Universidad Nacional de Mar del Plata.

Facultad de Humanidades.

Departamento de Ciencia de la Información.

Licenciatura en Bibliotecología y Documentación.

Seminario B: Temas principales con la Recuperación de la Información

COMPILACION Y ANALISIS BIBLIOGRAFICO

SRI: procesos preparatorios

2019

PRATONE, CAROLINA JEZABEL

Docente: Di Muro, Enzo

Compilación y análisis de bibliografía de nivel académico: sistemas de recuperación de información.

Este es el trabajo escrito para la catedra Seminario B con el docente Enzo Di Muro, de todas formas el trabajo final estará disponible en el siguiente link: https://carolinapratone.wixsite.com/srirecopilacion donde en él se establecen los hipervínculos y cuestiones estéticas.

Introducción.

Esta es una recopilación bibliográfica realizada para la catedra Seminario B: Temas principales de la Recuperación de Información, de la Licenciatura de Bibliotecología y Documentación de la Universidad Nacional de Mar del Plata en el marco de un trabajo final.

La misma se expresa en citas APA 6ta edición, y los descriptores fueron extraídos del tesauro online de la UNESCO (http://vocabularies.unesco.org/browser/thesaurus/es/). Todos los documentos fueron incluidos con la aprobación de ciertos requisitos, por ejemplo, estar disponible en acceso abierto y a texto completo, estar escrito en español, publicado entre los años 2005 y 2019, por último, que pertenezca a Iberoamérica.

Además, esta recopilación busca que cualquier usuario de la web tenga un recurso a mano de información confiable y verídica sobre la temática de Sistemas de Recuperación de Información y sus procesos preparatorios, para esto se dividió la recopilación en cuatro secciones:

- Conceptos Generales: se incorporaron aquellas referencias a todos los documentos que traten de forma general y marcando los conceptos principales, de forma académica y sencilla de comprender desde distintas aristas como las ciencias de la información, o la informática, la temática amplia de Sistemas de Recuperación de Información.
- Lenguaje de Marcado: es una manera de codificar formalmente los textos para que los sistemas los puedan reconocer, entre los más conocidos y tratados en este apartado están el HTML, SGML y XML.
- Operaciones con los textos: este apartado hace referencia a todos los procesos necesarios de preparación de documentos para poder hacerlos recuperables por los SRI.
- Procesos Semánticos: en este caso solo nos centramos en los procesos necesarios para extraer metadatos semánticos u ontológicos de los documentos digitales.

Recopilación Bibliográfica (Sistemas de Recuperación de información y los procesos preparatorios).

Conceptos Generales

La recuperación de la información es un área de investigación y desarrollo que se ocupa de la representación, almacenamiento, organización, acceso y visualización de información, la misma se puede ver desde dos lugares, los usuarios: aquellos que la recuperación de información les sirve para satisfacer sus necesidades de información; y desde los sistemas centrado resolver el problema de representar correctamente los documentos para luego cruzar los datos con las ecuaciones de búsqueda de los usuarios.

Desde este punto de vista se puede establecer que los procesos de recuperación de la información se dividen en dos:

Almacenamiento: selección de documentos, representación (metadatos), operaciones con los textos y generación de índices.

Búsqueda: expresión de la necesidad de información, reconocimiento de los componentes, ejecución, respuesta y redefinición en el caso de ser necesario.

En este caso se expone una serie de documentos donde la temática se aborda desde sus aspectos más generales e introductorios.

Fuente: Di Muro, E. & Sanllorenti, A.M. (2018) *Eje 1: Introducción y modelos de la Recuperación de Información.* Universidad de Mar del Plata – LICAD – Seminario B: Temas principales de la Recuperación de Información: Buenos Aires.

Bordignon, F. R., & Tolosa, G. H. (2007). Recuperación de información: un área de investigación en crecimiento. *Telematique*, *6*(1), 51-73. Disponible en: http://ojs.urbe.edu/index.php/telematique/article/view/1790/1711

El articulo realizado por investigadores de la Universidad de Lujan, Argentina, lo que expone es de forma introductoria el objeto de estudio de la recuperación de información y su alcance en la investigación.

DESCRIPTORES: BÚSQUEDA EN LÍNEA; RECUPERACIÓN DE INFORMACIÓN; SISTEMA DE INFORMACIÓN

Camaraza Monserrate, A. (2005). Recuperación de información: reflexiones epistémicas de una ciencia en su estado embrionario. *Acimed*, *13*(6). Disponible en: http://scielo.sld.cu/pdf/aci/v13n6/aci100605.pdf

En este artículo se puede conocer la historia del surgimiento de la recuperación de la información como campo formal del conocimiento y su evolución a lo largo de los estudios científicos, lo que nos lleva a incluirlo por su carácter introductorio y su análisis histórico como parte de la agenda científica.

DESCRIPTORES: PROCESAMIENTO DE LA INFORMACIÓN; SISTEMA DE INFORMACIÓN; RECUPERACIÓN DE INFORMACIÓN

Centelles, M. (2005). Taxonomías para la categorización y la organización de la información en sitios web. *Hipertext. net*, *3*. Disponible en: https://www.upf.edu/hipertextnet/numero-3/taxonomias.html

Encontramos en este articulo el concepto de creación de taxonómicas para la correcta descripción de los sitios web y como estos deben utilizarse, se incluye en esta sección por su carácter introductorio y esclarecedor de uno de los temas importantes de los SRI.

DESCRIPTORES: LENGUAJE DE INDEXACIÓN; BÚSQUEDA EN LÍNEA; TERMINOLOGÍA

Cuba, Y., & Batista, D. O. (2018). Los metadatos, la búsqueda y recuperación de información desde las Ciencias de la Información. *e-Ciencias de la información*, 8(2), 3-13. Disponible en: https://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S1659-41422018000200146&lang=es

El presente artículo marca conceptos introductorios en torno a las TIC (Tecnologías de la Información y de la Comunicación) y el acceso a la información, también analiza las

teorías concernientes a la representación de las necesidades de información (estrategias de búsqueda), la interacción usuario-sistema y la descripción en términos de metadatos de dichos sistemas, su principal valor se encuentra en la descripción completa y clara de la teoría en realización al termino metadato.

DESCRIPTORES: BASE DE DATOS; BÚSQUEDA EN LÍNEA; RECUPERACIÓN DE INFORMACIÓN

García, A. A. R. (2010). Elementos para reflexionar sobre el nuevo estándar para la descripción y acceso a recursos. *Biblioteca universitaria*, *13*(1), 55-63. Disponible en: http://revistas.unam.mx/index.php/rbu/article/viewFile/23925/22503

En este artículo se analiza los elementos de las nuevas reglas de la catalogación bibliográfica (RDA) entorno a la descripción en los sistemas de recuperación de información y como estos nuevos estándares corresponden al cambio paradigmático en el uso y gestión de la información sobre todo con la llegada de internet, la web semántica y las tecnologías 2.0. El mayor aporte se encuentra en las explicaciones breves y concretas con respecto al cambio paradigmático expuesto en los últimos tiempos con respecto a la descripción bibliográfica.

DESCRIPTORES: PROCESAMIENTO DE LA INFORMACIÓN; PROCESAMIENTO DE DATOS; CATALOGACIÓN

Gómez, F. L. (2007). Interoperabilidad en los Sistemas de Información Documental (SID): la información debe fluir. *Revista Códice, 3*(1), 23-39. Disponible en: https://cnb.gov.co/codices/online/Vol3-2007I.pdf

Este articulo tiene como principal interés informar conceptualmente la noción de interoperabilidad entre los sistemas de información, planteando un concepto clave para la recuperación de la información y enfocado especialmente en el ámbito bibliotecológico, lo que aporte es una de las distintas aristas centrales de los SRI y como se conjuga con el sistema documental conocido.

DESCRIPTORES: SISTEMA DE INFORMACIÓN EN LÍNEA; BASE DE DATOS

Herrero Solana, V., & Hassan-Montero, Y. (2006). Metodologías para el desarrollo de interfaces visuales de recuperación de información: análisis y comparación. *Information Research*, 11(3), 1-17. Disponible en: http://digibug.ugr.es/handle/10481/32693

El articulo propone una revisión de los diferentes modelos y algoritmos para la presentación de interfaces en la recuperación de información, lo que pretende es dar un amplio espectro teórico de las opciones existentes para la presentación de resultados de búsqueda en cualquier SRI.

DESCRIPTORES: RECUPERACIÓN DE INFORMACIÓN; ALGORITMO; INTERFAZ DE ORDENADORES

Kuna, H. D., Rey, M., Martini, E., Canteros, A., Rambo, A. R., Biale, C. O., & Zamudio, E. (2016). Avances en la Construcción de un Sistema de Recuperación de Información para Información Científica en Ciencias de la Computación. En *XVIII Workshop de Investigadores en Ciencias de la Computación* (WICC 2016, Entre Ríos, Argentina). Disponible en: http://sedici.unlp.edu.ar/handle/10915/53425

Este trabajo se centra en explicar desde el área de las Ciencias de la Computación la recuperación de la información científica y las herramientas que se utilizan para dicho fin, los desarrollos y la utilidad en el usuario de dicha información, explaya además la evaluación de dichas herramientas con un enfoque integral, esto aporta a la comprensión de los sistemas de recuperación de información, las búsquedas y la utilización consciente de metadatos.

DESCRIPTORES: SISTEMA DE INFORMACIÓN EN LÍNEA; LENGUAJE DE PROGRAMACIÓN; DIVULGACIÓN CIENTÍFICA

Kuna, H., Rey, M., Martini, E., Rambo, A., & Podkowa, L. (2015). Avances en el Desarrollo de un Sistema de Recuperación de Información para Publicaciones Científicas del Área de Ciencias de la Computación. *Revista Latinoamericana de Ingeniería de Software*, 3(1), 47-55. Disponible en: http://revistas.unla.edu.ar/software/article/view/406

El articulo presenta los avances de un proyecto de desarrollo de un metabuscador web especializado en las ciencias de la computación, marcando las mejoras en el ámbito de los algoritmos de rankeo, y consultas de usuarios. El valor de este articulo dentro de la recopilación se marca por su explicación en la evaluación de los distintos aspectos de una búsqueda web, las consultas, los metadatos, las respuestas y como se muestran y la interacción completa usuario – sistema.

DESCRIPTORES: SISTEMA DE INFORMACIÓN EN LÍNEA; ALGORITMO; DIVULGACIÓN CIENTÍFICA

Miralles, M. V. J. (2009). El catálogo: un recurso en expansión. *Anales de documentación: Revista de biblioteconomía y documentación,* 12, 69-91. Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=3000805

Se expone el conceto de catálogo bibliográfico como SRI específicamente, resaltando las ventajas de utilizar esta herramienta conocida y de amplio espectro en el mercado para la inclusión de metadatos que favorezcan la recuperación de información en una colección determinada y la visualización de registros bibliográficos como representaciones de documentos.

DESCRIPTORES: SERVICIO BIBLIOGRÁFICO; SISTEMA DE INFORMACIÓN EN LÍNEA; RECUPERACIÓN DE INFORMACIÓN

Pichinini, M. (2007). *Desarrollo de una guía de control para interfaces de búsqueda por materia en catálogos en línea de acceso web* (tesis de grado) Universidad Nacional de La Plata. Facultad de Humanidades y Ciencias de la Educación. Disponible en: http://www.memoria.fahce.unlp.edu.ar/tesis/te.26/te.26.pdf

El trabajo lo que desarrolla es un sistema de análisis para la evaluación de las búsquedas por catálogos en línea (OPACs) de distintas bibliotecas argentinas, este análisis se realiza a través de una serie de guías de control con estilo de checklist donde se pretende poner en contraposición la satisfacción de las búsquedas y recuperación de esta información por medio de un sistema clásico como lo son los catálogos de biblioteca.

DESCRIPTORES: SERVICIO BIBLIOGRÁFICO; SISTEMA DE INFORMACIÓN EN LÍNEA; INTERFAZ DE ORDENADORES

Rodríguez Leyva, P., Viltres Sala, H., Flores, P., & Amel, L. (2016). Componentes y funcionalidades de un sistema de recuperación de la información. *Revista Cubana de Ciencias Informáticas*, 10, 150-162. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2227-18992016000600012

Este articulo muestra principalmente el desarrollo y funcionalidades de Orion, un sistema de recuperación de información web especializado en ciencias de la información de origen Cubano, el artículo es valioso en cuento a su forma de explicar las funcionalidades básicas que debería tener cualquier sistema de recuperación de información, pero poniendo como ejemplo dicho SRI.

DESCRIPTORES: SISTEMA DE INFORMACIÓN EN LÍNEA; RECUPERACIÓN DE INFORMACIÓN; LENGUAJE DE PROGRAMACIÓN

Rodríguez Perojo, K., & Ronda León, R. (2006). Organización y recuperación de la información: un enfoque desde la perspectiva de la automatización. *Acimed*, *14*(1). Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352006000100004&lng=es&nrm=iso

En el presente artículo se tratan de forma teórica e introductoria la recuperación y organización de la información desde la automatización de la indización por medio de la lingüística y la explicación de los distintos modelos de recuperación de información, se incluye este articulo por su carácter introductorio a los conceptos y su claro desarrollo en los mismos.

DESCRIPTORES: INDEXACIÓN; APLICACIÓN INFORMÁTICA; RECUPERACIÓN DE INFORMACIÓN

Tolosa, G. H., & Bordignon, F. R. (2008). Introducción a la Recuperación de Información. Universidad de Lujan: Lujan. Disponible en: http://eprints.rclis.org/12243/

Este libro es el producto de una investigación documental realizada por docentes de la Universidad de Lujan, Argentina, el aporte por el cual se decidió incluir dicho material, se justifica en la simpleza con las cual se explican los principales conceptos referentes a la recuperación de información como concepto, y sistema, se explayan temas como la evaluación, procesamiento y modelos. Se considera un material básico en la comprensión de estas temáticas e introductorio para el estudio académico de los SRI.

DESCRIPTORES: SISTEMA DE INFORMACIÓN EN LÍNEA; RECUPERACIÓN DE INFORMACIÓN; CATALOGACIÓN

Lenguaje de Marcado

Un lenguaje de marcado es el conjunto de reglas que define todo aquello que es parte de un documento digital que no es el texto propiamente dicho, sino un agregado a modo de etiquetas o marcas, estos lengujes tienen una sintaxis propia que define la estructura de este.

Los lenguajes más conocidos son:

HTML (Hiper Text Marckup Language) describe en el texto tanto la estructura lógica del contenido (titulo, párrafos, de texto normal, enumeraciones, definiciones, citas, etc.) y dejar que la presentación final del documento hipertextual la realice el navegador.

XML (Extensible Markup Language) nació con el objetivo de adaptarlo a los documentos en internet. Este metalenguaje (ya que con el se puede definir el propio lenguaje de presentación) focaliza en la descripción de contenido.

Este apartado propone una serie de documentos donde se trate la temática desde sistemas informáticos y sistemas bibliotecarios para comprender el funcionamiento y aplicación de los mismos.

Fuente: Di Muro, E. (2018) *Eje 2: los procesos preparatorios y las operaciones con los textos*. Universidad de Mar del Plata – LICAD – Seminario B: Temas principales de la Recuperación de Información: Buenos Aires.

Báez, Y. A. (2016). *Recuperación de información en ficheros XES de gran dimensión mediante técnicas de indexación* (Tesis Doctoral) Universidad de Alicante, España. Disponible en: https://rua.ua.es/dspace/bitstream/10045/53889/1/tesis_aponte_baez.pdf

La tesis doctoral plantea un estudio detallado sobre indexación y sistemas estructurados, a su vez, define el sistema XES que está basado en el lenguaje XML generando un nuevo estándar para la descripción de contenidos bibliográficos, por otra parte, se definen los aspectos necesarios y tratamientos de la información para su posterior recuperación a través de los sistemas, por último, realiza una descripción exhaustiva de los experimentos realizados y los resultados positivos en materia de eficiencia en la generación de índices utilizando dicho lenguaje.

DESCRIPTORES: INDEXACIÓN AUTOMÁTICA; LENGUAJE DE PROGRAMACIÓN; ANÁLISIS AUTOMÁTICO DE TEXTOS

Blanco Encinas, A. M. (2008). *Aplicaciones de la web social: herramientas de participación en las bibliotecas universitarias españolas* (Tesis de Maestría) Universidad de Salamanca - Facultad de traducción y documentación, España. Disponible en: https://gredos.usal.es/handle/10366/115832

La tesis aporta una importante teorización sobre la web 2.0 y sus herramientas en post del uso de estas en bibliotecas universitarias, además del uso de estas en ideas de incorporación de nuevos productos a los tradicionales de las unidades de información, a favor de esta recopilación bibliográfica se incluyó por su contribución teórica-metodológica en las herramientas de etiquetado y catalogación social característica concluyente de la web 2.0 la inclusión del usuario productor de la información web.

DESCRIPTORES: INDEXACIÓN AUTOMÁTICA; CATALOGACIÓN

Carro, E. M., & Santa María, T. (2019). Biblioteca Electrónica Textual del teatro español (1868-1936) e investigación con grafos. *Revista de Humanidades Digitales*, *3*. Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=6861800

El articulo expone las experiencias reales de un grupo de trabajo encargado de generar una biblioteca digital con textos teatrales, donde plantean el proceso de trabajo con el lenguaje de marcado XML-TEI y la construcción de metadatos a partir de las primeras obras ingresadas. Como conclusión se plantean las ventajas y alcance de la utilización de este lenguaje de marcado.

DESCRIPTORES: BIBLIOTECA DIGITAL; LENGUAJE DE PROGRAMACIÓN; PROCESAMIENTO DE DATOS

Daudinot Founier, I. (2006). Descripción de los recursos de información en Internet: formato Dublín Core. *Acimed*, *14*(4). Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352006000400009

El articulo expone las características de las herramientas de búsqueda de información en internet, principalmente con su tratamiento de metadatos, aborda también las temáticas de formatos y se concreta en la caracterización del formato Dublín Core en HTML.

DESCRIPTORES: BÚSQUEDA EN LÍNEA; LENGUAJE DE PROGRAMACIÓN; CATALOGACIÓN

Daudinot Founier, I. (2006). Organización y recuperación de información en Internet: teoría de los metadatos. *Acimed*, *14*(5). Disponible en: http://scielo.sld.cu/scielo.php?pid=S1024-94352006000500006&script=sci_arttext&tlng=en

Este articulo aborda la temática de los recursos digitales, su descripción y posterior recuperación exclusivamente en internet, expone las características y funciones de los metadatos en lenguaje HTML y su relación con los tesauros y/o lenguajes controlados.

DESCRIPTORES: BÚSQUEDA EN LÍNEA; LENGUAJE DE PROGRAMACIÓN; CATALOGACIÓN

de Campos, L. M., Fernández Luna, J. M., Huete, J. F., & Vicente López, E. (2012) Personalización y Evaluación XML mediante la Simulación de Perfiles de Usuario y Juicios de Relevancia. En: *II Congreso Español de Recuperación de la Información*. Disponible en: https://www.researchgate.net/profile/Luis_De_Campos/publication/267685826_Personalizacion_y_Evaluacion_XML_mediante_la_Simulacion_de_Perfiles_de_Usuario_y_Juicios_de_Relevancia/links/546115b30cf2c1a63bff7cdd.pdf

Teniendo como premisa que dentro de los SRI con un lenguaje XML las expansiones de consulta generan resultados indeseados al distorsionar la consulta original del usuario, la cual mayormente equivale a la necesidad exacta de información se proponen la evaluación de los sistemas de recuperación de información mediante la realización de perfiles de usuarios y juicios de relevancia por conjuntos de consulta, es decir, la evaluación del mismo por tipo de usuario y tipo de consulta, de esta manera se logrará describir de forma más eficaz los documentos para posteriormente ser buscados y recuperados por el usuarios final.

DESCRIPTORES: BÚSQUEDA EN LÍNEA; LENGUAJE DE PROGRAMACIÓN; EVALUACIÓN DE LA TECNOLOGÍA

Duque Delgado, J. C., & Hoyos Palma, C. D. (2019). *Sistema de recomendación para mejorar la indexación de páginas web en el motor de búsqueda de Google*. (Tesis de grado) Universidad de San Buenaventura – Facultad de ingeniería, Colombia. Disponible en: http://45.5.172.45/handle/10819/7432

La tesis plantea una solución para el problema de la indexación de páginas webs propias, proporciona una herramienta para los usuarios creadores de contenidos el correcto ingreso de metadatos para la posterior lectura por parte del buscador web más

importante del mundo (Google) y de esta manera lograr alcanzar la mayor parte de lectores.

DESCRIPTORES: INDEXACIÓN AUTOMÁTICA; ANÁLISIS AUTOMÁTICO DE TEXTOS; APLICACIÓN INFORMÁTICA

Entralgo Amaro, R., & Aguiar Cedeño, J. D. (2009). Ventajas de la tecnología XML dentro de la Ciencias de la Información. *Ciencias de la Información*, *40*(3), 65-74. Disponible en: http://cinfo.idict.cu/index.php/cinfo/article/view/55

El articulo marca un recorrido histórico y comparativo entre el lenguaje HTML y XML, así también como del lenguaje SGML. Describe además las posibilidades del lenguaje XML y se centra en las aplicaciones derivadas XPointer, Xpath y XQuery, por último, marca las ventajas de la utilización de dicho lenguaje para las ciencias de la información y los sistemas de recuperación de información.

DESCRIPTORES: APLICACIÓN INFORMÁTICA; LENGUAJE DE PROGRAMACIÓN

Mata Vázquez, J., Crespo Azcárate, M., & Maña López, M. (2010). Un método eficaz de indexación para la recuperación de imágenes en archivos en formato pdf. *Procesamiento del lenguaje natural*, *45*, 21-29. Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=3643972

El articulo lo que marca es la incorporación de la recuperación de imágenes dentro de los documentos en formato PDF propone una solución desde la incorporación de metadatos e indexación para la posterior recuperación de dichos espacios de información visual. Marca una solución experimental sobre la metodología de indexación y su eficacia entre la recuperación.

DESCRIPTORES: LENGUAJE DE PROGRAMACIÓN; RECUPERACIÓN DE INFORMACIÓN; INDEXACIÓN AUTOMÁTICA

Menéndez, J. L. (2015). *Las herramientas gestionadas por las bibliotecas universitarias españolas en el servicio de la web: creación de la marca en línea* (Tesis Doctoral) Universidad Complutense - Facultad de Ciencias de la Información, Madrid, España. Disponible en: http://eprints.rclis.org/29201/

La tesis doctoral marca en primera instancia una importante recopilación teórica sobre la web, las herramientas, los lenguajes de marcado HTML, protocolos y ecosistema web en general, la web 2.0 y los productos que de esta se desprenden, blogs, redes sociales, aplicaciones para dispositivos móviles desde una recopilación exhaustiva de conceptos y definiciones claramente definidos. Luego habla de la creación de una marca (branding) y evaluación de la misma que, si bien no es el aporte fundamental a esta recopilación, se la incluyo por su amplio marco teórico sobre el lenguaje HTML.

DESCRIPTORES: LENGUAJE DE PROGRAMACIÓN; BÚSQUEDA EN LÍNEA

Senso, J. A. (2013). Sistemas de metadatos en recuperación de Información: propuesta de modelo de trabajo en el uso RDF sobre directorios LDAP. (Tesis Doctoral) Universidad de Granada - Departamento de Biblioteconomía y Documentación, Granada, España. Disponible en: http://digibug.ugr.es/handle/10481/27430

Se analiza el concepto de metadato y los sistemas actuales, se estudia además el lenguaje HTML y de los lenguajes Dublin Core Metadata Initiative (DC) y Resource Description Framework (RDF) desde una perspectiva de la ciencia de la información y su utilidad para el campo, por último, expone como ejemplo y evalúa el sistema de directorios LDAP en cuanto a la recuperación de información.

DESCRIPTORES: LENGUAJE DE PROGRAMACIÓN; RECUPERACIÓN DE INFORMACIÓN; CATALOGACIÓN

Operaciones con los Textos

Todas las operaciones tienen como objetivo confeccionar índices con la inclusión de términos de alto peso informativo, las operaciones posibles son:

Análisis sintáctico o lexicográfico: opera sobre todos los caracteres del texto y reconoce determinadas secuencias de caracteres como palabras o elementos indexables que sirven para la recuperación.

Palabras vacías: son de alta frecuencia de aparición y muy bajo significado, por lo tanto, se los elimina de las posibilidades de indexación.

Stemming: algoritmo de normalización lingüística por medio del cual las palabras se reducen a una unidad semántica conocida como *stem* en ingles o sema en español. Se realiza extrayendo los prefijos y sufijos de las palabras, manteniendo de esa forma una partícula menor que se relaciona directamente con la raíz de significado de la palabra con peso semántico propio.

Grupos nominales: grupos de palabras compuestos con peso semántico propio, dependiendo del idioma serán los grupos nominales.

En este espacio lo que se pretende es traer información actualizada sobre diferentes aplicaciones de estas operaciones para mejorar los SRI.

Fuente: Di Muro, E. (2018) *Eje 2: los procesos preparatorios y las operaciones con los textos.* Universidad de Mar del Plata – LICAD – Seminario B: Temas principales de la Recuperación de Información: Buenos Aires.

Álvarez García, J. C. (2010). *Clasificación automática de información en portales web mediante técnicas de clustering*. (Tesis de Doctorado) Universidad de Salamanca – Departamento de informática y automática, España Disponible en: https://gredos.usal.es/handle/10366/76382

En primera instancia esta tesis logra encuadrar un marco teórico claro en los modelos de recuperación de información, a su vez explica que las técnicas de clustering son de análisis de datos en las que se aplican las observaciones según su similitud. Este trabajo logra analizar las distintas reglas y técnicas del clustering para clasificar de forma automatizada grandes cantidades de información web, y lograr encontrar las mejores metodologías para la clasificación de documentos digitales incorporándolo en los distintos sistemas de recuperación de información. En este sentido el aporte de la tesis es clave para comprender las operaciones con los textos y la construcción de tesauros necesarios para la posterior automatización de la indexación.

DESCRIPTORES: LENGUAJE DE PROGRAMACIÓN; ANÁLISIS AUTOMÁTICO DE TEXTOS; INDEXACIÓN AUTOMÁTICA

Cabaleiro Barciela, B., & Peñas Padilla, A. (2012). Representación gráfica de documentos para extracción automática de relaciones. *Procesamiento del lenguaje natural*, 49, 57-64. Disponible en: https://rua.ua.es/dspace/bitstream/10045/23920/1/PLN_49_06.pdf

Este trabajo presenta un sistema de representación de documentos basados en el análisis semántico del lenguaje natural utilizado en los documentos, de esta forma pretende crear una representación a nivel múltiple para agilizar las tareas de procesamiento de la información en los distintos sistemas de recuperación de información.

DESCRIPTORES: LINGÜÍSTICA INFORMÁTICA; ANÁLISIS AUTOMÁTICO DE TEXTOS; INDEXACIÓN AUTOMÁTICA

Casali, A., Deco, C., & Beltramone, S. (2015). Asistente para Poblar Repositorios: Recopilación de Objetos Digitales y Extracción de Metadatos. *Versión Abierta Español–Portugués de la Revista Iberoamericana de Tecnología del aprendizaje*,3 (2), 93-100. Disponible en: http://rita.det.uvigo.es/VAEPRITA/201506/uploads/VAEP-RITA.2015.V3.N2.pdf#page=35

Este articulo plantea la estructura de un sistema de poblador de repositorios institucionales, es decir, un software que logre rastrear en la web documentos de relevancia educativa en acceso abierto, extraerlo, analizarlo, rescatar los metadatos y generar un registro y posterior publicación en un repositorio temático o institucional, para ello se emplean teorías informáticas y se analizan los prototipos ya existentes, la incorporación de este documento a la recopilación genera un recurso para la creación de un sistema propio en cualquier espacio de repositorios institucionales y por ende de servicio de recuperación de información.

DESCRIPTORES: ANÁLISIS AUTOMÁTICO DE TEXTOS; INDEXACIÓN AUTOMÁTICA; RECUPERACIÓN DE INFORMACIÓN

De la Calle Velasco, G. (2014). *Modelo basado en técnicas de procesamiento de lenguaje natural para extraer y anotar información de publicaciones científicas* (Tesis Doctoral), Universidad Politécnica de Madrid – Escuela Técnica Superior de Ingenieros Informáticos, España. Disponible en: https://core.ac.uk/download/pdf/33172955.pdf

Este trabajo plantea la problemática de identificar, clasificar y dar acceso a los distintos trabajos académicos en acceso abierto disponibles en la web, y propone como solución un modelo teórico y genérico para construir de índices a partir de la extracción de metadatos de los recursos académicos disponibles en internet, este modelo contempla desde las fuentes desde donde se realizaran las extracciones, hasta la clasificación y curación de contenido, tomando como base el lenguaje natural de los distintos artículos de la literatura científica.

DESCRIPTORES: ANÁLISIS AUTOMÁTICO DE TEXTOS; INDEXACIÓN AUTOMÁTICA; LENGUAJE DE PROGRAMACIÓN

Deco, C. (2009). *Mejora de la recuperación de información en bases de datos de texto utilizando recursos lingüísticos*. (Tesis Doctoral) Universidad Nacional de Rosario - Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Argentina. Disponible en: http://rephip.unr.edu.ar/handle/2133/6447

Esta tesis propone la redefinición de las búsquedas a través del refinamiento semántico de los conceptos expuestos por el usuario con el fin de optimizar los estándares de precisión de resultados y satisfacción de la necesidad informativa, para esto se utilizaron recursos lingüísticos y se construyeron mejoras en las estrategias de búsquedas. En este trabajo el refinamiento semántico se plantea a través de guías al usuario por los distintos conceptos similares a los ingresados en orden jerárquico, simulando un tesauro de estas características con ayudas claras para su comprensión. Esto permite el reordenamiento de los resultados según la jerarquía semántica de los documentos y no de la relevancia clásica.

DESCRIPTORES: ANÁLISIS AUTOMÁTICO DE TEXTOS; INDEXACIÓN AUTOMÁTICA; LENGUAJE DE PROGRAMACIÓN

Padilla Camacho, J. E. (2016). *Evaluación de sistemas de extracción de frases clave*. (Tesis de pregrado) Universidad Autónoma del Estado De México, México. Disponible en: https://core.ac.uk/download/pdf/132423119.pdf

El trabajo expone la comparación de distintos sistemas de extracción automática de frases claves sobre los documentos científicos, lo que llevo como resultado la valoración de que sistema funciona mejor según los objetivos planteados para cada una de las situaciones planteadas.

DESCRIPTORES: ANÁLISIS AUTOMÁTICO DE TEXTOS; INDEXACIÓN AUTOMÁTICA; EVALUACIÓN DE LA TECNOLOGÍA

Pérez Iglesias, J. (2012). *Predicción del rendimiento de consultas basado en rankings de documentos y nuevo marco de evaluación*. (Tesis Doctoral) Universidad Nacional de Educación a Distancia - Escuela Técnica Superior de Ingeniería Informática. Disponible en: http://e-spacio.uned.es/fez/eserv/tesisuned:IngInf-Jperez/Documento_1.pdf

La tesis cuenta con un marco teórico sobre sistemas de recuperación completo, e incorpora los conceptos y trabaja sobre la predicción de la calidad de la consulta, es que una técnica que se basa en estimar la calidad del conjunto de documentos recuperados en cualquier SRI, basados en la premisa de que un alto grado de dispersión de los valores de relevancia implica una mayor calidad de respuestas. Para esto propone un nuevo método de predicción encuadrado dentro de los denominados Post-Retrieval basados exclusivamente en los valores de relevancia que asigna una función de ranking en los documentos a partir de una consulta.

DESCRIPTORES: LENGUAJE DE PROGRAMACIÓN; EVALUACIÓN DE LA TECNOLOGÍA

Robles Morales, E. A., & Parra Baquero, A. N. (2017) *Prototipo de Herramienta de Software para la Extracción de Información en Documentos PDF Utilizando un Motor de Base de Datos nosql*. Universidad Distrital Francisco José De Caldas - Facultad De Ingeniería, Colombia. Disponible en: http://repository.udistrital.edu.co/bitstream/11349/7308/1/ParraBaqueroAndreaNathaly2017.pdf

Se plantean las bases de creación de un software para la extracción de metadatos dentro de los documentos en formato PDF, de esta forma se genera una indexación propia desde el lenguaje natural de cada documento digital. Es un aporte a la automatización de los procesos de descripción y representación de documentos.

DESCRIPTORES: LENGUAJE DE PROGRAMACIÓN; ANÁLISIS AUTOMÁTICO DE TEXTOS; EVALUACIÓN DE LA TECNOLOGÍA

Ryckeboer, H. E. J. L., Spositto, O. M., Barone, M. A. T., Blanco, G. E., Etcheverry, M. E., Procopio, G. E., & Quintana, F. H. (2019). *Implementación de un sistema de recuperación de la información*. (Tesis pregrado) Universidad Nacional de La Matanza - Departamento de Ingeniería e Investigaciones Tecnológicas, Buenos Aires. Disponible en: https://repositoriocyt.unlam.edu.ar/handle/123456789/310

El trabajo se basa en la implementación de un Sistema de Recuperación de información y su evaluación para lo cual se plantean las distintas técnicas que existen, y se toma el análisis para su uso de un modelo derivado del Vectorial, el de indexación semántica latente. Por último, plantea la estructura necesaria de dichos sistemas y la posterior forma de evaluar sus resultados teniendo como objetivo su pronta implementación en distintas Pymes. Marca un aporte practico y a su vez realiza un importante marco conceptual que lo hace valioso para esta recopilación.

DESCRIPTORES: RECUPERACIÓN DE INFORMACIÓN; EVALUACIÓN DE LA TECNOLOGÍA; PROGRAMACIÓN INFORMÁTICA

Sadeghi Hassanabadi, M. (2015). *Recuperación de información en Persa: revisión crítica y propuestas de mejora*. (Tesis Doctoral) Universidad de Valladolid, España. Disponible en: http://uvadoc.uva.es/handle/10324/16008

Esta tesis lo que marca es el desarrollo de un algoritmo que mejore los sistemas de recuperación de información en lengua Persa, lo que lograra que tanto el procesamiento de los textos para su indexación como el procesamiento de las búsquedas y por ende las respuestas incremente su grado de efectividad tanto en los SRI de software como los de recuperación web.

DESCRIPTORES: PERSA; RECUPERACIÓN DE INFORMACIÓN; PROGRAMACIÓN INFORMÁTICA

Viñas, E. F., Serrano-Guerrero, J., Olivas, J. A., de la Mata, J., & Soto, A. (2010). FOG: Arquitectura flexible para la generación automática de ontologías. En Actas del XV Congreso Español sobre Tecnologías y Lógica Fuzzy. ESTYLF. 525-530. Disponible en: https://www.researchgate.net/profile/Jesus_Serrano-Guerrero/publication/228789280_FOG_Arquitectura_flexible_para_la_generacion_automatica_de_ontologias/links/0fcfd50ded753a6a84000000.pdf

El articulo analiza el problema de las consultas en un sistema de recuperación de información y la falta de relación entre las palabras y el contexto de búsqueda, para lo cual propone la arquitectura de una plataforma llamada Fuzzy Ontology Generator (FOG) que permite establecer ontologías que establecen relaciones entre palabras creando conceptos complejos que logrará mejorar los resultados de cualquier consulta.

DESCRIPTORES: PROGRAMACIÓN INFORMÁTICA; ANÁLISIS AUTOMÁTICO DE TEXTOS; INDEXACIÓN AUTOMÁTICA

Procesos semánticos

La web semántica es un conjunto de actividades tendientes a la creación de tecnologías para publicar contenidos legibles por máquinas y basados en la idea de añadir metadatos semánticos y ontológicos a la web, para este proceso existen dos lenguajes:

RDF (Resource Descripton Framework) está diseñado para representar metadatos tales como autor, titulo, fechas de las páginas web, copyrigth, información sobre los documentos y disponer de un listado o catálogo de recursos compartidos, es un estándar para la descripción semántica.

OWL (Web Onyology Language) es un lenguaje donde propone el desarrollo de ontologías para la web semántica, una ontología es una estructura conceptual sistematizada y de consenso no solo para almacenar la información sino también para recuperarla, se definen los términos y sus relaciones.

En este apartado se buscaron documentos donde se expliquen los distintos procesos semánticos dentro de los SRI y su aplicación.

Fuente: Di Muro, E. (2018) *Eje 2: los procesos preparatorios y las operaciones con los textos*. Universidad de Mar del Plata – LICAD – Seminario B: Temas principales de la Recuperación de Información: Buenos Aires.

Arellano, F. F. M., & Ramírez, M. A. A. (2017). El papel de los metadatos en la Web Semántica. *Biblioteca universitaria*, 20(1), 3-10. Disponible en: http://biblio.unam.mx/rbu/index.php/rbu/article/view/171

El trabajo aborda teóricamente definiciones de metadatos y web semántica, su función y relación entre ellos, el complejo de la búsqueda de información y la necesidad de generar contextos donde las ontologías tengan sentido contextual de toda la información existente en internet.

DESCRIPTORES: ANÁLISIS AUTOMÁTICO DE TEXTOS; INDEXACIÓN; SEMÁNTICA

Chávez, S. M. D. J. U., Contreras, M. C. B., & Leyva, M. H. P. (2011). *Sistema de recuperación de información semántico*. (Tesis de grado) Universidad Autónoma Metropolitana Unidad Azcapotzalco - Licenciatura en Ingeniería en Computación, Disponible en: http://aisii.azc.uam.mx/ProyectosTerminales/Propuestas/PropuestaSelene.pdf

La tesis plantea la creación de un sistema de recuperación de información semántico, basado en la recopilación y análisis de los documentos existentes en internet, de esta forma el sistema debe rastrear documentos, procesar su lenguaje y extraer los incluidos dentro de una nómina de ontologías semánticas acorde a las temáticas tratadas. Y, por último, debe poder ser de fácil uso para el usuario final y el proceso de búsqueda y posteriormente recuperación de la información.

DESCRIPTORES: RECUPERACIÓN DE INFORMACIÓN; PROGRAMACIÓN INFORMÁTICA; SEMÁNTICA

Di Luca, M. A. (2019). Modelo basado en ontología para implementar Web Semántica que apoye la gestión de la información y el conocimiento. En: *II Jornadas de Doctorado en Informática,* 112-120. Disponible en: http://repositorio.ual.es:8080/bitstream/handle/10835/6381/actas-jdi2019.pdf?sequence=1&isAllowed=y#page=117

En el este trabajo se habla de la generación de un modelo basado en ontologías para implementar la web semántica en la gestión de la información almacenada en los repositorios digitales, de esta forma se utilizó un algoritmo de procesamiento del lenguaje natura para mejorar la calidad de la información almacenada, lo que derivó en la mejora de la calidad de los resultados a la hora de las búsquedas realizadas.

DESCRIPTORES: ONTOLOGÍA; SEMÁNTICA; SISTEMA DE INFORMACIÓN

González Lara, A. L. (2016). *Comparación de redes semánticas naturales mediante un modelo de recuperación de información* (Tesis Doctoral) Universidad Autónoma de Nuevo León, México. Disponible en: http://eprints.uanl.mx/15708/

Esta tesis marca una definición teórica de lo que se conoce como Redes Semánticas Naturales (RSN) es decir, representaciones de significados según cada lengua en un determinado contexto, por lo tanto se propuso la mejora de estas redes con un trabajo comparativo entre ellas y la formación de un nuevo sistema donde reúna todas ellas puedan ser utilizados según el contexto de la consulta, por lo tanto, es un sistema universal, capaz de detectar por cuál de estas redes optar a la hora de indexar determinados documentos según su origen semántico y lingüístico.

DESCRIPTORES: RECUPERACIÓN DE INFORMACIÓN, EVALUACIÓN DE LA TECNOLOGÍA, SEMÁNTICA

Hernández, A., Tomás, D., & Navarro Colorado, B. (2015). Una aproximación a la recomendación de artículos científicos según su grado de especificidad. *Procesamiento del lenguaje natural*, *55*, 91-98. Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=5560453

El articulo plantea la idea de un sistema de recomendación de documentos asumiendo que un usuario más experimentado necesitara información más específica de un determinado tema, mientras que otro con menos bagaje teórico necesitara documentos más generales o introductorios, por lo tanto, proponen un sistema basado no en perfiles de usuarios, sino en sistemas de tópicos extraídos desde el contenido de los documentos, de esta forma se establece por similitud y diferencias de tópicos las recomendaciones al usuario (modelo Topic Modelling)

DESCRIPTORES: EVALUACIÓN DE LA TECNOLOGÍA; SISTEMA DE INFORMACIÓN; DIVULGACIÓN CIENTÍFICA

Lezama, A. L., Tovar, M., Vilarino, D., & Pinto, D. (2016) Aproximaciones para la expansión semántica de consultas de un Sistema de Recuperación de Información Booleano. *Research in Computing Science,* 115, 117-131. Disponible en: https://pdfs.semanticscholar.org/cde0/93bc18d7511d6af6d316110f02c2ab4b24bc.pdf

El trabajo plantea una serie de experimentos con las expansiones de consulta de los sistemas de recuperación de información booleanos, esto conlleva a que las consultas formadas por palabras que integran conceptos y relaciones semánticas con las ontologías de dominio, lo que se expone es recuperar la información a partir del corpus de dominio de los conceptos ingresados para la búsqueda y la relación ontológica de dominio, de esta forma se mejoró la calidad de los resultados en relación a las aproximaciones originalmente utilizadas.

DESCRIPTORES: SEMÁNTICA; SISTEMA DE INFORMACIÓN; PROGRAMACIÓN INFORMÁTICA

López, T., & Alexander, D. (2016). *Estudio de la web semántica*. Universidad Católica de Pereira, Colombia. Disponible en: http://repositorio.ucp.edu.co/handle/10785/3647

En este trabajo monográfico lo que se explaya es la historia del traspaso de la web 1.0 a la 2.0 y a su vez explica y centra su atención en la web 3.0 o web Semántica donde la mayor parte de la información valiosa y accesible para la recuperación se encuentra en la denominada web obscura o profunda, desde donde se crean los metadatos accesibles por los motores de búsqueda, su incorporación a esta recopilación subyace en los aportes teóricos sobre el concepto de web semántica necesarios para comprender el sistema de recuperación predominante en los últimos tiempos, los buscadores web, así mismo explica los lenguajes de marcado y su relación con la web semántica.

DESCRIPTORES: INTERNET; SEMÁNTICA; SISTEMA DE INFORMACIÓN

Pabón, O. S., & González, M. E. D. S. M. (2014). Propuesta para Extender Semánticamente: el Proceso De Recuperación De Información. *Revista EIA*, 11(22), 51-65. Disponible en: https://www.redalyc.org/pdf/1492/149237906005.pdf

El documento plantea utilizar las tecnologías de la web semántica con el objetivo de asignar significado a los términos de búsqueda, es decir, reestructurar los procesos de los SRI para incorporar descripciones semánticas y aprovechas las tecnológicas de la web en las mejoras de los procesos de indexación y validación de metadatos para la posterior mejora de los resultados obtenidos.

DESCRIPTORES: RECUPERACIÓN DE INFORMACIÓN, EVALUACIÓN DE LA TECNOLOGÍA, SEMÁNTICA

Ráez, A. M., Galiano, M. C. D., Ortega, J. M. P., & López, L. A. U. (2009). Aplicación de la realimentación por relevancia con base textual para sistemas de recuperación visual. *Procesamiento del lenguaje natural*, *43*, 177-183. Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=3125871

Plantea el problema de los sistemas de recuperación de imagen en donde la precisión de los resultados suele ser muy poco precisas y su relevancia con las búsquedas casi nulas, plantea la fusión de metadatos textual y visuales para la caracterización de estos sistemas basados en contenidos, lo que llegaron a la conclusión que sistemas mixtos de descripción son más efectivos que de un solo tipo.

DESCRIPTORES: RECUPERACIÓN DE INFORMACIÓN; PROGRAMACIÓN INFORMÁTICA; INDEXACIÓN AUTOMÁTICA

Sala, H. V., Leyva, P. R., Febles, J. P., & Sentí, V. E. (2018) Modelo para la recuperación de información con anotación semántica. En: *VIII Congreso Internacional de Tecnologías y Contenidos Multimedia*, La Habana Disponible en: http://www.informaticahabana.cu/sites/default/files/ponencias2018/MUL33.pdf

El trabajo presenta un modelo de RI con anotaciones semánticas que permita identificar el contexto de búsqueda del usuario para luego ajustar el resultado al contexto de información, para ello se generaron algoritmos que procesan la

información almacenada, los metadatos semánticos y los documentos completos para luego comparar con los contextos de búsqueda de usuarios. Para luego utilizar un sistema tradicional de rankeo vectorial para la presentación de los resultados. Estas combinaciones de técnicas demostraron ser mayormente exitosas en los procesos de recuperación de información web.

DESCRIPTORES: RECUPERACIÓN DE INFORMACIÓN; PROGRAMACIÓN INFORMÁTICA; SEMÁNTICA

Suárez de la Torre, M. M., Castillo Ossa, L. F., Ríos Cardona, C., Muñoz, G. M., & Aranzazu Álvarez, J. (2009). Análisis, diseño e implementación de un agente deliberativo para extraer contextos definitorios en textos especializados. *Revista Interamericana de Bibliotecología*, 32(2), 59-84. Disponible en: http://eprints.rclis.org/16983/

El articulo expone los resultados no del todo satisfactorios sobre un proyecto de implementación de un sistema semiautomático para la extracción de contextos definitorios en literatura científica basados en el reconocimiento de forma sintáctica y semántica de las frases regulares en dichos documentos, explica además los pasos necesarios, los experimentos y metodologías para llegar al resultado de que no es suficiente para el refinamiento de las búsquedas la extracción de patrones lingüísticos de todas formas es un avance dentro de la indexación y la extracción de contextos semánticos.

DESCRIPTORES: PROGRAMACIÓN INFORMÁTICA; INDEXACIÓN AUTOMÁTICA; ANÁLISIS AUTOMÁTICO DE TEXTOS

Utrera Sust, E. B., Simón Cuevas, A., & Olivas, J. A. (2018). Análisis de tendencias en la personalización de los resultados en buscadores web. Revista Cubana de Ciencias Informáticas, 12(2), 126-146 Disponible en: http://scielo.sld.cu/scielo.php?pid=S2227-18992018000200009&script=sci_arttext&tlng=en

En este artículo se plantea una revisión completa sobre los modelos de personalización para buscadores web con un enfoque en la personalización de las búsquedas de usuarios distintos, para la investigación se realizaron distas etapas de evaluación y experimentación con los modelos existentes y se llegó a la conclusión de la necesidad de una personalización de los perfiles de usuarios por la forma de presentación de la consulta, el idioma y el contexto semántico de las búsquedas, si bien las conclusiones no son excluyentes y deja un marco de ampliación y continuación de la investigación el aporte teórico metodológico inicial es crucial para esta recopilación bibliográfica en post de entender estas nuevas líneas de investigación.

DESCRIPTORES: BÚSQUEDAS EN LÍNEA; EVALUACIÓN DE LA TECNOLOGÍA; SEMÁNTICA

Utrera Sust, E. U., Simón Cuevas, A., Olivas, J. A., & Romero, F. P. (2018). Aproximación a un modelo de recuperación de información personalizada basado en el análisis semántico del contenido. *Procesamiento del Lenguaje Natural, 61*, 31-38. Disponible en: http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/2018-61-3

El artículo plantea un nuevo modelo de recuperación de información basado en el procesamiento semántico, las conclusiones no son excluyentes y en primera instancia plantea una aproximación teórica a las mejoras a los modelos de reducción de sobrecarga de información basados en la premisa que la descripción precisa

sustentada en significado aproxima a los documentos con las consultas de los usuarios según su contexto de búsqueda. Por último, marca una necesidad de generar perfiles de información y de usuarios para la compatibilidad de estos espacios y reducir el ruido informacional de forma considerable.

DESCRIPTORES: BÚSQUEDAS EN LÍNEA; EVALUACIÓN DE LA TECNOLOGÍA; SEMÁNTICA

Viltres Sala, H., Rodríguez Leyva, P., Febles, J. P., & Estrada Sentí, V. (2018). Procesamiento Semántico de información en Sistemas de Recuperación de Información. Revista Cubana de Ciencias Informáticas, 12(1), 102-116. Disponible en: http://scielo.sld.cu/scielo.php?pid=S2227-18992018000100008&script=sci_arttext&tlng=en

El trabajo plantea el procesamiento semántico de información para mejorar la relevancia de los resultados dentro de cualquier SRI el mismo desarrolla tres componentes claves Rastreo-Indexación, Procesamiento y Presentación para poder identificar las necesidades de información, la selección de material a mostrar y por último la forma de presentar los resultados.

DESCRIPTORES: ANÁLISIS AUTOMÁTICO DE TEXTOS; EVALUACIÓN DE LA TECNOLOGÍA; SEMÁNTICA

Conclusiones

A lo largo de la recopilación bibliográfica se intentó brindarle al usuario investigador la información más actualizada dentro del periodo temporal de 2005 a 2019 en acceso abierto y español escrito en Iberoamérica, a raíz de esta investigación se llegaron a algunas conclusiones estadísticas que se considera de interés y relevancia con el fin de brindar más información al respecto.

Filiación institucional

Entendiendo este concepto como los datos brindados por cada autor en sobre su espacio laboral, se puede concluir que a nivel general las investigaciones se realizan desde universidades cubanas, argentinas y españolas en mayor medida, pero siendo la Universidad de las Ciencias Informáticas (CUBA) la que más investigadores posee, continuando con la Universidad Nacional de Misiones (Argentina) y la Universidad de Huelva (España) en tercer lugar, también se observan en menor medida universidades de México y Paraguay.

Siguiendo con el análisis se realizó una serie de gráficos por sección, continuando el orden de la recopilación, donde por cada ítem se puede concluir que para los textos de:

- Conceptos Generales: la universidad con mayores exponentes es de Argentina.
- Lenguajes de marcado: la universidad con mayores exponentes es de España.
- Operaciones con los textos: la universidad con mayores exponentes es de España.
- Procesos Semánticos: la universidad con mayores exponentes es de Cuba.

(ver anexo Gráficos: filiación institucional)

País de publicación

Otro dato que nos parece interesante tener el análisis de las publicaciones por países ya que no siempre coinciden los espacios de investigación con los países donde se publica la información, en este caso descubrimos el mayor país publicador es Argentina, seguido de Colombia y Costa Rica en tercer lugar.

Luego continuando con la lógica de la recopilación son centramos en los mayores países donde se publicó con respecto a cada sección planteada, y concluimos que en:

- Conceptos Generales: el mayor país es Argentina seguido de Cuba.
- Lenguajes de marcado: el mayor país es España seguido de Cuba.
- Operaciones con los textos: el mayor país es España seguido de Argentina.
- Procesos Semánticos: el mayor país es España seguido de Cuba y en tercer lugar, pero con poca diferencia México.

(ver anexo Gráficos: País de publicación)

Espacio de publicación

Dentro de la recopilación se tomó como premisa que los documentos estén disponibles en acceso abierto a texto completo, esto implico incluir artículos de revistas académicas, pero también tesis de grado o posgrado y actas de congresos o simposios, para lo cual se realizó entonces el análisis de esos espacios de publicación, en principio se expone que el 54% fueron artículos de revista, mientas que solo el 10% se publicaron en actas de congresos, por lo tanto en cuanto a tesis abarcan el 36% restante.

Luego se puede observar que las revistas donde más se publica sobre las temáticas son Procesamiento del lenguaje natural, Acimed, y la Revista Cubana de Ciencias Informáticas. En cuanto a las tesis se observó que provienen mayoritariamente de la Universidad de Salamanca (España), y por último se nombra las actas Actas del XV Congreso Español sobre Tecnologías y Lógica Fuzzy. ESTYLF; II Congreso Español de Recuperación de la Información; II Jornadas de Doctorado en Informática; VIII Congreso Internacional de Tecnologías y Contenidos Multimedia; y XVIII Workshop de Investigadores en Ciencias de la Computación.

(ver anexo Gráficos: Espacio de publicación)

Años de publicación

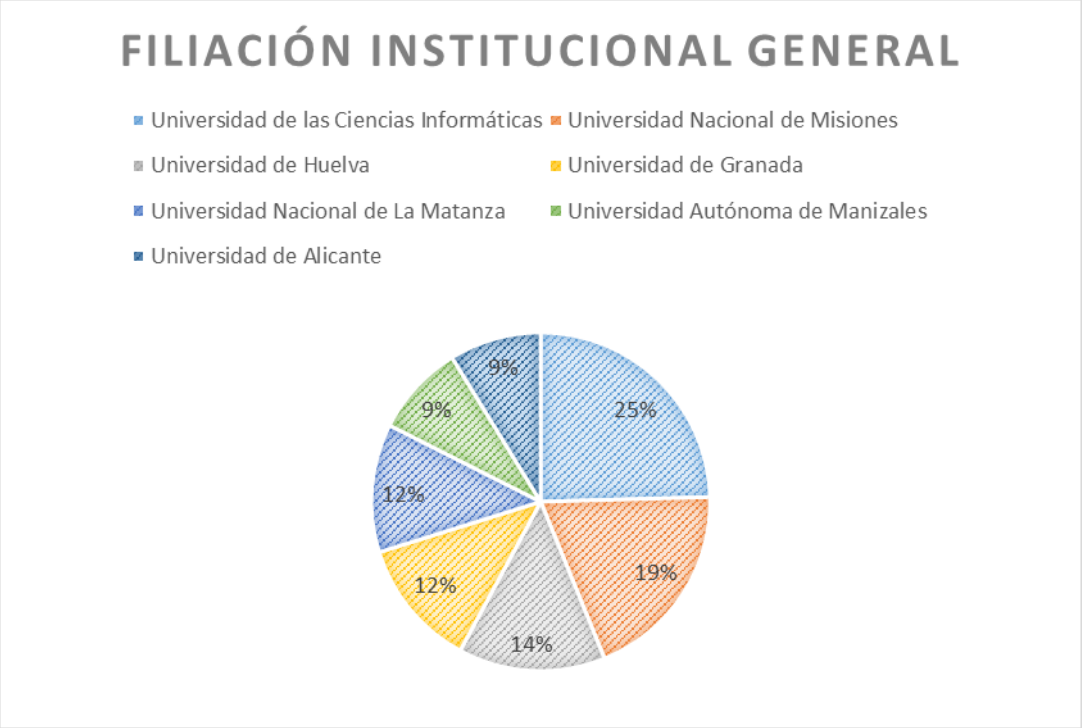
En este último apartado se evaluó los picos de publicación en el periodo mencionado anteriormente (2005-2019) donde se observó un pico máximo con 20 publicaciones en el 2016 y el año 2013 con una publicación, se posiciona en el año con menos trabajos para estas temáticas. En líneas generales se puede decir que la investigación de estas temáticas tuvo varios picos de gran investigación los años 2009 y 2010; los años 2015 y 2016; y por último los años 2018 y 2019 donde se observan los picos de publicación.

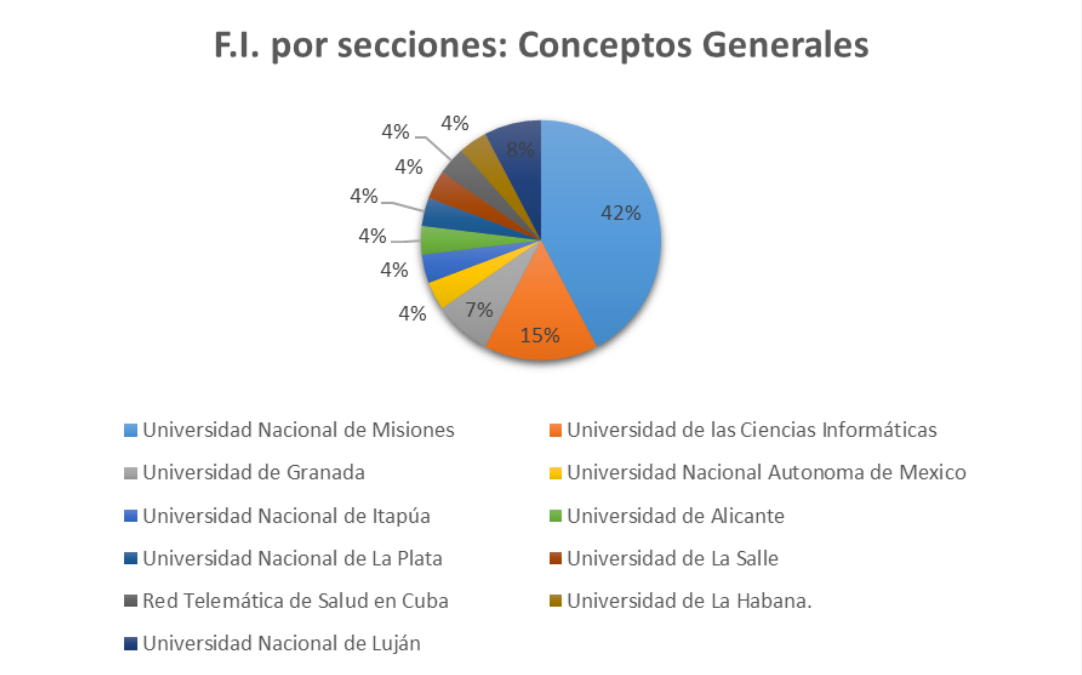
(ver anexo Gráficos: Años de publicación)

Anexo Gráficos

Filiación institucional

En este apartado se expondrán los gráficos en orden a como se leen en el apartado Conclusiones: Filiación institucional.



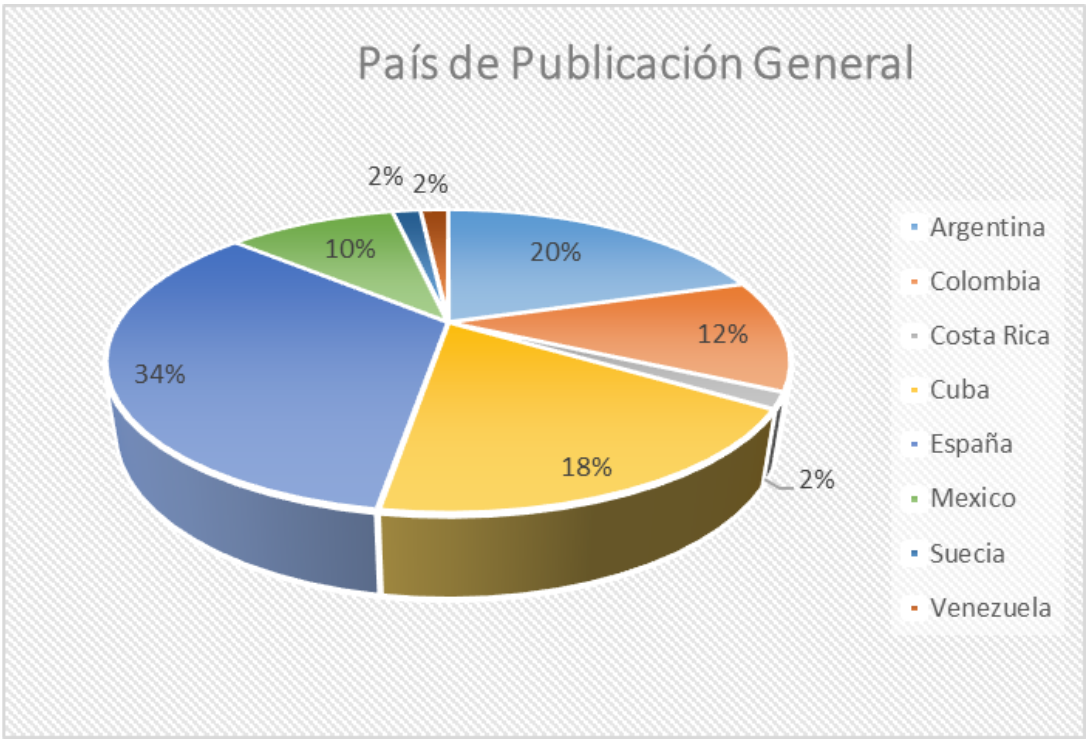


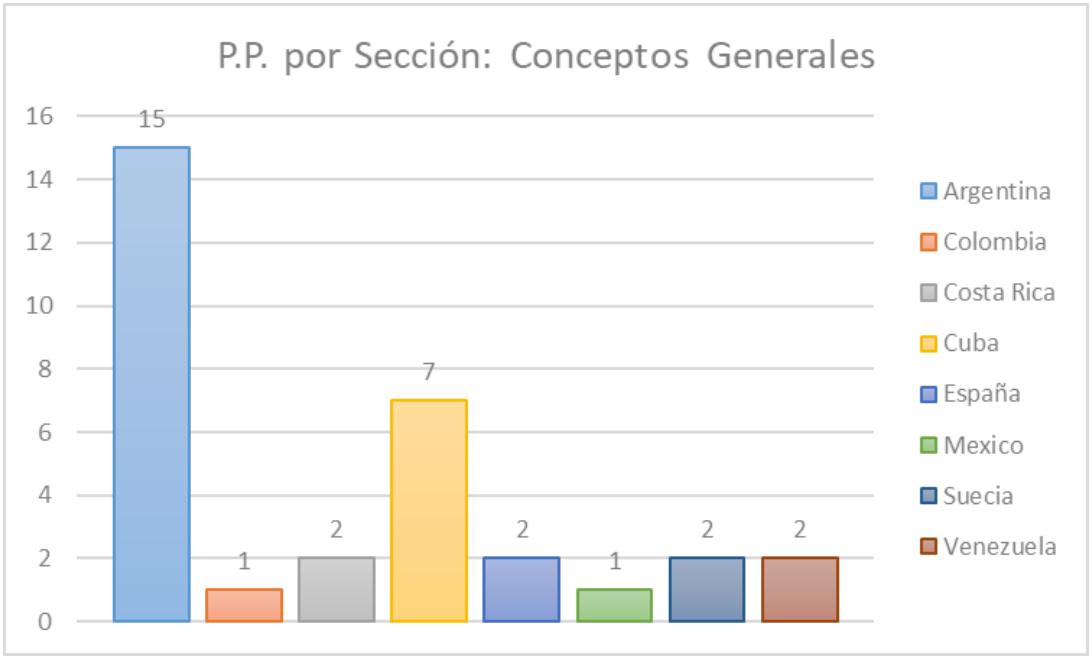

F.I. por secciones: Lenguaje de Marcado
10%
11%
5%
5%
26%
16%
5%
11%
11%
Biblioteca Nacional José Martí
Empresa de Tecnologías de la Información y Servicios Telemáticos Avanzados.
Universidad Complutense
Universidad de Alicante
Universidad de Granada
Universidad de Huelva
Universidad de Salamanca
Universidad de San Buenaventura
Universidad Internacional de La Rioja

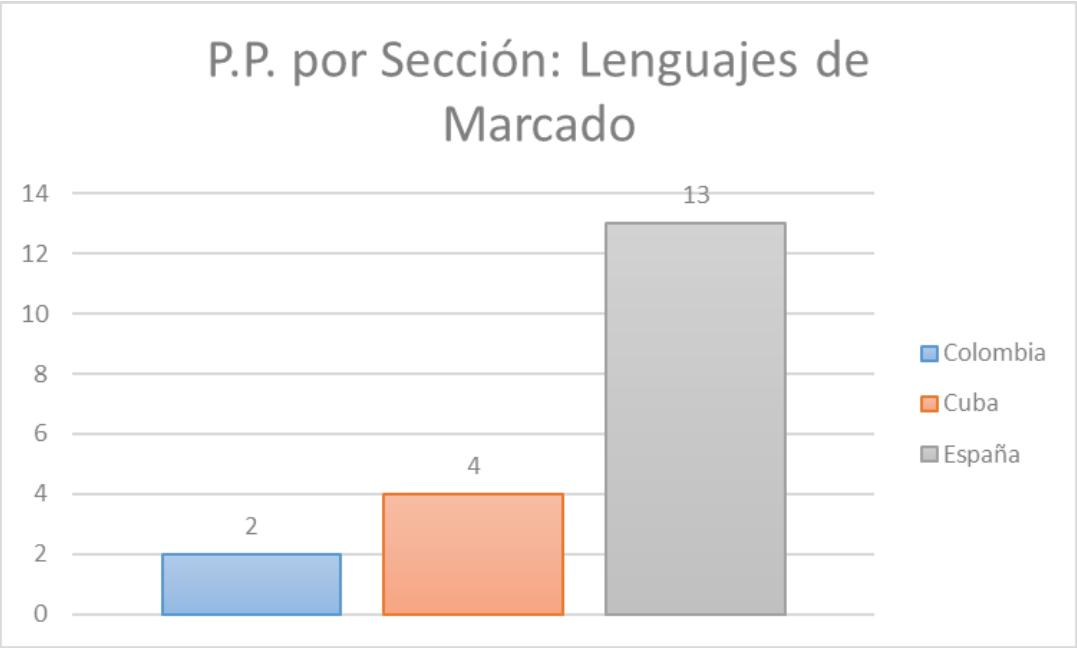

F.I. por secciones: Operaciones con los Textos
8%
16%
4%
20%
4%
4%
8%
4%
4%
28%
4%
Escuela Técnica Superior de Ingeniería Informática
Universidad Nacional de Rosario
Universidad Autónoma del Estado De México
Universidad de Huelva
Universidad de Salamanca
Universidad de Valladolid
Universidad Distrital Francisco José De Caldas
Universidad Nacional de Educación a Distancia
Universidad Nacional de La Matanza
Universidad Politécnica de Madrid

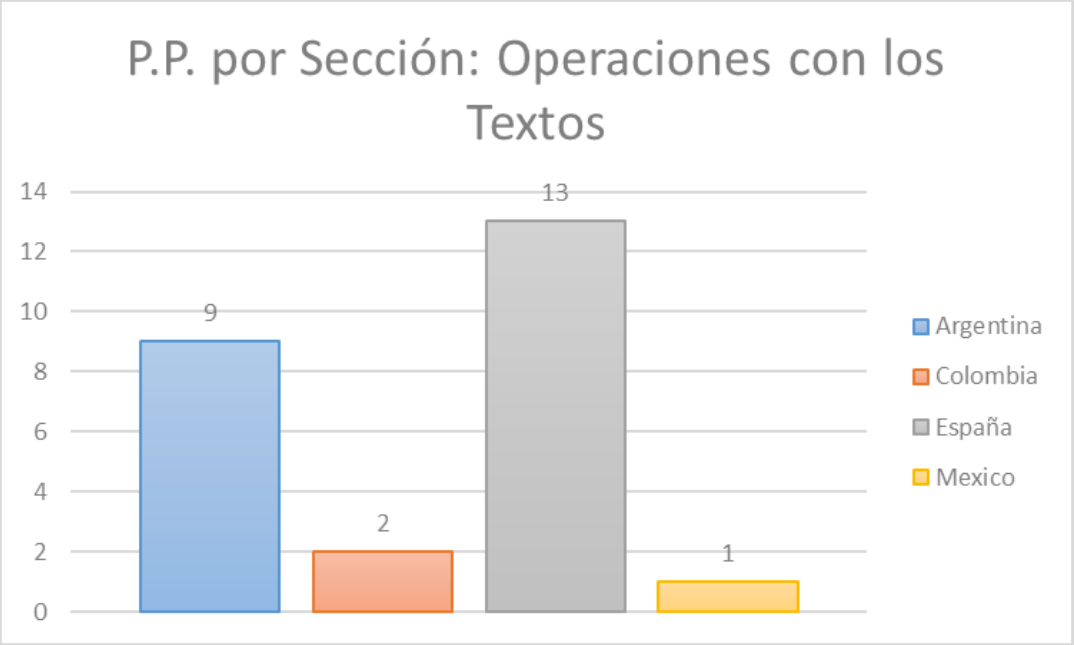
País de publicación

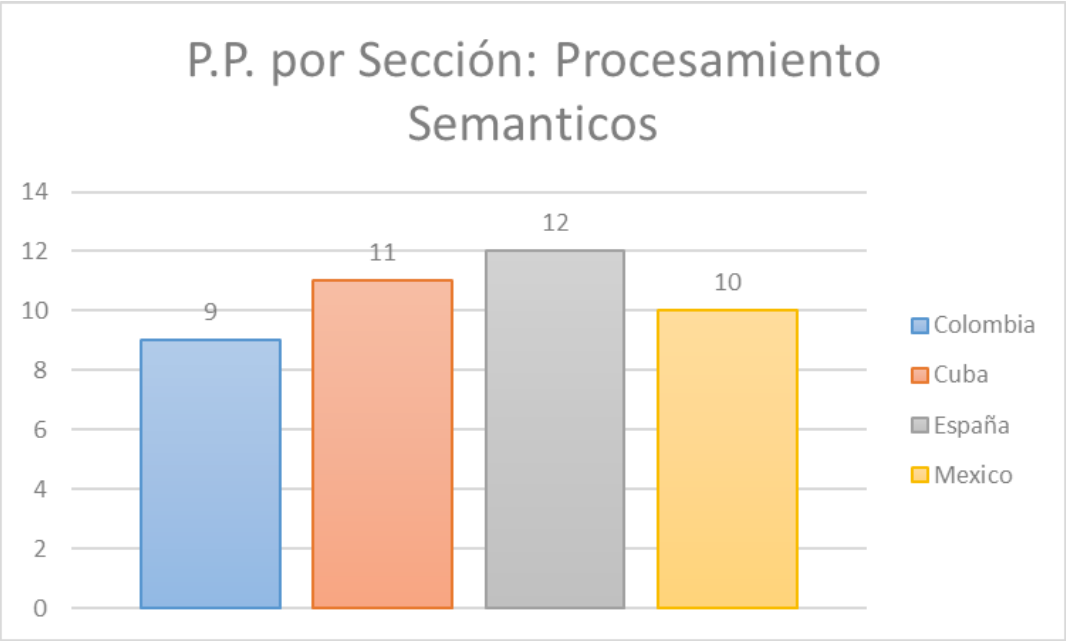
En este apartado se expondrán los gráficos en orden a como se leen en el apartado Conclusiones: País de publicación.





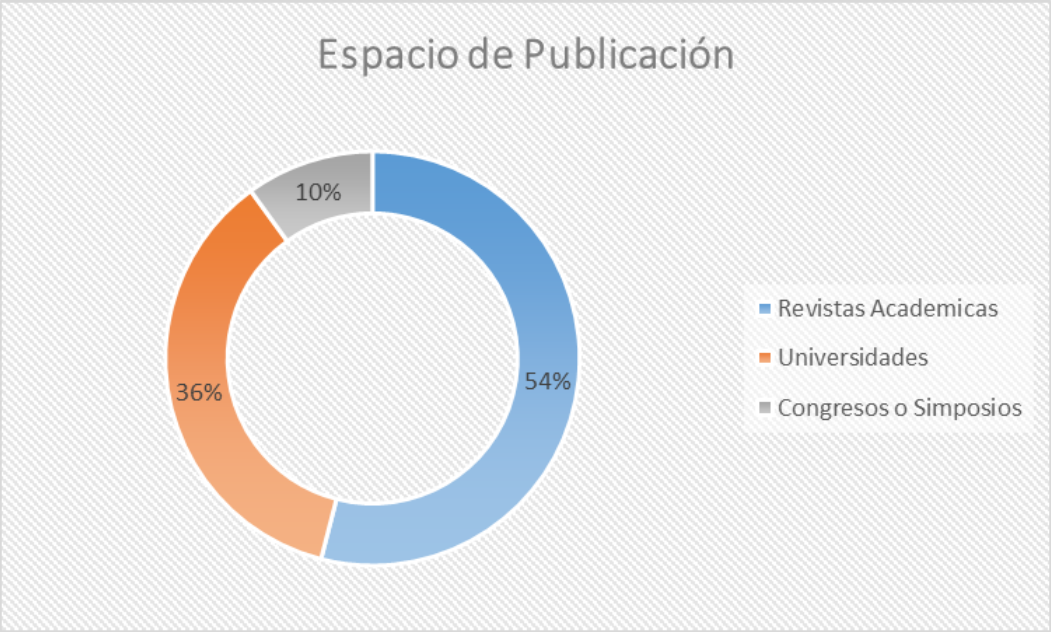

P.P. por Sección: Lenguajes de Marcado
14
12
10
8
6
4
2
0
2
4
13
Colombia
Cuba
España

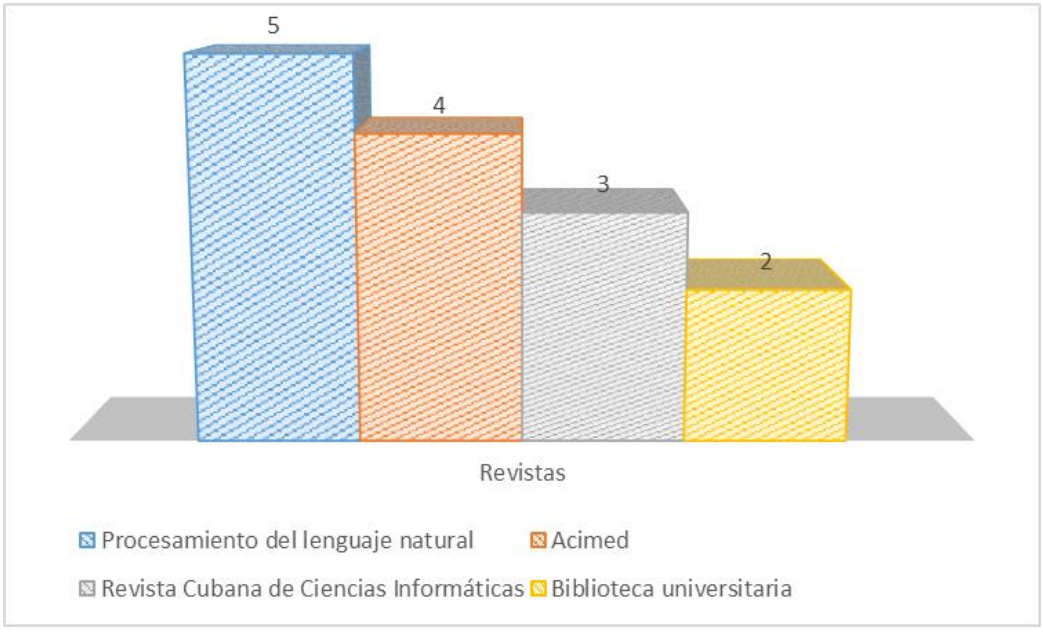

P.P. por Sección: Operaciones con los Textos
14
12
10
8
6
4
2
0
9
2
13
1
Argentina
Colombia
España
Mexico



Espacio de publicación

En este apartado se expondrán los gráficos en orden a como se leen en el apartado Conclusiones: Espacio de publicación.




5
4
3
2
Revistas
Procesamiento del lenguaje natural
Acimed
Revista Cubana de Ciencias Informáticas
Biblioteca universitaria

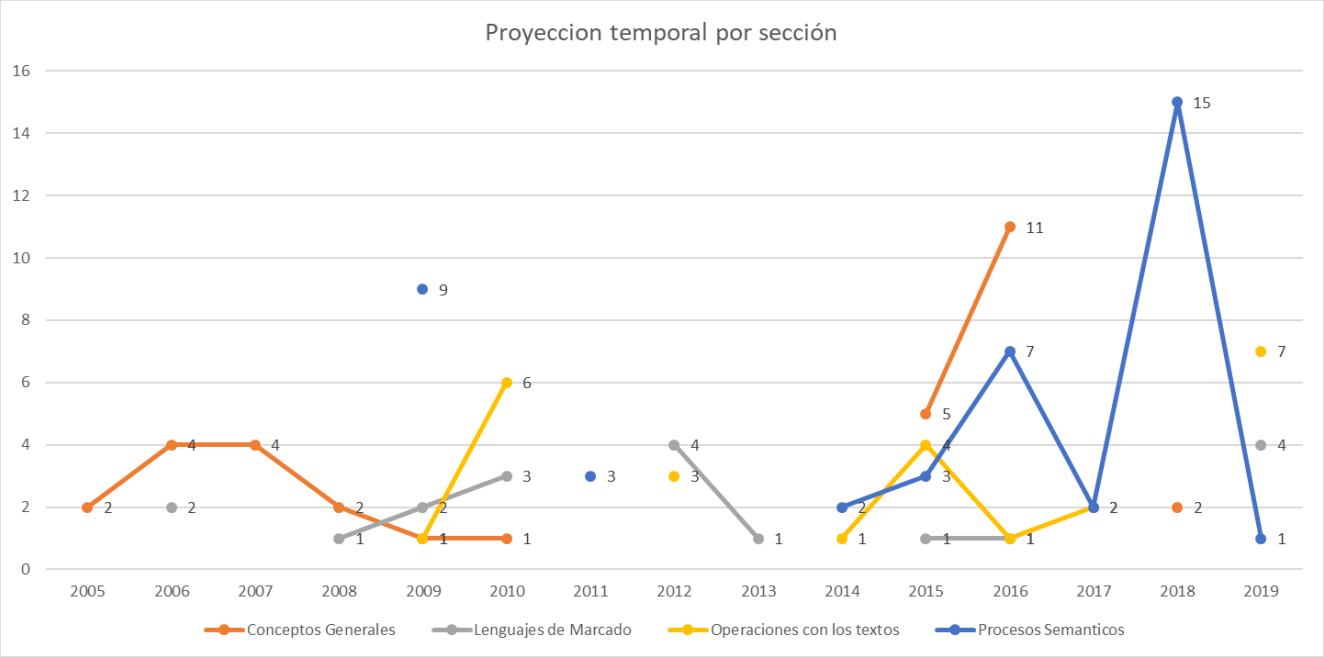

Universidades Editoras
Universidad de Salamanca 2
Universidad Nacional de Rosario 1
Universidad Autónoma del Estado De México 1
Universidad Católica de Pereira 1
Universidad de Alicante 1
Universidad de Lujan 1
Universidad de Valladolid 1
Universidad Nacional de Educación a Distancia 1
Universidad Nacional de La Matanza 1
Universidad Autónoma de Nuevo León 1
Universidad Autónoma Metropolitana Unidad Azcapotzalco 1
Universidad Complutense 1
Universidad de Granada 1
Universidad de San Buenaventura 1
Universidad Distrital Francisco José De Caldas 1
Universidad Nacional de La Plata 1
Universidad Politécnica de Madrid 1


Conferencias o Simposicios
1
1
1
1
1
Actas del XV Congreso Español sobre Tecnologías y Lógica Fuzzy. ESTYLF
II Congreso Español de Recuperación de la Información
II Jornadas de Doctorado en Informática
VIII Congreso Internacional de Tecnologías y Contenidos Multimedia
XVIII Workshop de Investigadores en Ciencias de la Computación

Años de publicación

En este apartado se expondrán los gráficos en orden a como se leen en el apartado Conclusiones: Años de publicación.





Índice de Autor

En este apartado se expondrá una lista de los autores correspondientes a la recopilación, para esto se expone el autor, la sección donde se encuentra y el título del artículo.

En el apartado "Sección" se encontrarán las iniciales:

CG: Conceptos Generales

LM: Lenguaje de Marcado

OP: Operaciones con los Textos

PS: Procesos Semánticos

Autor	Apariciones	Sección	Titulo
Aguiar Cedeño, J. D.	1	LM	Ventajas de la tecnología XML dentro de la Ciencias de la Información
Alexander, D.	1	PS	Estudio de la web semántica
Álvarez García, J. C.	1	OT	Clasificación automática de información en portales web mediante técnicas de clustering
Amel, L.	1	CG	Componentes y funcionalidades de un sistema de recuperación de la información
Aranzazu Álvarez, J.	1	PS	Análisis, diseño e implementación de un agente deliberativo para extraer contextos definitorios en textos especializados
Arellano, F. F. M.	1	PS	El papel de los metadatos en la Web Semántica.
Báez, Y. A.	1	LM	Recuperación de información en ficheros XES de gran dimensión mediante técnicas de indexación
Barone, M. A. T.	1	OT	Implementación de un sistema de recuperación de la información.
Batista, D. O.	1	CG	Los metadatos, la búsqueda y recuperación de información desde las Ciencias de la Información
Beltramone, S.	1	OT	Asistente para Poblar Repositorios: Recopilación de Objetos Digitales y Extracción de Metadatos
Biale, C. O.	1	CG	Avances en la Construcción de un Sistema de Recuperación de Información para Información Científica en Ciencias de la Computación
Blanco Encinas, A. M.	1	LM	Aplicaciones de la web social: herramientas de participación en las bibliotecas universitarias españolas
Blanco, G. E.	1	OT	Implementación de un sistema de recuperación de la información.
Bordignon, F. R.	2	CG CG	Recuperación de información: un área de investigación en crecimiento Introducción a la Recuperación de Información
Cabaleiro Barciela, B.	1	OT	Representación gráfica de documentos para extracción automática de relaciones
Camaraza Monserrate, A.	1	CG	Recuperación de información: reflexiones epistémicas de una ciencia en su estado embrionario
Canteros, A.	1	CG	Avances en la Construcción de un Sistema de Recuperación de Información para Información Científica en Ciencias de la Computación
Carro, E. M.	1	LM	Biblioteca Electrónica Textual del teatro español (1868-1936) e investigación con grafos
Casali, A.	1	OT	Asistente para Poblar Repositorios: Recopilación de Objetos Digitales y Extracción de Metadatos

Castillo Ossa, L. F.	1	PS	Análisis, diseño e implementación de un agente deliberativo para extraer contextos definitorios en textos especializados
Centelles, M.	1	CG	Taxonomías para la categorización y la organización de la información en sitios web
Chávez, S. M. D. J. U.	1	PS	Sistema de recuperación de información semántico
Contreras, M. C. B.	1	PS	Sistema de recuperación de información semántico
Crespo Azcárate, M.	1	LM	Un método eficaz de indexación para la recuperación de imágenes en archivos en formato pdf
Cuba, Y.	1	CG	Los metadatos, la búsqueda y recuperación de información desde las Ciencias de la Información
Daudinot Founier, I.	2	LM	Descripción de los recursos de información en Internet: formato Dublín Core
		LM	Organización y recuperación de información en Internet: teoría de los metadatos
de Campos, L. M.	1	LM	Personalización y Evaluación XML mediante la Simulación de Perfiles de Usuario y Juicios de Relevancia
De la Calle Velasco, G.	1	OT	Modelo basado en técnicas de procesamiento de lenguaje natural para extraer y anotar información de publicaciones científicas
de la Mata, J.	1	OT	FOG: Arquitectura flexible para la generación automática de ontologías
Deco, C.	2	OT	Asistente para Poblar Repositorios: Recopilación de Objetos Digitales y Extracción de Metadatos
		OT	Mejora de la recuperación de información en bases de datos de texto utilizando recursos lingüísticos
Di Luca, M. A.	1	PS	Modelo basado en ontología para implementar Web Semántica que apoye la gestión de la información y el conocimiento
Duque Delgado, J. C.	1	LM	Sistema de recomendación para mejorar la indexación de páginas web en el motor de búsqueda de Google
Entralgo-Amaro, R.	1	LM	Ventajas de la tecnología XML dentro de la Ciencias de la Información
Estrada Sentí, V.	1	PS	Procesamiento Semántico de información en Sistemas de Recuperación de Información
Etcheverry, M. E.	1	OT	Implementación de un sistema de recuperación de la información.
Febles, J. P.	2	PS	Modelo para la recuperación de información con anotación semántica
		PS	Procesamiento Semántico de información en Sistemas de Recuperación de Información
Fernández Luna, J. M.	1	LM	Personalización y Evaluación XML mediante la Simulación de Perfiles de Usuario y Juicios de Relevancia
Flores, P.	1	CG	Componentes y funcionalidades de un sistema de recuperación de la información
Galiano, M. C. D.	1	PS	Aplicación de la realimentación por relevancia con base textual para sistemas de recuperación visual
García, A. A. R.	1	CG	Elementos para reflexionar sobre el nuevo estándar para la descripción y acceso a recursos.
Gómez, F. L.	1	CG	Interoperabilidad en los Sistemas de Información Documental (SID): la información debe fluir
González Lara, A. L.	1	PS	Comparación de redes semánticas naturales mediante un modelo de recuperación de información
González, M. E. D. S. M.	1	PS	Propuesta Para Extender Semánticamente: el Proceso De Recuperación De Información

Hassan-Montero, Y.	1	CG	Metodologías para el desarrollo de interfaces visuales de recuperación de información: análisis y comparación.
Hernández, A.	1	PS	Una aproximación a la recomendación de artículos científicos según su grado de especificidad
Herrero-Solana, V.	1	CG	Metodologías para el desarrollo de interfaces visuales de recuperación de información: análisis y comparación.
Hoyos Palma, C. D.	1	LM	Sistema de recomendación para mejorar la indexación de páginas web en el motor de búsqueda de Google
Huete, J. F.	1	LM	Personalización y Evaluación XML mediante la Simulación de Perfiles de Usuario y Juicios de Relevancia
Kuna, H.	1	CG	Avances en el Desarrollo de un Sistema de Recuperación de Información para Publicaciones Científicas del Área de Ciencias de la Computación.
Kuna, H. D.	1	CG	Avances en la Construcción de un Sistema de Recuperación de Información para Información Científica en Ciencias de la Computación
Leyva, M. H. P.	1	PS	Sistema de recuperación de información semántico
Leyva, P. R.	1	PS	Modelo para la recuperación de información con anotación semántica
Lezama, A. L.	1	PS	Aproximaciones para la expansión semántica de consultas de un Sistema de Recuperación de Información Booleano
López, L. A. U.	1	PS	Aplicación de la realimentación por relevancia con base textual para sistemas de recuperación visual
López, T.	1	PS	Estudio de la web semántica
Maña López, M.	1	LM	Un método eficaz de indexación para la recuperación de imágenes en archivos en formato pdf
Martini, E.	2	CG CG	Avances en la Construcción de un Sistema de Recuperación de Información para Información Científica en Ciencias de la Computación Avances en el Desarrollo de un Sistema de Recuperación de Información para Publicaciones Científicas del Área de Ciencias de la Computación.
Mata Vázquez, J.	1	LM	Un método eficaz de indexación para la recuperación de imágenes en archivos en formato pdf
Menéndez, J. L.	1	LM	Las herramientas gestionadas por las bibliotecas universitarias españolas en el servicio de la web: creación de la marca en línea
Miralles, M. V. J.	1	CG	El catálogo: un recurso en expansión
Muñoz, G. M.	1	PS	Análisis, diseño e implementación de un agente deliberativo para extraer contextos definitorios en textos especializados
Navarro Colorado, B.	1	PS	Una aproximación a la recomendación de artículos científicos según su grado de especificidad
Olivas, J. A.	1	OT	FOG: Arquitectura flexible para la generación automática de ontologías
Olivas, J. A.	2	PS PS	Análisis de tendencias en la personalización de los resultados en buscadores web Aproximación a un modelo de recuperación de información personalizada basado en el análisis semántico del contenido
Ortega, J. M. P.	1	PS	Aplicación de la realimentación por relevancia con base textual para sistemas de recuperación visual
Pabón, O. S.	1	PS	Propuesta Para Extender Semánticamente: el Proceso De Recuperación De Información

Padilla Camacho, J. E.	1	OT	Evaluación de sistemas de extracción de frases clave
Parra Baquero, A. N.	1	OT	Prototipo de Herramienta de Software para la Extracción de Información en Documentos PDF Utilizando un Motor de Base de Datos nosql
Peñas Padilla, A.	1	OT	Representación gráfica de documentos para extracción automática de relaciones
Pérez Iglesias, J.	1	OT	Predicción del rendimiento de consultas basado en rankings de documentos y nuevo marco de evaluación
Pichinini, M.	1	CG	Desarrollo de una guía de control para interfaces de búsqueda por materia en catálogos en línea de acceso web
Pinto, D.	1	PS	Aproximaciones para la expansión semántica de consultas de un Sistema de Recuperación de Información Booleano
Podkowa, L.	1	CG	Avances en el Desarrollo de un Sistema de Recuperación de Información para Publicaciones Científicas del Área de Ciencias de la Computación.
Procopio, G. E.	1	OT	Implementación de un sistema de recuperación de la información.
Quintana, F. H.	1	OT	Implementación de un sistema de recuperación de la información.
Ráez, A. M.	1	PS	Aplicación de la realimentación por relevancia con base textual para sistemas de recuperación visual
Rambo, A.	1	CG	Avances en el Desarrollo de un Sistema de Recuperación de Información para Publicaciones Científicas del Área de Ciencias de la Computación.
Rambo, A. R.	1	CG	Avances en la Construcción de un Sistema de Recuperación de Información para Información Científica en Ciencias de la Computación
Ramírez, M. A. A.	1	PS	El papel de los metadatos en la Web Semántica.
Rey, M.	2	CG CG	Avances en la Construcción de un Sistema de Recuperación de Información para Información Científica en Ciencias de la Computación Avances en el Desarrollo de un Sistema de Recuperación de Información para Publicaciones Científicas del Área de Ciencias de la Computación.
Ríos Cardona, C.	1	PS	Análisis, diseño e implementación de un agente deliberativo para extraer contextos definitorios en textos especializados
Robles Morales, E. A.	1	OT	Prototipo de Herramienta de Software para la Extracción de Información en Documentos PDF Utilizando un Motor de Base de Datos nosql
Rodríguez Leyva, P.	2	CG PS	Componentes y funcionalidades de un sistema de recuperación de la información Procesamiento Semántico de información en Sistemas de Recuperación de Información
Rodríguez Perojo, K.	1	CG	Organización y recuperación de la información: un enfoque desde la perspectiva de la automatización
Romero, F. P.	1	PS	Aproximación a un modelo de recuperación de información personalizada basado en el análisis semántico del contenido
Ronda León, R.	1	CG	Organización y recuperación de la información: un enfoque desde la perspectiva de la automatización
Ryckeboer, H. E. J. L.	1	OT	Implementación de un sistema de recuperación de la información.
Sadeghi Hassanabadi, M.	1	OT	Recuperación de información en Persa: revisión crítica y propuestas de mejora

Sala, H. V.	1	PS	Modelo para la recuperación de información con anotación semántica
Santa María, T.	1	LM	Biblioteca Electrónica Textual del teatro español (1868-1936) e investigación con grafos
Senso, J. A.	1	LM	Sistemas de metadatos en recuperación de Información: propuesta de modelo de trabajo en el uso RDF sobre directorios LDAP
Sentí, V. E.	1	PS	Modelo para la recuperación de información con anotación semántica
Serrano-Guerrero, J.	1	OT	FOG: Arquitectura flexible para la generación automática de ontologías
Simón Cuevas, A.	2	PS	Análisis de tendencias en la personalización de los resultados en buscadores web
		PS	Aproximación a un modelo de recuperación de información personalizada basado en el análisis semántico del contenido
Soto, A.	1	OT	FOG: Arquitectura flexible para la generación automática de ontologías
Spositto, O. M.	1	OT	Implementación de un sistema de recuperación de la información.
Suárez de la Torre, M. M.	1	PS	Análisis, diseño e implementación de un agente deliberativo para extraer contextos definitorios en textos especializados
Tolosa, G. H.	1	CG	Introducción a la Recuperación de Información
Tolosa, G. H.	1	CG	Recuperación de información: un área de investigación en crecimiento
Tomás, D.	1	PS	Una aproximación a la recomendación de artículos científicos según su grado de especificidad
Tovar, M.	1	PS	Aproximaciones para la expansión semántica de consultas de un Sistema de Recuperación de Información Booleano
Utrera Sust, E. B.	2	PS	Análisis de tendencias en la personalización de los resultados en buscadores web
		PS	Aproximación a un modelo de recuperación de información personalizada basado en el análisis semántico del contenido
Vicente López, E.	1	LM	Personalización y Evaluación XML mediante la Simulación de Perfiles de Usuario y Juicios de Relevancia
Viltres Sala, H.	2	CG	Componentes y funcionalidades de un sistema de recuperación de la información
		PS	Procesamiento Semántico de información en Sistemas de Recuperación de Información
Viñas, E. F.	1	OT	FOG: Arquitectura flexible para la generación automática de ontologías
Zamudio, E.	1	CG	Avances en la Construcción de un Sistema de Recuperación de Información para Información Científica en Ciencias de la Computación

Índice por Descriptores

En este apartado se expondrá una lista de los descriptores utilizados para la representación intelectual de cada documento, en el mismo se encontrará la información de sección en la recopilación donde se encuentra, las veces que aparece y titulo del artículo.

En el apartado "Sección" se encontrarán las iniciales:

CG: Conceptos Generales

LM: Lenguaje de Marcado

OP: Operaciones con los Textos

PS: Procesos Semánticos

DESCRIPTOR	Apariciones	Sección	Titulo
ALGORITMO	2	CG	Avances en el Desarrollo de un Sistema de Recuperación de Información para Publicaciones Científicas del Área de Ciencias de la Computación.
		CG	Metodologías para el desarrollo de interfaces visuales de recuperación de información: análisis y comparación.
ANÁLISIS AUTOMÁTICO DE TEXTOS	13	PS	Análisis, diseño e implementación de un agente deliberativo para extraer contextos definitorios en textos especializados
		OT	Asistente para Poblar Repositorios: Recopilación de Objetos Digitales y Extracción de Metadatos
		OT	Clasificación automática de información en portales web mediante técnicas de clustering
		PS	El papel de los metadatos en la Web Semántica.
		OT	Evaluación de sistemas de extracción de frases clave
		OT	FOG: Arquitectura flexible para la generación automática de ontologías
		OT	Mejora de la recuperación de información en bases de datos de texto utilizando recursos lingüísticos
		OT	Modelo basado en técnicas de procesamiento de lenguaje natural para extraer y anotar información de publicaciones científicas
		PS	Procesamiento Semántico de información en Sistemas de Recuperación de Información
		OT	Prototipo de Herramienta de Software para la Extracción de Información en Documentos PDF Utilizando un Motor de Base de Datos nosql
		LM	Recuperación de información en ficheros XES de gran dimensión mediante técnicas de indexación
		OT	Representación gráfica de documentos para extracción automática de relaciones
		LM	Sistema de recomendación para mejorar la indexación de páginas web en el motor de búsqueda de Google
APLICACIÓN INFORMÁTICA	3	CG	Organización y recuperación de la información: un enfoque desde la perspectiva de la automatización
		LM	Sistema de recomendación para mejorar la indexación de páginas web en el motor de búsqueda de Google
		LM	Un método eficaz de indexación para la recuperación de imágenes en archivos en formato pdf

BASE DE DATOS	2	CG	Interoperabilidad en los Sistemas de Información Documental (SID): la información debe fluir
		CG	Los metadatos, la búsqueda y recuperación de información desde las Ciencias de la Información
BIBLIOTECA DIGITAL	1	LM	Biblioteca Electrónica Textual del teatro español (1868-1936) e investigación con grafos
BÚSQUEDA EN LÍNEA	9	PS	Análisis de tendencias en la personalización de los resultados en buscadores web
		PS	Aproximación a un modelo de recuperación de información personalizada basado en el análisis semántico del contenido
		LM	Descripción de los recursos de información en Internet: formato Dublín Core
		LM	Las herramientas gestionadas por las bibliotecas universitarias españolas en el servicio de la web: creación de la marca en línea
		CG	Los metadatos, la búsqueda y recuperación de información desde las Ciencias de la Información
		LM	Organización y recuperación de información en Internet: teoría de los metadatos
		LM	Personalización y Evaluación XML mediante la Simulación de Perfiles de Usuario y Juicios de Relevancia
		CG	Recuperación de información: un área de investigación en crecimiento
		CG	Taxonomías para la categorización y la organización de la información en sitios web
CATALOGACIÓN	6	LM	Aplicaciones de la web social: herramientas de participación en las bibliotecas universitarias españolas
		LM	Descripción de los recursos de información en Internet: formato Dublín Core
		CG	Elementos para reflexionar sobre el nuevo estándar para la descripción y acceso a recursos.
		CG	Introducción a la Recuperación de Información
		LM	Organización y recuperación de información en Internet: teoría de los metadatos
		LM	Sistemas de metadatos en recuperación de Información: propuesta de modelo de trabajo en el uso RDF sobre directorios LDAP
DIVULGACIÓN CIENTÍFICA	3	CG	Avances en el Desarrollo de un Sistema de Recuperación de Información para Publicaciones Científicas del Área de Ciencias de la Computación.
		CG	Avances en la Construcción de un Sistema de Recuperación de Información para Información Científica en Ciencias de la Computación
		PS	Comparación de redes semánticas naturales mediante un modelo de recuperación de información

EVALUACIÓN DE LA TECNOLOGÍA	10	PS	Análisis de tendencias en la personalización de los resultados en buscadores web
		PS	Aproximación a un modelo de recuperación de información personalizada basado en el análisis semántico del contenido
		PS	Comparación de redes semánticas naturales mediante un modelo de recuperación de información
		OT	Evaluación de sistemas de extracción de frases clave
		OT	Implementación de un sistema de recuperación de la información.
		LM	Personalización y Evaluación XML mediante la Simulación de Perfiles de Usuario y Juicios de Relevancia
		OT	Predicción del rendimiento de consultas basado en rankings de documentos y nuevo marco de evaluación
		PS	Procesamiento Semántico de información en Sistemas de Recuperación de Información
		PS	Propuesta Para Extender Semánticamente: el Proceso De Recuperación De Información
		OT	Prototipo de Herramienta de Software para la Extracción de Información en Documentos PDF Utilizando un Motor de Base de Datos nosql
INDEXACIÓN	2	PS	El papel de los metadatos en la Web Semántica.
		CG	Organización y recuperación de la información: un enfoque desde la perspectiva de la automatización
INDEXACIÓN AUTOMÁTICA	13	PS	Análisis, diseño e implementación de un agente deliberativo para extraer contextos definitorios en textos especializados
		PS	Aplicación de la realimentación por relevancia con base textual para sistemas de recuperación visual
		LM	Aplicaciones de la web social: herramientas de participación en las bibliotecas universitarias españolas
		OT	Asistente para Poblar Repositorios: Recopilación de Objetos Digitales y Extracción de Metadatos
		OT	Clasificación automática de información en portales web mediante técnicas de clustering
		OT	Evaluación de sistemas de extracción de frases clave
		OT	FOG: Arquitectura flexible para la generación automática de ontologías
		OT	Mejora de la recuperación de información en bases de datos de texto utilizando recursos lingüísticos
		OT	Modelo basado en técnicas de procesamiento de lenguaje natural para extraer y anotar información de publicaciones científicas
		LM	Recuperación de información en ficheros XES de gran dimensión mediante técnicas de indexación
		OT	Representación gráfica de documentos para extracción automática de relaciones
		LM	Sistema de recomendación para mejorar la indexación de páginas web en el motor de búsqueda de Google
		LM	Ventajas de la tecnología XML dentro de la Ciencias de la Información
INTERFAZ DE ORDENADORES	2	CG	Desarrollo de una guía de control para interfaces de búsqueda por materia en catálogos en línea de acceso web
		CG	Metodologías para el desarrollo de interfaces visuales de recuperación de información: análisis y comparación.
INTERNET	1	PS	Estudio de la web semántica
LENGUAJE DE INDEXACIÓN	1	CG	Taxonomías para la categorización y la organización de la información en sitios web

LENGUAJE DE PROGRAMACIÓN	16	CG	Avances en la Construcción de un Sistema de Recuperación de Información para Información Científica en Ciencias de la Computación
		LM	Biblioteca Electrónica Textual del teatro español (1868-1936) e investigación con grafos
		OT	Clasificación automática de información en portales web mediante técnicas de clustering
		CG	Componentes y funcionalidades de un sistema de recuperación de la información
		LM	Descripción de los recursos de información en Internet: formato Dublín Core
		LM	Las herramientas gestionadas por las bibliotecas universitarias españolas en el servicio de la web: creación de la marca en línea
		OT	Mejora de la recuperación de información en bases de datos de texto utilizando recursos lingüísticos
		OT	Modelo basado en técnicas de procesamiento de lenguaje natural para extraer y anotar información de publicaciones científicas
		LM	Organización y recuperación de información en Internet: teoría de los metadatos
		LM	Personalización y Evaluación XML mediante la Simulación de Perfiles de Usuario y Juicios de Relevancia
		OT	Predicción del rendimiento de consultas basado en rankings de documentos y nuevo marco de evaluación
		OT	Prototipo de Herramienta de Software para la Extracción de Información en Documentos PDF Utilizando un Motor de Base de Datos nosql
		LM	Recuperación de información en ficheros XES de gran dimensión mediante técnicas de indexación
		LM	Sistemas de metadatos en recuperación de Información: propuesta de modelo de trabajo en el uso RDF sobre directorios LDAP
		LM	Un método eficaz de indexación para la recuperación de imágenes en archivos en formato pdf
		LM	Ventajas de la tecnología XML dentro de la Ciencias de la Información
LINGÜÍSTICA INFORMÁTICA	1	OT	Representación gráfica de documentos para extracción automática de relaciones
ONTOLOGÍA	1	PS	Modelo basado en ontología para implementar Web Semántica que apoye la gestión de la información y el conocimiento
PERSA	1	OT	Recuperación de información en Persa: revisión crítica y propuestas de mejora
PROCESAMIENTO DE DATOS	2	LM	Biblioteca Electrónica Textual del teatro español (1868-1936) e investigación con grafos
		CG	Elementos para reflexionar sobre el nuevo estándar para la descripción y acceso a recursos.
PROCESAMIENTO DE LA INFORMACIÓN	2	CG	Elementos para reflexionar sobre el nuevo estándar para la descripción y acceso a recursos.
		CG	Recuperación de información: reflexiones epistémicas de una ciencia en su estado embrionario

PROGRAMACIÓN INFORMÁTICA	9	PS	Análisis, diseño e implementación de un agente deliberativo para extraer contextos definitorios en textos especializados
		PS	Aplicación de la realimentación por relevancia con base textual para sistemas de recuperación visual
		PS	Aproximaciones para la expansión semántica de consultas de un Sistema de Recuperación de Información Booleano
		OT	FOG: Arquitectura flexible para la generación automática de ontologías
		OT	Implementación de un sistema de recuperación de la información.
		PS	Modelo para la recuperación de información con anotación semántica
		OT	Recuperación de información en Persa: revisión crítica y propuestas de mejora
		PS	Sistema de recuperación de información semántico
		PS	Una aproximación a la recomendación de artículos científicos según su grado de especificidad
RECUPERACIÓN DE INFORMACIÓN	15	PS	Aplicación de la realimentación por relevancia con base textual para sistemas de recuperación visual
		OT	Asistente para Poblar Repositorios: Recopilación de Objetos Digitales y Extracción de Metadatos
		CG	Componentes y funcionalidades de un sistema de recuperación de la información
		CG	El catálogo: un recurso en expansión
		OT	Implementación de un sistema de recuperación de la información.
		CG	Introducción a la Recuperación de Información
		CG	Los metadatos, la búsqueda y recuperación de información desde las Ciencias de la Información
		PS	Modelo para la recuperación de información con anotación semántica
		CG	Organización y recuperación de la información: un enfoque desde la perspectiva de la automatización
		OT	Recuperación de información en Persa: revisión crítica y propuestas de mejora
		CG	Recuperación de información: reflexiones epistémicas de una ciencia en su estado embrionario
		CG	Recuperación de información: un área de investigación en crecimiento
		PS	Sistema de recuperación de información semántico
		LM	Sistemas de metadatos en recuperación de Información: propuesta de modelo de trabajo en el uso RDF sobre directorios LDAP
		LM	Un método eficaz de indexación para la recuperación de imágenes en archivos en formato pdf
		CG	Metodologías para el desarrollo de interfaces visuales de recuperación de información: análisis y comparación.
		PS	Propuesta Para Extender Semánticamente: el Proceso De Recuperación De Información

SEMÁNTICA	11	PS	Análisis de tendencias en la personalización de los resultados en buscadores web
		PS	Aproximación a un modelo de recuperación de información personalizada basado en el análisis semántico del contenido
		PS	Aproximaciones para la expansión semántica de consultas de un Sistema de Recuperación de Información Booleano
		PS	El papel de los metadatos en la Web Semántica.
		PS	Estudio de la web semántica
		PS	Modelo basado en ontología para implementar Web Semántica que apoye la gestión de la información y el conocimiento
		PS	Modelo para la recuperación de información con anotación semántica
		PS	Procesamiento Semántico de información en Sistemas de Recuperación de Información
		PS	Propuesta Para Extender Semánticamente: el Proceso De Recuperación De Información
		PS	Sistema de recuperación de información semántico
		PS	Una aproximación a la recomendación de artículos científicos según su grado de especificidad
SERVICIO BIBLIOGRÁFICO	2	CG	Desarrollo de una guía de control para interfaces de búsqueda por materia en catálogos en línea de acceso web
		CG	El catálogo: un recurso en expansión
SISTEMA DE INFORMACIÓN	7	PS	Aproximaciones para la expansión semántica de consultas de un Sistema de Recuperación de Información Booleano
		PS	Comparación de redes semánticas naturales mediante un modelo de recuperación de información
		PS	Estudio de la web semántica
		PS	Modelo basado en ontología para implementar Web Semántica que apoye la gestión de la información y el conocimiento
		CG	Recuperación de información: reflexiones epistémicas de una ciencia en su estado embrionario
		CG	Recuperación de información: un área de investigación en crecimiento
		PS	Una aproximación a la recomendación de artículos científicos según su grado de especificidad
SISTEMA DE INFORMACIÓN EN LÍNEA	7	CG	Avances en el Desarrollo de un Sistema de Recuperación de Información para Publicaciones Científicas del Área de Ciencias de la Computación.
		CG	Avances en la Construcción de un Sistema de Recuperación de Información para Información Científica en Ciencias de la Computación
		CG	Componentes y funcionalidades de un sistema de recuperación de la información
		CG	Desarrollo de una guía de control para interfaces de búsqueda por materia en catálogos en línea de acceso web
		CG	El catálogo: un recurso en expansión
		CG	Interoperabilidad en los Sistemas de Información Documental (SID): la información debe fluir
		CG	Introducción a la Recuperación de Información
TERMINOLOGÍA	1	CG	Taxonomías para la categorización y la organización de la información en sitios web