

**LAPORAN AKHIR PENELITIAN**



**Pengembangan Aplikasi Mobile Penerjemah Bahasa Indonesia-Jawa Berbasis  
Cloud**

**TIM PENGUSUL**

**Kristian Adi Nugraha, S.Kom.,MT  
Danny Sebastian, S.Kom.,MM,MT**

**DUTA WACANA**

**Informatika**

**UNIVERSITAS KRISTEN DUTA WACANA**

**November 2022**

## DAFTAR ISI

DAFTAR ISI.....	i
RINGKASAN.....	iii
BAB 1 PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan Penelitian .....	7
1.4 Manfaat Penelitian .....	7
1.5 Batasan Penelitian.....	2
1.6 Luaran Penelitian .....	2
BAB 2 TINJAUAN PUSTAKA .....	4
2.1 Pengolahan Bahasa Natural .....	5
2.2 Pra-Pemrosesan Teks.....	5
2.3 <i>Cross-lingual Information Retrieval (CLIR)</i> .....	5
2.4 <i>Cloud Computing</i> .....	6
BAB 3 METODOLOGI PENELITIAN.....	8
3.1 Roadmap Penelitian .....	8
3.2 Pengumpulan Data Awal .....	8
3.3 Pembangunan Sistem.....	9
3.4 Pengujian Sistem.....	9
3.5 Analisis Sistem .....	9
3.6 Penulisan Laporan.....	9
BAB 4 BIAYA DAN JADWAL.....	<b>Error! Bookmark not defined.</b>
4.1 Ringkasan Anggaran Biaya.....	<b>Error! Bookmark not defined.</b>
4.2 Jadwal Penelitian .....	<b>Error! Bookmark not defined.</b>

DAFTAR PUSTAKA ..... 19

LAMPIRAN..... 21



## RINGKASAN

Indonesia merupakan negara yang memiliki banyak ragam bahasa daerah. Namun sebagian besar bahasa daerah tersebut tidak terdokumentasi dengan baik, menjadi kendala bagi orang yang ingin mempelajari bahasa tersebut. Apabila kondisi tersebut dibiarkan terus menerus, maka bahasa daerah tersebut terancam mengalami kepunahan. Pada penelitian ini, penulis memiliki gagasan untuk membangun sebuah aplikasi penerjemah Bahasa Indonesia ke bahasa daerah, atau sebaliknya. Aplikasi tersebut dibangun menggunakan metode *Cross-lingual Information Retrieval* dengan arsitektur berbasis awan (*cloud computing*) yang memungkinkan untuk pembaruan data secara terus menerus tanpa harus mengubah perangkat yang sudah ada. Dengan adanya aplikasi ini, diharapkan dapat mempermudah seseorang dalam mempelajari sebuah bahasa daerah, sekaligus menjadi bentuk dokumentasi bahasa daerah tersebut dalam rangka pelestarian budaya agar tidak terancam punah.

Kata Kunci: aplikasi *mobile*, bahasa daerah, *cloud*, *cross-lingual information retrieval*, translasi



# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Indonesia merupakan negara yang memiliki banyak keragaman suku dan budaya. Berdasarkan data BPS tahun 2010, Indonesia memiliki 1340 suku dengan total jumlah bahasa daerah sebanyak 718 bahasa (Abdulah, 2021). Tidak semua bahasa daerah terdokumentasikan dengan baik, misalnya dalam bentuk buku dan kamus baik fisik maupun *digital*. Hal tersebut dapat menjadi kendala bagi seseorang yang ingin mempelajari bahasa tersebut, karena dokumentasi mengenai bahasa daerah tersebut sulit untuk ditemukan di toko buku, perpustakaan, atau *internet*. Sehingga untuk mempelajari bahasa daerah tersebut hanya dapat dilakukan dengan bertemu secara langsung kepada orang yang telah menguasai bahasa tersebut sebelumnya. Selain itu, bahasa daerah yang telah didokumentasikan oleh pihak ketiga juga masih memiliki beberapa kekurangan. Misalnya Bahasa Jawa yang terdapat pada aplikasi Google Translate, di mana aplikasi tersebut hanya memandang Bahasa Jawa sebagai satu jenis bahasa saja. Padahal dalam Bahasa Jawa terdapat dua jenis tingkatan yaitu Krama dan Ngoko yang memiliki daftar kata-kata yang berbeda. Akibatnya saat kita menerjemahkan Bahasa Indonesia ke Bahasa Jawa menggunakan Google Translate, maka Google Translate akan menggunakan kata-kata campuran yang diambil dari Bahasa Jawa Krama dan Bahasa Jawa Ngoko sekaligus yang seharusnya tidak dapat digunakan secara bersamaan. Apabila tidak terdapat langkah-langkah konkrit terkait pendokumentasian bahasa daerah, maka bahasa daerah tersebut akan terancam punah dan tidak dapat diwariskan pada generasi-generasi berikutnya.

Pada penelitian ini, penulis akan membangun sebuah aplikasi berbasis *mobile* untuk menerjemahkan kata-kata dalam Bahasa Indonesia berbentuk teks maupun suara yang akan diterjemahkan ke bahasa daerah. Aplikasi yang akan dibangun didesain agar dapat mengakomodasi berbagai macam bahasa daerah sekaligus, namun selama tahap pengembangan dan pengujian penulis akan menggunakan Bahasa Jawa sebagai objek penelitian. Dengan menggunakan teknologi *cloud computing*, maka seluruh data beserta pemrosesan terdapat pada *server* yang selalu *online*, dengan demikian apabila terdapat perubahan data seperti penambahan bahasa atau kosakata baru, maka seluruh aplikasi dapat terbaru seketika tanpa perlu pemasangan

atau pengunduhan ulang. Harapan penulis, aplikasi tersebut dapat digunakan sebagai sarana seseorang untuk mempelajari bahasa daerah, sehingga dapat menjadi salah satu jalan untuk menyelamatkan bahasa daerah dari kepunahan.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan yang telah dikemukakan sebelumnya, permasalahan yang akan diteliti dalam penelitian ini adalah sebagian besar bahasa daerah di Indonesia tidak terdokumentasi dengan baik, sehingga dapat terancam punah.

## 1.3 Batasan Penelitian

Batasan penelitian ini adalah:

- Bahasa daerah yang digunakan sebagai objek penelitian adalah Bahasa Jawa.
- Proses terjemahan terbatas dari Bahasa Indonesia ke bahasa daerah atau dari bahasa daerah ke Bahasa Indonesia

## 1.4 Luaran Penelitian

Beberapa Kontribusi nyata atau luaran yang dihasilkan dari penelitian ini dijelaskan pada Tabel 1.1 Rencana Target Capaian dibawah ini.

**Tabel 1.1 Rencana Target Capaian**

No	Jenis Luaran				Indikator Capaian		
	Kategori	Sub Kategori	Wajib	Tambahan	TS	TS+1	TS+3
1	Artikel ilmiah dimuat di jurnal	Internasional Bereputasi					
		Nasional Terakreditasi	√		accepted		
		Nasional Tidak Terakreditasi					
		Internasional Terindeks	√		accepted		

	Artikel ilmiah dimuat di prosiding	Nasional					
3	<i>Invited speaker</i> dalam temu ilmiah	Internasional					
		Nasional					
4	<i>Visiting lecturer</i>	Internasional					
5	Hak Kekayaan Intelektual (HKI)	Paten					
		Paten sederhana					
		Hak Cipta		√	accepted		
		Merk dagang					
		Rahasia dagang					
		Desain Produk Industri					
		Indikasi Geografis					
		Perlindungan varietas tanaman					
		Perlindungan topografi sirkuit terpadu					
6	Teknologi Tepat Guna						
7	Model/Purwarupa/Desain/Karya Seni/Rekayasa Sosial						
8	Buku Ajar (ISBN)						
9	Tingkat Kesiapan Teknologi (TKT)		√	3			

## **BAB 7**

### **KESIMPULAN DAN SARAN**

Penelitian dari tahap awal, yaitu pengumpulan data hingga pengujian sistem telah dapat dilaksanakan dengan baik dan lancar. Pada waktu yang tersisa, penulis akan menyelesaikan artikel seminar internasional yang berisi mengenai perancangan aplikasi secara keseluruhan.





## DAFTAR PUSTAKA

- Zong, Z., & Hong, C. (2018). On Application of Natural Language Processing in Machine Translation. *2018 3rd International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*. Huhhot, China: IEEE.
- Sebastian, D., & Nugraha, K. A. (2019). Text normalization for Indonesian abbreviated word using crowdsourcing method. *2019 International Conference on Information and Communications Technology (ICOIACT)* (pp. 529-532). Yogyakarta: IEEE.
- Abdulah, A. T. (2021, November 29). *Berapa Jumlah Bahasa Daerah di Indonesia? Ini Bahasa yang Terancam Punah*. Retrieved from Tribunnews: <https://m.tribunnews.com/pendidikan/2021/11/29/berapa-jumlah-bahasa-daerah-di-indonesia-ini-bahasa-yang-terancam-punah>
- Maryamah, M., Arifin, A. Z., Sarno, R., & Hasan, A. M. (2020). Adapting Google Translate using Dictionary and Word Embedding for Arabic-Indonesian Cross-lingual Information Retrieval. *2020 IEEE International Conference on Internet of Things and Intelligence System (IoT&IS)*. Bali, Indonesia: IEEE.
- Rahmani, A. (2017). Adapting google translate for English-Persian cross-lingual information retrieval in medical domain. *2017 Artificial Intelligence and Signal Processing Conference (AISP)*. Shiraz, Iran: IEEE.
- Nasharuddin, N. A., Abdullah, M. T., Kadir, R. A., & Azman, A. (2010). A review on the cross-lingual information retrieval. *2010 International Conference on Information Retrieval & Knowledge Management (CAMP)*. Shah Alam, Malaysia: IEEE.
- Li, X., Xu, C., Wang, X., Lan, W., Jia, Z., Yang, G., & Xu, J. (2019). COCO-CN for Cross-Lingual Image Tagging, Captioning, and Retrieval. *IEEE Transactions on Multimedia*, 2347 - 2360.
- Rodriguez, M. A., & Buyya, R. (2014). Deadline Based Resource Provisioning and Scheduling Algorithm for Scientific Workflows on Clouds. *IEEE Transactions on Cloud Computing*, 222 - 235.
- Junnila, V., & Laihonen, T. (2016). Information Retrieval With Varying Number of Input Clues. *IEEE Transactions on Information Theory*, 625 - 638.

- Xu, B., Bu, J., Chen, C., Wang, C., Cai, D., & He, X. (2015). EMR: A Scalable Graph-Based Ranking Model for Content-Based Image Retrieval. *IEEE Transactions on Knowledge and Data Engineering*, 102 - 114.
- Xu, B., Lin, H., Lin, Y., Ma, Y., Yang, L., Wang, J., & Yang, Z. (2018). Improve Biomedical Information Retrieval Using Modified Learning to Rank Methods. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* , 1797 - 1809.
- Mahalakshmi, M. P., & Fatima, N. S. (2021). Maximum Entropy Principle based Document Ranking with Term Selection Analysis for Cross-Lingual Information Retrieval . *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*. Tirunelveli, India: IEEE.
- Patel, J., Makvana, K., & Shah, P. (2019). Cross-lingual Information Retrieval: application and Challenges for Indian Languages. *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*. Bombay, India: IEEE.
- Wu, C., Toosi, A. N., Buyya, R., & Ramamohanarao, K. (2021). Hedonic Pricing of Cloud Computing Services. *IEEE Transactions on Cloud Computing*, 182 - 196.

