7C009 Literature Survey

* Can social media be reliably analysed to provide useful statistical data?

Nadkarni, Prakash M; Ohno-Machado, Lucila; Chapman, Wendy W. (2011) Natural language processing: an introduction. Journal of the American Medical Informatics Association. [Online] **18**(5), pp.544 - 551 [Accessed 19 October 2014] Available at:

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3168328/?tool=pmcentrez&report=abstract>

The paper gives a history of NLP, summarises NLP sub-issues, highlights a selection of NLP being used in the medical field, briefly describes some machine-learning techniques used for NLP, describes how newer NLP systems are designed with examples and concludes on whether NLP applications will become a commodity and how NLP would benefit from this. These claims are precise because of the wealth of references that are used to back themselves up, these references are mainly from around the same time period as the papers publication and come from reputable institutions and journal, therefore they act as their evidence and ensures the papers trustworthiness. The paper only gives a brief description with figures of the techniques/algorithms involved but does serve to introduce and give a basic understanding of what it covers.

Workman, T Elizabeth; Stoddart, Joan M. (2012)Cover ImageRethinking information delivery: using anaturallanguageprocessingapplication for point-of-care data discovery. Journal of the Medical Library Association **[Online] 100**(2), pp.113-120 [Accessed19 October 2014] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3324802/?tool=pmcentrez&report=abstract>

The reference analyses a specific NLP application in combination with an algorithm in its possible use help clinicians make decisions and the results were promising therefore it could be used for this purpose. The paper backs this up by testing this combination with three different methods providing qualitative results that show how promising this system is if used in this field, also the paper has reputable sources from around the same time period which shows they have done their research prior to writing the paper. The paper only briefly mentions NLP but goes into depth on the application, how it works and its possible uses.

Davide Calì, ; Antonio Condorelli; Santo Papa; Marius Rata; Luca Zagarella. (2011) Improving intelligence through use of Natural Language Processing. A comparison between NLP interfaces and traditional visual GIS interfaces. Procedia Computer Science. [Online] **5**, pp.920–925 [Accessed 19 October 2014] Available at: <http://ac.els-cdn.com.ezproxy.wlv.ac.uk/S1877050911004534/1-s2.0-S1877050911004534-main.pdf?\_tid=74f6bff4-5928-11e4-8936-00000aacb360&acdnat=1413899282

\_c766fd55b9e7cb48b0369fec812288fc>

The source analyses the difference between the older geographical information systems(GIS) that did not use NLP and the newer versions that do make use of NLP in the design of their user interface showing how NLP has helped improve said systems. This is backed up by testing traditional and NLP GIS with twenty four queries and three metrics to consider, this therefore gives qualitative evidence of NLP's usefulness in this field. The paper concludes that NLP is very useful for this field and illustrates that despite some improvement being needed, it is a promising alternative to traditional GIS.

Revuelta-Martínez, Alejandro;Rodríguez, Luis;García-Varea, Ismael;Montero, Francisco. (2013) Multimodalinteractionfor informationretrievalusing naturallanguage. Computer Standards & Interfaces. [Online] **35**(5), pp.428-441 [Accessed 19 October 2014] Accessed at: <http://ac.els-cdn.co

m.ezproxy.wlv.ac.uk/S0920548912001262/1-s2.0-S0920548912001262-main.pdf?\_tid=6f4bc884-592c-11e4-99bd-00000aab0f27&acdnat=1413900990\_b91fb6bd33d16e3437ff85363b145c20>

The source shows a design for information retrieval that enables the user to interact with databases using natural language processing with the aim of improving its ease of use particularly in the user interface side of the design.

Jaytrilok Choudhary; Deepak Singh Tomar. (2014) Semi-Automated Ontology building through Natural Language Processing. International journal of computer & technology. [Online] **13**(8), pp.4738-4746 [Accessed 20 October 2014] Available at: <http://ijctonline.com/ojs/index.php/ijct/article/view/908N>

The source introduces the reader to the field of ontology with its issue of requiring a large amount of effort for successful information retrieval and so a proposal to automate the process is discussed which makes use of NLP.

Bitter, Christian; Elizondo, David A; Yang, Yingjie. (2010) Natural language processing: a prolog perspective. Artificial Intelligence Review. [Online] **33**(1), pp. 151 - 173 [Accessed 21 October 2014] Available at: <http://search.proquest.com.ezproxy.wlv.ac.uk/docview/197995849?pq-origsite=summon>

The source gives an introduction to NLP from the side of Prolog a programming language used for developing NLP applications.

Tu, Kun; Cooper, David G; Siegelmann, Hava T. (2009) Memory reconsolidation for natural language processing. Cognitive neurodynamics. [Online] **3**(4), pp.365-372 [Accessed 21 October 2014] Available at: <http://wk6kg9sd8m.search.serialssolutions.com/?ctx\_ver=Z39.88-2004&ctx\_enc=info%3Aofi%2Fenc%3AUTF-8&rfr\_id=info:sid/summon.serialssolutions.com&rft\_val\_fmt=info:ofi/fmt:kev:mtx:journal&rft.genre=article&rft.atitle=Memory+reconsolidation+for+natural+language+processing&rft.jtitle=Cognitive+neurodynamics&rft.au=Tu%2C+Kun&rft.au=Cooper%2C+David+G&rft.au=Siegelmann%2C+Hava+T&rft.date=2009-12-01&rft.issn=1871-4080&rft.eissn=1871-4099&rft.volume=3&rft.issue=4&rft.spage=365&rft\_id=info:pmid/19862641&rft.externalDocID=19862641&paramdict=en-US>

The source presents a design for memory reconsolidation a type of NLP.

Ahmed, S., Pasquier, M., & Qadah, G. (2013). Key issues in conducting sentiment analysis on Arabic

social media text. 2013 9th International Conference on Innovations in Information Technology

(IIT)[Online] 10.1109/Innovations.2013.6544396, pp.72–77 [Accessed 29 October 2014] Available at: <http://ieeexplore.ieee.org.ezproxy.wlv.ac.uk/stamp/stamp.jsp?tp=&arnumber=6544396>

Wan, S., & Paris, C. (2014). Improving government services with social media feedback. In Proceedings of the 19th international conference on intelligent user interfaces. [Online] ACM. doi:10.1145/2557500.2557513, pp.27–36 [Accessed 30 October 2014] Available at: <Y6BR4ZMFv6Gd\_\_Mvhl5za27t19uT4uN\_\_5B\_CCwB5rxJ5fm1EyjKFLZ3qfLWlNVFhCT3EI40hSbcp41FwHcFMKohrg4LazYdvRnaJg6dJaYzIkKjXAKBP1kyrb5IcnliwzJg>

Dredze, M. (2012). How Social Media Will Change Public Health. IEEE Intelligent Systems. [Online] **27**(4), pp.81–84 [Accessed 30 October 2014] Available at:

<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6285937>

Alhanini, Y. (2011). The Enhancement of Arabic Stemming by Using Light Stemming and Dictionary-Based Stemming. Journal of Software Engineering and Applications, 04(09), pp.522–526.

Jivani, A. G. (2011). A Comparative Study of Stemming Algorithms. International Journal of Computer Technology and Applications, 2, 1930–1938. Retrieved from http://wlv.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwXZ25DQJBDEVHiAqQQIQ0MGgOzxUjVhSwDfgYh\_Qf4kUEQAeO3vO3Jdu5nK7B\_zGB362qSjIkT-1Z68hVYmak2hvLz\_H9L8AvB7ebz6Nbl\_t6e\_jPfwAv2\_Aq1sJmb-laZECJUwgwC3HsNTdJhaCwWmDHioiAYQRUNsEpM2zbx5PbW8SeZ3cBU0WnZKUFgTKUlMZsAIAABE1e0AQ2HQ

Sandhya, N., Lalitha, Y. S., Sowmya, V., Anuradha, K., & Govardhan, A. (2011). Analysis of Stemming Algorithm for Text Clustering. International Journal of Computer Science Issues (IJCSI), 8(5), 352–359. Retrieved from http://wlv.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwXV2xDQIxDIwQEyCBKFkgKI4TJ64RLwb4Bd52UrJ\_iZEoAA9gXePzXXMXAuZrin-c0F2IQzGSyqWxSrECWVjnuy-pqv2E738R\_HIIu\_E8hnW5r7dH\_PQDRANoEFEIabQGWmlociGiTOxbfbormUm9svlHNBTOW-U6kTZNDruhIp7C3i32OIcLiGV0gNzU7UqyLloEHTL79WftL-LfM8Q

Kumar, S., Hu, X., & Liu, H. (2014). A behavior analytics approach to identifying tweets from crisis regions. Cover Art: In Proceedings of the 25th ACM conference on Hypertext and social media - HT ’14 [Online]. Santiago, Chile 1-4 September. New York: ACM. [Accessed 6 November 2014]. Available at: <http://dl.acm.org/citation.cfm?id=2631775.2631814>

The paper discusses the issue of gathering tweets from crisis regions only so their analysis can be increasingly efficient as well as to tell whether a tweet is from a crisis region from the tweets information and its owner. It concludes that their method is successful in this task and details the planned future work. The paper includes a wealth of reputable sources on the topics covered that are appropriately referenced as well as their own data from the testing of their method such as in graph form which backs up its conclusion.

Chelaru, S., Herder, E., Naini, K. D., & Siehndel, P. (2014). Recognizing skill networks and their specific communication and connection practices. Cover Art: In Proceedings of the 25th ACM conference on Hypertext and social media - HT ’14 [Online]. Santiago, Chile 1-4 September. New York: ACM. [Accessed 6 November 2014]. Available at: <http://dl.acm.org/citation.cfm?id=2631775.2631801>

Wang, A., Hoang, C., & Kan, M.-Y. (2013). Perspectives on crowdsourcing annotations for natural language processing. Language Resources & Evaluation, **47**(1), pp.9–31 [Accessed 7 November 2014]. Available at: <http://wk6kg9sd8m.search.serialssolutions.com/?ctx\_ver=Z39.882004&ctx\_enc=info%3Aofi%2Fenc%3AUTF8&rfr\_id=info:sid/summon.serialssolutions.com&rft\_val\_fmt=info:ofi/fmt:kev:mtx:journal&rft.genre=article&rft.atitle=Perspectives+on+crowdsourcing+annotations+for+natural+language+processing&rft.jtitle=Language+Resources+and+Evaluation&rft.au=Wang%2C+Aobo&rft.au=Hoang%2C+Cong+Duy+Vu&rft.au=Kan%2C+MinYen&rft.date=20130301&rft.pub=Springer+Netherlands&rft.issn=1574020X&rft.eissn=15740218&rft.volume=47&rft.issue=1&rft.spage=9&rft.epage=31&rft\_id=info:doi/10.1007%2Fs10579-012-9176-1&rft.externalDBID=n%2Fa&rft.externalDocID=2013\_10579\_47\_1

\_9176&paramdict=en-US>

The source describes crowd sourcing as well as annotation programs, gives opinions on the possibilities crowd sourcing has for the future, it concludes that humans and computers can work together to solve problems where computers cannot do the task themselves and that while research into this area is relatively new it shows promise. It is backed up in its bibliography by reputable sources of information on the topics covered which are at the forefront of current research as well as qualitative evidence of their testing of several applications designed for the task at hand.