Synchronic LSC with Economics: Engel & Marx vs Smith

Jade Sillère
MSc&T Data and Economics for Public Policy
École Polytechnique, ENSAE & Telecom
jade.sillere@polytechnique.edu

Abstract

Building on the existing literature on Synchronic Lexical Semantic Change in the context of Economics - looking at whether the difference between contextualised semantic representations of the same key economics terms within the text of Adam Smith - a famous classical economist with ideas considered mainstream - are more similar than when compared with representations of the same terms in the corpus of Friedrich Engels & Karl Marx - a corpus recognized as the manifesto of the heterodox economic school of thought of marxist economics.

1 ML for NLP mini-project accompanying paper

1.1 Introduction

The domain of Lexical Semantic Change (LSC) is concerned with looking at how the meaning of a word evolves over time or between domains or both. The two key types are:

- Diachronic Lexical Semantic Change (Diachronic LSC)
  - Diachronic Lexical Semantic Change refers to looking at changes in the meaning of a word over time. This field gained popularity in NLP around 2016–2017 with key research produced by Fremann and Lapata (2016), Hamilton et al. (2016), and Schlechtweg et al. (2017).
- Synchronic Lexical Semantic Change (Synchronic LSC)
  - Synchronic Lexical Semantic Variation explores how word meanings differ across domains, genres, or communities of speakers.

This paper and experiment lie within the domain of Synchronic LSC. Research in Synchronic LSC originates back to the domain of word sense disambiguation looking at term extraction from domain-specific corpora and classification problems for Word Sense Disambiguation (WSD) (Maynard and Ananiadou [1998]; Chen and Al-Mubaid [2006]; Taghipour and Ng [2015]; Daille et al. [2016]). A notable related research paper is Ferrari et al. [2017] which uses cosine similarity between word vector embeddings to detect semantic shifts in identical words between domains. These studies all address how specific contexts can alter the meaning of a word — for instance, how the German noun Schnee typically means 'snow,' but can refer to 'beaten egg whites' in a recipe book. However, an emerging domain is exploring differences in meaning of identical words as well as differences in meaning over time (crossing over with Diachronic LSC), as measured by the semantic similarity of their embeddings, depending on political stances to predict political stances and track polarization (Azarbonyad et al. [2017];RODRIGUEZ et al. [2023]).

Through the mini-project and its associate paper, I will be seeking to build upon the paper titled *One Word, Two Sides: Traces of Stance in Contextualized Word Representations*. The paper looks at

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

determining whether vector representations reflect a higher similarity in word usage within a common stance than between different stances, specifically "whether contextual models are able to capture a difference between the representation of a word when it is used by people who are in favour vs. against a certain target" (Garí Soler et al. [2022b]).

1.2 Reference paper

The paper *One Word, Two Sides: Traces of Stance in Contextualized Word Representations* investigates whether contextualized word embeddings, reflect differences in how the same word is used by speakers with opposing stances on a topic closely related to the word - e.g., comparing a representation of "woman" from sentences in favour of the Feminist movement and within their corpus and comparing it to representations of those who are opposed (Garí Soler et al. [2022a]).

In the paper, they use stance-labelled datasets like SemEval2016, Covid19, P-stance, and ArgQ, to test whether the contextualised word embeddings produced by the contextual models explored are able to capture differences in meaning of the same word used by authors with a favourable view of the subject compared to those with a unfavourable view. To do this, they look at whether word representations are more similar within the same stance than between differing stances. Their methodology includes testing various contextual embedding models' (BERT, Context2Vec and À la carte) effectiveness on sense-annotated corpora to detect subtle semantic change, that is, whether they are successful in predicting a smaller between-corpus semantic similarity than within-corpus.

The results show that BERT's 768d *bert-base-uncased* model, especially embeddings from its 10th layer, consistently captures stance-based differences in word usage. While the differences are small, they are statistically significant and more pronounced for central, topical words (typically nouns). These findings suggest that contextual embeddings do encode traces of stance, even with limited data, and could be useful for applications like stance detection, disagreement identification, and dialogue analysis. The authors propose extending this approach to conversational contexts to measure conceptual alignment between speakers.

This paper is very similar to the literature described above which could be applied to understand or predict ideological or political stance of individuals, whereas this project seeks to look at an application of these models to understand whether differences in opposing economic school of thoughts seep into their use of economic key terms.

1.3 Experiment description

This project aims to look at whether differences of representations of economics terms within the Wealth of Nations by Adam Smith and within the Communist Manifesto by Karl Marx are significantly smaller than those than between texts of the aforementioned economists with opposing economic schools of thought. Therefore, for simplicity, I will be assuming that over the time elapsed between the publications of the two papers, these terms have not changed significantly in their meaning, such that I am not capturing LSC across time.

Additionally, contrary to the dominant literature in differences in semantic representation aimed at detecting or predicting political stances, I am focusing on exploring whether differences in economic school of thought change the meaning of key economic terms e.g., labour or capital. Therefore, contrary to the research paper looking at words which are quite strongly related to the ideological or political stance e.g., "woman" with pro-Feminists or anti-Feminists, here it is not so obvious here whether the school of thought would result in a significantly different contextualised meaning of the economist terms between economists than in their own corpus, which is what I seek to explore.

Similarly to the reference paper for this project, I will be making use of contextualised representations generated with the 768d *bert-based-uncased model* which was shown to be the best performing model when tested on labelled databases of texts with opposing stances. I will be applying this model to texts of economists deemed to have an irreconcilable economic school of thought. Additionally, similarly to the paper and other related literature, the measure for similarity of representations used will be the cosine similarity. Furthermore, I will also be using TF-IDF to find the most and least relevant economic key terms present in all corpus. Finally, although the default option in extracting embeddings from BERT is to use the last hidden layer of the trained model, this isn't always the best layer to use, and similarly to the paper, I will explore using different intermediate layers.

2 Descriptive analysis: TF-IDF

Similarly to what was done in the reference paper, to extract the key economic ideas which were central to each text, I used TF-IDF (Term Frequency–Inverse Document Frequency) to identify and rank the 200 most important words in The Wealth of Nations and The Communist Manifesto based on their relevance within each text. The process begins by lemmatizing (i.e., reducing words to their simplest forms) and cleaning both texts by removing common stop words, numbers, and punctuation, and normalizing white space. The "term frequency" (TF) part captures how often a word appears in a single document, while the "inverse document frequency" (IDF) penalizes terms that are common across both documents. This means that words unique or particularly significant in one document—but not in the other—will receive higher TF-IDF scores. As a result, the top-scoring words represent terms that are not only frequent but also central to each text and helps to filter out words, other than stop words, that are not informative of the text. From this process, I was able to extract the following list of unambiguously economic terms[1] which were among the 200 highest-scoring words according to TF-IDF in both texts, to ensure the words were mentioned frequently in both texts:

- Produce
- Trade
- Labour
- Labourer [2]
- Land
- Capital
- Market
- Manufacture
- Industry
- Work
- Government
- Private
- Wage
- Demand
- Exchange
- Subsistence

3 Calculating In-text and Between-text Semantic Similarities

I've used BERT's pre-trained model to extract contextual embeddings from BERT for all the instances of each of the words in the list above for both corpora. I then compute the in-text pairwise semantic similarities using the cosine distance for both texts and finally, the between-text semantic similarities of embeddings.

Before feeding both texts into BERT, I preprocess them. To do this, I load the full text and split it into paragraphs, which helps to preserve the structure of the text and help make local context more meaningful. I then tokenize the text into manageable windows based on BERT's input length constraints (as BERT can only handle up to 512 tokens at a time). Finally, I make use of stride, which is an overlap between consecutive windows, when a paragraph is too long and needs to be split into smaller windows (due to the token constraints), a generous overlap of 10 is used between consecutive windows to avoid losing information located at chunk boundaries and to maximize the chance that the model sees target words with enough context.

[1]For example, words which have a dual meaning were excluded. For example, good can either mean a product or a positive adjective.

[2]Given BERT uses subword tokenization and instead considers this word as labour + another token, this word is excluded from the list below.

For each target word, I identify all instances based on BERT's subword tokenization and extract the contextual embeddings using a fixed-size window of 5 tokens before and after the word. This window size was chosen to capture enough context while maintaining computational efficiency. Once I collect the contextualized embeddings for all instances of a word, I compute pairwise cosine similarity to assess how consistently the word is used semantically throughout the text.

I repeat this in-text pairwise semantic similarity extraction process for each word in The Communist Manifesto. Then, I perform a cross-text analysis by computing pairwise cosine similarity between instances of the same word across both texts, to evaluate how consistently each word is used semantically between them.

Finally, I use an Independent t-test along with a Mann-Whitney U test to test for whether the semantic similarities in-text for each corpus belong to the same population as those between text, that is, whether the semantic meaning of each economic term is significantly different within the text compared to between both texts. The independent t-test assumes normality and equal variances whereas the Mann-Whitney U test is more flexible and has no distributional assumptions. I use both methods to ensure robustness of the results to distributional assumptions. Additionally, the many tests conducted result in a higher chance of detecting a significant results, therefore, I apply a Bonferroni and FDR correction to p-values. These methods account for multiple comparisons and reduce the likelihood of Type I errors. Given I am conducting t-tests on 15 words for two bodies of text, and I am conducting two types of tests, this means I am essentially conducting 60 statistical tests. Therefore, these methods will account for this.

Finally, I repeat all the above steps using the penultimate layer in BERT instead of the final layer. The reason for doing so is that using the penultimate (before-last) layer of a neural network, rather than the final layer, is often preferred when extracting general-purpose features or performing analyses beyond the model's original task. This is because the final layer is typically tailored to the specific output task and therefore, discards much of the intermediate information the model has learned. In contrast, the penultimate layer captures high-level semantic or structural patterns without being biased by the model's decision-making process.

4 Results: Comparing In-text and Between-text Semantic Similarities

Overall, the results indicate significant differences in semantic similarity between within-text and between-text contexts for the majority of economic terms analyzed. This suggests the meaning behind seemingly neutral economics terms are not so neutral when used by economists with very different schools of thought. Additionally, both statistical tests agree, suggesting these results are more robust.

4.1 Results: Wealth of Nations

Note that for the following section, unless specified, the significance level considered is 5%. As is clear from the tables below, in The Wealth of Nations, both the independent t-tests and Mann–Whitney U-tests reveal consistently strong statistical significance ($p < 0.001$ after Bonferroni and FDR correction) for nearly all terms. This suggests that semantic similarity is meaningfully higher within the text than across corpora for almost all terms. Contextually, this implies the way in which economic terms are employed in the Wealth of Nations does not appear comparable to the way they are used in the Communist Manifesto. Only the term "wage" showed no significant difference, indicating its usage may be more consistent with that of the Communist Manifesto.

Table 1: T-statistics, P-values, Bonferroni-Adjusted and FDR-adjusted P-values for Independent t-test Within-text vs Between-text Semantic Similarity of Embeddings from Wealth of Nations.

Word	T-stat	P-value	Bonferroni P-value	FDR P-value
produce	154.17330000	0.00000000	0.00000000	0.00000000
trade	194.63120000	0.00000000	0.00000000	0.00000000
labour	150.78310000	0.00000000	0.00000000	0.00000000
land	68.23590000	0.00000000	0.00000000	0.00000000
capital	222.97320000	0.00000000	0.00000000	0.00000000
market	62.41780000	0.00000000	0.00000000	0.00000000
manufacture	35.19470000	0.00000000	0.00000000	0.00000000
industry	145.81850000	0.00000000	0.00000000	0.00000000
work	76.85590000	0.00000000	0.00000000	0.00000000
government	74.20750000	0.00000000	0.00000000	0.00000000
private	43.21820000	0.00000000	0.00000000	0.00000000
wage	0.01720000	0.98781184	1.00000000	0.98781184
demand	24.62020000	0.00000000	0.00000000	0.00000000
exchange	21.84490000	0.00000000	0.00000000	0.00000000
subsistence	6.28570000	0.00000000	0.00000000	0.00000000

Table 2: U-statistics, P-values, Bonferroni-Adjusted and FDR-adjusted P-values for Mann–Whitney U-tests Within-text vs Between-text Semantic Similarity of Embeddings from Wealth of Nations.

Word	U-stat	P-value	Bonferroni P-value	FDR P-value
produce	2627204890.00000000	0.00000000	0.00000000	0.00000000
trade	7561147942.00000000	0.00000000	0.00000000	0.00000000
labour	32212094685.00000000	0.00000000	0.00000000	0.00000000
land	3141089493.50000000	0.00000000	0.00000000	0.00000000
capital	6495454303.00000000	0.00000000	0.00000000	0.00000000
market	716732611.00000000	0.00000000	0.00000000	0.00000000
manufacture	9707907.00000000	0.00000000	0.00000000	0.00000000
industry	2362048810.00000000	0.00000000	0.00000000	0.00000000
work	1224402877.50000000	0.00000000	0.00000000	0.00000000
government	511484700.00000000	0.00000000	0.00000000	0.00000000
private	126710196.00000000	0.00000000	0.00000000	0.00000000
wage	78.00000000	0.82308591	1.00000000	0.82308591
demand	48983975.50000000	0.00000000	0.00000000	0.00000000
exchange	44806872.50000000	0.00000000	0.00000000	0.00000000
subsistence	20599472.50000000	0.00000000	0.00000000	0.00000000

4.2 Results: A Communist Manifesto

Similarly, in The Communist Manifesto, most terms—particularly "labour", "capital", "industry", and "government" exhibited highly significant differences, confirming that the document maintains a consistent meaning of key economic concepts within-text which differs significantly to that used when compared to the Wealth of Nations. A few terms such as "demand" and "trade" (for the latter, this is not the case for the FDR-adjusted p-value of the Independent T-test) showed non-significant differences, which could imply more variable or generalized usage of these terms.

Table 3: T-statistics, P-values, Bonferroni-Adjusted and FDR-adjusted P-values (60 tests) for Independent t-test Within-text vs Between-text Semantic Similarity of Embeddings from A Communist Manifesto.

Word	T-stat	P-value	Bonferroni P-value	FDR P-value
produce	3.78990000	0.00114451	0.01716765	0.00143064
trade	2.81930000	0.00544596	0.08168940	0.00583496
labour	9.79990000	0.00000000	0.00000000	0.00000000
land	3.06100000	0.00238647	0.03579705	0.00275362
capital	47.16270000	0.00000000	0.00000000	0.00000000
market	6.09640000	0.00000023	0.00000345	0.00000035
manufacture	7.48150000	0.00000000	0.00000000	0.00000000
industry	51.68140000	0.00000000	0.00000000	0.00000000
work	-5.07780000	0.00000048	0.00000720	0.00000065
government	17.49990000	0.00000000	0.00000000	0.00000000
private	41.28330000	0.00000000	0.00000000	0.00000000
wage	8.09160000	0.00000000	0.00000000	0.00000000
demand	1.06130000	0.30098189	1.00000000	0.30098189
exchange	6.73360000	0.00000000	0.00000000	0.00000000
subsistence	11.85840000	0.00000000	0.00000000	0.00000000

Table 4: U-statistics, P-values, Bonferroni-Adjusted and FDR-adjusted P-values for Mann–Whitney U-tests Within-text vs Between-text Semantic Similarity of Embeddings from A Communist Manifesto.

Word	U-stat	P-value	Bonferroni P-value	FDR P-value
produce	99607.00000000	0.00066500	0.00997500	0.00076731
trade	1448453.00000000	0.11762789	1.00000000	0.12602988
labour	99953712.50000000	0.00000000	0.00000000	0.00000000
land	3531682.50000000	0.00000063	0.00000945	0.00000079
capital	46578262.50000000	0.00000000	0.00000000	0.00000000
market	196815.00000000	0.00000001	0.00000015	0.00000002
manufacture	77505.00000000	0.00000000	0.00000000	0.00000000
industry	108169241.50000000	0.00000000	0.00000000	0.00000000
work	5389457.00000000	0.00000014	0.00000210	0.00000019
government	8360798.50000000	0.00000000	0.00000000	0.00000000
private	13250313.50000000	0.00000000	0.00000000	0.00000000
wage	8016.00000000	0.00000000	0.00000000	0.00000000
demand	21880.00000000	0.53528080	1.00000000	0.53528080
exchange	277183.00000000	0.00000000	0.00000000	0.00000000
subsistence	291989.50000000	0.00000000	0.00000000	0.00000000

These results together support the hypothesis that key economic terms are embedded in distinct semantic contexts within each text, reflecting their economic ideology. Additionally, the few economic terms in each text which seem reconcilable in their use with the other author are not the same between the two texts. This implies that even in the rare cases where there is some overlap in term usage, it does not involve the same terms in both directions. This asymmetry suggests that, among the terms analyzed, there is no such thing as a neutral economic key term—that is, no term whose meaning is universally shared across economic ideologies.

4.3 Results using the Penultimate Layer: Wealth of Nations

Table 5: T-statistics, P-values, Bonferroni-Adjusted and FDR-adjusted P-values for Independent t-test Within-text vs Between-text Semantic Similarity of Embeddings from Wealth of Nations.

Word	T-stat	P-value	Bonferroni P-value	FDR P-value
produce	132.9392	0.00000000	0.00000000	0.00000000
trade	142.4930	0.00000000	0.00000000	0.00000000
labour	111.8835	0.00000000	0.00000000	0.00000000
land	43.0238	0.00000000	0.00000000	0.00000000
capital	117.5366	0.00000000	0.00000000	0.00000000
market	33.7599	0.00000000	0.00000000	0.00000000
manufacture	24.1471	0.00000000	0.00000000	0.00000000
industry	97.6770	0.00000000	0.00000000	0.00000000
work	56.5666	0.00000000	0.00000000	0.00000000
government	61.1348	0.00000000	0.00000000	0.00000000
private	44.7714	0.00000000	0.00000000	0.00000000
wage	0.0155	0.98903855	1.00000000	0.98903855
demand	17.1447	0.00000000	0.00000000	0.00000000
exchange	3.8294	0.00013001	0.00195015	0.00015001
subsistence	3.7976	0.00014811	0.00222165	0.00015869

Table 6: U-statistics, P-values, Bonferroni-Adjusted and FDR-adjusted P-values for Mann–Whitney U-tests Within-text vs Between-text Semantic Similarity of Embeddings from Wealth of Nations.

Word	U-stat	P-value	Bonferroni P-value	FDR P-value
produce	2562424121.00	0.00000000	0.00000000	0.00000000
trade	6888655491.00	0.00000000	0.00000000	0.00000000
labour	29889434032.00	0.00000000	0.00000000	0.00000000
land	2846635268.50	0.00000000	0.00000000	0.00000000
capital	5459039826.50	0.00000000	0.00000000	0.00000000
market	608345930.00	0.00000000	0.00000000	0.00000000
manufacture	8667709.50	0.00000000	0.00000000	0.00000000
industry	2103772711.00	0.00000000	0.00000000	0.00000000
work	1131213585.50	0.00000000	0.00000000	0.00000000
government	485735819.50	0.00000000	0.00000000	0.00000000
private	127041938.50	0.00000000	0.00000000	0.00000000
wage	93.00	0.82308591	1.00000000	0.82308591
demand	45438448.50	0.00000000	0.00000000	0.00000000
exchange	36900535.50	0.21272192	1.00000000	0.22791634
subsistence	19370140.50	0.00000105	0.00001575	0.00000121

4.4 Results using the Penultimate Layer: A Communist Manifesto

As discussed earlier, these results should be interpreted the same way as the section above, the only difference here is the layer used for the embeddings is the penultimate layer of BERT's pre-trained model instead of the final layer.

Table 7: T-statistics, P-values, Bonferroni-Adjusted and FDR-adjusted P-values (60 tests) for Independent t-test Within-text vs Between-text Semantic Similarity of Embeddings from A Communist Manifesto.

Word	T-stat	P-value	Bonferroni P-value	FDR P-value
produce	4.4446	0.00024689	0.00370335	0.00028487
trade	4.2951	0.00003078	0.00046170	0.00004197
labour	9.7747	0.00000000	0.00000000	0.00000000
land	2.9383	0.00353179	0.05297685	0.00378406
capital	45.6676	0.00000000	0.00000000	0.00000000
market	6.4501	0.00000007	0.00000105	0.00000011
manufacture	7.0934	0.00000000	0.00000000	0.00000000
industry	42.1114	0.00000000	0.00000000	0.00000000
work	-3.7112	0.00022136	0.00332040	0.00027670
government	13.0559	0.00000000	0.00000000	0.00000000
private	34.3947	0.00000000	0.00000000	0.00000000
wage	6.2555	0.00000003	0.00000045	0.00000005
demand	2.0151	0.05721642	0.85824630	0.05721642
exchange	6.7124	0.00000000	0.00000000	0.00000000
subsistence	11.3956	0.00000000	0.00000000	0.00000000

Table 8: U-statistics, P-values, Bonferroni-Adjusted and FDR-adjusted P-values for Mann–Whitney U-tests Within-text vs Between-text Semantic Similarity of Embeddings from A Communist Manifesto.

Word	U-stat	P-value	Bonferroni P-value	FDR P-value
produce	104932.00	0.00006092	0.00091380	0.00007029
trade	1542919.00	0.00222772	0.03341580	0.00238684
labour	99281825.00	0.00000000	0.00000000	0.00000000
land	3532333.50	0.00000060	0.00000900	0.00000082
capital	45494096.50	0.00000000	0.00000000	0.00000000
market	197960.00	0.00000001	0.00000015	0.00000002
manufacture	74725.00	0.00000000	0.00000000	0.00000000
industry	102298326.50	0.00000000	0.00000000	0.00000000
work	5557546.00	0.00005545	0.00083175	0.00006931
government	7894068.50	0.00000000	0.00000000	0.00000000
private	12703990.00	0.00000000	0.00000000	0.00000000
wage	7201.00	0.00000007	0.00000105	0.00000011
demand	23923.00	0.15712363	1.00000000	0.15712363
exchange	278702.00	0.00000000	0.00000000	0.00000000
subsistence	287709.50	0.00000000	0.00000000	0.00000000

These results are for the most identical to those obtained using the final layer. The few differences are in the words which have inconsistent results between the t-test and the Man-Whitney U-test in both sets of results.

In these results, I find that in the Wealth of Nations, rather than "wage" being the only word found to be non-significant, I also find "exchange" to be non-significant here, but only in the case of the Mann-Whitney U-test. Similarly, in the first set of results, I found the words "demand" and "trade" (the latter to a lesser extent as the two tests did not agree on this point) to be non-significant, whereas in the second set, I find the words "demand" and "land" (the latter being barely significant in the case of the t-test). Therefore, these results are aligned, with the only results being inconsistent due to their p-values being close to the 5% threshold.

Therefore, to be conservative, I would assume all of these "marginally" significant words to be non-significant. And even in this conservative case, there are very few words which are not found to differ significantly in their meaning between their in-text usage and their between-text usage. For this reason, the overall conclusions have been confirmed when using the penultimate layer instead of the final layer, suggesting these results are robust to the layer selected from which to extract the contextual embeddings.

4.5 Conclusions

These results demonstrate that seemingly neutral economic terms are, in fact, deeply contextual and ideologically inflected. Using BERT's pre-trained contextual semantic embeddings, I have found evidence the way these terms are employed varies systematically depending on the economic ideology expressed in the text. For example, in both The Wealth of Nations and The Communist Manifesto, terms like labour, capital, industry, and government exhibit significantly different patterns of semantic similarity within the texts compared to across texts, with extremely low p-values even after stringent Bonferroni and FDR corrections. This suggests that the meaning of such terms is not fixed or universally understood, even within the discipline, but rather anchored to the economic school of thought and discourse in which they appear.

Moreover, while technical terminology is often assumed to be stable within a given domain, these findings reveal that even within a field, if the field can be further divided into groups with different ideologies or ways of thinking, key terms can take on distinct semantic roles. The stark within-text coherence is particularly evident in The Communist Manifesto where terms like private, wage, and manufacture contrasts strongly with the between-text semantic shifts. This reinforces the idea that technical language is shaped by narrative, purpose, and perspective. Thus, semantic similarity analysis using contextual embeddings offers a powerful lens for uncovering the implicit ideological structures embedded in expert language, beyond the general domain, towards .

References

Hosein Azarbonyad, Mostafa Dehghani, Kaspar Beelen, Alexandra Arkut, Maarten Marx, and Jaap Kamps. Words are malleable: Computing semantic shifts in political and media discourse. *CoRR*, abs/1711.05603, 2017. URL `http://arxiv.org/abs/1711.05603`.

Ping Chen and Hisham Al-Mubaid. Context-based term disambiguation in biomedical literature. In *Proceedings of the 19th International FLAIRS Conference*, Melbourne, Florida, 2006.

Béatrice Daille, Evelyne Jacquey, Gaël Lejeune, Luis Felipe Melo, and Yannick Toussaint. Ambiguity diagnosis for terms in digital humanities. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 4353–4359, Portorož, Slovenia, May 2016. European Language Resources Association (ELRA). URL `https://aclanthology.org/L16-1690/`.

Alessio Ferrari, Beatrice Donati, and Stefania Gnesi. Detecting domain-specific ambiguities: An nlp approach based on wikipedia crawling and word embeddings. In *2017 IEEE 25th International Requirements Engineering Conference Workshops (REW)*, pages 393–399, 2017. doi: 10.1109/REW.2017.20.

Aina Garí Soler, Matthieu Labeau, and Chloé Clavel. One word, two sides: Traces of stance in contextualized word representations. In Nicoletta Calzolari, Chu-Ren Huang, Hansaem Kim, James Pustejovsky, Leo Wanner, Key-Sun Choi, Pum-Mo Ryu, Hsin-Hsi Chen, Lucia Donatelli, Heng Ji, Sadao Kurohashi, Patrizia Paggio, Nianwen Xue, Seokhwan Kim, Younggyun Hahm, Zhong He, Tony Kyungil Lee, Enrico Santus, Francis Bond, and Seung-Hoon Na, editors, *Proceedings of the 29th International Conference on Computational Linguistics*, pages 3950–3959, Gyeongju, Republic of Korea, October 2022a. International Committee on Computational Linguistics. URL `https://aclanthology.org/2022.coling-1.347/`.

Aina Garí Soler, Matthieu Labeau, and Chloé Clavel. One word, two sides: Traces of stance in contextualized word representations. In Nicoletta Calzolari, Chu-Ren Huang, Hansaem Kim, James Pustejovsky, Leo Wanner, Key-Sun Choi, Pum-Mo Ryu, Hsin-Hsi Chen, Lucia Donatelli, Heng Ji, Sadao Kurohashi, Patrizia Paggio, Nianwen Xue, Seokhwan Kim, Younggyun Hahm, Zhong He, Tony Kyungil Lee, Enrico Santus, Francis Bond, and Seung-Hoon Na, editors, *Proceedings of the 29th International Conference on Computational Linguistics*, pages 3950–3959, Gyeongju, Republic of Korea, October 2022b. International Committee on Computational Linguistics. URL `https://aclanthology.org/2022.coling-1.347/`.

Diana Maynard and Sophia Ananiadou. Term sense disambiguation using a domain-specific thesaurus. In *Proceedings of the 1st Computational Linguistics UK Research Colloquium (CLUK1)*, 1998.

PEDRO L. RODRIGUEZ, ARTHUR SPIRLING, and BRANDON M. STEWART. Embedding regression: Models for context-specific description and inference. *American Political Science Review*, 117(4):1255–1274, 2023. doi: 10.1017/S0003055422001228.

Kaveh Taghipour and Hwee Tou Ng. One million sense-tagged instances for word sense disambiguation and induction. In *Proceedings of the Nineteenth Conference on Computational Natural Language Learning*, pages 338–344, Beijing, China, July 2015. Association for Computational Linguistics. doi: 10.18653/v1/K15-1037. URL `https://aclanthology.org/K15-1037/`.