

ANALYSIS OF THE MEANING SHIFT OF THE LOANWORD TECHNOLOGY IN INDONESIAN USING THE WORD EMBEDDING TECHNIQUE

Ahmad Fathoni¹, Hannatul Ma'we², M. Alfani Habib³, Fitriani⁴

^{1 4}Department of Information System, STMIK Syaikh Zainuddin Nahdlatul Wathan,
Anjani, East Lombok, Indonesia

^{2 3}Department of Informatics Engineering, STMIK Syaikh Zainuddin Nahdlatul Wathan,
Anjani, East Lombok, Indonesia

e-mail: fathoni338@gmail.com¹, hannatulmawa@gmail.com², alfanihabib20@gmail.com³,
fitri31121984@gmail.com⁴

Received : 01 September, 2025	Accepted : 07 October 2025	Published : 31 October 2025
-------------------------------	----------------------------	-----------------------------

Abstract

The development of information technology has encouraged the widespread adoption of technology related loanwords in the Indonesian language, particularly in digital communication. The intensive use of these loanwords across diverse contexts has led to semantic shifts in their meanings. This study aims to analyze semantic shifts in technology related loanwords in Indonesian by employing word embedding techniques as a computational linguistic approach. The research adopts a quantitative method based on the analysis of Indonesian digital text corpora representing language use across different time periods. The data are analyzed through text preprocessing, word embedding modeling and the measurement of semantic relationship changes using cosine similarity and nearest-neighbor analysis. The results indicate that several technology-related loanwords have undergone significant semantic shifts, including meaning expansion, changes in semantic associations and shifts in usage functions in line with technological developments and digital culture. These findings demonstrate that word embedding techniques are effective for empirically detecting semantic shifts in Indonesian. This study is expected to contribute methodologically to Indonesian semantic studies and to enrich corpus-based computational linguistics research.

Keywords: semantic shift, technology loanwords, semantics, word embedding, Indonesian language

1. INTRODUCTION

The development of information technology has brought fundamental changes to various aspects of human life, including language practices. Digital technology, particularly the internet, social media and computer-based platforms, has not only transformed the way humans communicate but also influenced the structure, vocabulary and meaning of language used in everyday life. Language, as a dynamic system, constantly adapts to the social, cultural and technological changes that surround it. In the Indonesian language context, the development of information technology has led to the influx of various foreign terms, particularly from English, related to the world of digital technology. Words like application, platform, content, upload, server and streaming are not only used in technical contexts but have also become part of the general vocabulary. The intense use of

these terms in various digital communication contexts has led to gradual and systematic changes in their meanings.

The phenomenon of meaning change, or semantic shift is an important object of study in linguistics, particularly in the branches of historical semantics and social semantics. Meaning shift occurs when the relationship between a language form and the concept it represents changes over time. In the digital era, meaning shift occurs more rapidly than in previous periods because language is widely used in open, interactive and cross-contextual digital spaces. Social media, online forums and other digital platforms are the primary spaces for innovation in meaning and the expansion of word functions.

Several recent studies have shown that language use in digital media significantly influences the dynamics of word meaning. Semantic studies of

Indonesian social media over the past five years have shown that many words have undergone expansions in meaning and changes in semantic associations due to the evolving context of digital communication. These shifts in meaning are not only individual but also form collective patterns observable in large-scale digital text corpora.

However, studies of meaning shifts in Indonesian are still dominated by qualitative-descriptive approaches, such as discourse analysis, lexicography or the study of specific literary and media texts. These approaches provide in-depth conceptual understanding but have limitations in observing meaning changes quantitatively and empirically, especially in the vast and ever-growing volume of digital language data.

With the development of computational linguistics, new approaches have emerged that enable data-driven analysis of word meaning. One approach widely used in international research is word embedding, which represents words in vector form based on the context in which they appear in a text corpus. Recent research has shown that word embedding and contextual language models are capable of detecting temporal changes in word meaning with a high degree of accuracy. This approach allows researchers to systematically observe shifts in word semantic relations across time periods.

However, the use of word embedding techniques to study the shifting meanings of technological loanwords in Indonesian remains relatively limited. Most computational linguistics research in Indonesia still focuses on natural language processing tasks such as text classification, sentiment analysis, or semantic similarity measurement, without directly linking them to semantic studies and meaning change. Thus, there is a research gap in integrating Indonesian semantic studies with digital corpus-based computational linguistic approaches.

Based on these conditions, this research positions itself at the intersection of Indonesian language and literature studies with information technology. This study aims to analyze the shifting meanings of loanwords “technology” in Indonesian using word embedding techniques as a methodological approach. The research focuses on how changes in the context of loanword usage are reflected in changes in semantic relations that can be measured quantitatively.

Theoretically, this research is expected to enrich the study of Indonesian semantics with an empirical approach based on digital corpus data. Methodologically, this research proposes the use of word embedding techniques as a tool for analyzing meaning shifts that can be replicated in other language research. Practically, the findings of this study are expected to contribute to the development

of Indonesian language studies in the digital era, particularly in understanding the dynamics of the meaning of technological loanwords that continue to evolve.

2. RESEARCH METHODOLOGY

Language is a social system that constantly changes as its speakers evolve. In the digital age, information technology plays a role as an external factor accelerating this change, particularly through social media and online communication platforms. [1]Explains that a network society is characterized by the rapid and massive flow of information, thus influencing the way individuals interact and speak.

In the context of Indonesian, several recent studies have shown that digital media contributes to changes in word meaning. [2]Found that language use on Indonesian social media encourages the expansion of meaning and changes in vocabulary connotation due to shifts in the context of digital communication. Similar findings were also reported by [3] who showed that Twitter user interactions form new meaning associations for certain words through repetitive and collective communication practices.

Another study by [4] confirmed that social media platforms like Instagram are a key platform for innovation in meaning, particularly in vocabulary related to technology and digital lifestyles. These findings suggest that information technology influences not only the frequency of word use but also the structure of meaning inherent in those words.

2.1 Technology Loanwords in Indonesian

Loanwords are the result of language contact that occurs due to cultural, economic and technological interactions. In Indonesian, loanwords from English dominate digital technology vocabulary. [5] States that loanwords undergo phonological, morphological and semantic adaptation to suit the recipient language system.

Recent linguistic research shows that technological loanwords tend to change meaning when used by non-specialist speakers. [6] Showed that technological vocabulary in Instagram memes has expanded in meaning and shifted in function from technical terms to means of social expression and digital humor. Furthermore, research by [7] found that the metaphorical use of technological vocabulary on social media leads to significant changes in semantic associations.

[3] Also noted that the term “technology” which was originally technical, experienced a generalization of meaning when used in everyday communication contexts on social media. This indicates that the use of “technology” in Indonesian is dynamic and heavily influenced by the digital social context.

2.2 Semantic Shift in Linguistic Studies



Semantic shift is a change in the relationship between a linguistic form and the concept it represents. [8] Classified semantic shift into several types, such as expansion, narrowing, amelioration and pejoration. In modern linguistic studies, semantic shift is understood as a process influenced by social, cultural and technological factors.

Over the past five years, semantic studies have shown that meaning shifts in the digital era are occurring more rapidly than in previous periods. [2] Demonstrated that the intensity of use of certain vocabulary on social media contributes to the collective formation of new meanings. [4] Also found that meaning changes on Instagram occur systematically through repeated usage patterns within digital communities.

These findings confirm that the shift in meaning in the digital era is not just an individual phenomenon, but a linguistic symptom that can be observed empirically through digital corpus data.

2.3 Word Embedding as a Computational Linguistics Approach

The distributional approach in linguistics states that the meaning of a word can be understood through the context in which it appears. This principle forms the basis for the development of word embedding techniques in computational linguistics. [9] Word2vec is used to determine the length of word dimensions in a sentence without even calculating TF-IDF.

Recent research has shown that word embedding techniques are effective in detecting shifts in word meaning over time. [10] Demonstrated that contextual language models can identify temporal changes in meaning with high accuracy. [11] Also confirmed that embedding-based approaches are the dominant method in semantic shift detection studies.

In the context of Indonesian, research by [12] and [13] shows that word embedding is effective for measuring semantic similarity, but studies specifically utilizing this technique to analyze the shift in meaning of loanwords are still very limited. This suggests an opportunity to develop semantic studies of Indonesian based on computational linguistics.

2.4 Previous Research and Research Position

International research on word embedding-based meaning shift has grown rapidly in the past five years. [14] Pioneered the study of diachronic word embeddings, which was further developed by [10] and [15] using a contextual modeling approach. These studies demonstrate that changes in word vector position can be used as a quantitative indicator of meaning shift.

In Indonesia, research on meaning change is still dominated by qualitative-descriptive approaches. [2], [3] and [6] studied meaning change in the context of social media, but have not utilized a digital corpus-based computational approach.

Meanwhile, research on Indonesian computational linguistics by [16] and [13] still focuses on natural language processing tasks without directly linking it to historical semantics studies.

Based on this review, it can be concluded that there is still a research gap in the study of the shift in meaning of loanwords, "technology," in Indonesian using a word embedding approach. This study aims to fill this gap by combining Indonesian semantics with computational linguistic techniques based on digital corpus data.

2.5 Research Design and Approach

This research employs an interdisciplinary approach, integrating the study of Indonesian language and literature with information technology, particularly in the field of Natural Language Processing (NLP). The approach employed is a mixed qualitative-quantitative approach, with an emphasis on quantitative analysis based on digital corpus data, interpreted qualitatively through a semantic perspective.

2.6 Data Sources and Research Corpus

The research data consists of a corpus of Indonesian-language digital texts containing loanwords from technology. The corpus was compiled from digital sources representing Indonesian language use in online spaces, including:

1. Online media (technological and digital lifestyle articles)
2. Social media (Twitter/X and Instagram)
3. Online forums and discussion platforms

The data span is divided into two temporal periods for the purposes of analyzing meaning shifts:

- Period I: 2015–2018
- Period II: 2020–2024

This division aims to capture the dynamics of changes in the meaning of technological loanwords before and after the intensification of digital media use in Indonesia. The selection of this time span follows the diachronic semantic analysis approach used by [10].

2.7 Data Collection Technique

Data collection was conducted using a web-based corpus collection technique with the following steps:

1. Determining an initial list of technology loanwords based on the Big Indonesian Dictionary and a list of popular technology terms.
2. Collecting digital texts containing these words from online sources using keyword-based crawling techniques.
3. Filtering the data to ensure the texts were in Indonesian and relevant to the technological context and everyday use.

The collected data is then saved in text format (.txt) and grouped by time period to facilitate temporal analysis.

2.8 Data Preprocessing



Before word embedding modeling is performed, the text data undergoes preprocessing to improve the quality of the analysis. These stages include:

1. Text normalization, such as lowercasing and removing non-alphabetic characters.
2. Tokenization, which separates the text into word units.
3. Stopword removal, which removes functional words that do not contribute to lexical meaning.
4. Stemming, which restores words to their base form using an Indonesian stemming algorithm.

This stage follows standard practice in Indonesian NLP as suggested by [13].

2.8 Word Embedding Modeling

Word representation modeling was performed using the Word2Vec technique with the Skip-gram architecture, as this model is effective in capturing semantic relationships between words based on their context of occurrence[9]. The word embedding model was trained separately for each time period to allow for comparison of meaning across time.

The main parameters used in the modeling include:

- Vector dimension: 200
- Context window size: 5
- Minimum word frequency: 10

After the training process, vector alignment was performed between periods to ensure that comparisons of word vector positions could be performed validly, following the method proposed by [14].

2.9 Meaning Shift Analysis Technique

The analysis of meaning shift was conducted by measuring the change in the position of word vectors between periods using the cosine distance metric. The greater the change in vector distance, the greater the indication of meaning shift.

The results of these quantitative measurements were then analyzed qualitatively by:

1. Identifying words that experienced significant vector changes.
2. Analyzing the context in which the words appeared in each period.
3. Interpreting the types of meaning shifts based on semantic theory, such as meaning expansion and connotation changes [8].

This approach allows integration between computational results and contextual linguistic interpretation.

2.10 Data Validity and Validity

To ensure the validity of the research, method triangulation was conducted by combining quantitative analysis based on word embedding and qualitative analysis based on semantic analysis. Furthermore, the reliability of the model was tested by comparing the findings with previous research

and manual evaluation of the context of word usage in the corpus.

This approach is expected to produce findings that are not only technically accurate, but also linguistically and culturally meaningful.

3. RESULT AND DISCUSSION

3.1 Identification of Technology Loanwords in the Corpus

Based on the results of digital corpus processing, this study identified several technology loanwords with high and stable frequency of occurrence across both observation periods (2015–2018 and 2020–2024). These words include viral, online, download, application, platform, account, content and streaming.

These results align with the findings of [3] who stated that technological vocabulary has become part of the everyday lexicon of Indonesians due to the intensity of digital media use. However, unlike previous descriptive research, this study goes further by analyzing changes in meaning quantitatively through word vector representation.

3.2 Interperiod Word Embedding Modeling Results

Word embedding modeling produced word vector representations for each period. Cosine distance analysis showed that several technology loanwords experienced significant changes in vector position, indicating a shift in meaning.

For example, the word "viral" showed a significant increase in cosine distance between the two periods. In the initial period (2015–2018), the word "viral" appeared more frequently in medical or biological contexts. However, in the 2020–2024 period, its context shifted to social media, content popularity, and digital phenomena. This change indicates semantic broadening, as categorized by [3].

This finding is consistent with research by [2] which shows that social media plays a significant role in shaping the new meanings of popular vocabulary in Indonesia. However, this study provides additional evidence in the form of quantitative indicators of this change.

1.3. Analysis of the Shift in Meaning of the Loanword Technology

a. Expansion of Meaning

In its early period, the term "platform" tended to appear in technical contexts, such as software systems or digital infrastructure. In the second period, the term expanded in meaning, becoming used to refer to spaces of social, political, and digital economic expression. This shift is reflected in the semantically closest words in the vector space, such as "community", "media" and "ecosystem."

This phenomenon supports [6] findings, which state that the term "technology" is often metaphorized in everyday use on social media. Thus, the results of this study strengthen the



argument that the loanword "technology" is no longer merely terminological but has become part of society's discursive practices.

Table 1: The Shift In The Meaning Of Loanwords In Indonesian Using Word Embingment

Word	Embedding	Shift
Mouse (English:Tikus)	"Right-click the mouse to bring up the menu"	Shifting from rodent to computer pointing device
Application	"Please update your mobile banking application"	From general "application of theory" to specific software for devices
Capsule (English:Capsule)	"Logistics deliveries now use automated underground capsules"	Shifting from drug containers to modern transportation or delivery units
Cloud (Equivalent to Cloud)	"Company data is securely stored in cloud computing"	From a collection of water vapor to an internet data storage server
Network	"We are fixing the 5G network disruption in this area"	From a rope/social structure to a digital telecommunications system
Board	"Be careful when installing the processor on the motherboard"	Shifting from wood to the main electronic circuit
Window	"Close all unused browser windows"	From a building air vent to a graphical display interface on a screen
Server	"The central server is experiencing a data overload"	From a waiter to a computer providing data services
Architecture	"This system uses the latest artificial intelligence architecture"	From the science of building design to the logical structure of software systems
Wall	"Don't forget to activate the firewall for security"	Shifting from the physical boundaries of buildings to a digital security system that protects data

b. Connotational Change

The word "account" exhibits a shift in connotation from a digital administrative term to a symbol of online social identity. Between 2020 and 2024, the word frequently appeared in the context of trust, reputation, and social relations. This shift in semantic associations suggests that information technology contributes to the construction of linguistic identity in the digital space.

This finding aligns with [4] who emphasized that social media facilitates the formation of new meanings through repeated communication practices. However, this study demonstrates that these changes can be objectively measured through changes in word vector distances.

c. Comparison with Previous Research

When compared with previous research in Indonesia, this study demonstrates a distinct methodological contribution. Research by [3] and [2] analyzed meaning change qualitatively using a discourse approach. Meanwhile, this study combines this approach with word embedding techniques, thus demonstrating how and to what extent meaning change occurs.

Internationally, these findings align with the diachronic word embedding approach proposed by [14][10]. However, this study extends the application of this approach to Indonesian, which remains relatively underexplored in computational linguistics studies.

d. Linguistic and Technological Implications

Linguistically, the results of this study demonstrate that the shift in meaning of loanwords (technology) is a systematic phenomenon that can be analyzed empirically. This enriches the study of Indonesian semantics with a big data-based linguistic data approach.

From an information technology perspective, this study demonstrates that word embedding techniques are not only relevant for natural language processing tasks but also effective as a linguistic analysis tool. The collaboration between linguistics and information technology in this research demonstrates the great potential of an interdisciplinary approach in developing Indonesian language and literature studies in the digital age.

e. Research Limitations

Despite producing significant findings, this study has several limitations. First, the corpus used is limited to specific online sources, thus not fully representing variations in Indonesian across all levels of society. Second, the word embedding model used is still static and does not fully capture the nuances of sentence context.

These limitations open up opportunities for further research to use more sophisticated contextual language models and expand the corpus data sources.



4. CONCLUSION

This research demonstrates that the development of information technology has a significant influence on language dynamics, particularly in the use and development of the meaning of loanwords (technology) in Indonesian. Using a computational linguistics approach based on word embedding techniques, this research successfully identified semantic shifts in a number of technological vocabulary words used intensively in the digital space.

The analysis reveals that loanwords (technology) have not only experienced an increase in frequency of use but also systematic changes in meaning, particularly in the form of expansion of meaning and changes in connotation. These findings indicate that digital media and online interactions act as new social spaces that collectively shape language meaning. Thus, language in the digital context is not static, but continuously adapts to technological developments and societal communication practices.

Methodologically, this study confirms the effectiveness of word embedding techniques as a quantitative tool for analyzing word meaning shifts. The integration of a computational approach with qualitative semantic analysis allows for a more objective and measurable disclosure of meaning changes. This approach enriches the study of Indonesian language and literature with an interdisciplinary perspective that connects linguistics and information technology.

Based on the research findings and limitations, several suggestions can be made as follows.

First, future research is recommended to expand the scope of the corpus by including more diverse data sources, such as digital literary texts, reader comments and other social media platforms, to provide a more comprehensive representation of Indonesian language usage.

Second, the use of more advanced contextual language models, such as contextualized word embeddings, can be considered to capture the more complex nuances of word meaning that depend on sentence context.

Third, from a linguistic and educational perspective, the results of this study can be used as material for developing digital literacy and updating vocabulary in Indonesian language learning. Understanding the shifting meanings of technological loanwords is expected to foster critical and adaptive language awareness in the digital age.

In general, this research is expected to become an academic reference for technology-based semantic studies of Indonesian, digital literacy studies and encourage closer collaboration between the fields of language and literature with information technology in the development of linguistic research in the future.

REFERENCES

- [1] M. Castells, *The Rise of the Network Society*, 2nd ed. Oxford, U.K.: Wiley-Blackwell, 2010.
- [2] D. Maharani, H. S. Simanjuntak, N. Cahyani, R. Hazizah, and Y. Sari, "Maknadalam era digital: Kajiansemantikterhadapbahasa di media sosial Indonesia," *Semantik: JurnalRisetIlmuPendidikan, Bahasa danBudaya*, vol. 3, no. 3, pp. 79–109, 2025. [Online]. Available: <https://journal.aspirasi.or.id/index.php/Semantik/article/view/1830>
- [3] D. G. Ayu, A. F. P. Hanafi, N. S. Rianita, N. M. Hidayat, and E. Nurhayati, "Semantikkebahasaan Indonesia dalamdinamikateknologidan media sosial Twitter," *Sabda: Jurnal Sastra dan Bahasa*, vol. 19, no. 2, pp. 210–224, 2024. [Online]. Available: <https://jurnal.anfa.co.id/index.php/sabda/article/view/2426>
- [4] E. Siringoringo, S. Situmorang, J. B. Sitorus, U. Y. Sihotang, and Y. Sari, "Perubahanmakna kata dalambahasa Indonesia pada media sosial Instagram," *JurnalPendidikanMotivasidan Bahasa Harapan*, vol. 4, no. 2, pp. 45–58, 2025. [Online]. Available: <https://ourhope.biz.id/ojs/index.php/JP/article/view/131>
- [5] A. Chaer, *LinguistikUmum*. Jakarta, Indonesia: RinekaCipta, 2015.
- [6] U. Malikha, "Transformasimaknakosakatabahasa Indonesia dalam meme Instagram," *JurnalTinta*, vol. 7, no. 2, pp. 134–148, 2025. [Online]. Available: <https://ejournal.alqolam.ac.id/index.php/jurnaltinta/article/view/2223>
- [7] Linguistik Indonesia, "Pergeseranmaknakosakata flora dalammetafora di media sosial Twitter: Kajiansemantik," *Linguistik Indonesia*, vol. 43, no. 1, pp. 67–82, 2025. [Online]. Available:



- https://ojs.linguistik-indonesia.org/index.php/linguistik_indonesia/article/view/418
- [8] S. Ullmann, *Semantics: An Introduction to the Science of Meaning*. Oxford, U.K.: Basil Blackwell, 1962.
- [9] H. Ma, A. Y. Husodo, and B. Irmawati, "PERFORMANCE COMPARISON OF NAIVE BAYES AND BIDIRECTIONAL LSTM ALGORITHMS IN BSI MOBILE REVIEW SENTIMENT ANALYSIS PERBANDINGAN KINERJA ALGORITMA NAÏVE BAYES DAN BIDIRECTIONAL," vol. 6, no. 1, pp. 159–171, 2025. Available: <https://jutif.if.unsoed.ac.id/index.php/jurnal/article/view/4178/727>
- [10] A. Kutuzov, E. Velldal, and L. Øvrelid, "Contextualized language models for semantic change detection: Lessons learned," *arXiv preprint*, 2022. [Online]. Available: <https://arxiv.org/abs/2209.00154>
- [11] S. Montanelli and F. Periti, "A survey on contextualised semantic shift detection," *arXiv preprint*, 2023. [Online]. Available: <https://arxiv.org/abs/2304.01666>
- [12] B. E. S. Dewi, "Pengukurankemiripankalimatbahasa Indonesia menggunakanFastText," *TRIDI: JurnalTeknologiInformasidan Digital*, vol. 3, no. 1, pp. 55–66, 2025. [Online]. Available: <https://banisalehjurnal.ubs.ac.id/index.php/tridi/article/view/99>
- [13] M. R. H. Aditya, M. Ilham, D. F. Surianto, and A. M. Mappalotteng, "Evaluasipengukuransemantiks sinonim KBBI menggunakan word embedding," *Jurnal Nasional TeknikElektro danTeknologiInformasi*, vol. 14, no. 1, pp. 1–10, 2025. [Online]. Available: <https://jurnal.ugm.ac.id/JNTETI/article/view/17117>
- [14] W. L. Hamilton, J. Leskovec, and D. Jurafsky, "Diachronic word embeddings reveal statistical laws of semantic change," *Proc. 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 1489–1501, 2016, doi: 10.18653/v1/P16-1141.
- [15] D. Card, "Substitution-based semantic change detection using contextual embeddings," *Proc. 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 635–646, 2023. [Online]. Available: <https://aclanthology.org/2023.acl-short.52>
- [16] B. E. S. Dewi, "Pengukurankemiripankalimatbahasa Indonesia menggunakanFastText," *TRIDI: JurnalTeknologiInformasidan Digital*, vol. 3, no. 1, pp. 55–66, 2025. [Online]. Available: <https://banisalehjurnal.ubs.ac.id/index.php/tridi/article/view/99>

