

PENGENALAN TEKNOLOGI HANDWRITING TEXT RECOGNITION: TRANSKRIBUS

Pada zaman dahulu, manusia menjadikan batu, lembaran daun, lembaran kulit kayu, hingga perkamen sebagai sarana komunikasi. Kita bisa menemukan beberapa coretan, lukisan dan tulisan yang berasal dari masa lalu, yang bahkan beberapa di antaranya masih belum diketahui artinya. Hingga kini, tulisan-tulisan kuno dijaga dan dilestarikan agar bisa dinikmati dengan bantuan paleografi. Paleografi menurut Kamus Besar Bahasa Indonesia adalah ilmu tentang tulisan kuno. Secara etimologis, kata paleografi berasal dari kata *paleos* (kuno) dan *grafein* (menulis, tulisan). Paleografi berkaitan erat dengan epigrafi, yaitu ilmu yang mempelajari aksara-aksara yang dipakai sebagai sarana penulisan prasasti.

Secara umum paleografi merupakan kajian tentang tulisan-tulisan kuno, termasuk ilmu membaca, menentukan waktu (tanggal), dan menganalisis tulisan-tulisan kuno yang ditulis di atas papyrus, tablet-tablet tanah liat, tembikar, kayu, perkamen (vellum) kertas, dan lontar (daun enau). Di Indonesia sendiri, seperti dikutip dari buku keluaran Departemen Pendidikan dan Kebudayaan RI yang berjudul *Pelestarian dan Modernisasi Aksara Daerah: Perkembangan Metode dan Teknik Menulis Aksara Jawa (1996/1997)*, penelitian paleografi diawali oleh AB Cohen Stuart pada tahun 1875 melalui buku "Kawi Oorkonden in Fasimile, Mer Inleiding en Transcriptie". Penelitian lainnya secara berturut-turut dilakukan oleh Kern (1917), JG de Casparis (1975) Brandes, Stein Callenfels, Krom, Bosch, Stutterheim, Poerbatjaraka, Pigeaud, Buchari, Ktut Ginarsa, AS Wibowo dan Sukarto K. Atmodjo.

Dalam dunia kearsipan, keahlian paleografi juga diperlukan untuk melacak waktu dan tempat asal tulisan kuno ketika naskah/arsip tersebut dihasilkan. Hal ini dikarenakan salah satu tugas utama paleografi dalam penelusuran adalah menentukan waktu penciptaan naskah/arsip. Penentuan waktu dan lokasi tersebut sangat krusial dalam menentukan peristiwa sejarah yang sedang terjadi. Tulisan tangan biasanya akan menunjukkan sebuah gaya penulisan tertentu, yang dapat dilacak menggunakan pengaruh tertentu yang dapat diidentifikasi waktunya. Dalam bukunya *Filologi Nusantara: Pengantar ke Arah Penelitian Filologi*, Harahap (2021:100) menyatakan bahwa dalam penelusuran paleografi, peneliti atau arsiparis bisa melakukan pengamatan terhadap anatomi tulisan, seperti ukuran, bahan, tinta, panjang, dan jarak baris dalam tulisan.

Selain penentuan waktu dan lokasi, tugas lain paleografi adalah penentuan tujuan penciptaan naskah/arsip yang juga sangat penting. Tujuan arsip tersebut diciptakan biasanya mengikuti fenomena sejarah yang melingkupinya, termasuk tren penggunaan gaya tulisan tertentu, penambahan gambar, dan sebagainya. Perubahan-perubahan yang muncul dalam arsip juga dapat menentukan perubahan sosial, ekonomi dan material hingga politik. Karena itulah paleografi juga semakin berkembang dengan penggunaan teknologi digital yang dapat membaca arsip kuno dengan lebih mudah. Teknologi tersebut dimaksudkan untuk mengetahui teks/naskah masa lalu memanfaatkan paleografi dengan sistem digital. Sistem ini memungkinkan peneliti untuk mengenali teks/naskah kuno secara langsung, baik isi, tujuan, dan lokasinya.

Teknologi ini dikenal dengan istilah *Handwriting Text Recognition*

(HTR) atau pengenalan teks tulisan tangan merupakan kemampuan komputer untuk mengenal teks tulisan dan memudahkan kita membaca tulisan dari arsip atau dokumen lama. Penggunaan komputer untuk memproses dan mencari gambar dari dokumen-dokumen bersejarah dengan menggunakan HTR ini sangat membantu peneliti, arsiparis, mahasiswa, institusi serta masyarakat umum untuk mengetahui masa lalu melalui teks tulisan tersebut.

Penelitian yang dilakukan semakin terbantu dengan keberadaan platform menggunakan HTR yang dapat diakses kapan pun, yakni Transkribus. Transkribus merupakan sebuah program *artificial intelligence* (AI) yang bisa mengenali naskah, transkripsi dan mencari dokumen historis, dari mana pun, kapan pun, dan dengan bahasa apa pun. Hal ini dapat mendukung pekerjaan atau penelitian yang terkait dengan dokumen historis.

Tidak hanya sebatas pada penelitian serta bidang kearsipan dan perpustakaan, Transkribus juga dapat digunakan untuk bidang pendidikan, ilmu pengetahuan, dan juga bisnis. Pada sektor bisnis, pemanfaatan AI ini dapat digunakan untuk analisis tata letak dan pemahaman dokumen, serta bekerja dengan dokumen dalam jumlah besar (baik dokumen historis maupun dokumen cetak).

Untuk kearsipan khususnya, Transkribus menawarkan beberapa hal yang dapat dimanfaatkan oleh pengguna Transkribus, mulai dari digitasi, pengenalan teks berbasis AI, transkripsi, hingga pencarian dokumen bersejarah – dari mana saja, dalam Bahasa apa saja, dan dari abad berapa pun. Penggunaan HTR dalam menganalisis sebuah gambar tulisan sangat berkontribusi dalam kearsipan karena memudahkan semua kalangan untuk membaca, mentranskripsi, dan menggali informasi dari dokumen

historis.

Dalam website readcoop.eu, disebutkan banyak fitur Transkribus yang dapat dimanfaatkan oleh pengguna, di antaranya:

- Pengguna dapat melatih model/pola pengenalan teks tertentu yang bisa mengenali jutaan dokumen yang ditulis, diketik, atau dicetak.
- Pengguna dapat memanfaatkan pengenalan teks tulisan bertenaga AI, analisis tata letak, dan pengenalan struktur.
- Pengguna dapat menggunakan editor transkripsi untuk menyalin dokumen historis secara manual atau bekerja dengan pengenalan bertenaga AI yang menggunakan model AI publik atau model yang dapat pengguna latih sendiri.
- Pengguna dapat mencari dokumen dengan pilihan pencarian lanjutan seperti alat pendeteksi kata kunci.
- Pengguna dapat berkolaborasi melalui dokumen, atur dalam koleksi dan manfaatkan pekerjaan satu sama lain.
- Pengguna dapat membagikan pekerjaannya dengan orang lain melalui portal baca dan pencarian atau pengguna dapat mengubah dokumen menjadi format
- Semua konten di Transkribus, yakni foto yang pengguna unggah, teks yang dikenali, model pengenalan yang terlatih, dan metadata yang sudah diunggah, berada di bawah kendali Uni Eropa dan Transkribus menyatakan mematuhi *General Data Protection Regulation* (GDPR).

Melalui teknologi HTR yang dapat diakses kapan pun dan di mana pun ini, pengguna dapat memperkaya bahan dengan metadata dan saling berkolaborasi. Dengan kemampuan teknologi yang semakin canggih, akses untuk menjelajah masa lalu melalui teks mudah dilakukan. ■