ULIL ALBAB : Jurnal Ilmiah Multidisiplin
Vol.5, No.11, Oktober 2026

Penerapan *Extreme Programming* (XP) dalam Pengembangan Sistem Pengolahan Rekomendasi Hasil Psikotes Terintegrasi Berbasis Web

Arip Priyanto[1], Tumini[2], Agmawarnida[3]
[1,2,3]Universitas Panca Sakti Bekasi, Indonesia
E-mail: arip.andaresta@gmail.com[1]

Article History:
Received: 01 Oktober 2026
Revised: 10 Oktober 2026
Accepted: 12 Oktober 2026

Kata kunci: *sistem rekomendasi; psikotes berbasis web; Extreme Programming; integrasi psikotes; rekrutmen karyawan.*

Abstrak: *Pengelolaan hasil psikotes dalam proses rekrutmen membutuhkan sistem yang mampu mengintegrasikan berbagai instrumen penilaian secara konsisten, efisien, dan dapat ditelusuri. Namun, pengolahan hasil psikotes yang terpisah berpotensi meningkatkan kompleksitas administrasi dan menyulitkan penyusunan rekomendasi kandidat. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi hasil psikotes berbasis web yang mengintegrasikan instrumen Kraepelin, Myers-Briggs Type Indicator (MBTI), dan Dominance, Influence, Steadiness, Compliance (DISC) pada PT Timur Satria Perkasa. Penelitian menggunakan pendekatan penelitian dan pengembangan dengan metode Extreme Programming (XP) melalui tahapan perencanaan, perancangan, pengkodean, dan pengujian secara iteratif. Data kualitatif digunakan untuk mengidentifikasi kebutuhan sistem, sedangkan data kuantitatif deskriptif digunakan untuk mengevaluasi fungsionalitas dan konsistensi perhitungan menggunakan kasus simulatif. Hasil penelitian menunjukkan bahwa sistem mampu mengintegrasikan pengolahan ketiga instrumen, menghasilkan skor dan kategori rekomendasi, serta menyediakan fitur pengelolaan data, pelaporan, konfigurasi parameter, dan pencatatan validasi profesional. Pengujian black-box menunjukkan bahwa seluruh fungsi utama berjalan sesuai dengan skenario yang ditetapkan, sedangkan pemeriksaan terhadap data simulatif memperlihatkan kesesuaian hasil perhitungan sistem dengan lembar skoring acuan. Temuan ini menunjukkan bahwa penerapan XP mendukung pengembangan purwarupa sistem psikotes yang terintegrasi, adaptif, dan memiliki konsistensi perhitungan pada skenario pengujian. Penelitian ini memberikan kontribusi praktis terhadap digitalisasi pengelolaan hasil psikotes sebagai instrumen pendukung keputusan rekrutmen, dengan tetap menempatkan interpretasi psikologis dan keputusan akhir pada tenaga profesional. Implementasi operasional*

memerlukan validasi lebih lanjut terhadap ketepatan rekomendasi, keamanan informasi, dan keandalan penyimpanan data.

PENDAHULUAN

Psikotes berfungsi sebagai sumber informasi tambahan mengenai karakteristik kandidat yang relevan dengan kebutuhan pekerjaan, mencakup aspek kemampuan, kecenderungan kepribadian, dan perilaku kerja (Society for Industrial and Organizational Psychology [SIOP], 2018). Pada PT. Timur Satria Perkasa, proses seleksi memanfaatkan tiga instrumen sekaligus, yaitu *Kraepelin, Myers-Briggs Type Indicator* (MBTI), dan *Dominance*, *Influence*, *Steadiness*, *Compliance* (DISC). Ketiga instrumen memiliki fungsi berbeda: Kraepelin membaca aspek kerja seperti kecepatan, ketelitian, kestabilan, dan ketahanan; MBTI menggambarkan kecenderungan preferensi kepribadian; sedangkan DISC menggambarkan pola perilaku dalam situasi kerja (Marston, 1928; Myers et al., 2018; Wibowo et al., 2022). Karena keluarannya tidak identik, hasil ketiga instrumen tersebut selama ini diproses melalui jalur yang terpisah, sehingga petugas *Human Resource* (HR) harus memeriksa, memindahkan, dan merangkum data secara manual pada beberapa tahapan. Kondisi ini membuat pekerjaan administratif berulang dan membatasi keterlacakan dasar rekomendasi yang dihasilkan.

Kesenjangan tersebut diperlakukan sebagai masalah pengembangan sistem, bukan sebagai bukti bahwa instrumen psikologis yang digunakan tidak valid. Fokus penelitian adalah menyediakan alat bantu yang menata masukan, perhitungan, penyimpanan, pencarian, pelaporan, dan penelaahan hasil secara lebih konsisten, tanpa melakukan standardisasi ulang instrumen maupun pengujian validitas prediktif terhadap kinerja kandidat (SIOP, 2018; International Organization for Standardization & International Electrotechnical Commission [ISO/IEC], 2023).

Extreme Programming (XP) dipilih sebagai kerangka pengembangan karena mendukung iterasi pendek dan perbaikan setelah suatu fungsi diuji (Beck, 1999; Dybå & Dingsøyr, 2008). Beck (1999) menempatkan komunikasi, kesederhanaan, umpan balik, keberanian melakukan perubahan, dan disiplin praktik sebagai unsur penting XP, sehingga rancangan dapat dibuat secukupnya untuk kebutuhan yang sedang dikerjakan dan disempurnakan berdasarkan temuan pengujian. Kajian sistematis Dybå dan Dingsøyr (2008) memperlihatkan bahwa pendekatan agile banyak dibahas dalam kaitannya dengan adaptasi, kolaborasi, dan siklus feedback yang lebih cepat, sedangkan Dingsøyr et al. (2012) menempatkan respons terhadap perubahan sebagai karakter penting metodologi agile. Dikert et al. (2016) mengingatkan bahwa penerapan agile pada skala organisasi dapat menghadapi persoalan koordinasi dan konsistensi praktik, sehingga penelitian ini membatasi klaim XP pada skala prototipe dan tidak menyamakannya dengan transformasi agile organisasi. Pressman dan Maxim (2020) menempatkan pengujian sebagai bagian yang diperlukan untuk memastikan perangkat lunak berperilaku sesuai spesifikasi, prinsip yang diterapkan melalui *Black-Box Testing* dan pemeriksaan hasil perhitungan pada dataset simulatif (ISO/IEC, 2023).

Pada konteks yang lebih dekat dengan objek penelitian, Jamil et al. (2024) menerapkan XP pada pengembangan sistem seleksi karyawan berbasis web menggunakan kerangka server dan basis data MySQL. Surianti dan Wijaya (2017) merancang sistem pakar psikotes untuk satu alat ukur, sedangkan Wibowo et al. (2022) mengklasifikasikan kepribadian calon karyawan menggunakan certainty factor dan forward chaining. Penelitian ini berbeda dari ketiganya karena berfokus pada pengolahan terpadu keluaran tiga instrumen sekaligus (Kraepelin, MBTI, dan DISC) melalui satu nilai integrasi berproporsi bobot, disertai ruang penelaahan profesional yang eksplisit di dalam sistem.

Berdasarkan uraian tersebut, penelitian ini mengembangkan purwarupa sistem berbasis web yang mengintegrasikan tiga instrumen psikotes menggunakan XP sebagai kerangka teknis, kemudian menguji konsistensi perhitungan dan fungsi aplikasi melalui *Black-Box Testing* dan pemeriksaan 50 kasus simulatif (Beck, 1999; Pressman & Maxim, 2020). Tujuan penelitian dirumuskan untuk: (1) menganalisis kondisi awal dan kebutuhan pengguna; (2) merancang dan mengembangkan model integrasi Kraepelin, MBTI, dan DISC menggunakan XP; (3) menganalisis kelayakan model secara teoretik dan empiris awal; serta (4) menguji, mengevaluasi, dan merevisi model melalui *Black-Box Testing* dan pemeriksaan konsistensi hasil. Sistem diposisikan sebagai alat bantu HR, bukan pengganti kewenangan psikolog atau tenaga profesional dalam interpretasi dan keputusan akhir rekrutmen (SIOP, 2018).

METODE PENELITIAN

Desain dan Pendekatan Penelitian

Penelitian ini merupakan penelitian pengembangan (*development research*) dengan luaran utama berupa purwarupa sistem berbasis web (Richey & Klein, 2007). Data kualitatif deskriptif digunakan untuk memahami proses yang berjalan dan merumuskan kebutuhan, sedangkan data kuantitatif deskriptif digunakan untuk mengevaluasi fungsi program dan konsistensi hasil perhitungan pada dataset simulatif (Creswell & Creswell, 2018; Wohlin et al., 2012). Kerangka penelitian terdiri atas penelitian pendahuluan, analisa kebutuhan, rancangan model, serta validasi, evaluasi, dan revisi; *Extreme Programming* (XP) berperan sebagai mekanisme teknis pembangunan perangkat lunak di dalam kerangka tersebut melalui *Planning*, *Design*, *Coding*, dan *Testing* yang berulang (Beck & Andres, 2004; Pressman & Maxim, 2020).

Tempat, Waktu, dan Sasaran Penelitian

Penelitian dilaksanakan pada konteks pemrosesan hasil psikotes di PT. Timur Satria Perkasa selama dua minggu pada Mei 2026. Sasaran penelitian dibagi menjadi pengguna, validator, proses yang dibangun, dan data uji sebagaimana dirangkum pada Tabel 1. Pengujian diterapkan pada skenario kandidat Operator HD 785 menggunakan dataset simulatif berisi 50 kasus dengan skor Kraepelin, MBTI, dan DISC yang lengkap; nama, identitas, dan skor dibuat untuk kebutuhan simulasi, bukan hasil asesmen individu nyata (Wohlin et al., 2012).

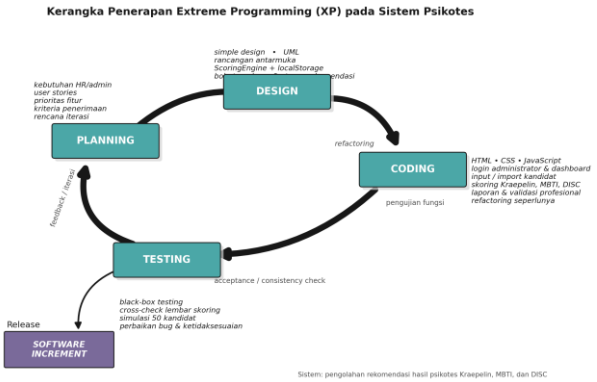
Tabel 1. Karakteristik Sasaran Penelitian

Sasaran	Karakteristik / Peran
HR/Administrator	Mengelola kandidat, skor, dan pelaporan
Psikolog/Pakar	Kompetensi interpretasi psikotes; validator bobot dan hasil
Proses rekrutmen	Pengolahan Kraepelin, MBTI, DISC (objek pengembangan)
Data uji	50 kasus kandidat Operator HD 785 (simulatif)

Kerangka *Extreme Programming* (XP)

Siklus XP diringkas menjadi empat tahap yang berulang: Planning menetapkan kondisi awal, kebutuhan pengguna, dan prioritas fitur; Design menerjemahkan kebutuhan menjadi alur proses, struktur data, antarmuka, dan *Unified Modeling Language* (UML); Coding merealisasikan rancangan menjadi purwarupa berbasis HTML, CSS, dan JavaScript; dan Testing memeriksa fungsi aplikasi melalui *Black-Box Testing* serta membandingkan hasil pengolahan dengan lembar skoring (Beck & Andres, 2004; Pressman & Maxim, 2020). Temuan pada Testing menjadi *feedback* yang dapat mengubah prioritas *Planning*, menuntut perbaikan Design, atau mengharuskan revisi Coding sebelum versi berikutnya diuji kembali, sebagaimana diilustrasikan

pada Gambar 1 (Beck & Andres, 2004).



Gambar 1. Kerangka Penerapan *Extreme Programming* (XP) pada Sistem Psikotes

Kriteria Operasional dan Bobot Integrasi Skor

Kriteria operasional memuat cara pengolahan tiga instrumen serta parameter integrasi. Skor Kraepelin dihitung sebagai rata-rata empat aspek (kecepatan, ketelitian, kestabilan, dan ketahanan) pada skala 0-100. Skor MBTI dibentuk dari keseimbangan preferensi *Extraversion-Introversion* serta kontribusi *Sensing*, *Thinking*, dan *Judging* melalui persamaan *Balance* = 100 - |E - 50| x 1,2 dan Skor M = (Balance x 0,15) + (S x 0,30) + (T x 0,25) + (J x 0,30). Skor DISC dibentuk dari jarak profil kandidat terhadap profil target simulatif jabatan Operator HD 785 (D=20, I=15, S=30, C=35) melalui persamaan Skor D = 100 - 0,50 x (|D-20| + |I-15| + |S-30| + |C-35|), dengan nilai dibatasi pada rentang 0-100. Ketiga skor tersebut kemudian digabungkan menjadi nilai integrasi melalui persamaan Skor Integrasi = (0,50 x Kraepelin) + (0,20 x MBTI) + (0,30 x DISC), sebagaimana dirangkum pada Tabel 2. Seluruh kriteria, target profil, proporsi bobot, dan batas nilai merupakan parameter operasional purwarupa untuk kebutuhan simulasi dan memerlukan validasi psikolog atau tenaga profesional yang berwenang sebelum digunakan pada seleksi aktual (Society for Industrial and Organizational Psychology [SIOP], 2018).

Tabel 2. Kriteria Operasional dan Bobot Integrasi Model

Instrumen	Perhitungan Skor (ringkas)	Bobot
Kraepelin	Rata-rata 4 aspek (0-100)	0,50
MBTI	Balance & preferensi terboboti	0,20
DISC	Jarak ke profil target jabatan	0,30

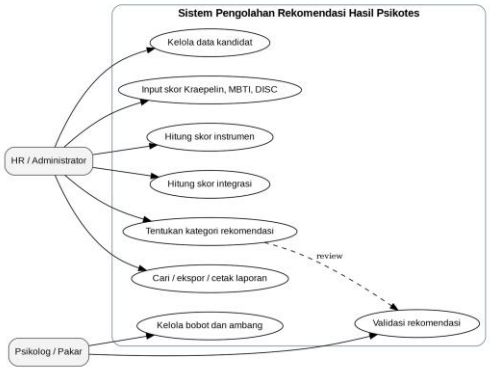
Nilai integrasi selanjutnya dipetakan ke empat kelas rekomendasi (Tabel 3): >=85 Sangat Direkomendasikan; 75-<85 Direkomendasikan; 65-<75 Dipertimbangkan; dan <65 Tidak Direkomendasikan. Kriteria tersebut merupakan parameter simulasi dan bukan norma psikologis universal (SIOP, 2018).

Tabel 3. Kriteria Rekomendasi Berdasarkan Skor Integrasi

Rentang Skor Integrasi	Kategori Rekomendasi
>= 85	Sangat Direkomendasikan
75 - < 85	Direkomendasikan
65 - < 75	Dipertimbangkan
< 65	Tidak Direkomendasikan

Perancangan Sistem

Rancangan sistem dituangkan melalui Use *Case Diagram, Activity Diagram, Sequence Diagram,* dan *Class Diagra*m (Object Management Group [OMG], 2017). *Use Case Diagram* (Gambar 2) memisahkan peran HR/Administrator, yang mengelola data, memasukkan skor, menjalankan perhitungan, mencari, dan mengekspor/mencetak laporan, dari peran Psikolog/Pakar, yang menelaah konfigurasi bobot-ambang serta memvalidasi rekomendasi. Arsitektur sistem memisahkan WebAppController, AuthManager, ScoringEngine, ScoringConfig, LocalStorageRepository, ImportService, ReportService, Candidate, dan ExpertValidation, sehingga perubahan antarmuka atau fungsi pendukung tidak mengubah formula skoring utama (Pressman & Maxim, 2020).



Gambar 2. Use Case Diagram Sistem Pengolahan Rekomendasi Hasil Psikotes

Implementasi

Rancangan direalisasikan menjadi purwarupa berbasis HTML, CSS, dan JavaScript. Fitur yang diimplementasikan meliputi login administrator, pengelolaan kandidat, masukan skor Kraepelin/MBTI/DISC secara manual maupun impor massal (CSV, Excel, dan salin-tempel), perhitungan skor otomatis, pencarian data, ekspor CSV, pencetakan laporan, pengaturan parameter bobot dan ambang, perubahan password, serta pencatatan penelaahan profesional. Data kandidat, konfigurasi, hash password, dan catatan validator disimpan pada localStorage browser, sedangkan sessionStorage menandai status sesi login (World Wide Web Consortium [W3C], 2021), sehingga satu fungsi skoring yang sama digunakan untuk masukan manual maupun impor massal guna menghindari duplikasi logika (Pressman & Maxim, 2020).

Teknik Pengujian dan Evaluasi

Evaluasi teknis dilakukan melalui dua jalur. Pertama, *Black-Box Testing* digunakan untuk memeriksa sepuluh fungsi utama aplikasi (login, perhitungan skor, penyimpanan, pencarian, impor massal, ekspor CSV, pencatatan validasi, pengaturan parameter, ubah password, dan logout/login ulang) dari sudut pandang masukan-keluaran (International Software Testing Qualifications Board [ISTQB], 2023; Pressman & Maxim, 2020). Kedua, nilai integrasi dan kategori rekomendasi program dibandingkan dengan lembar skoring pada 50 kasus simulatif sampai batas pembulatan, menggunakan parameter bobot dan ambang yang telah diverifikasi sama antara kedua sisi (Wohlin et al., 2012). Statistik deskriptif (rata-rata, minimum, maksimum, median, dan simpangan baku) dimanfaatkan untuk menggambarkan karakteristik dataset uji (Creswell & Creswell, 2018). Karena dataset bersifat simulatif dan validasi pakar aktual belum tersedia, hasil evaluasi hanya digunakan

untuk menyatakan konsistensi implementasi dan kelayakan teknis awal, bukan validitas psikologis, akurasi seleksi, atau efektivitas organisasi (SIOP, 2018; Wohlin et al., 2012).

HASIL DAN PEMBAHASAN

Implementasi Antarmuka dan Alur Sistem

Purwarupa akhir menyajikan alur login administrator, dashboard ringkasan kategori, input skor, halaman hasil, data kandidat, impor massal, validasi pakar, pengaturan parameter, dan ubah password. Gambar 3 menunjukkan halaman Hasil setelah proses skoring dijalankan: sistem menampilkan skor Kraepelin, MBTI, dan DISC sebelum skor integrasi, tipe MBTI dan profil DISC sebagai informasi deskriptif, kategori rekomendasi, kelebihan, area pengembangan, saran tindak lanjut, dan kesimpulan yang selalu disertai pernyataan bahwa hasil merupakan keluaran teknis prototipe dan memerlukan validasi profesional (Society for Industrial and Organizational Psychology [SIOP], 2018). Pemisahan ini penting agar label tipe kepribadian tidak disalahartikan sebagai aturan penerimaan kandidat (Zárate-Torres & Correa, 2023).



Sumber: Hasil pengembangan aplikasi, 2026

Gambar 3. Tampilan Hasil Perhitungan dan Interpretasi Skor

Statistik Deskriptif Dataset Simulasi

Pengujian menggunakan 50 kasus kandidat Operator HD 785 menghasilkan nilai integrasi rata-rata 74,32, median 74,85, simpangan baku 7,57, dengan rentang 52,01 sampai 86,65 (Tabel 4). Rata-rata dan median yang berdekatan menunjukkan pusat distribusi dataset uji berada di sekitar batas antara kategori Dipertimbangkan dan Direkomendasikan, sedangkan simpangan baku memperlihatkan variasi yang cukup untuk membawa kasus ke berbagai kategori. Tipe MBTI yang paling sering muncul adalah ESTJ (19 kasus) dan ISTJ (14 kasus), sedangkan profil DISC terbanyak adalah C-S (13 kasus) dan S-C (12 kasus); distribusi ini menggambarkan komposisi dataset uji, bukan prevalensi psikologis populasi karyawan (Creswell & Creswell, 2018; SIOP, 2018).

Tabel 4. Statistik Deskriptif Data Simulasi (n=50)

Aspek	Rata-rata	Min	Maks	SD
Kecepatan	74,35	43,75	92,64	11,54
Ketelitian	75,08	39,61	98,86	12,91
Kestabilan	73,25	41,70	96,50	12,30
Ketahanan	74,84	49,67	97,25	11,57
Skor Integrasi	74,32	52,01	86,65	7,57

Distribusi Kategori Rekomendasi

Distribusi 50 kasus menghasilkan 2 kasus (4%) Sangat Direkomendasikan, 23 kasus (46%) Direkomendasikan, 21 kasus (42%) Dipertimbangkan, dan 4 kasus (8%) Tidak Direkomendasikan (Tabel 5). Dua kategori tengah meliputi 88% dataset, menunjukkan bahwa sebagian besar kasus uji berada pada wilayah yang membutuhkan tindak lanjut, bukan pada label ekstrem, sekaligus memastikan keempat kategori teruji pada batas nilainya. Seluruh kolom rekomendasi manual pada dataset masih kosong dan status validasi tercatat Belum Divalidasi pada seluruh 50 kasus, sehingga tingkat kesesuaian dengan pertimbangan profesional belum dapat dihitung pada tahap ini (SIOP, 2018).

Tabel 5. Distribusi Kategori Rekomendasi pada 50 Kasus Uji

Kategori Rekomendasi	Jumlah	Persentase
Sangat Direkomendasikan	2	4%
Direkomendasikan	23	46%
Dipertimbangkan	21	42%

Hasil *Black-Box Testing* dan Konsistensi Perhitungan

Hasil *Black-Box Testing* pada sepuluh fungsi utama menunjukkan seluruh skenario berjalan sesuai expected result, atau 10 dari 10 fungsi berhasil (100%), sebagaimana dirangkum pada Tabel 6 (Pressman & Maxim, 2020). Pada pengujian konsistensi numerik, 50 dari 50 nilai integrasi dan kategori rekomendasi yang dihasilkan program sama dengan lembar skoring pembanding sampai batas pembulatan, tanpa ditemukan perbedaan skor integrasi maupun kategori. Temuan ini mendukung kesimpulan bahwa fungsi dan formula operasional telah diimplementasikan secara konsisten pada lingkungan pengujian, tetapi tidak digunakan untuk menyatakan bahwa sistem lebih akurat daripada psikolog, valid secara psikometrik, atau mampu memprediksi keberhasilan kerja (SIOP, 2018; Wohlin et al., 2012).

Tabel 6. Ringkasan Hasil *Black-Box Testing* Aplikasi

Fungsi	Skenario Uji	Status
Login admin	Kredensial benar/salah	Berhasil
Hitung skor	50 kasus vs skoring	50/50
Simpan data	Simpan & buka lagi	Berhasil
Cari data	Cari nama/kode	Berhasil
Impor massal	Baris valid/tidak	Berhasil
Ekspor CSV	Ekspor kandidat	Berhasil
Validasi pakar	Simpan rekomendasi	Berhasil
Parameter	Ubah bobot/ambang	Berhasil
Ubah password	Lama & baru valid	Berhasil
Logout/login	Tutup & buka sesi	Berhasil

Pembahasan

Pengembangan model didukung oleh kejelasan kesenjangan proses (tiga hasil psikotes yang masih diproses terpisah), struktur data dan formula yang transparan, arsitektur web yang fleksibel, serta mekanisme iterasi teknis XP yang memungkinkan penambahan impor massal, autentikasi, dan pengaturan parameter sebagai respons terhadap kebutuhan yang berkembang selama pengembangan (Beck & Andres, 2004; Dybå & Dingsøyr, 2008). Kekuatan model terutama terletak pada kemampuannya menggabungkan tiga sumber informasi tanpa menghilangkan komponen asalnya: pengguna tetap dapat menelusuri skor Kraepelin, informasi MBTI, profil

DISC, nilai integrasi, kategori, area pengembangan, saran, dan catatan validator, sehingga hasil akhir memiliki jejak data yang dapat diperiksa (Pressman & Maxim, 2020). Pemisahan rekomendasi sistem dari rekomendasi manual juga membuka ruang bagi penelitian lanjutan untuk mengukur agreement dan menganalisis kasus ketidaksesuaian secara transparan (SIOP, 2018).

Di sisi lain, kelemahan model berasal dari parameter yang masih bersifat simulatif dan penyederhanaan konstruk psikologis menjadi skor operasional 0-100, yang bermanfaat untuk implementasi dan pengujian program tetapi belum didukung kalibrasi psikometrik terhadap outcome kerja (Leutner et al., 2023; SIOP, 2018). Arsitektur localStorage dan sessionStorage pada sisi browser juga membatasi sinkronisasi multiuser, backup terpusat, dan audit log, sedangkan autentikasi satu akun administrator di sisi klien belum setara dengan keamanan aplikasi produksi (World Wide Web Consortium [W3C], 2021). Faktor penghambat lain adalah waktu pengembangan yang singkat, yang membatasi jumlah iterasi serta kedalaman uji usability, uji performa pada volume data besar, dan pengujian keamanan sistematis; kondisi ini perlu dipertimbangkan ketika menafsirkan hasil 100% pada skenario black-box, yang hanya mencakup sepuluh skenario fungsional yang secara eksplisit diuji (Pressman & Maxim, 2020; Wohlin et al., 2012).

Dibandingkan dengan Surianti dan Wijaya (2017) yang berfokus pada satu alat ukur, penelitian ini memperluas ruang pengolahan menuju integrasi tiga keluaran instrumen sekaligus. Berbeda dari Wibowo et al. (2022) yang menggunakan certainty factor dan forward chaining, penelitian ini memilih pendekatan deterministik berbasis ambang nilai integrasi, yang menguntungkan dari sisi reproduksibilitas perhitungan tetapi tidak berupaya merepresentasikan penalaran pakar yang kompleks. Dibandingkan dengan Jamil et al. (2024) yang menerapkan XP pada arsitektur server dan basis data MySQL untuk seleksi karyawan, purwarupa ini menggunakan penyimpanan lokal berbasis browser yang mempermudah distribusi demonstrasi tetapi membatasi penggunaan multiuser dan keamanan data. Leutner et al. (2023) menyediakan standar pembanding yang jauh lebih tinggi untuk evaluasi asesmen digital, dengan pengujian psikometri dan fairness pada ribuan penyelesaian, sehingga kesenjangan skala tersebut menjelaskan mengapa kesimpulan penelitian ini dibatasi pada konsistensi teknis, bukan validitas prediktif. Temuan Zárate-Torres dan Correa (2023) turut memperkuat keputusan agar tipe MBTI tidak dijadikan aturan tunggal keputusan seleksi, sejalan dengan desain sistem yang menempatkan MBTI sebagai informasi pendukung. Sejalan dengan kajian Raghavan et al. (2020) mengenai mitigasi bias pada seleksi berbasis algoritma, penerapan sistem pada seleksi aktual memerlukan analisis fairness lebih lanjut agar konfigurasi bobot dan ambang tidak menimbulkan kerugian yang tidak adil bagi kelompok kandidat tertentu.

Keterbatasan Penelitian

Penelitian ini dibatasi oleh penggunaan 50 kasus simulatif tanpa rekomendasi manual dan validasi pakar aktual, waktu pengembangan yang singkat sehingga hanya mencakup satu pengembang tanpa praktik tim seperti pair programming, pengujian yang baru dilakukan pada satu skenario jabatan (Operator HD 785), serta arsitektur penyimpanan dan autentikasi yang masih bersifat lokal di sisi browser. Karena itu, kesimpulan penelitian terutama berkaitan dengan konsistensi komputasi dan keberhasilan fungsi yang diuji, dan belum dapat digunakan untuk menyatakan bahwa sistem mempercepat proses seleksi, mengurangi bias, meningkatkan ketepatan seleksi, atau memprediksi keberhasilan kerja kandidat (Leutner et al., 2023; Raghavan et al., 2020; SIOP, 2018).

KESIMPULAN

Penelitian ini berhasil mengembangkan purwarupa sistem berbasis web yang mengintegrasikan pengolahan hasil Kraepelin, MBTI, dan DISC melalui *Extreme Programming* (XP), dengan siklus Planning, Design, Coding, dan Testing yang berulang sehingga perubahan kebutuhan dan temuan pengujian dapat ditelusuri kembali ke keputusan rancangan. Hasil pengujian menunjukkan bahwa sepuluh fungsi utama berhasil pada seluruh skenario *Black-Box Testing* (100%), sedangkan 50 dari 50 nilai integrasi dan kategori rekomendasi program konsisten dengan lembar skoring pembanding sampai batas pembulatan. Temuan tersebut mendukung kelayakan teknis awal model dan konsistensi implementasi formula operasional, tetapi belum dapat diperluas menjadi bukti validitas psikologis, akurasi seleksi, atau efektivitas organisasi karena validasi profesional aktual belum tersedia.

Secara praktis, sistem berpotensi membantu HR menata dokumentasi, perhitungan, dan pelaporan hasil psikotes secara lebih terintegrasi, dengan syarat penggunaannya tetap disertai penelaahan psikolog atau tenaga profesional yang berwenang. Penelitian lanjutan disarankan untuk melibatkan pakar guna memperoleh rekomendasi manual dan validasi aktual, memindahkan arsitektur ke basis data server dengan kontrol akses dan audit log, mengukur usability serta waktu proses pada pengguna nyata, dan menganalisis fairness sebelum sistem digunakan pada proses seleksi operasional. Dengan langkah tersebut, *Extreme Programming* dapat terus dimanfaatkan sebagai mekanisme iteratif untuk menyempurnakan sistem pengolahan rekomendasi psikotes tanpa mengambil alih kewenangan profesional dalam keputusan akhir rekrutmen.

REFERENSI

Beck, K. (1999). Extreme programming explained: Embrace change. Addison-Wesley.

Beck, K., & Andres, C. (2004). Extreme programming explained: Embrace change (2nd ed.). Addison-Wesley.

Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE Publications.

Dikert, K., Paasivaara, M., & Lassenius, C. (2016). Challenges and success factors for large-scale agile transformations: A systematic literature review. Journal of Systems and Software, 119, 87–108. https://doi.org/10.1016/j.jss.2016.06.013

Dingsøyr, T., Nerur, S., Balijepally, V., & Moe, N. B. (2012). A decade of agile methodologies: Towards explaining agile software development. Journal of Systems and Software, 85(6), 1213–1221. https://doi.org/10.1016/j.jss.2012.02.033

Dybå, T., & Dingsøyr, T. (2008). Empirical studies of agile software development: A systematic review. Information and Software Technology, 50(9–10), 833–859. https://doi.org/10.1016/j.infsof.2008.01.006

International Organization for Standardization & International Electrotechnical Commission. (2023). Systems and software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—Product quality model (ISO/IEC Standard No. 25010:2023). https://www.iso.org/standard/78176.html

International Software Testing Qualifications Board. (n.d.). Black-box testing. ISTQB Glossary. https://glossary.istqb.org/en_US/term/black-box-testing/2

Jamil, N., Harman, R., Amrizal, A., & Fauzi, R. (2024). Pendekatan sistem pakar forward chaining dengan extreme programming pada seleksi karyawan PT. Enka Mandiri Sukses. Jurnal Desain dan Analisis Teknologi, 3(1), 59–67. https://doi.org/10.58520/jddat.v3i1.53

Leutner, F., Codreanu, S.-C., Brink, S., & Bitsakis, T. (2023). Game based assessments of cognitive

ability in recruitment: Validity, fairness and test-taking experience. Frontiers in Psychology, 13, Article 942662. https://doi.org/10.3389/fpsyg.2022.942662

Marston, W. M. (1928). Emotions of normal people. Harcourt, Brace and Company.

Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (2018). MBTI manual for the Global Step I and Step II assessments (4th ed.). The Myers-Briggs Company.

Object Management Group. (2017). OMG Unified Modeling Language (OMG UML), version 2.5.1. https://www.omg.org/spec/UML/2.5.1

Pressman, R. S., & Maxim, B. R. (2020). Software engineering: A practitioner's approach (9th ed.). McGraw-Hill Education.

Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 469–481). Association for Computing Machinery. https://doi.org/10.1145/3351095.3372828

Richey, R. C., & Klein, J. D. (2007). Design and development research: Methods, strategies, and issues. Lawrence Erlbaum Associates.

Society for Industrial and Organizational Psychology. (2018). Principles for the validation and use of personnel selection procedures (5th ed.). Industrial and Organizational Psychology, 11(S1), 1–97. https://doi.org/10.1017/iop.2018.195

Surianti, S. (2017). Rancangan sistem pakar psikotes untuk penyeleksian penerimaan karyawan pada STMIK TIME Medan. Jurnal TIMES, 6(2), 69–78. https://doi.org/10.51351/jtm.6.2.2017588

Wibowo, G. S., Wahiddin, D., & Kusumaningrum, D. S. (2022). Sistem pakar klasifikasi dan saran pengembangan kepribadian calon karyawan dengan metode certainty factor dan forward chaining. Scientific Student Journal for Information, Technology and Science, 3(2), 190–196. https://journal.ubpkarawang.ac.id/mahasiswa/index.php/ssj/article/view/439

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). Experimentation in software engineering. Springer. https://doi.org/10.1007/978-3-642-29044-2

World Wide Web Consortium. (2016). Web storage (Second Edition). https://www.w3.org/TR/2016/REC-webstorage-20160419/

Zárate-Torres, R., & Correa, J. C. (2023). How good is the Myers-Briggs Type Indicator for predicting leadership-related behaviors? Frontiers in Psychology, 14, Article 940961. https://doi.org/10.3389/fpsyg.2023.940961