Personality profiling from text: Introducing Part-of-speech N-grams

William R. Wright and David N. Chin
{wrightwr, chin}@hawaii.edu

University of Hawaiʻi at Mānoa
Information and Computer Sciences Department
POST Building #317
1680 East-West Road
Honolulu, HI 96822

Abstract. A support vector machine is trained to classify the Five Factor personality of writers of free text. Writers are classified for each of the five personality dimensions as high/low with the mean personality score for each dimension used for the dividing point. Writers are also separately classified as high/medium/low with division points at one standard deviation above and below mean. The two-class average accuracy using 5-fold cross validation of 80.6% is much better than the baseline (pick most likely class) accuracy of 50%, but the 3-class accuracy is only slightly better (7.4%) than baseline because most writers fall into the medium class due to the normal distribution of personality values. Features include bag of words, essay length, word sentiment, negation count and part-of-speech n-grams. POS n-grams have not previously been used for personality prediction, so their consistently positive contribution (averaging 4.8% and 5.8% for the 2/3 class cases) is analyzed in detail. The information gain for the most predictive features for each of the five personality dimensions are presented and discussed.

Keywords: personality, classifier, part-of-speech n-grams, information gain, support vector machine

1 Introduction

The ability to profile user personality, particularly by inferring stable personality traits given an author's available writings promises a variety of useful applications, such as in organizational management, marketing, education, and social media. Some successful demonstrations have already been conducted [18], [11]. Since personality consists of stable differences in individual behavior, accurate predictions about someone's personality can be can be used to predict their preferences and future behavior with sufficient accuracy to be very useful. For example, the personality trait of openness may be predictive of receptiveness of users to user model based adaptations.

The personality predictions are performed by training models using various machine learning algorithms, such as Naive Bayes and support vector machines (SVMs), with features extracted from text written by a subject. Besides bag-of-words features, which are highly predictive throughout a homogenous sample but not very generalizable between populations, researchers have identified some additional features that are useful for personality inference. Word sentiment is strongly correlated to the personality of a writer [5], [9], as are quantity of words and punctuation.

Despite the strength of the correlations, they are neither of sufficient magnitude to explain significant differences between individuals' behavior along the various personality dimensions, nor to predict personality with as much accuracy as could be desired. Thus the search for additional features continues. Although the field acknowledges that features with linguistic basis hold promise, little has been done identify additional features of that nature. Meanwhile a rich set of natural language processing techniques are available for this purpose.

The search must focus on features that actually improve the accuracy of classifiers. Blindly exploring huge feature sets has its limitations: feature selection is a computationally intractable problem whose solutions become more difficult to approximate as one considers larger feature sets. Some theory to guide the search may illuminate appropriate features amongst the sea of features, and even suggest entirely new, untapped feature classes.

One appealing avenue to explore is that of part-of-speech (POS) n-grams: they describe deeply embedded grammatical structures that seem unlikely to vary for a given author [1]. One exploratory study [7] attempted to employ Dutch language part-of-speech n-grams to predict Myers-Briggs Type Indicator (MBTI) types. Despite its other successes, the study achieved only 51% precision when employing the n-grams. Rather than any deficiency in the choice of features, their lack of precision may be due to their use of MBTI, a commercial testing tool.

To take the effort forward, we train SVM classifiers on some texts to predict the personalities of authors. As much as possible, we replicate the features used in [13], part of a project to identify human threats within corporate networks. We offer the following contributions: we describe how to build a successful personality classifier, including feature extraction and selection. We introduce the use of POS n-grams to predict authors' Five Factor personality. Finally we report our results to inform others attempting to build a personality predictor.

2 Personality and text

Personality traits are consistent patterns in a person's behavior over time—behavior that is significant and varies between individuals. Below is a list of the five personality traits enumerated by the prevailing model of human personality (the Five Factor Model). Factors I and II are considered mainly interpersonal dimensions; they describe modes of interaction with others. The factors are in approximate ascending order of the degree to which they account for individual differences. (Descriptions below adapted from [4, 6].)

Factor I: ***Extraversion***. The Extravert approaches the world with energy, enthusiasm, lack of inhibition and a sense of adventure, especially when it comes to social engagement.

Factor II: ***Agreeableness***. The second interpersonal dimension, Agreeableness is what it sounds like, and a person can be high on Extraversion yet low on Agreeableness or any of the other dimensions.

Factor III: ***Conscientiousness***. This factor describes effectiveness in performing prescribed rote, repeated activities. Also assiduous following of rules. However assessments on this trait do not generally ask questions about a person's ethics [2].

Factor IV: ***Neuroticism***. Also called Emotional Stability, reversing the measure. Neurotic individuals tend to perceive events negatively, and to be very sensitive to such events, to lack confidence, and apt to cease action in the face of difficulty or to refuse action in anticipation of obstacles.

Factor V: ***Openness***. This dimension is related to qualities that nurture the mind, i.e. openness to considering unfamiliar ideas and participating in new experiences.

Clearly a very telling aspect of a person's behavior is verbal. The advent of computer technology, particularly digital storage and retrieval of text allows us to examine this aspect of behavior. Simple word frequency (bag-of-words counts) along with overall stem and word counts comprise some of the most intuitive and common features extracted from text. When relevant, such as in e-mail exchanges, speech acts may predict personality (e.g. the disagreeable person is apt to repeat demands without offering a variety of other speech acts) [13], and punctuation and word sentiment certainly do [18].

PD (Propositional Idea density)

3 Method

To create a predictor of personality from author text, one extracts predictive features from texts written by authors of known personalities. The known personality scores (as determined by self-report personality assessment questionnaires or human observer reports) then function as labels for supervised learning. For our data feature selection matters—skipping that step leads to failed or inferior models, notwithstanding occasional reports to the contrary from others working with different data sets.

When training for the binary classification task, one divides participants into two or more sets (discretization) for each personality dimension. Although regression of personality score is possible, perhaps even more useful for many applications, researchers tend to report their accuracy on the binary or 3-class classification task. In sequence, the typical steps, as also observed in this study, are:

- Collect text corpora from participants.
- Collect personality questionnaire scores.
- Extract predictive features from each author text.
- Perform feature selection.
- Train and test a binary classifier: Binary classification, Low/High classes on each personality dimension. Techniques include SVM, Naive Bayes classifier, NN, decision tree, etc.

A wide variety of text features have been shown to be predictive of writer personality, falling in the following categories:

- Word sentiment
- Speech acts (such as, negotiate, greet, deliver, remind)
- Punctuation (repetition, smileys)
- Bag of words
- Part of speech (POS) unigrams, n-grams

We extracted all of the above, excluding speech acts, which are more relevant for e-mail messages or other text intended as direct communication between participants in an activity.

4 Experimental details

Our participants are 2,588 university students in North America who each wrote freely for 20 minutes in English. If a writer stopped writing, the computer would stop the clock until typing resumed. The essays span 2005 through 2008, and the average of the essay word counts is 787. Each student also took the Five Factor Inventory, a personality questionnaire [8]. To preserve anonymity, the essays and personality scores are assigned ID numbers in place of participants' names.

4.1 Features extracted

We extracted features commonly used for personality prediction, and as well as POS n-grams. We tokenize each participant essay and extract the features of interest; some features are extracted using pre-existing tools, while others involve tools that we implemented for use in these experiments. Finally a script outputs the feature data into a format readable by WEKA [15], a suite of implemented machine learning algorithms.

Bag of words. After tokenizing each essay, we extracted the bag-of-words features, counting the frequency of words appearing in a list of the 20,000 most common words from TV and movie scripts, while excluding common names per the latest available U.S. Census data. These measures help us preserve the general usefulness of the features by denying inappropriate emphasis to unusual words or to particular people mentioned in the essays.

Vocabulary size. We counted tokens and stems. For our stemmer we chose the Porter stemming algorithm, which is well specified, fast, and thoroughly tested. We implemented Martin Porter's latest revision of his specifications, which he has constantly updated on his website [10], and tested our code on the list of 30,000 test words offered there. Then we counted the stems present in each essay (disregarding duplicate appearance of words with a common stem such as "run" and "running"). This count of unique stems reflects the variety of vocabulary the author employed in an essay.

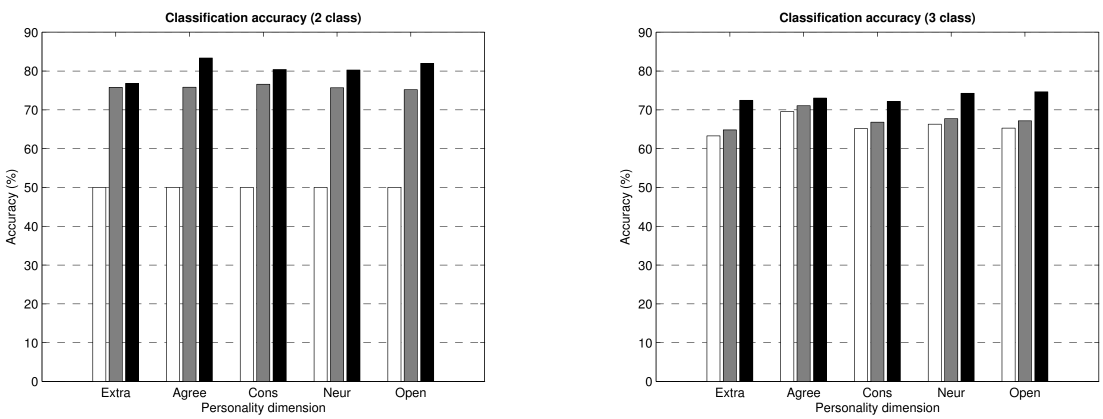
Word sentiment. We employed the sentiment polarity lexicon from UPitt [17] [14], counting positive/negative words, and word subjectivity (high/low).

Negations. We implemented our own count of negations including contractions (not, never, can't, won't, etc.)

Part of speech (POS) unigrams, n-grams. Used a POS tagger, an implementation of that discussed in [16]. Among the POS features is the count of pronouns that the previous study employed (pronouns are positively correlated with the neuroticism score). We also computed the POS n-grams.

4.2 Results of classification and discussion

Although Big Five personality assessments do not label people into binary classes (e.g. "Extraverted" vs. "Introverted"), binary classification of participants falling in High or Low ranges on each dimension is useful in decision making applications. Given the features we extracted and the known personality classes, we trained SVM classifiers to distinguish between authors falling in the High/Low (2-class) and in the High/Medium/Low (3-class) ranges on each personality dimension. To determine whether POS n-grams improve prediction accuracy, we trained with and without POS n-grams. Training with 5-fold cross validation was conducted with LibLINEAR (an SVM with no kernel popular for text mining) linked with the Weka platform.



(a) (b)

Figure 1: Classification accuracy: 2 class and 3 class.

The two-class bins of equal size (High and Low for each dimension) are the result of dividing participants according to their personality scores: above or below the median score. The additional 3-class partitioning follows the method of [13], dividing participants into three sets (High, Medium and Low) whose bounds are the scores falling one standard deviation above or below the mean personality score for each personality dimension.

As is evident in Fig. 1 the presence of POS n-grams indeed improved the accuracy of personality prediction (Table 1 provides the accuracy outcomes). Following [12], the significances of the improvements were calculated for a single 5-fold cross validation with the Binomial test. All improvements were statistically significant at $p < 0.0001$ except for the 2-class Extraversion ($p = 0.186$) and the 3-class Agreeableness ($p = 0.767$). Also as suggested by [3], we ran 50 replications of 5-fold cross validation with each repetition consisting of different randomly-selected membership for the 5 folds, then applied the paired Student t-test separately to each 5-fold cross validation, averaged the t-values and converted this to a significance value. These significance values agree with the previous Binomial test. Table 2 summarizes all the significance scores.

We repeat the use of these two tests to assess the utility of our classifiers over the baseline. Baseline predictions consist of assuming all participants fall in the majority class, which is arbitrary in the 2-class cases (the classes are of equal size) and is the Medium class in the 3-class cases. In both 2-class and 3-class cases for all five personality traits, the improvement over the baseline in terms of accuracy of the SVM classifiers both with and without n-gram features were statistically significant to $p < 0.0001$ in all cases as calculated for a single 5-fold cross validation.

Given this way of dividing participants into 3 classes, the Low and High classes represent participants falling at the extremes of the personality dimensions. In the case of Agreeableness, the available n-gram features are useful for predicting in which of the dichotomous 2-classes (High vs. Low, i.e. Extravert vs. Introvert) participants fall. However, the features available to us do not help much in predicting which participants fall in the outer extremes: the 3-class results show no significant improvement in classifier accuracy when we conduct training with POS n-gram features included. The opposite was true with Extraversion. POS n-gram features were significantly helpful in the 3-class case, but not in the 2-class case. A plot of the errors vs. personality scores shows that the Extraversion error distribution did not change much when POS n-gram features were added. On the other hand, the Agreeableness error distribution showed much higher error rates in the Medium Agreeableness region when POS n-grams were used. These errors negated the gains in the High and Low regions, which collectively are smaller than the Medium region.

	(a)	(b)	(c)	(c)-(a)	(c)-(b)	(d)	(e)	(f)	(f)-(d)	(f)-(e)
Personality Dimension	2-class baseline	2 class without POS n-grams $\pm$*stdev*	2-class with POS n-grams $\pm$*stdev*	Improvement over baseline	Improvement due to POS n-grams	3-class baseline	3-class without n-grams $\pm$*stdev*	3-class with n-grams $\pm$*stdev*	Improvement over baseline	Improvement due to POS n-grams
Extraversion	50.00	75.81±0.37	76.82±1.94	26.82	1.01	63.29	64.83±0.17	72.47±0.39	9.18	7.64
Agreeableness	50.00	75.83±0.41	83.36±0.65	33.36	7.53	69.55	71.07±0.14	73.04±1.47	3.49	1.98
Conscientiousness	50.00	76.60±0.36	80.41±0.86	30.41	3.81	65.15	66.82±0.12	72.20±0.41	7.06	5.39
Neuroticism	50.00	75.69±0.59	80.28±0.71	30.28	4.59	66.31	67.71±0.08	74.28±0.24	7.98	6.57
Openness	50.00	75.20±0.21	81.99±0.39	31.99	6.79	65.30	67.18±0.17	74.66±0.29	9.36	7.48
Average	50.00	75.83	80.57	30.57	4.75	65.92	67.52	73.33	7.41	5.81

Table 1: Classification accuracy (percentage).

An important goal of this project was to find new predictive features that may be generalizable to other populations. Among the features of highest predictive power (shown in Table A.1), several are already well known to indicate personality, for example the use of “me” and words with negative sentiment are positively correlated with higher neuroticism scores. The remaining features will

	With vs. without POS n-grams				Without POS n-grams vs. baseline*			
	2-class		3-class		2-class		3-class	
Personality Dimension	Binary test	Bradford-Brodley†	Binary test	Bradford-Brodley†	Binary test	Bradford-Brodley†	Binary test	Bradford-Brodley†
Extraversion	0.1860	0.4356	<0.0001	0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Agreeableness	<0.0001	0.0003	0.7670	0.2762	<0.0001	<0.0001	<0.0001	<0.0001
Conscienciousness	<0.0001	0.0052	<0.0001	0.0005	<0.0001	<0.0001	<0.0001	<0.0001
Neuroticism	<0.0001	0.0044	<0.0001	0.0004	<0.0001	<0.0001	<0.0001	<0.0001
Openness	<0.0001	0.0006	<0.0001	0.0019	<0.0001	<0.0001	<0.0001	<0.0001

*The "With POS n-grams vs. baseline" case is not shown here; all probabilities are $p < 0.0001$.

†Following [12].

Table 2: Probability of significant improvement in classifier accuracies.

ultimately fall in two categories: those that generalize to other populations, and those that do not. Some features, such as homework, hurricane, Arkansas and evacuate for Conscientiousness, are predictive for our participants due to their unique circumstances in time and geography. Our participants were students living near Hurricane Katrina during and soon after the event. Such features seem unlikely to generalize to populations spanning more varied geography and chronology. Others seem promising, such as "fascinating" and "lesbian", both of which bear positive correlations with Openness scores.

Among POS n-gram features, of interest is the (ADV ADJ to) collocation, wherein adverbs modify adjectives. Table 3, examples 4, 5 and 6 show "so" and "extremely" employed as intensifiers. One explanation may be that gregarious extraverts are apt to pile on words to drive their point home, rather than to deliberate about the choice of perhaps a stronger verb—or simply to speak in less emphatic terms. Features such as these may persist across time and other circumstances, rendering them useful for personality prediction among various populations.

Ex.																
(1)	that's	a	bad	habit	I	need	to	work	on	,	procrastination	.				
				N	PRO											
(2)	SAE's	are	so	damn	arrogant	and	have	nothing	to	be	arrogant	about	.			
											ADJ	PREP	PERIOD			
(3)	two	faces	that	can	be	perceived	as	kissing	or	just	looking	at	you	dead	on	.
														ADJ	PREP	PERIOD
(4)	I	am	so	excited	to	be	done	with	school!							
			ADV	ADJ	to											
(5)	It's	so	funny	to	me	.										
		ADV	ADJ	to												
(6)	After	a	busy	and	funfilled	weekend	it	is	extremely	boring	to	sit	in	front	of	a computer
									ADV	ADJ	to					

Table 3: POS n-gram examples.

5 Conclusion

Since their presence during training significantly improved classification test results results, indeed for this population grammatical structures as represented by POS n-grams are indicative of writer personality. Their consistency over the lifetime of a writer should render them useful for personality prediction among a variety of populations. In addition to the grammatical features, a few bag-of-words features emerge as possibly generalizable.

6 Future work

It may be possible to extend this work by including coarse-grained parts of speech (e.g. noun phrases) extracted by chunking tools. Adjustments in the feature selection methodology may improve the general usefulness of the resulting classifiers, reducing the nuisance of features that are only predictive of the current population or sample. Further testing of promising features may establish their generalizability to a variety of populations. Lastly, compelling explanations of why particular POS n-grams are indicative of personality would be of great interest in directing the exploration of new text features useful for personality prediction.

7 Acknowledgment

We extend grateful thanks to James W. Pennebaker for making this research possible by sharing the essays and personality scores.

A Appendix Tables

	$IG \cdot 10^2$	Category	Feature
Extraversion	0.834	POS n-gram	N PRO
	0.670	Bag-of-words	tonight
	0.596	Bag-of-words	we
	0.595	Bag-of-words	yay
	0.575	POS n-gram	MOD ADV N
	0.572	Bag-of-words	out
	0.552	Bag-of-words	have
	0.551	Bag-of-words	all
	0.551	POS n-gram	ADV ADJ to
	0.526	POS n-gram	PREP PRO
	0.525	POS n-gram	WADV PROP
	0.511	POS n-gram	VBZ PRO N
	0.504	Bag-of-words	am
	0.494	POS n-gram	VBG to
	0.482	POS n-gram	N PRO VBP
Agreeableness	0.669	POS	NS
	0.601	POS n-gram	ADJ PREP PERIOD
	0.591	Bag-of-words	family
	0.585	Bag-of-words	mom
	0.567	POS n-gram	PREP PRO CC
	0.563	POS n-gram	VBP ADV ADJ
	0.525	Bag-of-words	me
	0.437	POS n-gram	WDT VBZ WADV
	0.428	POS n-gram	ADJ PERIOD PRO
	0.428	POS n-gram	PREP NP WADV
	0.426	Bag-of-words	billions
	0.426	POS n-gram	VBD ADJR PRO
	0.426	Bag-of-words	italian
	0.412	POS n-gram	PROP NS
	0.410	Bag-of-words	patient
Conscientiousness	1.014	Bag-of-words	hurricane
	0.744	POS n-gram	PROP N
	0.693	Bag-of-words	mom
	0.625	Bag-of-words	go
	0.617	POS n-gram	ADJ to
	0.602	Bag-of-words	goodness
	0.601	Bag-of-words	week
	0.582	POS n-gram	RP ADV PRO
Conscientiousness (cont.)	0.582	Word sentiment	weaksubj
	0.570	Bag-of-words	excited
	0.509	Bag-of-words	homework
	0.491	Bag-of-words	arkansas
	0.489	Bag-of-words	nice
	0.483	Bag-of-words	evacuate
	0.466	POS n-gram	ADJ to VB
Neuroticism	0.658	Word sentiment	negative
	0.582	POS n-gram	N CC WADV
	0.544	Bag-of-words	and
	0.523	POS n-gram	VBP ADJ
	0.508	POS n-gram	COMMA NS VBP
	0.485	Bag-of-words	feel
	0.484	Bag-of-words	things
	0.468	Bag-of-words	game
	0.465	POS n-gram	PARENS PREP ADJ
	0.464	Bag-of-words	im
	0.464	POS n-gram	PRO VBP COMMA
	0.463	POS n-gram	PRO VBP ADJ
	0.455	Bag-of-words	saturday
	0.454	Bag-of-words	me
	0.450	Bag-of-words	everything
Openness	0.486	POS n-gram	ADV VB NS
	0.465	POS n-gram	CC PREP ADV
	0.465	Bag-of-words	lesbian
	0.443	POS n-gram	VBN PROP NS
	0.438	Bag-of-words	please
	0.435	POS n-gram	PERIOD CC VBP
	0.435	POS n-gram	VBP COMMA WADV
	0.427	POS n-gram	to N
	0.426	POS n-gram	PROP ADJS CC
	0.426	POS n-gram	N VBD PREP
	0.412	POS n-gram	VB COMMA WADV
	0.408	Bag-of-words	fascinating
	0.398	POS n-gram	VBD CD NS
	0.387	Bag-of-words	determine
	0.387	Bag-of-words	cursed

Note: An index defining each POS tag is available online: http://www.williamwright.info/downloads/pos_tags.pdf

Table A.1: The 15 most predictive features for each personality dimension. IG is Information Gain.

References

1. Shlomo Argamon, Casey Whitelaw, Paul Chase, Sobhan Raj Hota, Navendu Garg, and Shlomo Levitan. Stylistic text classification using functional lexical features. *Journal of the American Society for Information Science and Technology*, 58(6):802–822, 2007.
2. J. Block. The five-factor framing of personality and beyond: Some ruminations. *Psychological Inquiry*, 21(1):2–25, 2010.
3. Jeffrey P Bradford and Carla E Brodley. The effect of instance-space partition on significance. *Machine Learning*, 42(3):269–286, 2001.
4. P.T. Costa and R.R. McCrae. Neo pi-r professional manual. *Odessa, FL: Psychological Assessment Resources*, 396:653–65, 1992.
5. J. Golbeck, C. Robles, M. Edmondson, and K. Turner. Predicting personality from twitter. In *Privacy, security, risk and trust (passat), 2011 ieee third international conference on and 2011 ieee third international conference on social computing (socialcom)*, pages 149–156. IEEE, 2011.
6. O.P. John, R.W. Robins, and L.A. Pervin. *Handbook of personality: theory and research*. The Guilford Press, 2008.
7. Kim Luyckx and Walter Daelemans. Using syntactic features to predict author personality from text. In *Proceedings of Digital Humanities 2008 (DH 2008)*, pages 146–149, 2008.
8. Robert R McCrae, Paul T Costa, Jr, and Thomas A Martin. The neo–pi–3: A more readable revised neo personality inventory. *Journal of personality assessment*, 84(3):261–270, 2005.
9. J.W. Pennebaker and L.A. King. Linguistic styles: language use as an individual difference. *Journal of personality and social psychology*, 77(6):1296, 1999.
10. Martin Porter. The English (Porter2) stemming algorithm. http://snowball.tartarus.org/algorithms/english/stemmer.html, 2006. [Online; accessed 2-March-2013].
11. A. Roshchina, J. Cardiff, and P. Rosso. User profile construction in the twin personality-based recommender system. *Sentiment Analysis where AI meets Psychology (SAAIP)*, page 73, 2011.
12. Steven L. Salzberg and Usama Fayyad. On comparing classifiers: Pitfalls to avoid and a recommended approach. *Data Mining and Knowledge Discovery*, pages 317–328, 1997.
13. Jianqiang Shen, Oliver Brdiczka, and Juan Liu. Understanding email writers: Personality prediction from email messages. In *User Modeling, Adaptation, and Personalization*, pages 318–330. Springer, 2013.
14. Online source mpqa.cs.pitt.edu. Retrieved, October 2013.
15. Online source www.cs.waikato.ac.nz/ml/weka. Retrieved, October 2013.
16. Kristina Toutanova, Dan Klein, Christopher D Manning, and Yoram Singer. Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*, pages 173–180. Association for Computational Linguistics, 2003.
17. Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 347–354. Association for Computational Linguistics, 2005.
18. William Wright. Literature Review. http://www.williamwright.info/downloads/litreview.pdf, 2012. [Online; accessed 2-March-2013].