

KE ARAH MENDOKUMENTASIKAN BAHASA PERIBUMI MENGGUNAKAN TEKNOLOGI DIGITAL UNTUK GENERASI MASA DEPAN

**oleh
Sarah Samson Juan**

Timbalan Dekan Hubungan Industri dan Masyarakat
Fakulti Sains Komputer dan Teknologi Maklumat
Universiti Malaysia Sarawak

Pengenalan

Malaysia mempunyai masyarakat yang berbilang kaum dan boleh bertutur di dalam pelbagai bahasa (multilingual). Seperti yang dilaporkan oleh Ethnologue, katalog dalam talian untuk bahasa-bahasa dunia, terdapat lebih dari 100 bahasa hidup di negara ini. Namun begitu, kita sedia maklum terdapat bahasa-bahasa yang hampir pupus disebabkan oleh pemodenan dan pilihan bahasa untuk komunikasi. Sebagai salah satu inisiatif untuk melestarikan dan menghidupkan kembali bahasa-bahasa Malaysia, projek kami berhasrat mengumpulkan bahasa-bahasa menggunakan teknologi digital iaitu membangunkan sebuah sistem web pengurusan bahasa yang dipanggil, Kalaka.

Kalaka telah dibangunkan untuk membantu menguruskan pengumpulan data ucapan dan teks awal serta perincian penutur asli. Kami juga mengaktifkan teknologi untuk menyinkronkan data yang telah dikumpulkan menggunakan aplikasi mudah alih Kalaka, dimana aplikasi mudah alih ini digunakan untuk menyimpan suara dan teks semasa di kerja lapangan, ke sistem web Kalaka. Tambahan pula, sistem web ini boleh menghasilkan analisis data yang mudah difahami serta menghasilkan peta bahasa untuk Malaysia, yang pertama seumpamanya, untuk menggambarkan penyebaran bahasa yang dituturkan oleh masyarakat di negara ini. Sistem ini berpotensi berguna untuk ahli bahasa, penyelidik dan pelajar yang berminat untuk menyumbang dalam usaha memelihara bahasa dan

menghidupkan semula. Pada masa yang sama, orang ramai boleh melayari peta bahasa kami yang interaktif sambil mendengar suara pencerita dari rakaman audio yang disimpan dalam sistem.

Latar Belakang Pengumpulan Data Bahasa Menggunakan Teknologi Digital

Terdapat semakin banyak aplikasi digital yang digunakan untuk pemeliharaan dan revitalisasi bahasa. Perisian SIL International dan Payap Language menghasilkan perisian untuk membuat kamus. Perisian ini dipanggil WeSay (tersedia di: www.wesay.org) adalah perisian percuma yang boleh digunakan oleh bukan ahli bahasa untuk membina kamus dalam bahasa mereka sendiri (Albright & Hatton, 2007; Perlin, 2012)

Sebuah e-kamus untuk 16 bahasa peribumi di Taiwan telah dibuat oleh Majlis Peribumi Taiwan dengan tujuan untuk menyeragamkan sistem ortografi bahasa-bahasa ini. Kandungan kamus dalam talian berdasarkan bahan bercetak digital (Taiwan Indigenous Council, 2016). Selain itu, Zahwa, sebuah aplikasi mudah alih, membolehkan pengguna mengambil gambar dan video pendek (Bettinson & Bird, 2017). Selain itu, ia membolehkan pengguna melakukan rakaman suara sambil menggesek gambar dan video. Pembangun merancang aplikasi untuk mendokumentasikan resipi atau wacana prosedur. Pengguna boleh merakam arahan dalam bahasa mereka sendiri dan memberikan terjemahan.

Bird et al., (2015) mengembangkan aplikasi mudah alih sebagai platform untuk dokumentasi bahasa kolaboratif. Aplikasi ini, yang dipanggil Aikuma, dimana ia boleh menyimpan data pertuturan dwibahasa, di mana pengguna dapat merakam, bercakap semula dan melakukan terjemahan lisan. Data yang disimpan di dalam peranti mudah alih dapat diambil dari alat transkripsi berbasis penyemak imbasnya, yang dapat dihubungkan melalui antarmuka WiFi, Bluetooth atau USB.

Kesimpulan

Teknologi digital merupakan elemen penting di dalam memudahkan proses pengumpulan data bahasa. Teknologi ini mempunyai pelbagai fungsi sedia ada yang boleh membantu menganalisa data serta membuat aplikasi interaktif yang memberi manfaat kepada pengguna sistem. Setakat ini, inisiatif kami di dalam orojek Kalaka telah berjaya digunakan untuk mengumpul data bahasa melalui kursus Pengenalan kepada Linguistik untuk pelajar Tahun 1 di Universiti Malaysia Sarawak. Selain daripada itu, projek Kalaka telah dipilih untuk menyertai pertandingan Malaysia Technology Expo 2021 yang akan diadakan secara maya pada pertengahan bulan Mac akan datang.

Penghargaan

Setinggi-tinggi penghargaan diucapkan kepada Dewan Bahasa dan Pustaka Sarawak di atas jemputan pembentang kertas kerja dan pihak Universiti Malaysia Sarawak di atas sokongan menjalankan projek Kalaka.

Rujukan:

- Albright, E., & Hatton, J. (2007). WeSay: A tool for engaging native speakers in dictionary building. *Documenting and Revitalizing Austronesian Languages*.
- Bettinson, M., & Bird, S. (2017). *Developing a Suite of Mobile Applications for Collaborative Language Documentation*. <https://doi.org/10.18653/v1/w17-0121>
- Bird, S., Hanke, F. R., Adams, O., & Lee, H. (2015). *Aikuma: A Mobile App for Collaborative Language Documentation*. <https://doi.org/10.3115/v1/w14-2201>
- Perlin, R. (2012). WeSay, A Tool for Collaborating on Dictionaries with Non-Linguists. *Language Documentation & Conservation*, 6, 181–186.
<http://www.himalayanlanguages.org/files/perlin/wesay-perlin.pdf>
- Taiwan Indigenous Council. (2016). *Aboriginal Ethnic Language Dictionary*:
<http://e-dictionary.apc.gov.tw/Index.htm>