



Available online:

<http://journal.imla.or.id/index.php/arabi>

Arabi : Journal of Arabic Studies, 4 (1), 2019, 65-74

DOI: <http://dx.doi.org/10.24865/ajas.v4i1.145>

MANUSKRIP ARAB DI NUSANTARA DALAM TINJAUAN LINGUISTIK KORPUS

Nur Hizbullah, lin Suryaningsih, Zaqiatul Mardiah

Universitas Al Azhar Indonesia, Indonesia

E-mail : nurhz@uai.ac.id

Abstract

The history and development of Islam in Indonesia are enriched by the existence of manuscripts written in Arabic language or written in Arabic script, like Pegon or Jawi although they do not use Arabic. In the context of corpus linguistics, the manuscript is a proof of the existence and dynamics of real Arabic usage by Indonesian speakers. This paper describes several classifications of manuscripts written in Arabic and their urgency as the material of Arabic corpus data in Indonesia in the context of the development of multidisciplinary Arabic research. Furthermore, the manuscript will be mapped based on seven types of Arabic corpus in Indonesia. Based on the mapping, it is projected that the majority of Arabic manuscripts in the archipelago are categorized as a corpus of scientific works, the corpus of Islamic studies, and corpus of literary works. For this purpose, it is necessary to process those manuscripts into digital text material to be analyzed with corpus processing applications through three stages: image scanning, image conversion into text, and manual text verification.

Keywords: *Arabic-written manuscripts, Arabic in Indonesia, linguistic corpus, Arabic corpus*

Abstrak

Sejarah dan perkembangan Islam di Indonesia diperkaya dengan keberadaan manuskrip Arab, baik yang ditulis dalam bahasa Arab, maupun yang ditulis dengan aksara Arab seperti Pegon ataupun Jawi, meski tidak menggunakan bahasa Arab. Dalam konteks linguistik korpus, manuskrip tersebut merupakan suatu bukti riil keberadaan dan dinamika penggunaan bahasa Arab oleh penutur Indonesia. Tulisan ini mendeskripsikan beberapa klasifikasi manuskrip yang ditulis dalam bahasa Arab dan urgensinya sebagai bahan data korpus bahasa Arab di Indonesia dalam konteks pengembangan penelitian bahasa Arab multidisipliner. Lebih lanjut, manuskrip tersebut nantinya akan dipetakan berdasarkan tujuh tipe korpus bahasa Arab di Indonesia. Berdasarkan pemetaan tersebut, diproyeksikan bahwa mayoritas manuskrip berbahasa Arab di Nusantara dikategorikan sebagai korpus karya ilmiah, korpus studi Islam, dan korpus karya sastra. Untuk keperluan tersebut, diperlukan pengolahan manuskrip menjadi material teks digital untuk dapat dianalisis dengan aplikasi pengolah korpus melalui tiga tahap: pemindaian citra, konversi citra menjadi teks, dan verifikasi teks secara manual.

Kata Kunci: manuskrip bertulisan Arab, bahasa Arab di Indonesia, linguistik korpus, korpus bahasa Arab

Pendahuluan

Kehadiran Islam di Nusantara tak lepas dari proses dakwah Islam yang menyebar dari wilayah Timur Tengah melalui berbagai jalur, seperti perdagangan, pendidikan, maupun sosial-budaya. Dalam konteks pendidikan, proses penanaman ajaran dan nilai Islam yang luhur dalam kehidupan masyarakat Nusantara dilakukan melalui dakwah dan lembaga pendidikan. Dalam konteks sosial-budaya, proses dakwah dan pendidikan umat itu dilakukan oleh para ulama melalui berbagai pendekatan agar bisa masuk dan diterima oleh masyarakat setempat di mana mereka berdakwah. Proses itu kemudian berlangsung secara luas di berbagai penjuru Nusantara, dari wilayah barat sampai ke wilayah timur. Penyebaran dakwah Islam di Nusantara tak pelak membawa serta bahasa Arab sebagai bahasa pengantar agama Islam yang ikut menyebar luas di Nusantara. Penggunaan bahasa Arab dalam konteks dakwah Islam tidak hanya terbatas pada topik maupun wacana ritual keislaman semata, tetapi juga merambah sampai wacana keilmuan Islam dan juga keilmuan bahasa Arab itu sendiri. Hal itu tak terlepas dari sumber pustaka berbahasa Arab yang dibawa oleh para ulama dari negeri asalnya, Timur Tengah. Alhasil, penyebarluasan bahasa Arab berjalan seiring dengan penyebarluasan Islam ke tengah masyarakat Nusantara.

Sehubungan dengan hal itu, para ulama di Nusantara tidak hanya memasyarakatkan bahasa Arab melalui wacana keislaman yang mereka sebarluaskan melalui dakwah. Sebagian dari mereka bahkan ada yang produktif menulis karya maupun literatur dalam berbagai bidang, seperti keislaman, sejarah, sosial, budaya, dan tak ketinggalan sastra dengan menggunakan bahasa Arab. Selain menggunakan bahasa Arab, beberapa karya dan literatur ditulis dengan aksara Arab—meski tidak menggunakan bahasa Arab, tetapi menggunakan bahasa lokal. Tradisi intelektual yang sangat menarik itu telah berlangsung sekian lama dan menyebar luas di beberapa daerah. Banyak di antara karya-karya monumental itu yang kini menjadi manuskrip istimewa dan bahkan masih diajarkan oleh para ulama generasi terkini kepada masyarakat lokal. Meski beberapa karya dalam bahasa ataupun tulisan Arab tersebut ditulis oleh ulama keturunan Arab, ada pula karya yang ditulis oleh ulama asli pribumi, bukan keturunan Arab. Hal itu juga menjadi keistimewaan tersendiri dalam tradisi keilmuan dan keislaman di Indonesia, karena menggambarkan suatu kekayaan khazanah budaya yang unik dan patut dibanggakan oleh umat Islam di Indonesia khususnya.

Pembicaraan dan diskusi tentang “Nusantara” akan membuka wawasan dan informasi tentang perjalanan panjang rumpun Melayu dalam kaitannya dengan wilayah Nusantara. Aksara dan bahasa yang digunakan pada rumpun negara-negara Melayu akan saling beririsan dari satu fase perkembangan ke fase perkembangan berikutnya dengan sejumlah bahasa yang ada di Nusantara. Melihat luasnya bidang bahasan dan banyaknya ragam manuskrip Nusantara yang menggunakan variasi bahasa maupun aksara termasuk Arab, penting untuk terlebih dahulu dibuat semacam peta dan klasifikasi yang terperinci. Ragam dan sebab munculnya bahasa dan aksara yang digunakan oleh penutur di Nusantara berasal dari pertukaran ilmu para ulama Nusantara yang belajar di Timur Tengah lalu kembali ke tanah air membawa naskah berbahasa Arab atau menuliskan ide dan gagasannya dengan aksara dan bahasa Arab. Naskah-naskah tersebut kemudian dipelajari oleh masyarakat Indonesia yang tersebar di berbagai wilayah sebagai bahan kajian keilmuan. Adapun mengenai aksara Arab yang berbahasa Melayu dan bahasa daerah yang disebut sebagai aksara Jawi dan Pegon, tidak dapat diketahui secara persis kapan mulai ditemukan. Setiap pendapat memiliki bukti tersendiri sebagai dasar pijakannya. Ada yang mengatakan bahwa kira-kira tiga ratus tahun setelah penemuan aksara Arab di beberapa batu nisan barulah ditemukan bukti penggunaan aksara Arab pada tulisan yang berbahasa Melayu.

Bukti lain tulisan yang menggunakan aksara Arab-Melayu dapat dilihat pada peninggalan budaya dalam bentuk tulisan yang masih manual, yakni naskah atau manuskrip Melayu yang ditulis dengan tangan asli masyarakat Melayu. Peninggalan tulisan tangan tersebut tersimpan di berbagai negara. Tulisan tangan yang ditinggalkan mereka itulah yang dikatakan sebagai bentuk dinamika intelektual masyarakat yang hidup pada zaman lampau di Nusantara. Naskah Melayu

dapat dikatakan sebagai bukti bahwa aksara Arab-Melayu telah digunakan untuk menulis di Nusantara, sehingga keberadaan aksara Arab-Melayu telah menjadi suatu lambang budaya di kawasan kepulauan ini. Bahwa aksara Arab-Melayu cepat berkembang tiada lain karena masyarakat Nusantara menerima tulisan dan bacaan Arab ini langsung dari orang Arab yang datang ke Nusantara dalam konteks dakwah dan penyebaran Islam. Orang Arab mengajari masyarakat Melayu dan Nusantara cara menuliskan aksara Arab. Dengan demikian akhirnya orang Melayu menjadi mudah mengolaborasikan aksara Arab dengan bahasa Melayu sehingga terwujudlah sarana untuk menulis.

Masuk dan berkembangnya Islam di Nusantara tidaklah kaku ketika berhadapan dengan persoalan keagamaan sebagai ritual yang mewarnainya saja. Islam yang dekat dengan budaya dan kultur Arab hadir dengan membawa perangkat keilmuan yang tidak dapat dipisahkan dalam perjalanannya. Yang sangat urgen dan berdampak dalam proses islamisasi tersebut adalah aksara yang menjadi media penyampaian ilmu. Aksara yang menjadi titik tumpu dalam upaya mewakili maksud setiap makna ini menjadi hal yang sangat unik dan penting untuk dikaji. Betapa tidak, bahwa melalui proses panjang bernama alih aksaralah sebuah transformasi keilmuan dapat berjalan hingga dapat dibuktikan zaman hingga saat ini. Aksara yang digunakan pada naskah-naskah Nusantara ditulis dalam berbagai jenis tulisan, misal saja bahwa sejarah membuktikan adanya keterpengaruhannya komunitas besar yang sebelumnya sudah ada mewarnai Nusantara seperti munculnya tulisan dari India Selatan atau yang berasal dari tulisan Arab. Adapun tulisan yang berasal dari India Selatan, yaitu tulisan yang merupakan perkembangan dari tulisan Pallawa, sebut saja di antaranya adalah tulisan Sunda Kuno, Kawi, Batak, Makasar, Lampung, Jawa-Sunda, sedangkan yang bersumber dari tulisan Arab, yaitu tulisan yang merupakan perkembangan dari tulisan Arab, misalnya tulisan Arab-Melayu atau tulisan Jawi dan tulisan Pegon (Rahmawati, Nasution, & Nst, 2017).

Di antara kekhususan aksara Arab-Melayu yang dikenal dengan sebutan Jawi adalah kedekatannya dengan aksara Arab. Keduanya dianggap “berkolaborasi” dengan melakukan beberapa penyesuaian dan tambahan huruf. Artinya, aksara Jawi merupakan campuran aksara Arab yang terdiri dari 29 aksara yang dimulai dari “alif” sampai “ya” dan ditambah dengan lima aksara yang bukan aksara Arab, melainkan aksara yang diciptakan oleh orang Melayu sendiri. Penambahan aksara tersebut digunakan untuk variasi menjawab keperluan fonem Melayu yang lebih banyak dibandingkan fonem Arab itu sendiri. Aksara tambahan itu ialah “ca”, “nga”, “pa”, “ga” dan “nya”. Bentuk tempat aksaranya sama dengan aksara Arab namun ditambahkan dengan beberapa titik sebagai pembeda bunyi dan fungsinya. Sebagaimana yang dikatakan oleh Ikram, karena sistem fonologi bahasa Melayu tidak sama dengan sistem fonologi bahasa Arab, maka digunakan bantuan titik diakritik untuk menyatakan bunyi bahasa yang tidak ada di dalam bahasa Arab. Oleh karena itu, tidak semua huruf Arab dapat digunakan secara tepat untuk menuliskan bahasa Melayu, kecuali dengan melakukan beberapa penambahan titik dengan tidak mengubah bentuk huruf asalnya, seperti huruf p-c-g-ng-ny (Luthfi, 2016).

Penggunaan bahasa dan aksara Arab dalam manuskrip Nusantara memiliki sejarahnya sendiri. Pada masa berlangsungnya proses Islamisasi kepulauan Melayu dalam kurun waktu yang sangat panjang pada paruh kedua abad ke-13, banyak terjadi pengadopsian dan penggunaan kosa kata bahasa Arab dalam berbagai manuskrip (Suryaningsih, 2016). Salah satu bukti nyata yang menunjukkan hal itu adalah banyak ditemukannya historiografi Islam di semenanjung kawasan ini, termasuk di dalamnya mencakup wilayah Nusantara. Penggunaan istilah “Nusantara” sendiri sesungguhnya tidak dapat dipisahkan dari istilah “Melayu”, khususnya dalam kaitan dengan perkembangan aksara dan bahasa di kedua kawasan. Menurut hubungan sejarah dan letak geografis diduga kuat, bahwa bahasa Melayu merupakan bahasa yang awalnya digunakan di suatu daerah di Sumatera bagian Timur yang kemudian disebarluaskan oleh para imigran ke beberapa daerah sekitarnya, seperti Malaka, kepulauan Riau, dan kepulauan Lingga. Bahasa ini bahkan ditemukan sudah dipakai sejak zaman kerajaan Sriwijaya sebagai bahasa resmi, tidak hanya terbatas dalam

bidang administrasi kerajaan, tetapi juga sebagai bahasa pengantar dalam berbagai kegiatan keagamaan dan keilmuan. Bahasa Melayu pada masa kejayaan kerajaan Melayu di Malaka, Pasai, dan Aceh digunakan untuk menyusun dan menggubah karya sastra seperti puisi dan prosanya, yang semua produk tersebut dihasilkan di istana dan umumnya berupa karya sastra tertulis yang tergolong sastra rakyat seperti sastra lisan. Dalam kaitan dengan huruf Arab yang digunakan pertama kali di Nusantara, sejarah mencatat bahwa huruf Arab masuk bersamaan dengan masuknya Islam ke Nusantara. Aksara Arab ini menggantikan aksara sebelumnya yang digunakan oleh banyak masyarakat Nusantara, yaitu aksara Melayu kuno, yang diperoleh dari pengaruh aksara Pallawa dari agama (Roza, 2017). Berikut ini adalah salah satu contoh manuskrip dengan aksara Jawi yang diperoleh dari situs <https://jantungmelayu.com/2019/04/kolofon-dalam-manuskrip-melayu-riau-lingga/>.

Tabel 1. Susunan Aksara Jawi (“Abjad Jawi,” 2019)

ت ت	ta teh	ب ب	ba beh	ا ا	alif alif
ح ح	ha hah	ج ج	jim jeem	ث ث	tha theh
د د	dal dal	خ خ	kha khah	چ چ	cha tch eh
ز ز	zai zain	ر ر	ra reh	ذ ذ	dzal thal
ص ص	shad sad	ش ش	shin sheen	س س	sin seen
ظ ظ	dzo zah	ط ط	tho tah	ض ض	dhad dad
غ غ	nga	غ غ	ghain ghain	ع ع	ain ain
ق ق	qaf qaf	ف ف	pa	ف ف	fa feh
ل ل	lam lam	ك ك	gaf	ك ك	kaf kaf
و و	wau waw	ن ن	nun noon	م م	mim meem
ي ي	ya yeh	ه ه	ha heh	و و	va
A A	Jawi name Original name			ن ن	nya
	From Arabic		From Persian		Unique to Jawi

Gambar 1. Contoh ‘penggalan’ dari manuskrip dengan aksara Jawi



Sebagai sesuatu yang amat berharga, keberadaan manuskrip yang ditulis dalam bahasa Arab atau menggunakan aksara Arab merupakan bukti keberlangsungan sejarah panjang tradisi keilmuan dan keislaman di Indonesia (Mahfud & Zuhy, 2018). Di sisi lain, manuskrip tersebut menunjukkan kualifikasi dan otoritas keilmuan para penulisnya. Uniknya, manuskrip berbahasa Arab itu berada jauh dari tempat asal bahasa itu, yaitu kawasan Arab, dan berkembang begitu pesatnya sejak zaman dahulu sampai masa kini (Iswanto, 2015).

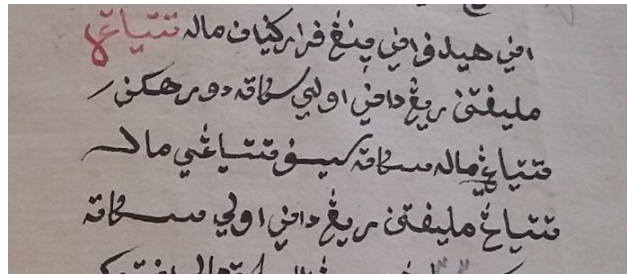
Menilik tentang islamisasi di tanah Jawa, adalah suatu kekhususan tersendiri karena Jawa merupakan jantung dan bukti sebuah transformasi besar-besaran Islam di Nusantara. Apapun yang terjadi dan mewarnai Jawa, menjadi penting untuk mengetahui dan mempelajari hal tersebut dengan seksama karena sebagian besar keilmuan Islam Arab tumpah ruah di telatah Jawa ini melalui berbagai macam cara. Aksara pegon merupakan satu aksara khusus yang terlahir dari produk akulturasi Arab dengan pribumi di Nusantara selain Jawi, dan aksara ini masih terjaga dan terlestarikan dari generasi ke generasi dengan upaya yang dilakukan oleh berbagai pihak. Di sisi lain, Pegon adalah huruf atau tulisan berbahasa Jawa yang ditulis dalam teks Arab (hijaiyah) (Astuti, 2014). Kromopawiro (1867: 1) mendefinisikan kata “pegon” berasal dari bahasa Jawa, *pego*, yang artinya “*ora lumrah anggone ngucapake*” (tidak lazim dalam pengucapan). Hal ini disebabkan karena banyaknya kata Jawa yang ditulis dengan tulisan Arab dan menjadi aneh ketika diucapkan. Menurut Pigeaud (dalam Pudjiastuti, 2006), menegaskan bahwa teks Jawa yang ditulis dengan aksara Arab disebut teks pegon artinya, sesuatu yang berkesan menyimpang. Penamaan ini disebabkan karena jumlah aksara yang diparalelkan dengan aksara Jawa lebih sedikit dari aksara Arab yang mejadi dasarnya.

Aksara Pegon ini dalam penggunaannya merupakan sebuah bagian dari tradisi sastra lokal masyarakat Islam di Jawa dalam mentransmisikan ajaran-ajaran Islam melalui teks. Selain berupa prosa, syair, dan undang-undang, teks-teks Pegon tersebut merupakan karya saduran ajaran Islam yang terdapat pada kitab-kitab karya ulama Timur Tengah abad pertengahan. Pola yang digunakan berupa pengembangan teks-teks Arab, seperti alih bahasa dan alih aksara ke dalam bahasa yang dapat dipahami oleh masyarakat setempat semisal Jawa. Masyarakat Islam di Jawa, terutama kalangan Islam tradisional/klasik sebagian besar sangat mengenal huruf Pegon dengan baik. Huruf ini sangat populer pasca masuknya Islam ke Nusantara. Beberapa dugaan muncul, salah satunya dari hasil kajian seorang cendekiawan asal Jawa bahwa Aksara Pegon uncul disebabkan karena alasan-alasan primordial dan politis. Arah dugaan ini seolah membuktikan “kolaborasi” antara huruf Arab dan bahasa Jawa dan telah menjadi barometer kemandirian Islam lokal di tanah Jawa sejak berabad silam lamanya. Dengan demikian, bisa dinyatakan secara konkret bahwa aksara Pegon adalah sebuah keniscayaan yang dipahami dan dipelajari secara turun temurun di kalangan Islam tradisional, terutama sekali para ulama sebagai pelaku dakwah dalam pendidikan tradisional (pesantren), sebagai produk akulturasi keilmuan dengan dunia Timur Tengah (Fikri, 2014).

Tabel 2. Aksara Pegon (“Abjad Pegon,” 2019)

ج jim	ث tha	ت ta	ب ba	أ alif
ڙ da	د dal	خ kha	چ cha	ح ha
ش shin	س sin	ز zai	ر ra	ذ dzal
ظ dzo	ڙ ta	ط tho	ض dhad	ص shad
ڦ pa	ڦ fa	ڱ nga	غ ghain	ع ain
م mim	ل lam	ڳ ga	ك kaf	ق qaf
ي ya	ه ha	و wau	ڻ nya	ن nun

Gambar. 2 Contoh 'penggalan' dari manuskrip dengan aksara Pegon dari situs <http://www.nubelanda.nl/warta/117-artikel/liputan/manuskrip-islam-nusantara/257-makna-pegon.html>



Metode Penelitian

Penelitian ini merupakan penelitian eksploratif yang terfokus pada data berupa manuskrip Arab dan variasi aksara turunannya, seperti Jawi dan Pegon sebagai bukti nyata keberlangsungan penggunaan aksara ini dalam sejarah penutur pribumi Nusantara hingga saat ini. Dengan demikian, metode yang akan digunakan adalah metode filologi. Data-data manuskrip akan diklasifikasikan secara analitis-deskriptif berdasarkan kategori korpus ilmiah. Lebih lanjut, eksplorasi terhadap manuskrip diarahkan kepada identifikasi permasalahan yang berkenaan dengan proses digitalisasi manuskrip menjadi material teks dokumen digital untuk dapat dianalisis dengan aplikasi pengolah korpus melalui tiga tahap, yaitu: pemindaian citra, konversi citra menjadi teks, dan verifikasi teks secara manual.

Hasil dan Pembahasan

Sejarah memperjelas bahwa pernah adanya hubungan intelektual secara langsung antara ulama-ulama Nusantara terkemuka kisaran tengah abad ke-15 M dengan para ulama Timur Tengah di mulai dengan adanya *halaqah* ilmiah seputar agama dan lain-lain. Keterpengaruhannya secara kultur, budaya dan bahasa yang di pakai selama berjalannya halaqah tersebut, menumbuhkan ciri khas tersendiri bagi karakter tulisan para ulama Nusantara kita dalam menuangkan ide dan pikiran mereka di tanah air. Tampak jelas bahwa tulisan Arab seperti yang kita kenal saat ini sebagai skrip huruf Pegon dan Jawi salah satu muatan teksnya terdiri dari beberapa kategori tulisan seperti yang dikenal dengan sebutan matan berpola *nazham* yang berbaris sebagaimana dikenal dengan sebutan syair dan narasi berupa komentar (Suryaningih, 2016).

Sejauh pengamatan yang pernah dilakukan oleh para filolog, naskah berbahasa dan beraksara Arab berjumlah ribuan dan tersebar di beberapa titik koleksi perpustakaan dan museum, baik di dalam ataupun luar negeri (Fathurrahman, 2000). Sebut saja untuk koleksi dalam negeri, terdapat kisaran 1000 buah naskah Arab di Perpustakaan Nasional (PNRI) Jakarta dan kisaran 400 naskah di Dayah Tanoh Abec, Seulimeum, Aceh. Naskah-naskah di Tanoh Abec mendapat perhatian khusus karena -dalam konteks keagamaan (Islam)- menggambarkan dasar pendidikan agama di Aceh pada abad ke-19. Selain di dalam negeri, salah satu bukti adanya koleksi manuskrip berbahasa dan beraksara Arab di luar negeri dengan tersedianya koleksi yang tersimpan antara lain di Universiteits Bibliotheek, Leiden, Belanda.

Untuk teks-teks keislaman ini setidaknya dapat dibagi menjadi tiga kelompok besar: (a) teks-teks yang memuat nilai keagamaan, seperti teks teologi (tauhid), fikih, tasawuf, dan lainnya; (b) teks-teks yang memuat nilai kesejarahan, seperti adat-istiadat, hokum, dan mitologi, mencakup teks-teks sejarah Islam, hukum-hukum (perundang-undangan) atau tatacara dan astrologi yang dipengaruhi oleh Islam serta kisah-kisah orang suci dalam Islam (ulama/wali), termasuk adat-istiadat, hukum dan mitologi, dan (c) teks-teks yang memuat nilai susastra, mencakup roman-roman atau epos-epos Islam, seperti cerita Menak Amir Hamzah, Roman Yusuf, atau cerita Johar Manikam (Iswanto, 2015).

Hal yang juga dianggap khusus pada manuskrip di Nusantara adalah terjadinya alih bahasa dan alih aksara dari teks asal. Alih bahasa dan aksara ini merupakan pergantian bahasa dan aksara

dari bahasa dan aksara lama (arkhais) yang terdapat di dalam naskah ke dalam bahasa dan aksara yang diketahui dan digunakan oleh masyarakat zaman sekarang. Tujuan utama alih bahasa dan aksara ini adalah membantu menjembatani teks lama dan tulisan yang menggunakan aksara kuno dengan pembaca masa kini dan mendokumentasikan naskah-naskah tersebut agar isi yang terdapat dalam naskah yang berupa ilmu pengetahuan serta kebudayaan dapat dipahami oleh masyarakat sehingga masyarakat dapat mengetahui adat, tradisi, dan kebudayaan di masa lampau (Tanjung, Nst, & Zulfadhli, 2017). Tradisi alih bahasa dan aksara ini ternyata juga marak dilakukan sejak masa lampau, sebut saja alih bahasa dan aksara dari teks asli berbahasa Arab menjadi teks Melayu. Selain proses akulturasi antar budaya yang terjadi pada proses alih bahasa dan aksara tersebut, faktor upaya penerjemahan teks satu ke teks lainnya adalah satu dari beberapa alasan terjadinya varian dalam teks yang kisi temanya sama. Oleh karena itu, pemetaan dan klasifikasi yang sangat beragam tersebut menunjukkan bahwa mayoritas manuskrip Arab yang muncul dengan ragam aksara dan bahasa di Nusantara terjadi akibat akulturasi, adaptasi, dan representasi budaya Arab dengan Nusantara melalui berbagai jalur dan kebutuhan saat itu. Dominasi bidang kajian di dalamnya dapat dikategorikan ke dalam bidang kebahasaan, keislaman, dan sastra.

Di sisi lain, upaya pelestarian ragam manuskrip di beberapa koleksi yang tersebar, berbagai naskah kuno kini juga telah dikumpulkan dalam katalog tercetak, disusun berdasarkan tema-tema kajian tertentu, maupun berdasarkan asal usul tempat ditemukannya naskah, ragam aksara dan bahasa. Katalog naskah kuno ini serupa dengan katalog di perpustakaan. Di antara contoh katalog yang dimaksud adalah Katalog Induk Naskah-naskah Nusantara Jilid 3-A dan 3-B; Fakultas Sastra UI (1997), Katalog Induk Naskah-naskah Nusantara Perpustakaan Nasional Republik Indonesia (1998), Katalog Induk Naskah-naskah Nusantara Jilid 5A: Jawa Barat Koleksi Lima Lembaga (1999), Direktori Edisi Naskah Nusantara (2000), Katalog Naskah-naskah Perpustakaan Puro Pakualaman (2005), dan Koleksi dan Katalogisasi Naskah Klasik Keagamaan Bidang Tasawuf (2013). Jauh sebelum katalog-katalog tersebut, Poerbatjaraka dkk telah menulis katalog naskah Nusantara dengan judul *Indonesische Handschriften* (1950) dan katalog karya Pigeaud *Literature of Java I-III* (1967-1970) (Pudjiastuti, 2006).

Selama ini, upaya pemetaan keberadaan naskah kuno sudah dilakukan oleh berbagai lembaga di Indonesia, seperti Masyarakat Pernaskahan Nusantara (Manassa) Jakarta, Islamic Manuscript Unit (ILMU) Pusat Pengkajian Islam dan Masyarakat (PPIM) UIN Syarif Hidayatullah Jakarta, Badan Litbang dan Diklat Kementerian Agama RI, dan Balai-balai di bawah koordinasinya, seperti Balai Penelitian Agama di Jakarta, Semarang, dan Makassar. Sebagian naskah-naskah kuno yang telah dikumpulkan dengan digitalisasi itu dibuatkan pangkalan data dengan nama *Thesaurus of Indonesian Islamic Manuscript* (T2IM). T2IM yang disponsori Puslitbang Lektor dan Khazanah Keagamaan Badan Litbang Agama dan Keagamaan RI.

Pada tahap awal, T2IM didasarkan pada katalog naskah, daftar naskah, dan pencatatan naskah yang telah diterbitkan. Alamat website T2IM dapat dilihat pada <http://tiim.ppim.or.id/>. Hasil digitalisasi naskah kuno Puslitbang Lektor Keagamaan terdapat pada <http://lektor.kemenag.go.id/manuskrip/>, sedangkan koleksi naskah kuno Perpustakaan Nasional RI dapat dilihat pada <http://www.perpusnas.go.id/collection-worksheets/manuskrip/>. Bahkan naskah kuno Nusantara di luar negeri juga dapat diakses dengan cara digital, antara lain pada website British Library, <http://britishlibrary.typepad.co.uk/asian-and-african/javanese/>.

Dalam pandangan keilmuan linguistik korpus, manuskrip-manuskrip tersebut merupakan "korpus" atau data kebahasaan riil dan faktual yang dapat dikaji dari berbagai sudut keilmuan kebahasaan maupun keilmuan lain. Semakin banyak manuskrip tersebut bisa ditemukan dan didokumentasikan, akan semakin kuat pula bukti perkembangan dan dinamika keilmuan dan keislaman di Indonesia.

Dalam kaitan proyek penyusunan korpus bahasa Arab di Indonesia yang digarap oleh IMLA (*Ittihad Mudarrisi Al-Lughah Al-'Arabiyah* atau Ikatan Pengajar Bahasa Arab di Indonesia), keberadaan manuskrip Arab sebagai korpus di Indonesia amat penting dan strategis, antara lain

dapat digunakan sebagai parameter produktivitas dan dinamika wacana keilmuan dan keilmuan klasik di Nusantara. Melihat luasnya jangkauan topik dan kekayaan ragamnya, korpus manuskrip Arab di Nusantara ini kelak dapat menyumbang data dalam jumlah sangat besar bagi koleksi korpus bahasa Arab di Indonesia. Berdasarkan kategorinya, korpus yang bersumber dari manuskrip Arab Nusantara ini dapat digolongkan sebagai korpus regional dan korpus historis (Nesselhauf, 2011). Dalam kaitan dengan rencana penyusunan korpus bahasa Arab di Indonesia, berikut klasifikasi korpus yang diproyeksikan muncul dan dapat dikoleksi berdasarkan ketersediaan bahan materialnya (Hizbullah & Muchlis, 2018).

1. Korpus karya ilmiah di bidang kebahasaan Arab, yang berisi kumpulan karya tulis ilmiah yang ditulis oleh akademisi Indonesia.
2. Korpus karya populer, yang memuat terbitan maupun karya populer berupa majalah, buletin, surat kabar, dan sejenisnya yang beredar di Indonesia.
3. Korpus karya sastra, yang mengoleksi karya-karya sastra dalam bahasa Arab ataupun bertulisan Arab meski tidak berbahasa Arab.
4. Korpus karya ilmiah keagamaan di bidang studi atau kajian Islam.
5. Korpus Alquran dan terjemahannya, yang diproyeksikan menjadi korpus data yang dapat dimanfaatkan sebagai objek kajian kebahasaan Alquran di Indonesia.
6. Korpus pembelajar atau korpus pendidikan, yang memuat berbagai data hasil pembelajaran oleh penutur Indonesia.
7. Korpus hadis dan terjemahannya.
8. Korpus dokumen yang berisi data terjemahan dalam bahasa Arab dan bahasa Indonesia.

Kedelapan klasifikasi tersebut akan ditambah dengan satu bank data suplemen, yaitu glosarium kata serapan Arab ke dalam bahasa Indonesia.

Korpus bahasa Arab di Indonesia yang akan didedikasikan sebagai “bank data” pertama khusus bahasa Arab di Indonesia ini akan menyimpan dan menyajikan secara komprehensif segala bentuk fakta kebahasaan riil tentang aktivitas berbahasa Arab penutur Indonesia dari masa ke masa, mulai dari masa klasik hingga masa modern. Seiring waktu, pekerjaan pendokumentasian, pengolahan, dan pengunggahannya ke web IMLA akan membuat jumlah maupun besaran data semakin meningkat dan bertambah besar. Pertambahan data korpus yang bisa diamati melalui dua pendekatan, yaitu jumlah arsip yang masuk dan volume data arsip, akan dapat menunjukkan korpus mana yang paling produktif dan dinamis dalam kemunculannya. Peta itu akan dapat mudah terlihat melalui tabel data koleksi web korpus yang akan terus diperbarui dan berkembang secara berkala.

Untuk kepentingan pendokumentasian, korpus-korpus tersebut perlu diolah dulu menjadi data digital dengan format teks tertentu untuk kemudian dikompilasikan dan diklasifikasikan dalam kesatuan data besar manuskrip Arab di Indonesia. Proses digitalisasi bahan data korpus, dalam hal ini, sedikit lebih jauh dari apa yang disebut dengan istilah alih media dan digitalisasi (Bustamam, 2017). Proses itu berujung pada produk citra digital dari halaman-halaman manuskrip. Lebih lanjut, proses pengolahan manuskrip sebagai bahan material sampai menjadi korpus berupa teks yang dapat dianalisis dengan aplikasi pengolah korpus dilakukan melalui tiga tahap: (1) pemindaian manuskrip menjadi citra digital, (2) konversi citra menjadi teks, dan (3) verifikasi teks secara manual.

Pada tahap kedua, proses yang dilakukan mengandalkan perangkat pemindai citra (scanner) yang terhubung dengan aplikasi pembaca dan pengonversi citra menjadi teks yang populer dengan istilah *optical character recognition* (OCR). Contoh dari aplikasi ini adalah ABBYY Fine Reader. Aplikasi semacam ini dapat membaca teks dalam bentuk citra dan mengonversinya menjadi teks yang dapat disunting ataupun diubahsuai melalui aplikasi pengolah kata. Dalam istilah lain, teknik ini disebut *scan-to-text* ‘pindai (langsung) menjadi teks’. Proses ini sejatinya amat sangat memudahkan proses pemerolehan bahan data korpus dari bahan data korpus yang masih berupa manuskrip konvensional ataupun yang sudah diolah menjadi citra digital. Akan tetapi, terdapat

tantangan tersendiri apabila manuskrip Arab yang dipindai tidak dapat langsung atau bahkan tidak dapat sama sekali diubah dari citra menjadi teks melalui aplikasi pemindai citra menjadi teks. Masalah yang sering terjadi adalah tingkat kejelasan atau keterbacaan teks yang terdapat dalam manuskrip. Semakin jelas citra karakter ataupun huruf pada teks, maka akan semakin jelas pula proses pemindaian dan konversi citra menjadi teks digital. Sebaliknya, jika citra memuat teks yang sudah sulit dibaca karena tulisan yang rusak apalagi hilang, maka dilakukanlah proses konvensional penulisan ulang teks yang terdapat dalam manuskrip ke dalam komputer agar mudah dikonversi menjadi format standar *.txt nantinya, sampai bisa diolah menggunakan aplikasi pengolah korpus semacam *WordSmith*, *AntCont*, *Sketch Engine*, dan sebagainya.

Tantangan lainnya dalam proses pemindaian menjadi citra dan konversi menjadi teks digital adalah pada pemrosesan karakter tertentu pada aksara Arab-Melayu (Jawi) yang sedikit berbeda dari aksara Arab baku. Perbedaan karakter itu ada pada konsonan /c/ yang ditandai dengan huruf چ , /g/ dengan huruf ك , /p/ dengan huruf ف , /v/ dengan huruf ذ , /ny/ dengan huruf ن , bunyi sengau /ng/ ditandai dengan huruf غ , serta bunyi vokal /e/ yang biasa ditandai dengan kombinasi konsonan tertentu dengan huruf ع (Ghani, Zakaria, & Omar, 2009). Banyak manuskrip Arab Nusantara yang bertulisan Arab namun tidak berbahasa Arab yang menggunakan karakter tersebut. Dengan demikian, pada proses konversi menuju teks digital, perlu penanganan khusus agar tidak menimbulkan masalah teknis yang lebih jauh pada teks yang sudah terkonversi.

Dalam konteks pemanfaatan korpus manuskrip Arab untuk penelitian multidisipliner, linguistik korpus sesungguhnya menyediakan semacam perangkat teknis untuk menelaah konten dari suatu teks dalam bahasa tertentu. Beberapa teknik analisis yang disediakan oleh subdisiplin ilmu ini, antara lain: daftar dan kalkulasi jumlah kata, frekuensi kemunculan kata dalam teks, penelusuran konkordansi kata dalam teks, dan penelusuran maupun pengujian konteks kata melalui konkordansi tersebut. Berangkat dari pengalaman bahasa Inggris, sebagai contoh, pemanfaatan teknik analisis linguistik korpus tersebut dapat dilakukan setidaknya pada tiga ranah besar penelitian, yaitu ranah linguistik dengan segala cabang kajiannya, ranah sastra dengan berbagai macam teksnya, dan ranah pengajaran bahasa yang mencakup semua proses dari hulu sampai hilir (Romer, 2006).

Namun demikian, tidak tertutup kemungkinan penelitian terhadap korpus manuskrip tersebut melalui bidang ilmu lain sesuai dengan konten manuskrip yang bersangkutan. Hanya saja, pendekatan linguistik korpus menawarkan teknik dan mekanisme khusus terkait penelusuran aspek tertentu dalam teks manuskrip dalam rangka mengungkap lebih jauh sisi-sisi kebahasaan teks yang bersangkutan.

Untuk kepentingan penyusunan korpus dari manuskrip Arab Nusantara, proses digitalisasi manuskrip yang dilakukan tidak berhenti hanya sampai pada bentuk citra (*image*) dari hasil pemindaian, tetapi juga sampai kepada bentuk teks digital dengan teknik dan mekanisme tertentu yang mengandalkan teknologi komputasi. Keberadaan dan ketersediaan model teks digital manuskrip Arab akan menguntungkan banyak pihak yang berkepentingan, khususnya dari segi pendokumentasian, karena format data digital dari suatu manuskrip lebih banyak ragamnya dan juga arah pemanfaatannya.

Simpulan

Keberadaan manuskrip Arab di Nusantara sungguh merupakan khazanah sekaligus karunia yang luar biasa bagi perkembangan keilmuan di tanah air. Perhatian dari banyak pihak terkait pelestarian, pendokumentasian, dan pengkajian terhadap manuskrip semakin meningkat terutama dari kalangan akademisi yang mencoba berkontribusi melalui pendekatan disiplin ilmu masing-masing. Ke depan, tentunya amat diharapkan bahwa temuan-temuan manuskrip Arab di Indonesia semakin banyak jumlahnya dan semakin berkontribusi penting bagi penambahan koleksi dan volume data korpus bahasa Arab di Indonesia. Di sisi lain, diharapkan semakin marak pula kerja sama kolaboratif antar pihak-pihak yang berkompeten untuk menjadikannya sebagai data ilmiah

yang perlu diolah dan dikaji lebih lebih luas dari berbagai sudut pandang keilmuan. Secara khusus, manuskrip Arab Nusantara diharapkan menjadi sesuatu yang membuka mata dunia bahwa Indonesia, meski jauh dari kawasan Arab, adalah bagian tak terpisahkan dari dinamika keilmuan dan keislaman yang integral dan komprehensif. Harus diakui pula bahwa Indonesia menjadi suatu kepingan penting dalam sejarah dan dinamika perkembangan peradaban Islam dan dunia sejak masa lalu sampai masa depan.[]

Daftar Rujukan

- Astuti, K. 2014. “Aspek Tasawuf dalam Manuskrip Beraksara Pegon: Kajian atas Serat Munjiyat Karya Kyai Saleh Darat dan Singir Parase Nabi Karya Anonim”, *Thesis*, Universitas Muhammadiyah Surakarta.
- Bustamam, R. 2017. “Eksplorasi dan Digitalisasi Manuskrip Keagamaan: Pengalaman di Minangkabau”, *Jurnal Lektur Keagamaan*, Vol. 15, No. 2.
- Fathurrahman, O. 2000. *Filologi dan Penelitian Teks-teks Keagamaan*. Bogor: t.tp.
- Fikri, I. 2014. “Aksara Pegon: Studi tentang Simbol Perlawanan Islam di Jawa pada Abad XVIII–XIX”, *Thesis*, UIN Semarang.
- Ghani, R. A., M. S. Zakaria., & K. Omar. 2009. “Jawi-Malay Transliteration”, *Proceedings*, International Conference on Electrical Engineering and Informatics.
- Hizbullah, N., & M. M. Muchlis. 2018. “Projected Characteristics and Content of Arabic Corpus in Indonesia”, *Advances in Social Science, Education and Humanities Research (ASSEHR)*, Vol. 154.
- Iswanto, A. 2015. “Kecenderungan Kajian Manuskrip Keislaman Di Uin Syarif Hidayatullah Jakarta”, *Al-Qalam*, Vol. 21, No. 1.
- Luthfi, K. M. 2016. “Kontekstualisasi Filologi Dalam Teks-Teks Islam Nusantara”, *Ibda’ Jurnal Kebudayaan Islam*, Vol. 14, No. 1.
- Mahfud, Sahal., & Halimi Zuhdy. 2018. “Kitâbah ‘Arab Pegon Khashâishuhâ wa Ishâmâtuhâ fî Tathwîr Ta’lîm al-Lughah al-‘Arabîyyah bi Indûnîsiyyâ”, *Jurnal Arabiyat*, Vol. 5, No. 2.
- Nesselhauf, N. 2011. “Corpus Linguistics: A Practical Introduction”, Retrieved from [http://www.as.uni-heidelberg.de/personen/Nesselhauf/files/ Corpus Linguistics Practical Introduction.pdf](http://www.as.uni-heidelberg.de/personen/Nesselhauf/files/Corpus_Linguistics_Practical_Introduction.pdf)
- Pudjiastuti, T. 2006. “Poerbatjaraka dan Manuskrip Islam”, *Thesis*, Universitas Indonesia.
- Rahmawati, N., B. Nasution., & M. I. Nst. 2017. “Alih Aksara dan Alih Bahasa Teks Kitab Fadhilah Syuhur: Menerangkan Bulan Ramadhan dan Kelebihan Ibadah-ibadah di Dalamnya dan Kelebihan Bulan Syawal Disusun oleh Imam Maulana Abdul Manaf Amin Al-Khatib”, *Bahasa Dan Sastra*, Vol. 5, No. 1.
- Romer, U. 2006. “Where the Computer Meets Language, Literature, and Pedagogy: Corpus Analysis in English Studies”, In A. & A. M.-W. Gerbig (Ed.), *How Globalization Affects the Teaching of English: Studying Culture Through Texts* (pp. 81–109). Lampeter: E. Mellen Press.
- Roza, E. 2017. “Aksara Arab-Melayu di Nusantara dan Sumbangsihnya dalam Pengembangan Khazanah Intelektual”, *Tsaqafah Jurnal Peradaban Islam*, Vol. 13, No. 1.
- Suryaningsih, I. 2016. “Sastra Islam dalam Manuskrip (Kajian Teks Beraksara Jawi sebagai Salah Satu Bahan Ajar Mahasiswa)”, *Prosiding Konferensi Nasional Bahasa Arab II*, Universitas Negeri Malang.
- Tanjung, R. A., B. Nst., & Zulfadhli. 2017. “Alih Aksara dan Alih Bahasa Teks Kitab Fadhilah Syuhur Menerangkan Kelebihan Bulan Muharram dan Kelebihan Hari Asyura Disusun oleh Imam Maulana Abdul Manaf Amin Al-Khatib”, *Jurnal Bahasa Dan Sastra*, Vol. 5, No. 1.