

## Supplementary Materials

Table S1: Examples of generated citation texts

<p><b>Citing paper title:</b> “A Mixture Model for Expert Finding”[1]</p> <p><b>Citing paper abstract:</b> “This paper addresses the issue of identifying persons with expertise knowledge on a given topic. Traditional methods usually estimate the relevance between the query and the support documents of candidate experts using, for example, a language model. However, the language model lacks the ability of identifying semantic knowledge, thus results in some right experts cannot be found due to not occurrence of the query terms in the support documents. In this paper, we propose a mixture model based on Probabilistic Latent Semantic Analysis (PLSA) to estimate a hidden semantic theme layer between the terms and the support documents. The hidden themes are used to capture the semantic relevance between the query and the experts. We evaluate our mixture model in a real-world system, ArnetMiner. Experimental results indicate that the proposed model outperforms the language models.”</p>
<p><b>Reference paper title:</b> “Retrieval evaluation with incomplete information”[2]</p> <p><b>Reference paper abstract:</b> “This paper examines whether the Cranfield evaluation methodology is robust to gross violations of the completeness assumption (i.e., the assumption that all relevant documents within a test collection have been identified and are present in the collection). We show that current evaluation measures are not robust to substantially incomplete relevance judgments. A new measure is introduced that is both highly correlated with existing measures when complete judgments are available and more robust to incomplete judgment sets. This finding suggests that substantially larger or dynamic test collections built using current pooling practices should be viable laboratory tools, despite the fact that the relevance information will be incomplete and imperfect.”</p>
<p><b>Citation Intent:</b> Method</p> <p><b>Original Citation Text:</b> “As a result, we use the method of pooled relevance judgments #REF together with human judgments.”</p> <p><b>Generated Citation Texts:</b> “We compare the cranfield evaluation methodology with the methods of #REF and #OTHERREF to estimate the semantic relevance of the support documents.”</p>
<p><b>Citing paper title:</b> “Temporally Graded Activation of Neocortical Regions in Response to Memories of Different Ages”[3]</p> <p><b>Citing paper abstract:</b> “The temporally graded memory impairment seen in many neurobehavioral disorders implies different neuroanatomical pathways and/or cognitive mechanisms involved in storage and retrieval of memories of different ages. A dynamic interaction between medial-temporal and neocortical brain regions has been proposed to account for memory’s greater permanence with time. Despite considerable debate concerning its time-dependent role in memory retrieval, medial-temporal lobe activity has been well studied. However, the relative participation of neocortical regions in recent and remote memory retrieval has received much less attention. Using functional magnetic resonance imaging, we demonstrate robust, temporally graded signal differences in posterior cingulate, right middle frontal, right fusiform, and left middle temporal regions in healthy older adults during famous name identification from two disparate time epochs. Importantly, no neocortical regions demonstrated greater response to older than to recent stimuli. Our results suggest a possible role of these neocortical regions in temporally dating items in memory and in establishing and maintaining memory traces throughout the lifespan. Theoretical implications of these findings for the two dominant models of remote memory functioning (Consolidation Theory and Multiple Trace Theory) are discussed.”</p>
<p><b>Reference paper title:</b> “Regional cortical thickness matters in recall after months more than minutes”[4]</p> <p><b>Reference paper abstract:</b> “The aim of this study was to determine the role of regional cortical thickness in recall of verbal material over an extended time period. MRI scans of healthy adults of varying ages were obtained. Verbal memory tests assessing recall after 5 min, 30 min, and a mean interval of 83 days were administered. A general linear model (GLM) of the effects of thickness at each vertex on the different memory indices was computed, controlling for gender, age, IQ, and intracranial volume. These analyses were repeated with hippocampal volume as an additional variable to be controlled for, to assess to which extent effects of cortical thickness were independent of hippocampal size. Minute effects of cortical thickness were observed with regard to shorter time intervals (5 and 30 min). However, even when controlling for the effects of hippocampal volume, higher recall across months was associated with thicker cortex of distinct areas including parts of the gyrus rectus, the middle frontal gyrus, the parieto-occipital</p>

---

sulcus and the lingual gyrus of both hemispheres. In addition, hemisphere-specific associations were found in parts of the right temporal and parietal lobe as well as parts of the left precuneus. This supports a unique and critical role of the thickness of distinct cortical areas in recall after months, more than after minutes.”

**Citation Intent:** Background

**Original Citation Text:** “Interestingly, increases in cortical thickness in this region have been associated with recall of verbal material after months, but not after minutes #REF.”

**Generated Citation Texts:** “In addition, cortical thickness has been shown to be associated with higher recall of verbal material over time #REF.”

---

**Citing paper title:** “A Unified Approach for Taxonomy-based Technology Forecasting”[5]

**Citing paper abstract:** “For decision makers and researchers working in a technical domain, understanding the state of their area of interest is of the highest importance. For this reason, we consider in this chapter, a novel framework for Web-based technology forecasting using bibliometrics (i.e. the analysis of information from trends and patterns of scientific publications). The proposed framework consists of a few conceptual stages based on a data acquisition process from bibliographic online repositories: extraction of domain relevant keywords, the generation of taxonomy of the research field of interests and the development of early growth indicators which helps to find interesting technologies in their first phase of development. To provide a concrete application domain for developing and testing our tools, we conducted a case study in the field of renewable energy and in particular one of its subfields: Waste-to-Energy (W2E). The results on this particular research domain confirm the benefit of our approach.”

**Reference paper title:** “Taxonomy based trend discovery of renewable energy technologies in desalination and power generation”[6]

**Reference paper abstract:** “Renewable energy (RE) technologies are increasingly viewed as critically important since the noticeable depletion of fossil fuel. Knowledge that facilitates forecasting the likely growth and consequences of emergent technologies is essential for wellinformed technology management. Acquiring and analyzing such knowledge is hampered by the amount of data available in publications. In order to elucidate the advance of technologies, we want to address questions like: “How many scientific articles have been published in solar energy recently?” Intelligent search techniques capable of grouping semantically similar concepts are therefore needed, such that e.g. the term ”parabolic trough” is subsumed under solar energy related technologies and hence articles about it should be included in the analysis. The novelty of this work is the deployment of a large, high quality RE-taxonomy for comprehensive trend discovery in publications and patents. We report interesting trends of renewables in two case studies: power generation and desalination techniques. While all major renewables —except geothermal—recently boomed in power generation in terms of publication volume, leading to a nearly equibalanced diversification, patents only reflect strong growth for wind and solar. Renewables in desalination, in particular reverse osmosis, are mainly solar and wind with a slight upward trend of biofuels in publications, hereas other renewables are still in experimental stage.”

**Citation Intent:** ResultComparison

**Original Citation Text:** “The findings for Biodiesel are consistent with earlier results reported in #REF even though a different corpus and a different keyword set was used.”

**Generated Citation Texts:** “These results are in line with the results of #REF, who found that the number of publications in the field of renewable energy was higher than the number in the literature of other renewable energy technologies.”

---

**Citing paper title:** “A Multi-Threading Architecture to Support Interactive Visual Exploration”[7]

**Citing paper abstract:** “During continuous user interaction, it is hard to provide rich visual feedback at interactive rates for datasets containing millions of entries. The contribution of this paper is a generic architecture that ensures responsiveness of the application even when dealing with large data and that is applicable to most types of information visualizations. Our architecture builds on the separation of the main application thread and the visualization thread, which can be cancelled early due to user interaction. In combination with a layer mechanism, our architecture facilitates generating previews incrementally to provide rich visual feedback quickly. To help avoiding common pitfalls of multi-threading, we discuss synchronization and communication in detail. We explicitly denote design choices to control trade-offs. A quantitative evaluation based on the system VISPLORE shows fast visual feedback during continuous interaction even for millions of entries. We describe instantiations of our architecture in additional tools.”

---

---

**Reference paper title:** “SimVis: Interactive visual analysis of large and time-dependent 3D simulation data”[8]

---

**Reference paper abstract:** “SimVis is a novel technology for the interactive visual analysis of large and complex flow data which results from computational fluid dynamics (CFD) simulation. The new technology which has been researched and developed over the last years at the VRVis Research Center in Vienna, introduces a new approach for interactive graphical exploration and analysis of time-dependent data (computed on large three-dimensional grids, and resulting in a multitude of different scalar/vector values for each cell of these grids). In this paper the major new technological concepts of the SimVis approach are presented and real-world application examples are given.”

---

**Citation Intent:** Background

---

**Original Citation Text:** “Previous analyses of the dataset have been conducted using the SIMVIS system #REF, which also implements the proposed architecture #OTHERREF.”

---

**Generated Citation Texts:** “In this paper, we compare our architecture with the Vrvis approach #REF, which provides fast visual feedback even when dealing with large data.”

---

---

**Citing paper title:** “Stream-Packing: Resource Allocation in Web Server Farms with a QoS Guarantee”[9]

---

**Citing paper abstract:** “Current web server farms have simple resource allocation models. One model used is to dedicate a server or a group of servers for each customer. Another model partitions physical servers into logical servers and assigns one to each customer. Yet another model allows customers to be active on multiple servers using load-balancing techniques. The ability to handle peak loads while minimizing cost of resources required on the farm is a subject of ongoing research. We improve resource utilization through sharing. Customer load is expressed as a multidimensional probability distribution. Each customer is assigned to a server so as to minimize the total number of servers needed to host all the customers. We use the notion of complementarity of customers in simple heuristics for this stochastic vector-packing problem. The proposed method generates a resource allocation plan while guaranteeing a QoS to each customer. Simulation results justify our scheme.”

---

**Reference paper title:** “Charging and Accounting for Bursty Connections”[10]

---

**Reference paper abstract:** “Statistical sharing over several timescales is a key feature of the Internet, and is likely to be an essential aspect of future ATM networks. In this chapter we describe how usage-sensitive pricing can encourage statistical sharing, and we provide a quantitative framework for the discussion of pricing issues in systems where statistical sharing is important. In particular we describe a simple charging and accounting mechanism for real-time bursty connections, based on the concept of an effective bandwidth. The mechanism performs the dual role of conveying information to the network that allows more efficient statistical sharing, and information to the user about the resource implications of differently policed connection requests. The resulting tariff takes a strikingly simple form: a charge  $a(x)$  per unit time, a charge  $b(x)$  per unit volume of traffic carried, and a charge  $c(x)$  per connection, where the triple  $(a(x); b(x); c(x))$  are fixed at the time of connection acceptance as a function of the connection contract  $x$ . This form of tariff is also able to reveal user preferences concerning delay tolerant traffic, and thus promises to provide a unified pricing model over a wide range of quality of service classes.”

---

**Citation Intent:** Method

---

**Original Citation Text:** “We use an existing notion of QoS #REF and improve it to ensure a high probability of resource availability to each client.”

---

**Generated Citation Texts:** “In #REF, we use a simple charging and accounting mechanism for real-time bursty connections, based on the concept of an effective bandwidth. ”

---

---

**Citing paper title:** “Retrieval of Images Using Mean-Shift and Gaussian Mixtures Based on Weighted Color Histograms”[11]

---

**Citing paper abstract:** “The topic of this paper is Content-Based Image Retrieval (CBIR) based on colors as a content image descriptor. The tool we developed to that purpose modelizes the colors of an image as a set of 2D Gaussian distributions based on weighted color histograms. Then, given a reference image proposed by a user, the system can automatically classify the image and provide the user with the most similar images to the reference image in its category. Experiments with Corel-1000 dataset demonstrate that our method outperforms the known implementations.”

---

**Reference paper title:** “Fuzzy indexing for Bag of Features scene categorization”[12]

---

---

**Reference paper abstract:** “This paper presents a novel Bag of Features (BoF) method for image classification. The BoF approach describes an image as a set of local descriptors using a histogram, where each bin represents the importance of a visual word. This indexing approach has been frequently used for image classification, and we have seen several implementations, but crucial representation choices —such as the weighting schemes —have not been thoroughly studied in the literature. In our work, we propose a Fuzzy model as an alternative to known weighting schemes in order to create more representative image signatures. Furthermore, we use the Fuzzy signatures to train the Gaussian Naïve Bayesian Network and classify images. Experiments with Corel-1000 dataset demonstrate that our method outperforms the known implementations.”

---

**Citation Intent:** ResultComparison

---

**Original Citation Text:** “Similar results have been obtained on Corel-1000 using a different method, using SIFT points as local image descriptors to build bags of features (BoF) over the whole database #REF.”

---

**Generated Citation Texts:** “Our results are in accordance with the results of #REF, which are consistent with those of #OTHERREF and #REF.”

---

---

**Citing paper title:** “Social networking and academic performance: A review”[13]

---

**Citing paper abstract:** “The ubiquitous use of social networking sites by students and the potential impacts of such use on academic performance are of both theoretical and practical importance. Hence, this paper addresses the question: how does the use of social networking sites influence academic performance? The present review synthesizes the empirical findings of the extant literature, via a systematic review, that examines the efforts that have been made to explicate the association between the use of social networking sites and academic performance. The review of 23 peer-reviewed papers highlights mixed findings regarding the relationship between social network use and academic performance—serving as a call for further research.”

---

**Reference paper title:** “Impact of social media usage on students academic performance in Saudi Arabia”[14]

---

**Reference paper abstract:** “The quantitative study to explore the most popular amongst Saudi students. The study establishes relation among the social media usage and academic grades. Poor time management and football also have negative impact on academic grades. Normality tests are to find relationship among social media usage and GPA scores. The hypothesis about existence of relationship among GPA and SN is not satisfied. Social media is a popular method for communication amongst university students in Saudi Arabia. However excessive social media use can raise questions about whether academic performance is affected. This research explores this question by conducting a survey on university students in Saudi Arabia in regards to social media usage and their academic performance. The survey also explored which social network is the most popular amongst Saudi students, what students thought about their social media usage and factors besides social media usage which negatively affect academic performance. The survey received 108 responses and descriptive statistics including normality tests i.e. scatter plots were used to examine the relationship between the average number of hours students spent of social media a week and GPA scores of the students. The results demonstrated that there was no linear relationship between social media usage in a week and GPA score. Students highlighted that besides social media use, time management is a factor which affects students ’studies negatively. The findings of the paper can be used to propose the effective plans for improving the academic performance of the students in such a way that a balance in the leisure, information exchange and academic performance can be maintained.”

---

**Citation Intent:** Background

---

**Original Citation Text:** “However, within this stream of work, the SNS use and academic performance link is unstable and widely contested because of the mixed findings within published results #REF.”

---

**Generated Citation Texts:** “Social media use is a popular method for communication among university students in the country #REF.”

---

---

**Citing paper title:** “Knowledge integration for improving performance in LVCSR”[15]

---

**Citing paper abstract:** “This paper presents a knowledge integration framework to improve performance in large vocabulary continuous speech recognition. Two types of knowledge sources, manner attribute and prosodic structure, are incorporated. For manner of articulation, six attribute detectors trained with an American English corpus (WSJ0) are utilized to rescore hypothesized phones

---

---

in word lattices obtained by a baseline ASR system. For the prosodic structure, models trained with an unsupervised joint prosody labeling and modeling (PLM) technique using WSJ0 are used in lattice rescoring. Experimental results on the American English WSJ word recognition task of the Nov92 test set show that the proposed approach significantly outperforms the baseline system that does not use articulatory and prosodic information. The results also demonstrate the effectiveness and usefulness of the PLM technique in constructing prosodic models for American English ASR.”

---

**Reference paper title:** “Improved Speech Recognition using Acoustic and Lexical Correlates of Pitch Accent in a N-Best Rescoring Framework”[16]

---

**Reference paper abstract:** “Most statistical speech recognition systems make use of segmentlevel features, derived mainly from spectral envelope characteristics of the signal, but ignore supra-segmental cues that carry additional information likely to be useful for speech recognition. These cues, which constitute the prosody of the utterance and occur at the syllable, word and utterance level, are closely related to the lexical and syntactic organization of the utterance. In this paper, we explore the use of acoustic and lexical correlates of a subset of these cues in order to improve recognition performance on a read-speech corpus, using word error rate (WER) as the metric. Using the features and methods described in this paper, we were able to obtain a relative WER improvement of 1.3% over a baseline ASR system on the Boston University Radio News Corpus.”

---

**Citation Intent:** Method

---

**Original Citation Text:** “The proposed knowledge-assisted #REF was evaluated on the 5kword speaker independent WSJ0 task of Nov92 #OTHERREF.”

---

**Generated Citation Texts:** “For the speech recognition task of the nov92 test set, we used the WSBJ0 word recognition system #REF, which has been shown to outperform the baseline ASR system.”

---

---

**Citing paper title:** “DeepProbLog: Neural Probabilistic Logic Programming”[17]

---

**Citing paper abstract:** “We introduce DeepProbLog, a probabilistic logic programming language that incorporates deep learning by means of neural predicates. We show how existing inference and learning techniques can be adapted for the new language. Our experiments demonstrate that DeepProbLog supports both symbolic and subsymbolic representations and inference, 1) program induction, 2) probabilistic (logic) programming, and 3) (deep) learning from examples. To the best of our knowledge, this work is the first to propose a framework where general-purpose neural networks and expressive probabilistic-logical modeling and reasoning are integrated in a way that exploits the full expressiveness and strengths of both worlds and can be trained end-to-end based on examples.”

---

**Reference paper title:** “Programming with a Differentiable Forth Interpreter”[18]

---

**Reference paper abstract:** “Given that in practice training data is scarce for all but a small set of problems, a core question is how to incorporate prior knowledge into a model. In this paper, we consider the case of prior procedural knowledge for neural networks, such as knowing how a program should traverse a sequence, but not what local actions should be performed at each step. To this end, we present an end-to-end differentiable interpreter for the programming language Forth which enables programmers to write program sketches with slots that can be filled with behaviour trained from program input-output data. We can optimise this behavior directly through gradient descent techniques on user-specified objectives, and also integrate the program into any larger neural computation graph. We show empirically that our interpreter is able to effectively leverage different levels of prior program structure and learn complex behaviours such as sequence sorting and addition. When connected to outputs of an LSTM and trained jointly, our interpreter achieves state-of-the-art accuracy for end-to-end reasoning about quantities expressed in natural language stories.”

---

**Citation Intent:** ResultComparison

---

**Original Citation Text:** “On the WAPs #OTHERREF, DeepProbLog reaches an accuracy between 96% and 97%, similar to #REF(96%).”

---

**Generated Citation Texts:** “Our results are in line with previous work #REF, where we show that we can effectively leverage prior procedural knowledge for neural networks to learn complex behaviours such as sequence sorting and addition.”

---

## References

1. J. Zhang, J. Tang, L. Liu, J. Li, A mixture model for expert finding, in: Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer, 2008, pp. 466-478.
2. C. Buckley, E.M. Voorhees, Retrieval evaluation with incomplete information, in: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval, 2004, pp. 25-32.
3. J.L. Woodard, M. Seidenberg, K.A. Nielson, S.K. Miller, M. Franczak, P. Antuono, K.L. Douville, S.M. Rao, Temporally graded activation of neocortical regions in response to memories of different ages, *Journal of Cognitive Neuroscience*, 19 (2007) 1113-1124.
4. K.B. Walhovd, A.M. Fjell, A.M. Dale, B. Fischl, B.T. Quinn, N. Makris, D. Salat, I. Reinvang, Regional cortical thickness matters in recall after months more than minutes, *Neuroimage*, 31 (2006) 1343-1351.
5. A. Henschel, E. Casagrande, W.L. Woon, I. Janajreh, S. Madnick, A unified approach for taxonomy-based technology forecasting, in: Business Intelligence Applications and the Web: Models, Systems and Technologies, IGI Global, 2012, pp. 178-197.
6. G. Dawelbait, T. Mezher, W.L. Woon, A. Henschel, Taxonomy based trend discovery of renewable energy technologies in desalination and power generation, in: PICMET 2010 TECHNOLOGY MANAGEMENT FOR GLOBAL ECONOMIC GROWTH, IEEE, 2010, pp. 1-8.
7. H. Piringer, C. Tominski, P. Muigg, W. Berger, A multi-threading architecture to support interactive visual exploration, *IEEE Transactions on Visualization and Computer Graphics*, 15 (2009) 1113-1120.
8. H. Doleisch, SimVis: Interactive visual analysis of large and time-dependent 3D simulation data, in: 2007 Winter Simulation Conference, IEEE, 2007, pp. 712-720.
9. J. Shahabuddin, A. Chrungoo, V. Gupta, S. Juneja, S. Kapoor, A. Kumar, Stream-packing: Resource allocation in web server farms with a qos guarantee, in: International Conference on High-Performance Computing, Springer, 2001, pp. 182-191.
10. F.P. Kelly, Charging and accounting for bursty connections, Citeseer, 1996.
11. M.A. Bouker, E. Hervet, Retrieval of images using mean-shift and gaussian mixtures based on weighted color histograms, in: 2011 Seventh International Conference on Signal Image Technology & Internet-Based Systems, IEEE, 2011, pp. 218-222.
12. W. Bouachir, M. Kardouchi, N. Belacel, Fuzzy indexing for bag of features scene categorization, in: 2010 5th International Symposium On I/V Communications and Mobile Network, IEEE, 2010, pp. 1-4.
13. T. Doleck, S. Lajoie, Social networking and academic performance: A review, *Education and Information Technologies*, 23 (2018) 435-465.
14. E. Alwagait, B. Shahzad, S. Alim, Impact of social media usage on students academic performance in Saudi Arabia, *Computers in Human Behavior*, 51 (2015) 1092-1097.
15. C.-Y. Chiang, S.M. Siniscalchi, S.-H. Chen, C.-H. Lee, Knowledge integration for improving performance in LVCSR, in: INTERSPEECH, 2013, pp. 1786-1790.
16. S. Ananthakrishnan, S. Narayanan, Improved speech recognition using acoustic and lexical correlates of pitch accent in a n-best rescoring framework, in: 2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07, IEEE, 2007, pp. IV-873-IV-876.
17. R. Manhaeve, S. Dumancic, A. Kimmig, T. Demeester, L. De Raedt, Deepproblog: Neural probabilistic logic programming, *Advances in Neural Information Processing Systems*, 31 (2018).
18. M. Bošnjak, T. Rocktäschel, J. Naradowsky, S. Riedel, Programming with a differentiable forth interpreter, in: International conference on machine learning, PMLR, 2017, pp. 547-556.