

Transliterasi Melayu ke Jawi Menggunakan Teknik Pemetaan Karakter

Dzulfikar Fatahillah*

Universitas Muhammadiyah Malang,
dzulfikarcode@gmail.com*

Abstrak

Transliterasi melayu ke jawi secara umum digunakan untuk penulisan kitab-kitab yang diterjemahkan dari bahasa arab dan kitab Islam yang dikembangkan ulama' di asia tenggara. Transliterasi melayu latin ke jawi saat ini menggunakan pendekatan kata arabic sehingga hasil mendekati kata arabic aslinya. Namun dalam perkembangan bahasa melayu banyak perubahan ejaan dan kata istilah baru. Sehingga belum menyelesaikan permasalahan homograf, sedikit vokal, ejaan yang tidak konsisten. Transliterasi pada penelitian ini menggunakan pendekatan ejaan melayu menggunakan metode pemetaan karakter berdasarkan bunyi huruf secara umum pada kata melayu. Penelitian ini mentransliterasi seluruh karakter/huruf tiap kata dari satu kalimat atau lebih. sehingga tidak ada huruf yang hilang dan opsi menambahkan tanda dalam huruf vokal. Dari pengujian yang dilakukan pada jawi baru menggunakan jawi versi Pusat Rujukan Persuratan Melayu (PRPM) menunjukkan tingkat jawi berkurang dan dengan menggunakan opsi tanda vokal sehingga pada opsi tersebut belum ditemukan adanya homograf dijawi.

Kata Kunci: Jawi, Pemetaan Karakter, Transliterasi

Abstract

Malay to jawi transliteration generally used to writing translated books from arabic and islamic book who developed by moeslem scholar in southeast asia. Malay to jawi transliteration currently using arabic word approach. That the result close to real arabic word. But in the development of Malay language many changes in the spelling and new terms. So have not yet solve homograph problems, less vowel, and spelling not consistency. Transliteration in this study using the Malay spelling approach using the character mapping method based on the sound of letters in general in Malay words. This research transliterates all characters / letters of each word from one or more sentences. so that no letters are lost and the option of adding a sign in vowels. From the tests conducted on the new jawi using the version of the Pusat Rujukan Persuratan Melayu (PRPM), the level of jawi is reduced and by using the vowel sign option so while using that option the homograph has not been found.

Keywords: Jawi, Character Mapping, Transliteration

1. Pendahuluan

Penulisan teks jawi telah ditemukan sejak abad ke 16 hingga saat ini. Penulisan jawi memiliki hubungan dengan perkembangan islam di asia tenggara[1]. Penulisan huruf jawi pada awalnya mengadopsi huruf arab yang memiliki 28 karakter dengan tambahan 6 چ,ك,غ,ز,ف,ن, character yang tidak terdapat pada bahasa arab. Jawi secara umum ditulis tanpa menggunakan harakat, dan huruf digunakan untuk sebagian vokal. transliterasi melayu ke jawi dapat dipahami sebagai merubah kata berdasarkan suara atau huruf dari karakter melayu (latin) ke jawi menggunakan pemetaan karakter, yang dalam ejaan dan pengucapan sama seperti melayu. Contoh kata "jawi" (melayu) ke "جاوي" (jawi).

Transliterasi dapat dipahami merupakan penyalinan dengan penggantian huruf dari abjad satu ke bentuk abjad yang lain [2]. Dalam penelitian ini memfokuskan transliterasi melayu ke jawi menggunakan pemetaan karakter berdasarkan bunyi huruf yang digunakan pada kata melayu.

Menurut Ghani [3] dalam penelitian Yonhendri[4] meneliti mesin transliterasi rumi jawi mengusulkan metode segmentasi suku kata latin, kata disegmentasi berdasarkan awalan, kata dasar, dan akhiran. Awalan dan akhiran langsung menuju pemetaan karakter dan kata dasar diproses dengan aturan jawi. Kemudian awalan kata dasar digabung kembali.

Segmentasi suku kata secara umum menggunakan teknik stemmer [5][6], hasil segmentasi berupa huruf, suku kata, digraf. Kemudian diolah menjadi imbuhan, kata, dan data mentah. Hasil segmentasi secara umum sebagian dilakukan rule checking, pencarian pada kamus kata [7], direct transliteration.

Bakar [8] menggunakan teknik buckwalter dalam transliterasi. Memetakan seluruh karakter yang unicode arabic dan melayu kedalam huruf dan simbol. Dalam kesimpulannya tidak semua huruf sering digunakan dalam malay corpora.

Transliterasi melayu ke jawi yang ada saat ini memiliki aturan penulisan jawi lama yang mengadopsi kata aslinya yakni kata arab. Transliterasi saat ini memiliki pemilihan karakter yang tidak konsisten dalam penggantian hurufnya seperti karakter yang memiliki kemiripan bunyi seperti "s" (ص, ش, س, ث) . juga bunyi "d", "h", dan lainnya. Sehingga sebagian kata hurufnya menggunakan huruf berdasarkan kata arab sedangkan yang lainnya acak.

Pada Tabel 1, jawi memiliki 5 bunyi vokal yang mengadopsi 3 huruf vokal arabic a,i,e,u,o (ا, و, ي, ي, ا) dan huruf vokal selain "a" "i" pada awal kata ditambahkan huruf alif "a" contoh "u" "او" karena tidak bisa memiliki bunyi tanpa huruf alif "a". huruf "e" pada huruf kedua pada kata terkadang di transliterasikan. Dan aturan-aturan lainnya yang menyebabkan permasalahan homograf, sedikit vokal, ejaan yang tidak konsisten.

Tabel 1. Contoh Vokal Jawi

Vokal		
Latin	Kata melayu	Kata jawi
a	ajar	اجر
i	sisi	سيسى
u	kuda	كورا
e	tempe	تيمبي
o	roda	رورا

Penelitian ini menggunakan metode pemetaan karakter menggunakan karakter unicode utf-8 berdasarkan bunyi huruf dengan dengan mentransliterasikan seluruh huruf secara langsung. setiap huruf merepresentasikan satu bunyi sehingga memiliki sifat unik. Huruf vokal memiliki tambahan opsi vokal sehingga dapat menjadi vokal unik yang sebelumnya 5 bunyi berdasarkan 3 huruf dan huruf "u" sama dengan "o" huruf "i" sama dengan "e". Teknik ini digunakan untuk mengurangi permasalahan diatas. Seluruh huruf dalam kata akan ditransliterasikan dan setiap huruf dan vokal memiliki bentuk abjad dan bunyi masing-masing yang saling memiliki perbedaan. Sehingga menjadi standar jawi moderen yang dapat menyesuaikan dengan ejaan baru seperti rumi baru [9] dan ejaan yang disempurnakan [2].

2. Metodologi

Pada bagian ini akan dijelaskan tentang kategorisasi karakter, pemetaan karakter dan arsitektur.

2.1 Kategorisasi Karakter

Karakter yang akan dipetakan terbagi 3 kategori yakni vokal, konsonan (b-z), karakter ganda.

2.1.1 Vokal

Tabel 2. Vokal

Vokal awal		Vokal	
Latin	jawi	Latin	jawi
a	ا	a	ا
i	اي	i	ي
u	او	u	و
e	اي	e	ي
o	او	o	و

Pada Tabel 2, bagian vokal terdapat 2 macam yakni vokal awal (initial vowel) dan vokal secara umum. Vokal awal adalah vokal yang terdapat pada awal kata. Vokal awal memiliki tambahan huruf alif ‘ا’ karena tidak dapat memiliki bunyi jika sebagai huruf tunggal. Sedangkan vokal secara umum adalah vokal pada tengah kata atau akhir.

2.1.2 Optional Vowel

Penelitian ini menemukan bentuk bunyi yang berbeda dalam classical arabic pada Al Qur’an dengan dialect hafs beberapa dari cetakannya memberikan tanda “U+06EB arabic empty centre high stop” sebagai tanda “imahlah” yang terdapat pada qur’an chapter... verse....yang dalam penelitian ini digunakan sebagai tanda vokal untuk memberikan sifat unik tiap karakter.

Tabel 3. Optional Vowel

Vokal awal		Vokal	
Latin	jawi	Latin	jawi
a	ا	a	ا
i	اي	i	ي
u	او	u	و
e	ايّ	e	ئي
o	اوّ	o	ويّ

Vokal opsional pada Tabel 3 merupakan suatu vokal yang digunakan sebagai opsi untuk mengurangi tingkat homograf antar vokal. sebagai contoh vokal i=e u=o untuk membedakannya penelitian ini menggunakan opsi tambahan tanda atau sign pada huruf.

2.1.3 Konsonan

Mendefinisikan seluruh konsonan melayu. Karakter “x” tidak digunakan dalam melayu dan digantikan dengan “eks”.

Tabel 4. Consonant

konsonan					
latin	jawi	latin	jawi	latin	jawi
b	ب	k	ك	s	س
c	چ	l	ل	t	ت
d	د	m	م	v	ف
f	ف	n	ن	w	و
g	گ	p	پ	y	ي
h	ه, هـ	q	ق	z	ز
j	ج	r	ر		

Pada Tabel 4, pembagian karakter pada konsonan karakter tunggal untuk mendefinisikan karakter konsonan tunggal yang memberikan bunyi tunggal. Bunyi tunggal yang dimaksud yakni merepresentasikan bunyi karakter jawi dengan bunyi karakter tunggal pada latin melayu.

2.1.4 Karakter Ganda

Tabel 5. Multichar

Karakter ganda			
latin	jawi	latin	jawi
ng	غ	th	ط
ny	ث	ts	ث
sh	ص	kh	خ
sy	ش	ch	ح
dh	ظ	‘d	ض
dz	ذ	gh	غ

Pada Tabel 5, karakter ganda mendefinisikan karakter jawi yang direpresentasikan menggunakan karakter ganda pada bunyi karakter latin melayu.

2.2 Pemetaan karakter

Pengkodean karakter terdapat enam jenis, enam jenis tersebut adalah kode Morse, kode Baudot, kode Hollerith, ASCII, EBCDIC, dan Unicode. Sebagian besar word editor menggunakan unicode. Utf-8 merupakan pengkodean karakter unicode yang paling sedikit dalam menyimpan data di memori komputer. Dalam penelitian ini menggunakan unicode utf-8 agar lebih mudah dalam transliterasi.

2.2.1 Unicode mapping

Pada Tabel 6, karakter yang akan dipetakan dibagi menjadi 2 yakni initial vowel dan allchar (consonant, vowel, multichar)

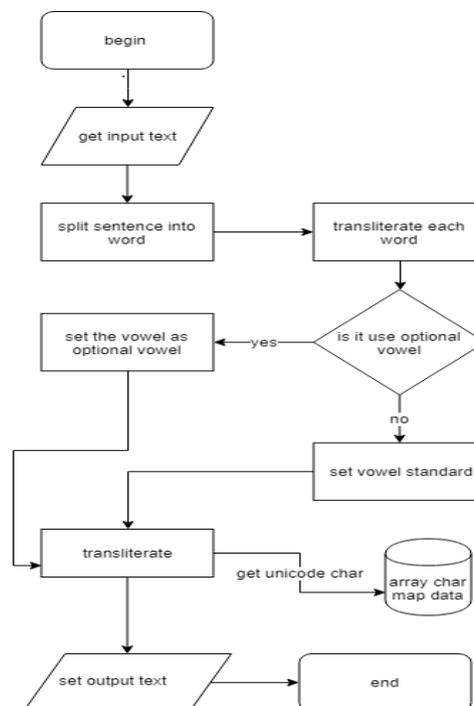
Tabel 6. Contoh Unicode Mapping

latin	unicode	jawi
ng	\u06A0	ڠ
ny	\u06BD	ڠ
v	\u06CF	ڤ
u	\u0627\u0648\u06EB	او
i	\u0627\u064A	اي
i	\u064A	ي

2.3 Sistem

2.3.1 Arsitektur program

Program dibangun menggunakan java 8 yang ditujukan pada aplikasi dekstop.



Gambar 1. Rancangan Flowchart Transliterasi

Rancangan Gambar 1 di atas menggambarkan proses transliterasi karakter, dimana input teks akan di pecah per-kata kemudian akan ditransliterasi tiap karakter dari kata tersebut. Kemudian terdapat pemilihan opsi optional vowel untuk memilih vokal dengan sign atau standar. karakter yang ditransliterasi sesuai dengan pembagian karakter yang telah di jelaskan

sebelumnya yang tersimpan dalam arrays char map. Kemudian akan di gabung dan ditampilkan kembali teks yang telah ditransliterasi.

2.3.2 Algoritma

Sistem memiliki beberapa tahapan-tahapan dalam mentransliterasi melayu ke jawi. Pada java, python dan pemrograman lainnya terdapat fungsi substring replace yang berfungsi merubah char tanpa perlu menggunakan n-gram, lebih efisien dan mudah dalam analisa data. Berikut tahapan-tahapan yang dilakukan:

```
start
  Read input
  Split sentences into arrays word
  For each word in sentences
    Replace first letter of arrays word
  if letter equal initial vowel in array initial characters map
  Replace all unreplaced letter left of arrays word
  if letter if letter equal allchar in array characters map
  Join word into sentences
  Print output
  Finish and return
```

3. Hasil

Penelitian ini mengajukan metode transliterasi sederhana dan proses yang mudah. Setiap kata yang dihasilkan memiliki sifat unik sehingga belum ditemukan homograf pada kata. Aturan Ejaan jawi dan pemilihan bunyi bentuk huruf berdasarkan melayu sehingga mengatasi masalah spelling not consistency.

Namun karena berdasarkan aturan melayu maka kata yang dihasilkan berbeda dengan kata yang berasal dari kata serapan arab. Sedikit berbeda dari jawi lama. Penelitian ini sebagai alternatif menyelesaikan permasalahan jawi yang ada.

4. Kesimpulan dan Saran

Transliterasi Jawi menggunakan metode diatas berfungsi dengan baik dan cukup mampu sebagai teks editor Jawi. Transliterasi dengan metode direct yang mentransliterasi tiap karakter kedalam teks jawi mampu mengurangi homograf secara umum dan untuk menangani homograf yang memiliki vokal sedikit dapat menggunakan opsi vokal yang disediakan.

Dalam transliterasi jawi terdapat beberapa bentuk penulisan yakni jawi baru, jawi lama, dan jawi yang ditulis berdampingan dengan tulisan arab. Untuk permasalahan homograf pada jawi baru telah dapat dikurangi menggunakan metode yang dipakai dalam penelitian ini. Namun untuk jawi lama atau kata arabic yang tanpa menggunakan harakat dapat menggunakan metode arabic transliteration custom charmap, buckwalter, corpus, behavior maupun kombinasi dengan menambahkan fungsi word suggestion untuk menyesuaikan kata yang beragam. Untuk menulis jawi berdampingan dengan arabic yang menggunakan harakat dapat pula menambahkan fitur harakat sebagai pembeda antara arabic dan jawi.

Referensi

- [1] E. U. Kratz, "Jawi spelling and orthography: A brief review," *Indonesia and the Malay World*, vol. 30, no. 86. Indonesia and the Malay World, pp. 21–26, 2002.
- [2] K. P. dan K. R. I. Badan Pengembangan dan Pembinaan Bahasa, "KBBI Kemendikbud," *Kemendikbud*, 2016. [Online]. Available: <https://kbbi.kemdikbud.go.id>.
- [3] R. A. Ghani, M. S. Zakaria, and K. Omar, "Jawi-Malay transliteration." 2009.
- [4] Yonhendri, "Enjin Transliterasi Rumi Jawi," Universiti Kebangsaan Malaysia, 2008.
- [5] S. Sulaiman, K. Omar, N. Omar, M. Z. Murah, and H. Abdul Rahman, "A Malay stemmer for Jawi characters," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011, vol. 7106 LNAI, pp. 668–676.
- [6] S. Sulaiman, K. Omar, N. Omar, M. Z. Murah, and H. A. Rahman, "The Effectiveness of a Jawi Stemmer for Retrieving Relevant Malay Documents in Jawi Characters," *TALIP*, vol. 13, no. 2. TALIP, pp. 1–21.

-
- [7] C. W. Shamsul, K. Omar, M. F. Nasrudin, and M. Z. Murah, "Machine Transliteration for Old Malay Manuscript." Selangor, Malaysia, pp. 19–25, 2012.
- [8] J. A. Bakar, K. Omar, M. F. Nasrudin, M. Z. Murah, and C. W. S. B. C. W. Ahmad, "Implementation of Buckwalter transliteration to Malay corpora." .
- [9] Dewan Bahasa dan Pustaka malaysia, "Pusat Rujukan Persuratan Melayu," 2018. [Online]. Available: <http://prpm.dbp.gov.my>.IEEE