos Ordenado	Valor de Tiempo	0..1
contiene (una colección ordenada de)	Datos Ordenado	Valor de Dato	1..n

Tabla 19. Relaciones contenidas dentro del modelo de conjunto de datos

- **Identificación de atributos:** A continuación se muestran los atributos del concepto “conjunto de dato” (Tabla 20). Los atributos para los otros conceptos (clases) del modelo de conjunto de datos se encuentran consignados dentro del Anexo D.

Nombre del Atributo	Descripción	Tipo
Tiempo de Inicio	Fecha y hora de inicio del intervalo que comprende el conjunto de datos	Fecha y Hora
Tiempo de Finalización	Fecha y hora de finalización del intervalo que comprende el conjunto de datos	Fecha y Hora
Referencia al Sistema de	Empleado para especificar la localización del sistema	Cadena

Tendencia	externo donde se almacenan los datos, si aplica.	
-----------	--	--

Tabla 20. Atributos de la clase conjunto de datos

- **Definición de instancias.**

Finalmente, una vez identificados los conceptos, relaciones, atributos del modelo, a continuación se procede a la labor de definición de instancias. Es en este punto donde se emplea la terminología extraída en el paso 1. A continuación se muestran los ejemplos donde se definen las instancias de las clases: propiedad de material producido (Tabla 21) y material consumido (Tabla 22), definidas dentro del modelo de desempeño de la producción y unas instancias pertenecientes a la clase especificación de etiqueta (Tabla 23) definida dentro del modelo de conjunto de datos. En el Anexo E se encuentran consignadas las instancias para todos los modelos desarrollados.

PROPIEDAD DE MATERIAL PRODUCIDO			
Nombre de la propiedad	Descripción	Valor	Unidad Medición Valor
Color	Color (promedio) en el agua debido a la presencia de metales, materia orgánica u otras sustancias disueltas o suspendidas en ella.	1	Ppm
Turbiedad	Apariencia (promedio) del agua no traslúcida, debida a la presencia de material en suspensión que obstruye el paso de la luz	1	NTU
pH	Define el nivel de acidez promedio del material producido	7.6	-

Tabla 21. Instancias de propiedad de material producido, planta Tulcán

- **Instancias de material consumido**

MATERIAL CONSUMIDO							
Clase de Material	Definición de Material	Lote de Material	Sub-Lote de Material	Descripción	Localización	Cantidad	Unidad de Medición
Coagulante	Sulfato Aluminio	LMSA-51	-	Coagulante empleado para acelerar la precipitación partículas suspendidas y coloides presentes en el agua.	Vertedero	47	Kg
Modificador de pH	Cal	LMCC-35	-	Sustancia empleada para mantener el pH del agua en niveles óptimos (7.2-7.6)	Zona de desinfección	23	Kg
Desinfectante	Cloro	LMCL-40	-	Agente con propiedades germicidas y bactericidas empleado para eliminar organismos patógenos	Zona de desinfección	28	Litros

Tabla 22: Instancias de material consumido, planta Tulcán

- **Instancias de la clase especificación de etiqueta**

ID Fuente Datos	Alias	Descripción	ID Equipo	Elemento Procedimiento	Unidad Medición	Tipo Dato	Dígitos Significativos
CAUDAL_ENT	Caudal de entrada	Caudal de entrada a la planta de tratamiento.	Medidor flujo entrada		Litros/Segundo	Real	R2.2
NIVEL_T1500	Nivel Tanque 2	Nivel en el tanque de almacenamiento de 1500Mts ³ .	Tanque Almacenamiento #2	Almacenamiento	Metros	Real	R2.2
PH_SAL	pH de Salida	Nivel de pH del agua en el agua tratada.	Medidor de pH.	Almacenamiento	-	Real	R2.2
DOSIF_CAL	Dosis de cal	Valor instantáneo de cal dosificada.	Dosificador Cal	Desinfección	Gramos/Segundo	Real	R2.2

Tabla 23: Instancias de la clase especificación de etiqueta, planta Tulcán

3.5 Paso 5: Implementación

La formalización de los modelos conceptuales anteriormente descritos fue llevada a cabo empleando el lenguaje OWL-DL y la herramienta de edición de ontologías Protégé 3.2. Los criterios tenidos en cuenta para la selección del lenguaje de codificación y la herramienta de edición se pueden encontrar en el Anexo F.

A continuación se presentan los diagramas de los modelos formalizados (archivos con extensión *.owl). Para la visualización de los diagramas se ha empleado el plug-in OntoViz incluido dentro de Protégé 3.2. Para que los diagramas sean más fáciles de entender únicamente se muestran los conceptos y relaciones que componen cada uno de los modelos. No obstante cada concepto contiene los atributos (propiedades de tipo de dato) establecidos por el estándar los cuales han sido descritos en el Anexo G.

La Figura 13 muestra el diagrama del modelo de desempeño de producción ISA-95 formalizado a través de OWL (archivo: desepeno_produccion.owl). Cabe anotar que debido a que el modelo de desempeño de producción hace referencia a elementos de los modelos de personal, materiales, equipos y de segmento del proceso, propuestos dentro del mismo estándar, dichos modelos también fueron objeto de formalización y pueden encontrarse en el Anexo G.

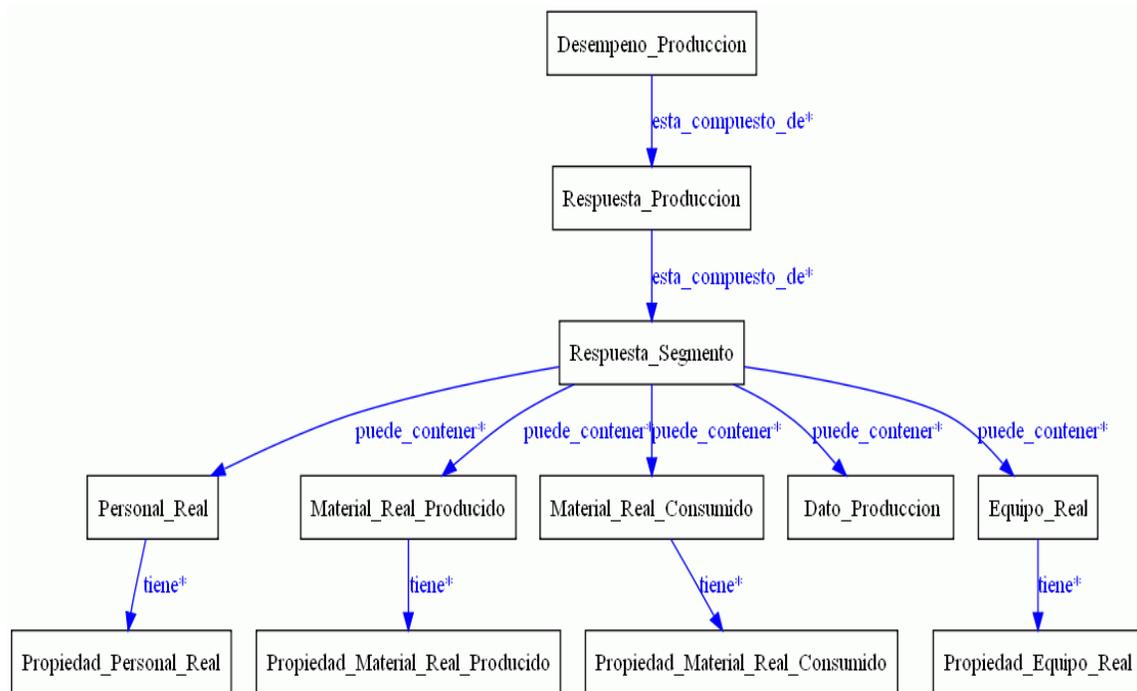


Figura 13. Diagrama del modelo de desempeño de producción ISA-95.00.02 formalizado en OWL

La Figura 14 muestra el diagrama del modelo de conjunto de datos formalizado en OWL (archivo conjunto_datos.owl).

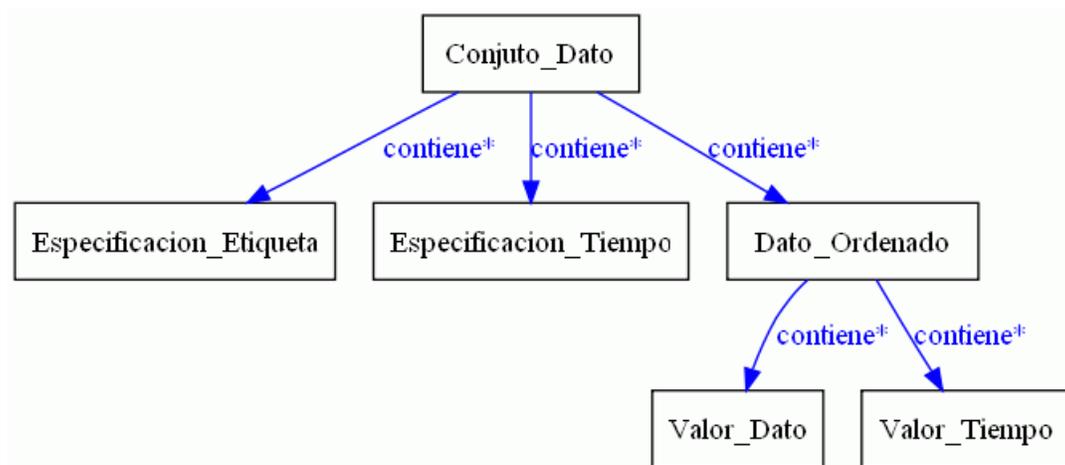


Figura 14. Diagrama del modelo del conjunto de datos ISA-88.00.04 formalizado en OWL

La Figura 15 muestra el diagrama del modelo de conjunto de datos formalizado en OWL (archivo: eventos.owl).

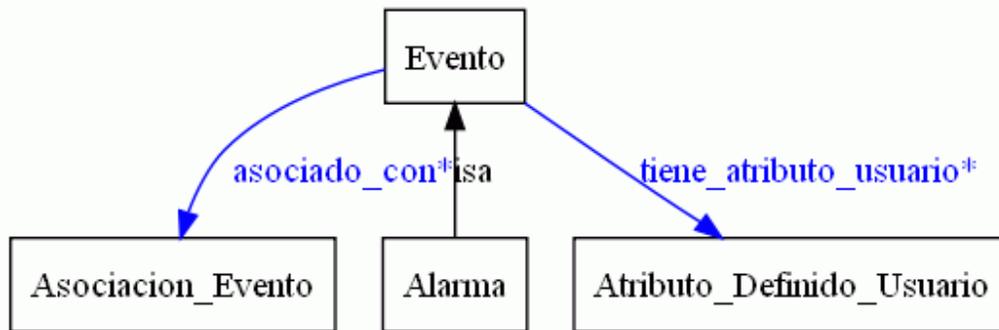


Figura 15: Diagrama del modelo eventos ISA-88.00.04 formalizado en OWL.

El enfoque para el manejo de los diferentes términos asociados a las fuentes de información, usuarios de la información y modelos formalizados (conflicto de sinónimos) encontrados en el caso de estudio (ver Tabla 9, Tabla 10, Tabla 11), consiste en que una vez formalizados en OWL cada uno de los modelos seleccionados de los estándares, entonces dichos modelos son importados dentro de otro modelo que contiene un conjunto de clases y propiedades que representan la terminología empleada por los modelos de bases de datos y los usuarios (operador y encargado producción). Una vez realizado este paso se procede a establecer la equivalencia entre términos a través de la utilización de los constructores `owl:equivalentClass` y `owl:equivalentProperty` que proporciona el lenguaje OWL. La Figura 16 muestra un fragmento de código en lenguaje OWL en el cual se define la equivalencia entre dos clases asociadas a los términos Histórico y Conjunto_Dato.

```
<owl:Class rdf:ID="Historico">
  <owl:equivalentClass rdf:resource="http://www.owl-ontologies.com/Ontology1226769574.owl#Conjunto_Dato"/>
</owl:Class>
```

Figura 16. Ejemplo de definición de equivalencia entre clases en el lenguaje OWL

La Figura 17 muestra un fragmento de código en lenguaje OWL en el cual se define la equivalencia entre dos atributos asociados a los términos Id_Tag e ID_Fuente_Dato.

```
<owl:DatatypeProperty rdf:ID="Id_Tag"
  <owl:equivalentProperty rdf:resource="http://www.owl-ontologies.com/Ontology1226769574.owl#Id_Fuente_Dato"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
```

Figura 17. Ejemplo de definición equivalencia entre propiedades en el lenguaje OWL

El enfoque anteriormente descrito permite que los modelos de los estándares que han sido formalizados puedan emplearse en otro caso de estudio, debido a que únicamente es necesario realizar modificaciones en el modelo que realiza la importación, ya sea agregando o eliminando propiedades y clases, además de establecer las respectivas equivalencias con los elementos definidos dentro del modelo base (modelo formalizado tomando como referencia los estándares internacionales).

De esta forma el modelo ontológico está conformado por los siguientes modelos:

1. Modelo de Conjunto de Datos (conjunto_datos.owl).

2. Modelo de Eventos (eventos.owl).
3. Modelo de Desempeño de Producción (desempeno_produccion.owl).

Debido a que el modelo de desempeño de producción hace referencia a elementos de los modelos de recursos propuestos en el estándar ISA–95.00.02, entonces el modelo ontológico comprende adicionalmente los siguientes modelos:

1. Modelo de Personal (personal.owl).
2. Modelo de Material (material.owl).
3. Modelo de Equipo (equipo.owl).
4. Modelo de Segmento Proceso (segmento_proceso.owl).

3.6 Paso 7: Refinamiento y estructuración de la ontología

Para la realización del refinamiento y estructuración del modelo ontológico creado se hizo uso del prototipo construido a través del cual se pudo probar la funcionalidad brindada por el modelo y con esto verificar si el modelo obtenido cumple con los objetivos planteados al inicio de su construcción.

Por esta razón dentro del capítulo de validación de este trabajo se procede a describir en mayor detalle la realización de este paso.

4 DESCRIPCIÓN DEL PROTOTIPO CREADO

Una vez desarrollado el modelo ontológico del dominio en cuestión, para su validación se procedió a la realización de un prototipo software que interpretara dicho modelo y poder con ello dar respuesta a las consultas por parte de los usuarios definidos. Por esta razón, en el presente capítulo se procede a la descripción del prototipo construido, los actores que se distinguen en el sistema, la arquitectura de implementación de la herramienta y detalles de implementación.

4.1 DESCRIPCIÓN DEL PROTOTIPO DESARROLLADO

El prototipo cuenta con una interfaz que despliega una terminología acorde al rol de cada usuario (operador y encargado de producción), que representa el dominio de conocimiento con respecto al proceso productivo. La terminología se emplea como medio para estructurar las solicitudes de información que se envían al modelo ontológico que establece el dominio común entre la información del proceso y los usuarios de la misma. Gracias al dominio común, el usuario puede obtener información de forma transparente sin importar la terminología empleada en los modelos de bases de datos y la ubicación de la misma, con lo cual se facilita la labor de interpretación de la información presentada.

El desarrollo del prototipo se efectuó bajo la plataforma Microsoft.NET. La selección de la plataforma estuvo orientada básicamente por aspectos de facilidad de aprendizaje y de desarrollo debido a la poca experiencia con alguna plataforma en particular. En este sentido .NET fue escogida frente Java debido a que presenta una curva de aprendizaje más favorable, además de ofrecer programas de desarrollo como Visual Studio.NET, que proponen una interfaz de usuario

mucho más “amigable” que el resto de IDEs (Entorno de Desarrollo Integrado) de desarrollo para otras plataformas [34].

Para el procesamiento del modelo ontológico obtenido se empleó la librería Jena para Java, la cual provee toda la funcionalidad para el manejo de ontologías, permitiendo operaciones de lectura, escritura, consultas e inferencias. Además de esto se empleó el plug-in D2RQ para Jena, a través del cual se establecieron una serie de mapeos entre el modelo ontológico y los modelos de bases de datos que contienen la información del proceso.

Finalmente, para la utilización de las librerías Java anteriormente descritas dentro del Framework de .NET, se utilizó la herramienta IKVM.NET, la cual permitió la conversión de las librerías Java de Jena y D2RQ (archivos *.jar) en DLLs que se emplearon dentro del Framework de .NET.

4.1.1 ACTORES DEL SISTEMA

En el desarrollo del presente trabajo de grado se identificaron los siguientes actores, como esenciales para el funcionamiento y gestión del prototipo.

- **Operador:** Usuario que interactúa con el prototipo a través de la realización de peticiones de información de históricos y eventos del proceso con el fin de lograr un mejor entendimiento de la situación actual del mismo.
- **Encargado de Producción:** Usuario que interactúa con el prototipo a través de la solicitud de información de históricos, eventos del proceso e información relacionada al desempeño de la producción conseguido en las sedes de producción.

- **Administrador:** Usuario encargado de la gestión de cuentas de usuarios (operador, encargado de producción, administrador) de acceso al sistema.

4.1.2 MARCO DE TRABAJO DE JENA

Jena2 constituye un marco de trabajo para Java gratuito y de código abierto desarrollado por los laboratorios HP, el cual cuenta con un conjunto de APIs (*Application Programming Interface*) y herramientas para trabajar con modelos RDF y OWL que permiten realizar operaciones básicas de lectura y escritura, y otras un poco más avanzadas a través del motor de consulta SPARQL (ver Anexo H para una descripción del lenguaje de consultas) y los razonadores basados en reglas con los que cuenta [35].

En la sección de implementación se describirán brevemente cada una las APIs de Jena sobre las cuales se soportó el desarrollo del prototipo.

4.1.3 MODELO CONCEPTUAL DEL PROTOTIPO

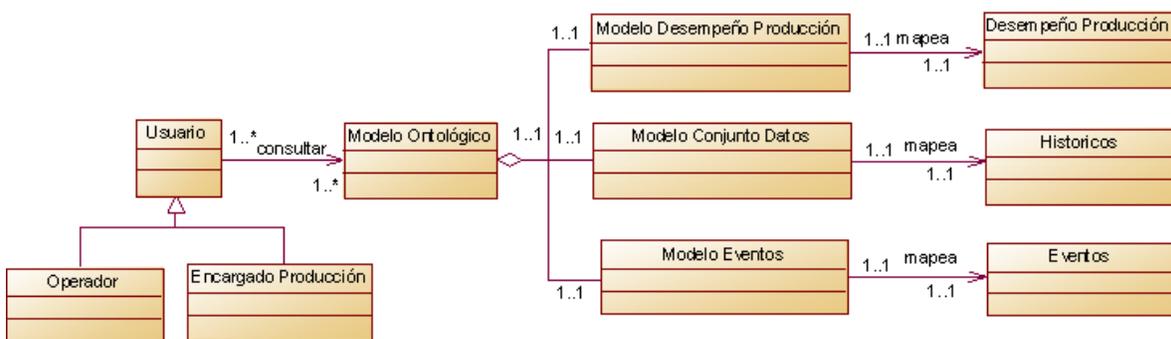


Figura 18. Modelo conceptual del prototipo construido

La Figura 18 presenta el diagrama conceptual del sistema, el cual comprende los siguientes elementos:

- **Usuario:** Representa a los usuarios que harán uso del prototipo, entre ellos los operadores y el encargado de producción.
- **Modelo Ontológico:** Representa el modelo que es consultado por los usuarios. Éste es el modelo que establece el dominio común entre la información generada por las aplicaciones SCADA y los usuarios de la información. El modelo está compuesto por los modelos formalizados de desempeño de producción, eventos y conjunto de datos.
- **Modelo Desempeño Producción:** Representa el modelo formalizado de desempeño de producción obtenido a partir del modelo conceptual del estándar ISA 95.02, terminología de los modelos de bases de datos y terminología de los usuarios. El modelo se mapea al modelo de bases de datos de desempeño de producción.
- **Modelo de Conjunto Datos:** Representa el modelo formalizado de conjunto de datos obtenido a partir del modelo conceptual del estándar ISA 88.04, terminología de los modelos de bases de datos y terminología de los usuarios. El modelo se mapea al modelo de base de datos de históricos.
- **Modelo de Eventos:** Representa el modelo formalizado de eventos basado en el modelo conceptual del estándar ISA 88.04, terminología de los modelos de bases de datos y terminología de los usuarios. El modelo se mapea al modelo de base de datos de eventos.

- **Desempeño Producción:** Representa el modelo de bases de datos de desempeño de producción que almacena la información consolidada del proceso.
- **Históricos:** Representa el modelo de bases de históricos que almacena la información de las variables del proceso.
- **Eventos:** Representa el modelo de bases de eventos que almacena los acontecimientos (eventos y alarmas) generados de la ejecución del proceso.

4.1.4 MODELOS DE BASES DE DATOS

Para el almacenamiento de los datos generados por las herramientas SCADA (iFIX y RSView32) asociadas a cada planta de tratamiento se creó un conjunto de modelos de bases de datos. En este aspecto cabe anotar que actualmente en las plantas de tratamiento no se efectúa un registro de los datos del proceso dentro de bases de datos, por lo cual en la labor de modelado para los propósitos de este trabajo se emplearon los términos manejados por sistema SCADA de la planta del Tablazo, los cuales hacen referencia a históricos de las variables del proceso, eventos y alarmas generados durante la ejecución del mismo.

Para conseguir una estructura adecuada de los modelos se emplearon aspectos del estándar ISA-88.00.04⁷. Éste estándar se seleccionó debido a que proporciona un modelo de referencia para el desarrollo de aplicaciones que almacenen y/o intercambien registros de producción, o que permitan recuperar, analizar y generar reportes de los datos de producción seleccionados [32], lo cual

⁷ Aunque la parte 4 del estándar está enfocado principalmente en procesos por lotes (*batch*), los modelos presentados pueden ser gran valor para otros tipos de procesos [32]

se ajusta claramente a los objetivos pretendidos con el desarrollo del modelo ontológico y del prototipo que lo va a procesar.

Dentro del modelo de registro de producción propuesto en ISA-88.00.04 se incluyen elementos tales como la respuesta de producción, definición de producto y solicitud de producción, los cuales están definidos dentro del estándar ISA-95.00.02. De dichos elementos resulta de gran interés para el desarrollo de este trabajo el de “respuesta de producción”, el cual se encuentra contenido dentro del modelo de desempeño de producción. La escogencia del modelo de desempeño de producción está basada en el hecho de que permite manejar información consolidada de los resultados de la producción obtenidos de la ejecución de uno o más procesos específicos [33].

Sobre los modelos de bases de datos de desempeño de producción desarrollados para las plantas, hay que señalar que éstos no incluyen todos los elementos sugeridos en el estándar. Esto se debe a que por un lado no aplican al caso de estudio o porque los modelos desarrollados están centrados en la información de material producido, material consumido y propiedad de material producido, las cuales pueden obtenerse a partir de la información de históricos del proceso, y además debido a que en el caso de estudio la información asociada a los recursos de equipos y personal, como por ejemplo el tipo, cantidad o propiedades asociadas, están establecidos de antemano.

A continuación se presentan los modelos de bases de datos desarrollados para cada planta. Cabe señalar que en las tablas que componen dichos modelos se han empleado algunos términos distintos a los propuestos dentro del estándar para hacer referencia a los nombres y atributos de las tablas. Esto con el fin de poder tener una heterogeneidad de términos que representen un mismo concepto (conflicto de sinónimos).

La Figura 19 muestra el modelo de bases de datos de históricos de Tablazo. En la Tabla 24 se presenta una descripción de cada una de las tablas que componen el modelo.

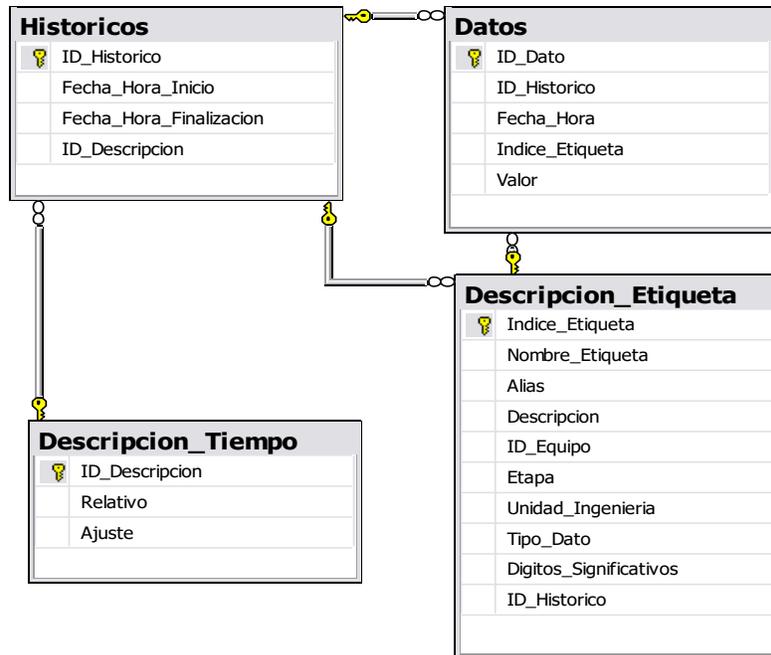


Figura 19: Modelo de bases de datos de históricos de Tablazo

TABLA	FUNCIÓN
Históricos	Contiene la información de un grupo de variables del proceso durante un periodo de tiempo.
Datos	Contiene la información de las variables del proceso en un instante de tiempo.
Descripción Etiqueta	Contiene información para la descripción de las variables del proceso registradas.
Descripción Tiempo	Contiene información para clarificar el atributo tiempo de las variables registradas.

Tabla 24. Descripción de tablas del modelo de históricos de Tablazo

La Figura 20 muestra las tablas que sirven para el registro de los eventos y alarmas ocurridos durante la ejecución del proceso de tratamiento de agua en la planta de tratamiento del Tablazo.

Eventos	
	ID_Entrada
	Fecha_Hora
	Tipo_Evento
	Subtipo_Evento
	ID_Equipo
	Valor
	Valor_Anterior
	Descripcion
	ID_Operador
	Etapa
	Clasificacion

Alarmas	
	ID_Entrada
	Fecha_Hora
	Tipo_Evento
	Subtipo_Evento
	ID_Equipo
	Valor
	Valor_Anterior
	Descripcion
	ID_Operador
	Etapa
	Clasificacion
	Evento_Alarma
	Limite_Alarma
	Tipo_Alarma
	Importancia

Figura 20: Tablas de Eventos y Alarmas de Tablazo

La Figura 21 presenta el modelo de bases de datos de desempeño producción para la planta de tratamiento del Tablazo. En la Tabla 25 se presenta una descripción de cada una de las tablas que componen el modelo.

TABLA	FUNCIÓN
Desempeño Tablazo	Contiene la información que describe el informe de desempeño de la producción.
Respuesta Producción	Contiene la información que identifica la respuesta de producción que conforma el informe de desempeño de la producción.
Respuesta Segmento	Contiene la información que describe el proceso productivo.
Producción Diaria	Contiene la información que define el producto y la cantidad

	fabricada.
Característica Producto	Contiene la información de las características del producto fabricado.
Insumo Consumido	Contiene la información que define el insumo y cantidad consumida dentro del proceso productivo.

Tabla 25. Descripción de tablas del modelo de desempeño producción Tablazo

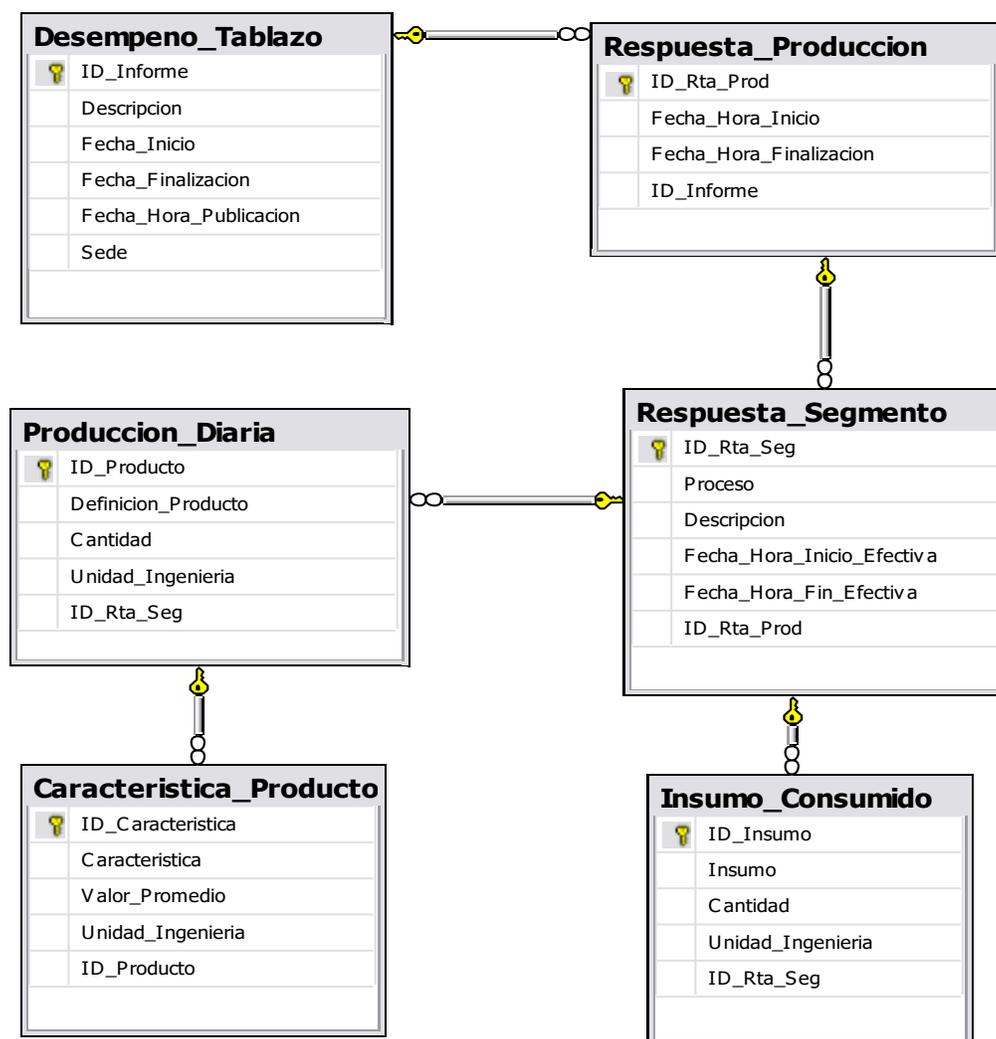


Figura 21. Modelo de la base de datos desempeño producción Tablazo

La Figura 25 muestra el modelo de bases de datos de históricos de Tulcán. En la Tabla 26 se presenta una descripción de cada una de las tablas que componen el modelo.

TABLA	FUNCIÓN
Históricos	Contiene la información de un grupo de variables del proceso durante un periodo de tiempo.
Registros	Contiene la información de las variables del proceso en un instante de tiempo.
Especificación Tag	Contiene información para la especificación de las variables del proceso registradas.
Especificación Tiempo	Contiene información para especificar el atributo tiempo de las variables registradas.

Tabla 26: Descripción de tablas del modelo de históricos de Tulcán

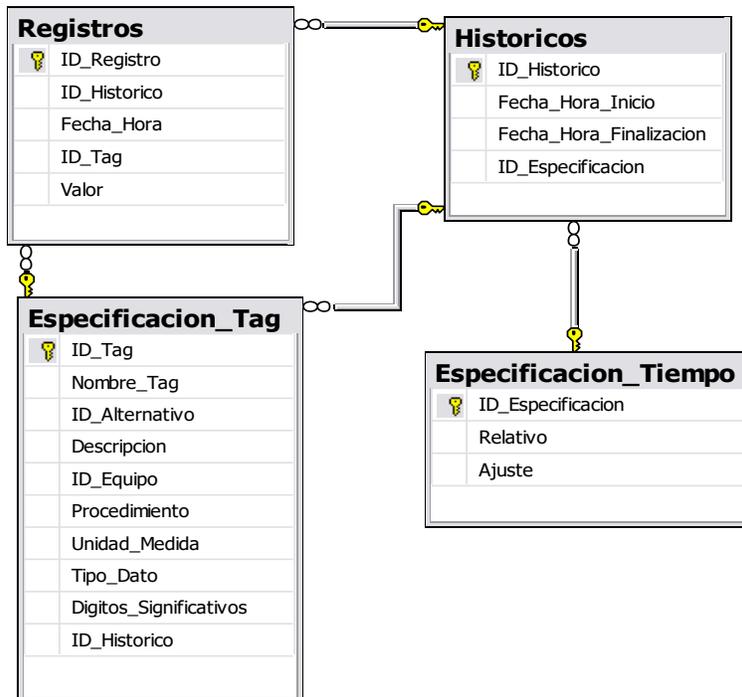


Figura 22. Modelo de bases de datos de históricos de Tulcán

La Figura 23 muestra las tablas que sirven para el registro de los eventos y alarmas ocurridos durante la ejecución del proceso de tratamiento de agua en la planta de tratamiento de Tulcán.

Eventos	
	ID_Evento
	Fecha_Hora
	Tipo_Evento
	Subtipo_Evento
	ID_Equipo
	Valor
	Valor_Pasado
	Descripcion
	ID_Empleado
	Procedimiento
	Categoria

Alarmas	
	ID_Alarma
	Fecha_Hora
	Tipo_Evento
	Subtipo_Evento
	ID_Equipo
	Valor
	Valor_Pasado
	Descripcion
	ID_Empleado
	Procedimiento
	Categoria
	Evento_Alarma
	Limite_Alarma
	Clase_Alarma
	Relevancia

Figura 23. Tablas de Eventos y Alarmas de Tulcán

La Figura 24 presenta el modelo de bases de datos de desempeño producción para la planta de tratamiento de Tulcán. En la Tabla 27 se presenta una descripción de cada una de las tablas que componen el modelo.

TABLA	FUNCIÓN
Desempeño Tulcan	Contiene la información que describe el informe de desempeño de la producción.
Respuesta Producción	Contiene la información que identifica la respuesta de producción que conforma el informe de desempeño de la

	producción.
Respuesta Segmento	Contiene la información que describe el proceso productivo.
Producto	Contiene la información que define el producto y la cantidad fabricada.
Propiedad Producto	Contiene la información de las propiedades del producto fabricado.
Insumo Químico	Contiene la información que define el insumo y cantidad consumida dentro del proceso productivo.

Tabla 27. Descripción de tablas del modelo de desempeño producción Tulcán

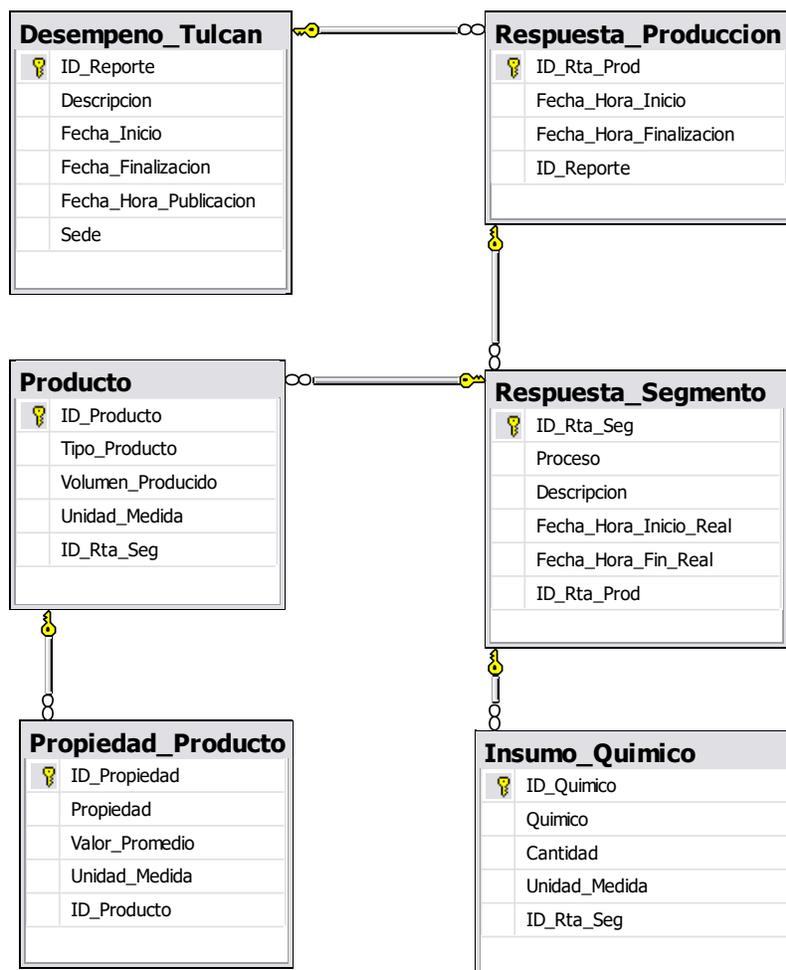


Figura 24. Modelo de la base de datos desempeño producción Tulcán

4.1.5 ARQUITECTURA DEL SISTEMA

La Figura 25 muestra la arquitectura del prototipo desarrollado, el cual está basado en la funcionalidad proporcionada por las APIs de Jena2.

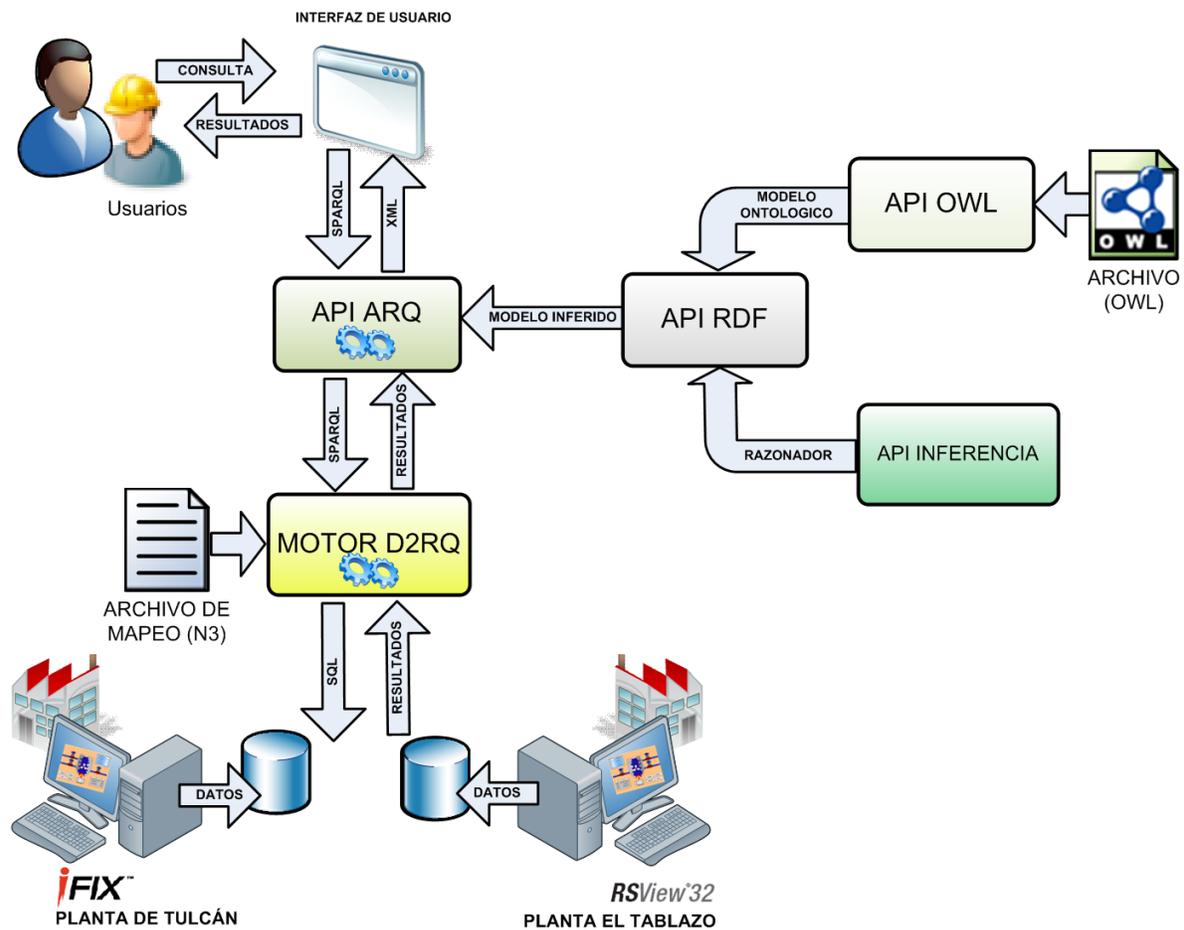


Figura 25. Arquitectura del prototipo desarrollado

A continuación se detallará cada uno de los elementos de la arquitectura del sistema:

- **INTERFAZ DE USUARIO**

La interfaz de usuario está basada en una aplicación Web ASP (Active Server Pages) desarrollada en lenguaje C# empleando la herramienta Microsoft Visual Web Developer 2008 Express. A través de ella el usuario puede autenticarse y navegar hasta la funcionalidad requerida (consulta de históricos, eventos o solicitud de reportes). Una vez en ella el usuario puede conformar las solicitudes de información con base en una serie de criterios de búsqueda presentados en los términos asociados a cada rol (operador o encargado de la producción). Dependiendo de la funcionalidad requerida los resultados son presentados como una tabla a través del uso del control GridView o como reportes empleando el control ReportViewer.

La razón de la escogencia de una aplicación Web frente a una aplicación de escritorio fue influenciada principalmente por el acceso a la información que proporcionan las aplicaciones Web independientemente de la ubicación del usuario, la cual en el caso de una posible implantación permitiría al encargado de la producción consultar la información del proceso ya sea desde las distintas aéreas operativas o administrativas de la empresa.

A continuación se describirán brevemente cada una las APIs de Jena2 sobre las cuales se soportó el desarrollo del prototipo. En ellas se describen cada una de las clases, interfaces y los metodos asociados empleados dentro del prototipo.

- **API RDF**

Es la API principal sobre la cual está fundamentado todo el marco de trabajo de Jena2, permitiendo al usuario la creación y manipulación de grafos RDF. Un grafo RDF es un diagrama de nodos y arcos etiquetados que representan

recursos, sus propiedades, y valores. En Jena2, a un grafo RDF se le denomina modelo y es representado por la interfaz Model [35].

- **API de Ontologías**

La API de Ontologías extiende la API RDF a través de un conjunto de clases e interfaces que permiten el manejo de ontologías. Mediante la API de Ontologías, Jena2 proporciona una interfaz de programación consistente para el desarrollo de aplicaciones, independiente del lenguaje ontológico (OWL, DAML, RDFS) que el usuario emplee [35]. En Jena el manejo de una ontología se hace a través de la interfaz OntModel [35].

- **API ARQ**

ARQ es el motor de consulta de Jena2 que soporta el lenguaje SPARQL desarrollado por el W3C⁸. SPARQL es un lenguaje de consulta definido en términos del modelo de datos de RDF (grafo), el cual puede emplearse para consultar cualquier fuente de datos que pueda mapearse en términos de RDF (como por ejemplo una ontología en lenguaje OWL) [35].

- **API de Inferencia**

La API de Inferencia de Jena2 comprende un conjunto de motores de inferencia (razonadores) basados en reglas que dan soporte a los lenguajes RDFS y OWL [35].

El razonador permite derivar información implícita de un modelo a partir de las instancias y descripciones de clases establecidas. Para poder emplear dicha información, el razonador de Jena2 construye un modelo (modelo inferido) que

⁸ <http://www.w3.org/>

adiciona al modelo de entrada (modelo declarado) la información derivada en forma de tripletas RDF. Una vez obtenido el modelo inferido, entonces se pueden realizar operaciones de consulta empleando los mismos mecanismos que ofrece Jena2 para los modelos a los cuales no se les haya asociado un razonador [35].

En relación al lenguaje OWL, Jena cuenta con tres configuraciones de razonador. Una configuración por defecto (full), que resulta la más completa de todas, y otras dos, un poco más ligeras y por lo tanto mucho más rápidas (mini y micro) [35].

La Tabla 28 muestra los constructores soportados por las distintas configuraciones del razonador OWL de Jena2.

Constructor ⁹	Soportado por
rdfs:subClassOf, rdfs:subPropertyOf, rdf:type	Todos
rdfs:domain, rdfs:range	Todos
owl:intersectionOf	Todos
owl:unionOf	Todos
owl:equivalentClass	Todos
owl:disjointWith	Full, mini
owl:sameAs, owl:differentFrom, owl:distinctMembers	Full, mini
owl:equivalentProperty, owl:inverseOf	Todos
owl:FunctionalProperty, owl:InverseFunctionalProperty	Todos
owl:SymmetricProperty, owl:TransitiveProperty	Todos
owl:someValuesFrom	Full
owl:allValuesFrom	Full, mini
owl:minCardinality, owl:maxCardinality, owl:cardinality	Full
owl:minCardinality, owl:maxCardinality, owl:cardinality	Full
owl:hasValue	Todos

Tabla 28. Constructores soportados por las distintas configuraciones del razonador OWL [35]

⁹ En el Anexo F aparece una descripción de cada uno de los constructores del lenguaje OWL.

De las tres configuraciones del razonador OWL, la empleada dentro del prototipo corresponde a la micro, debido a que cumple con los servicios de inferencia de clasificación (jerarquía inferida de conceptos) y de equivalencia entre clases y propiedades de manera mucho más rápida que las otras dos configuraciones.

- **MOTOR D2RQ**

EL motor D2RQ es un plug-in que se puede emplear dentro de la API de Jena2. El motor hace uso del lenguaje de mapeo D2RQ, el cual se emplea para describir la relación entre una ontología y un modelo de datos relacional. El lenguaje cuenta con un gran conjunto de constructores, entre ellos `d2rq:ClassMap` y `d2rq:PropertyBridge`, para realizar la labor de enlace. `ClassMap` permite realizar el mapeo entre una clase (`owl:Class`) de una ontología y tabla de una base de datos. `PropertyBridge` por su parte permite definir el mapeo entre una propiedad de tipo de datos (`owl:DatatypeProperty`) y una columna de una tabla o de una propiedad que relaciona objetos (`owl:ObjectProperty`) hacia dos columnas de dos tablas diferentes o columnas de una misma tabla que se empleen para almacenar las relaciones entre individuos (instancias) [36].

Mediante la interpretación de las relaciones o mapeos establecidos entre una ontología y una base de datos, el motor d2rq puede reescribir las solicitudes que se realicen sobre la API de Jena2 en consultas SQL contra una base de datos y enviar los resultados hacia las API para su posterior procesamiento [36].

La Figura 26 muestra un ejemplo donde se realiza el mapeo entre la clase “Evento” contenida dentro del modelo ontológico (más específicamente en `eventos.owl`) y la tabla “Eventos” perteneciente al modelo de bases de datos de Tulcán.

```
# INICIO MAPEO CLASE EVENTO & TABLA EVENTOS
map:evento a d2rq:ClassMap;
  d2rq:dataStorage map:database;
  d2rq:uriPattern "Eventos/@@Eventos.ID_Entrada@@";
  d2rq:class reporte:Evento;
.
```

Figura 26. Mapeo entre una clase y una tabla a través del lenguaje D2RQ

La Figura 27 muestra un ejemplo en el cual se efectúa el mapeo entre el atributo (propiedad de tipo de dato) Fecha_Hora asociado a la clase Evento y la columna del mismo nombre de la tabla Eventos.

```
# Columna Fecha_Hora
map:fecha_hora a d2rq:PropertyBridge;
  d2rq:belongsToClassMap map:evento;
  d2rq:property reporte_sin:Fecha_Hora;
  d2rq:property xsd:dateTime;
  d2rq:column "Eventos.Fecha_Hora";
.
```

Figura 27. Mapeo entre un atributo y una columna a través del lenguaje D2RQ

4.2 CASOS DE USO DE ALTO NIVEL

En la Figura 28 se muestran los casos con los cuales todos los actores del sistema (Administrador, Encargado de producción y Operador de planta) pueden interactuar. A través del prototipo, los usuarios pueden iniciar sesión, recuperar contraseña y cambiar contraseña.

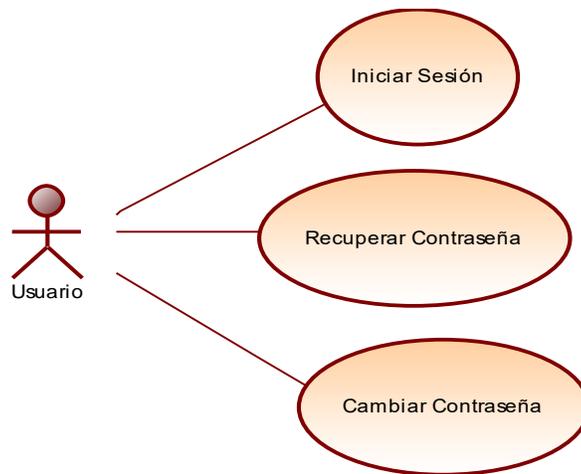


Figura 28. Casos de uso comunes a todos los actores del sistema

La Figura 29 muestra el caso de uso para el administrador del sistema. El administrador se encarga de la gestión de cuentas de usuario, a través de la creación y eliminación de las mismas.

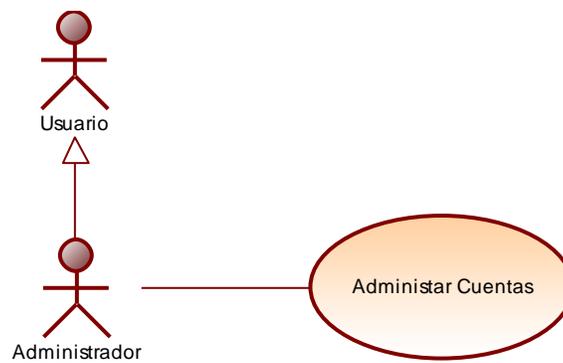


Figura 29. Diagrama de casos de uso para el administrador del sistema

La Figura 30 muestra los casos de uso para el operador de planta, el cual a través del sistema puede consultar los históricos y eventos generados de la ejecución del proceso de manufactura.

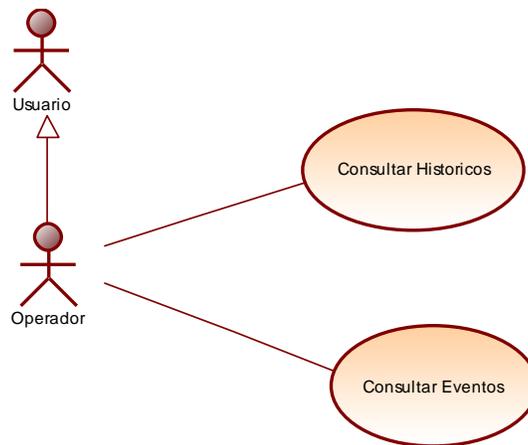


Figura 30. Diagrama de casos de uso para el operador de planta

La Figura 31 muestra los casos de uso para el encargado de producción. El encargado de producción, además de poder consultar la información de históricos y eventos, puede solicitar un reporte del desempeño de producción, así como una versión de simplificada de dicho reporte que le permite conocer las cantidades de insumos consumidos y de producto fabricado.

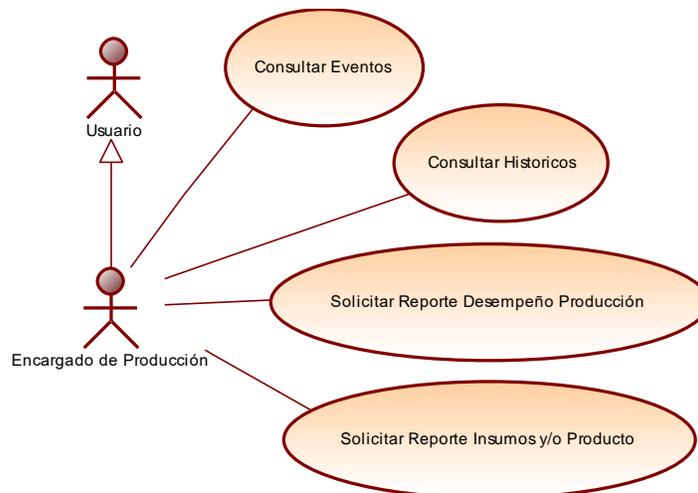


Figura 31. Diagrama de casos de uso para el encargado de producción

4.3 CASOS DE USO REALES

Un caso de uso real describe el diseño concreto del caso de uso. A continuación se presentará el diseño de dos de los casos de uso más importantes que corresponden a: (1) solicitar reporte de desempeño de la producción y (2) solicitar reporte de insumos y/o producto. Estos casos de uso son los más relevantes, debido a que muestran la funcionalidad del sistema para integrar la información de dos fuentes de información, asociadas a las bases de datos de cada planta de tratamiento de agua potable que manejan algunos términos diferentes para hacer referencia a la información que almacenan.

Los restantes casos de uso se encuentran consignados dentro del Anexo I de este trabajo.

4.3.1 Caso de uso real solicitar reporte de desempeño de la producción

CASO DE USO REAL: SOLICITAR REPORTE DE DESEMPEÑO DE LA PRODUCCIÓN	
Actor	Encargado de Producción.
Propósito	Obtener el reporte que resume el desempeño de producción diario para cada una de las plantas de tratamiento de agua potable.
Resumen	El encargado de la producción selecciona la fecha y sede de interés para obtener el reporte de desempeño de producción.

The screenshot shows a web application interface. At the top, there is a header bar with a circled 'A' and the word 'Principal' on the right. Below the header, there are two links: 'Cerrar Sesión' and 'Cuenta'. On the left side, there is a navigation menu with the following items: 'Principal', 'Iniciar Sesión', 'Información de Producción', 'Desempeño de la Producción' (highlighted with a circled 'B'), 'Insumos y/o Producto Fabricado', and 'Información del Proceso'. The main content area displays the text 'Bienvenido a la aplicación Supervisor1'.

C

Principal : Información de Producción : Desempeño de la Producción

[Cerrar Sesión](#) [Cuenta](#)

- Principal
 - Iniciar Sesión
 - Información de Producción
 - Desempeño de la Producción
 - Insumos y/o Producto Fabricado
 - Información del Proceso

Seleccionar Fecha: 2008-11-04 **D**

Seleccionar Sede: Planta el Tablazo **E**

Obtener Reporte **F**

G

[Cerrar Sesión](#) [Cuenta](#)

- Principal
 - Iniciar Sesión
 - Información de Producción
 - Desempeño de la Producción
 - Insumos y/o Producto Fabricado
 - Información del Proceso

REPORTE DE PRODUCCIÓN

Desempeño de la Producción

Descripción: Reporte diario de desempeño de la producción de la planta de tratamiento de agua potable "El tablazo"

Fecha Inicio: 2008-11-03

Fecha Finalización: 2008-11-04

Fecha Publicación: 2008-11-02 14:42:15

Sede: Planta el Tablazo

Proceso Productivo

Proceso: Potabilización de Agua

Descripción: Producción diaria, contiene información del

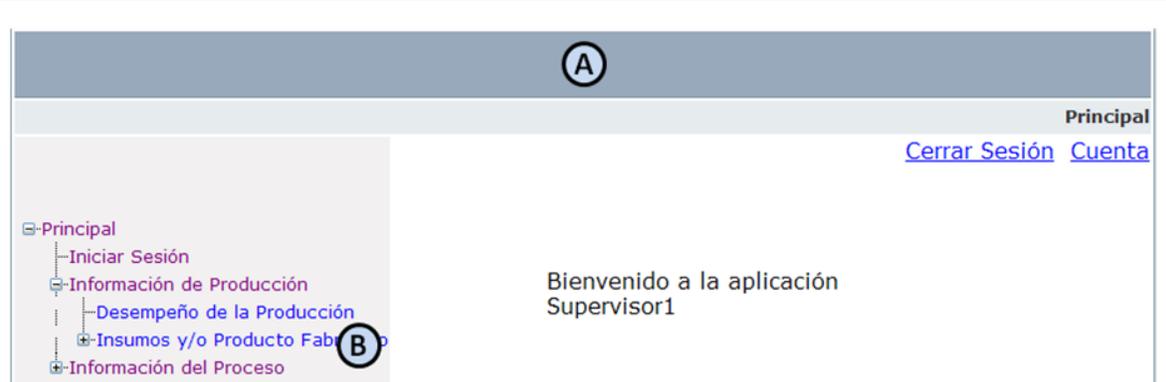
CURSO NORMAL DE LOS EVENTOS

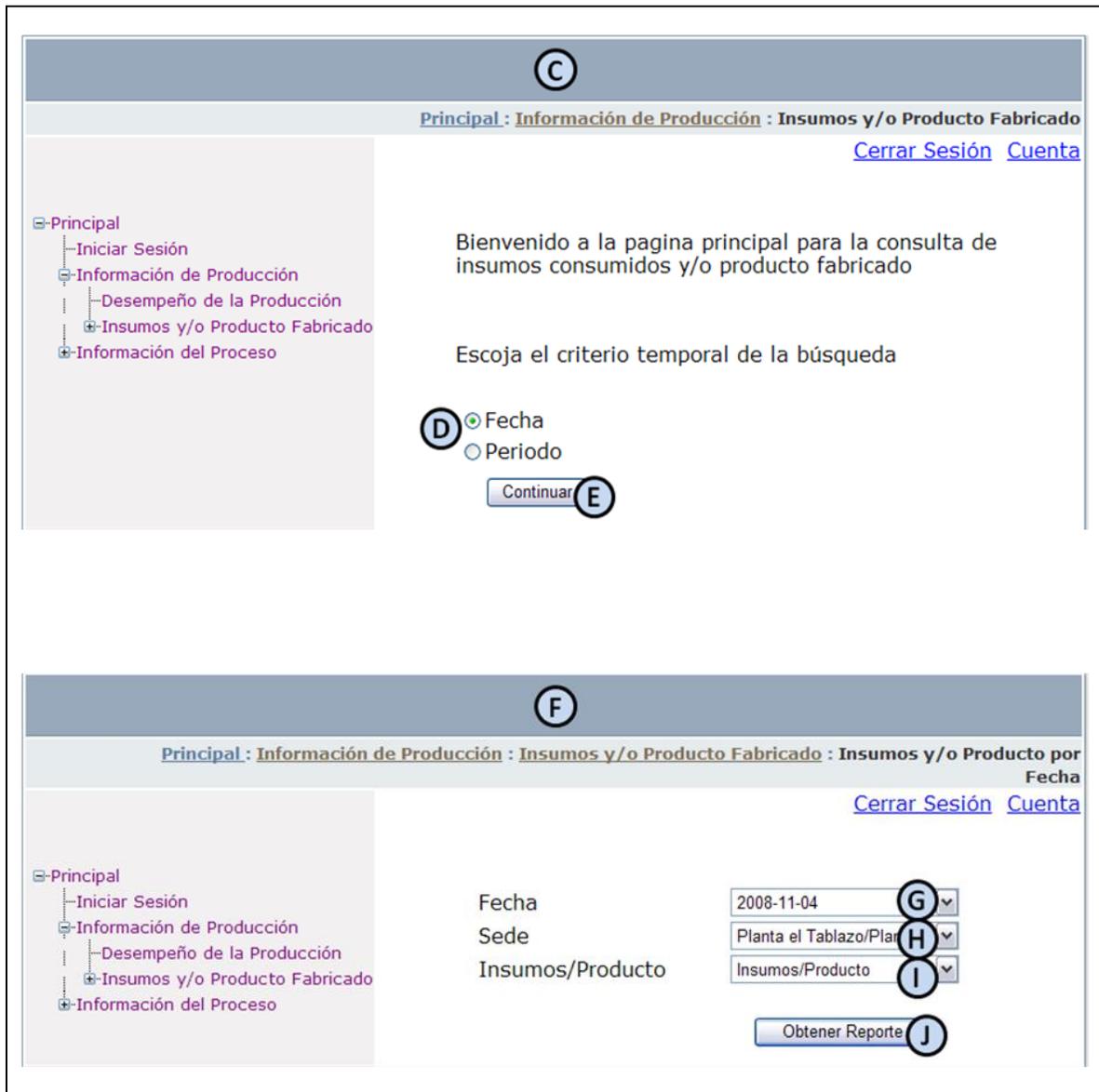
Acción del actor	Respuesta del sistema
1. El encargado de producción después de iniciar sesión se encuentra en la ventana de bienvenida de la aplicación [A].	
2. El encargado de producción da clic en el enlace de "Desempeño de la Producción" [B]	3. El sistema despliega una ventana que presenta los criterios de búsqueda del reporte de desempeño de la

	producción [C].
4. El encargado de producción selecciona una fecha de la lista desplegable [D].	
5. El usuario selecciona alguna de las sedes disponibles [E].	
6. El usuario da clic sobre el botón obtener reporte [F].	7. El sistema despliega una ventana que contiene el reporte de desempeño de producción obtenido teniendo como base los criterios de fecha y sede seleccionados [G].

Tabla 29. Caso de uso real solicitar reporte de desempeño de la producción

4.3.2 Caso de uso real solicitar reporte insumos y/o producto

CASO DE USO REAL: SOLICITAR REPORTE INSUMOS Y/O PRODUCTO	
Actores	Encargado de Producción.
Propósito	Obtener un reporte que contenga información de la cantidad de insumos consumidos y/o producto fabricado para una o varias sedes en una fecha específica o durante un periodo de tiempo.
Resumen	<p>El encargado de la producción solicita el reporte de insumos y/o producto fabricado con base en los siguientes aspectos:</p> <ul style="list-style-type: none"> • Selección de una fecha específica o un periodo de tiempo. • Selección del tipo de recurso de interés: insumos y/o material producido. • Selección de una o más sedes.
	



K

[Cerrar Sesión](#) [Cuenta](#)

Fecha	Sede	Color_Promedio	ph_Promedio	Lts_Cloro	Kg_Sulfato	Kg_Cal	Lts_Potabilizados
2008-11-04	Planta de Tulcán	1.2	7.1	25.3	9.5	10.3	11.002.100
2008-11-04	Planta el Tablazo	1.2	7.2	194.6	72.9	79.3	84.672.000
				219.9	82.4	89.6	95.674.100

L

[Cerrar Sesión](#) [Cuenta](#)

Principal : Información de Producción : Insumos y/o Producto Fabricado : Insumos y/o Producto por Periodo

Principal

- Iniciar Sesión
- Información de Producción
- Desempeño de la Producción
- Insumos y/o Producto Fabricado
- Información del Proceso

Fecha Inicio: M

Fecha Finalización: N

Sede: O

Insumos/Producto: P

Q

(R)								
							Cerrar Sesión Cuenta	
<ul style="list-style-type: none"> Principal - Iniciar Sesión - Información de Producción - Desempeño de la Producción - Insumos y/o Producto Fabricado 	Fecha	Sede	Lts_Potabilizados	Color_Promedio	ph_Promedio	Lts_Cloro	Kg_Sulfato	Kg_Cal
	2008-11-03	Planta el Tablazo	86.400.000	1.3	7.3	198.7	74.4	80.9
	2008-11-03	Planta de Tulcán	10.368.000	1.2	7.2	23.8	8.9	9.7
	2008-11-04	Planta el Tablazo	84.672.000	1.2	7.2	194.6	72.9	79.3
	2008-11-04	Planta de Tulcán	11.002.100	1.2	7.1	25.3	9.5	10.3
	2008-11-05	Planta el Tablazo	81.580.000	1.1	7.5	187.6	70.3	76.4
	2008-11-05	Planta de Tulcán	10.335.000	1.2	7.3	23.7	8.9	9.7
	2008-11-06	Planta el Tablazo	83.376.000	1.4	7.6	191.7	71.8	78.1
	2008-11-06	Planta de Tulcán	12.096.000	1.2	7.4	27.8	10.4	11.3
	2008-11-07	Planta el Tablazo	89.856.000	1.2	7.5	206.5	77.4	84.3
	2008-11-07	Planta de Tulcán	11.664.000	1.2	7.6	26.8	10.1	10.9
	2008-11-08	Planta el Tablazo	85.536.000	1.3	7.4	196.5	73.7	80.6
	2008-11-08	Planta de Tulcán	11.059.200	1.2	7.5	25.4	9.5	10.3
	2008-11-09	Planta el Tablazo	87.696.000	1.0	7.0	201.7	75.6	82.1
	2008-11-09	Planta de Tulcán	9.936.000	1.2	7.4	22.8	8.5	9.3
			675.576.300	1.2	7.3	1552.9	581.9	633.2

CURSO NORMAL DE LOS EVENTOS	
Acción del actor	Respuesta del sistema
1. El encargado de producción después de iniciar sesión se encuentra en la ventana de bienvenida de la aplicación [A].	
2. El encargado de producción da clic en el enlace de "Insumos y/o Producto Fabricado" [B].	3. El sistema despliega una ventana que presenta el criterio de búsqueda temporal del reporte de insumos y/o producto fabricado [C].
4. El encargado de producción selecciona un criterio de búsqueda temporal [D].	
5. El usuario da clic sobre el botón continuar [B].	6. Si el usuario ha seleccionado la opción "Fecha" entonces el sistema despliega la ventana para la solicitud del reporte de insumos y/o producto por fecha [F].
7. El usuario selecciona una fecha [G], una sede específica (Tablazo o Tulcán) o	

	ambas [H], y un ítem de material (insumos y/o material producido) [I].	
8.	El usuario da clic en el botón obtener reporte [J].	9. El sistema despliega una ventana que contiene el reporte con base en los criterios seleccionados [K].
CURSO ALTERNO		
Acción del actor		Respuesta del sistema
		6. Si el usuario ha seleccionado la opción “Periodo”, entonces el sistema despliega la ventana para la solicitud del reporte de insumos y/o producto por periodo [L].
7.	El usuario selecciona una fecha de inicio del periodo [M], una fecha de finalización del periodo [N], una sede específica (Tablazo o Tulcán) o ambas [O], y un ítem de material (insumos y/o material producido) [P].	
8.	El usuario da clic en el botón obtener reporte [Q].	9. El sistema despliega una ventana que contiene el reporte con base en los criterios seleccionados [R].

Tabla 30. Caso de uso real solicitar reporte insumos y/o producto

5 VALIDACIÓN DEL MODELO ONTOLÓGICO CONSTRUIDO

Dentro de este capítulo se procede a describir el proceso empleado para la validación del modelo ontológico desarrollado.

Para el proceso de validación se ha empleado el prototipo creado que proporciona los mecanismos para que el usuario pueda aprovechar la descripción explícita del dominio en cuestión y verificar con ello si el modelo ontológico obtenido cumple con los objetivos planteados al inicio de su construcción.

Cabe aclarar que en este proceso no se realizó una implantación del prototipo dentro del ambiente real, debido a que para este trabajo se consideró un escenario en el cual cada una de las plantas cuenta con un sistema SCADA, las cuales realizan el registro de la información dentro de sistemas de bases de datos. Esto no sucede en la realidad, debido a que sólo la planta de Tablazo cuenta con un sistema de supervisión, el cual no se encuentra funcionando en óptimas condiciones porque los datos que muestra no son en muchos casos los verdaderos y tampoco está configurado para realizar el registro de información en algún medio de forma automática, razón por la cual los operadores tienen que realizar dicho proceso de forma manual y aproximada en hojas de Excel.

Para la realización del proceso de validación se tuvieron en cuenta los siguientes aspectos de evaluación presentados en [31] :

- **Lo que se evalúa:** Es el modelo creado y funcionamiento.

- **Los que lo evalúan:** Los clientes (usuarios que van a usar el modelo ontológico), que determinarán si el modelo ontológico representa fielmente el dominio de conocimiento del caso de estudio específico.

5.1 PROCESO DE VALIDACIÓN

Para la realización del proceso de validación del modelo ontológico se emplearon las simulaciones que recrearon la situación normal de operación de las plantas y se indagó a los usuarios del sistema (operarios y encargado de producción) sobre el funcionamiento del prototipo desarrollado y la utilidad de la información obtenida a través de él.

Los servicios probados por cada uno de los usuarios fueron los siguientes:

- **Usuario1 (operario de planta)**
 1. Solicitar información de los históricos del proceso: Se realizó la búsqueda de la información actual y pasada asociada a las variables del proceso.
 2. Solicitar información de eventos: Se realizó la búsqueda de la información actual y pasada asociada a los acontecimientos generados dentro del proceso.
- **Usuario 2 (Encargado de producción o jefe de planta)**
 1. Solicitar reporte de desempeño de la producción: Se efectuó la petición de reporte de desempeño de la producción reciente y pasado de cada una de las plantas.

2. Solicitar reporte de insumos y/o producto: Se efectuó la petición de reportes acerca de las cantidades de insumos consumidos y producto fabricado en una o ambas plantas.

5.2 RESULTADOS

Después de la realización de las pruebas, los usuarios expresaron su punto de vista a través de una encuesta (ver Anexo J) acerca del funcionamiento del prototipo, llegando al consenso que el prototipo brinda una funcionalidad muy útil, debido a que permite conocer la información (actual y pasada) de cada planta de una forma clara y completa, lo cual en una situación real les permitiría tomar mejores decisiones relacionadas con las operaciones que se deben realizar dentro de las plantas.

A continuación se presentan las opiniones emitidas por parte de los usuarios del prototipo.

Usuario1 (operario de planta): Las opiniones proporcionadas por los operadores luego de probar el funcionamiento del prototipo se pueden interpretar de la siguiente manera:

- Se puede acceder a la información de cada planta en el momento en que se necesite.
- Se puede ver el comportamiento y los valores actuales de cada una de las variables del proceso de forma clara y entendible, lo que facilita determinar las acciones que sea necesario realizar según los datos presentados.

- Se evita el trabajo de estar llenando manualmente un archivo con los valores de las variables. En este proceso se dejaba de registrar por varios momentos el comportamiento de cada variable, por olvido de los operadores.
- Se ve claramente los aspectos que acompañan cada variable medida, como por ejemplo las unidades, ubicación y lo que significa esa medida.
- La búsqueda y presentación de la información se hace empleando palabras claras.

Usuario 2 (Encargado de producción o jefe de planta): El punto de vista del jefe de planta sobre el funcionamiento del prototipo e información presentada por él, es la siguiente:

- Es un buen diseño porque permite obtener información sobre el consumo real de los insumos utilizados en el proceso de potabilización de agua ya sea por día, semana, etc.
- La información de la cantidad de agua producida y de insumos consumidos puede servir de base en otros ámbitos, como por ejemplo el contable.
- Gracias al manejo de información pasada y actual se podría analizar el consumo de insumos en la planta y optimizar el proceso de compra de los mismos.

5.3 CONCLUSIONES DE LA VALIDACIÓN

A partir de los resultados obtenidos se pudo verificar que la funcionalidad proporcionada por el modelo ontológico permite establecer un dominio común entre las herramientas y el lenguaje de los usuarios. Esta afirmación se puede realizar, debido a que usuarios (operadores y jefe de planta) mostraron su satisfacción con los servicios proporcionados por el prototipo, ya que les facilita la obtención de la información requerida de una manera clara, ilustrativa y mucho más oportuna.

A pesar de que el prototipo construido no fue implantado dentro del proceso caso de estudio, a través del uso de simulaciones que recrearon el proceso real, permitió que los usuarios tuvieran la percepción de que la información obtenida era la originada del proceso real, por lo cual permite establecer una tendencia del impacto del proyecto si se implantara.

Finalmente, a partir de los resultados obtenidos se puede concluir que en el caso de una implantación, el prototipo podría traer las siguientes ventajas:

- Libraría a los usuarios de realizar búsquedas manuales (propensas a errores) de información puntual o de información relacionada con el fin de tener una comprensión de una determinada situación que le brinde el soporte adecuado para una toma de decisiones.
- La opción de buscar información pasada, permitiría realizar un seguimiento del proceso productivo, con lo cual se podrían identificar aspectos que afecten el funcionamiento del mismo, tales como fallas o desperfectos en los equipos.

- La información ilustrativa (basada en relaciones entre conceptos) que es presentada por el prototipo permitiría una comprensión rápida de una cierta situación dentro del proceso.
- La información del proceso podría servir como soporte en otro ámbito de la empresa (no contemplado en este trabajo) como es el contable encargado de realizar balances económicos del proceso.

6 CUMPLIMIENTO DE OBJETIVOS

6.1 CUMPLIMIENTO DE OBJETIVOS

6.1.1 Indicadores usados

Con el fin de expresar los resultados finales de cada uno de los objetivos propuestos en este trabajo de grado se presenta a continuación unos tipos de indicadores utilizados en la evaluación de los resultados y la forma correcta de interpretarlos.

Los indicadores de desempeño que evalúan, básicamente adoptan la forma de un cociente, en el cual, el denominador es un valor numérico que ayuda a efectuar la comparación con el logro obtenido así:

$$\text{Indicador} = \left(\frac{\text{Numerador}}{\text{Denominador}} \right) * \text{FactorEscala}$$

De esta forma se definen los siguientes modelos de indicadores que se deben personalizar y aplicar a los actores, productos, funciones, dependiendo del contexto del objetivo evaluado:

1. **Indicador de Cobertura (IC):** Determina la cantidad de elementos cobijados por un producto o estrategia.

$$\text{Cobertura} = \left(\frac{\text{No. de nodos Beneficiados con el Servicio}}{\text{Número de nodos que se esperaba Servir}} \right) * 100$$

2. **Indicador de Eficacia (IE):** Cumplir con los Requisitos definidos

$$Eficacia = \left(\frac{Recursos Ejercidos}{Recursos Asignados} \right) * 100$$

1. **Indicador de Eficiencia (IF):** Permite identificar la relación que existe entre las metas alcanzadas, tiempo y recursos consumidos con respecto a un estándar. Buen uso de los recursos.

$$Eficiencia = \left(\frac{Metas Alcanzadas}{Recursos Consumidos} \right) * 100$$

2. **Indicador de Calidad (IQ):** Están orientados a medir la satisfacción de los beneficiarios.

Indicador = Calificación entre (1: Mala (0%), 2: Regular (50%), 3: Buena (75%), 4: Excelente (100%))

Con el modelo de indicadores presentado anteriormente, en el siguiente apartado se desarrollaron un conjunto de indicadores adecuados que permiten evaluar adecuadamente el nivel de cumplimiento de cada uno de los objetivos.

6.1.2 Objetivos y sus indicadores de cumplimiento

En las siguientes tablas se especifican de arriba hacia abajo: los objetivos comprometidos en el proyecto, los productos esperados derivados de cada objetivo, los resultados obtenidos, los indicadores que evalúan el objetivo, los medios de verificación de los resultados y finalmente unas observaciones que permiten aclarar los resultados en cada objetivo.

Se debe tener en cuenta que se desarrolla una tabla por cada objetivo específico comprometido en la propuesta del proyecto “Modelo Ontológico para la integración de información entre dos aplicaciones industriales”.

No. De Objetivo:	1
Descripción del Objetivo:	Obtener un modelo ontológico que se adapte a las fuentes de información que se pretenden describir formalmente entre las aplicaciones iFIX™ y RSView@32™, con los requerimientos de información por parte del cliente interno.
Productos Esperados:	1. Modelo ontológico de integración de los dominios del conocimiento entre las aplicaciones iFIX™ y RSView@32™, con el dominio del conocimiento de las personas que toman las decisiones en las organizaciones productivas manufactureras.
Resultados Obtenidos:	1. Modelo ontológico de integración de los dominios del conocimiento entre las aplicaciones iFIX™ y RSView@32™, con el dominio del conocimiento de las personas que toman las decisiones dentro de un caso de estudio. 2. Desarrollo de una metodología ONTOPRIN para la creación de ontologías en el ámbito de la automatización industrial.
Indicadores (Escala *100):	<p>Eficacia</p> $IE1 = \frac{No\ Productos\ Obtenidos}{No\ Productos\ A\ Obtener} = 2/1 = 200\%$ <p>Calidad</p> <p>IQ1 = ¿La metodología ONTOPRIN facilitó el proceso de desarrollo de la ontología que permitiera unificar el dominio de conocimiento entre las herramientas SCADA seleccionadas y los usuarios del caso particular escogido= 4 = 100%?</p>

	<p>IQ2 = ¿La ontología creada permitió unificar el dominio de conocimiento entre las herramientas SCADA seleccionadas y los usuarios del caso particular escogido? = 4 = 100%.</p>
<p>Medios de Verificación:</p>	<ol style="list-style-type: none"> 1. En el capítulo 3 de este documento se encuentra la definición de la metodología ONTOPRIN. 2. En el capítulo 4 de este documento se encuentra la descripción del modelo ontológico construido empleando la metodología ONTOPRIN.
<p>Estrategias y/o Observaciones:</p>	<p>Estrategias</p> <ol style="list-style-type: none"> 1 Metodología ONTOPRIN: Debido a que el objetivo principal de este proyecto es la obtención de un modelo ontológico, su construcción debía estar basada en algún tipo de metodología. Para lograr esto inicialmente se realizó una exhaustiva revisión de la literatura acerca de las metodologías existentes. Sin embargo, debido a la poca información que guiara de manera efectiva cada uno de los pasos de desarrollo y teniendo en cuenta que en la mayoría de los casos cada metodología estaba enfocada a un problema particular, se vio la oportunidad de proponer una metodología que pueda emplearse en otros proyectos dentro del ámbito industrial. 2 Selección del formalismo de representación: Se realizó un estudio de los principales formalismos de representación de conocimiento, del cual se encontró que la Lógica Descriptiva (DL) era el formalismo que contaba con algoritmos de inferencia que garantizaban siempre una respuesta en un tiempo finito. 3 Selección del lenguaje de implementación: Se seleccionó el lenguaje OWL-DL, el cual ha surgido como la evolución de una serie de lenguajes para la representación de ontologías (OIL, DAML+OIL) que han tenido su fundamento en la Lógica Descriptiva. Debido a

	<p>esto, actualmente OWL es el principal lenguaje para la representación de ontologías, razón por la cual es soportado por un gran número de herramientas de edición y APIs de desarrollo.</p> <p>4. Selección del editor de ontologías: Se realizó una revisión de la literatura con el fin de definir una serie de criterios de comparación y para identificar las principales herramientas de edición de ontologías.</p> <p>Una vez realizado esto, se procedió a determinar los criterios para las herramientas evaluadas (Protégé-OWL, SWOOP, Altova Semantic-Works), los cuales permitieron fundamentar la selección de la herramienta Protégé para la implementación de la ontología en el lenguaje OWL.</p> <p>Observación</p> <p>1. La conceptualización del modelo ontológico al estar realizada tomando como referencia los modelos de conjunto de datos, eventos y desempeño de la producción propuestos en los estándares ISA-88.04 e ISA-95.02, hace que el modelo ontológico obtenido pueda aplicarse en otro proceso caso de estudio, siempre y cuando sólo se requiera la descripción explícita de la información asociada a los eventos, históricos de proceso y desempeño de la producción que fueron las categorías de información consideradas para la realización del modelo creado.</p>
--	--

No. De Objetivo:	2
Descripción del Objetivo:	Construir un prototipo básico software que permita validar el modelo ontológico desarrollado, en materia de mejoramiento de procesamiento y refinamiento de la información que se le suministra al cliente interno para que realice sus diferentes tareas de toma de decisiones.
Productos Esperados:	1 Prototipo software basado en el enfoque ontológico para la integración de la información generada por las aplicaciones de control y supervisión realizadas en iFIX™ y RSView@32™.
Resultados Obtenidos:	1 Prototipo software basado en el enfoque ontológico para la integración de la información generada por las aplicaciones de control y supervisión realizadas en iFIX™ y RSView@32™.
Indicadores (Escala *100):	<p>Eficacia</p> $IE1 = \frac{No\ Productos\ Obtenidos}{No\ Productos\ A\ Obtener} = 1/1 = 100\%$ <p>Calidad</p> <p>IQ1 = ¿La información presentada a los usuarios por el prototipo sirve de soporte para el desarrollo de toma de decisiones dentro del proceso? =4 = 100%.</p>
Medios de Verificación:	<ol style="list-style-type: none"> 1. En el capítulo 4 de este documento se encuentra la descripción del prototipo básico software que permite validar el modelo ontológico desarrollado. 2. En el capítulo 5 de este documento se encuentra la descripción del proceso de validación del modelo ontológico.
Estrategias y/o Observaciones:	<ol style="list-style-type: none"> 1. Selección de la plataforma de desarrollo: La escogencia de la plataforma sobre la cual fue desarrollado el prototipo fue realizada considerando aspectos de facilidad de aprendizaje y de desarrollo. En este sentido la plataforma .NET fue seleccionada frente Java debido a que presenta un curva de aprendizaje más favorable, además de ofrecer programas de desarrollo como Visual Studio.NET que proponen una

	<p>interfaz de usuario mucho más amigable que el resto de los IDE's de desarrollo para la plataforma Java.</p> <p>2. Selección de la API de desarrollo: Para la utilización del modelo ontológico creado dentro del prototipo fue necesario encontrar algún tipo de herramienta que permitiera procesar las descripciones realizadas dentro del modelo, con el fin de realizar tareas de inferencia y recuperación de información del mismo. Para esto inicialmente se realizó un estudio de las APIs (Application Programming Interface) para el procesamiento de ontologías más referenciadas dentro de la literatura (Jena2, OWL API y KAON2), el cual permitió identificar una serie de criterios de comparación útiles para el desarrollo del prototipo.</p> <p>Posteriormente se efectuó un conjunto de pruebas sencillas para observar el soporte brindado en materia de consulta de información y servicios de inferencia.</p> <p>Una vez analizados los criterios de comparación y los resultados de las pruebas, fue seleccionada la librería Jena2 debido a que brinda un soporte más completo en materia de consulta de información y de inferencia. Además de los aspectos anteriormente mencionados, un punto a favor de Jena2 fue la disponibilidad del plug-in (D2RQ) a través del cual se realizó el mapeo entre la ontología y las bases de datos que contenían la información del proceso enviada por las herramientas SCADA.</p> <p>3. Mecanismo de Interoperabilidad entre Java y .NET:</p> <p>Una vez seleccionada la API para el procesamiento de la ontología, fue necesario seleccionar un mecanismo que permitiera su uso dentro del entorno de .NET. Esto debido a que Jena2 es una API para Java. Para lograr</p>
--	---

	este propósito se empleó la herramienta IKVM, con la cual se realizó el proceso de conversión de los archivos JAR que componen el API de Jena2 en una librería (DLL) utilizable dentro del Framework de .NET.
--	---

No. De Objetivo:	3
Descripción del Objetivo:	Obtener un documento en el cual se consignen los diferentes pasos que se debieron realizar para la obtención del modelo ontológico básico aplicado al dominio de interés.
Productos Esperados:	1 Monografía que contenga toda la información debida al desarrollo del proyecto.
Resultados Obtenidos:	1 Monografía que contenga toda la información debida al desarrollo del proyecto 2 Realización del artículo "Metodología ONTOPRIN para la creación de ontologías en el ambiente industrial"
Indicadores (Escala *100):	Eficacia $IE1 = \frac{No\ Productos\ Obtenidos}{No\ Productos\ A\ Obtener} = 2/1 = 200\%$
Medios de Verificación:	1. Monografía de trabajo de grado
Estrategias y/o Observaciones:	Observación 1. Actualmente se encuentra en proceso de elaboración un segundo artículo, en el cual se consignarán los resultados obtenidos del proceso de creación del modelo ontológico mediante el seguimiento de los pasos de la Metodología ONTOPRIN y del prototipo creado para la validación del modelo.

7 CONCLUSIONES Y TRABAJO FUTURO

7.1 CONCLUSIONES

- La construcción de un modelo ontológico en el dominio de los procesos industriales, como el desarrollado en este trabajo, demostró ser una solución eficiente para establecer un dominio de conocimiento común y compartido, permitiendo inter-operar a los clientes internos con la información generada por las herramientas SCADA en sus propios términos.
- La utilización de los estándares internacionales ISA-88 e ISA-95 facilitó y al mismo tiempo proporcionó una base sólida para la etapa de conceptualización del modelo ontológico. Esto debido a que dichos estándares cuentan con modelos conceptuales y terminología para describir dichos modelos, los cuales han sido el resultado de un consenso de un gran grupo de expertos. Por esta razón, el modelo obtenido cumple con una de las características de una ontología, la cual es el de capturar conocimiento consensuado.
- El modelo ontológico creado no restringe su aplicación al ámbito de manufactura empleado en este trabajo, ni a las herramientas seleccionadas. Esto se debe a que el modelo cuenta con un componente basado en la conceptualización obtenida de los estándares industriales y otro componente basado en la conceptualización encontrada en el caso de estudio. Por esta razón, el modelo es altamente reutilizable, bastando con modificar el componente asociado al caso de estudio siempre y cuando sólo

se requiera la descripción explícita de información asociada a los eventos, históricos de proceso y desempeño de la producción que fueron los considerados para la creación del modelo.

- El enfoque modular seguido para la creación del modelo ontológico permite que éste pueda extenderse de una manera fácil para comprender otros modelos contenidos dentro de los estándares ISA-88 e ISA-95, dependiendo de los requerimientos de información por parte de los usuarios.
- El modelo ontológico creado puede aplicarse a otros casos de estudio en los que se requiera la descripción de información asociada a los eventos, históricos de proceso y desempeño de la producción sin la necesidad de realizar modificaciones en los modelos de bases de datos debido a que el modelo ontológico puede verse como una capa adicional que resuelve los conflictos semánticos entre las fuentes de información y los usuarios de dicha información.
- Se desarrolló una nueva metodología para la creación de ontologías en el campo de la automatización industrial denominada ONTOPRIN, la cual rescató los principales aportes de las metodologías existentes de creación de ontologías en otros ámbitos particulares y detalló con precisión los pasos necesarios para implementar ontologías específicas en el campo de la automatización industrial.
- A través del prototipo construido se pudo validar el modelo ontológico ya que se pudo establecer un medio de comunicación e interacción entre la información generada por las herramientas SCADA iFIX, RSview32 y sus

usuarios, para un caso de estudio particular como las Plantas de procesamiento de agua potable en la ciudad de Popayán.

- La utilización de herramientas de interoperabilidad entre JAVA y .NET como IKVM.NET facilitan la creación de aplicaciones basadas en ontologías dentro del ambiente de .NET. Plataforma sobre la cual hay pocos desarrollos en materia de librerías para el procesamiento de ontologías, en comparación a los disponibles para JAVA, la cual dispone de un gran número de APIs como es el caso de Jena2 que permiten programar, consultar y realizar inferencias sobre una ontología.
- Se elaboró un documento en el que se describe en detalle todo el proceso que se siguió para desarrollar un modelo ontológico en el campo de la automatización industrial, teniendo como gran logro la creación de una metodología llamada ONTOPRIN. Este documento a su vez sirve como marco de referencia para futuros trabajos que relacionen la creación de ontologías en el campo de los procesos de automatización industrial. Además, se elaboró también un artículo en el cual se consignan y describen los pasos de la metodología ONTOPRIN, con el objetivo de difundirla y de presentarla como una herramienta de ayuda para creación de ontologías en el campo de la automatización industrial.
- A partir de la validación de funcionamiento del prototipo se puede ver una tendencia de que, en caso de una posible implantación, el prototipo podría facilitar el desarrollo de las operaciones cotidianas de los usuarios.

7.2 TRABAJO FUTURO

- Emplear la metodología ONTOPRIN para el desarrollo de una ontología aplicada a otro caso de estudio. Para esto se debe considerar que el modelo ontológico creado en este trabajo puede re-utilizarse siempre y cuando requiera la descripción explícita de la información asociada a eventos, históricos de proceso y desempeño de la producción que fue la considerada para la creación del modelo.
- Extender el modelo ontológico construido a través de la formalización de otros modelos conceptuales propuestos dentro de los estándares ISA-88 e ISA-95 dependiendo de los requerimientos de descripción de información que se quieran obtener.
- Emplear el modelo ontológico que contiene la descripción explícita de algunos modelos del estándar ISA-95 (modelo de material, modelo de personal, modelo de equipos, modelo de segmento de proceso y modelo de desempeño de la producción), para la integración de información entre sistemas de control y de empresa (por ejemplo de sistemas de Planificación de Recursos Empresariales, ERP).
- Emplear el vocabulario común que proporciona el modelo ontológico creado como base de conocimiento y como medio de comunicación de un sistema multi-agente empleado para la distribución de tareas dentro de un ambiente industrial.
- Implantar el prototipo creado y validarlo en un entorno real.

REFERENCIAS BIBLIOGRÁFICAS

- [1]. A. Pakonen, "*Information Agent Technology in Process Automation System*," Tesis de Maestría, Helsinki University of Technology, Department of Automation and Systems Technology, 2004. Disponible en: http://www.automation.hut.fi/projects/proage/pubs/Thesis_Pakonen.pdf. Consultado: Septiembre 21 de 2007.
- [2]. T. Tommila, J. Hirvonen, L. Jaakkola y J. Peltoniemi, "*Next Generation of Industrial Automation: Concepts and architecture of a component-based control system*". VTT Research Notes 2303, 2005. Disponible en: <http://www.vtt.fi/inf/pdf/tiedotteet/2005/T2303.pdf>. Consultado: Septiembre 25 de 2007.
- [3]. J. Janne, "*Agent-Based Approach to Supervisory Information Services in Process Automation*," Tesis de Maestría, Helsinki University of Technology, Department of Automation and Systems Technology, 2006. Disponible en: http://www.automation.hut.fi/projects/proage/pubs/Thesis_Jussila.pdf. Consultado: Febrero 26 de 2008.
- [4]. T. Pirttioja, "*Agent-Augmented Process Automation System*," Tesis de Maestría, Helsinki University of Technology, Department of Electrical and Communications Engineering, 2004. Disponible en: <http://www.automationit.hut.fi/file.php?id=316>. Consultado: Marzo 20 de 2008.
- [5]. P. Ziegler y K. Dittrich, "Data Integration: Problems, Approaches, and Perspectives," en *Conceptual Modelling in Information Systems Engineering*, pp 39-58. Berlin: Springer, 2007.

- [6]. H. Stuckenschmidt y F. Van Harmelen, "Semantic Integration," en *Information Sharing on the Semantic Web, Advanced Information and Knowledge Processing*, ch. 1, pp. 3–23, Berlin: Springer, 2005.
- [7]. F. Hakimpour, A. Geppert, "Resolution of Semantic Heterogeneity in Database Schema Integration Using Formal Ontologies," *Information Technology and Management*, vol. 6, pp. 97-122, 2005.
- [8]. A. Poggi, D. Lembo, D. Calvanese, G. D. Giacomo, M. Lenzerini, y R. Rosati, "Linking Data to Ontologies," *Journal on Data Semantics X*, vol. 4900, pp. 133–173, 2008.
- [9]. F. Ruiz y J. R. Hilera, "Using ontologies in software engineering and technology," en *Ontologies for Software Engineering and Software Technology*, ch. 2, pp. 49–102, Berlin: Springer, 2006.
- [10]. Sistemas Scada. Disponible en:
<http://www.automatas.org/redes/scadas.htm>. Consultado: Marzo 15 de 2007.
- [11]. IFIX Technical Data Sheet, GE FANUC International, Inc, 2003. Disponible en: http://www.esac.com/Downloads/iFix_Tech_data_sheet.pdf. Consultado: Septiembre 7 de 2007.
- [12]. D. M. Sánchez, J. M. Cavero, E. M. Martínez, "The Road Toward Ontologies," en *Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems*, ch. 1, pp. 3–20, Springer US, 2007.
- [13]. J. Breis, "Un Entorno de Integración de Ontologías para el Desarrollo de Sistemas de Gestión de Conocimiento," PhD thesis, Universidad de Murcia,

- Departamento de Ingeniería de la Información y las Comunicaciones, 2003. Disponible en: http://www.tdr.cesca.es/TESIS_UM/AVAILABLE/TDR-1123105-134449//FernandezBreis.pdf. Consultado: Octubre 20 de 2006.
- [14]. D. Fensel, *Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce*. Berlin: Springer, 2nd ed., 2004.
- [15]. S. Grimm, P. Hitzler y A. Abecker, "Knowledge Representation and Ontologies: Logic, Ontologies and SemanticWeb Languages," en *Semantic Web Services: Concepts, Technologies, and Applications*, ch. 3, pp. 51–105, Berlin: Springer, 2007.
- [16]. D. Oberle, "Ontologies," en *Semantic Management of Middleware*, ch. 3, pp. 33–53, Spriner US, 2006.
- [17]. M. Pereira, Módulo 2f – metodologías para Criação de Ontologias. Disponible en: <http://www.inf.puc-rio.br/~casanova/INF2328-Topicos-WebBD/modulo2-Ontologias/modulo2f-ontologias-metodologias.PDF>
Consultado: Mayo 20 de 2008.
- [18]. V. Vanesa, I. Lee, y B. Hamidzadeh. "Towards Ontological Modelling of Historical Documents" ca 2003. Disponible en: <http://protege.stanford.edu/conference/2004/abstracts/Mirzaee.pdf>
Consultado: Septiembre 20 de 2007
- [19]. C. Partridge, "*The CEO Project: An Introduction. Technical Report 07/02, LADSEB-CNR*", Italia, 2002. Disponible en: http://www.boroprogram.org/bp_pipex/ladsebreports/ladseb_t_r_07-02.pdf
Consultado: Septiembre 20 de 2007.

- [20]. Ontologías. Disponible en:
<http://www.hipertexto.info/documentos/ontologias.htm>. Consultado: Julio 02 de 2007.
- [21]. O. Corcho, M. Fernández y A. Gómez. *Ontology Based Information Exchanged For Knowledge Management and Electronic Commerce*. Universidad Politécnica de Madrid. Disponible en:
<http://www.ontoweb.org/About/Deliverables/ppOntoweb.pdf>
Consultado: Junio 16 de 2007.
- [22]. A. Gomez, O. Corcho y M. Fernandez, *Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Belirn: Springer, 2004.
- [23]. O. Corcho, M. F. Lopez, A. G. Pérez. "Methodologies, tools and languages for building ontologies. Where is their meeting point?," *Data & Knowledge Engineering*, vol. 46, pp. 41–64, 2003.
- [24]. R. V. Guha, y D. Lenat. "Cyc: A Midterm Report". Disponible en:
http://www.cyc.com/doc/articles/midterm_report_1990.pdf. Consultado: Enero 10 de 2008.
- [25]. M. Fernández López, "Overview Of Methodologies For Building Ontologies". Disponible en:
<http://www.lsi.upc.edu/~bejar/aia/aia-web/4-fernandez.pdf>. Consultado: Abril 2 de 2008.
- [26]. E. Morales, "Metodología de Construcción usada en KACTUS". Disponible en:
<http://ccc.inaoep.mx/~emorales/Cursos/RdeC/node200.html>. Consultado: Enero 12 de 2008.

- [27]. F. Ruiz González, “*Definición de un Entorno para la Gestión del Mantenimiento de Software*,” Tesis Doctoral, Universidad de Castilla La Mancha, Departamento de Informática. Disponible en: <http://alarcos.inf-cr.uclm.es/doc/psgc/doc/lec/parte1/ruiz-cap5anexos.pdf>. Consultado: Febrero 2 de 2008.
- [28]. J. Contreras, y J. Martínez Comeche, “Tutorial ontologías”. Disponible en: http://www.sedic.es/gt_normalizacion_tutorial_ontologias.pdf
Consultado: Junio 7 de 2008.
- [29]. B. Biebow, S. Szulman, “TERMINAE: a method and a tool to build a domain ontology”, Université de Paris-Nord. Disponible en: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=3DF3A0EEB458219165324CEA26E461F5?doi=10.1.1.22.7575&rep=rep1&type=pdf>
Consultado: Enero 16 de 2008.
- [30]. N. Noy, D. McGuinness, “Ontology Development 101: A Guide to Creating Your First Ontology”. Disponible en: http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html. Consultado: Diciembre 18 de 2008
- [31]. C. Gómez Moreno, D. González Chana, “El uso de las ontologías para evaluar modelos conceptuales”. Disponible en: www.sinbad.dit.upm.es/docencia/grado/curso0506/25_10_2005_Ontologias_Texto.pdf. Consultado: Agosto 13 de 2008.
- [32]. ISA-88.com. ISA-88.04. 2006. Disponible en: <http://www.isa-88.com/subpages/technology/isa-88/isa-88-04.php?>
Consultado: Febrero 8 de 2009.

- [33]. ISA, ANSI/ISA–95.00.01–2000. Enterprise-Control System Integration Part 1: Models and Terminology. 2000. Consultado: Mayo 15 de 2008.
- [34]. M. Pino, “*Evaluación Comparativa de aplicaciones Web entre J2EE y Microsoft.NET,*” Tesis de grado, Universidad Católica de Temuco, Facultad de Ingeniería. 2006. Disponible en: <http://biblioteca.uct.cl/tesis/miguel-garrido/tesis.pdf> . Consultado: Agosto 2 de 2008.
- [35]. Jena – A Semantic Web Framework for Java. 2008 Disponible en: <http://jena.sourceforge.net/>. Consultado: Enero 11 de 2009.
- [36]. C. Bizer, R. Jörg Garbers, y O. Maresch. “The D2RQ Platform v0.5.1 - Treating Non-RDF Relational Databases as Virtual RDF Graphs”. 2007. Disponible en: <http://www4.wiwiss.fu-berlin.de/bizer/D2RQ/spec/>. Consultado: Febrero 9 2009.