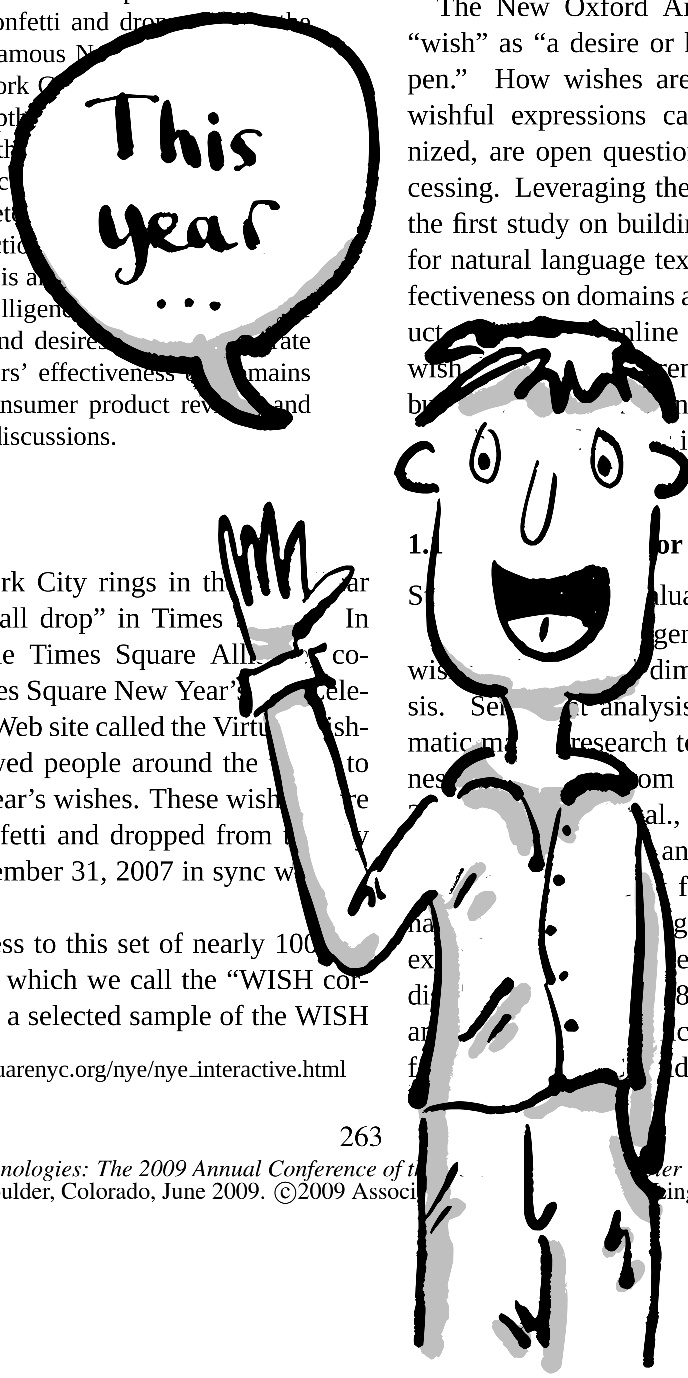
May All Your Wishes Come True: A Study of Wishes and How to Recognize Them

Andrew B. Goldberg, Nathanael Fillmore, David Andrzejewski
Zhiting Xu, Bryan Gibson, Xiaojin Zhu
Computer Sciences Department, University of Wisconsin-Madison, Madison, WI 53706, USA
{goldberg, nathanae, andrzeje, zhiting, bgibson, jerryzhu}@cs.wisc.edu

Abstract

A wish is “a desire or hope for something to happen.” In December 2007, people from around the world offered up their wishes to be printed on confetti and drop[illegible] the sky during the famous N[illegible] drop” in New York C[illegible] present an in-depth [illegible] of wishes. We th[illegible] source to conduc[illegible] general “wish det[illegible] text. Wish detecti[illegible] sentiment analysis a[illegible] ing business intelligen[illegible] world's wants and desire[illegible]rate the wish detectors' effectiveness [illegible]mains as diverse as consumer product rev[illegible] and online political discussions.



1 Introduction

Each year, New York City rings in th[illegible]ar with the famous “ball drop” in Times [illegible] In December 2007, the Times Square All[illegible] co-producer of the Times Square New Year's [illegible]celebration, launched a Web site called the Virtu[illegible]ishing Wall[1] that allowed people around the [illegible] to submit their New Year's wishes. These wish[illegible]re then printed on confetti and dropped from t[illegible]y at midnight on December 31, 2007 in sync w[illegible] ball drop.

We obtained access to this set of nearly 100[illegible] New Year's wishes, which we call the “WISH corpus.” Table 1 shows a selected sample of the WISH corpus. Some are far-reaching fantasies and aspirations, while others deal with everyday concerns like economic and medical distress. We analyze this first-of-its-kind corpus in Section 2.

The New Oxford American Dictionary defines “wish” as “a desire or hope for something to happen.” How wishes are expressed, and how such wishful expressions can be automatically recognized, are open questions in natural language processing. Leveraging the WISH corpus, we conduct the first study on building general “wish detectors” for natural language text, and demonstrate their effectiveness on domains as diverse as consumer product[illegible]uct [illegible]online political discussions. Such wish[illegible]remendous value in collecting b[illegible]nd public opinions. We dis[illegible] in Section 3, and experimen[illegible]

1.1 [illegible]or Work

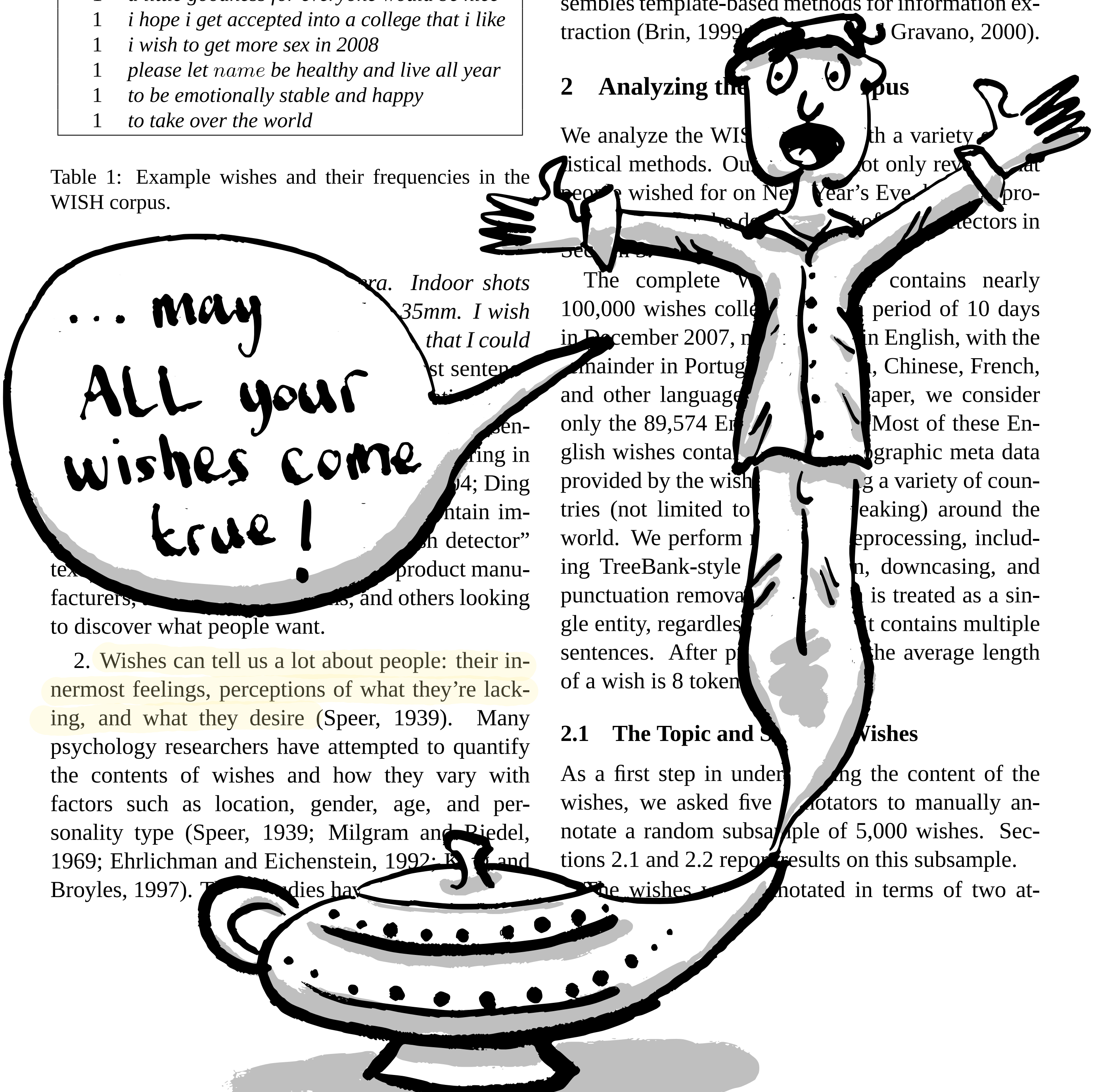
S[illegible]aluable in at least two aspects: [illegible]genre of subjective expression, wis[illegible]dimension to sentiment analysis. Se[illegible]t analysis is often used as an automatic ma[illegible]research tool to collect valuable busines[illegible]om online text (Pang and Lee, 2[illegible]al., 2005; Koppel and Shtrim[illegible]and Malouf, 2008). Wishes [illegible] focus of sentiment analysis, [illegible]g, by revealing what people ex[illegible]en, not just what they like or di[illegible]8; Hu and Liu, 2004). For exam[illegible]ct reviews could contain new f[illegible]der the following (real) prod-

[1] http://www.timessquarenyc.org/nye/nye_interactive.html

Human Language Technologies: The 2009 Annual Conference of t[illegible]er of the ACL, pages 263–271,
Boulder, Colorado, June 2009. ©2009 Associ[illegible]inguistics

514	*peace on earth*
351	*peace*
331	*world peace*
244	*happy new year*
112	*love*
76	*health and happiness*
75	*to be happy*
51	*i wish for world peace*
21	*i wish for health and happiness for my family*
21	*let there be peace on earth*
16	*i wish u to call me if you read this 555-1234*
16	*to find my true love*
8	*i wish for a puppy*
7	*for the war in iraq to end*
6	*peace on earth please*
5	*a free democratic venezuela*
5	*may the best of 2007 be the worst of 2008*
5	*to be financially stable*
1	*a little goodness for everyone would be nice*
1	*i hope i get accepted into a college that i like*
1	*i wish to get more sex in 2008*
1	*please let* name *be healthy and live all year*
1	*to be emotionally stable and happy*
1	*to take over the world*

Table 1: Example wishes and their frequencies in the WISH corpus.



...ra. *Indoor shots* ... *35mm. I wish* ... *that I could* ... st senten... sen- ... ring in ...4; Ding ... ntain im- ... h detector" tex... product manufacturers, ... s, and others looking to discover what people want.

2. Wishes can tell us a lot about people: their innermost feelings, perceptions of what they're lacking, and what they desire (Speer, 1939). Many psychology researchers have attempted to quantify the contents of wishes and how they vary with factors such as location, gender, age, and personality type (Speer, 1939; Milgram and Riedel, 1969; Ehrlichman and Eichenstein, 1992; K... and Broyles, 1997). T... udies ha...

with only dozens or hundreds of participants. The WISH corpus provides the first large-scale collection of wishes as a window into the world's desires.

Beyond sentiment analysis, classifying sentences as wishes is an instance of non-topical classification. Tasks under this heading include computational humor (Mihalcea and Strapparava, 2005), genre classification (Boese and Howe, 2005), authorship attribution (Argamon and Shimoni, 2003), and metaphor detection (Krishnakumaran and Zhu, 2007), among others (Mishne et al., 2007; Mihalcea and Liu, 2006). We share the common goal of classifying text into a unique set of target categories (in our case, wishful and non-wishful), but use different techniques catered to our specific task. Our feature-generation technique for wish detection resembles template-based methods for information extraction (Brin, 1999; ... Gravano, 2000).

2 Analyzing the ... pus

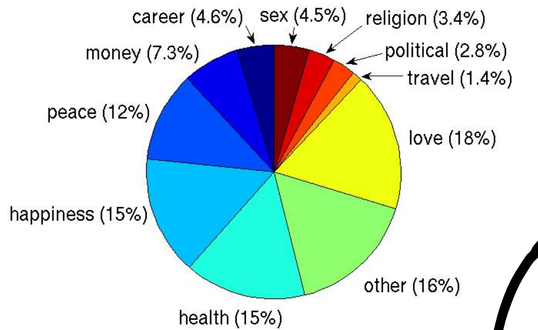
We analyze the WIS... th a variety ... istical methods. Ou... ot only rev... people wished for on New Year's Eve ... pro... ectors in Se...

The complete ... contains nearly 100,000 wishes colle... period of 10 days in December 2007, ... in English, with the remainder in Portug... , Chinese, French, and other language... aper, we consider only the 89,574 En... Most of these English wishes conta... graphic meta data provided by the wish... g a variety of countries (not limited to ... eaking) around the world. We perform ... eprocessing, including TreeBank-style ... n, downcasing, and punctuation remova... is treated as a single entity, regardles... t contains multiple sentences. After p... he average length of a wish is 8 token...

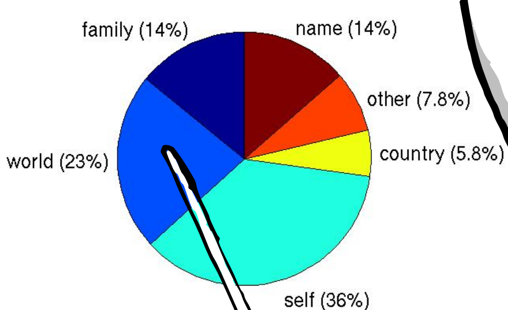
2.1 The Topic and S... Wishes

As a first step in under... ng the content of the wishes, we asked five ... notators to manually annotate a random subsample of 5,000 wishes. Sections 2.1 and 2.2 report results on this subsample.

The wishes w... notated in terms of two at-



(a) Topic of Wishes

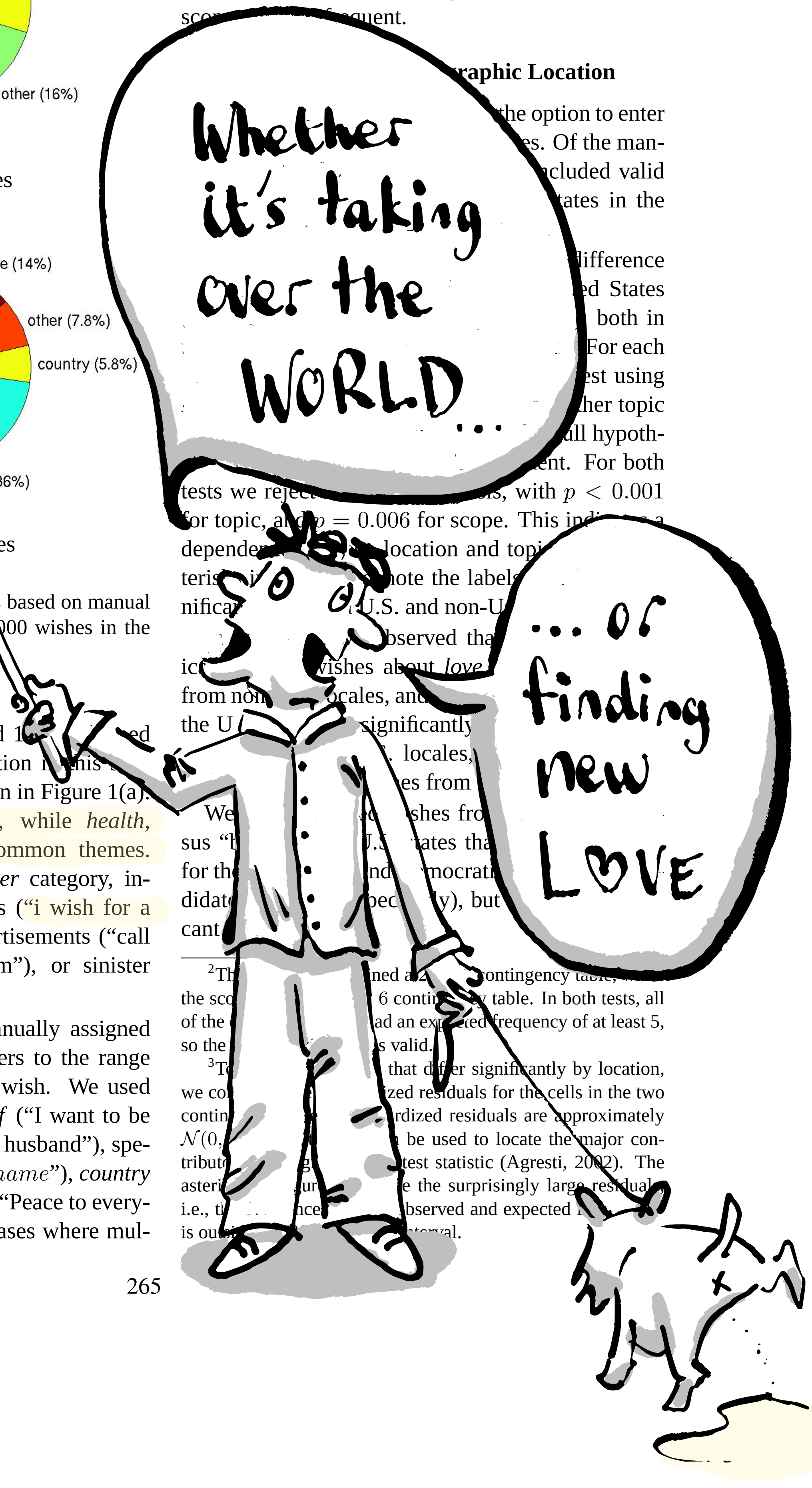


(b) Scope of Wishes

Figure 1: Topic and scope distributions based on manual annotations of a random sample of 5,000 wishes in the WISH corpus.

tributes: topic and scope. We used 1… …ed topic categories, and their distribution in this sample of the WISH corpus is shown in Figure 1(a). The most frequent topic is *love*, while *health*, *happiness*, and *peace* are also common themes. Many wishes also fell into an *other* category, including specific individual requests ("i wish for a new puppy"), solicitations or advertisements ("call me 555-1234", "visit *website*.com"), or sinister thoughts ("to take over the world").

The 5,000 wishes were also manually assigned a scope. The scope of a wish refers to the range of people that are targeted by the wish. We used 6 pre-defined scope categories: *self* ("I want to be happy"), *family* ("For a cure for my husband"), specific person by *name* ("Prayers for *name*"), *country* ("Bring our troops home!"), *world* ("Peace to everyone in the world"), and *other*. In cases where multiple scope labels applied, the broadest scope was selected. Figure 1(b) shows the scope distribution. It is bimodal: over one third of the wishes are narrowly directed at one's self, while broad wishes at the world level are also frequent. The in-between sco… …frequent.



…raphic Location

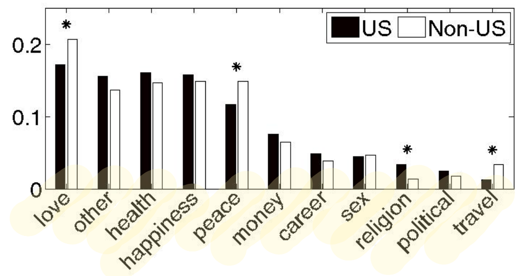
…the option to enter …es. Of the man- …cluded valid …tates in the …

…difference …ed States … both in … For each …est using …her topic …ull hypoth- …ent. For both tests we reje… …s, with $p < 0.001$ for topic, and $p = 0.006$ for scope. This ind… a dependen… location and topi… teris… …note the labels nifica… …U.S. and non-U… …bserved tha… ic… wishes about *love* … from no… …ocales, and… the U… significantly … locales, … …es from …

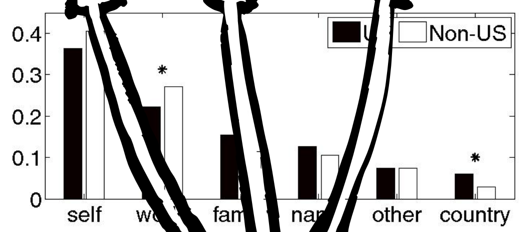
We… …shes fro… sus "… …U.S. …ates tha… for the … …nd… …mocrati… didat… …pec… …ly), but cant…

[2]Th… …ned a 2… contingency … the sc… 6 conti… …y table. In both tests, all of the … …ad an expected frequency of at least 5, so the … …s valid.

[3]T… … that differ significantly by location, we co… …ized residuals for the cells in the two contin… …rdized residuals are approximately $\mathcal{N}(0,$… … be used to locate the major contribut… …test statistic (Agresti, 2002). The asteri… …ur… …e the surprisingly large residuals, i.e., t… …nce… …bserved and expected … is ou… …terval.



(a) Wish topi s differ by Lo ation



(b) Wish ... y Location

Figure 2: C... distributi... tagged... diffe...

2

W e n a o on wis Schu quency to its ra plots rank reveals an app suggesting that w shows that low-occurre... hence learning might be hindered by data sparseness.

WHATEVER

IT

IS ...

2.4 Latent Topic Modeling for Wishes

The 11 topics in Section 2.1 were manually pre-defined based on domain knowledge. In contrast, in this section we applied Latent Dirichlet Allocation (LDA) (Blei et al., 2003) to identify the latent topics in the full set of 89,574 English wishes in an

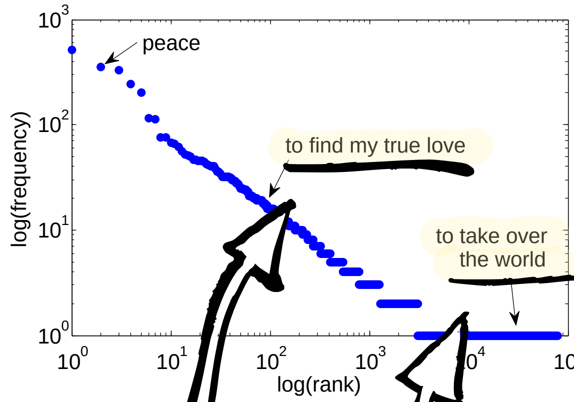


Figure 3: The ra vs. frequency plot f wishes, approximately obeyin pf's law. Note the og-log scale.

unsuper fashion. The al is to validate and comp t the study in S on 2.1.

y LDA to the es, we treated each indi- wish as a short ument. We used 12 topics, llapsed Gibbs S ling (Griffiths and Steyvers, 2004) for infere hyperparameters $\alpha = 0.5$ and 0.1 a Markov Chain Monte Carlo for

g 12 LDA topics are shown in Ta- form of the highest probability words) in each topic. We manually added criptors for readability. With LDA, it is to observe which words were assigned ch wish. For example, LDA as- words in th "world(8) peace(8) ends(4) in iraq(1) to come(1) home(1)" ics: peace and troops (topic numbers in es). Interestingly, these LDA topics largely th the pre-defined topics in Section 2.1.

lding Wish Detectors

We now st novel NLP task of wish detection, i.e., classifying in dual sentences as being wishes or not. Importantly, we ant our approach to transfer to domains other than ew Year's wishes, including consumer product re d online political discussions. It should be poin t that wishes are highly domain dependent. For example, "I wish for world peace" is a common wish on New Year's Eve, but is exceedingly rare in product reviews; and vice versa: "I want to have instant access to the volume" may occur in product reviews, but is an un-

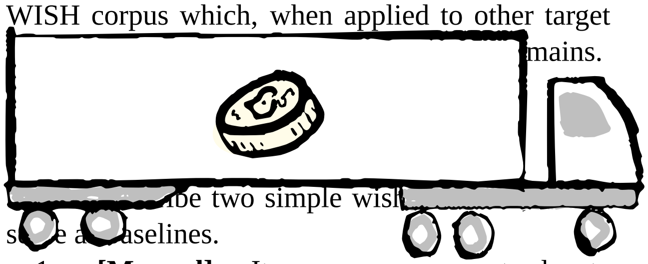
Topic	Summary	Top words in the topic, sorted by $p(\text{word}\|\text{topic})$
0	New Year	
1	Troops	
2	Election	
3	Life	
4	Prosperity	
5	Love	
6	Career	
7	Lottery	
8	Peace	
9	Religion	
10	Family	
11	Health	

MAY IT NOT GET STUCK AT THE PORT OF DOVER*!

Table 2: Wish topics learned from Latent Dirichlet Allocation. Words are sorted by $p(\text{word}|\text{topic})$.

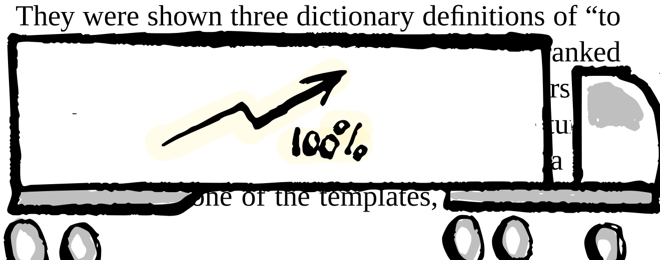
likely New Year's wish. For this initial study, we do assume that there are some labeled training data in the target domains of interest.

To transfer the knowledge learned from the out-of-domain WISH corpus to other domains, our key insight is the following: while the content of wishes (e.g., "world peace") may not transfer across domains, the ways wishes are expressed (e.g., "I wish for ___") may. We call these expressions *wish templates*. Our novel contribution is an unsupervised method for discovering candidate templates from the WISH corpus which, when applied to other target ... mains.



... be two simple wish ... s...e a...aselines.

1. **[Manual]**: It may seem easy to locate wishes. Perhaps looking for sentences containing the phrases "i wish," "i hope," or some other simple patterns is sufficient for identifying the vast majority of wishes in a domain. To test this hypothesis, we asked two native English speakers (not the annotators, nor affiliated with the project; no exposure to any of the wish datasets) to come up with text patterns that might be used to express wishes. They were shown three dictionary definitions of "to ... anked ... rs ... u ... a



... ne of the templates,

i wish ___
i hope ___
i want ___
hopefully ___
if only ___
don't believe ___ did ___
___ do want
i can has ___

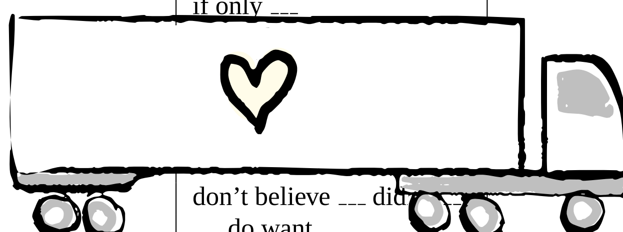
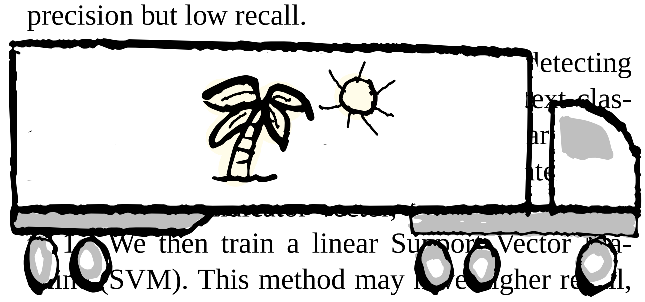


Table 3: Manual templates for identifying wishes.

classified as a wish. By varying the depth of the list, one can produce different precision/recall behaviors. Overall, we expect [Manual] to have relatively high precision but low recall.

... etecting ... ext clas- ... ar ... te



... ndicator vector, ... (1... We then train a linear Su... Vector ...a-... in (SVM). This method may ... igher re...l, but precision may suffer. For instance, the sentence "Her wish was carried out by her husband" is not a wish, but could be misclassified as one because of the word "wish."

Note that neither of the two baseline methods uses the WISH corpus.

* obligatory BREXIT reference

3.2 Automatically Discovering Wish Templates

We now present our method to automatically discover high quality wish templates using the WISH corpus. The key idea is to exploit redundancy in how the same wish content is expressed. For example, as we see in Table 1, both "world peace" and "i wish for world peace" are common wishes. Similarly, both "health and happiness" and "i wish for health and happiness" appear in the WISH corpus. It is thus reasonable to speculate that "i wish for ___" is a good wish template. Less obvious templates can be discovered in this way, too, such as "let there be ___" from "peace on earth" and "let there be peace on earth."

We formalize this intuition as a bipartite graph, illustrated in Figure 4. Let $W = \{w_1, \ldots, w_n\}$ be the set of unique wishes in the WISH corpus. The bipartite graph has two types of nodes: content nodes C and template nodes T, and they are generated as follows. If a wish w_j (e.g., "i wish for world peace") contains another wish w_i (e.g., "world peace"), we create a content node $c_1 = w_i$ and a template node t_1 ="i wish for ___". We denote this relationship by $w_j = c_1 + t_1$. Note the order of c_1 and t_1 is insignificant, as how the two combine is determined by the underscore in t_1, and $w_j = t_1 + c_1$ is just fine. In addition, we place a directed edge from c_1 to t_1 with edge weight count(w_j), the frequency of wish w_j in the WISH corpus. Then, a template node appears to be a good one if many heavy edges point to it.

On the other hand, a template is less desirable if it is part of a content node. For example, when w_j ="health and happiness" and w_i ="health", we create the template t_2 ="___ and happiness" and the content node $c_3 = w_i$. If there is another wish w_k ="i wish for health and happiness", then there will be a content $c_2 = w_j$. The template t_2 thus contains so words (since it matches c_2), and may ll in a new domain. We capture th s: if $\exists c' \in C$, and $\exists$ string s (s not or W) such that $c' = s + t$, we add om t to c' with edge weigh count

Based on such c ler s, we devised the following scheme for

(1)

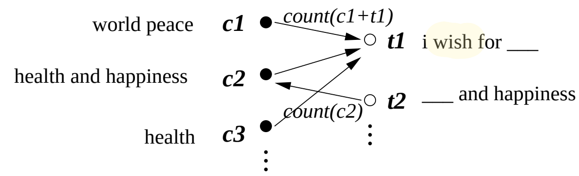


Figure 4: The bipartite graph to create templates.

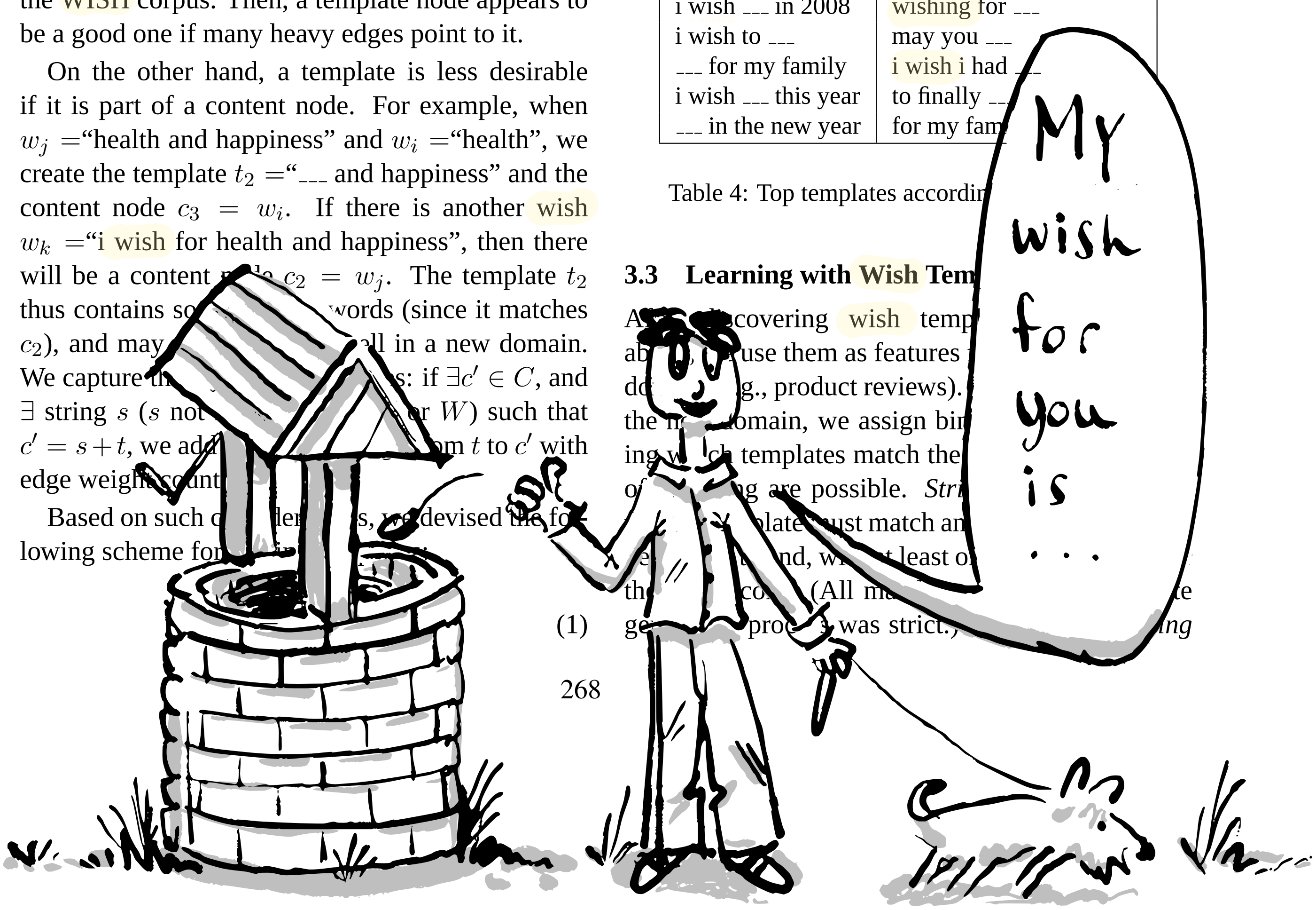
where in(t) is the in-degree of node t, defined as the sum of edge weights coming into t; out(t) is the out-degree of node t, defined similarly. In other words, a template receives a high score if it is "used" by many frequent wishes but does not match many frequent content-only wishes. To create the final set of template features, we apply the threshold score$(t) \geq 5$. This produces a final list of 811 templates. Table 4 lists some of the top templates ranked by score(t). While some of these templates still contain time- or scope-related words ("for my family"), they are devoid of specific topical content. Notice that we have automatically identified several of the manually derived templates in Table 3, and introduce many new variations that a learning algorithm can leverage.

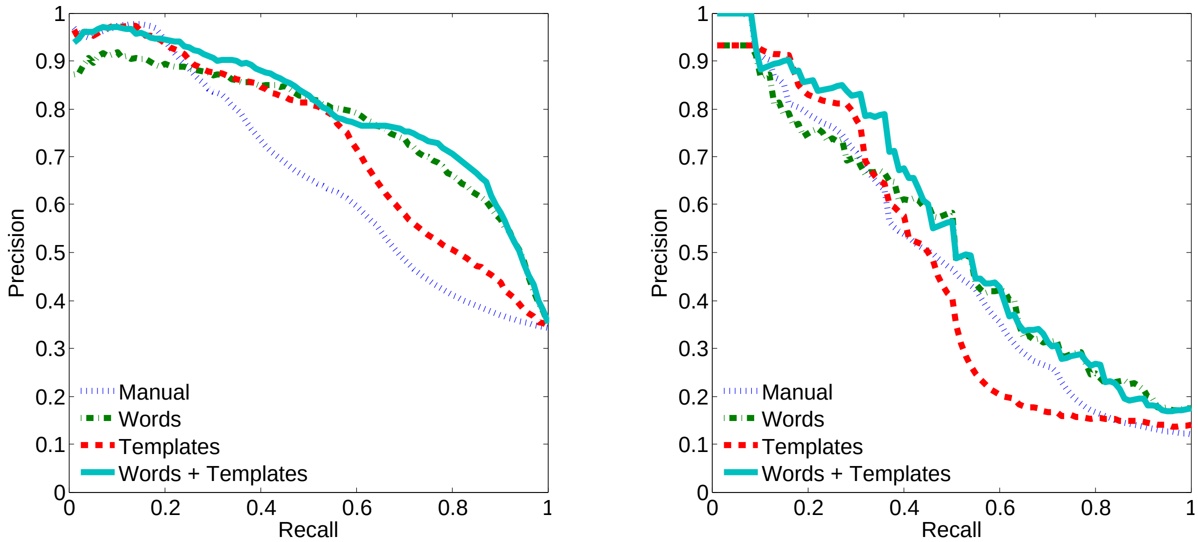
Top 10	Others in Top 200
___ in 2008	i want to ___
i wish for ___	___ for everyone
i wish ___	i hope ___
i want ___	my wish is ___
___ this year	please ___
i wish ___ in 2008	wishing for ___
i wish to ___	may you ___
___ for my family	i wish i had ___
i wish ___ this year	to finally ___
___ in the new year	for my fam

Table 4: Top templates accordin

3.3 Learning with Wish Tem

A covering wish temp ab use them as features do g., product reviews). the domain, we assign bin ing w ch templates match the of ng are possible. *Stri* plate must match an nd, w t least o th co (All m te ge proc was strict.) *ing*





THAT YOUR COUNTRY'S
COVID GRAPHS START
LOOKING LIKE THIS!

a

4

xpense
cision) al po added at recall 1.0, corresponding to applying an empty template that matches all sentences. For the SVM-based methods, we vary the threshold applied to the real-valued margin prediction to produce the curves. All curves are interpolated, and AUC measures are computed, using the techniques of (Davis and Goadrich, 2006).

4.1 Target Domains and Experimental Setup

We experimented with two domains, manually labeled at the sentence-level as wishes or non-wishes.[4] Example wishes are listed in Table 6.

Products. Consumer product reviews: 1,235 sentences selected from a collection of amazon.com and cnet.com reviews (Hu and Liu, 2004; Ding et al., 2008). 12% of the sentences are labeled as wishes.

Politics. Political discussion board postings: 6,379 sentences selected from politics.com (Mullen and Malouf, 2008). 34% are labeled as wishes.

We automatically split the corpora into sentences using MxTerminator (Reynar and Ratnaparkhi, 1997). As preprocessing before learning, we tokenized the text in the Penn TreeBank style, down-

[4]These wish-annotated corpora are available for download at http://pages.cs.wisc.edu/~goldberg/wish_data.

4.2 Results

Figure 5 shows the precision-recall curves for the Politics corpus. All curves are averages over 10 folds (i.e., for each of 100 evenly spaced, interpolated recall points, the 10 precision values are averaged). As expected, [Manual] can be very precise with low recall—only the very top few templates achieve high precision and pick out a small number of wishes with "i wish" and "i hope." As we introduce more templates to cover more true wishes, precision drops off quickly. [Templates] is similar,

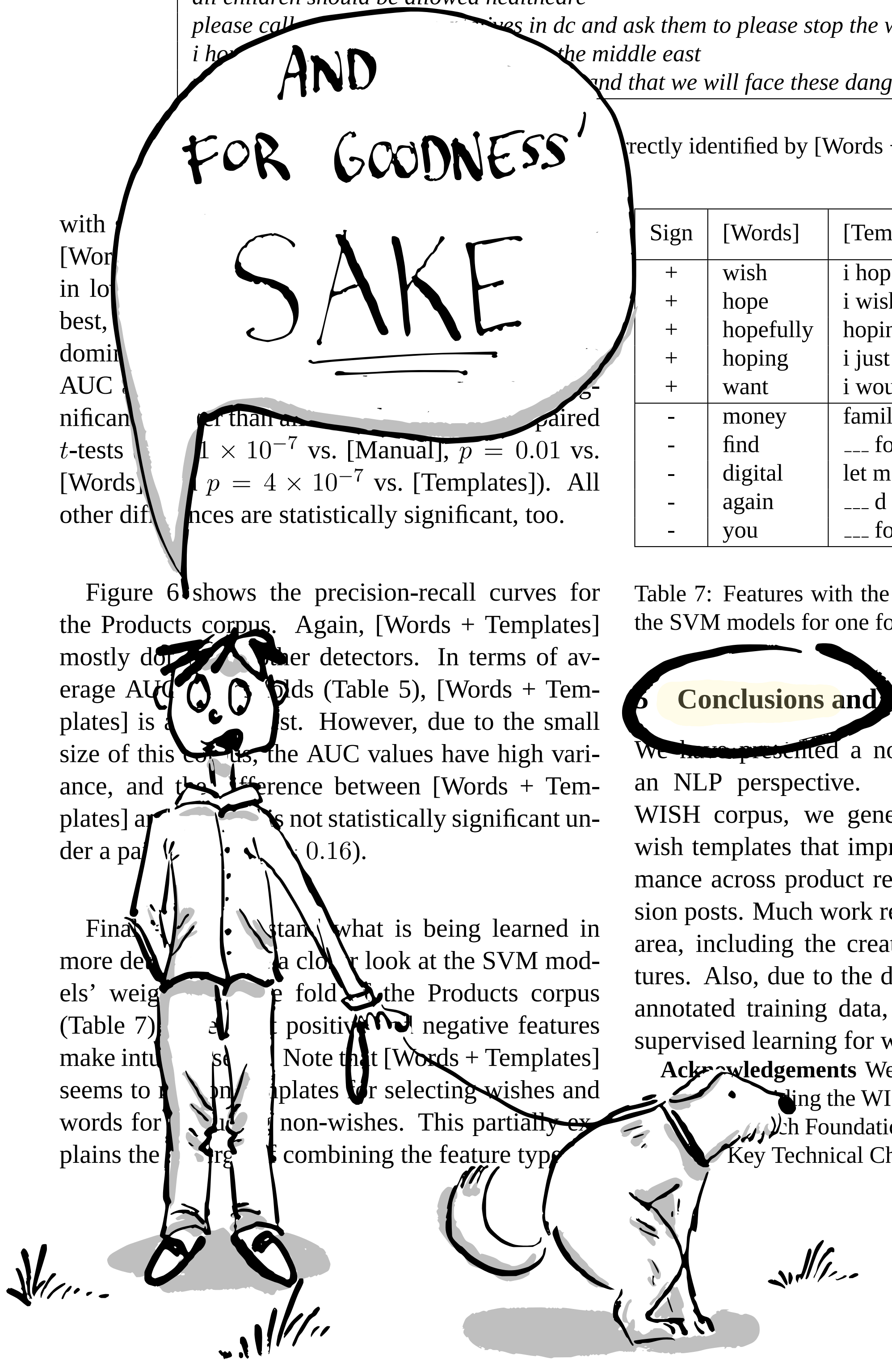
Corpus	[Manual]	[Words]	[Templates]	[Words + Templates]
Politics	0.67 ± 0.03	0.77 ± 0.03	0.73 ± 0.03	0.80 ± 0.03
Products	0.49 ± 0.13	0.52 ± 0.16	0.47 ± 0.16	0.56 ± 0.16

Table 5: AUC results (10-fold averages $\pm$ one standard deviation).

Products:
the only area i wish apple had improved upon would be the screen
i just want music to eminate from it when i want how i want
the dial on the original zen was perfect and i wish it was on this model
i would like album order for my live albums and was just wondering

Politics:
all children should be allowed healthcare
please call ... ives in dc and ask them to please stop the waste in iraq
i ho... the middle east
...nd that we will face these dangers in the future

...rectly identified by [Words + Templates].



with ... [Wor... in lo... best, ... domi... AUC ... g- nifican... er than al... paired t-tests ... 1×10^{-7} vs. [Manual], $p = 0.01$ vs. [Words], ... $p = 4 \times 10^{-7}$ vs. [Templates]). All other dif... nces are statistically significant, too.

Figure 6 shows the precision-recall curves for the Products corpus. Again, [Words + Templates] mostly do... her detectors. In terms of average AU... olds (Table 5), [Words + Templates] is a... st. However, due to the small size of this co... us, the AUC values have high variance, and th... ifference between [Words + Templates] a... s not statistically significant under a pa... 0.16).

Fina... stan... what is being learned in more de... a clo... r look at the SVM models' weig... e fold ... the Products corpus (Table 7)... positiv... negative features make intu... se... Note that [Words + Templates] seems to ... plates for selecting wishes and words for ... non-wishes. This partially explains the ... combining the feature typ...

Sign	[Words]	[Templates]	[Words + Templates]
+	wish	i hope ___	hoping ___
+	hope	i wish ___	i hope ___
+	hopefully	hoping ___	i just want ___
+	hoping	i just want ___	i wish ___
+	want	i would like ___	i would like ___
-	money	family ___	micro
-	find	___ forever	about
-	digital	let me ___	fix
-	again	___ d	digital
-	you	___ for my dad	you

Table 7: Features with the largest magnitude weights in the SVM models for one fold of the Products corpus.

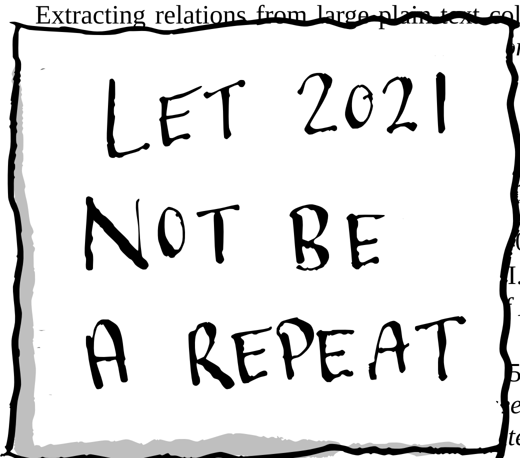
6 Conclusions and Future Work

We have presented a novel study of wishes from an NLP perspective. Using the first-of-its-kind WISH corpus, we generated domain-independent wish templates that improve wish detection performance across product reviews and political discussion posts. Much work remains in this new research area, including the creation of more types of features. Also, due to the difficulty in obtaining wish-annotated training data, we plan to explore semi-supervised learning for wish detection.

Acknowledgements We thank the Times Square Al... ...ing the WISH corpus, and the Wisconsin ...ch Foundation. AG is supported in part by ... Key Technical Challenges Grant.

References

Eugene Agichtein and Luis Gravano. 2000. Snowball: Extracting relations from large plain-text collections. ...nal Con-



... Wiley-

...003. Au- ...hor gen- ...01–412.

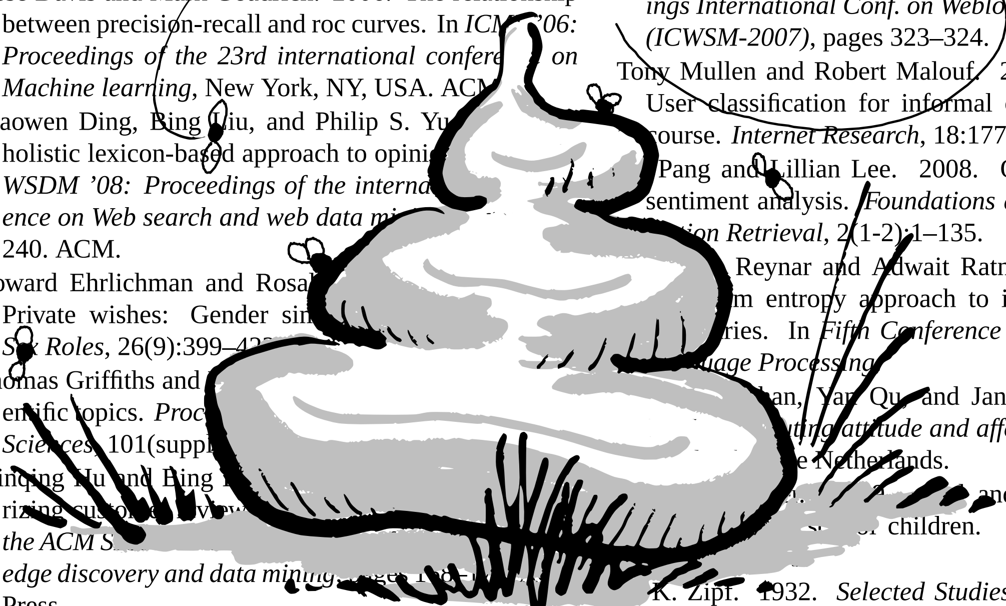
...I. Jordan. ...*Machine*

...5. Genre ...*edings of* ...*telligence (AAAI-05), Poster paper*.

Sergey Brin. 1999. Extracting patterns and relations from the world wide web. In *WebDB '98: Selected papers from the International Workshop on The World Wide Web and Databases*, pages 172–183. Springer-Verlag.

Jesse Davis and Mark Goadrich. 2006. The relationship between precision-recall and roc curves. In *ICML '06: Proceedings of the 23rd international conference on Machine learning*, New York, NY, USA. ACM...

Xiaowen Ding, Bing Liu, and Philip S. Yu... holistic lexicon-based approach to opini... *WSDM '08: Proceedings of the interna... ence on Web search and web data mi...* 240. ACM.



Howard Ehrlichman and Rosal... Private wishes: Gender sim... *Sex Roles*, 26(9):399–42...

Thomas Griffiths and ... entific topics. *Proc... Sciences*, 101(suppl...

Minqing Hu and Bing ... rizing customer review... *the ACM S... edge discovery and data mining*... Press.

Thorsten Joachims. 1999. Making large-scale svm learning practical. In B. Schölkopf, C. Burges, and A. Smola, editors, *Advances in Kernel Methods - Support Vector Learning*. MIT Press.

Laura A. King and Sheri J. Broyles. 1997. Wishes, gender, personality, and well-being. *Journal of Personality*, 65(1):49–76.

Moshe Koppel and Itai Shtrimberg. 2004. Good news or bad news? let the market decide. In *AAAI Spring Symposium on Exploring Attitude and Affect in Text*, pages 86–88.

Saisuresh Krishnakumaran and Xiaojin Zhu. 2007. Hunting elusive metaphors using lexical resources. In *Proceedings of the Workshop on Computational Approaches to Figurative Language*, pages 13–20, Rochester, New York, April. Association for Computational Linguistics.

Christopher D. Manning and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. The MIT Press, Cambridge, Massachusetts.

Rada Mihalcea and Hugo Liu. 2006. A corpus-based approach to finding happiness. In *Proceedings of AAAI-CAAW-06, the Spring Symposia on Computational Approaches to Analyzing Weblogs*.

Rada Mihalcea and Carlo Strapparava. 2005. Making computers laugh: Investigations in automatic humor recognition. In *Empirical Methods in Natural Language Processing*.

Norman A. Milgram and Wolfgang W. Riedel. 1969. Developmental and experiential factors in making wishes. *Child Development*, 40(3):763–771.

Gilad Mishne, Krisztian Balog, Maarten de Rijke, and Breyten Ernsting. 2007. Moodviews: Tracking and searching mood-annotated blog posts. In *Proceedings International Conf. on Weblogs and Social Media (ICWSM-2007)*, pages 323–324.

Tony Mullen and Robert Malouf. 2008. Taking sides: User classification for informal online political discourse. *Internet Research*, 18:177–190.

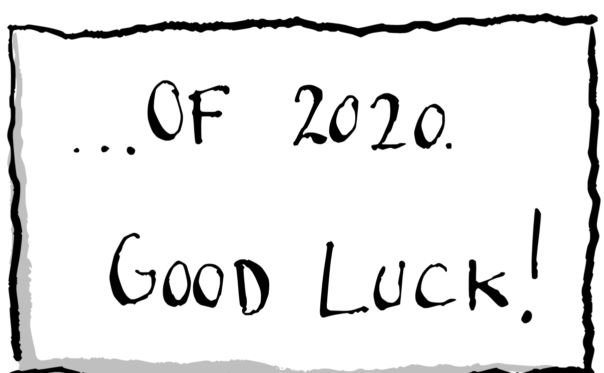
...Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Infor...ation Retrieval*, 2(1-2):1–135.

... Reynar and Adwait Ratnaparkhi. 1997. A ...m entropy approach to identifying sentence ...ries. In *Fifth Conference on Applied Natural ...uage Processing*.

...han, Yan Qu, and Janyce Wiebe, editors. ...*ing attitude and affect in text*. Springer, ...e Netherlands.

... and written wishes of ... children. *Child Development*, ...

...K. Zipf. 1932. *Selected Studies of the Principle of Relative Frequency in Language*. Harvard University Press.



2021-01-09 juanuys.com