

MODELO BASADO EN ARQUITECTURA OPEN Y LINKED DATA PARA EL
SECTOR TURÍSTICO DE NORTE DE SANTANDER

MATIAS HERRERA CACERES

UNIVERSIDAD AUTÓNOMA DE BUCARAMANGA
FACULTAD DE INGENIERÍA DE SISTEMAS
GRUPO DE INVESTIGACIÓN PRESERVACIÓN E INTERCAMBIO DIGITAL DE
INFORMACIÓN Y CONOCIMIENTO PRISMA
MAESTRÍA EN SOFTWARE LIBRE
BUCARAMANGA
2015

MODELO BASADO EN ARQUITECTURA OPEN Y LINKED DATA PARA EL
SECTOR TURÍSTICO DE NORTE DE SANTANDER

MATIAS HERRERA CACERES

Trabajo de grado presentado como requisito para optar el título de:
Magíster en Software Libre

Director
Eduardo Carrillo Zambrano
PhD.

UNIVERSIDAD AUTÓNOMA DE BUCARAMANGA
FACULTAD DE INGENIERÍA DE SISTEMAS
GRUPO DE INVESTIGACIÓN PRESERVACIÓN E INTERCAMBIO DIGITAL DE
INFORMACIÓN Y CONOCIMIENTO PRISMAMAESTRÍA EN SOFTWARE LIBRE
BUCARAMANGA
2015

Nota de aceptación:

Presidente del Jurado

Jurado

Jurado

Bucaramanga, 25 de junio de 2015

A DIOS, quien día a día me da sabiduría e ilumina el camino para seguir adelante y brindar soluciones para el bienestar de la sociedad, a mi esposa y mis hijos que son la fuente de mi inspiración, a mis padres que son mis mejores maestros, a amigos y profesores por toda la ayuda. Gracias a todos.

AGRADECIMIENTOS

A Dios

Por iluminarme, por darme salud y permitir que se llevara a cabo el proyecto exitosamente.

A mis Hijos

Matías y Santiago porque son el motor y la energía que permite que todas las metas sean posibles. Porque con Uds. todo tiene más sentido.

A mi Esposa

Deysi por soportar que tome parte de su tiempo para el proyecto y por aguantar traspasos a mi lado.

A mis Padres

Jeremías y Nohemí por sus consejos y buenos deseos.

Al Director

PhD. Eduardo Carrillo por todo su apoyo, por sus sabios consejos y por compartir sus conocimientos.

CONTENIDO

	Pág.
RESUMEN	18
INTRODUCCIÓN	21
1. MARCO TEORICO	23
1.1 ANTECEDENTES	23
1.1.1 Procolombia	27
1.2 POLITICA DE MERCADEO Y PROMOCION TURISTICA “COLOMBIA DESTINO TURISTICO DE CLASE MUNDIAL”	25
1.2.1 Matriz de lineamientos – política de mercadeo y promoción turística	29
1.2.2 Actores Involucrados	32
1.2.2 Marcas País en las Américas	32
1.2.3 Marca País Mundial	34
1.2.4 Visión Colombia 2019	35
1.2.5 El Conpes 3397	35
1.3 TECNOLOGIAS PARA EL TURISMO	35
1.3.1 Open Data Y Turismo	36
1.3.2 Open Linked Data y Turismo	38
1.3.3 Ontologías Para Turismo	40
1.3.4 Aplicaciones Móviles en turismo	41
1.4 HERRAMIENTAS DE SOFTWARE LIBRE PARA EL DESARROLLO	44
1.4.1 Herramienta para la implementación de Open Data (CKAN)	44

1.4.2 Herramienta Implementación Linked Data (Virtuoso Universal Server	45
1.4.3 Entorno para el desarrollo de Aplicaciones Móviles	48
1.5 MODELO DE NEGOCIO	50
1.5.1 Innovación Abierta	52
1.5.2 Innovación Abierta y Open Data	53
1.5.3 Ideas en el Ecosistema Open Data	54
1.5.4 Modelo de Negocio Canvas	56
1.6 ESTADO DEL ARTE	58
1.6.1 La Web Semántica	58
1.6.2 Linked Open Data: Ideas y Principios	60
1.6.3 La Nube de Datos Abiertos Enlazados	61
1.6.4 Tecnologías Open Linked Data	62
1.6.5 Escenarios para Linked Open Data	64
1.6.6 Linked Open Data. Guía de Diseño	65
1.6.7 Definición De La Estrategia Uri	67
1.6.8 Linked Open Data: Vocabularios Existentes	70
1.6.9 El ejemplo de la DBpedia	72
1.6.10 Ventajas del Linked Open Data	73
1.7 PLANTEAMIENTO DEL PROBLEMA Y JUSTIFICACION	92
1.8 OBJETIVOS	98
1.8.1 Objetivo General	98
1.8.2 Objetivos Específicos	98
1.9 BENEFICIOS	99

1.10 IMPACTO	102
2. METODO	104
2.1 ACTIVIDADES	104
3. RESULTADOS DE LA INVESTIGACIÓN	107
3.1 IMPLEMENTACION DE LA INICIATIVA LINKED OPEN DATA	107
3.1.1 Análisis Para Transformación A Linked Open Data	107
3.1.2 Panorama Actual De Los Datos En Las Organizaciones	109
3.1.4 Proceso de Integración de LOD a Arquitecturas Empresariales	112
3.1.4.1 Reemplazar el Enterprise Software Bus (ESB) por Protocolos de LOD en Servidores Externos.	113
3.1.4.2 Sustitución del Enterprise Software Bus (ESB) con Protocolos de Linked Open Data entre los Servidores de Publicación LOD de la Empresa	114
3.2 BUENAS PRÁCTICAS PARA LINKED OPEN DATA	115
3.2.1 Identificación de Fuentes de Datos	115
3.2.2 Modelado de Dominios Específicos	116
3.2.3 Migrar Datos Heredados a Vocabularios	117
3.3 PUBLICACIÓN	118
3.3.1 Patrones de Publicación para Datos Relacionales	118
3.3.2 Definición del Lenguaje de Mapeo desde RDB2RDF	119
3.3.3 Materializar los Datos	120
3.3.4 Patrón Publishing para Excel / Datos CSV	120
3.3.5 Patrón Publishing para Datos XML	121
3.3.6 Patrón Publishing para Datos no Estructurados	121

3.4 SERVIDOR Y ALMACENAMIENTO	123
3.5 MAPEO DE VOCABULARIO	125
3.6 CASO DE ESTUDIO: ENFOQUE DE OPEN Y LINKED DATA VINCULADO AL SECTOR TURÍSTICO DE NORTE DE SANTANDER	125
3.6.1 Condiciones del Entorno.	125
3.6.2 Plan de Desarrollo Turístico de Cúcuta	125
3.6.3 Plan Regional de Competitividad	126
3.6.4 Stakeholders	126
3.6.5 El Sector Turístico de Norte de Santander como parte de la Web de Datos	128
3.6.6 Modelo de Negocio de Datos Abiertos del Sector Turismo en Norte de Santander	129
3.6.7 El Proceso de Apertura de Datos	131
3.6.8 Fuentes de Datos Disponibles	134
3.6.9 Inventario de Contenidos	136
3.6.10 Introduciendo Un Modelo Común	138
3.6.11 El proceso de publicación de los datos	141
3.6.12 Factores Clave de Éxito para la Iniciativa LOD Turismo de Norte de Santander	144
3.6.13 Participación y Comunicación	145
3.6.14 Aspectos Legales, LOPD Y Licencias de Datos Abiertos	146
3.6.15 Metadatos Utilizado Para Turismo LOD Norte De Santander	148
3.6.16 Arquitectura, Herramientas Y Requisitos Del Proyecto.	148
3.6.17 Actores Principales del Sistema	151
3.6.18 Casos de Uso	151

3.6.19 Desarrollo e Implementación	156
3.6.20 Instalación y Configuración de Plataformas	157
3.6.21 Desarrollo del prototipo	163
4. DISCUSION	169
4.1 COMPARACION DE LOS RESULTADOS CON TRABAJOS EXISTENTES	169
4.2 APORTES DEL PROYECTO	169
5. CONCLUSIONES	171
6. RECOMENDACIONES Y TRABAJOS FUTURO	173
BIBLIOGRAFIA	174
ANEXOS	178

LISTA DE FIGURAS

	Pág.
Figura 1. Cifra equivalente a la suma de extranjeros no residentes en Colombia más el total de colombianos residentes en el exterior y los visitantes de cruceros 2013 y 2014	26
Figura 2. Cifras de turistas por países	26
Figura 3. Llegada de Extranjeros no residentes por departamento	27
Figura 4. Sectores Involucrados	32
Figura 5. Entes involucrados en marca País	32
Figura 6. Ranking Marca País de las Américas	33
Figura 7. Top de las 25 mejores marca País	34
Figura 8. Página oficial de datos abiertos gobierno de Navarra y aplicación móvil camino de Santiago en Galicia a partir de datos abiertos	36
Figura 9. Aplicación catálogo de datos de turismo Euskadi	37
Figura 10. Portal de Datos abierto de Cataluña	37
Figura 11. Diseño de servicios Open data	38
Figura 12. Modelo de Integración de Tecnología para Gestión Turística	39
Figura 13. Almacén de Datos DataHub	40
Figura 14. Portal de Turismo de la Rioja	40
Figura 15. Ingresos Globales en Aplicaciones	42
Figura 16. Aplicaciones Apps Innovaciones sociales	43
Figura 17. Arquitectura general de Virtuoso Universal Server	46
Figura 18. Arquitectura detallada de Virtuoso Universal Server	47
Figura 19. Las cinco etapas del ciclo de vida de una aplicación Multiplataforma	48

Figura 20. Jboss y Apache Córdoba	50
Figura 21. Modelos de Negocio para Datos	52
Figura 22. Representación del proceso de innovación abierto	53
Figura 23. Potencial del open data	54
Figura 24. Moovit, aplicación crowd sourced para el transporte público	55
Figura 25. Mapa Conceptual Marco Teórico	58
Figura 26. Pila de la Web semántica	60
Figura 27. Expansión de LOD cloud del 2008 al 2014	61
Figura 28. El crecimiento de los conjuntos de datos LOD en el Data Hub.	62
Figura 29. Pila de Tecnologías LOD	63
Figura 30. Mashup de Visualización de datos geo referenciados	64
Figura 31. Búsqueda en google usando semántica de datos enlazados	65
Figura 32. Características a tener en cuenta sobre Linked Data	66
Figura 33. Reglas para la persistencia de URIs	68
Figura 34. Ilustración de propiedades que define el Vocabulario Dublin Core	71
Figura 35. La arquitectura de la DBpedia	73
Figura 36. Cloud Linked Open Data	75
Figura 37. Conversión de Open Data a LOD en data.gov	76
Figura 38. Portal Data.Gov.Uk, estándares abiertos	77
Figura 39. Mapa de iniciativas OGD en España	78
Figura 40. Portal de turismo LOD de la Rioja	79
Figura 41. Portal de Datos Abiertos de Colombia	80
Figura 42. App Castilla de la Mancha	81

Figura 43. App Astroturismo de la Palma	82
Figura 44. Aplicaciones para Transporte	82
Figura 45. Aplicaciones para Alojamiento	83
Figura 46. Aplicaciones para Ocio y Restauración	83
Figura 47. Aplicaciones para Turismo Activo	84
Figura 48. Aplicaciones para Turismo Activo	84
Figura 49. Aplicaciones para Turismo Activo	85
Figura 50. App Kayak.com	87
Figura 51. Experiencia de usuario en la Rioja	89
Figura 52. Escenarios de actuación del proyecto Hermes	91
Figura 53. App el Aire de Madrid	92
Figura 54. Índice de Turismo y Viajes	94
Figura 55. Árbol de problemas	97
Figura 56. Árbol de Objetivos	99
Figura 57. Flujo de Trabajo de datos en una aplicación empresarial	108
Figura 58. Arquitectura Tradicional de Un sistema de Información de una Organización	109
Figura 59. Beneficios para las organizaciones al adoptar LOD	112
Figura 60. Implementación de LOD corporativo	113
Figura 61. Funcionamiento un servidor LOD para publicar sus propios datos	114
Figura 62. Mapeo de Datos de RDB a RDF	119
Figura 63. Transformación de datos	120
Figura 64. Sponger Inputs y Outputs	122

Figura 65. Modelo de publicación y registro de datos	124
Figura 66. Apuestas Productivas del Plan de competitividad	126
Figura 67. Stakeholder engagement in an open innovation scenario	127
Figura 68. Pasos para Apertura de Datos	132
Figura 69. Fuentes de datos del sector Turismo de Norte de Santander	135
Figura 70. Ontología LOD Turismo Norte de Santander	139
Figura 71. Proceso de Creación de Nuevo Conjunto de Datos	142
Figura 72. Proceso de Actualización de un conjunto de datos	142
Figura 73. Arquitectura del Proyecto LOD turismo de Norte de Santander	149
Figura 74. Caso de uso para Editores de Datos	152
Figura 75. Caso de Uso para Consumidores de Datos	155
Figura 76. Plataforma CKAN	157
Figura 77. Interfaz de dataset CKAN	158
Figura 78. Interfaz de ejecución del Repositorio Virtuoso	159
Figura 79. Entorno Web del Repositorio Virtuoso	160
Figura 80. Administración del Repositorio Virtuoso	161
Figura 81. Interfaz Web “Linked Data” Servidor Virtuoso	161
Figura 82. Interfaz para Cargar la ontología en el repositorio Virtuoso	162
Figura 83. Ontologías Cargadas en el repositorio	162
Figura 84. Mapa del Sitio	164
Figura 85. Mapa de contenidos de la aplicación	165
Figura 86. Relación entre Arquitectura de la Información y diseño	165
Figura 87. Wireframes de la AppMovil	166

Figura 88. Wireframes de la web	166
Figura 89. Aplicación AppMovil y Web	167
Fuente 90. Arquitectura de una aplicación Híbrida	168
Fuente 91. Acceso a LOD desde APP	168

LISTA DE CUADROS

	Pág.
Cuadro 1. Matriz de lineamientos – Política de mercadeo y promoción turística: “Promoviendo la clase mundial del turismo colombiano”	29
Cuadro 2. Visión Turística por Regiones	95
Cuadro 3. Beneficios Marginales Directos	100
Cuadro 4. Beneficios extras	101
Cuadro 5. Descripción los stakeholders	128
Cuadro 6. Actividad vs. Descripción	133
Cuadro 7. Categorías de Datos Sector Turístico	134
Cuadro 8. Sitios de Interés y Turísticos, Municipio de Arboledas N de S	136
Cuadro 9. Sitios de Interés y Turísticos, Municipio de Cacota N de S	137
Cuadro 10. Sitios de Interés y Turísticos, Municipio de Cachira N de S	137
Cuadro 11. Propiedad y label	140
Cuadro 12. Propiedad e identifier	140
Cuadro 13. propiedad y lat	141
Cuadro 14. Open Data Licenses que se pueden usar en el proyecto LOD del Turismo de Norte de Santander	147
Cuadro 15. Metadatos Utilizados para el Proyecto	148
Cuadro 16. Herramientas y tecnológicas utilizadas para la apertura de los Datos del sector turístico de Norte de Santander	150
Cuadro 17. Requisitos de hardware dentro del proyecto LOD Turismo Norte de Santander	150
Cuadro 18. Descripción de casos de uso para el editor de datos	152

Cuadro 19. Descripción de casos de uso para el consumidor de datos	155
Cuadro 20. Ventajas del repositorio Virtuoso	159
Cuadro 21. Comparaciones importantes	169

LISTA DE ANEXOS

	Pág.
Anexo A. Siglas	179
Anexo B. Glosario	181
Anexo C. Definición de propiedades LODObject	182
Anexo D. Documento RFD-LOD ontology	183
Anexo E. Inventario de Turismo	184
Anexo F. Código Fuente	185

RESUMEN

El sector del Turismo es uno de los de mayor futuro para el desarrollo económico del país, debido a la cantidad de empleos que genera y el impacto en las pequeñas y medianas empresas. La tendencia está encaminada a la agilización de este gran mercado de oportunidades para facilitar el acceso a las a la información presente y la masificación de los datos, permitiéndoles así la construcción de un entorno de mercado competitivo. El gobierno nacional ha formulado políticas con el fin de convertir a Colombia en un “destino turístico de clase mundial”, una estrategia para lograrlo es impulsando la “marca país” y que ha dado muy buenos resultados en los últimos años, aunque todavía son insuficientes.

Desde otro lado, el creciente uso de Internet durante los últimos años ha puesto de manifiesto la importancia de la misma para la ejecución de aplicaciones, utilizando como plataforma la Web. Las nuevas tecnologías e iniciativas están emergiendo para dar soporte al desarrollo y despliegue de aplicaciones y servicios, así como para la publicación de datos e información. En este sentido, iniciativas y movimientos como open data, y la evolución a Linked Data dentro de la nueva “Web de Datos”, le da valor a los elementos y recursos de la web buscando una mayor integración e interoperabilidad entre aplicaciones y facilitar acceso libre a los datos a través de estándares y modelos unificados.

Para el caso de la información turística En norte de Santander surge la necesidad de abordar problemas tales como la dispersión de la información, la heterogeneidad de las fuentes de datos, múltiples formatos, la cantidad de actores que intervienen en el proceso. Con el fin de proporcionar una solución a estas necesidades mediante la aplicación de la iniciativa Open Data y Linked Data, se realiza en este trabajo un estudio de estado del arte de la iniciativa, la integración de tecnologías emergentes, el fortalecimiento del turismo como un sector de importancia para el gobierno nacional, la visión de las ciudades inteligentes y la apuesta por una mayor visibilidad de la región, orientado específicamente al sector turístico. Para lograr de esta forma el acceso a la información y a los datos de una forma abierta, para incrementar las oportunidades de negocio, la innovación y el emprendimiento, favoreciendo la visibilidad regional e impulsando un entorno de datos abiertos y enlazados, como estrategia para los sectores productivos tanto por su carácter económico como por el creciente desarrollo de innovación abierta

Finalmente se muestra un subconjunto de datos del sector Turístico de Norte de Santander, convertido a la iniciativa LOD. Mediante la introducción de un esquema común. Los datos resultantes tiene que ajustarse a los principios establecidos en

Linked Open Data y se publicaron en un servidor web por lo que se puede acceder y consultar utilizando SPARQL.

Palabras Claves: Turismo de Norte de Santander, Open Data, Linked Open Data, Smart Sities, Web Semantica.

INTRODUCCIÓN

La información que se presenta en la World Wide Web (WWW) está generalmente diseñada para los seres humanos. Y en los últimos años el movimiento se ha dado hacia la generación de contenidos, especialmente con el surgimiento de la Web 2,0 y la revolución de las redes sociales, este contenido está diseñado especialmente en lenguajes como HTML5. Aunque en la versión de HTML5 se incluyen algunos conceptos de semántica de los datos, todavía es limitado y no ha casos de uso prometedores que permitan la interpretación de datos por las máquinas. Teniendo en cuenta que la información en la web esta por lo general sin anotación semántica, se dificulta ser interpretado por máquinas. Por lo que generalmente por ejemplo, los motores de búsqueda buscan coincidencias a través del texto. La Web Semántica introduce el tipo de anotación semántica que se requiere para el razonamiento automatizado basado en los datos en la web.

La visión de la Web Semántica se introdujo en 1998 cuando Sir Tim Berners-Lee publicó la hoja de ruta¹ donde se resume el concepto y la visión. La idea es que las máquinas puedan entender semánticamente el contenido en la web. Esto es posible mediante la adición de metadatos sobre la información de los datos. Estos metadatos se basan en esquemas a lo que se llamó las ontologías, que permite el razonamiento computacional basado en datos que fueron publicados en la web. Para materializar el proceso en entornos de producción y principalmente abordar la implementación en empresas tanto pública como privada. Nace la iniciativa de Linked Data que se basa en un conjunto de principios y que fue propuesto originalmente por Tim Berners-Lee en 2006, para aplicar la arquitectura de la web para el intercambio de datos estructurados a escala mundial². Ahora, se considera un movimiento mundial que ha recibido el apoyo de varias organizaciones, incluidos los gobiernos, las empresas, organizaciones sin ánimo de lucro, universidades, entre otros³. Para apoyar este movimiento, Tim Berners-Lee propuso un esquema de despliegue de cinco estrellas para clasificar datos abiertos.

Con la oportunidad tecnológica y los retos que imponen en el nuevo milenio, surge Open Data como movimiento mundial para abrir los datos al público dentro de las organizaciones públicas y privadas. Estos datos deben publicarse en los términos

¹ Tim Berners-Lee: SemanticWeb Road map. <http://www.w3.org/designissues/semantic.html>, 1998. consultado el 18-05-2015).

² Christian Bizer, Tom Heath, Kingsley Idehen, and Tim Berners-Lee. *Linked Data on the Web*. 2008.

³ Bizer, Christian; Heath, Tom; Berners-Lee, Tim. "Linked Data—The Story So Far". (2009): pp. 1–22. *International Journal on Semantic Web and Information Systems* 5.

y formatos estructurados específicos, asegurando que los datos están disponibles para cualquier persona y para cualquier propósito. Lo que significa que los datos deben estar en formatos legibles por máquina. Iniciática liderada principalmente por el gobierno de los estados Unidos y Reino Unido quienes dan origen al open government data.

Ante las restricciones que imponen algunos formatos con CSV, documentos PDF, etc. Y la idea de seguir avanzando hacia el esquema de las cinco estrellas propuesto por Tim Berners-Lee. Se empieza a implementar Linked Data para los datos abiertos lo que da origen a Linked Open Data que significa, exponer, compartir y conectar los datos relacionados en la web, a través de formatos abiertos y licencias abiertas⁴. Lo que permite crear un poderoso ecosistema donde los gobiernos, las empresas y los ciudadanos interactúan, comparten y consumir datos. Creando unos beneficios importantes que se pueden generar con la dinámica de intercambio de datos.

Por lo tanto, teniendo en cuenta el contexto estratégico que supone el sector turístico de Norte de Santander para el desarrollo económico de la región y considerando las iniciativas de Open data y Linked Data, que animan y potencian la reutilización de datos, información y modelos de conocimiento compartido, mejoran e impulsan este sector con un nuevo enfoque y brindan la posibilidad a muchos ciudadanos a generar nuevas ideas y modelo de negocios.

⁴ creativecommons.org. <http://creativecommons.org/licenses/by/3.0/legalcode>, 2013. consultado el 18-05-2015)..

1. MARCO TEORICO

1.1 ANTECEDENTES

Las tecnologías de la información y las comunicaciones han ido revolucionando diferentes áreas de la industria, El turismo es uno de los campos que está siendo fuertemente impactado para bien, por los cambios que está viviendo la web. En un principio, con la web 1.0, las personas se conectaban a la red y las empresas ofrecían la información que los clientes buscaban. La web 2.0 abrió las posibilidades de colaboración, las personas ahora comparten, se conectan con otras personas; hay un diálogo, no es comunicación unidireccional. Pero, en muy poco tiempo, las cosas han cambiado significativamente. La web 3.0, el crecimiento del móvil y el mundo de las aplicaciones ahora enriquecen la experiencia de las personas.

Esta evolución de la web 1.0 hasta la web 3.0 coincide en que en esta etapa se le añadirá significado a la Web. La Web 3.0 es el nacimiento de la web de datos "Data Web", ya que los formatos en que se publica la información en Internet son dispares, como XML, RDF y micro formatos; el reciente crecimiento de la tecnología SPARQL, permite un lenguaje estandarizado y una API para la búsqueda a través de bases de datos en la Red. La Web de datos permite un nuevo nivel de integración de datos y aplicación inter-operable, haciendo los datos tan accesibles y enlazables como las páginas Web. Esto ha permitido que muchos países empiecen a trabajar en iniciativas de Datos abiertos Open Government Data⁵. En el cual Colombia se ha involucrado y busca exponer sus datos al público para ser reutilizados por terceros como una forma de gobernabilidad y transparencia. A través de estos datos abiertos, personalizar la forma en que se presenta y se combina la información, la cantidad de Apis públicas de aplicaciones como Facebook, google, Twitter, etc. Y si a esto se añade un poco de lógica, mediante una RDF, u otras tecnologías como los mapas temáticos y algo de razonamiento basado en técnicas de inteligencia artificial, Internet podría estar cerca de alcanzar las aspiraciones iniciales de su inventor, Tim Berners-Lee. Integrar contenidos provenientes tanto de fuentes y servicios municipales como de fuentes de información y servicios externos, resultado que se conoce como "mashup".

Norte de Santander no es ajeno a la dinámica y beneficios que puede generar el turismo como motor de desarrollo en tanto que su ámbito de acción abarca todos los reglones de la economía y fortalece la dinámica regional. El departamento

⁵ "Iniciativa Open Government Data". <http://opengovernmentdata.org/>

posee las potencialidades requeridas para su desarrollo, sin embargo, se requiere del trabajo mancomunado de los sectores público y lo privado, cuyos esfuerzos apuntan a hacer de esta actividad una alternativa de desarrollo económico y social. El comportamiento del sector turismo a nivel mundial demuestra que sigue siendo una de las actividades más crecientes de las últimas décadas. A pesar de los acontecimientos que se han presentado en los últimos años, relacionados con terrorismo, desastres naturales, epidemias, conflictos políticos y económicos, su crecimiento ha sido constante durante el periodo del 2001 al 2014. Según la OMT⁶ la demanda turística expresada en número de viajeros se duplicará en 20 años alcanzando de un 4% a 6% anual convirtiéndose en uno de los sectores con mayores expectativas de crecimiento.

Con el fin de afianzar el turismo como un sector económico de importancia para la región la cámara de comercio, la Secretaria de Desarrollo Económico del Departamento, la Corporación Mixta de Promoción de Norte de Santander y Corponor⁷, con el apoyo del Ministerio de Comercio, Industria y Turismo (MICT) y el Fondo de Promoción Turística de Colombia (FPTC), se desarrolló en el año 2008 y 2009, un levantamiento técnico de los recursos y atractivos turísticos con los que cuenta el departamento. La ejecución de dicho proyecto permitió hacer presencia en 32 municipios de los 40 que componen el departamento y mediante el formato de levantamiento de atractivos subdividido entre bienes patrimoniales materiales e inmateriales, bienes de infraestructura, y bienes naturales, su ejecución logró consolidar una importante herramienta para aproximación de la vocación turística y los recursos con que se cuenta. A demás de identificar los modelos de negocios que puedan surgir.

Para afianzar estas estrategias la Gobernación de Norte de Santander ha firmado una serie de convenios que se han venido realizando y ejecutando desde el año 2000. En la actualidad se tiene el Convenio de Asociación de Competitividad para el departamento, firmado por el Ministerio de Comercio, Industria y Turismo, la Gobernación y la Corporación Mixta de Promoción de Norte de Santander, cuya fecha de la última firma se llevó a cabo el 19 de noviembre de 2009. El objetivo es: "Mejorar la productividad y competitividad del destino turístico del Departamento de Norte de Santander, con miras al fortalecimiento de los servicios turísticos, a través de una mayor integración de los eslabones de la cadena productiva del destino". (Convenio de Competitividad, 2009). La visión de este convenio para el departamento fue: Destino histórico-cultural y de compras. Durante este trabajo, el MCIT-FPT manifestó el interés de apoyar la formulación del Plan de Desarrollo Turístico de Norte de Santander. Como resultado de dicho proceso, se formula el

⁶ Organización Mundial del Turismo. [En línea]. Disponible en Internet: <<http://www2.unwto.org/es>>

⁷ Corporación Autónoma del Norte de Santander. [En línea]. Disponible en Internet: <<http://www.corponor.gov.co/>>

mismo para el periodo 2012-2016. Por su parte, el municipio de San José de Cúcuta elaboró el Plan Sectorial de Desarrollo Turístico 2011–2020 (Plan Convencional) del municipio, a través de un proyecto aprobado por el FPT. A partir de estos esfuerzos se logró crear la oficina de turismo del departamento y el consejo departamental de turismo, además de separar la secretaria de turismo de Cúcuta de la secretaria de cultura. Estas iniciativas buscan integrar todas aquellos agentes interesados por el turismo en el consejo departamental en los que participan COTELCO⁸, ANATO⁹, ACODRES¹⁰, la Corporación Mixta, alcaldías, PROCOLOMBIA¹¹, la Comisión Regional de Competitividad, la Oficina Territorial del Ministerio de Comercio, Industria y Turismo, la Cámara de Comercio, Sena, universidades, entre otros.

Según PROCOLOMBIA en el 2013 se registró la llegada a Colombia de un total de 2.594.809 viajeros, que incluye extranjeros no residentes en Colombia (1.726.300), el total de colombianos residentes en el exterior y los cruceristas que arribaron a los diferentes puertos del país. Al incluir las llegadas de viajeros por zonas de integración fronteriza (1.153.248 viajeros), reportada por la Dirección de Análisis Sectorial y Promoción de Turismo del MinCIT (Ministerio de Comercio, Industria y Turismo), se registró un ingreso total de viajeros de 3.748.057. A cierre de 2014 se registró la llegada a Colombia de un total de 2.879.543 viajeros, que incluye la llegada de extranjeros no residentes en Colombia (1.967.814), el total de colombianos residentes en el exterior y los cruceristas que arribaron a los diferentes puertos del país. Al incluir las llegadas de viajeros por zonas de integración fronteriza (1.313.200 viajeros), reportada por la Dirección de Análisis Sectorial y Promoción de Turismo del MinCIT, se registró un ingreso total de viajeros de 4.192.743.

⁸ Asociación Hotelera y Turística de Colombia. [En línea]. Disponible en Internet: <<http://www.cotelco.org/>>

⁹ Asociación Colombiana de Agencias de Viajes y Turismo. [En línea]. Disponible en Internet: <<http://www.anato.org/>>

¹⁰ Asociación Colombiana de la Industria Gastronómica. [En línea]. Disponible en Internet: <<http://acodres.com.co/>>

¹¹ Exportaciones, Turismo, Inversión marca País. [En línea]. Disponible en Internet: <<http://www.procolombia.co/>>

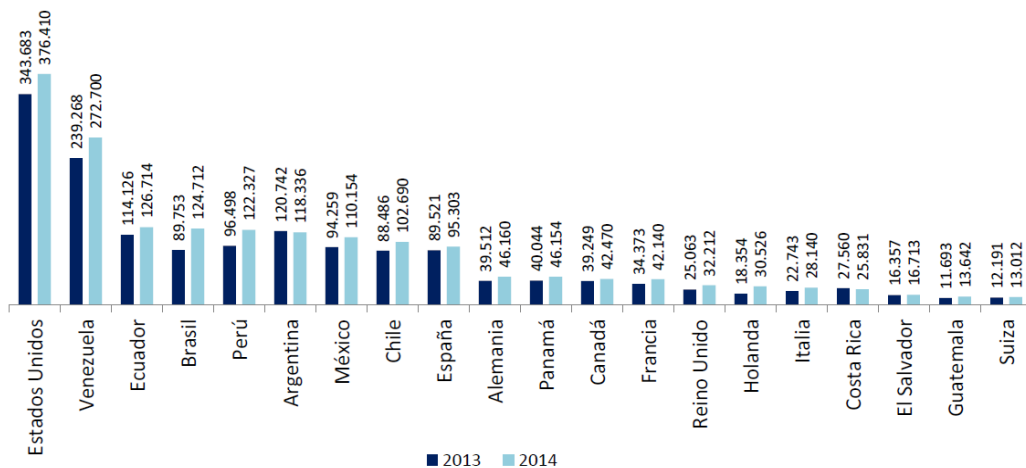
Figura 1. Cifra equivalente a la suma de extranjeros no residentes en Colombia más el total de colombianos residentes en el exterior y los visitantes de cruceros 2013 y 2014



Fuente: Procolombia. [En línea]. Disponible en Internet: <<http://procolombia.com>>

Estados Unidos y los países de América del Sur, son los mercados que emiten el mayor número de viajeros extranjeros no residentes a Colombia.

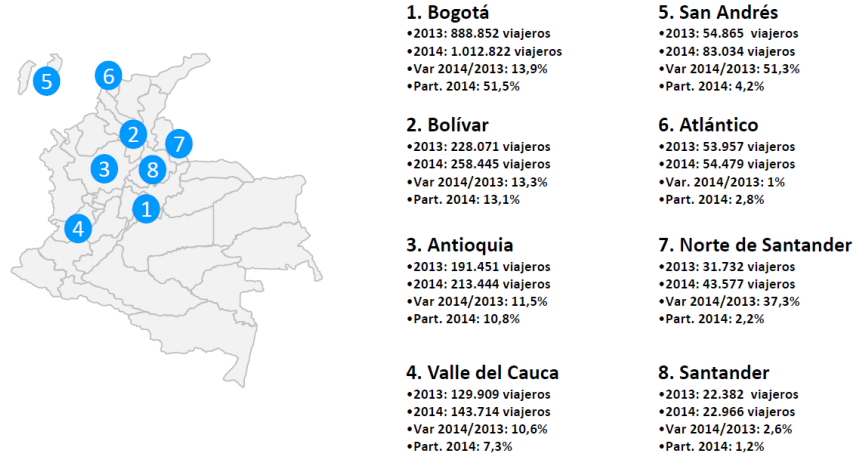
Figura 2. Cifras de turistas por países



Fuente: Procolombia. [En línea]. Disponible en Internet: <<http://procolombia.com>>

Teniendo en cuenta que la mayor afluencia de turistas son provenientes de Venezuela. Norte de Santander por ser la frontera más activa con este país, se considera una región con mayores oportunidades para poner en práctica modelos de negocios sostenibles y rentables en el turismo, hoy día Norte de Santander es el séptimo departamento con mayor afluencia de turística como lo indica PROCOLOMBIA.

Figura 3. Llegada de Extranjeros no residentes por departamento



Fuente: Procolombia. [En línea]. Disponible en Internet: <<http://procolombia.com>>

1.1.1 Procolombia. Proexport la entidad que nació a comienzo de los años 90 para impulsar las exportaciones dio un giro para adecuarse a las nuevas realidades de Colombia y del mundo. El pasado 13 de noviembre de 2014 fue el lanzamiento de Procolombia y es la entidad encargada de promover el Turismo, la Inversión Extranjera en Colombia, las Exportaciones no minero energético y la imagen del país.

A través de la red nacional e internacional de oficinas, ofrece apoyo y asesoría integral a los clientes, mediante servicios o instrumentos dirigidos a facilitar el diseño y ejecución de su estrategia de internacionalización, que busca la generación, desarrollo y cierre de oportunidades de negocios

1.2 POLITICA DE MERCADEO Y PROMOCION TURISTICA “COLOMBIA DESTINO TURISTICO DE CLASE MUNDIAL¹²”

Debido a la importancia del sector turismo en la economía nacional, el Ministerio de Comercio, Industria y Turismo, creó la Política de Mercadeo y Promoción Turística para apoyar y fomentar la adecuada distribución de los recursos

¹² COLOMBIA. MINISTERIO DE COMERCIO, INDUSTRIA Y TURISMO. “Política de mercadeo y promoción turística de Colombia”. [En línea] [Consultado Diciembre 4 del 2012] Disponible en Internet: <<https://www.mincomercio.gov.co/minturismo/publicaciones.php?id=195>>

destinados al Mercadeo del turismo en Colombia.

Estas iniciativas de mercadeo le ayudan al país a competir con los mercados internacionales ya que es necesario saber las tendencias en los turistas nacionales e internacionales y así poder conocer los requerimientos de la demanda en el sector y las falencias de los productos ofrecidos en el país. En un mundo globalizado y especialmente en el sector turismo es fundamental que el país cuente con herramientas que le permita tener información actualizada de los cambios que con el paso del tiempo se van presentando, las nuevas tendencias que los países desarrollados han presentado debido a que cada día los consumidores son más exigentes y con diversos gustos.

Durante los últimos años, el país ha desarrollado políticas turísticas, y que constituyen un factor fundamental para consolidar a Colombia en un país de destino turístico con niveles de competitividad, a pesar de los diferentes problemas de inseguridad el país fue fortaleciendo el turismo interno, antes de pensar en turismo receptivo. Ante esta oportunidad que brindaba la posibilidad del turismo receptivo, el gobierno trazo una estrategia marco que mejorara la competitividad de los productos y destinos turísticos como una apuesta a mediano y largo plazo. La idea es generar nuevos productos en turismo y que permita mayor competitividad que satisfagan las propuestas de un mercado internacional más exigente, es importante que los trabajos que se piensen realizar estas enfocados bajo las mismas políticas de calidad.

Esta consolidación del turismo interno y receptivo, el sector empezó a ganar un espacio significativo en las expectativas del país en cuanto al desarrollo económico, al evaluar indicadores como entrada de turistas, generación de empleo, ocupación hotelera y el PIB. Algunas de esas políticas son referentes destacados ya que en muchas se mencionan para su intervención directamente o brinda los espacios necesarios para su consolidación.

1.2.1 Matriz de lineamientos – política de mercadeo y promoción turística. Se describe a continuación

Cuadro 1. Matriz de lineamientos – Política de mercadeo y promoción turística: “Promoviendo la clase mundial del turismo colombiano”

TEMÁTICA	INVEST. DE MERCADOS	DESARROLLO DE PRODUCTOS	ACTIVIDADES Y HERRAMIENTAS DE PROMOCIÓN	COMERCIALIZACIÓN	ARTICULACIÓN DEL SECTOR
Objetivo Específico	Fortalecer las acciones en materia de investigación de mercados turísticos, que permitan contar con información crucial y el mantenimiento de Estadísticas fiables.	Promover los productos turísticos competitivos diferenciadores en la oferta turística nacional, acuerdos con las nuevas referencias y tendencias mundiales en turismo, así como los lineamientos de la política de competitividad turística.	Lograr un efectivo mercadeo y promoción del Turismo y los destinos nacionales, a través del desarrollo actividades y herramientas de promoción turística de alto impacto y eficiencia.	Impulsar estrategias para el fortalecimiento de la comercialización turística.	Articular la gestión interinstitucional de todas las Entidades públicas y privadas vinculadas al mercadeo y la promoción turística.
Lineamiento	Fortalecimiento de la Investigación de Mercados turísticos.	Diferenciación y desarrollo de productos turísticos prioritarios.	Fortalecimiento de las actividades y herramientas de promoción turística.	Facilitación de oportunidades de negocio para la oferta turística colombiana.	Articulación interinstitucional para el fortalecimiento de los mecanismos de mercadeo y promoción.

Cuadro 1. (Continuación)

TEMÁTICA	INVEST. DE MERCADOS	DESARROLLO DE PRODUCTOS	ACTIVIDADES Y HERRAMIENTAS DE PROMOCIÓN	COMERCIALIZACIÓN	ARTICULACIÓN DEL SECTOR
Acciones	<p>Elaborar estudios de segmentos de mercados, enfocados en los países prioritarios y en su dinámica de comportamiento turístico.</p> <p>Elaborar estudios y programas e identificar herramientas que apoyen la estructuración de planes de mercadeo turístico de los productos y destinos colombianos, según los segmentos o perfiles de turista.</p> <p>Investigar en detalle Las estrategias y acciones adelantadas por los países y destinos competidores en materia turística.</p>	<p>Promover la creación y diseño de productos turísticos diferenciados, que sean viables según las preferencias de la demanda y los mecanismos de comercialización pertinentes para cada mercado.</p> <p>Promover el desarrollo de productos turísticos innovadores en la oferta turística nacional, con conceptos y prestaciones ajustados a las preferencias de los turistas, tales como el enfoque hacia el turismo experiencial.</p> <p>Reposición para los productos turísticos de sol y playa y de cambio de clima que tienen preferencia en el mercado nacional y que requieren un fortalecimiento de sostenibilidad, incluyendo el ecoturismo y el turismo cultural, entre otros.</p>	<p>Desarrollar material promocional pertinente, impactante y de alta calidad, que contribuya con la conservación del ambiente y que responda al plan de mercadeo del destino y del producto.</p> <p>Ejecutar las acciones y actividades de mercadeo y promoción más eficientes según los mercados emisores, de manera que se ajusten a los perfiles y necesidades de los turistas.</p> <p>Promover los destinos y empresas certificadas en calidad turística, en concordancia con los parámetros del Sistema Nacional de Calidad.</p>	<p>Implementar herramientas de gestión de la comercialización de los destinos turísticos, en los mercados emisores nacionales e internacionales.</p> <p>Involucrar a las oficinas comerciales de Colombia en el exterior en la adopción de las estrategias de comercialización del turismo Colombiano.</p> <p>Promover nuevas estrategias de comercialización según el producto o destino a promocionar.</p>	<p>Definir competencias de las diferentes entidades del Estado para la gestión del mercadeo y promoción del turismo en los ámbitos nacional e internacional.</p> <p>Propiciar en las regiones, la articulación de las entidades territoriales del orden departamental y municipal y sus órganos de gestión del turismo.</p> <p>Involucrar a las entidades del sector privado del turismo, tanto aportantes como no aportantes de la contribución parafiscal, en el desarrollo y ejecución de la Política de Mercadeo y Promoción del Turismo.</p>

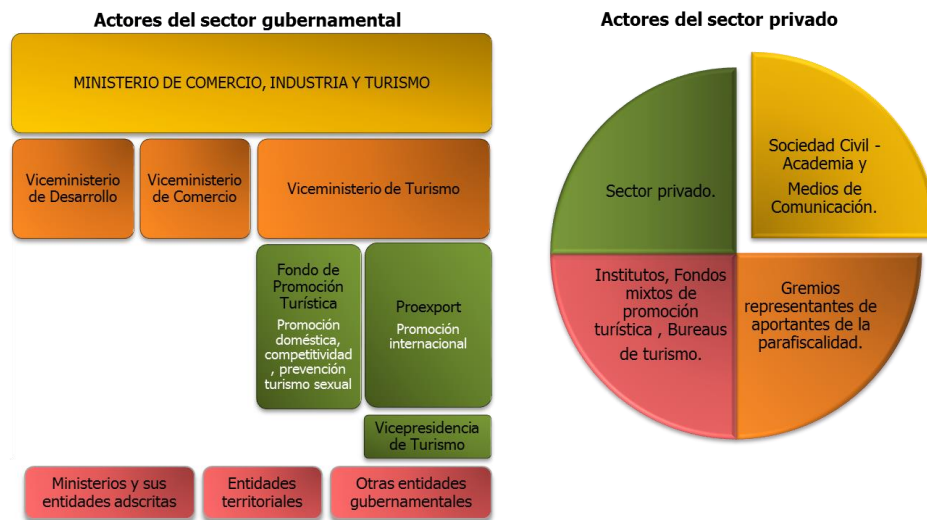
Cuadro 1. (Continuación)

TEMÁTICA	INVEST. DE MERCADOS	DESARROLLO DE PRODUCTOS	ACTIVIDADES Y HERRAMIENTAS DE PROMOCIÓN	COMERCIALIZACIÓN	ARTICULACIÓN DEL SECTOR
	Identificar las acciones y actividades de mercadeo y promoción más eficientes según los mercados emisores, de manera que se ajusten a los perfiles y necesidades de los turistas.	Especializar y promover los ocho productos "básicos" que concretan la oferta turística del país para el mercado nacional, a saber: sol y playa; historia y cultura; agroturismo; ecoturismo; etnoturismo; deportes y aventura; ferias y fiestas.	Fortalecer de manera permanente el componente turístico de la marca país.	Promover el uso de herramientas digitales, como parte de la estrategia de comercialización.	Promover el liderazgo compartido de las instituciones públicas, las comunidades receptoras y del sector privado, en la puesta en marcha de programas para el mercadeo y promoción del turismo.
	Profundizar el levantamiento de datos y estadísticas relevantes sobre el comportamiento del sector y de los destinos turísticos, que apoye la toma de decisiones de mercado acordes con las necesidades del turismo.	Desarrollar para el mercado internacional, productos turísticos viables para los segmentos de turismo de salud, ecoturismo, turismo cultural y de eventos, congresos y convenciones, en concordancia con las ventajas comparativas y competitivas del país y el Plan Sectorial de Turismo.	Crear programas de promoción y comunicación turística como apoyo a la estrategia para reducción de las advertencias de viaje de gobiernos extranjeros para viajar a Colombia o travel warnings, especialmente en mercados emisores potenciales	Desarrollar instrumentos y programas de asistencia técnica a la comercialización turística para mejorar la gestión de venta de los destinos y productos turísticos.	Reforzar la interlocución y trabajo conjunto del Ministerio de Comercio, Industria y Turismo, y las entidades territoriales en las estrategias de mercadeo y promoción turística, a partir de los Convenios de Competitividad.

Fuente: COLOMBIA. Ministerio de Comercio, Industria y Turismo. Fragmento de la matriz

1.2.2 Actores Involucrados. Las nuevas disposiciones crean nuevos grupos de trabajo para hacer de nuestro país un destino de talla mundial.

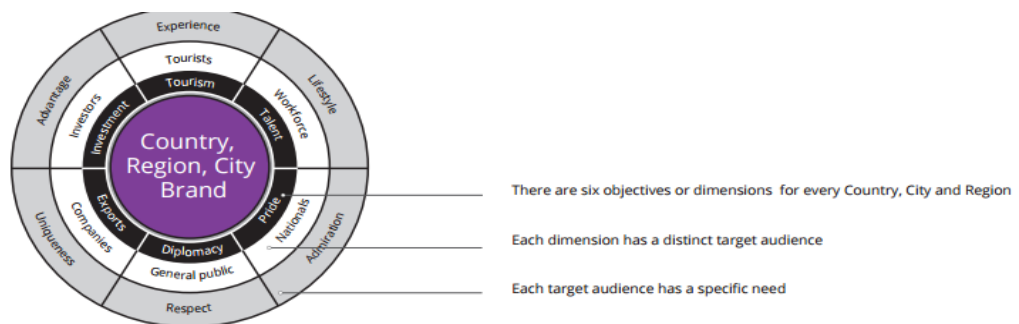
Figura 4. Sectores Involucrados



Fuente: COLOMBIA. Ministerio de Comercio, Industria y Turismo

1.2.2 Marcas País en las Américas. El desarrollo de una marca país busca incentivar que el mercado internacional para que compren productos en el país, visiten el país y fomenten la inversión. Pero además, una “marca país” es el bien más intangible de una nación y puede exceder incluso su performance económica, sus niveles de conocimiento e incluso su tamaño geográfico.





















Figura 5. Entes involucrados en marca País



Fuente: Bloom Consulting

Según el ranking mundial 2014-2015 de marca país, elaborado por la consultora de branding internacional Future Brand¹³, Colombia aparece en el puesto quince (15), perdiendo dos puestos, y aparece en Latinoamérica en el puesto trece (13)

Figura 6. Ranking Marca País de las Américas

#	Variation		World Rank	CBS Rating
1.	-	 United States of America	1.	AAA
2.	+1	 Canada	17.	BBB
3.	-1	 Mexico	20.	AA
4.	-	 Argentina	39.	A
5.	-	 Brazil	40.	BBB
6.	+2	 Chile	44.	A
7.	-	 Puerto Rico	45.	A
8.	+1	 Panama	47.	BBB
9.	+3	 Peru	52.	BBB
10.	-4	 Dominican Republic	54.	BBB
11.	-	 Costa Rica	58.	BBB
12.	+4	 Bahamas	59.	BBB
13.	-3	 Uruguay	61.	A
14.	-	 Aruba	65.	A
15.	-2	 Colombia	68.	BB
16.	+1	 Jamaica	71.	BB
17.	-2	 Guatemala	73.	BBB
18.	-	 Cuba	77.	B
19.	-	 Ecuador	79.	A
20.	-	 Barbados	83.	BBB



















Fuente: Future Brand.

¹³ Bloom Consulting Country Brand Ranking. Estudio sobre el posicionamiento de marca país en el mundo.

1.2.3 Marca País Mundial. El último estudio Country Brand Index 2014-2015, que ha elaborado Future Brand analiza 118 países de todo el mundo, y posiciona en lo más alto del ranking a aquellos que generan una percepción más optimista e influyente entre los consumidores. En definitiva, en definitiva lo que significa que. “La gente tiene que consumir los productos y servicios y preferirte como lugar para vivir y aprender”.

Los países que en estos momentos destacan en campos asociados a la modernidad, la innovación, la tecnología y la sostenibilidad aparecen en lo más alto en el ranking. También cabe destacar la importancia en la relación entre el poder de la marca país y el poder de las marcas que son originarias del lugar, que también es estrecha.

Figura 7. Top de las 25 mejores marca País

#	Variation		CBS Rating
1.	-	 United States of America	AAA
2.	-	 Spain	AA
3.	-	 Germany	AAA
4.	+9	 Hong Kong SAR, China	BBB
5.	-1	 France	A
6.	-	 Thailand	A
7.	+1	 Australia	BBB
8.	+4	 Macao SAR, China	AA
9.	+1	 China	BBB
10.	-3	 Italy	BBB
11.	-6	 United Kingdom	A
12.	-1	 Austria	AA
13.	+4	 Singapore	A
14.	+10	 Japan	A
15.	-1	 Switzerland	A
16.	-7	 Turkey	A
17.	+3	 Canada	BBB
18.	-	 Korea, Rep.	A
19.	-4	 Malaysia	BBB
20.	-1	 Mexico	AA
21.	-	 Netherlands	A
22.	+1	 Portugal	A
23.	+3	 India	BBB
24.	+4	 South Africa	AA
25.	-9	 Greece	A

Fuente: Future Brand.

1.2.4 Visión Colombia 2019. Un documento formulado en el 2005 y que propone las metas de crecimiento en el país, Según el documento unos de los sectores que mayor contribuirá al crecimiento económico del país será el turismo Como uno de los tres sectores potenciales en esos términos destaca.

“Colombia posee una gran diversidad turística. De acuerdo con el Ministerio de Cultura, han sido establecidas como patrimonios de Colombia 1128 localidades del país, repartidas en 294 municipios y 30 departamentos. Además, Colombia cuenta con ventajas comparativas que le permiten ofrecer una gran variedad de productos turísticos: ecoturismo, turismo rural y agrario; turismo histórico y cultural; turismo de negocios; de aventura y de sol y playa. A pesar de los avances presentados en los últimos años, es evidente que las ventajas comparativas no son suficientes para competir con éxito en los mercados internacionales. Para lograr un adecuado desarrollo del sector se requiere construir, sobre la base de las ventajas comparativas con que cuenta el país, ventajas competitivas. Esto llevará al país a adoptar estrategias de desarrollo para la actividad turística, mediante la potencialización de sus ventajas competitivas, el posicionamiento de la marca país y la generación de infraestructura adecuada para su desarrollo”¹⁴.

1.2.5 El Conpes 3397. Un documento suscrito en el 2005, y se elaboró a partir de las necesidades de infraestructura para el sector turismo en el marco del convenio de competitividad para el turismo, buscando impulsar varios proyectos regionales, además de integrar a los demás sectores como el SENA, Ministerio de ambiente y vivienda, Ministerio de Agricultura. Para buscar estrategias que brinde turismo de calidad.

1.3 TECNOLOGIAS PARA EL TURISMO

Las empresas del siglo XXI tienen retos y oportunidades para construir sobre las tecnologías un nuevo modelo de relación con sus clientes actuales y potenciales más eficiente, innovador y humano. Para ello se necesita una cultura empresarial totalmente novedosa en el que la captación, gestión e interpretación de la información sirva de inspiración a todo el negocio. Bajo esa mirada, los datos dejan de comprenderse como parte de un proceso para convertirse, directamente, en el sistema.

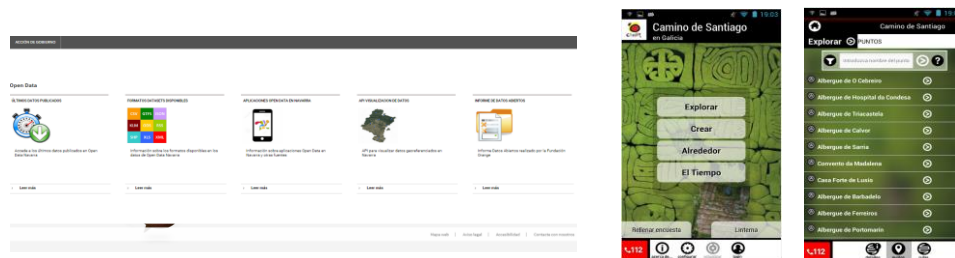
¹⁴ COLOMBIA. Ministerio de Comercio, Industria y Turismo. 2008. “PLAN SECTORIAL DE TURISMO 2008-2010 “COLOMBIA, DESTINO TURÍSTICO DE CLASE MUNDIAL”. [En línea] Disponible en Internet: <http://www.fontur.com.co/aymsite/aym_document/aym_normatividad/2008/PLAN_SECTORIAL_2008-2010.pdf con acceso el 10-05-2014>

Se apunta que el turismo como uno de los sectores donde mayor será el impacto de la revolución y donde una estrategia de open data, linked open data y bigdata será necesaria para enfrentar la gran hola digital. El viajero de hoy busca una experiencia personalizada, integrada y continuada, apoyada en recomendaciones, valores y motivaciones particulares expresadas de forma espontánea en Internet. Estos viajeros van dejando rastros a través de las webs oficiales y no oficiales sus preferencias personales a través del uso de las redes sociales; la posibilidad de buscar comentarios de personas sobre los próximos les ayuda a planificar sus visitas; con el crecimiento de la tecnología móvil y su popularización, los usuarios pueden construir recorridos únicos e irrepetibles con fotos y comentarios de lo que ve, escucha, siente y, en general, desea compartir. De esta forma la tecnología nos permite identificar casi de forma exacta las preferencias de los consumidores.

1.3.1 Open Data Y Turismo. La apertura y reutilización de los datos son elementos imprescindibles en la implementación de los destinos turísticos inteligentes, y posibilita el establecimiento de mecanismo para la colaboración y participación de los ciudadanos, por otro lado potencia la oferta de nuevos productos y servicios digitales, dinamizando la actividad comercial y económica y generando un valor para el turista y permite fortalecer la industria turística. Además de un impulso importante para la conformación de ciudades inteligentes. Algunas iniciativas de datos abiertos específicamente como por ejemplo:

Proyecto Información del Camino de Santiago: Muestra los mejores recursos turísticos que se pueden encontrar a lo largo del Camino de Santiago a su paso por Galicia (España), los cuales muestra en su aplicación para dispositivos móviles “Camino de Santiago en Galicia” pero adicionalmente y siguiendo la filosofía Open Data.

Figura 8. Página oficial de datos abiertos gobierno de Navarra y aplicación móvil camino de Santiago en Galicia a partir de datos abiertos



Fuente: Gobierno de Navarra [En línea]. Disponible en Internet: <http://www.gobiernoabierto.navarra.es/es/open-data>

Los datos turísticos de Euskadi, que permite obtener información geo localizada de alojamientos, restaurantes, etc.

Figura 9. Aplicación catálogo de datos de turismo Euskadi



Fuente: Open Data Euskadi

Figura 10. Portal de Datos abierto de Cataluña



Fuente: Open Data Cataluña. [En línea]. Disponible en Internet: <http://dadesobertes.gencat.cat/es>

El open data ofrece ventajas económicas con la posibilidad que se generen valores económicos a partir de los datos públicos como:

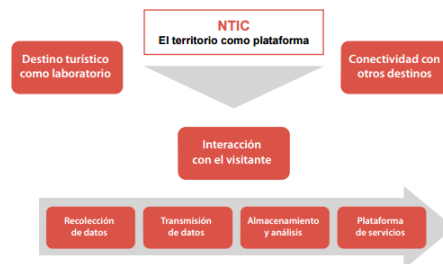
- Servicios y aplicaciones.
- Contenidos Digitales

Según el estudio de caracterización del sector infomediario¹⁵ en el 2012 la reutilización de la información pública en España genero un volumen de negocio anual de entre 330 y 550 millones de euros anuales, periodo en el que el sector empleó a una cifra de entre 3.600 a 4.400 personas para labores de reutilización de datos abiertos.

1.3.2 Open Linked Data y Turismo. La Web semántica y Linked Data son tecnologías que permiten dar un gran salto de la Web tradicional y resolver el problema de la no inferencia entre las relaciones conceptuales y navegar entre recursos. Esta es una de las grandes oportunidades que el linked data puede dar al turismo y más concretamente a los llamados “destinos inteligentes” y que al integrar con otras tecnologías puede garantizar un desarrollo sostenible y la posibilidad de avanzar hacia las Smart city o ciudades inteligentes.

Estas iniciativas apoyan el diseño de una oferta de servicios de ocio completa, interdependiente, complementaria, personalizada y orientada a una experiencia ininterrumpida. Pero además le permite, una gestión más eficiente y coordinada de los recursos públicos y privados que apoyan la actividad turística. Como se presenta en la siguiente figura.

Figura 11. Diseño de servicios Open data

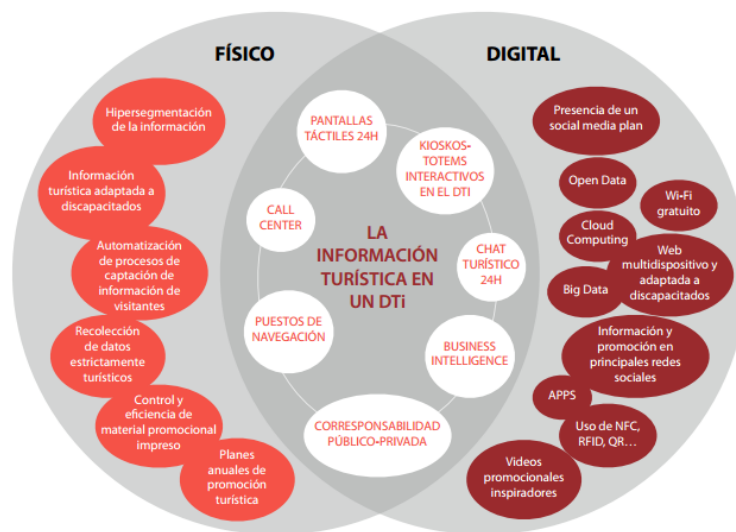


Fuente: Segittur. [En línea]. Disponible en Internet: <<http://www.segittur.es/>>

¹⁵ Ministerio de Industria, Energía y Turismo. 2012. Estudio de caracterización del Sector Infomediario en España. [En línea] Disponible en Internet: : <http://www.ontsi.red.es/ontsi/sites/default/files/info_sector_infomediario_2012_vfr.pdf> (con acceso el 12-05-2014)>

De esta manera se pueden generar destinos inteligentes que ofrecen soporte y conectan entre sí toda clase de soluciones digitales, desde mapas, audioguías y puntos de interés turístico basados en la geolocalización y la realidad aumentada, hasta aplicaciones móviles con horarios y opciones de transporte o webs que visualizan la información en tiempo real¹⁶.

Figura 12. Modelo de Integración de Tecnología para Gestión Turística



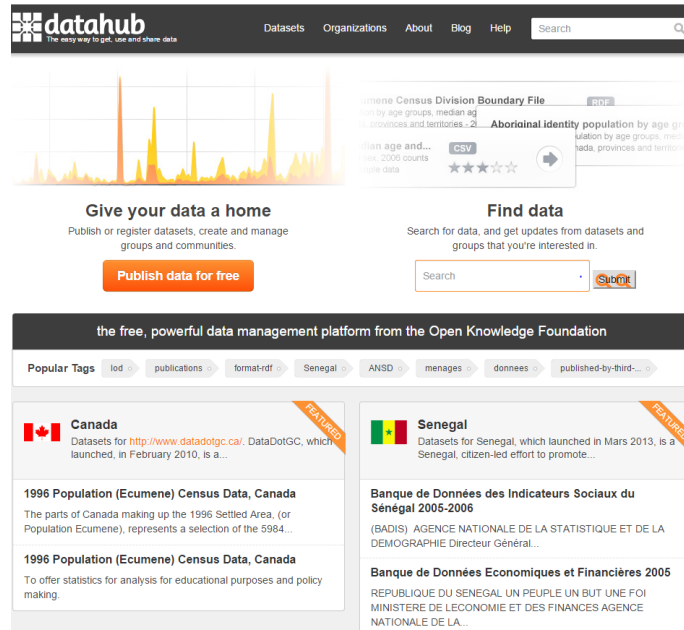
Fuente: Segittur. [En línea]. Disponible en Internet: <<http://www.segittur.es/>>

Un reto particular es de gestionar los múltiples conjuntos de datos que se liberan, en su mayoría desestructurados, es decir, no siguen estándares y se disponen en formatos propietarios o de difícil manejo. Son varias las iniciativas que se han desarrollado o se están desarrollando para gestionar los datos bajo el paradigma Linked Data, facilitando al usuario el acceso homogéneo a distintos servicios heterogéneos. Iniciativas de Linked data Turismo.

DataHub. Es un gran almacén de datos utilizado por Los gobiernos nacionales y locales, instituciones de investigación y otras organizaciones y empresas para la publicación de datos abiertos. El datahub ofrece libre acceso, lo que permite buscar datos, registrar datos publicados, crear y administrar grupos de conjuntos de datos y obtener actualizaciones de bases de datos y de los grupos que te interesan.

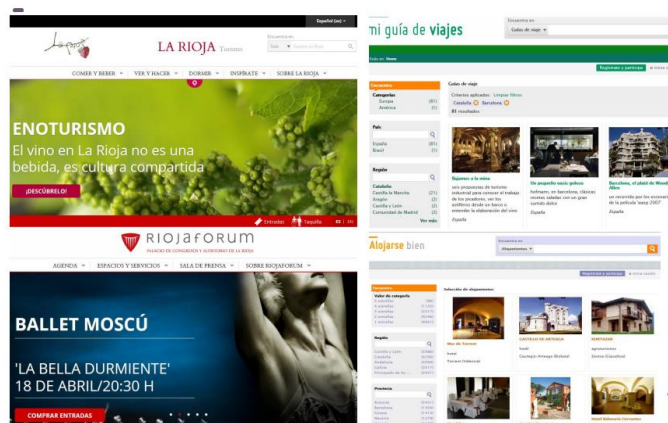
¹⁶ Clear Streets, visualize city's snow plows in real time. CartoDB, 2014

Figura 13. Almacén de Datos DataHub



Fuente: DataHub. [En línea]. Disponible en Internet: <<http://datahub.io/>>

Figura 14. Portal de Turismo de la Rioja



Fuente: Gnos

1.3.3 Ontologías Para Turismo. Existen una serie de especificaciones que han servido de base para el desarrollo de la mayoría de las ontologías para el turismo. Las dos más importantes son el tesoro de la Organización Mundial de Turismo (WTO) y la especificación de la Open Travel Alliance (OTA). La Organización

Mundial de Turismo (World Tourism Organization) ofrece un tesoro de conceptos turísticos y actividades de ocio para ayudar a la clasificación de la información relacionada con las actividades turísticas. La Open Travel Alliance (OTA) es una organización sin ánimo de lucro formada por organizaciones que representan a todos los segmentos de la industria del turismo (aerolíneas, empresas hoteleras, de alquiler de coches, etc.), junto a proveedores de tecnología y servicios. Su objetivo es facilitar la interoperabilidad en el desarrollo de software entre compañías del ámbito turístico. Su actividad principal es el desarrollo y mantenimiento de una biblioteca de esquemas XML para su uso en la industria del turismo. A partir de estas especificaciones existen numerosas ontologías en el dominio turístico, las más significativas son las siguientes: Harmonise¹⁷, QALL-ME Ontology¹⁸, DERI e-Tourism Ontology¹⁹, CDOTT²⁰, TAGA Travel Ontology²¹, GETESS Ontology²², Cruzar²³, Ontología ANOTA²⁴, etc.

1.3.4 Aplicaciones Móviles en turismo. Según el Informe Global sobre Contenidos para Móviles²⁵ realizado por Strategy Analytics, los ingresos globales obtenidos a través del móvil (smartphone & tablet) habrán alcanzado los 114.000 millones de euros en 2012, lo que significa un aumento del 17% respecto a 2011.

En el ecosistema móvil, las aplicaciones representan actualmente la segunda mayor fuente de ingresos por detrás del consumo de datos, tanto por pago de contenidos como por publicidad. En 2011 se descargaron más de 23.000 millones de aplicaciones en todo el mundo. En 2012 se estima un crecimiento del 38%, hasta alcanzar los 32.000 millones.

¹⁷ H. ONTOLOGYAN, The HarmoNET Ontology, 2011. http://www.ecca.at/harmonet/index.php?option=com_content&task=view&id=67&Itemid=63.

¹⁸ S. OU, V. PEKAR, C. ORASAN, C. SPURK, AND M. NEGRI, "Development and Alignment of a Domain-Specific Ontology for Question Answering," *Artificial Intelligence*, 2008, pp. 2221-2228.

¹⁹ K. SIORPAES, "OnTour System Design," *Tourism*, 2005. <http://etourism.deri.at/ont/>

²⁰ R. BARTA, C. FEILMAYR, B. PRÖLL, C. GRÜN, AND H. WERTHNER, "Covering the semantic space of tourism: an approach based on modularized ontologies," *Proceedings of the 1st Workshop on Context, Information and Ontologies*, 2009, pp. 1-8.

²¹ Y. ZOU, T. FININ, L. DING, H. CHEN, AND R. PAN, "Using semantic web technology in multi-agent systems: a case study in the taga trading agent environment," 2003.

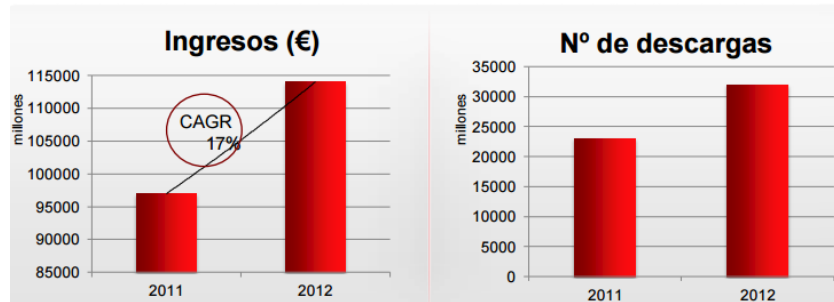
²² S. STAAB, C. BRAUN, I. BRUDER, A. DÜSTERHÖFT, A. HEUER, M. KLETTKE, G. NEUMANN, B. PRAGER, J. PRETZEL, H.-P. SCHNURR, R. STUDER, H. USZKOREIT, AND B. WRENGER, "GETESS - Searching the Web Exploiting German Texts," *Proceedings of the 3rd International Workshop on Cooperative Information Agents III/Workshop on Cooperative Information Agents*, London, UK: Springer-Verlag, 1999, pp. 113-124.

²³ I. MÍNGUEZ, D. BERRUETA, AND L. POLO, "CRUZAR: An Application of Semantic Matchmaking to e-Tourism."

²⁴ I. MURUA, E. LLADO, AND B. LLODR, "The Semantic Web for Improving Dynamic Tourist Packages Commercialisation," *Proceedings of the 13th ISPE International Conference on Concurrent Engineering (ISPE CE 2006)*.

²⁵ <https://www.strategyanalytics.com/>

Figura 15. Ingresos Globales en Aplicaciones



Fuente: Strategy analytics

Muchas de las actuales aplicaciones están enfocadas al sector turismo, como uno de los sectores de mayor crecimiento a nivel mundial, esto también obedece a maratones e iniciativas lideradas por muchas organizaciones, de las cuales podemos destacar:

Challenge Open Data Tourism Hack-at-Home²⁶. Es un proyecto enfocado a impulsar la creación de aplicaciones móviles para ayudar a las ciudades a gestionar los retos y beneficios que supone el turismo.

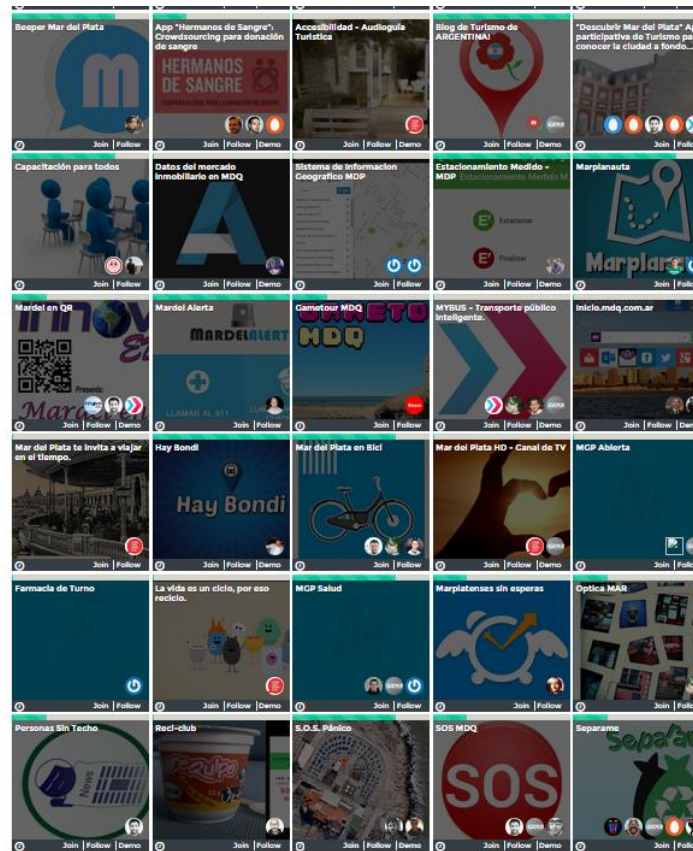
El proyecto, que se basa en los conceptos de open data y redes de sensores, permitirá que Ámsterdam, Barcelona, Berlín, Helsinki, París, Roma y Bolonia puedan beneficiarse del talento de los creadores de apps para encontrar soluciones en el campo del turismo. Uno de los principales retos es minimizar el impacto masivo que tiene el turismo para las ciudades y sus residentes. Para ello, el Open Data Tourism Hack-at-Home busca soluciones móviles que permitan conectar los residentes con los visitantes, personalizar las visitas turísticas, optimizar el tiempo de las visitas y mejorar la manera en que los turistas se mueven por la ciudad durante la estancia.

The App Date²⁷. Es un proyecto que incentiva el desarrollo de Apps, para diferentes sectores. Tiene presencia en 6 países, incluido Colombia donde se llevó un evento que sirvió además como plataforma para impulsar el emprendimiento a desarrollos de tecnologías accesibles para todos (incluyentes) en proyectos como Ayudapps de MinTIC.

²⁶ El Turismo Open Data hack-at-home. [En línea]. Disponible en <http://opencities.hackathome.com/>
²⁷ ¿Qué es the APP Date?. [En línea]. Disponible en Internet: <<http://www.theappdate.com/#step-1>>

Hackaton Ciudadano28. Iniciativa para crear Apps para innovar en la solución de problemas sociales.

Figura 16. Aplicaciones Apps Innovaciones sociales



Fuente: Hackaton Ciudadano

Proyectos que se han presentado en Santander-Colombia a través del programa Apps.co del Min TIC, donde se vienen incubando tres emprendimientos digitales²⁹ que le están apostando a ayudar a los empresarios del sector turismo a comercializar sus productos y a su vez, a mejorar la experiencia de los viajeros en Colombia.

²⁸ Primer Hackaton Ciudadano. [En línea]. Disponible en Internet: <<http://hackdash.org/dashboards/hcmdp2014>>

²⁹ Emprendimientos tic fortalecen a Colombia como destino turístico mundial. [En línea]. Disponible en Internet: <<http://appscounab.co/emprendimientos-tic-fortalecen-a-colombia-como-destino-turistico-mundial/>>

Ciudad. Travel. Es una plataforma Web y aplicación para realizar reservas de alojamiento y servicios de valor agregado como transporte de lujo y planes turísticos en una sola transacción ahorrando tiempo y dinero, enfocándose así a empresas, viajeros de negocios y ejecutivos.

Ummus. Aplicación que se ha convertido en aliado comercial de propietarios de viviendas turísticas rurales ubicadas en destinos únicos por su proximidad a sitios atractivos y experiencias en entornos naturales, brindando la posibilidad de reservar y pagar en línea a la hora de alquilar temporalmente un inmueble rural.

1.4 HERRAMIENTAS DE SOFTWARE LIBRE PARA EL DESARROLLO

El software open source es parte de un movimiento más amplio, casi una filosofía, que busca crear nuevos modelos de relación colaborando y compartiendo gracias a Internet. Por eso ahora más que nunca, en estos momentos de incertidumbre económica, está experimentando un crecimiento imparable. Le permite a las empresas un significativo ahorro de costes, gracias a que se evitan pagar licencias de programas, pudiendo disfrutar de las aplicaciones habituales en su alternativa SL.

Independencia tecnológica, al no estar atado al desarrollo propio de un fabricante. En el caso del SL existen comunidades robustas de desarrolladores, que garantizan la continuidad de los programas.

Innovación, al disponer de libertad de aportar con desarrollos propios al resto de la comunidad, lo que ofrece un beneficio global a todos.

1.4.1 Herramienta para la implementación de Open Data (CKAN³⁰). es la plataforma de datos, de código abierto, líder a nivel mundial. Es una solución completa de software lista para utilizar que hace los datos accesibles y utilizables al proveer herramientas para publicar, compartir, encontrar y usar los datos (incluyendo almacenamiento de datos y provisión de APIs de datos robustas). CKAN está orientada a proveedores de datos (gobiernos nacionales y regionales, compañías y organizaciones) que desean hacer sus datos abiertos y disponibles.

CKAN es utilizada por gobiernos y grupos de usuarios a nivel mundial y gestiona

³⁰ Acerca de. [En línea]. Disponible en Internet: <<http://beta.ckan.org/es/about>>

una variedad de portales de datos oficiales y comunitarios, incluyendo portales para gobiernos locales, nacionales e internacionales tales como data.gov.uk de Reino Unido, publicdata.eu de la Unión Europea; dados.gov.br de Brasil; además portales de los gobiernos de Dinamarca y Holanda, así como también sitios de ciudades y municipalidades en Estados Unidos, Reino Unido, Argentina, Finlandia y en otros lugares.

Tecnologías Usadas por CKAN

CKAN usa las siguientes tecnologías:

Postgresql³¹: es un Sistema de gestión de bases de datos relacional orientado a objetos y libre, publicado bajo la licencia BSD.

SqlAlchemy³²: Conjunto de herramientas y mapeador objeto-relacional (ORM) para el lenguaje de programación Python publicado bajo la licencia MIT³³.

Apache Solr³⁴: Es un motor de búsqueda de código abierto basado en la biblioteca Java del proyecto Lucene, con APIs en XML/HTTP y JSON, resaltado de resultados, búsqueda por facetas, caché, y una interfaz para su administración

Backend en Python³⁵: Es un lenguaje de programación interpretado cuya filosofía hace hincapié en una sintaxis que favorezca un código legible. Se trata de un lenguaje de programación multiparadigma, ya que soporta orientación a objetos, programación imperativa y, en menor medida, programación funcional.

Cuenta con FrontEnd desarrollado en CSS3, HTML5 y Java Script.

1.4.2 Herramienta Implementación Linked Data (Virtuoso Universal Server³⁶).

Virtuoso es una plataforma creada por Open Link Software concebida para la

³¹ Postgresql. En línea]. Disponible en Internet: <<http://www.postgresql.org.es/>>

³² Sqlal. En línea]. Disponible en Internet: <<http://www.sqlalchemy.org/>>

³³ Opensource. En línea]. Disponible en Internet: <<http://opensource.org/licenses/MIT.>>

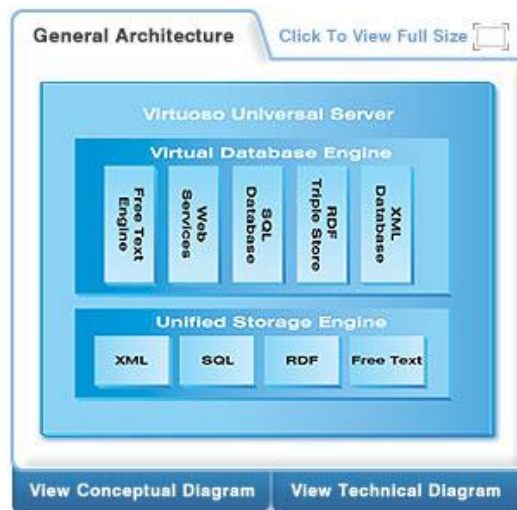
³⁴ Lucene. Apache. En línea]. Disponible en Internet: <<http://lucene.apache.org/solr/>>

³⁵ Python. En línea]. Disponible en Internet: <<https://www.python.org/>>

³⁶ Virtuoso. En línea]. Disponible en Internet: <<http://virtuoso.openlinksw.com/>>

integración de datos multi-modelos para empresas o particulares. Es una solución para gestión de procesos de negocio que implican SQL, RDF, XML y servicios web, de forma combinada. La arquitectura híbrida revolucionaria del servidor permite comunicar diferentes funcionalidades de tradicionalmente distintos servidores, dentro de un producto único, a través de las siguientes características:

Figura 17. Arquitectura general de Virtuoso Universal Server



Fuente: OpenLink.

- Gestión e integración de datos (SQL, XML y RDF).
- Integración de aplicaciones (servicios web y SOA).
- Gestión e integración de procesos (BPEL).
- Aplicaciones distribuidas colaborativas.

Y cubriendo las siguientes áreas:

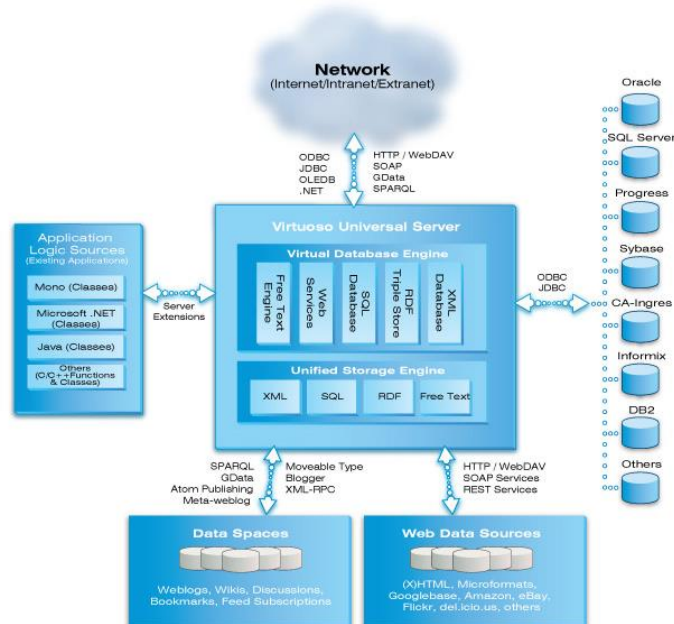
- Modelo Relacional.
- RDF.
- XML.
- Gestores de contenido de texto libre o indexado.
- Web Server de documentos.
- Servidor de datos vinculados.
- Web Application Server.
- Despliegue de servicios web (SOAP o REST).

Su motor de bases de datos provee conexión a fuentes XML, ODBC, JDBC, ADO .NET y OLE DB. A su vez, la biblioteca Virtuoso PL proporciona funciones para trabajar con distintos protocolos de Internet para clientes y servidores (POP3, FTP, NNTP, cliente LDAP o cliente SMTP).

Virtuoso es un motor que une orígenes heterogéneos permitiendo seguridad, transacciones (incluyendo commit de dos fases), y replicación. Su motor de consulta soporta vistas heterogéneas, procedimientos almacenados, cursors, y búsquedas de texto. Es capaz de acceder a orígenes externos a través de ODBC, así como a su propio modelo relacional.

Virtuoso suporta XML en un gran número de formas. Primeramente, contiene un almacenamiento nativo de XML, el cual no es relacional, y puede almacenar e indexar documentos XML, parseándolos o sin parsear. Además, es capaz de transferir datos desde base de datos relacionales hacia documentos XML, usando el mismo mapeo disponible en la cláusula FOR XML de Microsoft SQL Server. Por último, incluye una implementación de XPath para las consultas (que también puede ser usado sobre los datos relacionales si se transfieren a XML previamente). Además posee soporte de XSLT, ejecución de procedimientos almacenados a través de SOAP y WebDAV.

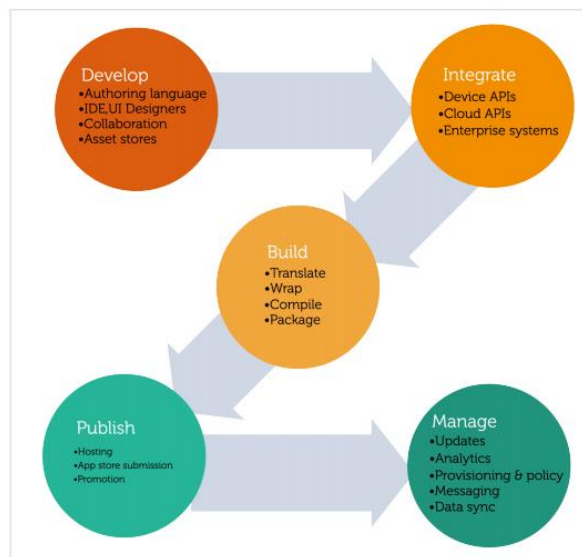
Figura 18. Arquitectura detallada de Virtuoso Universal Server



Fuente: OpenLink.

1.4.3 Entorno para el desarrollo de Aplicaciones Móviles. Según un estudio de visionmobile³⁷, los últimos años marca un punto de inflexión en las plataformas móviles. IOS de Apple y la plataforma Android de Google se han impulsado a alturas inimaginables, a partir de cientos de miles de desarrolladores. A raíz de la caída de herramientas como adobe Flash Lite y Java ME, aparecen una serie de herramientas multiplataforma de las cuales proporcionan los componentes necesarios durante las etapas de desarrollo. Una herramienta puede tener cinco componentes, que corresponden a las cinco etapas en la aplicación del ciclo de vida: desarrollar, integrar, publicar, desplegar y gestionar.

Figura 19. Las cinco etapas del ciclo de vida de una aplicación Multiplataforma



Fuente: Cross Platform Tools. [En línea]. Disponible en Internet: <<http://www.visionmobile.com/product/cross-platform-developer-tools-2012/>>

Adobe PhoneGap (Apache Córdoba³⁸). PhoneGap, una de las herramientas de desarrollo multiplataforma más populares, desarrollada por Nitobi, una empresa de desarrollo web y aplicaciones en Vancouver, Canadá, con alrededor de 20 empleados. El proyecto PhoneGap nació en un evento iPhone Dev Camp en 2008 como una forma de utilizar las habilidades de desarrollo web para ofrecer aplicaciones nativas. En 2009, ganó el concurso de Phone Gap Expo Launch Pad

³⁷ Vision Mobile. Cross-platform developers tolos 2012. 2013. [En línea]. (consultado el 18-05-2015). Disponible en Internet: <<http://www.visionmobile.com/product/cross-platform-developer-tools-2012/>>.

³⁸ <https://cordova.apache.org/>

Web 2.0, y el código fue liberado bajo una licencia permisiva de código abierto (MIT).

Tras la adquisición de Nitobi por parte de Adobe en octubre de 2011, el proyecto fue renombrado a Apache Córdoba, y el código fue donado a la Apache Software Foundation. Los socios y los principales contribuyentes al proyecto incluyen IBM, Sony Ericsson, Symbian, Palm, RIM y Sencha.

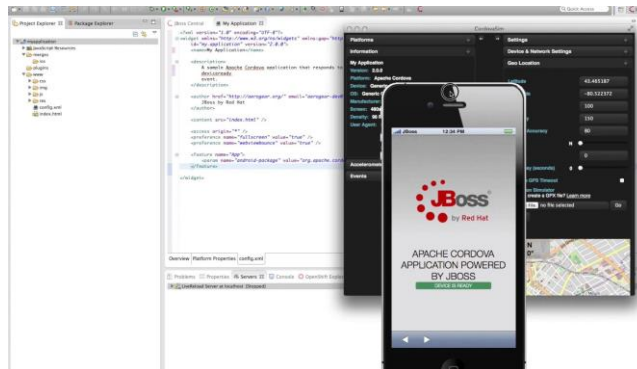
PhoneGap está dirigida a los desarrolladores web que deseen crear aplicaciones para teléfonos inteligentes nativos (iOS, Android, Blackberry, Symbian, Bada, Windows) y distribuirlos a través de las tiendas de aplicaciones nativas. Las aplicaciones PhoneGap son, construidas en, HTML, CSS y JavaScript desarrolladas en entornos como Dreamweaver, Netbeans, Jboss o Eclipse. El desarrollador puede utilizar JavaScript para llamar a las API de la plataforma nativa. El proceso de construir paquetes el código web, junto con la biblioteca PhoneGap, en un envoltorio de ejecución de la aplicación nativa. Las aplicaciones PhoneGap son por lo tanto aplicaciones web híbridas.

Jboos Developer Studio. JBoss Developer Studio³⁹ (JBDS) es un entorno de desarrollo creado y actualmente desarrollado por JBoss (una división de Red Hat) y Exadel. Integra y certifica tanto herramientas y componentes de tiempo de ejecución mediante la combinación de Eclipse, Eclipse Tooling, y la plataforma de aplicaciones JBoss Enterprise.

Las herramientas de desarrollo integradas se utilizan para construir aplicaciones web ricas utilizando código abierto tecnologías como JBoss Seam , JBoss Application Server , Hibérnate y JBoss jBPM. Esta herramienta permite el desarrollo de aplicaciones híbridas usando Apache Córdoba.

³⁹ <http://www.jboss.org/>

Figura 20. Jboss y Apache Córdoba



Fuente: Jboss.

Git⁴⁰. Git es un sistema de control de versiones distribuido gratuito y de código abierto diseñado para manejar todo, desde pequeños a grandes proyectos con rapidez y eficiencia. Git es fácil de aprender y tiene un tamaño compacto con un rápido rendimiento. Supera herramientas SCM como Subversión, CVS, Perforce y ClearCase con características como bifurcaciones locales, y múltiples flujos de trabajo.

Node JS⁴¹. Node.js es un entorno de programación en la capa del servidor, basado en el lenguaje de programación ECMAScript, asíncrono, con I/O de datos en una arquitectura orientada a eventos y basado en el motor V8 de Google. Fue creado con el enfoque de ser útil en la creación de programas de red altamente escalables, como por ejemplo, servidores web. Fue creado por Ryan Dahl en 2009 y su evolución está apadrinada por la empresa Joyent.

Android. Android⁴² es un sistema operativo basado en el núcleo Linux. Fue diseñado principalmente para dispositivos móviles con pantalla táctil, como teléfonos inteligentes o tablets; y también para relojes inteligentes, televisores y automóviles. Inicialmente fue desarrollado por Android Inc., empresa que Google respaldó económicamente y más tarde, en 2005, compró. Android combina las siguientes cualidades:

⁴⁰ <https://git-scm.com/>

⁴¹ <https://nodejs.org/>

⁴² <https://www.android.com/>

Plataforma realmente abierta. Es una plataforma de desarrollo libre basada en Linux y de código abierto.

Adaptable a cualquier tipo de hardware. Android no ha sido diseñado exclusivamente para su uso en teléfonos y tabletas. Hoy en día podemos encontrar relojes, cámaras, electrodomésticos y gran variedad de sistemas empotrados que se basan en este sistema operativo.

Portabilidad asegurada. Las aplicaciones finales son desarrolladas en Java lo que nos asegura que podrán ser ejecutadas en cualquier tipo de CPU, tanto presente como futuro. Esto se consigue gracias al concepto de máquina virtual.

Arquitectura basada en componentes inspirados en Internet. Por ejemplo, el diseño de la interfaz de usuario se hace en XML, lo que permite que una misma aplicación se ejecute en un móvil de pantalla reducida o en un TV.

Filosofía de dispositivo siempre conectado a Internet

Gran cantidad de servicios incorporados. Por ejemplo, localización basada tanto en GPS como en redes, bases de datos con SQL, reconocimiento y síntesis de voz, navegador, multimedia.

Aceptable nivel de seguridad. Los programas se encuentran aislados unos de otros gracias al concepto de ejecución dentro de una caja que hereda de Linux.

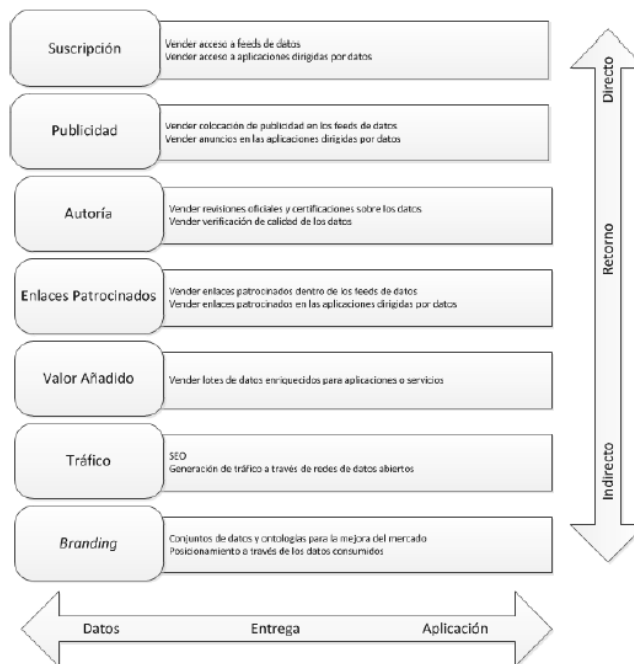
Optimizado para baja potencia y poca memoria. Android utiliza la Máquina Virtual Dalvik. Se trata de una implementación de Google de la máquina virtual de Java optimizada para dispositivos móviles.

Alta calidad de gráficos y sonido. Gráficos vectoriales suavizados, animaciones inspiradas en Flash, gráficos en 3 dimensiones basados en OpenGL. Incorpora codecs estándar más comunes de audio y vídeo, incluyendo H.264 (AVC), MP3, AAC, etc.

1.5 MODELO DE NEGOCIO

La conjunción de los principios de Linked Data y Open Data ha conllevado una corriente de actuación en las Administraciones Públicas con amplios beneficios tanto para los ciudadanos como para las empresas. La experiencia inicial es altamente enriquecedora para las partes implicadas y su tendencia es seguir creciendo. No obstante, también se presentan varias cuestiones abiertas que se habrá de ir resolviendo a la vez que la tecnología y la experiencia se incrementan.

Figura 21. Modelos de Negocio para Datos

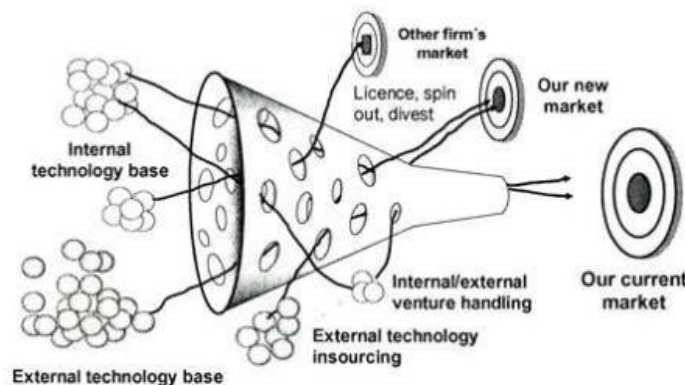


1.5.1 Innovación Abierta. Es necesario adaptar los sistemas de innovación organizacionales a los nuevos tiempos. Comprendiendo los cambios sociales, tecnológicos y principalmente el crecimiento del conocimiento y su democratización al mundo, se podría resumir que estamos ante una nueva sociedad del conocimiento que se aproxima a cambiar un paradigma “tradicional” que lleva años como líder, y que al parecer ahora está quedando atrás ante la rapidez con la cual se pueden ejecutar cualquiera de las etapas, así podría comenzar a quedar atrás la individualización de las empresas para innovar, para dar la bienvenida a un proceso abierto y

permeable al compartir conocimiento con otras personas.

Uno de los primeros investigadores que acuñó el término innovación abierta u “open innovation” fue Henry Chesbrough, director ejecutivo del Centro de Open Innovation de la Universidad de California-Berkeley. El profesor Chesbrough mantiene en que los modelos monolíticos de I+D restringen el flujo del capital intelectual de la organización, limitando las oportunidades para convertir en dinero dicho flujo de conocimiento.

Figura 22. Representación del proceso de innovación abierta



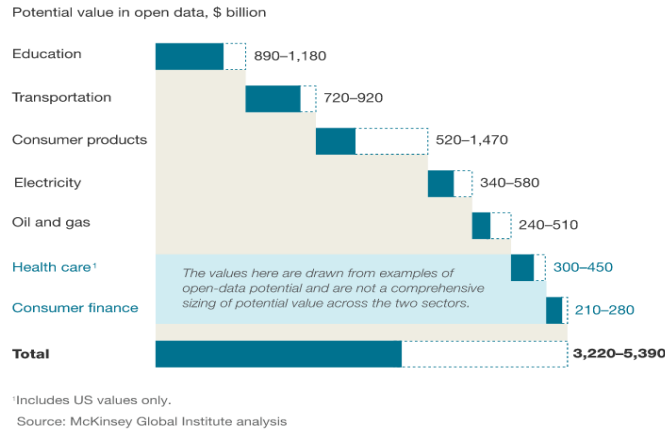
Fuente:

http://cvb.ehu.es/open_course_ware/castellano/social_juri/herrami_gestion/innovacionabierta.pdf

1.5.2 Innovación Abierta y Open Data. McKinsey Global Institute⁴³ publicaba hace unos días las conclusiones de su estudio sobre Open Data: la existencia de datos abiertos permiten mejorar la prestación de los servicios, la transparencia de los gobiernos, y las posibilidades de las empresas. Algo que, además, puede medirse en cifras: un incremento de 3 billones de dólares solo en Estados Unidos en el valor de siete sectores clave para nuestras sociedades: la educación, el transporte, los productos de consumo, la electricidad, el combustible y el gas, la medicina, y el sector financiero.

⁴³http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information?cid=other-eml-alt-mgi-mck-oth-2910. (consultado el 18-05-2015).

Figura 23. Potencial del open data



Cabe destacar que, de acuerdo al estudio, el uso de datos abiertos muestra un potencial económico casi inimaginable, que tiende a una mejora en la eficiencia de las sociedades, con nuevos productos y servicios adaptados a los ciudadanos, y unas mejoras notables respecto a la era actual, traducidas en un mayor ahorro en costes para las empresas, una mayor comodidad para los usuarios, y una mayor sostenibilidad en las ciudades.

1.5.3 Ideas en el Ecosistema Open Data. Se describe a continuación

Impulsar la Innovación. Las plataformas de datos abiertos pueden generar nuevas ideas, proyectos de investigación o aplicaciones innovadoras. Para activar esta parte, se debe apoyar en un buen plan de comunicación, incluyendo actividades como desafíos, hackatons o actividades de exploración de datos que puede fomentar la participación de grupos de usuarios que puedan descubrir los aspectos relevantes de los datos publicados. Estos usuarios pueden ser desarrolladores, científicos de datos, analistas, economistas, periodistas, gerentes, empresarios o personas con diferentes habilidades para analizar los datos y que pueden tener conocimiento sobre algunos de los conjuntos de datos publicados. De este modo, impulsar la innovación no es fácil y sencillo, ya que requiere esfuerzos importantes y una fuerte implicación de los principales actores que administran el ecosistema, pero al mismo tiempo de los usuarios como consumidores. En esta etapa se pueden apoyar de programas de innovación e emprendimiento como Apps.co⁴⁴.

⁴⁴ <https://apps.co/>

También existen otras iniciativas como por ejemplo European project Open Cities⁴⁵, algunos eventos y desafíos como (Hackatons) se organizaron, para estimular el proceso de difusión y adopción de la plataforma entre la comunidad de desarrolladores. Entre ellos, el turismo “Open Data Tourism Hack at Home⁴⁶” el cual fue organizado para ayudar a las ciudades europeas a encontrar nuevas formas de gestionar los grandes retos y beneficios de turismo hoy día⁴⁷. Los resultados fueron sorprendentes, con 21 aplicaciones de comunicaciones y 10 finalistas. La aplicación ganadora fue Moovit⁴⁸, que es la aplicación crowdsourced del mundo de más rápido crecimiento para el transporte público, con cerca de 2 millones de usuarios en más de 40 ciudades. Desafíos similares se pueden encontrar en la página web de Appcircus⁴⁹, como Philippine Transit,, BCN Apps&Cultura o El Cairo Transport App Challenge.

Figura 24. Moovit, aplicación croad sourced para el transporte público



Fuente: Moovit. En línea]. Disponible en Internet: <http://appcircus.com/apps/moovit-1>.

⁴⁵ es un proyecto co-fundada por la Unión Europea que tiene como objetivo validar la forma de bordar Abrir e impulsar la innovación. <http://opencities.net/>.(consultado el 18-05-2014).

⁴⁶ <http://opencities.hackathome.com/>

⁴⁷ <http://ckan.opendata.nets.upf.edu/organization/about/hackathome>

⁴⁸ <http://appcircus.com/apps/moovit-1>

⁴⁹ AppCircus is an international showcase for the world's latest apps. Emerging and established developers pitch their apps for a rare opportunity to compete for the best app at Mobile Premier Awards. <http://appcircus.com/challenges>. (consultado el 18-05-2014).

Teniendo en cuenta una amplia gama de usuarios, el proyecto Urban Data Challenge⁵⁰, cuyo objetivo son programadores, artistas, hackers y científicos de datos, para crear e inspirar visualizaciones basadas en datos abiertos en tiempo real. Un ejemplo similar es Visualizing.org⁵¹, que es una comunidad que ofrece canales de datos y organiza retos para crear visualizaciones basadas en datos abiertos.

Focos de Negocio. Las ideas que se generan dentro del ecosistema de los datos abiertos para el sector turismo, pueden dar lugar al desarrollo de ciertos modelos de negocio que pueden generar valor económico en los datos proporcionados. Por ejemplo, los nuevos productos o servicios, tales como una aplicación móvil en el mercado o un proyecto que pueden ser elegibles para la financiación pública. En el sector turismo, estas ideas pueden surgir en un espacio de colaboración entre estudiantes, profesores, investigadores y la comunidad empresarial. La interacción entre el desarrollo de nuevas ideas y los resultados de negocio, representa el enfoque de la innovación abierta en el turismo.

1.5.4 Modelo de Negocio Canvas. El Business Model Canvas, es una herramienta desarrollada por el consultor suizo Alexander Osterwalder en el año 2004 para su tesis doctoral para llegar a lo que hoy conocemos como método Canvas, que describe de manera lógica la forma en que las organizaciones crean, entregan y capturan valor. La metodología Canvas parte de la idea de la existencia de una propuesta de valor, la cual debe ser comunicada a los potenciales clientes. Esto conllevará la necesaria disponibilidad de recursos y obligará a establecer relaciones con agentes externos e internos.

Módulos Desarrollados

Segmentos de clientes. El objetivo es de agrupar a los clientes con características homogéneas en segmentos definidos y describir sus necesidades, averiguar información geográfica y demográfica, gustos, etc.

Propuestas de valor. El objetivo es de definir el valor creado para cada Segmento de clientes describiendo los productos y servicios que se ofrecen a cada uno. Para cada propuesta de valor hay que añadir el producto o servicio más importante y el nivel de servicio. Estas primeras dos partes son el núcleo del modelo de negocio.

⁵⁰ <http://urbandatachallenge.org/>

⁵¹ <http://www.visualizing.org/about>

Canales. Se resuelve la manera en que se establece contacto con los clientes.

Relación con el cliente. Aquí se identifican cuáles recursos de tiempo y monetarios se utiliza para mantenerse en contacto con los clientes.

Fuentes de ingresos. Este paso tiene como objetivo identificar que aportación monetaria hace cada grupo y saber de dónde vienen las entradas (ventas, comisiones, licencias, etc.).

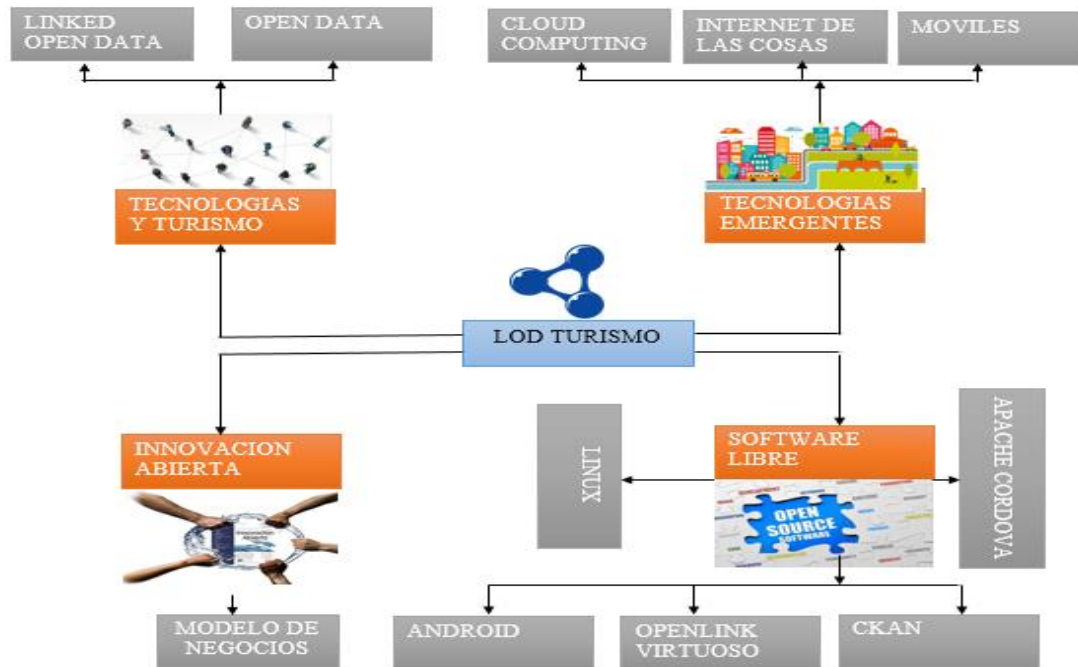
Recursos clave. Después de haber trabajado con los clientes, hay que centrarse en la empresa. Para ello, hay que utilizar los datos obtenidos anteriormente, seleccionar la propuesta de valor más importante y la relacionarse con el segmento de clientes, los canales de distribución, las relaciones con los clientes, y los flujos de ingreso.

Actividades clave. En esta etapa es fundamental saber qué es lo más importante a realizar para que el modelo de negocios funcione.

Asociaciones claves. En este apartado se describe a los proveedores, socios, y asociados con quienes se trabaja para que la empresa funcione.

Estructura de costos. Aquí se especifican los costos de la empresa empezando con el más alto. Luego se relaciona cada costo con los bloques definidos anteriormente, evitando generar demasiada complejidad.

Figura 25. Mapa Conceptual Marco Teórico



Fuente: Elaboración propia

1.6 ESTADO DEL ARTE

1.6.1 La Web Semántica. Con el transcurrir del tiempo se puede interpretar el término de la “web Semántica”, como una evolución de la web actual, para Gartner. Ya se vaticinaba que en el año 2005 la integración de aplicaciones empresariales estaría guiada por el uso de ontologías y la Web Semántica. La Web Semántica según la definición propuesta por el W3C se presenta como:

“Una web extendida, dotada de mayor significado, en la que cualquier usuario en Internet podrá encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida”.

La web semántica aparece para dar solución a las dificultades de la web actual. El contenido de la web actual esta creado para ser consumidos por personas, lo que dificulta que los agentes software puedan encontrar, interpretar, extraer y combinar la información ya disponible en la web. Y aquellas tareas que parecían sencillas son difíciles de implementar. Tim Berners-Lee, impulso a través del W3C

el desarrollo de la Web Semántica. Con esta iniciativa, se desarrollan los estándares, lenguajes que permitan la actualización de la web y proponer que los contenidos sean accesibles tanto a personas como a máquinas. A partir de esta nueva tecnología abre la posibilidad al desarrollo de aplicaciones que puedan desarrollar aplicaciones de forma automática que antes se realizaban de forma manual, como darle valor a la información de diferentes fuentes.

En la web actual es mucha la información que se encuentra publicada y una gran cantidad de servicios que se pueden acceder. Pero si explotar estos datos y servicios requiere de la intervención humana, se perdería las grandes oportunidades para realizar productos y emprendimientos es por eso que el consorcio W3C pretenda “guiar a la web hacia su máximo potencial” como herramienta universal y multipropósito. A través de la Web Semántica se permite explotar eficientemente el potencial de la web. Se han desarrollado estándares que son el soporte necesario para hacer realidad la misma. En la base se utilizan tecnologías estándar, como XML⁵², HTTP⁵³ y URIs⁵⁴, que han sido reutilizadas en la web actual. Pero se han creado nuevos componentes como, el RDF⁵⁵, RDF Schema⁵⁶ y OWL⁵⁷. Con estos lenguajes y modelos se pueden describir los datos de manera precisa y coherente.

Ontologías de la web semántica. La evolución de la web permite un mejor manejo en los datos y por ende la información. Actualmente existen grandes problemas de la web por la gran cantidad de información existente y la falta de mecanismos, control y procedimientos para la estructuración de los datos, el cual no permite que se aproveche completamente.

Las ontologías son una forma para organizar los conocimientos y contenido heterogéneo en el ámbito de la Web Semántica, las ontologías se utilizan como instrumentos para el tratamiento del lenguaje natural. Como mecanismo de comunicación entre los agentes de software. Con base en modelos de conocimiento.

⁵² [Http://www.w3.org/TR/xml11/](http://www.w3.org/TR/xml11/).

⁵³ Tim Berners-Lee, Roy Fielding, y Larry et al. Masinter. RFC 2616, Hypertext Transfer Protocol – HTTP/1.1, 1999.

⁵⁴ Tim Berners-Lee, Roy Fielding, y Larry Masinter. RFC 3986, Uniform Resource Identifier (URI): Generic Syntax. Request For Comments (RFC), 2005

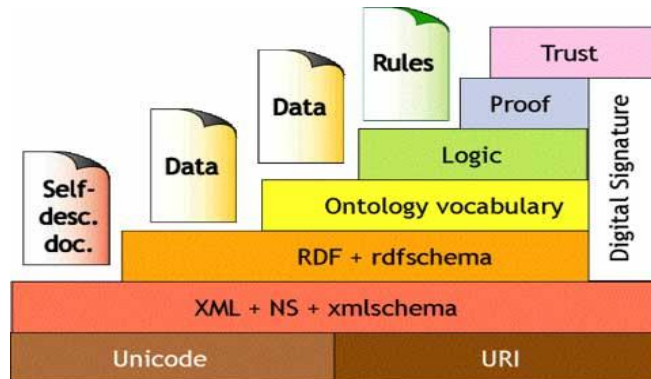
⁵⁵ G. Klyne y JJ. Carroll. Resource Description Framework (RDF): Concepts and Abstract Syntax. W3C recommendation, W3C, 2004. [Http://www.w3.org/TR/rdf-concepts/](http://www.w3.org/TR/rdf-concepts/).

⁵⁶ Dan Brickley y R.V. Guha. RDF Vocabulary Description Language 1.0: RDF Schema. W3C recommendation, W3C, 2004.

⁵⁷ Pascal Hitzler, Markus Krötzsch, Bijan Parsia, Peter F. Patel-Schneider, y Sebastian Rudolph. OWL 2 Web Ontology Language: Primer. Informe técnico, 2009. <http://www.w3.org/TR/owl2-primer/>.

Arquitectura para la Web Semántica. La infraestructura de la Web Semántica comprende un conjunto de tecnologías y lenguajes que se organizan en varias capas o niveles según lo que se ha denominado la Pila de la Web Semántica, como se muestra en la Figura 26.

Figura 26. Pila de la Web semántica



Fuente: W3C

Tecnologías de la web Semántica

1.6.2 Linked Open Data: Ideas y Principios. La idea de Linked Open Data (LOD) surgió en 2007, cuando se inició el proyecto Linking Open Data⁵⁸. Linked Open Data es sinónimo de los datos publicados en la web en un formato abierto legible para una máquina y que está interrelacionada con otras fuentes de datos. La existencia de LOD es un requisito muy importante para la realización de una visión global de la Web Semántica.

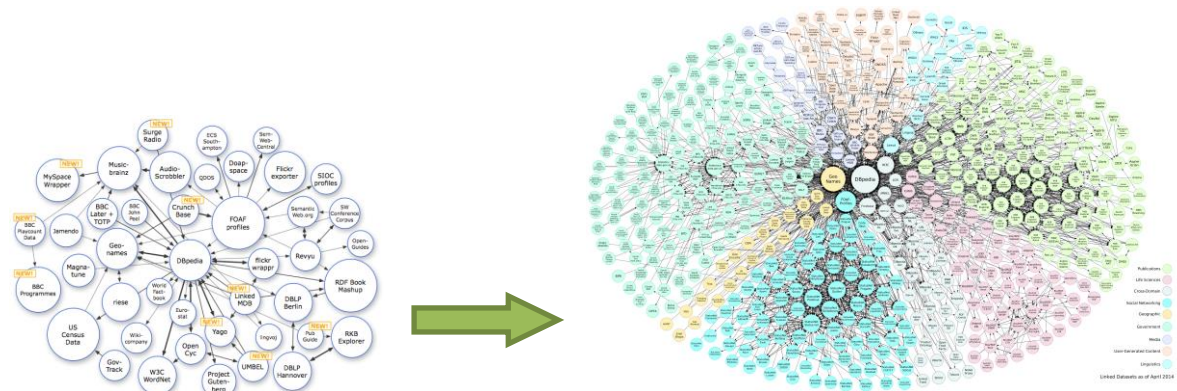
En 2008, en un artículo publicado por Christian Bizer⁵⁹ describe que la web se entiende como un espacio de información global que consiste no sólo de documentos vinculados, sino también de datos enlazados. Y que es el resultado de la maduración de los estándares de la tecnología de la Web Semántica, y por la creciente publicación de conjuntos de datos de acuerdo con los principios de Linked Data.

⁵⁸ W3C LOD Project. <http://esw.w3.org/topic/sweoig/taskforces/communityprojects/linkingopendata>, 2007. (consultado el 18-05-2014).

⁵⁹ Christian Bizer, Tom Heath, Kingsley Idehen, and Tim Berners-Lee. 2008. Linked Data on the Web. (consultado el 18-05-2014).

Una definición de Linked Data la dio Tim Berners-Lee⁶⁰, como parte de la Web Semántica. Surge con el fin de integrar grandes bases de datos a través de modelos de conocimientos compartidos u ontologías. Aunque existía fundamento teórico suficiente, en la práctica las aplicaciones eran muy escasas, debido a la dificultad para su implementación⁶¹. Es por ello, que tomando la visión de la Web Semántica en la parte de modelos y formatos de datos estandarizados y formalizados a través de conocimiento compartido, fue tomada para definir lo que actualmente se conoce como Linked Data. A través del cual se hace un manifiesto práctico de la Web Semántica para su implementación a la realidad, y se hace uso de los fundamentos de la semántica y de la web para proveer mecanismos para la publicación, y el consumo de datos y el desarrollo de aplicaciones. El problema de la web tradicional fue que se diseñó para enlazar documentos, pero no la estructura de los mismos y en la web semántica se diseñó la estructura de los datos pero no la forma como estos se deberían enlazar. Por lo tanto la web de datos no se centra tan sólo en publicar datos sino en hacer que los datos estén accesibles, tanto a las personas como a las máquinas.

Figura 27. Expansión de LOD cloud del 2008 al 2014



Fuente: University of Mannheim. En línea]. Disponible en Internet: <http://data.dws.informatik.uni-mannheim.de/lodcloud/>.

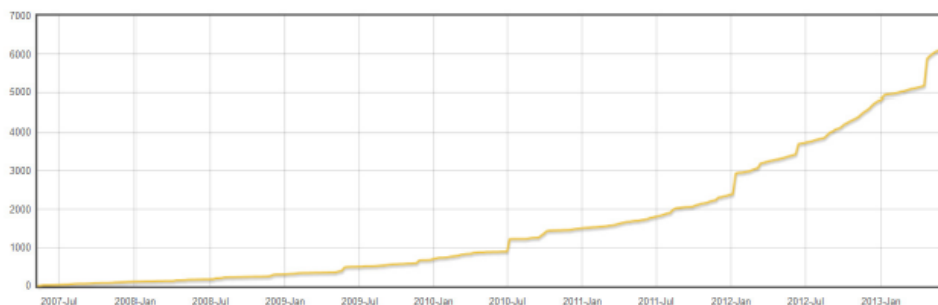
1.6.3 La Nube de Datos Abiertos Enlazados. Hay un crecimiento masivo de los datos publicados como LOD en los últimos años. Este desarrollo es visible en la nube LOD, un diagrama de datos enlazados en la web. Ha pasado de ser un

⁶⁰ Tim Berners-Lee, Wendy Hall, James Hendler, Kieron O'Hara, Nigel Shadbolt, y Daniel Weitzner. 2006. A Framework for Web Science. Foundations and Trends in Web Science. <http://www.nowpublishers.com/article/Details/WEB-00>. (consultado el 18-05-2014).

⁶¹ W3C. Semantic Web Case Studies and Use Cases. 2011. <http://www.w3.org/2001/sw/sweo/public/UseCases/s>. (consultado el 18-05-2014).

pequeño conjunto de fuentes de datos en un gráfico altamente interconectado y muy extenso. El desarrollo se representa en la figura 28. El crecimiento de los datos vinculados también es visible en un informe estadístico de la Data Hub [49], una plataforma de gestión de datos que permite a los proveedores de contenido registrar sus datos.

Figura 28. El crecimiento de los conjuntos de datos LOD en el Data Hub.



Fuente: The Data Hub. En línea]. Disponible en Internet: <http://datahub.io/stats>.

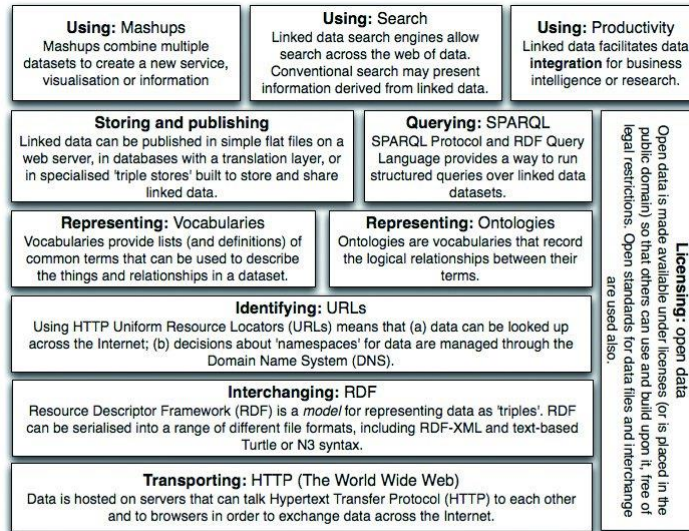
1.6.4 Tecnologías Open Linked Data. Se basa principalmente en las tecnologías de la web semántica. No existe una definición perfecta si un conjunto de datos publicada califica como Linked Open Data o no, en lugar existen varias normas y directrices que definen los requisitos que deben existir para los datos enlazados en la web. En muchos casos, los datos vinculados en la web se publican en varios formatos al mismo tiempo, como por ejemplo la DBpedia⁶².

Tim Davies creó una pila de tecnologías LOD⁶³ que es similar a la torta de capa de la Web Semántica. Se representa en figura 29.

⁶² <http://wiki.dbpedia.org/>

⁶³ Linked Open Data Stack. 2010. <http://www.opendataimpacts.net/2011/05/whats-in-the-linkedopen-data-stack/>, 2010. (consultado el 18-05-2014).

Figura 29. Pila de Tecnologías LOD



Elements of the Linked Open Data Puzzle (revision 2) - 2nd May 2011. CC BY-SA-NC
 Draft sketch by Tim Davies (@timdavies / tim@practicalparticipation.co.uk) for IKM Working Paper on Linked Open Data for Development. Comments welcome. Search 'linked open data puzzle' on <http://www.opendataimpacts.net> for latest version.

Idea based on Semantic Web Stack at http://en.wikipedia.org/wiki/Semantic_Web_Stack

Fuente: Platform Linked Data Nederland. En línea]. Disponible en Internet: <http://www.pilod.nl/wiki/Boek/BauerKaltenbock>

Esta pila LOD es una buena referencia para la comprensión de cómo LOD debe implementarse ya que muestra las tecnologías que se requieren y su relación. La World Wide Web (HTTP) sirve de partida para servir documentos LOD. Puesto que se utiliza HTTP, URLs (URIs) que definen las ubicaciones y recursos. RDF se utiliza para intercambiar datos y las ontologías se utilizan para describir esquemas.

A diferencia de la capa de la Web Semántica, XML no es un requisito y se menciona explícitamente que hay otras maneras para serializar datos RDF. Existen varios vocabularios en la web (a menudo en forma de una ontología) que se pueden volver a utilizar al desarrollar LOD. También se aborda el tema de almacenamiento de datos, donde se mencionan los archivos planos RDF junto triplestores. Sin embargo, la consulta de datos a través de SPARQL por lo general requiere la existencia de un triplestore con un punto final SPARQL.

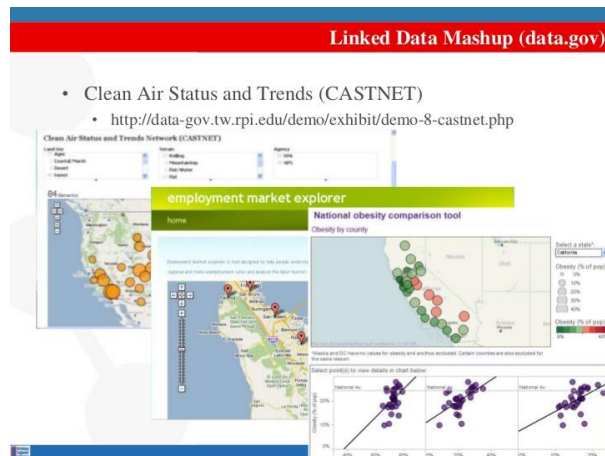
Puesto que la pila LOD es sólo para datos, en la capa lógica, no se relaciona los mecanismos de confianza. Para esto se introduce entonces una capa de "licencia", que pone de manifiesto la importancia de fijar una licencia abierta para LOD. Finalmente existe una capa de aplicación que introduce escenarios de uso.

1.6.5 Escenarios para Linked Open Data. Se describe a continuación

Mashups. Los Mashups combinan varias fuentes de datos para crear un conjunto de datos integrados. El resultado de los datos integrados es más útil que las fuentes de datos aisladas. Un ejemplo de un mashup es un mapa en línea que combina con una base de datos de imágenes. Si se conocen las posiciones geográficas de imágenes que se pueden visualizar en el mapa. Cuando se crearon los portales 2,0 surgieron muchas Web que ofrecen de datos generados por los usuarios, los mashups son cada vez más importantes. Para esto es importante la calidad de los datos ya que desde conjuntos de datos tienen que ser enlazados con alta precisión para crear datos útiles.

A iniciativa LOD es importante para mashups porque a diferencia de los mashups que utilizando conjuntos de datos fijos, mashups LOD funcionan con conjuntos de datos vinculados, abiertos y dinámicos. Por lo que ofrecen información precisa y actualizada. Al tener disponibilidad de LOD se permite la creación más fácil y rápida de mashups ya que los datos se puede acceder de forma unificada, mientras que los mashups estáticas a menudo tienen que manejar una variedad de formatos de datos.

Figura 30. Mashup de Visualización de datos geo referenciados

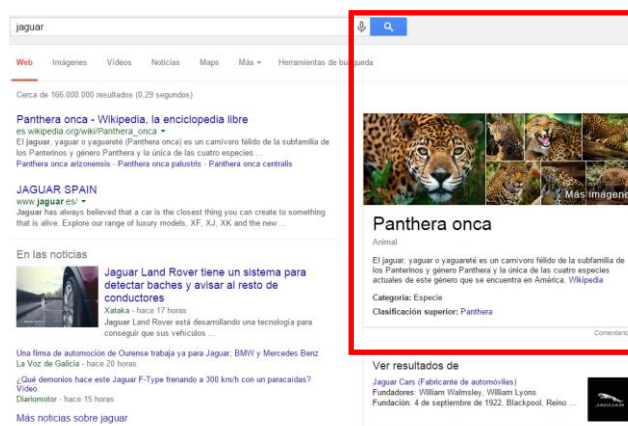


Fuente: <http://data.gov>

Búsqueda. Los motores de búsqueda tradicionales buscan en la web ocurrencia de los términos en sitios web o en los enlaces entre sitios web. Ellos no entienden semánticamente lo que el usuario está buscando y termina es comparando

cadena. Una web que le permitiera a los usuarios realizar preguntas podría ayudarle a obtener respuestas con mayor precisión. Sin embargo, además de la disponibilidad de Linked Open Data la realización de este tipo de motores de búsqueda necesita de la disponibilidad de los datos y los términos de licencia. De todas maneras, hoy día los motores de búsqueda ya utilizan datos vinculados en cierta medida. La figura 31 muestra una búsqueda en Google con la semántica. Se realizó la búsqueda "jaguar" y google devuelve no sólo el que describe la empresa Jaguar, sino también otra con el animal Jaguar, con el fin de indicarle al usuario lo que está buscando. Hay varias maneras de Google ya está utilizando los datos enlazados para mejorar los resultados de búsqueda⁶⁴.

Figura 31. Búsqueda en google usando semántica de datos enlazados



Fuente: Google.

1.6.6 Linked Open Data. Guía de Diseño. Linked Data se basa en cuatro principios básicos⁶⁵ y necesarios que ayudaran al crecimiento de la Web, definidos por Tim Berners-Lee.

1. Usar URIs para identificar los recursos publicados en la Web. Al nombrar los conceptos o cosas mediante URIs, se establece un identificador único, lo cual reduce problemas ya que no podrá haber errores de que dos recursos tengan el mismo URI.

⁶⁴ Thomas Steiner, Raphaël Troncy, and Michael Hausenblas. How google is using linked data today and vision for tomorrow. Linked Data in the Future Internet, 2010. <http://ceur-ws.org/Vol-700/Paper5.pdf>. (consultado el 19-05-2014).

⁶⁵ Wikipedia, «Datos Enlazados,». http://es.wikipedia.org/wiki/Datos_enlazados. (consultado el 19-05-2014).

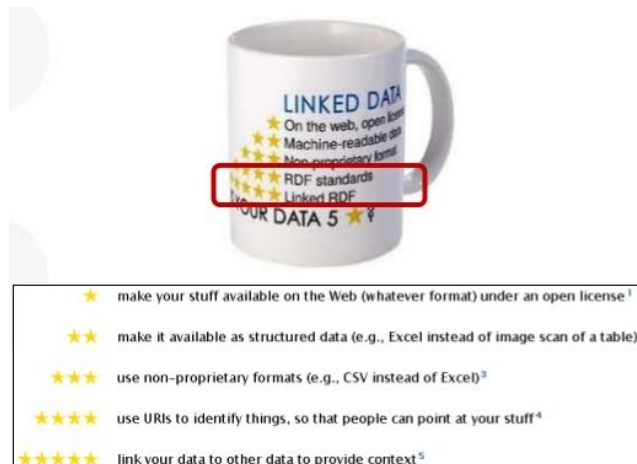
2. Usar URIs HTTP, la mayoría de los recursos son accedidos a través de HTTP, es por esto que se utiliza este protocolo, para asegurar el acceso a los recursos. Es decir proporcionar un camino simple para crear un nombre único global, se utiliza los URI HTTP como un medio de acceso a la información de otros recursos en la web.

3. Ofrecer información sobre los recursos usando RDF.

4. Incluir enlaces a otros URIs, de forma que se potencie el descubrimiento de información en la Web.

El gran valor de Linked Data se centra por tanto en brindar una oportunidad para utilizar los datos y explotar su valor mediante sistemas basados en conocimiento, con capacidad de procesar una cantidad masiva de datos obteniendo así soluciones más fiables y aproximadas a las necesidades y expectativas de los agentes. En la Figura 32, se pueden establecer una serie de características y criterios a satisfacer⁶⁶.

Figura 32. Características a tener en cuenta sobre Linked Data



Fuente: W3

⁶⁶ Tim Berners-Lee: Linked Data Design Issues. <http://www.w3.org/designissues/linkedata.html>. 2006. (consultado el 19-05-2014).

1.6.7 Definición De La Estrategia Uri. Para cumplir con altos estándares de calidad para la gestión de datos empresariales, se debe definir una política de representación de los datos persistentes que permita identificar cada elemento de datos del universo de datos empresariales. Esta política debe incluir los esquemas de direccionamiento para la localización de los recursos de datos en el ámbito empresarial. Una URI es un mecanismo relevante para definir un esquema de representación global de datos del negocio en el ámbito de la empresa.

Identificación de elementos de negocio como recursos referenciado por URIs. La primera recomendación en la construcción de una política de representación coherente y persistente es identificar los elementos de datos de negocios como los recursos, lo que se puede hacer referencia de forma individual. Para ajustarse a los principios de LOD, las URIs deben utilizarse como mecanismo de identificación para hacer referencia a las fuentes de información del negocio.

Uso de “minting HTTP URIs”. Son URIs que están bajo el control del que publica datos, es decir, utiliza las URIs de su dominio para los datos y documentos. La URI es un mecanismo que se puede utilizar para identificar diferentes objetos y conceptos. Algunos de estos objetos y conceptos pueden tener una existencia física como libros, por ejemplo, a través del ISBN. Otros conceptos son abstracto y representan cosas conceptuales tales como conceptos de ontología o elementos de datos. Existen diferentes esquemas de URI para representar un recurso: URIs basado en DNS (Domain Name Server), ARK (Archival Resource Key) y URIs basado en nombres y IDs tales como ISBN (International Standard Book Number), DOI (Digital Object Identifiers), códigos de barras, etc. Algunos de los esquemas descritos anteriormente pueden ser insuficiente para implementar características básicas de Linked Open Data tales como la edición, las referenciación y la interconexión. Por lo tanto, se recomienda principalmente el uso de URIs basado en protocolo HTTP y DNS (tales como URLs and ARK) con el fin de garantizar la visibilidad, accesibilidad y la reutilización de los elementos de negocio en aplicaciones externas y a terceros.

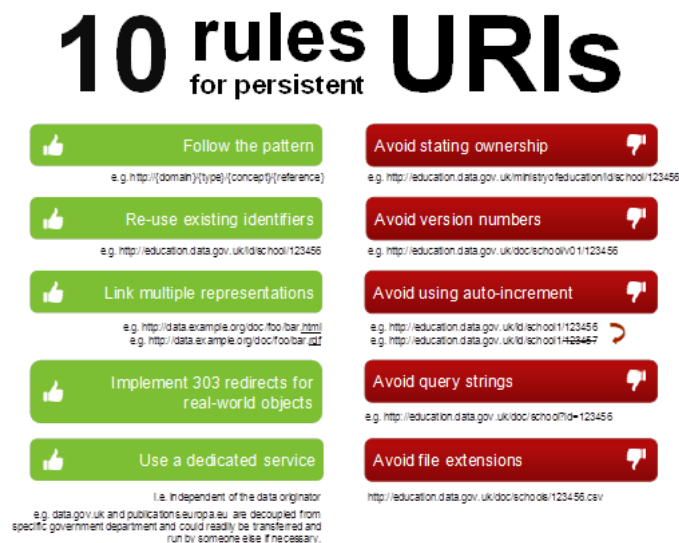
“Cool URIs”. Las URIs deben permanecer en el tiempo, en todo caso los usuarios las pueden cambiar, pero los identificadores deberían permanecer inalterados. De la misma forma, se debe proveer un mecanismo para que a través de una misma URI se pueda acceder a distinto contenido mediante negociación de contenido, extensiones, etc. Para el diseño conceptual de URIs, Sir Tim Berners Lee propone la idea de Cool URIs para garantizar que las URIs sean mantenibles, persistente y sencilla⁶⁷. Para garantizar URIs sostenibles, es importante diseñar un "cool" URI esquema que no cambia con el tiempo. Para ello hay que seguir las reglas

⁶⁷ <http://www.w3.org/TR/cooluris/>.

básicas, que se muestran en la Figura 33.

- Hacer que el URI sea independiente de la tecnología subyacente utilizada para generar o describir el recurso. Es decir evitar extensiones como .php, .cgi y .html en el URI path. Para saber que va devolver cuando se solicita un contenido (Sin extensión), se recomienda un mecanismo de negociación de contenido en el servidor HTTP que haga el outstream con el contenido apropiado.
- Asegúrese que la URI sea independiente de la ubicación física del archivo que describe el recurso. Hay que tener en cuenta que las direcciones físicas están sujetas a cambios
- Asegúrese que los metadatos de los recursos no estén incluidos en la URI debido a su evolución en el tiempo. Es decir, como lo estados, los derechos de acceso debido a que pueden cambiar con el tiempo.

Figura 33. Reglas para la persistencia de URIs



Fuente: JOINUP. En línea]. Disponible en Internet:
<https://joinup.ec.europa.eu/node/53858>

Uso de URI De-referenciable. Las personas relacionan mecánicamente una URI con URLs y esperan como respuesta una página a través del navegador Web. Desafortunadamente, la asociación de una URI a una página Web no siempre es

cierto y automática. Para algunas empresas, esta situación puede generar confusión y frustración. Para evitar tales confusiones, es muy recomendable proporcionar medios para resolver URIs, es decir, URIS que devuelvan respuestas significativas cuando se pega en la barra de direcciones del navegador. Una respuesta típica significativa podría ser una página HTML que contenga una descripción completa o parcial, incluyendo las propiedades del recurso correspondiente.

Recursos Independientes y Representación de Recursos. Crear objetos de negocio accesible y Dereferenceable a través del protocolo http puede generar una confusión conceptual entre el propio recurso y el documento que lo describe (la repuesta HTML por ejemplo, al momento de solicitar el recurso por HTTP). El recurso sí mismo como un elemento de datos de negocios debe ser identificado por un URI que es diferente de las posibles representaciones que se podría generar para describir el recurso (un HTML, RDF, XML O JSON, un documento en un lenguaje determinado usando tecnologías: PHP, HTML, etc.). La W3C propone dos soluciones técnicas para resolver la confusión anterior: uso hash URIs y el uso de 303 URIs:

Hash URIs. Esta solución consiste en utilizar fragmento URIs para hacer referencia a un recurso de negocio no-documento. Un fragmento URI es un URI que separa la parte que identifica el recurso de la ruta de ubicación en el servidor DNS utilizando el símbolo hash '#'. Por ejemplo, la referencia a un libro 2-253-09634-2 en una biblioteca puede disociarse de su descripción usando un hash URI de la siguiente manera: <http://www.mylibrary.com/books/about#2-253-09634-2>. Con este ejemplo, la biblioteca puede gestionar un repositorio de libros en un solo archivo grande en RDF que contiene todas las referencias a los libros y sus propiedades. Al acceder 2-253-09634-2, una consulta para seleccionar se puede ejecutar sobre el documento RDF para extraer el fragmento RDF correspondiente a 2-253-09634-2 triple. El servidor HTTP gestiona la de referenciación de la URI se aplicarán las reglas específicas de negocio para representar el fragmento RDF en la tecnología deseada (como JSON, HTML, XML, etc.).

303 URIs. Esta solución consiste en la implementación de un mecanismo de redirección representado por el HTTP response code 303 para indicar que el recurso ha sido identificado y el servidor está redirigiendo la solicitud a la descripción adecuada. En el ejemplo de la biblioteca, el URI podría ser <http://www.mylibrary.com/books/2-253-09634-2>. El servidor HTTP responderá a la petición a la URI por una redirección (Código HTTP 303) A una nueva ubicación; como <http://www.library.com/books/2-253-09634-2.about.html>, que proporciona la descripción del recurso solicitado.

Ambas técnicas tienen ventajas y desventajas como lo discute Sir Tim Berners Lee. Mientras que la técnica hash URI puede parecer restrictiva debido a que la URI hace parte de la raíz para diferentes recursos, la técnica 303 URIs introduce latencia en las solicitudes debido al mecanismo de redirección.

1.6.8 Linked Open Data: Vocabularios Existentes. Al crear propiedades para los datos RDF se recomienda utilizar un vocabulario que ya está definido y se utiliza en otras fuentes de datos. Esto hace que sea mucho más fácil para las máquinas interpretar la estructura de los datos publicados. Existen vocabularios para muchos casos de uso y están definidos en ontologías disponibles al público y ampliamente utilizados durante varias fuentes LOD importantes. Para reutilizar el vocabulario solo se debe definir un prefijo, enumerando el espacio de nombres y el uso de los elementos definidos allí. La siguiente lista describe las opciones de vocabularios que se pueden utilizar al crear LOD.

RDFS y OWL . El vocabulario predeterminado de RDFS⁶⁸ y OWL⁶⁹ contiene elementos que son conocidos y útiles al crear datos enlazados. El elemento `rdfs:label` se utiliza a menudo en las etiquetas de los recursos. Esta etiqueta puede ser utilizada cuando se despliegan datos de usuarios. Otro elemento importante es `owl:sameAs`. Este elemento puede ser usado para definir la similitud entre los recursos.

v Card. El formato vCard se diseñó para la definición de las tarjetas de visita electrónicas. Es ampliamente apoyado por los clientes de correo electrónico y, a menudo utilizado para transmitir datos de contacto entre las personas u organizaciones. Dado que este formato es muy popular, existe una codificación RDF⁷⁰ que se puede utilizar al crear LOD. Esta aplicación RDF define propiedades como "postalcode", "street-address" o "e-mail" que se puede utilizar para describir direcciones o datos de contacto.

FOAF The Friend-of-a-Friend. Define los términos para describir a las personas, sus actividades y sus relaciones con otras personas. Cuando la Web Semántica comenzó a surgir se crearon una gran cantidad de ontologías pero no muchas se utilizan ampliamente. La ontología FOAF luego se convirtió en una de las primeras ontologías ampliamente utilizadas. FOAF define un vocabulario RDF / XML para

⁶⁸ World Wide Web Consortium: RDFS. <http://www.w3.org/tr/rdf-schema/>, 2004. (consultado el: 23/05/2014).

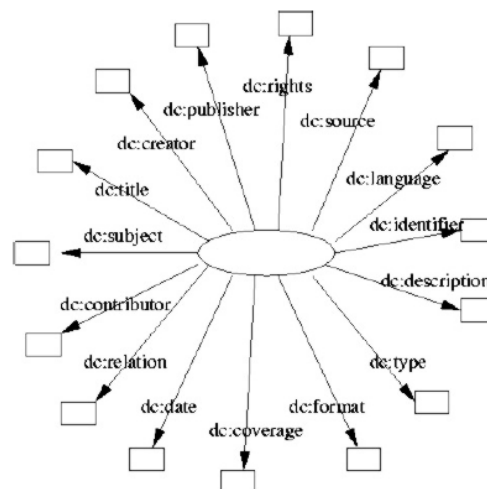
⁶⁹ World Wide Web Consortium: OWL. <http://www.w3.org/2004/owl/>, 2004. (consultado el: 23/05/2014).

⁷⁰ World Wide Web Consortium: vCard. <http://www.w3.org/submission/2010/subm-vcardrdf-20100120/>, 2010. (consultado el: 23/05/2014).

definir la información personal como nombre, página principal, buzón y relaciones con amigos y otras propiedades. La propiedad foaf: name se utiliza a veces como una alternativa a rdfs: label para etiquetado de recursos.

Dublin Core . Es una iniciativa de metadatos⁷¹ que proporciona propiedades que son adecuados para la descripción de los documentos. Atributos como título, autor, fecha, tema o editorial están disponibles. Dublin Core se ha convertido en un vocabulario muy popular. La figura 34 muestra las propiedades que están disponibles.

Figura 34. Ilustración de propiedades que define el Vocabulario Dublin Core



Fuente: Iniciativa Dublin Core. En línea]. Disponible en Internet: <http://dublincore.org/>.

WGS84 . Una ontología que define los términos para latitud y longitud y otra información sobre ubicación de cosas, usando WGS84 como un dato de referencia⁷².

BIBO. Ontología Bibliográfica (BIBO) ofrece conceptos y propiedades principales para describir citas y referencias bibliográficas (es decir, citas, libros, artículos, etc.).

⁷¹ Dublin Core Metadata Initiative. <http://dublincore.org/>, 2013. (consultado el: 23/05/2014).

⁷² World Wide Web Consortium Linked Data Cookbook. http://www.w3.org/2011/gld/wiki/linked_data_cookbook, 2011. (consultado el 19-05-2014).

1.6.9 El ejemplo de la DBpedia. Es una de las fuentes LOD más populares. Se puede ver como un ejemplo de buenas prácticas y por lo tanto se tomara de referencia y se describirá en detalle.

DBpedia es un proyecto comunitario que trata de construir una base de datos LOD grande que extrae información estructurada de la Wikipedia. La Wikipedia está disponible en más de 250 idiomas, con la edición de Inglés solo representa más de 1,95 millones de artículos a partir de 2007. El proyecto DBpedia ahora describe 3.770.000 cosas, de los cuales 2,35 millones se clasifican en una ontología coherente⁷³. El conjunto de datos consta de 1,89 miles de millones de tripletas RDF y versiones localizadas están disponibles en más de 100 idiomas.

Objetivos de DBpedia. En un artículo⁷⁴ los iniciadores del proyecto definen las siguientes contribuciones como objetivos de DBpedia:

- Desarrollar un marco para la extracción de información que convierte el contenido de Wikipedia en una base de conocimiento de varios dominios. Al acceder al artículo actualizado en vivo, y alimentar la base de conocimientos DBpedia refleja oportuna el estado real de la Wikipedia.
- Definen un identificador en la Web Dereferenceable para cada entidad DBpedia. Esto ayuda a superar el problema de la falta de identificadores de entidad que ha obstaculizado el desarrollo de la Web de los datos hasta el momento y sienta las bases para la interconexión de fuentes de datos en la Web.
- Publicar enlaces RDF apuntando desde DBpedia en fuentes de datos de otras Web y editores de datos para apoyar los vínculos con sus fuentes de datos. Esto ha dado lugar a la aparición de una red de datos alrededor de DBpedia.

El proyecto DBpedia es en sí mismo el "Núcleo de la Web de datos"⁷⁵. La principal

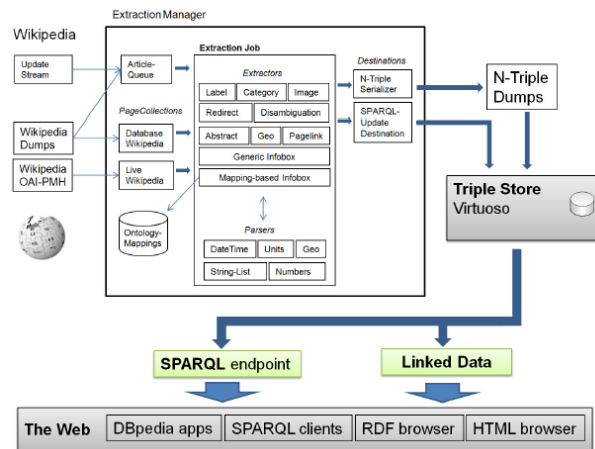
⁷³ About DBPedia. <http://dbpedia.org/about>, 2013. (consultado el 19-05-2014).

⁷⁴ Christian Bizer, Jens Lehmann, Georgi Kobilarov, Sören Auer, Christian Becker, Richard Cyganiak, and Sebastian Hellmann. Dbpedia-a crystallization point for the web of data. *Web Semantics: Science, Services and Agents on the World Wide Web*, 7(3):154–165, 2009.

⁷⁵ Soeren Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. DBpedia: A Nucleus for a Web of Open Data. In *The semantic web*, pages 722–735. Springer, 2007.

ventaja es la posibilidad de obtener resultados de consultas muy sofisticados a través de SPARQL. DBpedia también incluye una gran cantidad de datos de clasificación y por lo tanto puede ser utilizado para proporcionar el vocabulario y la taxonomía para otros conjuntos de datos enlazados.

Figura 35. La arquitectura de la DBpedia



Fuente: DBpedia.

1.6.10 Ventajas del Linked Open Data. Una de las iniciativas que tanto empresas públicas y privadas y entidades gubernamentales es abrir los datos, gran parte del futuro de la sociedad, está en la libertad de los datos y la calidad de los mismos y que estén disponibles para cualquier persona y empresa y que puedan ser aprovechados. El W3C está impulsando a nivel mundial la apertura de datos al público, y se establece como una filosofía que requiere que los datos estén disponibles y libres de restricciones de copyright, patentes o cualquier otro mecanismo que restrinja la utilización. El movimiento de datos abiertos comenzó su explosión en el año 2010 y sigue creciendo, gracias al respaldo de muchos gobiernos entre ellos el de Estados Unidos⁷⁶, el gobierno de la Gran Bretaña⁷⁷, en la Unión Europea⁷⁸, en España con muchos gobiernos regionales. En América Latina, se están posicionando iniciativas de Open Data en Chile⁷⁹, Brasil⁸⁰ y Colombia⁸¹.

⁷⁶ <http://data.gov>

⁷⁷ <http://data.gov.gb>

⁷⁸ <http://ec.europa.eu>

⁷⁹ <http://datos.gob.cl>

⁸⁰ <http://dados.gov.br>

⁸¹ <http://www.datos.gov.co/>

En la práctica Open Data ⁸² permite ofrecer la información a la sociedad de manera libre y sin restricciones. Estos datos procedentes de diferentes organizaciones, especialmente de la administración pública, busca que cualquier persona, organización o empresa puedan obtener conocimientos u oportunidades emprendedoras empresariales.

Existen una variedad de razones por la que publicar datos abiertos y enlazados proporciona beneficios y valor:

Open Innovation. Linked Open Data ofrece un alto potencial para la innovación. Los ciudadanos no tienen que pedir autorización para la utilización de datos, pues pueden hacerlo por sí mismos y ofrecer sus resultados a otros. Los datos no usados pueden ser analizados y mejorados.

Promoción del desarrollo económico. Dado que las empresas e industrias pueden utilizar Linked Open Data al igual que todos los demás, la pueden utilizar para aumentar su competitividad y la producción económica. El desarrollo de nuevos productos y puestos de trabajo, se pueden diseñar nuevos modelos de negocio. Esto también puede aumentar el atractivo como lugar de negocios y - por último pero no menos importante - la calidad de vida de los ciudadanos.

A través de la generación de destinos inteligentes se puede obtener conocimiento de cliente y desarrollo de productos. Puede utilizarse como herramienta para descubrir y medir la influencia de los clientes.

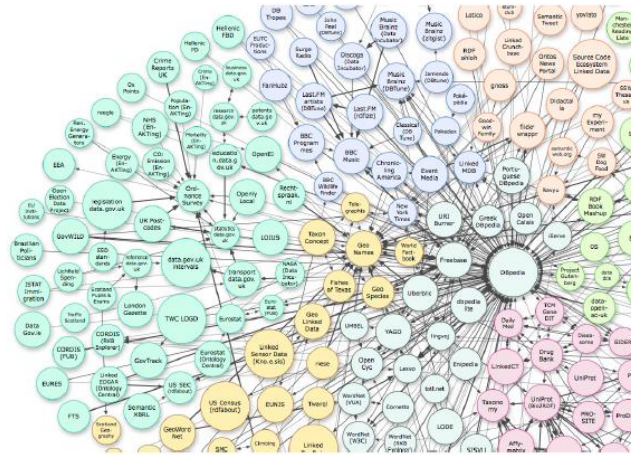
Con la implementación de LOD se brinda un avance importante al big data, ya que de esta forma no solo se expone los datos sino que se le da valor y significado tal como lo expresa Beth Noveck, anterior CTO gobierno de Obama y líder de su iniciativa de Open Government. “Tenemos que realizar la transición desde simplemente publicar cualquier cosa hacia publicar datos que permitan solucionar problemas reales”

1.6.11 Iniciativas Open Data y Linked Open Data. En el mundo son numerosas las iniciativas de open data, existen iniciativas que son exitosas como los casos de Estados Unidos, la Unión Europea, Gran Bretaña, España. La nube LOD también contiene datos de las agencias gubernamentales. Una versión codificada por

⁸² Computer World, 11 de enero de 2011, “Comienza el movimiento Open Data”, (consultado el 19-05-2014).

colores del diagrama está disponible en los que los datos vinculados gobierno son fácilmente visibles. Estos datos se representan en la figura 36 en el color turquesa. Por ejemplo, los datos de Eurostat, la agencia espacial de la NASA de los EE.UU. y el gobierno británico se liberan como Linked Open Data.

Figura 36. Cloud Linked Open Data



Fuente: GNOSS.

Estados unidos (Data.gov). Es el portal de datos abiertos de Estados Unidos, incluye páginas relativas a Data y Apps, Communities, Open Government, Learn, Semantic Web y Developers Corner. Los datos liberados son muy completo, pero aún no se libera en un formato que se ajusta a los principios de LOD. Sin embargo, el "TetherlessWorld Constellation" en el Rensselaer Polytechnic Institute ha creado el "Data-gov Wiki" como una extensión que integra los datos publicados en data.gov en la nube LOD. Han producido 6.460.000.000 triples que abarca temas como:

- El gasto del gobierno
- Los registros medioambientales
- Las estadísticas sobre el costo y el uso de los servicios públicos
- Los datos relacionados con la salud.

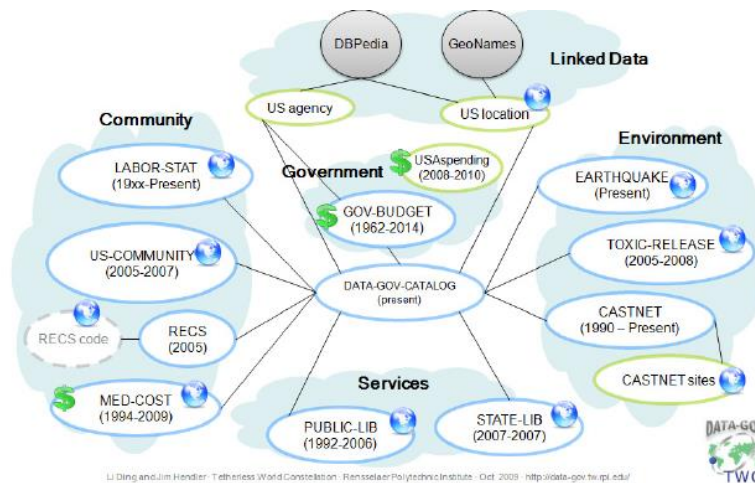
La nube de datos resultante se representa en la figura 37. Los datos están disponibles como archivos de datos o a través de SPARQL endpoints. Como se

describe en el artículo⁸³. Este proceso de conversión es bastante sencillo y altamente extensible. En primer lugar, los datos del gobierno en bruto se limpia-y se preservan a través de la representación basada en RDF.

En segundo lugar, estos conjuntos de datos se convierten usando URI Dereferenciable, por lo que tanto los conjuntos de datos y sus ontologías pueden extenderse a terceros.

El proceso de mejora se centra en la extracción de la semántica de los valores literales en los datos del gobierno en los URI significativas y la vinculación de los conjuntos de datos mediante la asociación de forma declarativa URI en diferentes conjuntos de datos.

Figura 37. Conversión de Open Data a LOD en data.gov



Fuente: Data.gov

Reino Unido (Data.gov.uk). El Reino Unido ha lanzado una variedad de Open Data Gobierno. Además también han adoptado los estándares de LOD. Contrariamente al portal data.gov EE.UU. implementan y liberan LOD automáticamente⁸⁴.

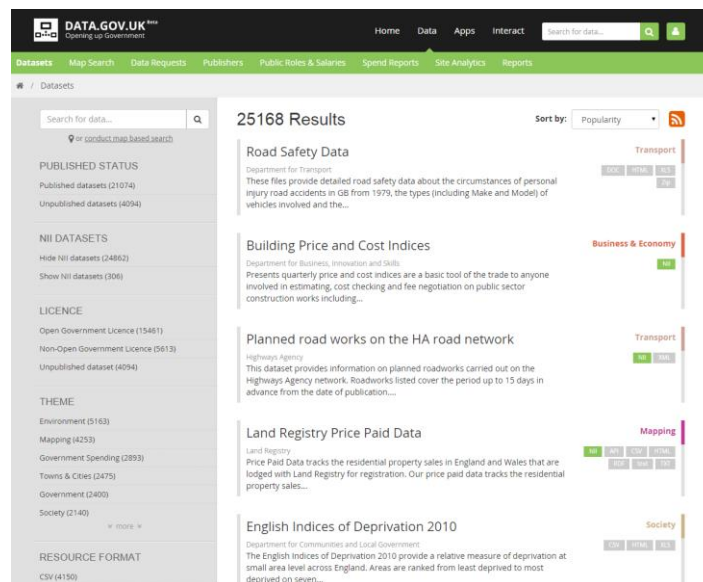
⁸³ Li Ding, Dominic DiFranzo, Alvaro Graves, James R Michaelis, Xian Li, Deborah L McGuinness, and Jim Hendler. Data-gov wiki: Towards linking government data. BCH+. 2010.

⁸⁴ UK Government Linked Data. <http://data.gov.uk/linked-data>, 2013. (consultado el 25-05-2014).

Legislación.gov.uk es un sitio web oficial del gobierno del Reino Unido. El sitio web ofrece acceso a toda la legislación publicada en el Reino Unido. Los textos de la legislación se pueden acceder y compartir por los usuarios. Todo en el portal está disponible como datos abiertos en los términos de una licencia abierta (que están usando Open Government License⁸⁵). Los datos se publican con anotaciones utilizando el micro formato RDFa y también está disponible en RDF. Si bien la creación del portal, el gobierno del Reino Unido siguió pautas LOD como la reutilización de vocabulario cada vez que fue posible y utilizando URI persistentes. Estándares del W3C se utilizaron para la liberación de datos. El gobierno del Reino Unido está utilizando datos vinculado como una tecnología subyacente, los datos también está disponible a través de una API para los desarrolladores que quieren tener acceso programático a los datos.

Los estándares abiertos en data.gov.uk. El Gobierno del Reino Unido se comprometió oficialmente en usar y re-uso de estándares abiertos y definiciones en el año 2010⁸⁶. Un artículo publicado en el mismo año se muestra las ventajas al usar Linked data en lugar de bases de datos convencionales.

Figura 38. Portal Data.Gov.Uk, estándares abiertos



Fuente: Data.Gov.Uk

⁸⁵ Open Government License. <http://www.nationalarchives.gov.uk/doc/open-governmentlicence/>, 2011. (consultado el 25-05-2014).

⁸⁶ United Kingdom Government Cabinet Office. Open source, open standards and re-use: government action plan. <https://www.gov.uk/government/publications/open-source-openstandards-and-re-use-government-action-plan>, 2010. (consultado el 25-05-2014).

España. En España son muchas las iniciativas a nivel de administración autonómica, local y a nivel nacional, aunque esta iniciativa a diferencia de las anteriores no cuenta con un proyecto general. Existen iniciativas exitosas como las del principado de Asturias, el país Vasco, Cataluña y Navarra y algunos ayuntamientos locales como los de Córdoba y Zaragoza. Mapa actual de las iniciativas de datos abiertos en España.

Figura 39. Mapa de iniciativas OGD en España



Fuente: Fundación CTIC. En línea]. Disponible en Internet: <http://datos.fundacionctic.org/2014/03/mapa-actual-de-las-iniciativas-open-data-en-espana/>

En una búsqueda sobre los datasets⁸⁷ de turismo en España se encontraron los siguientes proyectos:

Portal de Turismo de la Rioja. Es un portal LOD basado en tecnología Gnos.

⁸⁷ <http://datahub.io/>

Figura 40. Portal de turismo LOD de la Rioja



Fuente: <http://lariojaturismo.com/comunidad/larioja>

Otros proyectos tales como ayuntamiento de Zaragoza⁸⁸, Madrid⁸⁹, Cataluña⁹⁰, etc.

Brasil (OpenData.br). En Brasil el gobierno viene desarrollando herramientas para facilitar la colecta, almacenamiento, validación y utilización de informaciones sobre la acción gubernamental. Se creó un catálogo de información abierta y mejorar la gestión pública y facilitar el seguimiento por la sociedad.

En Colombia. En Colombia a través de la iniciativa gobierno en línea se impulsa una visión unificada de Estado colombiano, y la relación entre el estado y los ciudadanos se haga de manera transparente. Se busca que los ciudadanos vean al estado como una sola entidad, y que la relación sea más fácil y fluida. Es un modelo definido tanto para entidades del nombre central como territorial. A partir de la modernización del estado colombiano con el uso de la TIC, surgen proyectos como:

Gobierno abierto Colombia (Datos.gov.co). El artículo 230 del Plan Nacional de Desarrollo 2010 – 2014, presenta en su contenido: “Gobierno En Línea Como Estrategia De Buen Gobierno”. Donde se busca que todas las entidades de la

⁸⁸ <http://www.zaragoza.es/ciudad/risp/>

⁸⁹ <http://datos.madrid.es/portal/site/egob/>

⁹⁰ <http://dadesobertes.gencat.cat/es>

administración pública deberán adelantar las acciones señaladas por el Gobierno Nacional a través del Ministerio de las Tecnologías de la Información y las Comunicaciones para la estrategia de Gobierno en Línea. Esta estrategia liderada por el Programa Gobierno en Línea contemplará como acciones prioritarias el cumplimiento de los criterios establecidos al respecto, así como, las acciones para implementar la política de cero papel, estimular el desarrollo de servicios en línea del Gobierno por parte de terceros basados en datos públicos, la ampliación de la oferta de canales aprovechando tecnologías con altos niveles de penetración como telefonía móvil y televisión digital terrestre, la prestación de trámites y servicios en línea y el fomento a la participación y la democracia por medios electrónicos⁹¹.

En el portal de gobierno abierto encontramos un catálogo con datos de diferentes categorías como salud, vivienda, educación, transporte, etc. Y entidades del estado que se encargan de suministrar esta información del orden nacional y territorial. Se encuentra un espacio para desarrolladores con apps, Apis, protocolos y la tecnología usada para el despliegue.

Figura 41. Portal de Datos Abiertos de Colombia



Fuente: Gobierno Abierto de Colombia. En línea]. Disponible en Internet: <http://datos.gov.co>.

⁹¹ Ley No. 1450 por la cual se expide el plan nacional de desarrollo, 2010 – 2014, 16 de junio de 2011.

1.6.12 Las Apps y el turismo. Con la llegada de los Smartphone, las tabletas o las wearables technologies, las nuevas tecnologías se han puesto definitivamente al servicio del turista. Este nuevo turista digital, hiperconectado y multicanal, acostumbrado a usar sus dispositivos móviles y aplicaciones en su lugar de residencia para interactuar con el entorno y hacer su vida más cómoda, necesita disponer de conectividad y servicios móviles a lo largo de las distintas etapas del viaje. Por ello, la actividad turística ha generado en los últimos años la aparición de miles de aplicaciones relacionadas con productos, servicios y experiencias turísticas. Apps para buscar y reservar transporte, alojamiento o un lugar donde comer; apps para descubrir un destino; apps ligadas a la naturaleza, a la cultura, a la gastronomía; apps que ayudan a elegir compañero de viaje, a saber cómo están las olas para surfear o el viento para navegar, a traducir carteles, a obtener descuentos en entradas.

El uso de las aplicaciones turísticas ayuda en la preparación del viaje y en la interacción e integración del turista en el destino. Esta es una de las claves de los nuevos destinos turísticos inteligentes.

Crear aplicaciones que mejoren la experiencia turística y minimicen el impacto que en ocasiones puede tener el turismo de masas. El turismo es uno de los sectores con más potencial de crecimiento en la economía móvil. El año pasado se duplicó el número de reservar por móvil gracias a las aplicaciones y de los 1.500 millones de pasajeros que se esperan para 2015 en países como España, se calcula que el 90% utilizará terminales móviles. Viajeros que querrán encontrar en ellos toda la información de su destino, reservar hoteles, llamar a un taxi o planear su estancia. La mayoría de aplicaciones han sido creadas a partir de iniciativas de open cities. Algunas Apps para turismo en España⁹²:

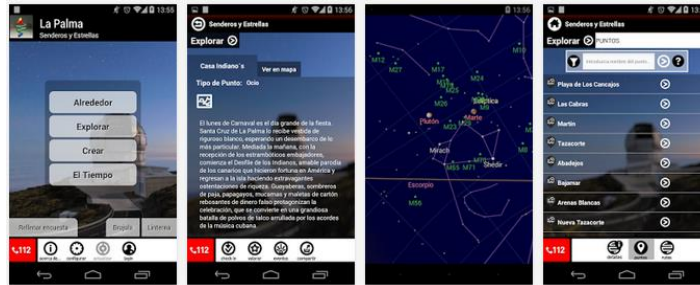
Figura 42. App Castilla de la Mancha



Fuente: Google Play

⁹² <https://play.google.com/store/apps/developer?id=SEGITTUR&hl=es>

Figura 43. App Astroturismo de la Palma

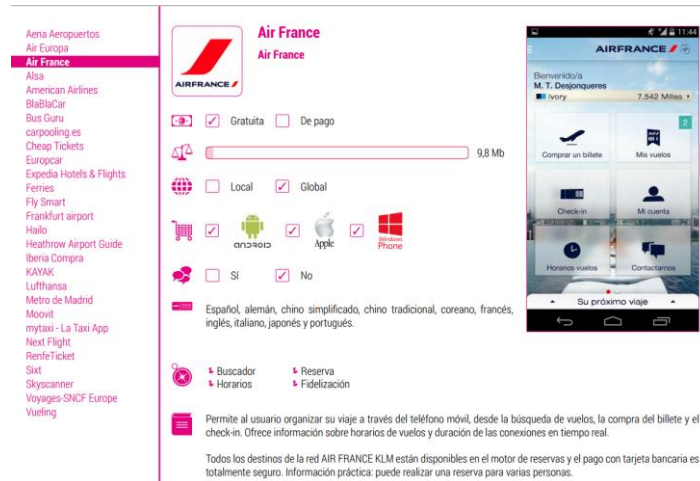


Fuente: Google Play

Según un estudio realizado por SEGITTUR⁹³ sobre las apps turísticas más influyentes a nivel mundial, según los productos más usados por los ciudadanos:

Transporte. Selección de aplicaciones relacionadas con el transporte, entre las que se incluyen comparadores, líneas aéreas, aeropuertos, trenes, autobuses, ferries, taxis y seguimiento de vuelo.

Figura 44. Aplicaciones para Transporte



Fuente: SEGITTUR.

⁹³ SEGITTUR. App Turística 2015. En línea]. Disponible en Internet: <http://www.segittur.es/opencms/export/sites/segitur/.content/galerias/descargas/documentos/2015-guia-apps1.pdf>. (consultado el 2-05-2015).

Alojamiento: Selección de aplicaciones que ayudan al turista a buscar y reservar alojamiento, entre ellas se incluyen buscadores de hoteles, alojamientos rurales, camping, hostales y una representación de cadenas hoteleras nacionales e internacionales.

Figura 45. Aplicaciones para Alojamiento



Fuente: SEGITTUR.

Ocio y Restauración: Las aplicaciones más representativas para buscar y reservar un lugar donde comer en cualquier lugar del mundo. Asimismo, incluye apps que permiten reserva de entradas, disfrute de parques temáticos, descuentos, así como diversas propuestas de ocio.

Figura 46. Aplicaciones para Ocio y Restauración



Fuente: SEGITTUR.

Turismo Activo: Selección de aplicaciones vinculadas con el turismo activo y herramientas que mejoran la experiencia del turista en sus prácticas deportivas de submarinismo, deportes de vela, esquí, senderismo, trekking o golf.

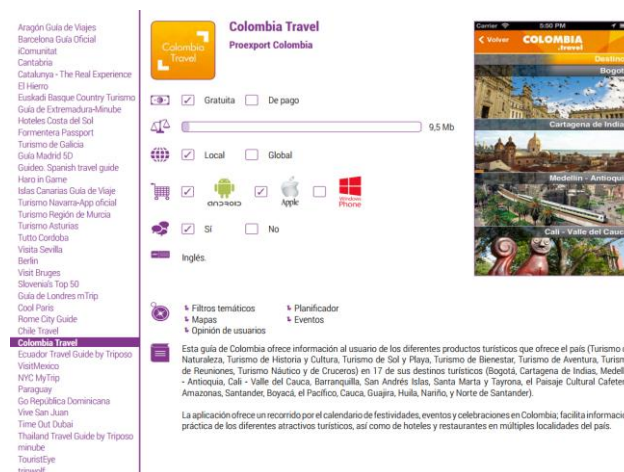
Figura 47. Aplicaciones para Turismo Activo



Fuente: SEGITTUR.

Guía de Destinos: apps que le ayudarán a conocer distintos destinos de España, así como importantes destinos internacionales: París, Londres, Berlín, Nueva York.

Figura 48. Aplicaciones para Turismo Activo



Fuente: SEGITTUR.

Utilidades. Aplicaciones que ayudan al turista en el destino. Apps que le sirven para convertir moneda, traducir, hacer fotografías, encontrar compañero de viaje, organizar la maleta, localizar wifi o controlar los gastos.

Figura 49. Aplicaciones para Turismo Activo



Fuente: SEGITTUR.

Una de las características de los dispositivos móviles más apreciadas por los usuarios es su interactividad, la posibilidad de personalizar la información, desarrollar un papel activo y dialogar con el emisor o con otros usuarios. Estas potencialidades se deben aprovechar por las aplicaciones móviles para el sector turístico.

1.6.13 Tecnologías relacionadas al Linked Data y Turismo. El impacto de las nuevas tecnologías en los hábitos turísticos tiene que ver con variables de distinto tipo: por un lado las que hacen referencia a las pautas de su uso con carácter general y, por otro, las que afectan al grado de utilidad y su capacidad de respuesta a los requerimientos del turista. Entender el papel de las nuevas tecnologías en su vinculación con las pautas turísticas significa, por tanto, descender al análisis de sus funcionalidades en los distintos momentos del proceso del viaje completo. Significa comprender la inserción de su uso en los distintos hitos del viaje y de los beneficios –de distinta naturaleza– que su uso aporta, o puede llegar a aportar, en cada uno de ellos.

En este sentido, es importante buscar la participación de las TIC, Es por esto que el primer paso para una completa comprensión del fenómeno de las TIC en la movilidad turística es encontrar las principales tendencias y las Nuevas Tecnologías éntrelas que tenemos el cloud computing, el big data, realidad aumentada y el Internet de las cosas.

Big data y turismo. En esencia, el Big Data posibilita el estudio y explotación inteligente de millones de bytes de información sobre toda clase de fenómenos y actividades -desde las variaciones atmosféricas hasta los patrones diarios de consumo de un colectivo concreto-, producida, difundida o almacenada a través de teléfonos móviles, redes sociales o, por ejemplo, máquinas conectadas al Internet de las cosas⁹⁴. Por su parte, la consultora Gartner prevé ya para 2015 la creación en todo el mundo de 4,4 millones de puestos de trabajo de perfil tecnológico ligados a esta disciplina⁹⁵. En general con el big data se pueden obtener los siguientes beneficios:

- Detección de tendencias y comportamientos.
- Conocimiento de cliente y desarrollo de producto.
- Análisis predictivo.
- Agilidad en la toma de decisiones.
- Optimización de procesos.
- Atracción y fidelización de clientes.

Una de las grandes oportunidades que el Big Data ofrece a la industria turística se encuentra en las Smart Cities⁹⁶ y, más concretamente, en los llamados destinos inteligentes. Se trata de territorios dotados de una infraestructura tecnológica suficientemente avanzada e integrada para garantizar su desarrollo sostenible y, a la vez, facilitar la experiencia e interacción entre el territorio y el visitante.

De esta forma, la captura, tratamiento, análisis y explotación de la información en un destino turístico permite a los agentes del sector una toma de decisiones más rápida y eficaz. Con datos, procedentes de teléfonos móviles, de las redes sociales, semáforos, pasos peatonales inteligentes⁹⁷, etc. Que facilita, un conocimiento más preciso de los hábitos y necesidades de los turistas.

⁹⁴ The Internet of Things. Agenda Digital Europea. Comisión Europea. <http://ec.europa.eu/digital-agenda/en/internet-things>.

⁹⁵ Gartner says Big Data creates Big Jobs: 4.4 Million IT Jobs Globally to Support Big Data By 2015. Gartner, 2012.

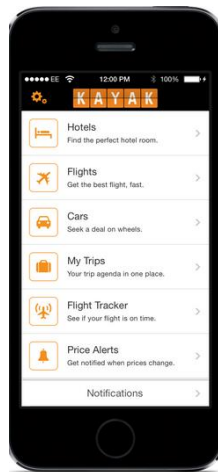
⁹⁶ <http://www.smartcities.es/>.

⁹⁷ The Internet of Everything for Cities. Cisco, 2013.

El concepto Big Data empieza a introducirse en el sector turístico ligado directamente a la gestión de la información y del conocimiento de las empresas y destinos. Aprovechar las oportunidades que presenta este nuevo escenario es un reto que exige no sólo una adaptación tecnológica, sino también una mejora en los procesos de gestión de la información. En este sentido, los nuevos modelos de excelencia turística en la gestión de destinos han entendido, en un escalón anterior a la adopción del enfoque Big Data, que hoy es el viajero quien diseña, realiza y controla.

Casos de éxito de implementación de Big Data en el turismo. Kayak.com⁹⁸ es un motor de búsqueda de viajes que ofrece a los usuarios la posibilidad de encontrar hoteles, vuelos, vacaciones y coches de alquiler a través de la búsqueda en cientos de webs de reservas.

Figura 50. App Kayak.com



Fuente: Kayak. En línea]. Disponible en Internet: <https://www.kayak.com/mobile>.

MakeMyTrip⁹⁹ es una compañía de viajes online. Ofrece a sus clientes la selección más amplia de productos y servicios de viajes de la India.

⁹⁸ At the Big Data Crossroads: turning towards a smarter travel experience. Amadeus, 2013.

Big shoes to fill – new Kayak tech boss on flights, mobile, Europe and more. Tnooz, 2014. <http://www.kayak.com/>

⁹⁹ MakeMyTrip, India - An emerging market travel technology pioneer. Amadeus, 2014.

Moovit¹⁰⁰ es una aplicación GPS de transporte público colaborativa y gratuita que permite a los usuarios del servicio viajar de un modo más inteligente. Informa sobre el estado del transporte público permitiendo planificar rutas que tengan en cuenta los diferentes medios disponibles y facilita la búsqueda de alternativas en situaciones de incidencia.

A demás de otras empresas como Landry's: detección de fraude de empleados, NH Hoteles: herramienta de medición de la satisfacción de los clientes, InterContinental Hotels Group: cómo hacer sentir a los clientes como en casa, etc.

Cloud Computing y Turismo. El cloud computing permite en el sector hotelero ejemplo, un esquema en el que los datos de éstos sean accesibles desde cualquier lugar del mundo sin necesidad de sistemas especiales es evidente que será de una gran utilidad. En el informe de conclusiones del IV Think Tank, organizado por SEGITTUR y Turismo, dedicado al Cloud Computing y el Turismo, se destaca que el ahorro de costes puede ser una de las principales ventajas de adoptar un modelo de implantación de computación en nube para las empresas turísticas. En su opinión, la situación económica actual invita a adoptar soluciones de computación en nube debido al ahorro de costes, que se sitúa entre el 30% y el 40% frente a las soluciones tradicionales.

En la era del turismo actual hay una serie de retos que brindaran oportunidades para muchas empresas en el momento que las permitan aprovechar el uso del cloud computing:

La movilidad. La cobertura de wifi, la energía y compartir.

Social Sharing. Las personas están obsesionadas por compartir, pues el hecho de viajar genera historia a los usuarios, y los nuevos canales han hecho posible la distribución de estos contenidos. El auge de las redes sociales y la facilidad de compartir sus experiencias en los canales más adecuados.

Contenido experiencial. Las personas comparten aquellas experiencias que lo apasionan, construyendo perfiles sociales que conforman su reputación social, los textos y sobre todos las imágenes aportan credibilidad y generan inspiración.

¹⁰⁰ Moovit. Transporte público inteligente. 2014.

Virialidad. No sirve de nada generar buen contenido si no llega a la audiencia adecuada.

Curación de contenido. Hay mucho contenido disponible, por lo que es ideal enseñar el contenido exclusivamente excelente.

Multidispositivos. La mayoría de personas en el mundo tiene un dispositivo conectado. El mundo actual es omnicanal, el online y el offline ya no existen, todo es lo mismo, además de aprovechar el internet de las cosas que permitirá que todo esté conectado.

Figura 51. Experiencia de usuario en la Rioja



Fuente: Minube. En línea]. Disponible en Internet: <http://www.forumturistic.com/big-data-big-opportunities>.

Todo lo que se hace genera datos, que se convierte en información y que la cual al ser analizada y procesada produce conocimiento, que luego para las organizaciones se convierte en ventas. En este nuevo entorno que propone la tecnología, se debe saber satisfacer las necesidades del viajero y aprovechar todo ese conocimiento.

Involucración del sector. Los profesionales del sector del turismo tienen una responsabilidad en este nuevo entorno tecnológico, siendo parte del movimiento de generar contenido y servir de inspiración, para que los viajeros puedan vivir su experiencia y contársela al mundo.

El internet de las Cosas y el Turismo. El Internet de las cosas va a suponer una auténtica revolución digital que cambiará considerablemente las formas en las que hoy día hacemos las cosas. Por supuesto esta revolución va a afectar a la

industria turística. Hay muchas empresas turísticas que ya están implantando aplicaciones del Internet de las cosas a diversos objetos para así facilitar las tareas de los turistas y mejorar sus experiencias en los destinos, hoteles y atracciones.

A través del Internet de las cosas, debido a la gran cantidad de equipos conectados se generan grandes volúmenes de datos, en este sentido ya están surgiendo interesantes debates sobre cómo tratar dichos datos para que permitan la toma de decisiones. Al aplicarlo al turismo estos datos serán de vital importancia. De una manera sencilla sabremos rápidamente, el perfil de turistas que visitan nuestra ciudad, qué monumentos visitan, cuanto tiempo emplean para las visitas, por qué calles realizan sus compras, hasta el perfil de la persona que se detiene delante de un escaparate o consume un determinado producto turístico. Y a través de estos datos se puede tener una imagen completa de la ciudad para conocer gustos y tendencias de los visitantes.

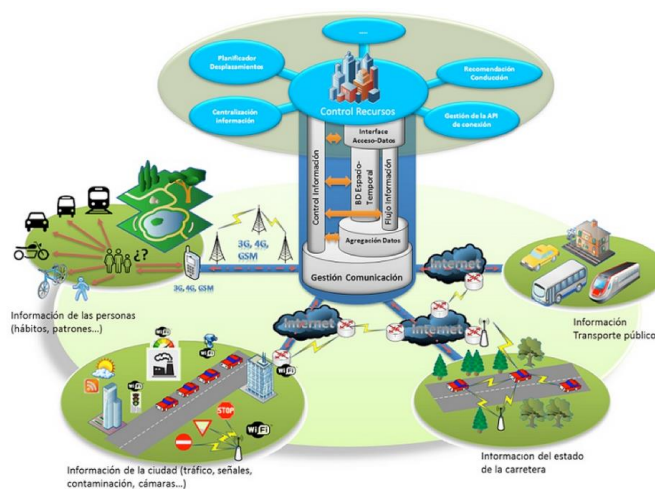
Para los ciudadanos esta información empleada correctamente ayudará a encontrar rápidamente lo que buscan y necesitan y por otro lado a los agentes turísticos a conocer a sus clientes y adaptar su oferta según sus gustos y necesidades. Por lo que el Internet de las cosas supone ya de por sí un cambio muy importante en la era digital e integrar el Internet de las cosas a los productos y servicios turísticos que es algo fundamental hacia un turismo inteligente¹⁰¹.

1.6.14 LOD y Turismo. Los servicios a través de LOD podría proporcionar un acceso más fácil y consistente a la información para permitir el desarrollo de nuevos servicios. Las aplicaciones de Smart City podrán aprovechar para obtener acceso a información integrada que incluye: mapas, estado del tráfico, las condiciones meteorológicas, estacionamiento, los niveles de contaminación, la energía, los sensores en tiempo real en los vehículos públicos y privados, los puntos de interés de la ciudad como museos, monumentos, restaurantes, hoteles, hospitales, etc., sino también datos estadísticos como accidentes de viaje, por la calle por año. Toda esta información, si bien estructurada y conectado se puede utilizar para ofrecer nuevos servicios a los ciudadanos. Sin embargo, la gestión eficiente de esta información no se pueden hacer a través de herramientas sencillas sino que se deben desarrollar herramientas específicas para integrar nuevos datos. Algunos de los proyectos de Smart cities y Linked open Data:

¹⁰¹ GUIDEO.<http://www.guideoapp.com/blog/internet-de-las-cosas-en-turismo-hacia-el-turismo-inteligente>. (consultado el 2-05-2015).

Proyecto Hermes102. El objetivo del proyecto es integrar los diferentes agentes y elementos de infraestructura de una Smart City en un sistema cooperativo y masivo que optimice los desplazamientos urbanos, minimice la emisión de gases contaminantes, maximice el bienestar de los ciudadanos y ofrezca nuevas oportunidades de negocio sobre la Smart City. Para ello, se combinarán y se ampliarán las últimas técnicas y tecnologías de captación, procesamiento, almacenamiento, consulta y publicación de datos, así como las técnicas de desarrollo de software en plataformas abiertas orientadas a servicios. A demás de la generación de nuevos modelos de negocio en el ámbito de las ciudades inteligentes aplicando la filosofía Linked Open Data a las soluciones desarrolladas.

Figura 52. Escenarios de actuación del proyecto Hermes



Fuente: Proyecto Hermes.

Otros proyectos como Smart City Málaga¹⁰³, promovida por ENDESA, se ha marcado como objetivo conseguir un ahorro energético del 20%, y una reducción de emisiones de más de 6.000 toneladas de CO2 al año, con el que se beneficiarán 300 clientes industriales, 900 servicios y 11.000 clientes domésticos.

El Ayuntamiento de Madrid tiene una aplicación para móvil bajo el nombre, “el aire de Madrid” que permite a los usuarios conocer en tiempo real la situación de la calidad del aire en la capital española, a través de los índices de calidad. Así, esta herramienta permite segmentar la aplicación por localización geográfica y tipo de indicador, además de ofrecer eco consejos a sus usuarios.

¹⁰² Proyecto Hermes. <http://madeirasic.us.es/hermes/>

¹⁰³ <http://www.smartcitymalaga.es/>.

Figura 53. App el Aire de Madrid



Fuente: Google Play.

El "Proyecto Rescátame104". En Salamanca, desarrollado conjuntamente por diversas entidades privadas y organismos públicos, que busca logra una "gestión sostenible del tráfico urbano o por medio de dos elementos clave: modelos de predicción y una red de sensores de calidad del aire.

Proyecto Turismo Inteligente105. Una plataforma de difusión multicanal orientada a promoción e información turística. Incluye la elaboración de un geo portal turístico y una App, ambos para uso de los ciudadanos.

1.7 PLANTEAMIENTO DEL PROBLEMA Y JUSTIFICACION

Colombia es un país con una gran riqueza y diversidad en su territorio esto lo convierte en un atractivo turístico natural, para quienes les gusta disfrutar de las posibilidades de conocer las culturas de los pueblos, las tradiciones, la diversidad en fauna y flora, etc.

Durante muchos años no se ha explotado a gran escala este potencial que permita fortalecer un campo económico importante para el país en el futuro, que contribuya en generación de ingresos y empleo.

El gobierno nacional ha planteado el propósito de convertir a Colombia en un destino turístico, de naturaleza, impulsado por el ministerio de industria y turismo y

¹⁰⁴ <http://rescatame.eu/>

¹⁰⁵ <http://www.tecnogeo.es/soluciones/smart-cities/turismo-inteligente/>

que se ha socializado en muchas regiones del país a través de la rueda nacional de negocios turísticos ¹⁰⁶ “Turismo Negocio”. Donde lo que se busca es generar oportunidades de negocios para pequeños empresarios del sector turístico nacional de tal forma que permite posicionar en el mercado nacional e internacional los productos turísticos de las regiones. Estas iniciativas han logrado que el país siga creciendo en turismo, según el diario el espectador¹⁰⁷ el número de turistas extranjeros en el 2001 que llegó a Colombia se citaba en 1'723.000 diez años después esa cifra aumentó a 2'767.319. Hace una década el ingreso por turismo era de U\$ 1.483 millones, el año 2010 fueron U\$2.797, pero a un seguimos por debajo de otros países de la región, En las evaluaciones realizadas Colombia ocupó puesto 77 entre 139 países en el 2011. Luego en el estudio de 2013 bajó al puesto 84. En el reporte¹⁰⁸, Colombia se ubica por debajo de países americanos competidores como Panamá (puesto 19), México (puesto 44), Costa Rica (puesto 47), Brasil (puesto 51), Puerto Rico (puesto 52), Chile (puesto 56), Uruguay (puesto 59), Argentina (puesto 61), Perú (puesto 73) y Ecuador (puesto 81).

A pesar que estas cifras son regulares para el país, lo comprometen a seguir trabajando y estableciendo nuevas estrategias para lograr consolidar la actividad turística como un campo de mayor importancia para el desarrollo económico¹⁰⁹. En el foro económico mundial, por su parte como se muestra en la figura 53, se detalla los indicadores que se consideran imprescindibles para medir la competitividad turística de los países y su posición a escala mundial; Dentro del cual se plantea indicadores para fortalecer la infraestructura tecnológica y la preparación para el turismo del futuro además como estrategias de promoción y divulgación que permiten atraer turistas extranjeros y consolidar una cultura turística de la sociedad colombiana.

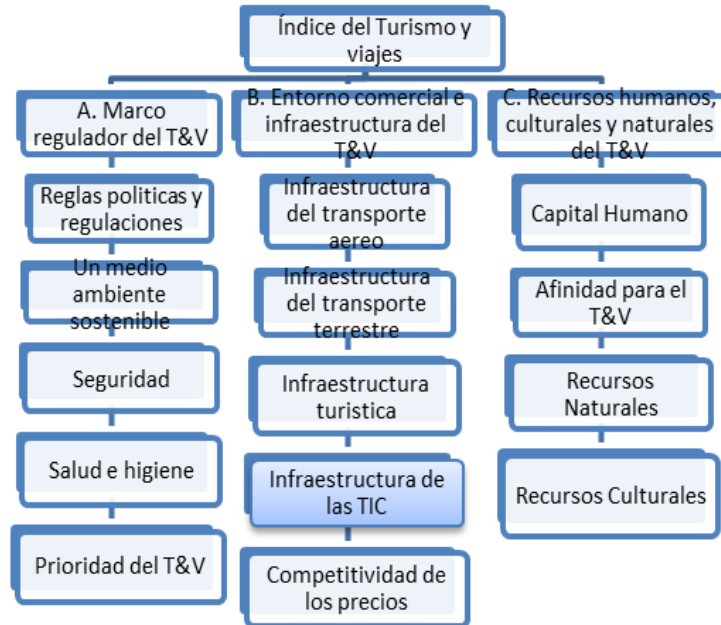
¹⁰⁶ Colombia es Pasión, “Turismo Negocia”. <http://www.turismonegocia.com.co/>

¹⁰⁷ El espectador, “Colombia ocupa el puesto 77 en competitividad en turismo”. <http://www.elespectador.com/noticias/nacional/articulo-302638-colombia-ocupa-el-puesto-77-competitividad-turismo>

¹⁰⁸ FORO ECONÓMICO MUNDIAL, Reporte de Competitividad de Viajes y Turismo, Ginebra 2013.

¹⁰⁹ PLAN SECTORIAL DE TURISMO 2011 – 2014 “Turismo: factor de prosperidad para Colombia.

Figura 54. Índice de Turismo y Viajes



Fuente: Indicadores de Competitividad Turística, Foro Económico Mundial.

Actualmente el turismo comercial se concentra fundamentalmente en las grandes ciudades y en los sitios turísticos ampliamente reconocidos a nivel nacional como internacional. Pero desconocemos muchas regiones y sitios del país con un alto potencial turístico, como se muestra en el cuadro 2. Los cuales no han podido consolidarse por diferentes razones, pero principalmente por la baja demanda de información, divulgación y promoción a través de infraestructuras tecnológicas. Hay que tener en cuenta que los servicios turísticos en estas zonas se manejan a través de pequeños empresarios o negocios familiares en los que se sigue manteniendo una amplia brecha digital y deficientes sistemas de divulgación y promoción a través de las TIC y que no permite que la información llegue fácilmente a turistas locales como internacionales, además que permita integrarse con aplicaciones de terceros, debido a que las tecnologías usadas no permiten un intercambio eficiente de contenidos. Desaprovechando el auge de tecnologías masivas y emergentes de bajos costos y amplios beneficios como equipos móviles y portables usados ampliamente en el mundo para el sector turismo y en donde se evidencian los grandes beneficios que generan a la economía mundial.

Cuadro 2. Visión Turística por Regiones

DEPARTAMENTO	VISIÓN/ TIPO PRODUCTO	DEPARTAMENTO	VISIÓN/ TIPO PRODUCTO
Amazonas	Turismo Comunitario / Turismo de Naturaleza	Guaviare	Turismo comunitario-turismo de naturaleza
Antioquia	Negocios, ferias y convenciones, turismo de compras, salud, agroturismo, turismo cultural, congresos y convenciones y ecoturismo	Huila	Ecoturismo, cultural, comunitario, convenciones, ferias y fiestas
Arauca	Ecoturismo, agroturismo, histórico-cultural y comercial	Magdalena	Ecoturismo
Atlántico	Negocios y cultura e historia	Meta	Trabajo de llano, naturaleza y folclor
Bogotá	Negocios, congresos, eventos y convenciones, compras, cultural	Nariño	Ecoturismo y turismo cultural
Bolívar	Ecoturismo, etnoturismo y cultura	Norte de Santander	Destino histórico, cultural y de compras.
Boyacá	Ecoturismo, cultural, comunitario, entrenamiento de alto rendimiento	Putumayo	Etnoturismo- turismo de naturaleza
Caldas	Ecoturismo y turismo cultural	Quindío	agroturismo
Caquetá	Turismo comunitario y turismo de naturaleza	Risaralda	Turismo rural y de negocios
Cartagena	Sol y playa, historia y cultura, congresos y convenciones	San Andrés y Providencia	Sol y playa y ecoturismo
Casanare	Agro turístico y cultural	Santa Marta	Ecoturismo
Cauca	Agroturismo, termalismo, cultural, aventura, mítico, étnico, arqueológico	Santander	Destino de aventura, salud y cultural

Fuente: Ministerio de Comercio, Industria y Turismo – Convenios de Competitividad Turística.

Actualmente existe una dispersión de los datos turísticos que se difunden a través de fuentes heterogéneas como archivos de Excel, bases de datos relacionales, texto tanto en sistemas web propios como a través de redes sociales, etc. Adicionalmente, una serie de datos que reposan en las organizaciones encargadas del sector turístico en Norte de Santander como la Secretaria de Desarrollo Económico, la Cámara de Comercio, las Alcaldías y entes territoriales, Corporación Mixta para el Fomento del Turismo, etc. Y que no se expone en ningún medio y que suponen una serie de oportunidades comerciales. Un reto importante para el sector turístico de Norte de Santander es conseguir el diseño

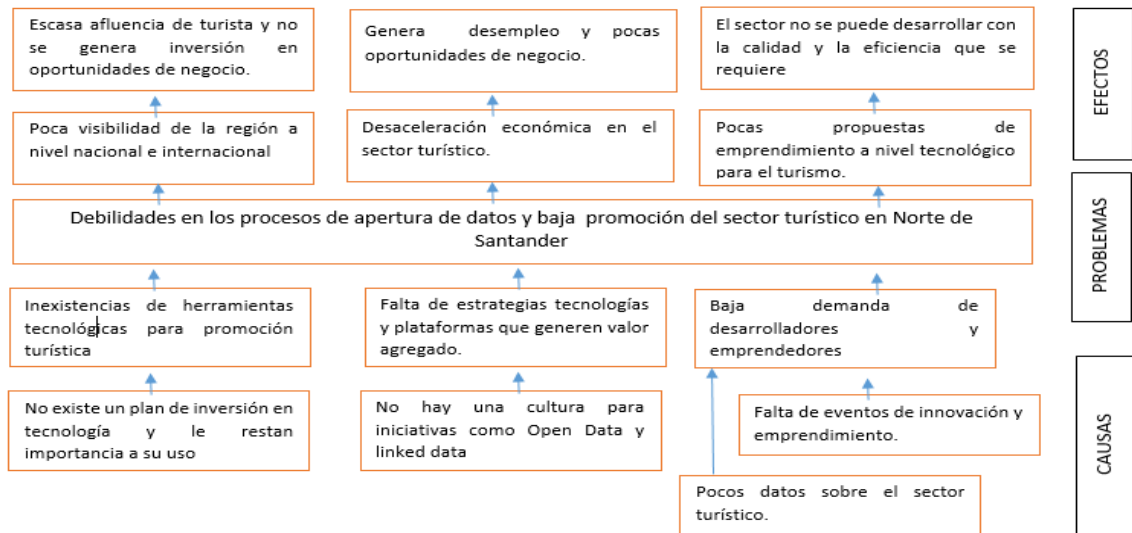
de un modelo eficiente para abrir datos a través fuentes abiertas y enlazadas con el fin que se pueda compartir, extender y reutilizar fácilmente constituye un problema tecnológico importante por resolver.

Una de las problemáticas importante es que la información se encuentra en formatos heterogéneos. Buscar, descargar y procesar la información de cada una de las fuentes para que sea reutilizable, requiere resolver diferentes problemas complejos, entre ellos, interpretar la información procedente de varias fuentes estructurada o no, y homogeneizar todas las fuentes. Además los datos disponibles y su formato pueden varían en el tiempo. En la actualidad mucha de la información estructurada se encuentra almacenada en bases de datos de tipo relacional, en documentos XML, HTML, etc. Estos esquemas de representación de información no presentan características adecuadas que permitirá el intercambio eficiente de información de la actividad turística y no permite forjar conexiones entre ellas y ofrecerles servicios basados en la Web para difundir y reutilizar esta información en nuevos contextos y para realizar razonamiento e inferencia ya que carecen de mecanismos para representar información que describa la semántica de los datos.

A pesar del crecimiento acelerado de las tecnologías, es muy poca la información que se encuentra sobre el turismo Norte santandereano, y no hay aplicaciones tanto Web como móviles, lo que genera un desconocimiento total de los ciudadanos, sobre sitios de interés y destinos turísticos, afectando profundamente la economía del turismo.

¿El diseño de un modelo basado en una arquitectura de datos abiertos y enlazados “Open Linked Data” permitirá que se creen nuevos servicios y productos para promocionar, comercializar e incentivar los datos abiertos sobre el turismo en Norte de Santander?

Figura 55. Árbol de problemas



Fuente: Elaboración propia

El uso de nuevas tecnologías e iniciativas basadas en Open Data y Linked Data, son de actualidad, y que se convierte en pieza clave para el desarrollo de nuevas aplicaciones y la capacidad de promocionar y comercializar nuevos servicios y productos en el sector turístico que le permite dar valor añadido a una actividad compleja y transversal en todos los sectores de la economía y en la cual hay un enorme potencial para el desarrollo económico de la región.

Los modelos semánticos a través de las arquitecturas de datos enlazados permiten la creación y distribución de datos los cuales se pueden compartir, reutilizar y extender para generar a las organizaciones mejorar los procesos internos de tratamientos de datos, enlace de datos y la publicación de los mismos los cuales incrementa el número de desarrolladores y emprendedores que interactúan con la organización quienes no están sujetos a formatos propios y pueden crear herramientas más generales, además para las organizaciones es el primer paso hacia la gestión de la información Turística en la nube.

Por otra parte, la investigación desarrollada en estas áreas se está produciendo múltiples publicaciones de alto valor científico, realización de conferencias nacionales e internacionales, desarrollo de nuevas técnicas, herramientas y transferencia tecnológica a empresas e instituciones de otros sectores tanto públicos como privados, las cuales han integrado diferentes fuentes de datos para

generar valor añadido y servicios inteligentes y contextualizados en sectores como la salud, educación, industria, etc. A través de Linked Data se dan aportes significativos. Puede asistir en la integración de información, ayudar a crear nuevos senderos al descubrimiento de información e influenciar el nivel de capacitación turística de los usuarios para consumir como promocionar servicios. Al igual que permite impulsar la marca regional y marca país convirtiendo a la región en un destino turístico apetecido para los ciudadanos de cualquier parte del mundo

En el campo de la Web Semántica se manejan distintas líneas de investigación que van desde modelos, protocolos, estándares hasta la ejecución de procesos más innovadores como es la integración de fuentes de datos de distintas fuentes o gestión de la información. Con esta investigación se busca incrementar las actividades de investigación en múltiples líneas y con distinto carácter: básica, aplicada, innovadora, etc. Para este caso la investigación se centra en la arquitectura de Linked Data al dominio del sector turístico, con el objetivo de crear un dominio digital que de acceso a la información que de estos procesos se produce y que puede generar gran trascendencia económica, por lo que implica que la tecnología que da soporte debe impulsar el conocimiento en el campo del turismo, incentivar desarrolladores y emprendedores impulsando y mejorar la cadena de valor tanto para la propia Administración como para los proveedores.

1.8 OBJETIVOS

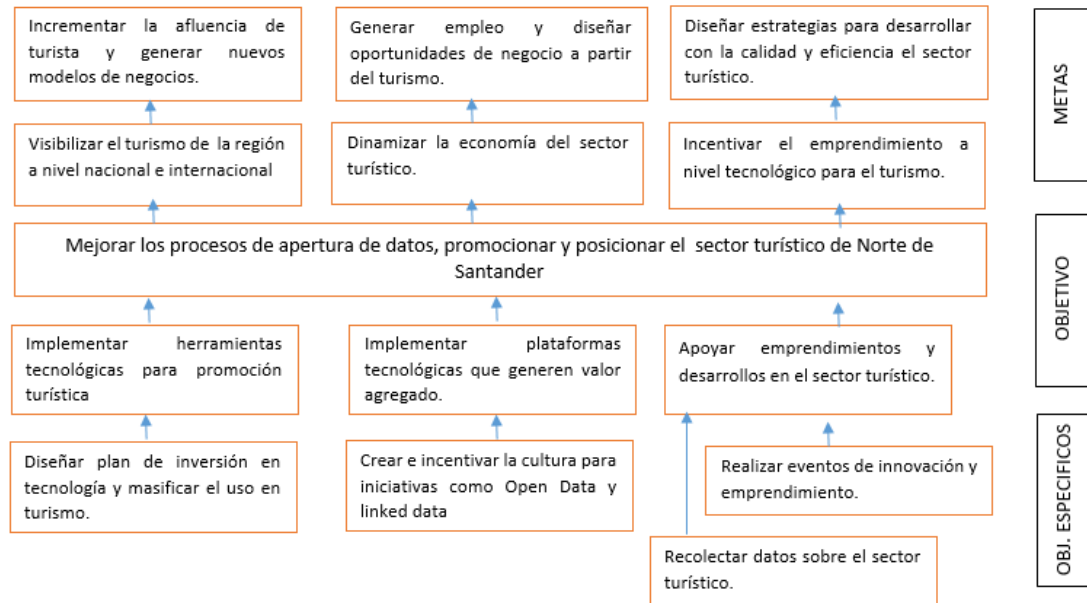
1.8.1 Objetivo General. Diseñar un modelo basado en arquitectura open y linked data para el sector turístico del Departamento Norte de Santander.

1.8.2 Objetivos Específicos. Se describe a continuación

- Elaborar un estado del arte sobre Open Linked Data y tecnologías asociadas y su implementación a nivel de Colombia y el mundo para la apertura de datos.
- Definir la estructura y los componentes del modelo que permita extracción, filtrado, enriquecimiento y publicación de los datos sobre el turismo de Norte de Santander a partir de las necesidades del negocio.

- Diseñar y documentar un prototipo App sobre los puntos turísticos de Norte de Santander que se obtenga a partir de la información publicada en la plataforma.

Figura 56. Árbol de Objetivos



Fuente: Elaboración propia.

1.9 BENEFICIOS

Los beneficios en la aplicación de los datos abiertos y enlazados en el sector turístico de Norte de Santander. Aunque el análisis se centrará en el proyecto objeto de esta tesis, se puede aplicar a cualquier otro sector económico o al mismo sector turístico en cualquier parte del mundo

Una de las principales razones para proponer el proyecto de datos abiertos enlazados en el sector Turístico de Norte de Santander, es la necesidad de mejorar el modelo de masificación de los datos, teniendo en cuenta la tendencia mundial para la liberación de datos públicos, especialmente en sectores turísticos de países con un amplio potencial turístico. Sin embargo, este es sólo uno de los beneficios marginales directos que aparecerá después de la aplicación de la proyecto. La lista de beneficios directos que fueron identificados para el proyecto es la siguiente:

Cuadro 3. Beneficios Marginales Directos

No.	Beneficio.	Como
1	Extender la masificación y transparencia.	Mediante la publicación de datos en diferentes formatos y estructurados para que los usuarios pueden descargar estos datos.
2	Aumentar la visibilidad internacional de la de la región	Dado que los datos están abiertos al mundo y el enfoque de los datos enlazados pueden mejorar la posición de la región a nivel nacional e internacional.
3	Mejorar la calidad de los datos a través de comentarios de los usuarios	A veces, la única manera de mejorar la calidad de los datos es a través de la retroalimentación de los usuarios. Por compartir recursos a través de la plataforma abierta de datos, la calidad de algunos datos internos se puede mejorar.
4	Apoyar el movimiento de datos abiertos y enlazados.	El movimiento de datos abiertos y enlazados es un valor importante para la sociedad. Los datos abiertos pueden ser parte de, y haciendo eso, el turismo puede aumentar el interés de más ciudadanos.
5	Difundir los datos en diferentes formatos, permite a los usuarios acceder a ellos cada vez que lo necesitan. Esto genera un mayor interés de los usuarios para utilizar la plataforma.	A veces, para un conjunto de datos específico, pueden existir diferentes usuarios que requieren los datos en diferentes formatos. Por ejemplo, un desarrollador puede desear conseguir los datos en JSON a través de un API y un analista de datos puede necesitar los datos en formato csv o Excel. A través de plataforma de datos abierta esto se puede hacer, el suministro de datos para todos los posibles perfiles de usuario.

A grandes rasgos, esos son los principales beneficios que el proyecto dará al sector turístico sin ningún esfuerzo adicional, que sólo la publicación de datos abiertos. Pero, teniendo en cuenta sólo esto, no hay ningún valor añadido que el ecosistema de datos abiertos y enlazados puede generar. De acuerdo con esto, se identificaron algunos beneficios adicionales, que puede dar más valor a los datos publicados. Que se describe en la siguiente lista:

Cuadro 4. Beneficios extras

No.	Beneficios	Esfuerzos Obligatorios
1	Apoyo a la Innovación	Crear un plan de comunicación y promoción para animar a los estudiantes y otras personas para explorar los datos, para obtener nuevas ideas o puntos de vista acerca de la los datos y la forma de crear más valor en él. Este beneficio requiere la diversidad de los datos en el sector turismo. Este plan de comunicación debe tener un enfoque especial en las aplicaciones con el fin que se desarrollen Apps con nuevos servicios.
2	Mejorar la conexión con la comunidad empresarial	Promover el portal de datos entre las empresas privadas y organizaciones externas. Las empresas pueden ayudar a los estudiantes a desarrollar nueva ideas o proyectos con la ayuda de esta plataforma y que den un enfoque de negocio para ellos. Se puede aprovechar algunas iniciativas de emprendimiento como apps.co
3	Mejorar el servicio al público objetivo.	Fomentar el desarrollo de nuevas aplicaciones con un enfoque social.
4	Apoyo a las actividades de investigación académica.	La plataforma de datos abiertos puede ayudar a los investigadores a trabajar directamente con los datos en bruto.
5	Mejorar el proceso de desarrollo a través eventos de cómo concursos o desafíos.	El desarrollo de aplicaciones puede ser más dinámico y más rápido si algunos eventos, desafíos o se crean hackatons. Estos permiten el desarrollo rápido a través de enfoques iterativos para desarrollar Apps.
7	Dar más valor y mejorar el uso de tecnologías que ya están implementado en turismo.	Las entidades del estado desarrollan aplicaciones que no son utilizadas. Usando los datos abiertos enlazados, la visibilidad se puede mejorar y puede ser más útil, para buscar datos que los usuarios requieran.

La mayor parte de los beneficios anteriores necesitan esfuerzos continuos que dependerán del apoyo de las organizaciones involucradas en el turismo de Norte de Santander.

1.10 IMPACTO

Impacto Social: Las iniciativas de datos abiertos y enlazados buscan brindar muchos beneficios a la sociedad y principalmente mejorar la calidad de vida. A través de datos abiertos muchas personas pueden construir aplicaciones principalmente innovaciones sociales que contribuyan a ofrecer servicios a la sociedad.

A través de la iniciativa ciudadanos se motivan a participar activamente en la mejora de los servicios públicos, no sólo con sugerencias, propuestas, etc., sino también con contenidos generados por ellos mismos o ideas e iniciativas creadas y promovidas por ellos o nuevas prestaciones (aplicaciones) desarrolladas a partir de los datos públicos liberados.

Al permitir ofrecer información relevante sobre el turismo de Norte de Santander, se puede visualizar mejor la región, esto conlleva a mejorar el flujo de turista y de esta forma estimular el desarrollo económico, contribuyendo a generar más puestos de trabajo.

Impacto para el sector de Turismo: A nivel de Colombia, los datos abiertos enlazados tiene un alcance más amplio que implica la independencia de la iniciativa, La idea principal es hacer más fácil el proceso para la agregación, la integración y el consumo de datos que se generen en diferentes instituciones. Algunos de los principales beneficios son los siguientes:

- Crear una red que permita que todas las instituciones involucradas con el turismo, puedan generar conjuntos de datos abiertos y enlazados que pueden fomentar la generación de nuevos proyectos e ideas, a nivel nacional.
- Integración y combinación de datos de turismo con el fin de crear más interés y la demanda del desarrollo de aplicaciones, ya que el tiempo para la conciliación de datos entre fuentes de datos se reduce.
- Utilizar diferentes idiomas para la descripción de datos, ya que los datos las conexiones se realizan a través de mapeos semánticos.

- Comparación de datos para mejorar la investigación y nuevos modelos de negocio.

Impacto Económico. Los datos abiertos generan la posibilidad de crear valor económico. Esto genera ideas de todos los sectores, elevando la productividad, que permitan la innovación, la aplicación de procesos impulsados por datos y la creación de nuevos productos, servicios y mercados y que se pueden aplicar en los siguientes sectores educación, transporte, consumo, electricidad, petróleo y gas, servicios de salud y de financiación de consumo. Una evidencia significativa del valor económico de los datos abiertos está relacionado con la creación de nuevos modelos de negocio que pueden generar ingresos futuros y el ahorro para los gobiernos y organizaciones. A demás la generación de empleos por la participación de los estudiantes, desarrolladores, investigadores, empresarios y otros grupos de interés. Entre ellos, se requiere un alto nivel de participación de los desarrolladores, ya que son los que pueden transformar los datos en las aplicaciones reales que pueden proporcionar servicios adicionales o, a su vez ser parte de un proyecto empresarial.

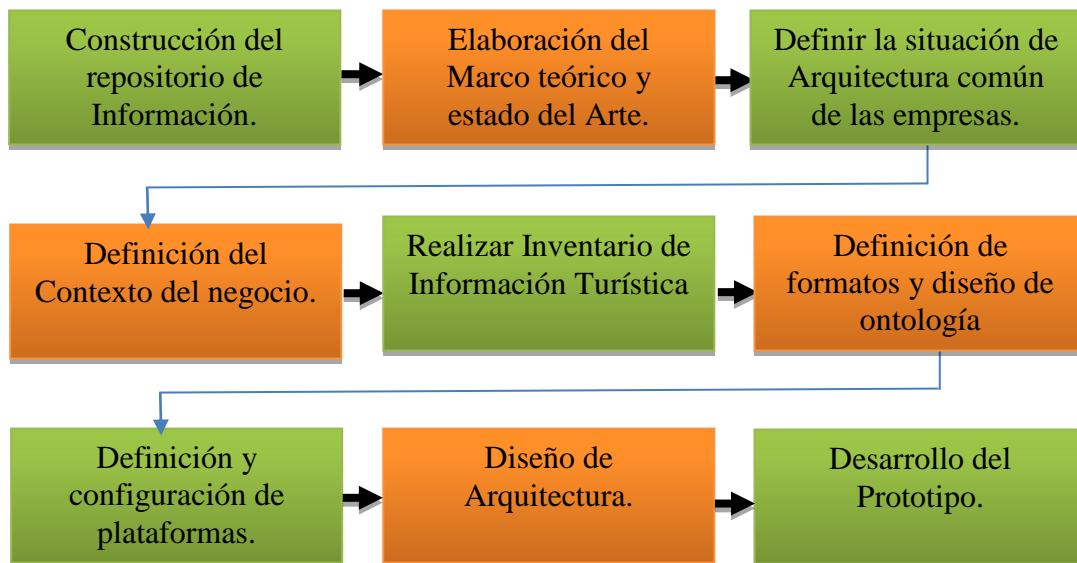
Impacto Ambiental. Reduce los costos relacionados con la gestión manual de la promoción de la ciudad (impresión, escritura, transporte).

Tecnológico. Las nuevas tendencias tecnológicas en este caso Open Linked Data, permite gestionar de manera eficiente la información con el fin de poner a disposición de la ciudadanía y diferentes actores interesados datos para el desarrollo de nuevos servicios.

2. METODO

2.1 ACTIVIDADES

A continuación se describen las actividades relacionadas con el desarrollo del proyecto.



Construcción de repositorio de información. Durante esta fase se desarrollarán actividades que comprende la recolección, consulta, selección y descarga de material bibliográfico, artículos científicos, libros, revistas, ponencias, etc. De bases de datos especializadas.

Elaborar el Marco Teórico y el Estado del Arte. Durante esta actividad se identificaron tecnologías, protocolos y servicios, conceptos técnicos, estándares y políticas de buenas prácticas definidas en la W3C, casos de estudio sobre la implementación de la arquitectura de datos enlazados Linked data, Open data, casos de éxito, algunos temas que involucra esta actividad se encuentran:

- La web semántica.
- Datos enlazados “Linked Data”.
- Datos Abiertos “Open data”.
- Datos enlazados y abiertos Open Linked Data.

- Cloud computing.
- Smart Cities.
- Internet de las Cosas.
- Turismo y tecnologías.

Definir la situación de Arquitectura común de las empresas. Definir la estructura y los componentes del modelo que permita extracción, filtrado, enriquecimiento y publicación de los datos sobre el turismo de Norte de Santander a partir de las necesidades del negocio.

Durante esta actividad se analiza la situación de las arquitecturas actuales de las organizaciones, con el fin de determinar una arquitectura común y a partir de la misma generar mecanismos y estrategias tanto conceptuales como tecnológicas que permitan la implementación de Open Linked Data.

Determinar las tecnologías que permitan la adquisición de contenidos desde diferentes fuentes de datos estructurados y no estructurados. La extracción, filtrado y edición de los mismos con el fin de darle enriquecimiento semántico y ofrecerlo como servicios a terceros.

Definición del Contexto del negocio. Identificar el contexto del sector turístico y establecer los actores involucrados y los intereses sobre la plataforma.

Realizar Inventario de Información Turística. Recolección de datos de inventario turístico de Norte de Santander, además de procesarlos.

Definición de formatos y diseño de ontología. Definir los formatos a utilizar y definir el esquema ontológico para la información.

Definición y configuración de plataformas. Identificación de las plataformas a utilizar y el proceso de instalación y configuración.

Diseño de Arquitectura. Definir la arquitectura de integración de datos, y el acceso a servicios.

Desarrollo del Prototipo. Diseñar y documentar un prototipo App sobre los puntos de interés turísticos de Norte de Santander que se obtenga a partir de la información publicada en la plataforma.

En esta fase el objetivo es realizar el diseño de un prototipo App para dispositivos móviles que obtenga la información de los puntos de interés turísticos. Durante esta fase se llevaran a cabo las siguientes actividades:

Identificación de requerimientos: En esta actividad cuales son las funcionalidades que debe satisfacer el prototipo de la App, además de la identificación de servicios y fuentes externas.

Diseño de Servicios: En esta actividad se realiza el modelado de interfaces, clases, entidades, acceso a servicios, etc. Usando un lenguaje de modelado unificado UML.

Prototipado: Usando lenguajes, técnicas y estándares se construye un prototipo que cumpla con los requerimientos y el diseño propuesto.

3. RESULTADOS DE LA INVESTIGACIÓN

RESUMEN. Actualmente, un gran número de almacenes de datos abiertos, catálogos y portales han ido surgiendo en diferentes ámbitos. En este capítulo se definió un modelo conceptual que permite la implementación de linked open data desde la etapa de adquisición, filtrado, enriquecimiento y publicación de datos en el sector de turismo de Norte de Santander, a partir del modelo se diseñó el despliegue de la plataforma, la arquitectura del portal y las herramientas que se han utilizado para la apertura de los datos. A demás información sobre los vocabularios u ontologías que se utilizaron y la generación de un servicio SPARQL Point para acceder a los datos.

Se diseñó un prototipo App de los puntos de interés turístico del departamento Norte de Santander, que son consumidos desde los datos publicados a través del portal.

3.1 IMPLEMENTACION DE LA INICIATIVA LINKED OPEN DATA

3.1.1 Análisis Para Transformación A Linked Open Data. Las iniciativas Linked Data, Open Data y Open Data Linked (LOD), han sido muy populares en la comunidad de la web semántica hoy día. Y varias de las iniciativas, como openspending.org¹¹⁰, están creciendo rápidamente con el fin de promover la apertura de los datos con mayor transparencia en las instituciones y diferentes sectores. Con la aparición del paradigma Linked Open Data (LOD) se democratizó el enfoque de la apertura de datos y la interconexión de los contenidos semejantes con relaciones equivalentes y que se encuentran en diferentes ubicaciones ya sean geográficamente o internamente en la organización.

A nivel de las organizaciones, el paradigma Linked Open Data (LOD) brinda nuevos desafíos científicos y técnicos para dar respuestas a las exigencias semánticas en la integración de datos de negocios. El impacto de la incorporación del paradigma LOD en las empresas genera cambios profundos que se genera en los procesos estratégicos de la empresa, especialmente en el proceso de transformación de los conjuntos de datos. Permitiendo, el enriquecimiento semántico y la interconexión de datos con el fin de optimizar el ciclo de vida de los datos de negocio, disminuyendo el tiempo de integración de datos y el coste. Por otra parte, cuando las organizaciones gestionan los datos semánticamente desde el inicio, es decir, desde su adquisición o creación, se requiere de menos tiempo y

¹¹⁰ <https://openspending.org/>

esfuerzos para procesar e integrar en las aplicaciones de negocio. Esta gestión semántica implica una serie de procedimientos y técnicas como la identificación de los datos, conexiones con otros datos, etc. usando estándares del W3C¹¹¹.

A partir de las técnicas de Linked Open Data se busca fomentar la creación de aplicaciones y servicios avanzados con el uso de datos heterogéneos que se generan en el que hacer diario del sector de turístico de Norte de Santander, en la práctica muchas de las fuentes de datos existentes en las organizaciones provienen:

- A partir de fuentes internas como CRM, ERP, DBMS, sistemas de archivos.
- a partir de fuentes externas, como correos electrónicos, páginas web, redes sociales, foros.

Como consecuencia de la cantidad de fuentes heterogéneas con las que se puede contar y especialmente en un sector que reúne muchos actores interesados tanto públicas como privadas, se hace necesario la inclusión de iniciativas como Linked Open Data, que brinda canales innovadores para consumir, explotar y rentabilizar los datos y sus activos. Para entender mejor la iniciativa, en la Figura 56. Se muestra el ciclo de vida de los datos semánticos desde la adquisición hasta el consumo final, que se puede implementar para dar solución no solo al sector turístico de Norte de Santander, sino a cualquier otra organización que quiera implementar la iniciativa LOD.

Figura 57. Flujo de Trabajo de datos en una aplicación empresarial



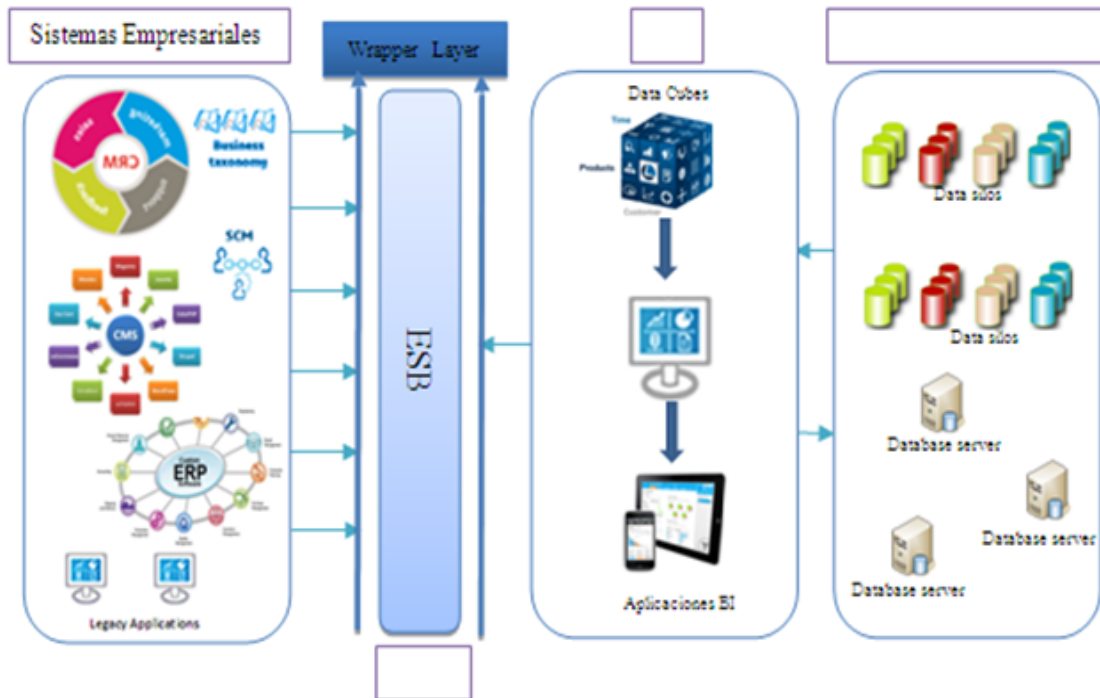
Fuente: Adaptación de LOD2¹¹².

¹¹¹ <http://www.w3.org/>

¹¹² Building Enterprise Ready Applications Using Linked Open Data. Amar-Djalil Mezaour, Bert Van Nuffelen, and Christian Blaschke. 2014

3.1.2 Panorama Actual De Los Datos En Las Organizaciones. La integración de datos y el uso eficiente de la información que tienen las organizaciones es uno de los principales retos actuales. Normalmente una organización maneja aplicaciones de diferentes proveedores y que se ejecutan en diferentes tecnologías, plataformas y se comunican a través de diferentes conexiones y protocolos al interior y exterior de la organización. A partir de estas aplicaciones se generan fuentes de datos heterogéneas, y conjuntos de archivos que conllevan enormes costos a la hora de integrarlos. Para gestionar esta complejidad, una empresa desde el ecosistema tecnológico es vista como un conjunto de aplicaciones interconectadas o parcialmente conectadas que gestionan los diferentes procesos de la empresa, donde la separación de las aplicaciones a menudo conlleva una reproducción de los mismos datos en diferentes formas. Además cada proceso manipula diferentes tipos de datos y produce nuevos datos en una forma estructurada o no estructurada en la figura 57 se ve una arquitectura clásica de una organización.

Figura 58. Arquitectura Tradicional de Un sistema de Información de una Organización



Fuente: Adaptación de Adaptación de LOD2.

En la actualidad existen enfoques tecnológicos, tales como la Integración de Aplicaciones Empresariales (EAI), en el que se crea una capa (middleware) entre las distintas fuentes de información haciendo uso de varios modelos arquitectónicos; por ejemplo: Event Driven Architecture (EDA) o Arquitectura Orientada a Servicios (SOA), que usualmente se implementan con servicios web y SOAP. Otros métodos comunes utilizados generalmente en las empresas incluyen DataWarehouse y Gestión de datos maestros.

Mensajes simples en XML y otros estándares B2B que aseguren el flujo de información a través de los sistemas internos de una manera sencilla y eficiente. Pero hay casos en los que no es suficiente, Para organizaciones que manejan grandes cantidades de información y servicios se les hace imposible describir procesos complejos con XML. Por lo que hay una evidente necesidad de proveer definiciones claras o semánticas a los datos para facilitar la integración en la capa de datos.

Sin embargo para las organizaciones la integración de los datos a nivel de la capa de datos no es una tarea sencilla de resolver y el paradigma Linked Data plantea soluciones a algunos de los problemas de la integración de datos, los dos enfoques tecnológicos planteados EAI y LOD no son contradictorios si no que son complementarios. La arquitectura SOA implementada en EAI trabaja orientada a servicios, mientras que LOD trabaja con enlaces a recursos (Datos, conjunto de datos, documentos, etc.). En las arquitecturas SOA se utilizan servicios personalizados especialmente para tareas específicas, en cambio la arquitectura LOD utiliza solo servicios (SPARQL, REST) y los enlaces a los recursos mencionados. Ambos métodos tienen el esfuerzo complementario de estandarización (metadatos vs servicios) y en esta tarea donde cada método es más adecuado para diferentes tareas, es por eso que el enfoque EAI- SOA se le puede sacar mayor provecho para tareas o servicios bien definidos y personalizados, mientras que el enfoque LOD está más dirigido a tareas innovadoras que implican la semántica como son (integraciones, mapeos, reportes, etc.).

3.1.3 Beneficios al Migrar de una Arquitectura Tradicional A Linked Open Data LOD. Los beneficios de la adopción del enfoque Linked Data en las organizaciones son multidimensional entre ellos tenemos:

- Abordar el problema de la heterogeneidad de los datos y la integración dentro de la organización.

- Crear cadenas de valor en el interior y a través de otras organizaciones. Generando la posibilidad para desarrollo de aplicaciones de terceros.
- Incrementar el valor de los datos existentes y crear nuevos conocimientos utilizando técnicas BI y análisis predictivo.
- Obtener ventajas competitivas al realizar una adaptación temprana de las tecnologías LOD.
- La implementación de Linked data sin la necesidad de cambiar la arquitectura existente.
- Dar significado a los datos de tal manera que permite la búsqueda de información pertinente.

Estos beneficios se detallan en la siguiente Figura tomado del informe de Deloitte " Open data: Driving growth, ingenuity and innovation"¹¹³

¹¹³Deloitte. <http://www2.deloitte.com/content/dam/Deloitte/uk/Documents/deloitte-analytics/open-data-driving-growth-ingenuity-and-innovation.pdf> (pág. 12). (consultado el: 26-03-2015).

Figura 59. Beneficios para las organizaciones al adoptar LOD

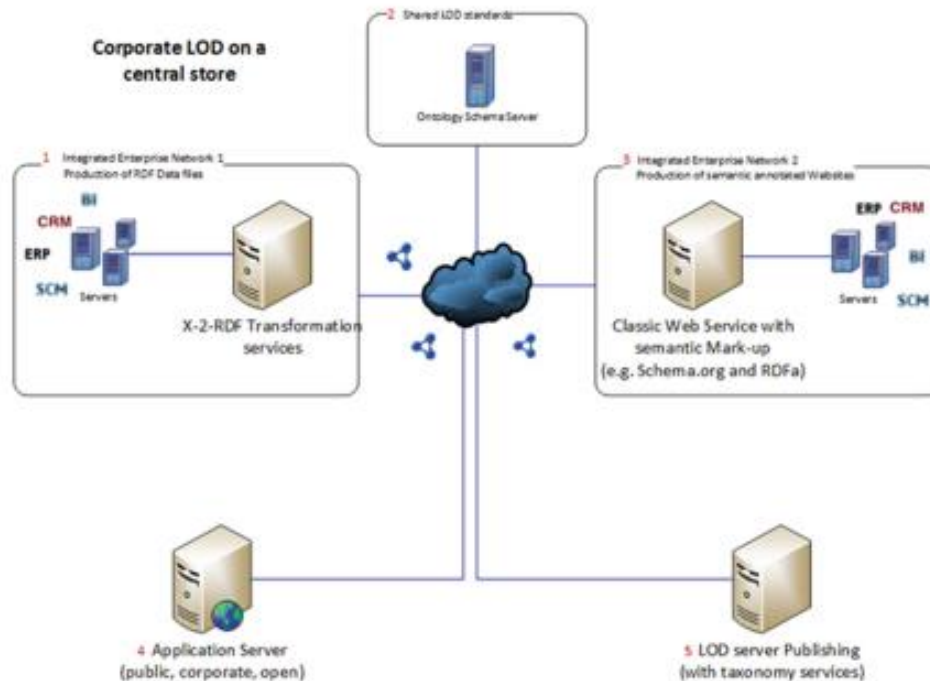


Fuente: Deloitte¹⁰⁹.

3.1.4 Proceso de Integración de LOD a Arquitecturas Empresariales. Cuando se adoptan los principios de LOD, la arquitectura de TI clásica empresarial (figura 59) se debe modificar de tal manera que permita trabajar a través en Internet superando las barreras técnicas de los formatos y las diferencias semánticas de intercambio y manipulación de datos. A continuación se describe las modificaciones que se deben llevar a cabo en una arquitectura tradicional para el procesamiento de los datos para LOD. Ver las figuras siguientes:

3.1.4.1 Reemplazar el Enterprise Software Bus (ESB) por Protocolos de LOD en Servidores Externos. Se describe a continuación

Figura 60. Implementación de LOD corporativo



En la figura 60 Se ilustra la arquitectura empresarial LOD, donde se reemplaza Enterprise Software Bus (ESB) de la arquitectura IT clásica con una nube LOD (cloud LOD). Esta arquitectura muestra dos tipos de publicación de datos, la empresa publica sus datos RDF en un servidor LOD externo (servidor 5, Figura 60) en los siguientes dos escenarios.

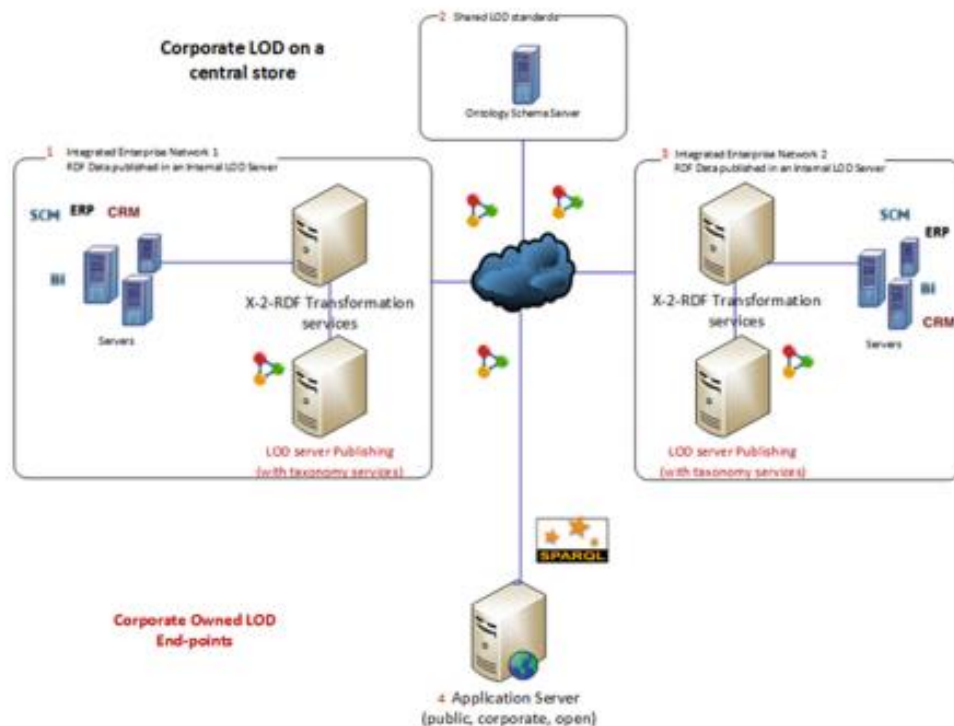
1. Un conjunto de datos RDF se generan desde diferentes fuentes de datos y subsistemas (caja 1 en la figura 60) y se transfiere a un servidor central de LOD externo.
2. Los metadatos se suben a un sitio web clásico utilizando anotaciones semánticas de metadatos como (RDFa, Schema.org¹¹⁴) en las páginas HTML (caja 3 de la Figura 59). Un servidor de LOD extrae estos metadatos, los organiza y los hace accesible como un servicio central (caja 4 en la Figura 60).

¹¹⁴ <https://schema.org/>

El servidor Ontology Schema (caja 2 en la Figura 60) almacena las ontologías usadas para capturar la semántica de LOD. Estas pueden ser estandarizadas o creadas para necesidades particulares. Otros servidores de aplicaciones o plataformas (servidor 4 en la Figura 60) pueden utilizar los servicios del LOD central para construir cálculos e informes específicos. A partir de estos datos generar nichos de mercado o encontrar nuevas oportunidades de negocio que se pueden crear a partir de las visualizaciones y agrupamiento de datos.

3.1.4.2 Sustitución del Enterprise Software Bus (ESB) con Protocolos de Linked Open Data entre los Servidores de Publicación LOD de la Empresa

Figura 61. Funcionamiento un servidor LOD para publicar sus propios datos



En la figura 61. El proceso fundamental se da cuando la empresa pone en funcionamiento un servidor LOD para publicar sus propios datos. Por otra parte, la ilustración permite la integración entre diferentes redes sin importar si pertenecen a diferentes áreas de la organización u otra organización. En la Figura 61. Esta arquitectura muestra dos tipos de publicación de servicios LOD (recuadro 1 y 3 en la Figura 61). El RDF publicado se pone en la plataforma de la empresa, en el

servidor (corporativo o empresarial). Otras plataformas de aplicaciones de servicios (servidor 4 en la Fig. 5) pueden utilizar los servicios LOD para construir cálculos e informes específicos. Tales aplicaciones de servicios pueden estar en una plataforma externa dedicada o pueden estar en una o más de las plataformas LOD de la organización. El servidor Ontology Schema (cuadro 2 en la Fig. 5) almacena las ontologías usadas para capturar la semántica de los datos abiertos enlazados. Estas pueden ser estándar o personalizadas.

3.2 BUENAS PRÁCTICAS PARA LINKED OPEN DATA

3.2.1 Identificación de Fuentes de Datos. La información corporativa se puede definir como los datos que utilizan y comparten los diferentes empleados, departamentos y procesos de una empresa. Dependiendo de la política de seguridad que haya definido la organización para la información, se puede acceder, procesar y publicar a través de diferentes aplicaciones de negocio del sistema TI. Hay que tener en cuenta que el alcance de los datos se puede extender a diferentes lugares (Departamentos internos y entidades, información de en otras ciudades y en otros países.).

Al integrar las tecnologías LOD en un sistema o aplicación existente en la organización, la primera recomendación es llevar a cabo un diagnóstico y auditoría sobre las diferentes fuentes de datos de negocios utilizados por la organización. Esta auditoría debe incluir los siguientes elementos:

- Clasificar las fuentes de datos del negocio de acuerdo a la importancia para la operación de los procesos de mercado estratégicos.
- Hacer una cartografía entre el flujo de trabajo de datos y las fuentes de datos identificadas para descubrir si falta, es redundante o incompleta la información intercambiada, el tipo de datos (estructurados, no estructurados), etc.
- Mapeo entre las Tablas y los formatos de datos nativos de la organización y los formatos estándares correspondientes (preferiblemente RDF de la W3C) y el impacto de pasar de un formato original a un formato estándar.

Con este diagnóstico y auditoría les permite a los arquitectos de datos comprender mejor el funcionamiento de las aplicaciones que tiene la organización y les ayuda

a evaluar el costo de la integración de la tecnología LOD. De acuerdo al esfuerzo requerido y el costo, la primera buena práctica consiste en la migración tanto como sea posible de los formatos nativos a los estándares, preferiblemente RDF-Like de la W3C. Esto facilita considerablemente la publicación, anotación y la interconexión de los datos empresariales.

Para cumplir con el criterio de apertura del paradigma LOD, la publicación de los datos corporativos es una de las principales recomendaciones en el proceso. Para ello, se debe definir un sistema de licencias que permita definir cómo los datos abiertos se pueden reutilizar y explotar por terceros usuarios, aplicaciones y servicios. Teniendo en cuenta el interés de la organización, deben establecerse acuerdos para abrir la cantidad de datos como sea posible y mantener un buen equilibrio entre la conservación de los datos empresariales confidenciales estratégicos, y el resto de los datos abiertos.

Por último, pero no menos importante, el sistema de licencias de datos abierto debe garantizar el principio de reutilización de los datos de las aplicaciones de terceros con el menor número de restricciones técnicas, financieras y jurídicas posibles. Una forma de lograr estos objetivos es proporcionar descripciones robustas de metadatos de los datos abiertos con vocabularios apropiados, como DCAT¹¹⁵, VoID¹¹⁶, DublinCore¹¹⁷, etc. Para que los datos abiertos y publicados sean comprensibles y recuperables, la descripción de metadatos deben proporcionar elementos clave como los derechos de autor y la licencia asociada, frecuencia de actualización de datos, formatos de publicación, procedencia datos, versión de los datos, descripción textual del conjunto de datos, punto de contacto cuando sea necesario para reportar inconsistencias o errores, etc.

3.2.2 Modelado de Dominios Específicos. Con el fin de transformar el modelo existente de una empresa en esquemas interoperables, las mejores prácticas se centran en el uso de vocabularios comunes. Usando términos de vocabularios existentes es más fácil para publicar y contribuye mucho en la reutilización y el intercambio de información continua de los datos empresariales.

Como primer paso, la estructura inherente de los datos heredados tiene que ser analizado. Si no existe una jerarquía especificada, a menudo puede ser creado por un experto en conocimiento de datos. Si la organización de los datos como tal no es posible, entonces sólo una lista de conceptos, básicamente un glosario, se

¹¹⁵ <http://www.w3.org/TR/vocab-dcat/>

¹¹⁶ <http://www.w3.org/TR/void/>

¹¹⁷ <http://dublincore.org/>

puede construir. Dependiendo de la complejidad de los datos y cómo se relacionan las entidades, diferentes esquemas de datos se pueden utilizar para expresarlas.

3.2.3 Migrar Datos Heredados a Vocabularios. La migración de un vocabulario existente a un esquema RDF varía en complejidad dependiendo de cada caso, pero hay pasos que son comunes en la mayoría de las situaciones. La transformación de los datos empresariales a RDF requiere:

- La traducción entre el modelo fuente y el modelo RDF es una tarea compleja y con muchos mapeos alternativos. Para minimizar los problemas, la solución que se utilice debe proteger la semántica.
- La entidad básica de un RDF es un recurso y todos los recursos tiene que tener un identificador único, una URI para este caso. Si los datos en sí no proporciona identificadores que se pueden convertir en URI, se debe desarrollar, una estrategia para la creación de URI para todos los recursos que van a ser generados.
- Preservar la denominación original, tanto como sea posible. Preservar la denominación original de entidades resulta en conversiones más claras y trazables. Nombres de propiedades duplicadas como prefijo con el nombre de la entidad de origen para que sean únicos.
- Usar soporte XML para todo los tipos de datos. Tipos de datos incorporados en esquema simples XML como xsd:date y xsd:integer son útiles para suministrar esquemas con información de propiedades.

Dominios específicos de datos, pueden ser modelados como vocabularios con Org¹¹⁸ o GoodRelations¹¹⁹. Sólo cuando los vocabularios existentes no cubren las necesidades se deben desarrollar nuevos esquemas. Los conjuntos de datos que se publicarán en la Web deben ser descritos con vocabularios de meta datos como VoID.

¹¹⁸ <http://www.w3.org/TR/2014/REC-vocab-org-20140116/>

¹¹⁹ <http://www.heppnetz.de/projects/goodrelations/>

Cuando no existen vocabularios adecuados para describir los datos del negocio, una posibilidad es desarrollar un diccionario skos¹²⁰ en vez de un modelo con RDFs. Este enfoque es más fácil de seguir en las organizaciones nuevas. Existen herramientas como PoolParty¹²¹ que soportan estas tareas. El más reciente estándar internacional para el desarrollo de diccionarios es la ISO 25964¹²². Esta norma proporciona las directrices y las buenas prácticas que se deben tener en cuenta.

Una vez que los datos están en este formato se puede cargar en triple store like Virtuoso y publicarlo internamente o en la Web.

3.3 PUBLICACIÓN

Los procedimientos para publicación de LOD empiezan con la identificación de fuentes de datos y la fase de modelado y continua con la descripción de los datos como RDF y el almacenamiento y servicio de los datos. Una variedad de herramientas de diferentes proveedores y con características especiales se han creado para asistir los diferentes aspectos de esta fase. De acuerdo a las necesidades específicas de cada caso de negocio y la naturaleza de los datos de la empresa. Se pueden crear patrones de publicación más sencillos.

3.3.1 Patrones de Publicación para Datos Relacionales. Las Bases de datos relacionales (RDB) son el principal activo de las organizaciones actualmente y seguirá siendo una fuente predominante de los datos en las mismas. Por lo tanto el interés de la comunidad científica^{123,124}, se ha centrado en torno al desarrollo de enfoques y técnicas de mapeo para pasar de RDB a datos RDF. Estos enfoques permitirán a las empresas:

- Integrar su RDB con otra fuente estructurada en RDB, XLS, CSV, etc. (o no estructurada HTML, PDF, etc.), por lo que deben convertir RDB a RDF y asumir cualquier otra fuente estructurada (o no estructurada), también puede estar en RDF.

¹²⁰ <http://www.w3.org/2004/02/skos/>

¹²¹ <http://www.poolparty.biz/>

¹²² <http://www.niso.org/schemas/iso25964/>

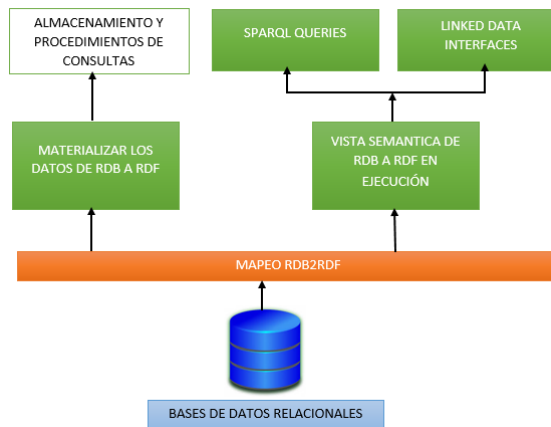
¹²³ <http://www.w3.org/2011/10/rdb2rdf-charter.html>

¹²⁴ <http://www.w3.org/DesignIssues/RDB-RDF.html>

- Integrar el RDB con el RDF en la web actual (Linked Data) por lo que debe convertir a RDF para luego poder vincular e integrar.
- Hacer que los datos RDB estén disponibles para SPARQL u otro RDF-base query, y/o para que otros puedan integrar con otras fuentes de datos (estructurados, RDF, no estructurado).

Dos aspectos importantes que deben tomarse a consideración y abordarse en la organización (ver Figura 62):

Figura 62. Mapeo de Datos de RDB a RDF

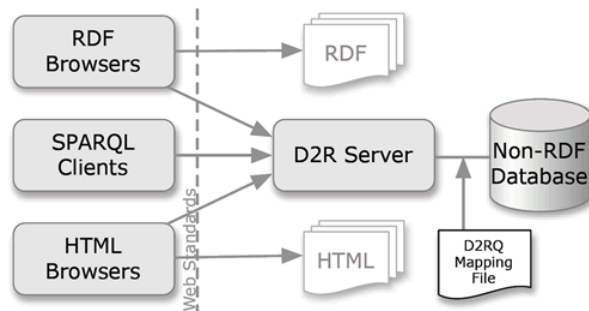


3.3.2 Definición del Lenguaje de Mapeo desde RDB2RDF. Mapeos automáticos proporcionados por herramientas como D2R¹²⁵ y Virtuoso RDF Views proporcionan un buen punto de partida, especialmente en los casos en que no hay una Ontología de dominio para mapear el esquema relacional. Sin embargo, es común la definición manual de los mapeos, para esto es necesario que los usuarios declaren la semántica del dominio en la configuración del mapeo y aprovechen la integración y la vinculación de Linked Data. R2RML¹²⁶, un lenguaje recomendado por la W3C para expresar los mapeos personalizados, se puede contar con varias herramientas como Virtuoso RDF Views y D2R.

¹²⁵ <http://d2rq.org/d2r-server>

¹²⁶ <http://www.w3.org/TR/r2rml/>

Figura 63. Transformación de datos



Fuente: BIZER, Chris. Cómo publicar datos enlazados en la Web. [En línea]. Disponible en Internet: <http://wifo5-03.informatik.unimannheim.de/bizer/pub/LinkedDataTutorial/>

3.3.3 Materializar los Datos. Una característica común de la herramienta RDB2RDF es la posibilidad de crear una "vista semántica" de los contenidos de la base de datos relacional. Es decir, una versión RDF que se genera desde la base de datos para ser publicada a través de un punto final SPARQL y una interfaz de Linked Data que trabaja directamente en la parte superior de la fuente de datos relacional, a partir de una "vista" virtual de la base de datos. Tal "vista semántica" garantiza el acceso hasta la información de actualización de los datos, lo cual es particularmente importante cuando los datos se actualizan con frecuencia. En contraste, la generación y almacenamiento RDF requiere la sincronización cuando ya estén modelados los datos.

3.3.4 Patrón Publishing para Excel / Datos CSV. Cuando los datos originales residen en formato Excel o CSV, describirlos en RDF sería un primer paso de un patrón de publicación, para luego alojarlo y servirlo en la web. LODRefine es una pila de componentes, muy adecuado para la automatización y facilitar el procedimiento "RDFizing". Su uso aporta un valor añadido directo al negocio:

- Capacidad de limpieza a gran escala de los datos originales del negocio.
- Las capacidades de conciliación, en caso de ser necesario, para encontrar datos similares en la nube LOD y hacerlos compatibles con los datos originales de negocio es decir con las fuentes Linked Data.
- Aumentar las capacidades, permitir agregar columnas desde DBpedia u otras

fuentes al conjunto de datos original, basada en la reconciliación de servicios mencionado anteriormente.

- Facilidades de extracción de entidades cuando residen en el interior del texto de las celdas.

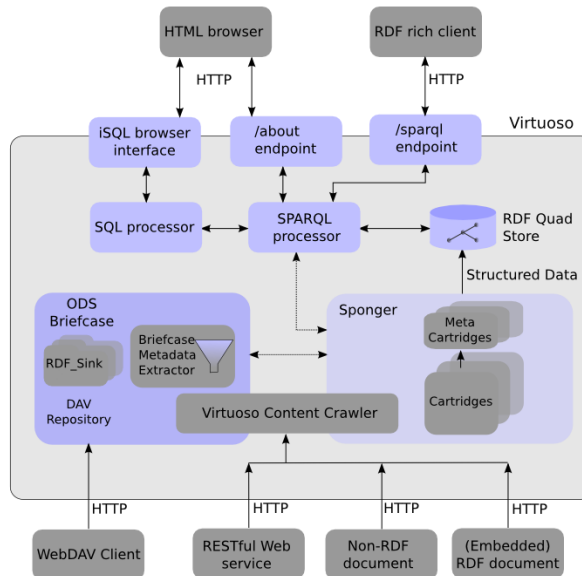
3.3.5 Patrón Publishing para Datos XML. Cuando los datos originales están en formato XML la solución adecuada es usar un XSLT para transformar el documento XML en un conjunto de tripletas RDF. Los archivos originales no cambian; sino que se crea un nuevo documento basándose en el contenido del existente. La idea básica es que las estructuras específicas son reorganizadas y se transforman en tripletas con un recurso determinado, en predicados y valores. La pila LOD2 soporta transformaciones XML a RDF / XML XSLT. Las tripletas resultantes se guardan como un archivo RDF / XML Graph y se pueden almacenar para su publicación.

3.3.6 Patrón Publishing para Datos no Estructurados. A pesar de la evolución de las instalaciones de almacenamiento, el entorno empresarial sigue manteniendo el paradigma de repositorios de contenidos no estructurados y semi-estructurados. Información corporativa básica y el conocimiento se almacena en una variedad de formatos tales como PDF , Archivos de texto, correos electrónicos, sitios Web tradicionales o anotaciones semánticas, que provienen de la Web 2.0 como las redes sociales o de API específicas como Geonames¹²⁷, Freebase¹²⁸ etc. Herramientas de extracción y generación de Linked Data y mapeo de instancias de datos a ontologías adecuadas es la ruta para producir datos RDF y facilitar la consolidación de la información de la empresa. Una Herramienta destacado de la pila LOD2 que facilitan la transformación de estos tipos de datos a grafos RDF es Virtuoso Sponger.

¹²⁷ <http://www.geonames.org/>

¹²⁸ <http://www.freebase.com/>

Figura 64. Sponger Inputs y Outputs



Fuente: Virtuoso Universal Server

Virtuoso sponger¹²⁹ es un middleware Linked Data que genera Linked Data a partir de una gran variedad de formatos no estructurados. Su funcionalidad básica se basa en Cartuchos, que cada uno proporciona la extracción de datos a partir de varias fuentes de datos y capacidades de mapeo a ontologías existentes. Las fuentes de datos pueden estar en RDFa format¹³⁰, GRDDL¹³¹, documentos de Microsoft, y microformatos¹³² o puede ser fuentes específicas y otros datos proporcionados por los proveedores de API. Los cartuchos son muy personalizables para permitir la generación de los datos enlazados estructurados desde prácticamente cualquier tipo de recurso.

El PoolParty Thesaurus Server¹³³ se utiliza para crear tesauros y otros vocabularios y ofrece la posibilidad de publicar al instante como HTML, además proporcionar versiones RDF legible por maquinas a través de la negociación de contenido. Esto significa que cualquier persona que utilice Poolparty puede convertir a normas W3C sin tener que saber nada acerca de aspectos técnicos de

¹²⁹ <http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/VirtSponger>

¹³⁰ <http://rdfa.info>

¹³¹ <http://www.w3.org/TR/grddl/>

¹³² <http://microformats.org/>

¹³³ <http://www.poolparty.biz>

la Web Semántica. El diseño de todas las páginas front-end de datos enlazados puede ser controlado por generador que puede utilizar sus propias hojas de estilo y crear diferentes vista sobre los datos con velocity templates.

Stanbol¹³⁴ es otra herramienta para extraer información de un CMS u otra aplicación web con el uso de una API y lo representa como RDF. NIF Stanbol pronto se convertirá en una recomendación del W3C.

DBpedia Spotlight¹³⁵ es una herramienta para anotaciones automáticas de menciones de texto de recursos DBpedia, que proporciona una solución para vincular las fuentes de información no estructurada en la nube Linked Open Data a través de la DBpedia.

3.4 SERVIDOR Y ALMACENAMIENTO

La fase de publicación generalmente implica los siguientes pasos:

1. Almacenar los datos en una Triple Store.
2. Ponerlos a disposición en un SPARQL endpoint.
3. Hacer las URIs Dereferenceable para que las personas y las máquinas pueden buscarlos como Web.
4. Proporcionarles volcado RDF de modo que los datos se pueden volver a utilizar fácilmente.

Los tres primeros pasos pueden ser plenamente abordados con un componente de pila LOD2 llamado Virtuoso, mientras que subir el archivo RDF a CKAN¹³⁶ sería el procedimiento para que el RDF quede público.

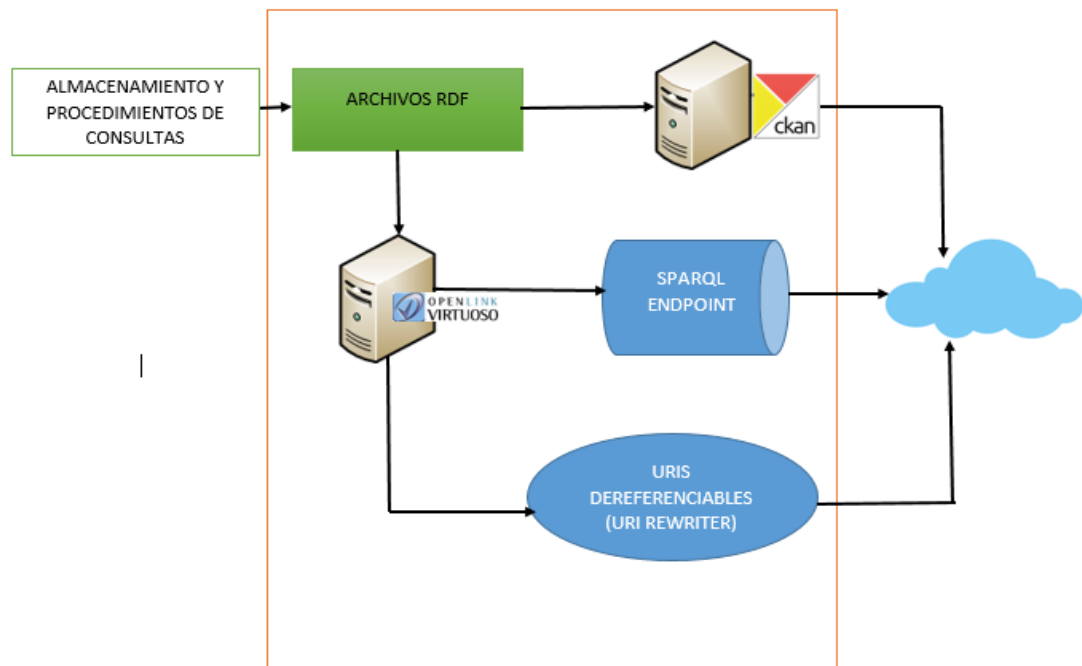
¹³⁴ <http://stanbol.apache.org/>

¹³⁵ <https://github.com/dbpedia-spotlight/dbpedia-spotlight/wiki>

¹³⁶ <http://ckan.org/>

OpenLink Virtuoso Universal Server es una arquitectura híbrida que puede ejecutarse como almacenamiento para múltiples modelos de datos, tales como datos relacionales, RDF, XML y documentos de texto. Virtuoso es compatible con una interfaz de gestión de repositorios y faceted browsing de los datos. Puede funcionar como un servidor de documentos Web, servidor de Linked Data y el servidor de aplicaciones Web. La versión de código abierto de Virtuoso se incluye en la pila LOD2 y es ampliamente utilizado para la carga de datos en su Quad Store, que ofrece una SPARQL endpoint y un mecanismo llamado URL- Rewriter hacer URIs Dereferenceable. De acuerdo con el cuarto paso, el intercambio de los datos en un DataHub ampliamente conocido como CKAN facilitará su descubrimiento de otras empresas y editores de datos. La funcionalidad de CKAN se basa en paquetes donde los conjuntos de datos pueden ser subidos. CKAN permite también actualizaciones, realiza un seguimiento de los cambios, las versiones y la información del autor. Se recomienda como buena práctica para acompañar a los conjuntos de datos con archivos de información (por ejemplo, archivos VOID) que contienen metadatos relevantes. Ver la figura 65.

Figura 65. Modelo de publicación y registro de datos



Fuente: Adaptación LOD2.

3.5 MAPEO DE VOCABULARIO

A veces, una empresa puede necesitar desarrollar una ontología propia para aplicar los principios de Linked Data. Mapeando los términos que se han utilizado para la publicación de los triples con términos que existen en los vocabularios existentes lo que facilitará el uso de los datos de la empresa en aplicaciones de terceros. Una herramienta que apoya este tipo de mapeo es R2R¹³⁷.

3.6 CASO DE ESTUDIO: ENFOQUE DE OPEN Y LINKED DATA VINCULADO AL SECTOR TURÍSTICO DE NORTE DE SANTANDER

El Sector Turístico de Norte de Santander quiere incentivar el uso de datos abiertos y enlazados con el fin de mejorar la oferta en desarrollo de servicios y aplicaciones de terceros, como parte del proceso natural de la transformación de la información y de iniciativas internacionales para el sector. De acuerdo con esto, el desarrollo de iniciativas de datos abiertos y enlazados debe ser considerada como una prioridad dentro de la estrategia global del sector, con el apoyo de la herramientas tecnológicas necesarias que puedan fomentar la creación de una plataforma que puede llegar a ser la columna vertebral del modelo de desarrollo económico actual. En las siguientes secciones se describen, los pasos que se consideraron para el desarrollo del caso de estudio de los datos abiertos y enlazados para el sector turismo en Norte de Santander. Además, el modelo de negocio que a través de esta estrategia se pueda generar.

3.6.1 Condiciones del Entorno. Durante los últimos años, el diseño e implementación de políticas turísticas ha sido un factor fundamental para mantener un esfuerzo sostenido hacia la consolidación y el desarrollo regional. El departamento se ha caracterizado por un turismo de frontera dada la colindancia con el vecino país de Venezuela, lo cual estimula los sectores básicos de la economía de la región. Una economía de frontera se caracteriza por tener una estructura y unos agentes que centran las actividades en el intercambio de bienes y servicios, de los cuales hacen parte el comercio, las finanzas, el transporte y el turismo.

3.6.2 Plan de Desarrollo Turístico de Cúcuta. San José de Cúcuta es el único que tiene Plan de Desarrollo Turístico 2011-2020. Realizado en el año 2010 y socializado en marzo del 2011, su objetivo era la identificación de la vocación turística de la ciudad, dejando a Cúcuta como la frontera de los congresos,

¹³⁷ <http://wifo5-03.informatik.uni-mannheim.de/bizer/r2r/>

convenciones y las compras. Este trabajo fue realizado por la firma Tourism Consulting.

3.6.3 Plan Regional de Competitividad. El Programa Nacional de Competitividad recoge la iniciativa que el departamento estructuró en 2008 y planteó en 2009, en la que se estableció con claridad una ruta estratégica de desarrollo y competitividad. El Plan Regional de Competitividad del Norte de Santander (PRCNS) es un documento dinámico y concertado entre los sectores público, privado y la sociedad civil, en donde se contemplan las necesidades del departamento en materia de competitividad con el fin de identificar las iniciativas y gestionarlas con el objetivo de generar un impacto positivo en esta materia. En mesas de trabajo se priorizaron las siguientes apuestas productivas, en las que el turismo hace parte activa:

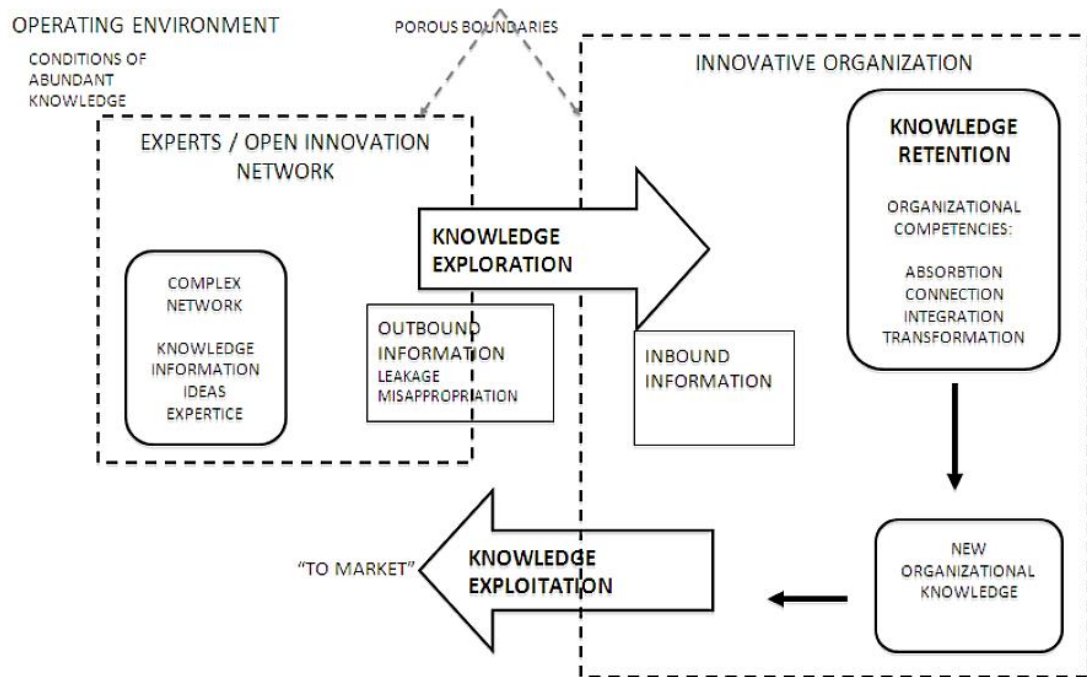
Figura 66. Apuestas Productivas del Plan de competitividad

APUESTAS PRODUCTIVAS PLAN REGIONAL DE COMPETITIVIDAD			
1	Palma de aceite	7	Cuero, calzado y sus manufacturas
2	Forestal, madera y muebles	8	Confecciones
3	Hortofrutícola	9	Turismo
4	Cacao - chocolate	10	Infraestructura - transversal
5	Carbón	11	Educación - transversal
6	Arcilla	12	Ciencia y tecnología - transversal

Fuente: Plan de competitividad.

3.6.4 Stakeholders. Incluyendo en el modelo de negocio para el sector turístico de Norte de Santander un ecosistema de datos abiertos basada en un escenario de innovación abierta, que puede crear tensiones estructurales debido a la información procedente de diversas fuentes. Debido a que son muchas las organizaciones que pueden intervenir, puede ser difícil de compartir la misma visión entre todos los grupos de interés. La razón principal es que la innovación abierta tiene un enfoque mucho más complejo que necesita más esfuerzos dentro del modelo actual, tales como una mejor comunicación, la participación y la interacción entre todas las partes interesadas. Todas las partes interesadas deben entender que los datos abiertos y la innovación abierta aparecen como una forma de crear no sólo el valor de la organización, sino también un valor social figura 67. En un sentido estratégico, estos enfoques dirigidos a la exploración, la agregación, la retención y la explotación de conocimiento, donde se necesita una constante participación de los interesados y la interacción.

Figura 67. Stakeholder engagement in an open innovation scenario



Fuente: Open Innovation¹³⁸

Teniendo en cuenta esto, en el apartado siguiente se escribe la lista de las posibles interesadas en el ecosistema abierto, y la forma como deben interactuar e integrarse.

¹³⁸ Robert Wayne Gould. "Open Innovation and Stakeholder Engagement." (2012), (pg. 6). J. Technol. Manag. Innov. 2012, Volume 7, Issue 3.

Cuadro 5. Descripción los stakeholders

Interesados	Intereses en el proyecto
Autor del proyecto	Que se cumpla con el objetivo del proyecto y con lo pactado en alcance.
Universidad Autónoma de Bucaramanga, Grupo prisma	Que se desarrollen proyectos de investigación de alto impacto para el país.
Estudiantes de pregrado y postgrados	Juegan un papel clave en el ecosistema de datos enlazados definido. Los estudiantes pueden participar directamente en iniciativas para impulsar la innovación como hackatons o concursos de aplicaciones, donde pueden explorar los datos para generar ideas que puedan crear nuevos servicios o productos. Para fomentar la participación de los estudiantes, estas iniciativas deben estar enfocadas a la enseñanza, la investigación, el espíritu empresarial, cultural o actividades externas.
Desarrolladores externos y analistas de aplicaciones	La idea principal del enfoque de la innovación abierta es conseguir que los innovadores externos pueden utilizar los datos. Desarrolladores y analistas pueden ser independientes de cualquier lugar local o nacional.
Comunidad empresarial y de negocios	La comunidad empresarial y de negocios es parte de los usuarios externos de la plataforma y es importante estimularlos para que se conviertan en participantes y colaboradores activos dentro del ecosistema a través de la comunicación bidireccional como se presenta en la (fig. 10). Esto se puede hacer a través de asociaciones de turismo.
Gobiernos Locales, nacionales e internacionales.	Los gobiernos locales, nacionales e internacionales también pueden ser beneficiados con el proyecto. A nivel local, la publicación de la información del sector turismo Linked Open Data pueden ayudar a abrir un nuevo canal de distribución de datos con diferentes organizaciones y empresas dedicadas al turismo. A nivel internacional, los datos en formatos abiertos y estandarizados permiten que la información pueda ser reutilizada por agentes externos.
Ciudadanos	Los ciudadanos también son beneficiarios del proyecto, ya que tiene un alto nivel de influencia dentro de la sociedad. Los principales resultados del proyecto están relacionados con la creación de nuevas ideas que puede aprovechar el uso de la plataforma, la creación de nuevos productos y servicios que busquen el beneficio de la sociedad.










3.6.5 El Sector Turístico de Norte de Santander como parte de la Web de Datos. Teniendo en cuenta el movimiento de datos enlazados que se está extendiendo a nivel mundial, lo que busca este proyecto es vincular la mayor parte

de los datos publicados sobre el turismo en Norte de Santander, con el propósito de apoyar la creación de redes de datos a nivel nacional. Esta iniciativa de datos abiertos enlazados va más allá del beneficio individual de cada institución. Esto significa que solo publicar información en formatos de datos abiertos no es suficiente, y el uso de datos en formatos enlazados deben ser promovidos como una buena práctica dentro de los esfuerzos de publicación de las diferentes instituciones tanto públicas como privadas y que manejen información de importancia para la sociedad . De acuerdo con esto, dentro del proyecto el caso de estudio del sector turismo del departamento, en el procesos de publicación deben mantenerse los conceptos de normalización (es decir, vocabularios u ontologías) y formatos de acuerdo a los estándares internacionales con el propósito que a futuro se puedan compartir.

3.6.6 Modelo de Negocio de Datos Abiertos del Sector Turismo en Norte de Santander

CANVAS

Generación de modelos de negocio

<p><i>Asociados claves</i> </p> <ul style="list-style-type: none"> - Asociaciones Adscritas al sector turismo. - Empresas. -Emprendedores. - Sector Público. - Centros de Investigación. 	<p><i>Actividades claves</i> </p> <ul style="list-style-type: none"> - Servicios externos. - Servicios de emprendimiento e innovación. - Comunicación y promoción. - Gestión de la economía. - Relaciones <p><i>Recursos claves</i> </p> <ul style="list-style-type: none"> - Portal Web. - Estructura de Red. - Hardware. 	<p><i>Propuesta de valor</i> </p> <ul style="list-style-type: none"> - Datos abiertos y enlazados. - Servicios a Terceros. - API 	<p><i>Relación con el cliente</i> </p> <ul style="list-style-type: none"> - Asociaciones Adscritas al sector turismo (transporte, agencias, etc). <p><i>Canales de distribución</i> </p> <ul style="list-style-type: none"> Empresas. Emprendedores - Sector Público. - Centros de Investigación. 	<p><i>Segmentos de clientes</i> </p> <ul style="list-style-type: none"> Hacketon. Desafíos y Retos -Redes sociales Posicionamiento google. -Medios masivos de Comunicación
<p><i>Estructura de costos</i> </p> <ul style="list-style-type: none"> Configuración e instalación de la plataforma. Formación de personal. Infraestructura 		<p><i>Flujos de ingresos</i> </p> <ul style="list-style-type: none"> Recursos del departamento. Recursos de regalías para ciencia y tecnología. Convocatorias Colciencias 		

@coworkinglemon

Teniendo en cuenta el modelo de negocio anterior sobre la implementación de la iniciativa LOD, podemos analizar cómo este proyecto puede evolucionar positivamente el modelo del sector turístico de Norte de Santander y de la siguiente manera:

De los recursos clave:

Plataforma LOD para:

- Mejora de la propuesta de valor o Incrementar la transparencia actual y la responsabilidad pública.
- El uso de datos abiertos para promover la innovación y la investigación.
- El uso de datos abiertos para generar nuevas oportunidades de negocio.
- El uso de datos abiertos para promover el desarrollo de aplicaciones.
- Mejora de la relación con el cliente o creación de una nueva forma de proporcionar datos de turismo a los emprendedores, empresas e investigadores, en formatos estructurados.
- Contribuir a la integración de los datos del sector turismo con datos vinculados.
- Apoyar el movimiento de datos abiertos y enlazados.
- **Las nuevas actividades clave:**
- Promoción y comunicación para el uso de la plataforma de datos abierta.
- Creación de retos, hackatons, expediciones de datos o actividades similares, a fomentar la exploración de los datos. actividades o emprendimiento a través del uso de la plataforma LOD.

Los socios o Creación de alianzas con otras instituciones:

- Nuevos socios de la comunidad de negocios a través del uso de la plataforma LOD.

- Nuevas oportunidades de ingresos o creación de nuevas ideas de negocio y de investigación.

Estructura de costos o Los costos de mantenimiento de la plataforma de datos abierta:

- El personal, la infraestructura, el desarrollo y el tiempo empleado en el proceso de publicación de nuevos datos.

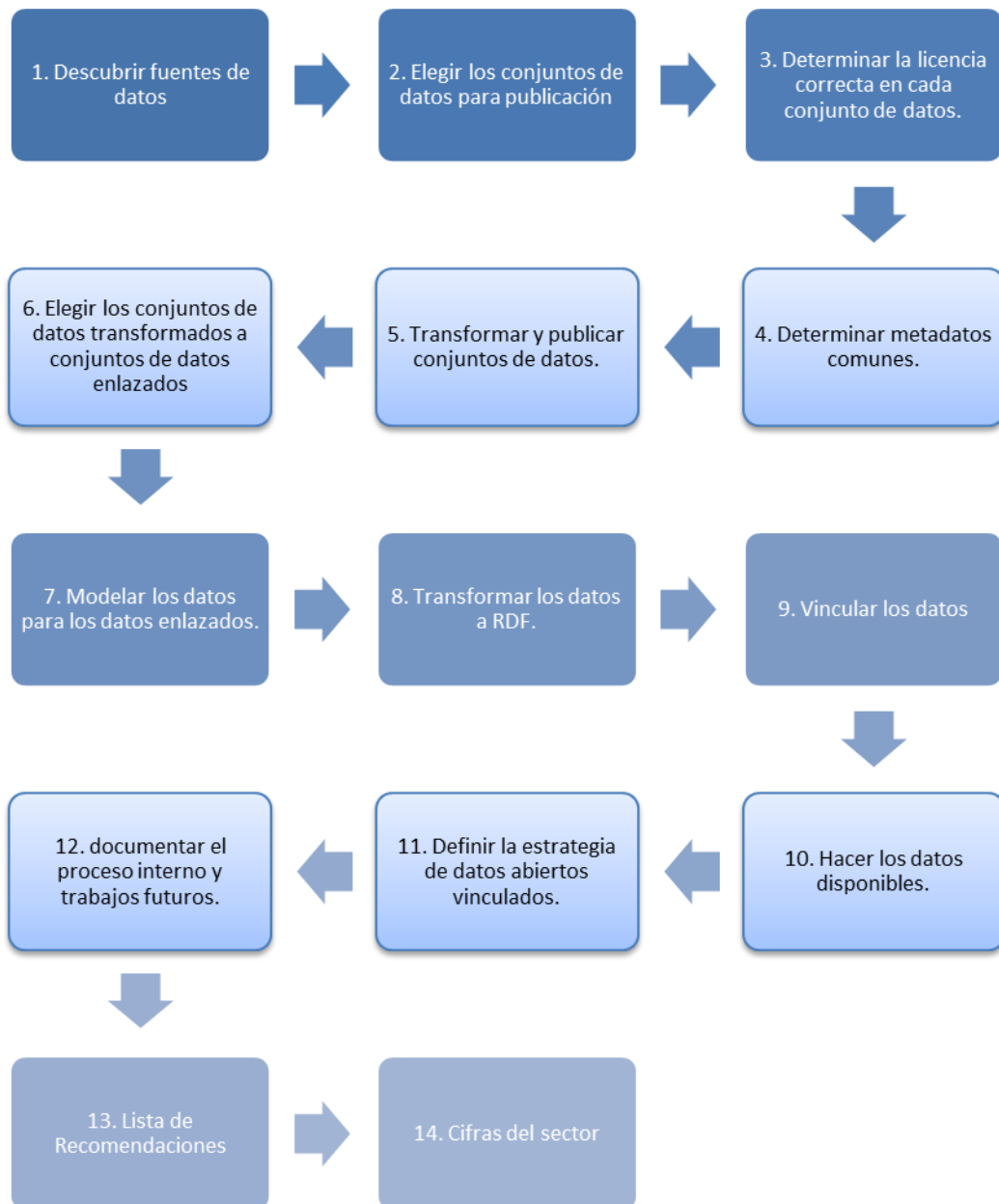
- Los costos de activos de comunicación y promoción.

3.6.7 El Proceso de Apertura de Datos. Para desarrollar el proyecto, las siguientes actividades fueron considerados, basado principalmente en los pasos recomendados de LOD, el "Open Data Handbook"¹³⁹ y las recomendaciones de "Linked Data: Evolving the Web into a Global Data Space"¹⁴⁰.

¹³⁹ Open Knowledge Foundation. "Open Data Handbook Documentation." Release 1.0.0,2012: pp. 6-7 (1.3.1).

¹⁴⁰ Heath, Tom; Bizer, Christian. "Linked Data: Evolving the Web into a Global Data Space.",.2011. pp. 72. Morgan & Claypool Publishers.

Figura 68. Pasos para Apertura de Datos



Fuente: Construcción propia.

Cuadro 6. Actividad vs. Descripción

No.	ACTIVIDAD	DESCRIPCIÓN
1.	Descubrir los diferentes fuentes de datos de Turismo en un nivel institucional	Descubre todos los datos conocidos y desconocidos de turismo a nivel institucional. Este proceso es necesario a fin de conocer qué datos puede generar un valor añadido para la entidad, considerándolo como un recurso sin explotar.
2.	Elegir los primeros conjuntos de datos que va a ser publicado.	Elija la lista de conjuntos de datos que serán publicados a principios del proyecto.
3	Determinar la licencia correcta para cada conjunto de datos.	Elija una licencia para cada conjunto de datos, a especificar cómo se deben utilizar los datos y los límites
4	Determinar los metadatos comunes que se utilizará para toda la conjuntos de datos	Elija los metadatos que se utilizará para publicar los datos. Los metadatos son parámetros que describen los datos, tales como nombre, autor, mantenedor, fuente de datos o calidad
5	Transformar y publicar el conjuntos de datos, en formatos abiertos y subirlos en un datos sistema de gestión	Transformar los datos (es decir, normalización,etc.) a los formatos primarios y publicarlas en un sistema de manejo de datos (es decir CKAN)
6	Elija los conjuntos de datos que deben ser transformados a los formatos de datos enlazados.	Elija la lista de conjuntos de datos que serán transformado a los datos enlazados
7	Modelar los datos para los datos enlazados	Los conjuntos de datos se modelan de acuerdo a vocabularios y ontologías específicos.
8	Transformar los datos al estándar RDF.	Los conjuntos de datos se transforman en formatos de datos enlazados
9	Vincular los datos a partir de los datos relacionados con otros sectores.	La vinculación de los datos a los conjuntos de datos similares de otros sectores.
10	Poner los datos disponibles y visibles a través de un SPARQL endpoint.	Hacer los conjuntos de datos pueden descubrirse a través de un SPARQL Point con RDF graph.
11	Definir la estrategia de datos abiertos enlazados para el sector turístico.	Se define la estrategia del proyecto para el sector turístico.

Cuadro 6. (Continuación)

No.	ACTIVIDAD	DESCRIPCIÓN
12	Documentar el proceso interno y flujo de trabajo para la apertura de datos futuros. Los procesos formales para la publicación de datos.	Los procesos formales para la publicación de datos están documentados.
13	Crear una lista de recomendaciones para una comunicación interna y plan de promoción para el proyecto.	Las recomendaciones para el desarrollo de un plan de comunicación y promoción están documentados. Este paso se centra en la promoción del uso del portal y la comunicación de los beneficios del proyecto.
14	Extender el sector turismo en cifras (es decir, hechos y cifras portal de datos abiertos) información que debe publicarse utilizando la plataforma de datos abierta enlazado	La información del sector turismo en cifras es transformada a los formatos de datos enlazados con el fin de crear un portal dinámico a través de la ayuda de la plataforma LOD.

3.6.8 Fuentes de Datos Disponibles. Como es un sector de alta importancia para el desarrollo del departamento se cuenta con una amplia variedad de tipos de datos que pueden traer grandes beneficios dentro del proyecto. Estos se pueden categorizar y priorizar según su alineamiento con la estrategia del plan de desarrollo. El siguiente es la lista de las categorías identificadas relacionadas con el sector turístico:

Cuadro 7. Categorías de Datos Sector Turístico

Tipo	Identificador
Zonas de acampada	ZONASLOD
Parques	PARQUESLOD
Zonas para perros	ZONAPERROLOD
Jardines infantiles	JARDININFANTILLOD
Piscinas públicas	PISCINASLOD
Museos	MUSEOSLOD
Estaciones Policía	ESTACIONPOLICIALOD
Escuelas Autos	ESCUELACONDUCCIONLOD

Cuadro 7. (Continuación)

Tipo	Identificador
Monumentos	MONUMENTOSLOD
Bibliotecas	BIBLIOTACASLOD
Jardines infantiles	JARDINESINFANTILESLOD
Estaciones de radio	EMISORASLOD
Museo	MUSEOLOD
Escuelas	ESCUELASLOD
Hoteles	HOTELESLOD
Clínicas	CLINICASLOD
Hospitales	HOSPITALESLOD
Tiendas	TIENDASLOD
Restaurantes	RESTAURANTESLOD
Sitios Naturales	SITIOSNAURALESLOD
Bibliotecas	BIBLIOTECASLOD
Parqueaderos	PARQUEADEROSLOD
Universidades	UNIVERSIDADESLOD
Cementerios	CEMENTERIOSLOD

Figura 69. Fuentes de datos del sector Turismo de Norte de Santander



Fuente: Elaboración propia.

Probablemente una de las fuentes de datos más interesantes de todo el sector turístico son los contenidos que se relaciona con el inventario de las categorías planteadas en el cuadro 7. Estos datos son realmente valiosos y debe ser parte de la plataforma de datos abierta. La Secretaria de desarrollo económico del departamento y la Corporación Mixta de Promoción de Norte de Santander cuentan con información de sitios turísticos. A continuación se muestra los sitios más relevantes y se puede encontrar en detalle en el Anexo E.

3.6.9 Inventario de Contenidos. Se describe a continuación

Cuadro 8. Sitios de Interés y Turísticos, Municipio de Arboledas N de S

Tipo de sitio	Nombre
Sitios Naturales	Cascada Chicagua Alto
	Cascada Chicagua Bajo
	Cueva El Buitron De Castro
	Piedra De Los Buitrones
	Rio Arboledas
	Sendero Turístico Puente Histórico Simón Bolívar Paso Libertador
	Cueva Del Indio
Patrimonio Cultural- Patrimonio Material	Tanque Aguas Termales
	Parque La Santísima Trinidad
	Plazoleta Nuestra Señora Del Rosario
	Hogar Juvenil Campesino
	Molino De Piedra (Muela)
	Molino De La Época Colonial Traído Por Los Españoles.
	Puente De Paso Simón Bolívar
	La Lomita De La Cruz
	Monumento De La Santísima Trinidad
	Templo De La Santísima Trinidad
	El Callejón La Primera Calle Empedrada
	Molino De Piedra - Vestigios De La Cultura Chitareros
Patrimonio Cultural - Patrimonio Inmaterial	Leyenda El Pozo De La Feliciania
	Leyenda Cerro De La Mucura
Hoteles	Hotel Y Restaurante Arboledas
	Hotel San Martin
Restaurantes	Restaurante La Arboleana
	Restaurante Y Cafeteria A Comer Sin Culpa
	Zona Refrescante Siravita

Fuente: Elaboración propia.

Cuadro 9. Sitios de Interés y Turísticos, Municipio de Cacota N de S

Tipo de sitio	Nombre
Sitios Naturales	Laguna De Cáota
	Piedra De Ojo
Patrimonio Cultural- Patrimonio Material	Óleo Sobre Lienzo De La Virgen De Los Dolores
	Escultura Del Sagrado Corazón De Jesús
	Camarín De La Virgen De Los Dolores
	Santuario De Nuestra Señora De Los Dolores
	Casa Cural
	Casa De La Cultura De Cacota
	Casona Carrillo
	Parque Principal De Cacota De Velasco
	Hacienda Hernández
	Molino Sucre
	Hacienda Villamizar
	Corral De Piedra
Patrimonio Cultural Patrimonio Inmaterial	Macho Rusio
	Arepa De Trigo
Hoteles	Restaurante Y Hospedaje Colonial
	Restaurante, Hospedaje Y Fuente De Soda Doña Angela
	Hostal Hacienda Cacota
Restaurantes	Restaurante Y Hospedaje Colonial
	Restaurante Mama Rosita
	Restaurante, Hospedaje Y Fuente De Soda Doña Angela

Fuente: Elaboración propia.

Cuadro 10. Sitios de Interés y Turísticos, Municipio de Cachira N de S

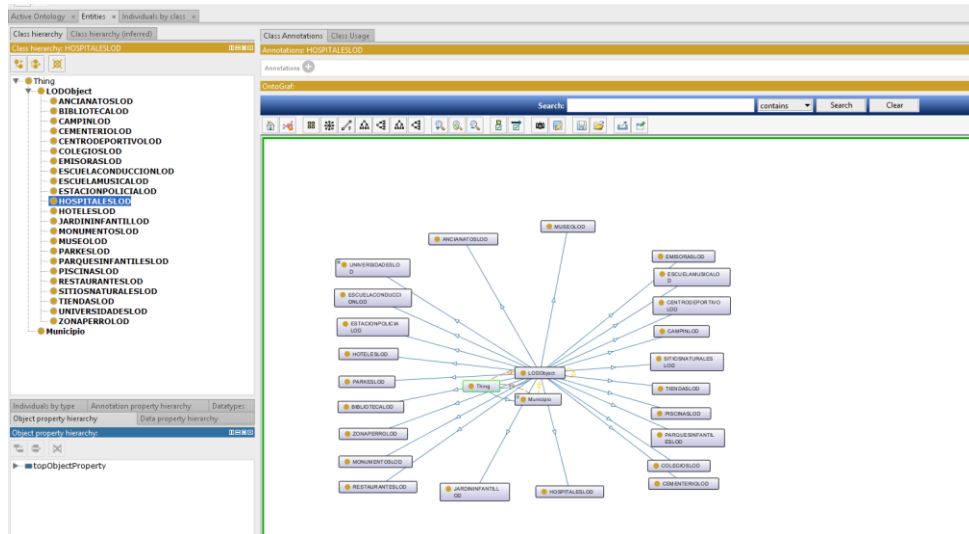
Tipo de sitio	Nombre
Sitios Naturales	La Mesa
	Laguna De Guerrero - Paramo De Guerrero
	Complejo De Cuevas Indígenas Vía Paramo De Guerrero
	Cascada La Caramba
	Cascada Quebrada Ramírez
	Complejo De Cuevas Indígenas Sitio La Mesa

Cuadro 10. (Continuación)

Tipo de sitio	Nombre
	Avistamiento De Aves (Cóndor)
	Siete Lagunas
Patrimonio Cultural- Patrimonio Material	Iglesia Nuestra Señora Del Carmen
	Parque Principal Municipio De Cachira
	Busto Del Sacerdote Víctor Manuel Montaguth
	Centro Recreacional Campo Alegre
	Monumento Del Arriero
	Iglesia De San Agustín
	Monumento De Don Emergido Méndez "Sute"
	Malecón De Cachira
	Parque Antonio José Ramírez
	Plaza De Toros Manzanares
	Templo De San Pedro Y San Pablo
	Parque La Carrera
Hoteles	Hotel Hildamas
	Apartamentos Ihb
	Residencia La Carrera
	Hospedaje La Esquina Del Buen Amigo
Restaurantes	Restaurante Doña Martha
	Restaurante Sason Cachirence
	Estadero El Rancho

3.6.10 Introduciendo Un Modelo Común. Ahora que se ha seleccionado la fuente de datos que se utilizará, se puede crear un modelo común para los datos RDF resultantes. Este tipo de modelo viene en forma de una ontología OWL. Se utilizó software protegé para diseñar la ontología. Hay una clases principal. La clase "LODObject" está diseñado como una clase base para los objetos de puntos de interés. La mayoría de los tipos de sitios individuales, tendrán a "LODObject" como una superclase. Sin embargo, habrá una clase distinta para cada tipo de punto de interés (POI) que se incluye. Estas clases distintas se definen como subclases de LODObject. Hay una subclase para cada tipo de POI. Los nombres de estas subclases son idénticos a los identificadores que se utilizan dentro del cuadro 7 "Categorías de Datos Sector Turístico" fuentes de datos disponibles (por ejemplo, "MUSEOLOD" para los museos). Puesto que cada POI está dentro de un determinado municipio de Norte de Santander y estos datos están disponible para muchos objetos hay otra clase que representa un municipio. Esta clase se llama "Municipio". No tiene subclases o superclases (a excepción de owl:Thing por supuesto). Las clases de la ontología y la jerarquía se presentan en la figura 70.

Figura 70. Ontología LOD Turismo Norte de Santander



Fuente: Elaboración Propia.

URIs. Una solución perfecta para definir un espacio de nombres URI para el caso de uso no es posible. Tal solución sería tener URI que apunten al servidor de la Gobernación de Norte de Santander o Alguna de las dependencias relacionadas (nortedesantander.gov.co). Sin embargo, ya que no tenemos control sobre este servidor no tendría demasiado sentido para definir este tipo de URI. El W3C LOD afirma que es importante seleccionar un DNS de la organización.

El espacio de nombres que se propone es:
<http://datos.nortedesantander.gov.co/LODTurismo>.

Para el caso de este proyecto se usó un espacio de nombre local. , con el fin de poder utilizarlo como un espacio de nombres para pruebas.

La definición de propiedades es un proceso complejo, pues se debe considerar si existe un vocabulario que se pueda utilizar, cada vez que sea posible especialmente para las propiedades.

Etiquetado de Individuos. Municipio y OGDOObject así como sus subclases utilizan el rdfs: property o label para etiquetar los individuos.

Una alternativa sería foaf: name (vocabulario FOAF), pero la propiedad rdfs: label es más utilizado para las personas en materia de etiquetado, por ejemplo, por DBpedia.

Cuadro 11. Propiedad y label

Property	Label
Type	Data Property
Namespace	http://www.w3.org/2000/01/rdf-schema#
Descripción	Etiqueta legible por humanos para describir un individuo
Domain	LODObject, Municipio
Data Type	String

Definición de propiedades para LODObject. Todos los puntos de interés dentro de LOD turismo Norte de Santander tienen una identificación adjunta. Este ID es de la forma TYPE.X, donde tipo indica el tipo de LOD turismo (es decir RESTAURANTELOD para restaurantes) y X es un número consecutivo adjunta al objeto individual. Esta decisión obedece a que se puede utilizar este ID en LOD, y puede ser compatible, cuando los datos de otras fuentes de otros sectores se quieran vincular con los resultados de esta tesis. El vocabulario Dublín Core contiene el "identifier" que se puede utilizar para este propósito.

Cuadro 12. Propiedad e identifier

Property	Identifier
Type	Data Property
Namespace	http://purl.org/dc/elements/1.1/
Descripción	Un identificador único para cada objeto Punto de Interés Turístico
Domain	LODObject
Data Type	String

Cuando se refieren a puntos de interés, la ubicación geográfica es muy importante. En el inventario realizado, Todos los puntos de interés del inventario turístico de Norte de Santander tienen sus coordenadas de la ubicación. La ontología WGS84¹⁴¹ que define los términos para la latitud (lat.) y longitud (long)

¹⁴¹ <http://www.w3.org/2003/01/geo/>

se adapta muy bien para este caso de estudio. Este vocabulario es utilizado ampliamente en la DBpedia.

Cuadro 13. propiedad y lat

Property	Lat
Type	Data Property
Namespace	http://www.w3.org/2003/01/geo/wgs84_pos#
Descripción	Latitud Geográfica de un Objeto
Domain	LODObject
Data Type	String

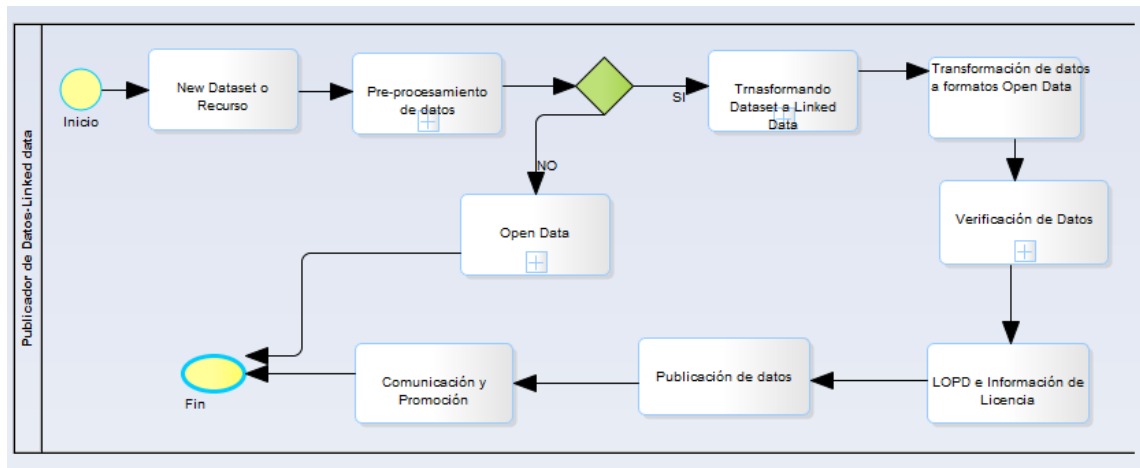
La información completa de las propiedades y los detalles de los vocabularios utilizados en el Anexo C.

3.6.11 El proceso de publicación de los datos. La publicación de datos debe seguir un proceso específico con el fin de garantizar la correcta difusión de los datos abiertos enlazados a través de la plataforma y deberá ser gestionado por personas con funciones específicas dentro del proceso. En las Figuras 71 y 72, dos procesos se muestran, uno para la creación de un nuevo conjunto de datos y la adición a un conjunto de datos existente. Estos procesos garantizan algunos aspectos importantes tales como:

- El pre-procesamiento de datos que incluirá técnicas para adaptar los datos de acuerdo con estructuras específicas.
- La publicación de datos enlazados para algunos conjuntos de datos.
- Los aspectos LOPD y licencias que seguirán técnicas especiales en el caso de que algún conjunto de datos tiene datos sensibles.
- incluirá un proceso de verificación de datos, relacionados con metadatos, calidad, recursos y procedencia.

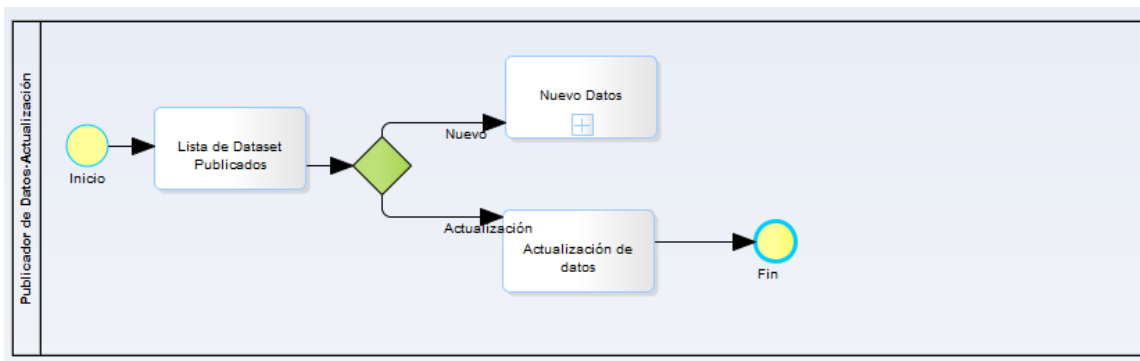
- Comunicación y promoción de los nuevos datos publicados para los usuarios de datos.
- La actualización de bases de datos ya publicados.

Figura 71. Proceso de Creación de Nuevo Conjunto de Datos



Fuente: Elaboración Propia.

Figura 72. Proceso de Actualización de un conjunto de datos



Fuente: Elaboración Propia.

Implementar Políticas de datos, con el fin de fomentar el desarrollo del proyecto y garantizar un flujo continuo dentro de los elementos del ecosistema, es fundamental que la secretaria de desarrollo económico implemente una política de

datos abiertos institucional, que busque estandarizar el proceso de publicación de datos dentro de la organización. Esta política debe ser regulada por los principales directivos de la organización.

Las siguientes son las recomendaciones:

- Automatizar el proceso para la obtención de datos actualizados de los almacenes de datos a través de diferentes formatos estructurados o no-estructurados, con el fin de apoyar la interoperabilidad y accesibilidad junto a la plataforma de datos abierta.
- Fortalecer la gestión de datos y liberar prácticas mediante la aclaración de las funciones y responsabilidades, comunicando el valor estratégico de los datos abiertos al interior de las partes interesadas y el público (ver tabla stakeholders).
- Los datos personales y sensibles deben ser asegurados adecuadamente para asegurar que la privacidad y confidencialidad están totalmente protegida.
- Crear una cultura de colaboración entre la estructura organizativa de la gobernación de Norte de Santander con el fin de ayudar a incorporar una cultura de datos abierta.
- Generar estrategias para comunicar y dar a conocer la política de datos abiertos a emprendedores de universidades y sociedad en general.
- Integrar el ecosistema de datos abiertos y la cultura dentro de la estrategia de la Gobernación.

Actividades. Orientar a los integrantes del sector turístico hacia los datos abiertos por defecto, Que significa que todos los datos públicos referente al sector deben publicarse abiertamente por defecto, al tiempo que reconoce que hay razones legítimas por las que algunos datos no pueden ser puestos libremente. De acuerdo a esto, los datos que se genera a partir de las fuentes públicas, deben:

- Los datos abiertos deben ser creados en formatos legibles para maquinas.
- Usar los estándares de datos, licencias abiertas.
- Estar documentado, incluyendo su esquema de datos y diccionarios.
- liberar de datos de alta calidad de forma oportuna, completa y precisa. Este incluye datos de liberación lo antes posible, permitir que los usuarios proporcionen información, y luego seguir haciendo revisiones para garantizar los más altos estándares de calidad conocidos de los datos abiertos.

3.6.12 Factores Clave de Éxito para la Iniciativa LOD Turismo de Norte de Santander. Con el fin de garantizar un buen rendimiento, son necesarios algunos factores clave de éxito para el seguimiento el proyecto durante el tiempo, por lo que podemos conocer algunos aspectos importantes, como el cumplimiento de los procesos establecidos para la publicación de nuevos datos y comprobar los antiguos conjuntos de datos publicados. Además, es importante hacer un seguimiento de si la cantidad de datos han sido enlazados publicados y el número de aplicaciones, ideas o proyectos que se han creado utilizando los datos han aumentado. La siguiente es la lista de los factores clave de éxito que fueron identificados como relevantes para este proyecto:

- Número de conjuntos de datos publicados en formatos abiertos.
- Número de conjuntos de datos publicados en formato de datos vinculados.
- Cumplimiento de las Estrellas es decir principios de cumplimiento Linked Data.
- Número de actividades de datos abiertos para promover el uso de los datos (Hacketones, concursos).
- El uso de formatos abiertos.
- Realización de la política de datos abiertos.

- Realización del plan de comunicación.
- Número de conjuntos de datos no actualizados.
- Tipo de datos publicados.
- Número de conjuntos de datos vinculados con otros sectores
- Número de aplicaciones Desarrolladas.
- Número de ideas de negocio desarrolladas y participando por ejemplo en Programas de emprendimiento como Apps.co.
- Número de proyectos desarrollados.
- Número de conjuntos de datos que son privado.
- Número de conjuntos de datos que son público.

3.6.13 Participación y Comunicación. La participación de los ciudadanos en el ecosistema de datos abierta es importante. El nivel de interacción representa parte del retorno de la inversión en los costos marginales del proyecto LOD para el turismo de Norte de Santander. Esta es la razón por la cual es necesario crear un plan de comunicación y promoción que estimulará la participación de los ciudadanos junto con el enriquecimiento de la cultura datos abiertos dentro de las instituciones participantes.

Recomendaciones para la creación del Plan:

- Crear o participar en eventos para promover la participación de los ciudadanos (locales, nacionales e internacionales).

- Comunicar la iniciativa y la estrategia de LOD del Turismo de Norte de Santander, entre todos los las partes interesadas.
- Promover el uso del portal LOD, con el fin de mejorar la visibilidad de la región y atraer turistas externos.
- Realizar una lista de los principales grupos de interés de la plataforma LOD que deben ser comunicadas y crear una matriz que explica el tipo de comunicación, los objetivos, y las posibilidades de cada uno de ellos en temas relacionados principalmente con el desarrollo de ideas de negocios, aplicaciones y proyectos.
- Usar técnicas de marketing para promover el uso de la plataforma LOD para turismo, a través de redes sociales.

3.6.14 Aspectos Legales, LOPD Y Licencias de Datos Abiertos. Los aspectos legales al publicar información en la Web son importantes y representan las regulaciones para los usuarios de datos y los editores de datos. Especialmente en el sector público, La liberación de los datos es un proceso que debe ser altamente apoyada por políticas públicas y las regulaciones que podrían garantizar los derechos de todas las entidades públicas a abrir sus datos mientras que el apoyo a los derechos de los ciudadanos para exigir la apertura de ciertos conjuntos de datos. La apertura de datos en el sector público es un movimiento delicado ya que hay muchos tipos de datos que podrían tener efectos adversos si no se libera los datos y se comunican de la forma correcta. Especialmente, estas políticas y reglamentos deben garantizar la protección de los datos personales de los ciudadanos y que especifique los mecanismos adecuados para asegurar que los datos no se utilizan de forma inapropiada.

Hay licencias de datos abiertos comunes se enumeran en la Tabla 14, y se describen según sus características de atribución y uso compartido. El campo "Requerir Atribución" describe si el usuario de datos debe hacer referencia al autor original de los datos y la Campo "LOC" describe si los trabajos derivados deben utilizar las condiciones de la licencia original. Teniendo en cuenta esto, el proyecto LOD para el Turismo de Norte de Santander debe comprobar cada conjunto de datos previamente antes de ser publicados y asignar la licencia más adecuada en función del contenido de cada conjunto de datos. Las licencias contempladas en el proyecto de datos abiertos vinculados son los siguientes:

- Open Database License¹⁴² (ODC-ODbL).
- Public Domain Dedication and License¹⁴³ (PDDL).
- Creative Commons Attribution¹⁴⁴ (CCA 3.0 Unported).

Cuadro 14. Open Data Licenses que se pueden usar en el proyecto LOD del Turismo de Norte de Santander

Open Data Licenses			
License	Requiere Atribución	LOC	Descripción
Open Database License (ODC-ODbL).	SI	SI	Para datos.
Public Domain Dedication and License (PDDL)	NO	NO	Dedicado a dominio público (todos los derechos exentos)
Attribution License (ODC-By)	SI	NO	Para datos.
Creative Commons CCZero (CC0 1.0 Universal (CC0 1.0))	NO	NO	Dedicado a dominio público (todos los derechos exentos).
Creative Commons Attribution 3.0 ¹⁴⁵ (CC-BY)	YES	NO	Creada originalmente para contenido (por ejemplo: texto, Imágenes, etc.), pero puede extenderse para los datos.
UK Open Government Licence 2.0 ¹⁴⁶ (OGL-UK-2.0)	YES	NO	Creado para ser utilizada principalmente en Reino Unido, pero se puede extender a otros países.

Fuente: Elaboración Propia.

¹⁴² <http://opendatacommons.org/licenses/odbl/>

¹⁴³ <http://opendatacommons.org/licenses/pddl/1-0/>

¹⁴⁴ <https://creativecommons.org/licenses/by/3.0/>

¹⁴⁵ <https://creativecommons.org/about/cc0>

¹⁴⁶ <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2/>

3.6.15 Metadatos Utilizado Para Turismo LOD Norte De Santander. Los metadatos son descripción de los datos, y es especialmente necesario en los datos abiertos. Estos le ayudan a los usuarios a conocer los detalles pertinentes, como la fuente de datos o detalles de la calidad. En LOD del turismo de Norte de Santander, algunos metadatos comunes fueron considerados para todos los conjuntos de datos. En el Cuadro 15, se muestra los metadatos utilizados para el proyecto:

Cuadro 15. Metadatos Utilizados para el Proyecto

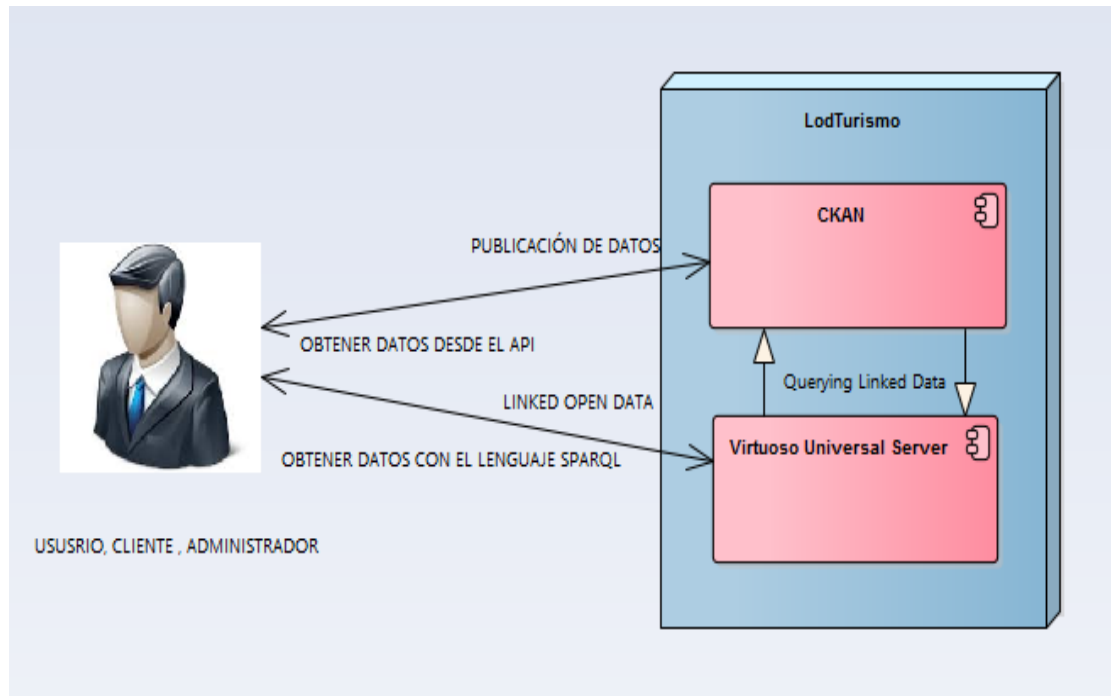
Metadatos para el Proyecto LOD Turismo Norte de Santander		
#	Nombre de Metadatos	Descripción
1	Titulo	Nombre del conjunto de datos
2	Descripción	Descripción del conjunto de datos
3	Autor	Autor
4	Mantenedor	Mantenedor Principal
6	Categoría	Categoría
7	Fuente de datos	Fuente/procedencia
8	Formato	Formato de conjunto de datos
9	Idioma	Idioma de los datos
10	Calidad	Calidad de los datos
11	Tablas complementarias relacionadas	Tabla complementarias relacionadas
12	Conjunto de datos relacionados	Los conjuntos de datos relacionados
13	Temporalidad	Frecuencia de actualización
14	Tipo de datos	Tipo de datos
15	Etiquetas/palabras claves	Palabras claves relacionadas
17	Organizaciones	Organizaciones afiliadas
18	Licencia	Descripción de la licencia
19	Recursos	Lista de recursos
20	Última actualización	Ultima fecha de actualización
21	Creado	Fecha de creación
22	Formato	Formato de los archivos
24	Mimetype	Mimetype formato
25	Estado	Estado del conjunto de datos(activo o no)

Fuente: Elaboración propia.

3.6.16 Arquitectura, Herramientas Y Requisitos Del Proyecto. Es importante definir una arquitectura que permita otras entidades o sectores que aún no han comenzado con un proyecto similar la utilicen como referencia. Las herramientas

que se utilizaron dentro de la arquitectura también se compartieron con el mismo propósito. La arquitectura que fue considerado para el proyecto de LOD para el turismo de Norte de Santander es la siguiente:

Figura 73. Arquitectura del Proyecto LOD turismo de Norte de Santander



Fuente: construcción propia.

Las herramientas que se utilizan para implementar la arquitectura anterior se muestran en el cuadro 16.

Cuadro 16. Herramientas y tecnológicas utilizadas para la apertura de los Datos del sector turístico de Norte de Santander

No	Tecnología	Razón	Propósito	Licencia
1	CKAN	Es una completa herramienta para datos abiertos que es ampliamente utilizado en todo el mundo	Gestión de datos sistema.	GNU Affero General Public License
2	OpenLink Virtuoso Universal Server.	Base de datos para los datos vinculados.	Base de datos RDF, SPARQL Point	OpenSource y Commercial Edition
3	Open Refine	Proporciona una funcionalidad para modelar datos vinculados con un plugin externo	Procesamiento de datos desordenados, transformar los datos a RDF	Open Source
4	Protégé	Es una completa herramienta para modelado de la ontología de Datos vinculados.	Ontología Editor, Linked Data	Open Source
6	Ubuntu	Es un operativo robusto basado en Linux del sistema, y se recomienda para CKAN.	Repositorio de Datos para hacerlo detectable.	Open Source

Fuente: Elaboración propia.

Como complemento a esta información, los siguientes son los requisitos de hardware para implementar la arquitectura anterior.

Cuadro 17. Requisitos de hardware dentro del proyecto LOD Turismo Norte de Santander

Hardware/Plataforma		
No.	Requerimiento	Descripción
1	Virtual Machine CKAN	Herramienta de base de datos predeterminada CKAN
2	Virtual Machine Virtuoso	Virtuoso (SPARQL Point)

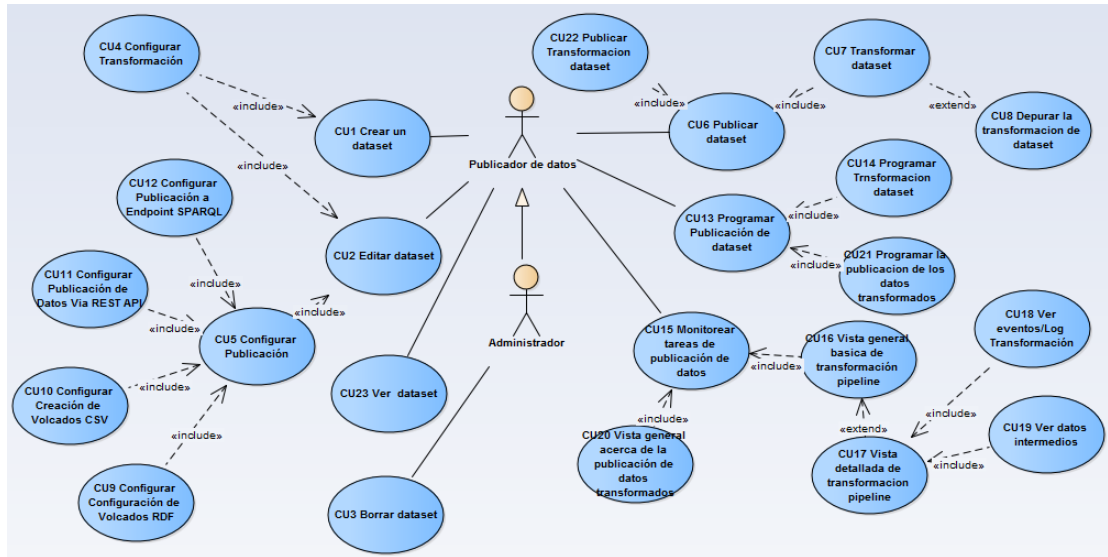
3.6.17 Actores Principales del Sistema. Actores principales que interactúan con el sistema.

- Proveedores de datos del gobierno, municipios, y otros sujetos que proporcionan datos.
- Los consumidores de datos del gobierno, municipios, organizaciones sin fines de lucro (ONG), los ciudadanos (público en general), empresas (PYME), los desarrolladores de aplicaciones que consumen los datos transformados.
- Los administradores de datos, conservadores de datos, etc. Responsables de la configuración de la plataforma LOD - generalmente empleados del proveedor de datos.
- Los administradores de TI, el personal responsable de la instalación, el mantenimiento y (parcialmente) la configuración de LOD- por lo general los empleados del proveedor de datos.

Entradas de datos. La entrada al sistema son datos de los administradores de almacenes de datos en entornos heterogéneos, con gran variedad de formatos y el empleo de una gran cantidad de diferentes tecnologías para acceder y procesar esos datos.

3.6.18 Casos de Uso. En este apartado, presentamos los casos de uso para los editores de datos y los consumidores de datos. Estos casos de uso pueden estar sujetos a cambio según las necesidades de los usuarios.

Figura 74. Caso de uso para Editores de Datos



Fuente: Elaboración propia.

Cuadro 18. Descripción de casos de uso para el editor de datos

Identificación	Nombre	Breve Descripción
CU1	Crear Dataset	El editor de datos puede crear nuevo dataset sobre los datos publicados previstos, de modo que pueda definir para cada conjunto de datos la información sobre los datos de origen, transformaciones previstas y maneras de cómo se deben publicar los datos transformados.
CU2	Editar / Administrar registro de datos	El editor de datos puede editar / administrar los registros del conjunto de datos.
CU3	Borrar registro de datos	El editor de datos puede borrar registro de datos obsoletos.

Cuadro 18. (Continuación)

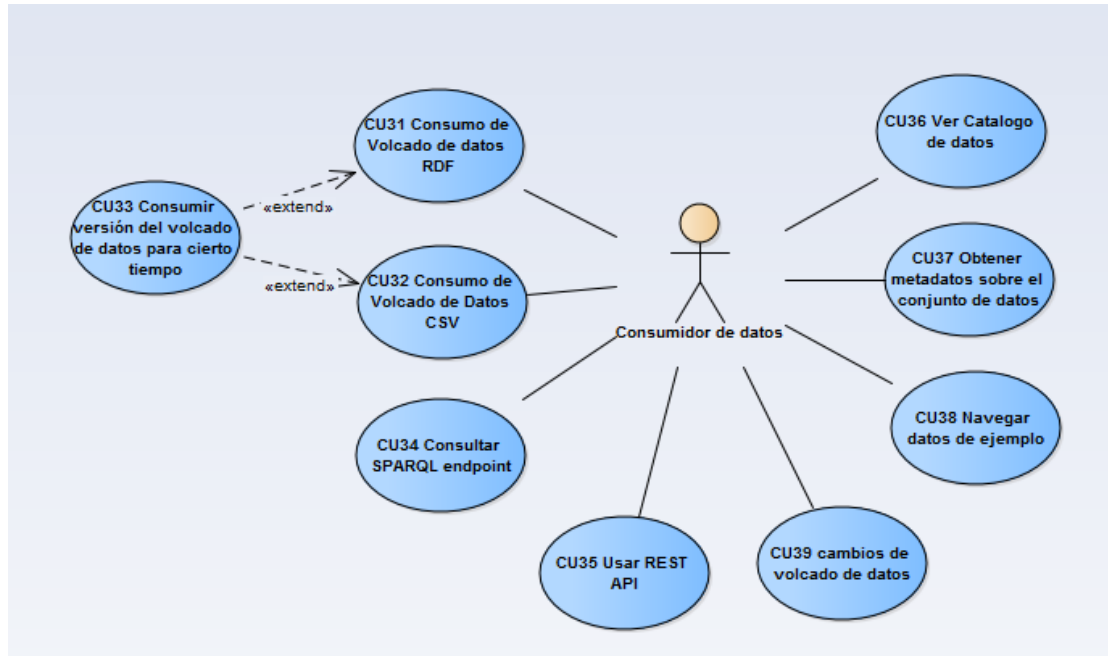
Identificación	Nombre	Breve Descripción
CU4	Configurar transformación	El editor de datos puede configurar la transformación de datos (limpieza, enlazar, enriquecimiento, la evaluación de la calidad, etc.).
CU5	Configurar publicación	El editor de datos puede configurar cómo se publica el conjunto de datos transformado, y, cómo el conjunto de datos puede ser consumido por los consumidores de datos.
CU6	Publicar conjunto de datos	El editor de datos puede publicar el conjunto de datos
CU7	Transformar conjunto de datos	El editor de datos puede transformar el conjunto de datos (el conjunto de datos se transforma, pero aún no se publica)
CU8	Depurar la transformación de datos.	El editor de datos puede depurar la transformación de datos, ver los resultados intermedios de la transformación, ver mensajes de depuración que ilustra lo que sucedió durante la transformación de datos.
CU9	Configure creación de volcado de datos RDF	El editor de datos puede configurar creación de volcado de datos en RDF del conjuntos de datos publicado.
CU10	Configurar la creación de volcado CSV.	El editor de datos puede configurar creación de volcado CSV de los conjuntos de datos publicados.
CU11	Configurar la publicación de datos a través de la API REST	El editor de datos puede configurar cómo se genera la API REST en la parte superior de los datos publicados, que los datos sean accesibles a través de la API REST, que los usuarios pueden utilizar la API REST, que los métodos para acceder a los datos estén disponibles para los consumidores de datos.

Cuadro 18. (Continuación)

Identificación	Nombre	Breve Descripción
CU12	Configurar la publicación del punto SPARQL	El editor de datos puede configurar cómo los consumidores de datos pueden conectarse a los puntos finales SPARQL de los datos publicados
CU13	Programar publicación de datos	El editor de datos puede automatizar el proceso de publicación, por lo que puede ejecutar todas las semanas o cada vez que cuando la nueva versión del conjunto de datos está disponible
CU14	Programación de Transformación de datos	El editor de datos puede automatizar la parte de transformación del proceso de publicación, por lo que puede ejecutar todas las semanas o cada vez que quiera la nueva versión del conjunto de datos.
CU15	Supervisar las tareas de publicación de datos	El editor de datos puede controlar la tarea de publicación para ver cómo se ejecuta la transformación de datos de publicación para el dataset.
CU16	Vista general básica de transformación pipeline.	El editor de datos puede controlar la publicación de la base de datos para ver si la publicación estaba bien, o hubo algunos errores.
CU17	Vista detallada de transformación pipeline.	El editor de datos puede ver la información general detallada sobre las transformaciones del conjunto de datos
CU18	Explorar registros y eventos de transformación.	El editor de datos puede ver los logs y eventos para ver en detalle lo que sucedió durante la transformación de datos
CU19	Ver datos intermedios.	El editor de datos puede ver algunos de los datos intermedios producidos para mirar como el conjunto de datos está siendo transformado.
CU20	Vista general acerca de la publicación de los datos transformados	El editor de datos puede estar informado sobre la publicación del conjunto de datos transformado, si hay algunos problemas o no
CU21	Programar la publicación de los datos RDF transformados.	El editor de datos puede automatizar la publicación de los datos transformados.
CU22	Publicar conjunto de datos transformado	El editor de datos puede publicar el conjunto de datos, que ya se ha transformado.

Fuente: Elaboración propia.

Figura 75. Caso de Uso para Consumidores de Datos



Fuente: Elaboración propia.

Cuadro 19. Descripción de casos de uso para el consumidor de datos

Identificación	Nombre	Breve Descripción
CU31	Consumir volcado de datos RDF	Como consumidor de datos puede descargar volcado de datos RDF, por lo que puedo cargar a mi almacén de datos y trabajar con ella
CU32	Consumir volcado de datos CSV	Como consumidor de datos puede descargar volcado de datos CSV, por lo que puedo cargar a mi almacén de datos y trabajar con ella
CU33	Consumir versión del volcado de datos para cierto tiempo.	Como consumidor de datos puede conseguir la versión de datos válida en un tiempo determinado.

Cuadro 19. (Continuación)

Identificación	Nombre	Breve Descripción
CU34	Consulta SPARQL endpoint.	Como consumidor de datos puede consultar datos RDF directamente utilizando SPARQL endpoint.
CU35	Uso de la API REST	Como consumidor de datos puede usar la API REST, para trabajar con los datos de la aplicación.
CU36	Ver catálogo de datos	Como consumidor de datos puede navegar y buscar en la lista de conjuntos de datos (catálogo de datos)
CU37	Obtener metadatos sobre el conjunto de datos	Como consumidor de datos puede obtener metadatos del conjunto de datos publicados.
CU38	Ver datos de ejemplo.	Como consumidor de datos puede ver datos de ejemplo para conseguir tener una idea de lo que está en el conjunto de datos.
CU39	Cambios de volcado de datos	Como consumidor de datos puede ser notificado (por ejemplo a través de RSS o Atom) cuando un volcado de datos se actualiza o cambia

Fuente: Elaboración propia.

3.6.19 Desarrollo e Implementación. La parte práctica del desarrollo e implementación del proyecto intentara poner en práctica los conceptos y criterios establecidos a lo largo de la investigación y con el uso de tecnologías de software libre, para al final construir una aplicación que permita consumir los datos RDF publicados a través de un servidor semántico Virtuoso Universal Server y con consultas SPARQL.

El resultado de las consultas pertenece a componentes específicos de los sitios turísticos de Norte de Santander entre los que se encuentra (hoteles, sitios de interés, restaurantes, cultura, patrimonio, etc.). Durante este proceso se desarrollan una serie de fases que permiten la terminación del proyecto, es decir la construcción de RDF y el proceso de consulta.

3.6.20 Instalación y Configuración de Plataformas. Se describe a continuación

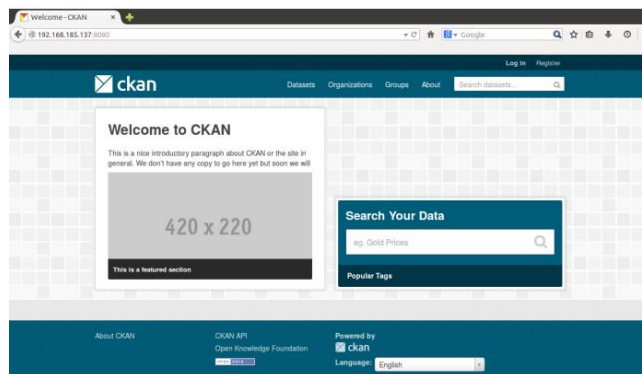
Plataforma de Open Data. La implementación de la plataforma se llevó a cabo a través de CKAN es un software Open Source que permite construir plataformas web de datos abiertos, que proporciona funcionalidades completas en catálogos de datos abiertos, visualización de datos y análisis, workflow, almacenamiento de los datos y personalización de la interfaz.

Los requerimientos de Software para la implementación de CKAN son:

- Máquina virtual en Virtual Box¹⁴⁷.
- SERVIDOR: Ubuntu server 12.04 de 64 bits¹⁴⁸
- Apache¹⁴⁹.
- Jetty¹⁵⁰.
- PostgreSQL¹⁵¹.
- Openjdk¹⁵².

De una vez instalado y configurado el servidor Ubuntu se procede a la instalación de CKAN, terminada la instalación se puede acceder a la página principal, para verificar que el proceso se llevó a cabo exitosamente

Figura 76. Plataforma CKAN



Fuente: CKAN

¹⁴⁷ <https://www.virtualbox.org/>

¹⁴⁸ <http://www.ubuntu.com/download/server>

¹⁴⁹ <http://httpd.apache.org/>

¹⁵⁰ <http://www.eclipse.org/jetty/>

¹⁵¹ <http://www.postgresql.org.es/>

¹⁵² <http://openjdk.java.net/>

Después se puede ingresar a la interfaz de administración y configuración de la plataforma y dataset.

Figura 77. Interfaz de dataset CKAN



Fuente: CKAN

Plataforma de LOD. La configuración de Triples Store y SPARQL endpoint, se llevó a cabo con OpenLink Virtuoso Universal Server.

El repositorio semántico VIRTUOSO, creado por OpenLink Software, tiene como principal característica la integración de SQL, RDF, XML y además servicios web. La integración de estos servicios permite la múltiple utilización del repositorio que además genera una arquitectura compuesta con la posibilidad de almacén de diversos datos.

A continuación se hace una descripción de las ventajas del repositorio VIRTUOSO.

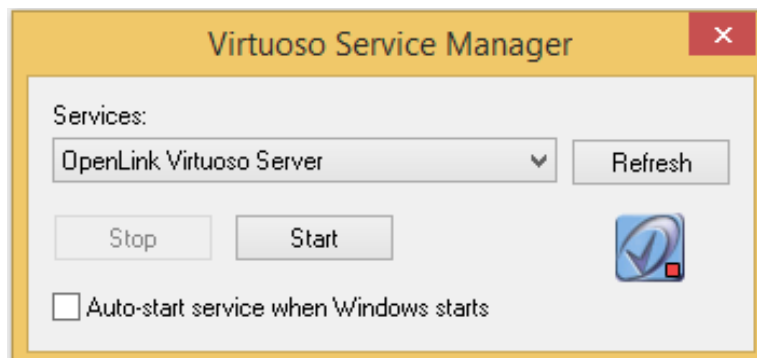
Cuadro 20. Ventajas del repositorio Virtuoso

Ventajas	Comentario
Agilidad y Rapidez	Virtuoso permite mezclar los mejores servicios de gestores de base de datos, con una infraestructura ligera y fácil de acceder.
Reducción de Costo	Debido a su licencia GPL reduce costos asociados a almacenar datos propietarios de aplicaciones.
Mayor Agilidad	Capacidad de mezclar y combinar los diferentes datos mediante estándares actuales como HTTP, SPARQL, y SQL.
Acceso por servicios web	Esta característica es importante para el manejo del repositorio remotamente
Versatilidad	Virtuoso tiene la capacidad para encajar de forma natural y segura hacia datos a través de fuentes internas y externas

Fuente: Virtuoso OpenLink.

Ejecución local del servidor VIRTUOSO. Después de descargar el repositorio Virtuoso de su página oficial¹⁵³, procedemos a instalar el software. Para poder utilizar virtuoso se debe instalar con la licencia con la que se descarga el instalador, una vez terminado de instalar, se ejecuta el servicio “Openlink Virtuoso Server”.

Figura 78. Interfaz de ejecución del Repositorio Virtuoso



Fuente: Virtuoso OpenLink

¹⁵³ <http://virtuoso.openlinksw.com/download/>

Si el servidor se ejecuta correctamente, para poder utilizar el entorno web del software virtuoso, se necesita abrir la siguiente dirección url en el navegador web: `http: localhost:8890`, esta dirección está formada por la dirección IP del servidor y el puerto 8890 que es el puerto por defecto por donde se escucha el servidor. La figura 77 muestra la interfaz de usuario del repositorio Virtuoso.

Figura 79. Entorno Web del Repositorio Virtuoso



Fuente: Virtuoso OpenLink.

Para poder utilizar los servicios del repositorio, se debe acceder a la pestaña Conductor en el menú lateral de la interfaz y autenticarse con el nombre de usuario: **dba** y contraseña: **dba** (Usuario y contraseña por defecto).

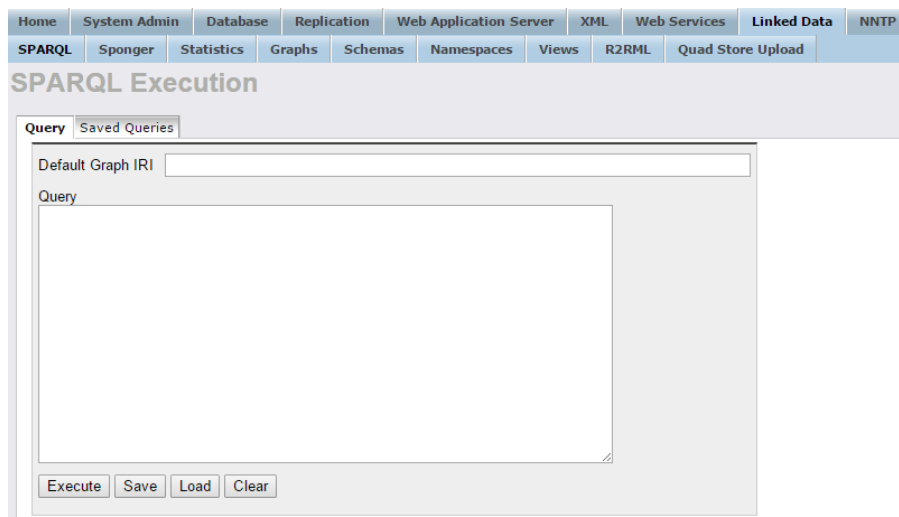
Figura 80. Administración del Repositorio Virtuoso



Fuente: OpenLink Virtuoso

El entorno web servidor Virtuoso. Después de autenticarse, se habilita la interfaz web del repositorio para el usuario por defecto dba. Para el alcance del proyecto, se utilizará el motor de gestión de tripletas (Triple Store), para eso, debemos acceder a la pestaña Linked_ Data de la interfaz web, cuando se abre el servicio, automáticamente se muestran todas las opciones para el manejo de tripletas, como se muestra la figura 80.

Figura 81. Interfaz Web “Linked Data” Servidor Virtuoso

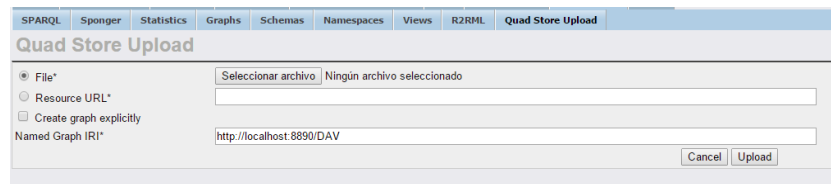


Fuente: OpenLink Virtuoso

Carga de modelo ontológico al repositorio Virtuoso. Después de ejecutar el entorno web del repositorio, el cual nos permite gestionar tripletes de información, se procede a cargar el modelo ontológico de los puntos turísticos de interés del departamento Norte de Santander.

Para esto procedemos a dar clic en la pestaña Quad Store Upload del entorno web Virtuoso en “LinkedData”. Nos aparecerá una ventana que nos permite cargar la ontología, seleccionando el archivo de texto plano (los archivos deben ser solo de tipo RDF, Turtle) especificando el nombre o IRI que identifica el grafo donde se cargara el modelo. Cuando se hayan especificado estos parámetros se procede a grabar el modelo, dando clic en Upload.

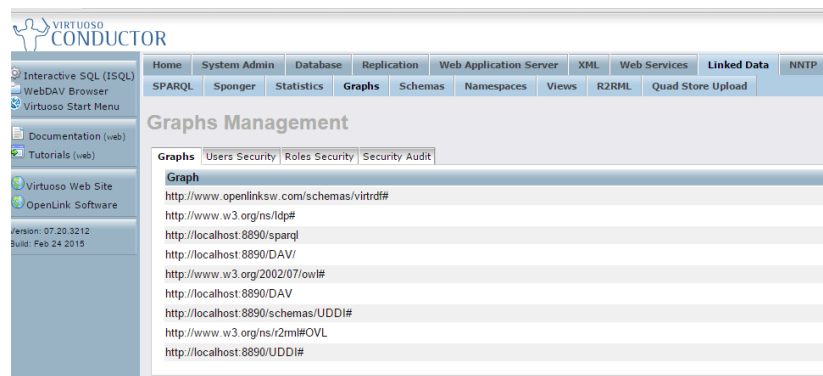
Figura 82. Interfaz para Cargar la ontología en el repositorio Virtuoso



Fuente: OpenLink Virtuoso.

Podemos confirmar que se ha cargado bien el modelo, accediendo a los grafos en la pestaña Graphs como muestra la figura 82.

Figura 83. Ontologías Cargadas en el repositorio



Fuente: OpenLink Virtuoso.

3.6.21 Desarrollo del prototipo. Se creó una aplicación Móvil y Web que visualiza los datos de puntos de interés. De la plataforma LOD de Turismo de Norte de Santander que se compone principalmente de datos de puntos de interés turístico para mostrarlos en un teléfono móvil o tablet.

Esta aplicación se desarrolló en Framework Apache Córdova lo que permite que se pueda importar a otras plataformas como IOS o Windows.

Web Site y App Móvil. La necesidad de información es, sin duda, una de las situaciones con la que se encuentra el turista en el momento de iniciar la planificación de un viaje. Esta necesidad inicial de información deberá ser complementada, a través de un sitio web o aplicación móvil, a la hora de realizar un viaje. Para ello, es de gran importancia definir la información turística que ofrece el sitio a los turistas con el objetivo de informarlo y orientarlo durante su estancia, o incluso, todas aquellas informaciones que le ayudaran a prepararlo de forma más precisa.

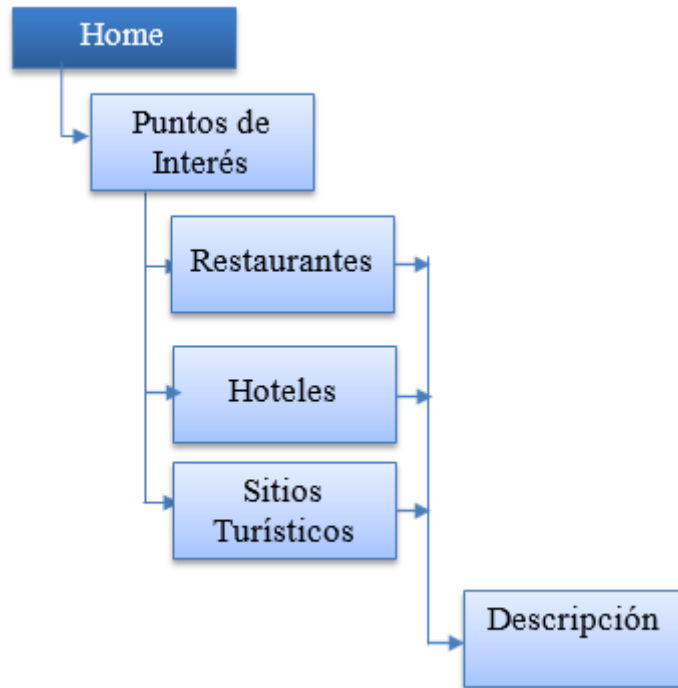
El turismo es uno de los sectores más representativos debido al alto número de empleos que genera y los diferentes negocios que en el intervienen, debido a esto se genera una serie de información, de importancia tanto a nivel local como internacional, y para lograr la globalización se plantea como imprescindible el uso de diferentes tecnologías. Como herramienta de promoción, Internet ofrece al sector turístico la facilidad de modificación y actualización de la información.

Requerimientos. La actividad principal de la aplicación es una vista con la lista de puntos de interés. Lo primero que se muestra es las categorías. El usuario puede seleccionar una categoría y puede ver la información más importante de los puntos de interés. Cuando se selecciona un Punto de Interés la aplicación cambia a una actividad que muestra la información a detalle.

Arquitectura de La Información Móvil. La arquitectura de la información (IA) es el fundamento de un producto móvil o sitio web. Un producto bien diseñado, con buen diseño visual puede fallar debido a una arquitectura de la información deficiente. Los productos móviles con verdadero éxito siempre tienen una arquitectura de información bien pensada. Con la arquitectura de la información debe asegurarse de que el usuario cuando llega al sitio web o aplicación móvil puede encontrar rápidamente la información.

Mapa del sitio. Crear el mapa del sitio es lo primero que se define en la arquitectura de la información, representan visualmente la relación de los contenidos a otros contenidos y proporcionan un mapa de cómo el usuario va a viajar por el espacio de información.

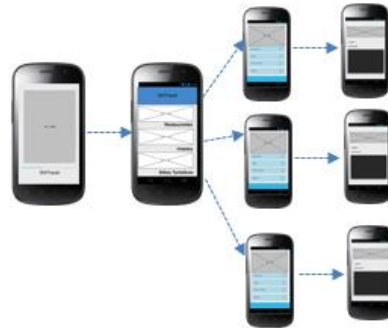
Figura 84. Mapa del Sitio



Fuente: elaboración Propia

Mapa de Contenido. Los mapas de contenido muestran el orden en que los usuarios viajan a través de la arquitectura de la información de un sitio, se puede ver en una pantalla los primeros elementos que se ofrece al usuario en una interfaz mucho más intuitiva, lo que indica inmediatamente qué tipo de contenido puede esperar el usuario. A continuación se presenta el mapa de contenido para la aplicación.

Figura 85. Mapa de contenidos de la aplicación

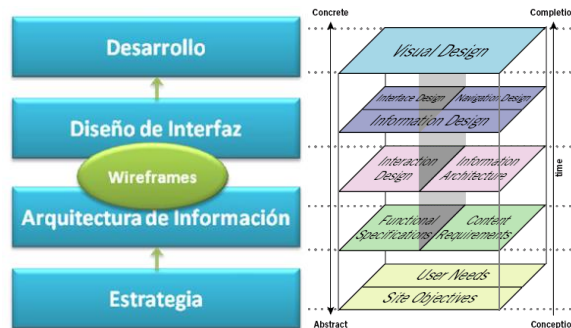


Fuente: Elaboración Propia

Wireframes. El Wireframe es una jerarquización de contenidos distribuida visualmente y una esquematización de la interfaz. Gráficamente, son estructuras muy simples y están enfocados a visualizar la distribución de los contenidos dentro de una pantalla.

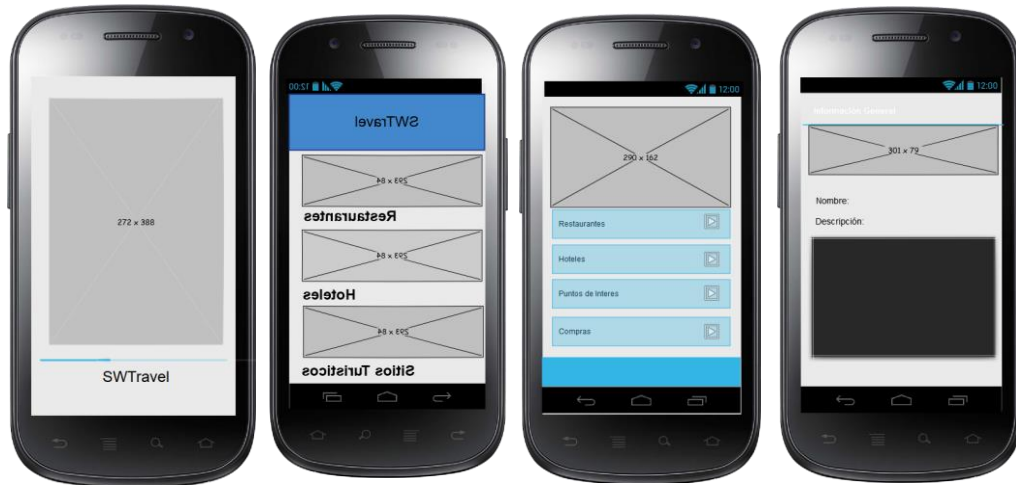
El Wireframe es el puente que une la Arquitectura de Información y el Diseño. Pasa de la “mentalidad estructural” de un mapa de contenidos, dónde decidimos y ordenamos los contenidos de la aplicación, a la emocionalidad del Diseño de Interfaz.

Figura 86. Relación entre Arquitectura de la Información y diseño



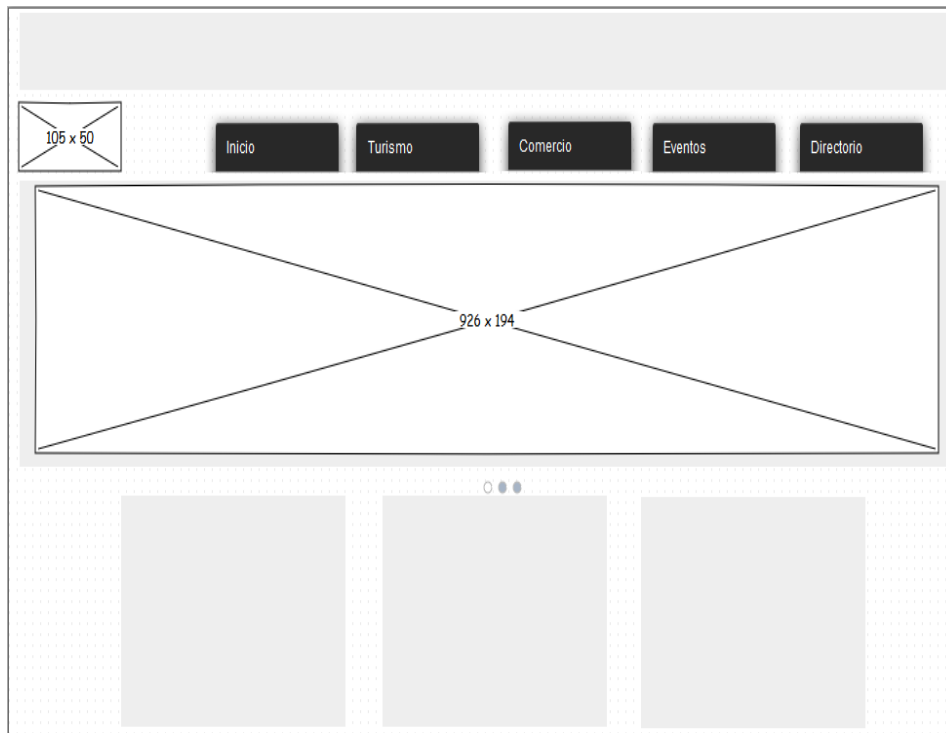
Fuente: ALCARAZ MARTÍNEZ, Rubén y RIBERA TURRÓ, Mireia. Mapas digitales y aplicaciones basadas en la localización: mejoras en su accesibilidad para las personas ciegas. [En línea]. Disponible en Internet: <<http://www.nosolousabilidad.com/>>

Figura 87. Wireframes de la AppMovil



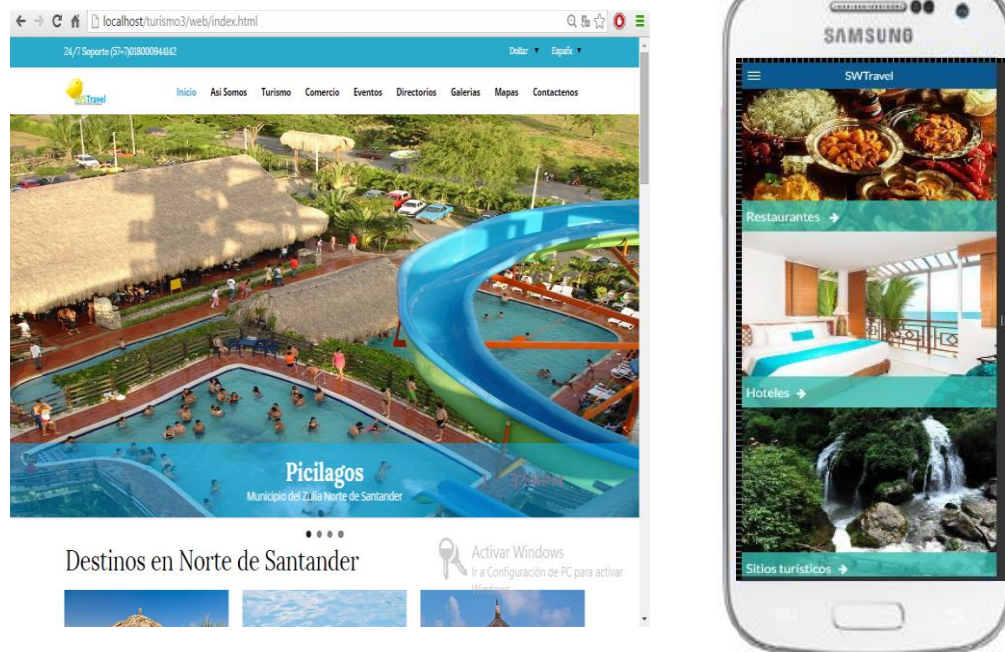
Fuente: Elaboración propia

Figura 88. Wireframes de la web



Fuente: Elaboración propia.

Figura 89. Aplicación AppMóvil y Web

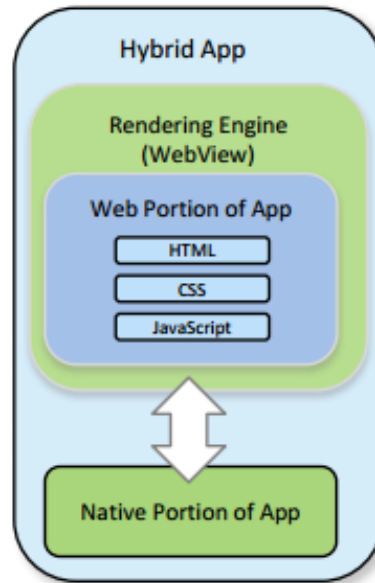


Fuente: Elaboración Propia

Codificación. Para el proceso de construcción de la aplicación móvil se toma el concepto de usar herramientas que permita le realización de una sola programación y después exportar a una plataforma deseada, para ello se utilizó la herramienta Apache Córdoba. Y para la parte móvil se utilizó una plantilla de Bootstrap con el fin de facilitar y agilizar la construcción.

Arquitectura para aplicaciones a partir de LOD. Las aplicaciones escritas con HTML5 + CSS3 + JS. Todo el contenido se carga directamente en la aplicación o en la página web según sea el caso. Teniendo en cuenta que la aplicación App es una aplicación híbrida, toma lo mejor de ambos Mundos Web y Nativo para brindar una solución eficiente. Esta puede estar encapsulada en un paquete en la aplicación o se descarga a través App Store o Market.

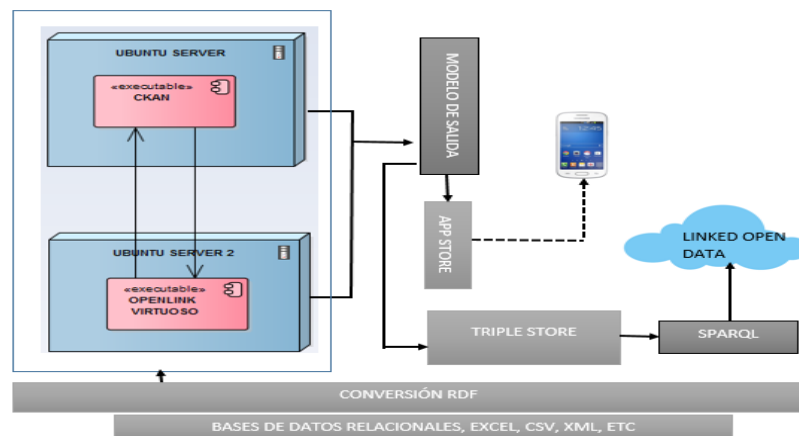
Figura 90. Arquitectura de una aplicación Híbrida



Fuente: IBM¹⁵⁴

El siguiente modelo representa la forma como las aplicaciones pueden acceder a los datos abiertos y enlazados.

Figura 91. Acceso a LOD desde APP.



Fuente: Elaboración propia

¹⁵⁴ <ftp://public.dhe.ibm.com/software/es/events/doc/start013/14/4-Definicion-de-una-arquitectura-para-aplicaciones-moviles.pdf>

4. DISCUSION

4.1 COMPARACION DE LOS RESULTADOS CON TRABAJOS EXISTENTES

Hay que tener en cuenta que Respecto de otro trabajo realizado en Colombia, a continuación se presentan algunas comparaciones importantes.

Cuadro 21. Comparaciones importantes

Trabajo	Comparativo
Observatorio Astronómico Nacional - Colombia. Universidad Nacional	Es un proyecto para manejar datos, diseñados exclusivamente para investigación, debido a eso no existe un plan de divulgación y una oportunidad para desarrolladores y ciudadanos en general.
Proyecto de Datos Abiertos Colombia	Tiene muy pocos registros, en cuyo caso resultan inútiles análisis de datos o aplicaciones. No tienen suficientes características. Hacen falta tomar más datos. No tienen suficiente granularidad.

4.2 APORTES DEL PROYECTO

Los aportes del proyecto: "MODELO BASADA EN ARQUITECTURA OPEN Y LINKED DATA PARA EL SECTOR TURÍSTICO DE NORTE DE SANTANDER". Deja importantes aportes para la comunidad de investigadores, desarrolladores, científicos de datos, ciudadanos, comunidad empresarial, emprendedora e innovadora, entidades públicas y privadas como:

- Documento con un estudio de estado de arte, técnicas, herramientas y métodos para transformación de datos a partir de arquitecturas tradicionales, implementación de mecanismos y políticas para la consolidación de una cultura de datos abiertos, herramientas y plataformas para llevar a cabo el proceso de despliegue, configuración y puesta en marcha.

- Un Modelo ontológico que permita la integración de datos sobre información turística, una aplicación web y un App que muestra sitios de interés y que sirven como objeto de estudio para nuevas implementaciones.
- Abierta la posibilidad de implementar la plataforma LOD que le permita mayor visibilidad de la ciudad en los entornos digitales para generar mayor competitividad y desarrollo económico.
- Este trabajo sirve de referente para que otros sectores del departamento como salud, energía, educación, agricultura implementen el proyecto y alternativas para otros departamentos.
- Este trabajo queda de guía para que innovadores e investigadores tengan una guía para desarrollar aplicaciones e implementar plataformas de LOD.

5. CONCLUSIONES

Son muchas las cosas que nos depara el futuro tecnológico entre ellos la visión de la Web Semántica que introduce mejoras significativas a la World Wide Web. Sin embargo, la brecha entre las tecnologías web tradicionales y tecnologías de la web semántica son muy grandes, por lo tanto, esto retrasa la adopción. Los conceptos que describen la Web Semántica tienen más de 16 años lo cual en términos tecnológico es un periodo largo, sin embargo, la puesta en práctica y masificación de esta iniciativa a un es poca. Sólo hay pocas aplicaciones semánticas y los datos que se pueden usar son pocos, principalmente Latinoamérica las iniciativas desarrolladas obedecen a trabajos de investigación que poco se han puesto en práctica. Otra situación es que no hay muchos datos codificados en formatos para la Web Semántica basados en RDF, porque existen muy pocas aplicaciones que permitan la generación de datos enlazados. Plataforma más notable en Linked Open Data, es la base de datos DBpedia y que ha ayudado a incentivar la iniciativa para usos reales. Hay miles de millones de triples disponibles que pueden ser procesados y utilizados por las aplicaciones de la Web Semántica. El conjunto de datos DBpedia se ha convertido en un gran éxito y es utilizado por muchas aplicaciones en todo el mundo para resolver diferentes problemas.

Se realizó un estudio sobre las iniciativas de datos abiertos y enlazados, que nos da una idea del estado del arte de este movimiento a través de este sector y las posibilidades en otros sectores. Algunos hallazgos importantes es que son pocos los trabajos realizados en Colombia. La iniciativa de datos abiertos en Colombia está en una fase inicial y falta mucho camino por recorrer, la información esta desactualizada, En cuanto a las herramientas, la mayoría de ellos han desarrollado sus propios portales de datos para gestionar sus datos y han dejado de utilizar plataformas como CKAN. CKAN es la plataforma de gestión de datos más utilizada en el mundo.

En este trabajo, la iniciativa de datos abiertos enlazados del sector turístico de Norte de Santander se explicó, considerando un ecosistema específico que da una visión completa de la propuesta de valor del proyecto, que le permite generar desarrollo económico y la visibilidad a la región y las posibilidades que de ella se desprenden y en el mismo tiempo por el valor añadido que el proyecto genera a través de los datos abiertos.

En esta tesis se ha demostrado que se puede trabajar en Linked Open Data y se puede lograr conversiones de datos de fuentes heterogéneas que son proporcionadas por muchas organizaciones y ciudadanos involucrados con el

sector turístico de Norte de Santander. Además que proporciona un enfoque más similar a lo que se hizo en los Estados Unidos y en el Reino Unido, donde la iniciativa de datos abiertos fue diseñada para apoyar Linked Open Data desde el comienzo. Este último enfoque podría llegar a ser más sostenible a largo plazo.

El valor añadido que se puede dar al portal de datos abiertos y enlazados es el resultado derivado de las ideas de negocios, proyectos o actividades de investigación, que pueden estar dirigidos a diferentes perfiles, como estudiantes, investigadores, emprendedores o ciudadanos. Estas ideas pueden generar valor económico en el proyecto, si pueden crear modelos de negocio innovadores. Estos modelos de negocio están limitados por la diversa cantidad y heterogeneidad de los datos que se proporciona en la plataforma de datos abierta. Esta heterogeneidad se relaciona con la cantidad de datos que se pueden publicar como Hoteles, Restaurantes, Universidades, etc.

Además, se generó modelo de negocio Canvas de la plataforma LOD, con el fin de saber cómo el proyecto de datos abiertos enlazados puede influir y evolucionar como parte del desarrollo económico de la región especialmente en el sector turístico. Básicamente, algunas partes pueden ser altamente influenciadas tales como la visibilidad local, nacional e internacional, la masificación de los datos, el espíritu empresarial y la innovación, nivel de sostenibilidad económica y el nivel de influencia en la sociedad. El último punto puede ser apoyado por los nuevos servicios que se pueden generar a través del uso de la plataforma, con el apoyo en el movimiento de innovación abierta y los programas de emprendedores como Apps.co.

Las plataformas de datos abiertos son inútiles si las personas no utilizan los datos. Por eso, la participación de los usuarios es importante con el fin de descubrir el valor oculto de los datos publicados. Para cumplir con este objetivo, se deben crear estrategias y un plan de comunicación. Un grupo relevante son los desarrolladores de aplicaciones, ya que pueden utilizar los datos para crear nuevos servicios a través de Internet o aplicaciones móviles. Eventos y concursos con enfoques rápidos e iterativos que pueden ayudar a los desarrolladores a crear nuevas soluciones de una manera más dinámica.

Finalmente hay problemas que condicionan la eficiencia y desarrollo de las iniciativas, tanto de datos abiertos como de datos enlazados y es la disponibilidad de los datos para la publicación y el compromiso de los actores encargados. Para evitar esto se debe hacer seguimiento continuo al proyecto y formular políticas internas de publicación de datos, además tratar de establecer una cultura dentro de la organización.

6. RECOMENDACIONES Y TRABAJOS FUTURO

Una parte de los datos de Turismo de Norte de Santander que está disponible se utilizó para la transformación LOD son muchos los datos que no fueron incluidos en el proceso de conversión todavía. Esto es debido a que el enfoque de esta tesis era mostrar cómo se podía implementar una estrategia de Datos abierto y enlazados y buscar las ventajas que en términos económico puede generar para la región. Los datos que se quedaron por fuera se pueden convertir en el futuro entre los que se encuentran más sitios de interés y estadísticas, etc. A demás de todas las rutas. Como por ejemplo pasajes naturales, senderismo, rutas de bicicletas, etc. También buscar opciones para generar consultas complejas.

También es posible añadir datos de otras ciudades en Linked Open Data utilizando o mejorando la ontología establecida en este proyecto. Por ejemplo, el turismo de los otros departamentos. A demás es importante trabajar sobre los otros sectores como agricultura, salud, energía.

Es posible aumentar el grado de interconexión a través de vínculos manuales para el uso de más pistas para establecer automáticamente dichos enlaces. Por ejemplo, sería posible detectar Puntos de interés turístico relacionados no sólo mediante la comparación de sus etiquetas, sino también mediante la detección de proximidad, debido a su posición geográfica (longitud / latitud y los datos de direcciones). Conociendo la ubicación y límites de los municipios que están disponibles en la plataforma LOD, también sería posible asignar una propiedad que une al objeto municipio correspondiente a los puntos de interés que no tienen ningún dato de direcciones adjuntas. Esto requeriría procesamiento geoespacial.

Implementar la plataforma en el datacenter de la gobernación y hacer realidad las políticas y procedimientos de LOD, Usar la plataforma para abrir y enlazar datos de otros sectores y buscar la forma de relacionar de tal forma que se pueda sacar provecho.

Explorar las API y las consultas a través del punto SPARQL endpoint, además configurar y diseñar proyectos para que se puedan tomar datos de sensores con el fin de brindar mayor información a los desarrolladores y ciudadanía en general.

BIBLIOGRAFIA

- Alexander, K., Cyganiak, R., Hausenblas, M., & Zhao, J. (2009, Abril). Describing Linked Datasets. Recuperado de <http://richard.cyganiak.de/2008/papers/void-ldow2009.pdf>
- Archer, P., Dekkers, M., Goedertier, S., & Loutas, N. (2013). Study on business models for Linked Open Government Data. Recuperado de http://ec.europa.eu/isa/documents/study-on-business-models-open-government_en.pdf
- Auer, S., Bryl, V., & Tramp, S. (Eds.). (2014). Linked Open Data--Creating Knowledge Out of Interlinked Data. Recuperado de <http://link.springer.com/book/10.1007%2F978-3-319-09846-3>
- Beckett, D., Berners-Lee, T., Prud'hommeaux, E., & Carothers, G. (2012) Turtle: Terse RDF Triple Language. Recuperado de <http://www.w3.org/TR/turtle/>
- Berendt, B., Hotho, A., & Stumme, G. (2002). Towards semantic web mining. In *The Semantic Web—ISWC 2002* (pp. 264-278). Springer Berlin Heidelberg.
- Berners-Lee, T., & Fischetti, M. (1999). *Weaving the web; the original design and ultimate destiny of the world wide web by its inventor (2 cassettes)* Harper Audio.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific american*, 284(5), 28-37.
- Berners-Lee, T., Hall, W., Hendler, J. A., O'Hara, K., Shadbolt, N., & Weitzner, D. J. (2006). A framework for web science. *Foundations and trends in Web Science*, 1(1), 1-130.
- Berrueta, D., & Phipps, J. (2008). *Best practice recipes for publishing RDF vocabularies-w3c working group note*. Recuperado de <http://www.w3.org/TR/2008/NOTE-swbp-vocab-pub-20080828/>

- Bizer, C., & Cyganiak, R. (2006, November). D2r server-publishing relational databases on the semantic web. Recuperado de <http://www.w3.org/2007/03/RdfRDB/papers/d2rq-positionpaper/>
- Bizer, C., Cyganiak, R., & Heath, T. (2007). How to publish linked data on the web. Recuperado de <http://events.linkedata.org/iswc2008tutorial/how-to-publish-linked-data-iswc2008-slides.pdf>
- Bizer, C., Heath, T., Idehen, K., & Berners-Lee, T. (2008). Linked data on the web (LDOW2008). *Proceedings of the 17th International Conference on World Wide Web*, Beijing, China. 1265-1266. doi:10.1145/1367497.1367760
- Cifuentes-Silva, F., Sifaqui, C., & Labra-Gayo, J. E. (2011). Towards an architecture and adoption process for linked data technologies in open government contexts: A case study for the library of congress of chile. *Proceedings of the 7th International Conference on Semantic Systems*, Graz, Austria. 79-86. doi:10.1145/2063518.2063529
- Data, L. (2014). Connect Distributed Data across the Web. Recuperado de <http://linkedata.org/>
- Decker, S., Fensel, D., Van Harmelen, F., Horrocks, I., Melnik, S., Klein, M. C., & Broekstra, J. (2000). Knowledge Representation on the Web. *Description Logics*, 33(6), 89-97. Recuperado de <http://www.cs.vu.nl/~frankh/postscript/DL00.pdf>
- Departamento Nacional de Planeación (2014). Plan sectorial de Turismo 2014-2018: Turismo para la construcción de la Paz. Recuperado de <http://www.mincit.gov.co/minturismo/descargar.php?id=71713>
- Fernández Carrasco, F. (2011). Propuesta de ontología para puntos de interés turístico. Recuperado de <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/8133/1/ffernandezcarTFC0611.pdf>

- Frosterus, M., Tuominen, J., & ONEN, E. H. (2014, November). Facilitating Re-use of Legal Data in Applications—Finnish Law as a Linked Open Data Service. *Legal Knowledge and Information Systems: JURIX 2014: The Twenty-Seventh Annual Conference*. 271(5), 115-124. doi: 10.3233/978-1-61499-468-8-115.
- Hartig, O., Bizer, C., & Freytag, J. (2009). Executing SPARQL queries over the web of linked data. *Proceedings of the 8th International Semantic Web Conference*, Chantilly, VA. 293-309. doi: 10.1007/978-3-642-04930-9_19
- Hartig, O., Sequeda, J., Taylor, J., & Sinclair, P. (2010). How to consume linked data on the web: Tutorial description. *Proceedings of the 19th International Conference on World Wide Web*, Raleigh, North Carolina, USA. 1347-1348. doi:10.1145/1772690.1772923
- Hyvönen, E., Tuominen, J., Alonen, M., & Mäkelä, E. (2014). Linked Data Finland: A 7-star model and platform for publishing and re-using linked datasets. Recuperado de <http://www.seco.tkk.fi/publications/2014/hyvonen-et-al-ldf-2014.pdf>
- Jacobs, J., & Linden, A. (2002). Semantic web technologies take middleware to the next level. Recuperado de <https://www.gartner.com/doc/366500/semantic-web-technologies-middleware-level>
- Kim, J., Kim, J., & Kim, C. (2007). Ontology-based user preference modeling for enhancing interoperability in personalized services. 45(56), 903-912. doi:10.1007/978-3-540-73283-9_98.
- Lopez-de-Ipina, D., Klein, B., Vanhecke, S., & Perez-Velasco, J. (2013). Towards ambient assisted cities and citizens. *Advanced Information Networking and Applications Workshops (WAINA), 2013 27th International Conference on*, 1343-1348. doi:10.1109/WAINA.2013.203
- Luna, J. A. G., Bonilla, M. L., & Torres, I. D. (2012). Metodologías y métodos para la construcción de ontologías. *Scientia et Technica*, 2(50), 133-140.
- Maali, F., Erickson, J., & Archer, P. (2014). Data catalog vocabulary (DCAT). Recuperado de <http://www.w3.org/TR/vocab-dcat/>

- Pattuelli, M. C. (2012). Personal name vocabularies as linked open data: A case study of jazz artist names. *Journal of Information Science*, 38(6), 558-565.
- Piedrabuena Moraleda, A., y Criado Fernández, L. (2013). Opendata, Oportunidad Escondida y Semilla de la Web Semántica. *Ruiderae: Revista de Unidades de Information. Description de Experiencias y Resultados Aplicados*. 1(6), 3-21.
- Schultz, A., Matteini, A., Isele, R., Bizer, C., & Becker, C. (2011). Ldif-linked data integration framework. Recuperado de http://ceur-ws.org/Vol-782/SchultzEtAl_COLLD2011.pdf
- Vilches-Blázquez, L. M., Corcho, O., González, A., Villazon-Terrazas, B., y Gómez, J. M. (2014). Combinando Open y Linked Data en el marco de las Smart Cities: Un caso de uso sobre alquiler de bicicletas. Recuperado de <http://www.innprontaciudad2020.es/index.php/es/documentacion-ficheros-relativos-al-proyecto/white-papers/91-combinando-open-y-linked-data-en-el-marco-de-las-smart-cities/download>
- Wayne Gould, R. (2012). Open innovation and stakeholder engagement. *Journal of technology management & innovation*, 7(3), 1-11.

ANEXOS

Anexo A. Siglas

SQL. Structured Query Language

URI. Universal Resource Identifier

W3C. World Wide Web Consortium

XML. Extensible Markup Language

XSD. XML Schema

XSL. Extensible Stylesheet Language

XSLT. XSL Transformation

OWL Web Ontology Language

RDF Resource Description Framework

RDFS RDF Schema

SPARQL SPARQL Protocol and RDF Query Language

LOD Linked Open Data

GML Geography Markup Language

HTML Hypertext Markup Language

JSON JavaScript Object Notation

CSV Comma Separated Values

DL Description Logics

FOAF Friend of a Friend

Anexo B. Glosario

Disponible en: <https://drive.google.com/open?id=0ByT4PZ-56D4sfkVNdWxRWIhnZTRBOFF4dUo1XzlhTkE2RGICcXR3bV8ycGY3UEhBR1dMcWc&authuser=0>

Anexo C. Definición de propiedades para LODObject

Disponible en: <https://drive.google.com/open?id=0ByT4PZ-56D4sfkVNdWxRWIhnZTRBOFF4dUo1XzIhTKE2RGICcXR3bV8ycGY3UEhBR1dMcWc&authuser=0>

Anexo D. Documento RFD-LOD Ontology

Disponible en: <https://drive.google.com/open?id=0ByT4PZ-56D4sfkVNdWxRWIhnZTRBOFF4dUo1XzIhTKE2RGICcXR3bV8ycGY3UEhBR1dMcWc&authuser=0>

Anexo E. Inventario de Turismo

Disponible en: <https://drive.google.com/open?id=0ByT4PZ-56D4sfkVNdWxRWIhnZTRBOFF4dUo1XzlhTke2RGICcXR3bV8ycGY3UEhBR1dMcWc&authuser=0>

Anexo F. Código Fuente.

Disponible en: **Carpeta código fuente**