
ADAPTASI SKALA LITERASI KECERDASAN ARTIFISIAL (SAIL4ALL) KE
DALAM BAHASA INDONESIA: STUDI PADA KARYAWAN PERUSAHAAN
RETAIL DI INDONESIA

Hendra Hutabarat¹, Inggit Djuningsih², Mela Astri³, Puji Astuti⁴, Sasmiatin Irawan⁵
^{1,2,3,4,5}Universitas Paramadina

Email: hutabarathendrawill@gmail.com¹, inggitdj@gmail.com², melaastri18@gmail.com³,
csukardjo@yahoo.com⁴, sasmiamuchtar@gmail.com⁵

ABSTRAK

Penelitian ini bertujuan untuk mengadaptasi *The Scale of Artificial Intelligence Literacy for All* (SAIL4ALL) ke dalam Bahasa Indonesia dan menguji validitas serta reliabilitasnya dalam konteks tenaga kerja sektor ritel non-teknis. Proses adaptasi dilakukan melalui prosedur *forward-back translation*, validasi pakar, uji keterbacaan, serta analisis psikometrik. Dari 56 item versi asli, 4 item dieliminasi karena ketidaksesuaian konteks, sehingga 52 item diuji lebih lanjut pada 600 responden dan hanya 560 responden yang memenuhi kriteria. Hasil analisis *item-rest correlation* menunjukkan bahwa 42 item memiliki korelasi $\geq 0,4$ dan layak digunakan dalam analisis lanjutan. *Exploratory factor analysis* (EFA) mengungkapkan empat faktor baru: Pemahaman AI, Etika AI, Konsep AI, dan Kapabilitas AI. Namun, hasil *confirmatory factor analysis* (CFA) menunjukkan bahwa model konstruk belum sepenuhnya fit, terutama pada dimensi Pemahaman AI dan Kapabilitas AI. Ketidaksesuaian ini mencerminkan pengaruh perbedaan budaya dan latar belakang pendidikan terhadap pemahaman konstruk literasi AI. Meskipun demikian, alat ukur versi adaptasi menunjukkan konsistensi internal yang tinggi ($\omega = 0,929$) dan memiliki potensi untuk digunakan secara terbatas dalam mengukur literasi AI berbasis fakta. Penelitian ini merekomendasikan pengembangan ulang struktur faktor serta uji validitas eksternal pada populasi yang lebih beragam untuk memperkuat generalisasi skala.

Kata Kunci: SAIL4ALL, Literasi Kecerdasan Artifisial, Adaptasi Alat Ukur, Validitas Konstruk, CFA.

ABSTRACT

This study aimed to adapt the Scale of Artificial Intelligence Literacy for All (SAIL4ALL) into Bahasa Indonesia and to examine its validity and reliability within the context of non-technical retail sector workers. The adaptation process followed standard procedures, including forward-back translation, expert validation, readability testing, and psychometric analysis. Of the 56 original items, 4 were removed due to contextual irrelevance, resulting in 52 items tested on 600 respondents, and only 560 respondents met the criteria. The item-rest correlation analysis identified 42 items with correlations ≥ 0.4 , which were retained for further analysis. Exploratory factor analysis (EFA) revealed a new four-factor structure: AI Understanding, AI Ethics, AI Concepts, and AI Capabilities. However, confirmatory factor analysis (CFA) showed that the overall model did not achieve satisfactory fit, particularly in the AI Understanding and AI Capabilities dimensions. These discrepancies indicate the influence of cultural and

educational differences on the interpretation of AI literacy constructs. Nevertheless, the adapted scale demonstrated high internal consistency ($\omega = 0.929$) and presents initial potential for assessing factual AI literacy in a non-technical Indonesian workforce. The study recommends further refinement of the factor structure and external validation across more diverse populations to enhance generalizability.

Keywords: *SAIL4ALL, Artificial Intelligence Literacy, Scale Adaptation, Construct Validity, CFA.*

PENDAHULUAN

Kecerdasan artifisial (*AI*) telah menjadi salah satu pilar utama dalam transformasi digital yang memengaruhi hampir semua aspek kehidupan manusia, mulai dari sektor industri, pendidikan, hingga layanan publik. *AI* tidak hanya menawarkan efisiensi dan automasi, tetapi juga mengubah cara manusia bekerja, berinteraksi, dan mengambil keputusan. Seiring dengan peningkatan eksponensial penggunaan *AI*, terdapat kebutuhan yang semakin mendesak untuk memastikan bahwa masyarakat umum, termasuk tenaga kerja di sektor non-teknis, memiliki pemahaman yang memadai tentang teknologi ini. Konsep literasi kecerdasan artifisial (*AI literacy*) kemudian muncul sebagai kerangka penting yang mencakup kemampuan individu untuk memahami, mengevaluasi, dan berinteraksi secara kritis dengan sistem *AI* (Long & Magerko, 2020). Literasi *AI* bukan hanya perihal teknis, tetapi juga mencakup aspek etika, sosial, dan kognitif dalam berinteraksi dengan *AI* (Ng, 2023; Ng et al., 2021). Studi oleh Tuomi (2023) juga menegaskan bahwa literasi *AI* adalah komponen fundamental dalam pengembangan kompetensi abad ke-21.

Di tengah tuntutan revolusi industri 5.0 dan transformasi digital, literasi *AI* menjadi landasan penting untuk memberdayakan individu agar dapat berpartisipasi secara aktif dalam ekosistem digital yang semakin kompleks. Beberapa studi menunjukkan bahwa pemahaman publik terhadap *AI* masih rendah dan dipenuhi dengan miskonsepsi, yang dapat menghambat adopsi teknologi secara bijak (West et al., 2019; Ng, 2023). Selanjutnya, survei oleh *World Economic Forum* (2021) menyoroti rendahnya kesiapan tenaga kerja global dalam memahami implikasi etis dan praktis dari penggunaan *AI*. Dalam konteks Indonesia, di mana penetrasi *AI* mulai merambah sektor-sektor operasional seperti ritel dan logistik, kesiapan tenaga kerja dalam memahami dan mengadopsi teknologi ini menjadi semakin krusial (Setiawan & Hidayat, 2022).

Sayangnya, sebagian besar alat ukur literasi *AI* yang ada saat ini masih berfokus pada populasi akademik atau pengguna teknologi tingkat lanjut, serta banyak yang menggunakan pendekatan berbasis persepsi diri (*self-report*), yang cenderung bias (Ng et al., 2021; Yigitcanlar et al., 2023). Djamaluddin et al. (2023) mencatat bahwa metode persepsi diri sering kali tidak mampu mengukur pengetahuan faktual secara akurat. Hal ini menciptakan kesenjangan antara kebutuhan lapangan dengan ketersediaan alat ukur yang valid dan reliabel untuk populasi umum, khususnya tenaga kerja di sektor non-teknis. Salah satu instrumen yang menjanjikan untuk mengatasi keterbatasan tersebut adalah SAIL4ALL (*The Scale of Artificial Intelligence Literacy for all*), yang dikembangkan untuk mengukur literasi *AI* secara objektif dan berbasis pengetahuan faktual (Ng, 2023) dengan kerangka konseptual Long dan Magerko (2020) serta diperkuat oleh pendekatan pedagogi teknologi cerdas berbasis data (Holmes et al., 2022).

Namun demikian, penggunaan SAIL4ALL secara luas masih terbatas karena instrumen ini dikembangkan dalam bahasa Inggris dan dikontekstualisasikan dalam budaya pengguna awalnya, sehingga penggunaannya di negara-negara non-Inggris seperti Indonesia memerlukan proses adaptasi bahasa dan budaya. Adaptasi alat ukur lintas budaya tidak hanya sebatas penerjemahan literal, tetapi juga mencakup validasi konseptual, semantik, dan teknis, agar makna item tetap utuh dalam konteks pengguna baru (Behr, 2017; van de Vijver & Leung, 2021). Menurut Hambleton et al. (2005), proses adaptasi juga harus memenuhi prinsip ekuivalensi linguistik dan psikometrik. Ditemukan pula dalam studi Chen et al. (2019) bahwa adaptasi tanpa validasi semantik yang menyeluruh dapat menurunkan daya prediksi alat ukur.

Beberapa penelitian terdahulu telah menekankan pentingnya adaptasi alat ukur berbasis fakta seperti SAIL4ALL dalam konteks global. Misalnya, Ng (2023) menyoroti potensi skala ini untuk digunakan pada populasi dewasa di luar lingkungan akademik, tetapi mengakui perlunya studi adaptasi yang ketat agar instrumen ini dapat digunakan secara luas. Studi lain oleh Tang et al. (2022) menunjukkan bahwa adaptasi linguistik tanpa validasi konteks dapat menghasilkan perbedaan signifikan dalam interpretasi responden, yang pada akhirnya menurunkan reliabilitas alat ukur. Holmes et al. (2022) juga mendukung pendekatan kontekstualisasi budaya dalam mengadaptasi alat ukur literasi berbasis teknologi untuk populasi non-Barat. Selain itu, penelitian oleh Soto et al. (2023) menunjukkan efektivitas adaptasi skala literasi *AI* SAIL4ALL dalam berbagai konteks budaya dan bahasa, termasuk validitas konstruk dan konsistensi internal instrumen.

Studi ini berfokus pada populasi karyawan PT X, sebuah perusahaan ritel nasional dengan struktur organisasi yang kompleks dan karyawan dari latar belakang pendidikan yang beragam. Dengan memilih konteks ini, penelitian memberikan kontribusi penting terhadap pemahaman bagaimana literasi *AI* dapat diukur secara faktual di lingkungan kerja non-teknis dan berbahasa Indonesia. Proses adaptasi dilakukan melalui pendekatan sistematis, termasuk penerjemahan terbalik, validasi isi oleh pakar, dan uji keterbacaan terhadap sampel terbatas sebelum implementasi penuh (Behr, 2017; Hambleton et al., 2005; van de Vijver & Leung, 2021).

Literatur sebelumnya menunjukkan bahwa literasi *AI* memiliki peran penting dalam meningkatkan adopsi teknologi secara kritis dan etis di tempat kerja (West et al., 2019; Ng et al., 2021; Schneider & Council, 2022). Dalam konteks adaptasi alat ukur, Behr (2017) dan Van de Vijver & Leung (2021) menekankan pentingnya pendekatan multi-tahap yang mencakup analisis semantik, kontekstualisasi budaya, dan uji empiris. Hambleton et al. (2005) menambahkan bahwa uji validitas lintas bahasa harus mempertimbangkan kesetaraan makna dan struktur konstruk asli. Studi validasi SAIL4ALL oleh Soto et al. (2023) juga memberikan referensi kuat bahwa pendekatan adaptasi yang ketat dapat menghasilkan alat ukur lintas budaya yang *robust* dan sesuai konteks

METODE PENELITIAN

Responden

Table 1. Tabel Demografi Responden

| Karakteristik | Kategori | <i>f</i> | % |
|----------------------|--------------------|-----------------|----------|
| Jenis Kelamin | Laki – Laki | 170 | 30,4% |
| | Perempuan | 390 | 69,6% |
| Usia | 18 - < 25 Tahun | 179 | 31,9% |
| | 25 - < 35 Tahun | 187 | 33,3% |
| | 35 - < 45 Tahun | 130 | 23,2% |
| | 45 - < 55 Tahun | 63 | 11,2% |
| | > 55 Tahun | 1 | 0,1% |
| Pendidikan | SMA Sederajat | 463 | 82,6% |
| | D1/D2/D3 (Diploma) | 12 | 3,0% |

| Karakteristik | Kategori | <i>f</i> | % |
|---------------|---------------|----------|-------|
| Jabatan | S1 (Sarjana) | 77 | 13,7% |
| | S2 (Magister) | 3 | 0,5% |
| | Manager | 16 | 2,8% |
| | Supervisor | 83 | 14,8% |
| | Staff | 461 | 82,3% |
| Total | | 560 | 100 |

Berdasarkan data demografi, jumlah responden dalam penelitian ini awalnya melibatkan 600 karyawan aktif di PT. X, tetapi hanya 560 yang memenuhi kriteria dan tidak termasuk dalam outlier. Mayoritas responden berjenis kelamin perempuan, yaitu sebanyak 390 orang (69,6%), sedangkan laki-laki berjumlah 170 orang (30,4%). Dari segi usia, responden terbanyak berada pada kelompok usia 25 - <35 tahun, yaitu 187 orang (33,3%), disusul oleh kelompok usia 18 - <25 tahun sebanyak 179 orang (31,9%), kelompok usia 35 - <45 tahun sebanyak 130 orang (23,2%), kelompok usia 45 - <55 tahun sebanyak 63 orang (11,2%), dan hanya 1 responden (0,1%) yang berusia di atas 55 tahun. Mayoritas responden memiliki tingkat pendidikan terakhir SMA/ sederajat, yaitu sebanyak 463 orang (82,6%). Sebanyak 77 orang (13,7%) lulusan sarjana (S1), 12 orang (3,0%) lulusan D1/D2/D3, dan hanya 3 orang (0,5%) yang berpendidikan magister (S2). Berdasarkan level jabatan, mayoritas responden menempati posisi sebagai staf sebanyak 461 orang (82,3%), diikuti oleh supervisor sebanyak 83 orang (14,8%), dan manajer sebanyak 16 orang (2,8%). Pengambilan sampel dilakukan dengan metode non-probability sampling melalui pendekatan *convenience sampling*, di mana kuesioner disebar ke seluruh cabang perusahaan tanpa acak. Hanya karyawan yang bersedia secara sukarela yang diikutsertakan, sebagaimana umum diterapkan dalam studi organisasi besar (Tang et al., 2022). Kriteria inklusi mencakup karyawan dewasa yang aktif bekerja tanpa batasan jenjang pendidikan atau posisi.

Instrumen

Skala yang digunakan pada penelitian ini adalah skala SAIL4ALL (*The Scale of Artificial Intelligence Literacy for all*), yang mencakup empat dimensi yaitu, Apa itu AI, Apa yang dapat dilakukan oleh AI, Bagaimana cara kerja AI, Bagaimana seharusnya AI digunakan, dengan jumlah 52 item. Berikut rangkuman item pada alat ukur SAIL4ALL pada tabel 2, yaitu:

Tabel. 2 Rangkuman Item pada alat ukur *Scale of Artificial Intelligence Literacy for All*
(SAIL4LL)

| Dimensi | Nomor Item | Jumlah Item |
|-----------------------------------|---|------------------------|
| Apa itu AI | 1,2,3,4,5*,6,7,8*,9*,10,11,12,13,14 | 14 |
| Apa yang dapat dilakukan AI | 15,16,17,18,19 20,21,22,23,24*,25,26,27,28,29,30 | 5 |
| Bagaimana cara kerja AI | ,31,32,33*,34,35,36,37*,38,39,40, 41,42 | 23 |
| Bagaimana seharusnya AI digunakan | 43,44,45,46,47,48,49,50,51,52 | 10 |

Catatan: **item Unfavorable*

Interpretasi dilakukan dengan cara menjumlahkan total skor. Total skor tersebut menggambarkan seberapa cakap kemampuan individu untuk memahami, mengevaluasi, dan berinteraksi secara kritis dengan sistem *AI* (Long & Magerko, 2020).

Desain

Pendekatan penelitian ini menggunakan metode kuantitatif dengan desain non-eksperimental yang bertujuan untuk mengevaluasi validitas dan reliabilitas suatu instrumen. Proses adaptasi tes mengacu pada prosedur yang dikemukakan oleh Hambleton et al. (2005), yang menekankan pentingnya kesetaraan makna dalam penerjemahan lintas budaya. Penelitian ini dilakukan dalam dua tahap utama. Tahap pertama melibatkan proses penerjemahan skala dari Bahasa Inggris ke Bahasa Indonesia dengan mempertimbangkan konteks budaya lokal. Tahap kedua adalah pengujian validitas dan reliabilitas instrumen yang telah diterjemahkan, guna memastikan bahwa alat ukur tersebut layak digunakan dalam konteks penelitian di Indonesia.

Teknik Analisis

Dalam penelitian ini, teknik analisis data yang digunakan mencakup beberapa tahapan penting untuk memastikan kualitas alat ukur. Pertama, dilakukan analisis *item-rest correlation* untuk menilai kualitas masing-masing item dalam skala. Tujuan dari teknik ini adalah mengidentifikasi item yang layak dipertahankan, perlu direvisi, atau dihilangkan karena tidak memenuhi standar kontribusi terhadap total skor. Batasan yang digunakan dalam menentukan

kelayakan item adalah nilai korelasi (r) lebih besar dari 0.4, yang menunjukkan kontribusi signifikan terhadap konstruk yang diukur (Zijlmans et al., 2018).

Selanjutnya, peneliti menerapkan teknik *exploratory factor analysis* (EFA) guna menyaring item lebih lanjut serta mengidentifikasi struktur faktor yang mendasari skala (Samuel, 2017). Setelah struktur faktor diperoleh, langkah berikutnya adalah menguji reliabilitas internal skala menggunakan teknik *Cronbach's alpha*, yang umum digunakan untuk menilai konsistensi internal instrumen.

Untuk menguji validitas konstruk, peneliti menggunakan *confirmatory factor analysis* (CFA), sebuah pendekatan statistik yang lazim digunakan untuk mengkonfirmasi struktur faktor teoritis dari suatu alat ukur (Atkinson et al., 2011). CFA memberikan akurasi statistik dalam pengujian validitas skala dan sangat membantu dalam proses pengembangan instrumen (Wibisono & Sasia, 2020; Widyatmoko et al., 2019). Beberapa indikator kelayakan model yang digunakan antara lain rasio $X^2/df < 2$, $p\text{-value} > .05$, $GFI > .90$, $RMSEA < .08$, $SRMR < .09$, $TLI > .95$, dan $CFI > .95$ (Brown, 2006; Cangur & Ercan, 2015; Hu & Bentler, 1999; Schumacker & Lomax, 2010). Seluruh proses analisis statistik dilakukan menggunakan perangkat lunak JASP versi 0.19.3 (JASP Team, 2025).

HASIL DAN PEMBAHASAN

Proses adaptasi diawali dengan penerjemahan skala SAIL4ALL (*The Scale of Artificial Intelligence Literacy for all*) ke dalam bahasa Indonesia (*forward translation*) oleh dua ahli dwibahasa, dilanjutkan dengan penerjemahan balik (*back translation*) oleh penerjemah independen. Pada tahap ini, dihasilkan dua versi terjemahan yang kemudian digabungkan kembali menjadi satu skala dalam Bahasa Inggris. Versi asli, hasil terjemahan ke Bahasa Indonesia (*forward translation*), serta hasil terjemahan balik ke Bahasa Inggris (*back translation*), diserahkan kepada seorang penelaah yang ahli di bidang psikologi dan memiliki kemampuan Bahasa Inggris yang baik. Ketiga versi tersebut dibandingkan dan dianalisis untuk memastikan kesamaan makna, meskipun penggunaan istilah mungkin berbeda (Hambleton et al., 2005). Skala dalam Bahasa Indonesia dikonfirmasi memiliki makna yang setara dengan versi Bahasa Inggris. Versi ini disebut sebagai versi pra-final.

Uji keterbacaan versi pra-final dilaksanakan terhadap sekelompok kecil partisipan ($n = 20$) untuk mengidentifikasi istilah yang membingungkan sebelum instrumen disebarkan secara luas. Proses ini memastikan bahwa versi adaptasi mempertahankan makna konseptual dari versi asli, sebagaimana direkomendasikan dalam studi validasi lintas budaya (Soto et al.,

2023). Semua responden menyatakan bahwa mereka memahami instruksi serta butir-butir yang terdapat dalam versi pra-final, sehingga tidak diperlukan revisi tambahan. Oleh karena itu, versi ini ditetapkan sebagai versi final dari skala dalam Bahasa Indonesia dan digunakan untuk proses pengumpulan data.

Pengumpulan data dilakukan secara daring dan luring melalui koordinasi dengan pihak internal perusahaan. Responden mengisi kuesioner dengan persetujuan sukarela melalui *informed consent*. Instrumen terdiri dari item pilihan skala likert 1 = Sangat Tidak Yakin – 5 = Sangat Yakin, yang menilai pemahaman faktual tentang AI, bukan persepsi diri, sebagaimana disarankan oleh Djamaluddin et al. (2023), untuk menghindari bias dalam pengukuran literasi teknologi. Setelah pengumpulan data dan analisis data dilanjutkan ke tahap berikutnya, yaitu pengujian *item-rest correlation*. Metode ini digunakan untuk menghitung korelasi antara masing-masing item dengan total skor dari item-item lainnya, guna mengetahui sejauh mana kontribusi setiap item terhadap keseluruhan instrumen (Zijlmans et al., 2018). Ringkasan hasil pengujian *item-rest correlation* disajikan dalam Tabel 3. Dari seluruh item yang dianalisis, terdapat sepuluh item dengan nilai r di bawah 0,4, yang menunjukkan kontribusi yang kurang signifikan terhadap alat ukur (Zijlmans et al., 2018), yaitu yaitu item P4(0,324), P5 (0,20), P6 (0,391), P7 (0,396), P8 (0,21), P9 (0,24), P24 (0,389), P33 (0,373), P37 (0,366), P38 (0,313) Oleh karena itu, sepuluh item tersebut digugurkan dari analisis selanjutnya karena tidak menunjukkan kontribusi yang memadai terhadap total skor skala.

Tabel. 3 Uji *item-rest correlation Scale of Artificial Intelligence Literacy for All (SAIL4LL)*.
 ($\omega = 0,929, M = 196,037, SD = 21,121$)

| Dimensi | Item | Pertanyaan | r- coreccted item total | Reliabilitas | Kesimpulan |
|-------------------------|------|---|-------------------------------|------------------|------------|
| Apa itu AI (M = 52,793) | 1 | Kecerdasan Buatan (AI) menggunakan algoritma untuk mempelajari dari data dan menjalankan tugas-tugas yang | 0,488 | $\alpha = 0,810$ | Diterima |

| Dimensi | Item | Pertanyaan | r- coreccted item total | Reliabilitas | Kesimpulan |
|---------------|------|--|-------------------------------|--------------|-------------|
| SD (6,668) | = | memerlukan kecerdasan | | | |
| | 2 | Penggunaan data dalam jumlah masif membantu beberapa algoritma AI I untuk meningkatkan kinerjanya. | 0,504 | | Diterima |
| | 3 | AI diaplikasikan untuk mengenali ucapan manusia | 0,498 | | Diterima |
| | 4 | Kemampuan untuk belajar dari pengalaman adalah fitur dari kecerdasan | 0,324 | | Dieliminasi |
| | 5 | Kecerdasan manusia adalah satu-satunya bentuk kecerdasan yang dapat dikategorikan kecerdasan yang sebenarnya | 0,209 | | Dieliminasi |
| | 6 | Kemampuan untuk menggunakan alat-alat dan memanipulasi lingkungan sekitar | 0,391 | | Dieliminasi |

| Dimensi | Item | Pertanyaan | r- corected item total | Reliabilitas | Kesimpulan |
|----------------|-------------|--|---------------------------------------|---------------------|-------------------|
| | | adalah suatu bentuk kecerdasan | | | |
| | 7 | <i>AI</i> dapat diprogram untuk menjalankan jangkauan tugas yang luas dengan tepat dan konsisten, melampaui kecerdasan spesifik dari kecerdasan anak-anak. | 0,396 | | Dieliminasi |
| | 8 | Berbagai hal terkait dengan kemanusiaan tidak memiliki tempat dalam <i>AI</i> | 0,219 | | Dieliminasi |
| | 9 | <i>AI</i> adalah sebuah teknologi tunggal | 0,241 | | Dieliminasi |
| | 10 | Computer Vision adalah contoh dari antardisiplin ilmu teknologi <i>AI</i> | 0,561 | | Diterima |
| | 11 | Narrow <i>AI</i> mendefinisikan algoritma dimana untuk memecahkan masalah-masalah tertentu | 0,628 | | Diterima |

| Dimensi | Item | Pertanyaan | r- corrected item total | Reliabilitas | Kesimpulan |
|--|------|--|-------------------------------|------------------|------------|
| | 12 | AI Berbasis Tugas yang spesifik disebut Narrow AI | 0,628 | | Diterima |
| | 13 | AI dapat dibagi kedalam spesifik sub bidang seperti General AI dan Narrow AI | 0,575 | | Diterima |
| | 14 | Sistem Narrow AI di rancang khusus untuk tugas dan domain tertentu. | 0,622 | | Diterima |
| Apa yang dapat dilakukan AI (M=16,389, SD = 3,731) | 15 | AI unggul dalam berkinerja dengan baik di lingkungan yang kompleks, seperti mengemudi di jalan yang ramai. | 0,458 | $\alpha = 0,763$ | Diterima |
| | 16 | Sebab termasuk kedalam topik lintas budaya, AI berlaku sama baiknya di seluruh negara | 0,488 | | Diterima |
| | 17 | Keputusan beresiko tinggi sebaiknya diserahkan kepada AI karena lebih netral dibandingkan manusia | 0,593 | | Diterima |

| Dimensi | Item | Pertanyaan | r- corrected item total | Reliabilitas | Kesimpulan |
|--|------|---|-------------------------------|------------------|------------|
| | 18 | Saat ini <i>AI</i> sudah sangat mampu menampilkan proses asosiasi rumit seperti yang manusia lakukan | 0,580 | | Diterima |
| | 19 | <i>AI</i> efisien dalam mengatasi permasalahan yang melibatkan emosi | 0,558 | | Diterima |
| Bagaimana cara kerja <i>AI</i> (M=87,079, SD=10,150) | 20 | Beberapa sistem <i>AI</i> dapat mewakili pola visual dan Audio | 0,538 | $\alpha = 0,890$ | Diterima |
| | 21 | Contoh dari gambaran sebuah pengetahuan mencakup decision trees dan bayesian networks | 0,522 | | Diterima |
| | 22 | Gambaran pengetahuan memainkan peran krusial dalam pembelajaran mesin dengan cara membuat representasi fitur dari data mentah | 0,531 | | Diterima |
| | 23 | Sistem berbasis-aturan adalah salah satu contoh bagaimana | 0,450 | | Diterima |

| Dimensi | Item | Pertanyaan | r- corrected item total | Reliabilitas | Kesimpulan |
|---------|------|---|-------------------------------|--------------|-------------|
| | 24 | komputer dapat bernalar Komputer hanya dapat bernalar dan membuat keputusan dengan cara yang identik seperti bagaimana yang dilakukan manusia. | 0,389 | | Dieliminasi |
| | 25 | <i>Machine learning</i> digunakan untuk memprediksi, menggroupkan, dan mengklasifikasikan data dalam jumlah besar. | 0,569 | | Diterima |
| | 26 | Deep learning adalah jenis <i>machine learning</i> | 0,565 | | Diterima |
| | 27 | Algoritma <i>machine learning</i> belajar dari data. | 0,551 | | Diterima |
| | 28 | Dalam <i>machine learning</i> , kumpulan data sering kali dibagi menjadi set pelatihan dan set pengujian. | 0,522 | | Diterima |
| | 29 | Pemilihan model merupakan langkah penting dalam proses pembelajaran mesin. | 0,611 | | Diterima |

| Dimensi | Item | Pertanyaan | r- corrected item total | Reliabilitas | Kesimpulan |
|----------------|-------------|--|--|---------------------|-------------------|
| | 30 | Data yang bias atau berpihak itu melestarikan stereotip sosial. | 0,494 | | Diterima |
| | 31 | Data itu tergantung penafsiran. | 0,450 | | Diterima |
| | 32 | Beberapa data yang digunakan dalam Kecerdasan buatan dibentuk dalam lingkungan budaya yang disetting, yang dapat memengaruhi hasil model yang menggunakan data tersebut. | 0,560 | | Diterima |
| | 33 | Karena data bersifat objektif, mesin model pembelajaran tidak bias | 0,373 | | Dieliminasi |
| | 34 | Data yang digunakan untuk melatih mesin model pembelajaran dapat bias tau memihak. | 0,465 | | Diterima |
| | 35 | Bias dalam data yang digunakan untuk melatih mesin model pembelajaran dapat | 0,535 | | Diterima |

| Dimensi | Item | Pertanyaan | r- corrected item total | Reliabilitas | Kesimpulan |
|---------|------|--|-------------------------------|--------------|-------------|
| | | menyebabkan hasil yang bias pula. | | | |
| | 36 | Pengawasan manusia diperlukan untuk memastikan bahwa sistem Kecerdasan buatan digunakan secara etis dan bertanggung jawab. | 0,468 | | Diterima |
| | 37 | Peran manusia dalam pengembangan Kecerdasan buatan terbatas pada pemantauan kinerja sistemnya. | 0,366 | | Dieliminasi |
| | 38 | <i>Autonomous driving</i> adalah satu bidang penerapan <i>AI</i> | 0,313 | | Dieliminasi |
| | 39 | Robot tidak hanya dapat bertindak terhadap dunia tetapi juga bereaksi. | 0,425 | | Diterima |
| | 40 | Mikrofon adalah sejenis sensor yang digunakan dalam robot. | 0,487 | | Diterima |
| | 41 | Sensor dapat membantu robot untuk | 0,500 | | Diterima |

| Dimensi | Item | Pertanyaan | r- corrected item total | Reliabilitas | Kesimpulan |
|--|------|---|-------------------------------|------------------|------------|
| | 42 | memahami lingkungannya. Sensor adalah perangkat yang dapat mendeteksi dan mengubah sifat fisik yang dapat diukur menjadi format digital. | 0,569 | | Diterima |
| Bagaimana seharusnya AI digunakan (M=40,074, SD = 6,659) | 43 | Untuk mencapai transparansi yang lebih luas, kode sumber, penggunaan data, dasar pembuktian untuk penggunaan AI, serta batasan dan tanggung jawab harus dikomunikasikan | 0,503 | $\alpha = 0,923$ | Diterima |
| | 44 | Kecerdasan Buatan yang diciptakan harus menyesuaikan dengan isu demokrasi dan kemasyarakatan dikomunikasikan | 0,474 | | Diterima |
| | 45 | Perlu untuk mengembangkan dan memperkuat aturan serta undang - undang, termasuk hak untuk mengajukan banding, | 0,602 | | Diterima |

| Dimensi | Item | Pertanyaan | r- corrected item total | Reliabilitas | Kesimpulan |
|---------|------|--|-------------------------------|--------------|------------|
| | | perbaikan atau pemulihan untuk solusi <i>AI</i> | | | |
| | 46 | <i>AI</i> harus memberikan informasi tentang alasan dan proses yang dapat menyebabkan potensi bahaya. | 0,670 | | Diterima |
| | 47 | Pengembang, perancang, lembaga, atau industri <i>AI</i> harus bertanggung jawab atas tindakan <i>AI</i> . | 0,685 | | Diterima |
| | 48 | Privasi harus dijamin melalui pendekatan desain <i>AI</i> , kontrol akses, kesadaran publik, dan hukum | 0,734 | | Diterima |
| | 49 | Pengembangan <i>AI</i> harus selaras dengan nilai-nilai kemanusiaan dan hak asasi manusia. | 0,701 | | Diterima |
| | 50 | <i>AI</i> yang terpercaya harus memiliki keandalan, akuntabilitas, serta terus melakukan proses pengawasan | 0,734 | | Diterima |

| Dimensi | Item | Pertanyaan | r- corrected item total | Reliabilitas | Kesimpulan |
|---------|------|---|-------------------------------|--------------|------------|
| | 51 | dan evaluasi integritas sistem AI seiring berjalannya waktu AI seharusnya tidak mengurangi atau menghancurkan tetapi menghormati, melestarikan atau bahkan meningkatkan martabat manusia. | 0,671 | | Diterima |
| | 52 | Manfaat AI tidak boleh mengancam keterikatan sosial dan harus menghormati orang maupun kelompok yang rentan | 0,683 | | Diterima |

Penentuan jumlah faktor pada analisis *faktor eksploratori* (EFA) dalam penelitian ini mengikuti kriteria *eigenvalue* ≥ 2 sebagaimana ditekankan DeVellis (2017), guna memastikan tiap faktor menyumbang variasi substansial terhadap total varians konstruk. Selain itu, butir-butir dengan muatan faktor (*factor loading*) di bawah ambang 0,30 dianggap tidak memadai lalu dieliminasi, sesuai anjuran Tavakol dan Wetzel (2020) agar item lemah tidak mengaburkan representasi faktor laten. Proses ekstraksi selanjutnya memakai rotasi ortogonal varimax yang bertujuan menyederhanakan pola korelasi antar butir sehingga interpretasi faktor menjadi lebih jelas. Teknik rotasi varimax digunakan dalam EFA ini. Berdasarkan factor loadings tiap item ditemukan 42 item yang memenuhi kriteria sehingga item yang diterima berjumlah 42. Hasil lebih detil dapat dibaca pada Tabel 4. Hasil EFA ini ditemukan sejalan dengan hasil item rest correlation, di mana terdapat 10 item yang menunjukkan performa

psikometrik di bawah standar. Oleh karena itu, hasil pengujian EFA dan item-rest correlation menjadi dasar argumen untuk mengeliminasi item-item tersebut dari alat ukur. sepuluh item tersebut tidak diikutsertakan dalam analisis lanjutan.

Table 4. Uji EFA pada alat ukur SAIL4LL

| Item | Faktor Loading 1 | Faktor Loading 2 | Faktor Loading 3 | Faktor Loading 4 |
|------|------------------|------------------|------------------|------------------|
| P29 | 0,744 | | | |
| P41 | 0,655 | | | |
| P42 | 0,646 | | | |
| P25 | 0,635 | | | |
| P28 | 0,622 | | | |
| P27 | 0,600 | | | |
| P26 | 0,556 | | | |
| P40 | 0,527 | | | |
| P32 | 0,498 | | | |
| P36 | 0,472 | | | |
| P35 | 0,429 | | | |
| P31 | 0,381 | | | |
| P39 | 0,375 | | | |
| P30 | 0,373 | | | |
| P20 | 0,371 | | | |
| P34 | 0,364 | | | |
| P22 | 0,355 | | | |
| P23 | 0,292 | | | |
| P50 | | 0,797 | | |
| P51 | | 0,782 | | |
| P48 | | 0,773 | | |
| P49 | | 0,773 | | |
| P52 | | 0,739 | | |
| P47 | | 0,694 | | |
| P46 | | 0,668 | | |
| P45 | | 0,489 | | |

| Item | Faktor Loading 1 | Faktor Loading 2 | Faktor Loading 3 | Faktor Loading 4 |
|------|------------------|------------------|------------------|------------------|
| P44 | | 0,477 | | |
| P43 | | 0.348 | | |
| P13 | | | 0,821 | |
| P14 | | | 0,821 | |
| P11 | | | 0,721 | |
| P12 | | | 0,697 | |
| P10 | | | 0,568 | |
| P2 | | | 0,469 | |
| P1 | | | 0,459 | |
| P3 | | | 0,348 | |
| P21 | | | 0,253 | |
| P17 | | | | 0,837 |
| P19 | | | | 0,711 |
| P18 | | | | 0,579 |
| P16 | | | | 0,506 |
| P15 | | | | 0,490 |

Berdasarkan hasil Exploratory Factor Analysis (EFA) terhadap 42 item skala adaptasi, diperoleh 4 faktor utama dengan *eigenvalue* > 1 dan didukung oleh pola *scree plot*. Namun, struktur faktor yang terbentuk menunjukkan perbedaan dibandingkan struktur teoritis skala aslinya. Beberapa item yang pada skala asal tergolong dalam satu dimensi justru mengelompok dengan item dari dimensi lain, serta membentuk konfigurasi baru yang secara konseptual memiliki makna berbeda. Oleh karena itu, peneliti menetapkan label baru untuk masing-masing faktor berdasarkan analisis isi item dan tema yang muncul

Table 5 : Rangkuman Item dengan label baru pada alat ukur *Scale of Artificial Intelligence Literacy for All (SAIL4LL)*

| Dimensi | Nomor Item | Jumlah |
|---------------------|---|--------|
| Pemahaman <i>AI</i> | 29,41,42,25,28,27,26,40,32,36,35,31,39,30,20, 34,22,23 | 18 |
| Etika <i>AI</i> | 50,51,48,49,52,47,46,45,44,43 | 10 |

| | | |
|-----------------------|-------------------------|---|
| Konsep <i>AI</i> | 13,14,11,12,10,2,1,3,21 | 9 |
| Kapabilitas <i>AI</i> | 17,19,18,16,15 | 5 |

Berdasarkan table di atas Faktor 1 diberi label Pemahaman *AI*, karena mencakup item-item yang berkaitan dengan komponen dan mekanisme penggunaan *AI*. Faktor 2 dinamai Etika *AI*, karena terdiri atas item-item yang digunakan untuk mengarahkan pengembangan, penerapan, dan penggunaan teknologi *AI* secara bertanggung jawab dan berkelanjutan. Faktor 3 diberi label Konsep *AI*, karena memuat item-item terkait kemampuan *AI* dalam memecahkan masalah layaknya manusia. Faktor 4 diberi label Kapabilitas *AI* karena item – item berkaitan dengan kinerja *AI* yang bisa meniru, mendukung atau menggantikan kecerdasan manusia. Penetapan label ini dilakukan untuk mencerminkan struktur faktor aktual yang muncul dari data, serta memastikan interpretasi hasil yang lebih relevan dengan konteks empiris. Setelah dilakukan struktur faktor yang baru dengan label dan item yang baru maka dilakukan selanjutnya Confirmatory Factor Analysis (CFA)

Table 6: *Goodness -Of- Fit Indices* dari hasil *Confirmatory Factor Analysis* (CFA) Skala SAIL4ALL

| Dimensi/Faktor | Chi - Square | Df | P | RMSEA | CFI | TLI | NNFI | GFI |
|-----------------------|--------------|-----|------|-------|-------|-------|-------|-------|
| Pemahaman <i>AI</i> | 818.702 | 135 | .001 | 0,095 | 0,793 | 0,765 | 0,765 | 0,845 |
| Etika <i>AI</i> | 281.093 | 35 | .001 | 0,112 | 0,909 | 0,883 | 0,883 | 0,903 |
| Konsep <i>AI</i> | 122.466 | 27 | .001 | 0,079 | 0,941 | 0,922 | 0,922 | 0,925 |
| Kapabilitas <i>AI</i> | 48.426 | 5 | .001 | 0,125 | 0,934 | 0,868 | 0,868 | 0,967 |
| <i>First Order</i> | 2840.055 | 813 | .001 | 0,067 | 0,796 | 0,784 | 0,784 | 0,788 |

Indeks *goodness-of-fit* dari hasil *Confirmatory Factor Analysis* (CFA) terhadap empat dimensi skala SAIL4ALL dan satu model *first-order* eksklusif. Secara umum, hasil menunjukkan bahwa tidak semua model dimensi menunjukkan kelayakan model (*model fit*) yang memadai.

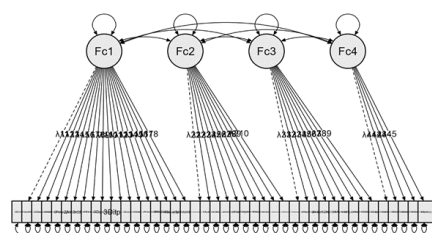
Salah satu indikator utama, yaitu nilai Chi-Square, menunjukkan signifikansi pada semua model ($p < .05$). Meskipun demikian, seperti dinyatakan oleh Brown (2015), nilai Chi-Square sangat sensitif terhadap ukuran sampel yang besar, sehingga interpretasi *fit* sebaiknya tidak hanya bergantung pada nilai ini. Indeks RMSEA (Root Mean Square Error of Approximation)

untuk sebagian besar model berada di atas batas ambang 0,08 yang disarankan oleh Hu dan Bentler (1999), yang menyatakan bahwa nilai $RMSEA \leq 0,06$ menunjukkan *good fit*, sementara nilai antara 0,08–0,10 dianggap *mediocre fit*, dan $> 0,10$ menunjukkan *poor fit*. Dalam hal ini: Dimensi Kapabilitas AI memiliki nilai RMSEA tertinggi (0,125), menunjukkan *poor fit*. Dimensi Etika AI juga berada di kategori kurang baik (0,112). Sementara Konsep AI (0,079) dan Model First-Order (0,067) mendekati batas acceptable fit.

Untuk CFI (Comparative Fit Index), standar nilai $\geq 0,90$ menunjukkan *acceptable fit* (Bentler, 1990). Hanya dimensi Etika AI (0,909), Konsep AI (0,941), dan Kapabilitas AI (0,934) yang melewati batas ini. Namun, model total (*first-order*) dan dimensi Pemahaman AI berada di bawah ambang tersebut, yaitu 0,796 dan 0,793, yang mengindikasikan *poor fit*. Nilai TLI/NNFI menunjukkan pola serupa. Menurut Hair et al. (2010), nilai TLI yang baik adalah $\geq 0,90$. Dimensi Konsep AI (0,922) dan Etika AI (0,883) mendekati nilai tersebut, sedangkan model total dan dimensi Pemahaman AI berada jauh di bawah standar. Indikator GFI (Goodness-of-Fit Index) pada umumnya lebih toleran. Nilai di atas 0,90 menunjukkan *fit* yang baik (Jöreskog & Sörbom, 1993). Dimensi Kapabilitas AI (0,967), Etika AI (0,903), dan Konsep AI (0,925) menunjukkan *fit* yang memadai berdasarkan indikator ini. Namun, model total hanya menunjukkan GFI sebesar 0,788. Berdasarkan temuan ini, dapat disimpulkan bahwa: Dimensi “Konsep AI” merupakan dimensi paling kuat secara struktural, dengan seluruh indeks fit menunjukkan nilai yang mendekati atau melampaui ambang batas yang disarankan. Dimensi “Kapabilitas AI” memiliki masalah serius dalam kelayakan model, sebagaimana ditunjukkan oleh nilai RMSEA yang tinggi, walaupun CFI dan GFI-nya masih tergolong baik. Model total (*first-order exclusive*) memperlihatkan *fit* yang lemah secara keseluruhan, yang mungkin mengindikasikan

bahwa struktur faktor yang dihipotesiskan perlu direvisi atau disederhanakan. Selanjutnya peneliti melakukan model plot

Model Plot



Gambar 1 : Model Plot First Order

Berdasarkan model plot diatas hasil analisis *Confirmatory Factor Analysis* (CFA) terhadap empat dimensi skala SAIL4ALL yakni Pemahaman *AI*, Etika *AI*, Konsep *AI*, dan Kapabilitas *AI* menunjukkan variasi tingkat kelayakan model (*goodness-of-fit*) di antara masing-masing dimensi. Secara visual, model CFA first-order memperlihatkan bahwa dimensi Konsep *AI* merupakan dimensi yang paling stabil, dengan faktor loading yang tinggi dan konsisten serta indeks fit statistik yang mendekati kriteria ideal (RMSEA = 0,079; CFI = 0,941).

Hal ini menunjukkan bahwa indikator pada dimensi ini secara kuat merefleksikan konstruk laten yang diukur. Sebaliknya, dimensi Kapabilitas *AI* tampak paling bermasalah, baik secara statistik (RMSEA = 0,125; TLI = 0,868) maupun secara visual, di mana indikator menunjukkan faktor loading yang tidak seragam dan jumlah indikator yang terlalu sedikit, sehingga mengurangi fleksibilitas model (Brown, 2015; Hair et al., 2010).

Sementara itu, dimensi Etika *AI* dan Pemahaman *AI* menunjukkan hasil yang beragam. Dimensi Etika *AI* masih memenuhi batas kelayakan model meskipun nilai RMSEA berada di atas ambang toleransi (0,112), yang menunjukkan bahwa model dapat diperbaiki melalui evaluasi indikator tertentu. Di sisi lain, dimensi Pemahaman *AI* meskipun terdiri dari indikator yang lebih banyak, memperlihatkan inkonsistensi faktor loading, yang secara statistik tercermin dalam nilai CFI yang rendah (0,793).

Temuan ini konsisten dengan pernyataan Hu dan Bentler (1999) bahwa nilai CFI di bawah 0,90 mengindikasikan model yang kurang memadai. Dengan mempertimbangkan hasil visual dan statistik, dapat disimpulkan bahwa struktur model *first-order* saat ini belum optimal untuk menjelaskan hubungan antar dimensi secara menyeluruh. Oleh karena itu, analisis ini tidak dilanjutkan ke tahap CFA second order, mengingat model dasar (*first Order*) belum memenuhi persyaratan kecocokan yang diperlukan. Keputusan ini diambil untuk menjaga validitas analisis dan menghindari interpretasi lanjutan yang berpotensi menyesatkan. Sebagaimana disarankan oleh Brown (2015), model high – order hanya dapat dianalisis jika model *first – order* telah menunjukkan kecocokan yang memadai.

Diskusi

Hasil analisis menunjukkan bahwa meskipun alat ukur versi adaptasi memiliki reliabilitas internal yang tinggi ($\omega = 0,929$), model konstruk yang dihasilkan melalui analisis CFA belum menunjukkan kecocokan yang optimal, baik pada level dimensi maupun model keseluruhan. Berbeda dengan hasil penelitian Ng (2023), di mana SAIL4ALL menunjukkan

struktur empat faktor yang stabil dan fit pada populasi mahasiswa dan profesional teknologi, versi adaptasi ini mengalami pergeseran struktur konstruk. Hasil EFA menghasilkan susunan dimensi baru, dan beberapa item berpindah ke faktor yang berbeda dari versi asli. Hal ini dapat dijelaskan oleh perbedaan latar belakang budaya dan pendidikan antara responden versi asli dan responden versi adaptasi. Ng mengembangkan alat ini dalam konteks negara maju dengan tingkat literasi digital tinggi, sementara dalam penelitian ini, responden didominasi oleh karyawan sektor ritel di Indonesia dengan latar pendidikan menengah dan paparan teknologi yang terbatas

Seperti ditegaskan oleh Hambleton et al. (2005) dan van de Vijver & Leung (2021), validitas konstruk dalam adaptasi alat ukur tidak hanya bergantung pada terjemahan bahasa, tetapi juga pada kesesuaian makna psikologis dalam budaya dan konteks lokal. Perbedaan pemahaman terhadap istilah seperti “algoritma”, “transparansi AI”, atau “etika mesin” menjadi hambatan dalam mempertahankan struktur faktor asli. Ini menunjukkan adanya kesenjangan konseptual antara alat ukur dan pengalaman nyata responden di lapangan. Meskipun demikian, penelitian ini memiliki kontribusi praktis penting, yaitu menyediakan versi awal alat ukur literasi *AI* yang bisa digunakan dalam konteks dunia kerja non-teknis di Indonesia. Hasil ini dapat menjadi dasar untuk menyusun intervensi pelatihan *AI* yang lebih tepat sasaran

KESIMPULAN DAN SARAN

Kesimpulan

Kesimpulan yang didapat dari adaptasi alat ukur SAIL4ALL ada 3, yaitu; (1) Hasil penelitian menunjukkan bahwa meskipun skala telah diterjemahkan dan diuji keterbacaannya, sejumlah item tetap sulit dipahami karena mengandung istilah teknis dan konteks yang asing bagi responden. Ini memperkuat pandangan Hambleton et al. (2005) bahwa adaptasi alat ukur lintas budaya harus mempertimbangkan ekuivalensi konseptual, bukan hanya terjemahan literal; (2) Sebagian besar responden dalam studi ini berasal dari sektor ritel dengan tingkat pendidikan menengah, yang berbeda jauh dari populasi pengembangan asli SAIL4ALL. Hal ini berdampak pada pemahaman terhadap istilah seperti “algoritma”, “machine learning”, atau “etika AI”. Seperti dijelaskan oleh Chen, Zhang, & Wang (2019), perbedaan tingkat literasi digital dapat menyebabkan bias pemahaman konstruk, sehingga struktur faktor asli tidak lagi sesuai; (3) Ketidaksesuaian model konstruk dalam CFA menunjukkan bahwa struktur alat ukur perlu dikaji ulang agar sesuai dengan cara berpikir dan pengalaman masyarakat Indonesia. Menurut Van de Vijver & Leung (2021), keberhasilan adaptasi hanya bisa dicapai jika alat ukur

mempertimbangkan nilai-nilai, norma, dan pengalaman budaya lokal yang membentuk cara individu merespon pertanyaan.

Saran

Hasil CFA yang belum menunjukkan model fit secara optimal, disarankan agar peneliti berikutnya melakukan revisi struktur konstruk melalui pengujian ulang EFA dan CFA pada populasi berbeda. Selain itu, perlu dilakukan external validation untuk menguji validitas konvergen dan prediktif dari skala. Validitas konvergen dapat diuji dengan skala *Digital Competence* (Ng, 2021) atau *Technology Readiness Index* (Parasuraman, 2000), sedangkan validitas prediktif dapat melibatkan variabel seperti minat mengikuti pelatihan *AI* atau frekuensi penggunaan sistem *AI* di tempat kerja

DAFTAR PUSAKA

- Atkinson, M., Atkinson, R. C., Smith, E. E., Bem, D. J., & Nolen-Hoeksema, S. (2011). *Introduction to psychology* (15th ed.). Cengage Learning.
- Behr, D. (2017). Assessing the use of back translation: The shortcomings of back translation as a quality testing method. *International Journal of Social Research Methodology*, 20(6), 573–584.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.
- Cangur, S., & Ercan, I. (2015). Comparison of model fit indices used in structural equation modeling under multivariate normality. *Journal of Modern Applied Statistical Methods*, 14(1), 152–167.
- Chen, F., Zhang, Y., & Wang, Z. (2019). Cultural adaptation in psychometric assessments: Challenges and strategies. *Journal of Cross-Cultural Psychology*, 50(3), 271–285.
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). SAGE Publications.
- Djamaluddin, R., Lestari, D. R., & Arifianto, A. (2023). Self-perception versus factual knowledge: A meta-review of technology literacy assessments. *Asian Journal of Technology and Society*, 5(1), 22–36.

- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson Education.
- Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Psychology Press.
- Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
- JASP Team. (2025). *JASP* (Version 0.19.3) [Computer software].
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific Software International.
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). <https://doi.org/10.1145/3313831.3376727>
- Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education*, 59(3), 1065–1078.
- Ng, W. (2023). SAIL4ALL: Developing and validating a standardized AI literacy instrument for all. *International Journal of Artificial Intelligence in Education*, 33(1), 45–67.
- Ng, W., Yuen, A. H. K., & Cheung, W. S. (2021). AI literacy in K-12 and beyond: Emerging concepts and design approaches. *British Journal of Educational Technology*, 52(5), 1981–1998.
- Parasuraman, A. (2000). Technology Readiness Index (TRI): A multiple-item scale to measure readiness to embrace new technologies. *Journal of Service Research*, 2(4), 307–320.
- Samuel, A. G. (2017). Exploratory factor analysis: A practical guide. *Behavior Research Methods*, 49, 200–211.
- Schneider, B., & Council, M. (2022). The role of AI literacy in workforce readiness: Implications for policy and education. *Technology in Society*, 68, 101902.
- Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). Routledge.
- Setiawan, B., & Hidayat, R. (2022). Strategi literasi teknologi bagi pekerja sektor ritel: Studi awal pada adopsi AI di Indonesia. *Jurnal Teknologi dan Masyarakat*, 10(2), 80–93.

- Soto, C. M., Ng, W., & Law, N. W. (2023). Validation of the SAIL4ALL instrument across linguistic and cultural contexts: A cross-national study. *International Journal of Artificial Intelligence in Education, 33*(2), 223–245.
- Tang, Y., Zhang, H., & Li, W. (2022). Cultural nuances in AI literacy assessment: A comparative study of Western and Asian samples. *Journal of Educational Computing Research, 60*(3), 512–534.
- Tavakol, M., & Wetzel, A. (2020). Factor analysis: A means for theory and instrument development in support of construct validity. *International Journal of Medical Education, 11*, 245–247.
- Tuomi, I. (2023). The future of learning in the age of AI: Challenges for education systems. *European Journal of Education, 58*(2), 147–164.
- van de Vijver, F. J. R., & Leung, K. (2021). Equivalence and bias: A review of concepts, models, and data analytic procedures. *Organizational Research Methods, 24*(1), 203–224.
- West, D. M., Allen, J. R., & Gorham, M. (2019). How artificial intelligence is transforming the world. *Brookings Institution Report*.
- Wibisono, S., & Sasia, W. (2020). Validitas konstruk dalam pengembangan skala psikologis: Studi konfirmasi menggunakan CFA. *Jurnal Psikologi, 47*(2),
- Widyatmoko, M. A., Susetyo, B., & Ratnasari, M. (2019). Penggunaan CFA dalam validasi skala pada penelitian sosial. *Jurnal Penelitian dan Evaluasi Pendidikan,*
- World Economic Forum. (2021). *Global competency index report*. World Economic Forum.
- Yigitcanlar, T., Kankanamge, N., & Corchado, J. M. (2023). The AI knowledge divide and public perception: Bridging the gap for inclusive technological advancement. *AI & Society*.