



**PROTOTIPO DE SOFTWARE PARA LA EVALUACIÓN DE PRINCIPIOS DE  
DATOS ABIERTOS CASO DE ESTUDIO**

**JOSÉ ALIRIO ROJAS CLAVIJO  
DIEGO FELIPE ARIZA AMADO**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS  
TRABAJO DE INVESTIGACIÓN TECNOLÓGICA  
BOGOTÁ D.C., COLOMBIA  
2019**

**PROTOTIPO DE SOFTWARE PARA LA EVALUACIÓN DE PRINCIPIOS DE  
DATOS ABIERTOS**

**JOSÉ ALIRIO ROJAS CLAVIJO  
DIEGO FELIPE ARIZA AMADO**

**TESIS PRESENTADA COMO REQUISITO PARCIAL PARA OPTAR AL TÍTULO  
DE: INGENIERO DE SISTEMAS Y COMPUTACIÓN**

**DIRECTOR (A):  
M.SC. JOHN ALEXANDER VELANDIA VEGA**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
TRABAJO DE INVESTIGACIÓN TECNOLÓGICA  
BOGOTÁ D.C., COLOMBIA  
2019**



## Atribución-NoComercial-SinDerivadas 2.5 Colombia (CC BY-NC-ND 2.5)

La presente obra está bajo una licencia:

### **Atribución-NoComercial-SinDerivadas 2.5 Colombia (CC BY-NC-ND 2.5)**

Para leer el texto completo de la licencia, visita:

<http://creativecommons.org/licenses/by-nc-nd/2.5/co/>

#### Usted es libre de:



Compartir - copiar, distribuir, ejecutar y comunicar públicamente la obra

#### Bajo las condiciones siguientes:



**Atribución** — Debe reconocer los créditos de la obra de la manera especificada por el autor o el licenciante (pero no de una manera que sugiera que tiene su apoyo o que apoyan el uso que hace de su obra).



**No Comercial** — No puede utilizar esta obra para fines comerciales.



**Sin Obras Derivadas** — No se puede alterar, transformar o generar una obra derivada a partir de esta obra.

## **Nota de aceptación**

Aprobado por el comité de grado en cumplimiento de los requisitos exigidos por la Facultad de Ingeniería y la Universidad Católica de Colombia para optar al título de Ingenieros de Sistemas.

---

M.Sc. John Alexander Velandia Vega  
Director de trabajo de grado

---

M.Sc. John Alexander Velandia Vega  
Coordinador de trabajo de grado

---

Firma del Jurado

Bogotá, junio de 2019

## **AGRADECIMIENTOS**

A Dios por ser siempre la guía que necesitamos en cada momento, a cada una de las personas que coadyuvaron en la consecución de este proyecto, a nuestros padres y familia por creer siempre en nosotros y mostrarnos su apoyo incondicional, al director de trabajo de grado MSc. John Velandia por estar siempre dispuesto a ser una guía y brindarnos sus conocimientos.

## CONTENIDO

	<b>PÁG.</b>
GLOSARIO.....	11
LISTA DE SÍMBOLOS Y ABREVIATURAS .....	11
<b>RESUMEN</b> .....	<b>12</b>
<b>ABSTRACT</b> .....	<b>13</b>
<b>INTRODUCCIÓN</b> .....	<b>14</b>
<b>1. GENERALIDADES</b> .....	<b>16</b>
1.1 ANTECEDENTES.....	16
1.2 PLANTEAMIENTO DEL PROBLEMA.....	17
1.2.1 Descripción del problema .....	17
1.2.2 Formulación del Problema.....	18
1.3 OBJETIVOS .....	19
1.3.1 Objetivo General .....	19
1.3.2 Objetivos Específicos.....	19
1.4 JUSTIFICACIÓN.....	19
1.5 ALCANCES Y LIMITACIONES .....	20
1.5.1 Alcances.....	20
1.5.2 Limitaciones .....	20
1.6 MARCO TEÓRICO.....	20
1.6.1 Datos Abiertos (Open Data) .....	21
1.6.2 Condiciones de los Datos Abiertos en América Latina .....	22
1.6.3 Comportamiento de los datos abiertos a nivel internacional .....	23
1.6.4 Actividad Nacional en Datos Abiertos .....	25
1.7 MARCO CONCEPTUAL .....	26
1.8 METODOLOGÍA .....	28
1.8.1 Tipo de Estudio .....	28
1.8.2 Fuentes de información .....	28
1.8.3 Diseño Metodológico .....	28
1.8.4 Estrategias de Comunicación .....	30
1.8.5 Recursos .....	30
1.8.6 Productos a entregar .....	31
<b>2. ANÁLISIS DE LOS PRINCIPIOS INTERNACIONALES Y LEYES COLOMBIANAS REFERENTE A LOS DATOS ABIERTOS</b> .....	<b>32</b>
2.1 INFORMACIÓN RECOPIADA .....	32
2.1.1 Carta de Datos Abiertos.....	33
2.1.2 Sunlight Foundation .....	34
2.1.3 Datos de Gobierno Abierto .....	35
2.1.4 Sr. Tim Berners Lee.....	36
2.1.5 Ley 1712 de 2014 .....	37
2.1.6 Carta Iberoamericana de Gobierno Abierto .....	38
2.1.7 Grupo Revolución de Datos.....	39
2.2 ANÁLISIS DE INFORMACIÓN RECOPIADA.....	40
2.2.1 Principios en Común.....	40

2.3	CONCLUSIÓN DEL ANÁLISIS .....	48
<b>3.</b>	<b>DISEÑO PROTOTIPO DE SOFTWARE .....</b>	<b>49</b>
3.1	DIAGRAMAS.....	49
3.1.1	Diagrama de despliegue .....	49
3.1.2	Diagrama de componetes.....	50
3.1.3	Diagrama de Actividades .....	51
3.1.4	Diagrama de Casos de Uso.....	52
3.2	REQUERIMIENTOS DE SOFTWARE.....	54
3.2.1	Requerimientos Funcionales .....	55
3.2.2	Requerimientos No Funcionales.....	55
3.3	MOCKUPS .....	60
<b>4.</b>	<b>ARQUITECTURA DE DATOS .....</b>	<b>62</b>
<b>5.</b>	<b>DEFINICIÓN DEL PROCESO DE EVALUACIÓN DE LA MADUREZ .....</b>	<b>64</b>
5.1	COMPLETOS.....	65
5.1.1	MEDICIÓN .....	65
5.2	PRIMARIOS .....	66
5.3	OPORTUNO.....	67
5.4	ACCESIBLE .....	68
5.4.1	Medición .....	69
5.5	PROCESABLE POR MÁQUINA .....	69
5.6	NO DISCRIMINATORIO .....	70
5.7	NO PROPIETARIO .....	72
5.8	LICENCIA ABIERTO.....	73
<b>6.</b>	<b>EVALUACIÓN Y ANÁLISIS DE RESULTADOS. ....</b>	<b>74</b>
6.1	RAPIDMINER.....	74
6.2	COMPLETO .....	75
6.3	PRIMARIO.....	75
6.4	OPORTUNO.....	76
6.5	ACCESIBLE .....	76
6.5.1	Análisis por tipo de acceso a SODA API.....	77
6.6	PROCESABLE POR MÁQUINA .....	78
6.7	NO DISCRIMINATORIO .....	79
6.8	NO PROPIETARIO .....	80
6.9	LICENCIA LIBRE .....	81
<b>7.</b>	<b>CONCLUSIONES. ....</b>	<b>82</b>
<b>8.</b>	<b>BIBLIOGRAFÍA .....</b>	<b>83</b>
<b>9.</b>	<b>ANEXOS .....</b>	<b>86</b>
	ANEXO A. MANUAL DEL USUARIO .....	87
	ANEXO B. MANUAL INSTALACIÓN RAPIDMINER SERVER .....	94
	ANEXO C. CREACIÓN DE INSTANCIA .....	103
	ANEXO D. INSTALACIÓN BASE DE DATOS Y TOMCAT .....	111

## LISTA DE TABLAS

TABLA 1. MATRIZ FODA .....	30
TABLA 2. PRESUPUESTO .....	31
TABLA 3. COMPARACIÓN LEY 1712 Y EL OPEN DATA CHÁRTER.....	41
TABLA 4. COMPARACIÓN LEY 1712 DE 2014 Y LOS PRINCIPIOS DE LA FUNDACIÓN SUNLIGHT .	42
Tabla 5. Comparación entre la Ley 1712 de 2014 y el Open Government Data	43
TABLA 6. COMPARACIÓN ENTRE LA LEY 1712 DE 2014 CON TIM BERNERS LEE .....	44
TABLA 7. COMPARACIÓN ENTRE LA LEY 1712 DE 2014 Y LA CARTA IBEROAMERICANA DE GOBIERNO ABIERTO.....	44
TABLA 8. COMPARACIÓN ENTRE LA LEY 1712 DE 2014 Y EL DATA REVOLUTION GROUP.....	46
TABLA 9. COMPARATIVO GENERAL.....	47
TABLA 10. CASO DE USO INGRESAR ESTRUCTURA DE DATOS.....	53
TABLA 11. CASO DE USO INGRESAR SET DE DATOS.....	54
TABLA 12. CASO DE USO VISUALIZACIÓN DE LOS RESULTADOS DEL ANÁLISIS.....	54
TABLA 13. DESCRIPCIÓN DE REQUERIMIENTO 1 .....	55
TABLA 14. DESCRIPCIÓN DE REQUERIMIENTO 2 .....	56
TABLA 15. DESCRIPCIÓN DE REQUERIMIENTO 3 .....	56
TABLA 16. DESCRIPCIÓN DE REQUERIMIENTO 4 .....	57
TABLA 17. DESCRIPCIÓN DE REQUERIMIENTO 5 .....	57
TABLA 18. DESCRIPCIÓN DE REQUERIMIENTO 6 .....	58
TABLA 19. DESCRIPCIÓN DE REQUERIMIENTO 7 .....	58
TABLA 20. DESCRIPCIÓN DE REQUERIMIENTO 5 .....	59
TABLA 21. DESCRIPCIÓN DE REQUERIMIENTO 7 .....	59
TABLA 22. ESTRUCTURA DE DATOS .....	63
TABLA 23. DESCRIPCIÓN DE LA MEDICIÓN DEL PRINCIPIO COMPLETO DE DATOS ABIERTO.....	65
TABLA 24. DESCRIPCIÓN PARA LA MEDICIÓN DE UN CONJUNTO DE DATOS PRIMARIOS .....	66
TABLA 25. DESCRIPCIÓN PARA LA MEDICIÓN DE UN CONJUNTO DE DATOS OPORTUNOS .....	67
TABLA 26. DESCRIPCIÓN PARA LA MEDICIÓN DE UN CONJUNTO DE DATOS ACCESIBLES .....	68
TABLA 27. DEFINICIÓN DEL LICENCIAMIENTO ABIERTO .....	73
TABLA 28. EVALUACIÓN DE RESULTADOS .....	74
TABLA 29. RESULTADO PRINCIPIO ACCESIBLE .....	78

## LISTA DE FIGURAS

PÁG.

FIGURA 1. MODELO EN ESPIRAL .....	29
FIGURA 2. PRINCIPIOS DE DATOS ABIERTOS Y GOBIERNO ABIERTO .....	32
FIGURA 3. DIAGRAMA DE DESPLIEGUE.....	50
FIGURA 4. DIAGRAMA DE COMPONENTES .....	51
FIGURA 5. DIAGRAMA DE ACTIVIDADES .....	51
FIGURA 6. DIAGRAMA DE CASOS DE USO .....	52
FIGURA 7. INICIAR APLICACIÓN .....	60
FIGURA 8. RESULTADOS DEL ANÁLISIS .....	61
FIGURA 9. GRÁFICAS DE RESULTADOS.....	61
FIGURA 10. CONFIGURACIÓN PRINCIPIO COMPLETO .....	66
FIGURA 11. RESULTADO PRINCIPIO COMPLETOS.....	66
FIGURA 12. CONFIGURACIÓN PRINCIPIO PRIMARIO .....	67
FIGURA 13. RESULTADO PRINCIPIO PRIMARIO.....	67
FIGURA 14. CONFIGURACIÓN PRINCIPIO OPORTUNO .....	68
FIGURA 15. RESULTADO PRINCIPIO OPORTUNO .....	68
FIGURA 16. CONFIGURACIÓN PRINCIPIO PROCESABLE POR MAQUINA .....	70
FIGURA 17. RESULTADO PRINCIPIO PROCESABLE POR MAQUINA .....	70
FIGURA 18. CONFIGURACIÓN PRINCIPIO NO DISCRIMINATORIO .....	72
FIGURA 19. RESULTADO PRINCIPIO NO DISCRIMINATORIO .....	72
FIGURA 20. CONFIGURACIÓN NO PROPIETARIO.....	73
FIGURA 21. RESULTADO PRINCIPIO NO PROPIETARIO.....	73
FIGURA 22. RESULTADO PRINCIPIO LICENCIA ABIERTO .....	74
FIGURA 23 RESULTADO PRINCIPIO COMPLETO .....	75
FIGURA 24. RESULTADO PRINCIPIO PRIMARIO.....	75
FIGURA 25. RESULTADO PRINCIPIO OPORTUNO .....	76
FIGURA 26. PROCESO DE ACCESIBLE.....	77
FIGURA 27. PARÁMETROS PARA LA SOLICITUD DE TENION.IO .....	77
FIGURA 28. RESTFUL SERVICES.....	77
FIGURA 29. PROCESABLE POR MÁQUINA .....	79
FIGURA 30. RESULTADO PRINCIPIO PROCESABLE POR MAQUINA .....	79
FIGURA 31. RESULTADO PRINCIPIO NO DISCRIMINATORIO .....	80
FIGURA 32. RESULTADO PRINCIPIO NO PROPIETARIO .....	80
FIGURA 33. RESULTADO PRINCIPIO LICENCIA LIBRE .....	81
FIGURA 34. ANEXO A - DATOS ABIERTOS .....	87
FIGURA 35. ANEXO A - SET DE DATOS.....	87
FIGURA 36. ANEXO A - SET DE DATOS DE CONTRATOS .....	88
FIGURA 37. ANEXO A - SELECCIÓN DEL TOKEN .....	88
FIGURA 38. ANEXO A - INGRESO DE ESTRUCTURA.....	89
FIGURA 39. ANEXO A - INGRESO DE DATOS .....	90
FIGURA 40. ANEXO A - RESULTADO GENERAL .....	91
FIGURA 41. ANEXO A - RESULTADO GENERAL 1 .....	91
FIGURA 42. ANEXO A - RESULTADO GENERAL 2 .....	92
FIGURA 43. ANEXO A - RESULTADO ESPECÍFICO 1 .....	92
FIGURA 44. ANEXO A - RESULTADO ESPECÍFICO 2 .....	93

FIGURA 45. ANEXO B - DESCARGA RAPIDMINER .....	94
FIGURA 46. ANEXO B - TERMINOS RAPIDMINER .....	94
FIGURA 47. ANEXO B - INICIO DE LA DESCARGA .....	95
FIGURA 48. ANEXO B - INSTALACIÓN RAPIDMINER 1 .....	96
FIGURA 49. ANEXO B - INSTALACIÓN RAPIDMINER 2 .....	96
FIGURA 50. ANEXO B - INSTALACIÓN RAPIDMINER 3 .....	97
FIGURA 51. ANEXO B - UBICACIÓN DE LA INSTALACIÓN.....	98
FIGURA 52. ANEXO B - LICENCIA RAPIDMINER 1 .....	98
FIGURA 53. ANEXO B - LICENCIA RAPIDMINER 2 .....	99
FIGURA 54. ANEXO B - LICENCIA RAPIDMINER 3 .....	99
FIGURA 55. ANEXO B - CONFIGURACIÓN DE PUERTOS.....	100
FIGURA 56. ANEXO B - CONFIGURACIÓN DE LA BASE DE DATOS.....	100
FIGURA 57. ANEXO B - FINALIZAR INSTALACIÓN.....	101
FIGURA 58. ANEXO B - INSTALACIÓN EXITOSA .....	101
FIGURA 59. ANEXO B - RAPIDMINER EN EL SERVIDOR .....	102
FIGURA 60. ANEXO C - SITIO AMAZON AWS .....	103
FIGURA 61. ANEXO C – CREACIÓN DE USUARIO 1. ....	103
FIGURA 62. ANEXO C - CREACIÓN DE USUARIO 2 .....	104
FIGURA 63. ANEXO C - CONSOLA AWS.....	104
FIGURA 64. ANEXO C - INICIO DE SESIÓN .....	104
FIGURA 65. ANEXO C - CREACIÓN DE INSTANCIA EC2.....	105
FIGURA 66. ANEXO C - INICIANDO INSTANCIA .....	105
FIGURA 67. ANEXO C - SELECCIÓN DE SISTEMA OPERATIVO.....	105
FIGURA 68. ANEXO C - CONFIGURACIÓN DE INSTANCIA 1 .....	106
FIGURA 69. ANEXO C - CONFIGURACIÓN DE INSTANCIA 2 .....	106
FIGURA 70. ANEXO C - CONFIGURACIÓN DE INSTANCIA 3 .....	107
FIGURA 71. ANEXO C - CONFIGURACIÓN DE INSTANCIA 4 .....	107
FIGURA 72. ANEXO C - REGLAS DE SEGURIDAD 1.....	108
FIGURA 73. ANEXO C - REGLAS DE SEGURIDAD 2.....	108
FIGURA 74. ANEXO C - REGLAS DE SEGURIDAD 2.....	109
FIGURA 75. ANEXO C - ESTADO DE LA INSTANCIA.....	109
FIGURA 76. ANEXO C - GENERAR ACCESO AL SERVIDOR.....	110
FIGURA 77. ANEXO D - DESCARGA MONGODB .....	111
FIGURA 78. ANEXO D - UBICACIÓN DE LOS DATOS 1 .....	112
FIGURA 79. ANEXO D - UBICACIÓN DE LOS DATOS 2.....	112
FIGURA 80. ANEXO D - EJECUTANDO SERVIDOR DE BASE DE DATOS.....	112
FIGURA 81. ANEXO D - CONFIGURACIÓN DE TOMCAT 1 .....	113
FIGURA 82. ANEXO D - CONFIGURACIÓN DE TOMCAT 2 .....	114
FIGURA 83. ANEXO D - SELECCIÓN DE LA VERSIÓN DE TOMCAT .....	114
FIGURA 84. ANEXO D - DESCARGA Y SELECCIÓN DE BINARIOS .....	115
FIGURA 85. ANEXO D - EJECUTAR TOMCAT.....	115
FIGURA 86. ANEXO D - CONSOLA TOMCAT .....	116

## GLOSARIO

**CEPAL:** La Comisión Económica para América Latina (CEPAL) fue establecida por la resolución 106 (VI) del Consejo Económico y Social, del 25 de febrero de 1948, y comenzó a funcionar ese mismo año. En su resolución 1984/67, del 27 de julio de 1984, el Consejo decidió que la Comisión pasara a llamarse Comisión Económica para América Latina y el Caribe.

**CSAIL:** La Ciencia de la Computación e Inteligencia Artificial de laboratorio - conocido como CSAIL - es el más grande laboratorio de investigación en el MIT y uno de los centros más importantes del mundo de la investigación en tecnología de la información.

**Estrategia Gel:** Gobierno en línea es el nombre que recibe la estrategia de gobierno electrónico (e-government) en Colombia, que busca construir un Estado más eficiente, más transparente y más participativo gracias a las TIC.

**ONU:** La Organización de las Naciones Unidas (ONU) es una organización internacional formada por 192 países independientes. Estos se reúnen libremente para trabajar juntos en favor de la paz y la seguridad de los pueblos, así como para luchar contra la pobreza y la injusticia en el mundo.

**WWW:** La World Wide Web (WWW) o red informática mundial<sup>1</sup> es un sistema de distribución de documentos de hipertexto o hipermedios interconectados y accesibles vía Internet. Con un navegador web, un usuario visualiza sitios web compuestos de páginas web que pueden contener textos, imágenes, vídeos u otros contenidos multimedia, y navega a través de esas páginas usando hiperenlaces.

## LISTA DE SÍMBOLOS Y ABREVIATURAS

## Abreviaturas

### Abreviatura Término

CEPAL	Comisión Económica para América Latina y el Caribe
CSAIL	MIT Ciencias de la Computación y el Laboratorio de Inteligencia Artificial
<i>GEL</i>	Gobierno en Línea
ONU	Organización de las Naciones Unidas
WWW	World Wide Web

## RESUMEN

Datos abiertos son todos aquellos datos que entidades públicas liberan para el uso, análisis y control por parte de terceros, para monitorear los bienes públicos del

Estado, en Colombia la aplicación de los datos abiertos son parte de la política gubernamental para generar control y monitoreo a las entidades públicas; la implementación la realiza el Ministerio de Tecnologías de la Información, con el apoyo de leyes sancionadas por el Congreso. Esta presenta una problemática en la cual las entidades no han tomado la responsabilidad de liberar los datos correcta y oportunamente lo cual dificulta las posibilidades de realizar análisis y monitoreo con información concreta, de allí se formuló como pregunta ¿Cómo se puede comprobar el cumplimiento de la Estrategia del Gobierno Nacional sobre los principios y leyes de los datos abiertos en Colombia de acuerdo con los principios establecidos por el Open Government Data? Como objetivo principal de esta investigación es la realización de un prototipo en el cual se pueda medir la madurez de la información que se encuentra en la plataforma datos.gov.co. Para lograr dicho prototipo se esquematizo con un análisis de la información existente sobre los datos abiertos y sus principios existentes, de esta forma realizó una comparación exhaustiva referente a lo que implementa el gobierno colombiano y así poder realizar el desarrollo del prototipo el cual analice la madurez de la implementación realizada. Como resultado de la implementación del prototipo se visualizó la falta de compromiso por las entidades por la información que es suministrada en la plataforma, para lograr el objetivo de difundir los datos y monitorear el desempeño de las instituciones.

**Palabras Clave:** Big Data, Datos Abiertos, Desarrollo de Software, Gobierno Abierto.

## **ABSTRACT**

Open data is all data that public entities release for the use, analysis and control by third parties, to monitor the public goods of the State, in Colombia the application of open data is part of the government policy to generate control and monitoring of

public entities; the implementation is carried out by the Ministry of Information Technologies, with the support of laws sanctioned by Congress. This presents a problem in which entities have not taken the responsibility to release data correctly and timely which hinders the possibilities of conducting analysis and monitoring with specific information, from there it was formulated as a question How can you verify compliance with the National Government Strategy on the principles and laws of open data in Colombia in accordance with the principles established by the Open Government Data? The main objective of this research is the realization of a prototype in which the maturity of the information found on the data.gov.co platform can be measured. To achieve this prototype, it was outlined with an analysis of the existing information on the open data and its existing principles, in this way it made an exhaustive comparison regarding what the Colombian government implements and thus be able to carry out the development of the prototype which analyzes the maturity of the implementation made. Because of the implementation of the prototype, there was a lack of commitment by the entities for the information that is provided in the platform, to achieve the objective of disseminating the data and monitoring the performance of the institutions.

Keywords: Big Data, Open Data, Software Development, Open Government,

## **INTRODUCCIÓN**

Los datos abiertos van ligados a políticas gubernamentales y estas a políticas internacionales que buscan innovación, modernización de los gobiernos, el gobierno colombiano toma como pilares del gobierno abierto los siguientes: Transparencia y

Acceso a la Información Pública, Rendición de Cuentas, Participación Ciudadana, Colaboración e Innovación Pública y Ciudadana<sup>1</sup>

Para el cumplimiento de estos pilares el gobierno ha establecido la Estrategia Gobierno en Línea “GEL” a través del Decreto 1078 de 2015 artículo 2.2.9.1.1.1; el cual tiene como propósito lograr que los ciudadanos cuenten con servicios en línea de alta calidad, impulsar el empoderamiento y la colaboración de los ciudadanos con el Gobierno, encontrar diferentes formas para que la gestión en las entidades públicas sea óptima gracias al uso estratégico de la tecnología, y garantizar la seguridad y la privacidad de la información.

Los datos abiertos han llevado a la revolución en las TIC’s ya que estas tecnologías han logrado que el manejo de grandes volúmenes de información sea de alguna manera más tolerable y por consiguiente la calidad de información que es tratada sea de mayor impacto al momento de realizar su análisis y evaluación de estas.

Sin embargo, esto es apenas una primera fase constituida por el desarrollo de portales para datos abiertos, pero es necesario que las entidades que publican los datos estimen el nivel de calidad del portal como de los datos vinculados que contiene, por medio de un prototipo de madurez para portales de datos abiertos, el cual permita optimizar constantemente dichas publicaciones, en favor de los consumidores y su satisfacción por los datos obtenidos.

En ese sentido, este trabajo de grado se divide en varios capítulos los cuales contextualizan en el primero de ellos, todas las generalidades que se abordaron en el anteproyecto como son la problemática, objetivos, justificación, metodología usada, marco teórico, entre otros. A partir del capítulo dos, se resuelven los objetivos específicos planteados.

En consecuencia, el capítulo dos da cuenta del estudio referente a los distintos principios internacionales y leyes colombianas concernientes a los datos abiertos, donde se hace la comparación entre los encontrados, para así determinar cuál es la ley internacional con la cual se comparará la colombiana. Seguidamente, está el capítulo tercero donde se presenta el diseño del prototipo de software propuesto con sus respectivos diagramas, Casos de uso, requerimientos de software y mockups necesarios para el desarrollo del prototipo.

Luego en el capítulo cuarto se puede evidenciar la arquitectura de datos para el software propuesto. Ya en el capítulo quinto se define el proceso para la evaluación de la madurez de los datos abiertos en Colombia, los cuales se analizan con base en los criterios del Open Government Data y el estudio general de sus principios, los cuales dan pie al capítulo siguiente, donde se evalúan y analizan los resultados

---

<sup>1</sup> CENTRO LATINOAMERICANO DE ADMINISTRACIÓN PARA EL DESARROLLO – CLAD. *Carta Iberoamericana de Gobierno Abierto. XVII Conferencia Iberoamericana de Ministras y Ministros de Administración Pública y Reforma del Estado.* 2016. (21). p. 24. Disponible en: [https://www.sfp.gov.py/sfp/archivos/documentos/CIGA\\_2016%20final\\_knawtdh9.pdf](https://www.sfp.gov.py/sfp/archivos/documentos/CIGA_2016%20final_knawtdh9.pdf)

a partir de la teoría de los principios expuestos por el Open Government Data y su posterior validación principio por principio contra lo efectuado en Colombia, lo cual sirvió de base para la implementación del prototipo propuesto.

Por último, se presentan las conclusiones pertinentes al estudio efectuado.

## **1. GENERALIDADES**

### **1.1 ANTECEDENTES**

Los datos abiertos, es una temática que se ha tratado ya bastante tiempo, en la creación u reestructuración de la World Wide Web por Tim Berners-Lee alrededor del 1994 cuando creaba la W3C, el cual dio unos lineamientos para la reestructuración de la data que tendría la World Wide Web, pero de igual forma fue solo uno de los pioneros en la reestructuración de la data.

Con el apoyo de la información levantada en la primera fase del presente proyecto, podemos definir una guía base de principios que tienen estructuración en las políticas implementadas por el gobierno nacional. Los datos han tomado relevancia en las políticas gubernamentales, ya que estos sirven para establecer y centralizar los datos que son de acceso público para fomentar a empresas privadas, públicas o fundaciones la realizar el análisis de esta información para la creación de empleos, veeduría de procesos de entidades del estado.

## 1.2 PLANTEAMIENTO DEL PROBLEMA

### 1.2.1 Descripción del problema

En el contexto de esta investigación, se definen los datos abiertos, como todos los datos principales sin procesar ubicados en formatos estándar que proveen su acceso y admiten su reutilización, además de estar custodiados por las entidades públicas los cuales son colocados a disposición de la ciudadanía, de manera espontánea y sin limitaciones, “con el fin de que terceros puedan reutilizarlos y crear servicios derivados de los mismos”<sup>2</sup>.

En el ámbito internacional la implementación de las políticas gubernamentales para los datos abiertos, es variado según el enfoque que los diferentes gobiernos le han dado a la implementación de los datos abiertos, entidades como World Economic Forum, International Finance Corporation, Unión Internacional de Telecomunicaciones y Waseda University Institute of e-Government<sup>3</sup> califican de manera furtiva los esfuerzos económicos, tecnológicos y sociales para así dar un posicionamiento a los países y ser reconocidos de manera global por sus grandes avances en diferentes campos.

En el gobierno australiano se impulsaron leyes como: Ley de Libertad de Información modificada en el 2010<sup>4</sup>, lo cual implemento la Publicación de información, teniendo en cuenta la privacidad de esta y garantizó que la información personal fuera tratada de conformidad con la ley<sup>5</sup>.

Por su parte, el gobierno canadiense buscó centrar su estrategia en: Open data, Open Information y Open Dialogue, la cual busca integrar los diferentes sectores del gobierno y ha puesto en marcha un proyecto piloto llamado Portal de Datos

---

<sup>2</sup> MINISTERIO DE LAS TECNOLOGÍAS Y LA INFORMACIÓN MINTIC. *Colombia es cuarto en el mundo y primero en América en datos abiertos*. 2015. p. 9. Disponible en: <http://www.mintic.gov.co/portal/604/w3-article-14303.html>

<sup>3</sup> VIVE DIGITAL MINTIC. *Índices internacionales - Estrategia GEL*. 2015. p. 13. Disponible en: <http://estrategia.gobiernoenlinea.gov.co/623/w3-propertyvalue-7653.html>

<sup>4</sup> Ley de 2010 que modificó la Ley de Libertad de Información (Reforma)

<sup>5</sup> OMPI. *Australia. Freedom of Information Amendment (Reform) Acto 2010*. 2012. p. 7. Disponible en: [http://www.wipo.int/wipolex/es/text.jsp?file\\_id=215466](http://www.wipo.int/wipolex/es/text.jsp?file_id=215466)

Abiertos en 2011 que en la actualidad cuenta con más de 272.000 conjuntos de datos procedentes de 20 departamentos y que ya ha dado lugar a más de 100.000 descargas de conjuntos de datos desde su lanzamiento<sup>6</sup>.

En el área nacional, el gobierno colombiano, como estrategia e-government, busca hacer el país más eficiente, transparente y participativo por medio del uso de las TIC, con la estrategia “Gobierno en Línea”<sup>7</sup>, donde se plantea un marco de referencia para el cumplimiento de esta estrategia y se compone de: Estrategia TI, Gobierno TI, Información, Sistemas de Información, Servicios Tecnológicos y Uso y Apropiación, cada una de estas integran lineamientos, roles, normatividad, indicadores e instrumentos para su adopción<sup>8</sup>. La ley 1712 del 2014 da los estándares para que esta estrategia se cumpla y logre sus objetivos, ya que esta estipula quienes están obligados a realizar la Publicación de información y los principios que estas deben cumplir<sup>9</sup>.

Los datos abiertos en Colombia han tenido gran impulso por la estrategia del gobierno y las leyes que los fomentan, la implementación datos.gov.co y el apoyo del Ministerio de las TIC que ha realizado un gran acompañamiento en el proceso de familiarización con la plataforma, la estrategia del gobierno y como resultado de esta, el país se ha posicionado en el 4° lugar a nivel internacional en la implementación de datos abiertos<sup>10</sup>.

Ante esta situación de tener grandes avances en la implementación de las estrategias gubernamentales y leyes que lo fomenten, esto no comprueba o valida que la información que allí se encuentre cumpla los principios estipulados internacionalmente o los que promueve la Ley 1712 de 2014 de transparencia y del derecho de acceso a la información pública.

## 1.2.2 Formulación del Problema

De acuerdo con lo anterior, esta propuesta busca evaluar los datos abiertos que está implementando el gobierno colombiano, para comprobar si cumplen con los principios internacionales que se han establecidos y las políticas que ha diseñado el mismo gobierno. Los resultados de esta evaluación crearán valor a las entidades públicas y privadas que colaboren en los procesos relacionados a la aplicación de datos abiertos en Colombia. De acuerdo con los planteamientos anteriores, surge la siguiente pregunta de investigación: ¿Cómo se puede comprobar el cumplimiento

<sup>6</sup> IGLESIAS, Carlos. *OGP Member Countries are betting on Open Data*. Web Whisperer. 2012. Disponible en: <http://carlosiglesias.es/blog/2012/07/ogp-member-countries-bet-on-open-data/>

<sup>7</sup> VIVE DIGITAL MINTIC. *Conoce la Estrategia de Gobierno en Línea*. 2014. Disponible en: <http://estrategia.gobiernoenlinea.gov.co/623/w3-propertyvalue-7650.html>

<sup>8</sup> CEPAL. *El desafío hacia el gobierno abierto en la hora de la igualdad*. Documento de proyecto. 2012. p. 124. Disponible en: [http://repositorio.cepal.org/bitstream/handle/11362/3969/S2012004\\_es.pdf;jsessionid=061A7C0F0092CAFDD6CA6FAF6F2463F7?sequence=1](http://repositorio.cepal.org/bitstream/handle/11362/3969/S2012004_es.pdf;jsessionid=061A7C0F0092CAFDD6CA6FAF6F2463F7?sequence=1)

<sup>9</sup> CONGRESO DE LA REPÚBLICA. *Ley 1712 de marzo 6 de 2014*. “Por medio de la cual se crea la Ley de Transparencia y del Derecho de acceso a la información pública nacional y se dictan otras disposiciones”.

<sup>10</sup> MINISTERIO DE LAS TECNOLOGÍAS Y LA INFORMACIÓN MINTIC. Op. Cit. p. 6.

de la Estrategia del Gobierno Nacional sobre los principios y leyes de los datos abiertos en Colombia de acuerdo con los principios establecidos por el Open Government Data?

### **1.3 OBJETIVOS**

#### **1.3.1 Objetivo General**

Implementar un prototipo de software que permita validar el cumplimiento de los principios y leyes de los datos abiertos en Colombia a partir de los principios establecidos por el Open Government Data.

#### **1.3.2 Objetivos Específicos**

- Analizar los principios de los datos abiertos y las leyes colombianas que involucran políticas sobre el acceso a la información.
- Diseñar arquitectura de prototipo de software.
- Implementar el prototipo de software de acuerdo con la arquitectura propuesta.
- Evaluar los datos recolectados de contratación existentes en la plataforma de datos.gov.co para establecer la madurez de este tipo de información.

### **1.4 JUSTIFICACIÓN**

A nivel internacional, la cantidad de datos disponibles cobra unas dimensiones de gran envergadura, lo que ha causado la necesidad de integrar y clasificar dichos datos, permitiendo su publicación de forma tal que posteriormente a ello, se pueda utilizar e intercambiar por otros usuarios. En este contexto de datos masivos, los cuales son publicados por una empresa y luego utilizados por otra o cualquier ciudadano, resulta ser importante puesto que adquiere gran valor social y económico a partir de la reutilización de dichos datos.

De hecho, de eso se trata el “open data” o datos abiertos, que se fundamenta en la publicación de datos organizados para potenciar su consumo futuro. Estos datos se publican por medio de portales de datos abiertos, como es el caso de datos.gov.co, los cuales sirven para tenerlos disponibles para los consumidores de datos. “Cuando estos datos abiertos provienen del sector público, se emplea el término datos públicos (“public data”), concepto que está relacionado directamente con la

información de carácter gubernamental generada por instituciones públicas, puesta a disposición de la ciudadanía con el afán de fortalecer principalmente la transparencia y la rendición de cuentas”<sup>11</sup>. En ese sentido, y de acuerdo con la calidad del portal de datos abiertos en el que se encuentren publicados, se puede garantizar la madurez del mismo portal, permitiendo la reutilización de estos, para conseguir valor agregado a quien los necesita, los use o los descargue. De esta forma, es clara la importancia de definir mecanismos adecuados para evaluar la madurez de un portal de datos abiertos y poder compararlos con los lineamientos y principios de uno de tantos que se encuentran a nivel global, como es en este caso el Open Government Data.

## **1.5 ALCANCES Y LIMITACIONES**

### **1.5.1 Alcances**

Dentro de los alcances de la investigación, se encuentran:

- La realización de un prototipo que sea capaz de evaluar los principios seleccionados.
- El uso de la información para suplir el prototipo, no se encuentra ligada a licencias de propietarias, lo cual beneficia al proyecto con un constante flujo de datos.

### **1.5.2 Limitaciones**

Las limitaciones encontradas, son:

- El prototipo por desarrollar busca analizar y evaluar el entorno colombiano, primariamente, pero respecto a la estructuración de este, será escalable a diferentes entornos mundiales, en donde se tengan en cuenta los principios internacionales de los datos abierto.
- Este proyecto será realizado en el transcurso de cinco meses, lo que comprende el periodo académico de la Universidad Católica de Colombia, en donde se entregará para su evaluación y aceptación.

## **1.6 MARCO TEÓRICO**

---

<sup>11</sup> OVIEDO, Edgar. *Modelo de madurez para portales de datos abiertos e incorporación a la norma técnica nacional de Costa Rica*. Universidad de Alicante. Tesis Doctoral. Alicante – España. 2016. p. 7.

A continuación, se elabora el marco teórico el cual da cuenta de, los principales antecedentes de estudios similares, además, muestra los autores más destacados que ilustran el tema, los cuales sirvieron para demarcar mejor la investigación.

### 1.6.1 Datos Abiertos (Open Data)

Los datos abiertos son aquellos datos que son puestos a disposición del público y deben cumplir con las siguientes características<sup>12</sup>:

- Completos
- Primarios
- Oportunos
- Accesibles
- Susceptibles de procesamiento computacional automatizado
- No-discriminatorios
- En formatos abiertos
- Libres de licencia

Estos son conocidos como los 8 principios de gobierno de datos abiertos de Larry Lessig<sup>13</sup>, el cual estipula que dan base a la aplicación de los datos abiertos en los gobiernos, ya que estimula la ejecución de los ideales del pensamiento “Open Government”, por su lado, la Sunlight Foundation<sup>14</sup> reconoce dos condiciones adicionales, la permanencia de los datos (disponibilidad a largo plazo) y costo accesible o nulo<sup>15</sup>.

La Carta Internacional De Datos Abiertos establece que la aplicación de datos abiertos en el gobierno, benefician aspectos económicos y sociales para la sociedad y se puede lograr por medio del apoyo a políticas públicas, el apoyo intersectorial, la muestra al público del uso de los recursos, vigilar el impacto que esto tendría y la mejor toma de decisiones, a la vez definen 6 principios que son necesarios para el acceso a los datos, su publicación y su uso:

- Abiertos por defectos
- Oportunos y Exhaustivos
- Accesibles y Utilizables

---

<sup>12</sup> OPEN GOV DATA. *Request for Comments*. 2007 [En línea] Disponible en: [http://wiki.opengovdata.org/index.php/Main\\_Page](http://wiki.opengovdata.org/index.php/Main_Page)

<sup>13</sup> LESSIG, Larry. *8 Principles of Open Government Data* Disponible en: [https://public.resource.org/8\\_principles.html](https://public.resource.org/8_principles.html)

<sup>14</sup> SUNLIGHT FOUNDATION. *Ten Principles for Opening Up Government Information*. Agosto 11 de 2010. Disponible en: <https://sunlightfoundation.com/wp-content/uploads/2016/11/Ten-Principles-for-Opening-Up-Government-Data.pdf>

<sup>15</sup> LAMBOT, S., SLOB, E., VAN BOSCH, I., STOCKBROECKX, D., & VANCOLOOSTER, M. *Modeling of ground penetrating radar for accurate-characterization of subsurface electric properties*. *IEEE Transactions on Geoscience and Remote Sensing*. 42(11), pp. 2555-2568. Disponible en: <https://doi.org/10.1109/TGRS.2004.834800>

- Comparables e Interoperables
- Para mejorar la Gobernanza y la participación Ciudadana
- Para el Desarrollo incluyente y la Innovación<sup>16</sup>.

La calificación de datos abiertos normalmente era valorada por el sistema de 5 estrellas de Tim Berners-Lee, donde los datos se calificaban con una estrella si simplemente estaban disponibles en la WWW, y las calificaciones aumentaban conforme eran legibles por sistemas automatizados, se encontraban en formatos aptos para la Web semántica y finalmente en forma de “datos ligados”<sup>17</sup>.

Harlan Yu y David Robinson<sup>18</sup> señalan que los datos abiertos pueden ser clasificados no en un solo eje sino en dos. El primero es similar al generalmente usado y en la nomenclatura de este trabajo va de “inerte” (información ya registrada que solamente puede ser leída) a “adaptable”, que puede ser explotada en diversos formatos, usos y combinaciones, y para ello debe estar estructurada, ser susceptible de lectura automatizada, y disponible para su descarga en grandes volúmenes. El segundo eje se refiere a la utilidad e impacto; va desde la entrega de servicios hasta la rendición de cuentas.

### 1.6.2 Condiciones de los Datos Abiertos en América Latina

La Comisión Económica para América Latina y el Caribe CEPAL, como organismo de la ONU era el sitio donde los acuerdos logrados buscaban enfocar a los gobiernos de los diferentes países, para enfatizar en su conexión, para que ofrecieran información, transacciones y servicios de gobierno en línea, para mejorar el énfasis y calidad del sector público y fuera intensiva la utilización de las Tecnologías de Información y Comunicación (TIC) para su modernización y el fortalecimiento de los servicios en línea para la ciudadanía, además de entregar mayor accesibilidad en vías de agilizar, optimizar, flexibilizar, transparentar y abaratar procesos y/o actividades del sistema público<sup>19</sup>. En ese sentido, es interesante visualizar lo que la CEPAL, indica al respecto:

Las TIC, por tanto son las herramientas que están produciendo los cambios más radicales en la gestión pública ya que permiten el desarrollo de aplicaciones cada vez más complejas, necesariamente apoyadas por arquitecturas dedicadas, especialmente diseñadas para trabajar de la manera más óptima, integrando sistemas, utilizando las mejores herramientas de gestión y desarrollando modelos adecuados a las necesidades de Gobierno, creando plataformas compatibles que

<sup>16</sup> OPEN DATA CHARTER. *International Open Data Charter*. 2015. Disponible en: <http://opendatacharter.net/principles-es/>

<sup>17</sup> TIM, B. *Linked Data – Design Issues*. 2009. Disponible en: <https://www.w3.org/DesignIssues/LinkedData.html>

<sup>18</sup> YU, Harlan., & ROBINSON, David. *The New Ambiguity of “Open Government”*. SSRN Electronic Journal. 2012. Disponible en: <https://www.uclalawreview.org/pdf/discourse/59-11.pdf>

<sup>19</sup> NASER, A., RAMÍREZ, A. *Plan de gobierno abierto*. 2014. Disponible en: [http://www.cepal.org/ilpes/noticias/paginas/3/54303/Plan\\_de\\_Gobierno\\_Abierto.pdf](http://www.cepal.org/ilpes/noticias/paginas/3/54303/Plan_de_Gobierno_Abierto.pdf)

resuelven temas como la interoperabilidad, compatibilidad, acceso, seguridad, entre otras<sup>20</sup>.

Del mismo modo, la utilización de las TIC's en la gestión pública constituye pilares fundamentales para la modernización y eficacia del Estado, ayudan al control interno y externo aportando transparencia al sector público, disminuye costos al compartir recursos, ayudan a la descentralización acercando el gobierno a los ciudadanos y facilita la participación ciudadana en los procesos de tomas de decisiones<sup>21</sup>.

Las consecuencias e impactos en la atención a los ciudadanos se expresan en al menos lo siguientes aspectos:

- Los usuarios pueden recibir atención sin restricción horaria y no importando en qué lugar geográfico se encuentren.
- Los usuarios tienen acceso a información pública en forma simple, oportuna, clara y transparente.
- Los usuarios pueden resolver sus problemas a través de un contacto único con el Estado; aunque se trate de requerimientos que involucren a más de una institución.
- Los usuarios no están obligados a presentar ningún documento o certificar información que se encuentre disponible en formato electrónico por alguna institución pública.
- Los usuarios pueden hacer transacciones financieras en forma electrónica, (por ejemplo, pagar un certificado de nacimiento, o recibir el pago por un servicio prestado un organismo público si se es un proveedor del estado<sup>22</sup>.

### **1.6.3 Comportamiento de los datos abiertos a nivel internacional**

En esta sección se visualizará en el contexto internacional la implementación de los datos abiertos, y las iniciativas que se han impulsado.

---

<sup>20</sup> CEPAL. *El gobierno electrónico en la gestión pública*. CEPAL. 2011. p. 5. Disponible en: [http://www.cepal.org/ilpes/noticias/paginas/3/54303/EI\\_Gobierno\\_electronico\\_en\\_la\\_gestion\\_Publica.pdf](http://www.cepal.org/ilpes/noticias/paginas/3/54303/EI_Gobierno_electronico_en_la_gestion_Publica.pdf)

<sup>21</sup> NASER, A., RAMÍREZ, A., Op. Cit., p. 9

<sup>22</sup> Ibid, p. 12

1.6.3.1 Australia. Este país por medio de sus iniciativas ha logrado la interoperabilidad de sus diferentes entidades públicas y ha obtenido un compromiso propio de realizar la iniciativa de datos abiertos como algo propio del gobierno dando así la posibilidad de ser más transparente para el pueblo y obtener mayor credibilidad<sup>23</sup>.

Australia está dividida en seis estados, dos territorios continentales y otros territorios menores. Los estados son Nueva Gales del Sur (NSW), Queensland (QLD), Australia Meridional (SA), Tasmania (TAS), Victoria (VIC) y Australia Occidental (WA). Los dos territorios continentales son el Territorio del Norte (NT) y el Territorio de la Capital Australiana (ACT). Hay 3 niveles de gobierno: federal, estado y local. Cada estado es independiente (constitución, gobierno, parlamento y leyes propias), los territorios funcionan similar a los estados con la excepción de que el Commonwealth Parliament puede reescribir la legislación de sus parlamentos.

### **Iniciativas Gubernamentales**

Los compromisos del Gobierno australiano a la acción sobre la open datan han sido:

- La reforma a la Ley de FOI y el establecimiento de la Oficina australiana de Information Commissioner;

El sitio oficial de su catálogo de datos es data.gov.au; en este sitio se encuentra estructurada la estrategia del gobierno australiano para centralizar los datos abiertos y estandarizarlos<sup>24</sup>.

1.6.3.2 Canadá. Este país por medio de su implementación de tres pilares fundamentales para la fundamentación de datos abiertos<sup>25</sup> y como caso de transparencia se puede observar, como por medio de los datos abiertos se descubrió un fraude de \$3.2 billones de dólares que se realizaba en este país<sup>26</sup>.

Canadá se compone de 10 estados y 3 territorios. Hay 3 niveles de gobierno: federal, provincia y local. Cada provincia es independiente (constitución, gobierno, parlamento y leyes propias), los territorios no tienen jurisdicción propia y sus

---

<sup>23</sup> OFFICE OF THE AUSTRALIAN INFORMATION. *Protecting information rights – advancing information policy*. 2012. Disponible en: <https://www.oaic.gov.au/resources/about-us/corporate-information/key-documents/languages/Spanish/OAIC-brochure-Spanish.pdf>

<sup>24</sup> Sitio oficial de Australia para datos abiertos *About – data.gov.au* - 2014

<sup>25</sup> ORGANIZACIÓN PARA LA COOPERACIÓN Y EL DESARROLLO ECONÓMICOS – OCDE. *Un gobierno abierto: fomentar el diálogo con la sociedad civil*. 2003. Disponible en: <http://www.oecd-ilibrary.org/docserver/download/9789264019959-sum-es.pdf?expires=1508516822&id=id&accname=guest&checksum=9612B8688511DEC0C3D669E6FE107848>

<sup>26</sup> EAVES, D. *Case Study: How Open data saved Canada \$3.2 Billion*. 2010. Disponible en: <https://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse>

poderes dependen de lo que les delegue el gobierno federal. El gobierno local se ejerce en la forma de condados (countries, distrits, etc.) y municipalidades (cities, towns, parishes, etc.).

### Iniciativa Expansion of Open Government

En marzo de 2011, el Gobierno de Canadá anunció su compromiso continuo hacia la transparencia en la declaración statement about the expansion of open government, el cual establece que la expansión del gobierno abierto se perseguirá a través de 3 vertientes principales:

- Open Data: Ofrecer datos gubernamentales de forma más útil para permitir a los ciudadanos, el sector privado y las ONGs apalancarlos en formas innovadoras y de valor agregado.
- Open Information: Liberar información proactivamente y a medida que se va generando, así será más fácil de encontrar y acceder para los canadienses.
- Open Dialogue: Otorgar a los canadienses una voz más fuerte en las prioridades y políticas del gobierno, y expandir su compromiso a través de tecnologías Web 2.0.

Proyecto Piloto - Government of Canada Open Data Portal (data.gc.ca).

El gobierno canadiense produce y adquiere datos en áreas como salud, medio ambiente, agricultura y recursos naturales. El objetivo del Open Data Portal es crear oportunidades socioeconómicas que promuevan la participación informada del público al expandir el acceso a datos del gobierno federal.

Según esta iniciativa, Open Data se define como: la práctica que hace que los datos procesables por computador (machine-readable) estén a libre disposición, por medio de portales, metadatos, y herramientas de búsqueda para ser reutilizados por gobiernos, ciudadanos, organizaciones de voluntariado, y el sector privado en formas nuevas y no anticipadas<sup>27</sup>.

#### 1.6.4 Actividad Nacional en Datos Abiertos

La Constitución de 1991 establece a Colombia como una República Unitaria que se divide administrativa y políticamente en 33 divisiones: 32 departamentos, los cuales son gobernados desde sus respectivas ciudades capitales, y Bogotá. Los departamentos forman regiones geográficas, culturales y económicas. En Colombia los recursos pasan de la nación al departamento y de este al municipio; como distrito capital, Bogotá es el único municipio que recibe directamente de la nación<sup>28</sup>.

---

<sup>27</sup> LATHROP, D., & RUMA, L. *Open Government*. 2003

<sup>28</sup> WIKIPEDIA. *Departamentos de Colombia*. 2017. Disponible en: [https://es.wikipedia.org/wiki/Departamentos\\_de\\_Colombia](https://es.wikipedia.org/wiki/Departamentos_de_Colombia)

## Iniciativas

Estrategia Gobierno en Línea o GEL: Promover el uso de los Datos Abiertos, para la transparencia en los procesos del gobierno, fortalecer sectores emergentes y que usen las TIC's para dar mayores resultados y activar la colaboración ciudadana<sup>29</sup>.

### Ley 1712 de 2014

El Gobierno colombiano por medio de la ley 1712 del 2014, la ley de transparencia y del derecho de acceso a la información pública, en el Artículo 3° establece los principios que deben cumplirse para la transparencia y acceso a la información:

- Principio de transparencia
- Principio de buena fe
- Principio de facilitación
- Principio de no discriminación
- Principio de gratuidad
- Principio de celeridad
- Principio de eficacia
- Principio de la calidad de la información
- Principio de la divulgación proactiva de la información
- Principio de responsabilidad en el uso de la información<sup>30</sup>

La Ley 1712 estipula quienes están obligados a la disposición de la ley: Toda entidad pública, incluyendo las pertenecientes a todas las ramas del poder público, en todos los niveles de la estructura estatal, central o descentralizada por servicios o territorialmente, en los órdenes nacional, departamental, municipal y distrital; los órganos, organismos y entidades estatales independientes o autónomos y de control; las personas naturales y jurídicas, públicas o privadas, que presten función pública, los partidos o movimientos políticos y los grupos significativos de ciudadanos, las entidades que administren instituciones parafiscales y tienen la obligación de responder de buena fe, de manera veraz, oportuna y adecuada a las solicitudes de información pública, los sujetos obligados también deben publicar proactivamente la información pública y además, implementar procedimientos archivísticos que garanticen la disponibilidad en el tiempo de documentos electrónicos auténticos<sup>31</sup>.

## 1.7 MARCO CONCEPTUAL

Teniendo en cuenta los objetivos del presente trabajo, en el cual se busca medir el cumplimiento de los principios y leyes de los datos abiertos en Colombia, es necesario conocer algunos conceptos relacionados con la investigación, incluidos

---

<sup>29</sup> VIVE DIGITAL – MINTIC Op. Cit., p. 3

<sup>30</sup> CONGRESO DE LA REPÚBLICA. *Ley 1712 de 2014*. Op. Cit., Artículo 3.

<sup>31</sup> *Ibid.*, Artículo 5.

la terminología, las instituciones o fundaciones, los procesos o tecnologías que se utilizan.

- **Dato:** Representación de una cierta entidad del mundo real en alguna forma de símbolo<sup>32</sup>
- **Información:** Son todos aquellos datos que contribuyen a reducir la incertidumbre o riesgo de un evento; o bien, faciliten la comunicación o adquisición de conocimientos que permitan ampliar o precisar los que se posean respecto de una materia determinada<sup>33</sup>.
- **Open Government (O-gov) Gobierno Abierto:** Define la forma de relacionarse entre la Administración Pública y los ciudadanos, que se caracteriza por el establecimiento de canales de comunicación y contacto directo entre ellos<sup>34</sup>.
- **Open Data Datos abiertos:** Los datos abiertos son datos que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, y que se encuentran sujetos, cuando más, al requerimiento de atribución y de compartirse de la misma manera en que aparecen<sup>35</sup>.
- **Big Data:** Se refiere a una combinación de un enfoque para informar la toma de decisiones con visión analítica derivada de los datos, y un conjunto de tecnologías que permiten que permite que la percepción que se derivan de vista económico, a veces diversas fuentes muy grandes, de los datos. La digitalización de comercio y comunicaciones, y la aparición y crecimiento de los medios de comunicación social son algunas de las tendencias que han creado la oportunidad de utilizar a gran escala, los datos de grano fino para entender los sistemas, el comportamiento y el comercio; mientras que la innovación en la tecnología hace que sea económicamente viable utilizar esa información para tomar decisiones informadas y mejorar los resultados<sup>36</sup>.
- **WebService:** Es un servicio ofrecido por una aplicación que expone su lógica a clientes de cualquier plataforma mediante una interfaz accesible a través de la red utilizando tecnologías (protocolos) estándar de Internet<sup>37</sup>.
- **Business Intelligence (BI):** La Inteligencia de Negocios BI Business Intelligence es una herramienta bajo la cual diferentes tipos de organizaciones pueden soportar la toma de decisiones basadas en información precisa y oportuna; garantizando la generación del conocimiento necesario que permita escoger la alternativa que sea más conveniente para el éxito<sup>38</sup>.

---

<sup>32</sup> ALFA-REDI. *Las Bases de Datos*. Lima – Perú. s.f. p. 1. Disponible en: [http://www.alfaredi.org/sites/default/files/articles/files/ossa\\_0.pdf](http://www.alfaredi.org/sites/default/files/articles/files/ossa_0.pdf)

<sup>33</sup> Ibid., p. 2

<sup>34</sup> LLINARES, Javier. *Open government: La idea*. 2007. Disponible en: <http://www.javierllinares.es/2007/12/29/open-government-la-idea/>

<sup>35</sup> ARGIS FOR DESKTOP. *¿Qué son los datos ráster?*. 2016. Disponible en: <http://desktop.arcgis.com/es/arcmap/10.3/manage-data/raster-and-images/what-is-raster-data.htm>

<sup>36</sup> AKRED, JOHN. *¿Qué son los grandes datos?*. 2014. Disponible: <https://datascience.berkeley.edu/what-is-big-data/>

<sup>37</sup> BESTEIRO, M. RODRIGUEZ, M. *Web Services* s.f. Disponible: <http://www.ehu.eus/mrodriguez/archivos/csharp/pdf/ServiciosWeb/WebServices.pdf>

<sup>38</sup> ROSADO, Alveiro, RICO, Dewar. *Inteligencia de Negocios: Estado del Arte*. Pereira. Revista Scientia et Technica. Universidad Tecnológica de Pereira. 2010. (44) pp. 321-326

- **RapidMiner:** Programa informático que permite la creación de procesos utilizados para el análisis y minería de datos, en la cual se permite el encadenamiento de operadores para así poder obtener una creación rápida de procesos de minería de datos.

## **1.8 METODOLOGÍA**

Se realizará un análisis de la información por medio de tablas descriptivas, las cuales dan un contexto general de cómo se encuentra el panorama internacional respecto a los principios que se han establecido y cuáles de esos principios aplica el gobierno colombiano en la implementación de su política de datos abiertos

### **1.8.1 Tipo de Estudio**

Los objetivos establecidos para el desarrollo de esta investigación concluyen con un prototipo, en el cual se analiza y se realiza un exhaustivo proceso de evaluación sobre los principios y políticas de datos abiertos que se están aplicando en Colombia.

### **1.8.2 Fuentes de información**

Los recursos teóricos de esta investigación son aquellos que nos permiten tener buenas bases investigativas para la ejecución del documento final, por tal razón el uso de Base de Datos indexadas, aplicados para la búsqueda de Journals y Papers, unas de las usadas son:

- CiteSeerX
- DBPL: Computer Science Bibliography

De igual forma, motores de búsqueda tales como Google Scholar, IEEE Explore; en donde con patrones de búsqueda es posible filtrar documentación respectiva sobre Open Data y Open Government.

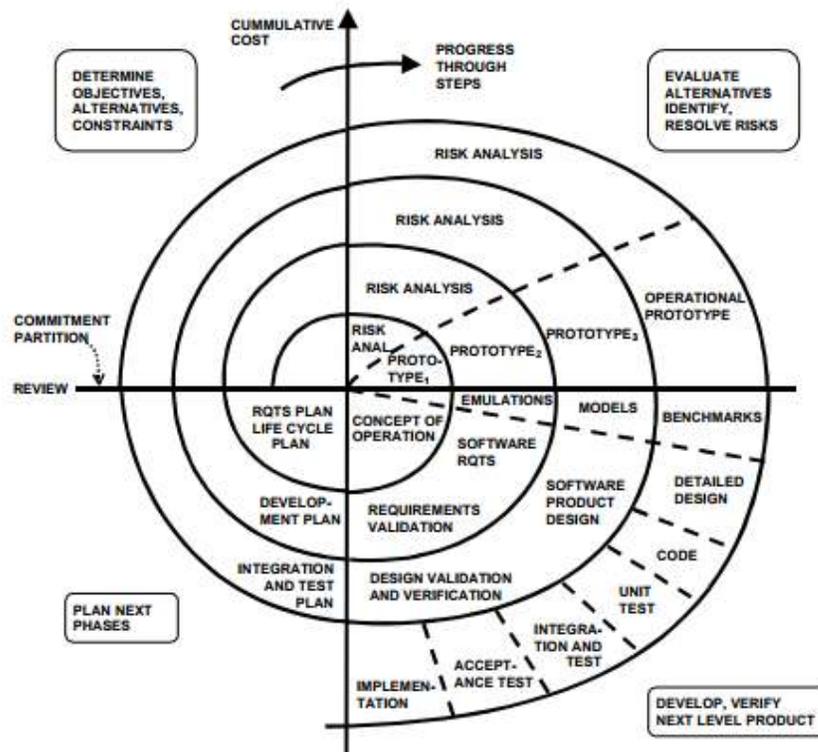
### **1.8.3 Diseño Metodológico**

La metodología que será usada para el desarrollo del prototipo se basará en el modelo de desarrollo en espiral, en el cual se reducen el riesgo de manera intensiva y se pueden mitigar rápidamente si alguno de ellos pasase, cada espiral o iteración del modelo, permiten refinar y madurar el prototipo a través de este modelo.

El modelo de desarrollo en espiral es un generador de modelo de procesos basado en riesgos. Se caracteriza por tener un enfoque cíclico para un crecimiento incremental del grado de decisión e implementación del sistema, mientras se reduce el riesgo de este<sup>39</sup>.

El modelo en espiral cuenta con cuatro fases: Planificación, Análisis de Riesgos, Ingeniería y Evaluación. Durante el ciclo de vida del proyecto se realizan iteraciones, llamadas espirales<sup>40</sup>

Figura 1. Modelo en espiral



Fuente: BOEHM, Barry. p. 2

En la espiral se comienza por la fase de planificación, se reúnen los requerimientos y se miden los riesgos. Cada espiral siguiente se construye sobre la espiral base. Los requerimientos son reunidos durante la fase de planeación. En la fase de análisis de riesgo, se lleva a cabo un proceso para identificar los riesgos y sus posibles soluciones y al final de esta fase se crea un prototipo.

El desarrollo es llevado a cabo en la fase de ingeniería y al final de esta, se llevan a cabo las pruebas en la fase de evaluación, necesarias para determinar si el

<sup>39</sup> BOEHM, Barry. *Spiral Development: Experience, Principles, and Refinements: Spiral Development Workshop*. Pittsburgh. 2000. p. 6. Disponible en: <https://www.sei.cmu.edu/reports/00sr008.pdf>

<sup>40</sup> BOEHM, Barry. *Ibid*, p. 2

prototipo cumple con los objetivos o si debe ser sometido a una iteración de ciclo de vida.

1.8.3.1 Matriz Foda. En la siguiente matriz se muestran las Fuerzas (Factores Críticos Positivos), Oportunidades (Aspectos positivos que podemos aprovechar), Debilidades (Factores Críticos Negativos) y Amenazas (Aspectos Negativos externos).

Tabla 1. Matriz FODA

	<b>POSITIVOS</b>	<b>NEGATIVOS</b>
<b>INTERNO</b>	<p><b>Fortalezas:</b></p> <ul style="list-style-type: none"> <li>Iteraciones o espirales del ciclo de vida</li> <li>Análisis de riesgo constante y sus posibles soluciones</li> <li>Maduración del proyecto en cada espiral</li> <li>Software producido desde el primer espiral</li> </ul>	<p><b>Debilidades:</b></p> <ul style="list-style-type: none"> <li>Análisis de riesgos requiere gran experiencia</li> <li>Tiempos de espirales largos si se realiza un mal análisis</li> </ul>
<b>EXTERNO</b>	<p><b>Oportunidades:</b></p> <ul style="list-style-type: none"> <li>Tiempos de espiral cortos</li> <li>Análisis de riesgos analizado en paralelo y con un inicio de base pequeña</li> </ul>	<p><b>Amenazas:</b></p> <ul style="list-style-type: none"> <li>Proyectos con maduración extensa si no se planifican bien</li> </ul>

#### 1.8.4 Estrategias de Comunicación

Como estrategias de comunicación, este proyecto socializará el trabajo de grado por medio de una exposición clara sobre el proyecto, a su vez, se requiere elaborar un artículo científico que posteriormente podrá ser enviado a un congreso ya sea nacional o internacional. Por último, la memoria que se genera mediante este proyecto también es vista como estrategia de comunicación.

#### 1.8.5 Recursos

Como recursos se considera todo el equipo y personal que se requiere para el desarrollo del proyecto, de acuerdo con las metas establecidas y visualización del proyecto.

- Humanos. Dos personas que son los encargados de llevar a cabo el proyecto en su totalidad, diseñando el software mencionado y adaptando las plataformas requeridas.
- Físicos. Se contará con las instalaciones de la universidad y bibliotecas públicas para trabajar las tareas designadas, y en los domicilios de los autores.
- Técnicos. Se necesita un equipo portátil por persona para que cada uno desarrolle por temas lo designado a cada uno.
- Financieros. Para la ejecución del proyecto se financiarán el valor de los transportes requeridos para las reuniones estipuladas que requiera el equipo investigador, para el levantamiento de información, aclaración de dudas y seguimiento del proyecto. Adicionalmente, se requiere financiar el valor de fotocopias o adquisición de libros para la documentación del proyecto.
- Los costos estimados del proyecto y que serán financiados por parte de los dos integrantes del proyecto, son los siguientes:

Tabla 2. Presupuesto

CONCEPTO	COSTOS UNITARIOS	COSTO TOTAL PROYECTO	% INCIDENCIA
Transportes	\$80.000	\$320.000	12%
Fotocopias	\$4.000	\$48.000	1%
Equipos de computo	\$0	\$0	0%
Software	\$0	\$0	0%
Mano de obra (recurso y tiempo)	\$750.000	\$1.500.000	86%
Diseños	\$100.000	\$100.000	1%
<b>TOTAL</b>	<b>\$934.000</b>	<b>\$1.968.000</b>	<b>100%</b>

Fuente: Elaboración propia

### 1.8.6 Productos a entregar

Se entregará un prototipo de software que dé cuenta del cumplimiento de los principios y leyes de datos abiertos establecidos por el Open Government Data para evaluarlos con los lineamientos expuestos por Colombia.

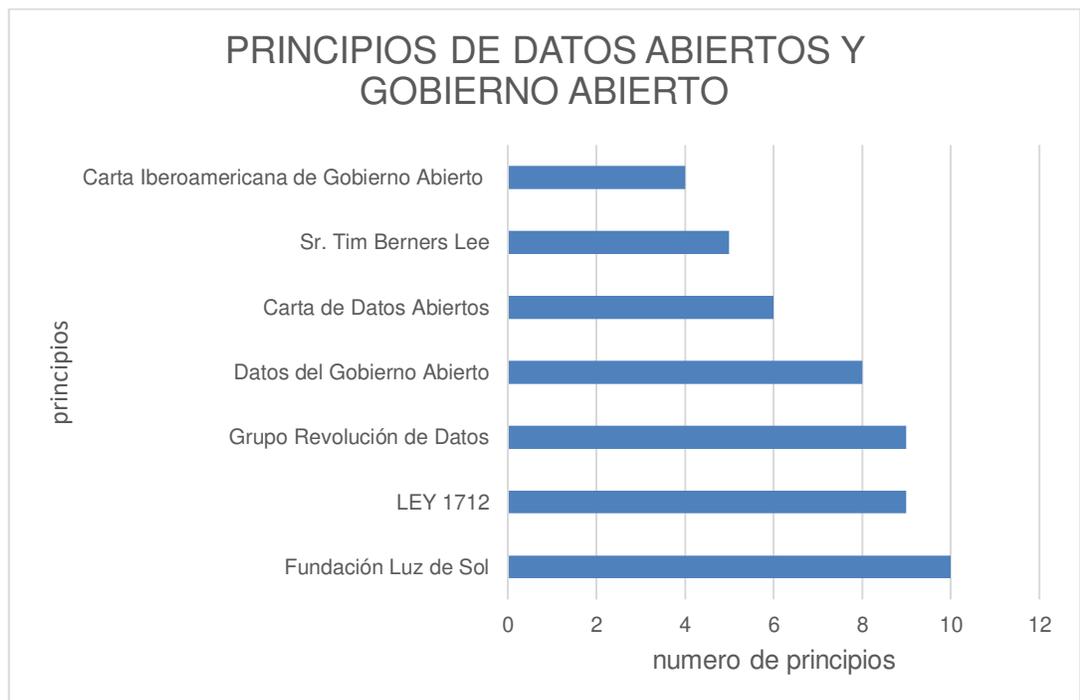
## 2. ANÁLISIS DE LOS PRINCIPIOS INTERNACIONALES Y LEYES COLOMBIANAS REFERENTE A LOS DATOS ABIERTOS

La recolección de información se realizó de forma controlada, realizando filtración de literatura de journals, papers, o sitios web especializados en la temática de Open Data y Open Government, para poder así consolidar toda aquella información de manera visible y ser interpretada de forma ágil y permitir realizar un análisis eficaz de la información que sea recolectada.

### 2.1 INFORMACIÓN RECOPIADA

La información recopilada de las diferentes fuentes, referente a los principios de open data se puede visualizar en la Figura 2.

Figura 2. Principios de datos abiertos y gobierno abierto



Fuente: Elaboración propia

En la figura 2 se observa el ponderado de los principios recolectados de acuerdo con el análisis efectuado anteriormente.

### 2.1.1 Carta de Datos Abiertos

El origen de esta carta de datos abierto se da en la cumbre G8 realizada en Irlanda de Norte en el año 2013, en donde se estableció como acuerdo unilateral los líderes del G8 firmaron la Carta Abierta de datos del G8, que describe un conjunto de principios básicos de datos abiertos de cinco.

Esta carta estipula 6 principios que son base para el acceso, liberación y uso de los datos abiertos<sup>41</sup>:

- **Abiertos por Defecto.** Deben ser para el uso público, el cual permitirá el acceder a ellos y poder realizar un procesamiento debido de la información, todo esto apoyado por políticas gubernamentales que impulsen el desarrollo, gestión y supervisión; así como la creación de cultura de apertura lo cual facilite la implementación de los datos abiertos en los gobiernos.
- **Oportunos y Completos.** Se asignarán recursos tanto financiero, técnico y humano para la implementación y liberación de los datos abiertos, todo eso por medio de listas públicas, que sean completas y de libre acceso para su uso y conocimiento.
- **Accesibles y Utilizables.** Toda información liberada, deberá ser centralizada, el cual facilite el uso del sitio, además que sea visible y accesible en un solo lugar, deberán ser gratuitos y encontrarse bajo una licencia abierta y sin restricciones, así como asegurar que estos puedan accederse por diferentes sectores de la población para así facilitar su divulgación y conocimiento de estos.
- **Comparables e Interoperables.** Estos deben aparte de estar centralizados, estos deben ser coherentes, interoperables y que faciliten el uso de ellos tanto de forma individual por una persona o por medio de máquinas para poder darle un mayor uso a las políticas abiertas que se han implantado para impulsar la cultura de apertura del gobierno.
- **Para mejorar la Gobernabilidad y Participación Ciudadana.** Son todas aquellas medidas que realiza el gobierno para impulsar el consumo de datos abiertos, de igual forma que estos impulsen la transparencia de la gobernabilidad para un mayor control y conocimiento por parte de la población, de tal forma que pueda ejercer un control sobre el consumo de los datos abiertos y un mayor control de la información que producen sus entidades.
- **Para el Desarrollo Inclusivo e Innovación.** Crear e impulsar proyectos que implemente el uso de los datos abiertos, no solo para un mayor control del gobierno si no que esta fuente de datos sea un banco de datos que impulsen a

---

<sup>41</sup> OPEN DATA CHARTER. *Principios*. 2015 disponible: <https://opendatacharter.net/principles-es/>

la sociedad a generar mayores formas de investigación independiente a partir de la información recopilada por el gobierno y puesta a disposición de la sociedad para su uso

### **2.1.2 Sunlight Foundation**

Organización de los Estados Unidos de América, no partidista y sin fines de lucro el cual busca por medio de la tecnología, los datos abiertos hacer del gobierno y la política más responsable y transparente para todos<sup>42</sup>. Por tal motivo, propone los siguientes principios para poder impulsar los datos abiertos en el gobierno:

- **Integridad.** Los conjuntos de datos publicados por el gobierno deben ser lo más completos posible. Toda la información en bruto a partir de un conjunto de datos debe ser liberado al público, salvo en la medida necesaria para cumplir con la ley federal con respecto a la divulgación de información personal identificable. Si se hace, permitirá a los usuarios a entender el alcance de la información disponible y examinar cada elemento de datos con el mayor nivel de detalle posible
- **Primacía.** Los conjuntos de datos publicados por el gobierno deben ser datos de fuente primaria. Esto incluye la información original recogida por el gobierno, la difusión pública permitirá a los usuarios verificar que la información se recogió correctamente y registra con precisión.
- **Oportunos.** Los conjuntos de datos publicados por el gobierno deben estar a disposición del público en el momento oportuno. Siempre que sea posible, la información recogida por el gobierno debe ser puesto en libertad tan pronto como se recolecta y se recoge. Se debe dar prioridad a los datos cuya utilidad es sensible al tiempo. Actualizaciones de la información en tiempo real serían maximizar la utilidad que el público puede obtener de esta información.
- **Fácil acceso Físico y Electrónico.**
- Los datos que sean publicados deben ser tan accesibles como sea posible, teniendo en cuenta que entre más fácil sea su acceso estos pueden incrementar ya sea por medios físicos o electrónicos, de ahí que se deba facilitar la accesibilidad a los datos en cada uno de los casos.
- **Procesable por máquina.** Los datos abiertos, que sean entregados, podrán ser procesados por máquinas y algoritmos, todo esto para tener un máximo uso de la información que es entregada. Por tal razón se deben recibir archivos que

---

<sup>42</sup> LAMBOT, S. et. al. 2004. Op. Cit. p. 17

sean en formato abierto y garantizar que se mantengan por el tiempo de la misma forma, así se garantiza que sean procesables por máquinas y algoritmos.

- No Discriminatorio. Los datos que se alojen en el repositorio deberán ser accesibles por todo tipo de persona o de quien lo solicite; no limitar la información por tipo de usuario, segmentos o ente que necesite usar la información que allí se encuentre alojada.
- Uso de estándares de propiedad común. Este principio nos indica el uso de estándares propietarios en vez del uso de estándares abiertos limitaran el uso y la adopción de la cultura de datos abiertos, por consiguiente, se deben establecer limitaciones para el uso de estos estándares propietarios.
- Licenciamiento. La imposición de "Términos de Servicio", en la información que allí es publicada no permite la máxima apertura. De ahí que debe ser etiquetado como información pública y no debe contar con ningún tipo de restricción de distribución.
- Permanencia. Este principio nos indica, que toda información Publicada debe ser perdurable en el tiempo y es un deber mantener versiones, actualizaciones de la información. Así como garantizarla en el transcurso del tiempo y tener un seguimiento apropiado de la información.
- Costos de Uso. El manejo de los costos debe realizarse por medio del gobierno, para así mitigar que los gastos de mantenimiento de la infraestructura sean mínimos, para que así se pueda generar una mayor manutención en la información abierta, que se encuentra disponible.

### **2.1.3 Datos de Gobierno Abierto**

- Grupo de trabajo, el cual analiza y complementa a la "Open Knowledge Foundation" el cual en sus análisis ha definido los principios que rigen los datos abiertos para el gobierno<sup>43</sup>.
- Completos. Todos los datos se ponen a disposición del público. Los datos que se publican son datos que no están sujetos a limitaciones de privacidad, seguridad o privilegio válidos.
- Primarios. Los datos se recogieron desde la fuente, con el más alto nivel posible de granularidad, no en formas agregadas o modificadas.

---

<sup>43</sup> OPEN GOV DATA. Op. Cit. p. 1.

- Oportuno. Los datos se recogieron como en la fuente, con el más alto nivel posible de granularidad, no en formas agregadas o modificadas.
- Accesible. Los datos están disponibles para la gama más amplia de usuarios para la más amplia gama de propósitos.
- Procesable por máquina. Los datos se estructuran razonablemente para permitir el procesamiento automatizado.
- No discriminatorio. Los datos están disponibles para cualquier persona, sin necesidad de registro.
- Licencia abierta. Los datos no están sujetos a ningún derecho de autor, patentes, marcas o regulación secreto comercial, privacidad razonable, se puede permitir restricciones de seguridad y privilegios.

#### **2.1.4 Sr. Tim Berners Lee**

El señor Tim Berners Lee inventó la World Wide Web en 1989, además dirige el Consorcio World Wide (W3C) del foro para el desarrollo técnico de la Web. Fundó la Fundación Web cuya misión es que la WWW sirva a la humanidad, y fue cofundador del Instituto Open Data en Londres. Su grupo de investigación CSAIL planea volver a descentralizar la Web. Tim pasa mucho tiempo luchando por los derechos como la privacidad, la libertad y la apertura de la Web, en lo cual propone un sistema de 5 estrellas para poder clasificar la información abierta que se encuentra en la WWW<sup>44</sup>.

- 1 estrella. Una estrella: ofrecer los datos en cualquier formato, aunque sean difíciles de manipular, como un pdf o una imagen escaneada.
- 2 estrellas. Dos estrellas: entregar los datos de manera estructurada, como en un archivo Excel con extensión xls.
- 3 estrellas. Tres estrellas: significa entregar los datos en un formato que no sea propietario, como csv en vez de Excel.
- 4 estrellas. Cuatro estrellas: usar URIs (que es una dirección web de un dato que sirve para enlazarlo con otros datos) para identificar cosas y propiedades, de manera que se pueda apuntar a los datos. Requiere usar un estándar RDF.

---

<sup>44</sup> BIBLIOTECA DEL CONGRESO NACIONAL DE CHILE. *Las cinco estrellas de los datos abiertos*. Consultado en febrero de 2017. Disponible en: <http://datos.bcn.cl/es/informacion/las-5-estrellas>

- 5 estrellas. Cinco estrellas: vincular sus datos con los de otras personas, dotándolos de contexto. En la práctica, a que la información entregada también apunte a otras fuentes de datos. Por ejemplo, si publico información sobre un libro dentro de mis datos, enlazar descripciones del libro que pertenezcan a otros publicadores (por ejemplo, DBPedia)<sup>45</sup>.

### 2.1.5 Ley 1712 de 2014

El Gobierno colombiano por medio de la Ley 1712 del 2014, ley de transparencia y del derecho de acceso a la información pública, en el Artículo 3° establece los siguientes principios que deben cumplirse para la transparencia y acceso a la información<sup>46</sup>:

- Principio de Transparencia. Principio conforme al cual toda la información en poder de los sujetos obligados definidos en esta ley se presume pública, en consecuencia, de lo cual dichos sujetos están en el deber de proporcionar y facilitar el acceso a la misma en los términos más amplios posibles y a través de los medios y procedimientos que al efecto establezca la ley, excluyendo solo aquello que esté sujeto a las excepciones constitucionales y legales y bajo el cumplimiento de los requisitos establecidos en esta ley.
- Principio de Buena Fe. En virtud del cual todo sujeto obligado, al cumplir con las obligaciones derivadas del derecho de acceso a la información pública, lo hará con motivación honesta, leal y desprovista de cualquier intención dolosa o culposa.
- Principio de no discriminación. De acuerdo con el cual los sujetos obligados deberán entregar información a todas las personas que lo soliciten, en igualdad de condiciones, sin hacer distinciones arbitrarias y sin exigir expresión de causa o motivación para la solicitud.
- Principio de Gratuidad. Según este principio el acceso a la información pública es gratuito y no se podrá cobrar valores adicionales al costo de reproducción de la información.
- Principio de Celeridad. Con este principio se busca la agilidad en el trámite y la gestión administrativa. Comporta la indispensable agilidad en el cumplimiento de las tareas a cargo de entidades y servidores públicos.

---

<sup>45</sup> BERNERS-LEE, T. *Gov 2.0 Expo 2010: Open, linked data for a global community*. 2010. Disponible en: <https://www.youtube.com/watch?v=ga1aSJXCFe0>

<sup>46</sup> PROCURADURÍA GENERAL DE LA NACIÓN. *Preguntas frecuentes de la Ley de transparencia y del derecho al acceso a la información 1712 de 2014*. Septiembre 25 de 2014. Disponible en: <https://www.procuraduria.gov.co/portal/media/file/PREGUNTAS.pdf>

- Principio de Eficacia. El principio impone el logro de resultados mínimos en relación con las responsabilidades confiadas a los organismos estatales, con miras a la efectividad de los derechos colectivos e individuales.
- Principio de la Calidad de la Información. Toda la información de interés público que sea producida, gestionada y difundida por el sujeto obligado, deberá ser oportuna, objetiva, veraz, completa, reutilizable, procesable y estar disponible en formatos accesibles para los solicitantes e interesados en ella, teniendo en cuenta los procedimientos de gestión documental de la respectiva entidad
- Principio de la Divulgación proactiva de la información. El derecho de acceso a la información no radica únicamente en la obligación de dar respuesta a las peticiones de la sociedad, sino también en el deber de los sujetos obligados de promover y generar una cultura de transparencia, lo que conlleva la obligación de Publicar y divulgar documentos y archivos que plasman la actividad estatal y de interés público, de forma rutinaria y proactiva, actualizada, accesible y comprensible, atendiendo a límites razonables del talento humano y recursos físicos y financieros.
- Principio de Responsabilidad en el uso de la información. En virtud de este, cualquier persona que haga uso de la información que proporcionen los sujetos obligados, lo hará atendiendo a la misma.

### **2.1.6 Carta Iberoamericana de Gobierno Abierto**

La carta dispone de 4 pilares fundamentales, en los cuales esta dispone que cada uno de ellos contribuya al desarrollo del otro de una forma independiente; por lo tanto, estos pilares son los que rigen el Gobierno abierto<sup>47</sup>.

- Transparencia y Acceso a la información Pública. Se debe orientar toda política gubernamental para impulsar la apertura de la información, de igual

---

<sup>47</sup> CENTRO LATINOAMERICANO DE ADMINISTRACIÓN PARA EL DESARROLLO – CLAD. *Carta Iberoamericana de Gobierno Abierto. XVII Conferencia Iberoamericana de Ministras y Ministros de Administración Pública y Reforma del Estado.* 2016. (21). p. 24. Disponible en: [https://www.sfp.gov.py/sfp/archivos/documentos/CIGA\\_2016%20final\\_knawtdh9.pdf](https://www.sfp.gov.py/sfp/archivos/documentos/CIGA_2016%20final_knawtdh9.pdf)

forma el gobierno debe favorecer y plantear formas y medios en los que las políticas de acceso y divulgación de información.

- Rendición de Cuentas Públicas. Los gobiernos deben plantear medios por los cuales el pueblo pueda realizar un control público de las entidades gubernamentales, así como se debe brindar toda la información referente.
- Participación Ciudadana. Como pilar del gobierno abierto, se deben dar todos los métodos y formas para la ejecución formal del mecanismo en los cuales la ciudadanía aporte al gobierno como medio de control de la gestión que este realiza.
- Colaboración e Innovación Pública y Ciudadana. El pilar de la colaboración se entiende como la generación de nuevos espacios de encuentro, diálogo y trabajo que favorezcan la co-creación de iniciativas y coproducción de nuevos servicios públicos, así como el desarrollo o emprendimiento por medio de la información pública que ha sido recogida, y puesta a disposición del público, ya sea para su consulta o manipularla para el emprendimiento.

### **2.1.7 Grupo Revolución de Datos**

Grupo conformado en el año 2014, por citación del secretario de la ONU Ban Ki-Moon para la búsqueda de la revolución de los datos que deben integrarse en el desarrollo sostenible y lo cual fue de gran impacto y de importancia ya que es un tema que es transversal a las políticas o leyes gubernamentales<sup>48</sup>.

- Calidad de la Información e Integridad. Todos los datos que sean usados deben ser de calidad, se debe asegurar que si son recolectados por terceros deben cumplir con un marco de aseguramiento; de ahí también se garantiza que los datos son íntegros que no han sido modificados o alterados.
- Desagregación de Datos. Los datos que sean recolectados deben ser lo más desagregados posibles, así como se debe proteger la privacidad y los derechos humanos en toda información que sea recolectada.
- Puntualidad de los Datos. La recolección de datos de calidad es importante para la política de datos abiertos, pero que los datos sean publicados sean recientes para generar una mayor gestión en el periodo en el que son publicados.
- Apertura y Transparencia de los Datos. Los sets de datos deben ser accesibles para todo tipo de persona, organización; con alguna política que proteja la privacidad y seguridad e integridad de estos, de igual forma deben

---

<sup>48</sup> IEAG. *A World that Counts*. Mobilising the data revolution for sustainable development. UN Report. 2015. Disponible en: [www.undatarevolution.org](http://www.undatarevolution.org)

disponerse de manuales y herramientas para el correcto manejo de la información de una forma independiente.

➤ Conservación y Utilización de los Datos. Las políticas de datos abierto aparte de promover, recoger y divulgar la información recolectada de forma abierta para cualquier tipo de persona deben fomentar la creación de comunidades de información, las cuales fomenten la implementación de herramientas para el uso de aquella información divulgada, así como formalizar cada proceso en el manejo de estos ya sean Webinar's, Conferencias, Cursos entre otros.

➤ Privacidad y Protección de Datos. La recopilación de información es algo demasiado importante, para ello se aplican marcos de seguridad, desagregación de la información ya sea por temáticas, pero esto no garantiza que la información que sea recopilada se encuentre segura; de ahí que los gobiernos deben aplicar e implementar las leyes internacionales las cuales dan lineamientos para proporcionar, recoger, curar y analizar datos sin violentar la libertad de funcionamiento y protección de recriminación.

➤ Manejo de Datos e Independencia. Las principales entidades que manejen la información recolectada de las naciones se les deben otorgar mayor financiamiento y lineamientos claros para que toda aquella información no se vea afectada por lineamientos políticos, debe mejorar su transparencia y rendición de cuentas.

➤ Capacidad y Recursos de los Datos. Es responsabilidad de todos los gobiernos garantizar un sistema estadístico nacional eficaz, en el cual sean medibles los estándares y expectativas, así como la inclusión de nuevas inversiones, ya sean recursos humanos, infraestructura, arquitectura del sistema; de igual forma el fomentar las oportunidades que se presentan por tener los sets de datos abiertos y por consiguiente un medio para un desarrollo sostenible.

➤ Derechos de los Datos. Se establece al igual que deben protegerse los derechos humanos, en los datos abiertos se deben proteger los derechos de las personas como el derecho a la identidad, privacidad, propiedad de datos de carácter personal, libertad de expresión, a la participación no discriminación. Es un deber del estado disponer de leyes, normas en los cuales promulgue el buen uso de la información que es recopilada y de igual forma estipular el tipo de control que debe ejercer para la protección de la persona.

## **2.2 ANÁLISIS DE INFORMACIÓN RECOPIADA**

### **2.2.1 Principios en Común**

De lo observado de cada una de las fuentes consultadas anteriormente, se encuentran en común entre la Ley 1712 de 2014 que es la guía para la

implementación de los datos abiertos en Colombia y, el Open Data Chárter, lo siguiente según se visualiza en la tabla 3:

Tabla 3. Comparación Ley 1712 y el Open Data Chárter

		Open Data Charter - Carta de Datos Abiertos					
		Abierto por defecto	Oportuna y completa	Accesible y utilizable	Comparables e interoperables	Para mejorar la gobernanza y la participación ciudadana	Para el desarrollo y la innovación inclusivos
Ley 1712 de 2014	Transparencia	X					
	Buena fe		X				
	No Discriminación	X			X		
	Gratuidad	X		X	X		
	Celeridad		X		X		
	Eficacia		X	X		X	
	Calidad de la Información		X		X		X
	Divulgación proactiva de la información				X	X	X
	Responsabilidad en el uso de la información				X		X

Fuente: Elaboración propia

Como se observa en la Tabla 3, de los 6 principios o lineamientos que estipula The Open Data Chárter para los datos abiertos, en común con la Ley 1712 de 2014 del gobierno colombiano, se encuentran que los 6 principios se encuentran distribuidos, de una manera uniforme, pero el principio número 4 de The Open Data Chárter que indica que los datos deben ser comparables e interoperables es aquel en el cual se tienen mayor cantidad de ítems en común con la Ley 1712 del 2014.

Ahora en la Tabla 4, se presenta el comparativo entre la Ley 1712 de 2014 y la Fundación Luz de Sol (Sunlight Foundation); donde se evidencia que de los principios 5 al 9 son los que más ítems en común tienen relación con los lineamientos y bases que da la ley 1712 del 2014, en los cuales se especifican que los datos no deben estar atados a un solo tipo de archivo, de igual forma no deben discriminar y deben ser de fácil acceso para cualquier persona o tipo de entidad.

Tabla 4. Comparación Ley 1712 de 2014 y los principios de la Fundación Sunlight

		Fundación Sunlight									
		Completo	Primarios	Oportunos	Facilidad de acceso físico y electrónico	Procesable por maquinas	No discriminatorios	Uso de normas de propiedad común	Licencia	Permanencia	Costos de uso
Ley 1712 del 2014	Transparencia						X	X	X		
	Buena fe										
	No Discriminación					X	X		X		
	Gratuidad										
	Celeridad										
	Eficacia			X	X						
	Calidad de la Información	X				X		X		X	X
	Divulgación proactiva de la información		X		X	X	X	X	X	X	X
	Responsabilidad en el uso de la información										

Fuente: Elaboración propia

En la siguiente Tabla 5, se muestra la comparativa entre Los Datos del Gobierno Abierto (The Open Government Data) y la Ley 1712 del 2014, evidenciándose una variante ya que de los lineamientos que expone The Open Government 3 de sus 8 principios se relacionan con los principios que da la Ley 1712 del 2014 de Colombia

Tabla 5. Comparación entre la Ley 1712 de 2014 y el Open Government Data

		Open Government Data							
		Completos	Primarios	Oportunos	Accesibles	Procesables por maquina	No Discriminatorio	No Propietario	Licencia Libre
Ley 1712 de 2014	Transparencia				X	X		X	X
	Buena fe								
	No Discriminación		X	X		X	X		X
	Gratuidad								
	Celeridad	X	X		X				
	Eficacia	X		X	X				
	Calidad de la Información	X	X			X	X	X	X
	Divulgación proactiva de la información	X				X		X	X
	Responsabilidad en el uso de la información	X	X	X	X	X		X	X

Fuente: Elaboración propia

En la tabla 6 se evidencia la comparación de la Ley 1712 de 2014 y la propuesta de Tim Berners Lee, donde se visualiza que la categorización de la quinta estrella que propone, como lineamiento de los datos abiertos es la que más enfoque tiene con el gobierno colombiano, ya que esta busca la interoperabilidad en la información con demás personas, entidades o institutos que hagan uso de la información que allí se encuentre.

Tabla 6. Comparación entre la Ley 1712 de 2014 con Tim Berners Lee

		Tim Berners Lee				
		1 Estrella	2 Estrella	3 Estrella	4 Estrella	5 Estrella
Ley 1712 de 2014	Transparencia	X	X	X	X	X
	Buena fe					X
	No Discriminación		X	X	X	X
	Gratuidad	X	X	X	X	X
	Celeridad				X	X
	Eficacia			X	X	X
	Calidad de la Información		X	X	X	X
	Divulgación proactiva de la información	X	X	X	X	X
	Responsabilidad en el uso de la información					X

Fuente: Elaboración propia

La comparación siguiente es la efectuada entre la Ley 1712 de 2014 y la Carta Iberoamericana de Gobierno Abierto la cual estipula 4 lineamientos para la apertura de la política de datos abiertos, de los cuales en la Ley 1712 la que mayor impacto tiene es el lineamiento numero 4 ya que este se alinea al ideal a la estrategia GEL, en la cual la aplicación de los datos abiertos influye en los procesos de validación y gestión de control institucional por parte de entes externos al gobierno, lo cual impulsaría la Urna de Cristal (Programa del Gobierno Colombiano para la transparencia del Gobierno), ver Tabla 7.

Tabla 7. Comparación entre la Ley 1712 de 2014 y la Carta Iberoamericana de Gobierno Abierto

		Carta Iberoamericana de Gobierno Abierto			
		Transparencia y Acceso de Información Pública	Rendición de Cuentas Públicas	Participación Ciudadana	Colaboración e Innovación Pública y Ciudadana
Ley 1712 de 2014	Transparencia	X	X	X	X
	Buena fe				X
	No Discriminación	X	X		X
	Gratuidad	X	X	X	X
	Celeridad			X	X
	Eficacia		X	X	X
	Calidad de la Información		X	X	X
	Divulgación proactiva de la información		X		X
	Responsabilidad en el uso de la información		X		X

Fuente: Elaboración propia

En seguida, en la tabla 8 se muestra la investigación impulsada por la ONU, nos evidencia la importancia de los datos abiertos para un desarrollo sostenible para el gobierno; en referencia a la ley 1712 del 2014, el principio 7 nos estipula que la gobernanza debe ser un eje central de la aplicación de políticas abiertas en un gobierno, lo cual facilita la gestión y control del gobierno, pero sin descuidar factores de la apertura política

Tabla 8. Comparación entre la Ley 1712 de 2014 y el Data Revolution Group

		Data Revolution Group								
		Calidad de la Información e Integridad	Desagregación de Datos	Puntualidad de los Datos	Apertura y Transparencia de los Datos)	Conservación y Utilización de los Datos	Privacidad y Protección de Datos	Manejo de Datos e Independencia	Capacidad y Recursos de los Datos	Derechos de los Datos
L E Y  1 7 1 2  D E L  2 0 1 4	Transparencia								X	
	Buena fe						X			
	No Discriminación							X		
	Gratuidad				X					
	Celeridad			X						
	Eficacia					X				
	Calidad de la Información	X	X							
	Divulgación proactiva de la información					X		X		
Responsabilidad en el uso de la información							X	X	X	

Fuente: Elaboración propia

De acuerdo a lo expuesto anteriormente de las Tablas 3 a la 8, se realizó la comparación una a una entre la Ley 1712 de 2014 y cada una de las organizaciones a nivel mundial que tienen principios o lineamientos para los datos abiertos en línea, por lo que se presentará seguidamente en la Tabla 9, la matriz general de todas ellas, para llegar a una conclusión final.

Tabla 9. Comparativo General

		OPEN DATA CHARTER					FUNDACIÓN SUNLIGHT						OPEN GOVERNMENT DATA					TIMBERNERS LEE					CARTA IBEROAMERICANA				DATA REVOLUTION GROUP																			
		Abierto por defecto	Oportuna y completa	Accesible y utilizable	Comparables e interoperables	Para mejorar la gobernanza y la participación ciudadana	Para el desarrollo y la participación inclusiva	Completo	Primario	Oportunos	Facilidad de acceso físico y electrónico	Procesables por máquina	No discriminatorio	Uso de normas de propiedad intelectual	Licencia	Permanencia	Costos de uso	Completos	Primarios	Oportunos	Accesibles	Procesables por máquina	No discriminatorio	No propietario	Licencia Libre	1 estrella	2 estrella	3 estrella	4 estrella	5 estrella	Transparencia y acceso de	Rendición de cuentas públicas	Participación Ciudadana	Colaboración e innovación pública y ciudadana	Calidad en la información e	Desagregación de datos	Puntualidad de los datos	Apertura y transparencia de los	Conservación y utilización de los	Privacidad y protección de datos	Manejo de datos e	Capacidad y recursos de los	Derechos de los datos			
Ley 1712 de 2014	Transparencia	X										X	X	X						X	X			X	X	X	X	X	X	X	X	X	X									X				
	Buena Fe		X																																							X				
	Discriminación	X			X						X	X		X				X	X		X	X			X		X	X	X	X	X	X	X											X		
	Gratuidad	X		X	X																				X	X	X	X	X	X	X	X	X				X									
	Celeridad		X		X													X	X		X								X	X																
	Eficacia		X	X		X			X	X								X	X	X								X	X	X																
	Calidad de información		X		X		X	X			X		X		X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X											
	Divulgación proactiva de la información				X	X	X		X		X	X	X	X	X	X	X	X				X			X	X	X	X	X		X		X													
Responsabilidad en el uso de la información				X		X											X	X	X	X	X		X	X					X		X											X	X	X		

Fuente: Elaboración propia

Como se observa en la tabla 9 y en la información recopilada anteriormente, en los principios de Tim Berners y en la Carta Iberoamericana sólo presentan 5 y 4 principios respectivamente, que aunque coincidan varios cruces con la Ley 1712 de 2014, no categorizan fundamentalmente los principios de la ley colombiana, entre tanto, la Fundación Sunlight aunque tiene 10 categorías o principios solo tiene 21 cruces con la Ley 1712; con respecto a los 9 principios del Data Revolution Group sólo cruza 13 veces con los 9 principios nacionales, y por último, el Open Data Chárter solo tiene 6 principios que por sus definiciones, aunque crucen los lineamientos de cada uno son disímiles.

Debido a lo anterior, es evidente que la Ley Colombiana (con 9 principios) y el Open Government Data (con 8 principios), son los que mayor afinidad tienen entre sus definiciones y el enfoque que se le da a cada uno, como se pudo observar en el numeral anterior.

### **2.3 CONCLUSIÓN DEL ANÁLISIS**

Del análisis anterior, en referencia a los principios que se han aplicado a nivel internacional los cuales varían la perspectiva según la organización, su finalidad u objetivo, de cada uno de los principios que se encuentran mencionados y analizados, todos buscan una apertura de la información al público y que toda aquella información sea gratuita, confiable y que el acceso de la información se realice de la forma más simple y eficaz para cualquier tipo de usuario.

La selección de los principios en los que se basa este proyecto son los estipulados por el “Open Government Data”, los que establecen unos principios que van acorde a un desarrollo sostenible, los cuales son más concordantes a la implementación que ha realizado el gobierno colombiano, donde busca que estas plataformas y políticas abiertas ayuden para el mismo desarrollo del país.

### **3. DISEÑO PROTOTIPO DE SOFTWARE**

Una vez expuesto el análisis anterior, y establecido que los principios que se van a utilizar son los del Open Government Data, para lograr establecer si los lineamientos nacionales expuestos por la Ley 1712 de 2014 son coherentes con ellos y analizar su correcta utilización por los usuarios, se plantea en este capítulo el diseño del prototipo propuesto con base en lo siguiente:

La metodología usada para el desarrollo del software es la metodología en espiral, la cual se mencionó anteriormente, adicionalmente en el desarrollo de software se suele utilizar UML(Unified Modeling Language), el cual permite modelar un sistema por medio de diferentes diagramas, permitiendo mostrar su comportamiento y configuraciones por medio de ellos (Se puede observar el Anexo A, evidenciándose capítulos 3 y 4).

#### **3.1 DIAGRAMAS**

La realización de los diagramas permitira tener un visión general del sistema en desarrollo, tanto en su configuración, funcionalidades y conexiones con sistemas externos. Por lo mismo, se realizó la selección de los siguientes diagramas ya que con ellos es posible mostrar de manera general el alcance del prototipo a realizar y su estructura principal como software.

##### **3.1.1 Diagrama de despliegue**

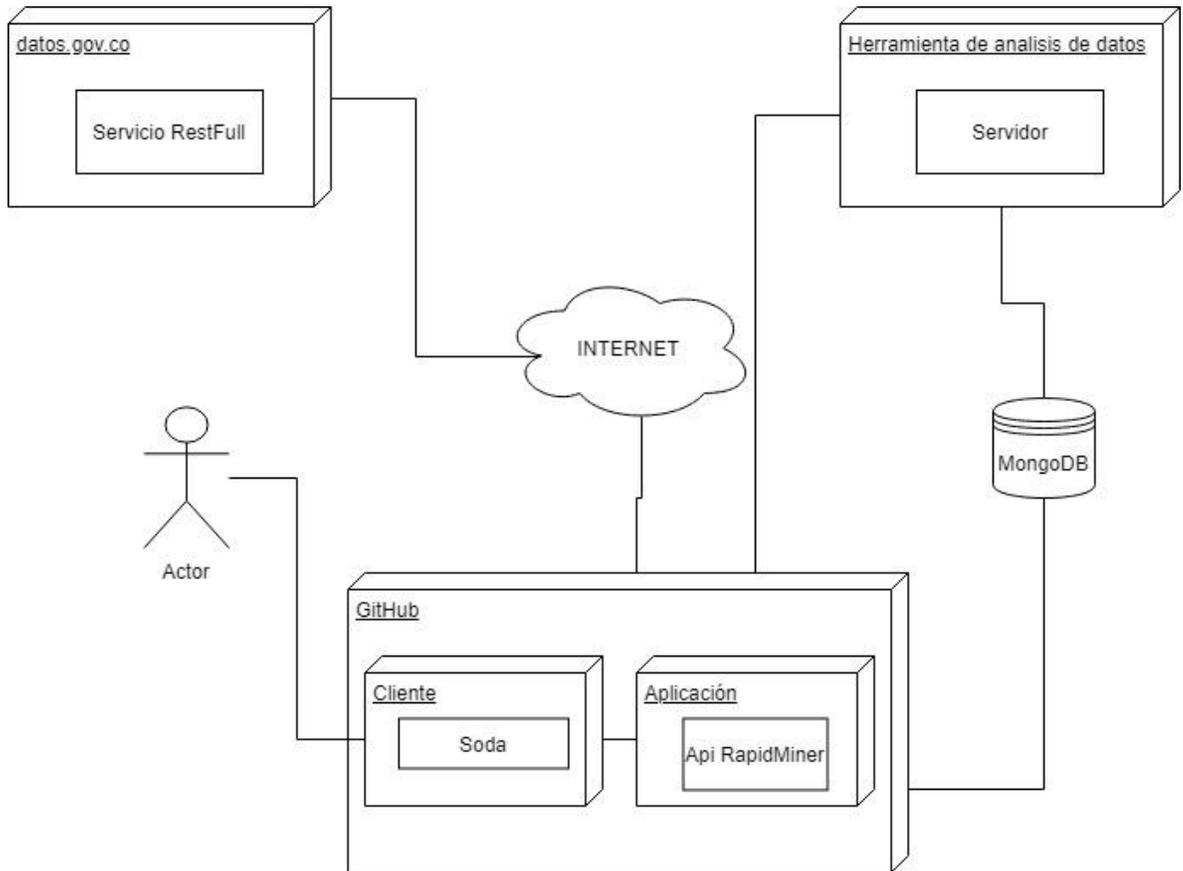
Los diagramas de despliegue son utilizados para modelar el hardware necesario para el despliegue de la de la herramienta y las relaciones entre los componentes. De esta manera será más fácil ver que equipos serán usados para el perfecto funcionamiento de la aplicación.

Según se observa en la Figura 3, el usuario cuando ingresa a la aplicación, lo primero que hará es extraer la información presente en datos.gov.co por medio de un cliente que se conectará con un servicio web.

Una vez el cliente obtiene conexión, procede a la extracción de la información. Dicha información es almacenada en una base de datos no relacional. La aplicación analiza la información y entrega unos resultados. Una vez se tiene la información almacenada, esta es extraída por RapidMiner (Ver anexo B), quien se encarga de procesar la información. Una vez se procesa y se obtiene un resultado, este es mostrado en la aplicación.

Se decidió utilizar una base de datos no relacional debido a que los sets de datos presentes en datos.gov.co no tienen la misma estructura. Cuando se habla de estructura, hace referencia a los encabezados o variables que conforman los sets de datos.

Figura 3. Diagrama de despliegue



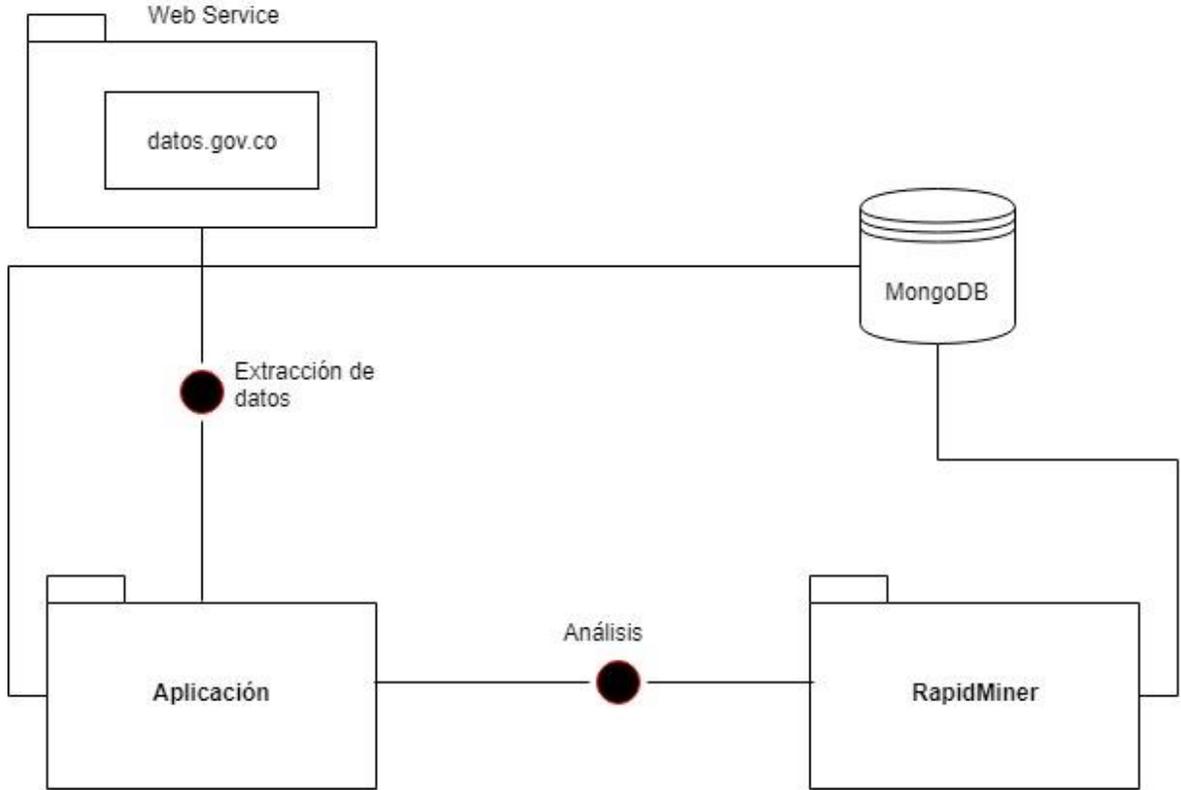
### 3.1.2 Diagrama de componenetes

El diagrama de componentes muestra cómo se divide un sistema en sus componentes y las dependencias entre ellos.

Siguiendo con el diseño de la arquitectura, el diagrama de componentes permite apreciar la interacción entre los componenetes del sistema, mostrando la relación y la utilización de distintos métodos que posteriormente llevan a unas funciones que cumple el sistema.

En la Figura 4, se quiere representarse a grandes rasgos el comportamiento de los componentes que comprende el sistema.

Figura 4. Diagrama de Componentes



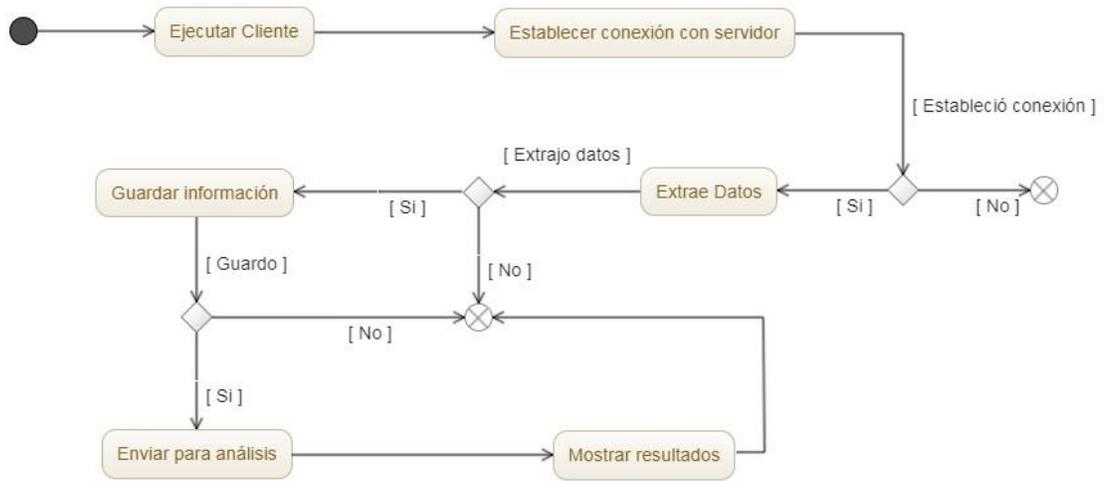
### 3.1.3 Diagrama de Actividades

El diagrama de actividades muestra el proceso que cumple el software a través una serie de acciones que conllevan a la culminación del proceso. Estas acciones son realizadas tanto por los usuarios como los sistemas.

Para modelar la secuencia de acciones que comprende el sistema, se utiliza el diagrama de actividades.

La ejecución del cliente permite establecer conexión con el servidor de donde se van a extraer los datos (`datos.gov.co`). Una vez se extraen son almacenados y analizados para después mostrar un resultado el análisis realizado. Ver Figura 5.

Figura 5. Diagrama de Actividades



### 3.1.4 Diagrama de Casos de Uso

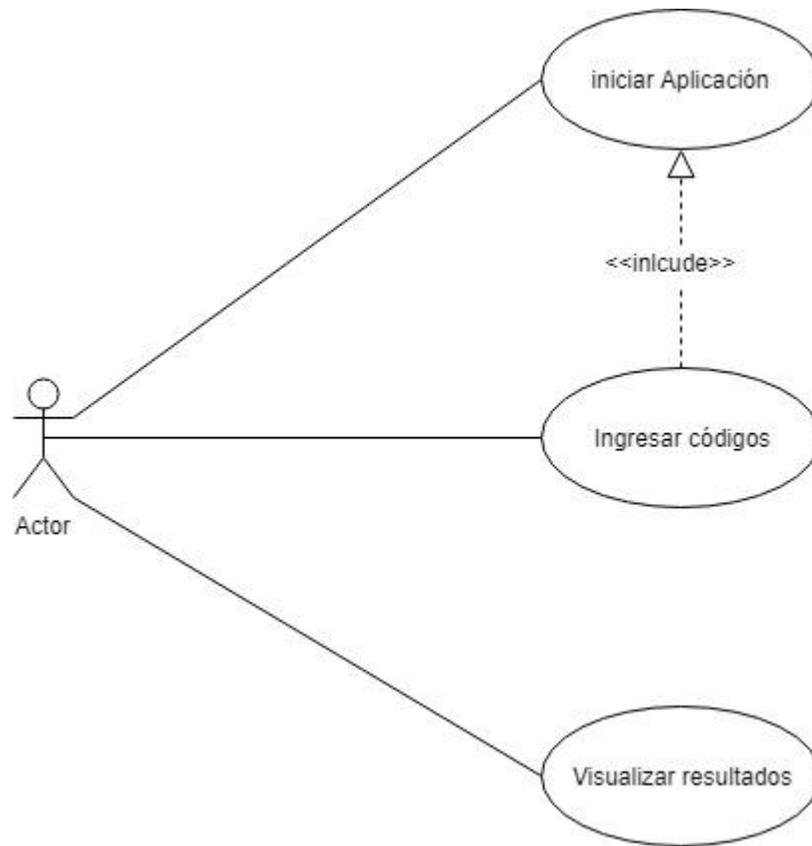
El diagrama de casos de uso permite ver la interacción del usuario con las funcionalidades del sistema.

Los diagramas de casos de uso son otro de los tantos diagramas que se pueden utilizar para modelar un sistema utilizando UML. Este diagrama describe cómo opera el usuario (Actor) con el sistema. Además, muestra el orden de interacción de los elementos con el usuario.

La interacción que tiene el usuario con el sistema es modelado y mostrado en la Figura 6. Las actividades que puede realizar el usuario son representadas y su orden de ejecución puede ser apreciado en la figura.

Cabe resaltar que, solo hay un caso de uso, pues el aplicativo es sencillo para el uso por lo cual no se ve la necesidad de la realización de más casos de uso en el desarrollo e implementación del prototipo, por lo cual solo se presenta un diagrama.

Figura 6. Diagrama de Casos de Uso



Los casos de uso que se utilizaron y desarrollaron para el prototipo del software propuesto son:

Tabla 10. Caso de uso ingresar estructura de datos

Caso de uso: Ingresar estructura de datos	
Actor: Usuario de la aplicación	
Curso normal	Alternativas
1. El usuario debe ir al sitio de datos.gov.co	
2. Buscar el set de datos al cual se le quiere sacar la estructura para ser almacenada.	
3. Ir a la URL del set de y copiar los últimos 7 caracteres que aparecen en ella.	
4. Abrir el software e ir a la pestaña ingreso de estructura.	
5. En el primer input pegar los dígitos que se copiaron.	
6. En el segundo input poner la URL de datos.goc.co	
7. Oprimir el botón.	

8. Si todo salió bien, el sistema mostrara un mensaje de que almaceno la estructura.	8.1 De lo contrario, el sistema mostrara un mensaje de que no se pudo almacenar la estructura.
--	--

Tabla 11. Caso de uso ingresar set de datos

<b>Caso de uso:</b> Ingresar set de datos	
<b>Actor:</b> Usuario de la aplicación	
<b>Curso normal</b>	<b>Alternativas</b>
1. El usuario debe ir al sitio de datos.gov.co	
2. Buscar el set de datos al cual se le quiere sacar la estructura para ser almacenada.	
3. Ir a la URL del set de y copiar los últimos 7 caracteres que aparecen en ella.	
4. Abrir el software e ir a la pestaña ingreso de datos.	
5. En el primer input pegar los dígitos que se copiaron.	
6. En el segundo input poner la URL de datos.goc.co	
7. Oprimir el botón.	
8. El sistema valida si la estructura del set de datos coincide con las estructuras que están almacenadas.	
9. Si la estructura del set de datos cumple con las almacenadas, el sistema almacena el set de datos y muestra un mensaje de que se almaceno.	9.1 Si el set no cumple con la estructura, el sistema impedirá que se almacene el set de datos y mostrará un mensaje de advertencia.

Tabla 12. Caso de uso Visualización de los resultados del análisis

<b>Caso de uso:</b> Visualización de los resultados del análisis	
<b>Actor:</b> Usuario de la aplicación	
<b>Curso normal</b>	<b>Alternativas</b>
1. Abrir el software e ir a la pestaña Madurez.	
2. Cuando se ingresa a la pestaña, el sistema empezara a cargar los gráficos y tablas resultantes del análisis.	2.1 Si no hay datos, el sistema no mostrara nada en la interfaz.

### 3.2 REQUERIMIENTOS DE SOFTWARE

Una vez vistos los diagramas, se muestran los requerimientos pertinentes, así:

### **3.2.1 Requerimientos Funcionales**

1. El sistema debe tener conexión con RapidMiner.
2. El sistema debe poder conectarse a la página datos.gov.co por medio de Soda Api.
3. El sistema debe almacenar la información extraída.
4. El sistema debe tener una interfaz donde se muestren los resultados del análisis realizado usando gráficos.
5. El sistema tendrá definidas dos estructuras de datos de acuerdo a las necesidades.
6. El sistema debe permitir ingresar los sets de datos de contratos.
7. El sistema contara con 3 interfaces. Ingreso de datos, ingreso de estructuras y madurez.

### **3.2.2 Requerimientos No Funcionales**

8. El sistema almacenara la información en MongoDB
9. El sistema debe ser desarrollado en Java
- 10 El análisis de los datos será realizado en RapidMiner
- 11 El sistema debe ser desplegado en un servidor de Amazon

Tabla 13. Descripción de requerimiento 1

DESCRIPCIÓN DE REQUERIMIENTOS			
<b>Nombre del proyecto:</b> Evaluación de principios de datos abiertos		<b>Fecha:</b>	
Dependencia:		<b>Usuario:</b>	
<b>Cargo del usuario:</b>			
Analista responsable:		Requerimiento Nro. 1	
<b>Descripción de requerimiento:</b>			
El sistema que se va a diseñar debe tener conexión con rapidmainer que permite analizar la información extraída de datos.gov.co.			
El resultado del análisis debe ser embebido en el sistema.			
Tipo de requerimiento	Funcional	X	No Funcional
<b>Restricción de requerimiento:</b>			
Para realizar el análisis es necesario programar los procesos necesarios para tal fin			

Tabla 14. Descripción de requerimiento 2

DESCRIPCIÓN DE REQUERIMIENTOS			
<b>Nombre del proyecto:</b> Evaluación de principios de datos abiertos		<b>Fecha:</b>	
Dependencia:		<b>Usuario:</b>	
<b>Cargo del usuario:</b>			
Analista responsable:		Requerimiento Nro. 2	
<b>Descripción de requerimiento:</b>			
Para poder extraer la información de datos.gov.co se usará el api soda que hará más fácil la extracción de la información.			
Tipo de requerimiento	Funcional	X	No Funcional
<b>Restricción de requerimiento:</b>			

Tabla 15. Descripción de requerimiento 3

DESCRIPCIÓN DE REQUERIMIENTOS			
<b>Nombre del proyecto:</b> Evaluación de principios de datos abiertos		<b>Fecha:</b>	
Dependencia:		<b>Usuario:</b>	
<b>Cargo del usuario:</b>			
Analista responsable:		Requerimiento Nro. 3	
<b>Descripción de requerimiento:</b>  La información extraída debe ser almacenada para posteriormente ser utilizada para el análisis.			
Tipo de requerimiento	Funcional	X	No Funcional
<b>Restricción de requerimiento:</b>			

Tabla 16. Descripción de requerimiento 4

DESCRIPCIÓN DE REQUERIMIENTOS			
<b>Nombre del proyecto:</b> Evaluación de principios de datos abiertos		<b>Fecha:</b>	
Dependencia:		<b>Usuario:</b>	
<b>Cargo del usuario:</b>			
Analista responsable:		Requerimiento Nro. 4	
<b>Descripción de requerimiento:</b>  Una vez realizado el análisis de la información, el resultado del análisis se debe poder ver en una interfaz.			
Tipo de requerimiento	Funcional	X	No funcional
<b>Restricción de requerimiento:</b>  El resultado debe ser apoyado con gráficos que ayuden a entender los resultados obtenidos del análisis de la información.			

Tabla 17. Descripción de requerimiento 5

DESCRIPCIÓN DE REQUERIMIENTOS			
<b>Nombre del proyecto:</b> Evaluación de principios de datos abiertos		<b>Fecha:</b>	
<b>Versión:</b> 1.0			
Analista responsable:		Requerimiento Nro. 5	
<b>Descripción de requerimiento:</b>  Se deben definir las estructuras de datos encontradas en los sets de datos de contratos. En este caso se encontraron dos diferentes estructuras.			
Tipo de requerimiento	Funcional	X	No funcional
<b>Restricción de requerimiento:</b>  Si la estructura de los datos no ha sido definida, el sistema no debe permitir que se almacene dicho set de datos.			

Tabla 18. Descripción de requerimiento 6

DESCRIPCIÓN DE REQUERIMIENTOS			
<b>Nombre del proyecto:</b> Evaluación de principios de datos abiertos		<b>Fecha:</b>	
<b>Versión:</b> 1.0			
Analista responsable:		Requerimiento Nro. 6	
<b>Descripción de requerimiento:</b>  El sistema debe permitir ingresar los sets de datos de contratos presentes en datos.gov.co.			
Tipo de requerimiento	Funcional	X	No funcional
<b>Restricción de requerimiento:</b>  Esto se debe comprobar con las estructuras definidas y que acepta el sistema.			

Tabla 19. Descripción de requerimiento 7

DESCRIPCIÓN DE REQUERIMIENTOS			
<b>Nombre del proyecto:</b> Evaluación de principios de datos abiertos		<b>Fecha:</b>	
<b>Versión:</b> 1.0			
Analista responsable:		Requerimiento Nro. 7	
<b>Descripción de requerimiento:</b>			
El sistema debe contar con 2 interfaces:			
Ingreso de datos: Esta ventana contará con dos inputs y un botón que permitirá guardar la información. El primer input será para poner el código que tiene cada set de datos y el otro será para poner la URL del datos.gov.co.			
Madurez: En esta interfaz se mostrará el resultado del análisis de los datos.			
Tipo de requerimiento	Funcional	X	No funcional
<b>Restricción de requerimiento:</b>			

Después de analizar los requerimientos de redefinieron algunos de los ya mencionados. La razón es porque con el cambio de estos requerimientos, se hace más eficiente y dinámica la aplicación. Los cambios son:

Tabla 20. Descripción de requerimiento 5

DESCRIPCIÓN DE REQUERIMIENTOS			
<b>Nombre del proyecto:</b> Evaluación de principios de datos abiertos		<b>Fecha:</b>	
<b>Versión:</b> 2.0			
Analista responsable:		Requerimiento Nro. 5	
<b>Descripción de requerimiento:</b>			
El sistema debe permitir el ingreso de las estructuras que tienen los diferentes sets de datos, es decir que ya no es necesario definir solo las estructuras de contratos como se había definido antes.			
Tipo de requerimiento	Funcional	X	No funcional
<b>Restricción de requerimiento:</b>			
Las estructuras deben ser almacenadas para poder compararlas a la hora de estar ingresando los sets de datos.			

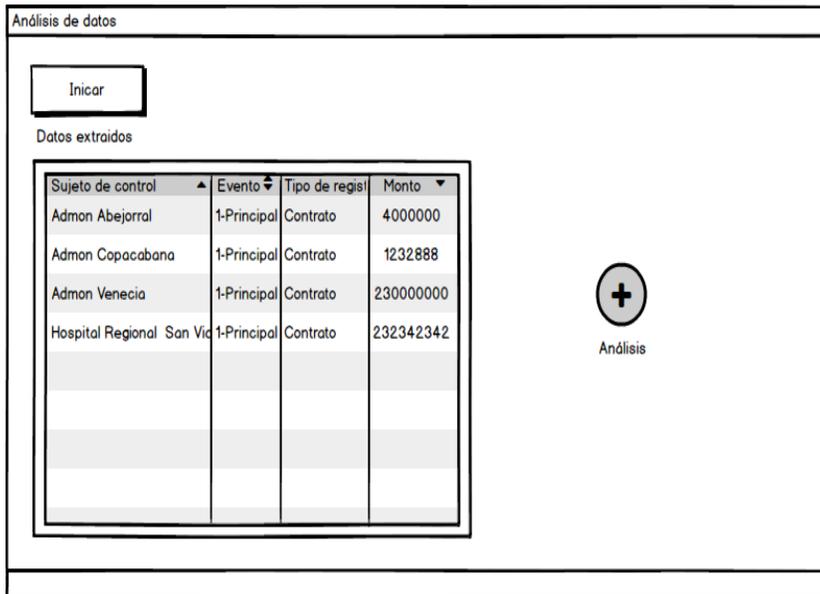
Tabla 21. Descripción de requerimiento 7

DESCRIPCIÓN DE REQUERIMIENTOS			
<b>Nombre del proyecto:</b> Evaluación de principios de datos abiertos		<b>Fecha:</b>	
<b>Versión:</b> 2.0			
Analista responsable:		Requerimiento Nro. 7	
<b>Descripción de requerimiento:</b>			
El sistema debe contar con 3 interfaces:			
<ul style="list-style-type: none"> <li>· Ingreso de datos: Esta ventana contará con dos inputs y un botón que permitirá guardar la información. El primer input será para poner el código que tiene cada set de datos y el otro será para poner la URL del datos.gov.co.</li> <li>· Ingreso de estructura: Esta ventana será similar a la del ingreso de datos, contara con dos inputs que serán los mismos de la ventana anterior, pero a la hora de oprimir el botón, este debe guardar la estructura que comprende dicho set de datos.</li> <li>· Madurez: En esta interfaz se mostrará el resultado del análisis de los datos.</li> </ul>			
Tipo de requerimiento	Funcional	X	No funcional
<b>Restricción de requerimiento:</b>			

### 3.3 MOCKUPS

El acercamiento a las interfaces que tendrá el sistema se hace por medio de los mockups. De esta manera es fácil tener una previsualización a grandes rasgos de cómo quedara el sistema. Se puede observar en las Figuras 7, 8 y 9.

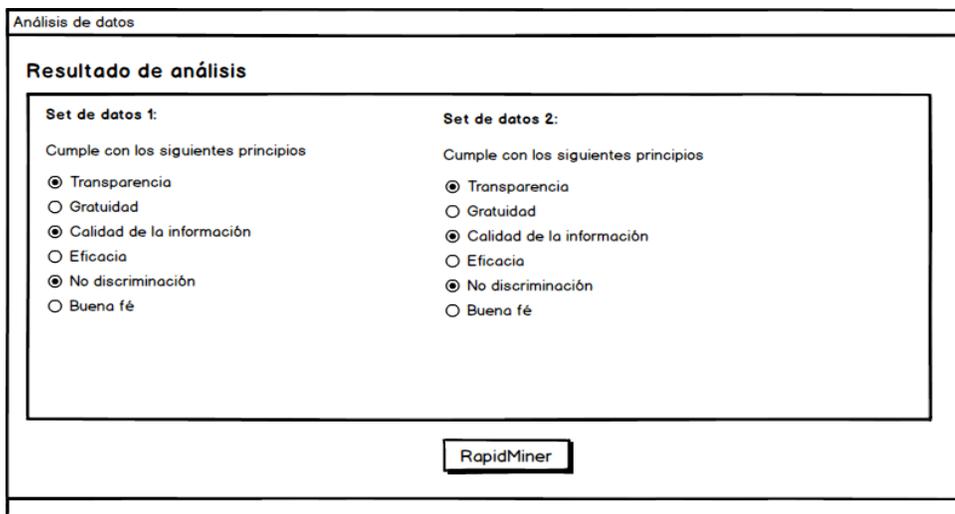
Figura 7. Iniciar aplicación



En esta primera pantalla se muestra un listado de los sets de datos almacenados previamente, junto con un botón que al presionarse inicia el análisis de la información.

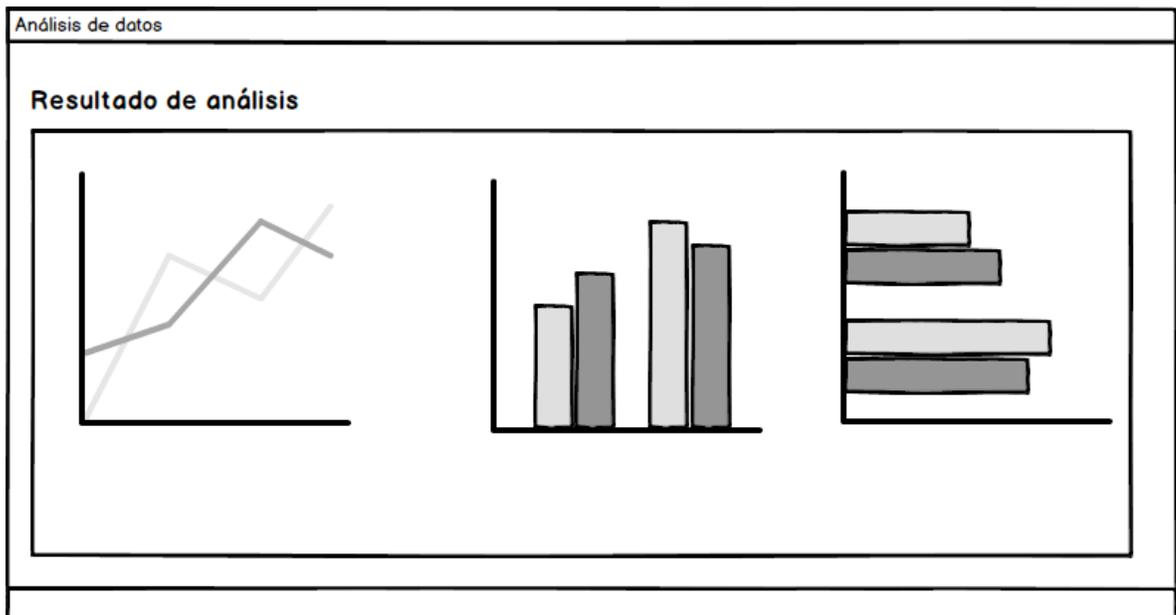
A continuación, en la Figura 8 se muestra el resultado del análisis, en él se puede apreciar si el set de datos cumple con los principios establecidos.

Figura 8. Resultados del Análisis



Al seleccionar un set de datos de la lista general, se visualiza de forma específica cómo se comporta ese set de datos en referencia a los principios seleccionados.

Figura 9. Gráficas de Resultados



Para algunos resultados se mostrarán algunas graficas las cuales se podrán visualizar en la pantalla como se observa en la Figura 9. Del mismo modo, en la exposición de los resultados, ya sea de forma general o individual, se visualizarán tablas, gráficos los cuales facilitarán el entendimiento de la información allí consignada.

Con los mockups anteriores se pretende dar un acercamiento al aspecto visual que tendrá el sistema en desarrollo.

#### 4. ARQUITECTURA DE DATOS

Para el sistema en desarrollo se ha definido una estructura de datos la cual será usada a lo largo del desarrollo y que hace parte de los procesos que realizará el sistema.

Los tipos de datos utilizados serán los siguientes:

- String: Cadenas de caracteres.
- Integer: Números enteros.
- Float: Números decimales.
- Boolean: Lógicos (True, False).
- Date: Fechas.

Basándose en los sets de datos que se extraen de datos.gov.co en cuanto a contratos, se definieron dos estructuras para el almacenamiento de los datos. Estas estructuras se definen como listas en el lenguaje de programación Java.

Tabla 22. Estructura de datos

Variable	Estructuras		Tipos de Datos				
	Estructura 1	Estructura 2	String	Integer	Float	Boolean	Date
codigo_contrato	X		X				
codigo_del_proyecto	X		X				
c_digo_contrato		X	X				
c_digo_del_proyecto		X	X				
disponibilidades_presupuestales	X	X					
Evento	X	X	X				
fecha_inicio	X	X					X
fecha_suscripci_n		X					X
fecha_suscripcion	X						X
identificacion_contratista	X		X		X		
identificaci_n_contratista		X			X		
identificacion_interventor	X				X		
identificaci_ninterventor		X			X		
no_contrato_interventor	X	X					
nombre_contratista	X		X				
nombrecontratista		X	X				
nombre_del_interventor	X	X	X				
nombre_del_proyecto	X	X	X				
objeto_del_contrato	X	X	X				
plazo_estimado	X	X		X			
proceso_de_contratacion	X	X					
proceso_de_contrataci_n		X					

registros_presupuestales	X	X			X		
sector_del_proyecto	X	X	X				
sujeto_de_control	X	X	X				
tipo_de_registro	X	X	X				
tipo_interventor	X	X	x				
tipologia	X		X				
tipolog_a		X	X				
valor_contrato	X	X			X		
valor_del_proyecto	X				X		
valor_ejecutado_del_proyecto	X				X		
valor_ejecutadodel_proyecto		X			X		

En la tabla 10 se pueden apreciar las variables que componen cada una de las estructuras de los datos y además el tipo de dato de cada variable. La razón por la que se definieron estas dos estructuras es porque al momento de revisar los datos de contratos, se encontraron diferencias entre los campos que se usaban en la estructura de los datos, algunos sets de datos tenían más campo o tenían nombres diferentes. Sin embargo, cabe resaltar que el software permite el ingreso de cualquier tipo de estructura de datos, pero por motivos de alcance se limita a la sección contratos y se realizó el análisis de los resultados con los tipos de estructuras anteriores. (Ver Anexo A).

## 5. DEFINICIÓN DEL PROCESO DE EVALUACIÓN DE LA MADUREZ

Como se dijo anteriormente la madurez de los datos abiertos, se realizará con base en los criterios de Open Government Data, en donde se estipulan los siguientes principios para los datos abiertos: completos, primarios, oportunos, accesibles, procesables por máquina, no discriminatorio, no propietario, licencia libre.

## 5.1 COMPLETOS

En la Tabla 11 se describe el principio número uno de datos abiertos del Open Government Data

Tabla 23. Descripción de la medición del principio completo de datos abierto

Titulo	Completo
Definición	La proporción de datos almacenados frente al potencial de "100% completado"
Medición	Una medida en la ausencia de valores en blanco (NULOS) o la presencia de valores no en blanco
Alcance	0-100%, en toda la data extraída
Unidad de medida	Porcentaje

### 5.1.1 MEDICIÓN

- • ListWise deletion

Este método aplicable sobre información faltante (Missing Values), consiste en la eliminación en el set de información que contenga algún dato faltante, para poder realizar el análisis de la información completa; la principal virtud Listwise es la simplicidad. Si puede resolver un problema de falta de datos descartando sólo una pequeña parte de la muestra, el método puede ser muy efectivo. Sin embargo, incluso en esa situación, uno debe explorar los datos para asegurarse de que los casos desechados no son excesivamente influyentes. (Schafer & Graham, 2002).

El ListWise deletion, se aplicará de la siguiente forma:

$f$  = Cantidad de filas incompletas

$C$  = Cantidad de filas completas

$T$  = Total de filas

$$\% \text{ de datos incompletos} = \frac{f * 100}{T}$$

$$\% \text{ de datos completos} = \frac{C * 100}{T}$$

La razón de aplicar este método es la facilidad en la cual la muestra donde se podrá realizar un análisis mucho más exhaustivo es reducida y concisa; esto permite tener acceso a la información oportuna

De igual manera se anexa la configuración que se realiza en el motor de análisis, como se evidencia en las siguientes imágenes, para la evidencia de la estructuración y modificación que se realizar para esta métrica:

Figura 10. Configuración Principio Completo

attribute name	function expressions
numCells	[count(_id.\$oid)]*[average(numColumns)]
PerCompleteCells	(1-([sum(incomplt)]/(numCells)))*100
% Complete rows	(1-([sum(incompleteRows)]/[count(_id.\$oid)]))*100

En la siguiente imagen se evidencia, el % de datos completos del set que se tienen.

Figura 11. Resultado Principio Completos

Row No.	% Complete rows
1	0.607

## 5.2 PRIMARIOS

Tabla 24. Descripción para la medición de un conjunto de datos primarios

Titulo	Primario
Definición	La proporción de datos almacenados, pero son al mismo tiempo datos para la fuente principal que los genera
Medición	Una medida cualitativa y tendrá el nombre de la organización de origen de los datos

Alcance	El nombre de la entidad que actualiza los datos
Unidad de Medida	Texto

Para la medición de este principio expuesto en la Tabla 12, se realiza por medio de la extracción de los metadatos donde se tomará el propietario del set de datos, es decir, quien realizó la creación de la información y se realiza el siguiente análisis:

*Si ("Información del Propietario del Set) = ! (Vacío), entonces  
Resultado: Paso  
Si no:  
Resultado: No Paso*

Se anexará la configuración que se realiza en el motor de análisis y el resultado que se obtiene del mismo.

Figura 12. Configuración Principio Primario

attribute name	function expressions
Result	if(Owner=="", "No Paso", "Paso")

El resultado que se obtiene es el siguiente:

Figura 13. Resultado Principio Primario

Row No.	Owner	Category	Result
1	Contraloria General de Antioquia	Organismos de Control	Paso

### 5.3 OPORTUNO

Tabla 25. Descripción para la medición de un conjunto de datos oportunos

Título	Oportuno
Definición	El grado en que los datos representan la realidad desde el punto requerido en el tiempo
Medición	Diferencia de tiempo
Alcance	Cualquier elemento de fecha sobre el tiempo de actualización

Unidad de medida	Tiempo
------------------	--------

Para la medición se usará la función `date_diff()`, del motor de análisis de datos en el cual se compara la fecha actual y la fecha en la que fueron creados o actualizado el set de datos, usando las fechas de creación del set de datos y la fecha de actual para realizar el cálculo.

*date\_diff(emision, Hoy)*

De igual forma se muestra la configuración que se realiza en el motor de análisis y el resultado obtenido

Figura 14. Configuración Principio Oportuno

attribute name	function expressions
Today	<code>date_now()</code>
Difference	<code>(((date_diff(Broadcast_Date,Today))/1000)/86400)/30</code>
Result	<code>if(Difference&lt;1,"Paso","No Paso")</code>

El resultado que se obtiene es el siguiente:

Figura 15. Resultado Principio Oportuno

Row No.	Name Set	Broadcast_Date	Today	Difference	Result
1	CONTRATOS FEBRERO 2016	Mar 3, 2016	Apr 22, 2019 10:33:30 AM COT	38.181	No Paso
2	CONTRATOS ABRIL 2016	May 5, 2016	Apr 22, 2019 10:33:30 AM COT	36.081	No Paso

## 5.4 ACCESIBLE

Tabla 26. Descripción para la medición de un conjunto de datos accesibles

Titulo	Accesibles
Definición	La posibilidad de que la gente tenga acceso a la información disponible, que incluye que es claro y simple
Medición	Las pautas de accesibilidad del contenido web
Alcance	La validación de las pautas de accesibilidad del contenido web (WCAG) 2,0
Unidad de medida	A,AA o AAA

Para la medición de este principio estipulado en la Tabla 14, es necesario:

#### 5.4.1 Medición

- Pautas de Accesibilidad para el contenido Web (WCAG) 2.0. Son una serie de recomendaciones para hacer el contenido de la Web más accesible, para personas con alguna discapacidad física, que dificulte la navegación en la Web, por medio de estas se desea disminuir la brecha que la web pueda generar a cualquier persona que desee acceder a la misma. (W3C, 2016). Las pautas son las siguientes.

**Perceptible:** Toda información y componente que tenga el aplicativo, deben ser simples para que el usuario pueda interactuar con la información allí existente de forma concisa.

**Operable:** La navegación del aplicativo debe ser intuitiva, debe generar al usuario que lo use la facilidad de acceder a las funciones principales de forma eficaz.

**Comprensible:** Todo lo que se encuentre en el aplicativo, para el usuario debe ser fácil de comprender y entender.

**Robusto:** La interoperabilidad del aplicativo, sin importar S.O que se utilice las funcionalidades no debería verse limitadas o restringidas.

- Niveles de Conformidad. Son criterios que debe cumplir o satisfacer un aplicativo con los cuales se evalúa la implementación del sistema frente a los principios o pautas establecidas para medir la accesibilidad

**Nivel A:** Tiene en total 25 criterios, los cuales se enfocan a contenido estructural y visual del aplicativo, lo cual favorezca una navegación simple, rápida y oportuna.

**Nivel AA:** Tiene en total 13 criterios, los cuales están enfocados al acceso multimedia, manejo de sugerencias en la aplicación

**Nivel AAA:** Tiene un total de 23 criterios, los cuales enfatiza en mejorar la experiencia del usuario del aplicativo, dando opciones de ayudas, varios lenguajes

- Método de acceso a la información SODA API. Comprende el Restful Services: SODA API, provee una serie de operaciones tales como: adicionar, actualizar, borrar un set de datos; de igual forma mantenimiento de la metadato y políticas de privacidad, la creación e importación de set de datos.

#### 5.5 PROCESABLE POR MÁQUINA

Titulo	Procesable por Maquina
--------	------------------------

Definición	La posibilidad de que los datos sean manejados por una máquina para análisis
Medición	La posibilidad de que los datos sean manejados por una máquina para análisis
Alcance	El procedimiento realizado para sus datos de extracción, análisis y evaluación
Unit of Measure	Texto

Para la confirmación de este principio; se adicionará un campo, en el cual nos indique que fue extraído exitosamente, lo cual nos indique que el set de datos en este momento se encuentra en la base de datos no relacional. Lo cual nos indique que ya fue procesable por maquina en el proceso de extracción.

*Si (Numero \_Token)!= Vacío entonces  
Resultado: Paso  
Si no:  
Resultado: No Pas*

Se anexa la configuración que se realiza en el motor, de igual forma la respuesta que este genera,

Figura 16. Configuración Principio Procesable por Maquina

attribute name	function expressions
Response	if(Numero_Token=="", "No Paso", "Paso")

La respuesta que se obtiene es la siguiente:

Figura 17. Resultado Principio Procesable por Maquina

Row No.	Response
1	Paso

## 5.6 NO DISCRIMINATORIO

Titulo	No Discriminatorio
Definición	The information can be accessed freely without registration on the website

Medición	El código de respuesta de una petición http
Alcance	Una muestra de 1000 solicitudes al servidor datos.gov.co, compruebe cuántos son correctas
Unidad de Medida	Porcentaje

Medición: Esta no es una métrica en la cual se pueda aplicar un método estadístico y tener una estimación de esta, por tanto, se realizó un desarrollo propio para realizar una serie de peticiones al servidor y tener el número de peticiones que fueron aceptadas.

Código:

```

1. public static int nondiscrimynatory(String token) throws JSONExcept
   ion {
2.     int error=0;
3.     int er=0;
4.     for (int i=1; i<1000; i++) {
5.         String urlQueryString =https://www.datos.gov.co/api
   /views/++token+/rows.json?accessType=DOWNLOAD;
6.         int json = 0;
7.
8.         try {
9.             URL url = new URL(urlQueryString);
10.            HttpURLConnection connection = (HttpURLConnection
   ion) url.openConnection();
11.            connection.setDoOutput(true);
12.            connection.setInstanceFollowRedirects(false);
13.            connection.setRequestMethod("GET");
14.            connection.setRequestProperty("Content-
   Type", "application/json");
15.            connection.setRequestProperty("charset", "utf-
   8");
16.            connection.connect();
17.            json=connection.getResponseCode();
18.
19.            if(json==200) {
20.                error=error+1;
21.            } else {
22.                System.out.println(json);;
23.            }
24.        } catch (IOException ex) {
25.            ex.printStackTrace();
26.        }
27.    }
28.

```

```
29. return error;
```

Con este código realizado en lenguaje java, se realizan 1000 peticiones, al servidor en donde se encuentra el set de datos y se almacenan las peticiones exitosas, las cuales permiten saber la cantidad de conexiones exitosas al servidor, de esta forma se puede medir que la información sea accesible sin ser usuario registrado de la plataforma y sacar el % de éxito de las peticiones realizadas.

*Si (Numero %) >60 entonces  
Resultado: Paso  
Si no:  
Resultado: No Paso*

Se anexa la configuración del motor para la interpretación del desarrollo realizado

Figura 18. Configuración Principio No Discriminatorio

attribute name	function expressions
Result	if((Numero*100/1000)>=60,"No Paso","Paso")

El resultado que se obtiene es el siguiente:

Figura 19. Resultado Principio No Discriminatorio

Row No.	Numero	Result
1	9	Paso

## 5.7 NO PROPIETARIO

Titulo	No propietario
Definición	El conjunto de datos no es propietario, no debe estar bajo restricción
Medición	El nombre del propietario del conjunto de datos
Alcance	Campo en los metadatos, que nos permite identificar al propietario del conjunto de datos
Unidad de Medida	Texto

Medición: De la extracción de los metadatos se obtiene el campo “attribution”, el cual permite conocer el propietario del set de los datos y por consiguiente es quien puede editarlo, actualizarlo o eliminarlo.

La métrica de esta no se cumple, porque tiene un dueño o propietario, aunque permite el acceso a la información de múltiples maneras y el análisis de los datos. De igual manera se anexa la configuración del motor de análisis, para la lectura e interpretación del campo

Figura 20. Configuración No Propietario

attribute name	function expressions
Result	if(Owner=="", "Paso", "No Paso")

Y el resultado de que se obtiene:

Figura 21. Resultado Principio no Propietario

Row No.	Owner	Result
1	Contraloria General de Antioquia	No Paso

## 5.8 LICENCIA ABIERTO

Tabla 27. Definición del licenciamiento abierto

Titulo	Licencia libre
Definición	La información es gratuita, no tiene copyright, patente o privacidad del mismo tipo
Medición	La licencia que protege los derechos de autor del conjunto de datos
Alcance	El set de datos
Unidad de Medida	Texto

Medición: Para la evaluación de este criterio se usará la METADA del set de datos, la cual contiene información sobre el set de datos, con la cual se sabe la licencia que cubre el set de datos extraído, de igual forma, saber que permite realizar sobre el set de datos que cubre.

Esta se divide en 6 tipos de licencia:

1. Atribución
2. Atribución – Sin Derivar
3. Atribución – No comercial – Sin Derivar
4. Atribución – No comercial
5. Atribución – No comercial – Compartir igual
6. Atribución – Compartir igual

El resultado que se obtiene del análisis es el siguiente

Figura 22. Resultado Principio Licencia Abierto

Row No.	License	License ID
1	Creative Commons Attribution   Share Alike 4.0 International	CC_40_BY_SA

## 6. EVALUACIÓN Y ANÁLISIS DE RESULTADOS.

En el capítulo anterior, se estudió la forma teórica en que se accede y se manejan los datos, en este acápite se muestra como principio por principio se validan los datos para el prototipo que se usará, con base en los siguientes sets de datos:

Tabla 28. Evaluación de resultados

ID Set de datos	NOMBRE DEL SET	CANTIDAD DE DATOS
j3bg-66aw	CONTRATOS FEBRERO 2016	12939
mpka-keq9	CONTRATOS ABRIL 2016	11925
ccgi-m8az	CONTRATACION DICIEMBRE 2018	8020

### 6.1 RAPIDMINER

Programa informático, el cual permite realizar análisis de grandes volúmenes de datos y minería de datos, el cual proporciona operadores para el análisis de datos, procesos de entradas y salida; preprocesamiento de datos y visualización, en las

últimas versiones incluye integraciones con diferentes lenguajes como R el cual permite realiza preprocesamiento estadístico y estocástico de la información.

## 6.2 COMPLETO

Como se explicó en la sección anteriormente, por medio del método ListWise se obtendrá la cantidad de datos completos.

Figura 23 Resultado principio Completo

Row No.	% Complete rows
1	0.760

Como se puede observar el porcentaje de datos completos es bajo a la cantidad de información que es recolectada, esto implica que los métodos que usa el propietario del set de datos no satisfacen, la prioridad de tener la información completa, esto perjudica el realizar un análisis o procedimientos con dicha información.

## 6.3 PRIMARIO

Como ya se explicó, por medio de la extracción de los metadatos se obtendrán campos de información del set de datos.

*Si (Información del Propietario del Set) != (Vacio) entonces  
 Resultado: Paso  
 Si no:  
 Resultado: No Paso*

En la tabla se puede observar, información del set de datos tales como su propietario que es aquel que tiene derechos sobre la misma de poder eliminar, actualizar o editar el set, la categoría en la que pertenece y el resultado del seudocódigo anterior.

Figura 24. Resultado Principio Primario

Row No.	Owner	Category	id	Result
1	Contraloria General de Antioquia	Organismos de Control	j3bg-66aw	Paso
2	Contraloría General de Antioquia	Organismos de Control	ccgi-m8az	Paso
3	Contraloria General de Antioquia	Organismos de Control	mpka-keq9	Paso

Por el seudocódigo, se puede concluir que los sets de datos cumplen el principio, el propietario es el encargado de la creación, actualización o eliminación de la información, que esta manipulación debe ser controlada para evitar duplicados, información faltante o fallos en los mismos.

## 6.4 OPORTUNO

Como se explicó, por medio de la extracción de los metadatos se obtendrán campos de información del set de datos.

Se usará la función `date_diff()`, del motor de análisis de datos en el cual se compara la fecha actual y la fecha en la que fueron creados el set de datos.

*date\_diff(emision, Hoy)*

Este es el tiempo que ha pasado desde la emisión del set de datos a la fecha de hoy, se realiza la conversión de (ms) a (meses) y se obtiene el siguiente resultado.

Por medio de este, toda la información de los metadatos selecciona el nombre del set de datos y la fecha en la cual fueron creados, al obtener la información se puede concluir que este principio no se cumple, ya que los meses son superior a 1, lo cual indica el tiempo desde su creación.

Figura 25. Resultado Principio Oportuno

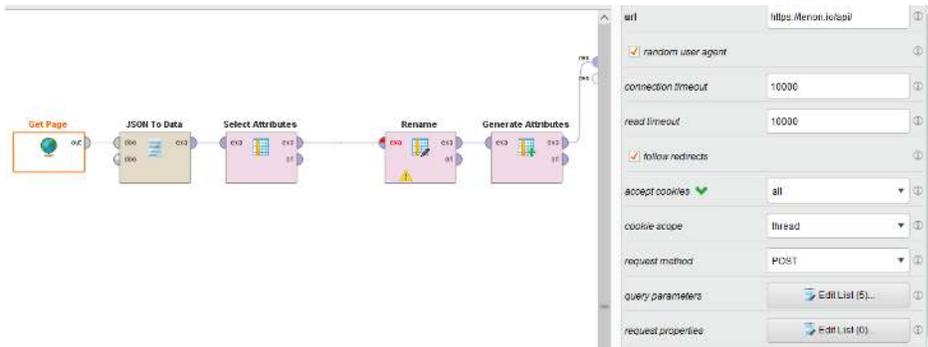
Row ...	Broadcast_...	Name Set	Broadcast_...	Today	Difference	Result
1	2016-03-03	CONTRATOS FEBRERO 2016	Mar 3, 2016	Apr 22, 2019 ...	38.196	No Paso
2	2019/01/15	CONTRATACION DICIEMBRE 2018	Jan 15, 2019	Apr 22, 2019 ...	3.262	No Paso
3	2016-05-05	CONTRATOS ABRIL 2016	May 5, 2016	Apr 22, 2019 ...	36.096	No Paso

Los meses obtenidos del análisis, indican que los sets de datos no se van actualizando como es debido, esto genera una desinformación en el set de datos lo cual limita los análisis y el potencial que tiene la información abierta para su procesamiento y uso de estos.

## 6.5 ACCESIBLE

Por medio del método externo, se obtendrá la calificación del sitio [www.datos.gov.co](http://www.datos.gov.co) frente a los principios de accesibilidad WCAG

Figura 26. Proceso de Accesible



La API la cual permite realizar el análisis de accesibilidad se llama tenon.io, esta admite analizar desde “A,” “AA” o “AAA”, los cuales son los niveles de conformidad del WCAG, así como se observa en la Figura 27.

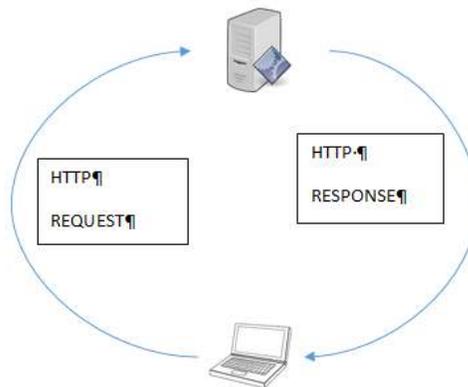
Figura 27. Parámetros para la solicitud de tenon.io

query key	query value
key	f1ccc43a81e543ba71671e07efd26251
url	https://www.datos.gov.co
projectID	DEFAULT_PROJECT
level	AAA
certainty	40

### 6.5.1 Análisis por tipo de acceso a SODA API

Para el acceso de la información desde “datos.gov.co” como se explicó el acceder a la información RESTFUL SERVICES, es como se muestra en la Figura 28.

Figura 28. Restful Services



Para la realización de este análisis, se estudia el código con el cual es posible acceder a la información alojada en la plataforma web “datos.gov.co”. La extracción se realiza por medio del api en la cual se define el dominio del cual se extraen los sets de datos y se definen los tokens de los sets de datos a extraer.

```
Soda2Consumer consumer
    = Soda2Consumer.newConsumer("https://" + domain);
    ClientResponse response = null;
```

```
response = consumer.query(token, HttpLowLevel.JSON_TYPE, "SELECT
    * LIMIT 2000");
```

La implementación de la Soda API, facilita el acceso a la información en la plataforma diseñada por el gobierno colombiano, por tal motivo la aplicación de este principio es válida.

Tabla 29. Resultado Principio Accesible

Fallos	Correctos	Total	Respuesta	Creador
5	69	74	Paso	Contraloría General de Antioquia

El análisis de este principio se divide en dos partes, primero los lineamientos WCAG, lo cual indica que el sitio está adaptado para facilitar el acceso de personas de cualquier tipo, segundo, el API que facilita la plataforma para poder acceder a la información, se encuentran facilidades para ese acceso, pero al ser un API en desarrollo presenta errores, lo cual afecta el desempeño de esta con bugs o momentos de falencia al acceder a la información.

## 6.6 PROCESABLE POR MÁQUINA

Como se explicó en la sección 10.5, por medio de la extracción de la información de la plataforma. Se válida la que es procesable por máquina.

Figura 29. Procesable por máquina



En este proceso se genera el nuevo campo, el cual indica que la información fue procesada correctamente, por consiguiente, el principio se cumple eficazmente.

Figura 30. Resultado Principio Procesable por Maquina

Row No.	Numero_To...	Response
1	j3bg-66aw	Paso
2	mpka-keq9	Paso
3	unsz-yhtr	Paso
4	ut88-hgk8	Paso
5	mtni-n7pq	Paso
6	mg7n-epjz	Paso
7	ccgi-m8az	Paso

La comparación es realizada referente al número de tokens, lo que indica que el set fue ingresado y procesado por máquina para ser adicionado en la base de datos.

*Si (Numero \_Token)≠! Vacio entonces*

*Resultado: Paso*

*Si no:*

*Resultado: No Paso*

El ser procesable por maquina permite identificar que la información se ha procesado de manera automática al ser extraída, esto facilita a futuro los análisis que se realicen por medio de máquinas, ya que se puede acceder a ella.

## 6.7 NO DISCRIMINATORIO

Como se explicó, por medio del código anexo se visualiza la cantidad de peticiones que fueron correctas, por cada uno de los id ingresados para validar.

Figura 31. Resultado Principio No Discriminatorio

Row No.	Numero	Numero_To...	Result
1	9	ccgi-m8az	Paso
2	9	j3bg-66aw	Paso
3	9	mg7n-epjz	Paso
4	9	mpka-keq9	Paso
5	9	mtni-n7pq	Paso
6	9	unsz-yhtr	Paso
7	9	ut88-hgk8	Paso

El acceso a la información es posible sin ser usuario registrado de la plataforma y su validación respectiva es la siguiente:

*Si (Numero %) >60 entonces  
Resultado: Paso  
Si no:  
Resultado: No Paso*

## 6.8 NO PROPIETARIO

Por medio de la extracción de los metadatos, se visualiza el propietario de los sets de datos, este es el que puede actualizarlo, eliminarlo o remplazar en la totalidad toda la información que estos contienen.

Pero, los datos no deben pertenecer a ningún propietario, debe permitirse acceder a ellos, sin tener ningún tipo de restricción, de igual forma, el objetivo de este principio es poder acceder a ellos de manera libre de restricciones.

Figura 32. Resultado Principio No Propietario

Row No.	Owner	Id	Result
1	Contraloria General de Antioquia	j3bg-66aw	No Paso
2	Contraloría General de Antioquia	ccgi-m8az	No Paso
3	Contraloria General de Antioquia	mpka-keq9	No Paso

Al tener un propietario, automáticamente los sets de datos no cumplen el criterio de no tener propietario ya que este es quien puede actualizar, editar o eliminar los campos concretamente.

## 6.9 LICENCIA LIBRE

Por medio de la extracción de los metadatos, se visualiza la licencia que cobija los sets de datos

Figura 33. Resultado Principio Licencia Libre

Row No.	id	License	License ID
1	j3bg-66aw	Creative Commons Attribution   Share Alike 4.0 International	CC_40_BY_SA
2	ccgi-m8az	Creative Commons Attribution   Share Alike 4.0 International	CC_40_BY_SA
3	mpka-keq9	Creative Commons Attribution   Share Alike 4.0 International	CC_40_BY_SA

El principio dice que la información debe estar bajo licencia abierta, para su manipulación. Las licencias Creative Commons son licencias “Copyleft” las cuales buscan proteger quien generó la información. Hay diferentes versiones de estas, las cuales permiten realizar diferentes modificaciones, siempre y cuando se referencie al autor de la información, no es necesaria la remuneración económica al autor; en caso de no cumplir con los criterios de uso de la licencia el autor puede demandar y la ley colombiana lo protege.

Una vez culminado el desarrollo del proyecto, es necesario alojarlo en un servidor para que pueda ser accedido desde cualquier parte. En este caso se alojará en un servidor en Amazon. (Ver Anexo C).

## 7. CONCLUSIONES.

- La implementación del prototipo de software realiza las mediciones parciales pertinentes para efectuar una evaluación de las métricas a evaluar.
- De los resultados obtenidos de un total de 32884 datos del set, solo el 76% de los sets de datos se encuentran completos, esto conlleva a analizar para una futura fase los mecanismos en los cuales se está realizando la recolección y procesamiento de la data, la limitante de no tener la data completa limita las posibilidades de poder manipularla y obtener resultados exitosos en donde esta se pueda usar.
- El desarrollo del prototipo se estructura, con la posibilidad de ingresar diferentes estructuras y así no limitar el proyecto a un solo tipo de estructura de datos, lo cual permite la diversificación de estructuras a analizar y poder analizar diferentes categorías en la colección de datos abiertos del gobierno colombiano.
- La consulta de información referente a los principios de los datos abiertos, para la realización del análisis de estos, presento demasiadas fuentes las cuales generan ambigüedad en la definición de estos, las 7 guías encontradas sobre datos abiertos presentan lineamientos similares por los datos abiertos sin importar la fuente, lo cual dificulta el lineamiento al momento de establecer las fuentes de forma clara.
- La madurez de la implementación del gobierno colombiano frente a políticas de datos abiertos es baja aun por la falta de compromiso de las entidades públicas frente a la necesidad de publicar de forma constante la información, de igual forma, validar el contenido de esta, para así tener información actualizada y eficaz para realizar un análisis a mayor profundidad.

## 8. BIBLIOGRAFÍA

AKRED, JOHN. ¿Qué son los grandes datos? 2014. Disponible: <https://datascience.berkeley.edu/what-is-big-data/>

ALFA-REDI. *Las Bases de Datos*. Lima – Perú. s.f. p. 1. Disponible en: [http://www.alfa-redi.org/sites/default/files/articles/files/ossa\\_0.pdf](http://www.alfa-redi.org/sites/default/files/articles/files/ossa_0.pdf)

ARGIS FOR DESKTOP. ¿Qué son los datos ráster? 2016. Disponible en: <http://desktop.arcgis.com/es/arcmap/10.3/manage-data/raster-and-images/what-is-raster-data.htm>

BERNERS-LEE, T. *Gov 2.0 Expo 2010: Open, linked data for a global community*. 2010. Disponible en: <https://www.youtube.com/watch?v=ga1aSJXCFe0>

BESTEIRO, M. RODRIGUEZ, M. *Web Services* s.f. Disponible: <http://www.ehu.es/mrodriguez/archivos/csharp/pdf/ServiciosWeb/WebServices.pdf>

BIBLIOTECA DEL CONGRESO NACIONAL DE CHILE. *Las cinco estrellas de los datos abiertos*. Consultado en febrero de 2017. Disponible en: <http://datos.bcn.cl/es/informacion/las-5-estrellas>

BOEHM, Barry. *Spiral Development: Experience, Principles, and Refinements: Spiral Development Workshop*. Pittsburgh. 2000. p. 6. Disponible en: <https://www.sei.cmu.edu/reports/00sr008.pdf>

CENTRO LATINOAMERICANO DE ADMINISTRACIÓN PARA EL DESARROLLO – CLAD. *Carta Iberoamericana de Gobierno Abierto. XVII Conferencia Iberoamericana de Ministras y Ministros de Administración Pública y Reforma del Estado*. 2016. (21). p. 24. Disponible en: [https://www.sfp.gov.py/sfp/archivos/documentos/CIGA\\_2016%20final\\_knawtdh9.pdf](https://www.sfp.gov.py/sfp/archivos/documentos/CIGA_2016%20final_knawtdh9.pdf)

CEPAL. *El desafío hacia el gobierno abierto en la hora de la igualdad*. Documento de proyecto. 2012. p. 124. Disponible en: [http://repositorio.cepal.org/bitstream/handle/11362/3969/S2012004\\_es.pdf;jsessionid=061A7C0F0092CAFDD6CA6FAF6F2463F7?sequence=1](http://repositorio.cepal.org/bitstream/handle/11362/3969/S2012004_es.pdf;jsessionid=061A7C0F0092CAFDD6CA6FAF6F2463F7?sequence=1)

CEPAL. *El gobierno electrónico en la gestión pública*. CEPAL. 2011. p. 5. Disponible en:

[http://www.cepal.org/ilpes/noticias/paginas/3/54303/El\\_Gobierno\\_electronico\\_en\\_la\\_gestion\\_Publica.pdf](http://www.cepal.org/ilpes/noticias/paginas/3/54303/El_Gobierno_electronico_en_la_gestion_Publica.pdf)

CONGRESO DE LA REPÚBLICA. *Ley 1712 de marzo 6 de 2014*. “Por medio de la cual se crea la Ley de Transparencia y del Derecho de acceso a la información pública nacional y se dictan otras disposiciones”.

EAVES, D. *Case Study: How Open data saved Canada \$3.2 Billion*. 2010. Disponible en: <https://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse>

IEAG. *A World that Counts*. Mobilising the data revolution for sustainable development. UN Report. 2015.

IGLESIAS, Carlos. *OGP Member Countries are betting on Open Data*. Web Whisperer. 2012. Disponible en: <http://carlosiglesias.es/blog/2012/07/ogp-member-countries-bet-on-open-data/>

LAMBOT, S., SLOB, E., VAN BOSCH, I., STOCKBROECKX, D., & VANCOLOOSTER, M. *Modeling of ground penetrating radar for accurate-characterization of subsurface electric properties*. *IEEE Transactions on Geoscience and Remote Sensing*. 42(11), pp. 2555-2568. Disponible en: <https://doi.org/10.1109/TGRS.2004.834800>

LATHROP, D., & RUMA, L. *Open Government*. 2003

LESSIG, Larry. *8 Principles of Open Government Data* Disponible en: [https://public.resource.org/8\\_principles.html](https://public.resource.org/8_principles.html)

LLINARES, Javier. *Open government: La idea*. 2007. Disponible en: <http://www.javierllinares.es/2007/12/29/open-government-la-idea/>

MINISTERIO DE LAS TECNOLOGÍAS Y LA INFORMACIÓN MINTIC. *Colombia es cuarto en el mundo y primero en América en datos abiertos*. 2015. p. 9. Disponible en: <http://www.mintic.gov.co/portal/604/w3-article-14303.html>

NASER, A., RAMÍREZ, A. *Plan de gobierno abierto*. 2014. Disponible en: [http://www.cepal.org/ilpes/noticias/paginas/3/54303/Plan\\_de\\_Gobierno\\_Abierto.pdf](http://www.cepal.org/ilpes/noticias/paginas/3/54303/Plan_de_Gobierno_Abierto.pdf)

OFFICE OF THE AUSTRALIAN INFORMATION. *Protecting information rights – advancing information policy*. 2012. Disponible en: <https://www.oaic.gov.au/resources/about-us/corporate-information/key-documents/languages/Spanish/OAIC-brochure-Spanish.pdf>

OMPI. *Australia. Freedom of Information Amendment (Reform) Acto 2010*. 2012. p. 7. Disponible en: [http://www.wipo.int/wipolex/es/text.jsp?file\\_id=215466](http://www.wipo.int/wipolex/es/text.jsp?file_id=215466)

OPEN DATA CHARTER. *International Open Data Charter*. 2015 Disponible en: <http://opendatacharter.net/principles-es/>

OPEN DATA CHARTER. *Principios*. 2015 Disponible: <https://opendatacharter.net/principles-es/>

OPEN GOV DATA. *Request for Comments*. 2007 [En línea] Disponible en: [http://wiki.opengovdata.org/index.php/Main\\_Page](http://wiki.opengovdata.org/index.php/Main_Page)

ORGANIZACIÓN PARA LA COOPERACIÓN Y EL DESARROLLO ECONÓMICOS – OCDE. *Un gobierno abierto: fomentar el diálogo con la sociedad civil*. 2003. Disponible en: <http://www.oecd-ilibrary.org/docserver/download/9789264019959-sum-es.pdf?expires=1508516822&id=id&accname=guest&checksum=9612B8688511DEC0C3D669E6FE107848>

OVIEDO, Edgar. *Modelo de madurez para portales de datos abiertos e incorporación a la norma técnica nacional de Costa Rica*. Universidad de Alicante. Tesis Doctoral. Alicante – España. 2016. p. 7.

PROCURADURÍA GENERAL DE LA NACIÓN. *Preguntas frecuentes de la Ley de transparencia y del derecho al acceso a la información 1712 de 2014*. Septiembre 25 de 2014. Disponible en: <https://www.procuraduria.gov.co/portal/media/file/PREGUNTAS.pdf>

ROSADO, Alveiro, RICO, Dewar. *Inteligencia de Negocios: Estado del Arte*. Pereira. Revista Scientia et Technica. Universidad Tecnológica de Pereira. 2010. (44) pp. 321-326

Sitio oficial de Australia para datos abiertos *About – data.gov.au* - 2014

SUNLIGHT FOUNDATION. *Ten Principles for Opening Up Government Information*. Agosto 11 de 2010. Disponible en: <https://sunlightfoundation.com/wp-content/uploads/2016/11/Ten-Principles-for-Opening-Up-Government-Data.pdf>

TIM, B. *Linked Data – Design Issues*. 2009. Disponible en: <https://www.w3.org/DesignIssues/LinkedData.html>

VIVE DIGITAL MINTIC. *Índices internacionales - Estrategia GEL*. 2015. p. 13. Disponible en: <http://estrategia.gobiernoenlinea.gov.co/623/w3-propertyvalue-7653.html>

VIVE DIGITAL MINTIC. *Conoce la Estrategia de Gobierno en Línea*. 2014. Disponible en: <http://estrategia.gobiernoenlinea.gov.co/623/w3-propertyvalue-7650.html>

YU, Harlan., & ROBINSON, David. *The New Ambiguity of "Open Government"*. SSRN Electronic Journal. 2012. Disponible en: <https://www.uclalawreview.org/pdf/discourse/59-11.pdf>

WIKIPEDIA. *Departamentos de Colombia*. 2017. Disponible en: [https://es.wikipedia.org/wiki/Departamentos\\_de\\_Colombia](https://es.wikipedia.org/wiki/Departamentos_de_Colombia)

## 9. ANEXOS

## ANEXO A. MANUAL DEL USUARIO

El siguiente manual tiene como objetivo guiar en el uso del software diseñado para evaluar la madurez de los datos extraídos de cualquier página web que use el API Soda.

Es necesario contar con unos id únicos que se le asignan a cada set de datos. El dominio del cual se extraerá la información será [www.datos.gov.co](http://www.datos.gov.co). Los datos por evaluar serán todos los relacionados con contratos.

### 1. Obtener ID

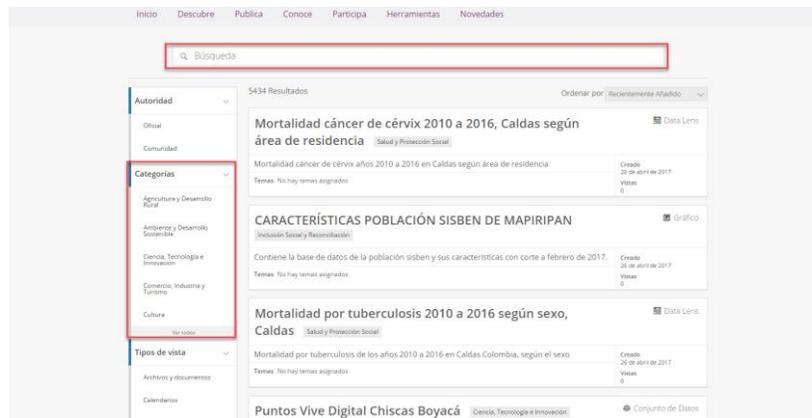
Para obtener los id hay que ir a la siguiente dirección: [www.datos.gov.co](http://www.datos.gov.co) y una vez dentro de la página, ir a la sección **Descubre**.

Figura 34. Anexo A - Datos abiertos



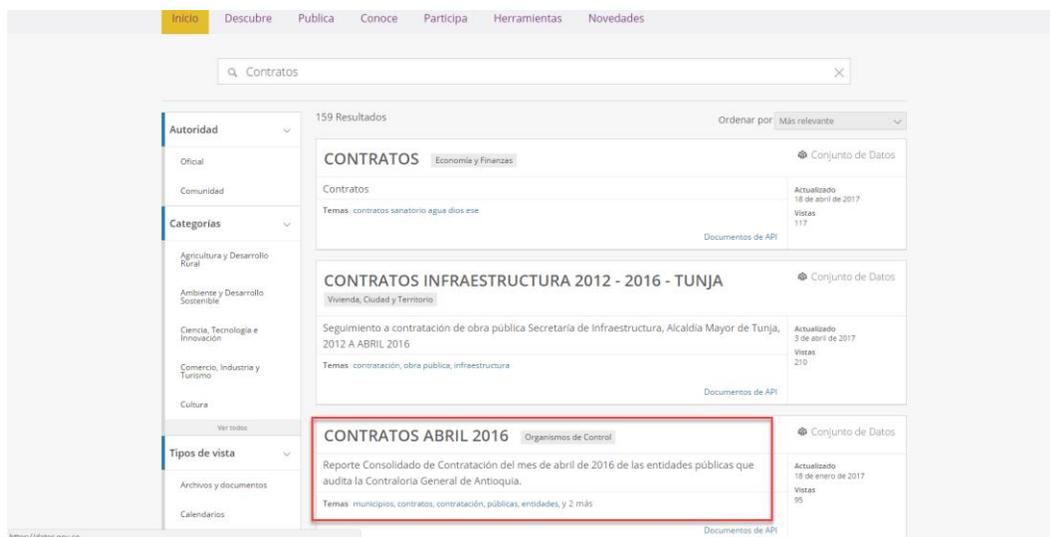
Dentro de la sección **Descubre**, hay que buscar los sets de datos de contratos. La forma de buscar es usando el cuadro de búsqueda que aparece en la página o buscando dentro de las categorías, la categoría **Organismos de control**, como se observa en la figura siguiente:

Figura 35. Anexo A - Set de datos



Luego de buscar todo lo relacionado con contratos, en la lista de resultados se da clic al set de datos de interés.

Figura 36. Anexo A - Set de datos de contratos



Al dar clic en el set de datos, aparece una pantalla como la que se ve a continuación. Estando ahí, hay que ir a la URL y copiar el código que aparece al final de la dirección.

Figura 37. Anexo A - Selección del token

SUJETO DE CONTROL	EVENTO	TIPO DE REGISTRO	CÓDIGO CONTRATO	IDENTIFICACIÓN CONTRATISTA
1 Admon Abajerral	1-Principal	Contrato	063-2016CD	1128437093
2 Admon Abajerral	1-Principal	Contrato	068-2016CD	29109698
3 Admon Abajerral	1-Principal	Contrato	072-2016CD	70726045
4 Admon Abajerral	1-Principal	Contrato	074-2016CD	1022033622
5 Admon Abajerral	1-Principal	Contrato	029-2016MC	21422379
6 Admon Abajerral	1-Principal	Contrato	065-2016CD	1022036133
7 Admon Abajerral	1-Principal	Contrato	066-2016CD	70785250
8 Admon Abajerral	1-Principal	Contrato	093-2016CD	43765760
9 Admon Abajerral	1-Principal	Contrato	030-2016MC	71223412
10 Admon Abajerral	1-Principal	Contrato	032-2016MC	70786350
11 Admon Abajerral	1-Principal	Contrato	033-2016MC	21675598
12 Admon Abajerral	1-Principal	Contrato	069-2016CD	1036621634
13 Admon Abajerral	1-Principal	Contrato	082-2016CD	70786664
14 Admon Abajerral	1-Principal	Contrato	086-2016CD	1022033528
15 Admon Abajerral	1-Principal	Contrato	089-2016CD	1040039343
16 Admon Abajerral	1-Principal	Contrato	092-2016CD	1022034138
17 Admon Abajerral	1-Principal	Contrato	026-2016MC	15379690
18 Admon Abajerral	1-Principal	Contrato	067-2016CD	15453351
19 Admon Abajerral	1-Principal	Contrato	070-2016CD	1027623907
20 Admon Abajerral	1-Principal	Contrato	077-2016CD	1022033865
21 Admon Abajerral	1-Principal	Contrato	078-2016CD	43765760

Luego de copiar el ID del set de datos, hay que ir al software y proceder con la descarga y posterior análisis.

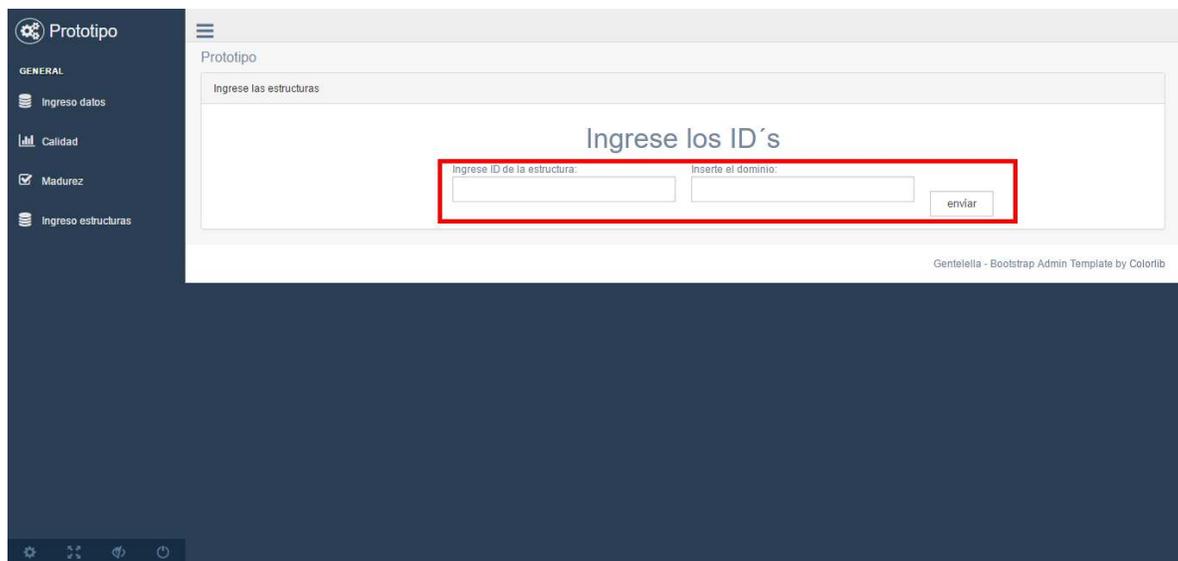
## 2. Definir estructuras

Para poder almacenar los sets de datos, estos deben cumplir con determinadas estructuras para poder ser almacenados. Si al momento de ingresar el set de datos aparece un error, es porque el set no cumple con ninguna estructura ya almacenada. De ser así, es posible ingresar diferentes estructuras.

Para ingresar las estructuras, es necesario contar con los ID obtenidos. Para conseguir los ID, hay que seguir el punto número 1.

Una vez se tiene el ID del set, hay que ir la opción del menú, **Ingreso estructura**. Aquí el procedimiento es similar al proceso de ingresar los sets. En el primer campo, hay que poner el ID que se obtuvo y en el siguiente, hay que poner el dominio de donde está el set de datos.

Figura 38. Anexo A - Ingreso de estructura



Completados los campos, se presiona el botón enviar para así poder ingresar la estructura. Una vez se ingresa la estructura, ya es posible ingresar el set de datos.

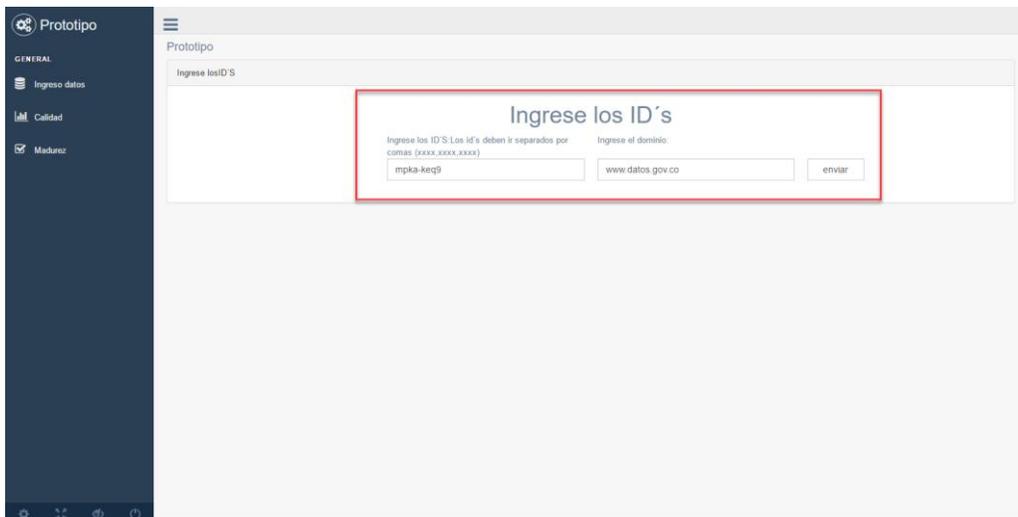
### 3. Uso del prototipo

Teniendo el ID o los ID es necesario ir a la siguiente dirección para acceder al prototipo.

Link: <http://localhost:8080/prototipoWEB/html/ingresoDatos.jsp>

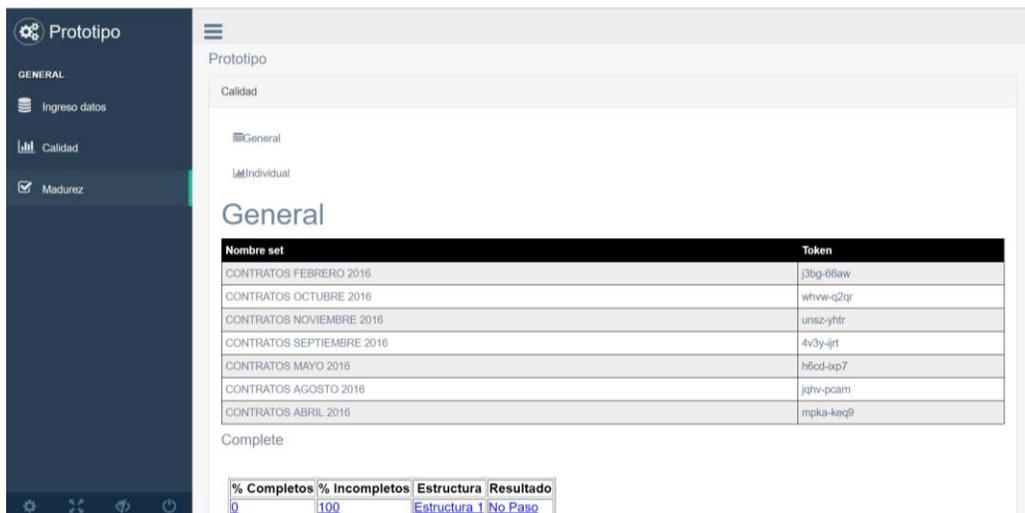
En la primera pantalla, hay dos campos de texto. En el primer campo de texto se ponen los ID obtenidos separados por comas si es más de uno. En el otro campo, hay que poner la dirección web de donde se obtienen los datos. Luego de completar los campos, hay que dar clic en el botón **Enviar**.

Figura 39. Anexo A - Ingreso de datos



Luego de descargar la información, ya se puede ir a ver el análisis realizado por el Software. Para ver el análisis que realiza en cuanto a madurez, hay que ir a la sección **Madurez**. (Observar figura siguiente).

Figura 40. Anexo A - Resultado general



En esta sección es posible ver un reporte general de la información extraída y que está presente en la base de datos. También es posible ver un reporte específico para cada set de datos.

Es posible ver un listado de todos los principios que se analizan en el software con su respectivo resultado

Figura 41. Anexo A - Resultado general 1



Timely

Fecha de emision (aaaa-mm-dd)	Nombre del Set	Resultado
<a href="#">2016-03-03</a>	<a href="#">CONTRATOS FEBRERO 2016</a>	<a href="#">Paso</a>
<a href="#">2016-11-04</a>	<a href="#">CONTRATOS OCTUBRE 2016</a>	<a href="#">Paso</a>
<a href="#">2016-12-12</a>	<a href="#">CONTRATOS NOVIEMBRE 2016</a>	<a href="#">Paso</a>
<a href="#">2016-10-03</a>	<a href="#">CONTRATOS SEPTIEMBRE 2016</a>	<a href="#">Paso</a>
<a href="#">2016-06-05</a>	<a href="#">CONTRATOS MAYO 2016</a>	<a href="#">Paso</a>
<a href="#">2016-09-05</a>	<a href="#">CONTRATOS AGOSTO 2016</a>	<a href="#">Paso</a>
<a href="#">2016-05-05</a>	<a href="#">CONTRATOS ABRIL 2016</a>	<a href="#">Paso</a>

Accessible

Fallos	Correctos	Totales	Respuesta
5	69	74	<a href="#">Paso</a>

Figura 42. Anexo A - Resultado general 2



Timely

Fecha de emision (aaaa-mm-dd)	Nombre del Set	Resultado
<a href="#">2016-03-03</a>	<a href="#">CONTRATOS FEBRERO 2016</a>	<a href="#">Paso</a>
<a href="#">2016-11-04</a>	<a href="#">CONTRATOS OCTUBRE 2016</a>	<a href="#">Paso</a>
<a href="#">2016-12-12</a>	<a href="#">CONTRATOS NOVIEMBRE 2016</a>	<a href="#">Paso</a>
<a href="#">2016-10-03</a>	<a href="#">CONTRATOS SEPTIEMBRE 2016</a>	<a href="#">Paso</a>
<a href="#">2016-06-05</a>	<a href="#">CONTRATOS MAYO 2016</a>	<a href="#">Paso</a>
<a href="#">2016-09-05</a>	<a href="#">CONTRATOS AGOSTO 2016</a>	<a href="#">Paso</a>
<a href="#">2016-05-05</a>	<a href="#">CONTRATOS ABRIL 2016</a>	<a href="#">Paso</a>

Accessible

Fallos	Correctos	Totales	Respuesta
5	69	74	<a href="#">Paso</a>

Figura 43. Anexo A - Resultado específico 1

Non-Proprietary

Propietario	ID	Resultado
Contraloria General de Antioquia	<a href="#">j3bg-66aw</a>	No Paso
Contraloria General de Antioquia	<a href="#">whvw-q2qr</a>	No Paso
Contraloria General de Antioquia	<a href="#">unsz-yhtr</a>	No Paso
Contraloria General de Antioquia	<a href="#">4v3y-ijrt</a>	No Paso
Contraloria General de Antioquia	<a href="#">h6cd-ixp7</a>	No Paso
Contraloria General de Antioquia	<a href="#">jqhv-pcam</a>	No Paso
Contraloria General de Antioquia	<a href="#">mpka-keq9</a>	No Paso

License Free

ID	Nombre de la Licencia	Id de la licencia
<a href="#">j3bg-66aw</a>	<a href="#">Creative Commons Attribution   Share Alike 4.0 International</a>	<a href="#">CC 4.0 BY SA</a>
<a href="#">whvw-q2qr</a>	<a href="#">Creative Commons Attribution   Share Alike 4.0 International</a>	<a href="#">CC 4.0 BY SA</a>
<a href="#">unsz-yhtr</a>	<a href="#">Creative Commons Attribution   Share Alike 4.0 International</a>	<a href="#">CC 4.0 BY SA</a>
<a href="#">4v3y-ijrt</a>	<a href="#">Creative Commons Attribution   Share Alike 4.0 International</a>	<a href="#">CC 4.0 BY SA</a>

Figura 44. Anexo A - Resultado específico 2

Non-Discriminatory

Numero	Numero_Token	Resultado
9	<a href="#">j3bg-66aw</a>	Paso
9	<a href="#">whvw-q2qr</a>	Paso
9	<a href="#">unsz-yhtr</a>	Paso
9	<a href="#">4v3y-ijrt</a>	Paso
9	<a href="#">h6cd-ixp7</a>	Paso
9	<a href="#">jqhv-pcam</a>	Paso
9	<a href="#">mpka-keq9</a>	Paso

Non-Proprietary

Propietario	ID	Resultado
Contraloria General de Antioquia	<a href="#">j3bg-66aw</a>	No Paso
Contraloria General de Antioquia	<a href="#">whvw-q2qr</a>	No Paso
Contraloria General de Antioquia	<a href="#">unsz-yhtr</a>	No Paso
Contraloria General de Antioquia	<a href="#">4v3y-ijrt</a>	No Paso

## ANEXO B. MANUAL INSTALACIÓN RAPIDMINER SERVER

### 1. Instalación de RapidMiner Server

Descargar RapidMiner Server

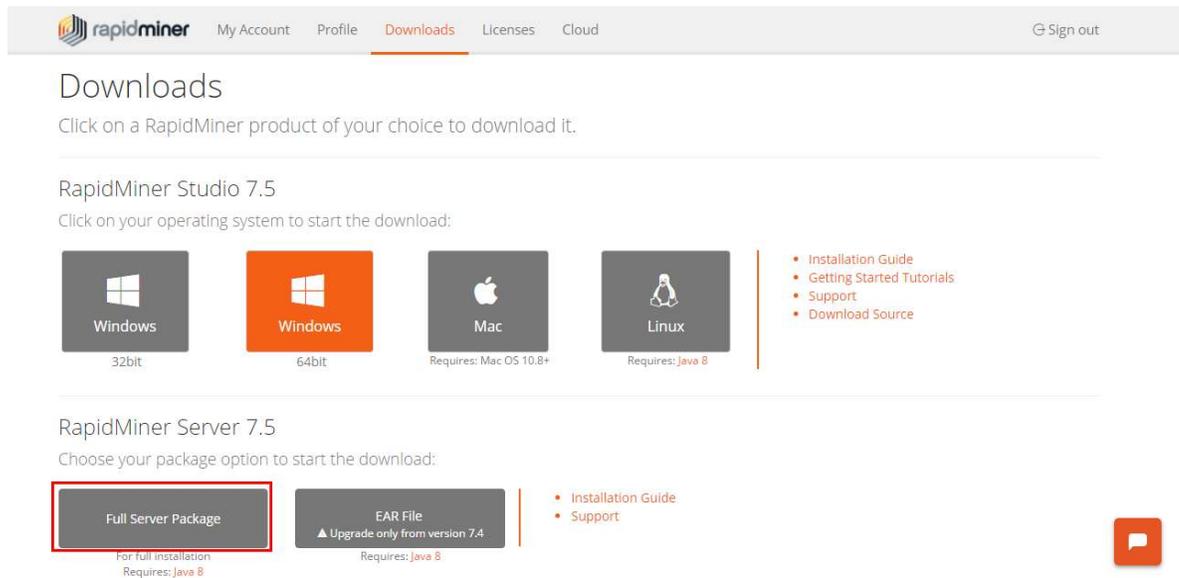
Para la instalación de RapidMiner Server, lo primero que hay que hacer es descargar los archivos necesarios para la instalación. Para ello hay que ir al siguiente link.

<https://my.rapidminer.com/nexus/account/index.html#downloads>

Para descargar estos archivos es necesario estar registrado y haber iniciado sesión, de lo contrario no será posible descargarlos.

Una vez dentro hay que ir a la sección de RapidMiner Server y dar clic en **Full Server Package**.

Figura 45. Anexo B - Descarga RapidMiner



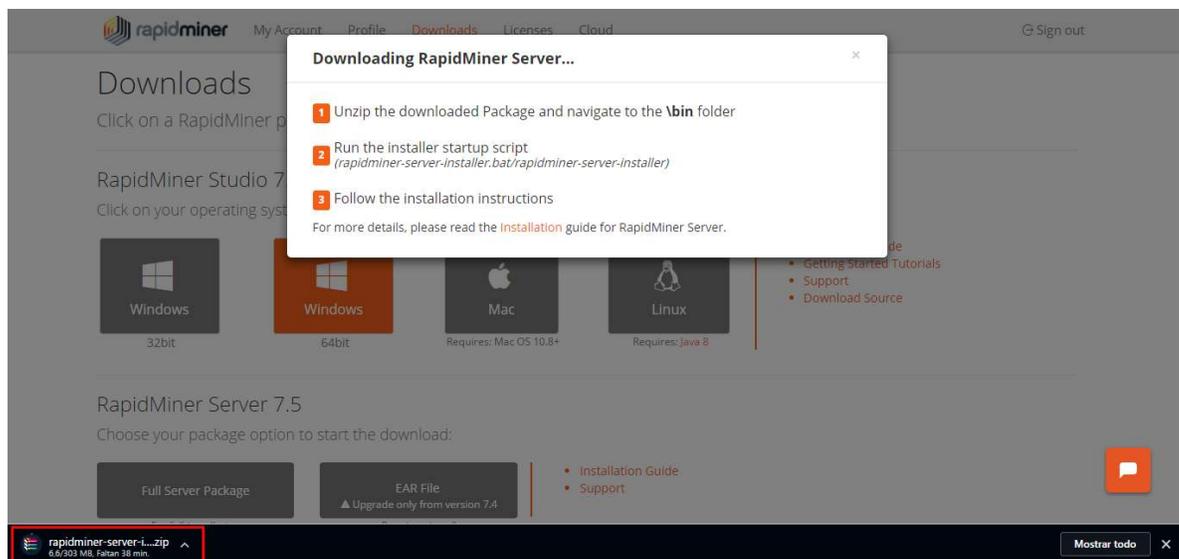
Luego de dar clic en el botón, la siguiente pantalla será para aceptar los términos de la licencia.

Figura 46. Anexo B - Terminos RapidMiner



Enseguida de aceptar los términos aparecer una nueva pantalla y en seguida iniciará la descarga.

Figura 47. Anexo B - Inicio de la descarga



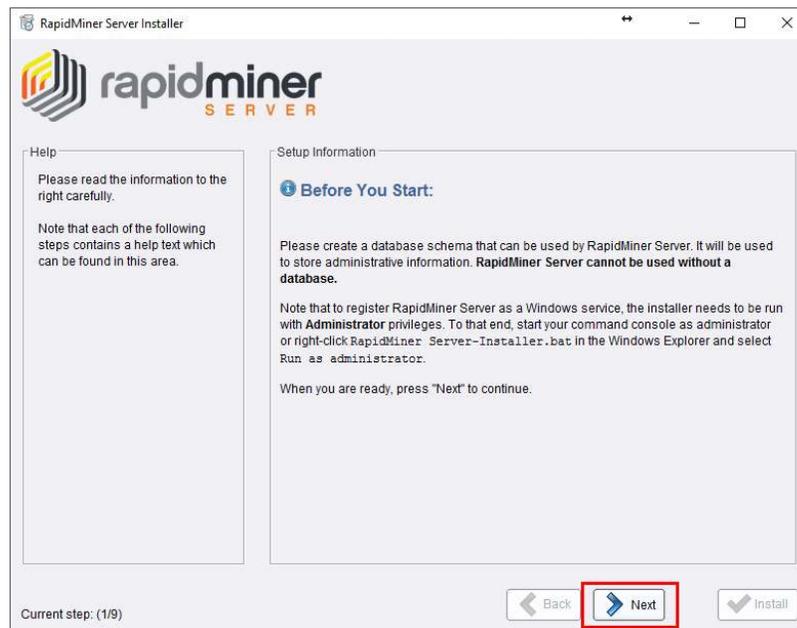
Al terminar la descarga, lo siguiente que hay que hacer es descomprimir el archivo descargado del cual se obtendrá una lista de archivos.

bin	28/04/2017 12:53 ...	Carpeta de archivos	
data	28/04/2017 12:53 ...	Carpeta de archivos	
lib	28/04/2017 12:53 ...	Carpeta de archivos	
EULA.txt	28/10/2016 2:46 p....	Documento de tex...	24 KB

Ahora hay que entrar en la carpeta **bin** y luego ejecutar el archivo **rapidminer-server-installer.bat**.

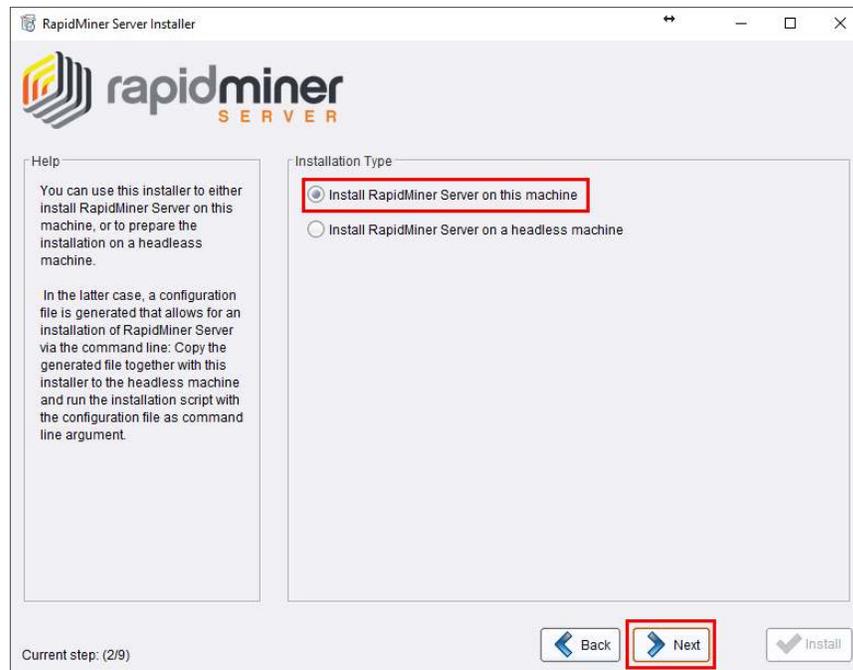
En la siguiente pantalla lo que hay que hacer es dar clic en el botón **Next**.

Figura 48. Anexo B - Instalación RapidMiner 1



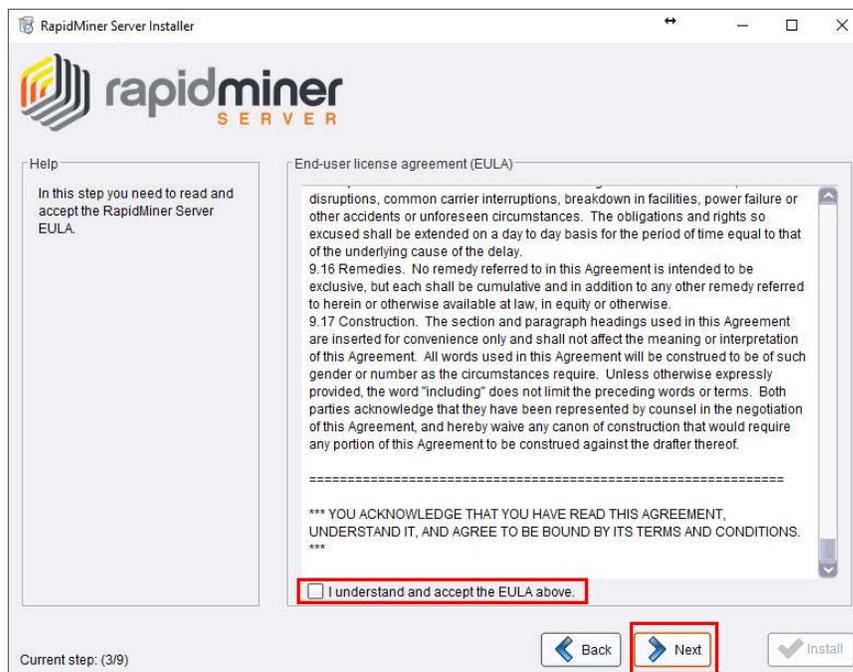
En la siguiente pantalla se escoge la primera opción y luego clic en el botón **Next**.

Figura 49. Anexo B - Instalación RapidMiner 2



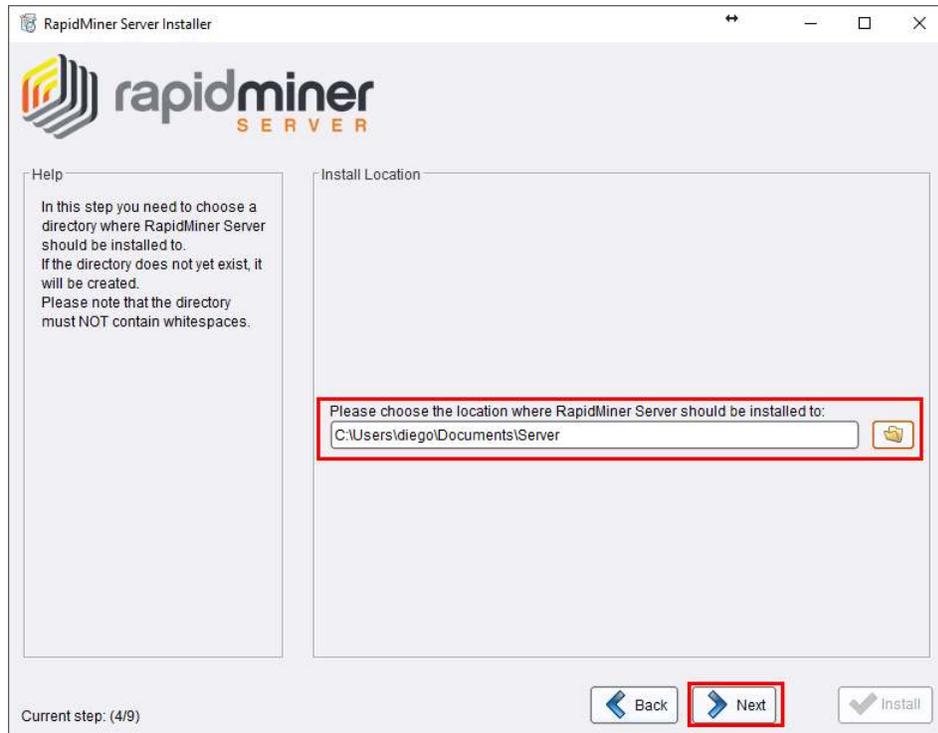
Lo siguiente que hay que hacer es dar clic en el check y luego hay que dar clic en el botón **Next**.

Figura 50. Anexo B - Instalación RapidMiner 3



Ahora hay que escoger la ubicación donde se instalara el servidor.

Figura 51. Anexo B - Ubicación de la instalación

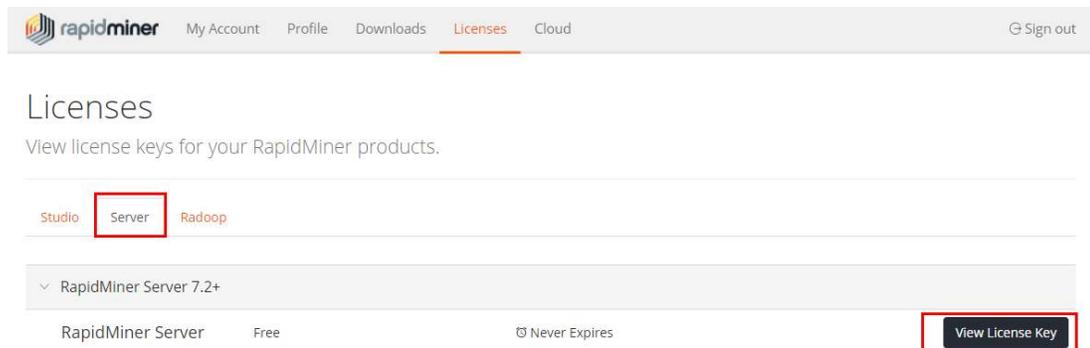


El siguiente paso es copiar la licencia, la cual se encuentra en la página en el siguiente link.

<https://my.rapidminer.com/nexus/account/index.html#licenses/rapidminer-studio>

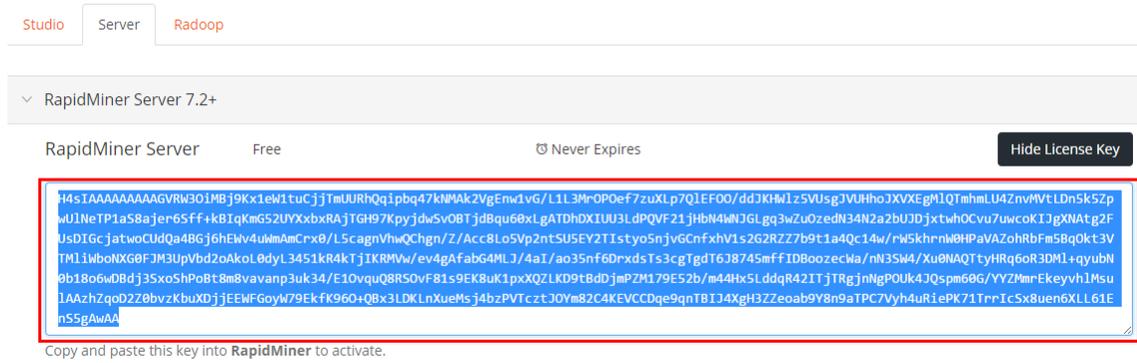
Una vez dentro de la página hay que ir a la pestaña del **Server**>>**View License Key**.

Figura 52. Anexo B - Licencia RapidMiner 1



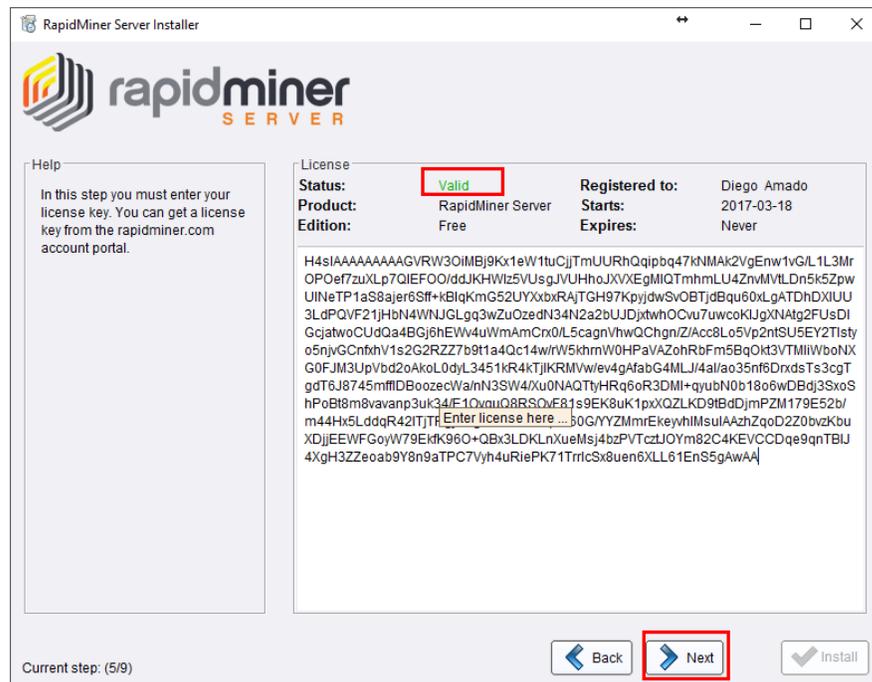
Luego de dar clic en el botón aparecerá la licencia, la cual hay que copiar e ir de nuevo a la instalación.

Figura 53. Anexo B - Licencia RapidMiner 2



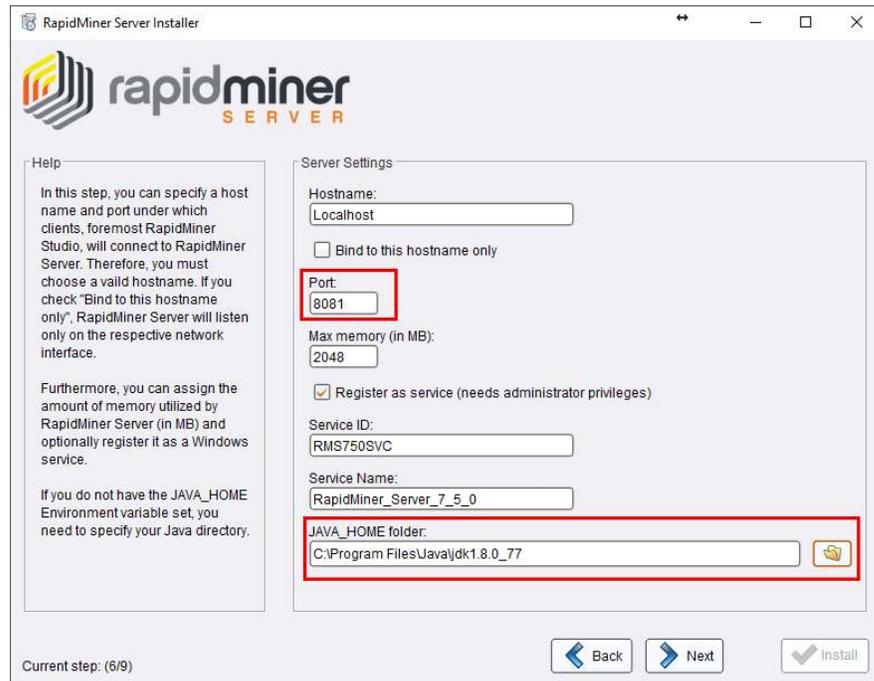
Luego de copiar la licencia, se debe pegar en el espacio presente en la siguiente pantalla y una vez se hace esto se puede ver que la licencia ha sido validada y se presiona el botón **Next**.

Figura 54. Anexo B - Licencia RapidMiner 3



En siguiente paso es definir el puerto del servidor y seleccionar la ruta donde se ha instalado java.

Figura 55. Anexo B - Configuración de puertos



Lo siguiente es asignar una base de datos MySQL y el driver de conexión.

Figura 56. Anexo B - Configuración de la base de datos



En la siguiente pantalla solo hay que dar clic en **Next** porque no se configurara servicio de email.

Hay que dejar todo por efecto y clic en finalizar.

Figura 57. Anexo B - Finalizar instalación

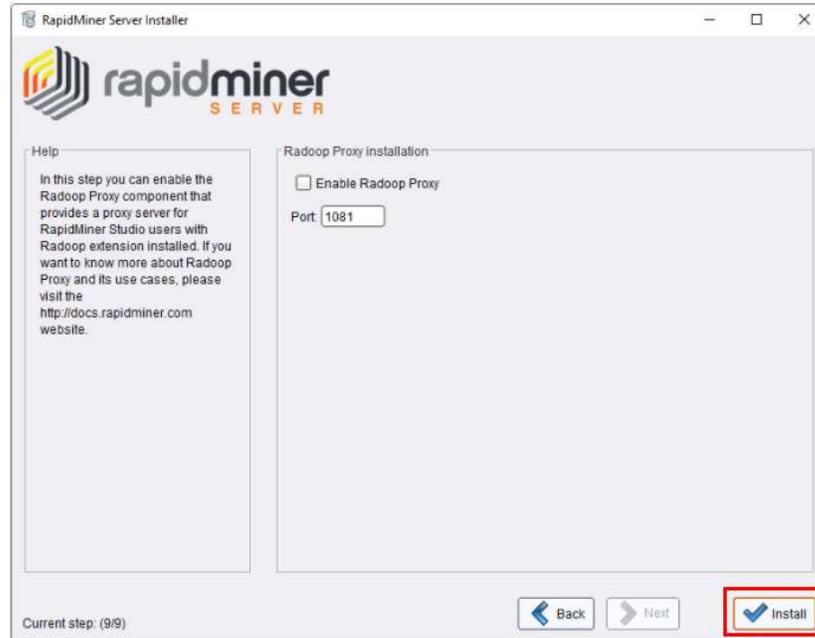
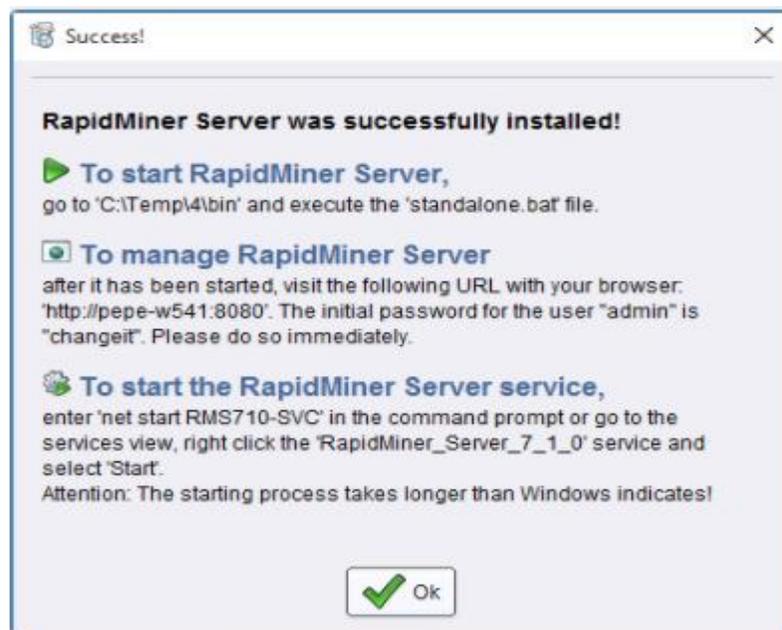


Figura 58. Anexo B - Instalación exitosa

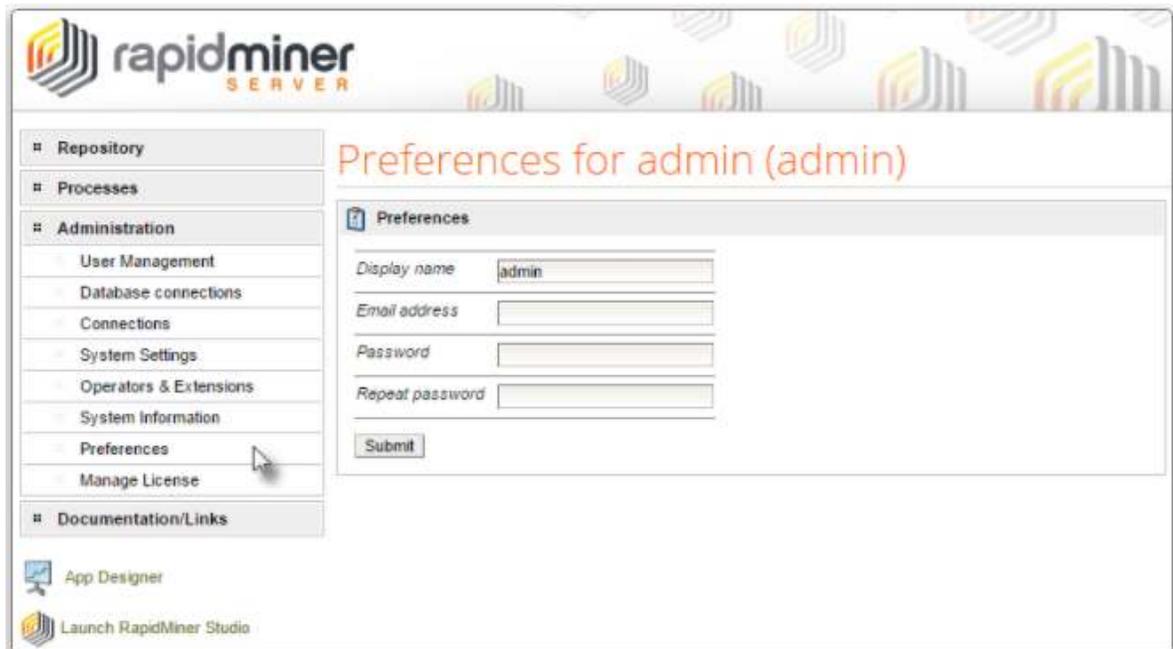


Ahora solo hay que pasar a configurar. Las credenciales por defecto son las siguientes:

- Usuario: admin
- Contraseña: changeit

Se recomienda cambiar la contraseña para mayor seguridad.

Figura 59. Anexo B - RapidMiner en el servidor



## ANEXO C. CREACIÓN DE INSTANCIA

Creación de instancia en Amazon AWS

### 1. Creación de cuenta

Si no se tiene una cuenta, es necesario crearla, para ello hay que ir al siguiente link <https://aws.amazon.com/es/>. Estando en la página, hay que ir al botón **Abre una cuenta gratuita**.

Figura 60. Anexo C - Sitio Amazon AWS



Al presionar el botón, aparece una pantalla donde se tendrá que poner el correo con el cual se quiere crear la cuenta y se selecciona la opción **Soy usuario nuevo** y se presiona el botón.

Figura 61. Anexo C – Creación de usuario 1.

amazon  
webservices

Inicia sesión o crea tu cuenta AWS

¿Cuál es su correo electrónico o teléfono móvil?

Mi e-mail o teléfono móvil:

cuentanueva@hotmail.com

Soy un usuario nuevo.

Soy usuario registrado y mi contraseña es:

Identificarse (servidor seguro)

¿Has olvidado tu contraseña?

Amazon Aurora Base de datos de clase empresarial a una décima parte del coste.

"10 veces más rápida que nuestro entorno de MySQL. ¡Simplemente funcional!"

Alfresco

Más información

El siguiente formulario que aparece debe ser diligenciado para completar el registro y una vez se llenen todos los campos, se presiona el botón **Crear cuenta**.

Figura 62. Anexo C - Creación de usuario 2

amazon  
webservices

### Credenciales de inicio de sesión

Utilice el formulario siguiente para crear credenciales de registro que podrá utilizar para AWS y Amazon.com.

Mi nombre es:

Mi dirección de e-mail es:

Escribela de nuevo:

Nota: ésta es la dirección de correo electrónico que usamos para comunicarnos consigo sobre su cuenta

Escribe una contraseña nueva:

Escribela de nuevo:

Con esto ya se tiene la cuenta creada. Para ingresar a la cuenta, hay que ir al link ya mencionado y se va a la opción **Mi cuenta**>>**Consola de administración de AWS**.

Figura 63. Anexo C - Consola AWS



En este caso, se ponen las credenciales y se selecciona la opción **Soy usuario registrado y mi contraseña es** y se da clic en el botón.

Figura 64. Anexo C - Inicio de sesión

amazon  
webservices

### Inicio de sesión o crea tu cuenta AWS

¿Cuál es su correo electrónico o teléfono móvil?

Mi e-mail o teléfono móvil:

Soy un usuario nuevo.

Soy usuario registrado y mi contraseña es:

[¿Has olvidado tu contraseña?](#)

Amazon Aurora | Base de datos de clase empresarial a una décima parte del coste.

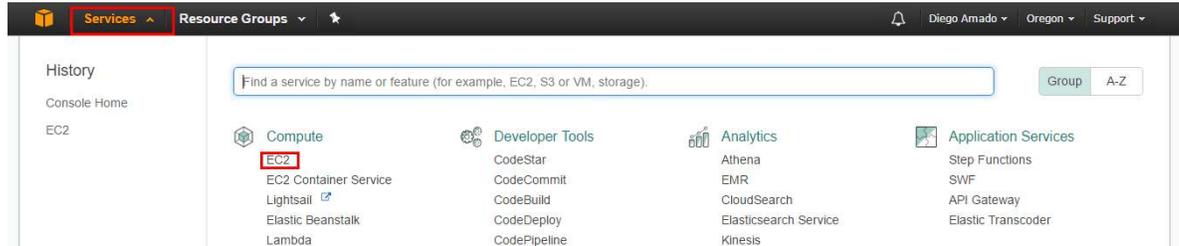
"10 veces más rápida que nuestro entorno de MySQL. ¡Simplemente funcional!"

Alfresco

## 2. Creación de instancia

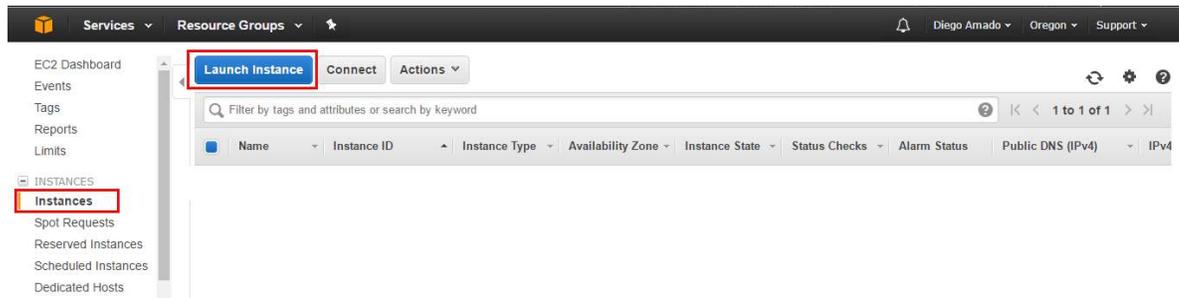
Después de crear la cuenta e ingresar, aparece una nueva pantalla en la que hay que ir a la opción de **Services>>EC2**.

Figura 65. Anexo C - Creación de instancia EC2



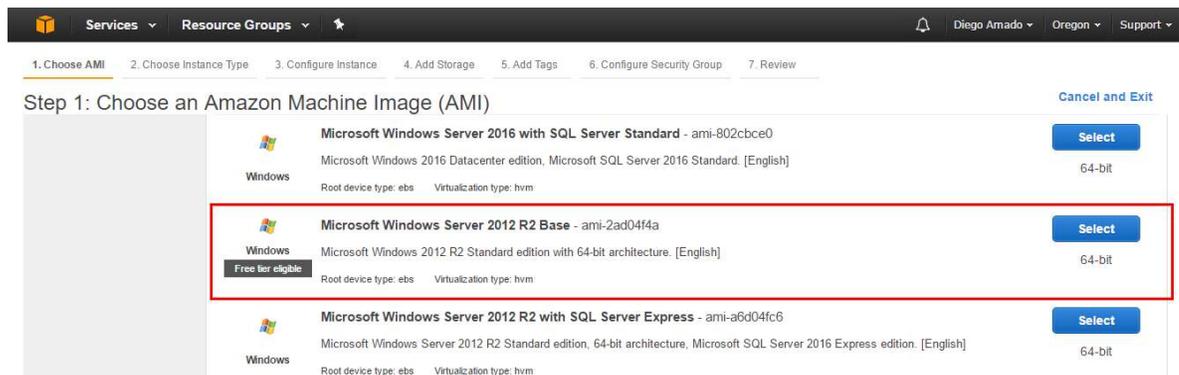
Estando en **EC2** hay que ir a **Instances>>Launch Instance**.

Figura 66. Anexo C - Iniciando instancia



El primer paso es seleccionar el sistema operativo que se necesite de la lista. EL que se usara para desplegar el proyecto es **Microsoft Windows Server 2012 R2**. Cuando se encuentre el sistema operativo elegido, se presiona el botón **Select**.

Figura 67. Anexo C - Selección de sistema operativo



En las siguientes cuatro pantallas se deja todo por defecto. En esta que es una capa gratuita la capacidad que dan es de 30 GB.

Figura 68. Anexo C - Configuración de instancia 1

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

### Step 2: Choose an Instance Type

Amazon EC2 provides a wide selection of instance types optimized to fit different use cases. Instances are virtual servers that can run applications. They have varying combinations of CPU, memory, storage, and networking capacity, and give you the flexibility to choose the appropriate mix of resources for your applications. [Learn more about instance types and how they can meet your computing needs.](#)

Filter by: All instance types Current generation [Show/Hide Columns](#)

Currently selected: t2.micro (Variable ECUs, 1 vCPUs, 2.5 GHz, Intel Xeon Family, 1 GiB memory, EBS only)

	Family	Type	vCPUs	Memory (GiB)	Instance Storage (GiB)	EBS-Optimized Available	Network Performance	IPv6 Support
<input type="checkbox"/>	General purpose	t2.nano	1	0.5	EBS only	-	Low to Moderate	Yes
<input checked="" type="checkbox"/>	General purpose	t2.micro <small>Free tier eligible</small>	1	1	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	General purpose	t2.small	1	2	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	General purpose	t2.medium	2	4	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	General purpose	t2.large	2	8	EBS only	-	Low to Moderate	Yes

[Cancel](#) [Previous](#) [Review and Launch](#) [Next: Configure Instance Details](#)

Figura 69. Anexo C - Configuración de instancia 2

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

### Step 3: Configure Instance Details

Configure the instance to suit your requirements. You can launch multiple instances from the same AMI, request Spot instances to take advantage of the lower pricing, assign an access management role to the instance, and more.

**Number of instances**  [Launch into Auto Scaling Group](#)

**Purchasing option**  Request Spot instances

**Network**  [Create new VPC](#)

**Subnet**  [Create new subnet](#)

**Auto-assign Public IP**

**Domain join directory**  [Create new directory](#)

**IAM role**  [Create new IAM role](#)

**Shutdown behavior**

**Enable termination protection**  Protect against accidental termination

[Cancel](#) [Previous](#) [Review and Launch](#) [Next: Add Storage](#)

Figura 70. Anexo C - Configuración de instancia 3

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

### Step 4: Add Storage

Your instance will be launched with the following storage device settings. You can attach additional EBS volumes and instance store volumes to your instance, or edit the settings of the root volume. You can also attach additional EBS volumes after launching an instance, but not instance store volumes. [Learn more](#) about storage options in Amazon EC2.

Volume Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Throughput (MB/s)	Delete on Termination	Encrypted
Root	/dev/sda1	snap-09a34d5f6fdb68ed2	30	General Purpose SSD (GP2)	100 / 3000	N/A	<input checked="" type="checkbox"/>	Not Encrypted

[Add New Volume](#)

Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage. [Learn more](#) about free usage tier eligibility and usage restrictions.

[Cancel](#) [Previous](#) [Review and Launch](#) [Next: Add Tags](#)

Figura 71. Anexo C - Configuración de instancia 4

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

### Step 5: Add Tags

A tag consists of a case-sensitive key-value pair. For example, you could define a tag with key = Name and value = Webservers. A copy of a tag can be applied to volumes, instances or both. Tags will be applied to all instances and volumes. [Learn more](#) about tagging your Amazon EC2 resources.

Key (127 characters maximum)	Value (255 characters maximum)	Instances	Volumes
This resource currently has no tags.			
Choose the Add tag button or click to add a Name tag. Make sure your IAM policy includes permissions to create tags.			

[Add Tag](#) (Up to 50 tags maximum)

[Cancel](#) [Previous](#) [Review and Launch](#) [Next: Configure Security Group](#)

En el siguiente paso se habilitan los puertos que se usaran. En este caso los puertos que se habilitarán serán los siguientes:

- Apache: 80.
- Apache tomcat: 8080.
- Rapidminer: 8081.
- MongoDB: 27017.
- MySql: 3306.

Para agregar un nuevo puerto hay que dar clic en **New Rule** y al dar clic se habilita otro espacio en donde hay que escoger el tipo y el puerto.

Figura 72. Anexo C - Reglas de seguridad 1

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

**Step 6: Configure Security Group**  
A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow Internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more](#) about Amazon EC2 security groups.

**Assign a security group:**  Create a new security group  
 Select an existing security group

**Security group name:**   
**Description:**

Type	Protocol	Port Range	Source
RDP	TCP	3389	Custom 0.0.0.0/0

**Add Rule**

**Warning**  
Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only.

[Cancel](#) [Previous](#) [Review and Launch](#)

Figura 73. Anexo C - Reglas de seguridad 2

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

**Security Group**  
Firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server each your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more](#) about Amazon EC2 security groups.

**Assign a security group:**  Create a new security group  
 Select an existing security group

**Security group name:**   
**Description:**

Type	Protocol	Port Range	Source
Custom TCP Rule	TCP	8080	Custom 0.0.0.0/0
Custom TCP Rule	TCP	8080	Custom 0.0.0.0/0

**Add Rule**

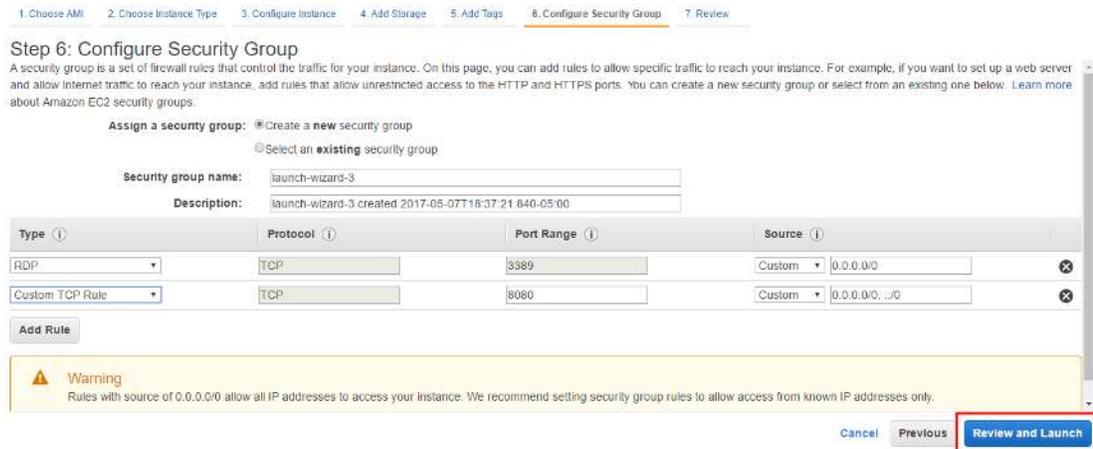
**Warning**  
Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only.

[Cancel](#) [Previous](#) [Review and Launch](#)

Esto hay que repetirlo hasta habilitar todos los puertos mencionados.

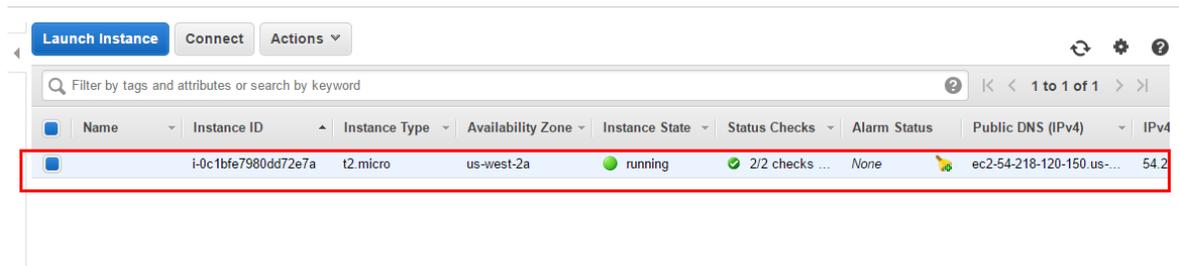
Una vez habilitados los puertos solo hay que dar clic en **Review and Launch** para iniciar la instancia.

Figura 74. Anexo C - Reglas de seguridad 2



Lo siguiente que aparecerá es la instancia creada y esta tardara un tiempo en ejecutarse.

Figura 75. Anexo C - Estado de la instancia

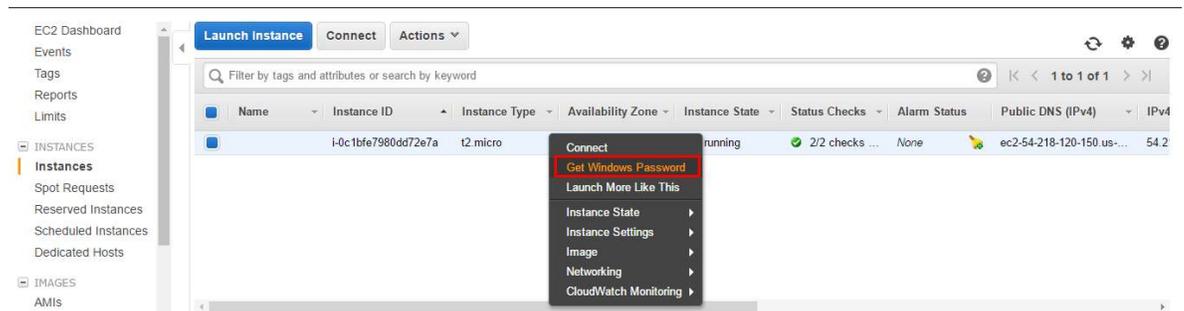


### 3. Obtener clave para acceso remoto

Una vez se ha ejecutado la instancia lo siguiente que hay que hacer es obtener las credenciales para acceder al servidor por escritorio remoto.

Para ello hay que ir a la instancia y dar clic derecho y seleccionar la opción **Get Windows Password**.

Figura 76. Anexo C - Generar acceso al servidor



Se busca el archivo .pem para poder descryptar la clave y se presiona el botón **Decrypt Password** y enseguida se generan las credenciales para el acceso remoto.

## ANEXO D. INSTALACIÓN BASE DE DATOS Y TOMCAT

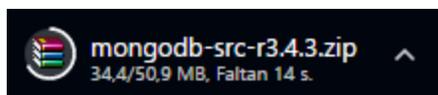
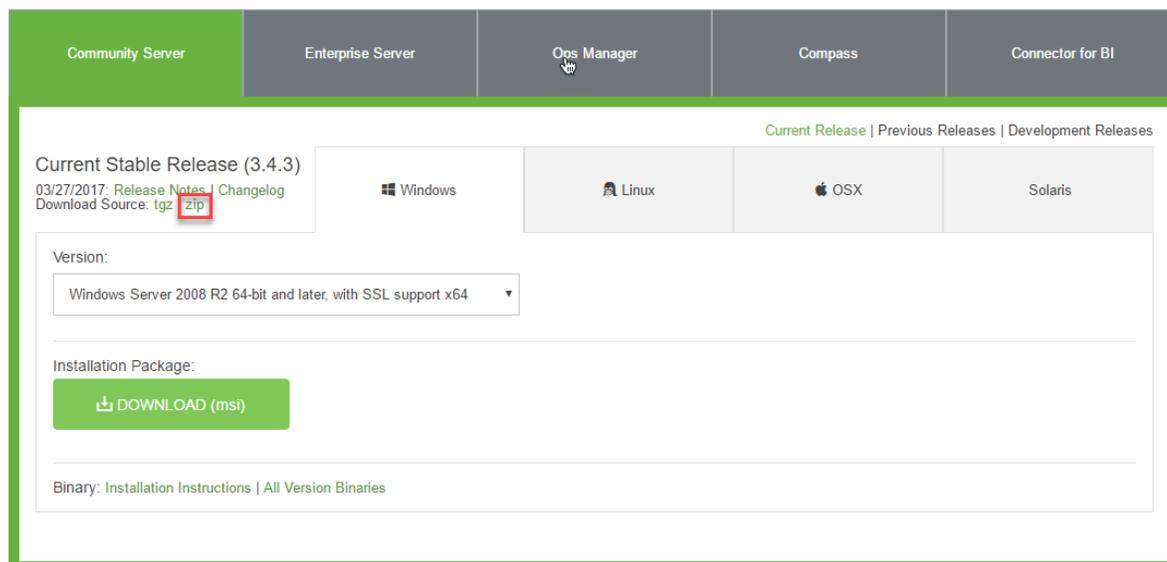
### 1. MongoDB

MongoDB es el motor de base de datos que se usa para el desarrollo del proyecto. Para la instalación es muy sencilla, lo primero que hay que hacer es descargar los archivos, para ello hay que ir al siguiente link:

<https://www.mongodb.com/download-center?jmp=nav#community>

Luego de entrar en el link, en la pantalla que aparece hay que ir a descargar los archivos en un comprimido. Hay que dar clic en el link "zip" que está resaltado en la siguiente imagen. Una vez se da clic, comenzará la descarga.

Figura 77. Anexo D - Descarga MongoDB



Luego de que se termine la descarga hay que descomprimir el archivo. La descompresión genera una carpeta con el nombre de **mongodb-src-r3.4.3** y dentro de esta carpeta hay otra con el mismo nombre y a esta es necesario cambiar el nombre y dejar solo **mongodb**.

Luego de cambiar el nombre es necesario que esta carpeta se ubique en el disco **C** en la carpeta **archivos de programa**.

Una vez este la carpeta en la nueva ubicación, en el disco **C** se crea una nueva carpeta con el nombre de **data** y dentro de esta carpeta crear otra con el nombre de **db**.

Figura 78. Anexo D - Ubicación de los datos 1

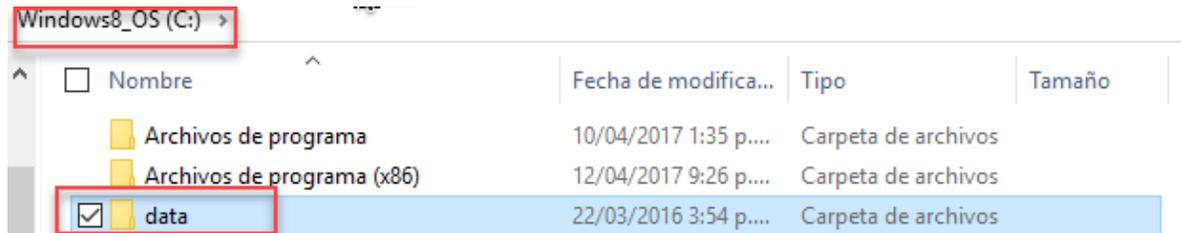
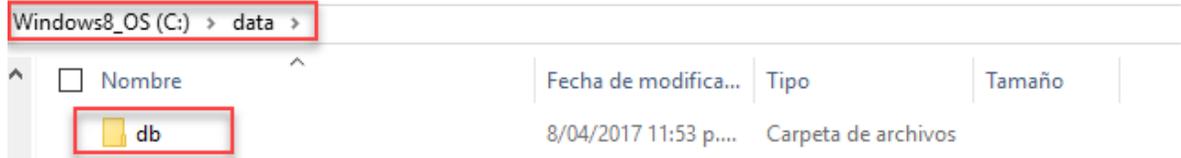


Figura 79. Anexo D - Ubicación de los datos 2



Con la creación de estas carpetas se tiene lo necesario para la utilización de **mongoDB** como base de datos.

## 2. Ejecutar Servidor

Para ejecutar el servidor, es necesario el uso de del CMD (En caso de Windows). Una vez se tenga abierto el CMD, hay que desplazarse a la ubicación donde se guardó la carpeta de mongodb y dentro, seguir la siguiente ruta:

```
C:\Program Files\MongoDB\Server\3.2\bin
```

Para desplazarse hasta dicha ruta es necesario usar el comando **cd**. Una vez estando en la ubicación, es necesario ejecutar el archivo **mongod.exe**.

Figura 80. Anexo D - Ejecutando servidor de base de datos

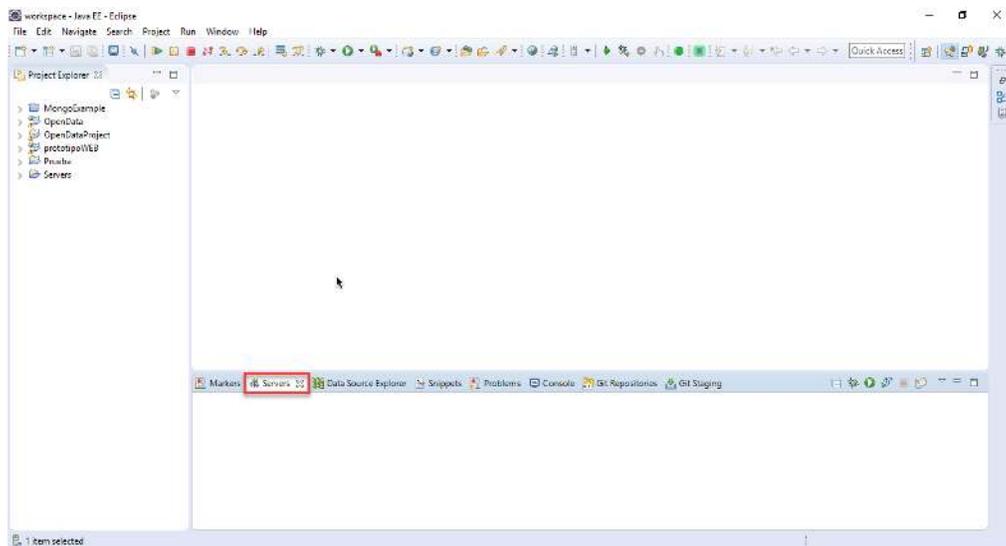
```
Símbolo del sistema - mongod.exe
C:\Program Files\MongoDB\Server\3.2\bin>mongod.exe
2017-04-16T17:24:37.298-0500 I CONTROL [initandlisten] MongoDB starting : pid=7976 port=27017 dbpath=C:\data\db\ 64-bit
host=DiegoAmado
2017-04-16T17:24:37.301-0500 I CONTROL [initandlisten] using the x86_64 Windows version of MongoDB
2017-04-16T17:24:37.303-0500 I CONTROL [initandlisten] Detected data files in C:\data\db\ created by the 'wiredTiger' s
storage engine, so setting the active storage engine to 'wiredTiger'.
2017-04-16T17:24:37.305-0500 I CONTROL [initandlisten] git version: e2ee9ffc9f95a94fad76802e28cc978718bb7a30
2017-04-16T17:24:37.307-0500 I CONTROL [initandlisten] allocator: tcmalloc
2017-04-16T17:24:37.309-0500 I CONTROL [initandlisten] modules: none
2017-04-16T17:24:37.311-0500 I CONTROL [initandlisten] build environment:
2017-04-16T17:24:37.316-0500 I CONTROL [initandlisten] Detected data files in C:\data\db\ created by the 'wiredTiger' s
storage engine, so setting the active storage engine to 'wiredTiger'.
2017-04-16T17:24:37.318-0500 I CONTROL [initandlisten] Detected unclean shutdown - C:\data\db\mongod.lock is not empty.
2017-04-16T17:24:37.320-0500 I CONTROL [initandlisten] Detected unclean shutdown - C:\data\db\mongod.lock is not empty.
2017-04-16T17:24:37.351-0500 I - [initandlisten] Detected unclean shutdown - C:\data\db\mongod.lock is not empty.
2017-04-16T17:24:37.357-0500 W - [initandlisten] Detected unclean shutdown - C:\data\db\mongod.lock is not empty.
2017-04-16T17:24:37.364-0500 W STORAGE [initandlisten] Recovering data from the last clean checkpoint.
2017-04-16T17:24:37.388-0500 I STORAGE [initandlisten] wiredtiger_open config: create,cache_size=2G,session_max=20000,e
viction=(threads_max=4),config_base=false,statistics=(fast),log=(enabled=true,archive=true,path=journal,compressor=snapp
y),file_manager=(close_idle_time=100000),checkpoint=(wait=60,log_size=2GB),statistics_log=(wait=0),
2017-04-16T17:24:39.371-0500 I NETWORK [HostnameCanonicalizationWorker] Starting hostname canonicalization worker
2017-04-16T17:24:39.371-0500 I FTDC [initandlisten] Initializing full-time diagnostic data capture with directory 'C
:\data\db\diagnostic.data'
2017-04-16T17:24:39.431-0500 I NETWORK [initandlisten] waiting for connections on port 27017
```

De esta manera ya se puede utilizar mongo como base de datos. El puerto por el que escucha es el 27017.

### 3. Descarga e instalación de Apache Tomcat

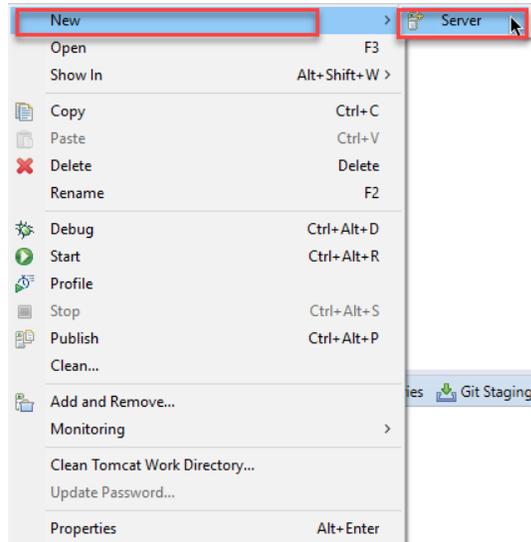
Apache tomcat es el servidor que se va a utilizar para alojar la aplicación que se está desarrollando. Una forma de descargar Apache Tomcat es utilizando el IDE Eclipse. Entando en el IDE, es necesario ir a la pestaña **Servers**.

Figura 81. Anexo D - Configuración de Tomcat 1



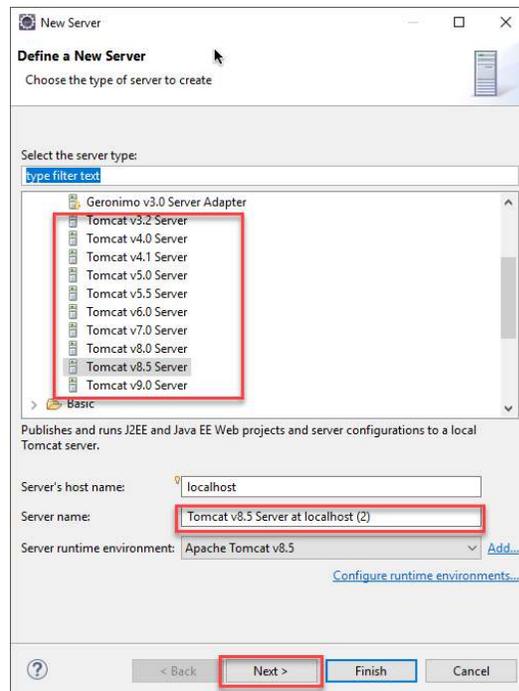
Estando en la pestaña se clic derecho y dentro de las opciones que se despliegan, se escoge New>>Server.

Figura 82. Anexo D - Configuración de Tomcat 2



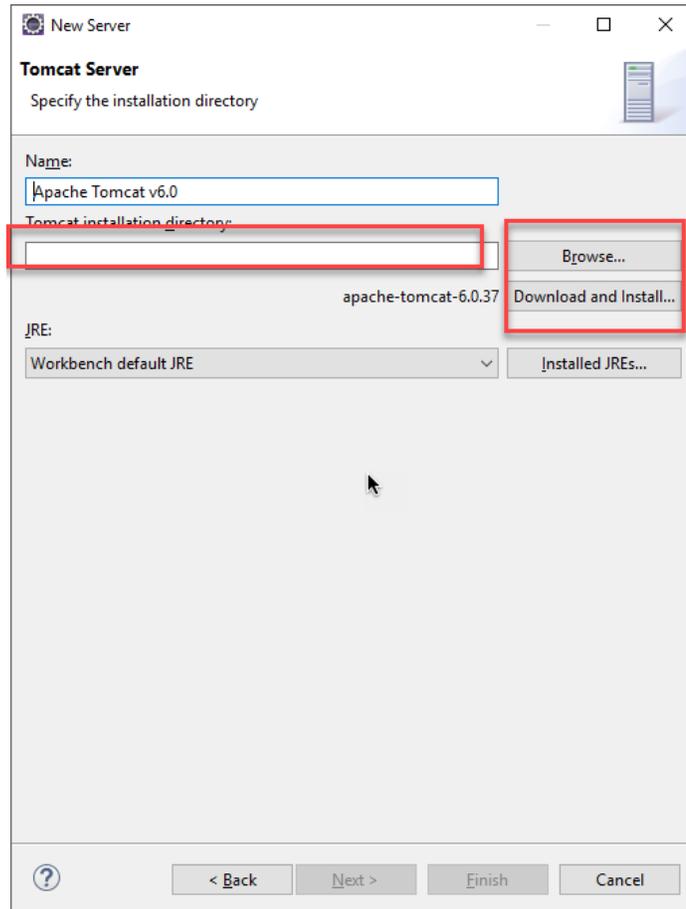
En la siguiente ventana que aparece, se debe escoger la versión que se desea instalar y se le asigna un nombre al servidor y en seguida se da clic en **next**.

Figura 83. Anexo D - Selección de la versión de Tomcat



Una vez se presiona el botón, esto muestra otra ventana, en donde se podrá escoger la ruta donde se guardó el Tomcat si ya se había descargado o de lo contrario da la opción de descargar los archivos con solo presionar el botón **Download and Install**. Una vez se hace esto, comenzara la descarga del servidor.

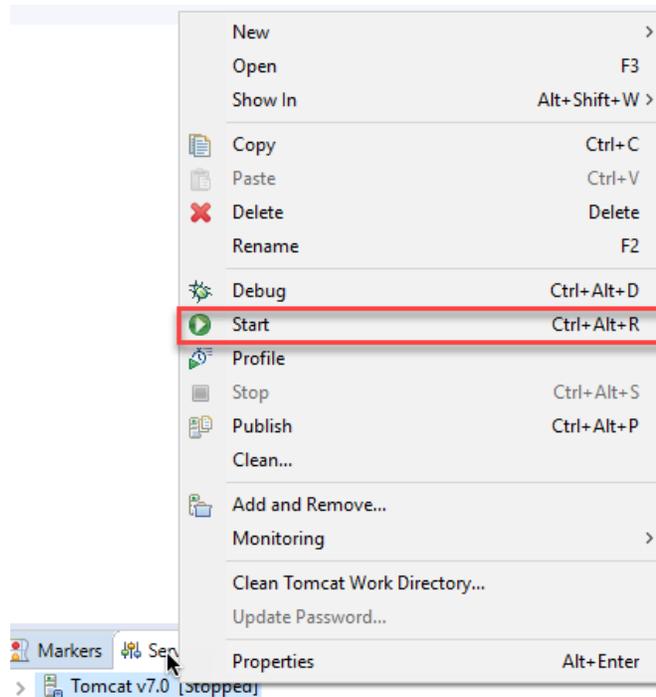
Figura 84. Anexo D - Descarga y selección de binarios



Cuando se han descargado los archivos solo hay que dar clic en el botón de **Finish** y ya quedara instalado el servidor.

Para ejecutarlo solo hay que ir a la pestaña de **Servers** y dar clic derecho en el servidor y clic en **Start**.

Figura 85. Anexo D - Ejecutar Tomcat



Cuando se ejecuta el servidor lo siguiente que se puede ver información acerca de la ejecución de este.

Figura 86. Anexo D - Consola Tomcat

```
Tomcat v8.5 Server at localhost [Apache Tomcat] C:\Program Files\Java\jre1.8.0_121\bin\javaw.exe (16/04/2017, 7:19:02 p.m.)
abr 16, 2017 7:19:09 PM org.apache.catalina.startup.VersionLoggerListener log
INFORMACIÓN: Command line argument: -Djava.endorsed.dirs=C:\xampp\tomcat\tomcat-8.5.13\endorsed
abr 16, 2017 7:19:09 PM org.apache.catalina.startup.VersionLoggerListener log
INFORMACIÓN: Command line argument: -Dfile.encoding=Cp1252
abr 16, 2017 7:19:09 PM org.apache.catalina.core.AprLifecycleListener lifecycleEvent
INFORMACIÓN: La biblioteca nativa de Apache Tomcat basada en ARP que permite un rendimiento óptimo en entornos de desarrollo no ha sido hallada en java
abr 16, 2017 7:19:11 PM org.apache.coyote.AbstractProtocol init
INFORMACIÓN: Initializing ProtocolHandler ["http-nio-8080"]
abr 16, 2017 7:19:14 PM org.apache.tomcat.util.net.NioSelectorPool getSharedSelector
INFORMACIÓN: Using a shared selector for servlet write/read
abr 16, 2017 7:19:14 PM org.apache.coyote.AbstractProtocol init
INFORMACIÓN: Initializing ProtocolHandler ["ajp-nio-8009"]
abr 16, 2017 7:19:14 PM org.apache.tomcat.util.net.NioSelectorPool getSharedSelector
INFORMACIÓN: Using a shared selector for servlet write/read
abr 16, 2017 7:19:14 PM org.apache.catalina.startup.Catalina load
INFORMACIÓN: Initialization processed in 6986 ms
abr 16, 2017 7:19:14 PM org.apache.catalina.core.StandardService startInternal
INFORMACIÓN: Arrancando servicio Catalina
abr 16, 2017 7:19:14 PM org.apache.catalina.core.StandardEngine startInternal
INFORMACIÓN: Starting Servlet Engine: Apache Tomcat/8.5.13
abr 16, 2017 7:19:18 PM org.apache.jasper.servlet.TldScanner scanJars
INFORMACIÓN: Al menos un JAR, que se ha explorado buscando TLDs, aún no contenía TLDs. Activar historial de depuración para este historiadore para una c
abr 16, 2017 7:19:18 PM org.apache.coyote.AbstractProtocol start
INFORMACIÓN: Starting ProtocolHandler ["http-nio-8080"]
abr 16, 2017 7:19:18 PM org.apache.coyote.AbstractProtocol start
INFORMACIÓN: Starting ProtocolHandler ["ajp-nio-8009"]
abr 16, 2017 7:19:18 PM org.apache.catalina.startup.Catalina start
INFORMACIÓN: Server startup in 4444 ms
```