Memorias del Congreso Internacional de Investigación Academia Journals Morelia 2019

Elibro Online con ISSN
1946-5351, Volumen 11, No. 2, 2019

Morelia, Michoacán, México
15 al 17 de mayo, 2019

www.AcademiaJournals.com

ACADEMIA JOURNALS
OPUS PRO SCIENTIA ET STUDIUM

© Academia Journals 2019

	Título	Autores	Primer Autor	Pág.
MRL0348	Obtención de materiales electrohilados polímero-nanoparticulas con propiedades bactericidas	Meléndez-Estrada Idahli Alejandra Muñoz-Escobar Antonio de Jesús Garibay-Alvarado Jesús Alberto Lang-Salas Michelle Gerardo Cuevas-González Juan Carlos Donohué-Cornejo Alejandro Espinosa-Cristóbal León Francisco Reyes-López Simón Yobanny	Meléndez-Estrada	1789
MRL0744	OBTENCIÓN DE FIBRAS DE ZIRCONIA COMO MATERIAL ADSORBENTE	Idahli Alejandra Melendez Estrada Ana Karen Ávila-Martínez José Hafid Roque-Ruiz Jonatan Torres Pérez Nahúm A. Medellín-Castillo Simón Yobanny Reyes-López	Melendez Estrada	1795
MRL0071	Introducción a las Industrias 4.0	C. Silvia Ibianyi Méndez García C. Daniela Pineda López MCC. Salvador Martínez Pagola M.T.E. Eric León Olivares Ing. Juan Alejandro Arrieta Zúñiga	Méndez García	1801
MRL0029	Identificación de hongos resistentes a mercurio, extraídos de la presa “El Pedernalillo” en el Estado de Zacatecas	MC Rubén Octavio Méndez Márquez QFB Daniela Lizbeth Troncoso Vázquez QFB Priscila Alejandra Carrillo Escobedo	Méndez Márquez	1808
MRL0522	SISTEMA PARA VERIFICAR LA AUTENTICIDAD DE LOS TRABAJOS ENTREGADOS EN FORMATO DIGITAL PARA OBTENER EL GRADO DE LICENCIATURA EN EL INSTITUTO TECNOLÓGICO DE ACAPULCO BASADO EN EL ALGORITMO DE MIYER	ING. CRISOL ANGELINA MENDIOLA PIZA M.T.I. ELOY CADENA MENDOZA M.T.I JUAN MIGUEL HERNÁNDEZ BRAVO M.T.I RAFAEL HERNANDEZ REYNA	MENDIOLA PIZA	1814
MRL0548	Revisión literaria concerniente al análisis semántico de requerimientos de un producto de software escritos en lenguaje natural mediante el método de Procesamiento de Lenguaje Natural	M. en S.E. Carlos Octavio Mendívil Vázquez Dr. Luis Corral	Mendívil Vázquez	1820

ISSN 1946-5351
Vol. 11, No. 2, 2019

© Academia Journals 2019

Revisión literaria concerniente al análisis semántico de requerimientos de un producto de software escritos en lenguaje natural mediante el método de Procesamiento de Lenguaje Natural

M. en S.E. Carlos Octavio Mendívil Vázquez[1], Dr. Luis Corral[2]

***Resumen*— La ingeniería de requerimientos representa una actividad de suma importancia durante el desarrollo de un producto de software. Por ello, el proceso de recolección, manejo y liberación de requerimientos adquiere gran relevancia a lo largo del desarrollo de software hasta su entrega final. Con el fin de encontrar un área de investigación abierta para contribuir en el manejo de requerimientos, se presenta una revisión literaria concerniente al análisis semántico de requerimientos escritos en lenguaje natural mediante el método de Procesamiento de Lenguaje Natural (NLP). Como resultado, se propone una clasificación de las metodologías empleadas para llevar a cabo el análisis semántico en requerimientos en formato textual en lenguaje natural. Asimismo, se encuentra que las propuestas basadas en combinaciones de diferentes métodos de NLP son las predominantes, ya que son versátiles y han presentado mejoras en sus resultados en los últimos años, particularmente cuando se emplea aprendizaje profundo (Deep Learning).**

***Palabras clave*—Requerimientos de software, NLP, Análisis semántico**

Introducción

La ingeniería de requerimientos representa una actividad de suma importancia durante el desarrollo de un producto o proyecto de software. Particularmente en el caso de proyectos de software, el conjunto de requerimientos simboliza la entrada para el resto de los procesos y etapas del Ciclo de Vida de Desarrollo de Software (SDLC por sus siglas en inglés). A continuación se enlistan algunas de las etapas de un SDLC general en modelo "V" que dependen directa o indirectamente de la ingeniería de requerimientos, los detalles de cada etapa se pueden encontrar en (Jolliffe, 2010):

- Análisis de Requerimientos del cliente
- Diseño y Arquitectura a nivel sistema
- Evaluación de los Requerimientos de SW
- Diseño y Arquitectura de SW
- Implementación del SW
- Pruebas de Integración de HW & SW
- Pruebas de Integración a nivel sistema
- Pruebas y Validación a nivel sistema.

Por lo tanto el estudio de la ingeniería de requerimientos puede ser una estrategia para ahorrar tiempo y dinero en el mismo proyecto así como obtener resultados de mayor calidad en procesos subsecuentes involucrados en el SDLC en Gral (Pohl, 2010). Particularmente, el manejo de requerimientos presenta una oportunidad de estudio en la ingeniería de requerimientos (Li, Yue, Ali, & Zhang, 2017a). De acuerdo a (Chemuturi, 2013) las actividades realizadas en el manejo de requerimientos son las siguientes:

- Documentación de Requerimientos
- Análisis de requerimientos
- Rastreabilidad de requerimientos
- Priorizar requerimientos
- Aceptación de requerimientos
- Controlar cambios de requerimientos
- Comunicar el progreso de la implementación de los requerimientos.

Hay dos factores en común de las actividades listadas anteriormente, el primero consiste en identificar el significado del requerimiento en sí, es decir, el ingeniero debe llevar a cabo un análisis semántico de los requerimientos con el objetivo de comprender el contenido del requerimiento y así sucesivamente tomar las decisiones necesarias para llevar a cabo cada actividad. Por ejemplo, para determinar las relaciones de rastreabilidad

[1] M. en S.E. Carlos Octavio Mendívil Vázquez es Estudiante de Posgrado en CIATEQ A. C., Querétaro, Querétaro. cs.mendivil@gmail.com**(autor corresponsal)**
[2] Dr. Luis Corral ITESM, lrcorralv@tec.mx

© Academia Journals 2019

que existen entre un requerimiento y otros artefactos de software (documentos de diseño, código fuente, etc.), así como para analizar el impacto de un cambio en un grupo de requerimientos. El segundo factor es que los requerimientos que son objeto de estas actividades son capturados tradicionalmente de manera textual en lenguaje natural (NL por sus siglas en inglés).

Descripción del Método

Con el fin de encontrar un área de investigación abierta para contribuir en el manejo de requerimientos, se ha realizado un análisis de la literatura concerniente al análisis semántico de requerimientos escritos en lenguaje natural mediante el método de procesamiento de lenguaje natural (NLP por sus siglas en inglés). La búsqueda se ha llevado a cabo en el periodo elegido de 2012-2018 en las siguientes bases de datos de artículos de investigación:

- Springer
- Science Direct
- IEEE

Al principio se emplearon los términos de búsqueda "semantic" y "nlp", pero los resultados arrojaban artículos no relevantes al tema de requerimientos, en consecuencia, se decidió agregar el término "requirement". Posteriormente se identificó que los artículos más relevantes mencionaban recurrentemente el término de similitud semántica, por lo cual se tomó la decisión de emplear "semantic similarity" para acotar aún más el número de resultados. En la Tabla 1 se muestran los resultados obtenidos por cada base de datos:

Base de Datos	Artículos Relevantes encontrados
Springer	23
Science Direct	15
IEEE	21

Tabla 1. Resultados de búsqueda bibliográfica en bases de datos de artículos de investigación.

Posteriormente se procedió a seleccionar los artículos relevantes para la investigación. Como resultado se obtuvo una clasificación con respecto a las metodologías empleadas para llevar a cabo el análisis semántico en requerimientos en formato textual en lenguaje natural. Así mismo también se añadieron referencias, que aunque se salen del periodo establecido de búsqueda (2012-2018) o no contienen todos los términos de búsqueda, se consideró pertinente agregarlas a la lista de artículos relevantes del conjunto bibliográfico de la presente.

Comentarios Finales

Se descartaron de lista final algunos artículos cuyo término de búsqueda no era aplicado en el contexto relevante para esta investigación, por ejemplo, el término "requirement" no es empleado como objeto de análisis, si no como una forma de identificar una característica pertinente al estudio que el mismo artículo aborda.

Resumen de resultados

Las clasificaciones de los resultados se pueden observar a continuación:

- Vector de similitud con análisis estadístico (Croft, Coupland, Shell, & Brown, 2013)
- Modelos combinados originales (Wu, Huang, & Yang, 2017)
- Wordnet como base de la propuesta (Matsuoka & Lepage, 2011) (Sagala, Wati, Budi, & Hidayanto, 2018)
- Ontología como base de la propuesta (Jayatilleke & Lai, 2013) (Y. Wang, 2016)
- NLP combinado con otros métodos (Rago, Marcos, & Diaz-Pace, 2016), (Arora, Sabetzadeh, Briand, & Zimmer, 2015), (Hajri, Goknil, Briand, & Stephany, 2018), (Arora, Sabetzadeh, Goknil, Briand, & Zimmer, 2015), (Priya & Anupriya, 2013), (Mahmoud & Williams, 2016), (Li, Yue, Ali, & Zhang, 2017b), (Lucassen, Dalpiaz, van der Werf, & Brinkkemper, 2016), (Unterkalmsteiner, Gorschek, Feldt, & Lavesson, 2016), (Guo, Cheng, & Cleland-Huang, 2017), (Effa Bella, Gervais, Bendraou, Wouters, & Koudri, 2018), (McZara, Sarkani, Holzer, & Eveleigh, 2015), (Misra, 2016)

© Academia Journals 2019

Mientras que en la

Tabla 2 se puede observar el número de ocurrencias por cada clasificación:

Clasificación	No. De Ocurrencias
Vector de similitud con análisis estadístico	1
Wordnet	2
Ontología como base de la propuesta	2
Modelos combinados	14

Tabla 2 - Resultados de recopilación bibliográfica

A continuación, se describe brevemente cada clasificación:

- **Vector de similitud con análisis estadístico**: se basa en emplear vectores para determinar el grado de similitud entre los términos analizados y así determinar su relación estadísticamente.
- **Wordnet**: hace referencia un grupo de palabras en inglés relacionadas semánticamente. Por lo general es usada como una ontología.
- **Ontologías** propuestas se emplean como base para determinar la relación semántica entre los requerimientos, a veces auxiliado por otros métodos para complementar los resultados como por ejemplo análisis estadístico.
- **Modelos combinados:** Se puede considerar como una serie de metodologías o procesos cuyo objetivo es abstraer el contenido semántico del texto y transformarlo a un formato procesable por algoritmos computacionales empleando dos o mas herramientas o metodologías complementarías entre sí. Es decir, busca ofrecer una solución alterna a sólo emplear Wordnet, ontologías o métodos estadisticos, por ejemplo propone una combinación de Deep Learning y Feature Engineering Model.

Conclusiones

Como se puede observar hay un bajo número de resultados relacionados con análisis estadístico, Wordnet (grupo de palabras) y propuesta de ontologías. Esto se debe principalmente a que son soluciones que son difíciles de trasladar a otro dominio. Por otra parte, también es visible el predominio de las propuestas basadas en combinaciones de diferentes métodos de NLP, ya que son versátiles y han presentado mejoras en sus resultados en los últimos años, particularmente cuando se emplea "Deep Learning". Asimismo, se ha identificado que la mayoría de los requerimientos son capturados de manera textual en lenguaje natural en inglés.

Recomendaciones

Por lo tanto, en el presente se propone contribuir a la ingeniería de requerimientos mediante el desarrollo de un proceso que permita emplear NLP basada en "Deep Learning" que concierne a la rastreabilidad requerimientos, ya que ésta representa un campo de oportunidad de estudio (B. Wang, Peng, Li, Lai, & Wang, 2018).

Referencias

Arora, C., Sabetzadeh, M., Briand, L., & Zimmer, F. (2015). Automated Checking of Conformance to Requirements Templates Using Natural Language Processing. *IEEE Transactions on Software Engineering*, *41*(10), 944–968. https://doi.org/10.1109/TSE.2015.2428709

Arora, C., Sabetzadeh, M., Goknil, A., Briand, L. C., & Zimmer, F. (2015). Change impact analysis for natural language requirements: An NLP approach. In *Requirements Engineering Conference (RE), 2015 IEEE 23rd International* (pp. 6–15). IEEE.

Chemuturi, M. (2013). Introduction to Requirements Engineering and Management. In M. Chemuturi, *Requirements Engineering and Management for Software Development Projects* (pp. 1–11). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-5377-2_1

Croft, D., Coupland, S., Shell, J., & Brown, S. (2013). A fast and efficient semantic short text similarity metric. In *Computational Intelligence (UKCI), 2013 13th UK Workshop on* (pp. 221–227). IEEE.

Effa Bella, E., Gervais, M.-P., Bendraou, R., Wouters, L., & Koudri, A. (2018). Semi-Supervised Approach for Recovering Traceability Links in Complex Systems. In *2018 23rd International Conference on Engineering of Complex Computer Systems (ICECCS)* (pp. 193–196). Melbourne, VIC: IEEE. https://doi.org/10.1109/ICECCS2018.2018.00030

Guo, J., Cheng, J., & Cleland-Huang, J. (2017). Semantically Enhanced Software Traceability Using Deep Learning Techniques. In *2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE)* (pp. 3–14). Buenos Aires: IEEE. https://doi.org/10.1109/ICSE.2017.9

Hajri, I., Goknil, A., Briand, L. C., & Stephany, T. (2018). Change impact analysis for evolving configuration decisions in product line use case models. *Journal of Systems and Software*, *139*, 211–237. https://doi.org/10.1016/j.jss.2018.02.021

Jayatilleke, S., & Lai, R. (2013). A Method of Specifying and Classifying Requirements Change. In *2013 22nd Australian Software Engineering Conference* (pp. 175–180). Hawthorne, Victoria, Australia: IEEE. https://doi.org/10.1109/ASWEC.2013.29

Jolliffe, G. (2010). Cost-efficient methods and processes for safety relevant embedded systems (CESAR)–an objective overview. In *Making Systems Safer* (pp. 37–50). Springer.

Li, Y., Yue, T., Ali, S., & Zhang, L. (2017a). Enabling automated requirements reuse and configuration. *Software & Systems Modeling*. https://doi.org/10.1007/s10270-017-0641-6

Li, Y., Yue, T., Ali, S., & Zhang, L. (2017b). Enabling automated requirements reuse and configuration. *Software & Systems Modeling*. https://doi.org/10.1007/s10270-017-0641-6

Lucassen, G., Dalpiaz, F., van der Werf, J. M. E. M., & Brinkkemper, S. (2016). Improving agile requirements: the Quality User Story framework and tool. *Requirements Engineering*, *21*(3), 383–403. https://doi.org/10.1007/s00766-016-0250-x

Mahmoud, A., & Williams, G. (2016). Detecting, classifying, and tracing non-functional software requirements. *Requirements Engineering*, *21*(3), 357–381. https://doi.org/10.1007/s00766-016-0252-8

Matsuoka, J., & Lepage, Y. (2011). Ambiguity spotting using wordnet semantic similarity in support to recommended practice for software requirements specifications. In *Natural Language Processing andKnowledge Engineering (NLP-KE), 2011 7th International Conference on* (pp. 479–484). IEEE.

McZara, J., Sarkani, S., Holzer, T., & Eveleigh, T. (2015). Software requirements prioritization and selection using linguistic tools and constraint solvers—a controlled experiment. *Empirical Software Engineering*, *20*(6), 1721–1761. https://doi.org/10.1007/s10664-014-9334-8

Misra, J. (2016). Terminological inconsistency analysis of natural language requirements. *Information and Software Technology*, *74*, 183–193. https://doi.org/10.1016/j.infsof.2015.11.006

Pohl, K. (2010). *Requirements engineering: fundamentals, principles, and techniques*. Springer Publishing Company, Incorporated.

Priya, G. K., & Anupriya, G. (2013). Clustering sentence level-text using fuzzy hierarchical algorithm. In *Human Computer Interactions (ICHCI), 2013 International Conference on* (pp. 1–8). IEEE.

Rago, A., Marcos, C., & Diaz-Pace, J. A. (2016). Assisting requirements analysts to find latent concerns with REAssistant. *Automated Software Engineering*, *23*(2), 219–252. https://doi.org/10.1007/s10515-014-0156-0

Sagala, T. W., Wati, T., Budi, N. F. A., & Hidayanto, A. N. (2018). Analysis and Implementation Measurement of Semantic Similarity Using Content Management Information on WordNet. In *2018 International Conference on Advanced Computer Science and Information Systems (ICACSIS)* (pp. 337–342). IEEE.

Unterkalmsteiner, M., Gorschek, T., Feldt, R., & Lavesson, N. (2016). Large-scale information retrieval in software engineering - an experience report from industrial application. *Empirical Software Engineering*, *21*(6), 2324–2365. https://doi.org/10.1007/s10664-015-9410-8

Wang, B., Peng, R., Li, Y., Lai, H., & Wang, Z. (2018). Requirements traceability technologies and technology transfer decision support: A systematic review. *Journal of Systems and Software*, *146*, 59–79. https://doi.org/10.1016/j.jss.2018.09.001

Wang, Y. (2016). Automatic semantic analysis of software requirements through machine learning and ontology approach. *Journal of Shanghai Jiaotong University (Science)*, *21*(6), 692–701. https://doi.org/10.1007/s12204-016-1783-3

Wu, D., Huang, J., & Yang, S. (2017). A Joint Model for Sentence Semantic Similarity Learning. In *2017 13th International Conference on Semantics, Knowledge and Grids (SKG)* (pp. 120–125). Beijing, China: IEEE. https://doi.org/10.1109/SKG.2017.00027