

Облікова картка ДіР



I. Загальні відомості

Державний обліковий номер: 0217U001370

Державний реєстраційний номер: 0115U000324

Особливі позначки: відкрита

Дата реєстрації: 07-02-2017

II. Етап виконання ДіР

Номер етапу: 1

Назва етапу: Методи та засоби автоматичної систематизації динамічного науково-освітнього ресурсу глобального електронного інформаційного простору

Початок етапу: 01.2015

Закінчення етапу: 12.2016

Вид звітного документа: Остаточний звіт

III. Відомості про виконавця ДіР

Повне найменування юридичної особи (або ПІБ фізичної особи): Національний технічний університет України "Київський політехнічний інститут", факультет прикладної математики

Код за ЄДРПОУ: 02070921

Місцезнаходження:

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

Розмір організації:

Телефон: (044) 236-32-02

IV. Відомості про співвиконавців ДіР

V. Відомості про замовника ДіР

Повне найменування юридичної особи (або ПІБ фізичної особи): Інститут енергозбереження та енергоменеджменту Національного технічного університету України "Київський Політехнічний Інститут"

Код за ЄДРПОУ: 00027677

Місцезнаходження: пр. Перемоги 37, м. Київ, Київ, 03056, Україна

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

Розмір організації:

Телефон: 0442048428

VI. Джерела, напрями та обсяги фінансування ДіР

Підстава для проведення ДіР: 34 - договір (замовлення) з центральним органом виконавчої влади, академією наук (головними розпорядниками бюджетних коштів на проведення НДДКР)

Напрямок фінансування: 2.2 - прикладні дослідження і розробки

Джерела фінансування

7713 - кошти держбюджету

Код програмної класифікації видатків і кредитування (КПКВК): 2201040

Фактичний обсяг фінансування (тис. грн.): 459.39

VII. Відомості про ДіР

Назва роботи українською:

Методи та засоби автоматичної систематизації динамічного науково-освітнього ресурсу глобального електронного інформаційного простору

Назва роботи англійською:

Methods and means for automatic systematization of dynamic scientific and educational resource of global electronic informational space

Реферат українською:

Об'єкт дослідження: процес доступу до вмісту динамічного інформаційного ресурсу в науково-освітній сфері. Предмет дослідження: основані на використанні теоретичного апарату інтелектуального аналізу даних програмні засоби для автоматичного упорядкування глобального інформаційного ресурсу та для наступного високоефективного пошуку інформації у ньому в ході науково-освітньої діяльності. Мета роботи полягає у створенні інформаційно-алгоритмічних засад автоматичної систематизації електронних інформаційних ресурсів науково-освітнього спрямування та розробці програмних засобів - прототипів розподіленої інформаційно-пошукової системи.

Реферат англійською:

The object of study: access process to content of the dynamic information resource in scientific and educational spheres. Subject of research: software tools based on the using of the theoretical apparatus of data mining for automatic ordering of the global information resource and the following highly effective information search in it during scientific and educational activities. Objective is to create the information and algorithmic foundations of automatic systematization of electronic information resources of scientific and educational direction and to develop software tools - distributed information retrieval system prototypes.

Індекс УДК: 681.3.06:004.92, 681.3

Коди тематичних рубрик: 81.14.11.05

Керівники роботи

Власне Прізвище Ім'я По-батькові: Тарасенко Володимир Петрович

Науковий ступінь:

Вчене звання:

Ідентифікатор ORCID ID:

Додаткова інформація:

VIII. Наукова (науково-технічна) продукція (НТП)

Назва НТП українською: Методи та засоби автоматичної систематизації динамічного науково-освітнього ресурсу глобального електронного інформаційного простору

Назва НТП англійською: Methods and means for automatic systematization of dynamic scientific and educational resource of global electronic informational space

НТП, яку передбачалося створити:

Причини, через які НТП не було створено:

Отримані результати:

Галузь застосування: М.80

Реєстраційний номер картки технології:

Опис НТП: Реалізовано принципово нову архітектуру програмних засобів автоматичної систематизації науково-освітнього сегменту глобального електронного інформаційного простору. Орієнтованість на науково-освітню галузь дозволила використати особливості відповідних даних, у першу чергу той факт, що їх зазвичай можна віднести до класу частково структурованих. Останнє призводить до можливості використання методів структурування, спрямованих на обробку даних цього класу, зокрема методів на базі відбору інформації (Information Extraction). Також були створені методи та відповідні програмні засоби, які реалізують освітньо-орієнтований метод відбору інформації, що розглядає шукані ресурси з точки зору контекстно-атрибутивної моделі.

Соціально-економічна спрямованість НТП:

Вплив НТП на довкілля:

Впровадження НТП: Не впроваджено

Практична реалізація НТП

Початок етапу: 2018

Закінчення етапу:

Споживачі продукції: освітні заклади, науково-дослідні установи

Перспективні ринки: Україна

Характер співробітництва з інвестором

Потрібний обсяг інвестицій, тис. грн.: 2

Права, що надаються інвестору після завершення роботи:

Наявність бізнес-плану:

Техніко-економічне обґрунтування:

Потенціальний обсяг продажу, тис. грн.: 0

Очікуваний термін окупності (років):

Додаткова інформація:

IX. Бібліографічний опис

1. Ingwersen, Peter, and Kalervo Järvelin. The turn: Integration of information seeking and retrieval in context. Vol. 18. Springer Science & Business Media, 2006.
2. Sarawagi, Sunita. Information Extraction. – Mumbai : Indian Institute of Technology, 2008.
3. Jurafsky D, Martin JH. Speech and language processing an introduction to natural language processing, computational linguistics, and speech recognition. – Prentice-Hall, Inc., 2009.
4. Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze, An introduction of information retrieval – Cambridge UP – 2009. – P.58
5. Jenny Rose Finkel, Trond Grenager, and Christopher Manning, Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling – Computer Science Department Stanford University – 2005. – P.7
6. Muslea, Ion, Steven Minton, and Craig A. Knoblock. "Hierarchical wrapper induction for semistructured information sources." Autonomous Agents and Multi-Agent Systems 4.1-2 (2001): 93-114.
7. Rajesh, T., et al. "Data extraction and label assignment for web databases." Compusoft 4.4 (2015): 1628.
8. Active Learning for Incremental POI Extraction and Pairing. 2016. PhD Thesis. National Central University.
9. Krallinger, Martin, et al. "Evaluation of text-mining systems for biology: overview of the Second BioCreative community challenge." Genome biology 9.2 (2008): 1.
10. Cunningham, Hamish. "GATE, a general architecture for text engineering." Computers and the Humanities 36.2 (2002): 223-254.
11. Zikopoulos, Paul, and Chris Eaton. Understanding big data: Analytics for enterprise class hadoop and streaming data. McGraw-Hill Osborne Media, 2011.
12. Максаков А. Сравнительный анализ алгоритмов классификации и способов представления Web-документов / Алексей Максаков // Труды РОМИП. – 2005. – с. 63-73.
13. Jing Han, Meina Song, and Junde Song. "A Novel Solution of Distributed Memory NoSQL Database for Cloud Computing", In ICIS 2011, 10th IEEE/ACIS International Conference on Computer and Information Science, 2011.
14. Nishtha Jatana, Sahil Puri, Mehak Ahuja, Ishita Kathuria, Dishant Gosain, "A Survey and Comparison of Relational and Non-Relational Database," International Journal of Engineering Research & Technology (IJERT), vol. I, no. 6, 2012.
15. S. Weber, NoSQL Databases, University of Applied Sciences HTW Chur, Switzerland, 2010.
16. "MapReduce Debates and Schema-Free", Woohyun Kim, Coord, March 3 2010.
17. "Data-Intensive Text Processing with MapReduce", Jimmy Lin and Chris Dyer, Tutorial at the 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL HLT 2010), June 2010, Los Angeles, California.
18. Borthakur, Dhruba. "HDFS architecture guide." HADOOP APACHE PROJECT <http://hadoop.apache.org/common/docs/current/hdfs design.pdf> (2008): 39.
19. Vora, Mehul Nalin. "Hadoop-HBase for large-scale data." Computer science and network technology (ICCSNT), 2011 international conference on. Vol. 1. IEEE, 2011.
20. Mitzenmacher, Michael. "Distributed, compressed Bloom filter Web cache server." U.S. Patent No. 6,920,477. 19 Jul. 2005.
21. Mobasher, Bamshad. "Data mining for web personalization." The adaptive web. Springer Berlin Heidelberg, 2007. 90-135.
22. Baeza-Yates, Ricardo, and Paolo Boldi. "Web structure mining." Advanced Techniques in Web Intelligence-I. Springer Berlin Heidelberg, 2010. 113-142.
23. Mobasher, Bamshad. "Web usage mining." Web data mining: Exploring hyperlinks, contents and usage data 12 (2006).
24. Замятін Д.С., Петрашенко А.В., Тафля Ф.В. Веб-орієнтовані програмні засоби пошуку текстової інформації // Збірник наукових праць II Всеукраїнської науково-практичної конференції молодих учених та студентів "Інтелектуальні технології в системному програмуванні". – Хмельницький. – 2013. – С. 248-259.
25. Петрашенко А.В., Навка О.С. Способи подання тезаурусу в базах даних пошукових систем // Збірник тез доповідей VI наукової конференції магістрантів та аспірантів "Прикладна математика та комп'ютеринг". – Київ. – 2014. – С. 94-97.
26. B.F. Cooper, A. Silberstein, E. Tam, R. Ramakrishnan, R. Sears Benchmarking cloud serving systems with YCSB. Proceedings of the 1st ACM Symposium on Cloud Computing – SoCC '10, ACM, New York, NY, USA (2010), pp. 143-154
27. Задача класифікації [Електронний ресурс] / Media Wiki, April 01 2015, Ukraine, Kyiv – Режим доступу: https://uk.wikipedia.org/wiki/Задача_класифікації. – (01.04.2015).
28. Лифшиц Ю. Классификация текстов / Юрий Лифшиц. – 2005. – с. 2-5.
29. Д.Сэлмон. Сжатие данных, изображений и звука [Текст]: Москва: Техносфера, 2004. – 368с
30. The Discrete Cosine Transform (DCT) [Електронний ресурс] – Режим доступу: <http://www.cs.cf.ac.uk/Dave/Multimedia/node231.html>
32. Hafed, Ziad M., and Martin D. Levine. "Face recognition using the discrete cosine transform." International Journal of Computer Vision 43.3 (2001): 167-188.
33. Кроновер Р. Фракталы и хаос в динамических системах / Р. Кроновер. – М. : Постмаркет, 2000. – 352 с.
34. Трунова О.В. Использование фрактального анализа для исследования динамики сложных систем / О.В. Трунова, И.С. Скитер // Математическое и имитационное моделирование систем: Восьмая

международная науч.-практ. конф., 24-28 июня 2013 г.; тезисы докл. - Чернигов, ЧНТУ, 2013. - С. 296-299. 35. Тарасенко В.П., Тесленко О.К. Комп'ютерний сервіс для виявлення запозичень у студентських роботах. Журнал "Вища школа", 2015, №2 (127), с.99-107. 36. Щербина О.А. Організація навчання викладачів на робочому місці при впровадженні інформаційно-комунікаційних технологій в університеті // О.А. Щербина // Збірник наукових праць "Педагогічний процес: теорія і практика" . - Вип. 5-6 (50-51) / гол. ред. Сисоева С. О. - К.: Видавництво ВП "Едельвейс", 2015 С. 78-83. 37. Щербина О.А. Experience API - новий стандарт програмних засобів для е-навчання та приклади його практичного використання / О.А. Щербина // Інформаційні технології і засоби навчання, 2016. - Том 53, №3. - С. 150-163. 38. Класифікація документів [Електронний ресурс] / Media Wiki, April 01 2015, Ukraine, Kyiv - Режим доступу: https://uk.wikipedia.org/wiki/Класифікація_документів. - (01.04.2015). 39. Ali M. Alakeel. A Guide to Dynamic Load Balancing in Distributed Computer Systems [Текст]. - IJCSNS. - 2010. 40. Wayne Jansen Timothy Grance. Guidelines on Security and Privacy in Public Cloud Computing. [Текст]. - NIST Draft Special Publication 800-144 - 2011. 41. Ram Prasad Padhy, P Goutam Prasad Rao. Load balancing in cloud computing systems. [Текст]. - Thesis from National Institute of Technology, Rourkela. - 2011. 42. Орлова М.М. Спосіб захисту даних в мережах LTE / М.М. Орлова, В.С. Гудименко // Комп'ютерно-інтегровані технології: освіта, наука, виробництво. Видавництво Луцького НТУ. - 2015. - № 19. - С. 108 - 114. 43. Шитий Д.В. Спосіб резервного копіювання гостей операційних систем в Hyper-V / Д.В. Шитий // Комп'ютерно-інтегровані технології: освіта, наука, виробництво. Видавництво Луцького НТУ. - 2015. - № 20. - С. 97 - 100. 44. М.Н. Орлова, Д.В.Шитий Спосіб распределения нагрузки на ресурсы сетевых узлов Journal of Qafqaz University (Baku): Mathematics and Computer Science 2015. - Vol. 3, № 1. - P. 57 - 61. 45. Yiming Yang. A Reexamination of Text Categorization methods / Yiming Yang, Xin Lui // 22nd International Conference on Research and Development in Information Retrieval (SIGIR'99). - 1998. - pp. 42-49. 46. Lewis D.D. A Classification of two learning algorithms for text categorization / D.D. Lewis and M. Ringuette // Third Annual Symposium on Document Analysis and Information Retrieval (SDAIR'94) . - 1994. - pp. 81-93. 47. Weigend A.S. Exploiting Hierarchy on Text Categorization / A.S. Weigend, E.D. Weiner and J.O. Peterson // Information Retrieval, 1(3). - 1999. - pp.193-216. 48. V. Abramova, J. Bernardino, P. Furtado Which NoSQL database? A performance overview Open J. Databases, 1 (2014), pp. 17-24. 49. Wei-ping, Zhu, L. I. Ming-Xin, and Chen Huan. "Using MongoDB to implement textbook management system instead of MySQL." Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on. IEEE, 2011. 50. Wei, Jianwen, et al. "Analysis farm: A cloud-based scalable aggregation and query platform for network log analysis." Cloud and Service Computing (CSC), 2011 International Conference on. IEEE, 2011. 51. Redmond, Eric, and Jim R. Wilson. Seven databases in seven weeks: a guide to modern databases and the NoSQL movement. Pragmatic Bookshelf, 2012. 52. Abadi, Daniel J., et al. "Materialization strategies in a column-oriented DBMS." 2007 IEEE 23rd International Conference on Data Engineering. IEEE, 2007. 53. Михайлюк, А. Ю. Організація системи автоматичного реферування тексту на основі нечіткої ієрархічної нейронної мережі [Текст] / А. Ю. Михайлюк, О. С. Михайлюк, Л. М. Огнівчук, В. П. Тарасенко // Радіоелектронні і комп'ютерні системи. - 2016. - №7.-С. XX 20 стор. 54. Щербина О.А. Нові засоби для оцінювання компетентностей в Moodle // О.А. Щербина // Інформаційні технології і засоби навчання, 2016. - Том 55, №5. - С. 96-104. 55. Петрашенко А.В., Голубовський О.І. Аналіз засобів семантичного аналізу документів в Internet // Збірник тез доповідей VIII наукової конференції магістрантів та аспірантів "Прикладна математика та комп'ютинг". - Київ. - 2016. - С. 146-150. 56. Index term [Електронний ресурс] / Media Wiki, April 01 2015, Ukraine, Kyiv - Режим доступу: http://en.wikipedia.org/wiki/Index_term. - (01.04.2015). 57. Петрашенко А.В., Тафля Ф.В. Метод класифікації Web-сторінок на основі структурного аналізу // Збірник тез доповідей VII наукової конференції магістрантів та аспірантів "Прикладна математика та комп'ютинг". - Київ. - 2015. - С. 141-145. 58. Тесленко О.К., Степанюк М.Ф. Аналіз реалізації основних алгоритмів нечіткого пошуку// Збірник тез доповідей сьомої наукової конференції магістрантів та аспірантів "Прикладна математика та комп'ютинг" (ПМК-2015), Київ, НТУУ "КПІ", 15-17 квітня 2015 р-с. 166-170 59. Broder, Andrei Z., et al. "Syntactic clustering of the web." Computer Networks and ISDN Systems 29.8 (1997): 1157-1166. 60. Вагин В.Н., Головина Е.Ю., Загорянская А.А., Фомина М.В. Достоверный и правдоподобный вывод в интеллектуальных системах - М:

Физматлит, 2004 - 704 стр. 61. K. Jain and R. C. Dubes, Algorithms for Clustering Data. Prentice-Hall, 1988. 62. L. X. Li, Z. J. Shao and J. X. Qian, "An optimizing Method based on Autonomous Animals: Fish Swarm Algorithm", System Engineering Theory & Practice, 2002. 63. McCallum A. A Comparison of Event Models for Na?ve Bayes Text Classification / A. McCallum, K. Nigam // AAAI-98 Workshop on Learning for Text Categorization. - 1998. - pp. 1-8. 64. Manning C.D. Introduction to Information Retrieval / C.D. Manning, P. Raghavan and H. Sch?tze. - Cambridge University Press, 2008. - p. 506. 65. Dumais S.T. Inductive Learning Algorithms and representations for text categorization / S.T. Dumais, J. Platt, D. Heckerman and M. Sahami // Proceedings of the Seventh International conference on Information and Knowledge Management (CIKM'98). - 1998. - pp. 148-155. 66. Na?ve Bayes classifier [Електронний ресурс] / Media Wiki, April 01 2015, Ukraine, Kyiv - Режим доступу: http://en.wikipedia.org/wiki/Naive_Bayes_classifier. - (01.04.2015). 67. Joachims T. Text Categorization with Support Vector Machines: Learning with Many Relevant Features. // Proceedings of ECML-98, 10th European Conference on Machine Learning - 1998. 68. Yang Y. An Evaluation of Statistical Approaches to Text Categorization / Journal of Information Retrieval, 1999 - V.1 - pp. 67--88. 69. Yang Y., Liu X. A re-examination of text categorization methods. // Proc. of Int. ACM Conference on Research and Development in Information Retrieval (SIGIR-99), 1999 - pp. 42-49. 70. Петрашенко А.В., Замятін Д.С., Максименко І.Г. Кластеризація текстових документів за допомогою модифікованого методу Concept Indexing // Збірник наукових праць ІІ Всеукраїнської науково-практичної конференції молодих учених та студентів "Інтелектуальні технології в системному програмуванні". - Хмельницький. - 2013. - С. 149-156. 71. Афонин А.А., Крейнс М.Г. Кластеризация текстовых коллекций: помощь при содержательном поиске и аналитический инструмент // Интернет-порталы: содержание и технологии, вып. 4. - М.: Просвещение, 2007. - с. 510-537. 72. Cruz I.F. Measuring Structural Similarity Among Web Documents / Measuring Structural Similarity Among Web Documents / I.F. Cruz , S. Borisov, M.A. Marks, T.R. Webb // 7th International Conference on Electronic Publishing. - 1998. - pp.513-524. 73. Sunita Sarawagi, Information Extraction - Indian Institute of Technology, Indian Institute of Technology - 2008. - P.117 74. Christopher D. Manning, Prabhakar Raghavan, Hinrich Sch?tze, An inntodaction of information retrieval - Cambridge UP - 2009. - P.58 75. Jenny Rose Finkel, Trond Grenager, and Christopher Manning, Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling - Computer Science Department Stanford University - 2005. - P.7 76. Eugene Agichtein, Venkatesh Ganti, Mining Reference Tables for Automatic Text Segmentation - Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining - 2004. - P.20-29 77. Петрашенко А.В., Лось В.С. Алгоритми виділення формальних атрибутів у текстових документах // Збірник тез доповідей VII наукової конференції магістрантів та аспірантів "Прикладна математика ти комп'ютинг". - Київ. - 2015. - С. 137-140. 78. Петрашенко А.В., Хижняк В.С. Статистичний метод визначення формальних атрибутів у текстових даних // Збірник тез доповідей VII наукової конференції магістрантів та аспірантів "Прикладна математика ти комп'ютинг". - Київ. - 2015. - С. 146-150. 79. K-nearest neighbor algorithm [Електронний ресурс] / Media Wiki, April 01 2015, Ukraine, Kyiv - Режим доступу: http://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm. - (01.04.2015). 80. Дюк В., Самойленко А. Data Mining: учебный курс. - Питер, 2001. 81. Arthur, D. and Vassilvitskii, S. (2007). "k-means++: the advantages of careful seeding". Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics Philadelphia, PA, USA. pp. 1027-1035. 82. K. Jain and R. C. Dubes, Algorithms for Clustering Data. Prentice-Hall, 1988. 83. L. X. Li, Z. J. Shao and J. X. Qian, "An optimizing Method based on Autonomous Animals: Fish Swarm Algorithm", System Engineering Theory & Practice, 2002. 84. S. L. Ukasik and AK. S?lawomirZ, "Firefly algorithm for Continuous Constrained Optimization Tasks", 1st International Confernce on Computational Collective Intelligence, Semantic Web, Social Networks and Multiagent Systems, Springer-Verlag Berlin, Heidelberg, pp.169-178, 2009. 85. Петрашенко А.В., Бабенко Я.В. Аналіз методів автоматизованого реферування новинних сюжетів українською мовою // Збірник тез доповідей VI наукової конференції магістрантів та аспірантів "Прикладна математика ти комп'ютинг". - Київ. - 2014. - С. 90-93. 86. .NET Framework [Електронний ресурс] / Media Wiki, March 26 2015, Ukraine, Kyiv - Режим доступу: <http://uk.wikipedia.org/wiki/.NET>. - (26.03.2015). 87. Mukherjee S. Automatic Discovery of Semantic Structures in HTML Documents / S. Mukherjee, G. Yang, W. Tan, I.V. Ramakrishnan // 7th International Conference on

Document Analysis and Recognition (ICDAR'2003). - 2003. - pp.245-249. 88. Stepanek J. Comparing web pages in terms of inner structure / J. Stepanek, M. Simkova. - 2013. - p.5. 89. Hunspell [Електронний ресурс] / Media Wiki, April 03 2015, Ukraine, Kyiv - Режим доступу: <https://en.wikipedia.org/wiki/Hunspell>. - (03.04.2015). 90. Information retrieval [Електронний ресурс] / Media Wiki, April 03 2015, Ukraine, Kyiv - Режим доступу: https://en.wikipedia.org/wiki/Information_retrieval. - (03.04.2015). 91. Contingency table [Електронний ресурс] / Media Wiki, April 05 2015, Ukraine, Kyiv - Режим доступу: https://en.wikipedia.org/wiki/Contingency_table. - (05.04.2015). 92. Confusion matrix [Електронний ресурс] / Media Wiki, April 07 2015, Ukraine, Kyiv - Режим доступу: https://en.wikipedia.org/wiki/Contingency_table. - (07.04.2015).

Х. Заключні відомості

Керівник юридичної особи

Дичка Іван Андрійович
Д. Т. Н.

Перелік осіб-виконавців

Єфімов Є.В.

Агапова О.А.

Балицький Ю.С.

Бояринова Ю.Є.

Горюшко Г.О.

Дробязко І.П.

Замятін Д.С.

Клятченко Я.М.

Михайлюк О.С.

Наливайчук Микола Васильович

Наливайчук О.Ю.

Орлова М.М.

Петрашенко А.В.

Плахотний М.В.

Потапова К.Р.

Романкевич В.О.

Сапсай Т.Г.

Тарасенко-Клятченко О.В.

Тесленко О.К.

Хіцко Я.В.

Відповідальний за підготовку облікових документів

Телефон

Реєстратор

**Керівник відділу УкрІНТЕІ, що є
відповідальним за реєстрацію наукової
діяльності**

