

Tipo de artículo: Artículo de datos
Temática: Ingeniería y gestión de software
Recibido: 18/05/2018 | Aceptado: 08/12/2018

Repositorio de datos para investigaciones en gestión de proyectos

Database Repository for research on Project Management

Pedro Y. Piñero Pérez, Iliana Pérez Pupo*, Claudia C. Rivero Hechavarría, Claribel Rojas Lusardo, Rosel González Sosa, Surayne Torres López

Grupo de Investigaciones en Gestión de Proyectos, Universidad de las Ciencias Informáticas, Carretera a San Antonio de Baños Km 2 1/2, Boyeros, La Habana, Cuba. {ppp, iperez, ccrivero, rsosa, storres}@uci.cu

* Autor para correspondencia: iperez@uci.cu

Resumen

Los repositorios de datos permiten el desarrollo de investigaciones, el descubrimiento de conocimiento y la prueba de nuevos algoritmos de aprendizaje. El presente trabajo tiene como objetivo presentar un repositorio para el desarrollo de investigaciones en gestión de proyectos. En el trabajo los autores presentan un conjunto de procesos técnicas para la gestión del repositorio. El repositorio está compuesto por 18 bases de datos que cubren diferentes áreas de conocimiento en gestión de proyectos. Los datos reflejan el comportamiento de proyectos de desarrollo de software. En la sección de resultados y discusión los autores referencias un conjunto de tesis de doctorado, tesis de maestría y artículos que emplearon los datos del repositorio para diferentes investigaciones. Los autores evalúan la calidad percibida de los investigadores aplicándole a los mismos un cuestionario que evalúa las variables pertinencia y usabilidad.

Palabras clave: descubrimiento de conocimiento, gestión de proyectos, gestión de software, repositorio de datos

Abstract

Data repositories allow the sharing, reuse and localization of data for learning and discovery of new knowledge within organizations. In this work authors presents a repository for investigations in project management. This repository consists of 18 databases covering the most of project management areas. All data are from software projects and can be use in different investigations. The paper presents to processes and techniques to manage the databases in repository. In section results and discussion authors presents different references of phd thesis, master

degree thesis and papers that used repository's databases. Authors evaluate the quality perceived of researchers by applying a questionnaire to evaluate two variables pertinence and usability.

Keywords: *knowledge discovery, project management, repository of data, software management.*

Introducción

La ingeniería del software constituye una de las industrias emergentes que mayor crecimiento e impacto ha experimentado en años recientes. El desarrollo de los proyectos de esta industria respecto a la calidad y las buenas practicas está marcado por la existencia de numerosos estándares y libros desarrollados por diversos autores, entre los que se destacan, entre otros, el PMBOK (PMI, 2017), la ISO 21500 (Stellingwerf & Zandhuis, 2013) y CMMI (SEI, 2015). Pero a pesar de la gran cantidad de guías persisten dificultades como se refleja en los Reportes CHAOS del Standish Group (Johnson, Gesmer, Poort, & Mulder, 2016).

En los reportes realizados entre el 2011 y el 2015 se señala que: aproximadamente el 19% de los proyectos fracasan, que alrededor del 52% de los proyectos son renegociados por variaciones en su alcance, costo o calidad y que solo el 29% de los proyectos fueron exitosos (Hastie & Wojewoda, 2015). En este escenario aprender de la experiencia de proyectos anteriores, ayudaría a disminuir el elevado número de proyectos cancelados o renegociados. En particular la minería de datos es una de las técnicas que los autores de este trabajo recomiendan en este sentido, algunas de estas mismas técnicas aparecen como recomendación a partir de la versión 6 del PMBOK (PMI, 2017).

En el contexto de esta investigación entendemos a la minería de datos como un campo interdisciplinario donde convergen técnicas estadísticas, matemáticas, inteligencia artificial, el reconocimiento de patrones y bases de datos para el descubrimiento de conocimiento no trivial que reside de manera implícita en los datos (Gorunescu, 2011). Pero para garantizar el desarrollo de investigaciones empleando estas técnicas es preciso contar con suficientes datos que en este campo en particular generalmente provienen de sistemas de información que no están públicos dificultando el acceso a la información.

Existen diferentes repositorios de bases de datos internacionales para el desarrollo de investigaciones (Kolisch & Sprecher, 1996). Uno de los repositorios más reconocidos es el UCI Repository (Catherine Blake, 1998)(C Blake & Merz, 2015) para aprendizaje automático que incluye numerosas bases de datos de disímiles campos de aplicación. El PSPLIB (Kolisch & Sprecher, 1996) por su parte es especializado en bases de datos para la asignación de recursos,

útil en la prueba de técnicas de optimización. En (Pérez Pupo, 2011) (Santiesteban et al., 2016) se hace un análisis de repositorios de datos y se identifica que no cubren las necesidades de investigaciones en la temática de gestión de proyectos.

En este trabajo se propone un repositorio para el desarrollo de investigaciones en gestión de proyectos que incluye 18 bases de datos asociadas a proyectos de software. En la sección materiales y métodos se propone un conjunto de procesos asociados a la construcción y gestión del repositorio. En esta sección también se explica la estructura del repositorio y las bases de datos que lo conforman. En la sección resultados y discusión se presentan un conjunto de investigaciones que han empleado el repositorio en la validación de los resultados.

Materiales y métodos o Metodología computacional

En esta sección se presenta una propuesta de modelo para la construcción y gestión de repositorios de datos para el desarrollo de investigaciones. En la Figura 1 se describen los procesos que intervienen en la construcción del repositorio y que fueron desarrollados inicialmente en la tesis de maestría de Iliana Pérez (Pérez Pupo, 2011).

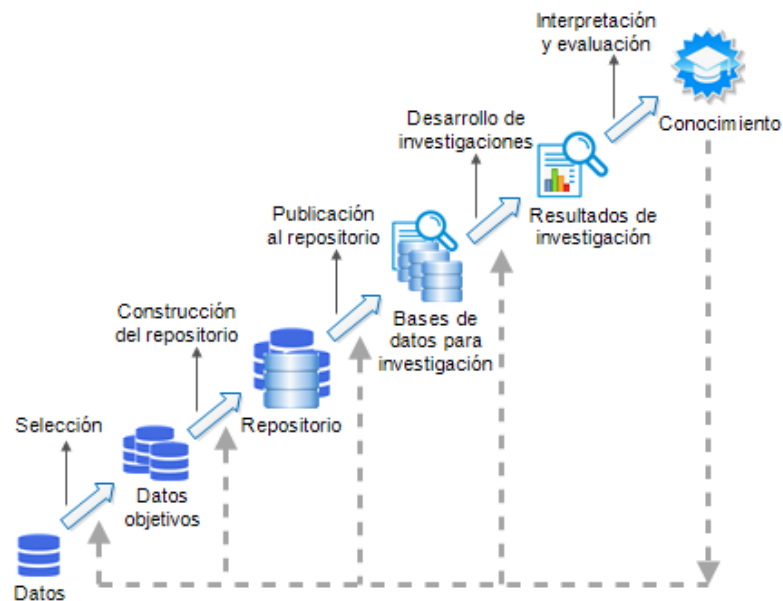


Figura 1. Procesos que intervienen en la gestión del repositorio. Fuente: elaboración propia.

La construcción del repositorio para investigaciones en gestión de proyectos tiene sus inicios en el 2008 con el desarrollo de la plataforma del sistema de información GESPRO (Piñero et al., 2010) (Piñero et al., 2015). Se decide

desarrollar el sistema de información en gestión de proyectos para: apoyar la innovación, la toma de decisiones y el desarrollo de investigaciones que permiten elevar los resultados en la ejecución de proyectos. Se inicia desde ese momento, la recopilación de datos que podrían ser empleados posteriormente para el desarrollo de investigaciones. Los procesos que intervienen en la gestión del repositorio, como muestra la Figura 1 son:

- Selección de los datos: tiene como objetivo seleccionar las fuentes de datos y caracterizar los datos dependiendo del contexto. Se proponen indicadores para la calidad del dato y se planifica el proceso de construcción del repositorio. Como salida, este proceso genera un conjunto de datos seleccionados para formar el repositorio.
- Construcción de base de datos. En este proceso se definen las áreas temáticas o áreas de conocimiento de gestión de proyectos que va a representar cada base de datos, e incluye los siguientes subprocesos:
 - Preprocesamiento de los datos: se aplican técnicas de limpieza (Escobar Pompa, 2015) como integración, transformación, reducción, anonimización, pseudonimización, imputación, verificación de datos y selección de atributos.
 - Construcción de base de datos: se seleccionan los atributos que conformarán la base de datos para investigaciones (Ril Valentín, 2012) (Medina Rodríguez, 2012). En este proceso se debe tener en cuenta que el diseño de la base de datos debe facilitar la recuperación rápida de los registros.
 - Actualización de la base de datos: se perfecciona la base de datos incorporando nuevos registros, modificando y/o eliminando otros. Como salida se obtiene una base de datos con mejores condiciones para la investigación.
- Publicación de bases de datos en el repositorio: es el proceso mediante el cual se revisa y se aprueba la publicación de las bases de datos en el repositorio cumpliendo con la estructura establecida (en la siguiente sección) para facilitar las consultas y explotación de las mismas. Como salida, se obtiene el repositorio conformado por cada una de las bases de datos que representarán las áreas de conocimiento de gestión de proyectos.
- Desarrollo de investigaciones: es donde el o los investigadores utilizan los datos descargados del repositorio para realizar y validar su investigación. Como salida, en este proceso se generan nuevos datos, que es el resultado de la aplicación de los algoritmos o métodos de la investigación al juego de datos inicial. Las bases de datos empleadas en la investigación deben ser debidamente referenciadas. La salida a este proceso son los resultados de la investigación.

- Interpretación y evaluación: en este paso se analizan los resultados del uso de las bases de datos. Se suben al repositorio los resultados de las investigaciones, los cuales podrán ser utilizados en comparaciones de técnicas y algoritmos en otras investigaciones. En este proceso la salida es el conocimiento generado en las investigaciones realizadas sobre los datos del repositorio para la gestión de proyectos.

Para la gestión y publicación del repositorio, son importantes los siguientes roles:

- Gestor del repositorio: personal encargado de mantener el orden y limpieza de los datos en el repositorio.
- Investigador: es todo aquel que utilizará los datos para el desarrollo y validación de su investigación.

Estructura del repositorio

En el repositorio, cada base de datos está representada por un fichero .name y .data, similar a la estructura del UCI Repository (Blake 1998) (Blake y Merz 2015):

- El fichero .name representa la estructura de la base de datos y debe contener la siguiente información:
 - Título: nombre de la base de datos, debe ser representativo al área de conocimiento de la gestión de proyectos correspondiente a los datos que la contiene.
 - Fuente: entidad o institución que provee los datos. Aquí se deberá ofrecer la mayor cantidad posible de datos de contacto e identificación; ejemplo, dirección, teléfonos, correos, etc.
 - Citado: la forma en que debe ser citada la base de datos, no deberá faltar el nombre ni la fecha en que se proveyeron los datos.
 - Historial de uso: se listan todas las investigaciones realizadas con esa base de datos, referenciando cada una de las publicaciones.
 - Información relevante: es donde se describen los atributos que conforman la base de datos.
 - Cantidad de registros de la base de datos.
 - Cantidad de atributos.
 - Valores de atributos ausentes: se indica el término o valor que se utilizará en los casos de valores ausentes.
 - Información sobre datos anómalos: en los atributos que sea necesario, se especifica el valor a partir del cual se considera que es un dato anómalo o fuera de rango.
- El fichero .data es un fichero texto en formato 'csv' que contiene los datos de cada uno de los registros y los atributos en el mismo orden que se explicaron en el fichero .name.

La versión actual del repositorio está compuesta por las siguientes bases de datos:

Asociadas a la temática de evaluación de proyectos y programa:

1. Base de datos de gestión de proyectos: contiene información general de proyectos que puede ser usada para la estimación de tiempo y costo de proyectos.
2. Base de datos de evaluación de proyectos: contiene información de los indicadores en los diferentes cortes evaluativos de los proyectos con datos mezclados.
3. Base de datos de evaluación de proyectos hard: contiene información de los indicadores en los diferentes cortes evaluativos de los proyectos con datos duros. Está basada en indicadores con valores nominales que evalúan a los proyectos.
4. Base de datos de evaluación borrosa de proyectos: contiene información de los indicadores en los diferentes cortes evaluativos de los proyectos con datos borrosos.
5. Base de datos de gestión de indicadores: contiene información de indicadores generales de proyectos.
6. Base de datos de gestión de indicadores del cliente: contiene información de indicadores generales de programas de proyectos asociados a clientes. Relaciona indicadores orientados al cliente.
7. Base de datos de gestión de indicadores de la organización: contiene información de indicadores generales de control de organizaciones orientadas a proyectos. Relaciona indicadores orientados a la organización.
8. Base de datos de gestión de indicadores de programas: contiene información de indicadores generales de programas de proyectos asociados a temáticas. Relaciona indicadores orientados a programas.

Asociadas a la temática de asignación a los proyectos de recursos humanos y no humanos:

9. Base de datos de gestión de tareas: contiene información general sobre las tareas y su ejecución.
10. Base de datos de asignación de recursos a tareas: contiene información general sobre la asignación de recursos a las tareas.
11. Base de datos de gestión de recursos humanos y materiales: contiene información general sobre la asignación de los recursos humanos y no humanos.

Asociadas a la temática de gestión de alcance:

12. Base de datos de gestión de requerimientos: contiene información general sobre la gestión de requerimientos.
13. Base de datos de gestión de cambios: contiene información general sobre la gestión de cambios.

Asociadas a la temática de gestión de riesgos:

14. Base de datos de gestión de riesgos: contiene información general sobre la gestión de riesgos.
15. Base de datos de gestión de desviaciones: contiene información general sobre la gestión de desviaciones.

Asociadas a la temática de gestión de recursos humanos:

16. Base de datos de gestión de competencias de usuarios: relaciona los usuarios y las competencias que estos presentan. Está basada en 4 indicadores que arrojan una evaluación del usuario, la cual se almacena con términos lingüísticos (Torres López, 2015).
17. Base de datos de gestión de los indicadores de los recursos humanos del proyecto: contiene información general sobre la gestión de los indicadores de los recursos humanos del proyecto. Relaciona indicadores que arrojan una evaluación cualitativa y cuantitativa del recurso humano por proyectos (Torres López 2011).
18. Base de datos de gestión de los indicadores de los recursos humanos de la organización: contiene información general sobre la gestión de los indicadores de los recursos humanos de la organización. Relaciona indicadores que arrojan una evaluación cualitativa y cuantitativa de los recursos humanos que almacena.

A continuación, se exponen las características generales de las bases de datos que conforman el repositorio.

Tabla 1. Características generales de las bases de datos del repositorio.

Base de Datos	Tipo de Datos	Cantidad de Atributos	Cantidad de Instancias	Cantidad de Proyectos
1. Gestión de Proyectos	Mezclado	49	136	134
2. Evaluación de Proyectos	Mezclado	84	8 430	667
3. Evaluación de Proyectos Hard	Mezclado	19	8 430	667
4. Evaluación Borrosa de Proyectos	Mezclado	19	8 430	667
5. Gestión de Indicadores	Mezclado	93	34 810	996
6. Gestión de Indicadores del Cliente	Mezclado	16	387	0
7. Gestión de Indicadores de la Organización	Mezclado	16	87	0
8. Gestión de Indicadores de Programas	Mezclado	16	87	0
9. Gestión de Tareas	Mezclado	26	12 563	103
10. Asignación de Recursos y Tareas	Mezclado	38	9 569	88
11. Gestión de Recursos Humanos y Materiales	Mezclado	28	612	35
12. Gestión de Requerimientos	Mezclado	17	1 652	36
13. Gestión de Cambios	Mezclado	20	593	44
14. Gestión de Riesgos	Mezclado	19	2 379	345
15. Gestión de Desviaciones	Mezclado	10	2 506	330
16. Gestión de Competencias de Usuarios	Mezclado	11	10 452	0

17. Gestión de los Indicadores de los Recursos Humanos del Proyecto	Mezclado	9	6 539	84
18. Gestión de los Indicadores de los Recursos Humanos de la Organización	Mezclado	8	3 149	0

En la Tabla 2 se muestra un breve análisis en el que se manifiesta la relación de las bases de datos del repositorio con 9 áreas de conocimiento de la gestión de proyecto propuestas por el PMBOK y la ISO 21500.

Tabla 2. Relación entre las bases de datos del repositorio y las áreas de conocimiento de la gestión de proyectos.

Áreas de conocimiento	Bases de datos																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Alcance	x				x				X	x	x	x	x		x			
Tiempo	x	x	x	x	x				X	x	x	x	x	x	x			
Costos	x	x	x	x	x	x	x	x	X	x	x					x		
Calidad			x		x		x	x				x	x					
Recursos humanos	x	x		x	x	x	x	x	X		x			x		x	x	x
Riesgos														x	x			
Adquisiciones					x					x	x							
Interesados	x	x				x												
Integración			x	x	x		x	x		x	x			x			x	x

Resultados y discusión

En esta sección se presentan las experiencias en el uso del repositorio. Se organiza esta sección considerando los siguientes conjuntos de indicadores:

- Indicadores asociados al nivel de explotación del repositorio en investigaciones de doctorado y maestría.
 - Cantidad de investigaciones de maestría emplean el repositorio.
 - Cantidad de investigaciones de doctorado emplean el repositorio.
 - Cantidad de artículos científicos que emplean el repositorio.
- Indicadores asociados a la calidad percibida de usuarios finales del repositorio.

El repositorio ha permitido el aumento de las investigaciones asociadas a la gestión de proyectos ver Figura 2. Se relacionan a continuación investigaciones realizadas que emplearon el repositorio en la validación de resultados.

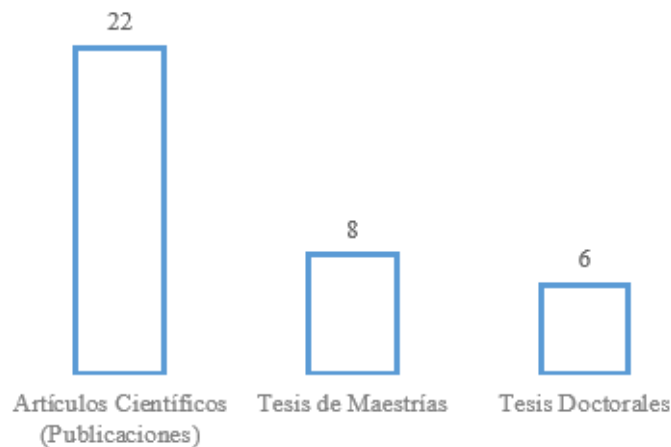


Figura 2. Investigaciones realizadas empleando el repositorio de gestión de proyectos.

Investigaciones de doctorado terminadas que han utilizado el repositorio en la validación:

1. Modelo de control de proyectos basado en análisis geo-referencial para la ayuda a la toma de decisiones (Jiménez Moya, 2017).
2. Modelo para el aseguramiento de ingresos en organizaciones orientadas a proyectos basado en minería de datos anómalos (Castro Aguilar, 2017b)
3. Modelo para el análisis de factibilidad de proyectos de software en entornos de incertidumbre (Peña Abreu, 2017).
4. Modelo para el control de la ejecución de proyectos basado en soft computing (Lugo García, 2015).
5. Modelo de evaluación de competencias a partir de evidencias durante la gestión de proyectos (Torres López, 2015).
6. Modelo ontológico de recuperación de información para la toma de decisiones en gestión de proyectos (Fernández Hernández, 2016).

Investigaciones de maestría que han utilizado información del repositorio:

1. Modelo para el control de la ejecución de proyectos basado en indicadores y lógica borrosa (Lugo García, 2012).
2. Modelo para análisis de factibilidad en la evaluación de proyectos de software (Peña Abreu, 2012).
3. Modelo para la evaluación de la composición de equipos de proyectos informáticos (Rodríguez Stiven, 2012).
4. Modelo para la evaluación de competencias como problemas multietiquetas a partir de sistemas basados en casos (Perdomo Alonso, 2013).

5. Algoritmo de sumarización lingüística como apoyo a la toma de decisiones en gestión de proyecto (Piñera Trinchet, 2013).
6. Descubrimiento de conocimiento a partir de la relación rasgos de la personalidad-rendimiento laboral en proyectos informáticos (Pacheco Rodríguez, 2014).
7. Sistema basado en técnicas de soft computing para la evaluación de la ejecución de proyectos (Bermudez Peña, 2015).
8. Procedimiento para la aplicación de test de personalidad como apoyo a la gestión de recursos humanos en proyectos informáticos (López 2018).

Respecto a la cantidad de artículos científicos que emplean el repositorio, se generaron un total de 22 publicaciones como se muestra a continuación:

1. A New Neutrosopic Cognitive Map with Neutrosophic Sets on Connections, Application in Project Management (Hasan Al-Subhi, Pérez Pupo, García Vacacela, Piñero Pérez, & Yelandy Leyva, 2018).
2. Extensions to Linguistic Summaries Indicators based on Neutrosophic Theory, Applications in Project Management Decisions (Pérez Pupo, Piñero Pérez, García Vacacela, Santos Acosta, & Yelandy Leyva, 2018)
3. Descubrimiento de resúmenes lingüísticos para ayuda a la toma decisiones en gestión de proyectos (Pupo, Acosta, Vacacela, Pérez, & Ramírez, 2018).
4. Construcción de resúmenes lingüísticos a partir de rasgos de la personalidad y el desempeño en el desarrollo de software (Pupo, Gómez, Varona, Pérez, & Vacacela, 2018).
5. Sistema clasificador borroso basado en algoritmos genéticos para evaluar el estado de ejecución de proyectos. (Rodríguez Rodríguez, Peña Abreu, Castro Aguilar, Pérez, & Yobanis, 2017)
6. Project Control and Computational Intelligence: Trends and Challenges. (García, Peña, Pérez, & Pérez, 2017)
7. PRODANALYSIS, un Sistema para el Aseguramiento de Ingresos Basado en Minería de Outliers (Aguilar et al. 2016).
8. Plataforma para aseguramiento de ingresos, aplicación en gestión de proyectos y telcos (Castro Aguilar, Pérez Pupo, Piñero Pérez, & García Vacacela, 2016).
9. Rough Sets for Human Resource Competence Evaluation and Experiences (López, Aguilar, Pupo, Pérez, & Diéguez, 2016).
10. Experiencias usando algoritmos genéticos en la planificación de proyectos (García Vacacela, Pérez Pupo, Villavicencio, Piñero, & Beovides, 2016).

11. Ecosistema de Software GESPRO-16.05 para la Gestión de Proyectos (Sosa González, Pérez Pupo, García, Peñaherrera, & Piñero Pérez, 2016).
12. Elección entre una metodología ágil y tradicional basado en técnicas de soft computing (Boaventura José, Peña Herrera, Verdecia Vicet, & Fustiel Alvarez, 2016).
13. SIGESPRO: Sistemas de Información Geográfica para controlar proyectos (Moya, Ernesto, León Companioni, Piñero Pérez, & Romillo Tarke, 2016).
14. An algorithm evaluation for discovering classification rules with gene expression programming (Guerrero-Enamorado, Morell, Noaman, & Ventura, 2016).
15. Experiencias en la integración de procesos en las organizaciones orientadas a proyectos de software. (Villavicencio Bermúdez, Peña Abreu, Burneo Valareso, & Pérez Pupo, 2016)
16. Red neuronal multicapa para la evaluación de competencias laborales. (López, Cuza, Pérez, & Diéguez, 2016)
17. Rough Sets for Human Resource Competence Evaluation and Experiences. (López, Aguilar, et al., 2016)
18. Aplicación de la minería de datos anómalos en organizaciones orientadas a proyectos. (Castro Aguilar, Pérez Pupo, Piñero Pérez, Martínez, & Cruz Castillo, 2016).
19. Método para el aseguramiento de ingresos en entornos de desarrollo de software. (Castro Aguilar, Pérez Pupo, Piñero Pérez, & García Vacacela, 2016)
20. Creation and evaluation of software teams a social approach. (Torres, Piñero, Piñero, & Capretz, 2015)
21. Técnicas formales y de inteligencia artificial para la gestión de recursos humanos en proyectos informáticos. (Torres-López et al., 2014)
22. Gestión de recursos humanos para centros de desarrollo de sistemas de información. (López, García, Pérez, & Fuentes, 2014).

A continuación, se muestra el estudio realizado sobre calidad percibida de los investigadores que han utilizado datos del repositorio. Los investigadores involucrados suman 57 personas entre estudiantes de maestría, estudiantes de doctorado e investigadores. Se diseñan una encuesta asociada al uso del repositorio que evalúa las siguientes variables: pertinencia del repositorio y la usabilidad. Las preguntas son formuladas de forma tal que las respuestas corresponden a alguno de los siguientes términos lingüísticos $LBTL = \{ninguno, muy\ bajo, bajo, medio, alto, muy\ alto, perfecto\}$. Las respuestas de los investigadores a cada pregunta son agregadas y computadas empleando la técnica de computación con palabras 2-tuplas (Herrera & Martínez, 2000).

Los resultados obtenidos se muestran a continuación.

Tabla 3. Análisis de criterios respecto a la variable pertinencia.

	Valoración
¿En qué grado su investigación estuvo relacionada con los datos del repositorio?	Alto, 0.02
Evalúe el nivel de cubrimiento de áreas de conocimiento en gestión de proyectos.	Alto, -0.43
¿En qué medida el repositorio facilitó la validación de los resultados de su investigación?	Alto, 0.04
Indique el impacto que pueda tener el repositorio en redes de colaboración.	Alto, 0.01
Indique el nivel de exclusividad del repositorio considerando la existencia de otros repositorios similares públicos para su uso.	Medio, -0.02
¿En qué grado considera el nivel de impacto del repositorio?	Alto,- 0.03

Tabla 4: Análisis de criterios respecto a la usabilidad.

	Valoración
¿Estructura del repositorio es adecuada para facilitar su explotación?	Muy alto, -0.02
Evalúe el nivel de accesibilidad del repositorio en la red.	Medio, -0.43
Evalúe el nivel de complejidad en el trabajo con el repositorio.	Bajo, 0.36
Indique el nivel de facilidad para incorporación de nuevas bases de datos al repositorio	Alto , 0.3
Evalúe el nivel de satisfacción del sistema de ayuda y/o descripción del repositorio	Medio, 0.29
Evalúe el nivel de calidad de los datos respecto a la cantidad de valores ausentes.	Medio, 0.4
Evalúe el nivel de consistencia de los registros de datos	Alto , -0.3

En general se manifiesta un alto nivel de satisfacción en el uso del repositorio. Respecto a la variable pertinencia el criterio con menor puntuación fue el nivel de exclusividad del repositorio elemento que se debe a la existencia del repositorio PSPLIB para el desarrollo de investigaciones asociadas a la resolución de problemas de optimización. Pero realmente no se ha encontrado por parte de los autores de este trabajo otros repositorios públicos para el desarrollo de investigaciones en gestión de proyectos de software. Respecto a la variable usabilidad los criterios con más baja evaluación fueron: el nivel de accesibilidad del repositorio, la calidad de los datos respecto a la cantidad de valores ausentes y la satisfacción respecto a la ayuda en el uso del repositorio. Son elementos que se deben continuar trabajando por parte de los autores para garantizar mayores niveles de uso del mismo.

Conclusiones

La construcción del repositorio de investigaciones en gestión de proyectos posibilita el desarrollo de disímiles investigaciones en esta área temática. En particular el repositorio presentado en esta investigación contiene 18 bases

de datos asociadas a proyectos de software y cubren la mayoría de las áreas de conocimiento de gestión de proyectos con las plantean el PMBOK y la ISO 21500. En general se han desarrollado más de 20 publicaciones que usan los datos del repositorio, así como tesis de doctorado y maestría. Los investigadores beneficiados con el uso del repositorio manifiestan un alto grado de satisfacción con el mismo, considerando tanto la pertinencia como la usabilidad. Los elementos que los autores consideran que se deben continuar trabajando para lograr mejores resultados son: el nivel de accesibilidad del repositorio, la calidad de los datos respecto a la cantidad de valores ausentes y la satisfacción respecto a la ayuda en el uso del repositorio.

Referencias

- (PMI), P. M. I. (2017). *A guide to the project management body of knowledge (PMBOK guide) Sixth Edition / Project Management Institute*. Project Management Institute, Inc. Newtown Square, Pennsylvania 19073-3299 USA.
- Aguilar, G. F. C., Pupo, I. P., Piñero, P. R., Pérez, P. Y. P., & Sosa, R. (n.d.). Plataforma para aseguramiento de ingresos, aplicación en gestión de proyectos y telcos.
- Bermudez Peña, A. (2015). Sistema basado en técnicas de soft computing para la evaluación de la ejecución de proyectos.
- Blake, C. (1998). UCI repository of machine learning databases. *Http://www. Ics. Uci. Edu/~mlearn/MLRepository. Html*.
- Blake, C., & Merz, C. (2015). UCI repository of machine learning databases, Department of Information and Computer Science, University of California, Irvine, CA, 1998. *URL:< Http://www. Archive. Ics. Uci. Edu/ml*.
- Boaventura José, C., Peña Herrera, E., Verdecia Vicet, P., & Fustiel Alvarez, Y. (2016). Elección entre una metodología ágil y tradicional basado en técnicas de soft computing. *Revista Cubana de Ciencias Informáticas, 10*, 145–158.
- Castro Aguilar, G. F. (2017a). *Modelo para el aseguramiento de ingresos en organizaciones orientadas a proyectos basado en minería de datos anómalos*. Universidad Católica de Santiago de Guayaquil, La Habana, Cuba.
- Castro Aguilar, G. F. (2017b). *Modelo para el aseguramiento de ingresos en organizaciones orientadas a proyectos basado en minería de datos anómalos*. Universidad Católica de Santiago de Guayaquil, La Habana, Cuba.
- Castro Aguilar, G. F., Pérez Pupo, I., Piñero Pérez, P. Y., & García Vacacela, R. (2016). Método para el aseguramiento de ingresos en entornos de desarrollo de software. *Revista Cubana de Ciencias Informáticas, 10*, 43–57.
- Castro Aguilar, G. F., Pérez Pupo, I., Piñero Pérez, P. Y., Martínez, N., & Cruz Castillo, Y. (2016). Aplicación de la minería de datos anómalos en organizaciones orientadas a proyectos. *Revista Cubana de Ciencias Informáticas, 10*, 195–209.
- Escobar Pompa, M. E. (2015). Procedimiento de limpieza de datos en el GESPRO.
- Fernández Hernández, A. (2016). *Modelo ontológico de recuperación de información para la toma de decisiones en gestión de proyectos*. Universidad de Granada, Ciudad de Granada, España.

- García, J. A. L., Peña, A. B., Pérez, P. Y. P., & Pérez, R. B. (2017). Project Control and Computational Intelligence: Trends and Challenges. *International Journal of Computational Intelligence Systems*, 10(1), 320–335.
- García Vacacela, R., Pérez Pupo, I., Villavicencio, N., Piñero, P. Y., & Beovides, S. (2016). Experiencias usando algoritmos genéticos en la planificación de proyectos. *Revista Cubana de Ciencias Informáticas*, 10, 71–86.
- Gorunescu, F. (2011). *Data Mining: Concepts, models and techniques* (Vol. 12). Springer Science & Business Media.
- Guerrero-Enamorado, A., Morell, C., Noaman, A. Y., & Ventura, S. (2016). An algorithm evaluation for discovering classification rules with gene expression programming. *International Journal of Computational Intelligence Systems*, 9(2), 263–280.
- Hasan Al-Subhi, S., Pérez Pupo, I., García Vacacela, R., Piñero Pérez, P., & Yelandy Leyva, M. (2018). A New Neutrosophic Cognitive Map with Neutrosophic Sets on Connections, Application in Project Management. *Neutrosophic Sets and Systems*, 21.
- Hastie, S., & Wojewoda, S. (2015). Standish group 2015 chaos report-q&a with Jennifer Lynch. *Retrieved*, 1(15). Retrieved from <https://www.infoq.com/articles/standish-chaos-2015>
- Herrera, F., & Martínez, L. (2000). A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 8(6), 746–752.
- Jiménez Moya, G. E. (2017). *Modelo de control de proyectos basado en análisis geo-referencial para la ayuda a la toma de decisiones*. Universidad Católica de Santiago de Guayaquil, La Habana, Cuba.
- Johnson, J., Gesmer, L., Poort, J., & Mulder, H. (2016). Chaos Report 2016. *Special Chaos Report on Digital Transformation Project*.
- Kolisch, R., & Sprecher, A. (1996). PSPLIB- A project scheduling problem library. *European Journal of Operation Research*, Elsevier, 96, 205–216.
- López, S. T., Aguilar, G. F. C., Pupo, I. P., Pérez, P. P., & Diéguez, L. P. (2016). Rough Sets for Human Resource Competence Evaluation and Experiences. *Applied Mathematics*, 7(12), 1317–1325. doi:10.4236/am.2016.712116
- López, S. T., Cuza, M. L. A., Pérez, P. Y. P., & Diéguez, L. A. P. (2016). Red neuronal multicapa para la evaluación de competencias laborales. *Revista Cubana de Ciencias Informáticas*, 10, 210–223.
- López, S. T., García, J. A. L., Pérez, P. Y. P., & Fuentes, A. P. (2014). Gestión de recursos humanos para centros de desarrollo de sistemas de información. *Ciencias de La Información*.
- Lugo García, J. A. (2012). Modelo para el control de la ejecución de proyectos basado en indicadores y lógica borrosa.
- Lugo García, J. A. (2015). *Modelo para el control de la ejecución de proyectos basado en soft computing*. Universidad de las Ciencias Informáticas, La Habana, Cuba.
- Medina Rodríguez, M. (2012). Propuesta de modelo de un almacén de datos para el control del tiempo y los costos de proyectos.
- Moya, J., Ernesto, G., León Companioni, A., Piñero Pérez, P. Y., & Romillo Tarke, A. (2016). SIGESPRO: Sistemas de Información Geográfica para controlar proyectos. *Revista Cubana de Ciencias Informáticas*, 10(2), 181–195.
- Pacheco Rodríguez, E. (2014). Descubrimiento de conocimiento a partir de la relación rasgos de la personalidad–rendimiento laboral en proyectos informáticos.

- Peña Abreu, M. (2012). Modelo para análisis de factibilidad en la evaluación de proyectos de software.
- Peña Abreu, M. (2017). *Modelo para el análisis de factibilidad de proyectos de software en entornos de incertidumbre*. Universidad de las Ciencias Informáticas, La Habana, Cuba.
- Perdomo Alonso, A. (2013). Modelo para la evaluación de competencias como problemas multietiquetas a partir de sistemas basados en casos.
- Pérez Pupo, I. (2011). Propuesta de metodología para el diseño e implantación de repositorios de activos de software reutilizables.
- Pérez Pupo, I., Piñero Pérez, P., García Vacacela, R., Santos Acosta, O., & Yelandy Leyva, M. (2018). Extensions to Linguistic Summaries Indicators based on Neutrosophic Theory, Applications in Project Management Decisions. *Neutrosophic Sets and Systems*, 21.
- Piñera Trinchet, E. E. (2013). Algoritmo de sumarización lingüística como apoyo a la toma de decisiones en gestión de proyecto.
- Piñero, P., Torres, S., Menéndez, J., Gonzáles, R., Dexter, H., Lugo, J. A., & Others. (2015). Solución de software XEDRO GESPRO v13.05. Centro Nacional de Registro de Derecho de Autor de Cuba, La Habana.
- Piñero, P., Torres, S., Menéndez, J., Gonzáles, R., Dexter, H., & others. (2010). Paquete de Herramientas para la Gestión de Proyectos GESPRO. Centro Nacional de Registro de Derecho de Autor de Cuba, La Habana.
- Pupo, I. P., Acosta, O. S., Vacacela, R. G., Pérez, P. Y. P., & Ramírez, E. C. (2018). Descubrimiento de resúmenes lingüísticos para ayuda a la toma de decisiones en gestión de proyectos. *Revista Cubana de Ciencias Informáticas*, 12, 163–175.
- Pupo, I. P., Gómez, P. L., Varona, E. P., Pérez, P. Y. P., & Vacacela, R. G. (2018). Construcción de resúmenes lingüísticos a partir de rasgos de la personalidad y el desempeño en el desarrollo de software. *Revista Cubana de Ciencias Informáticas*, 12, 135–150.
- Ril Valentín, E. B. (2012). Descubrimiento de conocimiento a partir de lecciones aprendidas documentadas en los procesos de cierre de proyectos informáticos.
- Rodríguez Rodríguez, C. R., Peña Abreu, M., Castro Aguilar, G. F., Pérez, P., & Yobanis, P. (2017). Sistema clasificador borroso basado en algoritmos genéticos para evaluar el estado de ejecución de proyectos. *Revista Cubana de Ciencias Informáticas*, 11(3), 174–189.
- Rodríguez Stiven, E. (2012). Modelo para la evaluación de la composición de equipos de proyectos informáticos.
- Santiesteban, A. M., Castro, F., Rojas, C., C, R., Piñero, P., Pérez, I., & Rodríguez, O. (2016). Repositorio para el desarrollo de investigaciones en Gestión de Proyectos.
- SEI, S. E. I. (2015). *CMMI for Development version 1.3*.
- Sosa González, R., Pérez Pupo, I., García, R., Peñaherrera, E., & Piñero Pérez, P. Y. (2016). Ecosistema de Software GESPRO-16.05 para la Gestión de Proyectos. *Revista Cubana de Ciencias Informáticas*, 10, 239–251.
- Stellingwerf, R., & Zandhuis, A. (2013). *ISO 21500 Guidance On Project Management: A Pocket Guide (Best Practice)*. Van Haren.
- Torres López, S. (2015). *Modelo de evaluación de competencias a partir de evidencias durante la gestión de proyectos*. Universidad de las Ciencias Informáticas, La Habana, Cuba.
- Torres, S., Piñero, Y., Piñero, P., & Capretz, L. F. (2015). Creation and evaluation of software teams-a social approach. *arXiv Preprint arXiv:1512.00787*.

- Torres-López, S., Lugo-García, J. A., Piñero-Pérez, P. Y., Torres-Quñones, K. M., Perdomo-Alonso, A., Cuza-García, B., & Aldana-Cuza, M. L. (2014). Técnicas formales y de inteligencia artificial para la gestión de recursos humanos en proyectos informáticos. *Revista Cubana de Ciencias Informáticas*, 8(3), 43–55.
- Villavicencio Bermúdez, N., Peña Abreu, M., Burneo Valareso, S., & Pérez Pupo, I. (2016). Experiencias en la integración de procesos en las organizaciones orientadas a proyectos de software. *Revista Cubana de Ciencias Informáticas*, 10, 171–185.