
Standar dan Kriteria Piranti Lunak ICR (Intelligent Character Recognition) ver. 1.2

Oleh: Tim Review TI KPU

1. Pendahuluan

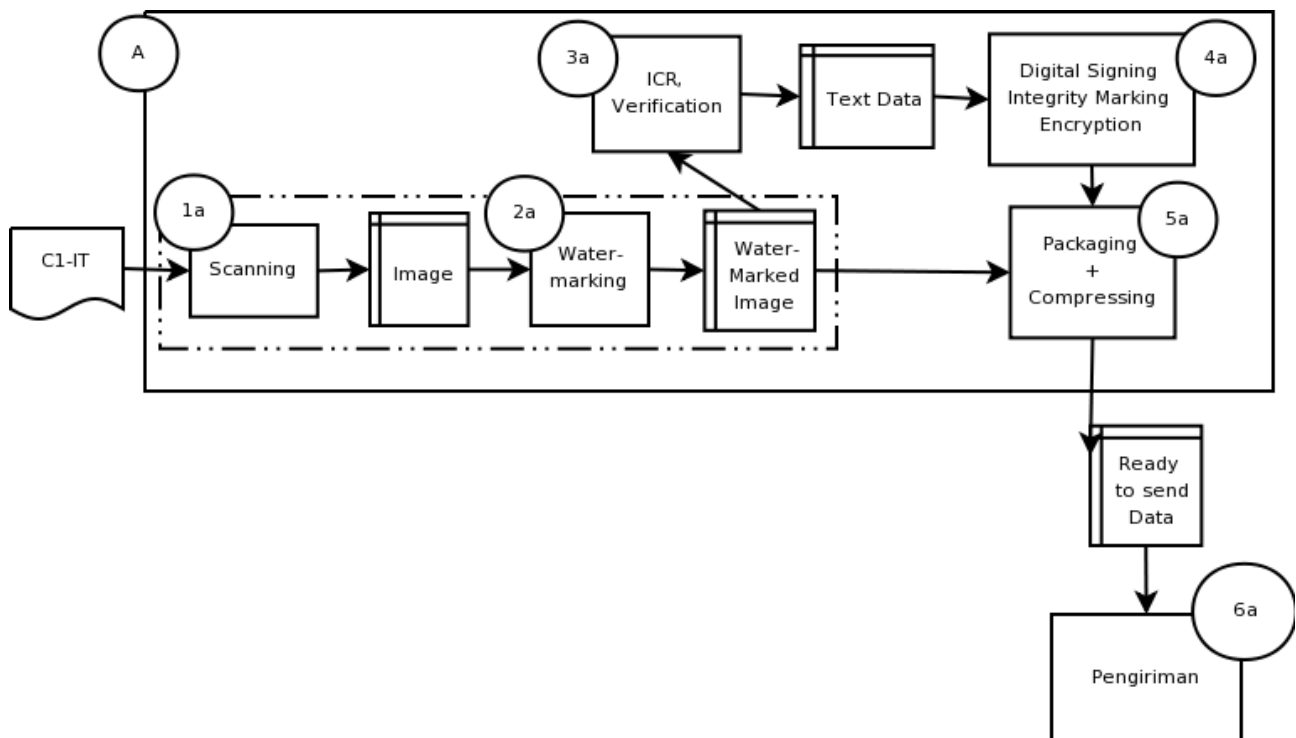
Untuk membantu proses penghitungan suara pada Pemilu 2009, dimanfaatkan mesin pemindai berbasis teknologi Intelligent Character Recognition (ICR) dalam input data hasil penghitungan sebelum dikirim ke sistem tabulasi KPU Pusat. Adapun tujuan utama yang ingin dicapai dengan penggunaan mesin pemindai yang berbasis ICR tersebut adalah:

1. Mempercepat proses perhitungan suara
2. Memperoleh salinan dokumen elektronik yang otentik dan teramankan

Agar ke-2 tujuan utama tersebut tercapai maka perlu ditetapkan standar dan kriteria yang harus dicapai oleh piranti pendukung, khususnya piranti lunak ICR (*Intelligent Character Recognition*)

2. Gambaran Umum Proses

Secara sederhana proses-proses penyalinan data ke dokumen elektronik dan penyiapan berkas data untuk dikirim ke sistem integrator seperti gambar di bawah.



Gambar 1. Skema Pemrosesan Form C1-IT DPR

Ada 5 (lima) proses utama dimulai dari otorisasi operator, proses pemindaian dokumen form C1-TI menghasilkan dokumen elektronik, pembacaan karakter hasil pemindaian (oleh ICR), koreksi dan verifikasi hasil ICR, dan penyiapan berkas untuk pengiriman

3. Standar & Kriteria Proses

1. Proses Otorisasi

- i. Tersedia fasilitas untuk otentikasi dan otorisasi operator yang berwenang
- ii. Tersedia isian kode Provinsi (PP), Kota/Kabupaten (KK), Kecamatan (CC), Desa/Kelurahan (DD), TPS (TT).
- iii. Isian kode Propinsi, Kota/Kabupaten, dan Kecamatan dilakukan secara otomatis (kode lokasi TPS sudah didefinisikan oleh KPU sampai tingkat kecamatan)

2. Proses Pemindaian

- i. Tersedia utility untuk melakukan penyesuaian (adjustment) default master template dengan lembar cetakan form dokumen C1-TI
- ii. Proses pemindaian berkas form C1-TI dilakukan per-TPS (form C1-TI terdiri dari

8 lembar)

- iii. Hasil pemindaian berupa: 8 (delapan) file citra dalam format png (300 dpi, black-white). Dari tiap lembar form C1-TI dihasilkan 1 file citra.
- iv. Masing-masing file hasil pemindaian dibubuhkan langsung watermarking untuk mencegah dan mengamankan dari usaha modifikasi file citra tsb. Lihat Lampiran 1 untuk proses watermarking.
- v. Format penamaan file hasil pemindaian berupa 10 (sebelas) digit gabungan kode Provinsi, Kota/Kab, Kecamatan, Desa/Kelurahan, TPS diikuti identifikasi nomor urut lembar dari dokumen C1-TI. Contoh sebagai berikut:

PPKKCCDDTT_a.png

Keterangan:

PP: 2 digit kode provinsi	a: lembar pertama dokumen C1-TI
KK: 2 digit kode kota/kab	b: lembar ke-2 dokumen C1-TI
CC: 2 digit kode kecamatan	c: lembar ke-3 dokumen C1-TI
DD: 2 digit kode desa/kelurahan
TT: 2 digit kode TPS	h: lembar ke-8 dokumen C1-TI

3. Proses Pembacaan Karakter

- i. Untuk menjaga konsistensi penamaan file citra sesuai nomor urut dari dokumen C1-TI, maka penamaan file tidak berdasarkan urutan pemindaian oleh scanner, tetapi menggunakan pembacaan identifier yang unik dari tiap-tiap lembar form dokumen C1-TI. Misal: lembar pertama dengan kotak kode lokasi TPS pada kanan atas; lembar ke-2 sd ke-8 dengan nomer partai kiri atas.
- ii. Hasil pembacaan karakter tersimpan dalam betuk teks file yang *readable* untuk digunakan dalam proses selanjutnya (koreksi dan penyimpanan dalam format .XML atau .JSON)
- iii. Tersedia *log-file* untuk tiap hasil pembacaan karakter yang digunakan untuk proses koreksi

4. Koreksi dan Verifikasi

- i. Memiliki kemampuan untuk menampilkan hasil pemindaian dan hasil

pembacaan karakter secara bersamaan untuk memudahkan koreksi kesalahan (*error*).

- ii. Dua proses yakni: koreksi dan verifikasi terpenuhi. Proses koreksi fokus kepada memastikan kesamaan antara hasil pembacaan karakter dengan hasil pemindaian. Sementara proses verifikasi fokus kepada bahwa keseluruhan obyek telah terkoreksi dan kelengkapan file hasil ICR terpenuhi
- iii. Memiliki kemampuan untuk pemisahan (*segregation*) antara proses pemindaian dengan proses koreksi dan verifikasi. Contoh: Pemindaian dilakukan pada PC yang terhubung dgn scanner, Koreksi dan Verifikasi dapat dilakukan di PC yang lain dengan jumlah yang lebih banyak.
- iv. Setiap koreksi terhadap hasil pembacaan karakter terekam pada log-file. Informasi log-file untuk setiap perubahan mencakup: time stamp, operator, supervisor, < semua kejadian perubahan – sebelum dan sesudah >
- v. File hasil ICR setelah dikoreksi dan diverifikasi disimpan dalam format xml dengan format penamaan: PPKKCCDDTT.xml dan terenkripsi. Lihat Lampiran 1 untuk proses enkripsi, Lampiran 2 Untuk format .xml ataupun .json
- vi. Format penamaan log-file PPKKCCDDTT.log. Log-file terakhir (menyertai file hasil ICR yang telah dikoreksi dan diverifikasi) terenkripsi.

5. **Penyiapan Berkas Pengiriman**

- i. Berkas file untuk pengiriman dalam bentuk zip file yang terdiri dari:
 - ii. - 8 file citra hasil pemindaian dan telah dibubuhi watermarking (refer ke 2.v)
 - iii. - 1 file hasil ICR dalam format xml dan terenkripsi (refer ke 4.v)
 - iv. - 1 file log-file terenkripsi (refer ke 4.vi)
- v. Format penamaan berkas yang akan dikirim: PPKKCCDDTT.zip