



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

SISTEMA DE CAPTURA Y BÚSQUEDA DE METADATOS EN  
LÍNEA CON BASE EN LA NORMA TÉCNICA PARA LA  
ELABORACIÓN DE METADATOS GEOGRÁFICOS DEL INEGI

T E S I S

QUE PARA OBTENER EL TÍTULO DE:  
INGENIERA GEOMÁTICA

PRESENTA:

ALEXIS JAZMÍN PALOMARES LÓPEZ

DIRECTOR DE TESIS

M.I. HÉCTOR DANIEL RESÉNDIZ LÓPEZ



CIUDAD UNIVERSITARIA, Cd. Mex., 2017



*Dedicado a las personas más importantes en mi vida:  
mis padres Teresa y Daniel y mi hermano Edgar.  
Gracias por cuidar de mí, por sus consejos,  
y hasta por los regaños.  
Gracias por su apoyo incondicional y  
por darle sentido a mi vida.*



# Agradecimientos

Agradezco a mi alma máter: la Universidad Nacional Autónoma de México por todo el apoyo brindado durante mi formación académica desde la prepa, por todo lo que nos ofrece: sus instalaciones, actividades culturales, recreativas, deportivas que me hicieron sentir a esta casa de estudios como mi segundo hogar y porque en ella conocí a personas que siguen siendo una parte importante de mi vida.

También quiero agradecer a aquellas personas que fueron fundamentales en mi proceso de formación profesional: a todos mis profesores de la Facultad de Ingeniería. Gracias por compartir sus conocimientos y parte de sus experiencias y por todas las enseñanzas brindadas durante esta etapa universitaria.

De igual forma quiero agradecer a un grupo de personas que fueron importantes durante mi formación profesional: al grupo de becarios del Programa de Tecnología en Cómputo (PROTECO), en especial a los chicos que fueron mis instructores durante mi etapa como *prebe*.

Agradezco al Dr. Luis Chias Becerril por haberme brindado la oportunidad de pertenecer a la Unidad en Geotecnología Inteligente en Transporte y Sustentabilidad (GITS), porque aquí he conocido a personas con quienes he tenido el honor de trabajar y también de quienes he podido aprender.

Aprovecho para agradecer de forma especial a los compañeros de la Unidad GITS que de forma directa hicieron posible la realización de este trabajo:

- A la Ing. Marisol Mancera Cedillo quien ha compartido conmigo sus conocimientos acerca de los metadatos geográficos y quien propuso la creación de este sistema.
- Al Ing. Jesús Olvera Ramírez, de quien he tenido el honor de aprender acerca de normatividad geoespacial y quien también ayudó en la revisión y corrección del primer capítulo.
- Al Lic. Juan Carlos Cortés Ortíz, la “mente maestra” detrás del sistema, pues fue él quien lo diseñó y quien me introdujo en el uso de las tecnologías con que se desarrolló. También por su invaluable aporte para la escritura y revisión del tercer capítulo.
- Al Ing. Jonathan Landin Zaragoza quien empezó a escribir el código del sistema y quien también me ofreció consejos para la elaboración de diagramas y mejorar la apariencia de este escrito y me ayudó con la investigación de conceptos para escribir el primer capítulo así como en la revisión de este trabajo escrito.
- A mi jefe, el Lic. Rodrigo Jiménez del Valle, quien me apoyó durante todo este proceso de escritura del trabajo. Gracias por la motivación y por hacerme sentir parte de este equipo y también por ayudarme en la revisión del contenido de todos los capítulos.

- Y a quien ha sido mi guía durante todo este proceso de escritura, mi director de tesis: el M.I. Héctor Daniel Reséndiz López. Gracias por todos los aportes, sugerencias, críticas, comentarios y la bibliografía proporcionada para el desarrollo de este trabajo.

También quiero agradecer a mis sinodales: Dra. Clemencia Santos Cerquera, M.G. Juan Manuel Núñez Hernández, Ing. Erik de Valle Salgado y M.C.T María Elena Osorio Tai por el tiempo que dedicaron a la revisión de este trabajo y por los comentarios y sugerencias hechas para la mejora del mismo.

Finalmente a mis amigos de la facultad: Diego Martínez Montes, Sandra Liliana Ocampo Cervantes y Tania Osiris Avilés Martínez. Gracias por su apoyo y por hacerme sentir acompañada durante este camino. Gracias por los buenos y malos momentos vividos a su lado. Los quiero.

# Índice general

Agradecimientos	III
Índice de figuras	VII
Índice de tablas	XI
Lista de siglas y acrónimos	XIII
Introducción	1
<b>1. Estado del arte</b>	<b>5</b>
1.1. Marco teórico . . . . .	5
1.1.1. Teoría general de sistemas . . . . .	6
1.1.1.1. Sistemas . . . . .	6
1.1.1.2. Teoría general de sistemas y el enfoque de sistemas . . . . .	7
1.1.1.3. Tipos de sistemas . . . . .	7
1.1.1.4. Estructura de los sistemas . . . . .	8
1.1.2. Modelo DIKW . . . . .	11
1.1.3. Capas de datos geográficos . . . . .	15
1.1.3.1. Formatos para datos geográficos . . . . .	18
1.1.3.2. Interoperabilidad . . . . .	19
1.1.4. Metadatos geográficos . . . . .	20
1.1.4.1. Federal Geographic Data Committee (FGDC) . . . . .	21
1.1.4.2. International Organization for Standardization (ISO) . . . . .	23
1.1.5. Infraestructuras de datos espaciales . . . . .	27
1.1.5.1. Geoportales . . . . .	30
1.2. Marco Normativo en México . . . . .	30
1.2.1. Ley del SNIEG . . . . .	31
1.2.2. Norma Técnica para la elaboración de Metadatos Geográficos . . . . .	36
1.2.3. Transparencia y datos abiertos: Norma Técnica para el acceso y publicación de Datos Abiertos de la Información Estadística y Geográfica de Interés Nacional	37
1.3. Software, sistemas y servicios para capturar y catalogar metadatos . . . . .	40
<b>2. Requerimientos del sistema</b>	<b>49</b>
2.1. Justificación y utilidad en la Geomática . . . . .	49
2.2. Análisis del problema . . . . .	52
2.2.1. Análisis FODA . . . . .	55

2.3. Documentación de requerimientos . . . . .	61
2.3.1. Objetivos y alcance del sistema . . . . .	62
2.3.2. Listado de requerimientos . . . . .	66
<b>3. Diseño del Sistema Gestor de Metadatos</b>	<b>81</b>
3.1. Arquitectura de datos . . . . .	82
3.1.1. Modelo de datos . . . . .	83
3.1.2. Tecnologías de bases de datos empleadas . . . . .	92
3.2. Arquitectura tecnológica de la aplicación . . . . .	93
3.2.1. Arquitectura de software por capas . . . . .	94
3.2.2. Tecnologías de desarrollo empleadas . . . . .	97
3.3. Infraestructura de hardware de la aplicación . . . . .	102
<b>4. Resultados</b>	<b>105</b>
<b>5. Conclusiones</b>	<b>115</b>
<b>Anexos</b>	<b>119</b>
<b>A. Clasificación de los sistemas</b>	<b>121</b>
<b>B. Estructura de los sistemas</b>	<b>125</b>
<b>C. Elementos de la Norma Técnica para la Elaboración de Metadatos Geográficos (NTM)</b>	<b>129</b>
<b>D. Ejemplos de sistemas para capturar y/o publicar metadatos geográficos</b>	<b>135</b>
<b>Referencias Bibliográficas</b>	<b>141</b>

# Índice de figuras

1.1. Ejemplo de caja negra para obtener un florero de barro. Elaboración propia, 2015. . . . .	9
1.2. Diagrama de bloques (o de estructura) de un sistema. Adaptado de: (Moreno Bonett, 2010, p.25) . . . . .	10
1.3. Diferentes tipos de estructuras de sistemas. Adaptado de: Ramírez (2002, p.23-25) . . . . .	11
1.4. La jerarquía DIKW en forma de pirámide. Fuente: Olvera Ramírez et al. (2014) . . . . .	13
1.5. La jerarquía DIKW en forma de bipirámide. Fuente: Olvera Ramírez et al. (2014) . . . . .	16
1.6. Línea del tiempo de las principales normas de metadatos geográficos a nivel internacional. Obtenida de: Olvera Ramírez (2014) . . . . .	22
1.7. Diagrama de las secciones de la norma CSDGM. Adaptado de: Federal Geographic Data Committee (2000) y Olvera Ramírez et al. (2014) . . . . .	24
1.8. Diagrama de la Norma ISO 19115 y sus dos partes. Fuente: elaboración propia con datos de Di y Deng (2014) . . . . .	25
1.9. Clasificación de geoportales. Fuente: Maguire y Longley (2005) . . . . .	31
1.10. Subsistemas Nacionales de Información y sus componentes. Tomada de: Instituto Nacional de Estadística y Geografía (2015a, p.7) . . . . .	32
1.11. Componentes del Subsistema Nacional de Información Geográfica y del Medio Ambiente. Tomada de: Instituto Nacional de Estadística y Geografía (2015a, p.8) . . . . .	33
1.12. Componentes del SNIEG. Tomada de: Instituto Nacional de Estadística y Geografía (2015a, p.9) . . . . .	34
1.13. Conceptos relacionados con Transparencia. Fuente: elaboración propia con información de Ferrer-Sapena et al. (2011) y López del Consuelo (2015) . . . . .	38
1.14. Fuentes nacionales de datos abiertos. Tomado de: Chias Becerril y Reséndiz López (2015). . . . .	41
1.15. Fuentes internacionales de datos abiertos. Tomado de: Chias Becerril y Reséndiz López (2015). . . . .	41
1.16. Los tres roles que completan la arquitectura de servicios web. Adaptado de: (Fu y Sun, 2010, p.52) . . . . .	43
1.17. Relación entre los principales estándares para metadatos geoespaciales. Traducido de Litwin y Rossa (2011, p.73). . . . .	44
1.18. Interfaz del SICAM, desarrollado por INEGI. Cada una de las pestañas contiene un formulario para el llenado de cada una de las secciones de la NTM . . . . .	45
1.19. Interfaz del buscador de metadatos del Clearinghouse del Instituto Nacional de Estadística y Geografía (INEGI) . . . . .	46
1.20. Ejemplo de un metadato encontrado a través del Clearinghouse del INEGI . . . . .	47

2.1.	Ejemplo de servicios de descubrimiento implementados por INEGI. Tomada de: Instituto Nacional de Estadística y Geografía (2015a, p.15)	50
2.2.	Etapas para la resolución del problema del sistema de metadatos.	52
2.3.	Comparación entre aplicaciones y servicios de catálogo de metadatos. Fuente: elaboración propia con datos de Pitney Bowes (2016), Federal Geographic Data Committee (2014), Díaz et al. (2014) y sitios web oficiales de las aplicaciones.	54
2.4.	Estructura del análisis FODA. Adaptado de Picón (2015)	56
2.5.	Análisis tipo FODA de situaciones alrededor de los sistemas de metadatos en México.	58
2.6.	Estrategias FA, FO, DA y DO en torno a la creación y utilización de sistemas de metadatos en México.	59
2.7.	Costo del personal. <b>A</b> =Arquitecto de software, <b>B</b> =Programador Jr. 1, <b>C</b> =Programador Jr. 2. Fuente: elaboración propia.	66
2.8.	Costo por uso de los equipos de cómputo. <b>A</b> =Arquitecto de software, <b>B</b> =Programador Jr. 1, <b>C</b> =Programador Jr. 2. Fuente: elaboración propia.	66
2.9.	Diagrama de bloques del Sistema Gestor de Metadatos.	70
2.10.	Diagrama de casos de uso del Sistema Gestor de Metadatos.	72
3.1.	Diagrama Conceptual del Sistema Gestor de Metadatos (SIGEM).	82
3.2.	Ejemplo de un data mart con esquema de estrella. Tomado de: García (2013)	84
3.3.	Ejemplo de un data mart con esquema de copo de nieve. Tomado de: García (2013)	85
3.4.	Modelo multidimensional del SIGEM.	86
3.5.	Modelo relacional de la dimensión <b>'dataset'</b> .	87
3.6.	Modelo relacional de la dimensión <b>'date_related'</b> .	88
3.7.	Modelo relacional de la dimensión <b>'address_book'</b> .	88
3.8.	Modelo relacional de la dimensión <b>'geographical_location'</b> .	89
3.9.	Modelo relacional de la dimensión <b>'reference_system'</b> .	89
3.10.	Modelo relacional de la dimensión <b>'quality_information'</b> .	90
3.11.	Modelo relacional de la dimensión <b>'distribution'</b> .	91
3.12.	Modelo relacional de la dimensión <b>'contact_info_metadata'</b> .	91
3.13.	Subgrupos y comandos del lenguaje SQL. Tomado de: W3schools (2016).	93
3.14.	Diagrama de una arquitectura de capas. Tomado de: Pressman (2010, p.215).	95
3.15.	Paquetes que conforman el sistema.	96
3.16.	Tecnologías de desarrollo del SIGEM.	98
3.17.	Infraestructura de hardware para el SIGEM.	102
3.18.	Flujo de trabajo de la aplicación.	103
3.19.	Flujo de trabajo de la aplicación con el servidor web y el de bases de datos alojados en un mismo equipo.	104
4.1.	Página de inicio de sesión.	105
4.2.	Página de búsqueda de metadatos.	106
4.3.	Ejemplo de búsqueda de metadatos.	107
4.4.	Vista inicial de la página de captura de metadatos.	107
4.5.	Ejemplo de uno de los formularios de captura.	108
4.6.	Lista de opciones disponibles al editar un metadato por primera vez.	108
4.7.	Ejemplo de un mensaje de validación.	109
4.8.	Mensaje indicando que ya se ha guardado el metadato.	109
4.9.	Lista de opciones disponibles al editar un metadato ya guardado en la base de datos.	110

4.10. Mensaje de correcta actualización de un metadato . . . . .	110
4.11. Mensaje de error lanzado cuando el nombre técnico no es cambiado. . . . .	110
4.12. Opción “Descargar avances” seleccionada. . . . .	111
4.13. Ventana emergente para la selección de metadatos en formato XML. . . . .	111
4.14. Ejemplo de un metadato mostrado a través del portal geográfico de la Unidad Académica de Geotecnología en Infraestructura, Transporte y Sustentabilidad (U-GITS). . . . .	112
4.15. Ejemplo de metadatos en formato HTML. . . . .	112
4.16. Ejemplo de metadatos en formato PDF. . . . .	113
4.17. Ejemplo de metadato en formato XML. . . . .	114
A.1. Clasificación de los sistemas según la relación con su ambiente. Elaboración propia con base en Moreno Bonett (2010, pp.19-22). . . . .	121
A.2. Clasificación de los sistemas considerando al tiempo como variable. Elaboración propia con base en Moreno Bonett (2010, pp.19-22). . . . .	122
B.1. Grafo de flujo derivado del diagrama de bloques de la Figura 1.2. . . . .	125
B.2. Extracto de un diagrama de PERT. (Fuente: ADA Consultores (2015)). . . . .	126
B.3. Diagrama de Pert. Adaptado de: Free-Logistics (2015) . . . . .	127
C.1. Elementos de la primera sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b) . . . . .	129
C.2. Elementos de la segunda sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b) . . . . .	130
C.3. Elementos de la tercera sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b) . . . . .	130
C.4. Elementos de la cuarta sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b) . . . . .	131
C.5. Elementos de la quinta sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b) . . . . .	131
C.6. Elementos de la sexta sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b) . . . . .	132
C.7. Elementos de la séptima sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b) . . . . .	132
C.8. Elementos de la octava sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b) . . . . .	133
C.9. Elementos de la novena sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b) . . . . .	133
D.1. Software MetaD para la captura de metadatos. Tomada de: Leal Villamil (2011). . . . .	135
D.2. Software CatMDEdit para la captura de metadatos. Imagen obtenida de: Grupo de Sistemas de Información Avanzados de la Universidad de Zaragoza (IAAA) y GeospatiumLab (GSL) (2010). . . . .	136
D.3. Ejemplo de un metadato consultado a través de la aplicación ArcCatalog. Tomada de: Bocchino y Konopka (2011). . . . .	137
D.4. Ejemplo de un servicio de catálogo implementado con Geonetwork para el geoportal de la Infraestructura de Datos Espaciales (IDE) de Uruguay. . . . .	138
D.5. Ejemplo de un servicio de catálogo implementado con MapInfo Manager. Imagen tomada de: Pitney Bowes (2016). . . . .	138

D.6. Imagen que ilustra el funcionamiento del servicio CSW provisto por la aplicación <i>pycsw</i> . Imagen tomada de: Open Source Geospatial Foundation (2015). . . . .	139
--	-----

# Índice de tablas

1.1. Características de los sistemas espaciales. Fuente: Krzanowski y Raper (2001) . . . . .	8
1.2. Comparación de la cadena DIKW según Zeleny(1987) y Ackoff(1989). Fuente: Rowley (2007, p.167). . . . .	12
1.3. Contraparte de los elementos del modelo DIKW. Fuente: Bernstein (2009). . . . .	14
1.4. Lista de formatos ráster y vectoriales para capas de datos geográficos. Fuente: (Sherman, 2008, p.91-92) . . . . .	18
1.5. Resumen de las secciones de la norma CSDGM del FGDC. Fuente: Elaboración propia con datos del Federal Geographic Data Committee (1998) . . . . .	22
1.6. Elemento núcleo de la norma ISO 19115. <b>O</b> = elemento obligatorio, <b>C</b> = elemento condicional, <b>Opc.</b> = elemento opcional. Fuente: Olvera Ramírez et al. (2014) . . . . .	24
1.7. Ejemplos de perfiles de la norma ISO 19115 . . . . .	26
1.8. Categorización de los documentos regulatorios publicados por INEGI en el marco de la LSNIEG. . . . .	35
1.9. Componentes de los metadatos para un conjunto de datos publicado como Datos Abiertos. Fuente: (Instituto Nacional de Estadística y Geografía, 2014, Artículo 9) . . . . .	39
2.1. Matriz de requerimientos de datos . . . . .	67
2.2. Matriz de requerimientos de aplicación. . . . .	68
2.3. Flujo normal del caso de uso “Iniciar sesión” . . . . .	72
2.4. Flujo normal del caso de uso “Editar metadato” . . . . .	73
2.5. Flujo normal del caso de uso “Descargar avance” . . . . .	73
2.6. Flujo normal del caso de uso “Guardar metadato” . . . . .	74
2.7. Flujo normal del caso de uso “Buscar metadato” . . . . .	75
2.8. Flujo normal del caso de uso “Exportar metadato” . . . . .	76
2.9. Flujo normal del caso de uso “Actualizar metadato” . . . . .	76
2.10. Flujo normal del caso de uso “Cargar metadato” . . . . .	77
2.11. Matriz de requerimientos tecnológicos. . . . .	78
2.12. Etapas de desarrollo del Sistema Gestor de Metadatos. . . . .	79



# Lista de siglas y acrónimos

<b>ANIM</b>	Atlas Nacional Interactivo de México
<b>CEN</b>	European Committee for Standardization
<b>CONABIO</b>	Comisión Nacional para el Conocimiento y Uso de la Biodiversidad
<b>CONAPO</b>	Consejo Nacional de Población
<b>CSDGM</b>	Content Standard for Digital Geospatial Metadata
<b>CSW</b>	Catalogue Service for the Web
<b>DCMI</b>	Dublin Core Metadata Initiative
<b>DENUE</b>	Directorio Estadístico Nacional de Unidades Económicas
<b>DIKW</b>	Data, Information, Knowledge and Wisdom
<b>ESRI</b>	Environmental Systems Research Institute
<b>FTP</b>	File Transfer Protocol
<b>GWT</b>	Google Web Toolkit
<b>HTTP</b>	Hypertext Transfer Protocol
<b>IDE</b>	Infraestructura de Datos Espaciales
<b>IG - UNAM</b>	Instituto de Geografía de la UNAM
<b>IGAC</b>	Instituto Geográfico Agustín Codazzi
<b>IGN</b>	Instituto Geográfico Nacional de España
<b>IM</b>	Instant Messaging
<b>INEE</b>	Instituto Nacional para la Evaluación de la Educación
<b>INEGI</b>	Instituto Nacional de Estadística y Geografía
<b>INSPIRE</b>	Infrastructure for Spatial Information in Europe
<b>IPGH</b>	Instituto Panamericano de Geografía e Historia
<b>IRC</b>	Internet Relay Chat
<b>ISO</b>	International Organization for Standardization
<b>LSNIEG</b>	Ley del Sistema Nacional de Información Estadística y Geográfica
<b>NAP</b>	Perfil Norteamericano de Metadatos

<b>NEM</b>	<b>N</b> úcleo <b>E</b> spañol de <b>M</b> etadatos
<b>NTM</b>	<b>N</b> orma <b>T</b> écnica para la <b>E</b> laboración de <b>M</b> etadatos <b>G</b> eográficos
<b>OCDE</b>	<b>O</b> rganización para la <b>C</b> ooperación y el <b>D</b> esarrollo <b>E</b> conómicos
<b>OGC</b>	<b>O</b> pen <b>G</b> eospatial <b>C</b> onsortium
<b>RTM</b>	<b>R</b> equirements <b>T</b> raceability <b>M</b> atrix
<b>SCT</b>	<b>S</b> ecretaría de <b>C</b> omunicaciones y <b>T</b> ransportes
<b>SEDATU</b>	<b>S</b> ecretaría de <b>D</b> esarrollo <b>A</b> grario, <b>T</b> erritorial y <b>U</b> rmano
<b>SEMARNAT</b>	<b>S</b> ecretaría del <b>M</b> edio <b>A</b> mbiente y <b>R</b> ecursos <b>N</b> aturales
<b>SIATL</b>	<b>S</b> imulador de <b>F</b> lujos de <b>A</b> gua de <b>C</b> uencas
<b>SICAM</b>	<b>S</b> istema de <b>C</b> aptura de <b>M</b> etadatos
<b>SIG</b>	<b>S</b> istema de <b>I</b> nformación <b>G</b> eográfica
<b>SIGEM</b>	<b>S</b> istema <b>G</b> estor de <b>M</b> etadatos
<b>SMDB</b>	<b>S</b> istema <b>M</b> anejador de <b>B</b> ases de <b>D</b> atos
<b>SMTP</b>	<b>S</b> imple <b>M</b> ail <b>T</b> ransfer <b>P</b> rotocol
<b>SNIEG</b>	<b>S</b> istema <b>N</b> acional de <b>I</b> nformación <b>E</b> stadística y <b>G</b> eográfica
<b>SNIGMA</b>	<b>S</b> ubsistema <b>N</b> acional de <b>I</b> nformación <b>G</b> eográfica y del <b>M</b> edio <b>A</b> mbiente
<b>U-GITS</b>	<b>U</b> nidad <b>A</b> cadémica de <b>G</b> eotecnología en <b>I</b> nfraestructura, <b>T</b> ransporte y <b>S</b> ustentabilidad
<b>UML</b>	<b>U</b> nified <b>M</b> odeling <b>L</b> anguage
<b>WCS</b>	<b>W</b> eb <b>C</b> overage <b>S</b> ervice
<b>WFS</b>	<b>W</b> eb <b>F</b> eature <b>S</b> ervice
<b>WMS</b>	<b>W</b> eb <b>M</b> ap <b>S</b> ervice
<b>WWW</b>	<b>W</b> orld <b>W</b> ide <b>W</b> eb

# Introducción

La información geográfica puede ser empleada para sustentar la toma de decisiones en cuestiones relacionadas con el territorio, abarcando áreas como: transporte, desarrollo de comunidades, agricultura, respuesta a emergencias, manejo de recursos naturales, entre otras. Sin embargo, que dichas decisiones resulten o no adecuadas dependerá en parte de la calidad de los datos con que se cuente, es decir, por ejemplo: que éstos cubran el área geográfica de interés, que la temporalidad de los datos sea consistente, que estén estructurados y sobre todo que estén adecuadamente documentados y que de ellos se haga una correcta interpretación, refiriéndose este último punto a la creación de metadatos.

Los metadatos tienen la finalidad de brindar información acerca de un producto y/o servicio y por medio de ellos se puede conocer los criterios de calidad de dichos productos o servicios. Lo mismo aplica para los metadatos de capas de información geográfica; pues éstos son la documentación de las capas geográficas y permiten conocer por ejemplo: quién o quiénes elaboraron la capa, cuándo, los insumos empleados, el procedimiento seguido para la elaboración de la misma, el área geográfica que cubierta, el propósito de la creación de los datos, cómo se distribuyen, cómo acceder a ellos y además proporcionan una descripción de los datos contenidos en la capa.

Debido a la creciente importancia de contar con datos geográficos, se han implementado iniciativas dedicadas a la creación de infraestructuras tecnológicas que permitan ya sea, compartir este tipo de información para que pueda ser descargada y utilizada o bien, para que los interesados sepan que ya existen determinados conjuntos de datos y cómo adquirirlos. A estas infraestructuras tecnológicas se les conoce como Infraestructuras de Datos Espaciales, a las cuales se accede por medio de geoportales.

Otra forma de compartir este tipo de información es mediante la iniciativa de datos abiertos. En México, existe un portal web <sup>1</sup> en el que distintas dependencias gubernamentales publican datos abiertos de diversos temas como: educación, energía, medio ambiente, salud, finanzas, entre otros. Dentro de estos grupos de datos se encuentra la categoría **geoespacial**, en la que se publican datos en formato KML, KMZ, Shapefile, CSV, entre otros formatos. Sin embargo, a pesar de que dichos datos están disponibles para su libre uso, varios de éstos sólo proveen el acceso a la capa sin mostrar ni el metadato ni el diccionario de variables de la misma, por lo que puede resultar un tanto complicado hacer uso de dichas capas pues no es posible conocer a detalle el contenido de las mismas.

Sin embargo, este no es el único sitio en el que se publican datos geoespaciales generados por la administración pública, pues en los portales web de algunas otras dependencias gubernamentales como la **Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO)** o la **Secretaría del Medio Ambiente y Recursos Naturales (SEMARNAT)** también se publica información geoespacial y, al menos estas dos dependencias, también publican el metadato de la capa. En este punto existe otro inconveniente: de acuerdo con la NTM del INEGI, en su segundo artículo,

---

<sup>1</sup><https://datos.gob.mx/> [Consultado en enero de 2017]

todas las Unidades de Estado o terceros, que por encargo de alguna unidad de estado, generen información geográfica de interés nacional deben documentar dicha información bajo los lineamientos de esta norma; pero ésto no siempre ocurre pues normalmente estos metadatos se entregan bajo los lineamientos del **Content Standard for Digital Geospatial Metadata** (CSDGM) o con un metadato generado por algún **Sistema de Información Geográfica** (SIG), por lo que se incumple con la norma del INEGI.

En un contexto más local, dentro de la U-GITS del **Instituto de Geografía de la UNAM** (IG - UNAM), se trabaja entre otras cosas, en la generación de capas de información geográfica así como en portales para la publicación de dicha información y de acuerdo con la norma del INEGI previamente citada, estas capas deben ser documentadas bajo los términos de dicha norma pues estos proyectos se han realizado por encargo de unidades de estado como la **Secretaría de Comunicaciones y Transportes** (SCT) o el **Instituto Nacional para la Evaluación de la Educación** (INEE). Si bien, el INEGI puso a disposición del público un software denominado **Sistema de Captura de Metadatos** (SICAM) gracias al cual se pueden capturar metadatos geográficos siguiendo los parámetros de la NTM, éste resultaba insuficiente para las necesidades de la U-GITS por lo que se decidió desarrollar un sistema propio de captura de metadatos al que se denominó **Sistema Gestor de Metadatos** (**SIGEM**) para que cumpliera principalmente con las siguientes características: funcionar bajo una plataforma web, proporcionar un repositorio centralizado de metadatos y que los metadatos puedan ser consultados como un servicio web sin necesidad de ingresar al sistema, lo cual, además, facilitaría su publicación dentro de los portales creados por este grupo de trabajo.

Finalmente, el sistema aquí presentado es el resultado de un proceso de toma de decisiones que de forma muy general puede ser dividido en dos partes. La primera parte, involucró al grupo de Normatividad Geoespacial de la U-GITS, pues las personas que lo conforman detectaron la necesidad de contar con un nodo de propio de metadatos, lo que a su vez implicaba contar también con una interfaz que permitiera a los encargados de generar los metadatos ingresarlos al nodo de metadatos. Por otro lado, la segunda etapa de este proceso involucra al grupo de investigación y desarrollo de software de la U-GITS, ya que el personal que lo conforma tuvo a su cargo el desarrollo de la interfaz de captura y muestra de metadatos así como del diseño de la base de datos, los cuales en conjunto conforman el nodo de metadatos de la U-GITS. Posteriormente, en el capítulo 2, se menciona el rol que jugó cada persona involucrada en el desarrollo de este sistema.

## Objetivos

En este trabajo se presenta el **Sistema Gestor de Metadatos** (SIGEM) buscando cumplir con los siguientes objetivos:

- Establecer la importancia del uso de los metadatos geográficos.
- Exponer el marco teórico que fundamenta la necesidad de contar con herramientas generadoras de metadatos geográficos.
- Conocer los elementos de la **Norma Técnica para la Elaboración de Metadatos Geográficos** (NTM) del INEGI.
- Explicar el proceso de identificación de requerimientos, diseño y elaboración del SIGEM.
- Reconocer la tecnología de desarrollo y la arquitectura tecnológica necesaria para el funcionamiento del SIGEM.

- Dar a conocer la utilidad y beneficios del SIGEM.

## Estructura capitular

A continuación se resume el contenido de cada uno de los capítulos de este trabajo:

En el capítulo 1 se aborda el marco teórico sustenta la realización del SIGEM. Se comienza por mostrar conceptos relacionados con la teoría general de sistemas para posteriormente abordar el modelo DICS (o Data, Information, Knowledge y Wisdom, en inglés) a partir del cual queda de manifiesto la importancia de la utilización de metadatos; pero no sólo para capas de datos geográficos sino también para cualquier tipo de datos e información. En la segunda parte de este capítulo se hace mención de las principales normas de metadatos geográficos a nivel mundial, así como del marco normativo en México relacionado con la ubicación de información geoespacial el cual incluye la ley del Sistema Nacional de Información Estadística y Geográfica (SNIEG) del INEGI y las normas para metadatos geográficos que existen en México. Finalmente, en la tercera parte de este capítulo se mencionan algunos de los sistemas, software y servicios que existen para documentar información geoespacial.

El 2o capítulo está dedicado a describir el proceso que concluyó en la decisión de crear el SIGEM, así como el levantamiento de requerimientos para este mismo sistema.

El capítulo 3 muestra la parte tecnológica del sistema. En él se aborda el modelo lógico de datos, la arquitectura de software del sistema y la infraestructura de hardware necesarios para la puesta en marcha del sistema. Además se explica de manera general cuáles fueron las tecnologías (herramientas de construcción, lenguajes de programación, marcos de trabajo o *frameworks*) que permitieron desarrollar el SIGEM.

En el capítulo 4 se presenta la parte gráfica del sistema, es decir, los principales elementos que conforman la vista del sistema, que son: el buscador de metadatos, el formulario de captura, el cargador de metadatos en formato XML y ejemplos de metadatos exportados en formatos HTML, PDF y XML.

Por último, el capítulo 5 está dedicado a las conclusiones en torno a la utilidad del sistema, el proceso de creación del mismo y las mejoras que podrían implementarse en el SIGEM como trabajo a futuro.



# Capítulo 1

## Estado del arte

Durante los últimos años, alrededor del mundo (principalmente en países con alto desarrollo como Estados Unidos, Reino Unido, España, etc.), a través de un conjunto de políticas y del empleo de las nuevas tecnologías informáticas y geoespaciales, se han desarrollado *Infraestructuras de Datos Espaciales* o IDE nacionales, las cuales son resultado de un gran esfuerzo interdisciplinario, brindando beneficios económicos a las naciones que las implementan, pues permiten poner a disposición de las instancias de gobierno y del público en general datos geográficos de interés nacional que pueden servir para desarrollar proyectos o nuevas políticas públicas, además de ayudar a evitar la réplica de esfuerzos en la generación de los datos, entre otros varios beneficios que se abordarán.

Este capítulo está dividido en tres secciones:

1)Marco teórico, 2)Marco Normativo en México y 3)Software, sistemas y servicios para capturar y catalogar metadatos.

En la primera se tocará el tema de IDE's, así como otros conceptos que es importante tener en cuenta para entender de forma más clara cómo están constituidas tanto física como conceptualmente. En la segunda sección se hará referencia a leyes y normas del INEGI para regular la publicación de datos e información geográfica de interés nacional y de datos abiertos. Finalmente, en la tercera sección se mencionarán otros sistemas y servicios con los que también es posible capturar y catalogar metadatos geográficos, aunque estos están basados en otras normas de metadatos.

### 1.1. Marco teórico

Si bien el tema que atañe a esta tesis es el sistema de captura y búsqueda de metadatos, en esta sección se ha decidido tomar como punto de referencia las infraestructuras de datos espaciales. Una parte fundamental de ellas son los datos que contiene; pero debido a la naturaleza y al uso que se hace de los mismos es de vital importancia asegurar y dar a conocer la calidad de estos datos, lo cual se logra en parte gracias a la documentación de éstos, a través de los metadatos.

En el primer apartado de esta sección se hace un resumen de algunos conceptos de la teoría general de sistemas, lo cual ayudará a comprender la forma en que está estructurado este sistema de captura de metadatos geográficos, así como la forma en que está estructurada una IDE, pues como se verá más adelante, ésta es equiparable a un sistema.

Ya se mencionó que una parte fundamental de una IDE son los datos que ésta contiene, por lo que también se incluye un apartado acerca de capas de datos geográficos en el que se explica cómo

se forman y algunos de sus formatos de representación. Una vez que se cuenta con los datos, éstos deben ser documentados para su inclusión en una IDE, razón por la que se incluye otro apartado acerca de metadatos geográficos: qué son, cómo surgieron y algunas normas que regulan la creación de los mismos.

Por otro lado, como menciona Olvera Ramírez et al. (2014), los términos *dato* e *información* suelen utilizarse de forma indistinta. Más adelante, en el apartado del modelo Datos, Información, Conocimiento y Sabiduría (DICS) o bien **D**ata, **I**nformation, **K**nowledge and **W**isdom (DIKW), por sus siglas en inglés, se mostrará la diferencia entre estos dos conceptos, así como la cadena que lleva de *datos* a *sabiduría*, pasando por *información* y *conocimiento*. Así mismo, se pondrá de manifiesto la importancia del uso de metadatos, al menos para la documentación de capas de datos geográficos.

Una vez que se tienen los datos, los metadatos, además de otros elementos, es posible comenzar a integrar una Infraestructura de Datos Espaciales, por ello en el último apartado de esta sección se presenta precisamente este tema: un poco de askinshistoria acerca de ellas, cuáles son sus componentes así como de algunas de la tendencias actuales para implementarlas y acceder a ellas a través de un geoportal.

### 1.1.1. Teoría general de sistemas

Uno de los objetivos de este trabajo es la presentación de un sistema de captura y búsqueda de metadatos de datos geoespaciales en línea; pero ¿qué y por qué es un sistema? Debido a esto se incluye en este capítulo un apartado acerca de la Teoría general de sistemas, con lo que se responde a qué es un sistema, además de incluirse algunas de las clasificaciones de sistemas que existen. Finalmente, también se incluye un apartado acerca de algunos de los tipos de estructuras de sistemas que existen.

En el capítulo 3 se explica el diseño del sistema de captura de metadatos, es decir, el modelo de datos y su arquitectura, por lo que al tener nociones de esta teoría general de sistemas es posible entender qué tipo de sistema es, cómo está estructurado y, finalmente poder responder a la pregunta: ¿por qué es un sistema?

#### 1.1.1.1. Sistemas

En la literatura existen varias definiciones para el concepto de *sistema*, sin embargo la gran mayoría de ellas coinciden en definirlo como un conjunto o una colección de elementos; sin embargo los sistemas no sólo son eso. Para que un sistema sea considerado como tal, todos sus elementos deben estar relacionados de forma directa o indirecta bajo ciertas reglas o criterios, es decir, estar organizados. De lo anterior puede concluirse que un sistema se define o caracteriza por sus elementos y sus relaciones.

En este punto cobra relevancia el concepto de sistema de interés (*system of interest*) (Haskins et al., 2010, p.9. Citado por Moser (2014)), pues un sistema no sólo se define por sus elementos y relaciones, sino que también se definen según el interés del observador y de un objetivo o propósito con el que el sistema debe cumplir. Entonces, el sistema de interés de una persona puede ser visto como un elemento del sistema de interés de otra persona. Ésto confiere a los sistemas una característica jerárquica, lo cual no se refiere a que haya una cadena de instrucciones de arriba a abajo, sino que un sistema puede a su vez ser un elemento o un subsistema de otro sistema. Con ello, la jerarquía de sistemas (Moser, 2014, p.12) queda de la siguiente manera: subsistema, sistema (de interés) e hipersistema (sistema de sistemas). Como un ejemplo de esto: si un motor es un sistema de interés, un carro es un hipersistema y la cámara de combustión un subsistema (Moser, 2014, p.13).

Otra característica de los sistemas es la integridad, lo cual se traduce a que un sistema forma una unidad especial al considerarla con su entorno. El entorno es todo aquello que queda fuera de la frontera del sistema; por lo tanto, el mismo sistema de interés determina la frontera del sistema. Finalmente, un sistema puede pensarse como un “Conjunto de cosas que relacionadas entre sí ordenadamente contribuyen a determinado objeto” (Real Academia Española, 2014).

#### 1.1.1.2. Teoría general de sistemas y el enfoque de sistemas

Los principios de la Teoría General de Sistemas fueron propuestos por el biólogo Ludwig Von Bertalanffy, cuando en 1925 publicó su trabajo sobre sistemas abiertos. Aunque en un inicio este planteamiento no fue muy bien recibido por la comunidad científica, con el tiempo fue adoptado por especialistas de diversas ramas del conocimiento.

Para plantear los principios de la Teoría General de Sistemas, Bertalanffy (1987) se dio cuenta que a pesar del alto y creciente grado de especialización que hay en la ciencia moderna, lo que lleva al surgimiento de cada vez más ramas de conocimiento, existe un fenómeno que resulta por demás interesante, pues “... han surgido problemas y concepciones similares en campos muy distintos, independientemente.”(p.30), lo que lo llevó a pensar que “existen modelos, principios y leyes aplicables a sistemas generalizados o a sus subclases, sin importar su género particular, la naturaleza de sus elementos componentes y las relaciones o fuerzas que imperen entre ellos”(p.32).

Como consecuencia de lo planteado anteriormente, Bertalanffy (1987) llegó a una teoría con principios aplicables a todos los sistemas en general. Esta teoría general de los sistemas es una disciplina y a la vez una herramienta gracias a la cual es posible tener una mejor aproximación o una mejor representación de la realidad. Arnold y Osorio (1998), propone que esta teoría general de los sistemas ofrece “un ambiente adecuado para la interrelación y comunicación fecunda entre especialistas y especialidades”. Por lo que Johansen Bertoglio (1982) mencionó que “la Teoría General de Sistemas es un corte horizontal que pasa a través de todos los diferentes campos del saber humano, para explicar y predecir la conducta de la realidad” (p.14).

Por otro lado, el enfoque elementalista o reduccionista es un método en el que se busca reducir el todo a sus partes fundamentales y estudiar cada una de estas partes de manera independiente. Es precisamente este enfoque el que ha permitido un gran avance en el desarrollo de las ciencias naturales (principalmente), gracias a lo cual se da el alto grado de especialización de la ciencia moderna mencionado anteriormente; pero este enfoque proporciona una pobre comprensión de la realidad. De forma contraria, el enfoque sistémico o integrado, según Drucker (1972) (citado en Moreno Bonett (2010, p.15)) “considera una gran cantidad de actividades y procesos, antes inconexos, como parte de un todo integral, más grande, no es algo que en sí mismo sea tecnológico, (sino) una manera de observar al mundo y a nosotros mismos”.

Este enfoque sistémico llegó a ser un nuevo paradigma gracias al cual, los problemas se estudian desde un punto de vista diferente, pues en este enfoque se considera que los objetos de estudio (o sistemas) tienen propiedades y cualidades que no pueden ser entendidas o explicadas hasta que son vistas en conjunto con su entorno, es decir, en su totalidad.

#### 1.1.1.3. Tipos de sistemas

En la literatura existen varias formas de clasificar los sistemas. Por ejemplo, Moreno Bonett (2010, pp.19-22), hace referencia a tres formas de clasificación: 1) considerando la relación de los sistemas con su ambiente o ámbito; 2) considerando el tiempo y 3) considerando la conectividad, unicidad y estabilidad estructural entre las partes del sistema.

En el anexo A se muestran diagramas, ejemplificando estos tipos de clasificación de sistemas.

## Sistemas espaciales

Los sistemas espaciales son de especial interés en las ciencias de la Tierra. En general, estos sistemas están formados por objetos espaciales, es decir objetos localizados en el espacio. Entonces, los sistemas espaciales se definen como “una combinación de elementos físicos selectivos y los patrones de asociaciones que ellos forman” (Krzanowski y Raper, 2001, p.64). Un par de ejemplos de sistemas espaciales son: una ciudad o una red de carreteras.

Todos los sistemas espaciales comparten cuatro características comunes, las cuales se describen en la tabla 1.1:

**Tabla 1.1.** Características de los sistemas espaciales. Fuente: Krzanowski y Raper (2001)

Característica	Ejemplo
Están compuestos de objetos espaciales	Una red de carreteras está compuesta de carreteras. Las ciudades están compuestas de calles, edificios, parques, etc.
Los objetos espaciales tienen propiedades designadas	Las carreteras tienen longitud, tamaño, tipo, número. Los edificios tienen direcciones, nombres, ubicaciones.
Los objetos espaciales están interrelacionados	Las carreteras están conectadas e intersecan unas con otras. Los edificios están cerca, lejos, unos frente a otros.
Un sistema espacial tiene características designadas que permiten la diferenciación entre varios sistemas similares.	Una red de carreteras tiene densidad de carreteras, capacidad de tráfico, etc. Las ciudades tienen un número de construcciones, porcentaje de espacio libre, etc.

Debido a que los objetos que conforman un sistema espacial tienen, como principal característica, una ubicación en el espacio, éstos son susceptibles de ser mostrados en mapas, cartas topográficas, croquis, etc. mediante el empleo de capas de datos geográficos (en formato vectorial o ráster). Esto puede traducirse a que las capas de datos geográficos son la representación de elementos o de sistemas espaciales completos.

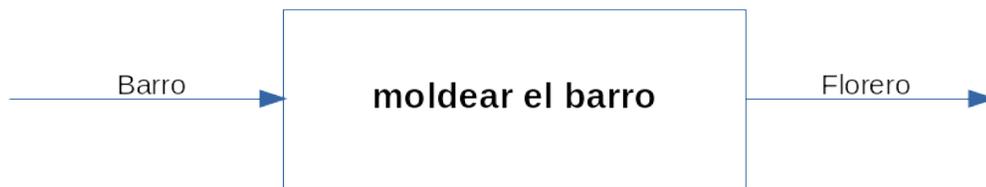
Para este trabajo, lo anterior tiene relevancia debido a que una vez que se cuenta con estas capas de datos, es importante para las personas que vayan a utilizarlas conocer a detalle lo que éstas representan y, en muchos casos, el proceso a través del cual fueron generados. Esta información puede ser obtenida a partir de los metadatos.

### 1.1.1.4. Estructura de los sistemas

Al inicio de la sección de sistemas se mencionó que todos los elementos de un sistema están relacionados bajo ciertas reglas. Entonces la forma en que estas relaciones mantienen unidos a los

elementos del sistema es a lo que se le llama estructura del sistema. La representación de la estructura de un sistema se puede hacer mediante un *diagrama de bloques*

En un diagrama de bloques se muestran relaciones entre ellos mediante su representación como nodos y cada uno de estos nodos es una *caja negra*. A su vez una caja negra es una forma de representación del funcionamiento de todo un sistema o de un subsistema, pues en ella se indican qué transformaciones u operaciones son aplicadas a los insumos del sistema; pero no se indica cómo se transforman estos insumos sino solamente que existen tales transformaciones. En la figura 1.1 se muestra, como ejemplo, una caja negra cuya salida es un florero de barro. En ésta, el insumo o entrada es el barro y la salida del subsistema es el florero. Así mismo se muestra que la operación llevada a cabo es la de “moldear el barro”; pero nunca se especifica la forma en que hay que moldearlo.

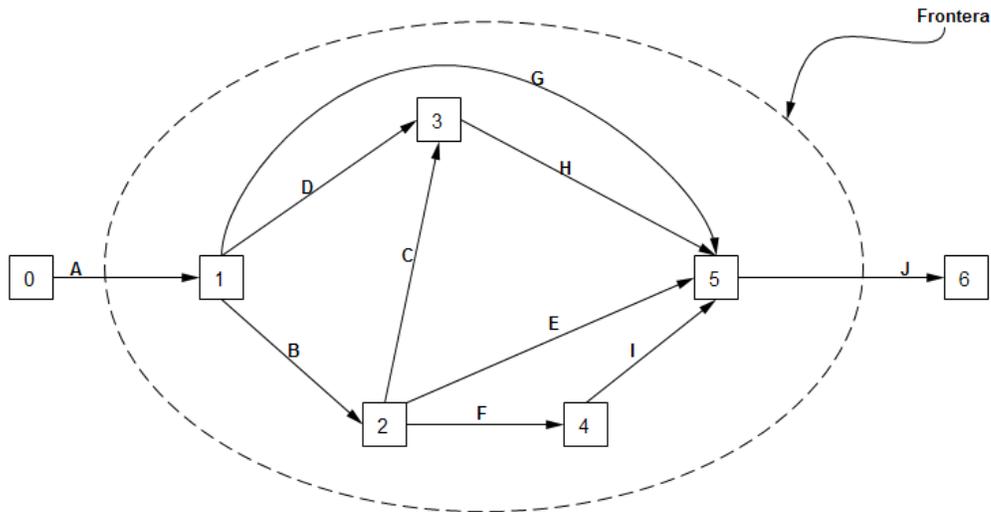


**Figura 1.1.** Ejemplo de caja negra para obtener un florero de barro. Elaboración propia, 2015.

Según (Moreno Bonett, 2010, p.24) Las cajas negras deben cumplir con lo siguiente :

1. *Estabilidad.* Se refiere a que el subsistema va a operar según lo esperado. Esta propiedad no se cumple, por ejemplo, si la caja negra opera con valores fuera de los rangos establecidos o con una combinación de insumos no considerada previamente.
2. *Independencia.* Esta propiedad hace referencia a que la función de transformación de la caja negra no cambia al conectarse con otra caja negra.
3. *Compatibilidad.* Con esta propiedad se señala que la salida de una caja negra es compatible con o sirve como entrada para otra caja negra.

Entonces, el diagrama de bloques sirve para representar cajas negras estables, independientes y compatibles. En la figura 1.2 se tiene el diagrama de bloques de un sistema. Los nodos 1 a 5 (encerrados en un cuadro) son cajas negras que forman parte del sistema y, por ejemplo, puede verse que la salida del nodo 2 (C, E o F) es el insumo de las cajas negras 3, 4 y 5. Mientras que la salida del nodo 0 (A) es el insumo de todo el sistema, la salida del nodo 5 (J) representa la salida de todo el sistema y es también la entrada de la caja negra 6; pero ésta ya no forma parte del sistema.



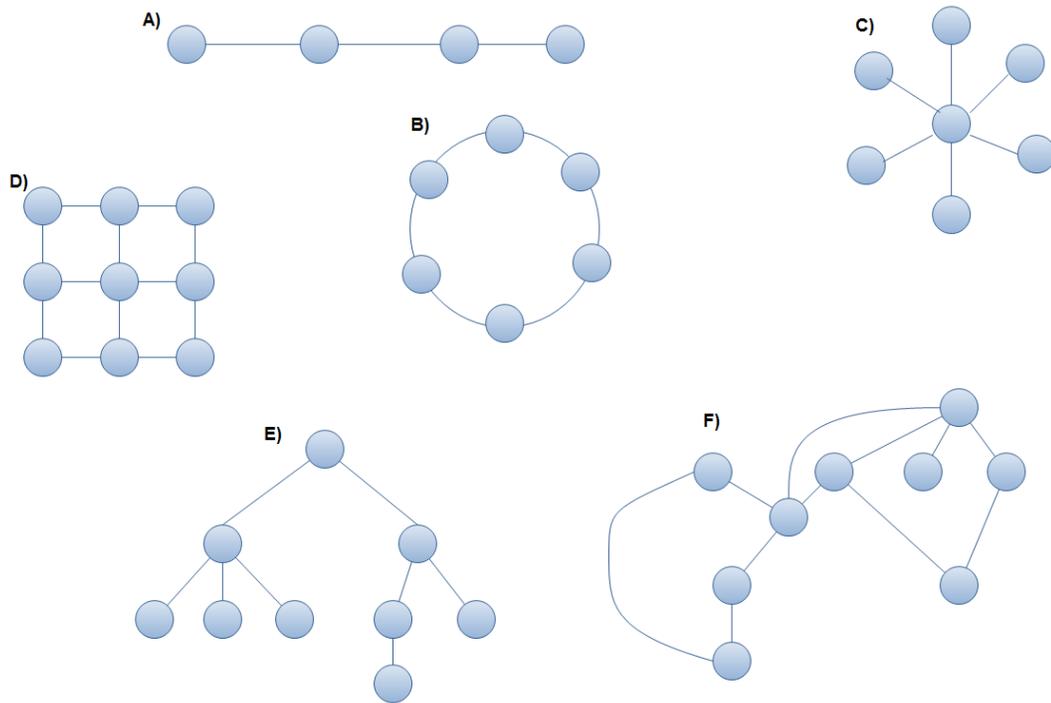
**Figura 1.2.** Diagrama de bloques (o de estructura) de un sistema. Adaptado de: (Moreno Bonett, 2010, p.25)

Además de diagramas de bloques, existen otras formas de representar la estructura de un sistema como un *grafo de flujo* o diagramas PERT. Estos dos tipos de diagramas son explicados con mayor detalle en el anexo B.

Gracias a este tipo de diagramas (de bloques, grafo de flujos y PERT) es posible identificar varios tipos de estructuras de sistemas, como:

- A) *Lineal*. Un ejemplo de éstos son los sistemas de producción en cadena.
- B) *Circular*. Los elementos se encuentran uno después del otro; pero no existe un principio o fin de la secuencia.
- C) *Centralizada*. Los elementos se encuentran unidos a uno que se denomina central.
- D) *Matricial*. Se asocia a la idea de tener varias estructuras lineales unidas.
- E) *Jerárquica*. Un ejemplo de estas estructuras son los organigramas.
- F) *Descentralizada*. En este tipo de estructuras no existen secuencias, elementos centrales o dependencia. Es conocida como estructura en red.

Estas estructuras se muestran en la figura 1.3.



**Figura 1.3.** Diferentes tipos de estructuras de sistemas. Adaptado de: Ramírez (2002, p.23-25)

### 1.1.2. Modelo DIKW

La información de tipo geográfica puede ser utilizada para la toma de decisiones o el planteamiento de políticas públicas por parte de los funcionarios del gobierno. Estas decisiones pueden influir de manera positiva o negativa en el bienestar de los ciudadanos dependiendo del sustento que se tenga (datos e información), es decir, que pueden tomarse decisiones con conocimiento y sentido común o, quizá, fuera de toda lógica.

En 1989, Russell L. Ackoff publicó *“From data to wisdom”*, texto en el que habla acerca de la cadena DIKW con la cual se explica el proceso por el cual es posible llegar de los datos a la sabiduría. De acuerdo con Rowley (2007), aunque Ackoff fue de los primeros en mencionar esta jerarquía también otros autores lo han hecho; y la mayoría de ellos identifica a los datos, la información y el conocimiento como parte de la jerarquía. No todos ellos hablan acerca de la sabiduría. Por ejemplo, esta misma autora menciona que Milan Zeleny, en su artículo *“The wisdom hierarchy: representations of the DIKW hierarchy”* publicado en 1987 pone como la cumbre de la jerarquía a la *ilustración*, mientras que para Ackoff la cumbre es precisamente la *sabiduría*. En la tabla 1.2 se muestra una comparación de las definiciones que cada uno de estos autores (Ackoff y Zeleny) hizo de los términos de la cadena DIKW.

Rowley (2007) también menciona que Bellinger, Castro y Mills en 2004 hicieron notar que el entendimiento, propuesto por Ackoff, no debería ser considerado como un nivel diferente porque

el entendimiento apoya la transición de cada etapa a la siguiente. Ellos sugieren que pasar de datos a información involucra el 'entendimiento de las relaciones', pasar de informa-

**Tabla 1.2.** Comparación de la cadena DIKW según Zeleny(1987) y Ackoff(1989). Fuente: Rowley (2007, p.167).

	<b>Zeleny</b>	<b>Ackoff</b>
<i>Datos</i>	Saber nada	Símbolos
<i>Información</i>	Saber qué	Datos que son procesados para ser útiles; da respuestas a la preguntas ¿quién?, ¿qué?, ¿dónde? y ¿cuándo?
<i>Conocimiento</i>	Saber cómo	Aplicación de datos e información; responde a la pregunta del cómo
<i>Entendimiento</i>		Apreciación del porqué
<i>Sabiduría</i>	Saber por qué	Entendimiento evaluado
<i>Ilustración</i>	Alcanzar el sentido de la verdad, el sentido de lo bueno y lo malo, y tenerlo socialmente aceptado, respetado y sancionado	

ción a conocimiento involucra el 'entendimiento de patrones', y pasar de conocimiento a sabiduría involucra el 'entendimiento de principios' (Rowley, 2007, pp.166).

Esto queda bien ilustrado en la figura 1.4, misma en la que se identifica que los procesos involucrados para pasar a los niveles más altos de la pirámide incluyen a los niveles inferiores. Los elementos de la jerarquía DIKW serán definidos en las siguientes secciones tomando como referencia lo mencionado en la obra publicada por Ackoff en 1989.

La base de la pirámide está conformada por los datos. Éstos, de acuerdo con Rowley (2007),

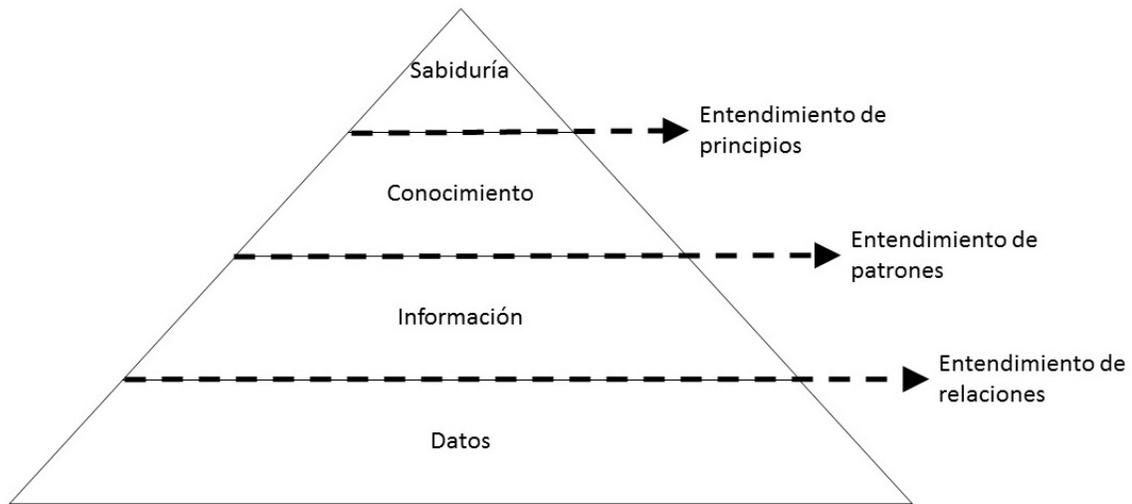
son definidos como símbolos que representan propiedades de objetos, eventos y su entorno. Son producto de la *observación*. Pero no son útiles hasta que están en una forma usable (es decir, relevante). La diferencia entre datos e información es funcional, no estructural (p.166).

Debido a que los datos son el resultados de observaciones, entonces no se deben a suposiciones. Además carecen de significado o valor pues están fuera de contexto alguno y no han sido interpretados. También tienen la cualidad de ser fáciles de comunicar, capturar y almacenar.

Respecto a la organización y el nivel de procesamiento que éstos llevan hay ciertas discrepancias, pues en el resumen hecho por Rowley (2007) se menciona que los datos están desorganizados y sin procesar, mientras que en Olvera Ramírez et al. (2014) se dice que los datos deben estar organizados y normalizados.

El segundo nivel de la pirámide está conformado por la información, la cual

está contenida en descripciones, respuestas a preguntas que comienzan con palabras como quién, qué, cuándo y cuánto. Los sistemas de información generan, almacenan, recuperan y procesan los datos. La información se infiere de los datos" (Rowley, 2007, p.166).



**Figura 1.4.** La jerarquía DIKW en forma de pirámide. Fuente: Olvera Ramírez et al. (2014)

También se dice que la información es el resultado de procesar datos con un fin específico, por lo que suele decirse también que al contrario de los datos, la información ya tiene un significado, pues ya ha sido interpretada y entendida por el receptor; pero el concepto de *significado* aplicado a la información resulta en algo subjetivo, porque dependiendo del contexto en el que es recibida y usada es el receptor quien decide si el mensaje recibido es información valiosa o si se trata de datos sin un significado particular.

Por otro lado, en el contexto de los sistemas de información y considerando también los pensamientos de quienes reciben la información, todos los datos que son captados tienen alguna estructura desde el momento en el que son recolectados o almacenados, por lo que es el significado y no la estructura lo que ayuda a diferenciar los datos de la información.

En resumen, para que la información sea considerada como tal se requiere que haya consenso de significado (Olvera Ramírez et al., 2014). Además la información es vista como datos organizados y estructurados lo que los hace valiosos, útiles, relevantes y los dota de significado.

El tercer tercer lugar de la pirámide lo ocupa el conocimiento. De acuerdo con Rowley (2007), éste implica

saber cómo, y es lo que hace posible la transformación de información a instrucciones. El conocimiento puede ser obtenido por transmisión desde otro que la tiene, por instrucción o extrayéndola de experiencias' (p.166)

Según la investigación realizada por Rowley (2007, pp.172-173) varios autores distinguen entre dos tipos de conocimiento: el conocimiento tácito y el conocimiento explícito. El conocimiento tácito es aquel inherente a la mente humana, pues se forma a través de experiencias, combinado con creencias y valores. Por otro lado, el conocimiento explícito es aquel que ya ha sido documentado en libros, artículos, reportes, etc. Esto también es útil, pues hace posible el control de los sistemas.

Pero el conocimiento no sólo es saber cómo, sino por qué, por lo que suele definirse al conocimiento con base en la información: como información combinada con entendimiento y capacidad (Rowley, 2007, p.173), porque permite al individuo la toma de decisiones con base en una creencia personal que tiene un sustento o está justificada, pues llegar a ella requirió la síntesis y entendimiento de varias fuentes de información.

El último nivel de la pirámide es la sabiduría, la cual se define como:

la habilidad de incrementar la efectividad. La sabiduría agrega valor, lo cual requiere la función mental que llamamos juicio. Los valores éticos y estéticos que esto implica son inherentes al actor y son únicos y personales (Rowley, 2007, p.166).

Se trata del nivel más alto de la jerarquía DIKW y es alcanzada después de mucho procesar datos, información y conocimiento. Según Olvera Ramírez et al. (2014) la sabiduría es la “Capacidad de discernir y juzgar lo que es correcto y lo que no lo es, para diferenciar lo que es bueno de lo que no lo es...”(p.19). Entonces, involucra un juicio ético y la creencias personales; pero la sabiduría también implica la capacidad de ver más allá, de prever. Es por ello que Olvera Ramírez et al. (2014) menciona que la sabiduría es creadora de futuro.

Ackoff era un consultor en administración y profesor de ciencias de la administración en Wharton School. Se especializaba en investigación de operaciones y teoría organizacional, por lo que de acuerdo con (Bernstein, 2009, p.68-69), para plantear este modelo piramidal, Ackoff se basó sólo en modos específicos de datos, información, conocimiento y sabiduría, pues éstos estaban más enfocados en la administración de organizaciones. De manera casi incidental, estos conceptos resultaron de gran valor para la bibliotecología y ciencias de la información.

Pero así como Ackoff planteó todo un proceso a través del cual es posible generar conocimiento a partir de la obtención de datos crudos, Bernstein (2009), comenzó por plantearse si existe lo opuesto a conocimiento y si de manera análoga al conocimiento también existe una serie de etapas que rodean a este opuesto al conocimiento y así llegó a identificar la contraparte de cada uno de los elementos del modelo DIKW. El resumen de su planteamiento se muestra en la tabla 1.3.

**Tabla 1.3.** Contraparte de los elementos del modelo DIKW. Fuente: Bernstein (2009).

Elemento DIKW	Contraparte
<i>Datos</i>	<b>Ausencia o necesidad de datos; falta de datos.</b> Para que los datos sean datos, éstos deben ser verdaderos. Si son erróneos, no son datos
<i>Información</i>	<b>Error y los términos en inglés <i>misinformation</i> y <i>disinformation</i></b> Un error es una equivocación cometida de manera desapercibida o por ignorancia. El término <i>misinformation</i> se refiere a mala información o falsos supuestos. <i>Disinformation</i> se refiere a la difusión de falsa información con el propósito de influir en la opinión de aquellos que la reciben.
<i>Conocimiento</i>	<b>Ignorancia</b>
<i>Sabiduría</i>	<b>Insensatez; estupidez. (<i>Folly</i> y <i>stupidity</i>, en inglés.)</b> La estupidez implica la falta de inteligencia, de capacidad mental así como falta de sabiduría, aunque también está relacionada a la no adaptación al cambio o a respuestas drásticas y radicales no informadas (por datos). La insensatez tiene las connotación de falacia y malentendido que resultan en algo devastador. La insensatez proviene del error y lo exagera.

Para Bernstein (2009), estos opuestos a los términos del modelo de Ackoff no forman una pirámide, pues para él, no siempre existe una “cadena” que lleve de la falta de datos a la insensatez

o estupidez, ya que los errores se pueden dar incluso contando con los datos o la información correctos ya sea por una mala interpretación o un razonamiento no lógico, mientras que el opuesto a la información (*misinformation* o *disinformation*) no necesariamente se dan por la falta de datos, sino que también pueden darse por distorsionar o comunicar de forma incorrecta la información o el conocimiento, incluyendo la información y el conocimiento que está bien entendido de origen.

Aunque Olvera Ramírez et al. (2014) plantean esta antítesis de la jerarquía DIKW como una bipirámide (ver figura 1.5) en la que, para estar en alguno de sus niveles, es necesario el uso de algún nivel de inteligencia; también plantean que en realidad este modelo no es tan sencillo debido a dos factores: el primero radica en la habilidad que se tiene para producir información relevante (no sólo datos derivados) a partir de los datos así como la habilidad para interpretar la información y producir conocimiento ligado al desconocimiento de la naturaleza de los datos como de la información. El segundo factor es “el grado de experiencia, el nivel de abstracción y la capacidad de entendimiento que puedan tener quienes son usuarios de datos, información y conocimiento”(Olvera Ramírez et al., 2014, p.22).

Al inicio de esta sección se mencionó la importancia de la información geográfica para el bienestar de los ciudadanos. Una forma de evitar la toma de decisiones que carezcan de sabiduría, es que los datos e información (de tipo geográfica, en este caso) que se emplean para apoyar esta toma de decisiones estén debidamente documentados, gracias a lo cual pueden reducirse las malinterpretaciones, pues se sabe lo que representan los datos y, por tanto, es posible dárselos un buen uso.

La forma de documentar los datos de esta índole, es a través del empleo de metadatos geográficos, para los cuales existen diferentes estándares (mismos que se abordan más adelante en este mismo capítulo). Además existen diferentes aplicaciones que facilitan la captura de estos metadatos. En el presente trabajo se plantea otra solución, además de las ya existentes, para la captura de metadatos geográficos.

### 1.1.3. Capas de datos geográficos

Actualmente el uso de Sistemas de Información Geográfica (SIG) se ha vuelto muy importante; pues son herramientas que ayudan en la gestión de datos y la toma de decisiones, principalmente en temas como administración territorial o medio ambiente (por poner sólo algunos ejemplos) ya que permiten procesar y analizar grandes cantidades de datos relacionados con aspectos territoriales a la vez que muestran a los usuarios una representación efectiva del territorio que sería muy difícil o imposible de conseguir con otros medios. Si se toma en cuenta lo mencionado en una de las secciones anteriores, entonces los SIG sirven para la representación de *sistemas espaciales*. (Las características de los sistemas espaciales se encuentran en la tabla 1.1)

Considerando lo anterior, Krzanowski y Raper (2001, p.66) define los siguientes conceptos:

- A) *Fenómeno espacial*. Es un fenómeno del mundo real definido en el tiempo y en el espacio, es decir, tiene un atributo de posición.
- B) *Entidad espacial*. Es el fenómeno espacial más pequeño que puede ser representado de forma discreta

En este contexto, las capas de datos geográficos son la abstracción de fenómenos y entidades espaciales, los cuales al ser representados en sistemas digitales, suelen llamarse objetos espaciales. Estas capas de datos geográficas son el insumo de los sistemas de información geográfica para empezar a realizar análisis de los datos para obtener información. Aunque en la actualidad existe un gran número de formatos digitales para codificar las capas de datos geográficos, éstos se dividen en dos grandes grupos: capas de datos vectoriales y capas de datos ráster.

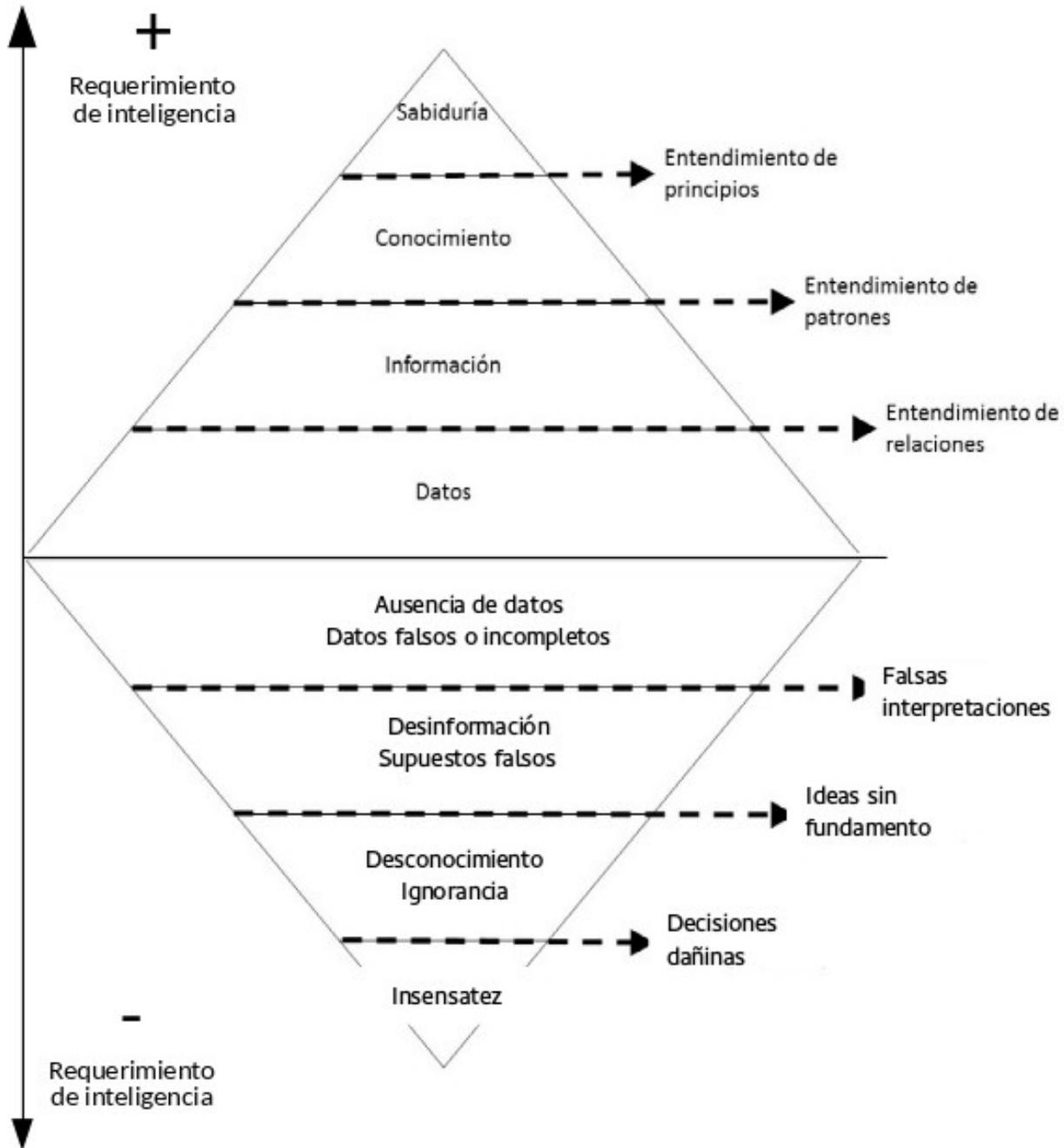


Figura 1.5. La jerarquía DIKW en forma de bipirámide. Fuente: Olvera Ramírez et al. (2014)

En las capas de datos vectoriales los objetos espaciales se clasifican de la siguiente manera (Decker (2001, p.12) y (Krzanowski y Raper, 2001, p.67)):

- **Puntos (dimensión 0)**: representan una posición  $x,y$  o longitud,latitud a la cual se vincula otra información. Debido a que esta posición está almacenada como un atributo, normalmente se les denomina elementos de dimensión 0.
- **Líneas (dimensión 1)**: se representan como líneas rectas que unen un par de puntos, las cuales se denominan polilíneas cuando se conforman de múltiples segmentos. Todas las líneas se miden en términos de su longitud, por lo tanto son elementos de dimensión 1.
- **Polígonos (dimensión 2)** son regiones bidimensionales cerradas que pueden contener o ser intersectadas por puntos y líneas. Se representan como polilíneas que se cierran en sí mismas.

También existen las capas de datos ráster. En éstas, el manejo de la geometría es más sencillo que en las vectoriales pues suelen utilizarse “mosaicos” con celdas de forma cuadrada, principalmente, en el que cada uno de ellos tiene un valor asignado. Cada uno de estos valores puede ser el resultado de una observación o un valor registrado por un sensor. Estos valores pueden ser manejados por la computadora como un arreglo de datos. Las coordenadas  $-x,y$  o *longitud,latitud* de cada celda o pixel (si se trata de una imagen) se pueden calcular a partir de saber el tamaño del arreglo y el tamaño que representan los pixeles en el mundo real así como las coordenadas de una de las esquinas del ráster.

Todos los objetos espaciales tienen dos componentes (Gomasasca, 2009, p.484):

- *posicional*, el cual define gráfica y geoméricamente la posición, la forma y la topología de los objetos representados por primitivos geométricos como puntos, líneas, polígonos y pixeles (por ejemplo: fronteras, ríos, lagos, etc.). Esta componente puede ser descrita por medio de atributos alfanuméricos (como sus topónimos, nombre de la unidad administrativa o direcciones); pero principalmente por sus coordenadas geográficas obtenidas de un sistema de referencia.
- *descriptiva*, son declaraciones alfanuméricas que describen propiedades no espaciales de los objetos espaciales, es decir, sus atributos. En sistemas digitales, estos atributos pueden ser numéricos, alfanuméricos o fechas (por ejemplo: la población de un país, los topónimos, etc.)

Dos conceptos que hay que tener muy presentes cuando se utilicen capas de datos geográficos son la escala y la exactitud (Decker, 2001, pp.16-19). La escala se refiere a la proporción entre las unidades medidas en el mapa y a las unidades medidas sobre la superficie de la Tierra. Mientras más pequeña es esta proporción más grande es la escala, por ejemplo: una escala 1:2,500 es mayor que una escala 1:250,000. La escala está asociada con el nivel de exactitud o detalle de los datos. Las escalas grandes suelen tener mayores niveles de exactitud que las escalas pequeñas. La exactitud se refiere a cuán cerca están los objetos de representar su posición real. Una forma de determinar la exactitud es obtener la diferencia entre las coordenadas que un objeto tiene en un mapa con sus coordenadas reales. En la gran mayoría de los casos, hay un desfase en los valores de las coordenadas; pero la existencia de estos errores no significa que los datos son inútiles. Para ello debe establecerse un valor de la tolerancia para este tipo de errores, el cual se determina considerando el uso y/o tipo de análisis que vaya a realizarse con los datos además de la escala de los mismos.

La calidad de un producto cualquiera está relacionada con el grado en que dicho producto cubre o no las necesidades del usuario. En el caso de los datos espaciales, la forma de determinar su calidad es a través de la correcta documentación, es decir, de la correcta elaboración de sus metadatos. En

ellos, debe quedar establecido el propósito y el proceso por el cual fueron elaborados los datos, esto incluye la escala cartográfica de la fuente de los datos, además de la precisión manejada para los datos resultantes. Pero no sólo estos dos elementos se deben documentar en los metadatos, también debe especificarse si los datos son de tipo ráster o vectoriales, su cobertura geográfica y en el caso de los datos vectoriales también debe especificarse los atributos representados en la capa. Además de éstos, también hay otros elementos de las capas de datos geográficos que deben representarse; pero éstos se mencionan en el apartado de metadatos geográficos.

### 1.1.3.1. Formatos para datos geográficos

El formato de las capas de datos geográficos es algo que también debe documentarse en los metadatos, pues está relacionado con la forma en que se distribuyen los datos, además de ser un elemento que permite a los usuarios decidir con qué software se va a realizar el análisis de los mismos o si requerirán hacer conversiones de formato e incluso es un elemento que determina la calidad de los mismos. Es por ello que tanto este, como el apartado 1.1.3.2 están relacionados con los formatos para capas de datos geográficos.

Como se mencionó en un párrafo anterior existe una gran cantidad de formatos de representación para las capas de datos geográficos, ya sean de tipo ráster o de tipo vectorial, algunos de los cuales se muestran en la tabla 1.4.

**Tabla 1.4.** Lista de formatos ráster y vectoriales para capas de datos geográficos. Fuente: (Sherman, 2008, p.91-92)

Formatos vectoriales	Formatos Ráster
Shapefiles (.shp) de ESRI	GeoTIFF (.tif).
Formato de intercambio de GPS (.gpx)	TIFF con world files (.tif, .tfw).
SDTS para datos vectoriales	SDTS para datos ráster.
Geography Markup Language (.gml).	Erdas Imagine(.img).
Keyhole Markup Language (.kml).	Portable Network Graphics con world file (.png, pgw).
GeoJSON.	Modelos digitales de elevación (.dem).
Formato vectorial de GRASS.	net CDF.

Hoy día muchos formatos son libres, por lo que estos formatos pueden ser utilizados en muchos SIG de escritorio, principalmente en aquellos que son de código fuente abierto<sup>2</sup> como QGIS, gvSIG, OpenJUMP, entre otros. Algunos otros formatos son cerrados<sup>3</sup>, lo que obliga a los usuarios a usar el software o el API<sup>4</sup> del propietario del formato. Para el caso de aquellos formatos que son abiertos existe una biblioteca de conversión de formatos, de código abierto, que viene integrada al instalar alguno de los software de SIG previamente mencionados, aunque también puede instalarse de manera independiente. Esta biblioteca es GDAL/OGR (Geospatial Data Abstraction Library) y a la fecha da soporte para más de 200 formatos.

Antes de la versión 2.0 solían ser dos bibliotecas: la primera, GDAL, contenía funciones para dar soporte y convertir a formatos ráster, mientras que OGR para dar soporte y conversión para

<sup>2</sup>(*Open Source*, en inglés) se refiere a aquellos programas cuyo código fuente se encuentra disponible para todo público, para ser estudiado y/o modificado y/o redistribuido.

<sup>3</sup>contrariamente a los programas *Open Source*, se refiere a aquellos programas cuyo código fuente no es puesto a disposición del público

<sup>4</sup>Application Programming Interface por sus siglas en inglés. Se trata de un conjunto de funciones que conforman una biblioteca y que permite a dos software comunicarse entre sí.

datos de tipo vectorial. Ahora GDAL da soporte para 142 formatos y OGR, a 84. A partir de la versión 2.0 ambas bibliotecas han sido integradas en una sola (GDAL, 2015). Este proyecto fue iniciado y dirigido por Frank Warmerdan y en Febrero de 2006 se volvió un proyecto de la Open Source Geospatial Foundation (OSGeo). A través de su página de internet se puede acceder a la lista completa de los formatos que soporta, de los software que utilizan esta biblioteca y de todo lo relacionado con el proyecto.

### 1.1.3.2. Interoperabilidad

“Con la *interoperabilidad* se pretende la capacidad o posibilidad de acceder a diferentes sistemas y datos espaciales, también organizados con diferentes herramientas de software, reconociendo y adaptando diferentes formatos, modelos espaciales, procesamientos, aplicaciones sin modificar su contenido ni su calidad” (Gomasasca, 2009, p.546).

Gracias a la necesidad que existe por compartir y usar la información entre diferentes actores, incluso a nivel internacional, se vuelve imprescindible normalizarla; pero éste no es el único requisito para agilizar el acceso a la misma, pues de igual forma resulta necesario tenerla en un formato neutro, es decir, que su uso sea independiente de cualquier software o sistema, tecnológicamente hablando; por lo que la existencia de un marco normativo se vuelve una condición necesaria para ayudar a lograr la interoperabilidad de la información geográfica.

Actualmente existen dos grandes organismos internacionales que se encargan de crear estándares para formatos y servicios de información geográfica. El primero es la **I**nternational **O**rganization for **S**tandardization (ISO) y el segundo es el **O**pen **G**eospatial **C**onsortium (OGC). ISO fue creada en 1947 y en 1994 surgió el Comité Técnico para Información Geográfica y Geomática, es decir, el comité técnico 211 (ISO/TC 211 *Geographic Information/Geomatics*). Por otro lado, el OGC fue fundado en 1994 y es una organización sin fines de lucro conformada por miembros comerciales, gubernamentales, académicos y empresas de los sectores público y privado.

Según Erba et al. (2012, p.231), los principales objetivos del Comité Técnico 211 de ISO son:

- Incrementar la comprensión y uso, así como la disponibilidad, acceso, integración y distribución de la información geográfica;
- Promover el uso eficiente, eficaz y económico de la información geográfica digital y los sistemas de hardware y software relacionados;
- Contribuir a un enfoque unificado para solucionar los problemas ecológicos y humanitarios globales.

Estos objetivos quedan de manifiesto en el conjunto de normas creadas por este comité técnico, conocido como la familia de normas ISO 19100, que sirven para facilitar la interoperabilidad de datos geoespaciales. Actualmente se compone de 68 miembros, de los cuales 39 son miembros participantes<sup>5</sup> y 29 son miembros observadores<sup>6</sup>. Entre estos miembros se encuentran países como Australia, Canadá, Chile, Estados Unidos, etc. (como miembros principales) y países como Colombia, Cuba, Grecia, etc. (como miembros observadores).

El OGC es una organización internacional encargada de crear normas con el objetivo de facilitar la interoperabilidad para datos geográficos así como de servicios geoespaciales y basados en localización.

---

<sup>5</sup>Se refiere a los miembros que participan activamente votando acerca de los estándares en varias etapas de su desarrollo.

<sup>6</sup>Son los miembros que pueden observar los estándares que están siendo desarrollados, ofreciendo comentarios y consejos.

En el año de 1999 se estableció un acuerdo de colaboración entre ambas organizaciones con el fin de evitar la duplicidad de esfuerzos y de acuerdo con Erba et al. (2012),

“En el marco de este acuerdo el OGC ha adoptado varias normas ISO/TC 211 como especificaciones abstractas en las que basa sus especificaciones de implementación. De forma recíproca, algunos estándares desarrollados originalmente por OGC se han llevado al ISO/TC 211, para publicarse finalmente como normas ISO internacionales”(p.232).

Respecto a las capas de datos geográficos, el OGC definió un formato de datos neutral, es decir, que no está ligado a ningún software de SIG en específico, este formato es el GML. Dicho formato está basado en el estándar XML(*eXtensible Markup Language*) para facilitar el que los productores compartan sus datos y que los usuarios accedan más fácilmente a ellos, también vía Web. Es importante aclarar que GML no sólo es un lenguaje para la representación de los datos, sino para la geometría, sus características y atributos.

#### 1.1.4. Metadatos geográficos

Para poder comprender qué son los metadatos geográficos, cuáles son sus elementos y comprender su utilidad, resulta imprescindible comenzar con la definición de *metadato*. En general, los metadatos suelen definirse como datos acerca de los datos, es decir, son datos que sirven para definir las características de un objeto o recurso.

Olvera Ramírez et al. (2014) mencionan que los esquemas de metadatos surgen en el dominio de las bibliotecas. Estos esquemas sirven para describir objetos o recursos de las bibliotecas; pero también ayudan en la localización, identificación y consulta de estos mismos materiales. Pensando en este mismo ejemplo de la biblioteca, las fichas bibliográficas constituyen los metadatos de los libros y/o revistas disponibles en una biblioteca.

Haciendo una analogía entre una biblioteca y una IDE, los datos geográficos constituirían el equivalente de los libros y revistas de la biblioteca y para poder localizarlos y consultarlos resulta necesario el uso de los metadatos; pero debido al origen y naturaleza de estos datos, se requieren metadatos muy específicos que permitan la búsqueda de estos recursos dentro de algún repositorio bajo criterios muy concretos como: el título del recurso, su cobertura geográfica, categoría a la que pertenecen los datos, palabras clave, nombre de la institución que genera el recurso, etc.

De acuerdo con Olvera Ramírez et al. (2014, p.96) el uso de metadatos geográficos ofrece las siguientes ventajas:

- Aumentan la oportunidad con que los datos se publican y se dan a conocer.
- La calidad de los datos puede declararse y caracterizarse.
- Se declara el objetivo de la creación de los datos y su uso pretendido.
- Se proporciona una descripción precisa y discreta de los datos.
- Los metadatos pueden compartirse sin que implique el suministro de los datos.
- Se proporciona una descripción del proceso de generación de los datos y su linaje.
- Inducen a normalizar los procesos de trabajo y a la creación de una cultura de documentación organizacional.

Al igual que con las capas de datos geográficos, existen distintos organismos que se encargan de crear normas en las que especifican los elementos que deben incluir los metadatos geográficos. Sin importar la norma que se ocupe para la documentación de datos espaciales, la mayoría de ellas comparten un conjunto de elementos núcleo, lo que facilita la interoperabilidad. Según Nogueras-Iso y Muro-Medrano, Pedro R. Zaragoza-Soria (2005, p.15), algunos de los elementos que debe contener este núcleo común son:

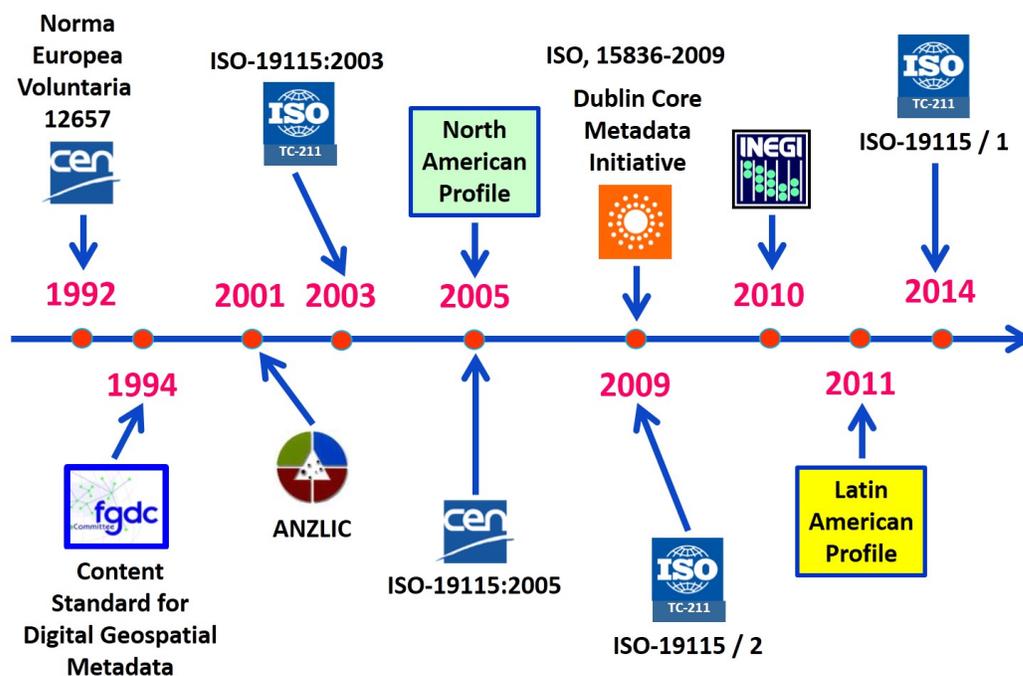
- Sección de referencia que provea información de la autoría del metadato, fecha de creación y última actualización, nombre y versión del estándar de metadatos o el lenguaje en que se captura el metadato.
- Sección con información de identificación. Ésta contiene datos como: título del conjunto de datos, quien o quienes originan los datos, área geográfica que abarca el conjunto de datos, palabras claves, datos de acceso y restricción de los datos o el propósito de la creación de los mismos.
- Una sección de calidad de datos, en la que se expresen las especificaciones de calidad que requiere el formato de metadatos, detalles de exactitud posicional y de atributos o en la que se expliquen las fuentes y el proceso que dieron origen al conjunto de datos.
- Secciones que contengan información acerca del modelo de datos usado para codificar los datos espaciales (vectorial o ráster) u otros posibles métodos de georreferencia indirecta (por ejemplo: geocodificación) e información acerca del sistema de referencia espacial empleado (datum, elipsoide, proyección cartográfica).
- Información acerca de las entidades y atributos de las entidades del conjunto de datos espaciales así como una explicación acerca de la codificación del dominio de un atributo (si es necesario).
- Información acerca de la distribución, por ejemplo: quién distribuye los datos, los formatos usados, disponibilidad y precio de los datos, etc.

La creación de metadatos para documentar información geográfica se remonta a la década de los años 90 (Ver figura 1.6). En las siguientes secciones se hará mención de dos organismos dedicados a la creación de estándares para información geográfica, así como de las normas que se crearon para metadatos geográficos.

#### 1.1.4.1. Federal Geographic Data Committee (FGDC)

La declaración para la creación de este organismo fue publicada en la circular A-16 el 19 de octubre de 1990 por medio de la Oficina de Administración y Presupuesto del gobierno de los Estados Unidos. Este comité, a cargo del Servicio Geológico de los Estados Unidos (USGS) promueve el desarrollo, uso, compartición y distribución de datos geoespaciales, provenientes de autoridades federales, estatales, locales y del sector privado. A este esfuerzo de publicación de datos se le conoce como Infraestructura Nacional de Datos Espaciales (NSDI, por sus siglas en inglés).

Fue este comité el que creó la primera norma para documentar datos geográficos. La primera versión fue publicada el 1994. En 1998 se publicó la segunda versión, la cual es la norma federal para metadatos geográficos vigente en Estados Unidos. El nombre de esta norma es *Content Standard for Digital Geospatial Metadata* (CSDGM) o *Norma de Contenido para Metadatos Geoespaciales Digitales*, en español.



**Figura 1.6.** Línea del tiempo de las principales normas de metadatos geográficos a nivel internacional. Obtenida de: Olvera Ramírez (2014)

La CSDGM se compone de 11 secciones. Las primeras 8 secciones (de la 0 a la 7) son las secciones principales, mientras que las tres restantes (de la 8 a la 10) son de soporte por lo que están incluidas en más de una sección. En estas últimas secciones se debe incluir la información para citas, de periodo de tiempo y de contacto. En la tabla 1.5 se hace un resumen del contenido de cada una de las secciones principales de la norma CSDGM.

**Tabla 1.5.** Resumen de las secciones de la norma CSDGM del FGDC. Fuente: Elaboración propia con datos del Federal Geographic Data Committee (1998)

Sección	Contenido
0. Metadatos	Datos acerca del contenido, calidad, condición y otras características de los datos. Esta sección constituye el punto de partida y contiene las secciones 1 a la 7.
1. Información de identificación	Información básica acerca del conjunto de datos. Debe contener datos como: descripción y propósito, fecha a la cual corresponden los datos, el estado del conjunto de datos, frecuencia de mantenimiento y actualización, coordenadas extremas, tema, palabras clave, restricciones de acceso y de uso.

Tabla 1.5 (continuación)

Sección	Contenido
2. Información de calidad de datos	Información acerca de la calidad del conjunto de datos. Debe contener un reporte de la exactitud de los atributos (evaluación de la exactitud de la identificación de entidades y asignación de valores a los atributos), de la consistencia lógica, de completitud, de exactitud posicional. Además debe incluirse el linaje de los datos y el porcentaje de la cobertura de nubes (si aplica).
3. Información de organización de datos espaciales	Debe contener datos acerca de los mecanismos usados para representar información espacial en el conjunto de datos (puntos, vectores o ráster. También deben incluirse las referencias espaciales indirectas, si se usaron).
4. Información de referencia espacial.	Esta sección debe contener datos acerca del marco de referencia espacial del conjunto de datos como: la definición del sistema de coordenadas horizontal (ej. Coordenadas geográficas o proyección cartográfica, datum, elipsoide, etc.), definición del sistema de coordenadas vertical
5. Información de entidades y atributos	Esta sección servirá para explicar el significado de las entidades, atributos y el valor de los atributos del conjunto de datos.
6. Información de distribución.	Esta sección debe contener información acerca del distribuidor y opciones para obtener el conjunto de datos así como el tiempo que tarde en llegar la información solicitada.
7. Información de referencia de metadatos.	Esta sección debe contener datos como: fecha de creación y actualización del metadato, datos de contacto para el metadato, restricciones de uso y acceso al metadato, nombre y versión del estándar del metadato, así como datos acerca de la seguridad del metadato.

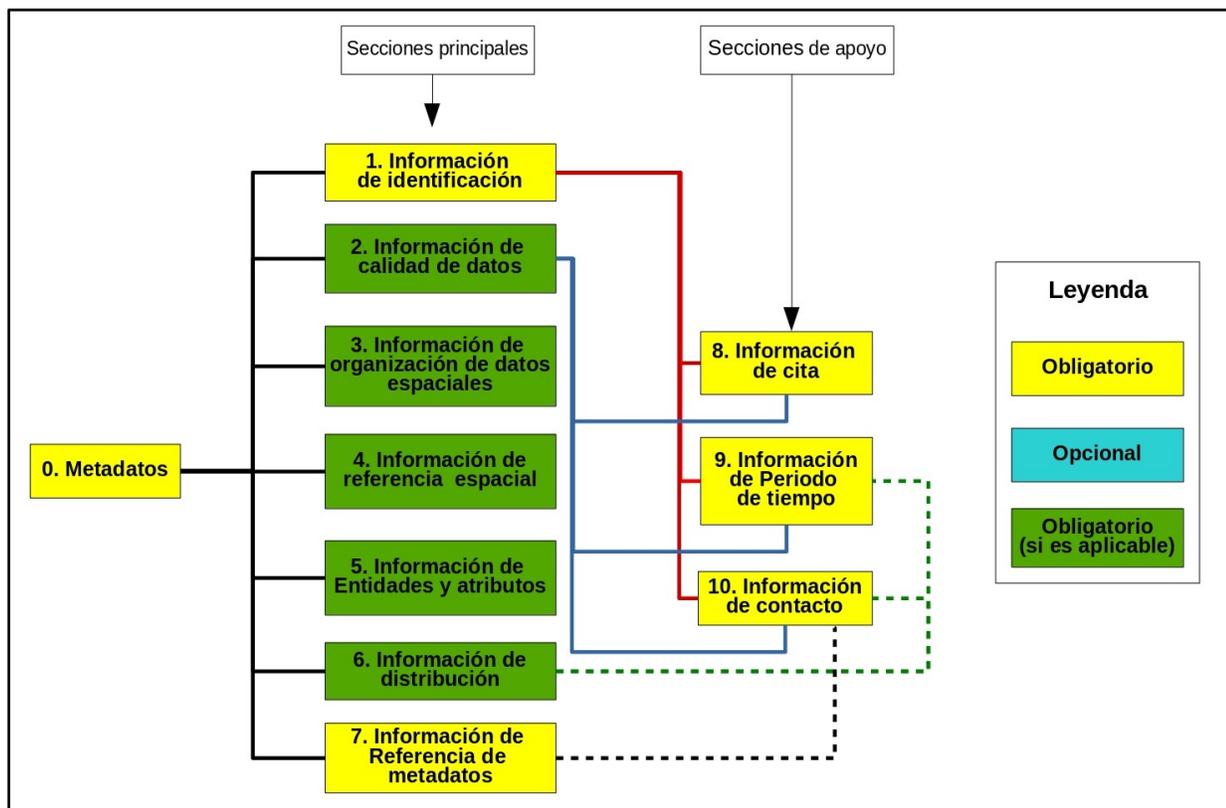
En 1994 Susan Stitt de la USGS Biological Resources Division preparó una serie de diagramas con el fin de facilitar el entendimiento de la estructura, así como el llenado de los elementos de esta norma. En la figura 1.7 se muestra un ejemplo de estos diagramas y el resto de ellos puede encontrarse en Federal Geographic Data Committee (2000).

#### 1.1.4.2. International Organization for Standardization (ISO)

En la sección de interoperabilidad se hizo mención acerca de esta organización internacional y de su comité técnico 211 para información geográfica y geomática, así como de las funciones principales de este comité y la familia de normas 19100.

En mayo de 2003 este comité publicó el estándar ISO 19115:2003 y para su creación se tomaron en cuenta otros estándares, como el CSDGM del FGDC o la prenorma 12657 del European Committee for Standardization (CEN). Este estándar define los elementos que lo componen así como la condicionalidad y relación entre estos elementos.

El estándar ISO 19115 define el esquema para la descripción de recursos de datos espaciales y servicios relacionados. Se aplica para catalogar y describir por completo los metadatos. Las reglas incluidas en este estándar pueden ser usadas para describir datos espaciales, series de datos espaciales y objetos individuales y sus atributos (Litwin y Rossa, 2011, p.70).



**Figura 1.7.** Diagrama de las secciones de la norma CSDGM. Adaptado de: Federal Geographic Data Committee (2000) y Olvera Ramírez et al. (2014)

La norma ISO 19115:2003 se compone de 387 elementos, de los cuales, 22 son elementos núcleo, lo que significa que son altamente recomendables, no así obligatorios. (Ver tabla 1.6).

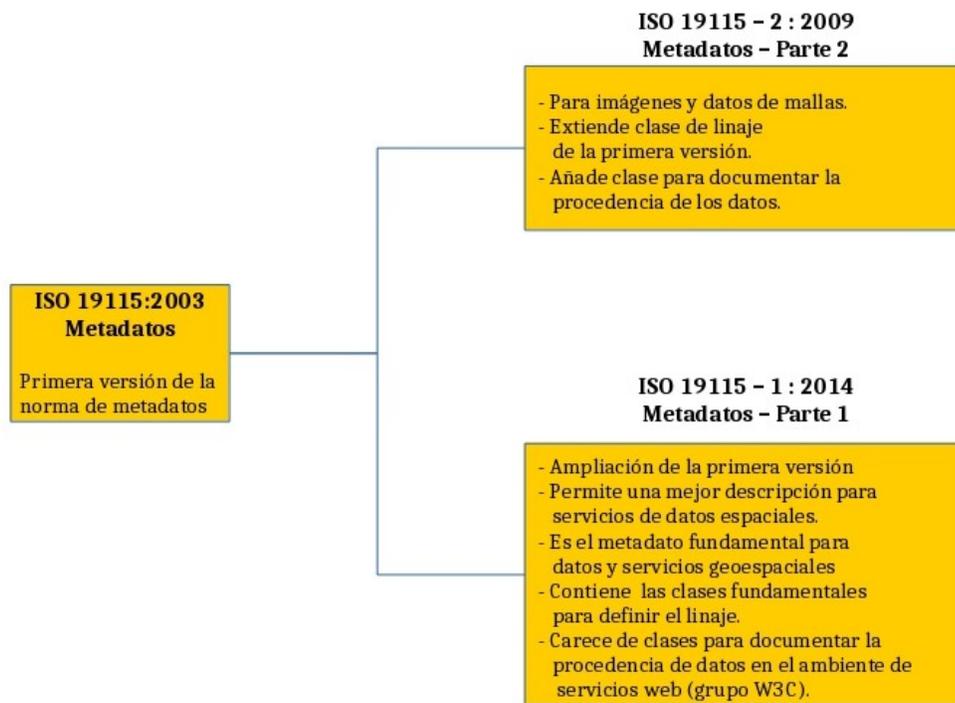
**Tabla 1.6.** Elemento núcleo de la norma ISO 19115. **O** = elemento obligatorio, **C** = elemento condicional, **Opc.** = elemento opcional. Fuente: Olvera Ramírez et al. (2014)

Título del conjunto de datos espaciales (O)	Tipo de representación espacial (Opc.)
Fecha de referencia del conjunto de datos (O)	Sistema de referencia(Opc.)
Parte responsable del conjunto de datos (Opc.)	Linaje (Opc.)
Localización geográfica del conjunto de datos (por cuatro coordenadas o por identificador geográfico) (C)	Recurso en línea (Opc.)
Idioma del conjunto de datos espaciales (O).	Identificador del archivo de metadatos (Opc.)
Conjunto de caracteres del conjunto de datos (C)	Nombre de la norma de metadatos (Opc.)
Categoría del tema del conjunto de datos (O)	Versión de la norma de metadatos (Opc.)
Resolución espacial del conjunto de datos (Opc.)	Idioma de los metadatos (C)

Tabla 1.6 (continuación)

Resumen descriptivo del conjunto de datos (O)	Conjunto de caracteres de los metadatos (C)
Formato de distribución (Opc.)	Punto de contacto para los metadatos (O)
Información adicional de la extensión del conjunto de datos espaciales (vertical y temporal) (Opc.)	Fecha de creación de los metadatos (C)

Actualmente la norma ISO 19115 se encuentra dividida en dos partes (Ver figura 1.8). Estas dos partes surgieron como revisión a la primera norma publicada en 2003 y desde noviembre de 2014 se está trabajando en una revisión de la segunda parte de la norma para homogeneizar la descripción de la procedencia con el modelo descrito en el estándar W3C Prov standard (para la web).



**Figura 1.8.** Diagrama de la Norma ISO 19115 y sus dos partes. Fuente: elaboración propia con datos de Di y Deng (2014)

En la norma ISO 19115 no se brinda ninguna especificación de implementación, es decir, de cómo organizar el contenido de la norma en un registro formal. Debido a esto en 2007 el comité técnico 211 publicó la norma ISO 19139:2007 *Metadatos: Esquema de Implementación XML*, en la que se define la forma de codificar los metadatos creados de acuerdo con la norma ISO 19115 en notación XML, a través del *Esquema de Definición XML* (XSD, por sus siglas en inglés) del consorcio W3C. De igual forma especifica el único modelo UML<sup>7</sup> apropiado para la interpretación del modelo abstracto de la ISO 19115. Es importante mencionar que este modelo de implementación UML se realizó conforme

<sup>7</sup>Unified Modeling Language (UML) o Lenguaje Unificado de Modelado (en español). Para mayor referencia respecto a este lenguaje, ver la sección de *Casos de uso en la página 71*

a los establecido en la norma ISO 19103:2005 *Lenguaje de esquema Conceptual*. El estándar ISO 19103 “Especifica el Perfil UML para información espacial. Provee las reglas y las pautas para el uso de diagramas conceptuales UML en los estándares ISO para información geoespacial” (Litwin y Rossa, 2011, p.52).

En el mundo existen diferentes grupos que han trabajado en iniciativas para crear sus propias normas de metadatos para información geoespacial, las cuales son aplicables o cumplen con las exigencias de grupos o países de determinadas regiones geográficas. Algunas de estas normas están basadas en la especificación ISO 19115 a las que además, se les agregan otros elementos que pueden provenir de alguna otra norma. A este tipo de normas se les denomina *Perfil de metadatos*. Entonces un perfil puede definirse como un conjunto de elementos de un estándar de metadatos al que se le añaden otros elementos no incluidos en la norma base, por lo que un perfil es “una adaptación de una norma para una comunidad específica de usuarios...” (Olvera Ramírez et al., 2014).

En la tabla 1.7 se muestran algunos ejemplos de perfiles de la norma ISO 19115, entre los cuales se encuentra la propia norma de metadatos del INEGI.

**Tabla 1.7.** Ejemplos de perfiles de la norma ISO 19115

Perfil	Descripción
Perfil Norteamericano de Metadatos (NAP)	Resultado de esfuerzos independientes por parte de Estados Unidos y Canadá, iniciados en 2005. Adopta los 22 elementos núcleo de la ISO 19115. Se adquiere mediante pago.
Perfil Latinoamericano de Metadatos (LAMP)	Su desarrollo comenzó desde el año 2006 por parte del Instituto Panamericano de Geografía e Historia (IPGH) en conjunto con el Grupo Consultivo del ISO/TC 211), el Instituto Geográfico Agustín Codazzi (IGAC), de Colombia, Programa GeoSur y con el auspicio de la CAF (Corporación Andina de Fomento). Publicado en 2011 por IPGH e IGAC. Se compone de 8 secciones principales y 3 secciones de soporte, además de anexos con especificaciones para crear perfiles y extensiones. Está disponible sin costo
Perfil de Metadatos de la Infraestructure for Spatial Information in Europe (INSPIRE)	Se define en dos documentos. El primer documento define el conjunto de elementos del perfil y el segundo es la documentación del perfil de metadatos de INSPIRE. Este segundo documento está basado en las normas ISO 19115 e ISO 19119 (de Servicios). El desarrollo de este perfil está basado en los 22 elementos núcleo de la ISO 19115 además de contar con las mismas secciones recomendadas, por la misma norma base, para la creación de perfiles, excepto la información acerca del sistema de referencia.
Norma Técnica para la elaboración de Metadatos Geográficos (NTM).	Esta norma fue publicada por el INEGI (Instituto Nacional de Estadística y Geografía) el 24 de diciembre de 2010. Consta de 149 elementos, de los cuales 25 son obligatorios y el resto son opcionales. En esta norma se ocupan los elementos núcleo de la norma ISO 19115:2003, así como algunos otros del CSDGM (o FGDC-STD-001-1998).

Resulta obvio que no sólo existen metadatos para documentar información de índole geográfica, tomando en cuenta el contexto en el que surge el concepto de metadato, sino que existen otros

esquemas de metadatos para documentar otra clase de recursos. Uno de estos esquemas se encuentra a cargo de la **Dublin Core Metadada Initiative** (DCMI), creada en 1995 en Dublin, Ohio y se trata de una “organización dedicada a promover la adopción generalizada de estándares de metadatos interoperables y el desarrollo de vocabularios de metadatos especializados para describir recursos que permiten sistemas de descubrimiento de información más inteligentes”(Dublin Core Metadata Initiative, 2010).

El estándar Dublin Core actualmente es una norma internacional utilizada para documentar toda clase de recursos como: archivos, servicios, publicaciones, programas, páginas web, etc. Consta de 15 elementos y puede ser codificado en formato XML basado en el Marco de Descripción de Recursos (RDF, *Resource Description Framework*).

### 1.1.5. Infraestructuras de datos espaciales

La importancia de los datos e información geográfica radica principalmente en que “cualquier actividad humana o fenómeno natural, puede estudiarse o preverse mediante la cartografía adecuada” (Mas et al., 2012, p.32), por lo que el disponer de los datos e información adecuados podrían, por ejemplo facilitar la toma de decisiones cuando ocurren desastres naturales o en cuestiones de políticas públicas que beneficien a la población de cualquier país.

Dicha información también puede contribuir al desarrollo económico de un país. Por ejemplo, según Olvera Ramírez et al. (2014), en Estados Unidos “los productos de datos espaciales de doce funciones federales clave facilitaron el progreso de la economía nacional en sectores que incluyen desde los bienes raíces y las aseguradoras, hasta la agricultura y la defensa por un valor de 3.5 trillones de dolares”(p.55).

Puede ocurrir que la información generada por un organismo sea necesaria para alguna otra persona u organización que gastará recursos en generar de nuevo esta información, al desconocer que ya existía o no tener acceso a la misma.

A raíz de la ocurrencia de este tipo de situaciones en todo el mundo se comenzaron a realizar esfuerzos con el fin de que los organismos del sector público que generaran información geográfica la compartieran a través de lo que se conoce actualmente como IDE.

Las IDE tienen su origen en la Orden Ejecutiva 12906 firmada por el presidente William Clinton de los Estados Unidos en 1994 en la que se ordena la creación de la Infraestructura Nacional de Datos Espaciales (NSDI, por sus siglas en inglés) con el fin de “apoyar aplicaciones de datos geoespaciales de los sectores público y privado en áreas como el transporte, desarrollo de comunidades, agricultura, respuesta a emergencias, manejo ambiental y tecnologías de la información” (Office of the President, 1994, p.1).

A partir de la publicación de la Orden Ejecutiva 12906 en varios países comenzaron a realizarse esfuerzos para desarrollar su propia IDE. Actualmente las IDE no sólo tienen un enfoque nacional, sino que han ido abarcando otros ámbitos como el estatal, municipal, local o internacional. Algunos ejemplos son: la IDE de España, la IDE de Uruguay, la IDE de Chile, etc.

En el ámbito internacional y de forma específica en la Unión Europea, con la creación de la creación de la directiva INSPIRE en 2007 se establece la creación de una IDE en la Comunidad Europea, que obliga a los países miembros de la Unión Europea a implantar una IDE Nacional. Ejemplos de éstas son las IDE de Holanda, de Suecia, de Alemania o de España. Esta directiva está basada en los siguientes principios (INSPIRE, 2007):

- Los datos deben ser recolectados sólo una vez y resguardados donde se puedan mantener de la manera más efectiva.

- Debe ser posible combinar información espacial de diferentes fuentes en toda Europa y compartirla con muchos usuarios y aplicaciones.
- Debe ser posible que la información recogida en un nivel o escala sea compartida a todos los niveles o escalas; detallada, para investigaciones exhaustivas; general para fines estratégicos.
- La información geográfica necesaria para el buen gobierno a todos los niveles debe estar fácilmente disponible y ser transparente.
- Debe ser fácil de encontrar la información geográfica que está disponible, la forma en que se puede utilizar para satisfacer una necesidad particular, y bajo qué condiciones puede ser adquirida y utilizada.

No sólo en Europa se han hecho esfuerzos por desarrollar IDEs. La Global Spatial Data Infrastructure Association “es una organización que mantiene una serie de recursos y actividades (sitio web, foro y congreso anual) como punto de encuentro en el ámbito mundial de proyectos e iniciativas IDE.”(Abad Power et al., 2012, p.52)

Si bien la finalidad de construir una IDE es la de poner a disposición de los usuarios un conjunto de datos espaciales, éstos no son el único componente de las IDE, pues también se requiere de mecanismos de búsqueda y entrega de la información a los usuarios, los mandatos legales que permitan que se destinen recursos públicos para la creación de las mismas, entre otros elementos.

En la siguiente lista se muestran los elementos que, según Abad Power et al. (2012, p.44-45), son necesarios para que exista una IDE:

- **Los datos.** Éstos se van a poner al alcance de los usuarios de la IDE bajo las restricciones de uso que decida el propietario.
- **El hardware y software.** En lo concerniente al software, esto incluye los sistemas operativos, bases de datos, aplicaciones, etc. En lo que respecta al hardware se debe verificar que la infraestructura tecnológica y el equipamiento sean adecuados, es decir, que tengan un rendimiento acorde a las necesidades establecidas. Este equipamiento incluye el número adecuado de servidores ya sea que estos pertenezcan a la organización encargada de crear la IDE o que sean rentados a una empresa que brinde ese servicio. También debe verificarse que el suministro de energía sea el adecuado, así como la conexión a internet.
- **Los metadatos.** Los metadatos constituyen el medio a través del cuál se hace una descripción de los datos disponibles en la IDE. Son una parte importante, pues representan una ayuda para los usuarios a la hora de localizar los recursos que estén buscando y decidir si estos recursos pueden o no cumplir con sus necesidades.
- **Las tecnologías.** Permitirán buscar, acceder y explotar los datos de la IDE. Estas tecnologías se refieren a los servicios geoespaciales, cuyas especificaciones son definidas por el OGC (especificaciones de implementación) y la ISO. Los servicios que más se utilizan en una IDE son los siguientes:
  - *Web Map Service (WMS).* Corresponde a una especificación del OGC y su primera versión fue liberada en Abril de 2000 y la versión 1.1, en Junio de 2001. Es un servicio de visualización de mapas e imágenes en web. Este servicio soporta varias operaciones, una de las cuales es *GetMap*, la cual da como resultado una imagen en formato gif, jpeg, png o wbmp. Ya que sólo es un servicio de visualización no provee información de localización ni

del sistema de referencia espacial. Este servicio permite visualizar datos; pero no permite descargarlos. De manera opcional implementa la operación *GetFeatureInfo*, que permite recuperar información asociada a elementos de una capa.

- **Web Feature Service (WFS).** También corresponde a una especificación del OGC. La versión 1.0.0 fue liberada en Mayo de 2002 y la versión 1.1.0, en Mayo de 2005. También se encuentra definido en la norma ISO 19142:2010. Permite publicar, acceder, consultar y descargar datos vectoriales (geometrías y atributos) por lo regular en formato GML. También permite recuperar o modificar (crear, actualizar y eliminar) los datos que se requieran.
  - **Web Coverage Service (WCS).** Al igual que los dos anteriores, es un estándar del OGC. En la norma ISO 19123 se establece el esquema conceptual para las características espaciales de la cobertura. Con este servicio se puede publicar y acceder a entidades continuas o de cobertura. Ejemplos de este tipo de datos son las capas ráster, redes irregulares triangulares (TIN) y coberturas de puntos, y polígonos, entre otros.
  - **Catalogue Service for the Web (CSW).** Definido por el OGC, este servicio permite a los usuarios consultar la información o los servicios disponibles dentro de la IDE. Se profundizará en este tipo de servicios dentro de la sección 1.3 de este mismo capítulo.
- **Los estándares.** Se refiere a las especificaciones técnicas que deben seguirse respecto al modelado y normalización de los datos, así como de los metadatos. También a las especificaciones técnicas referentes a la implementación de servicios geoespaciales. En general, son todas aquellas normas o estándares (del OGC, ISO o recomendaciones propias de un país o región) que facilitan la interoperabilidad técnica y semántica.
  - **Los acuerdos.** Se refiere a los acuerdos institucionales que evitan la duplicidad de datos, gastos y esfuerzos.
  - **El personal.** Se refiere al grupo de personas que ayudan a la creación y el mantenimiento de una IDE. En Olvera Ramírez et al. (2014) le llaman *capital Intelectual* a “un orden diferente de personas preparadas meticulosamente para el desarrollo de una IDE”(p.61). Pero la formación de este capital intelectual debería establecerse en el marco legal con el establecimiento de políticas para la formación y capacitación del mismo.
  - **Esquema organizativo.** Encargado de la coordinación de la IDE.
  - **El marco legal.** En este deben considerarse aspectos como la propiedad de los datos, derechos de autor, protección de la privacidad, derechos de acceso a los datos (hechos con dinero público), licencias de uso y precios.
  - **Las políticas.** Estas deben estar definidas por los gobiernos para regular y fomentar el uso de la información geográfica.
  - **Los usuarios.** Son quienes consultan y descargan la información de la IDE. Tienen la posibilidad de incorporar datos y opiniones para el mejoramiento de la misma.

Al inicio de este capítulo se redactó que una IDE puede ser vista como (y en realidad es) un sistema, pues cada una de las partes que la componen son de especial importancia y al igual que un sistema, una IDE tiene que verse como un todo y no como la suma de sus partes pues hacer de lado

a alguno de estos elementos puede repercutir en el funcionamiento de la IDE y en el abandono de proyectos de este tipo.

Hasta ahora se ha explicado cuál es la utilidad y los elementos que componen una IDE y así mismo, la importancia económica de los datos geográficos; pero no se ha dado una definición de lo que es una IDE. En (Olvera Ramírez et al., 2014) se define una IDE como:

una colección básica pertinente de tecnologías, políticas y disposiciones institucionales que facilitan la disponibilidad y el acceso a los datos espaciales. Una IDE incluye datos geográficos y atributos, documentación suficiente (metadatos), un medio para descubrir, visualizar y evaluar los datos (catálogos y cartografía por la red) y algún método para proporcionar acceso a los datos geográficos (p.25).

Finalmente, “Para que una IDE sea funcional, también debe incluir los acuerdos organizativos necesarios para coordinarla y administrarla a una escala local, regional, nacional o transnacional” (Olvera Ramírez et al., 2014, p.25).

#### 1.1.5.1. Geoportales

Entre los varios componentes de una IDE, dos componentes que cobran especial importancia cuando se toca el tema de geoportales son los metadatos y las tecnologías, independientemente del hardware o software que se utilice ya que un geoportal debe constituir una “puerta de entrada” a la información geográfica, así como a otros servicios. Por ello, como mencionan Maguire y Longley (2005, p.4), los geoportales son un elemento clave de las IDE, cuando no, el corazón de los sistemas de IDE.

Maguire y Longley (2005, p.7) también dicen que los geoportales sirven para controlar el uso comercial de los datos, pues facilitan la compraventa de datos y servicios geográficos gracias a que brindan facilidades de búsqueda y consulta de estos mismos datos y servicios a través de almacenes o catálogos de metadatos.

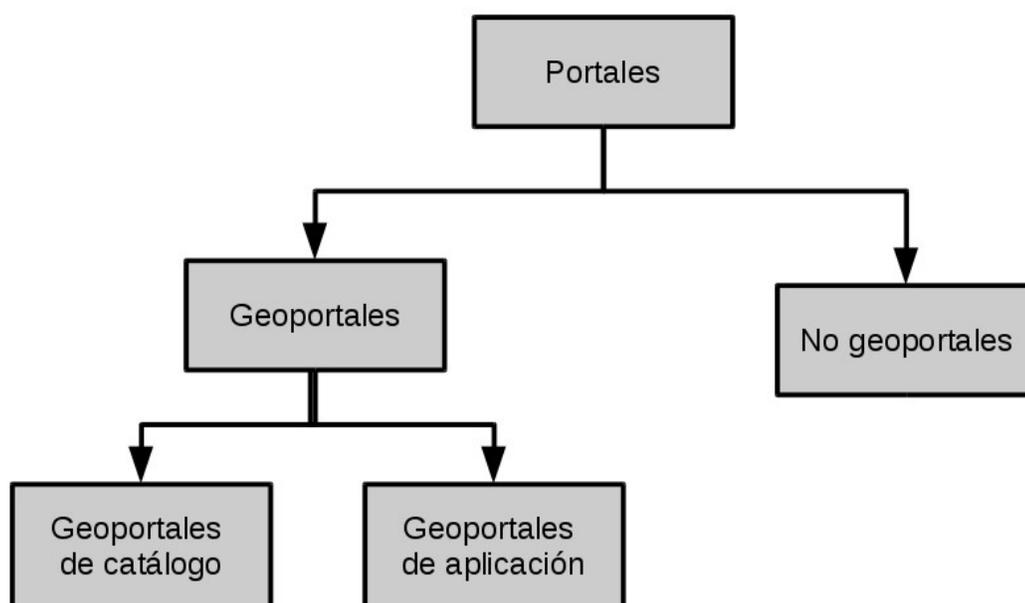
Ya sea que un geoportal sea de catálogo (aquel tipo de geoportal que se concentra solamente en organizar y administrar el acceso a la información geográfica) o de aplicación (aquellos que proveen servicios web geográficos como: servicios de mapeo o servicios de generación de rutas, etc. (ver Figura 1.9)) debe contar con un servicio de catálogo para publicar y acceder a los metadatos, los cuales permitan a los usuarios descubrir los servicios y datos disponibles en el geoportal de una IDE.

Sin importar la arquitectura del sistema, la vista de un geoportal corresponde con la de una página web que contiene un conjunto de ligas a otras páginas o servicios que contienen documentación tanto del mismo geoportal como de la información que contiene. Además se vale de un servidor de mapas para la administración de metadatos, el mapeo de la información, geocodificación, descarga de datos, etc.

Finalmente, estos servicios de que se compone un geoportal permiten a sus usuarios **publicar, buscar, descubrir** y **consumir** datos y servicios geoespaciales.

## 1.2. Marco Normativo en México

En México, el organismo encargado de la normatividad geográfica es el INEGI. Fue creado en el año 1983 y en 1985 concentró su sede en la ciudad de Aguascalientes. En 2004, el INEGI dio a conocer una iniciativa para la creación de una IDE para México, razón por la cual ya había empezado a trabajar en varios proyectos de normas técnicas, las cuales tienen como propósito establecer las



**Figura 1.9.** Clasificación de geoportales. Fuente: Maguire y Longley (2005)

reglas para la creación de datos geográficos, así como para la documentación de los mismos. Todas las normas publicadas por INEGI tienen la finalidad de facilitar tanto la interoperabilidad de los datos, como la interoperabilidad institucional (Olvera Ramírez et al., 2014).

En las siguientes secciones se hará mención de un par de normas técnicas que ha publicado INEGI, una de las cuales es de especial interés para este trabajo, ya que es la norma por la que se rigió la elaboración del SIGEM, motivo de esta tesis. Así mismo se abordará, aunque no de manera extensa, la Ley del Sistema Nacional de Información Estadística y Geográfica (LSNIEG), la cual encabeza el marco legal en cuestiones geográficas en México.

### 1.2.1. Ley del SNIIEG

En el año 2004, INEGI dio a conocer una iniciativa para la creación de una IDE para México, conocida como IDEMEX. Posteriormente en 2006, en la reforma hecha al artículo 26 constitucional, se establece que “El Estado contará con un Sistema Nacional de Información Estadística y Geográfica (SNIIEG)” (Secretaría de Gobernación, 2006), razón por la cual el 16 de Abril de 2008 se publicó la LSNIEG, que entró en vigor el 15 de julio del mismo año.

En esta misma norma, quedó establecido que el SNIIEG debería ser normado y coordinado por un organismo con autonomía técnica y de gestión, motivo por el cual en 2008, INEGI se convierte en un organismo autónomo y, a la fecha, es el organismo encargado de la coordinación del SNIIEG.

“La Ley del SNIIEG trajo consigo otro concepto diferente de IDEMEX” (Olvera Ramírez et al., 2014, p.137) y, actualmente, el SNIIEG es considerado por INEGI como una IDE (Instituto Nacional de Estadística y Geografía, 2015a). Entonces, el SNIIEG es el “conjunto de Unidades organizadas a través de los Subsistemas, coordinadas por el Instituto y articuladas mediante la Red Nacional de Información, con el propósito de producir y difundir la Información de Interés Nacional” (Art. 2, Fracción XIII de LSNIEG en Instituto Nacional de Estadística y Geografía (2008)).

Esta definición del SNIIEG resulta incompleta si no se sabe qué son los Subsistemas ni la Infor-

mación de Interés Nacional o la Red Nacional de Información.

La Red Nacional de Información quedó definida en el artículo 2 Fracción XII de la LSNIEG como el conjunto de procesos de intercambio y resguardo de información, para apoyar por un lado las actividades de coordinación del Sistema y de sus Subsistemas y por otro la prestación del Servicio Público de Información a toda la sociedad.

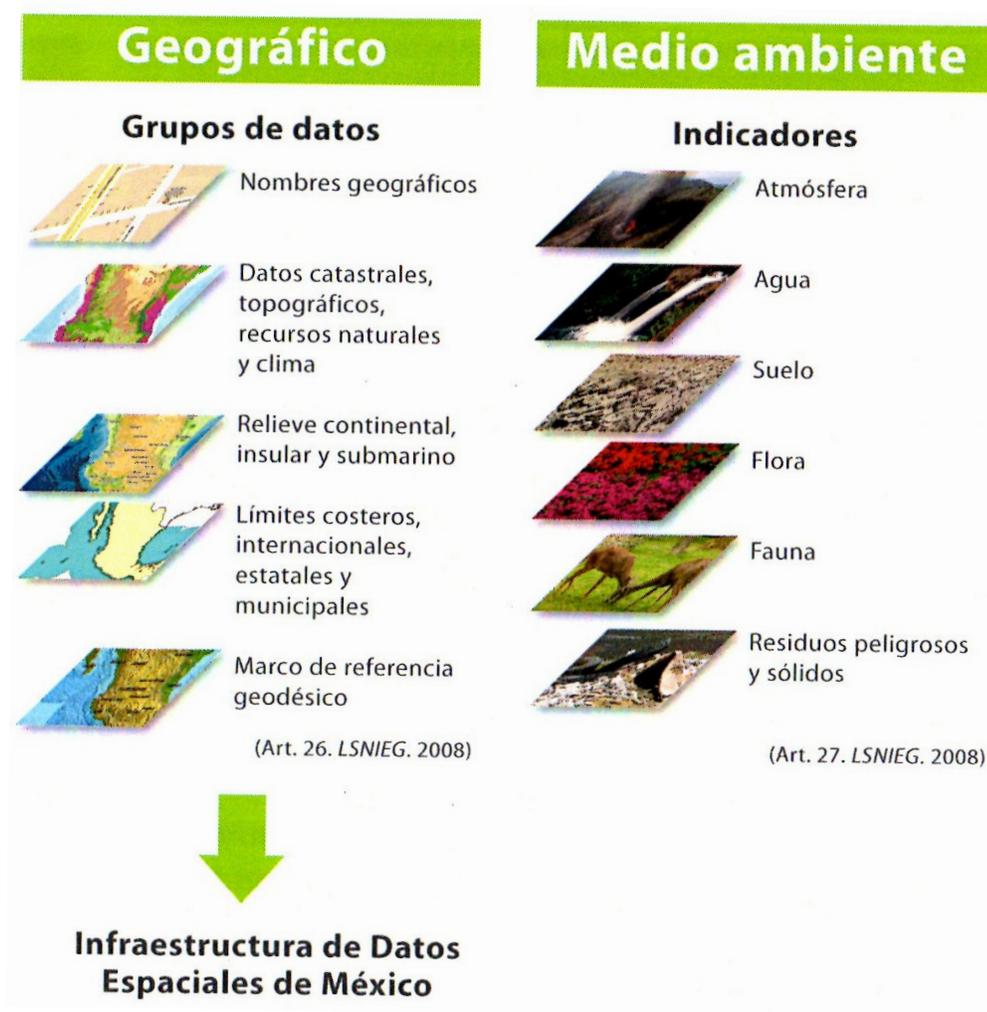
Por Subsistemas se entenderá a los Subsistemas Nacionales de Información, los cuales son los componentes del Sistema enfocados a producir información de una determinada clase o respecto de temas específicos (art. 2 Fracción XIV LSNIEG ). En Instituto Nacional de Estadística y Geografía (2015a) se menciona que el SNIEG se compone de cuatro subsistemas (ver figura 1.10), los cuales son:

- 1) Subsistema Nacional de Información Demográfica y Social,
- 2) Subsistema Nacional de Información Económica,
- 3) Subsistema Nacional de Información Geográfica y del Medio Ambiente y
- 4) Subsistema Nacional de Información de Gobierno, Seguridad Pública e Impartición de Justicia.



**Figura 1.10.** Subsistemas Nacionales de Información y sus componentes. Tomada de: Instituto Nacional de Estadística y Geografía (2015a, p.7)

Para el **S**ubsistema Nacional de **I**nformación **G**eográfica y del **M**edio **A**mbiente (SNIGMA) se generarán grupos de datos e indicadores como nombres geográficos o de relieve continental, insular y submarino o indicadores de temas como atmósfera, agua, suelo, entre otros (ver figura 1.11). A la componente geográfica también se le denominará “Infraestructura de Datos Espaciales de México” (art. 26 LSNIEG ).



**Figura 1.11.** Componentes del Subsistema Nacional de Información Geográfica y del Medio Ambiente. Tomada de: Instituto Nacional de Estadística y Geografía (2015a, p.8)

Por otro lado, la información generada por el INEGI será considerada de Interés Nacional. Además, de acuerdo con el artículo 78 de la LSNIEG también será considerada información de interés nacional aquella que:

I. Se trate de los siguientes temas, grupos de datos o indicadores: población y dinámica demográfica; salud; educación; empleo; distribución de ingreso y pobreza; seguridad pública e impartición de justicia; gobierno; vivienda; sistema de cuentas nacionales; información financiera; precios; trabajo; ciencia y tecnología; telecomunicaciones y radiodifusión; atmósfera; biodiversidad; agua; suelo; flora; fauna; residuos peligrosos y residuos sólidos; marco de referencia geodésico; límites costeros, internacionales, estatales y municipales; datos de relieve continental, insular y submarino; datos catastrales, topográficos, de recursos naturales y clima, y nombres geográficos, o bien se trate de temas que sean aprobados por unanimidad por el Consejo Consultivo Nacional;

II. Resulte necesaria para sustentar el diseño y la evaluación de las políticas públicas de alcance nacional.

III. Sea generada en forma regular y periódica.

IV. Se elabore con base en una metodología científicamente sustentada.

También podrá ser considerada como Información de Interés Nacional la que resulte necesaria para prevenir y, en su caso, atender emergencias o catástrofes originadas por desastres naturales, y aquella que se deba generar en virtud de un compromiso establecido en algún tratado internacional.

Dado que el SNIEG es considerado por INEGI como una IDE, el sistema cuenta con los siguientes cuatro componentes: Marco Organizacional, Marco Regulatorio, Marco Temático y el Marco Tecnológico (ver figura 1.12).



**Figura 1.12.** Componentes del SNIEG. Tomada de: Instituto Nacional de Estadística y Geografía (2015a, p.9)

El Marco Organizacional, que es la componente social, se refiere a los órganos colegiados de participación y de consulta, estos son: el Consejo Consultivo Nacional, los Subsistemas Nacionales de Información y el INEGI. En la figura 1.10 se muestra un cuadro con los componentes de cada uno de los Subsistemas Nacionales de Información. El Subsistema Nacional de Información Geográfica y del Medio Ambiente se compone de 8 Comités Técnicos Especializados, que incluyen 43 Unidades de Estado. Estos comites técnicos son:

- Información geográfica básica (9 Unidades del Estado)
- Información catastral y registral (7 Unidades del Estado)
- Agua (7 Unidades del Estado)
- Uso de suelo, vegetación y recursos forestales (7 Unidades del Estado)
- Emisiones, residuos y sustancias peligrosas (7 Unidades del Estado)

- Cambio climático (6 Unidades del Estado)
- Sector energético (10 Unidades del Estado)
- Desarrollo regional y urbano (9 Unidades del Estado)

Respecto al Marco Regulatorio, que es el componente político, éste incluye los instrumentos para la ordenación y regulación de las actividades necesarias para la planeación, programación, producción y difusión de la Información de Interés Nacional. Estas actividades se llevarán a cabo conforme a lo establecido en el Programa Estratégico del Sistema Nacional de Información Estadística y Geográfica; El Programa Nacional de Estadística y Geografía, y el Programa Anual de Estadística y Geografía.

Esta componente política, previamente citada, cubre dos rubros. El primero es la normatividad para la coordinación del SNIEG para que los integrantes del SNIEG participen en forma organizada. El segundo rubro es la normatividad técnica, con la cual se regula el diseño, captación, producción, actualización, organización, procesamiento, integración y compilación de la información con el fin de mejorar la homogeneidad y calidad de la información que producen las Unidades del Estado, razón por la que INEGI ha publicado, hasta ahora, 10 documentos regulatorios (ver tabla 1.8 ).

**Tabla 1.8.** Categorización de los documentos regulatorios publicados por INEGI en el marco de la LSNIEG.

<b>Categoría</b>	<b>Normas</b>
Acuerdos	– Acuerdo para el uso del Catálogo de los Términos Genéricos de las Formas del Relieve Submarino.
Lineamientos	– Lineamientos para el intercambio de información catastral con fines estadísticos y geográficos.
Normas técnicas	– Norma técnica para el Sistema Geodésico Nacional.
	– Norma técnica para la generación, captación e integración de datos catastrales y registrales con fines estadísticos y geográficos
	– Norma técnica sobre domicilios geográficos.
	– Norma técnica sobre elaboración de metadatos geográficos.
	– Norma técnica sobre estándares de exactitud posicional.
Procedimientos administrativos.	– Norma Técnica para la Generación de Modelos Digitales de Elevación con fines geográficos.
	– Norma Técnica para el Registro de Nombres Geográficos Continentales e Insulares con fines Estadísticos y Geográficos.
	– Norma para la autorización de levantamientos aéreos y exploraciones geográficas en el territorio nacional.

Pero según lo establecido por la LSNIEG en su artículo 33 no sólo el INEGI será el encargado de la normatividad geográfica, sino también aquellas Unidades del Estado cuando desarrollen actividades relacionadas con la producción, integración, conservación y difusión de Información de Interés Nacional. Estas Unidades de Estado deben proponer, en tiempo y forma, al Comité Ejecutivo que corresponda, los proyectos de normas técnicas y metodologías que, en el ámbito de sus funciones, sean necesarias para la realización de las Actividades tomando en cuenta los estándares nacionales e internacionales y las mejores prácticas en la materia.

Respecto a lo establecido en este mismo artículo y como menciona (Olvera Ramírez et al., 2014) la expresión “tiempo y forma” no tiene ningún sentido, ya que no se especifica ni el tiempo ni la

forma en el inexistente Reglamento de la Ley del SNIEG y por lo tanto no se les puede obligar a estas Unidades de Estado a cumplir con esta disposición.

La tercera gran componente del SNIEG es la componente geográfica, es decir, el Marco Temático, el cual integra el grupo de datos del SNIGMA; pero es a través del Marco Geoestadístico Nacional que es posible la integración de la información estadística de los otros Subsistemas Nacionales de Información con los otros indicadores del medio ambiente. Los grupos de datos que integran al SNIGMA pueden verse en la figura 1.11.

La última componente es la tecnológica, es decir, el Marco Tecnológico. Éste se encuentra integrado por todos los servicios que permiten el descubrimiento, visualización, descarga e intercambio de la información producida por las unidades de estado.

Ejemplos de los servicios para el descubrimiento de información estadística y geográfica son: el Registro Estadístico Nacional, el Centro Distribuidor de Metadatos (Clearinghouse), el Registro Nacional de Información Geográfica.

Entre los servicios de visualización y consulta se encuentran: el Mapa Digital de México (tanto la versión en línea, como la versión de Escritorio), Simulador de Flujos de Agua de Cuencas (SIATL), el Atlas Nacional Interactivo de México (ANIM), el Directorio Estadístico Nacional de Unidades Económicas (DENUE) o el Banco de Información Económica.

También se cuentan con sitios de descarga de información geográfica. Se pueden descargar datos acerca de: Recursos Naturales, Topografía, Continuo de Elevaciones Mexicano, datos de la Red Geodésica Nacional Activa o del Catálogo Único de Claves de Áreas Geoestadísticas Estatales, Municipales y Localidades.

De igual forma cuenta con servicios WMS de ortofotos, datos de relieve, de la Red Carretera Nacional, de las Estaciones Geodésicas de la Red Geodésica Nacional Pasiva, de imágenes satelitales, etc.

### 1.2.2. Norma Técnica para la elaboración de Metadatos Geográficos

Ya ha quedado establecida la importancia de los metadatos dentro de las IDE y también se abordaron de manera breve dos normas de metadatos muy usadas a nivel internacional; pero poco se mencionó acerca de la norma de metadatos para México.

La Norma Técnica de Elaboración de Metadatos Geográficos (NTM) fue publicada por INEGI en el Diario Oficial de la Federación el 24 de diciembre de 2010 y como se mencionó anteriormente, es un perfil de la norma internacional ISO 19115:2003. Esta norma es “de observancia obligatoria para el Instituto y para las Unidades del Estado que generen grupos de datos geográficos que sean determinados como información de interés nacional o sirvan para realizar ésta” (Instituto Nacional de Estadística y Geografía, 2010, Artículo 2). Esta norma entró en vigor al día siguiente de su publicación y se consideró que las Unidades de Estado debían adoptarla en un plazo no mayor a un año a partir de su entrada en vigor.

Para la elaboración de esta norma se tomaron en cuenta la mayoría de los elementos del CSDGM con el propósito de aminorar el impacto de transición entre una norma y otra (Instituto Nacional de Estadística y Geografía, 2015b). También se incluyeron algunos elementos definidos por la DGGyMA. La NTM consta de 9 secciones (5 son obligatorias, 3 son condicionales y 1 es opcional) dentro las cuales se encuentran distribuidos sus 149 elementos (ver figuras C.1 a C.9). Los elementos obligatorios son aquellos que deben ser incluidos en el metadato. Los elementos opcionales son aquellos que pueden, o no, ser incluidos dentro del metadato. Finalmente, los elementos condicionales son aquellos que se vuelven obligatorios cuando se cumple una condición durante el llenado del metadato. En la norma se especifican las condiciones que deben cumplirse para que estos elementos condicionales se

vuelvan obligatorios.

El llenado de un metadato geográfico puede volverse una tarea compleja, por lo que en 2015 INEGI publicó el documento “Guía Metodológica para la Generación e Integración de Metadatos Geográficos conforme a la Norma Técnica para la elaboración de Metadatos Geográficos (NTM)” en el cual se proporcionan ejemplos de llenado para cada uno de los elementos de la norma, así como una breve explicación acerca de los mismos, con el que la tarea de llenar un metadato puede volverse un poco más sencilla.

Las secciones con que cuenta la norma son:

1. Identificación del conjunto de datos espaciales o producto (O)
2. Fechas relacionadas con el conjunto de datos espaciales o producto (O)
3. Parte responsable del conjunto de datos espaciales o producto (O)
4. Localización geográfica del conjunto de datos espaciales o producto (C)
5. Sistema de referencia (C)
6. Calidad de la información (O)
7. Entidades y atributos (C)
8. Distribución (Opc)
9. Información del contacto para los metadatos (O)

En el anexo C se presentan los diagramas que permiten apreciar los elementos que conforman cada una de las secciones de la NTM.

### **1.2.3. Transparencia y datos abiertos: Norma Técnica para el acceso y publicación de Datos Abiertos de la Información Estadística y Geográfica de Interés Nacional**

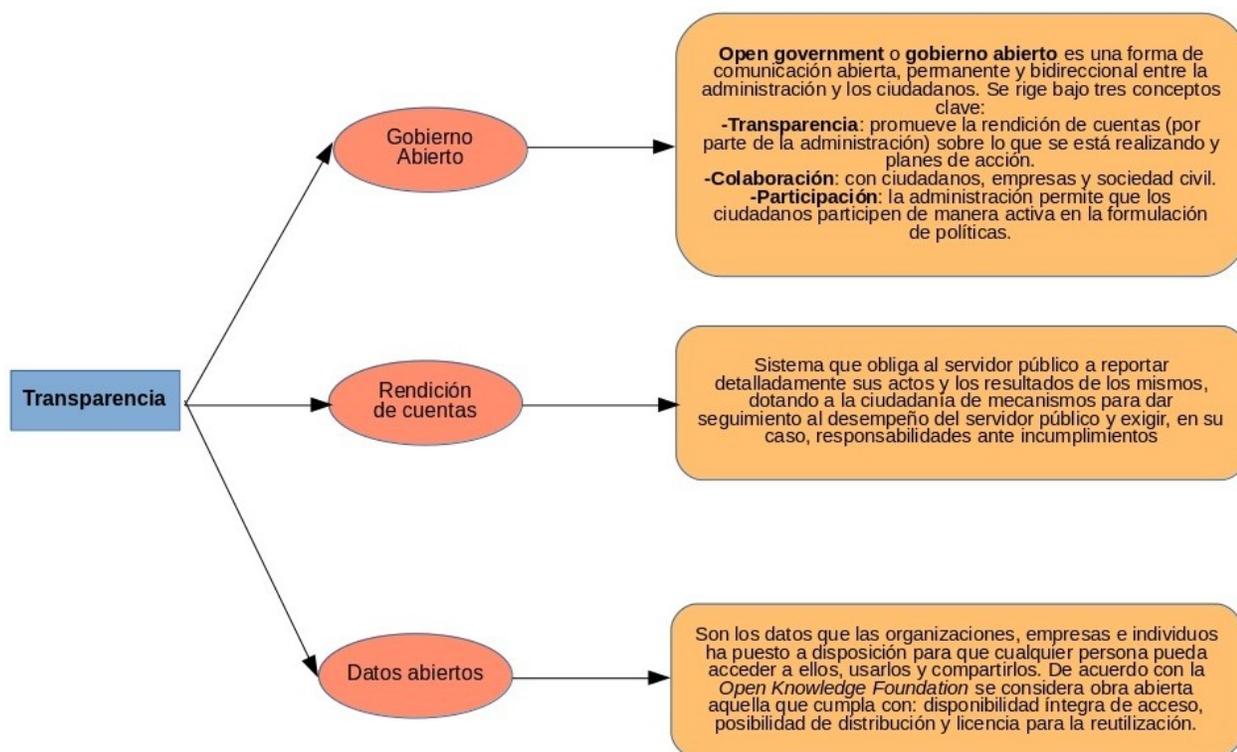
El INEGI es el organismo encargado de la coordinación del SNIEG, el cual cuenta con los siguientes principios rectores: accesibilidad, transparencia, objetividad e independencia, de acuerdo con lo establecido en el artículo 3 de la Ley del SNIEG. Es aquí donde resalta el tema de la transparencia, el cual es cada vez más sonado en México, en parte debido a la publicación de la *Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental* como resultado de la adhesión de México al *Pacto Internacional de Derechos Civiles y Políticos* con el que “el país se incorporó a la cultura del acceso a la información y rendición de cuentas” (Hernández Cruz, 2015).

El 4 de mayo de 2015, fue publicado en el Diario Oficial de la Federación la *Ley General de Transparencia y Acceso a la Información Pública*, en la que se establece la creación del *Sistema Nacional de Transparencia, Acceso a la Información Pública y Protección de Datos Personales* (del cual el INEGI es parte) con el objetivo de fortalecer la rendición de cuentas del Estado, la instauración de gobiernos abiertos y la transparencia proactiva. Además deberá establecer los lineamientos para la implementación de la *Plataforma Nacional de Transparencia* que permita cumplir con los procedimientos y obligaciones de los sujetos obligados, atendiendo las necesidades de accesibilidad de los usuarios (Secretaría de Gobernación, 2015).

Y ¿qué es *Transparencia*? Según Naessens (2010, p.2121), la *Transparencia* debe involucrar tres aspectos:

- Es un atributo o cualidad (algo que una persona tiene por el hecho de ser persona)
- Información clara y precisa, es decir, información comprensible, sin ambigüedades, para ser debidamente procesada;
- Capacidad de comprensión, vigilancia y comunicación, que resultan necesarias para la plena realización de la persona.

Con base en estos tres atributos, Naessens (2010) define la transparencia como el “atributo o cualidad que permite tener más información clara y precisa sobre una persona o algo, lo que redundaría en el aumento de nuestras capacidades de comprensión, vigilancia y comunicación”(p.2121). Este concepto suele aplicarse a instituciones públicas o a la forma en que los servidores públicos deben actuar, lo cual va ligado a conceptos como el derecho y el acceso a la información. La Organización para la Cooperación y el Desarrollo Económicos (OCDE) definió la *Transparencia* como “El proceso por el cual la información de las condiciones existentes permite ser accesible para la toma de decisiones y acciones, de manera visible y entendible” (citado en De León (2008, p.1).



**Figura 1.13.** Conceptos relacionados con Transparencia. Fuente: elaboración propia con información de Ferrer-Sapena et al. (2011) y López del Consuelo (2015)

Por otro lado *Datos Abiertos* u *Open Data* es una idea que también está relacionada con el concepto de *Transparencia* y es un movimiento que parte de la base de que la información de la Administración es patrimonio de todos los ciudadanos, puesto que el gobierno genera información con financiamiento de los impuestos que pagan los ciudadanos. Un punto en el que coinciden algunos autores con respecto a este tema es que los datos deben publicarse en formatos legibles e interoperables, para que éstos puedan ser reutilizados y procesados sin ningún problema.

De acuerdo con Ferrer-Sapena et al. (2011, p.263), en 2009, la W3C dio a conocer recomendaciones para la publicación de datos abiertos. Estas recomendaciones incluyen que los datos deben publicarse en formatos interoperables (como xml, rdf o csv); que se publiquen catálogos en línea para que los usuarios sepan lo que se ha publicado, además de incluir los metadatos (documentación) para conocer la calidad de estos datos y; que los datos sean enriquecidos en formato xhtml, el cual combina la sintaxis del html (para mostrar los datos) con la de xml (para describirlos). De igual forma, poner a disposición del público los datos, no tiene sentido si éstos no son usados para generar conocimiento.

Respecto a la información geográfica, ésta también puede publicarse por los gobiernos y otras instituciones bajo la iniciativa de *Datos Abiertos*, razón por la cual (en parte) el 4 de diciembre de 2014 el INEGI publicó en el Diario Oficial de la Federación la **Norma Técnica para el acceso y publicación de Datos Abiertos de la Información Estadística y Geográfica de Interés Nacional**. Esta norma consta de nueve elementos (ver tabla 1.9), los cuales son obligatorios y fueron establecidos como los elementos mínimos para un metadato de *Datos Abiertos*. Sin embargo, esta misma norma indica que un metadato puede ser extendido con elementos de otras normas técnicas del SNIEG para “una mayor precisión en diferentes contextos”.

En el artículo 5 de esta norma se establecen los requisitos que se deben cumplir para que un conjunto de datos publicado por una Unidad del Estado sea considerado como *Datos Abiertos*. Estos requisitos incluyen que la información sea:

- de carácter público,
- gratuita,
- publicada sin restricciones de acceso a los usuarios,
- de libre uso (con la única condición de citar la fuente de origen),
- publicada en archivos digitales de formatos abiertos,
- estructurada o semi-estructurada (para su fácil interpretación y procesamiento por equipos electrónicos),
- publicada con su correspondiente metadato y deberán contener el tema que describen a detalle,
- publicada con el mayor nivel de desagregación posible y que provengan de la fuente de origen,
- actualizada y publicada en cuanto sea generada,
- conservada en el tiempo junto con sus diferentes versiones.

**Tabla 1.9.** Componentes de los metadatos para un conjunto de datos publicado como Datos Abiertos. Fuente: (Instituto Nacional de Estadística y Geografía, 2014, Artículo 9)

Elemento	Descripción
Título	Nombre descriptivo de la base de datos que facilita la búsqueda, identificación y entendimiento del conjunto de datos. Equivalente a dct:title, referenciado por DCAT y definido por "La Iniciativa de Metadatos Dublin Core".

Tabla 1.9 (continuación)

Elemento	Descripción
Descripción	Una explicación de los datos, con suficiente detalle para que los usuarios puedan entender su contexto y determinar si es de su interés. Equivalente a <code>dct:description</code> , referenciado por DCAT y definido por Dublin Core.
Palabra clave	Etiquetas o palabras clave que facilitan a los usuarios la búsqueda del conjunto de datos. Es importante considerar el uso de términos tanto técnicos como no técnicos. Equivalente a <code>dcat:keyword</code> , definido por DCAT.
Última modificación	La fecha más reciente en que se cambió, modificó o actualizó el conjunto de datos. Equivalente a <code>dct:modified</code> , referenciado por DCAT y definido por Dublin Core.
Publicador	La entidad responsable de la publicación del conjunto de datos. Equivalente a <code>dct:publisher</code> , referenciado por DCAT y definido por Dublin Core.
Punto de contacto	Información de contacto relevante al conjunto de datos. Equivalente a <code>dcat:contactPoint</code> , definido por DCAT.
Identificador	El identificador debe ser único dentro del catálogo de datos de la entidad y en el mejor de los casos, presente como parte del URI del conjunto de datos. Equivalente a <code>dct:identifier</code> , referenciado por DCAT y definido por Dublin Core.
Distribución(es)	Conecta un conjunto de datos con sus distribuciones disponibles. Equivalente a <code>dcat:distribution</code> , definido por DCAT.
Términos de Libre Uso	Liga directa al documento que establece los derechos de libre uso del conjunto de datos. Equivalente a <code>dct:license</code> , referenciado por DCAT y definido por Dublin Core.

Actualmente existen diversas entidades de gobierno que publican información geográfica, aunque no toda la información publicada incluye metadatos. En México, un par de sitios en los que la que se puede encontrar los metadatos de la información descargada son el del INEGI y CONABIO.

Algunas fuentes de datos abiertos tanto de México como del mundo se muestran a continuación en las figuras 1.14 y 1.15.

### 1.3. Software, sistemas y servicios para capturar y catalogar metadatos

El objetivo de esta sección es el de hacer mención de algunas aplicaciones que permiten la captura y búsqueda de metadatos, así como el de describir qué son los servicios de catálogos para metadatos geográficos. Para ello parece conveniente comenzar con la definición de los conceptos que integran el título de esta sección, es decir, software y servicios (web). Se hace una breve mención acerca de los sistemas, pues ya se dedicó una parte de este texto a los mismos en el apartado 1.1.1.

El primer paso es definir el concepto de *Software*. De acuerdo con (Espinosa Prendes y Amorós Poveda, 2001, p.1) los términos *software* y *programa* se refieren a lo mismo y pueden considerarse como la parte intangible de un sistema informático o como un conjunto de instrucciones que indican



Figura 1.14. Fuentes nacionales de datos abiertos. Tomado de: Chias Becerril y Reséndiz López (2015).



Figura 1.15. Fuentes internacionales de datos abiertos. Tomado de: Chias Becerril y Reséndiz López (2015).

a la máquina (*hardware*) lo que debe hacer. Estas instrucciones abarcan actividades que van desde poner en marcha el equipo (software de sistema o conocido también como el sistema operativo) hasta satisfacer ciertas necesidades de los usuarios (software específico como editores de texto, hojas de cálculo, etc.).

Por otro lado, Jacobson et al. (2000) hablan acerca de un *sistema software*. Éste, al igual que cualquier otro sistema es un conjunto de elementos relacionados entre sí y, aunque redundante, es también un software (el conjunto de instrucciones o el código escrito por programadores). Pero no sólo eso, sino que también incluye todos los “artefactos que se necesitan para representarlo en una forma comprensible por máquinas u hombres, para las máquinas, los trabajadores y los interesados”(Jacobson et al., 2000, p.18). Dichos artefactos pueden ser diagramas UML (para la representación de modelos de casos de uso, de diseño, etc.) o cualquier otra documentación producida o utilizada por los trabajadores en el desarrollo del sistema.

Existen también los llamados servicios web; pero antes de definirlos se explicará a qué se refiere el término *Web*, así como el término *internet*. *Internet* es “una red masiva de redes que conecta millones de computadoras en todo el mundo” Fu y Sun (2010, p.4). Éstas se comunican entre sí por medio de protocolos como **H**ypertext **T**ransfer **P**rotocol (HTTP), **S**imple **M**ail **T**ransfer **P**rotocol (SMTP), **F**ile **T**ransfer **P**rotocol (FTP), **I**nternet **R**elay **C**hat (IRC), **I**ntant **M**essaging (IM), Telnet y P2P (peer-to-peer).

La *World Wide Web (WWW)*, es un “sistema de documentos de hipertexto interrelacionados y programas que pueden ser accedidos via la Internet, usando principalmente HTTP” (Fu y Sun, 2010, p.4). La *web* es la “cara” del internet ” (Citado por (Fu y Sun, 2010, p.4), con base en Douglas (2008)).

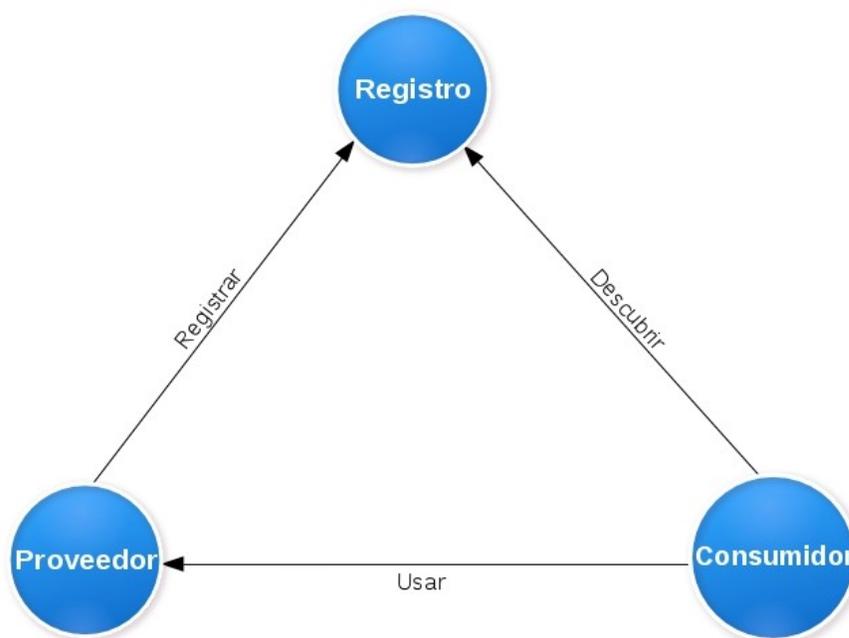
Entonces, de acuerdo con Fu y Sun (2010, p.51), los servicios web son programas que corren sobre servidores web y presentan interfaces de programación hacia otros programas en la web, es decir, que a diferencia del software que corre en una computadora personal (por ejemplo), los servicios web son programas que corren en un servidor y son accedidos por otros programas que comúnmente se encuentran en otras máquinas. Éstas intercambian información a través de la *Web*.

Fu y Sun (2010, p.52) identificaron tres roles dentro de la arquitectura de servicios web: el proveedor, el consumidor y el servicio(ver figura 1.16). El proveedor ofrece el servicio, el consumidor usa el servicio y el registro funciona como intermediario entre el proveedor y el consumidor, pues éste último puede consultar en el registro la disponibilidad de servicios. Tanto estos tres elementos, así como la forma en que se diseñan y operan los servicios *web*, los convierte también en *sistemas software*.

Cualquier *software* (o *sistema software* es creado para cumplir un propósito específico y dada la importancia de los metadatos para documentar las capas de datos geográficos así como para encontrar recursos dentro de una IDE, es que actualmente existen varias aplicaciones que ayudan a solucionar el problema de creación de metadatos para conjuntos de datos geográficos, algunas de las cuales también permiten integrarlos dentro de catálogos de metadatos.

Por otro lado, como se mencionó en la sección 1.1.5 de esta tesis, los catálogos de metadatos son una parte importante de las IDE; pero poco se mencionó con respecto a qué son los catálogos de metadatos así como los servicios CSW.

El CSW es un tipo de servicio web definido por el OGC cuya finalidad es permitir a los usuarios la localización de datos y servicios geográficos mediante la implementación de dos de sus principales operaciones: *GetRecords* y *GetRecordById*. La primera operación devuelve todos los registros que coincidan con los criterios seleccionados por el usuario, mientras que la segunda permite ver y/o analizar un solo metadato con mayor detalle. En un servidor CSW, un usuario puede buscar metadatos bajo ciertos criterios como: palabras clave, nombres de lugares, por datos que no rebasen cierta



**Figura 1.16.** Los tres roles que completan la arquitectura de servicios web. Adaptado de: (Fu y Sun, 2010, p.52)

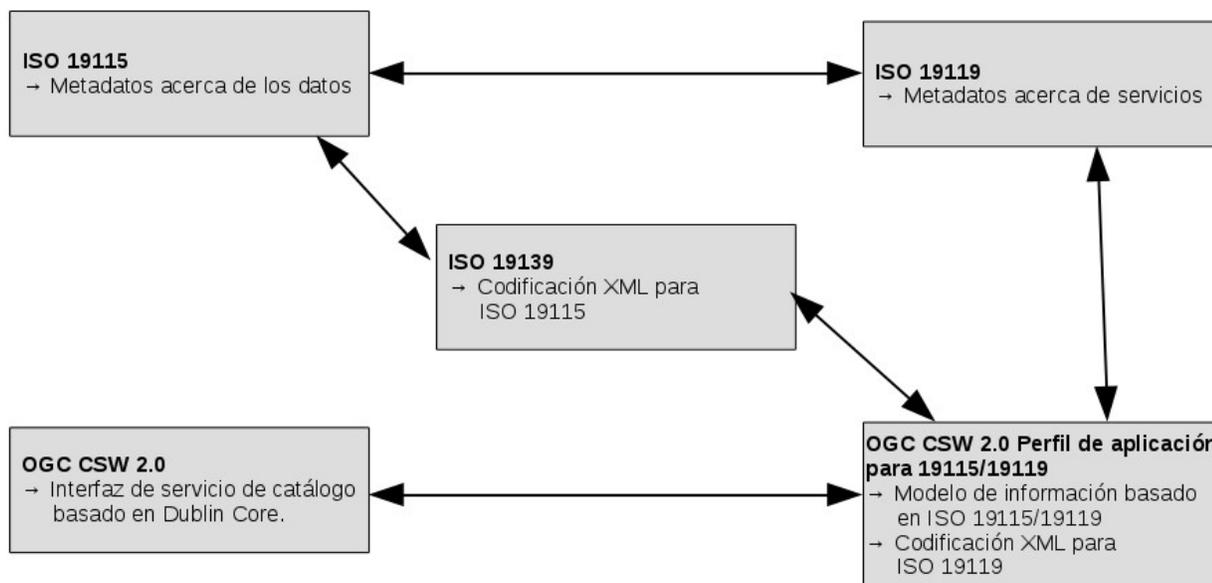
extensión geográfica, etc. Cuando el usuario va a ingresar un valor para alguno de estos criterios se ejecuta la operación *GetDomain*, la cual devuelve una lista de todos los valores registrados para ese elemento en todos los metadatos. Por ejemplo, si un usuario busca metadatos por palabras clave, entonces la operación *GetDomain* mostraría una lista de todas las palabras claves registradas en todos los metadatos. El número de elementos a través de los cuales se puede realizar la búsqueda de metadatos se define mediante la operación *GetRecords*. La operación *Transaction* permite insertar borrar o actualizar un metadato. La operación *Harvest* permite programar una consulta periódica a otro catálogo o a un conjunto de servicios web (*Capabilities*).

Según las operaciones que se implementen, existen dos tipos de servicios CSW: el CSW de sólo lectura y el CSW transaccional (Fu y Sun, 2010, p.74). El primero sólo soporta operaciones de consulta de metadatos como: *GetCapabilities*, *DescribeRecord*, *GetRecords*, *GetRecordById* y *GetDomain*; mientras que el CSW transaccional soporta operaciones que permiten la lectura y escritura de metadatos, además de publicar, editar y borrar metadatos por medio de las operaciones *Transaction* y *Harvest*.

Existe una operación que se implementa en todos los servicios geoespaciales definidos por el OGC y esta es *GetCapabilities*. Dicha operación permite obtener la información de las propiedades de un servicio como: el nombre del mismo, el estándar OGC que cumple, la versión, quién publica el servicio, la zona que cubre, las restricciones de uso, las operaciones que el servicio puede realizar, etc. Estos datos son importantes para la documentación de servicios geoespaciales mediante el estándar ISO 19119.

De acuerdo con (Litwin y Rossa, 2011, p.72), el estándar ISO 19119 así como 19115 y 19139 deben incluir especificaciones para el servicio CSW y para un perfil de aplicaciones del servicio CSW para las normas 19115 y 19119, ya que éstos se relacionan con las capacidades de publicación y descubrimiento de información y servicios geoespaciales y definen cómo deben construirse y operar

los servicios de catálogo. En la figura 1.17 se muestra un diagrama con la relación entre los estándares para metadatos geoespaciales.



**Figura 1.17.** Relación entre los principales estándares para metadatos geoespaciales. Traducido de Litwin y Rossa (2011, p.73).

En el anexo D se muestran ejemplos de algunas de aplicaciones que permiten la captura de metadatos, así como aplicaciones que permiten la implementación de servicios de catálogo basados en el estándar CSW del OGC. Hacer mención de todas las que existen en este momento no es el propósito de este trabajo, sino simplemente el de mostrar cómo son y las capacidades que en general ofrecen este tipo de aplicaciones.

Además de los servicios CSW existen los llamados *Clearinghouses* o Centros Distribuidores de Metadatos, los cuales son también catálogos de metadatos. El primer *Clearinghouse* para metadatos fue creado en Estados Unidos en 1994 a instancias del FGDC. Los *clearinghouses* son medios a través de los cuales se ponen a disposición los metadatos, para facilitar la localización de los datos geoespaciales. De acuerdo con Vergara Hernández (2006), los *clearinghouses* se componen principalmente de :

- Las instancias que contienen la información y manejan los metadatos.
- Servicios que hacen posible el acceso a la información.
- Software y los metadatos necesarios en Internet.

por lo que contribuyen a “minimizar la duplicación de esfuerzos en la búsqueda y colecta de los datos espaciales y fomentar las asociaciones de productores y usuarios con necesidades comunes”(Olvera Ramírez et al., 2014, p.114).

Un gran número de aplicaciones para capturar metadatos y de servidores CSW trabajan con el estándar ISO 19115 o perfiles del mismo; pero no específicamente con la Norma Técnica para la Elaboración de Metadatos Geográficos (NTM) del INEGI. Es por ello que el INEGI se dio a la tarea de desarrollar un sistema que permitiera la captura de metadatos de datos espaciales bajo

su propia norma: el SICAM –Sistema de Captura de Metadatos– (ver figura 1.18). Esta aplicación permite crear y validar metadatos contemplando la totalidad de los elementos de la NTM. También permite la conexión a un servidor para almacenar los metadatos dentro de una base de datos, la cual deberá tener una estructura conforme a la NTM. En caso de requerir una, el INEGI podrá proporcionar el esquema de la base de datos (Instituto Nacional de Estadística y Geografía, 2008). Si no se cuenta con conexión a Internet, los metadatos se almacenan dentro del sistema de archivos de la computadora en que esté instalada la aplicación. Esta aplicación puede ser descargada en [http://www.inegi.org.mx/geo/contenidos/metadatos/aplicaciones\\_conversion.aspx](http://www.inegi.org.mx/geo/contenidos/metadatos/aplicaciones_conversion.aspx)

**Figura 1.18.** Interfaz del SICAM, desarrollado por INEGI. Cada una de las pestañas contiene un formulario para el llenado de cada una de las secciones de la NTM

En lo que respecta a catálogos de metadatos, el INEGI cuenta con un *Clearinghouse* o Centro distribuidor de metadatos, el cual es un catálogo de metadatos (conforme a la NTM); pero no un servicio CSW. Sirve para localizar metadatos de los datos espaciales publicados por INEGI bajo criterios como: Escala cartográfica, por tipo de producto, por tema o texto, por fecha (de creación de los insumos o de inicio de verificación de campo, fin de verificación del campo o de referencia del producto) o por extensión geográfica o entidad federativa. (Ver figuras 1.19 y 1.20).

Por último, es conveniente mencionar, que anterior a la elaboración de la NTM del INEGI, los metadatos geográficos se elaboraban siguiendo el estándar CSDGM del FGDC de Estados Unidos. Lo anterior llevó al INEGI a la creación del Sistema de Conversión de Metadatos (SICOM) con el fin de facilitar la transición a la nueva norma. Ésta es una aplicación que procesa un metadato de la norma del FGDC (en formato XML), a partir del cual crea otro metadato (también en formato XML) pero conforme a la NTM (Instituto Nacional de Estadística y Geografía, 2008).

INEGI INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

Inicio | Contacto | Siguenos: RSS Twitter Facebook YouTube

Estadística Geografía Investigación Productos y Servicios Acerca del INEGI

Inicio > Geografía > Metadatos de geografía >

## Metadatos de Geografía

### Centro distribuidor de metadatos (Clearinghouse)

NTM

Buscar Reiniciar

Producto: Seleccione...

Escala cartográfica: Seleccione...

Tema o texto: Ejemplo: Ortofoto

Entidad: Todo el País

Tipo de fecha: Seleccione...

al

Derechos Reservados © INEGI | Versión: 201609271202

Figura 1.19. Interfaz del buscador de metadatos del Clearinghouse del INEGI

The screenshot displays the INEGI website's 'Clearinghouse' section. The header includes the INEGI logo and navigation menus for 'Estadística', 'Geografía', 'Investigación', 'Productos y Servicios', and 'Acerca del INEGI'. A search bar is present. The main content area is titled 'Centro distribuidor de metadatos (Clearinghouse)' and features a vertical 'Filtrar Metadatos' sidebar. The record shown is for 'Conjunto de Datos Vectoriales D14 de la Serie Topográfica Escala 1:1 000 000 (HUATULCO)'. It includes a list of sections (1-9) and detailed information under section 1, such as title, purpose, description, language, and category.

**Centro distribuidor de metadatos (Clearinghouse)**  
NTM

Regresar | Exporta: PDF | TXT | XML | HTML | RTF

### Conjunto de Datos Vectoriales D14 de la Serie Topográfica Escala 1:1 000 000 (HUATULCO)

Sección 1: Identificación del conjunto de datos espaciales o producto  
 Sección 2: Fechas relacionadas con el conjunto de datos espaciales o producto  
 Sección 3: Parte responsable del conjunto de datos espaciales o producto  
 Sección 4: Localización geográfica del conjunto de datos espaciales o producto  
 Sección 5: Sistema de referencia  
 Sección 6: Calidad de la información  
 Sección 7: Atributos  
 Sección 8: Distribución  
 Sección 9: Información del contacto para los metadatos

#### 1. Identificación del conjunto de datos espaciales o producto

**1.1 Título del conjunto de datos espaciales o producto:** Conjunto de Datos Vectoriales D14 de la Serie Topográfica Escala 1:1 000 000 (HUATULCO)

**1.2 Propósito:** La riqueza de información de la Carta Topográfica escala 1:1 000 000, la hace útil como base para una gran variedad de aplicaciones, como planeación, estudios ambientales, evaluación de infraestructura, evaluación de cambios, etc. Al proporcionarse en formato digital, los Datos Topográficos pueden ser complementados con otros tipos de información espacial para realizar diversos tipos de análisis y modelaje en sistemas de información geográfica.

**1.3 Descripción del conjunto de datos espaciales o producto:** Los Conjuntos de Datos Vectoriales de la Carta Topográfica escala 1:1 000 000, contienen la información sobre los diversos rasgos geográficos presentes en la Carta Topográfica impresa, como curvas de nivel, hidrografía, vías de comunicación, localidades, entre otros. Estos rasgos son representados digitalmente por un componente geométrico (puntos, líneas o áreas), y un componente descriptivo (los atributos del rasgo). Los topónimos o nombres geográficos se proporcionan en un producto separado como conjunto de datos toponímicos. LOS METADATOS ESTÁN SIENDO ACTUALIZADOS CONFORME A LA NORMA TÉCNICA PARA LA ELABORACIÓN DE METADATOS GEOGRÁFICOS (NTM).

**1.4 Idioma del conjunto de datos espaciales o producto:** ES - Español.

**1.5 Categoría del tema del conjunto de datos espaciales o producto:** 8. Base de imágenes de mapas de la cobertura de la tierra: Mapas base.

Figura 1.20. Ejemplo de un metadato encontrado a través del Clearinghouse del INEGI



## Capítulo 2

# Requerimientos del sistema

Uno de los propósitos de este trabajo de investigación es mostrar el proceso de diseño y desarrollo de un sistema de captura, almacenamiento, edición y búsqueda de metadatos, basado en los lineamientos de la NTM publicada por INEGI en 2010. Este proyecto es resultado de un requerimiento de la U-GITS del IG - UNAM a partir de su necesidad de documentar la información generada y publicada en los portales geográficos que forman parte de sus proyectos geoespaciales y aunque originalmente este sistema de metadatos se desarrolló para cubrir las necesidades de este grupo de trabajo, el resultado obtenido también cuenta con características que podrían permitir su uso por parte de otros usuarios y otras dependencias; pero este tema se aborda más adelante.

El presente capítulo se divide en tres partes. En las primeras dos se presenta la justificación del desarrollo de este sistema. Además, mediante un análisis FODA (de Fortalezas, Oportunidades, Debilidades y Amenazas) se identificaron las necesidades que debía cubrir este sistema; la tercera parte del capítulo está dedicada a definir los requerimientos del sistema de metadatos mediante casos de uso, que sirvieron para identificar las respuestas del sistema ante las peticiones de los usuarios y, por tanto, ayudar a definir su arquitectura y la forma en que fue programado.

### 2.1. Justificación y utilidad en la Geomática

Para poder justificar por qué un sistema de este tipo es útil en el campo de la geomática, es importante entender antes el concepto de *Geomática*. Gomasca (2009, p.vii) dice que se trata de un neologismo empleado para nombrar al conjunto de disciplinas y técnicas interrelacionadas empleadas para el estudio de la superficie de la Tierra y sus ambientes, es decir, para recolectar, integrar, modelar, analizar, recuperar, transformar, desplegar y distribuir datos georreferenciados con buena precisión y en formato digital.

Algunas de las disciplinas y técnicas que abarca la Geomática son la geodesia, cartografía, fotogrametría, percepción remota, topografía, entre otras; pero las que resultan fundamentales son las ciencias de la computación; pues proveen las bases teóricas del estudio de los sistemas y algoritmos que permiten integrar, modelar, analizar, recuperar, transformar, desplegar y distribuir los datos recolectados a través de las otras disciplinas que componen la geomática.

Es precisamente, de las ciencias de la computación que se desprende parte del nombre de esta disciplina. *Geomática* se deriva de dos vocablos, siendo el primero de ellos *geos*: Tierra y *mática*: informática. En Gomasca (2009, p.1) se menciona que el término *Geomática* fue creado en la década de los 80's en la universidad de Laval en Canadá basados en la idea de que los avances en la computación estaban revolucionando la forma en que recolectaban los datos, así como en la forma

de representarlos.

A partir del tratamiento que se le da a los datos, es posible obtener información y dado que en la *Geomática* estos datos están referidos a algún punto sobre la superficie de la Tierra, a la información derivada se le denomina GeoInformación o Información Geoespacial (**GI** por sus siglas en inglés). Cada vez es mayor la cantidad de datos espaciales generados y mayor la necesidad de Información Geoespacial que sirva en la resolución de problemas que requieran de conocimiento del espacio; pero darle un uso adecuado requiere de varios factores, siendo uno de ellos, contar con una infraestructura tecnológica adecuada para procesarlos, publicarlos y/o distribuirlos.

Sin embargo, el uso de los datos no sólo depende del factor tecnológico, pues para poder procesar dichos datos y obtener de ellos Información Geoespacial, es necesario conocer varios elementos referentes a los mismos, como su contenido, significado, precisión, exactitud, la forma en que fueron obtenidos, entre otros aspectos. Al conocer y entender estos elementos, es posible determinar tanto la calidad de los datos como la de la información generada a partir de ellos y de esta manera saber si serán o no útiles en la resolución de problemas que requieran de este tipo de datos.

Como ya se mencionó en el capítulo anterior, cuando estos datos son distribuidos, lo ideal es que también incluyan su metadato, pues este es el medio a través del cual se dan a conocer las características que determinan su calidad y en parte la calidad de los resultados derivados de éstos. Es por esta razón que se han creado diferentes normas para definir, de manera formal, los elementos que deben llevar estos metadatos y razón por la cuál también, se ha detectado la necesidad de desarrollar sistemas para capturarlos, almacenarlos y distribuirlos.

Entonces el contar con esta clase de sistemas así como la publicación de los metadatos, al menos desde el punto de vista de la geomática, facilita el acceso de los usuarios a la información geográfica, con lo que además es posible la utilización de la misma para otras aplicaciones que redunden en beneficios sociales o económicos de un determinado territorio.

En México, el INEGI cuenta con el SNIEG, el cual en Noviembre de 2015 fue dado a conocer como la IDE de México. En la sección 1.2.1 se abordó la estructura de esta IDE de manera muy general. El SNIGMA genera e integra grupos de datos geográficos así como indicadores, siendo los primeros considerados como el marco temático de la IDE de México (ver figuras 1.11 y 1.12). La componente tecnológica de esta IDE cuenta con geoservicios de descubrimiento, uno de los cuales es el Centro Distribuidor de Metadatos, basado en la NTM, del cual también ya se hizo mención en el capítulo anterior. En la figura 2.1 se muestra un diagrama con estos geoservicios de descubrimiento.



**Figura 2.1.** Ejemplo de servicios de descubrimiento implementados por INEGI. Tomada de: Instituto Nacional de Estadística y Geografía (2015a, p.15)

Por otro lado, al inicio de la administración del actual presidente, Enrique Peña Nieto, se dio a conocer el Programa para un Gobierno Cercano y Moderno (Gobierno de la República, 2013b) como parte del Plan Nacional de Desarrollo 2013-2018. Con este programa, se busca, entre otras cosas, acelerar la inserción de México en la Sociedad de la Información y del Conocimiento. En la Estrategia Digital Nacional, también dada a conocer a inicios de la administración actual, se establecen los medios para acelerar dicha inserción, es decir, que en este documento se plasman las acciones que el gobierno planea implementar o bien, que ya se encuentra implementando para aumentar “la digitalización de México, para que con ello se maximice su impacto económico, social y político en beneficio de la calidad de vida de las personas” (Gobierno de la República, 2013a, p.15).

Esta Estrategia Digital Nacional se centra en cinco puntos, los cuales reconoce como objetivos. Éstos son: Transformación Gubernamental, Economía Digital, Educación de Calidad, Salud Universal y Efectiva y Seguridad Ciudadana. Además se identifican los cinco habilitadores para alcanzar los objetivos de esta estrategia, los cuales son: Conectividad, Inclusión y Habilidades Digitales, Interoperabilidad, Marco Jurídico y Datos Abiertos.

Como ya se explicó en el capítulo anterior el concepto de Datos Abiertos surge con la idea de que la información en manos de la administración pública es de todos los ciudadanos, pues fue generada con sus impuestos. Se reconoce también la necesidad de publicarlos en formatos abiertos e interoperables para que puedan ser utilizados sin restricciones de ningún tipo. México cuenta con un portal para la publicación de datos abiertos<sup>8</sup> generados por las diferentes dependencias de gobierno. Estos datos abarcan un conjunto de temas, que incluyen: Salud, Seguridad y justicia, Energía y Medio ambiente, Educación, Economía, entre otros. Una categoría de datos existente en este portal, es la Geoespacial, a través de la cual se pueden descargar datos georreferenciados provenientes no sólo del INEGI, sino también de otras instituciones como la Secretaría de Desarrollo Agrario, Territorial y Urbano (SEDATU), CONABIO, Consejo Nacional de Población (CONAPO), etc. Algunos de los datos publicados a través de este portal no cuentan con metadatos y algunos otros no incluyen una liga hacia su diccionario de campos, por lo que a pesar de cumplir con el objetivo de que sean públicos y reutilizables, se hace más complicado el proceso de obtener información a partir de ellos.

De aquellos conjuntos de datos que sí tienen metadatos algunos no están bajo la NTM o peor aún carecen de algunos elementos que facilitarían la comprensión del contenido de las capas de datos geográficos que se descargan y su consiguiente reutilización. Con base en la experiencia de la U-GITS documentando capas de datos geográficos, se ha detectado que algunos elementos de la norma no son lo suficientemente claros o que por ejemplo en algunos casos no aplica su llenado y sin embargo resultan ser obligatorios. Aunado a esto, el uso del SICAM, si bien permite el llenado de metadatos utilizando la norma del INEGI, no siempre resulta una aplicación amigable y menos si se tiene poco conocimiento acerca de la norma de metadatos que emplea o de lo que significan los elementos de la misma. Situaciones similares a esta propician que los usuarios prefieran el uso de otras herramientas para generar metadatos geográficos, que no cumplen con el estándar del INEGI. Éstas, entre otras situaciones que se muestran en la figura 2.5, crean la necesidad de contar con otros sistemas de metadatos que faciliten la captura de metadatos y que además permitan la conexión con catálogos de metadatos basados todos en la misma norma para crear redes de metadatos en las que sea más fácil y eficiente la búsqueda de metadatos y de información geográfica.

---

<sup>8</sup><http://datos.gob.mx/>

## 2.2. Análisis del problema

En este trabajo de investigación se propone una solución informática para el problema de almacenamiento y publicación de metadatos, por lo menos para la U-GITS. La solución empleada para cubrir esta necesidad fue el diseño y desarrollo de un sistema de captura, almacenamiento, edición y búsqueda de metadatos en línea que sigue los lineamientos de la NTM del INEGI, el cual fue denominado Sistema Gestor de Metadatos (SIGEM). Para la creación del mismo, se siguió la metodología mostrada en la figura 2.2, la cual consta de cuatro etapas. En el resto del capítulo se da una explicación de lo realizado en cada una de las etapas de esta metodología.

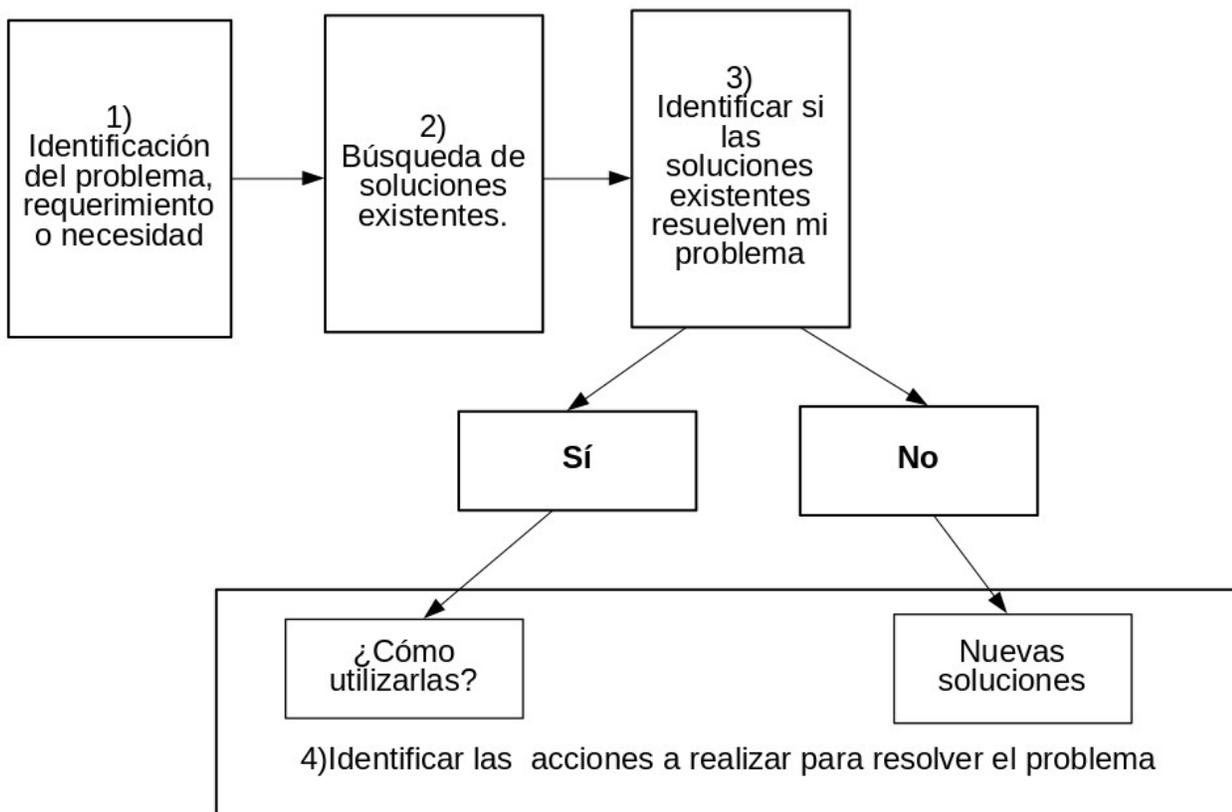


Figura 2.2. Etapas para la resolución del problema del sistema de metadatos.

### Etapa 1: Identificación del problema, requerimiento o necesidad.

La primera etapa consistió en identificar cuál es el problema que se quiere resolver. En este aspecto, la U-GITS, entre otras actividades, realiza proyectos en los que desarrolla portales web para la publicación de capas de datos geográficos (por medio de servidores WMS) para instituciones como el INEE o la SCT y dada la naturaleza de estos proyectos, los datos publicados en los portales desarrollados son considerados como información de interés nacional de acuerdo con los lineamientos de la LSNIEG. Es por esta razón que dichas capas deben ir acompañadas de su respectivo metadato, bajo la NTM.

Con base en la experiencia obtenida a partir de estos proyectos se detectó, entre otras cosas, lo siguiente:

- *Los formatos de archivo de los metadatos generados con el SICAM no son óptimos para el manejo de un portal de datos geográficos web.* Con el SICAM, los metadatos generados se almacenaban en archivos con formato XML, PDF, RTF, TXT y HTML, pues son los formatos en que este sistema almacena los metadatos dentro de la computadora del usuario. Almacenar los metadatos en más de uno de estos formatos provocaba que se tuviera información repetida y dificultaba la modificación así como la búsqueda de los mismos.
- *Existe gran cantidad de campos vacíos en el formato de metadato de la NTM.* No siempre se llenan todos los elementos del metadato debido a la falta de información, porque no se cuenta con ese elemento o porque la capa de datos no lo requiere. En cualquiera de esos tres casos, los metadatos suelen quedar con muchos espacios vacíos lo que puede dificultar su uso y confundir a usuarios no experimentados.
- *Se tienen diferentes versiones de un mismo metadato.* Cuando es necesario realizar cambios en las capas publicadas en un servidor de datos geoespaciales, dichos cambios también deben reflejarse en los metadatos de estas capas. Ésto provocaba que se tuvieran diferentes versiones de un mismo metadato y podía llegar a resultar complicado encontrar la última versión de los mismos.
- *Se requieren catálogos de metadatos.* Relacionado con el punto anterior, al no contar con un catálogo de metadatos resultaba más difícil rastrear los metadatos dentro del sistema de archivos, lo que implicaba un mayor tiempo en la modificación de los metadatos que así lo requirieran.

Aunque gestionar los metadatos con el SICAM<sup>9</sup> representa la ventaja de contar con una aplicación que permita su captura y exportación bajo los lineamientos de la NTM, no tiene la capacidad de estructurar la información en un repositorio centralizado que permita la búsqueda, rápido acceso y publicación de los metadatos para los portales desarrollados por la U-GITS. Por estas razones se pensó en desarrollar un sistema que permitiera la creación, edición y búsqueda de metadatos que cubran los aspectos descritos en la NTM del INEGI y que además se apegue a las directrices de desarrollo de sistemas de la U-GITS. Los requerimientos para este sistema se describen más adelante dentro de este capítulo.

Finalmente, a pesar de que el sistema presentado en este trabajo de tesis fue pensado inicialmente para la U-GITS, también se busca que pueda ser utilizado por otras dependencias o institutos que se encarguen de la generación y/o publicación de información geográfica.

## **Etapa 2: Búsqueda de soluciones ya existentes.**

Como resultado de la primera etapa del proceso mostrado en la figura 2.2 se detectó que era necesario contar, con un sistema de almacenamiento de metadatos que permita la captura y edición de los mismos para su posterior publicación dentro de los portales de datos geográficos desarrollados por la U-GITS.

En la figura 2.3, se muestra un cuadro comparativo entre varias aplicaciones y sistemas de captura y catálogos de metadatos que son utilizados actualmente. Se han tomado en cuenta los siguientes aspectos: las normas de metadatos con que trabajan, si son de código abierto o no, la plataforma (o sistema operativo) sobre la que se puede ejecutar la aplicación, su lenguaje de desarrollo, si es de escritorio o web y si se conecta con (o es) un servidor CSW.

<sup>9</sup>[http://www.inegi.org.mx/geo/contenidos/metadatos/aplicaciones\\_conversion.aspx](http://www.inegi.org.mx/geo/contenidos/metadatos/aplicaciones_conversion.aspx)

Software	Normas	Código Abierto	S.O.	Lenguaje de desarrollo	Ambiente de la aplicación	CSW	Últ. versión
MetaD	ISO 19115 ISO 19119 ISO 19139 Dublin Core INSPIRE	✗	Windows	Java	Escritorio	✓	4.0.2
ISO Metadata Editor "IME"	ISO 19115 ISO 19139	✗	Windows Linux	Java	Escritorio	✗	4.1 (2006)
ESRI ArcCatalog	ISO 19115 ISO 19139 CSDGM NAP INSPIRE	✗	Windows	XSL C++ Visual Basic	Escritorio	✗	ArcGIS 10.4.1
ESRI Geoportal Server	ISO 19115 ISO 19139 CSDGM NAP INSPIRE Dublin Core	✓	Multiplata forma	Java, Javascript, XSL, Flex, Silverlight	En línea	✓	1.2.6 (2015)
CatMDEdit	ISO 19115/19119/ 1939 NEM CSDGM Dublin Core	✓	Multiplata forma	Java XSL	Escritorio	✗	5.0 (2014)
Geonetwork	ISO 19115/19119 Dublin Core CSDGM	✓	Multiplata forma	Java Javascript	En línea	✓	3.0.4 (2016)
MapInfo Manager	ISO 19115	✗	Windows	C C++ Visual Basic	En línea	✓	
deegree	INSPIRE Dublin Core ISO 19115	✓	Multiplata forma	Java	En línea	✓	3.3.18 (2016)
Geoserver	CSDGM Dublin Core ISO 19115 ISO 19139 ISO 19119	✓	Multiplata forma	Java	En línea	✓	2.9.1 (2016)
pycsw	CSDGM Dublin Core ISO 19115 ISO 19139	✓	Multiplata forma	Python	En línea	✓	2.0 (2016)

**Figura 2.3.** Comparación entre aplicaciones y servicios de catálogo de metadatos. Fuente: elaboración propia con datos de Pitney Bowes (2016), Federal Geographic Data Committee (2014), Díaz et al. (2014) y sitios web oficiales de las aplicaciones.

De acuerdo con lo mostrado en la figura 2.3, un punto importante a destacar es que el 60 % de los software presentados son de código fuente abierto y se distribuyen de manera gratuita (aunque la primera condición no obliga a cumplir con la segunda). Se observa también que el 60 % de las aplicaciones mostradas son aplicaciones que funcionan en línea, mismas que permiten la conexión con servidores de catálogo (CSW) o bien lo son, como el caso de: Geonetwork o pycsw. Se destaca también el caso de MetaD que a pesar de ser una aplicación de escritorio, permite la conexión con servidores de catálogo. Además el 70 % de estas aplicaciones mostradas tienen la capacidad de ser ejecutadas en más un sistema operativo lo cual está relacionado con las tecnologías de desarrollo así como el (los) lenguaje(s) de desarrollo, pues vemos que en todos estos casos uno de los lenguajes de desarrollo es java (excepto en pycsw, que emplea python).

### 2.2.1. Análisis FODA

Esta sección del capítulo es la tercera etapa de la metodología mostrada en la figura 2.2. En ésta se muestra el proceso a través del cuál se llegó a la conclusión de crear un sistema de metadatos propio de la U-GITS; pero antes de abordarlo será conveniente considerar lo siguiente: qué es un análisis FODA y cuáles son los elementos que conforman a un sistema de información.

El análisis FODA es una herramienta de análisis que surge del contexto de la administración. Pickton y Wright (1998, p.101) dicen que se trata de una herramienta para categorizar factores ambientales internos y externos a una organización (pero también puede ser aplicado para proyectos, productos, etc) y es un auxiliar para la formulación de estrategias y toma de decisiones. Para el análisis FODA se consideran cuatro elementos: fortalezas, debilidades, oportunidades y amenazas. Éstos a su vez se clasifican en factores internos y externos, los cuales son:

- Factores internos
  - Las características que hacen a una organización o producto especial y le dan una ventaja frente a los competidores. Éstas son las **fortalezas**.
  - Los puntos débiles que lo ponen en desventaja frente a los competidores, es decir: las **debilidades**.
- Factores externos
  - Las **oportunidades** que se deben aprovechar y que servirían para colocar a la organización o producto en una posición favorable.
  - Las **amenazas** o factores que ponen en riesgo a la organización o un producto.

Finalmente, el análisis FODA puede organizarse en forma de matriz, como la mostrada en la figura 2.4.

		( + )	( - )
		Fortalezas	Debilidades
Interno	¿En qué eres bueno?		¿Qué puedes mejorar?
	¿Tienes algo que te diferencie?		¿Tienes menos ventajas que otros?
		Oportunidades	Amenazas
Externo	¿Qué oportunidades tienes a tu alcance?		¿Qué te podría distraer?
	¿De qué tendencias te puedes beneficiar?		¿Qué hace tu competencia mejor que tú?

Figura 2.4. Estructura del análisis FODA. Adaptado de Picón (2015)

Otro de los aspectos a considerar antes de abordar la siguiente etapa de resolución del problema de metadatos en la U-GITS es el de los elementos que conforman a un sistema de información. De acuerdo con Peña Ayala (2006, p.7), éstos son: financieros, administrativos, humanos, materiales y tecnológicos. Éstos se explican a continuación.

- **Financieros.** Permiten la adquisición, contratación y mantenimiento de los demás recursos que integran el sistema de información.
- **Administrativos.** Es la estructura orgánica de objetivos, lineamientos, funciones, procedimientos, departamentalización, dirección y control de actividades que sustentan la creación y uso de los sistemas.

Un aspecto fundamental de los sistemas de información son los datos que van a ser procesados para convertirse en productos de información, por lo que dentro de los elementos administrativos se deben contar aquellos procesos que permitan la adquisición y procesamiento de los datos.

- **Humanos.** Se divide en dos grupos:
  - *Técnicos.* Poseen los conocimientos para el desarrollo de sistemas. Pueden ser: administradores, líderes de proyecto, analistas, programadores, operadores, capturistas y aquellas personas encargadas de la adquisición y procesamiento de datos y de la creación de los productos de información.

- *Usuarios.* Las personas interesadas en el manejo de información mediante estos sistemas. Son los usuarios finales.
- **Materiales.** Elementos físicos que soportan el funcionamiento del sistema de información, por ejemplo: local de trabajo, instalación eléctrica, medios de comunicación, mobiliario, maquinaria, etc.
- **Tecnológicos.** Conjunto de conocimientos, experiencias, metodologías y técnicas que orientan la creación, operación y mantenimiento de un sistema. Dentro de los elementos tecnológicos también se cuenta el software, pues a través de éste es posible implementar los procedimientos que permiten el procesamiento de los datos para la generación de productos de información.

Estos elementos que conforman los sistemas de información se mencionan en esta sección ya que formarán parte del análisis FODA que se llevó a cabo como parte de la tercera etapa de la metodología mostrada en la figura 2.2.

### **Etapas 3: Identificar si las soluciones existentes resuelven mi problema.**

Al analizar el cuadro comparativo de la figura 2.3 es fácil observar que existen muchas soluciones al problema de implementar catálogos de metadatos o, una opción más avanzada, servicios de catálogo y, aunque el 80 % de estas aplicaciones pueden implementar más de una norma de metadatos además de la norma ISO 19115, ninguna de estas permite el uso de la norma de metadatos del INEGI.

Lo anterior sólo presenta un inconveniente, y ése es que de acuerdo con lo señalado en la Norma Técnica para la elaboración de Metadatos Geográficos del INEGI, todas las unidades de estado que generen información geográfica de interés nacional están obligadas a documentar dicha información bajo los términos de esta norma, **por lo que en realidad ninguna de tales aplicaciones podrían utilizarse por alguna unidad de estado en México** y tampoco para documentar la información espacial utilizada por la U-GITS en sus proyectos geoespaciales.

En la figura 2.5, se muestra un cuadro con el análisis FODA de las principales fuerzas alrededor de los sistemas de metadatos en México. En éste se muestran los elementos internos (Fortalezas y Debilidades) y elementos externos (Oportunidades y Amenazas). Estos cuatro elementos, a su vez, serán una lista categorizada tomando como referencia los elementos que conforman los sistemas de información.

Por otro lado, la matriz FODA puede ser usada como una herramienta para la formulación de cuatro tipos de estrategias. Las estrategias de fortalezas y oportunidades (estrategias FO) son aquellas en las que se busca utilizar las fortalezas para aprovechar las oportunidades externas. Las estrategias de debilidades y oportunidades (estrategias DO) buscan aminorar o superar las debilidades internas para poder aprovechar las oportunidades externas. Con las estrategias de fortalezas y amenazas (estrategias FA) se busca aprovechar las fortalezas para evitar o disminuir las repercusiones de las amenazas externas. Finalmente, las estrategias de debilidades y amenazas (estrategias DA) buscan disminuir las debilidades internas para evitar o disminuir las amenazas externas. (Rojas Enriquez, 2016)

En la figura 2.6 se muestra un cuadro con el listado de las estrategias FO, DO, FA y DA derivado del FODA de la figura 2.5 y a continuación se explica con más detalle en qué consisten estas estrategias.

## Análisis FODA de los sistemas generadores de metadatos en el contexto mexicano.

(+)

(-)

		Fortalezas	Debilidades
Interno	(+)	<p><b>Elementos humanos</b> F1. Se cuenta con capital humano capacitado que permite el desarrollo de sistemas informáticos especializados.</p> <p><b>Elementos tecnológicos</b> F2. INEGI desarrolló el SICAM. Es una herramienta que se apega a la NTM. F3. INEGI desarrolló el SICOM, una herramienta que permite la conversión de metadatos del estándar CSDGM a NTM.</p> <p><b>Elementos financieros</b> F4. El gobierno mexicano ha realizado inversiones para el desarrollo de herramientas informáticas que faciliten la gestión de los metadatos de acuerdo a la NTM. Estas son el SICAM y el SICOM y son de libre uso.</p> <p><b>Elementos administrativos</b> F5. Existen dependencias que generan una gran cantidad de datos geográficos: INEGI, CONAPO, CONABIO, SEMARNAT, etc. F6. Implementación del Programa para un Gobierno Cercano y Moderno, que incluye la puesta en marcha de la Estrategia Digital Nacional así como la publicación de Datos Abiertos del gobierno federal y gobiernos locales (fomento a la publicación de datos geoespaciales en manos de dependencias gubernamentales que necesitarán ser documentados).</p>	<p><b>Elementos humanos</b> D1. Falta de personal capacitado en temas referentes a metadatos geográficos.</p> <p><b>Elementos tecnológicos</b> D2. Sólo existe el SICAM para documentar con la NTM. D3. Puede resultar complicado instalar el SICAM o SICOM cuando se presenta un problema con la versión de Java instalada en el equipo. D4. SICAM no es una herramienta web (lo que facilitaría su uso; pues los usuarios finales ya no tendrían que preocuparse por el aspecto de la instalación). D5. No hay soluciones basadas en la NTM que permitan crear o conectarse con catálogos de metadatos.</p> <p><b>Elementos administrativos</b> D6. Pocas instituciones capacitan personal en temas referentes a metadatos geográficos y/o IDEs. D7. Poca interoperabilidad del SICAM (para poder utilizar otras normas de metadatos).</p> <p><b>Elementos materiales</b> D8. En las organizaciones hay pocos o nulos recursos de infraestructura dedicada a la gestión y servicio de los metadatos.</p> <p><b>Elementos financieros</b> D9. Generar datos geográficos y documentarlos considerando todos los elementos de la NTM (en especial de la sección 6) implicaría destinar una mayor cantidad de recursos (humanos y financieros) a la creación de los mismos.</p>
	(-)	Oportunidades	Amenazas
Externo	(+)	<p><b>Elementos tecnológicos</b> O1. Existe una gran cantidad de tecnologías (herramientas y marcos de trabajo) que permitirían la creación de un sistema de metadatos con base en la NTM. O2. Cada vez se cuenta con más herramientas de software libre para la creación de cualquier tipo de sistema.</p> <p><b>Elementos administrativos</b> O3. Sólo INEGI cuenta con un catálogo de metadatos basado en la NTM (Otras unidades podrían crear catálogos de metadatos basados en la NTM y unificar catálogos). O4. Las IDE (y todo lo que éstas comprenden) se están volviendo un tendencia mundial. En cuanto a metadatos, esto implica la creación de catálogos de metadatos para las mismas, pues los metadatos posibilitan el correcto intercambio de datos entre instituciones y países. O5. Existe gran cantidad de datos e información que requiere metadatos, por lo que se requieren profesionistas especialistas en el tema.</p> <p><b>Elementos materiales</b> O6. Cada vez más hay equipos con mayor capacidad de procesamiento y almacenamiento.</p> <p><b>Elementos financieros</b> O7. Muchas de las herramientas para creación de sistemas informáticos están disponibles de manera gratuita.</p>	<p><b>Elementos tecnológicos</b> A1. Existen varias soluciones para capturar metadatos geográficos que no están basadas en la NTM. A2. Muchas de estas aplicaciones permiten el uso de más de un estándar de metadatos. A3. Muchas herramientas de metadatos existentes (no basadas en la NTM) permiten la creación y/o conexión con catálogos de metadatos.</p> <p><b>Elementos administrativos</b> A4. Se generan capas y se hace poca documentación de las mismas, en parte porque el llenado de la NTM puede resultar complicado para los usuarios. A5. Algunas dependencias de gobierno como CONABIO documentan con normas distintas a la NTM, como CSDGM o ISO 19115, pues las aplicaciones disponibles actualmente permiten llenar automáticamente varios elementos de otras normas. A6. La NTM puede resultar insuficiente o poco adecuada en algunos de sus elementos.</p> <p><b>Elementos materiales</b> A7. No se han realizado esfuerzos suficientes (para la creación de infraestructura y mucho menos de administración y mantenimiento de la misma), para la implementación de redes de metadatos a nivel nacional.</p> <p><b>Elementos financieros</b> A8. Muchas de las soluciones basadas en otras normas son de libre uso y/o de software libre (lo que facilita el acceso y uso de las mismas). A9. En instituciones como SEMARNAT, INEE, entre otras, se gastan grandes cantidades de dinero para el pago de licencias comerciales de SIG. Este dinero podría invertirse en soluciones de software libre dedicadas también a la creación de redes de metadatos basadas en la NTM.</p>
	(-)		

Figura 2.5. Análisis tipo FODA de situaciones alrededor de los sistemas de metadatos en México.

	Fortalezas	Debilidades
Oportunidades	<p>Estrategias FO (maximizar fortalezas y oportunidades)</p> <ul style="list-style-type: none"> <li>• F1-O1-O2-O6. Difusión sobre la utilidad de los metadatos así como de los catálogos de metadatos de Información Geográfica.</li> <li>• F2-F3-F4-O3-O4. Promover el uso de la NTM.</li> <li>• F5-F6-O2-O3-O4-O5-O6. Aprovechar la tendencia de creación de IDE.</li> </ul>	<p>Estrategias DO (minimizar debilidades aprovechar oportunidades)</p> <ul style="list-style-type: none"> <li>• D2-D3-D6-D7-O1-O2. Considerar el SICAM para hacer nuevas aplicaciones.</li> <li>• D4-D5-O3-O4. Desarrollo de un sistema web de metadatos.</li> <li>• D1-D6-O3-O4. Dar mayor difusión acerca de la importancia de los metadatos.</li> </ul>
Amenazas	<p>Estrategias FA (maximizar fortalezas y minimizar amenazas)</p> <ul style="list-style-type: none"> <li>• F1-A1-A2-A3-A4. Diseño y desarrollo de una aplicación de metadatos basada en la NTM que permita la conexión con catálogos de metadatos.</li> <li>• F1-A6. Modificar la NTM.</li> </ul>	<p>Estrategias DA (minimizar debilidades y amenazas)</p> <ul style="list-style-type: none"> <li>• D2-D3-D7-A1-A2-A5. Promover la creación de una herramienta de metadatos que permita el manejo de más de un estándar de metadatos.</li> <li>• D4-A3. Promover el uso y creación de aplicaciones web de captura y catalogación de metadatos.</li> </ul>

**Figura 2.6.** Estrategias FA, FO, DA y DO en torno a la creación y utilización de sistemas de metadatos en México.

#### Estrategias FO (maximizar fortalezas y oportunidades)

- *F1-O1-O2-O6. Difusión sobre la utilidad de los metadatos así como de los catálogos de metadatos de información geográfica.* Informar tanto a funcionarios de las unidades del Estado como a desarrolladores acerca de la importancia de contar con datos y catálogos de metadatos con el propósito de contar con sistemas de captura y catálogos (o redes de catálogos) de metadatos

que ayuden a los usuarios a encontrar la información de su interés y así evitar la duplicidad de esfuerzos en la generación de IG que puede ya existir.

- *F3-F4-O3-O4. Promover el uso de la NTM.* Promover el uso de metadatos (basados en la NTM) dentro de las unidades del Estado encargadas de la generación de información geográfica de interés nacional, por lo que sería importante contar con herramientas o sistemas que permitan la generación de estos metadatos.
- *F5-F6-O2-O3-O4-O5-O6. Aprovechar la tendencia de creación de IDE.* En el mundo, la implementación de IDE, así como la creación de programas de Datos Abiertos en el sector gubernamental se está volviendo una tendencia. Bajo estos marcos y para el caso de México puede pensarse en la creación de redes de catálogos de metadatos, unificados bajo una misma norma, que ayuden a los usuarios a encontrar la información que están buscando.

#### **Estrategias DO (minimizar debilidades y maximizar oportunidades)**

- *D2-D3-D7-O1-O2. Considerar el SICAM para hacer nuevas aplicaciones con mejoras en su diseño* para que su instalación no dependa de otros programas instalados en la computadora del usuario, así como hacer que su interfaz sea más intuitiva y su uso resulte más sencillo.
- *D4-D6-O2. Considerar el SICOM para nuevas herramientas* que sean más fáciles de utilizar y que los metadatos, ya existentes, basados en la CSDGM puedan ser integrados de una forma más sencilla dentro de catálogos de metadatos basados en la NTM.
- *D4-D5-O3-O4. Desarrollo de un sistema web de metadatos* que permita la creación y conexión con catálogos de metadatos (bajo los lineamientos de la NTM) permitiría un mayor acceso a este tipo de herramientas así como la publicación de los metadatos, lo que resultaría en un mayor acceso a los datos.
- *D1-D6-O3-O4. Dar mayor difusión acerca de la importancia de los metadatos.* Promover actividades y eventos de difusión para mostrar la importancia de contar con IDEs, así como de documentar la información dentro de las mismas. De igual forma, fomentar la creación de redes de catálogos basados en la NTM, así como promover su uso dentro de las dependencias que aún no lo ocupan como norma de metadatos.

#### **Estrategias FA (maximizar fortalezas y minimizar amenazas).**

- *F1-A1-A2-A3-A4. Diseño y desarrollo de una aplicación de metadatos basada en la NTM que permita la conexión con catálogos de metadatos.* Fomentar la creación de redes de catálogos basados en la NTM, así como promover el uso de la NTM dentro de las dependencias que aún no lo ocupan como norma de metadatos.
- *F1-A6. Modificar la NTM.* Hacer modificaciones en la NTM para simplificar su uso. De igual forma sería conveniente rediseñar las herramientas disponibles para que el llenado de metadatos sea una tarea menos complicada y/o tediosa.

#### **Estrategias DA (minimizar debilidades y amenazas)**

- *D2-D3-D7-A1-A2-A5. Promover la creación de una herramienta de metadatos que permita el manejo de más de un estándar de metadatos.* Contar con una herramienta que permita el uso de más de un estándar de metadatos (incluida la NTM) podría facilitar que aquellas dependencias cuyos metadatos están hechos con otras normas puedan hacer sus catálogos compatibles con el estándar de INEGI y apegarse así a lo establecido en la NTM.

- *D4-A3. Promover el uso y creación de aplicaciones web de captura y catalogación de metadatos.* Contar con aplicaciones web de captura y catalogación de metadatos basados en la NTM facilitaría el acceso de los usuarios y podría propiciar que aquellos usuarios que utilizan otras normas de metadatos comiencen a utilizar el estándar del INEGI.

En general y, de acuerdo con lo mostrado a lo largo de esta sección (incluyendo las matrices FODA) **es necesario mejorar la cultura de documentación que existe en el país**. Para ello se identifican tres líneas de acción: 1) mejorar la difusión acerca de la importancia de los metadatos para proveer información acerca de la calidad de los datos espaciales; 2) la creación de catálogos de metadatos e incluso de redes de catálogos que permitieran hacer más eficiente el descubrimiento de información geoespacial y 3) una herramienta de creación y/o edición de metadatos que resulte sencilla de usar para que los usuarios o aquellas personas encargadas de la documentación de este tipo de información encuentren esta tarea más sencilla de realizar.

#### **Etapas 4: Identificar las acciones a realizar para resolver el problema.**

Si bien, documentar capas de datos geográficos se puede hacer bajo el uso de los estándares internacionales existentes para crear metadatos, en el caso de México por disposición oficial las unidades de estado deben documentar los datos que generan y publican siguiendo los lineamientos marcados en la NTM del INEGI. Para tal propósito fue creado el SICAM; sin embargo también hacen falta otros sistemas que permitan la creación de catálogos y que además funcionen en la web, pues esto podría permitir un mayor acceso a y uso de este tipo de herramientas.

De igual forma se mencionaron otros puntos como dar una mayor difusión a la NTM así como de la importancia de los metadatos. Aunque esto en realidad concierne mayoritariamente al INEGI, pues es el encargado de la coordinación del SNIEG, también se pueden realizar esfuerzos desde la academia, que pueden ir desde incluir temas acerca de IDE o normatividad espacial en los planes de estudio de carreras relacionadas con el tema como: Geografía, Geomática o Geoinformática (por ejemplo) hasta realizar proyectos de investigación en la materia o bien eventos de difusión como: congresos, mesas, coloquios, etc.

En esta tesis se propone la creación de un sistema de captura de metadatos geográficos siguiendo la NTM del INEGI. Además este sistema contará con un buscador para los metadatos ya capturados y también tendrá la opción de exportar los metadatos en formato HTML, PDF y XML (este formato se apegará al esquema XML del INEGI). Es importante mencionar que no se tratará de una aplicación de escritorio, sino de un sistema web.

## **2.3. Documentación de requerimientos**

El desarrollo de software es un proceso que requiere de varias etapas. La primera de ellas y que es fundamental para este proceso, es la definición de los requerimientos del sistema, pues a partir de ellos puede estimarse el alcance del proyecto, cuantificando el trabajo requerido para la creación del producto.

Los requerimientos son aquellas capacidades o condiciones con que cumple un producto (o sistema) y que permiten a sus usuarios resolver un problema, lograr un objetivo o bien, cumplir con una especificación determinada (Schwalbe, 2010, p.179).

Schwalbe (2010, pp.181-182) menciona que existen diversas formas de reunir los requerimientos para un sistema como: realizar entrevistas (muy efectivas pero muy costosas), aplicar encuestas (son

efectivas sólo cuando las respuestas son honestas), observación (en especial para la mejora de un producto o procedimiento), entre otras.

En la siguiente subsección se aborda el proceso de identificación de requerimientos para finalmente mostrar el listado de requerimientos con que deberá cumplir este sistema, mismos que se listarán por medio de una matriz de requerimientos adaptada de una **Requirements Traceability Matrix (RTM)**(Schwalbe, 2010, p.182). Dichos requerimientos se clasificarán en cuatro rubros: requerimientos de datos, requerimientos de la aplicación, requerimientos tecnológicos y de organización. Conforme se avance en el desarrollo de esta sección se irá explicando en qué consiste cada uno de estos rubros.

### 2.3.1. Objetivos y alcance del sistema

Obtener el listado de requerimientos del sistema es un proceso que consta de varias etapas. La primera de ellas es la de identificar las necesidades de los usuarios, en este caso la U-GITS del IG - UNAM, lo cual documenta desde el inicio de este capítulo y específicamente en el inicio de la sección 2.2.

INEGI puso a disposición del público una herramienta para la captura de metadatos geográficos llamada SICAM, la cual (como se ha mencionado) funciona bajo los lineamientos de la NTM. El SICAM durante cierto tiempo ha sido utilizado por los miembros de la Unidad GITS para la documentación de capas geográficas procesadas y generadas por ellos mismos con el fin de cumplir, en diversos proyectos geotecnológicos en los que ha participado, con la misma NTM y la LSNIEG, en la que se señala que la información geográfica de interés nacional publicada por unidades del estado debe llevar su metadato siguiendo los parámetros de esta norma.

Conforme la cantidad de información a procesar y documentar crecía, se hacía más importante la necesidad de contar con un repositorio centralizado de metadatos de las capas geográficas que se publican a través de los portales geográficos desarrollados por la U-GITS.

Entonces en resumen, el requerimiento se define como: **contar con una herramienta para la creación y administración de metadatos geográficos con base en la NTM del INEGI y un repositorio centralizado para dichos metadatos. Ambos componentes de la aplicación deben permitir la rápida consulta, modificación y publicación, así como evitar problemas de versionamiento de estos metadatos y servir para que la U-GITS se apegue a lo establecido en la LSNIEG.**

La segunda etapa para obtener los requerimientos del sistema consistió en definir la forma en que la situación actual de la U-GITS le permitió cubrir sus necesidades. Ésto se determinó por medio de las estrategias FODA, pues para plantearlas se requirió tomar en cuenta, situaciones internas. Si bien, al final del análisis se concluyó que la mayor parte de estas estrategias competen al INEGI y otras pueden abordarse desde la academia, algunas de éstas pueden ser consideradas y adaptadas al contexto de la U-GITS. Dichas estrategias son:

- *F1-O1-O2-O6. Difusión sobre la utilidad de los metadatos así como de los catálogos de metadatos de Información geográfica.*
- *F3-F4-O3-O4. Promover el uso de la NTM.*
- *D2-D3-D7-O1-O2. Considerar el SICAM para hacer nuevas aplicaciones con mejoras en su diseño.*
- *D4-D5-O3-O4. Desarrollo de un sistema web de metadatos.*

- *F1-A1-A2-A3-A4. Diseño y desarrollo de una aplicación de metadatos basada en la NTM que permita la conexión con catálogos de metadatos.*
- *D4-A3. Promover el uso y creación de aplicaciones web de captura y catalogación de metadatos.*

Las anteriores estrategias fueron consideradas debido a que: **a)** la U-GITS cuenta con un grupo dedicado a la investigación en normatividad geoespacial que ofrece capacitaciones internas y externas en materia de uso de estándares para metadatos geográficos (entre otros temas) y que **b)** la U-GITS también cuenta con un grupo de investigación y desarrollo en materia de software, mismo que se encarga de desarrollar herramientas y servicios para aprovechar datos georreferenciados generados por otros grupos de la U-GITS o externos.

Estos dos factores son importantes porque el grupo de investigación en normatividad geoespacial detectó la necesidad de contar con una herramienta de creación de metadatos que ofreciera capacidades adicionales a las que ya posee el SICAM y además, es este mismo grupo el que tiene las facultades para difundir y capacitar en el uso de la NTM. Además, el grupo de investigación y desarrollo de software posee las facultades para implementar algunas de las estrategias FODA seleccionadas como aquella en la que se propone la mejora del SICAM (como que su instalación no dependa de otros programas instalados) y de manera más específica, aquella en la que se plantea el desarrollo de un sistema web de metadatos, pues esta resultaría la forma más fácil de acceder a un repositorio centralizado de metadatos y a su vez evitarle problemas de instalación a los usuarios finales.

Por otro lado, a través de la figura 2.3 se observó que actualmente ya existen varios productos que dan solución al problema de creación de catálogos. De este cuadro comparativo pueden destacarse algunos puntos clave como que gran parte de estas aplicaciones (7 de 10) fueron desarrolladas con el lenguaje de programación Java. Esto es importante debido a que esta tecnología posee características que permiten a los desarrolladores crear sistemas que funcionan en una gran variedad de sistemas operativos y también en ambientes web, de escritorio e incluso móviles. Otro aspecto favorable es que más de la mitad (6 de 10) de estas aplicaciones se distribuyen bajo licencias de software libre, por lo que son de código abierto y pueden gozar de gran popularidad. Otro aspecto de estas aplicaciones de metadatos que resulta fundamental es que todos los productos presentados, sean aplicaciones de escritorio o servidores CSW, se apegan a estándares internacionales, razón por la que manejan al menos el estándar ISO 19115 (aunque algunas otras manejan también los estándares ISO 19139 o 19119 así como el CSDGM, entre otros). Es debido a esto que ninguna de las aplicaciones presentadas puede ser empleada directamente para la U-GITS y por lo que se optó por desarrollar un sistema propio.

Una vez que se decidió crear un sistema de metadatos propio, lo que siguió fue definir algunos puntos importantes antes de comenzar a desarrollarlo mediante un método conocido como 5W1H, útil para dar respuesta a preguntas acerca de un problema, proceso, producto, etc. para determinar, por ejemplo, cómo desarrollarlo o mejorarlo. Estas preguntas son: qué (what), quién (who), dónde (where), cuándo (when) por qué (why), y cómo (how), aunque para esta última se pueden emplear las preguntas cuánto cuesta y cuánto tiempo lleva desarrollarlo. Para definir las características de este sistema de metadatos, las respuestas a estas preguntas se muestran a continuación.

## WHAT

**¿Qué?:** Un sistema web que permita la captura (previa validación) y edición de metadatos para las capas de información geográfica generadas y publicadas por la U-GITS a través de sus portales. Así

mismo permitirá la creación de un catálogo de metadatos que acceda a un repositorio centralizado. Este sistema también deberá devolver metadatos en formatos interoperables como HTML o PDF; pero principalmente en formato XML. Asimismo debe ser capaz de incorporar al repositorio metadatos que provengan de fuentes externas (en formato XML), siempre y cuando éstos se apeguen a la norma de metadatos geográficos del INEGI.

## WHY

**¿Por qué?:** Debido a que el tipo de capas publicadas por la U-GITS en sus proyectos pueden considerarse información de interés nacional, éstas deben documentarse utilizando la NTM del INEGI. También se han llegado a presentar situaciones en las que más de una persona está involucrada en la creación y edición de estos metadatos, por lo que contar con un sistema web que acceda a una base de datos centralizada permitirá a todos aquellos involucrados en esta tarea encontrar los metadatos que buscan de una forma más sencilla y de igual forma ingresar al sistema sin necesidad de instalar programas adicionales en sus equipos de cómputo. Bastará con que tengan un navegador de internet.

## WHERE

**¿Dónde se usará?:** Principalmente en las instalaciones de la U-GITS, aunque debido a que será web se puede acceder a él desde cualquier lugar.

**¿Dónde se desarrollará?:** Dentro de las instalaciones de la U-GITS, en el IG - UNAM. De igual forma, el sistema estará alojado en los servidores de este mismo grupo de trabajo.

## WHEN

**¿Cuándo se desarrollará?:** El sistema ya se encuentra en funcionamiento. Fue desarrollado entre Junio de 2013 y Diciembre de 2014.

**¿Cuándo se usará?:** El sistema comenzó a utilizarse en 2015 para apoyar la impartición de la 2a edición del Diplomado en sistemas de información geográfica con software libre, en el módulo dedicado a normatividad geoespacial.

## WHO

**¿Quién lo desarrollará?:** Para el desarrollo del sistema, en primer lugar, se requirió definir los marcos de trabajo y tecnologías a emplear para definir la arquitectura del sistema, la base de datos del mismo y el programa de trabajo a seguir. Estas actividades fueron coordinadas por el responsable de Innovación tecnológica del grupo de investigación y desarrollo de software de la U-GITS, Lic. Juan Carlos Cortés Ortiz. También se requirió de personal para la escritura del código de la aplicación y programación de las funcionalidades del sistema, actividades realizadas por el Ing. Jonathan Landin Zaragoza, por Alexis Jazmín Palomares López y por el Lic. Juan Carlos Cortés Ortiz.

**¿Quién lo usará?:** El sistema será utilizado por personal perteneciente a la U-GITS que tenga a su cargo la documentación de capas de información geográfica (personal del grupo de normatividad geoespacial) así como personal del grupo de Geografía de la Seguridad Vial de esta unidad, también por personal del grupo de investigación y desarrollo de software para que los metadatos capturados sean publicados a través de los sistemas desarrollados por este grupo de trabajo.

## HOW

**¿Cómo?** Para poner en marcha el sistema se requirió un servidor para alojar la aplicación web y

otro servidor para alojar la base de datos. Ambos pueden encontrarse en el mismo equipo si no es necesario que la aplicación permita el acceso a demasiados usuarios concurrentes. Además, por tratarse de una aplicación web, los servidores en que sea alojada deben tener asignada una IP pública para que puedan ser accedidos desde cualquier sitio.

### HOW MUCH

**¿Cuánto tiempo requirió desarrollarlo?:** Desarrollar el sistema tomó aproximadamente 1990 horas de trabajo repartidas en un periodo de alrededor de un año y medio: entre Junio de 2013 y Diciembre de 2014.

**¿Cuánto cuesta?** Para determinar el costo de desarrollo del sistema, deben considerarse tanto los costos directos (de los materiales o equipo empleados directamente en la realización del mismo y mano de obra), como los costos indirectos (entre los cuales puede considerarse por ejemplo: el pago de la luz, el sueldo de empleados administrativos, renta, publicidad, etc.). Para este ejercicio sólo se tomarán en cuenta los costos directos, es decir, el sueldo de los involucrados directamente en el desarrollo y de los equipos de cómputo utilizados. Considerando lo anterior, el costo directo aproximado ejercido para la realización de este sistema es de \$210,388.05. Este costo se estimó con base en lo siguiente:

- **Nómina de personal involucrado.** Para la realización de este sistema, estuvieron directamente involucradas 3 personas. El perfil profesional del encargado de este proyecto corresponde al de un arquitecto de software. SG Software Guru (2017), dice que estos profesionales ganan aproximadamente \$35,000.00 mensuales. Estimando que la jornada laboral es de 8 horas al día y que un mes tiene alrededor de 21 días laborables (de lunes a viernes) se tiene que se trabajan alrededor de 168 horas al mes. Al dividir el sueldo mensual entre las 168 horas laborables, un profesional con este perfil ganaría alrededor de \$208.33 por hora. Por otro lado, Ramírez González (2013) dice que un programador junior de Java gana alrededor de \$15,000.00 pesos al mes, por lo que se estima gana alrededor de \$89.29 por hora. Debido a que el desarrollo de este sistema se llevó a cabo de manera paralela con otros proyectos, para el cálculo del costo se estimó el número de horas que cada participante invirtió en cada una de las etapas para estimar el pago de salarios por el desarrollo de este sistema. En este rubro se obtuvo un estimado de \$203,630.95. Más adelante, en la tabla 2.12 se incluye la descripción de las etapas del desarrollo del sistema.
- **Costo del equipo de cómputo.** Para el desarrollo del sistema, contar con equipo de cómputo resulta imprescindible. Si bien, se adquirieron computadoras para cada involucrado, éstas no fueron utilizadas de forma exclusiva para el desarrollo del sistema, por lo que para estimar este costo se consideró, de acuerdo con Gobierno de la República (2012), que el tiempo de vida útil del equipo de cómputo es de alrededor de 3 años. Si se considera que éste se utiliza durante 8 horas al día, sólo en días laborables (igualmente se consideraron 21 días por mes), son 168 horas al mes durante las cuales se hace uso de este equipo. Entonces, durante los 3 años de vida útil del equipo, se estima que éste se utilizará cerca de 6048 horas; por lo que si se divide el costo del equipo entre el total de horas se obtiene que el costo por hora de uso de las computadoras del responsable del proyecto y de un programador es de \$3.52 (el costo de adquisición del equipo fue de aproximadamente \$21,266.00 –de una Workstation HP Z620 –, mientras que el costo por hora de uso del equipo de cómputo del programador 2 es de \$3.35 (el costo de adquisición de este equipo –Workstation HP Z200 –fue de alrededor de \$20,241.00). Finalmente el total del costo por uso de los equipos de cómputo fue de alrededor de \$6,757.10. La descripción de los equipos utilizados se encuentra en la sección de requerimientos tecnológicos.

En las figura 2.7 y 2.8 se muestra una tabla con el detalle del cálculo del costo estimado de desarrollo del sistema.

Etapa	Inicio	Fin	Días trabajados	Número de horas dedicadas			Subtotal
				A	B	C	
Definición de la arquitectura de datos	03/06/13	04/06/13	2	16	0	0	\$3,333.33
Definición de la arquitectura de software	05/06/13	11/06/13	5	40	0	0	\$8,333.33
Programación de la Bibl. del modelo de datos	12/06/13	16/07/13	16	6	128	24	\$14,821.43
Programación de la Bibl. de persistencia	17/07/13	31/07/13	11	6	55	55	\$11,071.43
Programación de la Bibl. de negocio	02/09/13	17/09/13	11	6	55	55	\$11,071.43
Programación del módulo de presentación	02/10/13	12/09/14	183	138	101	915	\$119,464.29
Capacitación	17/09/14	17/09/14	1	0	0	4	\$357.14
Implementación de nuevos requerimientos	03/10/14	16/12/14	52	6	16	364	\$35,178.57
Totales			281	218	355	1417	\$203,630.95

**Figura 2.7.** Costo del personal. **A**=Arquitecto de software, **B**=Programador Jr. 1, **C**=Programador Jr. 2. Fuente: elaboración propia.

Etapa	Inicio	Fin	Días trabajados	Número de horas dedicadas			Subtotal
				A	B	C	
Definición de la arquitectura de datos	03/06/13	04/06/13	2	16	0	0	\$56.26
Definición de la arquitectura de software	05/06/13	11/06/13	5	40	0	0	\$140.65
Programación de la Bibl. del modelo de datos	12/06/13	16/07/13	16	6	128	24	\$551.49
Programación de la Bibl. de persistencia	17/07/13	31/07/13	11	6	55	55	\$398.56
Programación de la Bibl. de negocio	02/09/13	17/09/13	11	6	55	55	\$398.56
Programación del módulo de presentación	02/10/13	12/09/14	183	138	101	915	\$3,902.63
Capacitación	17/09/14	17/09/14	1	0	0	4	\$13.39
Implementación de nuevos requerimientos	03/10/14	16/12/14	52	6	16	364	\$1,295.56
Totales			281	218	355	1417	\$6,757.10

**Figura 2.8.** Costo por uso de los equipos de cómputo. **A**=Arquitecto de software, **B**=Programador Jr. 1, **C**=Programador Jr. 2. Fuente: elaboración propia.

Es importante considerar también lo siguiente: cuando la vida útil de los equipos de cómputo termina, no se recupera nada por venderlos y también que aunque se obtuvo el costo directo aproximado de producción aún no se determina el precio de venta del sistema, pues para esto sería importante considerar impuestos, el valor de la utilidad que se espera obtener, costos de imprevistos, entre otros.

### 2.3.2. Listado de requerimientos

Posterior a definir las características principales del sistema, deben definirse los requerimientos del sistema. Al inicio de la sección 2.3 se mencionó que son cuatro rubros los que se tomarían en cuenta

para definirlos. Éstos son: requerimientos de datos, requerimientos de la aplicación, requerimientos tecnológicos y de organización.

Los requerimientos deben documentarse para ser analizados y gestionados. Una forma de documentarlos es mediante una RTM, la cual es una lista de requerimientos con varios atributos. De acuerdo con Schwalbe (2010, p.182), esta lista también debe incluir el estado de dichos requerimientos, con el propósito de verificar que los requerimientos son abordados.

### Requerimientos de datos

Este rubro de requerimientos cubre dos aspectos: el tipo (y formato) de datos que entrarán en el sistema y dónde serán almacenados. Estos dos aspectos conllevan a la definición de un modelo de datos que mejor satisfaga las necesidades de almacenamiento de datos.

En la tabla 2.2 se muestra la lista de requerimientos de datos para este sistema. La matriz de requerimientos de datos de este sistema está adaptada de una RTM y contiene tres columnas. En la primera columna se indica la **fuer**te de cada requerimiento (es decir, lo que dio origen al mismo), en la segunda se especifica el **requerimiento** en sí y en la tercera se da una breve **descripción** de lo que permitirá hacer ese requerimiento.

Esta misma estructura se utilizará para las tablas de los otros tipos de requerimiento.

**Tabla 2.1.** Matriz de requerimientos de datos

Fuente	Requerimiento	Descripción
Todos los metadatos ingresados se encontrarán en un repositorio de metadatos centralizado.	Creación de una base de datos para almacenar los metadatos	El sistema contará con una base de datos en la que se almacenarán todos los metadatos creados por todos los usuarios con el privilegio de crearlos.
Evitar que comiencen a haber inconsistencias en los elementos introducidos en la base de datos.	Carga y descarga de metadatos en formato XML	El sistema no permitirá que se almacenen metadatos incompletos, por lo que se ofrecerá la opción al usuario de descargar el avance del metadato que esté capturando en formato XML (con base en la NTM). También se ofrecerá al usuario la opción de abrir en el formulario de captura un metadato que provenga de un archivo XML, mismo que posteriormente podrá ser guardado en la base de datos.
Los usuarios podrán capturar metadatos	Formulario para captura de metadatos	El sistema mostrará a los usuarios (con permiso) un formulario para la captura de los elementos del metadato.
Los usuarios podrán guardar un metadato en la base de datos		Dentro del formulario de captura de metadatos, el sistema dará al usuario la opción de guardar el metadato en la base de datos.

## Requerimientos de la aplicación

Este rubro cubre aspectos referentes a la funcionalidad del sistema como: la definición de los procesos que serán aplicados a los datos de entrada o cuestiones relacionadas con el mejoramiento de la experiencia del usuario con el sistema. Para la definición de requerimientos en este rubro, además de utilizar la RTM también se hará uso de un diagrama de bloques así como de casos de uso, con lo que quedarán explicados con mayor detalle las funcionalidades que debe ofrecer este sistema.

**Tabla 2.2.** Matriz de requerimientos de aplicación.

Fuente	Requerimiento	Descripción
Que los usuarios puedan acceder al repositorio de metadatos de una manera sencilla y sin importar donde se encuentren.	Creación de un sistema web	Se creará un sistema de captura y búsqueda de metadatos en línea. Estará instalado en un servidor y los usuarios accederán a él a través de un navegador de internet.
No todos los usuarios del sistema podrán editar un metadato.	Página de inicio de sesión	El sistema contará con una página de inicio de sesión y una vez que los usuarios se identifiquen a través de esta página recibirán los permisos asignados a su tipo de rol en el sistema.
	Menú de opciones	El sistema mostrará una barra con los menús disponibles para el usuario (búsqueda, carga y edición de metadatos).
Los usuarios podrán capturar metadatos	Formulario para captura de metadatos	El sistema mostrará a los usuarios (con permiso) un formulario para la captura de los elementos del metadato.
Los usuarios podrán guardar un metadato en la base de datos		Dentro del formulario de captura de metadatos, el sistema dará al usuario la opción de guardar el metadato en la base de datos.
Los usuarios podrán verificar cuáles son los elementos obligatorios que faltan por llenar en el metadato.	Validar metadatos	El sistema mandará mensajes al usuario indicando cuáles son los elementos obligatorios que le falta por llenar.
Obtener los metadatos en formato XML, PDF o HTML.	Capacidad de procesar peticiones internas o externas de exportar un metadato en formato PDF, XML o HTML.	El sistema contará con un buscador que permita a los usuarios encontrar los metadatos ya capturados. Por cada metadato arrojado por el buscador, el sistema dará al usuario la opción de exportarlo en formato PDF, XML o HTML.
	Buscador de metadatos	

Tabla 2.2 (continuación)

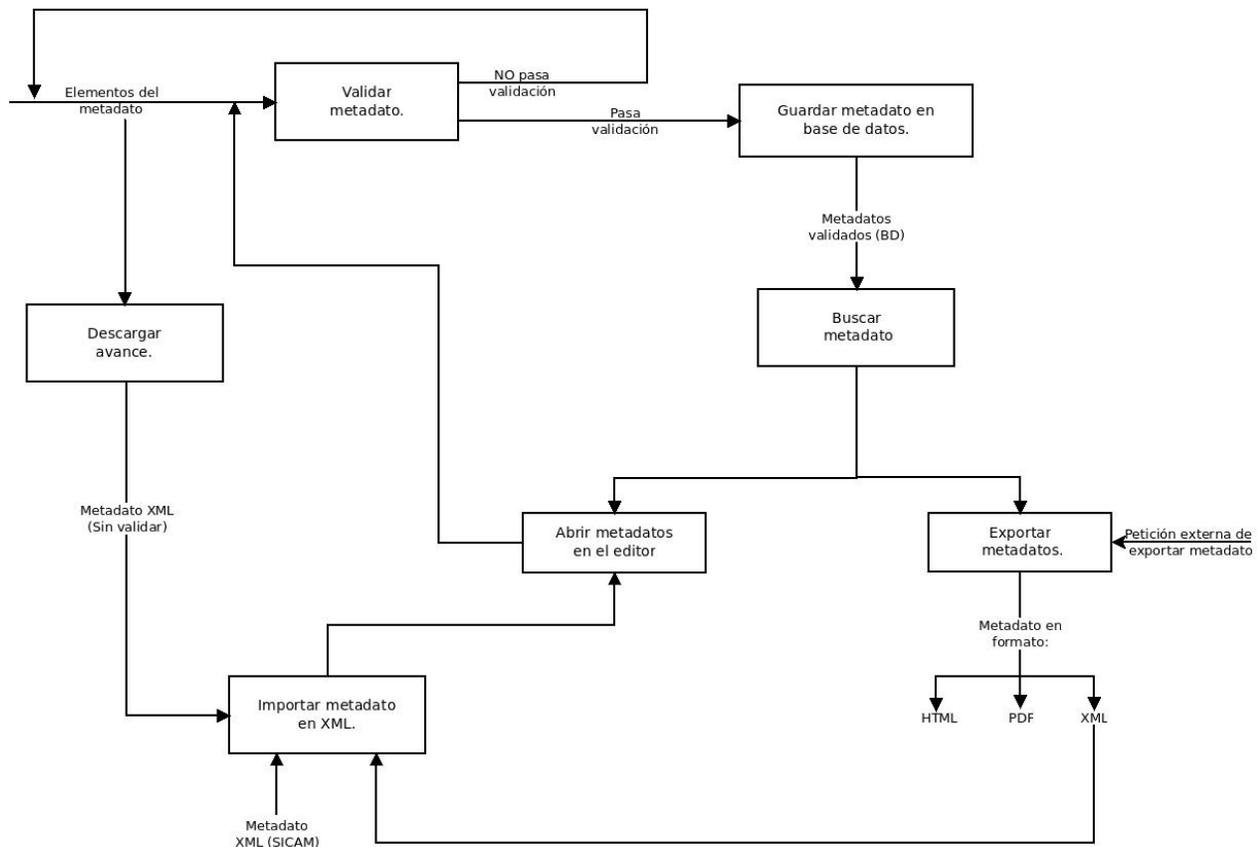
Fuente	Requerimiento	Descripción
Buscar metadatos ya existentes en el sistema	Buscador de metadatos	El sistema contará con un buscador que permita a los usuarios encontrar metadatos (dentro del sistema) por nombre del conjunto de datos, fecha de creación del metadato, por palabra clave o por nombre del usuario que capturó el metadato.
Permitir que los usuarios editen metadatos que ya se encuentren en la base de datos.	Formulario para captura de metadatos	El sistema contará con un buscador que permita a los usuarios encontrar los metadatos ya capturados. El usuario podrá abrir los metadatos arrojados por el buscador en el formulario de edición de metadatos y guardar los cambios hechos en el mismo.
	Buscador de metadatos	
Permitir que los usuarios creen un nuevo metadato a partir de otro que ya existe en la base de datos.	Formulario para captura de metadatos	El sistema contará con un buscador que permita a los usuarios encontrar los metadatos ya capturados. El usuario podrá abrir los metadatos arrojados por el buscador en el formulario de edición de metadatos y guardar los cambios hechos como un nuevo metadato (siempre y cuando el metadato esté completo).
	Buscador de metadatos	

### Funcionalidades requeridas del sistema

En la sección 1.1.1.4 se mencionó lo que son las cajas negras y que además éstas se utilizan para construir diagramas de bloques que ilustren la estructura de un sistema. En la figura 2.9 se muestra el diagrama de bloques de este Sistema Gestor de Metadatos (SIGEM). Cada una de las cajas negras representa una de las funcionalidades que tendrá el sistema, mientras que los arcos representan las entradas y salidas de esas cajas negras o funcionalidades. A través de este diagrama se ilustran las funcionalidades del sistema con base en las cuales puede pensarse también en la forma de desarrollarlo.

Como puede observarse en el diagrama, el sistema web cuenta con las siguientes funcionalidades.

- *Validar metadato.* Validará que el metadato que se está llenando a través del formulario tenga al menos todos sus campos obligatorios debidamente llenados. Si el metadato NO pasa la validación, en el formulario se indicarán los elementos que falta por llenar.
- *Guardar metadatos en la base de datos.* Una vez que el metadato que se está llenando pasa la



**Figura 2.9.** Diagrama de bloques del Sistema Gestor de Metadatos.

validación, éste es guardado dentro de la base de datos.

- *Buscar metadatos.* El sistema contará con un buscador sencillo que permita encontrar los metadatos ya almacenados en la base de datos.
- *Exportar metadatos.* Una vez que el buscador arroja los resultados encontrados, ofrecerá al usuario dos opciones, una de las cuales consiste exportar el metadato en formato PDF, HTML y/o XML. El formato XML se apega al generado con el SICAM del INEGI, permitiendo así la interoperabilidad de ese formato de metadatos.
- *Abrir metadatos en el editor.* Esta es la segunda opción para los resultados arrojados por el buscador de metadatos: editar o crear un nuevo metadato a partir de uno ya existente.
- *Importar metadato en formato XML.* Esta funcionalidad permite leer archivos XML cuya estructura sea igual a la de los archivos generados por el SICAM y los abrirá en el editor de metadatos, donde el usuario podrá validarlo y guardarlo en la base de datos.
- *Descargar avance.* El sistema permitirá descargar el avance del metadato que se esté editando en un archivo XML, mismo que podrá ser cargado posteriormente para continuar con el llenado del mismo.

### Casos de uso

Si bien, en este diagrama se muestran las funcionalidades del sistema, no se muestra la forma en que éstas operan. A través de los casos de uso se da mayor detalle acerca de éstas, aunque tampoco se describen los procesos internos.

Otra forma de reunir los requerimientos que debe cubrir el sistema es mediante el modelado de casos de uso, pues a través de ellos se identifican y modelan eventos del sistema, quiénes los inician o participan en ellos y cómo el sistema debe responder a ellos; pero ¿qué son los casos de uso? Rumbaugh et al. (2000) definen un caso de uso como la “Especificación de las secuencias de acciones, incluyendo secuencias variantes y secuencias de error, que pueden ser efectuadas por un sistema, subsistema o clase por interacción con autores externos”(p.157). Entonces, los casos de uso sirven para explicar una parte del comportamiento de un sistema; pero no revelan la estructura interna del mismo.

De acuerdo con Jacobson et al. (2000, pp.55-78), los casos de uso no sólo sirven para identificar los requerimientos del sistema, sino que también sirven para definir la forma en que será desarrollado el sistema, pues son la guía para el diseño de la arquitectura ya que los casos de uso deben encajar en la arquitectura y la arquitectura debe permitir el desarrollo de todos los casos de uso.

“El Lenguaje Unificado de Modelado (UML, por sus siglas en inglés) es un lenguaje de modelado visual que se usa para especificar, visualizar, construir y documentar artefactos de un sistema de software” (Rumbaugh et al., 2000, p.3). Este lenguaje fue creado principalmente para apoyar en los procesos de desarrollo orientados a objetos. Además es de propósito general, por lo que sus componentes no están directamente relacionados con ningún lenguaje de programación.

Rumbaugh et al. (2000), conocidos por ser los creadores de UML, definen las Vistas de UML como subconjuntos de UML que modelan construcciones que representan un aspecto de un sistema. Estas vistas son clasificadas en 3 categorías: las vistas de Clasificación estructural son aquellas en las que se representan los elementos del sistema y las relaciones entre estos elementos; las vistas de Comportamiento dinámico sirven para describir el comportamiento de un sistema en el tiempo; finalmente, las vistas de Gestión del modelo agrupan los elementos del sistema en unidades jerárquicas, normalmente denominadas *paquetes*.

Cada una de las vistas tiene asociado un grupo de conceptos, mismos que en conjunto con los elementos de UML sirven para la construcción de los diagramas asociados a cada vista. Por ejemplo, dentro de las vistas de clasificación estructural se encuentra la vista de casos de uso, la cual está ligada con los conceptos: caso de uso, actor, asociación, extensión, inclusión, generalización de casos de uso. A su vez, estos conceptos sirven para la construcción de diagramas de casos de uso. Es gracias a esta clasificación de vistas y diagramas, que con UML se pueden realizar diferentes representaciones de un mismo sistema, por lo que se pueden tener diferentes enfoques temáticos del mismo sistema.

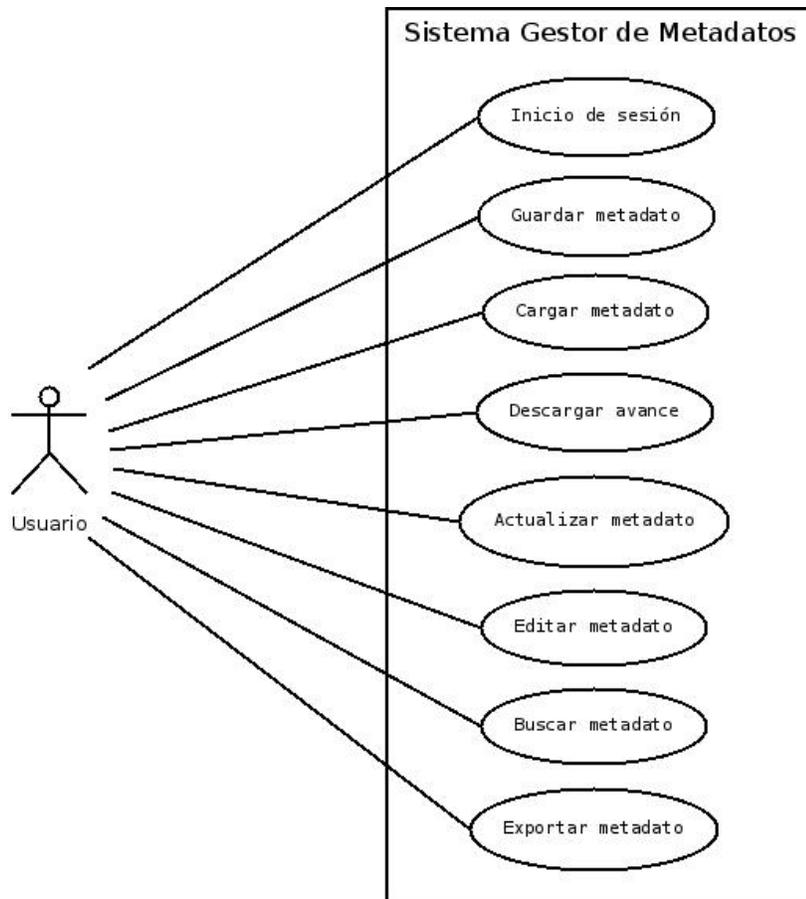
Aunque existen otras vistas, y por tanto otros diagramas que pueden ser construidos utilizando UML, en el párrafo anterior sólo se mencionó el diagrama de casos de uso por estar directamente relacionado con el tema de esta sección; pero en el siguiente capítulo se utilizarán otros diagramas para mostrar la arquitectura del sistema.

En este tipo de diagramas los casos de uso se representan con una elipse dentro de la cual va el nombre del caso de uso. Estas elipses se conectan mediante líneas continuas con los actores.

Los actores son, según Rumbaugh et al. (2000, p.56), la idealización de cualquier cosa (puede tratarse de una persona, de un proceso o de otro sistema) que interactúe con un sistema o una parte del sistema. En un caso de uso, los actores pueden ayudar a encontrar y definir la forma en que un usuario real puede llegar a interactuar con el sistema.

Los casos de uso de Sistema Gestor de Metadatos son: Inicio de sesión, Guardar metadato, Cargar

metadato, Descargar avance, Actualizar metadato, Editar metadato, Buscar metadato y Exportar metadato. Éstos se muestran en la figura 2.10.



**Figura 2.10.** Diagrama de casos de uso del Sistema Gestor de Metadatos.

En las siguientes tablas se muestra con mayor detalle la secuencia normal para cada uno de los casos de uso, sin tomar en cuenta las situaciones en que el sistema manda mensajes de error o no actúa de acuerdo a lo esperado por el usuario.

**Caso de uso: Iniciar sesión**

*Actores:* Usuario

*Descripción:* El usuario registrado podrá iniciar sesión en el sistema

*Precondiciones:* El usuario debe estar dado de alta en el Sistema gestor de Usuarios.

**Tabla 2.3.** Flujo normal del caso de uso “Iniciar sesión”

Acciones de los actores	Respuesta del sistema
1.- El usuario, a través de un navegador de internet, ingresa a la siguiente liga: <a href="http://www.gits.igg.unam.mx/sigem">http://www.gits.igg.unam.mx/sigem</a>	2.- Muestra la página de inicio de sesión

Tabla 2.3 (continuación)

Acciones de los actores	Respuesta del sistema
3.- El usuario ingresa el correo electrónico y la contraseña con la que se registró en el sistema.	
4.- Presiona el botón “Iniciar sesión”	5.- Muestra la página de inicio del sistema (Buscador de metadatos). Si el usuario cuenta con el permiso de edición de metadatos podrá ver en la parte superior de la ventana la opción “Editor”.

### Caso de uso: Editar metadato

*Actores:* Usuario

*Descripción:* El usuario podrá comenzar a llenar un metadato.

*Precondiciones:* Haber ejecutado con éxito el caso de uso “Iniciar sesión” y además contar con permisos de edición de metadatos.

Tabla 2.4. Flujo normal del caso de uso “Editar metadato”

Acciones de los actores	Respuesta del sistema
1.- El usuario presiona sobre la opción “Editor” que se encuentra en la barra de opciones (barra negra en la parte superior de la pantalla)	2.- Muestra nueve pestañas, cada una con el nombre de una sección de la NTM. Cada una de estas pestañas contendrá un formulario vacío con los elementos de la NTM que le correspondan. También muestra el botón “Opciones”.
3.- El usuario presiona sobre alguna de las 9 pestañas, según la sección que quiera llenar primero.	4. Despliega el formulario correspondiente a dicha sección.
5.- El usuario comienza a ingresar los elementos correspondientes a dicha sección de la NTM.	6.- Muestra mensajes en el extremo derecho de la pantalla del usuario si es necesario que éste vaya llenado otros elementos de otras secciones o de esa misma sección.

### Caso de uso: Descargar avance

*Actores:* Usuario

*Descripción:* El usuario podrá descargar el avance del metadato que esté llenando antes de guardarlo en la base de datos.

*Precondiciones:* Haber ejecutado con éxito el caso de uso “Editar metadato”.

Tabla 2.5. Flujo normal del caso de uso “Descargar avance”

Acciones de los actores	Respuesta del sistema
1.- El usuario presiona sobre el botón “Opciones” que se encuentra en el extremo derecho de su pantalla	2.- Despliega la lista de opciones disponibles para ese metadato.

Tabla 2.5 (continuación)

Acciones de los actores	Respuesta del sistema
3.- El usuario presiona sobre la opción “Descargar avances”.	4.- El sistema muestra mensajes sobre aquellos elementos obligatorios que falta por llenar.
	5.- Inicia la descarga del metadato en formato XML en el equipo del usuario.
	6.- Muestra en el navegador del usuario un mensaje indicando que el metadato se ha descargado.
7.- El usuario cierra la ventana en la que se muestra el mensaje (puede seguir editando el metadato si lo desea).	

### Caso de uso: Guardar metadato

*Actores:* Usuario

*Descripción:* El usuario podrá guardar el metadato que esté llenando en la base de datos.

*Precondiciones:* Haber ejecutado con éxito el caso de uso “Editar metadato”.

Tabla 2.6. Flujo normal del caso de uso “Guardar metadato”

Acciones de los actores	Respuesta del sistema
1.- El usuario presiona sobre el botón “Opciones” que se encuentra en el extremo derecho de su pantalla	2.- Despliega la lista de opciones disponibles para ese metadato.
3.- El usuario presiona sobre la opción “Guardar metadato”.	4.- El sistema busca que al menos todos los elementos obligatorios de todas las secciones estén debidamente llenados (Validación del metadato).
	4a.- <b>Termina caso de uso si</b> el sistema muestra mensajes sobre aquellos elementos obligatorios que falta por llenar.
	5.- Muestra en el navegador del usuario un mensaje indicando que el metadato se ha guardado.
6.- El usuario cierra la ventana en la que se muestra el mensaje.	7. Redirige al usuario a la página de inicio (Buscador de metadatos).

### Caso de uso: Buscar metadato

*Actores:* Usuario

*Descripción:* El usuario podrá buscar aquellos metadatos que ya se encuentren almacenados en la base de datos.

*Precondiciones:* Haber ejecutado con éxito al menos el caso de uso “Iniciar sesión” (puede haber ejecutado también “Editar metadato”).

Tabla 2.7. Flujo normal del caso de uso “Buscar metadato”

Acciones de los actores	Respuesta del sistema
1.- El usuario presiona sobre la opción “Inicio” que se encuentra en el extremo izquierdo de su pantalla dentro de la barra negra	
1a.- El usuario no ejecuta el paso 1 si ya se encuentra en la página de inicio	
	2.- Redirige al usuario a la página de inicio.
3.- ( <b>Este paso es opcional.</b> ) El usuario ingresa el nombre del metadato que quiere encontrar (este elemento corresponde al nombre del conjunto de datos) en el cuadro de texto “Nombre del metadato”.	
4.- ( <b>Este paso es opcional.</b> ) El usuario ingresa una o más palabras en el cuadro de texto “Palabra clave” si quiere encontrar metadatos que contengan dicha palabra clave.	
5.- ( <b>Este paso es opcional.</b> ) El usuario ingresa una o más palabras en el cuadro de texto “Usuario” si quiere encontrar metadatos creados por dicho usuario.	
6.- ( <b>Este paso es opcional.</b> ) El usuario ingresa una fecha en el campo "Desde" si quiere encontrar metadatos creados a partir de dicha fecha.	
7.- ( <b>Este paso es opcional.</b> ) El usuario ingresa una fecha en el campo "Hasta" si quiere encontrar metadatos creados hasta esa fecha.	
8.- El usuario presiona el botón “Filtrar”.	9.- El sistema arroja una lista de no más de 15 elementos con los metadatos encontrados, mostrando los siguientes elementos para cada uno de estos metadatos: Nombre del metadato, nombre del usuario que creó el metadato, fecha de creación del metadato y un conjunto de botones para exportar o editar ese metadato. También muestra el botón “Más resultados” en la parte inferior de la lista de resultados.
10.- ( <b>Este paso es opcional</b> ) El usuario presiona el botón “Más resultados” si quiere consultar el resto de los metadatos encontrados a través del buscador.	11. El sistema añade a la lista inicial de resultados otros 15 metadatos encontrados (si es que los hay).

**Nota:** el usuario puede repetir el paso 10 cuantas veces sea necesario hasta que el sistema le indique que ya no hay más resultados disponibles y el botón “Más resultados” desaparezca.

### Caso de uso: Exportar metadato

*Actores:* Usuario

*Descripción:* El usuario podrá obtener un metadato en formato PDF, HTML o XML.

*Precondiciones:* Haber ejecutado con éxito el caso de uso “Buscar metadato”.

**Tabla 2.8.** Flujo normal del caso de uso “Exportar metadato”

Acciones de los actores	Respuesta del sistema
1.- Cuando el usuario encuentra el metadato que quiere exportar presiona sobre el botón “pdf” si quiere exportar el metadato en formato PDF; sobre el botón “xml” para el formato XML; o sobre el botón “html” si desea exportar el metadato en formato HTML.	2.- El sistema abre una nueva ventana en el navegador del usuario y muestra en ella el metadato en el formato elegido por el usuario.

### Caso de uso: Actualizar metadato

*Actores:* Usuario

*Descripción:* El usuario podrá hacer cambios en uno de los metadatos que ya se encuentren en la base de datos.

*Precondiciones:* Haber ejecutado con éxito el caso de uso “Buscar metadato”.

**Tabla 2.9.** Flujo normal del caso de uso “Actualizar metadato”

Acciones de los actores	Respuesta del sistema
1.- Cuando el usuario encuentra el metadato que quiere actualizar presiona el botón de edición de metadatos que se encuentra en el extremo derecho del metadato deseado.	2.- El sistema redirige al usuario al editor de metadatos.
	3.- Carga el contenido del metadato dentro de los formularios del editor de metadatos.
4.- El usuario realiza los cambios que necesite sobre el metadato.	
5.- El usuario presiona sobre el botón “Opciones” que se encuentra en el extremo derecho de su pantalla.	6.- Despliega la lista de opciones disponibles para ese metadato.
7.- Presiona sobre la opción “Actualizar metadato”.	8.- El sistema busca que al menos todos los elementos obligatorios de todas las secciones estén debidamente llenados (validación del metadato).
	8a.- <b>Termina caso de uso si</b> el sistema muestra mensajes sobre aquellos elementos obligatorios que falta por llenar.

Tabla 2.9 (continuación)

Acciones de los actores	Respuesta del sistema
	9.- Muestra en el navegador del usuario un mensaje indicando que el metadato se ha actualizado.
10.- El usuario cierra la ventana en la que se muestra el mensaje.	11.- Redirige al usuario a la página de inicio (Buscador de metadatos).

### Caso de uso: Cargar metadato

*Actores:* Usuario

*Descripción:* El usuario podrá abrir en el editor de metadatos un metadato externo, siempre que el metadato se encuentre en formato XML.

*Precondiciones:* Haber ejecutado con éxito el caso de uso “Iniciar sesión”.

Tabla 2.10. Flujo normal del caso de uso “Cargar metadato”

Acciones de los actores	Respuesta del sistema
1.- El usuario presiona sobre la opción “Cargar metadato” que se encuentra en la barra negra en la parte superior de su pantalla.	2.- El sistema abre un cuadro con un formulario en el que el usuario puede ingresar su metadato.
3.- Presiona y botón examinar y busca el archivo XML que contiene el metadato que desea cargara en el sistema.	
4.- Presiona el botón “Enviar”	5.- El sistema procesa el archivo seleccionado por el usuario.
	6.- Si el archivo no contiene errores, el sistema redirige al usuario a la página de edición de metadatos.
	7.- El sistema carga el contenido del metadato dentro de los formularios del editor de metadatos.

## Requerimientos tecnológicos

En este rubro se definen aspectos relacionados con el software y el hardware empleado tanto para el desarrollo del sistema, así como para la puesta en marcha del mismo.

**Tabla 2.11.** Matriz de requerimientos tecnológicos.

Fuente	Requerimiento	Descripción
Se requiere de equipo de cómputo para el desarrollo y prueba del sistema que ofrezca estabilidad y buen rendimiento.	Estaciones de trabajo	Debido a que el equipo de cómputo no sólo se utiliza para el desarrollo de este proyecto, sino para otras aplicaciones (principalmente web) muchas veces estos equipos serán utilizados como servidores, al menos de forma temporal, durante la etapa de desarrollo y de pruebas de los sistemas, por lo que se requiere de computadoras que ofrezcan buenas prestaciones principalmente buena velocidad de procesamiento y buena cantidad de memoria RAM.
Con base en las funciones definidas para el sistema se debe elegir la tecnología (de software libre) que mejor cubra las necesidades de desarrollo y que se apegue a buenas prácticas.	JVM (Java Virtual Machine)	Cada uno de los equipos involucrados en el desarrollo de este sistema deberá tener instalado el software esencial para el desarrollo del sistema, el cual incluye: la plataforma Java, un sistema controlador de versiones y una herramienta para la construcción de proyectos Java.
	Sistema Controlador de Versiones (Git)	
	Herramienta para la construcción de proyectos Java (Maven).	
	IDE de desarrollo	

En la siguiente lista se muestran los equipos de cómputo utilizados para el desarrollo de este sistema así como el sistema operativo empleado en cada uno de estos equipos.

- **Servidor del sistema controlador de versiones y Responsable del proyecto.** Workstation HP Z620. Procesador Intel Xeon CPU ES-1607 0 @ 3.00 GHz. 16 GB de memoria RAM. 500 GB de almacenamiento. Sistema operativo: GNU Linux/Debian 7.
- **Programador 1.** Workstation HP Z620. Procesador Intel Xeon CPU ES-1607 0 @ 3.00 GHz. 16 GB de memoria RAM. 500 GB de almacenamiento. Sistema operativo GNU Linux/Debian 7.
- **Programador 2.** Workstation HP Z200. Procesador Intel Core i5 CPU 750 @ 2.67GHz. 8 GB de memoria RAM. 500 GB de almacenamiento. Sistema operativo GNU Linux/Fedora 19.

### Requerimientos de organización

En este rubro se cubren aspectos relacionados con el personal involucrado y con la definición de etapas de desarrollo del sistema.

El desarrollo del sistema involucró a tres personas. Sus roles fueron:

- *Responsable del proyecto:* Lic. Juan Carlos Cortés Ortiz. (JCCO)
- *Programador:* Ing. Jonathan Landin Zaragoza. (JLZ)
- *Programador:* P.I. Alexis Jazmín Palomares López. (AJPL)

En la tabla 2.12 se muestran las etapas de desarrollo, en qué consistieron, así como el personal involucrado en cada una de ellas.

**Tabla 2.12.** Etapas de desarrollo del Sistema Gestor de Metadatos.

Etapa	Descripción	Personal
Definición de la arquitectura de datos	Se creó el modelo relacional (lógico) de la base de datos del sistema; se estableció el <b>Sistema Manejador de Bases de Datos (SMDB)</b> a emplear y por último se implementó el modelo lógico en el SMDB seleccionado.	JCCO
Definición de la arquitectura de software del sistema.	Se definió la arquitectura de software para el sistema. También se definieron las tecnologías de desarrollo que mejor se adecuaban a los lineamientos de la arquitectura seleccionada. Por último, utilizando dichas tecnologías, se crearon los módulos del sistema (en ese momento sólo estaban especificadas las dependencias y la estructura de cada módulo) y se cargaron en el servidor del sistema de control de versiones.	JCCO
Programación de la Biblioteca del modelo de datos	En esta etapa se crearon clases que representan las tablas de la base de datos.	JCCO, JLZ, AJPL.
Programación de la Biblioteca de Persistencia	Esta etapa estuvo dedicada a la programación de las funciones encargadas de almacenar datos en la base de datos así como de extraerlos.	JCCO, JLZ, AJPL.
Programación de la Biblioteca de negocio	En este periodo se programaron las funciones que implementan la lógica del negocio. Estas funciones acceden a las que se encuentran en la Biblioteca de Persistencia.	JCCO, JLZ, AJPL.
Programación del módulo de presentación.	Durante este periodo se escribió el código para armar la interfaz de usuario de la aplicación, funciones de respuesta a eventos activados por el usuario y las funciones para exportar metadatos a través de peticiones externas.	JCCO, JLZ, AJPL.
Capacitación acerca del uso del sistema.	Se convocó a todo el personal de la U-GITS para mostrarles el funcionamiento del sistema. Al tiempo que se llevó a cabo esta sesión de capacitación se reunieron nuevos requerimientos para el sistema.	AJPL

Tabla 2.12 (continuación)

<b>Etapa</b>	<b>Descripción</b>	<b>Personal</b>
Implementación de nuevos requerimientos en el sistema	Se hicieron modificaciones en el sistema resultado de las observaciones hechas por el personal de la U-GITS durante la sesión de capacitación.	JCCO, JLZ, AJPL.

En el siguiente capítulo se abordará con mayor detalle las características de las tecnologías de desarrollo empleadas en este sistema.

## Capítulo 3

# Diseño del Sistema Gestor de Metadatos

Una vez definidos los requerimientos del sistema, el siguiente paso en la metodología de desarrollo fue el diseño del software, que de acuerdo con Pressman (2010, p.183) debe proporcionar detalles acerca de los siguientes puntos:

- **Arquitectura del sistema o producto.** El modelo de la arquitectura debe mostrar las relaciones existentes entre los principales componentes del software (módulos, objetos, clases, etc.). También debe basarse en patrones de diseño, es decir, en estructuras de diseño previamente probadas en la solución de otros problemas similares, pues esto permite la reutilización de bloques de construcción del sistema, lo que permitiría ahorrar tiempo para futuros diseños.  
La arquitectura de un sistema debe permitir cubrir los requerimientos del sistema (como los obtenidos con base en los casos de uso) y los modelos de dicha arquitectura deben ser comprendidos por los desarrolladores y otras partes interesadas (Jacobson et al., 2000, pp.55-78).
- **Interfaces.** En el diseño de interfaces se describe la forma en que el software se comunica con otros sistemas y con los usuarios finales. Ésto involucra el diseño de interfaces de usuario (la parte gráfica del sistema que los usuarios finales pueden ver); interfaces externas que tienen que ver con otros sistemas, dispositivos, redes o consumidores de la información del sistema (por ejemplo las APIs de algunos servicios web) e interfaces internas que se refieren a la forma en que se comunican entre sí los elementos internos del sistema.
- **Componentes.** Los componentes son partes funcionales de un sistema. Éstos tienen su propia lógica de procesamiento así como estructuras de datos que le permiten implementar la lógica de procesamiento. Visto de otro modo, es un módulo del sistema. Al llevar a cabo el diseño de componentes, además de definir esta lógica de procesamiento y las estructuras de datos propias de cada componente, también se deben incluir las interfaces que permiten el acceso a cada módulo. Los componentes forman la arquitectura.

Otro aspecto a considerar para el diseño del sistema son los datos, pues “En muchas aplicaciones de software, la arquitectura de los datos tendrá una influencia profunda en la arquitectura del software que debe procesarlo” (Pressman, 2010, p.199). A nivel de la arquitectura, el diseño de los datos se centra en archivos o bases de datos; pero a nivel de componentes, el diseño de datos se centra en la estructuras de datos que se requieren para implementar objetos de datos locales.

A continuación en la figura 3.1 se muestra la forma en que se conceptualizó la estructura del SIGEM. Si bien, este diagrama no representa la verdadera arquitectura del sistema, su intención es



contenidos dentro de los metadatos mediante el desarrollo de un modelo de datos que posteriormente se implementó dentro de un SMDB.

### 3.1.1. Modelo de datos

En esta sección se presenta el modelo de datos diseñado para el SIGEM, que es un **modelo relacional** estructurado con base en elementos de un modelo de datos **multidimensional**.

En primer lugar, un modelo relacional, según Tomlinson (2003, p.129), es aquel en el que los datos se almacenan en tablas asociadas lógicamente unas con otras mediante atributos compartidos. Estos modelos pueden ser de dos categorías: modelos lógicos o modelos físicos (Allen y Terry, 2006, p.7).

Los **modelos lógicos** se forman a partir de **entidades**, las cuales son abstracciones de elementos de la realidad y que tienen características de interés para quien modela los datos, o bien le ayudan a comprender el contexto que quiere modelar. Estas abstracciones pueden corresponder a una persona, lugar, cosa o concepto. Por ejemplo, de acuerdo con el interés para una organización, un conjunto de entidades de interés pueden ser: staff, proyecto, departamento (o área) y para otra quizá sean: productos, órdenes de compra, ventas, vendedores, etc. En el contexto que atañe a esta tesis, una entidad puede ser un *metadato*, o quizá un *punto de contacto para el metadato*.

Todas las entidades cuentan con **atributos**, los cuales son las características que definen a la entidad. Éstos deben tener un tipo de dato asignado (por ejemplo: texto, número o fecha). En el ejemplo de la entidad *punto de contacto para el metadato*, un atributo será el nombre del contacto y el tipo de dato de este atributo será texto.

Los atributos de una entidad pueden ser compartidos con otra entidad mediante una **relación** por medio de atributos que tengan en común. Estas relaciones tienen una **cardinalidad** y por medio de ella se especifica si uno o más elementos de una entidad podrán ser compartidos con otras entidades. Por ejemplo: un metadato tendrá asignado sólo un sistema de referencia horizontal y un sistema de referencia horizontal estará sólo en un metadato. En este caso, la cardinalidad de la relación será de 1 a 1 (1:1). Existen otros tipos de cardinalidad, como 1 a muchos (1:n) o muchos a muchos (n:n).

Las entidades tendrán **llaves**, es decir que tendrán uno o más atributos que servirán para identificar de manera única a cada instancia de una entidad. Estas son las llaves primarias. Pero también existen las llaves foráneas, que básicamente son atributos que permiten llevar a cabo las relaciones entre entidades.

Hasta ahora se han mencionado algunos de los elementos que componen a un modelo lógico de datos; pero también existen los **modelos físicos**. Éstos, representan la implementación de un modelo lógico en una base de datos. En este tipo de modelo, las entidades, por ejemplo, ya podrían convertirse en tablas y los atributos, en columnas de estas tablas; pero dicha conversión no siempre puede ser directa, pues para construir un modelo físico deben considerarse algunas restricciones o los tipos de datos soportados por el manejador de bases de datos en el que vaya a implementarse el modelo. En esta sección se mostrará el modelo lógico de datos del SIGEM.

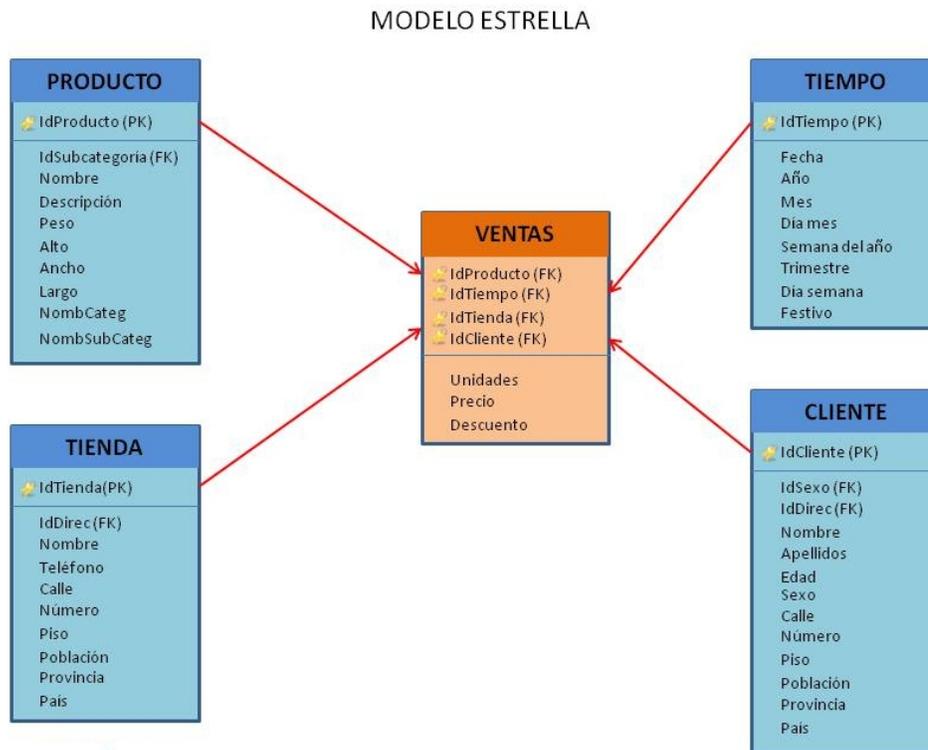
Todos los modelos de datos (ya sean lógico o físicos), deben trazarse con base en **reglas del negocio**, que en resumen, son los requerimientos y restricciones que deben contemplarse en el modelo de datos. Para este caso, las reglas del negocio se obtuvieron de la INEGI del INEGI.

Por otro lado, existe lo que se denomina **modelo de datos multidimensional**. Este tipo de modelado surge en el contexto de Business Intelligence (BI), que sirve para apoyar la toma de decisiones estratégicas para una organización. Dentro de este contexto, se emplean los Data Warehouse, los cuales son almacenes de grandes volúmenes de datos. En estos almacenes se concentra información acerca de la organización y debido a la gran cantidad de datos que puede llegar a

almacenar suelen dividirse en almacenes de datos más pequeños, a los cuales se les denomina data marts. Dentro de los data marts se almacena información acerca de un área o proceso específico de una organización, es decir, que representan sólo una parte del contexto completo que rodea a una organización. Los data marts, a su vez, suelen ser bases de datos diseñadas mediante el modelo multidimensional, a través del cual se representan las particularidades de los procesos que ocurren en una organización. El modelo multidimensional consta de dos principales elementos: hechos y dimensiones.

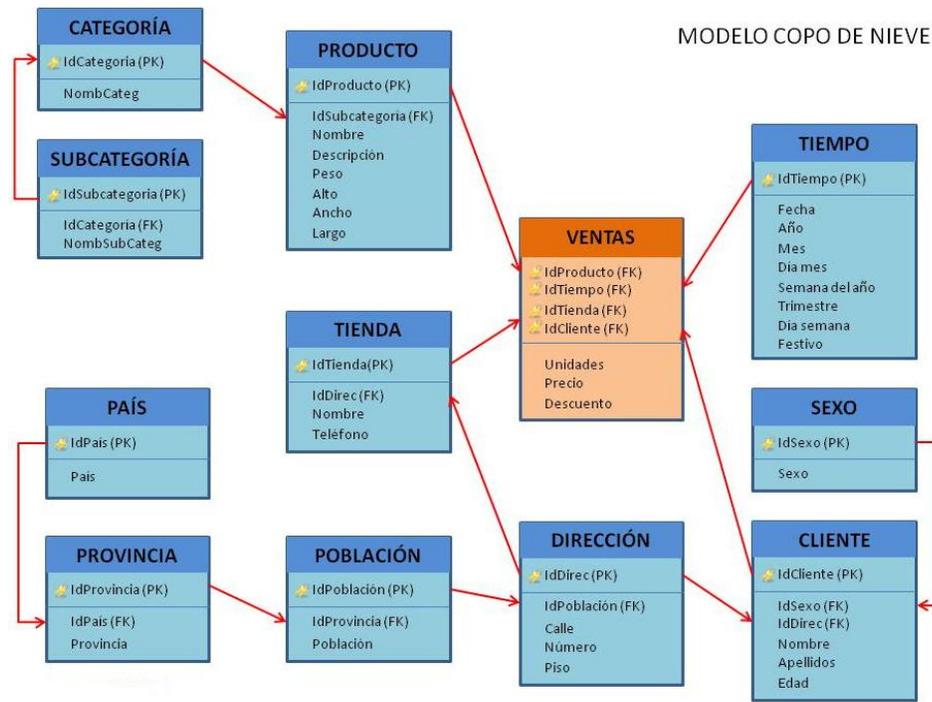
En el modelo multidimensional, los hechos representan el objeto de análisis, mientras que las dimensiones son los factores a través de los cuales se analizan los hechos o bien, son “información complementaria necesaria para la presnetación de los datos a los usuarios” (Rodríguez Sanz, 2010, p.78), como: nombres, zonas, descripciones, etc. En el modelo multidimensional, las tablas de hechos no se relacionan entre sí, por lo que sólo podrán relacionarse con las tablas de dimensiones. Las tablas de dimensiones contienen, por lo general, una llave primaria simple y un conjunto de atributos que describen la dimensión; mientras que las tablas de hechos están compuestas de múltiples llaves foráneas, cada una de las cuales hace referencia a una de las dimensiones. Debido a esto, los esquemas más comunes del modelo multidimensional son los esquemas de estrella y de copo de nieve.

En el esquema de estrella, la tabla de hechos se relaciona con sus dimensiones; pero las tablas de dimensiones no se relacionan con ninguna otra tabla que no sea la de hechos (ver figura 3.2).



**Figura 3.2.** Ejemplo de un data mart con esquema de estrella. Tomado de: García (2013)

En el esquema de copo de nieve, cada punta de la estrella puede, a su vez, formar otras estrellas, es decir que las tablas de dimensiones pueden relacionarse con otras tablas de dimensión; pero de jerarquía menor (ver figura 3.3). El modelo de datos del SIGEM pertenece al esquema de copo de nieve.



**Figura 3.3.** Ejemplo de un data mart con esquema de copo de nieve. Tomado de: García (2013)

Para este sistema, la tabla de hechos hace referencia a los metadatos geográficos. Además, cada una de las dimensiones de este modelo representa una de las secciones de la NTM, por lo que el modelo de datos del SIGEM será como el mostrado en la figura 3.4. En este primer modelo, la tabla de hechos toma el nombre **catalogue** y será utilizada para crear un buscador de metadatos. Debido a esto, además de contener llaves foráneas hacia las tablas de dimensiones, también incluye el nombre del conjunto de datos (atributo **name**), así como la descripción del conjunto de datos (atributo **description**). Por otra parte, las tablas de dimensión son:

- **dataset:** contiene los elementos correspondientes a la sección 1 (Identificación del conjunto de datos).
- **date\_related:** contiene los elementos correspondientes a la sección 2 (Fechas relacionadas con el conjunto de datos).
- **address\_book:** contiene los elementos correspondientes a la sección 3 (Parte responsable del conjunto de datos espaciales o producto).
- **geographical\_location:** contiene los elementos correspondientes a la sección 4 (Localización geográfica del conjunto de datos espaciales o producto).
- **reference\_system:** contiene elementos correspondientes a la sección 5 (Sistema de referencia).
- **quality\_information:** contiene los elementos correspondientes a la sección 6 (Calidad de la información).

- **entity\_attribute**: contiene los elementos correspondientes a la sección 7 (Entidades y atributos).
- **distribution**: contiene los elementos correspondientes a la sección 8 (Distribución).
- **contact\_info\_metadata**: contiene los elementos correspondientes a la sección 9 (Información del contacto para los metadatos).

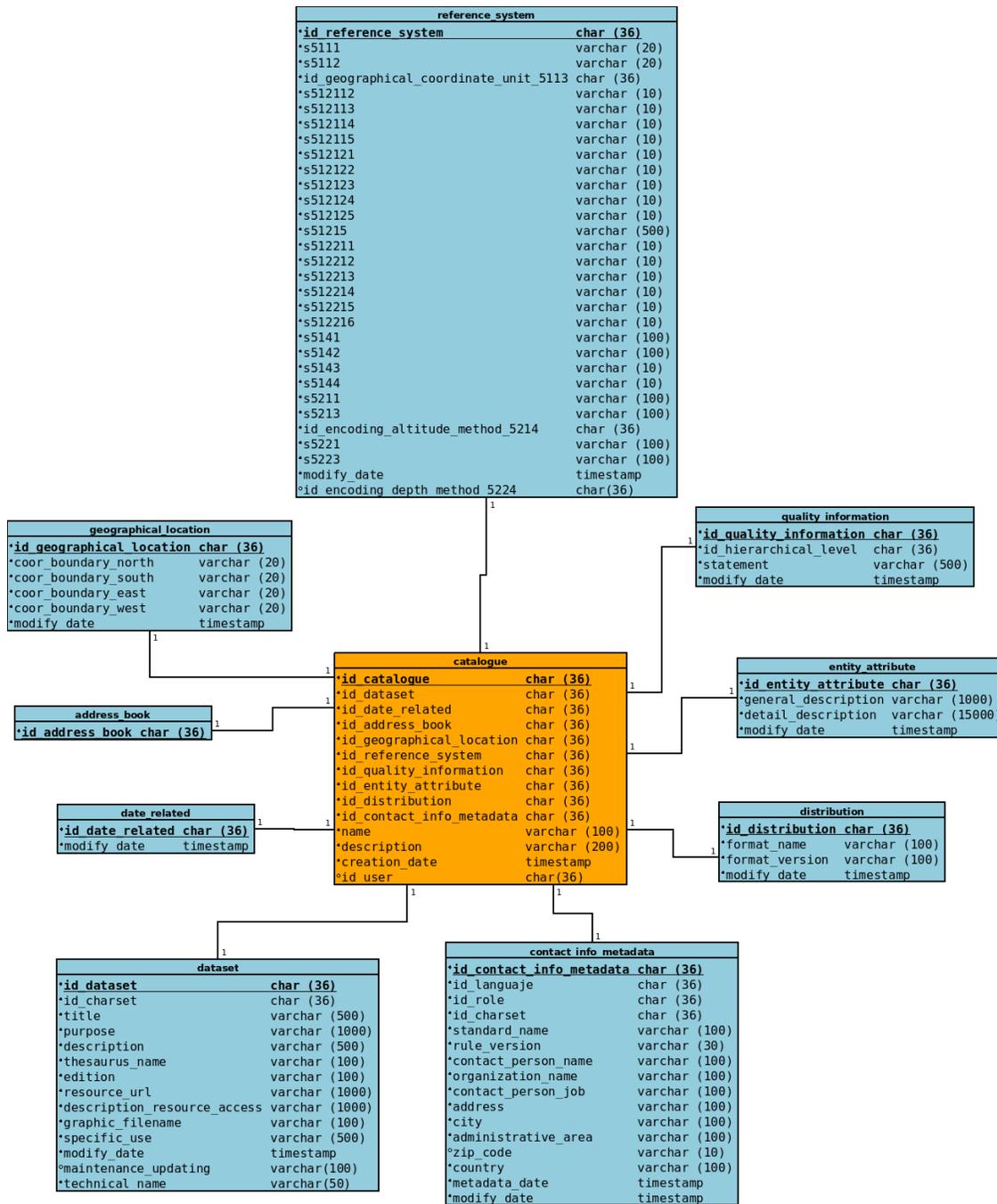


Figura 3.4. Modelo multidimensional del SIGEM.

En el modelo mostrado en la figura 3.4 se observa que la configuración del modelo de datos del SIGEM corresponde con un esquema en estrella; pero ésto en realidad no es así, ya que cada tabla de dimensión se relaciona, a su vez, con otras tablas que permiten almacenar todos los elementos correspondientes a una sección de la NTM. Es por esto que la estructura completa del modelo de datos del SIGEM corresponde con un esquema de copo de nieve. En los diagramas de las figuras 3.5 a 3.12 se muestra el modelo relacional de cada una de las dimensiones. En estos diagramas no se incluyó la dimensión ‘entity\_attribute’, porque se compone de una sola tabla, la cual ya se muestra en la figura 3.4.

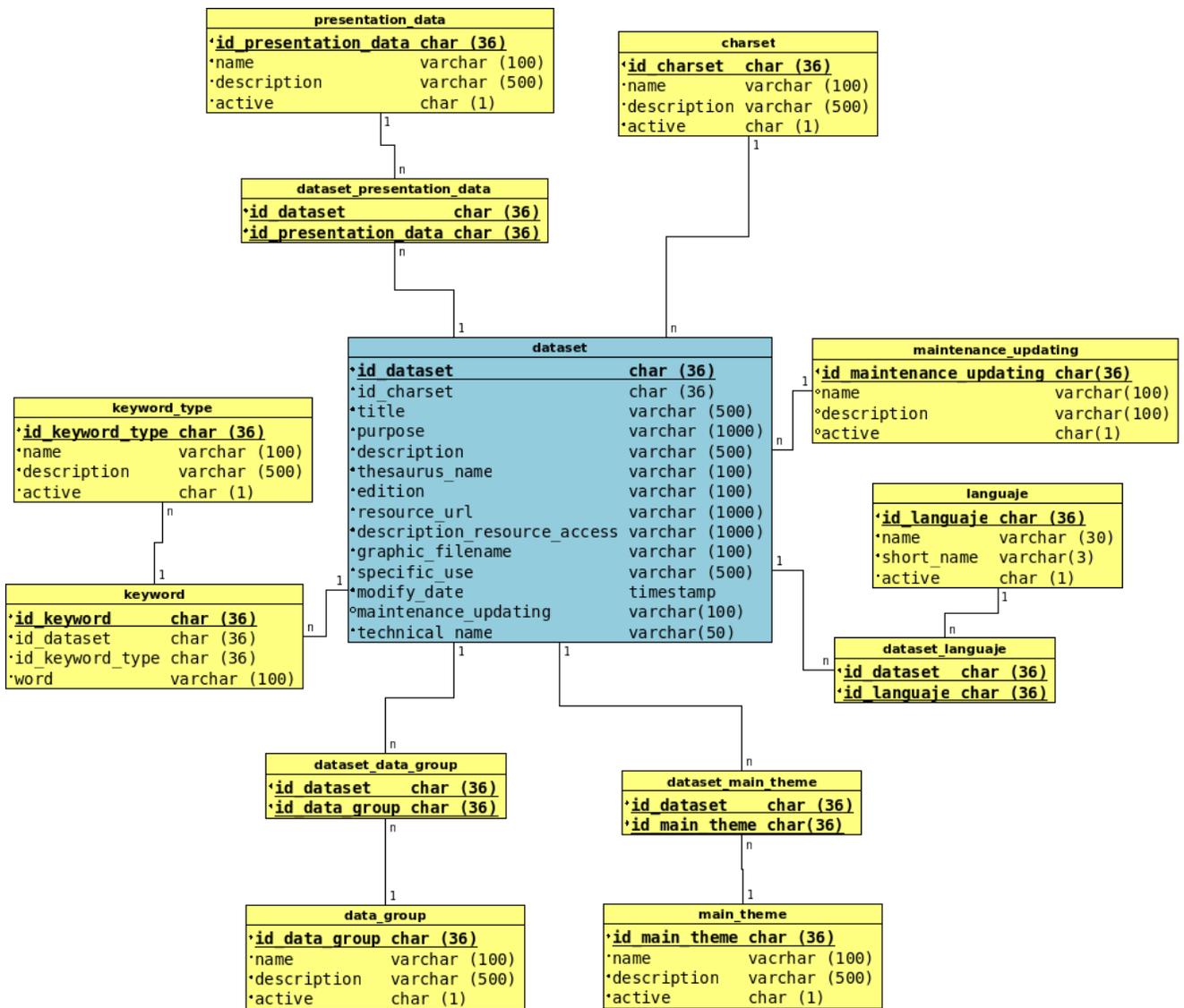


Figura 3.5. Modelo relacional de la dimensión ‘dataset’.

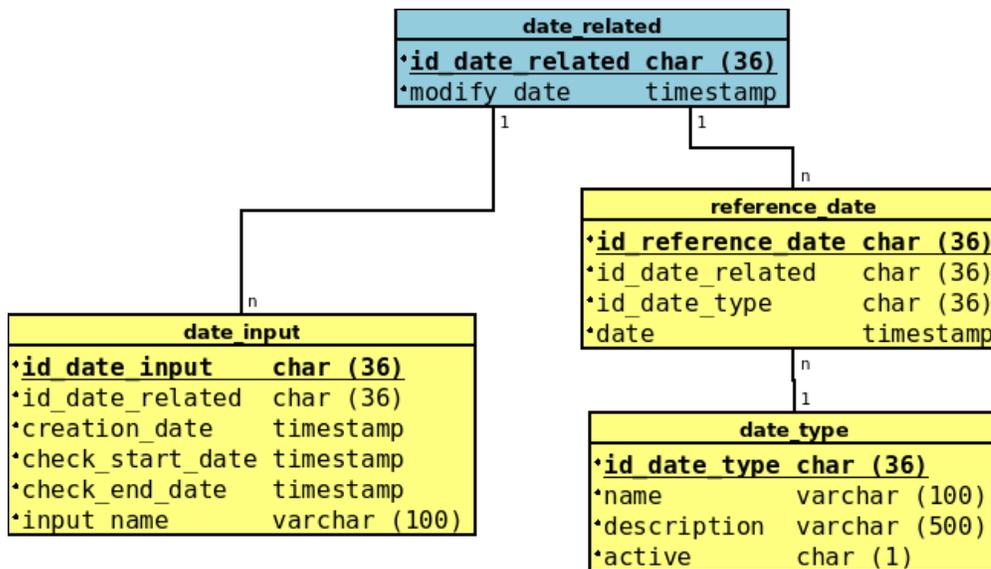


Figura 3.6. Modelo relacional de la dimensión ‘date\_related’.

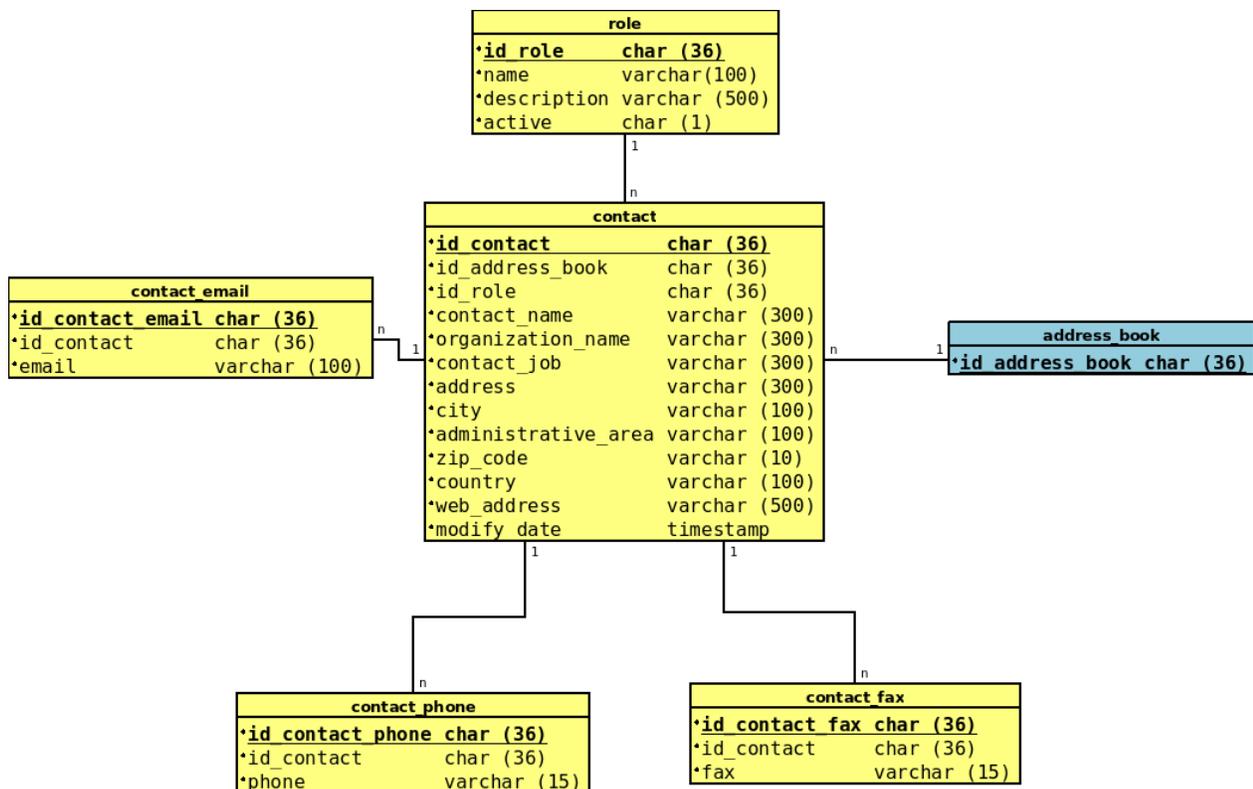


Figura 3.7. Modelo relacional de la dimensión ‘address\_book’.

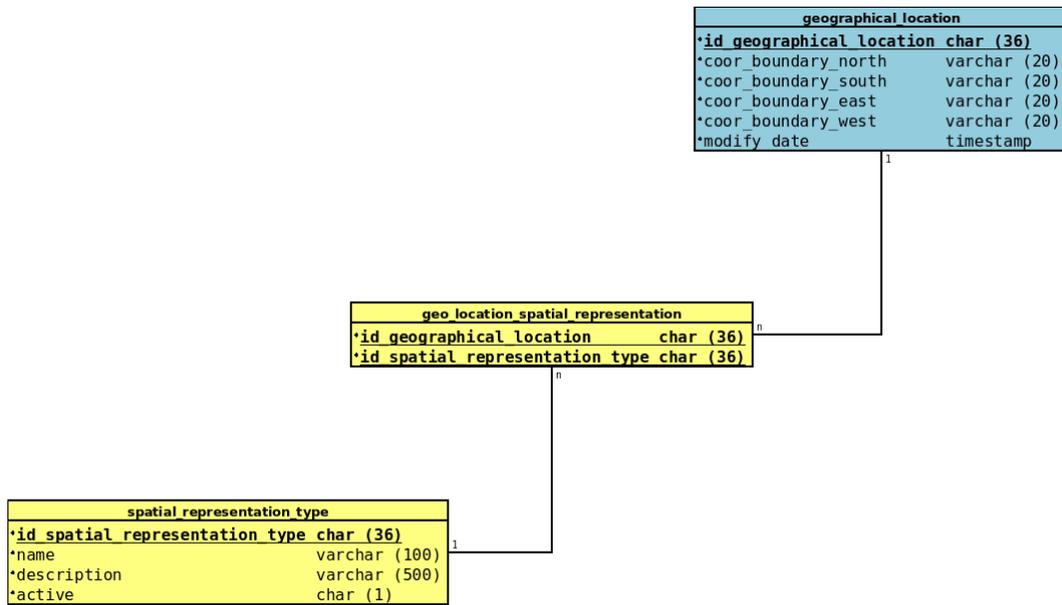


Figura 3.8. Modelo relacional de la dimensión ‘geographical\_location’.

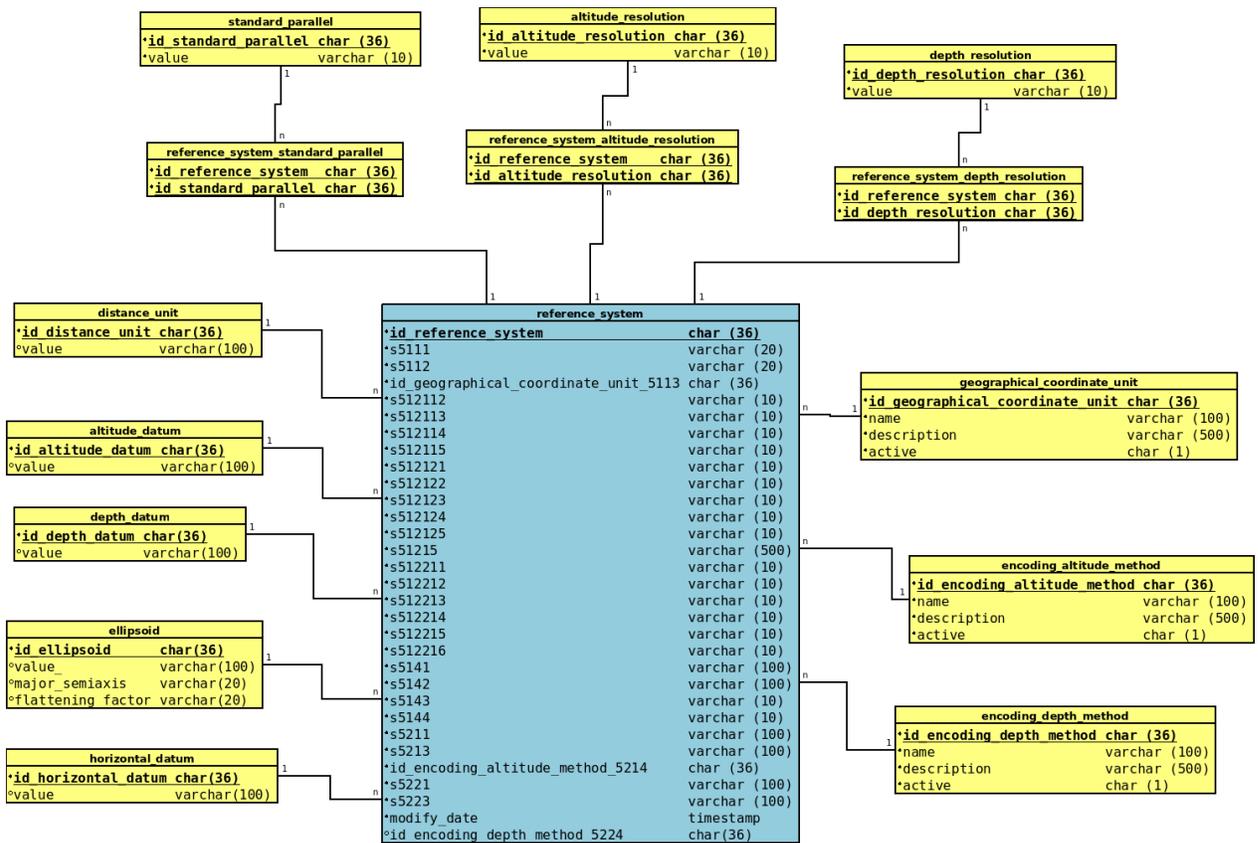


Figura 3.9. Modelo relacional de la dimensión ‘reference\_system’.

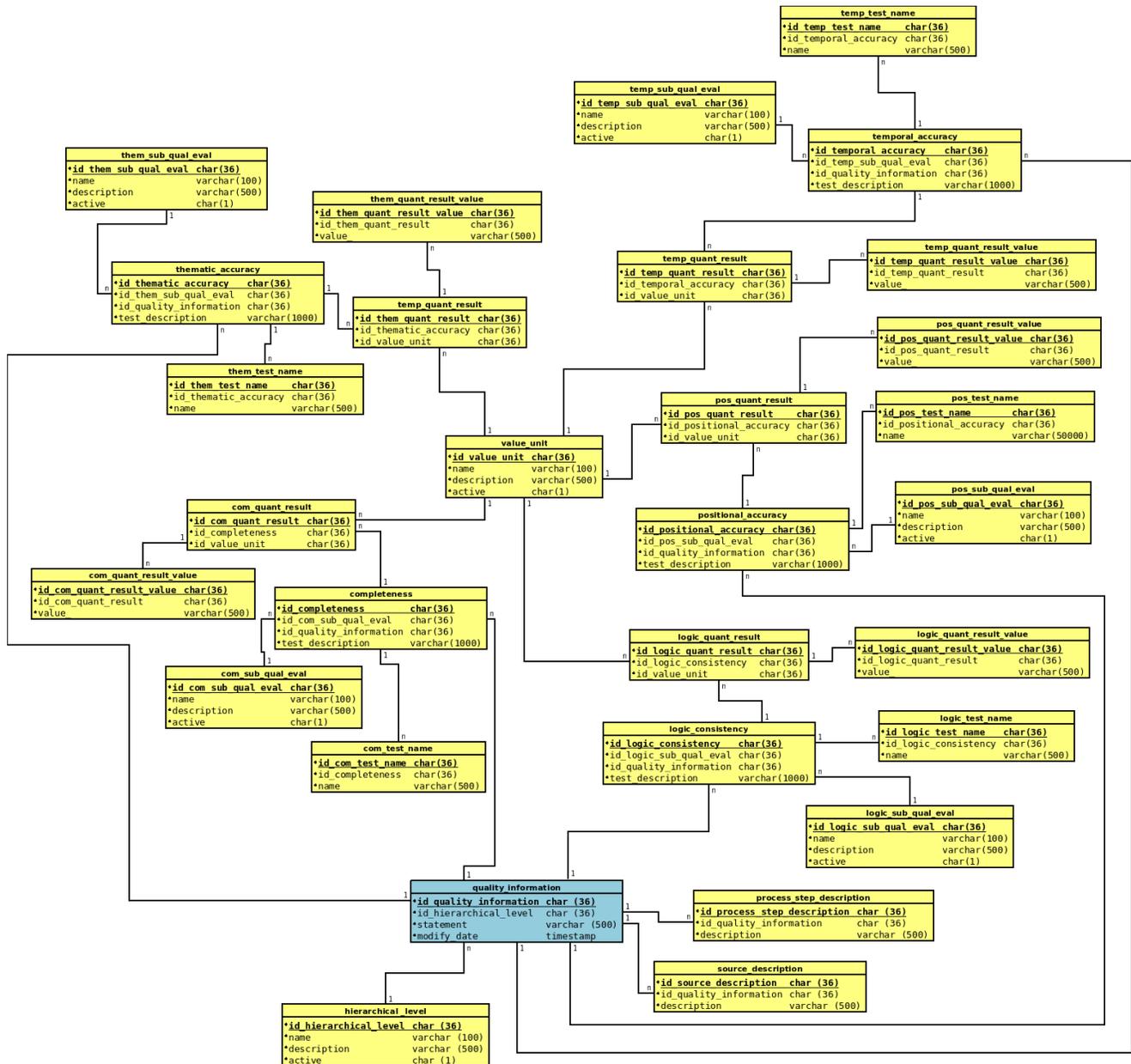


Figura 3.10. Modelo relacional de la dimensión 'quality\_information'.

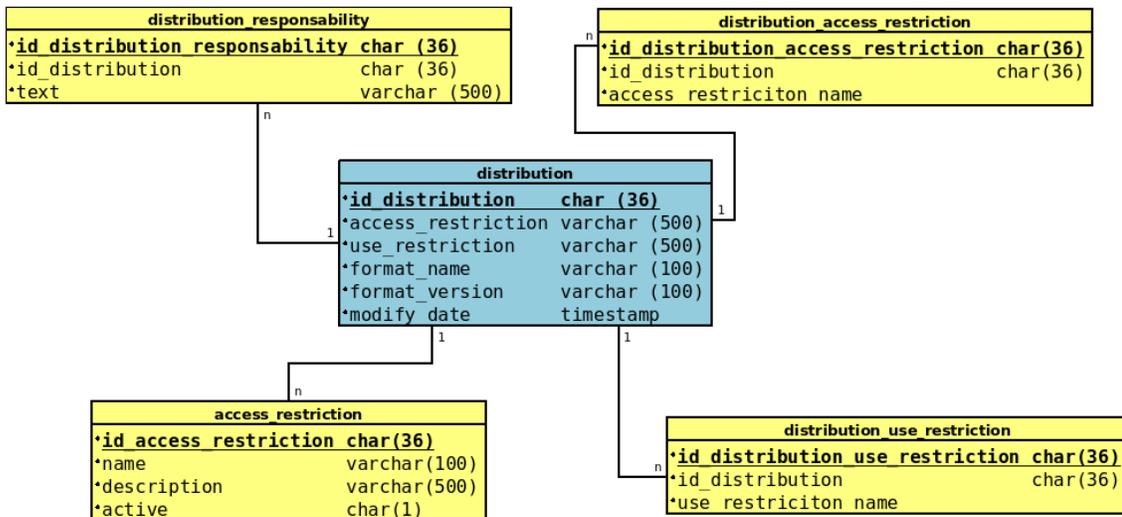


Figura 3.11. Modelo relacional de la dimensión ‘distribution’.

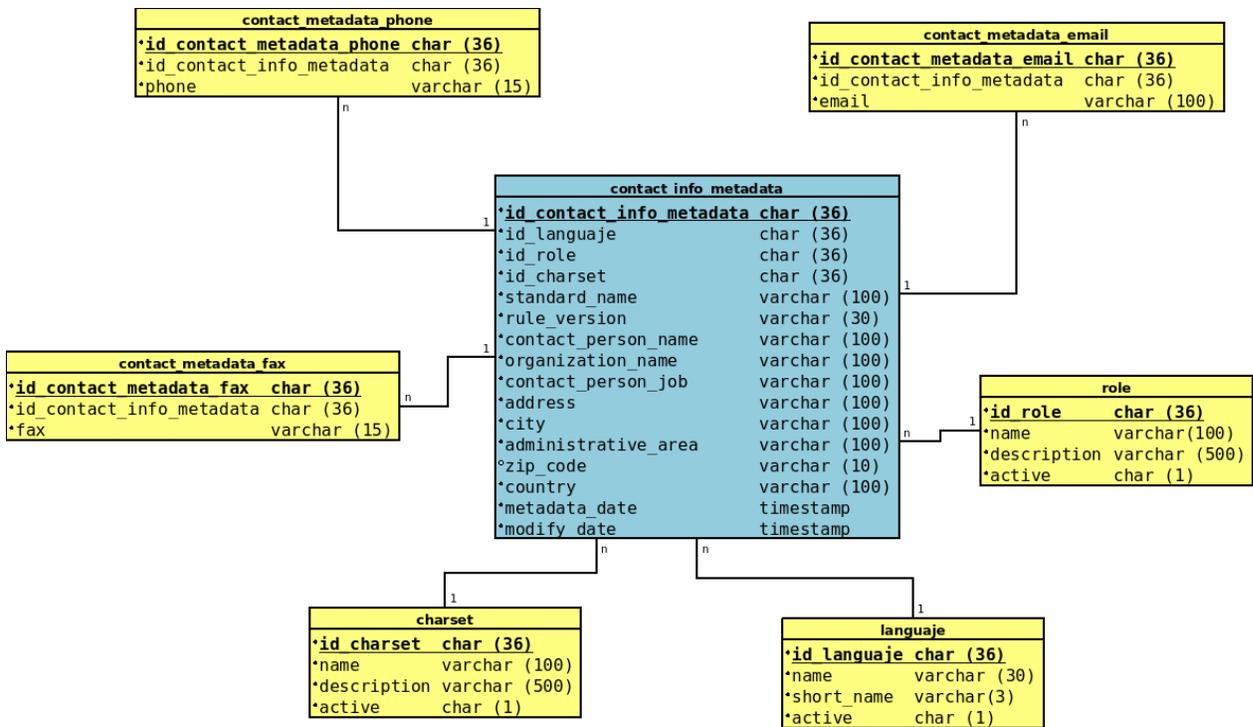


Figura 3.12. Modelo relacional de la dimensión ‘contact\_info\_metadata’.

Finalmente, en cada uno de los modelos relacionales mostrados, los nombres de las tablas que rodean a las tablas de dimensión (las tablas amarillas) dan una idea de los elementos de la NTM que representan, aunque estos nombres están en idioma inglés. Las tablas que contienen los atributos ‘name’, ‘description’ y ‘active’ representan catálogos de términos ya definidos en la norma. Por ejemplo, la tabla ‘charset’ (ver figura 3.5 o 3.12) hace referencia al elemento 1.13 de la NTM, que es, **Conjunto de caracteres**. En esta sección de la norma se especifica que el dominio para

este elemento es MD\_Código\_del\_conjunto\_de\_caracteres y en éste se especifican los valores que pueden tomar estos elementos. Lo mismo ocurre con otras tablas como ‘**presentation\_data**’ (figura 3.5), que contiene el conjunto de valores posibles para el elemento 1.10 Forma de representación de los datos, o para la tabla ‘**access\_restriction**’ (figura 3.11) que tiene los posibles valores para los elementos 8.1 y 8.2 de la NTM: Restricciones de acceso y Restricciones de uso, respectivamente.

### 3.1.2. Tecnologías de bases de datos empleadas

Es importante recordar que los modelos mostrados en las figuras 3.5 a 3.12 son lógicos, no físicos, por lo que no todas las llaves foráneas del modelo lógico, lo son en la implementación de la base de datos (modelo físico). Ésto se debe principalmente a que el llenado de algunos elementos de la norma es opcional y tener llaves foráneas hacia las tablas donde se guardan dichos elementos redundaría en obligar a los usuarios finales a llenar dichos elementos, aún cuándo no se disponga de información para ello. Para una referencia rápida de los elementos con que debe contar cada sección: cuáles son obligatorios, cuáles opciones y cuáles están condicionados, se pueden consultar las figuras C.1 a C.9.

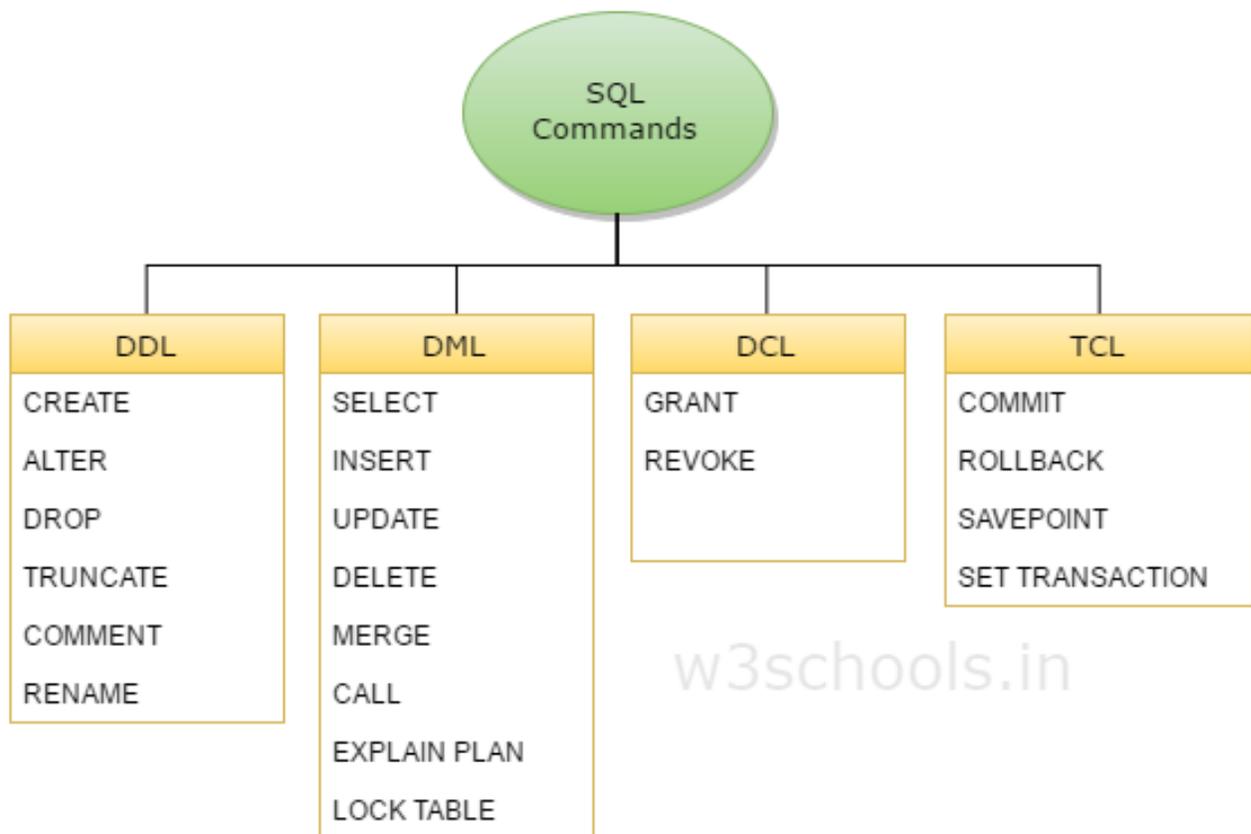
Con lo anterior, se puede ver que a pesar de contar con un modelo lógico de datos que este sistema emplea para almacenar los datos, éste se tuvo que modificar debido a los requerimientos propios del sistema y porque a su vez, éste influyó en la forma en que el sistema se desarrolló.

El modelo físico de datos para este sistema consta de las mismas tablas que las mostradas en el modelo lógico de la sección anterior, aunque como ya se mencionó, no fue posible conservar todas las relaciones entre estas tablas. Con base en este modelo fue posible crear la base de datos que alimentó al sistema. La base de datos se implementó mediante un SMDB. Éstos son programas o software que controlan “la organización, almacenamiento, recuperación, seguridad e integridad de los datos en una base de datos”(Master magazine, 2016). Dentro de un SMDB las bases de datos se crean usando el lenguaje de definición de datos (DDL, por sus siglas en inglés), el cual sirve para definir estructuras de datos, entre ellas las tablas, esquemas o índices. Algunas de las sentencias de este lenguaje son: CREATE, DROP, ALTER, TRUNCATE, COMMENT, RENAME. El lenguaje de manipulación de datos (DML, por sus siglas en inglés) puede ser usado para consultar y actualizar la base de datos. Éste incluye algunos comandos como INSERT, SELECT, UPDATE o DELETE. Por su parte, el lenguaje de consulta estructurada (SQL, por sus siglas en inglés) contiene tanto las sentencias de DDL como de DML. Pero éstos no son los únicos dos lenguajes soportados por SQL y por tanto por un SMDB, también existe el DCL y TCL. El primero tiene sentencias para controlar el acceso y permisos de la base de datos y sus comandos son GRANT y REVOKE. El segundo se encarga de las transacciones de la base de datos y sus comandos son, entre otros: COMMIT y ROLLBACK. En la figura 3.13 se puede consultar la lista de comandos SQL.

Para el SIGEM, el SMDB que se utilizó fue **PostgreSQL**. La razón principal de utilizar este SMDB fue para que el SIGEM se ajustara a las líneas de desarrollo de software de la U-GITS. En la U-GITS, entre otras cosas, se desarrollan portales para la visualización y consulta de datos e información geográficos. Éstos son almacenados en bases de datos mediante el manejador PostgreSQL<sup>10</sup>, a través del cual pueden manipularse grandes cantidades de datos y además soporta una alta concurrencia de usuarios. Otra de las características por las que se utiliza PostgreSQL en los sistemas de la U-GITS, es que ofrece una extensión que da soporte al manejo de objetos geográficos. Esta extensión es PostGIS<sup>11</sup> y tanto PostgreSQL como PostGIS son opciones de código fuente abierto, es decir, que se distribuyen bajo licencias de software libre y además está disponible sin costo. Cabe aclarar que para todos los sistemas creados por la U-GITS se utilizan tecnologías de software libre.

<sup>10</sup>[http://www.postgresql.org.es/sobre\\_postgresql](http://www.postgresql.org.es/sobre_postgresql)

<sup>11</sup><http://www.postgis.net/>



**Figura 3.13.** Subgrupos y comandos del lenguaje SQL. Tomado de: W3schools (2016).

## 3.2. Arquitectura tecnológica de la aplicación

El diseño de un sistema informático es un “proceso de etapas múltiples en el que, a partir de los requerimientos de información, se sintetizan las representaciones de los datos y la estructura del programa, las características de la interfaz y los detalles del procedimiento” (Pressman, 2010, p.206) y como se mencionó al inicio del este capítulo, el diseño de un sistema debe representar de forma clara, entre otros elementos, la arquitectura del sistema.

La arquitectura de un sistema de software debe proporcionar una visión completa del sistema, pues muestra los componentes del sistema, las propiedades de estos componentes, así como la forma en que éstos se integran para formar un todo. En otras palabras: la arquitectura es la representación de la estructura del sistema.

El diseño de la arquitectura g meuiará el proceso de desarrollo del sistema software, ya que en él deben tenerse identificados los componentes y los conectores del sistema. Los conectores son interacciones entre los elementos del sistema, que pueden tomar la forma de un llamado a un procedimiento, consultas a bases de datos, tuberías o pipes (para procesar flujos de datos) o la reacción a un evento (Garlan y Shaw, 1993, p.6).

Por otro lado, un sistema de software puede verse desde diferentes perspectivas, es decir, que para su representación puede tomarse sólo un tipo específico de componentes del sistema, lo que permite contar con varias representaciones del sistema. Por ejemplo: se puede tener una representación del sistema donde los componentes mostrados pueden ser entidades de procesamiento y los conectores

serían interfaces que proveen la capacidad de usar o pasar datos a ciertos componentes; o bien, que los componentes sean elementos físicos del sistema (o hardware) en los que reside el software, mientras que los conectores, igualmente, serían interfaces que permitan la comunicación entre estos componentes (Pressman, 2010, p.212).

Los dos ejemplos anteriores no son las únicas formas de representar la arquitectura de un sistema. En realidad existe lo que se conoce como estilos de arquitectura, los cuales definen familias o grupos. Un estilo arquitectónico determina:

el vocabulario de componentes y conectores que pueden ser usados en instancias de ese estilo, junto con un grupo de *restricciones* acerca de cómo pueden ser combinados. Éstas pueden incluir restricciones topológicas o descripciones arquitectónicas (p. ej., sin ciclos). Otras restricciones –es decir, aquellas que tienen que ver con la semántica de ejecución– pueden también ser parte de la definición del estilo. (Traducido de Garlan y Shaw (1993, p.6))

El estilo arquitectónico empleado para el desarrollo del SIGEM es el de arquitectura en capas, el cual se aborda en la siguiente sección.

### 3.2.1. Arquitectura de software por capas

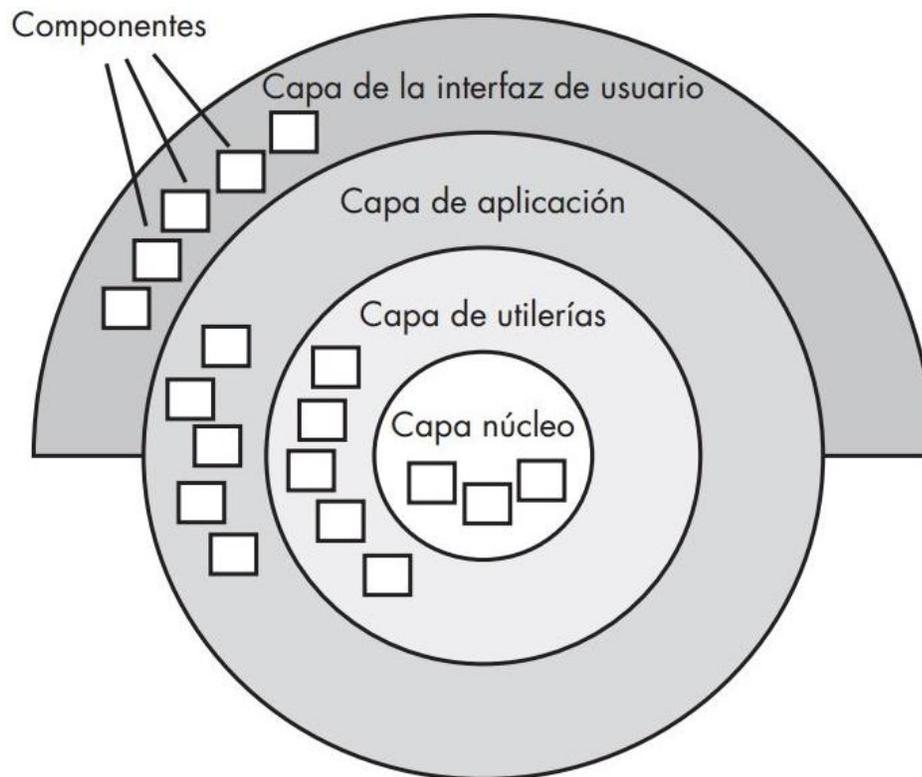
Los sistemas con arquitectura en capas representan la solución a un problema. Tal solución puede ser dividida en una secuencia de “pasos incrementales” (Garlan y Shaw, 1993, p.11). Cada uno de estos pasos, representa una capa y cada una de éstas ejecutará ciertas operaciones. La capa externa tiene componentes que permiten la comunicación con los usuarios, mientras que las capas internas proporcionan funciones para implementar la lógica del negocio del sistema y la capa más interna realiza la interfaz con el sistema operativo (ver figura 3.14).

Dentro de las principales características de este tipo de arquitecturas es que su estructura es jerárquica: cada capa proveerá un servicio a la capa superior y a su vez, será cliente de la capa inferior. Cada capa debe implementar técnicas y protocolos que permitan la comunicación con otras capas; pero con la restricción de que una capa sólo podrá comunicarse, como máximo, con dos capas adyacentes (la capa superior y/o la inferior). Otras de las características de este tipo de arquitectura son:

- **Permite el mejoramiento de la aplicación**, porque los cambios hechos en la funcionalidad de una de las capas deberán afectar sólo a dicha capa y en caso de requerirse un cambio más radical deberá afectar, a lo más, a las dos capas adyacentes a ésta, es decir, la capa inferior y la capa superior.
- **Permite el reuso** de las capas, es decir, que se pueden tener diferentes implementaciones de una misma capa y éstas pueden ser usadas de manera intercambiable, siempre que las capas adyacentes soporten los mismos protocolos de comunicación.

Existen diferentes ejemplos de este estilo arquitectónico (en capas). Uno de ellos es la arquitectura de 3 capas (o 3-tier, en inglés), la cual es muy común para aplicaciones web y que además se empleó para el desarrollo de este sistema. Las tres capas de que se compone esta arquitectura a nivel de software son:

1. **Capa de datos.** Dentro de esta capa se encuentran los métodos que interactúan directamente con la base de datos, es decir que a través de ellos se ejecutan las operaciones de consulta, actualización, creación y borrado de los elementos de las tablas de la base de datos del sistema.



**Figura 3.14.** Diagrama de una arquitectura de capas. Tomado de: Pressman (2010, p.215).

2. **Capa lógica.** Esta capa recibe datos que le llegan desde la capa de presentación, los procesa y lo envía a la capa de datos para que sean almacenados en la base de datos. De igual forma, hace consultas a la base de datos, procesa los datos recibidos y los envía a la capa de presentación para mostrarlos a los usuarios. Esta capa funge como intermediaria entre la capa de datos y la capa de presentación y dentro de ella se encuentran métodos que permiten implementar la lógica del negocio, por lo que dentro de ella también se hacen validaciones de los datos recibidos desde la capa de presentación.
3. **Capa de presentación.** Esta capa representa la interfaz de usuario del sistema. A través de ella se recolectan los datos ingresados por los usuarios y se envían a la capa lógica. Así mismo recibe datos que son enviados desde la capa lógica y se encarga de mostrar los resultados a los usuarios. Dentro de esta capa se encuentran las páginas web de la aplicación, mismas que pueden ser consultadas a través de cualquier navegador web.

En el contexto de desarrollo de aplicaciones Java un paquete es un directorio o un contenedor de archivos. Dichos archivos son, por lo general, clases que se relacionan entre sí para cumplir con una función específica del sistema. Al final, todos estos paquetes se ensamblan en un solo archivo ejecutable que será desplegado en un servidor de aplicaciones, lo que permitirá que la aplicación esté disponible en la web y pueda ser consumida a través de navegadores web. En la figura 3.15, se muestran los paquetes que conforman este archivo ejecutable, en donde cada uno de ellos desempeñará una tarea muy específica dentro del sistema.

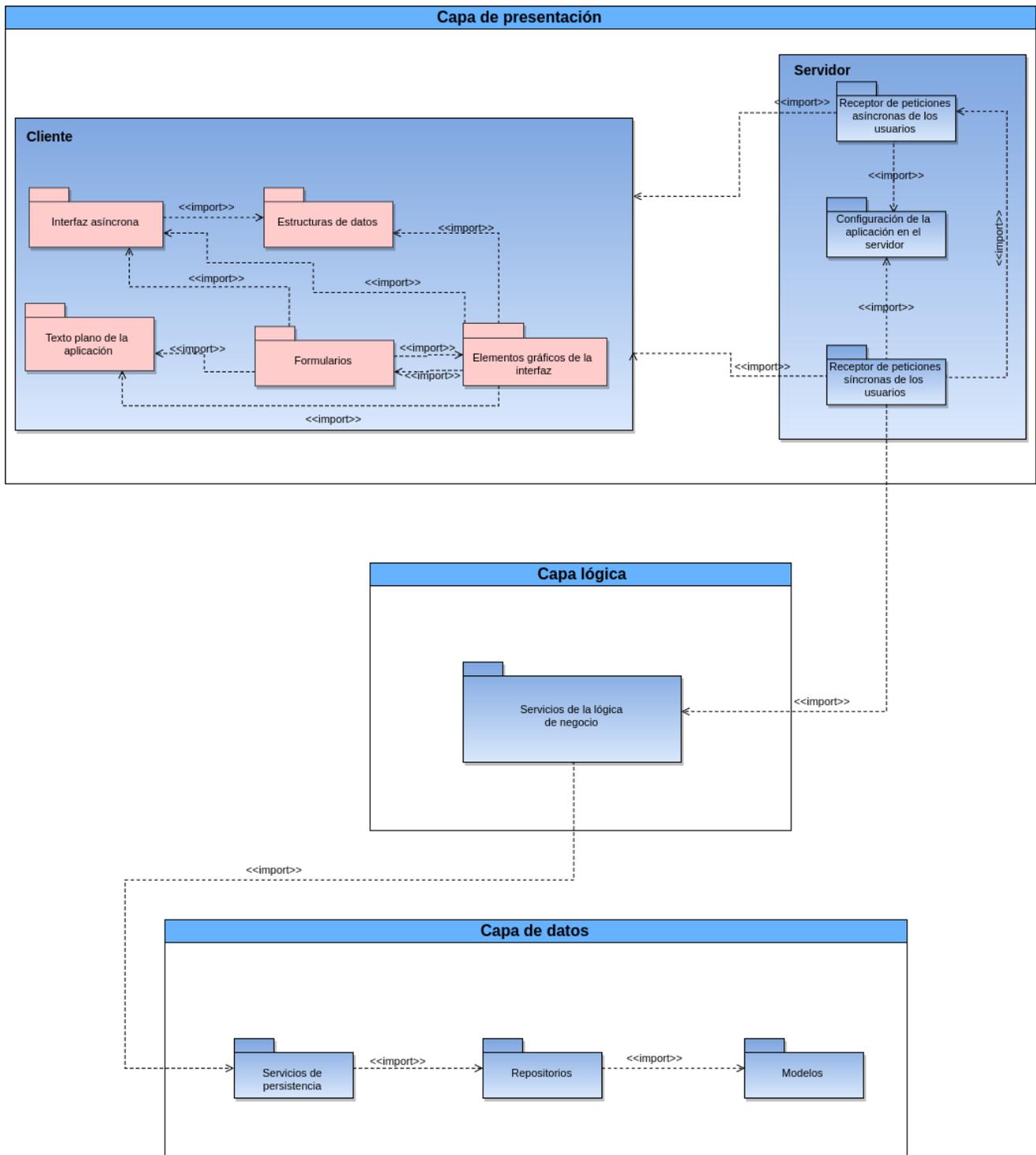


Figura 3.15. Paquetes que conforman el sistema.

Para el desarrollo de la capa de datos se utilizaron los marcos de trabajo *Spring Data* y *Spring framework* (los cuales se abordarán en la sección 3.2.2). Debido a éstos se generaron los tres paquetes que conforman esta capa, que son: **Modelos**, **Repositorios** y **Servicios de persistencia**.

En el paquete **Modelos** se lleva a cabo el *mapeo* del modelo relacional de datos a un modelo de objetos. Esto es necesario debido a que los metadatos son almacenados dentro de tablas en una base de datos que cumplen con un modelo relacional; pero todos los paquetes de la aplicación trabajan con objetos de Java. De forma general, cada una de las clases de este paquete es la representación de una tabla de la base de datos, mientras que los atributos de estas clases corresponden con las columnas de las tablas que representan.

El paquete **Repositorios** contiene interfaces que interaccionan con las clases del paquete **Modelos**, brindándoles un conjunto de funcionalidades para gestionar el contenido de la base de datos mediante operaciones de consulta, edición, inserción y borrado de registros de la base de datos.

El paquete de **Servicios de persistencia** contiene una clase que implementa una interfaz de comunicación entre la capa lógica y la capa de datos del sistema.

Dentro de la capa lógica se encuentra el paquete **Servicios de la lógica de negocio** y contiene una clase que implementa una interfaz de comunicación entre la capa de datos y la capa de presentación. Su función es procesar los datos que recibe y enviarlos a las capas adyacentes (ya sea la capa de datos o la de presentación) con la estructura que éstas requieran. Ésto significa, por ejemplo, que si se realiza una consulta a la base de datos esta capa lógica recibirá los datos resultantes desde la capa de datos y los estructurará de forma tal, que puedan ser utilizados en la capa de presentación y así ser mostrados al usuario; o, por el contrario, si se desea guardar o editar un metadato de la base de datos, este paquete recibirá los datos desde la capa de presentación, los procesará y les dará la estructura necesaria para que puedan ser utilizados por el paquete **Servicios de persistencia** y así ser almacenados en la base de datos.

Para el desarrollo de la capa de presentación se utilizó el marco de trabajo GWT. Ya que GWT se basa en la arquitectura cliente-servidor, una parte del código se ejecuta en el cliente (navegadores web), mientras que otra, se ejecuta en el servidor. Debido a esto se utiliza la estructura de paquetes que se describe a continuación. Los paquetes que son ejecutados en el cliente son: **Formularios**, que contiene los formularios de captura para cada una de las secciones de la norma de metadatos; **Elementos gráficos de la interfaz** que añade más elementos gráficos a los formularios de captura; **Texto plano de la aplicación**, es decir todos los textos y mensajes que serán mostrados en todas las páginas web que conforman la aplicación; **Interfaz asíncrona** en la que se definen las peticiones o llamadas que se pueden realizar al servidor y **Estructuras de datos**, que contiene un conjunto de clases auxiliares que ayudarán a manejar los datos recibidos desde el servidor para presentarlos al usuario o bien, datos recolectados de los formularios para enviarlos al servidor. Los paquetes que son ejecutados en el servidor son: **Configuración de la aplicación en el servidor** en el que se definen las clases encargadas de procesar los datos recibidos desde o enviados al servidor; el **Receptor de peticiones asíncronas de los usuarios** que se encargará de procesar peticiones hechas al servidor resultantes de eventos que ocurran dentro de las páginas web de la aplicación (como guardar un metadato o editarlo); **Receptor de peticiones síncronas de los usuarios** que se encargará de procesar las peticiones de exportar metadatos en formato HTML, PDF o XML (ver figura 3.15).

### 3.2.2. Tecnologías de desarrollo empleadas

Para el desarrollo de este sistema se seleccionaron las tecnologías que permitiesen implementar la arquitectura seleccionada y, debido a que cada capa de la arquitectura cumple con una función específica, las tecnologías empleadas para el desarrollo del sistema varían en cada una de estas

capas. En la figura 3.16 se muestran los marcos de trabajo, bibliotecas o tecnologías empleados en el desarrollo de este sistema.



Figura 3.16. Tecnologías de desarrollo del SIGEM.

Estas tecnologías se dividen en 6 rubros, los cuales son: Sistema de control de versiones, lenguajes para el desarrollo de software, herramientas de gestión y construcción de proyectos, marcos de trabajo, bibliotecas (API's) y contenedor de servlets.

### Sistema de control de versiones

Los Sistemas de Control de Versiones (o VCS por sus siglas en inglés) resultan muy útiles para el desarrollo de software en general, pues son herramientas que sirven para conservar el registro de todos los cambios hechos sobre determinados archivos y permiten a los programadores, en caso de ser necesario, volver a una versión específica del proyecto.

El sistema de control de versiones empleado para el desarrollo de este sistema es **Git**<sup>12</sup>. Entre los motivos por los que se seleccionó **Git** para el desarrollo del SIGEM se encuentran: que es una herramienta gratuita que se distribuye con licencia de software libre y que además tiene muchas opciones de configuración y cuyos comandos son fáciles de aprender, por lo que los colaboradores pueden comenzar a utilizarlo de forma rápida. Además es una tecnología cuya efectividad se encuentra ampliamente probada y actualmente es usada por empresas y proyectos como: Google, facebook, Microsoft, Netflix, PostgreSQL, KDE, Qt, entre otros.

### Lenguajes para el desarrollo de software

#### Java <sup>13</sup>

<sup>12</sup><https://git-scm.com>

<sup>13</sup><https://www.oracle.com/java/index.html>

Es un “un lenguaje de programación y una plataforma informática”(Oracle Corporation, 2012, p.11) que se empleó para el desarrollo completo del sistema. Si bien, existen cuatro tipos de plataforma de Java, cada una de ellas proporciona la máquina virtual de Java (Java Virtual Machine, JVM) y una interfaz de programación de aplicaciones (o API, por sus siglas en inglés). La JVM es un programa para un software y hardware específico que ejecuta aplicaciones escritas en lenguaje Java, mientras que una API es una biblioteca que ofrece un conjunto de componentes de software (como: clases, métodos, funciones, etc.) que pueden ser usados para crear otros componentes o aplicaciones.

La plataforma donde se ejecuta el SIGEM es la **Java Platform, Standard Edition (Java SE)** versión 1.7, que proporciona las funciones núcleo del lenguaje de programación Java como: los tipos de datos y/u objetos básicos de Java, así como características avanzadas que permiten el manejo de redes, acceso a bases de datos, desarrollo de aplicaciones con interfaz gráfica y control de seguridad de aplicaciones.

Algunas de las ventajas de usar Java como lenguaje y plataforma de desarrollo son:

- Es un lenguaje de programación de propósito general, por lo que sirve para desarrollar una gran variedad de aplicaciones y para un gran número de dispositivos que van desde computadoras de escritorios hasta teléfono inteligentes, tabletas electrónicas o consolas de video juegos.
- Existe mucha documentación respecto a su uso.
- Las aplicaciones Java son portables, lo cual significa que son independientes del sistema operativo en que se desarrollen, por lo que un mismo programa puede correr en diferentes sistemas, característica crucial para el desarrollo de aplicaciones web (IBM Corp., 2015).
- Es gratuito.
- Hay un gran número de herramientas que facilitan escribir aplicaciones Java así como corregirlas. Ejemplos de éstas herramientas son Eclipse o NetBeans y además son gratuitas.
- Existe un gran número de bibliotecas y marcos de trabajos creados para Java, que permiten potenciar y robustecer las aplicaciones escritas en este lenguaje y además están ampliamente probados y cuentan con el respaldo de grandes compañías o fundaciones como: Apache, Google, Facebook, entre otras (Mohan, 2016).

#### HTML<sup>14</sup>

Al tratarse de una aplicación web el uso de HTML es inherente al desarrollo de la misma. El lenguaje de marcado de hipertexto (HTML, por sus siglas en inglés) se usa para definir la estructura y contenido de páginas web y es interpretado por navegadores web.

#### CSS<sup>15</sup>

CSS (Cascading Style Sheets) es un lenguaje que sirve para definir la presentación de documentos en lenguaje de marcado (como HTML, XML, XHTML, etc.). Fue utilizado para personalizar los estilos de las páginas web de la aplicación.

### Gestión y construcción de proyectos

Dentro de este rubro, la tecnología empleada en todos los módulos es **Maven**<sup>16</sup>, la cual es una herramienta utilizada principalmente para la construcción de proyectos de Java, aunque de acuerdo

<sup>14</sup><https://www.w3.org/standards/webdesign/htmlcss>

<sup>15</sup><https://www.w3.org/standards/webdesign/htmlcss>

<sup>16</sup><https://maven.apache.org>

con O'Brien et al. (2008, p.1), también es una herramienta para la administración de proyectos. Esto quiere decir, que adicional a la capacidad de construcción de proyectos también puede, por ejemplo, generar reportes, generar sitios web con la información del proyecto o facilitar la comunicación entre los miembros de un grupo de trabajo, todo esto por medio de plugins. Como herramienta de construcción de proyectos, se enfoca en: la compilación, empaquetado, prueba y distribución de proyectos.

Para el SIGEM se utilizó con el fin de automatizar el proceso de resolución de dependencias Java, ejecución de pruebas unitarias y el empaquetado de la aplicación en un solo archivo ejecutable, mismo que es ejecutado en un servidor de aplicaciones o contenedor de servlets.

## Marcos de trabajo

Un marco de trabajo, o mejor conocido como *framework*, es una aplicación “semi-completa”, reusable que puede especializarse para producir aplicaciones a la medida (Fayad y Schmidt, 1997, p.32). Éstos ayudan a encapsular detalles técnicos difíciles de implementar, detrás de una interfaz más estable. Ésto, también permite la reutilización de elementos del sistema pues con ellos se crean componentes genéricos que pueden ser reutilizados en otras aplicaciones. El uso de un *framework* puede llegar a dictar la forma en que es desarrollada una aplicación, ya que suelen apegarse a patrones de diseño.

Los marcos de trabajo empleados para el desarrollo de este sistema son:

- **Hibernate**<sup>17</sup>. Este marco de trabajo es una implementación de la Persistence API de Java y se emplea para llevar a cabo el mapeo del modelo relacional de datos a un modelo de objetos, es decir, que mediante éste se generan clases Java con base en las tablas de la base de datos.
- **Spring**. Este marco de trabajo consta de un amplio conjunto de módulos o proyectos, sin embargo, para el desarrollo del SIGEM se utilizó SpringData<sup>18</sup> y Spring Framework<sup>19</sup>. SpringData está dedicado a facilitar el acceso a las tecnologías de acceso a datos como: bases de datos relacionales y no relacionales, servicios de datos basados en la nube, entre otros. Gracias a este marco de trabajo es posible automatizar la mayor parte de las consultas que deben realizarse a la base de datos (aunque también permite definir consultas personalizadas) por medio de los repositorios –de la capa de datos–.

Spring Framework<sup>20</sup> permite, entre otras cosas, la inyección de dependencias. Gracias a esta característica se puede definir a las interfaces de comunicación entre capas como servicios que serán consumidos por las capas adyacentes. Por ejemplo, la interfaz de los servicios de persistencia (ver figura 3.15) será un servicio consumido por la capa lógica, mientras que la interfaz de los servicios de la lógica de negocio será un servicio consumido por la capa de presentación.

- **Google Web Toolkit (GWT)**<sup>21</sup>. Es un conjunto de herramientas de desarrollo, utilidades y widgets utilizados para generar interfaces web complejas con contenido dinámico.

<sup>17</sup><http://hibernate.org/orm/>

<sup>18</sup><http://projects.spring.io/spring-data/>

<sup>19</sup><http://docs.spring.io/spring-framework/docs/current/spring-framework-reference/html/overview.html#overview-modules>

<sup>20</sup><http://docs.spring.io/spring-framework/docs/current/spring-framework-reference/html/overview.html#overview-modules>

<sup>21</sup><http://www.gwtproject.org>

Las ventajas que representó el uso de este marco de trabajo para el desarrollo del SIGEM fueron:

- Es gratuito y de código abierto.
  - Facilita el desarrollo de aplicaciones web complejas.
  - No es necesario contar con conocimientos avanzados de javascript y puede incluso omitirse el uso de este lenguaje durante el desarrollo de la aplicación.
  - El principal lenguaje de desarrollo es Java.
  - Se encarga de que la aplicación sea compatible con varios navegadores web.
- **GWT-Bootstrap**<sup>22</sup>. *Bootstrap*<sup>23</sup> es un framework creado en 2010 por Twitter para el desarrollo de interfaces web con CSS y Javascript. Actualmente es mantenido por GitHub. Este framework incluye plantillas de diseño basadas en HTML y CSS para tipografía, formularios, botones, tablas, listas, menús desplegables, entre otros elementos. **GWT-Bootstrap** es un biblioteca de GWT que provee los elementos gráficos definidos en el framework *Bootstrap*.

### Bibliotecas de clases (API's)

Una Interfaz de programación de aplicaciones (API, por sus siglas en inglés) es código que permite a dos software comunicarse entre sí. Por medio de una API se proveen clases y/o funciones que pueden servir para cubrir un requerimiento o solucionar un problema específico, por ejemplo: crear interfaces gráficas, publicar datos geográficos en la web, etc. El uso de éstas permite aumentar o potenciar las funcionalidades del sistema, además de agilizar el desarrollo pues permiten reutilizar código existente y ya probado en la solución de determinados problemas. Las APIs empleadas para desarrollar el SIGEM fueron **Xstream**<sup>24</sup> y **Itext**<sup>25</sup>.

Debido a que este sistema no recibe únicamente los datos que los usuarios capturen a través de los formularios, sino que también acepta como insumo metadatos en formato XML, se requirió implementar en el sistema una funcionalidad que permitiera la lectura de estos archivos y con base en su contenido llenar los formularios de captura. Para cumplir con lo anterior se utiliza la biblioteca (o API) gratuita para Java **XStream**. Además también fue empleada para obtener metadatos en formato XML con base en el contenido que ya se encuentra en la base de datos y para descargar el avance de la captura de un metadato en este formato.

**Itext** es una biblioteca para Java que sirve para manipular documentos PDF, ésto incluye desde crearlos hasta extraer elementos de ellos, copiar páginas, combinar documentos, encriptarlos, entre otros. Para el SIGEM se emplea una versión gratuita de esta biblioteca para exportar los metadatos ya existentes a este formato.

### Contenedor de servlets

Un servlet es un programa Java que se ejecuta dentro de un servidor web y se encarga de procesar peticiones de clientes web por medio del protocolo HTTP (Oracle Corporation, 2011). Los servlet se encuentran dentro de un **contenedor de servlets** que a su vez se encuentra dentro de un servidor web. Para que un servlet procese una petición HTTP, ésta primero debe ser recibida por el servidor

---

<sup>22</sup><https://gwtbootstrap.github.io/#>

<sup>23</sup><http://getbootstrap.com>

<sup>24</sup><https://x-stream.github.io>

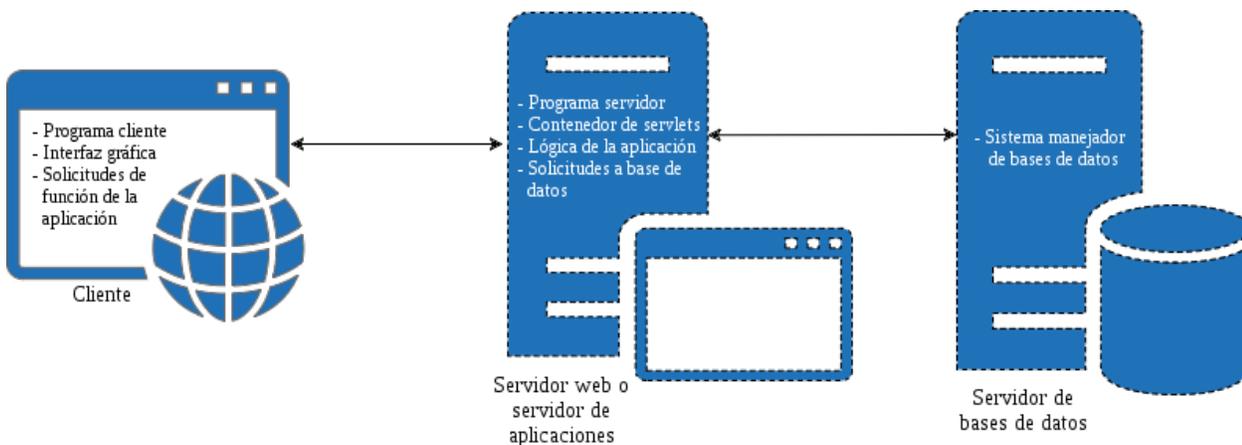
<sup>25</sup><http://itextpdf.com>

web que, en segundo lugar, la dirige al contenedor de servlets, el cual se encarga de seleccionar el servlet que procesará la petición y una vez que se ha procesado la petición la respuesta se envía de vuelta al servidor web, que finalmente envía la respuesta al cliente.

**Apache Tomcat**<sup>26</sup> es un contenedor de servlets desarrollado en Java, de código abierto y de distribución gratuita. Además de lo anterior otras de las ventajas de Tomcat es que por sí mismo puede ser un servidor web o bien, ser usado en conjunto con otros servidores web. **Apache Tomcat** es el contenedor de servlets en el que se aloja y ejecuta el SIGEM. Fue seleccionado debido a factores como: es ligero, fácil de usar y gratuito.

### 3.3. Infraestructura de hardware de la aplicación

Para el desarrollo de este sistema se utilizó la arquitectura de tres capas y se mostró la función de cada una de ellas dentro del sistema. Pues bien, algo similar ocurre con el hardware. En la figura 3.17 se muestran los principales elementos de esta arquitectura; pero orientados al hardware. Éstos son: clientes, un servidor web (físico) y un servidor de bases de datos.



**Figura 3.17.** Infraestructura de hardware para el SIGEM.

Los clientes de la aplicación podrán ser cualquier computadora (sin importar el sistema operativo) que tenga acceso a Internet y un navegador web que tenga habilitado el uso de javascript. GWT ofrece compatibilidad con varios navegadores web. De éstos, los más recomendables son Google Chrome, Firefox y Opera, mientras que uno de los menos recomendables es Internet Explorer, pues ofrece soporte limitado para estándares web abiertos.

El servidor web o servidor de aplicaciones se encargará de recibir y dar respuesta a las peticiones de los usuarios. Dentro de este servidor se instala el contenedor de servlets en el que se ejecuta la aplicación.

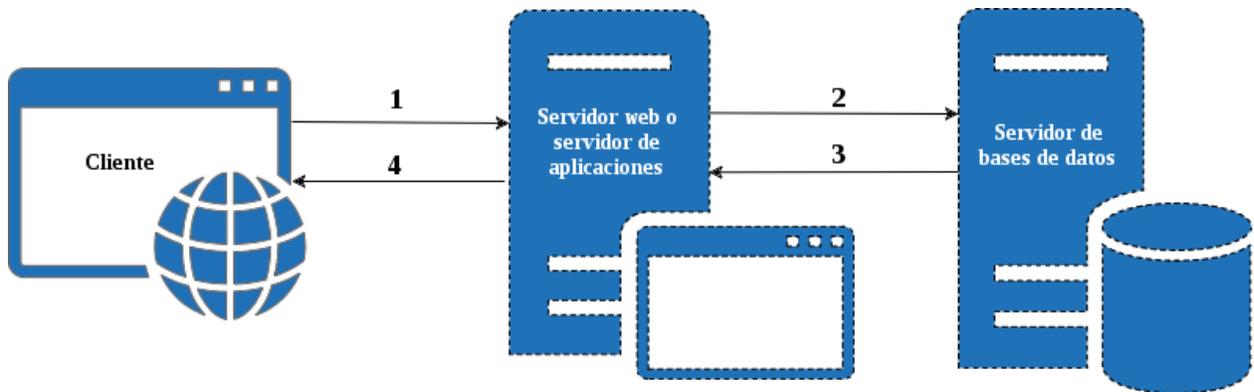
El servidor de bases de datos tiene instalado el SMDB en el que se almacenan todos los metadatos del sistema.

El flujo de trabajo normal de la aplicación consta de los siguientes pasos (ver figura 3.18):

1. El cliente hace una petición al servidor;

<sup>26</sup><https://tomcat.apache.org>

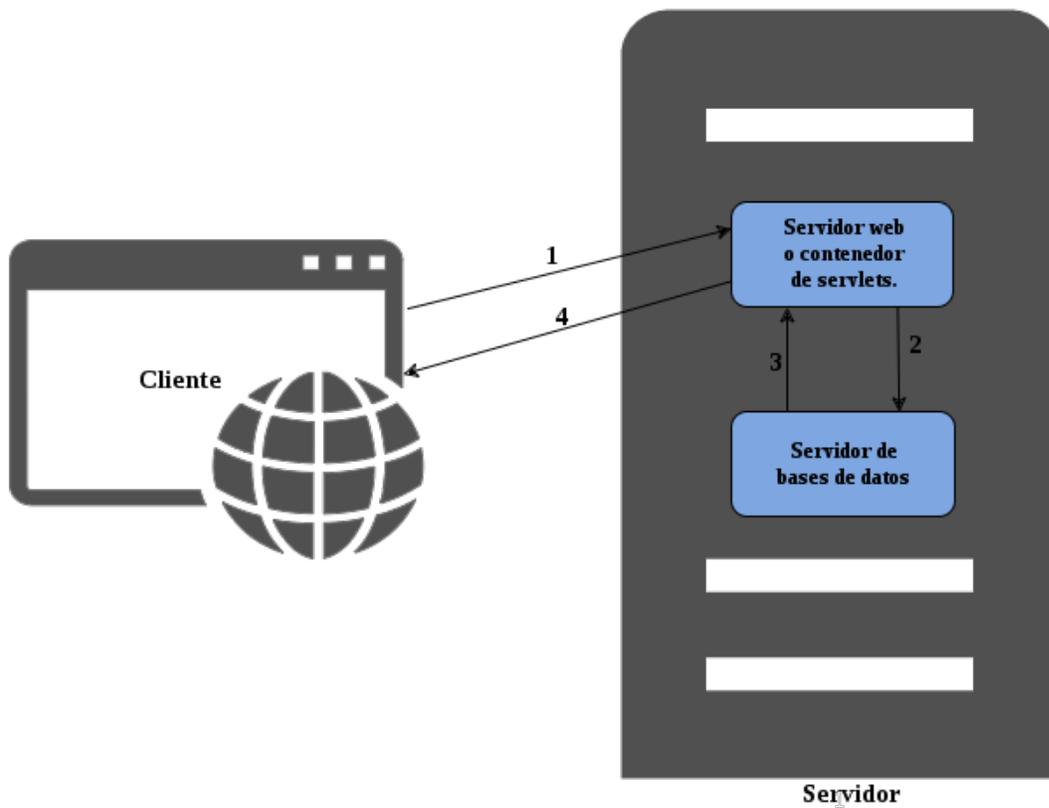
2. el servidor web recibe la petición del cliente, procesa los datos recibidos y realiza una consulta a la base de datos;
3. el servidor de bases de datos envía el resultado de la consulta del vuelta al servidor web, donde es procesado y finalmente
4. enviado al cliente.



**Figura 3.18.** Flujo de trabajo de la aplicación.

Este tipo de infraestructura no es la única, pues la división entre el servidor de aplicaciones y el de bases de datos puede ser meramente lógica, por lo que ambos pueden estar instalados en el mismo equipo. El flujo de trabajo no cambia con esta implementación. Esta infraestructura es similar a la mostrada en la figura 3.19.

El siguiente capítulo estará dedicado a mostrar el funcionamiento de la interfaz gráfica del sistema.



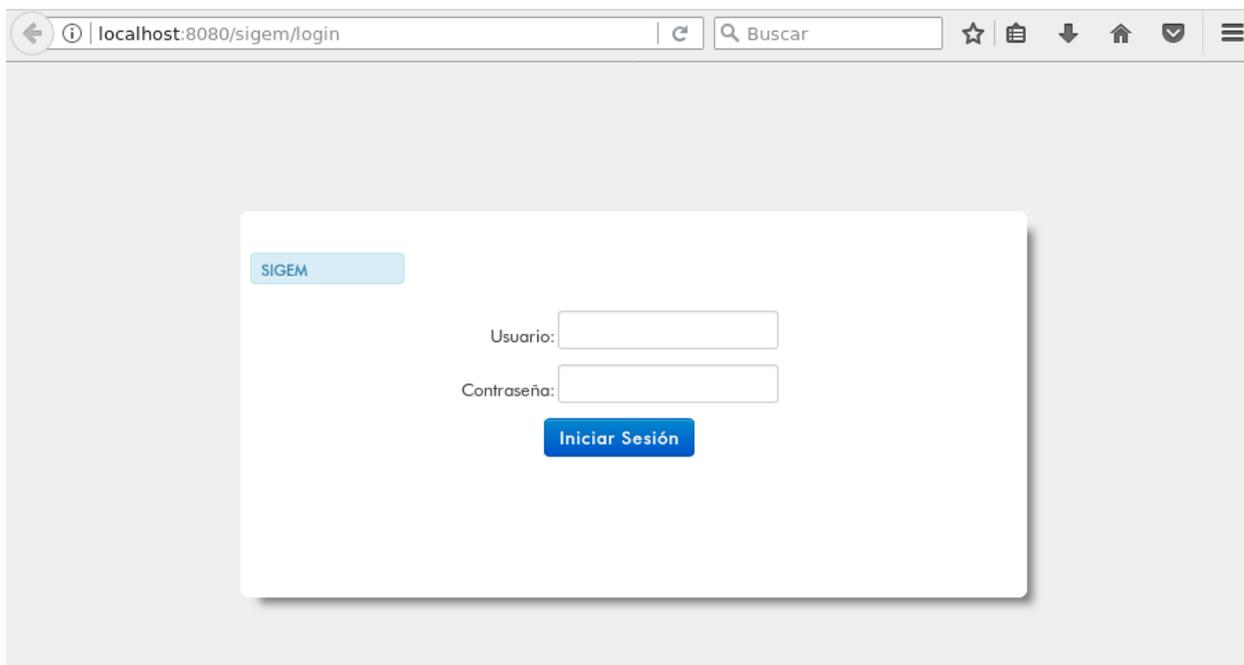
**Figura 3.19.** Flujo de trabajo de la aplicación con el servidor web y el de bases de datos alojados en un mismo equipo.

## Capítulo 4

# Resultados

El capítulo anterior estuvo dedicado a mostrar la estructura interna del sistema, es decir: el modelo de datos, la arquitectura de software y la infraestructura de hardware del mismo. Este capítulo estará dedicado a mostrar la interfaz gráfica, sus componentes y las funcionalidades del sistema.

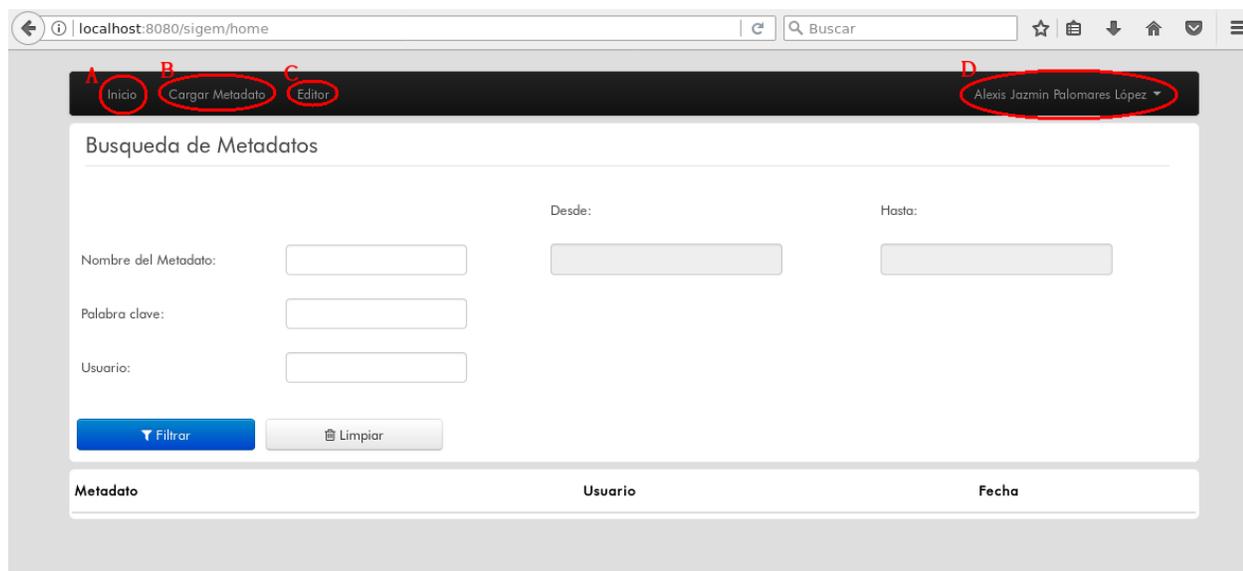
Para acceder al sistema se debe ingresar a través de un navegador de internet. GWT ofrece compatibilidad con los siguientes navegadores: Firefox, Mozilla, Safari, Opera e Internet Explorer; de éstos, el menos recomendable es Internet Explorer debido a que puede presentar incompatibilidades con los principales estándares abiertos. Una vez que se haya ingresado la dirección de la aplicación, el primer paso es iniciar sesión en el sistema (ver figura 4.1).



**Figura 4.1.** Página de inicio de sesión.

Una vez que se ha iniciado sesión e ingresado al sistema, lo primero que puede verse es la página de búsqueda de metadatos. En la parte superior de esta página se encuentra el menú de opciones del

sistema (barra de color negro en la figura 4.2). La opción *Inicio* (A) redirigirá al usuario a la página de búsqueda de metadatos. La opción *Cargar metadato* (B) desplegará una ventana emergente que permitirá al usuario seleccionar un metadato en formato XML y cargar su contenido en la página de edición de metadatos. La opción *Editor* (C) redirigirá al usuario a la página de edición de metadatos. La opción que se encuentra en la extrema derecha (D) sirve para cerrar la sesión del usuario.



**Figura 4.2.** Página de búsqueda de metadatos.

La segunda sección de esta página de inicio es el buscador. Para listar todos los metadatos existentes en el sistema se debe presionar el botón **Filtrar**; también se tiene disponible la opción de buscar metadatos bajo los siguientes criterios: nombre del metadato, palabra clave del metadato, nombre del usuario que creó el metadato y rango de fecha de creación del metadato. Por ejemplo, en la figura 4.3 se muestran los metadatos creados por el usuario Jazmin. Es importante aclarar que para realizar las búsquedas no es necesario escribir palabras completas, basta con escribir un indicio de la palabra, ya sea del nombre del usuario, de la palabra clave o del nombre del metadato (no se autocompletan).

La tercera sección de esta página de inicio es la tabla de resultados. Ésta se llena una vez seleccionados los criterios de búsqueda y después de presionar el botón **Filtrar**. La tabla de resultados se divide en cuatro columnas (ver figura 4.3): el nombre del metadato, el nombre del usuario que capturó el metadato, la fecha de captura del metadato y, en la última columna, hay cuatro botones. Los primeros tres botones sirven para exportar los metadatos en formato PDF, HTML o XML.

La tercera opción del menú del sistema es **Editor**. Al presionar sobre este botón el usuario es redirigido a la página de captura de metadatos. Al principio aparecen todos los formularios colapsados (ver figura 4.4) y al presionar sobre el nombre de cada una de las secciones, éstas se van desplegando, mostrando así los formularios, para que el usuario pueda capturar los metadatos (como en la figura 4.5).

Inicio Cargar Metadato Editor Alexis Jazmin Palomares López

### Busqueda de Metadatos

Desde: Hasta:

Nombre del Metadato:

Palabra clave:

Usuario:

[Filtrar](#) [Limpiar](#)

Metadato	Usuario	Fecha
Clasificación de la Red Carretera (DGDC)	Alexis Jazmin Palomares López	2017-2-09
Corredores carreteros	Alexis Jazmin Palomares López	2016-5-26
Cruces fronterizos de autos	Alexis Jazmin Palomares López	2015-7-06
Plazas de cobro (IMT)	Alexis Jazmin Palomares López	2015-7-06
Puentes	Alexis Jazmin Palomares López	2015-7-06
Rutas carreteras (DGST)	Alexis Jazmin Palomares López	2015-7-06
Centros SCT (IMT)	Alexis Jazmin Palomares López	2015-6-12
Curvas de nivel en el municipio Benito Juárez_ Prueba2	Alexis Jazmin Palomares López	2015-6-11
Marcas de kilometraje	Alexis Jazmin Palomares López	2015-6-11

Figura 4.3. Ejemplo de búsqueda de metadatos.

Inicio Cargar Metadato Editor Alexis Jazmin Palomares López

[Opciones](#)

- 1.- Identificación del conjunto de datos espaciales o producto
- 2.- Fechas relacionadas con el conjunto de datos espaciales o producto
- 3.- Unidad del estado responsable del conjunto de datos espaciales o producto
- 4.- Localización geográfica del conjunto de datos espaciales o producto
- 5.- Sistema de referencia
- 6.- Calidad de la información
- 7.- Entidades y atributos
- 8.- Distribución
- 9.- Información del contacto para los metadatos.

Figura 4.4. Vista inicial de la página de captura de metadatos.

Conforme el usuario captura el metadato, en el extremo derecho de la pantalla, se van mostrando mensajes indicándole al usuario de elementos que le falte por llenar (ver figura 4.5).

The screenshot shows a metadata capture form with the following elements:

- Coordenada lÍmite al Norte:** Input field with value 32.713766 and a green checkmark.
- Tipo de representaci3n espacial:** A list of options:
  - 1. Vector: Los datos vectoriales se utilizan para representar datos espaciales
  - 2. RÁster: Los datos malla o cuadrícula se utilizan también para representar datos espaciales
  - 3. Tabla del texto: Datos de texto o tabulares que se utilizan para representar datos espaciales
  - 4. TIN: Red de triángulos irregulares
  - 5. Modelo estéreo: Visión tridimensional formada por la intersección de los rayos homólogos de un solapamiento de un par de imágenes consecutivas
  - 6. Vídeo: Escena de una grabación de vídeo
- 5.- Sistema de referencia:** A section header with a blue background.
- Sistema de Referencia Horizontal:**
  - Coordenadas geográficas:**
    - Resolución de latitud: 0.0174532925199433
    - Resolución de longitud: 0.0174532925199433
    - Unidades de coordenadas geográficas: Grados decimales (with a green checkmark)
  - Coordenadas planas:**
    - Proyección cartográfica
    - Sistema de Coordenadas de cuadrícula
    - Modelo geodésico
  - Nombre del datum horizontal:** World Geodetic System (with a green plus button)
  - Nombre del elipsoide:** World Geodetic System (with a green plus button)
  - Semieje mayor:** 6378137.00
  - Factor de denominador de achatamiento:** 298.2572235630
- Sistema de referencia vertical:** (partially visible)

On the right side, there is a blue box with the text "Opciones" and a list of instructions:

- Debe llenar los datos de localización geográfica de la sección 4
- Las fechas de referencia del conjunto de datos espaciales ni los nombres de los insumos (Sección 2) pueden repetirse, de lo contrario se conservará únicamente la PRIMERA ocurrencia
- Debe llenar la sección 5 (Sistema de referencia)
- Debe llenar los datos de Entidades y atributos (sección 7).

Figura 4.5. Ejemplo de uno de los formularios de captura.

Una vez que el usuario ha terminado de capturar todos los elementos obligatorios de un metadato, puede guardarlo en la base de datos. Para hacer esto, en la esquina superior derecha se encuentra un botón con el texto **Opciones**. Al presionar este botón se muestran las opciones: **Guardar metadato** y **Descargar avance** (ver figura 4.6).

The screenshot shows the metadata capture form with the following elements:

- Inicio Cargar Metadato Editor:** Navigation links at the top.
- Alexis Jazmin Palomares López:** User name at the top right.
- 1.- Identificación del conjunto de datos espaciales o producto:** Section header with a blue background.
- Nombre técnico:** Input field with value riesgo\_sequia and a green checkmark.
- Opciones:** A dropdown menu with the following options:
  - Guardar metadato
  - Descargar avances

Figura 4.6. Lista de opciones disponibles al editar un metadato por primera vez.

Al presionar sobre la opción **Guardar metadato** primero se realiza una validación y si aún faltan elementos por llenar, se muestran mensajes indicando este hecho (ver figura 4.7).

The screenshot shows a web form with two main sections: "7.- Entidades y atributos" and "8.- Distribución". Under section 7, there are two fields. The first field, "Descripción general de entidades y atributos:", contains the text: "El layer contiene datos sobre la red nacional de carreteras pavimentadas, clasificadas por administración, número de carriles, tipo de tránsito, tipo de pavimento y nombre del tramo carretero. Datos nacionales escala 1:10,000." To the right of this text is a green checkmark. The second field, "Cita del detalle de entidades y atributos:", is empty. To the right of this field is a red error message box that says: "Este campo es obligatorio si el tipo de representación espacial (en la sección 4) corresponde a Vector y/o Tabla de texto".

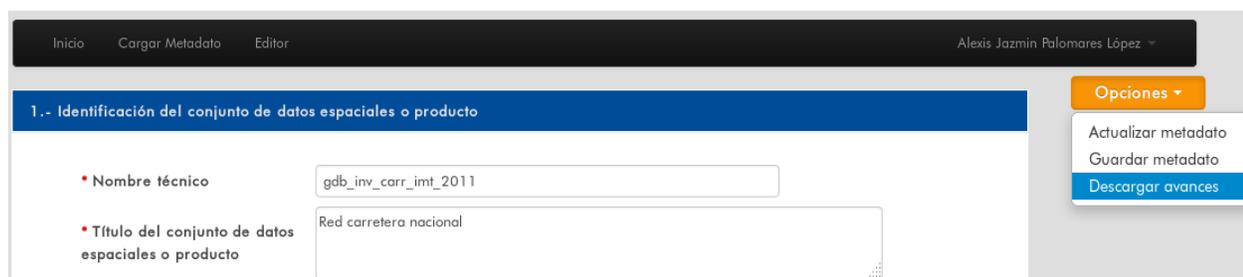
Figura 4.7. Ejemplo de un mensaje de validación.

Por otro lado, si se han llenado todos los elementos obligatorios del metadato, se muestra un recuadro verde con un mensaje indicando que el metadato ha sido guardado exitosamente (ver figura 4.8). Al cerrar este mensaje el usuario es redirigido a la página de inicio (búsqueda de metadatos).

The screenshot shows a web form titled "1.- Identificación del conjunto de datos espaciales o producto". It has a "Opciones" button in the top right. The form contains several fields: "Nombre técnico" with the value "gdb\_inv\_carr\_imt\_2011" and a green checkmark; "Título del conjunto de datos espaciales o producto" with the value "Red carretera nacional" and a green checkmark; "Propósito" with a long text description and a green checkmark; "Descripción del conjunto de datos espaciales o producto" with a long text description and a green checkmark; and "Idioma del conjunto de datos espaciales o producto" with "Español" selected and a green checkmark. A green message box is overlaid on the form, stating: "Se ha guardado exitosamente el metadato Red carretera nacional".

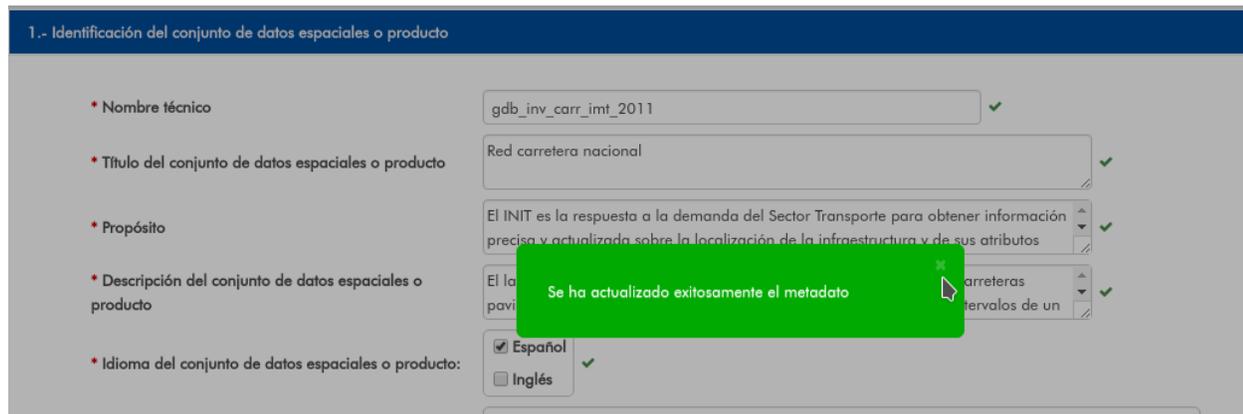
Figura 4.8. Mensaje indicando que ya se ha guardado el metadato.

Dentro de la página de inicio ya se puede consultar el metadato que se acaba de guardar. Si se requiere cambiar algún elemento del metadato, el primer paso es buscarlo en la página de inicio. Una vez que el metadato es encontrado se debe presionar el último botón de la derecha en la lista de metadatos. Con esto se abre la página de captura de metadato; pero en este caso, todos los formularios ya se encuentran llenos con el contenido del metadato. Cuando se abre un metadato, ya guardado en la base de datos, dentro la página de edición, del botón **Opciones** se desprenderán tres opciones: Actualizar metadato, Guardar metadato y Descargar avance (ver figura 4.9).



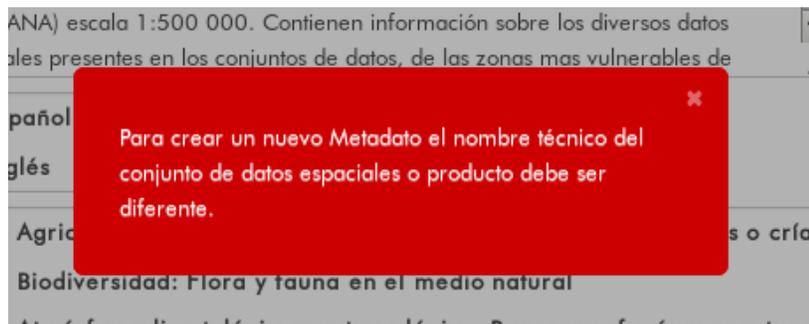
**Figura 4.9.** Lista de opciones disponibles al editar un metadato ya guardado en la base de datos.

Después de hacer los cambios necesarios en el contenido del metadato se presiona la opción **Actualizar metadato**. Con esto, nuevamente se realiza la validación del contenido y si no falta ningún elemento por llenar el sistema muestra un mensaje indicando la correcta actualización del metadato (ver figura 4.10):



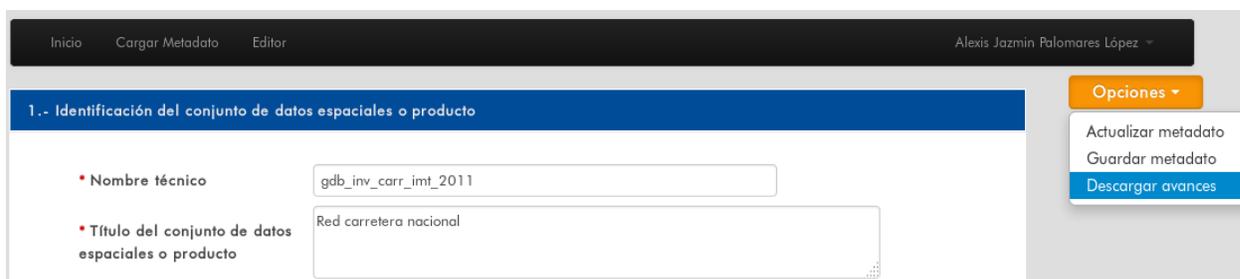
**Figura 4.10.** Mensaje de correcta actualización de un metadato

Si por el contrario, el usuario presiona la opción **Guardar metadato**, se crea un nuevo metadato a partir del original; pero sólo con la restricción de que el campo *Nombre técnico* sea diferente del metadato original. Si no se cumple con esta restricción (además de la de llenar todos los elementos obligatorios del metadato), el sistema no permitirá que el usuario guarde el nuevo metadato y arrojará un mensaje de error (ver figura 4.11).



**Figura 4.11.** Mensaje de error lanzado cuando el nombre técnico no es cambiado.

Por último, si el usuario presiona sobre la opción **Descargar avance** (ver figura 4.12), se descarga en la computadora del usuario el contenido del metadato que en ese momento esté capturando, en formato XML. Esto es de utilidad debido a que el sistema no permite guardar metadatos incompletos, entonces si el usuario desea conservar el avance que lleve, debe presionar sobre este botón. Hecho esto, puede continuar con la captura del metadato, regresar a la página de inicio o cerrar sesión en el sistema.



**Figura 4.12.** Opción “Descargar avances” seleccionada.

Para recuperar el avance se presiona el botón **Cargar metadato** de la barra negra que se encuentra en la parte superior de la página. Con esto, se abre una ventana emergente al centro de la pantalla, la cual contiene dos botones (figura 4.13). Con el primero (de color gris) se abre un explorador de archivos, el cual permite seleccionar el archivo XML que contenga el avance del metadato. Una vez seleccionado el archivo se presiona el botón azul **Enviar**.



**Figura 4.13.** Ventana emergente para la selección de metadatos en formato XML.

Si el archivo seleccionado no tiene ningún cambio en su estructura original, su contenido podrá verse en los formularios de la página de captura de metadatos y este metadato, una vez completado podrá ser guardado dentro del sistema. A través de esta opción, también puede cargarse en el sistema cualquier metadato creado con el SICAM.

Este sistema tiene la capacidad de recibir y procesar peticiones externas para exportar metadatos en formato PDF, XML o HTML. Esto permite que los metadatos guardados en este sistema puedan ser publicados dentro del portal geográfico de la U-GITS, tal como se muestra en la figura 4.14.

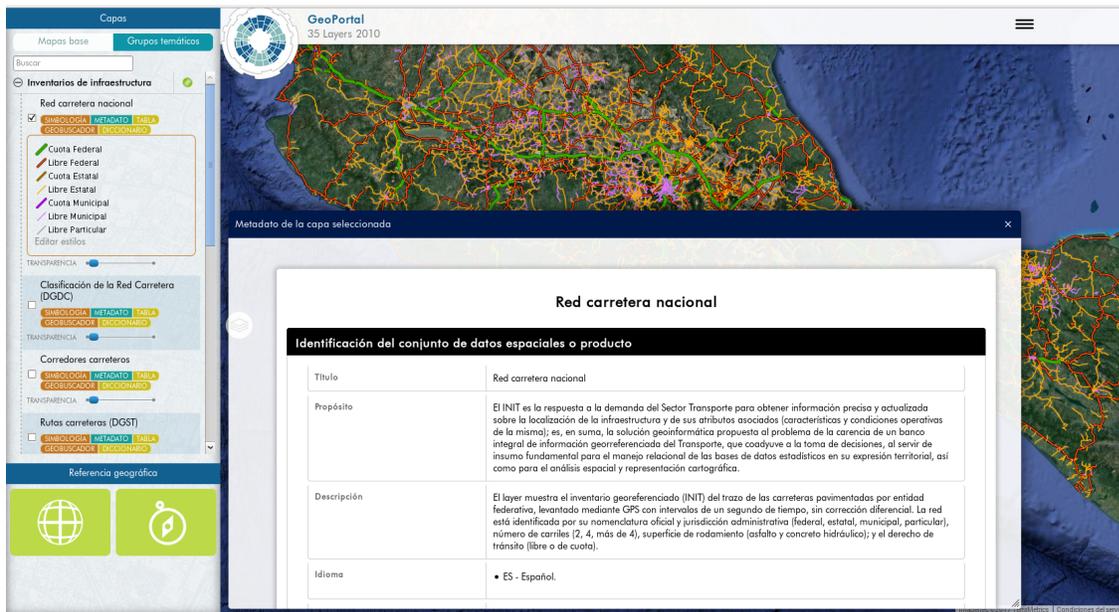


Figura 4.14. Ejemplo de un metadato mostrado a través del portal geográfico de la U-GITS.

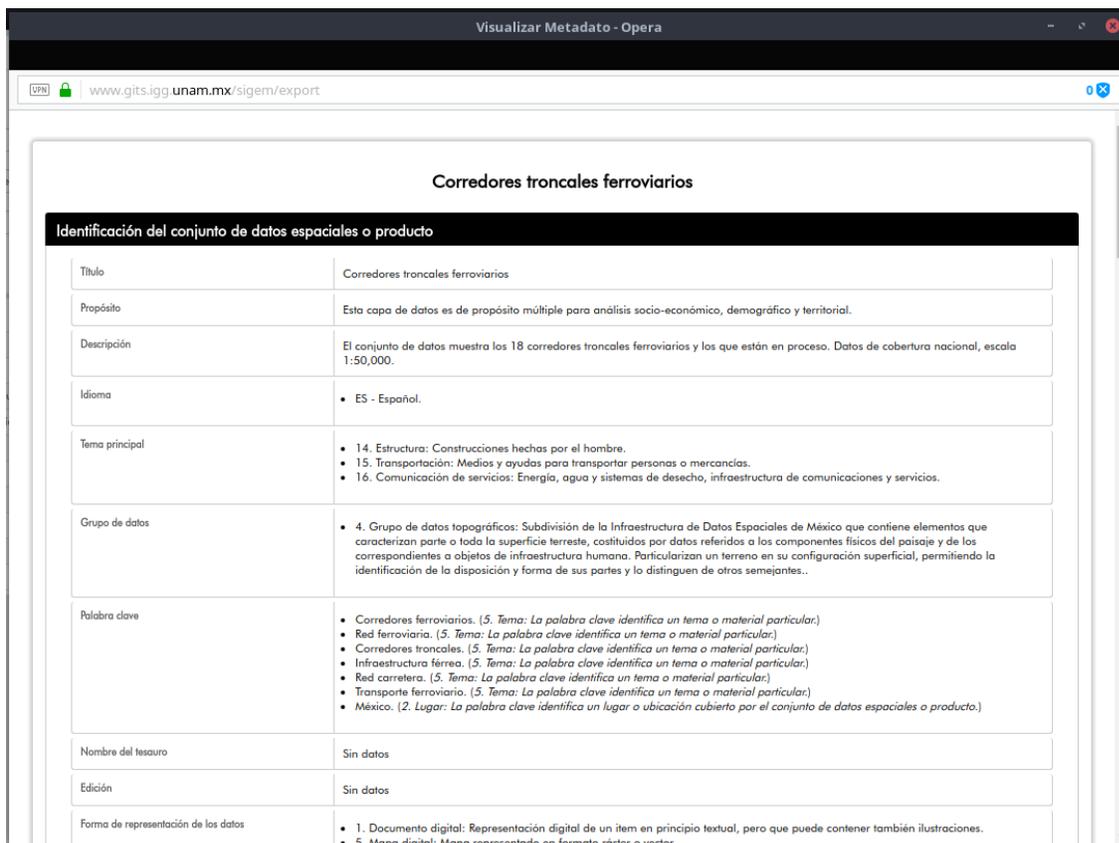


Figura 4.15. Ejemplo de metadatos en formato HTML.

Terracerías y caminos no pavimentados - Opera

UPN [www.gits.igg.unam.mx/sigem/export](http://www.gits.igg.unam.mx/sigem/export)

Terracerías y caminos no pavimentados 2 / 6

Unidad GITS - Sistema Generador de Metadatos - Terracerías y caminos no pavimentados

Identificación del conjunto de datos espaciales o producto	
<b>Título</b>	Terracerías y caminos no pavimentados
<b>Propósito</b>	Esta capa de datos es de propósito múltiple para análisis socio-económico, demográfico y territorial.
<b>Descripción</b>	La capa muestra los datos vectoriales de las carreteras no pavimentadas junto con las terracerías del Continuo Nacional de Vías de Comunicación, escala 1:250,000 a nivel nacional.
<b>Idioma</b>	- ES - Español.
<b>Tema principal</b>	- 14. Estructura: Construcciones hechas por el hombre. - 15. Transportación: Medios y ayudas para transportar personas o mercancías.
<b>Grupo de datos</b>	- 4. Grupo de datos topográficos: Subdivisión de la Infraestructura de Datos Espaciales de México que contiene elementos que caracterizan parte o toda la superficie terrestre, constituidos por datos referidos a los componentes físicos del paisaje y de los correspondientes a objetos de infraestructura humana. Particularizan un terreno en su configuración superficial, permitiendo la identificación de la disposición y forma de sus partes y lo distinguen de otros semejantes..
<b>Palabra clave</b>	- Carreteras pavimentadas (5. Tema: La palabra clave identifica un tema o material particular.) - Terracerías (5. Tema: La palabra clave identifica un tema o material particular.) - Red carretera (5. Tema: La palabra clave identifica un tema o material particular.) - Derecho de tránsito (5. Tema: La palabra clave identifica un tema o material particular.) - Jurisdicción (2. Lugar: La palabra clave identifica un lugar o ubicación cubierto por el conjunto de datos espaciales o producto.) - México (2. Lugar: La palabra clave identifica un lugar o ubicación cubierto por el conjunto de datos espaciales o producto.) - Vías de comunicación (5. Tema: La palabra clave identifica un tema o material particular.)
<b>Nombre del tesoro</b>	Sin datos
<b>Edición</b>	Sin datos
<b>Forma de representación de los datos</b>	- 1. Documento digital: Representación digital de un ítem en principio textual, pero que puede contener también ilustraciones. - 5. Mapa digital: Mapa representado en formato ráster o vector. - 11. Tabla digital: Representación de hecho o figuras sistemáticamente representadas, especialmente en columnas para su utilización directa por los usuarios.
<b>Url del recurso</b>	<a href="http://www.gits.igg.unam.mx/idea">http://www.gits.igg.unam.mx/idea</a>
<b>Descripción del acceso al recurso</b>	Geoportal de la Infraestructura de Datos Espaciales de la Unidad GITS.
<b>Frecuencia de mantenimiento</b>	10. Desconocido: La frecuencia de mantenimiento de los datos no se conoce..
<b>Conjunto de caracteres</b>	4. Utf8: Formato de Transferencia UCS de tamaño variable de 8-bit, basado en ISO/IEC 10646.
<b>Nombre del archivo gráfico</b>	<a href="https://www.gits.igg.unam.mx/idea/img/cat_thumbs/inegi_camino_50_2000.jpg">https://www.gits.igg.unam.mx/idea/img/cat_thumbs/inegi_camino_50_2000.jpg</a>
<b>Uso específico</b>	Para uso en los proyectos que genera la Unidad GITS.

Figura 4.16. Ejemplo de metadatos en formato PDF.

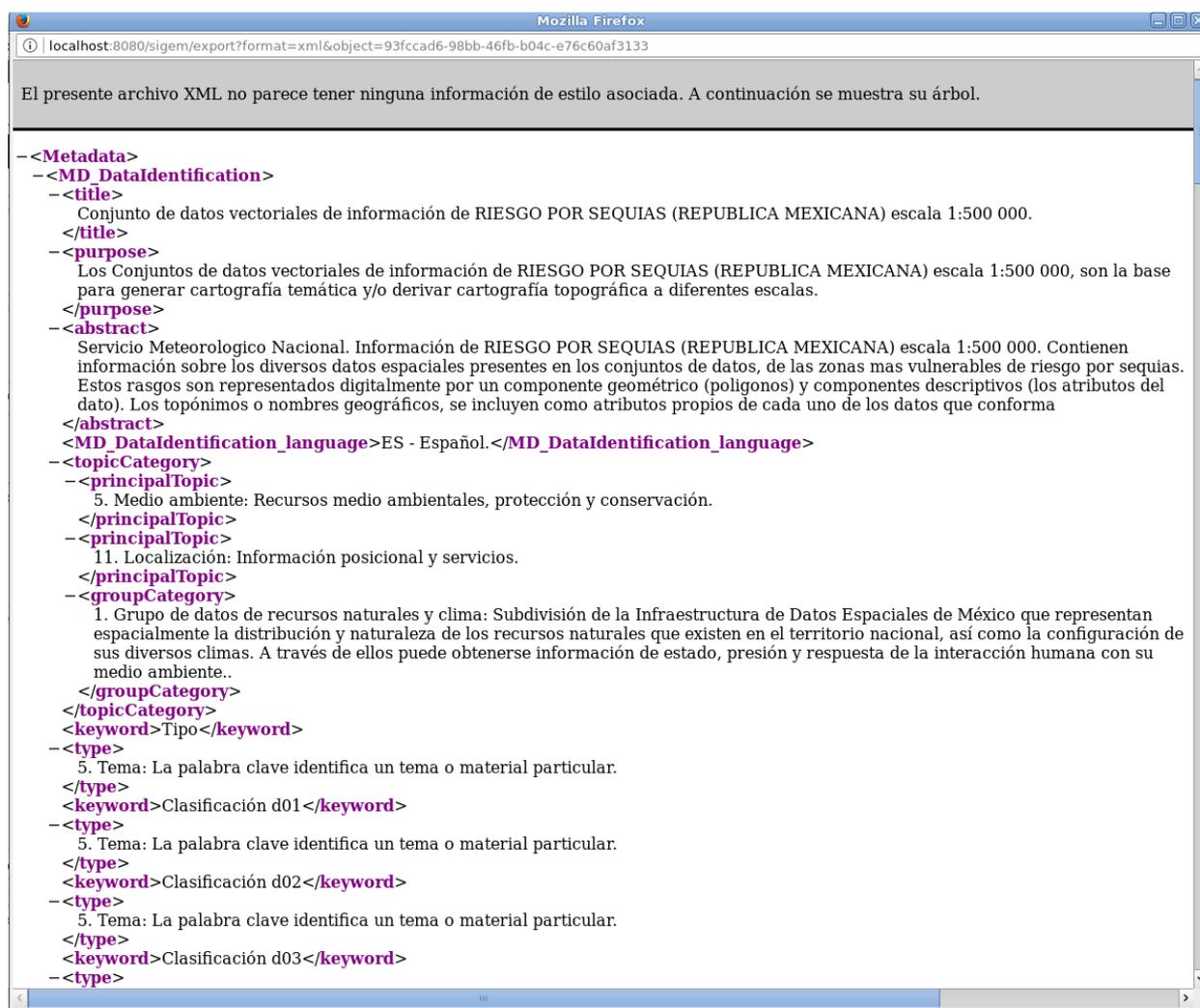


Figura 4.17. Ejemplo de metadato en formato XML.

Con respecto a la visualización dentro de diferentes dispositivos, cabe mencionar que el diseño de la interfaz gráfica del sistema no consideró en su inicio técnicas de diseño responsivo, por lo tanto, se recomienda el uso de computadoras de escritorio, portátiles y tabletas, quedando descartados los teléfonos celulares debido a que el espacio de visualización de estos dispositivos es más limitado y no permite el despliegue de algunos componentes de la interfaz gráfica.

Para el caso de las tabletas con sistema operativo iOS, éstas pueden presentar problemas con la descarga del avance en la captura de metadatos en formato XML, pues este tipo de dispositivos presenta limitantes con el tipo de archivos que se pueden descargar.

## Capítulo 5

# Conclusiones

En este capítulo se muestran las conclusiones derivadas de la realización de este trabajo, las cuales se dividen en tres rubros: el primero de ellos está dedicado a los logros obtenidos con este trabajo; el segundo rubro está dedicado a las mejoras que en un futuro podrían realizarse al sistema; finalmente, el tercer rubro está dedicado a mostrar la importancia de los metadatos geográficos y de la norma de metadatos así como de algunos aspectos relacionados con el plan de estudios del ingeniero geomático.

### Logros

- El sistema descrito y documentado en este trabajo cumple con los requerimientos planteados inicialmente, entre ellos, proveer un repositorio centralizado para los metadatos de las capas de información geográfica procesadas y publicadas a través de los portales de la U-GITS, así como implementar una herramienta web de captura de metadatos geográficos con base en la NTM del INEGI y cumplir así con lo establecido en esta norma. En otras palabras, esta herramienta forma ya parte de los proyectos y portales de la U-GITS.
- A diferencia de otros productos como los presentados en la figura 2.3 éste sí emplea la norma de elaboración de metadatos del INEGI, aunque también es la única norma que emplea. Respecto a la tecnología, además de utilizar Java, también emplea GWT y Bootstrap que sirven, entre otras cosas, para la implementación de la interfaz gráfica y Spring que permitió implementar la arquitectura de software seleccionada. Estas cuatro tecnologías, en conjunto con las demás abordadas en la sección 3.2.2, permitieron la implementación tanto de la aplicación para captura y búsqueda de metadatos como del servicio de consulta de metadatos, ambos en una plataforma web (que además permite su uso por parte de varios usuarios concurrentes), lo que facilita el acceso y uso por parte de cualquiera de sus usuarios, pues éstos únicamente requerirán de una computadora con acceso a internet y un navegador web que permita ejecutar javascript, ambos presentes en casi la totalidad de los equipos de cómputo actuales. Aunado a lo anterior, este sistema utiliza formatos de entrada y de salida interoperables como XML, HTML y PDF, que pueden ser abiertos por cualquier navegador web y/o editor de textos así como cualquier lector de pdf, mismos que de igual forma, ya se encuentran instalados en casi cualquier computadora actual.
- Otra de las ventajas de este sistema es que además de la búsqueda de metadatos completos con base en el nombre del conjunto de datos, palabra clave o por nombre del usuario que

crea el metadato; el modelo de datos también permitiría responder otra clase de “preguntas” relacionadas por ejemplo con los insumos con que se crearon las capas (como nombre o fecha de creación de éstos), o cuántos metadatos corresponden con algún tema en específico o cuántos de los recursos se distribuyen en un formato específico, entre otras que permitan mostrar estadísticas básicas sobre los conjuntos de datos o bien, de uso del sistema.

- Aunque no se mencionó a lo largo de este trabajo; ya se han hecho algunas actualizaciones al sistema, gracias a las cuales también se pudo hacer uso de la Norma Técnica para el acceso y publicación de Datos Abiertos de la Información Estadística y Geográfica de Interés Nacional así como exportar metadatos bajo el esquema ISO 19139.

## Trabajo futuro

- Por otro lado, si bien, el sistema ya cumple con los requerimientos planteados para la U-GITS, puede robustecerse con algunos elementos que hagan la captura de metadatos una tarea más sencilla, por ejemplo, el llenando de manera automática algunas partes del metadato como: el nombre y descripción de los atributos que componen la capa (campo *Cita del detalle de entidades y atributos* correspondiente a la sección 7 de la NTM), o los datos referentes al sistema de referencia espacial de la capa; ya sea leyendo estos datos directamente desde la capa de origen (por lo regular en formato SHP) o de la base de datos espacial en la que se alojan todas estas capas. Así mismo también podría resultar conveniente que se llevara una “agenda” para manejar los datos de los puntos de contacto o de las unidades del estado encargadas de los conjuntos de datos espaciales, de modo que el llenado de estos datos también pueda automatizarse.
- Otra de las cuestiones que pueden mejorarse en el sistema es que permita el uso y manejo de más estándares de metadatos, por ejemplo el de CSDGM del FGDC, las normas ISO 19115 o Dublin Core, por poner algunos ejemplos.
- Por ahora, más del 90 % de las capas publicadas por la U-GITS son de datos que abarcan toda la República Mexicana, por lo que la localización del conjunto de datos espaciales no constituye un requisito para la búsqueda de los metadatos de estas capas. Sin embargo, podría resultar de utilidad incluir también un mapa en el buscador en el que el usuario seleccione el área geográfica de los datos, para las capas que no son de cobertura nacional.

## Consideraciones finales

- La construcción y puesta en marcha de sistemas de información geoespacial, o de sistemas de información en general, es una tarea que requiere de una adecuada planeación pues deben conocerse las necesidades que éste debe cubrir y verificar que ayude a cumplir con los objetivos del grupo de trabajo, entre otras cuestiones. Debido a lo anterior, se debe considerar que los ingenieros geomáticos, como especialistas en el tratamiento de información georreferenciada, pueden llegar a ser líderes de proyectos orientados a la construcción de este tipo de sistemas y por tanto deben estar plenamente involucrados en todas las etapas de planeación para que el desarrollo del mismo llegue a buen fin. Si bien, existen diferentes metodologías de desarrollo, puede considerarse la estrategia mostrada en la figura 2.2 y que fue abordada a lo largo del capítulo 2, en la que se consideran algunas cuestiones como la identificación de necesidades,

principios de planeación estratégica (por medio de un análisis y definición de estrategias FODA) para concluir con la definición de requerimientos que permitirán estimar el alcance del sistema.

- Los metadatos también pueden vincularse a una cuestión de *ética cartográfica*. Considerando a la ética como los principios de conducta que guían las prácticas de un individuo o grupo profesional, que en este caso se refiere al de los generadores de cartografía (Harley, 2002, p.199); una *ética cartográfica* cobra relevancia debido a dos aspectos principales: el primero de ellos está orientado al aspecto tecnológico, ya que con las nuevas tecnologías es mucho más fácil infringir los derechos de copia del material cartográfico (y en general de cualquier tipo de obra) o bien, de la tecnología con que se genera la cartografía; así mismo tiene que ver con las herramientas con que se produce la información pues es necesario considerar hasta qué punto se debe confiar en la automatización de los procesos que ayudan en la generación cartografía, que van desde la creación de insumos hasta la obtención de los mapas. De igual forma se relaciona con los medios por los cuales este material es puesto a disposición del público o no.

El segundo de estos aspectos está directamente relacionado con el contenido de la información. Debido a que la información geográfica puede ser utilizada para la toma de decisiones estratégicas, al menos desde una perspectiva gubernamental, es decir, que está relacionada con el ejercicio del poder, por lo que, de acuerdo con Harley (2002) es importante que quienes generen cartografía (que en este punto ya es una forma de conocimiento) con base en esta información estén conscientes de las consecuencias que ésta podría tener al ser publicada e incluso cuando es omitida (el autor cita el ejemplo de que el Servicio Geológico de los Estados Unidos no pone en su cartografía los sitios donde se vierten residuos nucleares). Este aspecto también es importante porque quienes generan cartografía pueden ser relegados al rol de representar en un mapa datos e información generada por un tercero por lo que puede no quedar claro con qué propósito se generó la información que se está publicando.

Entonces, los metadatos son una forma de abordar esta *ética cartográfica*, pues por medio de ellos el generador información geográfica y de cartografía puede declarar: el propósito con que se generó dicha información además del uso específico que éste le dio a la información (sección 1 de la NTM); con qué insumos fue generada y los procesos aplicados a los insumos para llegar al resultado final (secciones 2 y 6 de la NTM); la extensión espacial y proyección cartográfica de la información (secciones 4 y 5 de la NTM); las pruebas de calidad realizadas a los datos antes de publicarlos (sección 6 de la NTM); el detalle de la información (sección 7 de la NTM); los medios por los que ésta será distribuida, el formato de distribución, si tiene restricciones de uso y/o acceso por cuestiones de licencias o derechos de autor, así como la responsabilidad del generador o custodio de la información en la distribución de la misma (sección 8 de la NTM).

- Entonces, el trabajo presentado a lo largo de este texto constituye un esfuerzo de la U-GITS para publicar información de interés nacional; pero sobre todo de calidad y apegándose a los lineamientos de la NTM. Si bien, por el momento el estado en el que se encuentra el sistema cumple con los requerimientos planteados por la U-GITS, el sistema tiene la flexibilidad de agregar elementos o funcionalidades, como las mencionadas en la sección de *Trabajo futuro*, para una versión posterior.
- Finalmente, respecto a los conocimientos requeridos para el desarrollo de este sistema así como de este trabajo de investigación resulta muy importante que los ingenieros geomáticos además de los conocimientos propios de la materia como: topografía, geodesia, sistemas de información geográfica, percepción remota, cartografía, etc. (que para el contexto de este trabajo brindan

las bases para la creación de capas y productos de información geográfica así como su correcta documentación) tengan también nociones, aunque no lleguen a ser especialistas, de administración de proyectos y programación, por lo que fue fundamental haber contado con estas asignaturas en el plan de estudios de la carrera.

Sin embargo, apegándose al concepto de geomática y teniendo en cuenta la fuerte influencia que esta disciplina tiene del campo de las ciencias de la computación resulta también de importancia contar con conocimientos más profundos de uso y modelado de bases de datos, bases de datos espaciales, programación orientada a objetos, servicios geoespaciales e incluso de ingeniería de software, por lo que éstas son cuestiones que pueden considerarse para futuros planes de estudios. Así mismo falta reforzar cuestiones relacionadas con normatividad geoespacial tanto nacional como internacional, contar con mayor conocimiento de las normas elaboradas por el INEGI o por organismos como el OGC y el comité técnico 211 de la ISO, entre otros.

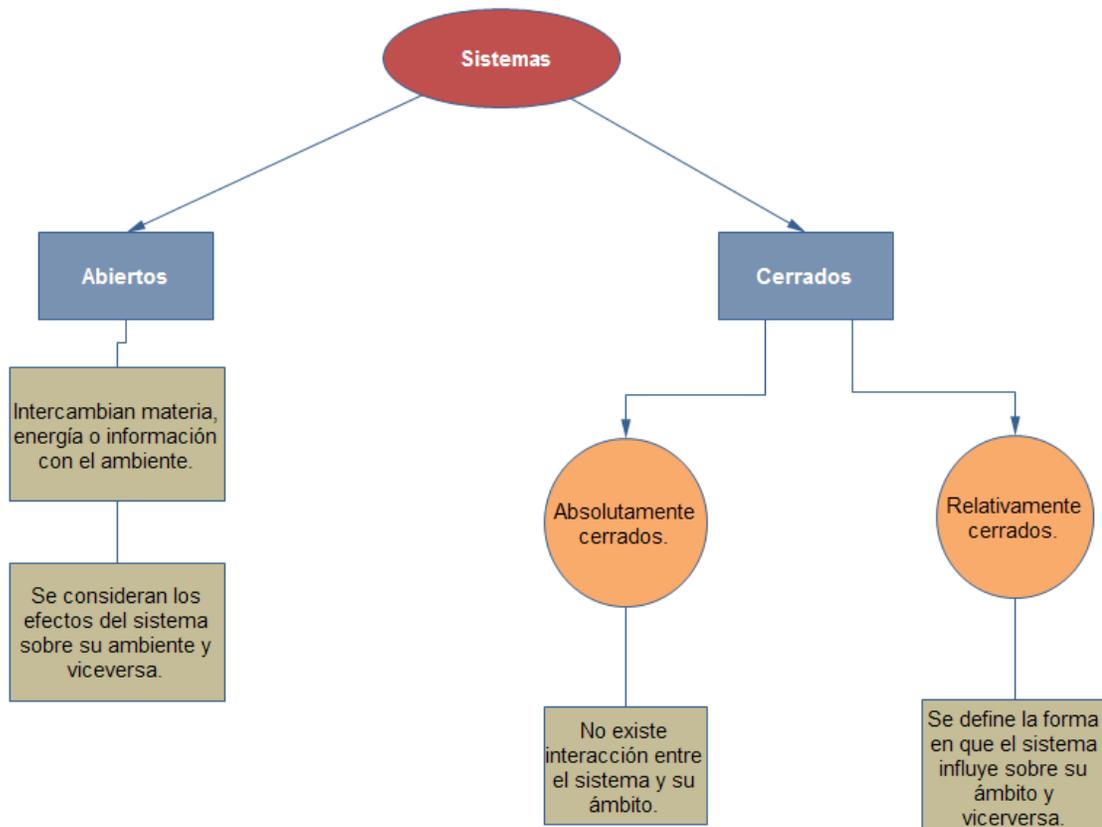
# Anexos



# Apéndice A

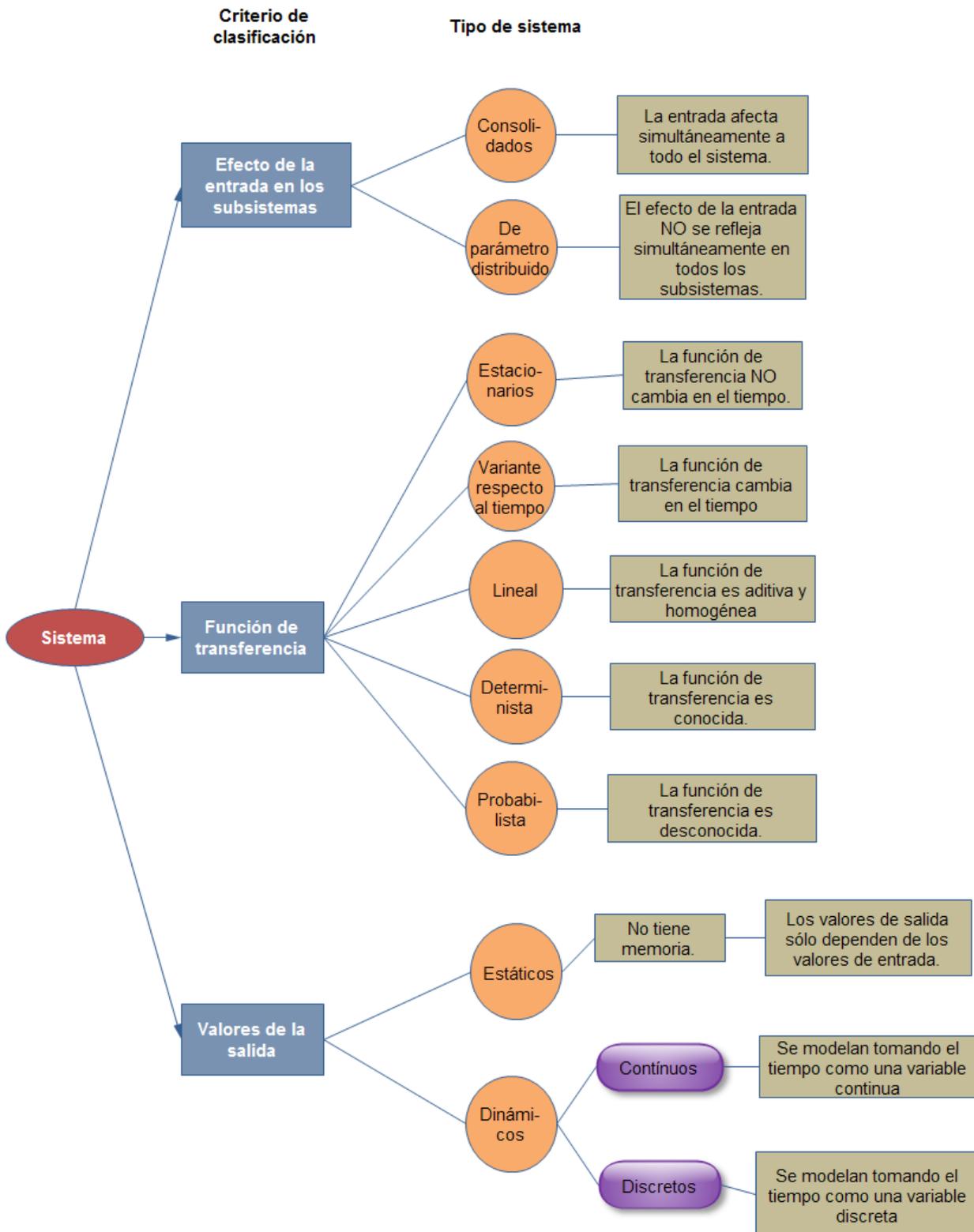
## Clasificación de los sistemas

La primera forma de clasificación se lleva a cabo considerando la relación de los sistemas con su ambiente o ámbito, lo que lleva a la clasificación mostrada en la figura A.1



**Figura A.1.** Clasificación de los sistemas según la relación con su ambiente. Elaboración propia con base en Moreno Bonett (2010, pp.19-22).

Por otro lado, existe otra clasificación considerando el tiempo, pues al hacer un análisis de los sistemas, es posible determinar la salida de los mismos en cualquier instante. Es por ello que en estas clasificaciones el tiempo se vuelve una variable independiente de las funciones de entrada y de salida. Vea la Figura A.2.



**Figura A.2.** Clasificación de los sistemas considerando al tiempo como variable. Elaboración propia con base en Moreno Bonett (2010, pp.19-22).

La tercera forma lleva a la clasificación de los sistemas en orgánicos y no-orgánicos. De acuerdo con Moreno Bonett (2010) un “sistema orgánico es un todo que se autodesarrolla pasando por diferentes etapas de compilación y diferenciación” (p.19). Las principales características que distinguen a un sistema orgánico de otro no orgánico son:

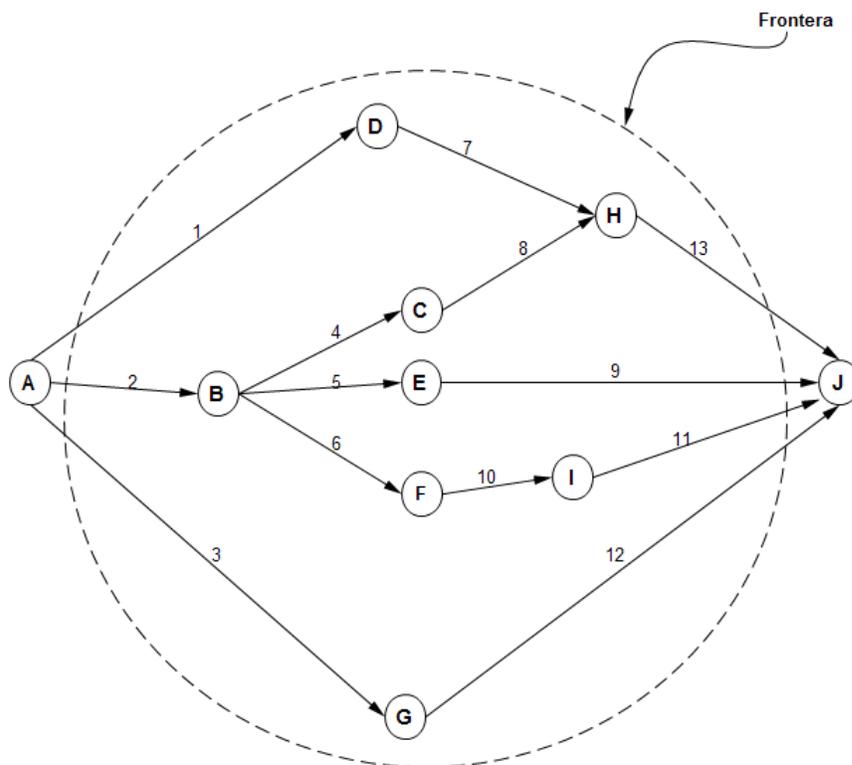
1. Presencia de conexiones no sólo estructurales, sino también genéticas entre sus elementos.
2. Existencia de coordinación y subordinación entre sus elementos.
3. Existencia de mecanismos de control.
4. Las principales propiedades de sus partes se determinan por las leyes y la estructura del todo.
5. La actividad de una cualquiera de sus parte refleja en una mayor actividad del todo.
6. La transformación de una cualquiera de sus partes implica una transformación del todo.



## Apéndice B

# Estructura de los sistemas

Un *grafo de flujos*, que también sirve para mostrar la estructura de un sistema, al igual que el diagrama de bloques, muestra relaciones entre nodos; pero en este caso los nodos representan las variables del sistema y los arcos representan las transformaciones que sufren los nodos. Si se quisiera obtener un grafo de flujos a partir del diagrama de bloques mostrado en la figura 1.2, en cada uno de los nodos estarían representadas la variables del sistema, es decir, de las letras A a la J. Los arcos del grafo representan las transformaciones aplicada a dichas variables. Un ejemplo de este tipo de diagrama se muestra en la figura B.1.



**Figura B.1.** Grafo de flujo derivado del diagrama de bloques de la Figura 1.2.

Aunado a estas dos formas de representación de los sistemas (diagrama de bloques y grafo de

flujos) también existe lo que se conoce como diagramas de PERT y, aunque éstos están más orientados hacia la gestión de proyectos, también constituyen una forma de representación de sistemas.

PERT(Program Evaluation and Review Techniques) es un método que se enfoca a la duración de los proyectos, es decir, sirve para obtener un estimado del tiempo total que podría tomar la realización de un proyecto con base en la duración de cada una de las tareas que involucra el proyecto. Para el cálculo de la duración de una actividad se emplea la siguiente fórmula (Project Management Institute, 2008, p.133):

$$t_E = \frac{(t_O - 4t_M + t_P)}{6}$$

Donde:

$t_E$ = Duración esperada.

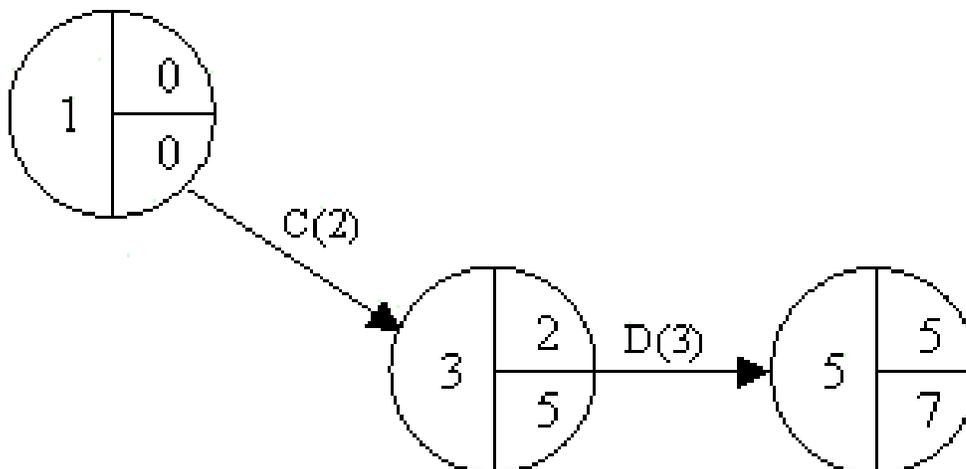
$t_O$ = La mejor duración, basada en el mejor escenario (Optimista) .

$t_M$ =Duración más probable (Más probable).

$t_P$ =La peor duración, basada en el peor de los escenarios (Pesimista).

En un diagrama de PERT se muestran las relaciones que hay entre las tareas del proyecto. Al igual que con los dos tipos de diagramas anteriores cuenta con nodos y arcos. Los arcos representan las tareas a realizar junto con su duración, mientras que los nodos representan periodos durante la ejecución del proyecto; es decir, en cada nodo se muestra el instante de inicio y el instante de final para dos actividades consecutivas.

En la figura B.2 se muestran las actividades C y D (arcos del grafo). El nodo 1 representa el momento de inicio, por lo que el instante de inicio y el instante de término para este nodo es 0. Después del instante de inicio, comienza la actividad C cuya duración es de 2 (La duración es el número entre paréntesis). En el nodo 3 (el número del nodo se encuentra a la izquierda en cada uno) el periodo va del instante 2 al instante 5 y está ligado con la actividad D cuya duración es 3.

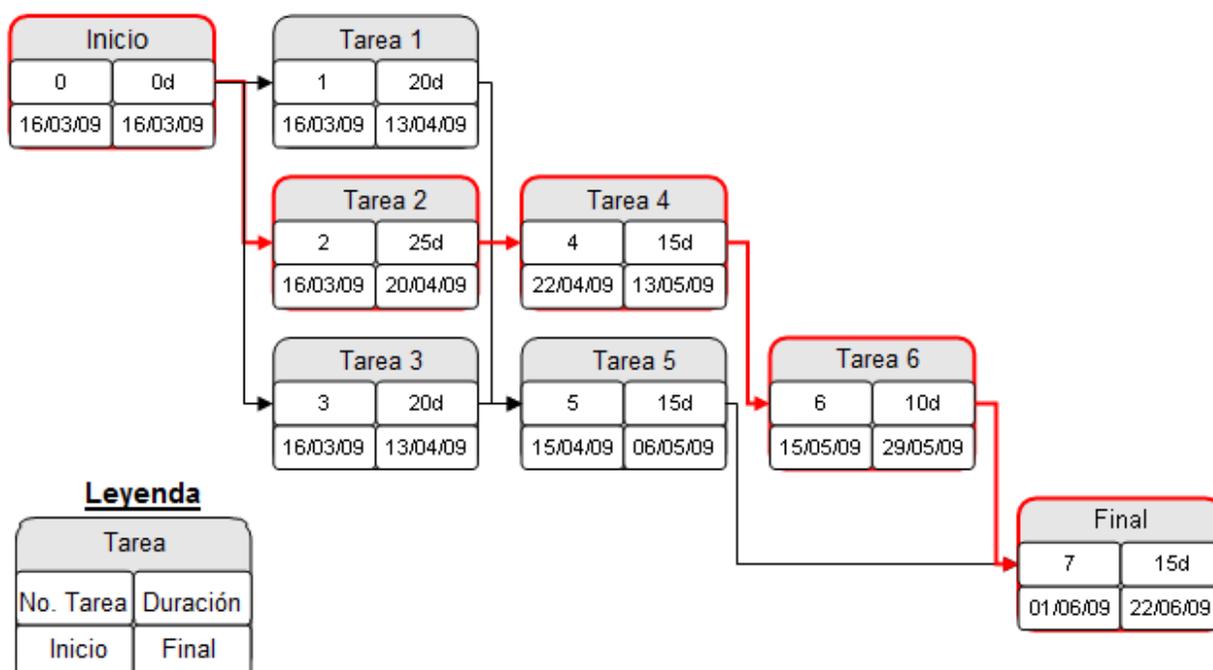


**Figura B.2.** Extracto de un diagrama de PERT. (Fuente: ADA Consultores (2015)).

Existen otras formas más sencillas de elaborar los diagramas de PERT, como la de la Figura B.3 en la que cada nodo lleva integrados los siguientes datos: número de tarea, duración, fecha inicio y fecha de término y es más fácil de interpretar que el diagrama del ejemplo anterior.

La razón de presentar este tipo de diagramas está relacionada con el contenido de los capítulos 2 y 3. En el capítulo 2 se presenta un diagrama de caja negra de este sistema de captura de metadatos, el cual es también un diagrama de bloques como el que se explicó en este apartado. Entonces, la elaboración del diagrama de caja negra del sistema permitirá identificar el tipo de estructura del mismo. Los diferentes tipos de estructuras de sistemas se mencionan más adelante en este mismo apartado.

Por otro lado, aunque la elaboración de diagramas de PERT obedece más a un fin de administración de proyectos, el incluir uno en este trabajo ayudará a identificar las etapas de la elaboración de este sistema, las cuales abarcan desde la identificación de requerimientos hasta la puesta en marcha del sistema.

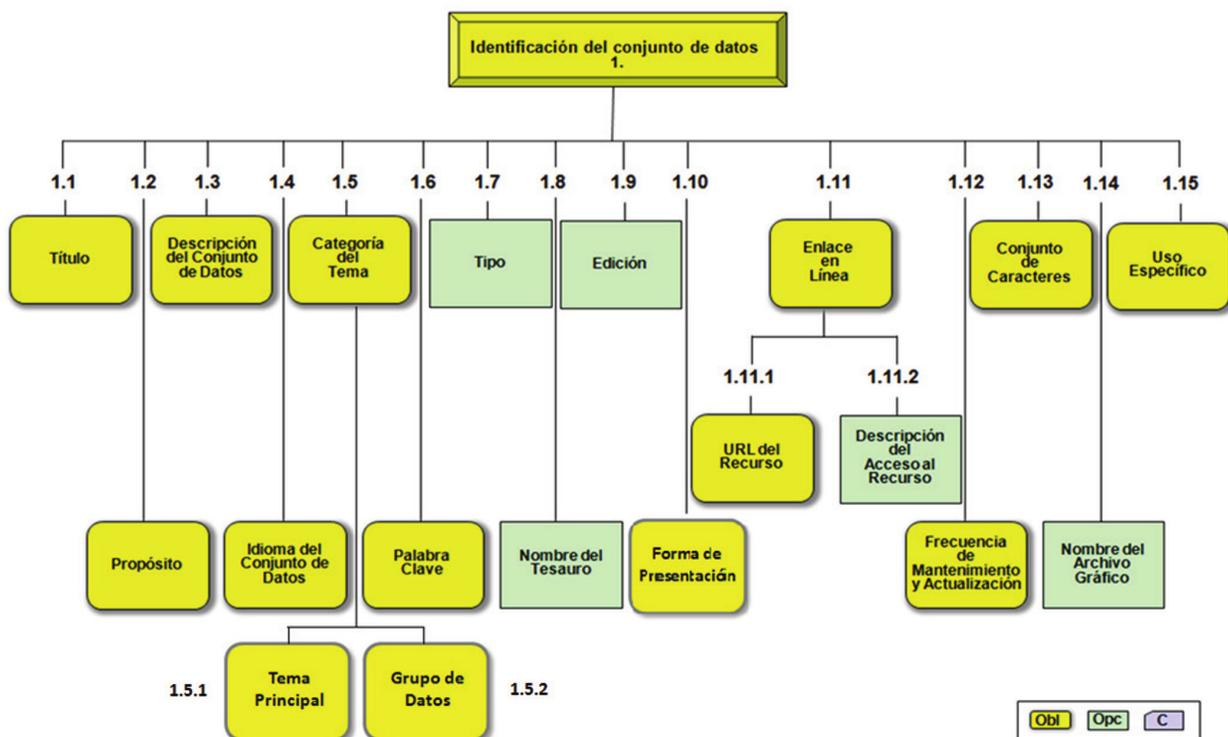


**Figura B.3.** Diagrama de Pert. Adaptado de: Free-Logistics (2015)



## Apéndice C

### Elementos de la NTM



**Figura C.1.** Elementos de la primera sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b)

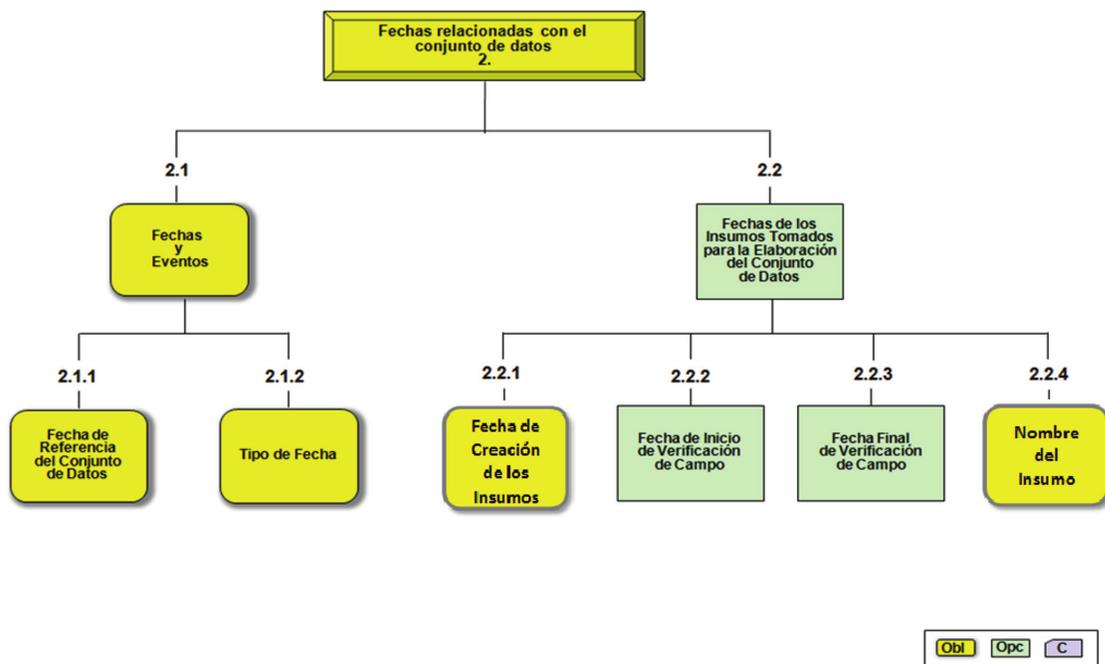


Figura C.2. Elementos de la segunda sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b)

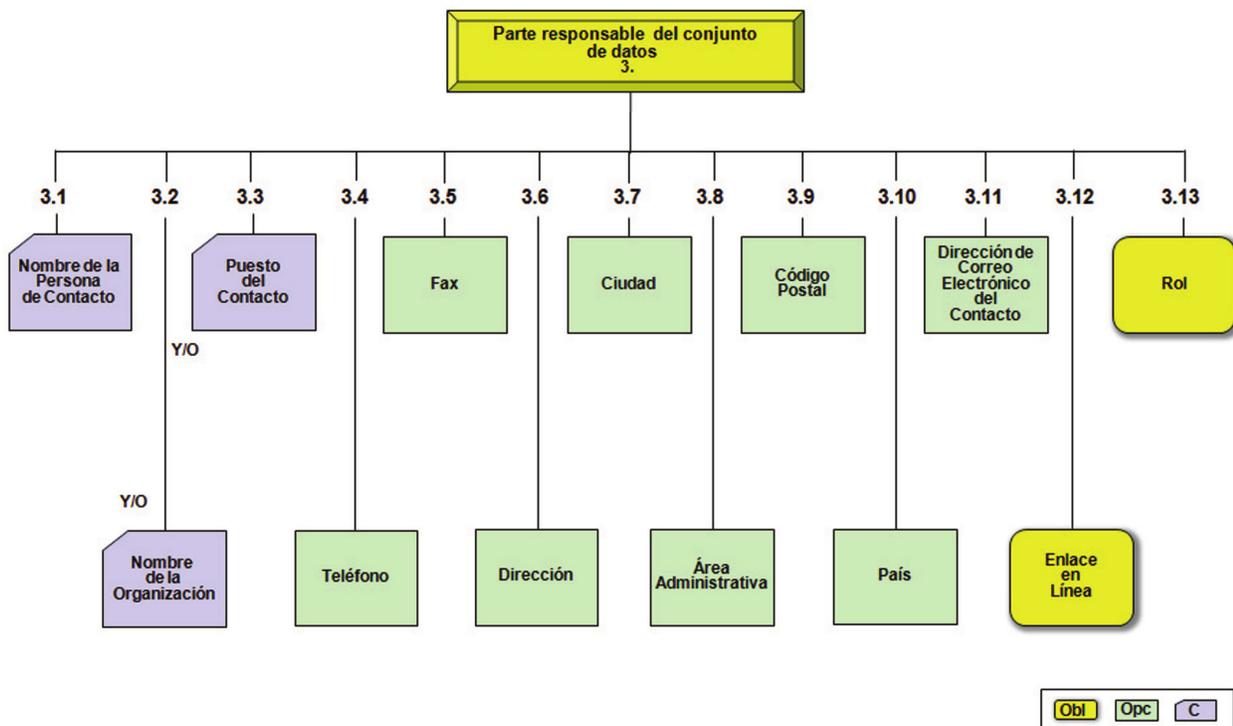
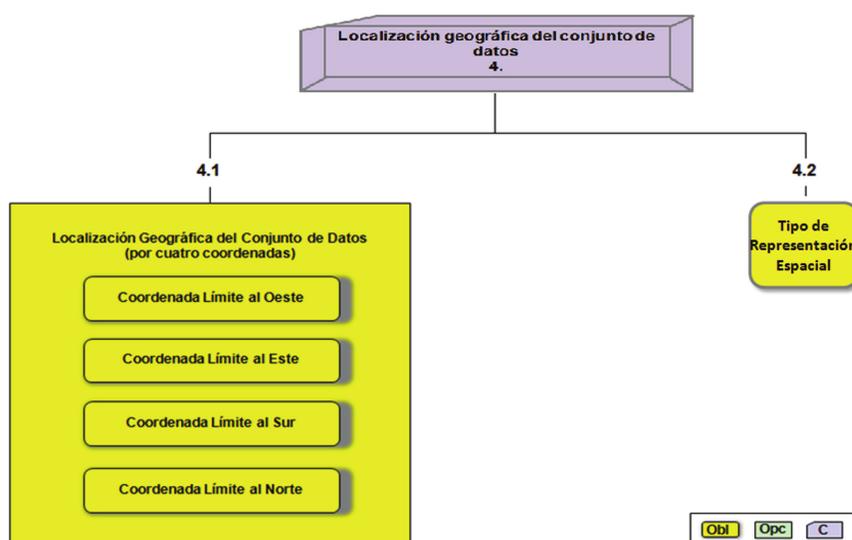
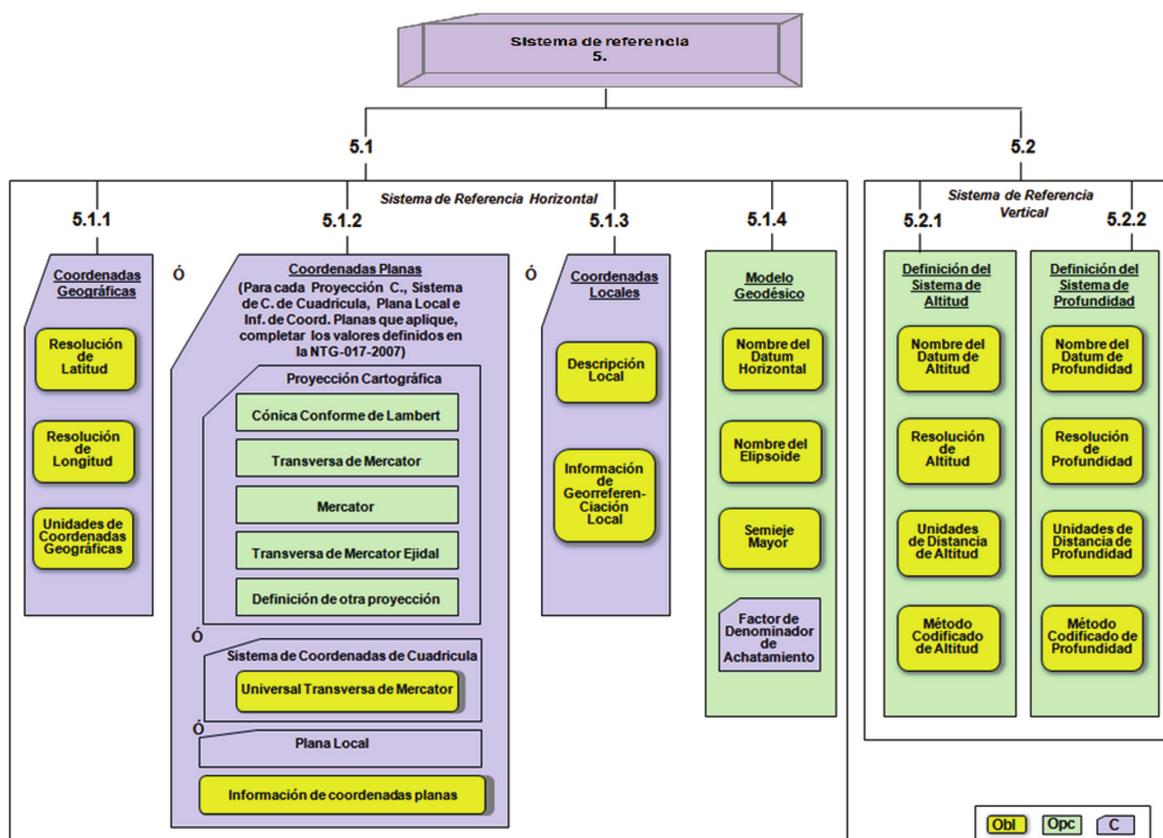


Figura C.3. Elementos de la tercera sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b)



**Figura C.4.** Elementos de la cuarta sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b)



**Figura C.5.** Elementos de la quinta sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b)

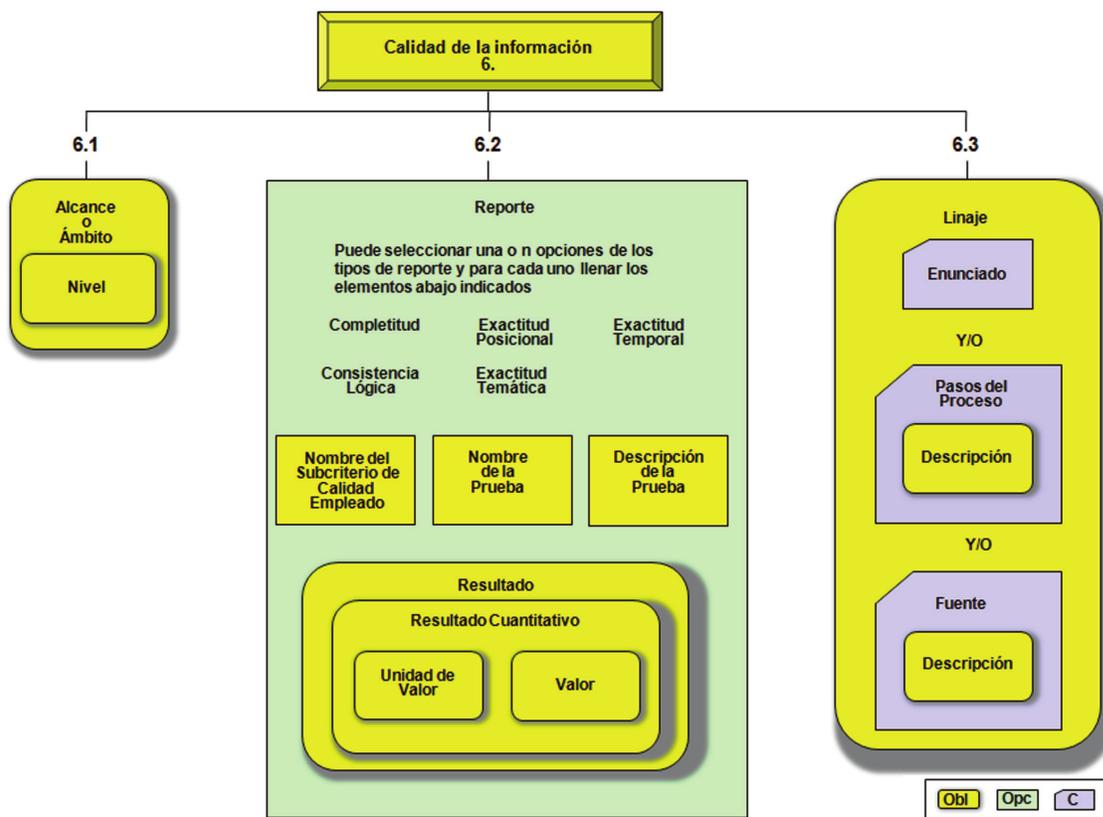


Figura C.6. Elementos de la sexta sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b)

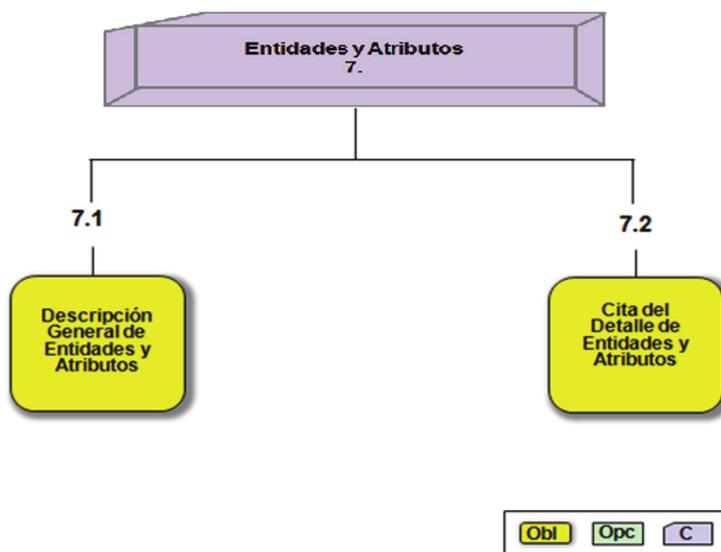
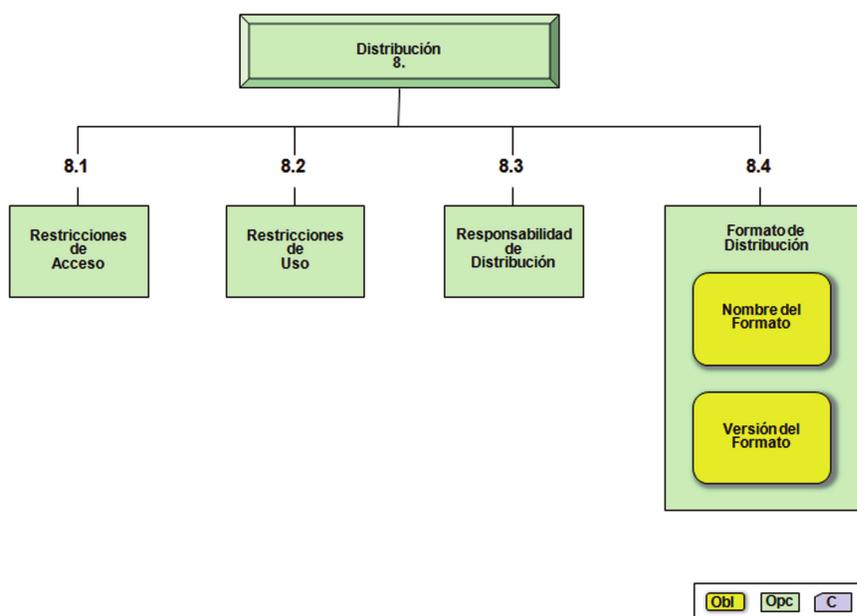
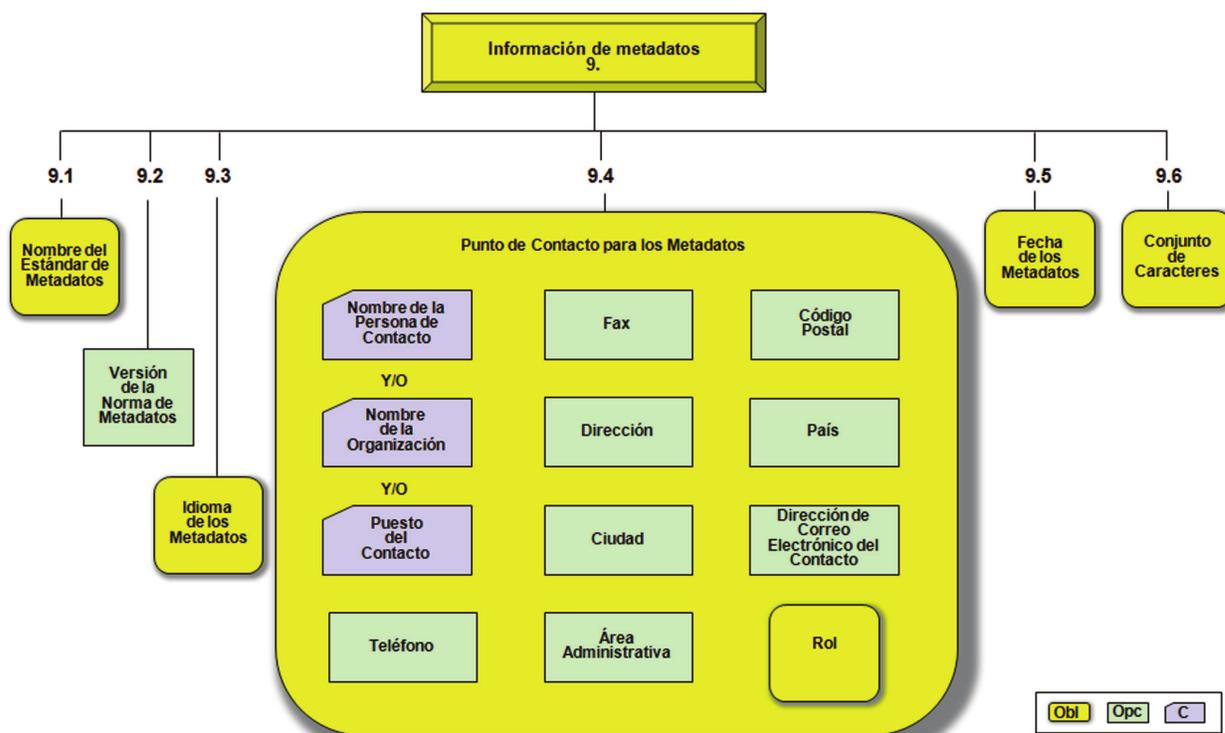


Figura C.7. Elementos de la séptima sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b)



**Figura C.8.** Elementos de la octava sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b)



**Figura C.9.** Elementos de la novena sección de la NTM. Tomado de: Instituto Nacional de Estadística y Geografía (2015b)



## Apéndice D

# Ejemplos de sistemas para capturar y/o publicar metadatos geográficos

Algunas aplicaciones que permiten la captura de metadatos geográficos son:

- **MetaD**<sup>27</sup>. Esta aplicación está destinada a la captura de metadatos geográficos y utiliza el Perfil IDEC (Infraestructura de Dades Espacials de Catalunya) de la norma ISO 19115 así como ISO 19139. Permite importar metadatos con el esquema de la norma del FGDC e ISO 19119 (para metadatos de servicios web geoespaciales). Se descarga a través del portal de la IDE de Cataluña y funciona en el sistema operativo Windows. (Ver figura D.1)

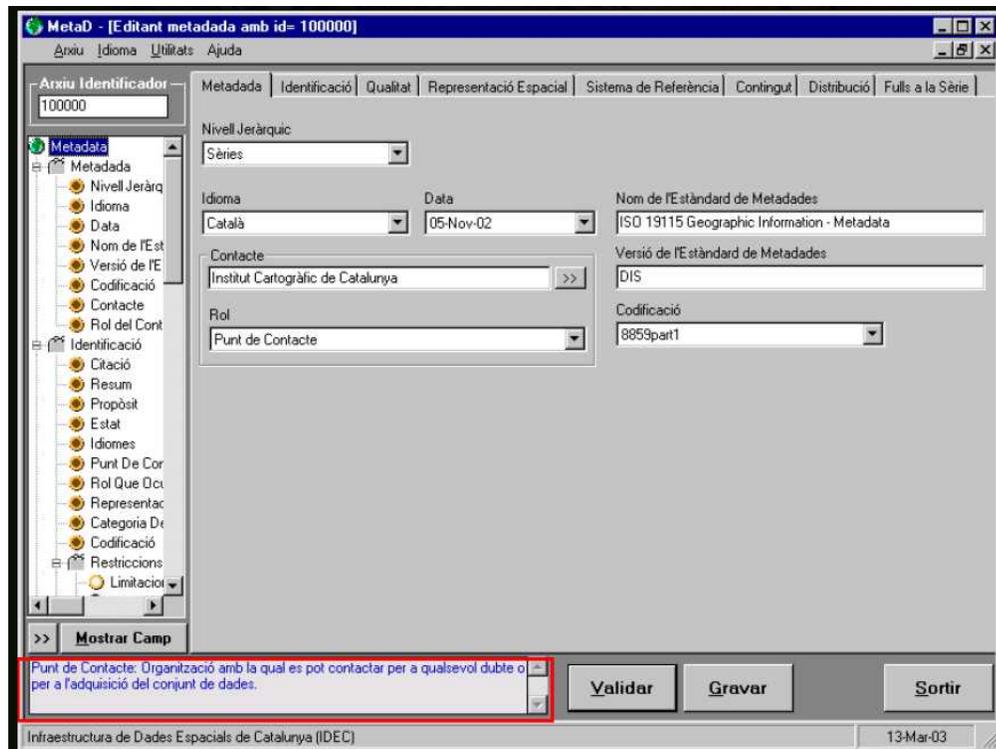
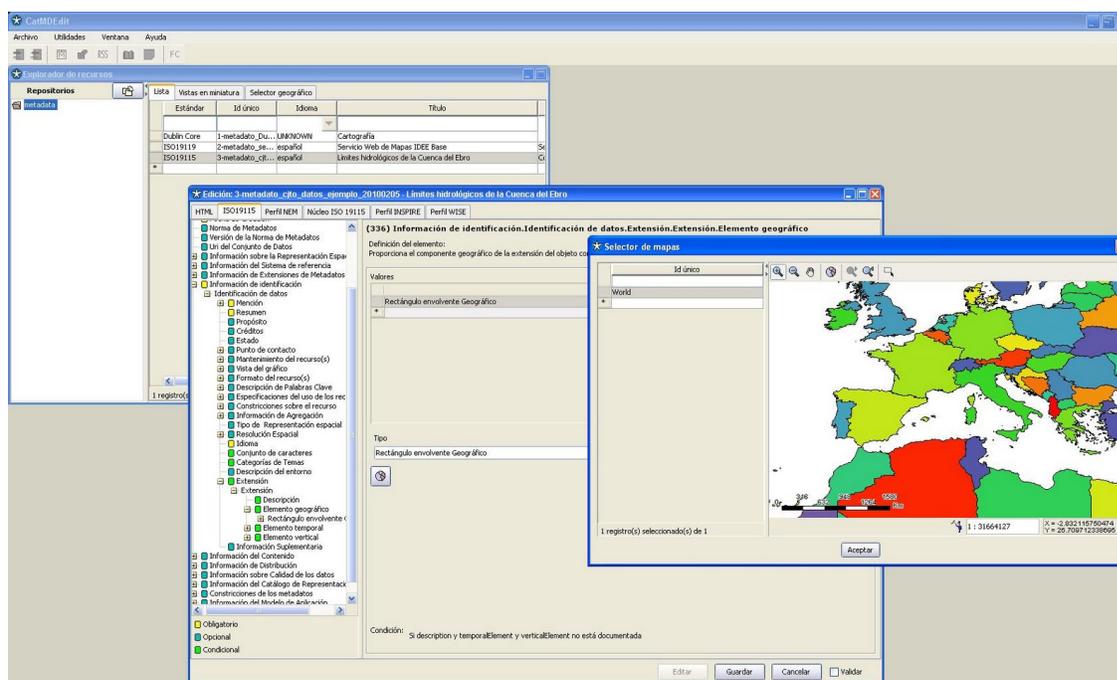


Figura D.1. Software MetaD para la captura de metadatos. Tomada de: Leal Villamil (2011).

<sup>27</sup><http://www.geoportal-idec.net/geoportal/cas/participar/publicar-metad/metad/index.jsp>

- CatMDEdit**<sup>28</sup>. Es un aplicación desarrollada por iniciativa del Instituto Geográfico Nacional de España (IGN). Permite la captura de metadatos para diferentes tipos de recursos, aunque está dedicada principalmente a productos geográficos. Uno de sus principales beneficios es el de trabajar con varias normas de metadatos, como: ISO 19115:2003, ISO 19139, perfil del Núcleo Español de Metadatos (NEM) basado en el núcleo de ISO 19115, ISO 19119, ISO 15836 (Dublin Core), CSDGM, Perfil de metadatos de la directiva INSPIRE y MARC 21 (ISO 2709) para datos bibliográficos (Ver figura D.2). Uno de sus puntos fuertes, además de trabajar con un gran número de formatos para metadatos es que está desarrollado en lenguaje java, lo que le permite ser una herramienta multiplataforma.

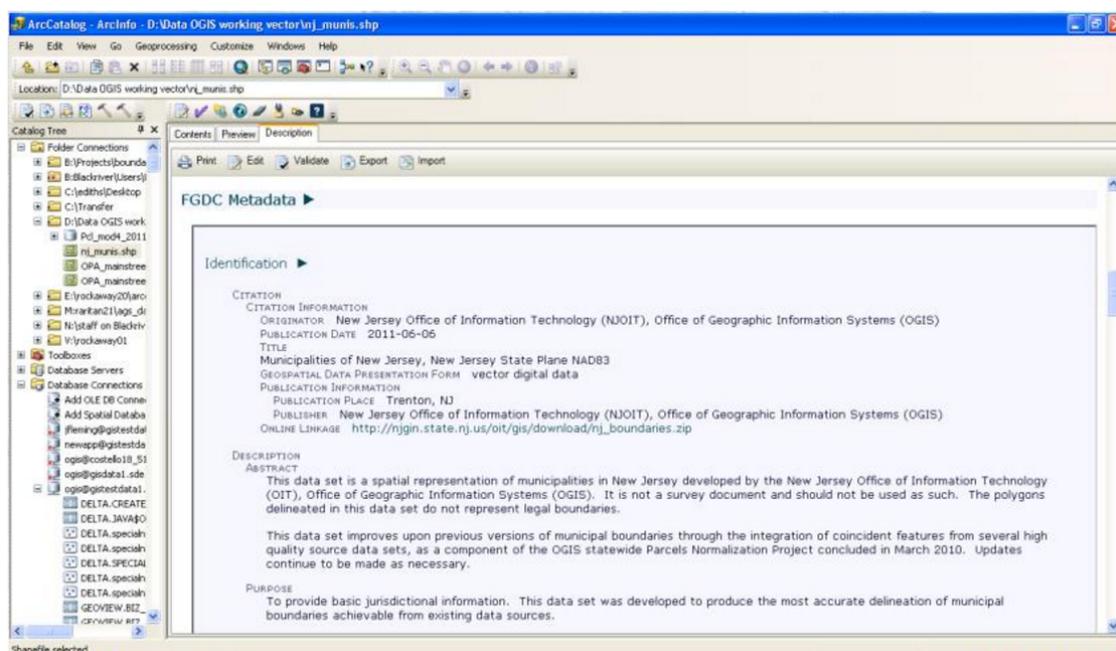


**Figura D.2.** Software CatMDEdit para la captura de metadatos. Imagen obtenida de: Grupo de Sistemas de Información Avanzados de la Universidad de Zaragoza (IAAA) y GeoSpatiumLab (GSL) (2010).

- ArcGIS**<sup>29</sup> Esta es un sistema desarrollado por la empresa Environmental Systems Research Institute (ESRI). Se conforma de una gran número de aplicaciones que permiten la creación y administración de datos geográficos en un gran número de formatos, así como bases de datos geoespaciales. Además de contar con una versión de escritorio, también cuenta con una versión en línea, con la que es posible crear portales para consultar y/o compartir datos espaciales en línea. En lo que respecta a metadatos geográficos, una de las aplicaciones desarrolladas para este sistema es *ArcCatalog*, a través de la cual pueden consultarse y editarse los metadatos de las capas de datos geográficos que se crean y administran con este sistema. Algunas de las norma de metadatos geográficos que se utilizan es esta aplicación son: FGDC CSDGM, el perfil de la directiva INSPIRE, ISO 19139 y el NAP basado en la ISO 19115. (Ver tabla 1.7 y figura D.3)

<sup>28</sup><http://catmmedit.sourceforge.net/index.html>

<sup>29</sup><https://blogs.esri.com/esri/arcgis/2011/01/03/a-new-approach-for-metadata-with-arcgis-10/>



**Figura D.3.** Ejemplo de un metadato consultado a través de la aplicación ArcCatalog. Tomada de: Bocchino y Konopka (2011).

Existen también herramientas que además de permitir la captura de metadatos, también permiten a los usuarios ir creando sus propios catálogos de metadatos o conectarse a otros catálogos, incluso de servidores CSW. Ejemplos de estas aplicaciones son:

- **Geonetwork**<sup>30</sup>. Esta es una aplicación de software libre además de ser una de las más utilizadas para la implementación de servicios de catálogo. Tiene un editor de metadatos que brinda soporte para los estándares ISO 19115, 19119 y Dublin Core. (Ver figura D.4)
- **MapInfo Manager**<sup>31 32</sup>. Este es un producto desarrollado por la empresa Pitney Bowes y una de sus principales aplicaciones es la brindar el servicio CSW. El catálogo puede ser accedido por otra de las aplicaciones desarrolladas por este grupo, MapInfo Pro. Este servicio se apega principalmente al estándar ISO 19115. (Ver figura D.5)
- **pycsw**<sup>33</sup>. Esta aplicación también es un servicio CSW desarrollado en lenguaje Python y es multiplataforma. A pesar de no tener una interfaz como Geonetwork o MapInfo Manager se puede acceder a este servidor a través de aplicaciones “cliente” como QGIS o CatMDEdit. Implementa los estándares Dublin Core, ISO 19115, ISO 19139, ISO 19119, FGDC CSDGM, etc. Además de dar soporte a operaciones de consulta como GetRecords o GetCapabilities, también da soporte para operaciones de CSW Transaccional, por lo permite la inserción, borrado y edición de metadatos dentro del servidor. (Ver figura D.6)

<sup>30</sup><http://geonetwork-opensource.org/>

<sup>31</sup><http://www.pitneybowes.com/au/location-intelligence/geographic-information-system/mapinfo-manager.html?tab1>

<sup>32</sup><http://www.pb.com/docs/US/pdf/Microsite/MapInfo-Professional/MapInfo-Manager-Data-Sheet.pdf>

<sup>33</sup><http://pycsw.org/>

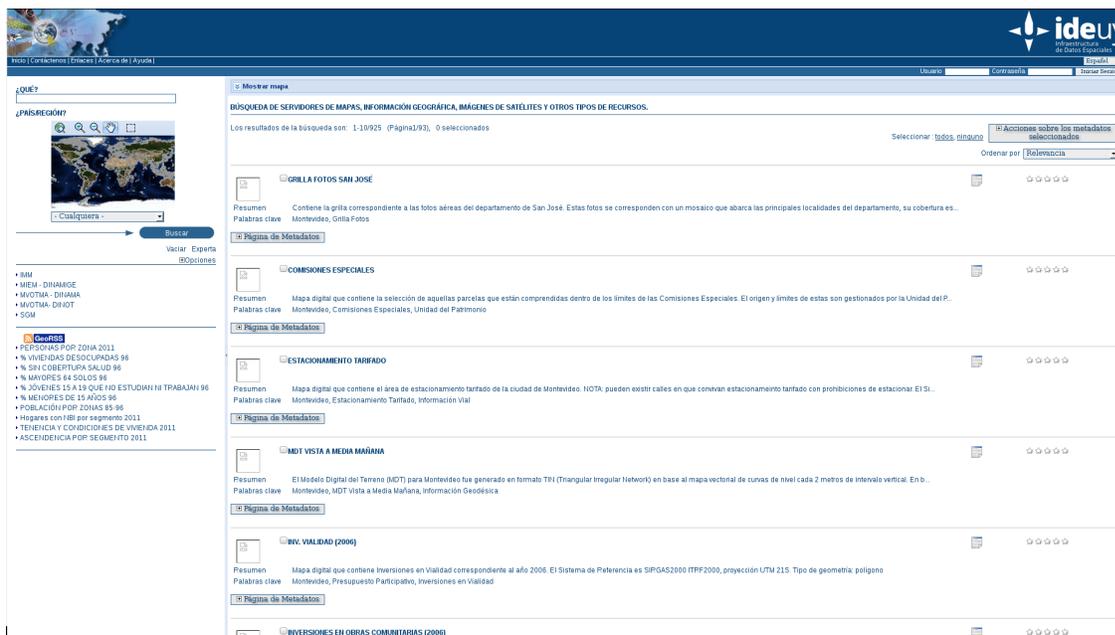


Figura D.4. Ejemplo de un servicio de catálogo implementado con Geonetwork para el geoportel de la IDE de Uruguay.

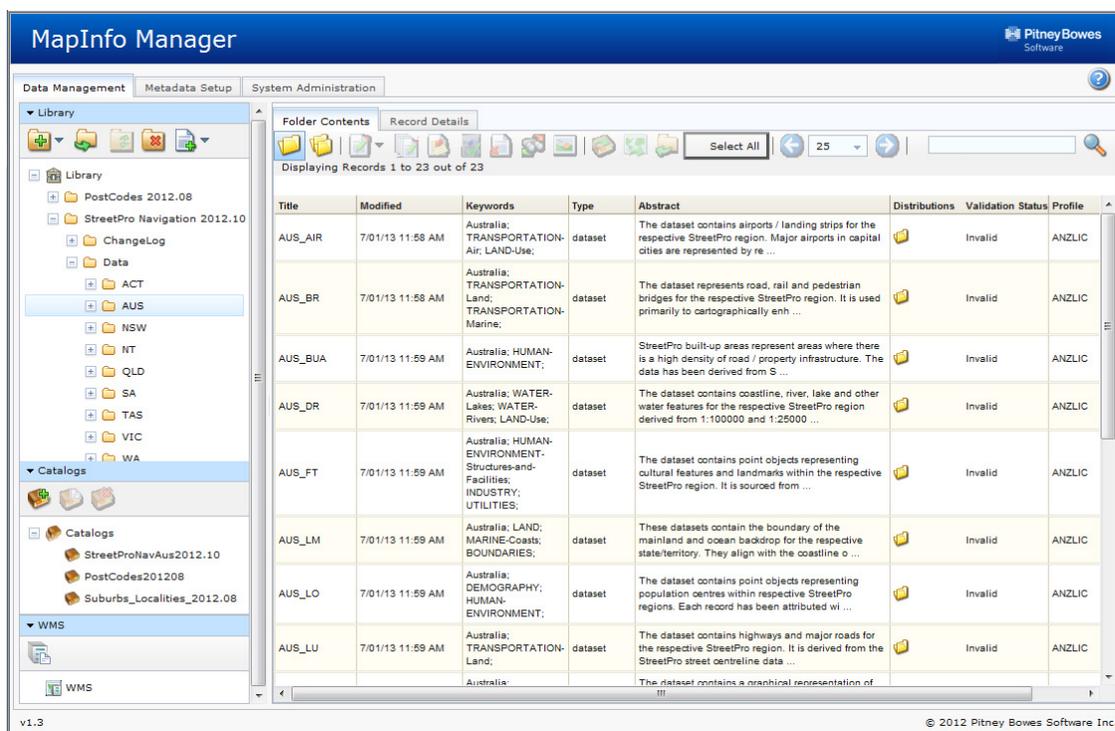
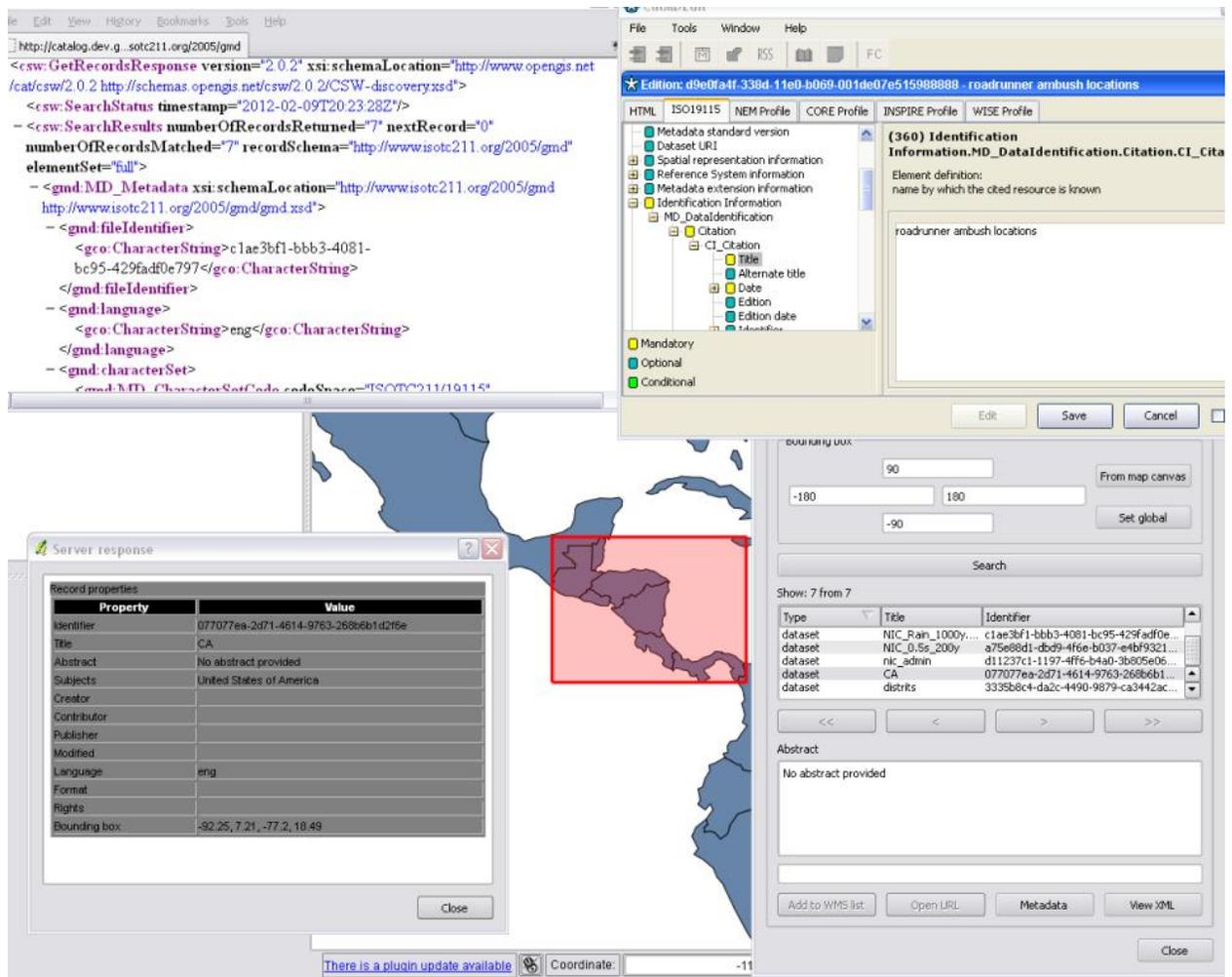


Figura D.5. Ejemplo de un servicio de catálogo implementado con MapInfo Manager. Imagen tomada de: Pitney Bowes (2016).



**Figure D.6.** Imagen que ilustra el funcionamiento del servicio CSW provisto por la aplicación *pycsw*. Imagen tomada de: Open Source Geospatial Foundation (2015).



# Referencias Bibliográficas

- Abad Power, P., Bernabé, M. A., y Rodríguez Pascual, A. F. (2012). Compartir: la solución está en las infraestructuras de datos espaciales. En *Fundamentos de la infraestructuras de datos espaciales*, capítulo 2, páginas 41–52. UPM Press, Madrid. Descargado de: [http://redgeomatica.rediris.es/Libro\\_Fundamento\\_IDE\\_con\\_pastas.pdf](http://redgeomatica.rediris.es/Libro_Fundamento_IDE_con_pastas.pdf) [Último acceso: 10 de Noviembre de 2015].
- ADA Consultores (2015). Diagrama PERT. Disponible en: <http://www.iusc.es/recursos/gesproy/textos/03.03.02.htm> [Último acceso: 14 de Agosto de 2015].
- Allen, S. L. y Terry, E. (2006). *Beginning relational data modeling*. Apress, 2a edición. Descargado de: <https://link-springer-com.pbidi.unam.mx:2443/content/pdf/10.1007%2F978-1-4302-0015-4.pdf> [Último acceso: 12 de Octubre de 2017].
- Arnold, M. y Osorio, F. (1998). Introducción a los Conceptos Básicos de la Teoría General de Sistemas. *Cinta de Moebio. Revista de Epistemología de Ciencias Sociales*, 3:1–8. Obtenido de <http://www.redalyc.org/pdf/101/10100306.pdf> [Último acceso: 1 Junio 2015].
- Bernstein, J. H. (2009). The Data-Information-Knowledge-Wisdom hierarchy and its antithesis. *North American Symposium on Knowledge Organization*, 2(1):68–75. Descargado de: [http://academicworks.cuny.edu/kb\\_pubs/11/](http://academicworks.cuny.edu/kb_pubs/11/) [último acceso: 2 de Febrero de 2016].
- Bertalanffy, L. v. (1987). *Teoría general de los sistemas, Fundamentos, desarrollos, aplicaciones*. Fondo de Cultura Económica, México. 6a reimp.
- Bocchino, J. y Konopka, E. (2011). Metadata in ArcGIS 10. Presentación publicada en: <http://www.slideserve.com/grace/metadata-in-arcgis-10> [Último acceso: 6 de Octubre de 2017].
- Chias Becerril, L. y Reséndiz López, H. D. (2015). Características de los datos geográficos. Documento presentado durante la 2a edición del Diplomado en Sistemas de Información Geográfica con Software Libre en Instituto de Geografía, UNAM.
- De León, P. (2008). Hacia un concepto de transparencia: orígenes e importancia. *Sitio digital CABI (Central American Business Intelligence)*, enero de. Disponible en: <http://cabi.com/blackbox/wp-content/uploads/downloads/2012/08/Transparencia1.pdf> [Último acceso: 4 de Octubre de 2017].
- Decker, D. (2001). *GIS data sources*. John Wiley, Nueva York. Consultado a través de: <http://librunam.dgbiblio.unam.mx:8991> [Último acceso: 3 de Febrero de 2016].

- Di, L. y Deng, M. (2014). ISO TC211 standards on Provenance for Earth science. *AGU Fall Meeting Abstracts*, página B5. Disponible en: <http://adsabs.harvard.edu/abs/2014AGUFMIN34B..05D> [Último acceso: 28 de Octubre de 2015].
- Díaz, A., Arias de Reyna, M., y Sanz, J. (2014). Servidores — Panorama del SIG Libre. Disponible en: <http://panorama-sig-libre.readthedocs.org/es/latest/servidores/> [Último acceso: 18 de Agosto de 2016].
- Douglas, J. (2008). 15 years of the world wide web. MSN Tech and Gadgets. En: <http://tech.uk.msn.com/features/article.aspx?cp-documentid=8197321> [Consultado en: 30 de Junio de 2009].
- Drucker, P. F. (1972). *Tecnología, administración y sociedad*. Fondo de Cultura Económica, México.
- Dublin Core Metadata Initiative (2010). DCMI: Dublin Core Metadata Initiative Resources. Disponible en: <http://dublincore.org/resources/> [Último acceso: 5 de Octubre de 2015].
- Erba, D. A., Duarte, A. N., y Stiefel, M. L. (2012). Interoperabilidad y estandarización de la información geográfica. En *Fundamentos de la infraestructuras de datos espaciales*, capítulo 17, páginas 229–237. UPM Press, Madrid. Descargado de: [http://redgeomática.rediris.es/Libro\\_Fundamento\\_IDE\\_con\\_pastas.pdf](http://redgeomática.rediris.es/Libro_Fundamento_IDE_con_pastas.pdf) [Último acceso: 22 de Septiembre de 2015].
- Espinosa Prendes, M. P. y Amorós Poveda, L. (2001). Accesibilidad en aplicaciones informáticas. Disponible en: <http://tecnologiaedu.us.es/nweb/htm/pdf/paz10.pdf> [Último acceso: 6 de Febrero de 2016].
- Fayad, M. y Schmidt, D. C. (1997). Object-oriented application frameworks. *Communications of the ACM*, Vol. 40(Núm. 10):pp. 32–38. Disponible en: [https://www.researchgate.net/publication/215617675\\_Object-Oriented\\_Application\\_Frameworks](https://www.researchgate.net/publication/215617675_Object-Oriented_Application_Frameworks) [Último acceso: 1 de Febrero de 2017].
- Federal Geographic Data Committee (1998). Content Standard for Digital Geospatial Metadata. Descargado de: [https://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/base-metadata/v2\\_0698.pdf](https://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/base-metadata/v2_0698.pdf) [Último acceso: 26 de Octubre de 2015].
- Federal Geographic Data Committee (2000). Content Standard for Digital Geospatial Metadata Workbook. Descargado de: [https://www.fgdc.gov/metadata/documents/workbook\\_0501\\_bmk.pdf](https://www.fgdc.gov/metadata/documents/workbook_0501_bmk.pdf) [Último acceso: 12 de Octubre de 2017].
- Federal Geographic Data Committee (2014). ISO Metadata Editor Review — Federal Geographic Data Committee. Disponible en: <https://www.fgdc.gov/metadata/iso-metadata-editor-review> [Último acceso: 18 de Agosto de 2016].
- Ferrer-Sapena, A., Peset, F., y Aleixandre-Benavent, R. (2011). Acceso a los datos públicos y su reutilización: open data y open government. *El profesional de la información*, Vol. 20(Núm. 3):pp. 260–269. Disponible en: <http://eprints.rclis.org/20997/2/92741636q145x727.pdf> [Último acceso: 4 de Octubre de 2017].
- Free-Logistics (2015). PERT Chart Diagram and Method | Free-Logistics. Disponible en: <http://www.free-logistics.com/pert-chart-diagram-and-method/> [Último acceso: 14 de Agosto de 2015].

- Fu, P. y Sun, J. (2010). *Web GIS: principles and applications*. Esri Press, California.
- García, J. (2013). Diseño data warehouse: hechos y dimensiones; modelo estrella VS copo de nieve. Disponible en: <http://mundodb.es/disenio-data-warehouse-hechos-y-dimensiones-modelo-estrella-vs-copo-de-nieve> [Último acceso: 20 de Octubre de 2016].
- Garlan, D. y Shaw, M. (1993). An introduction to software architecture. *Advances in Software Engineering and Knowledge Engineering*, 1(Enero):1-39. Disponible en: <http://repository.cmu.edu/cgi/viewcontent.cgi?article=1720&context=compsci> [Último acceso: 4 de Noviembre de 2016].
- GDAL (2015). GDAL: GDAL - Geospatial Data Abstraction Library. Disponible en: <http://www.gdal.org/> [Último acceso: 9 de Septiembre de 2015].
- Gobierno de la República (2012). Parámetros de estimación de vida útil. *Diario Oficial de la Federación*. (15 de Agosto). Disponible en: [http://dof.gob.mx/nota\\_detalle.php?codigo=5264340&fecha=15/08/2012](http://dof.gob.mx/nota_detalle.php?codigo=5264340&fecha=15/08/2012) [Último acceso: 23 de Enero de 2017].
- Gobierno de la República (2013a). Estrategia Digital Nacional. Disponible en: <http://cdn.mexicodigital.gob.mx/EstrategiaDigital.pdf> [Último acceso: 31 de mayo de 2016].
- Gobierno de la República (2013b). Programa para un Gobierno Cercano y Moderno 2013-2018. *Diario Oficial de la Federación*. (30 de Agosto) Disponible en: [http://www.dof.gob.mx/nota\\_detalle.php?codigo=5312420&fecha=30/08/2013](http://www.dof.gob.mx/nota_detalle.php?codigo=5312420&fecha=30/08/2013) [Último acceso: 31 de Mayo de 2016].
- Gomasasca, M. a. (2009). *Basics of geomatics*. Springer Netherlands. Descargado a través de: [http://librunam.dgbiblio.unam.mx:8991/F/?func=find-b&find\\_code=WRD&request=gomasasca&local\\_base=MX001](http://librunam.dgbiblio.unam.mx:8991/F/?func=find-b&find_code=WRD&request=gomasasca&local_base=MX001) [Último acceso: 31 de Agosto de 2015].
- Grupo de Sistemas de Información Avanzados de la Universidad de Zaragoza (IAAA) y GeoSpatialiumLab (GSL) (2010). BlogIDEE: Nueva versión de CatMDEdit. Disponible en: <http://blog-idee.blogspot.mx/2010/08/nueva-version-de-catmdedit.html> [Último acceso: 12 de Febrero de 2016].
- Harley, J. B. (2002). *The new nature of maps: Essays in the history of cartography*. Johns Hopkins University Press, Baltimore.
- Haskins, C., Forsberg, K., Krueger, M., Walden, D., y Hamelin, R. D. (2010). A guide for system life cycle processes and activities (No. INCOSE-TP-2003-002-03.2). *INCOSE Systems Engineering Handbook*, 3.2.
- Hernández Cruz, A. (2015). El derecho a la información en el marco de la Ley General de Transparencia. *SDPnoticias.com*, 1 Abril 2015. Disponible en: <http://www.sdpnoticias.com/nacional/2015/04/01/el-derecho-a-la-informacion-en-el-marco-de-la-ley-general-de-transparencia> [Último acceso: 23 Diciembre 2015].
- IBM Corp. (2015). IBM Knowledge Center - Advantages of Java. Disponible en: [https://www.ibm.com/support/knowledgecenter/ssw\\_aix\\_71/com.ibm.aix.performance/advantages\\_java.htm](https://www.ibm.com/support/knowledgecenter/ssw_aix_71/com.ibm.aix.performance/advantages_java.htm) [Último acceso: 22 de Febrero de 2017].
- INSPIRE (2007). About INSPIRE. Disponible en: <http://inspire.ec.europa.eu/inspire-principles/9> [Último acceso: 13 de Octubre de 2017].

- Instituto Nacional de Estadística y Geografía (2008). Aplicaciones de conversión y captura conforme a la Norma Técnica para la elaboración de Metadatos Geográficos (NTM). Disponible en: [http://www.inegi.org.mx/geo/contenidos/metadatos/aplicaciones\\_conversion.aspx](http://www.inegi.org.mx/geo/contenidos/metadatos/aplicaciones_conversion.aspx) [Último acceso: 24 Enero 2016].
- Instituto Nacional de Estadística y Geografía (2008). Ley del Sistema Nacional de Información Estadística y Geográfica. En Diario Oficial de la Federación (16 de Abril). Disponible en: [http://www.diputados.gob.mx/LeyesBiblio/ref/lsnieg/LSNIEG\\_orig\\_16abr08.pdf](http://www.diputados.gob.mx/LeyesBiblio/ref/lsnieg/LSNIEG_orig_16abr08.pdf) [Último acceso: 24 de Noviembre de 2015].
- Instituto Nacional de Estadística y Geografía (2010). Norma técnica para la elaboración de metadatos geográficos. *Diario Oficial de la Federación*. (24 de Diciembre). Descargada de: [http://www.inegi.org.mx/geo/contenidos/normastecnicas/doc/norma\\_tecnica\\_sobre\\_elaboracion\\_de\\_metadatos\\_geograficos.pdf](http://www.inegi.org.mx/geo/contenidos/normastecnicas/doc/norma_tecnica_sobre_elaboracion_de_metadatos_geograficos.pdf) [Último acceso: 6 de Diciembre de 2015].
- Instituto Nacional de Estadística y Geografía (2014). Norma técnica para el acceso y publicación de Datos Abiertos de la información estadística y geográfica de interés nacional. *Diario Oficial de la Federación*. (4 de Diciembre). Descargada de: [http://www.dof.gob.mx/nota\\_detalle.php?codigo=5374183&fecha=04/12/2014](http://www.dof.gob.mx/nota_detalle.php?codigo=5374183&fecha=04/12/2014) [Último acceso: 14 de Enero de 2016].
- Instituto Nacional de Estadística y Geografía (2015a). El SNIEG como la infraestructura de datos espaciales de México. Folleto distribuido en la edición 2015 del Latin American Geospatial Forum, Ciudad de México.
- Instituto Nacional de Estadística y Geografía (2015b). Guía Metodológica para la Generación e Integración de Metadatos Geográficos conforme a la Norma Técnica para la elaboración de Metadatos Geográficos (NTM). Disponible en: [http://internet.contenidos.inegi.org.mx/contenidos/productos/prod\\_serv/contenidos/espanol/bvinegi/productos/nueva\\_estruc/702825076962.pdf](http://internet.contenidos.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825076962.pdf) [Último acceso: 6 de Diciembre de 2015].
- Jacobson, I., Booch, G., y Rumbaugh, J. (2000). *El proceso unificado de desarrollo de software*. Addison-Wesley, Madrid.
- Johansen Bertoglio, O. (1982). *Introducción a la teoría general de los sistemas*. Limusa, México.
- Krzanowski, R. y Raper, J. (2001). *Spatial evolutionary modeling*. Oxford University Press, Oxford.
- Leal Villamil, J. (2011). INGENIERIA FORESTAL: Software MetaD. Disponible en: <http://ingenieriaforestalut.blogspot.mx/2011/06/software-metad.html> [Último acceso: 12 de Febrero de 2016].
- Litwin, L. y Rossa, M. (2011). *Geoinformation Metadata in INSPIRE and SDI: Understanding. Editing. Publishing*. Springer Science & Business Media. Descargado a través de: <https://link.springer-com.pbidi.unam.mx:2443/book/10.1007%2F978-3-642-15862-9> [Último acceso: 28 de Octubre de 2015].
- López del Consuelo, M. (2015). Transparencia, datos abiertos y gobierno abierto. Disponible en: <http://es.slideshare.net/MnicaLpez12/transparencia-datos-abiertos-y-gobierno-abierto> [Último acceso: 23 de Diciembre de 2015].

- Maguire, D. J. y Longley, P. A. (2005). The emergence of geoportals and their role in spatial data infrastructures. *Computers, environment and urban systems*, Vol. 29(Núm. 1):3–14. Descargado a través de: <http://www.sciencedirect.com/science/article/pii/S0198971504000456> [Último acceso: 17 de Noviembre de 2015].
- Mas, S., Rodríguez, A. F., y Bernabé, M. A. (2012). Compartir datos geográficos. En *Fundamentos de la infraestructuras de datos espaciales*, capítulo 1, páginas 31–40. UPM Press, Madrid. Descargado de: [http://redgeomatica.rediris.es/Libro\\_Fundamento\\_IDE\\_con\\_pastas.pdf](http://redgeomatica.rediris.es/Libro_Fundamento_IDE_con_pastas.pdf) [Último acceso: 10 de Noviembre de 2015].
- Master magazine (2016). Definición de DBMS - Significado y definición de DBMS. Disponible en: <http://www.mastermagazine.info/termino/4544.php> [Último acceso: 24 de Octubre de 2016].
- Mohan, S. (2016). All You Need to Know About Java – Advantages and Disadvantages. Disponible en: <https://www.weblinesindia.com/blog/all-you-need-to-know-about-java-advantages-and-disadvantages/> [Último acceso: 22 de Febrero de 2017].
- Moreno Bonett, A. (2010). *Apuntes del curso teoría general de sistemas. Capítulo 1. Teoría general de Sistemas*. Universidad Nacional Autónoma de México, Facultad de Ingeniería, 2a edición.
- Moser, H. A. (2014). *Systems engineering, systems thinking, and learning : a case study in space industry*. Springer, New York. Descargado a través de: <https://link-springer-com.pbidi.unam.mx:2443/book/10.1007%2F978-3-319-03895-7> [Último acceso: 7 de Agosto de 2015].
- Naessens, H. (2010). Ética pública y transparencia. En *XIV Encuentro de Latinoamericanistas Españoles: congreso internacional*, páginas 2113–2130. Universidade de Santiago de Compostela, Centro Interdisciplinario de Estudios Americanistas Gumersindo Busto; Consejo Español de Estudios Iberoamericanos. Disponible es: <https://halshs.archives-ouvertes.fr/halshs-00531532/document> [Último acceso: 4 de Octubre de 2017].
- Nogueras-Iso, J. y Muro-Medrano, Pedro R. Zaragoza-Soria, F. J. (2005). *Geographic information metadata for spatial data infrastructures : resources, interoperability, and information retrieval*. Springer, Verlag, Berlin. Descargado a través de: <https://link-springer-com.pbidi.unam.mx:2443/book/10.1007%2F3-540-27508-8> [Último acceso: 23 de Octubre de 2015].
- O'Brien, T., Casey, J., Fox, B., Snyder, B., Zyl, J. V., y Redmond, E. (2008). *Maven: The Definitive Guide*. Sonatype Inc.
- Office of the President (1994). Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure. Descargado de : <http://www.archives.gov/federal-register/executive-orders/pdf/12906.pdf> Último acceso: 9 de Noviembre de 2015.
- Olvera Ramírez, J. (2014). Taller de Metadatos geográficos. Documento presentado durante el Taller de metadatos geográficos en el Instituto de Geografía de la UNAM.
- Olvera Ramírez, J., Sara Gutiérrez, C. A., Mancera Cedillo, M., Reséndiz López, H. D., y Chias Becerril, L. (2014). *Infraestructuras de datos espaciales y normatividad geográfica en México : una perspectiva actual*. UNAM, Instituto de Geografía, México, D.F.

- Open Source Geospatial Foundation (2015). pycsw — OSGeo-Live 9.0 Documentation. Disponible en: [http://live.osgeo.org/en/overview/pycsw\\_overview.html](http://live.osgeo.org/en/overview/pycsw_overview.html) [Último acceso: 12 de Febrero de 2016].
- Oracle Corporation (2011). Servlet (Java EE 6). Disponible en: <http://docs.oracle.com/javaee/6/api/javax/servlet/Servlet.html> [Último acceso: 16 de Diciembre de 2016].
- Oracle Corporation (2012). Your First Cup: An Introduction to the Java EE Platform. Disponible en: <http://docs.oracle.com/javaee/6/firstcup/doc/firstcup.pdf> [Último acceso: 16 de Diciembre de 2016].
- Peña Ayala, A. (2006). *Ingeniería de software: una guía para crear sistemas de información*. Instituto Politécnico Nacional, México. Descargado de: [https://www.google.com.mx/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwiGy5uS\\_-7WAhUBQyYKHa1KBbUQFgguMAA&url=http%3A%2F%2Feducagratis.c1%2Fmoodle%2Fmod%2Fresource%2Fview.php%3Fid%3D4694&usg=AOvVaw2sv3Vn6hx4MD0ve-qhhiUu](https://www.google.com.mx/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwiGy5uS_-7WAhUBQyYKHa1KBbUQFgguMAA&url=http%3A%2F%2Feducagratis.c1%2Fmoodle%2Fmod%2Fresource%2Fview.php%3Fid%3D4694&usg=AOvVaw2sv3Vn6hx4MD0ve-qhhiUu) [Último acceso: 8 de Julio de 2016].
- Pickton, D. W. y Wright, S. (1998). What's SWOT in strategic analysis? *Strategic Change*, 7(April):101–109. Descargado de: [http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1099-1697\(199803/04\)7:2%3C101::AID-JSC332%3E3.0.CO;2-6/pdf](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1099-1697(199803/04)7:2%3C101::AID-JSC332%3E3.0.CO;2-6/pdf) [Último acceso: 5 de julio de 2016].
- Picón, P. (2015). Análisis DAFO (FODA). Disponible en: <https://ingenierosquimicos.wordpress.com/2015/06/11/analisis-dafo-foda/> [Último acceso: 6 de julio 2016].
- Pitney Bowes (2016). MapInfo Manager. Disponible en: <http://www.pitneybowes.com/hk/location-intelligence/geographic-information-system/mapinfo-manager.html> [Último acceso: 12 de Febrero de 2016].
- Pressman, R. S. (2010). *Ingeniería del software. Un enfoque práctico*. McGraw-Hill, 7a edición.
- Project Management Institute (2008). *Guía de los Fundamentos para la Dirección de Proyectos (Guía del PMBOK)*. Project Management Institute, Inc., Pennsylvania, 4a edición.
- Ramírez, L. A. (2002). Teoría de Sistemas. Disponible en: <https://es.slideshare.net/magnoraymundodiaz/teora-de-sistemas-16817870> [Último acceso: 7 de Agosto de 2015].
- Ramírez González, J. A. (2013). Ingresos de un desarrollador Java Jr. Disponible en: <http://www.recluit.com/ingresos-de-un-desarrollador-java-jr/> [Último acceso: 23 de enero de 2017].
- Real Academia Española (2014). DLE: sistema - Diccionario de la lengua española - Edición del Tricentenario. Disponible en: <http://dle.rae.es/?id=Y2AFX5s> [Fecha de consulta: 29 Enero 2016].
- Rodríguez Sanz, M. (2010). *Análisis y diseño de un data mart para el seguimiento académico de alumnos en un entorno universitario*. Tesis, Universidad Carlos III de Madrid. Disponible en: [https://e-archivo.uc3m.es/bitstream/handle/10016/9856/PFC\\_Miguel\\_Rodriguez\\_Sanz.pdf?sequence=6&isAllowed=y](https://e-archivo.uc3m.es/bitstream/handle/10016/9856/PFC_Miguel_Rodriguez_Sanz.pdf?sequence=6&isAllowed=y) [Último acceso: 20 de Octubre de 2010].
- Rojas Enriquez, H. (2016). ESTRATEGIAS FA FO DA DO - Clase V. Disponible en : [clasev.net/v2/mod/resource/view.php?id=21480](http://clasev.net/v2/mod/resource/view.php?id=21480) [Último acceso: 1 de junio de 2016].

- Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of information science*, Vol. 33(Núm. 2):163–180. Descargado de: <http://inls151f14.web.unc.edu/files/2014/08/rowleydikw.pdf> [Último acceso: 1 de Septiembre de 2015].
- Rumbaugh, J., Jacobson, I., y Booch, G. (2000). *El lenguaje unificado de modelado. Manual de referencia*. Addison-Wesley, Madrid.
- Schwalbe, K. (2010). *Information Technology Project Management, Revised*. Course Technology Press, Boston, MA, United States, 6a edición.
- Secretaría de Gobernación (2006). DECRETO por el que se declaran reformados los artículos 26 y 73 fracción XXIX-D de la Constitución Política de los Estados Unidos Mexicanos. En Diario Oficial de la Federación (7 de Abril). Disponible en : [http://www.dof.gob.mx/nota\\_detalle.php?codigo=2142855&fecha=07/04/2006](http://www.dof.gob.mx/nota_detalle.php?codigo=2142855&fecha=07/04/2006) [Último acceso: 4 de Octubre de 2017].
- Secretaría de Gobernación (2015). Ley General de Transparencia y Acceso a la Información Pública. *Diario Oficial de la Federación*. (4 de Mayo) Disponible en: [http://www.dof.gob.mx/nota\\_detalle.php?codigo=5391143&fecha=04/05/2015](http://www.dof.gob.mx/nota_detalle.php?codigo=5391143&fecha=04/05/2015) [Último acceso: 12 de Enero de 2016].
- SG Software Guru (2017). Estudio de Salarios SG 2016 | SG. Disponible en: <https://sg.com.mx/revista/53/estudio-salarios-sg-2016#.WIY1fLicG9t> [Último acceso: 19 de enero de 2017].
- Sherman, G. E. (2008). *Desktop GIS : mapping the planet with open source tools*. Pragmatic Bookshelf, Raleigh, North Carolina.
- Tomlinson, R. F. (2003). *Thinking about GIS : geographic information system planning for managers*. Esri Press, Redlands, California.
- Vergara Hernández, S. (2006). Metadatos y Clearinghouse. *Mati. Sobre la letra digital* [www.mati.unam.mx](http://www.mati.unam.mx). Disponible en: [http://www.mati.unam.mx/index.php?option=com\\_content&task=view&id=206&Itemid=51](http://www.mati.unam.mx/index.php?option=com_content&task=view&id=206&Itemid=51) [Último acceso: 23 de Enero de 2016].
- W3schools (2016). MySQL What is DDL, DML and DCL? Disponible en: <http://www.w3schools.in/mysql/ddl-dml-dcl/> [Último acceso: 24 de Octubre de 2016].