







outcomes in treating AD. The work reported here demonstrates the value of a TM approach in distilling knowledge from research findings related to AD that are reported in text. Better knowledge of the availability and impact of treatments should eventually lead to treatment and care regimens that achieve the goal of improved quality of life in older adults and caregivers [3].

The present results show that a carefully labeled training set corpus can form a good basis for subsequently automated entity recognition within medical research publications. While the present research focused only on AD it seems likely that similar results should be obtained with other medical syndromes and contexts.

In future research, it would be good to use a larger corpus, a larger set of entity types, and more extensive test data in showing that the methods introduced here can be scaled up to more extensive distillations of knowledge within the research literature. NER would then need to be followed by additional analyses to identify useful treatment guidelines and options, so that useful clinical evidence can be synthesized from large volumes of research literature.

**Table 2: Example of unseen Interventions**

Search criteria	Intervention
Normal search	unfamiliar music, dance therapy, massage, environmental cueing, Chinese medicine, coral calcium, animal-assisted therapy, multi-sensory therapy,
Past week	physical therapy, intranasal insulin therapy
Past month	physical therapy, intranasal insulin therapy, antiepileptic drugs, light therapy, flashing light therapy, mouse's spine
Past year	flashing light therapy, routine screening, biogen therapy, nilvadipine, BACE inhibitors

## 6. ACKNOWLEDGMENTS

This work was partially supported by the Petchra Pra Jom Klao Doctoral Scholarship of King Mongkut's University of Technology Thonburi.

## 7. REFERENCES

- [1] Campos, D., Matos, S., and Oliveira, J. L. 2013. Gimli: open source and high-performance biomedical name recognition. *BMC bioinformatics*, 14(1), 54.
- [2] Campos, D., Oliveira, J. L., and Matos, S. 2012. Biomedical named entity recognition: a survey of machine-learning tools. INTECH Open Access Publisher.
- [3] Chan, J. H. Digital information and communication in the ageing society. In *Proceedings of the 6th International Conference on Applications of Digital Information and Web Technologies* 2015. DOI= <http://dx.doi.org/10.3233/978-1-61499-503-6-3>.
- [4] Cohen, A. M., and Hersh, W. R. 2005. A survey of current work in biomedical text mining. *Briefings in bioinformatics*, 6(1), 57-71.
- [5] Eltyeb, S., and Salim, N. 2014. Chemical named entities recognition: a review on approaches and applications. *Journal of cheminformatics*, 6(1), 17.
- [6] Finkel, J. R., Grenager, T., and Manning, C. Incorporating non-local information into information extraction systems by gibbs sampling. In *Proceedings of the 43rd annual meeting on association for computational linguistics*. (June. 2005), 363-370. Association for Computational Linguistics.
- [7] Kongburan, W., Padungweang, P., Krathu, W., and Chan, J. H. Semi-automatic construction of thyroid cancer intervention corpus from biomedical abstracts. In *Eighth International Conference on Advanced Computational Intelligence*, ICACI, (February. 2016), 150-157.
- [8] Lafferty, J., McCallum, A., and Pereira, F. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the eighteenth international conference on machine learning*, ICML, Vol. 1, (June. 2001), 282-289.
- [9] Leaman, R., and Gonzalez, G. BANNER: an executable survey of advances in biomedical named entity recognition. In *Pacific symposium on Biocomputing*, Vol. 13, (January. 2008), 652-663.
- [10] Rocktäschel, T., Weidlich, M., and Leser, U. 2012. ChemSpot: a hybrid system for chemical named entity recognition. *Bioinformatics*, 28(12), 1633-1640.