

PROYECTO DE INVESTIGACIÓN  
DATOS INFORMATIVOS

TIPO DE CONVOCATORIA

Proyecto Interno  Proyecto Semilla  Proyecto Junior  Proyecto Multi e Interdisciplinario

Fecha de presentación (dd/mm/aa): 26/07/18

Título del proyecto:

Semántica embebida con JSON-LD para generar información enriquecida en resultados de búsqueda de Recursos Educativos Abiertos

TIPOS DE INVESTIGACIÓN

Investigación básica

Investigación aplicada

DEPARTAMENTO(S) Y/O INSTITUCIÓN:

1. DICC

LÍNEA(S) DE INVESTIGACIÓN (verificable en el SAEW):

1. DICC A4-L3 Sistemas de Información

RESUMEN DE INFORMACIÓN DEL DIRECTOR Y COLABORADORES

Director

Apellidos y nombres	No. de Cédula	HSS	Departamento	Título de mayor nivel y mención.	No. ORCID
Navarrete Rueda Rosa del Carmen	1001303849	10	DICC	PhD en Informática	0000-0002-1435-5000

Codirector (Se aplica para todos los proyectos, el codirector será a su vez colaborador)

Apellidos y nombres	No. de Cédula	HSS	Departamento	Título de mayor nivel y mención.	No. ORCID
Montenegro Armas Carlos Estalesmit	1704448818	6	DICC	MSc. en Ciencias de la Computación e Informática	0000-0002-6909-4215

Colaboradores

Apellidos y nombres	No. de identificación	HSS	Institución	Título de mayor nivel y mención.	No. ORCID
Luján Mora Sergio	AAK002366 (pasaporte)	-----	Universidad de Alicante	PhD en Informática	

\* HSS = Horas Semana Semestre

## PROYECTO DE INVESTIGACIÓN

Proyecto Interno  Proyecto Semilla  Proyecto Junior  Proyecto Multi e Inter Disciplinario

Investigación Básica  Investigación Aplicada

**DEPARTAMENTO(S) Y/O INSTITUTOS:**  
1. Departamento de Informática y Ciencias de la Computación

**LINEA(S) DE INVESTIGACIÓN:**  
1. DICC A4-L3 Sistemas de Información  
2.

<b>CAMPO DEL CONOCIMIENTO (Ver Anexo A: Detalle de los campos del conocimiento)</b>		
Campo amplio	Campo detallado	Campo específico
Información y Comunicación (TIC)	Información y Comunicación (TIC)	Software y desarrollo y análisis de aplicativos

<b>DISCIPLINA CIENTÍFICA (Marque X, solamente una opción)</b>	
Ciencias Naturales y Exactas	
Ingeniería y Tecnologías	X
Ciencias Médicas	
Ciencias Agrícolas	
Ciencias Sociales	
Humanidades	

<b>OBJETIVO SOCIOECONÓMICO (Marque X, solamente una opción)</b>	
Exploración y explotación del medio terrestre	
Ambiente	
Exploración y explotación del espacio	
Transporte, telecomunicaciones y otras infraestructuras	
Energía	
Producción y tecnología industrial	
Salud	
Agricultura	
Educación	X
Cultura, ocio, religión y medios de comunicación	
Sistemas políticos y sociales, estructuras y procesos	
Defensa	
Avance general del conocimiento: I+D financiada con los Fondos Generales de Universidades (FGU)	
Avance general del conocimiento: I+D financiados con otras fuentes	

<b>Alcance Territorial (Marque X, solamente una opción)</b>			
Institucional		Nacional	
Parroquial		Internacional	X
Cantonal		No definido	
Provincial			



<b>1 Proyecto de Investigación</b>
<b>Título</b> (mínimo 10 palabras): <b>Semántica embebida con JSON-LD para generar información enriquecida en resultados de búsqueda de Recursos Educativos Abiertos</b>
<b>Resumen del proyecto</b> (máximo 200 palabras) <p>La UNESCO promueve el uso de Recursos Educativos Abiertos (REA) para apoyar una educación de calidad inclusiva y equitativa a nivel global. La disponibilidad de REA en la Web crece constantemente; sin embargo, sus potenciales usuarios no buscan REA en sitios web específicos, sino directamente en la Web usando motores de búsqueda como Google o Bing. Los resultados desplegados por los motores de búsqueda no exhiben información relevante (p.ej., características educativas, licencia de propiedad intelectual o propiedades de accesibilidad). Esto constituye una barrera para que los usuarios encuentren los REA que requieren y, consecuentemente, desalienta su uso. Para superar esta limitación se plantea embeber anotaciones semánticas en REA, que puedan ser interpretadas por los motores de búsqueda para desplegar información enriquecida. En este proyecto se propone definir un conjunto de descriptores para el marcado semántico embebido de REA con el uso de la especificación JSON-LD y vocabulario de Schema y otros estándares complementarios. La recomendación tomará como base los descriptores de Schema más utilizados en la Web, definidos a partir de un análisis cuantitativo del despliegue de marcado con JSON-LD en corpus de rastreo web de gran escala. Esta investigación contribuirá con una primera perspectiva sobre el estado de adopción de JSON-LD el campo educativo y con base al conjunto de descriptores, se desarrollará un prototipo que automatice la generación de marcado semántico con JSON-LD para REA.</p>
<b>Palabras clave</b> (4-6):  JSON-LD, Schema, marcado semántico embebido, Recursos Educativos Abiertos, análisis de corpus de rastreo web

<b>2 Objetivos, limitaciones, hipótesis y resultados esperados de esta propuesta de investigación</b>
---

### 2.1 Objetivos

#### 2.1.1 Objetivo General

Desarrollar un esquema de semántica embebida con JSON-LD y vocabulario estandarizado para enriquecer información de resultados de búsqueda de Recursos Educativos Abiertos.

#### 2.1.2 Objetivos Específicos

- Analizar la extensión de uso de la especificación JSON-LD mediante el análisis cuantitativo de corpus de rastreo web.
- Identificar los descriptores del vocabulario Schema relevantes al contexto educativo, de mayor uso en los corpus de rastreo web.
- Definir una recomendación para el marcado semántico embebido de REA utilizando JSON-LD y vocabulario seleccionado de Schema o complementario a Schema.
- Desarrollar un prototipo generador automático de scripts para el marcado de REA con JSON-LD.

#### 2.2 Limitaciones (Aspectos que quedan fuera del alcance del Proyecto de Investigación)



- a. El presente proyecto no abarca la evaluación del impacto del uso del marcado semántico propuesto, debido a que se necesitarían grandes volúmenes de publicación a nivel global y seguimiento de la evolución en años sucesivos.
- b. El presente proyecto no incluye el despliegue de la aplicación que genera el script de marcado, sino únicamente la construcción del prototipo.

### 2.3 Hipótesis (Responden al problema de investigación)

Los usuarios de REA que busquen dichos recursos en la Web podrán visualizar información relevante a datos generales, valor educativo, licencia y características de accesibilidad de los recursos, a partir de información extraída por los motores de búsqueda desde anotaciones semánticas embebidas que utilicen JSON-LD.

### 2.3 Detalle de los resultados esperados (con relación a los objetivos)

- a. Resultados del uso del marcado en datos estructurados que usen JSON-LD con base en el análisis cuantitativo de corpus de rastreo web de gran escala.
- b. Resultados del vocabulario de Schema más utilizado, relevante al contexto educativo, con base en el análisis cuantitativo de corpus de rastreo web de gran escala.
- c. Recomendación de las propiedades del vocabulario de Schema y de otros vocabularios complementarios, para el marcado semántico de REA utilizando JSON-LD.
- d. Prototipo de una aplicación que genere scripts JSON-LD para el marcado de REA con base en las propiedades recomendadas.

<b>3</b>	<b>Relevancia de la propuesta de investigación y su relación con la(s) líneas de investigación</b>
----------	--

#### Relevancia de la propuesta de investigación

El World Wide Web Consortium (W3C) [1], organismo rector de la Web, genera recomendaciones y estándares que aseguran el crecimiento de la Web a largo plazo. La finalidad de estos estándares es que sean acogidos en forma global y aplicados en todos los dominios sobre los cuáles se publica en la Web. Sin embargo, en la realidad, las personas que publican contenidos en la Web no tienen necesariamente el conocimiento para hacer uso de tales estándares y tecnologías [2]. Por tal razón, es relevante poner dichos estándares y tecnologías al alcance de los usuarios finales a través de aplicaciones que faciliten su uso.

En este proyecto, se propone que la tecnología referida al uso de anotaciones semánticas embebidas en el contenido web pueda ser utilizada por los usuarios que producen contenido educativo en Recursos Educativos Abiertos. Para ello, se propone crear una aplicación que permita a los usuarios generar el código de marcado embebido en JSON-LD, sin requerir conocimientos de esta tecnología sino únicamente del contexto educativo del contenido que publican.

Dado que JSON-LD es una especificación establecida como recomendación de la W3C desde enero de 2014 [3], es relevante conocer en qué medida ha sido adoptado por la comunidad y en qué dominios se está utilizando. Para ello, en esta investigación se parte del análisis del despliegue de esta tecnología en corpus web de gran escala provenientes de rastreo (*crawling*) web, lo que asegura que la muestra del volumen de páginas web analizadas sea representativa [4].

Por lo anotado, la relevancia de esta investigación está en dirimir la pertinencia de la aplicación de una tecnología web reconocida como recomendación de la W3C, en el contexto de la educación.

Por otra parte, el potencial uso de los REA para expandir las oportunidades de educación amplía los horizontes para el intercambio de conocimientos y la colaboración entre educadores, instituciones y países. La UNESCO recomienda la integración de los REA en apoyo al Objetivo de educación de calidad [5], como parte de los Objetivos de desarrollo del milenio propuesto por las Naciones Unidas en la Agenda 2030 [6] y que ha sido adoptada por gobiernos en todo el mundo, incluido el gobierno de Ecuador, como política pública [7].

Uno de los desafíos que obstaculizan la integración de los REA es la capacidad de los usuarios para encontrar, reutilizar, crear y compartir REA [8], por lo cual, el contexto para el cuál se realiza esta investigación, es relevante.



#### Relación con la línea de investigación

La línea de investigación “Sistemas de Información” (*Information System*) se ha tomado de la clasificación del conocimiento en el dominio de la Computación, propuesta por la ACM [9]. En el ámbito de Sistemas de Información se inscribe un subdominio “*World Wide Web*”, que considera las tecnologías relacionadas a la descripción de datos (*semantic web description languages, markup languages*) y los motores de búsqueda (*Web Search engines and Web crawling*). Este proyecto aborda una tecnología de la web semántica, la especificación JSON-LD para el marcado semántico y el uso de Schema, un vocabulario consensuado para la Web. Plantea también el problema de la búsqueda de REA en la Web realizada a través de motores de búsqueda. Además, analiza corpus web extraídos mediante *web crawling* para descubrir datos estructurados que usen el formato JSON-LD. Adicionalmente, se considera el subdominio de Computación aplicada a la Educación, también reconocido en la clasificación ACM, dado que se aborda el uso de las tecnologías web, objeto del proyecto, en REA.

## 4 Impacto de la investigación

### 4.1 Impacto Social (máximo 250 palabras)

En los Objetivos de Desarrollo Sostenible acordados por las Naciones Unidas se incluye expresamente un objetivo relacionado a la educación que señala “Garantizar una educación inclusiva y equitativa de calidad y promover oportunidades de aprendizaje permanente para todos” [10]. El movimiento de REA tiene sus raíces en el derecho fundamental del acceso a oportunidades de educación de calidad para todos; por ello, los REA procuran el soporte para una educación de calidad, equitativa, inclusiva, abierta y participativa [5].

El fin último de este proyecto de investigación es el de utilizar la tecnología web de marcado semántico con JSON-LD para enriquecer la visualización de resultados de búsqueda en la Web a través de la inclusión de información relevante que guíe a todas las personas, incluidas las personas con discapacidad, a encontrar los contenidos educativos que requieren. De esta forma se busca eliminar la barrera de localización de REA y promover su uso para configurar una estrategia dirigida a extender las oportunidades de educación a todos los sectores de la sociedad.

De lo expuesto se desprende el impacto social de este proyecto, dado que la sociedad en su conjunto se beneficia del uso de los REA. Por ejemplo, los REA liberados por instituciones académicas de prestigio mundial pueden facilitar a las personas adultas mayores el acceso a educación para toda la vida [11]; pueden convertirse en contenidos complementarios de calidad para los estudiantes de todos los niveles educativos y en material de referencia para los profesores, quienes están facultados a adaptarlos conforme a sus necesidades.

### 4.2 Impacto Económico (máximo 250 palabras)

Los REA se publican bajo licencias de propiedad intelectual que permiten un libre acceso a los contenidos, sin restricciones para su libre uso, distribución o modificación [12]. En tal sentido, el uso de REA en contextos formales educativos puede reducir el costo relacionado a los libros de texto producidos comercialmente, que pueden ser sustituidos por textos publicados como REA, eliminando la barrera de asequibilidad económica para los estudiantes [13]. Además, los beneficios de extender la educación a amplios sectores de la población pueden generar un impacto económico al mejorar su preparación para el empleo a costos reducidos [13]. En particular, la población creciente de adultos mayores puede enrolarse en programas de educación que los mantenga como individuos activos en la sociedad [11]. Debido al impacto social y económico que genera, la iniciativa de REA ha sido respaldada en muchos países como política de gobierno, por ejemplo, India [14], Brasil [15] y países de África [16].

### 4.3 Impacto Político (máximo 250 palabras)

Muchos gobiernos en el mundo, incluido el Ecuador, han incorporado políticas relacionadas a gobierno abierto [17], uso de software y conocimiento abiertos [18]. En este sentido, también tiene cabida el movimiento de REA que busca llevar la educación con recursos abiertos a una amplia mayoría de la población. Este proyecto contribuye a mejorar la expectativa de uso de los REA para educadores y potenciales estudiantes, al apoyar su localización en y recuperación en la Web. Además, con la inclusión de



semántica embebida, es posible extraer información estructurada que puede ser explorada como datos abiertos.

#### 4.4 Impacto Científico (máximo 250 palabras)

La W3C pretende enfocar estrictamente sus esfuerzos formales de estandarización en las especificaciones que tienen más probabilidades de obtener consenso dentro del consorcio y cuyo uso puede ser extensivo en la Web real [19]. En 2014, el formato JSON-LD fue aceptado con el estado de “W3C Recommendation” [3]; por tanto, es pertinente revisar el estatus del uso que se le ha dado en la Web. Este proyecto de investigación es relevante ya que aportará a la comunidad científica un análisis cuantitativo de la extensión de adopción de este formato, a partir de corpus de rastreo web de dominio público que contienen millones de páginas web [20]. De la misma forma, el conjunto de vocabularios que forma Schema, se ha convertido en el estándar de facto para el marcado incorporado [21]. En este proyecto se obtiene también los descriptores del vocabulario de Schema más utilizados en el marcado con JSON-LD y se define aquellos que puedan ser aplicables a la descripción de los REA. Además, de encontrarse la ausencia de descriptores para documentar propiedades específicas, se complementa con descriptores de otros vocabularios. Esto podría originar un conjunto de propiedades que pueden ser propuestas para Schema (*pending status proposal*) [22], ya que la educación y particularmente los REA constituyen un campo de aplicación reciente, constituyendo un tema de interés para la comunidad científica que trabaja sobre vocabularios consensuados.

#### 4.5 Otro Impacto (máximo 250 palabras)

5	Productos esperados
---	---------------------

Tipo de Producto:	Marcar con una “X”
a. Publicaciones científicas y/o patente (obligatorio);	X
b. Disertación a la comunidad politécnica;	X
c. Trabajo de titulación de acuerdo a lo que establece el Reglamento de Régimen Académico y la Normativa Interna de la EPN;	
d. Aplicación tecnológica construida o implementada;	
e. Perfil de proyecto de mayor impacto científico, técnico, pedagógico o de innovación.	

6	Descripción, metodología y diseño del proyecto
---	--

#### 6.1 Descripción, metodología y diseño del proyecto (Máximo dos carillas)

El enfoque central del proyecto de investigación es validar el uso de la estrategia de marcado semántico embebido con la especificación JSON-LD, con base en un conjunto de descriptores de un vocabulario estandarizado, para enriquecer la información de los resultados que presentan los motores de búsqueda en la Web [23], aplicado a Recursos Educativos Abiertos (REA).

La motivación para elegir los REA proviene de las características inherentes a estos contenidos educativos que, al ser publicados con licencias abiertas de propiedad intelectual, pueden estar al alcance de todos los potenciales usuarios. Las universidades más importantes del mundo, así como grandes coaliciones entre gobierno y academia publican REA [12], lo cual ofrece la posibilidad de aumentar de forma significativa el acceso a la educación de calidad a través de la Web, cubriendo una mayor diversidad de necesidades del educador y el alumno. Esto ha sido reconocido por la UNESCO que plantea al movimiento de REA como una estrategia para sustentar el Objetivo 4 de la Agenda de Desarrollo Sostenible 2030, que insta a la



comunidad internacional a "Garantizar una educación de calidad inclusiva y equitativa y promover oportunidades de aprendizaje permanente para todos" [10].

Se ha optado por trabajar con la especificación JSON-LD porque es el formato de marcado más reciente incorporado por la W3C como recomendación [3] y constituye una forma clara y simple de embeber información en una página web; se implementa en un script y no interfiere con el código HTML. Además, a diferencia de los microdatos, el JSON-LD se ejecuta en segundo plano y no es visible, por lo que no necesita configuración para el resultado en las páginas web. JSON-LD es completamente compatible con JSON tradicional por lo cual se puede generar Linked Data o un grafo RDF, permitiendo que el contenido web pueda incluso ser explorado por herramientas de la web semántica [24].

Al momento, los estudios realizados sobre la temática del proyecto son escasos; se ha encontrado el uso de semántica embebida con JSON-LD en campos como la creación de repositorios documentales abiertos [25] o la generación de estructuras RDF en datos de gobierno abierto [26], pero no se ha encontrado información que relacione el uso de la tecnología de marcado embebido con JSON-LD al campo educativo, por lo cual, los aportes que se obtengan del proyecto serán de interés para los investigadores y los profesionales del área.

Para el proyecto se utiliza una metodología de investigación mixta, un enfoque que combina métodos de investigación cuantitativa y cualitativa [27]. Esta metodología permite la comprensión de un fenómeno de interés, en este caso la adopción de una tecnología web, que no puede ser completamente entendido utilizando únicamente un método cuantitativo o cualitativo.

En el proyecto se reconocen las siguientes fases:

- Fase I. Identificación de los corpus de rastreo web de dominio público que puedan ser utilizados para extraer la información estructurada.
- Fase II. Definición de un procedimiento de extracción del corpus de datos estructurados con formato JSON-LD desde el corpus original.
- Fase III. Análisis cuantitativo del corpus para reconocer la incidencia del formato JSON-LD y los descriptores del vocabulario Schema y otros, utilizados en el marcado.
- Fase IV. Análisis cualitativo del despliegue y pertinencia de los descriptores en el corpus como base para recomendar un conjunto de descriptores para el marcado. Se analiza también los descriptores faltantes (no presentes en el corpus) que son requeridos para la descripción de los REA.
- Fase V. Desarrollo y pruebas de la aplicación prototipo para generar scripts de marcado con JSON-LD y el vocabulario seleccionado para el marcado de los REA.

En la Fase I, se considera los corpus web de dominio público ofrecidos por Common Crawl [20] que realiza el rastreo de miles de millones de páginas web (p.ej., en la edición de noviembre de 2015 tenía 1.770.525.212 páginas web). Se verificará los corpus web de años sucesivos que pudieran estar disponibles para su procesamiento.

En la Fase II, se considera usar la nube de Amazon (EC2), un servicio web que proporciona capacidad informática en la nube, dado que la extracción de datos estructurados debe realizarse de forma distribuida para finalizar esta tarea en un tiempo razonable. Se prevé utilizar el marco de Hadoop para procesar los datos en el volumen que tienen en los corpus [28]. Dado que los corpus Common Crawl ya están divididos en archivos comprimidos, se debe idear una forma de extracción para aplicar a las páginas individuales un extractor que obtenga n-quads [29].

En la Fase III, a partir de estos n-quads que se obtengan en el corpus se realiza el análisis cuantitativo del despliegue de la especificación JSON-LD en las páginas web [30]. Los n-quads permiten la representación de una secuencia de términos RDF que contienen el sujeto, el predicado, el objeto y una etiqueta opcional vacía o un IRI. También se realiza el análisis de las propiedades de Schema y otros vocabularios, utilizadas en el marcado semántico, mediante una tokenización de los n-quad.

En la Fase IV, se realizará un análisis cualitativo de los descriptores más utilizados de Schema y otros vocabularios y los dominios en que se utilizan (con base en los resultados obtenidos de la fase previa). Se analiza la pertinencia de los descriptores (en términos de la aplicabilidad y utilidad descriptiva) para la comprensión del usuario potencial y la interpretación semántica por parte del motor de búsqueda. Se revisarán propiedades que permitan describir las generalidades del REA, sus cualidades educativas, la licencia con la que ha sido publicado y adicionalmente, la accesibilidad que ofrece el recurso. Si no se tienen propiedades completas o suficientes, se debe examinar otros estándares de vocabularios adicionales [31] [32].



En la Fase V, se desarrollará una aplicación prototipo que permita al usuario generar un marcado semántico automático del REA con JSON-LD con base en el conjunto de propiedades recomendadas. El usuario no debe requerir ningún conocimiento técnico, sino únicamente del contexto educativo en relación con el REA en el cual se inserte el marcado semántico. El script puede ser insertado directamente en el código fuente de la página, por lo que esa tarea no se incluirá en la aplicación. Las actividades de esta fase se corresponden con el ciclo de desarrollo de software: definición de requerimientos, diseño, construcción y pruebas. Al ser un prototipo, no se considera el despliegue de la aplicación.

## Referencias

- [1] W3C, "About W3C," 2018. [Online]. Available: <https://www.w3.org/Consortium/mission>. [Accessed July 2018].
- [2] L. F. Sikos, *Web Standards: Mastering HTML5, CSS3, and XML*, Springer Science, 2011.
- [3] W3C, "JSON-LD 1.0," 2014. [Online]. Available: <https://www.w3.org/TR/json-ld/>. [Accessed July 2018].
- [4] V. Suchomel and J. Pomikálek, "Efficient web crawling for large text corpora," in *Proceedings of the seventh Web as Corpus Workshop*, Lyon, France, 2012.
- [5] UNESCO, "Ljubljana OER Action Plan 2017 adopted to support quality open educational resources," 2017. [Online]. Available: <https://en.unesco.org/news/ljubljana-oer-action-plan-2017-adopted-support-quality-open-educational-resources>. [Accessed July 2018].
- [6] PNUD Ecuador, "Objetivos de Desarrollo Sostenible," 2018. [Online]. Available: <http://www.ec.undp.org/content/ecuador/es/home/sustainable-development-goals.html>. [Accessed July 2018].
- [7] Andes, "Agencia pública de noticias del Ecuador y Suramérica," 2018. [Online]. Available: <https://www.andes.info.ec/es/noticias/actualidad/15/ecuador-adopta-como-politica-publica-la-agenda-2030-para-el-desarrollo-sostenible>. [Accessed July 2018].
- [8] UNESCO and Slovenia Government, "2ND WORLD OER CONGRESS LJUBLJANA OER ACTION PLAN 2017," Ljubljana, Slovenia, 2017.
- [9] ACM, 2012. [Online]. Available: <https://www.acm.org/publications/class-2012>. [Accessed July 2018].
- [10] Naciones Unidas, "Objetivos de Desarrollo Sostenible," 2015.
- [11] R. Navarrete and S. Luján-Mora, "OER in lifelong learning for older adults," in *The Visually Disabled and the Elderly in the Age of IC Technologies*, S. A. N. O. M. Pilar Munuera Gómez, Ed., Lenovo, 2018, pp. 31-58.
- [12] D. Wiley, T. J. Bliss and . M. McEwen, "Open Educational Resources: A Review of the Literature," in *Handbook of Research on Educational Communications and Technology*, New York, Springer, 2014, pp. 781-789.
- [13] D. Annand and T. Jensen, "Incentivizing the Production and Use of Open Educational Resources in Higher Education Institutions," *International Review of Research in Open and Distributed Learning*, vol. 18, no. 4, 2017.
- [14] A. Thakran and R. C. Sharma, "Meeting the challenges of higher education in India through open educational resources: policies, practices, and implications.," *OER Knowledge Cloud*, vol. 24, no. 37, 2016.
- [15] N. Allen, "Brazil adopts Open Licensing in National Textbook Program," SPARC, 2017. [Online]. Available: <https://sparcopen.org/news/2017/brazil-adopts-open-licensing-national-textbook-program/>. [Accessed July 2018].
- [16] Africa, "OER Africa," 2018. [Online]. Available: <https://www.oerafrica.org/>. [Accessed July 2018].
- [17] SENPLADES, "Catálogo de datos abiertos gubernamentales," 2017. [Online]. Available: <http://sni.gob.ec/datosabiertos>. [Accessed Julio 2018].
- [18] Open Knowledge International, "Open Knowledge Ecuador," 2014. [Online]. Available:



- <https://ec.network.okfn.org/>. [Accessed Julio 2018].
- [19] W3C, "W3C Recommendation Track Readiness Best Practices," 2016. [Online]. Available: <https://www.w3.org/Guide/standards-track/>. [Accessed July 2018].
- [20] Common Crawl Foundation, "Common Crawl," 2016. [Online]. Available: <http://commoncrawl.org/>. [Accessed July 2018].
- [21] R. Meusel, C. Bizer and H. Paulheim, "A Web-scale Study of the Adoption and Evolution of the Schema.org Vocabulary over Time," in *ACM 5th International Conference on Web Intelligence, Mining and Semantics (WIMS '15)*, New York, NY, USA, 2015.
- [22] L. Yu, "Schema.org and Semantic Markup," in *A Developer's Guide to the Semantic Web*, Berlin, Heidelberg, Springer Berlin Heidelberg, 2014, p. 475–515.
- [23] K. Haas, P. Mika, P. Tarjan and R. Blanco, "Enhanced Results for Web Search," in *34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Beijing, China, 2011.
- [24] T. D. Huynh, D. T. Michaelides and L. Moreau, "PROV-JSONLD: A JSON and Linked Data Representation for Provenance," in *International Provenance and Annotation Workshop (IPAW 2016)*, McLean, Virginia, USA, 2016.
- [25] M. J. O'Connor, M. Martínez-Romero, A. L. Egyedi, D. Willrett, J. Graybeal and M. A. Musen, "An Open Repository Model for Acquiring Knowledge About Scientific Experiments," in *European Knowledge Acquisition Workshop (EKAW 2016)*, Bologna, Italy, 2016.
- [26] S. Theocharis and G. Tsihrintzis, "RDF serialization from JSON data: The case of JSON data in Diavgeia.gov.gr," in *7th International Conference on Information, Intelligence, Systems & Applications (IISA 2016)*, Chalkidiki, Greece, 2016.
- [27] V. Venkatesh, S. A. Brown and H. Ba, "Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems," *MIS Quarterly*, vol. 37, no. 1, pp. 21-54, 2013.
- [28] H. Mühleisen and C. Bizer, "Web data commons - Extracting structured data from two large web corpora," in *CEUR Workshop Proceedings*, Heraklion, Greece, 2012.
- [29] W3C, "RDF 1.1 N-Quads," 2014. [Online]. Available: <https://www.w3.org/TR/n-quads/>. [Accessed July 2018].
- [30] P. Mika and T. Potter, "Metadata statistics for a large web corpus," in *CEUR Workshop Proceedings*, Heraklion, Greece, 2012.
- [31] R. Navarrete and S. Luján-Mora, "Metadata in Open Educational Resources websites: a review from the perspective of disabled users' requirements," in *Proceedings of the 6th International Conference on Education and New Learning Technologies (EDULEARN 2014)*, Barcelona, Spain, 2014.
- [32] P. Barker and L. Campbell, "Learning Resource Metadata Initiative: using schema.org to describe open educational resources," in *OCWC Glob. Conf. 2014 - Open Educ. a Multicult. World*, Ljubljana, Slovenia, 2014.

## 7 Infraestructura, equipos y fondos adicionales.

### 7.1 Infraestructura y equipos

Se requiere infraestructura para procesamiento distribuido. Esta se solicitará a CEDIA o a la Universidad de Alicante, por intermedio del PhD Sergio Luján Mora, colaborador externo del proyecto y catedrático de dicha universidad.

### 7.2 Breve justificación del equipo requerido

No se solicita equipamiento como parte del proyecto.

### 7.3 Fondos Adicionales

No existen fondos adicionales.





## VICERRECTORADO DE INVESTIGACIÓN Y PROYECCIÓN SOCIAL

UNIDAD DE INVESTIGACIÓN

PRESUPUESTO PROYECTOS DE INVESTIGACIÓN



AÑO 1

Título del proyecto

Semántica embebida con JSON-LD para generar información enriquecida en resultados de búsqueda de Recursos Educativos Abiertos

Lista de Items		Cantidad	Unidad	Precio Unitario Referencial	Precio Total Referencial	Precio Unitario Referencial +Aporte IESS	Precio Total Referencial con IVA + Aporte del IESS
<b>1 Contratación de servicios personales por contrato</b>							
1.1	Ayudantes de investigación	5	mes	\$ 133.67	\$ 668.35	\$ 157.20	\$ 785.98
1.2	Prestación de servicios profesionales (Homologado Escala de remuneración de servidores publicos)		mes	\$ -	\$ -	\$ -	\$ -
<b>Subtotal 1</b>				<b>\$ 133.67</b>	<b>\$ 668.35</b>	<b>\$ 157.20</b>	<b>\$ 785.98</b>
Lista de Items		Cantidad	Unidad	Precio Unitario Referencial sin IVA	Precio Total Referencial sin IVA	Precio Unitario Referencial con IVA	Precio Total Referencial con IVA
<b>2 Maquinaria y equipo especializado</b>							
2.1	Item 1 (Detallar nombre de la maquinaria y equipos solicitado)			\$ -	\$ -	\$ -	\$ -
2.2	Item 2 (Detallar nombre de la maquinaria y equipos solicitado)			\$ -	\$ -	\$ -	\$ -
2.3	Item 3 (Detallar nombre de la maquinaria y equipos solicitado)			\$ -	\$ -	\$ -	\$ -
2.4	Item 4 (Detallar nombre de la maquinaria y equipos solicitado)			\$ -	\$ -	\$ -	\$ -
2.5	Item 5 (Detallar nombre de la maquinaria y equipos solicitado)			\$ -	\$ -	\$ -	\$ -
<b>Subtotal 2</b>				<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>
<b>3 Equipo informático</b>							
3.1	Item 1 (Detallar nombre de la maquinaria y equipos solicitado)			\$ -	\$ -	\$ -	\$ -
3.2	Item 2 (Detallar nombre de la maquinaria y equipos solicitado)			\$ -	\$ -	\$ -	\$ -
3.3	Item 3 (Detallar nombre de la maquinaria y equipos solicitado)			\$ -	\$ -	\$ -	\$ -
3.4	Item 4 (Detallar nombre de la maquinaria y equipos solicitado)			\$ -	\$ -	\$ -	\$ -
3.5	Item 5 (Detallar nombre de la maquinaria y equipos solicitado)			\$ -	\$ -	\$ -	\$ -
<b>Subtotal 3</b>				<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>
<b>4 Insumos y reactivos</b>							
4.1	Item 1 (Detallar nombre de los insumos y reactivos)			\$ -	\$ -	\$ -	\$ -
4.2	Item 2 (Detallar nombre de los insumos y reactivos)			\$ -	\$ -	\$ -	\$ -
4.3	Item 3 (Detallar nombre de los insumos y reactivos)			\$ -	\$ -	\$ -	\$ -
4.4	Item 4 (Detallar nombre de los insumos y reactivos)			\$ -	\$ -	\$ -	\$ -
4.5	Item 5 (Detallar nombre de los insumos y reactivos)			\$ -	\$ -	\$ -	\$ -
<b>Subtotal 4</b>				<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>
<b>5 Literatura especializada</b>							
5.1	Item 1 ( Detallar nombre del libro)			\$ -	\$ -	\$ -	\$ -
5.2	Item 2 ( Detallar nombre del libro)			\$ -	\$ -	\$ -	\$ -
5.3	Item 3 ( Detallar nombre del libro)			\$ -	\$ -	\$ -	\$ -
5.4	Item 4 ( Detallar nombre del libro)			\$ -	\$ -	\$ -	\$ -
5.5	Item 5 ( Detallar nombre del libro)			\$ -	\$ -	\$ -	\$ -
<b>Subtotal 5</b>				<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>
<b>6 Salidas de campo y de muestreo</b>							
6.1	Pasajes al interior			\$ -	\$ -	\$ -	\$ -
6.2	Viaticos y subsistencias al interior			\$ -	\$ -	\$ -	\$ -
<b>Subtotal 6</b>				<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>
<b>7 Ponencias nacionales, capacitaciones y/o visitas técnicas</b>							
7.1	Pasajes al interior			\$ -	\$ -	\$ -	\$ -
7.2	Viaticos y subsistencias al interior			\$ -	\$ -	\$ -	\$ -
<b>Subtotal 7</b>				<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>
<b>8 Ponencias en el exterior, capacitaciones, y/o visitas técnicas</b>							
8.1	Pasajes al exterior	1	Unidad	\$ 1,700.00	\$ 1,700.00	\$ 1,904.00	\$ 1,904.00
8.2	Viaticos al exterior	5	Días	\$ 240.00	\$ 1,200.00	\$ 240.00	\$ 1,200.00
<b>Subtotal 8</b>				<b>\$ 1,940.00</b>	<b>\$ 2,900.00</b>	<b>\$ 2,144.00</b>	<b>\$ 3,104.00</b>
<b>9 Pago de inscripciones</b>							
9.1	Pago de inscripciones al interior			\$ -	\$ -	\$ -	\$ -
9.2	Pago de inscripciones al exterior	1	Unidad	\$ 810.00	\$ 810.00	\$ 1,109.70	\$ 1,109.70
<b>Subtotal 9</b>				<b>\$ 810.00</b>	<b>\$ 810.00</b>	<b>\$ 1,109.70</b>	<b>\$ 1,109.70</b>
<b>10 Pago de publicaciones y patentes</b>							
10.1	Pago de publicaciones			\$ -	\$ -	\$ -	\$ -
10.2	Pago de publicaciones al exterior			\$ -	\$ -	\$ -	\$ -
10.2	Pago de patentes			\$ -	\$ -	\$ -	\$ -
<b>Subtotal 10</b>				<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>
<b>TOTAL</b>				<b>\$ 2,883.67</b>	<b>\$ 4,378.35</b>	<b>\$ 3,410.90</b>	<b>\$ 4,999.68</b>



**VICERRECTORADO DE INVESTIGACIÓN Y PROYECCIÓN SOCIAL**  
**UNIDAD DE INVESTIGACIÓN**  
**PRESUPUESTO PROYECTOS DE INVESTIGACIÓN**



Título del proyecto
Semántica embebida con JSON-LD para generar información enriquecida en resultados de búsqueda de Recursos Educativos Abiertos

**Presupuesto consolidado sin IVA**

AÑO	Contratación de servicios personales por contrato	Maquinaria y equipo especializado	Equipo informático	Insumos y reactivos	Literatura especializada	Salidas de campo y de muestreo	Ponencias nacionales, capacitaciones y/o visitas técnicas	Ponencias en el exterior, capacitaciones, y/o visitas técnicas	Pago de inscripciones	Pago de publicaciones y patentes	Total sin IVA
1	\$ 668.35	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ 2,900.00	\$ 810.00	\$ -	\$ 4,378.35
2	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -
3	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -
<b>TOTAL</b>	<b>\$ 668.35</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ 2,900.00</b>	<b>\$ 810.00</b>	<b>\$ -</b>	<b>\$ 4,378.35</b>

**Presupuesto consolidado con IVA**

AÑO	Contratación de servicios personales por contrato	Maquinaria y equipo especializado	Equipo informático	Insumos y reactivos	Literatura especializada	Salidas de campo y de muestreo	Ponencias nacionales, capacitaciones y/o visitas técnicas	Ponencias en el exterior, capacitaciones, y/o visitas técnicas	Pago de inscripciones	Pago de publicaciones y patentes	Total con IVA
1	\$ 785.98	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ 3,104.00	\$ 1,109.70	\$ -	\$ 4,999.68
2	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -
3	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -
<b>TOTAL</b>	<b>\$ 785.98</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ -</b>	<b>\$ 3,104.00</b>	<b>\$ 1,109.70</b>	<b>\$ -</b>	<b>\$ 4,999.68</b>

## DECLARACIÓN FINAL

### TIPO DE PROYECTO

Proyecto Interno  Proyecto Semilla  Proyecto Junior  Proyecto Multi e Interdisciplinario

### TIPO DE INVESTIGACIÓN

Investigación básica  Investigación aplicada

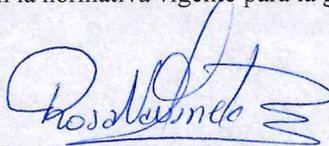
### TÍTULO DEL PROYECTO

Semántica embebida con JSON-LD para generar información enriquecida en resultados de búsqueda de Recursos Educativos Abiertos

### DECLARACIÓN DEL DIRECTOR DEL PROYECTO

El equipo de investigadores, representado por el Director del Proyecto declara lo siguiente:

- Que el presente proyecto es una obra original de este equipo de investigadores y por tanto, asumimos la completa responsabilidad legal en caso de que un tercero alegue la titularidad de los derechos intelectuales del proyecto, exonerando a la EPN de cualquier acción legal que se derive por esta causa.
- Que el presente proyecto no ha sido presentado en ninguna convocatoria de otra institución pública o privada solicitando el financiamiento total del presupuesto. El incumplimiento será causal para que la propuesta sea descalificada de la convocatoria de la EPN.
- Que todos los bienes adquiridos en el proyecto permanecerán bajo la custodia y responsabilidad del director de proyecto.
- Que aceptamos que si el proyecto genera algún producto o procedimiento susceptible de obtener de derechos de propiedad intelectual, de los cuales se deriven beneficios, éstos serán compartidos entre los investigadores y la EPN.
- Que aceptamos conocer y cumplir con la normativa vigente para la gestión de proyectos de investigación.



Firma del Director del Proyecto  
Nombre: Rosa del Carmen Navarrete Rueda  
C.I.: 1001303849

### DECLARACIÓN DEL JEFE DE DEPARTAMENTO

Esta propuesta ha sido avalada por el Consejo del Departamento de Informática y Ciencias de la Computación, en sesión del día 30 de julio de 2018 mediante resolución No. 127.032.30-07-2018.

Las instalaciones, incluyendo personal, edificios, equipo y recursos financieros están a disposición del proponente y sus colaboradores de acuerdo con las especificaciones que se encuentran en esta propuesta.



Firma del Jefe del Departamento  
Nombre: Dra. Myriam Peñafiel  
C.I.: 1705828711



\*Se debe adjuntar el acta en el que conste la resolución que avala la propuesta de proyecto