

Collaborative by Design, Fragmented in Delivery: A Scoping Review of National Digital Talent Programmes

Lia Puspitasari^{1,2*}[0000-0002-5742-2074], Alfira Sofia¹[0000-0002-3123-864X], Lili Adi Wibowo¹[0000-0002-4405-8744], Agus Rahayu¹[0000-0002-8737-8043]

¹Faculty of Economics and Business Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

²Ministry of Communications and Digital of Indonesia, Bandung, Indonesia

*Corresponding Email: puspitasari.lia@upi.edu

Abstract. Governments across income groups now fund national programmes to build digital talent, yet whether they work, and how anyone would know, remains unclear. This scoping review maps how their performance is conceptualised, measured and governed. Following established scoping review guidance, we searched Scopus and screened 265 records against criteria fixed in advance, retaining 25 peer-reviewed reports published between 2018 and 2026. Data were charted using a 67-variable form; 98 governance mechanisms and 105 performance indicators were then coded to evaluation and collaborative governance frameworks and synthesised thematically. Performance is defined explicitly in 14 of 25 reports, measured mainly by self-report (44%), and reported without any time horizon in 60%. One report uses a comparison group and none estimates a counterfactual. Only three measure how much of the intended population was reached, and all three find that most was not. Governance mechanisms are coded as barriers in 33% of design-stage units but 70% of delivery-stage units. No report with an enabling reading attains the highest appraisal rating, so the negative pattern is not an artefact of weak evidence. We propose, as hypotheses requiring empirical validation rather than as a validated extension, that reach, financing structure, post-programme filters and regime status be modelled as variables rather than assumed.

Keywords: Digital Talent Policy, Collaborative Governance, Programme Reach, Performance Measurement, Public Workforce Development

1 Introduction

Digital transformation is changing what employers ask of workers faster than education and training systems adjust. Analyses of large job posting corpora show artificial intelligence restructuring the composition of skill demand rather than simply raising it [1], with measurable effects on hiring behaviour already visible in national labour markets [2]. At the same time, access to technology has not translated evenly into capability, and the relationship between internet skills, engagement and tangible outcomes differs enough across settings that instruments developed in high-income countries cannot be assumed to travel [3]. The result is a policy problem that is simultaneously economic, educational and distributional, that is urgent because the

adjustment is already under way, and that few governments can address through any single ministry.

Three concepts define the object of this review and its boundaries. A national digital talent development programme is understood here as an intervention that a state initiates, mandates, funds or regulates, that operates at national or sectoral rather than institutional scale, and that seeks to build the digital capabilities of a working-age population; firm-level talent management and single-institution training provision therefore fall outside it. Programme performance is treated as an open construct rather than a settled metric, since it may denote competence acquired, services delivered, labour market position or system-level capability, and the review charts which of these a study adopts rather than imposing one in advance. Collaborative governance, following the integrative framework of Emerson [4], denotes arrangements in which public agencies and non-state actors deliberate and act jointly on a shared public mandate. The review concerns the intersection of the three: how performance is understood, and how collaboration operates, in publicly mandated programmes at system scale.

Practice has expanded quickly. Across income groups, states now initiate, mandate, fund or regulate national digital talent and digital skills programmes, typically delivering them through partnerships with industry, universities, training providers and multilateral bodies, which makes such programmes an instance of collaborative governance rather than of conventional public provision. The public sector is not only the convenor in these arrangements but also a beneficiary, since digital government capability depends on the workforce the programmes are meant to build [5]. Research has developed alongside this expansion, though unevenly and largely in parallel streams. One stream documents implementation: public sector digital strategies can be adopted without being enacted, producing decoupling between stated transformation and actual practice [6], and strategy documents frequently reproduce the status quo rather than commit to change [7]. Another stream examines the collaborations themselves and reports that performance management inside collaborative governance remains under-theorised and is rarely linked back to how the collaboration was designed [8], [9]. The two streams have seldom been brought together.

The measurement problem is consequently deeper than an absence of data; it is one of heterogeneity in what is measured and of fragmentation in how the pieces relate. What "performance" means for a programme of this kind is unsettled: it can refer to competence acquired, to services delivered, to labour market position, or to system-level capability, and evaluation frameworks such as CIPP [10] distinguish these but do not adjudicate between them. Each candidate measure carries known problems. Self-assessed digital competence diverges systematically from demonstrated competence, and unevenly across social groups [11]. Labour market effects of training emerge over multi-year horizons and can point in the wrong direction in the interim [12]. Participation cannot be treated as revealed preference, because awareness and learning costs shape who enrolls [13] and reducing those costs measurably changes take-up [14]. Implementation capacity in the delivering bureaucracy determines who is reached at all [15]. Even the standardisation instruments most often proposed as a remedy, occupational standards and

competency-based qualification frameworks, have been shown to deliver limited value in the systems that adopted them most extensively [16]. The evidence is thus dispersed across constructs, disciplines and units of analysis that do not readily align.

Existing reviews do not resolve this, because none takes system-level programme performance as its object. The closest syntheses examine adjacent territory: outcomes of digital skills for young people rather than for the workforce [17]; the digital competence of pre-service teachers [18]; digital health training within a single profession [19]; digitalised health workforce information systems rather than talent development itself [20]; and assessment instruments within one clinical discipline [21]. The talent management literature, meanwhile, has advanced considerably but remains anchored at the level of the firm [22] or the individual institution [23]. None of these addresses programmes that operate at national or sectoral scale under a public mandate, and none treats governance arrangements and measurement practice as joint objects of analysis. Evidence at that level is dominated by institutional reports rather than peer-reviewed synthesis.

A scoping review is the appropriate response to an evidence base with these properties. Where the literature is conceptually heterogeneous, methodologically diverse and growing rapidly, and where the prior question is what has been studied rather than how large an effect is, the purpose of synthesis is to map the extent, range and nature of that literature and to identify where evidence is absent, not to estimate the effect of an intervention [24]. A systematic review with meta-analysis would require a stable outcome definition and broadly comparable designs; the preceding paragraphs indicate that this literature supplies neither, nor pooling across such divergent constructs would obscure the very variation that needs to be described. A scoping approach also permits governance mechanisms and performance indicators to be charted side by side within the same set of studies. We therefore follow Arksey & O'Malley [25] as refined by Levac et al. [26] and operationalized through JBI methodology [27], reporting in accordance with PRISMA-ScR [28].

What distinguishes this review from the syntheses above is that it treats governance arrangements and measurement practice as a single object rather than as two literatures that happen to describe the same programmes. The contribution is threefold. Empirically, the review provides the first synthesis of peer-reviewed evidence on the performance of publicly mandated digital talent programmes at system scale, charted against an explicit and reproducible scheme. Analytically, it separates what this literature measures from what it governs and shows that the two are misaligned in a specific and consequential way, since collaboration is documented where programmes are designed while difficulty accumulates where they are delivered, and the indicators in use are least equipped to detect it. Theoretically, it identifies configurations that the dominant collaborative governance framework does not accommodate and proposes a bounded extension to it. For policy and for strategic management practice, the review supplies a vocabulary for distinguishing the parts of a programme that are underperforming from those that are merely unmeasured.

The aim of this review is to map the peer-reviewed evidence on the performance of national and system-scale digital talent development programmes and on the collaborative arrangements through which they are delivered, and to specify where

that evidence is absent. Four questions follow. First, what is the extent, nature and distribution of peer-reviewed empirical literature published between 2015 and 2026 on national and system-scale digital talent development programmes? Second, how is programme performance conceptualised and operationalised, and where do the reported indicators sit across the four CIPP dimensions and across individual, organisational and sectoral or national outcome levels? Third, which governance and collaboration mechanisms, mapped to the eight components of the integrative framework for collaborative governance, are reported as enablers or barriers to programme performance? Fourth, what methodological, theoretical, contextual and measurement gaps follow, and what research agenda do they imply for strategic management and public policy? Each question carries a corresponding objective, stated in terms of what the review sets out to produce: to chart a descriptive map of the evidence base and its distribution; to build a taxonomy of the performance indicators reported and to locate that taxonomy across evaluation dimensions and outcome levels; to assemble a matrix of governance mechanisms against their reported valence; and to derive an evidence gap map, and a research agenda, from the preceding three.

2 Method

2.1 Design and Reporting

We conducted a scoping review following Arksey & O'Malley [25], as refined by Levac et al. [26] and operationalised through the JBI methodology for scoping reviews [29]. A scoping review was the appropriate design because our aim was to map the extent, range and nature of a conceptually heterogeneous evidence base and to identify gaps, rather than to estimate the effect of an intervention [24]. Reporting follows PRISMA-ScR [28]. An a priori protocol was written before screening began. Deviations from the protocol were dated and logged as they occurred; the single substantive amendment is reported in Section 2.3.

2.2 Eligibility Criteria

Eligibility followed the Population–Concept–Context (PCC) mnemonic [30] and is set out in full in Table 1. In brief, we included peer-reviewed articles and review articles, published in English between 2015 and 2026, that examined digital talent or digital skills development programmes initiated, mandated, funded or regulated by public actors and operating at national, subnational, sectoral or cross-national scale, and that addressed either programme performance or the governance arrangements linked to delivery or outcomes. Context was recorded as a variable rather than applied as a filter, and grey literature was excluded. Exclusion codes were hierarchical, so that a record breaching several criteria received a single, highest-priority code.

Table 1 *Eligibility criteria (PCC) and hierarchical exclusion codes.*

Element	Included	Excluded (code)
Population	Digital talent or digital skills development programme, policy or scheme that is (a) initiated, mandated, funded or regulated by a public actor, or delivered through an explicit public–private partnership, and (b) operating at national, subnational, sectoral or cross-national scale	Single-organisation or course-level provision without public mandate at system scale (E1); K-12 pupils or citizen digital literacy without a workforce objective (E2); no identifiable programme (E12); programme without a digital competence domain (E14)
Concept	Conceptualisation, assessment or measurement of programme performance, outcomes, impact or effectiveness; governance, coordination or collaboration arrangements linked to delivery or outcomes	Programme description without a performance or governance dimension (E3); instrument validation without programme context (E5); technology or system adoption rather than talent development (E13)
Context	Any country, World Bank income group or economic sector	Setting not identifiable (E6)
Sources	Peer-reviewed articles and review articles; quantitative, qualitative and mixed-methods designs; English; 2015–2026	Editorials, letters, commentaries, notes, book chapters, theses, grey literature (E7); non-English (E8); pre-2015 (E9); full text not retrieved (E10); no empirical basis (E15)

Note. Where a record breached more than one criterion, a single code was recorded following a fixed priority order: source eligibility first, then population, then concept, then context. Counts for each code are reported in Fig. 1.

2.3 Information Sources and Search Strategy

We searched Scopus on 1 August 2026 using the string reported in Table 2, which were developed by principal researcher. The string combines four Boolean blocks: the object (digital talent or skills), the intervention (development programme), the level and governance setting, and the analytical lens (performance, evaluation or governance). The search returned 265 records and required no calibration, since the yield fell within the target range of 100–450 specified in advance; the string was therefore never re-tuned after its results had been observed.

The protocol specified Scopus as database searched. Prisma-ScR checklists allowed full electronic strategy at least 1 database [28]. Therefore, coverage is limited to Scopus-indexed, English-language articles, which we report as a principal limitation rather than a comprehensive map. The decision to search a single database was taken on 1 August 2026, after the Scopus search returned a corpus within the pre-specified target range, on the judgment that a second full screening and charting cycle was disproportionate to the expected gain. That was a resource judgement, not an evidential one. We did not measure the overlap between Scopus and any other

database and we make no claim about it anywhere in this article. The likely direction of the resulting bias can be stated without being quantified; Scopus indexes public administration, management and economics journals well and education journals less completely than ERIC, while ProQuest would add dissertations and grey literature that the protocol excluded by design. The most probable consequence is under-representation of education-specific and non-English evidence, which would most affect the geographic map in Section 3.5 where coverage is already sparsely.

Table 2 Search string executed in Scopus, 1 August 2026.

Field	String
TITLE-ABS-KEY	("digital talent*" OR "digital skill*" OR "digital competenc*" OR "digital literac*" OR "digital workforce" OR "digital human capital" OR "ict skill*" OR "e-skill*" OR "tech* talent") AND ("talent development" OR "skill* development" OR "workforce development" OR "human capital development" OR "capacity building" OR upskill* OR reskill* OR "training program*" OR "training programme*" OR "development program*" OR "development programme*" OR "training initiative*" OR bootcamp* OR "digital academy" OR "vocational training" OR "scholarship program*") AND (national* OR government* OR "public sector" OR "public polic*" OR "public-private" OR "public private partnership*" OR ecosystem* OR "multi-stakeholder*" OR "cross-sector*" OR "state-led" OR "country-level") AND (performance OR outcome* OR impact* OR effectiveness OR evaluat* OR employab* OR "success factor*" OR governance)
Limits	PUBYEAR > 2014; DOCTYPE = article, review; LANGUAGE = English
Yield	265 records

2.4 Selection of Sources of Evidence

Screening proceeded in two stages. At each stage two reviewers ([L], [A]) screened independently under deliberately asymmetric decision thresholds, and a third reviewer ([LA]) adjudicated disagreements. At title and abstract, one reviewer applied a recall-first threshold and the other a precision-first threshold; ambiguity arising from the *absence* of information was resolved in favour of inclusion, since a relevant record lost at this stage cannot be recovered. At full text the burden of proof was reversed: inclusion required explicit positive evidence for each criterion. No "maybe" category was permitted at either stage, and every full-text exclusion was recorded with a verbatim supporting quotation and page number. The asymmetry was adopted because the consequences of the two error types are not symmetrical. At title and abstract a relevant record wrongly excluded is unrecoverable, since it never reaches the stage at which the evidence needed to judge it becomes available, whereas a record wrongly included costs only the time to read it. Setting one reviewer to resolve missing information ambiguity toward inclusion and the other toward exclusion ensured that the adjudicator saw every genuinely borderline record rather than only

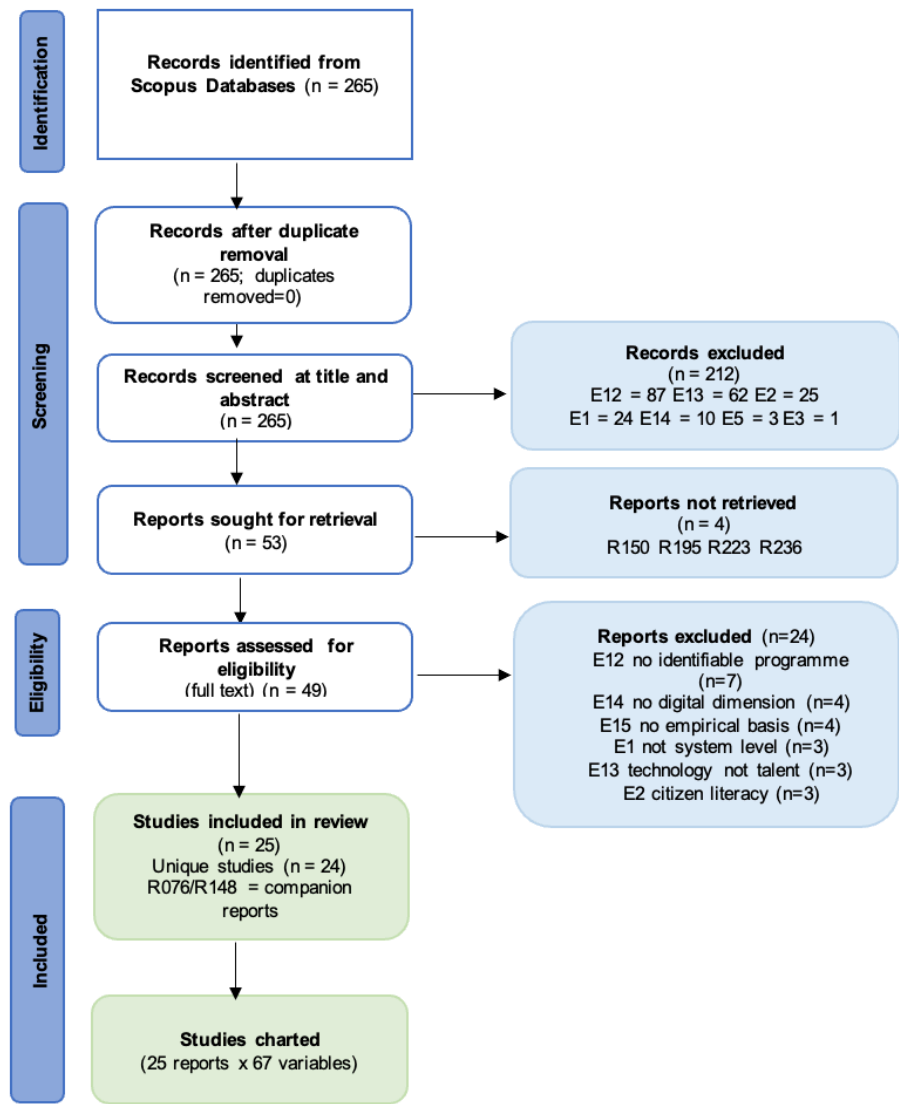
those on which two similarly calibrated reviews happened to differ. At full text the evidence is present, so the burden of proof was reversed with it. The design carries two acknowledged cost; it depressed Cohen’s kappa through the bias index, and it raises adjudication load. The third reviewer determined 48 of 265 decisions at title and abstract (18%) and 15 of 49 at full text (31%).

Agreement was 81.9% at title and abstract and 69.4% at full text ($\kappa = 0.43$ at both stages), with 48 and 15 disagreements adjudicated respectively. Because the thresholds were asymmetric by design, we also report the bias and prevalence indices and the prevalence-adjusted bias-adjusted kappa, which affect κ independently of observed agreement [31]; κ bands follow Landis & Koch [32]. Fig. 1 presents the flow of records through the review. The full cell counts are reported in Table 3 since the summary coefficients alone are not interpretable. The decisive feature is that the third cell is zero at both stages; disagreement was entirely one-directional, which is the signature of the asymmetric design and depresses kappa independently of how much the reviewers agreed. We do not claim this disposes of the concern. At full text PABAK is 0.3878, which is modest by any standard. What protects the full-text decisions is not the coefficient but the evidence rule: inclusion required explicit positive evidence for every criterion, quoted verbatim with page number, and each of the 15 disagreements was adjudicated on that quoted evidence rather than on reviewer impression.

Table 3. Reviewer agreement, full cell counts

	Title and abstract	Full text
Both include	25	21
Reviewer 1 include, Reviewer 2 exclude	48	15
Reviewer 1 exclude, Reviewer 2 include	0	0
Both exclude	192	13
Observed agreement	0.8189	0.6939
Cohen’s Kappa	0.4301	0.4262
PABAK	0.6377	0.3878
Bias index	0.1811	Not computed
Prevalence index	0.6302	Not computed
Adjudications by third reviewer	48 (18% of 265)	15 (31% of 49)

Note. Cell counts are reported because kappa is depressed by the asymmetric threshold design; PABAK is reported alongside following Byrt et al [33].



*Exclusion codes are defined in Table 1. Search Executed 1 August 2026

Fig. 1. PRISMA-ScR flow of the search and selection process. Adapted from Page et al.[34] and reported in accordance with PRISMA-ScR [28]. Exclusion codes are defined in Table 1.

2.5 Data Charting

We charted data using a 67-variable form developed from the JBI template [29], covering bibliographic details, study context, programme characteristics, methodology, performance constructs, governance mechanisms and author-stated

gaps. Charting was treated as an iterative process [26]: the form was piloted on five reports and revised before full extraction. Where a source did not report a variable we recorded "NR" rather than inferring a value, so that non-reporting itself became analysable. Each empirical report was appraised with the Mixed Methods Appraisal Tool [35]; ratings are reported descriptively only, and no report was excluded on methodological grounds. We report the division of labour on charting and coding precisely, because the two were not equivalent. Screening was independently duplicated at both stages. Charting and classification of indicators and mechanism were not: one researcher (L) performed the initial classification for all 25 reports, and a second researcher (A) verified every classification against the verbatim extract and its page reference, with disagreements resolved by discussion and the extract itself treated as the arbiter. This is verification of a single coder's work, not independent double-coding, and it does not yield an inter-rater coefficient. We therefore report no agreement statistic for the indicator and mechanism classifications and make no claim of independent duplication for them. What supports the classifications instead is auditability: every one of the 105 indicators and 98 mechanism is tied to a verbatim quotation with a page number.

2.6 Synthesis and Analysis

Analysis combined three strands. The first was a descriptive numerical summary of study, programme and methodological characteristics. The second was deductive coding: performance indicators were coded to the four dimensions of the CIPP evaluation model [10] and to three outcome levels, and governance mechanisms to the eight components of the integrative framework for collaborative governance [4], each mechanism receiving a valence of enabler, barrier, ambiguous or mixed. Because the design and delivery contrast reported in Section 3.4 rests on a classification that was not previously stated, we set out the rule here. Each of the 14 descriptive themes was assigned to the design stage, the delivery stage, or neither, on a single criterion fixed before the valence figures were computed: a theme belongs to design if the mechanism it describes operates before or at the point at which programme architecture, partnership structure, mandate or funding envelope is settled, and to delivery if it operates at or after the point at which training is provided to participants. Themes describing conditions wholly outside the programme boundary were assigned to neither. Design comprises DT1, DT2, DT4, and DT5, holding 48 units; delivery comprises DT3, DT6, DT7, and DT8, holding 27 units. The remaining six themes describe conditions that are not specific to either stage and are assigned to neither. The two stage categories together cover 75 of the 98 mechanism units, or 77%. The contrast is therefore a statement about those 75 units and not about the corpus as a whole, and report it as such.

The third was a three-stage thematic synthesis [36] applied to the 98 verbatim governance-mechanism extracts and 105 performance indicators charted from the included reports: line-by-line coding, grouping of codes into descriptive themes, and generation of analytical themes that extend beyond the primary studies. Mechanisms that could not be mapped to either framework were retained as inductive codes and

are reported as candidate framework extensions. We examined theme robustness through a saturation curve, cross-tabulation by World Bank income group, and comparison of mechanism valence against MMAT rating.

Gap statements were derived from cross-tabulations using a decision rule fixed in advance: cells containing no studies were classed as evidence-absent, one or two studies as evidence-scarce, and three or more as evidence-present.

3 Results

3.1 Characteristics of The Evidence Base

The searches yielded 265 records, of which 25 reports met all eligibility criteria (Fig. 1). Two of these reports, Choudhary and Bansal [37], [38], analyse the same corpus of 86 primary studies and are treated as companion publications, so the 25 reports represent 24 unique studies. All counts below are reported at the level of the report. Because the two companion reports describe a single programme and could inflate any pattern they both support, we tested their influence directly. Excluding the companion report reduces the mechanism corpus from 98 units to 94 and the indicator corpus from 105 to 101. The design-stage barrier proportion moves from 33% to 32% and the delivery-stage proportion from 70% to 68%. The excluded report contributes 4% of mechanism units and contributes to four descriptive themes (DT3, DT5, DT8, and DT9) without being the sole support for any theme. No finding reported below depends on retaining it. The rest of the descriptive profile is equally stable. Reports measuring reach remain 3, reports with a comparison group remain 1, and no enabler-valenced report reaches the highest appraisal in either case. Reports defining performance explicitly move from 14 to 13, self-report as the dominant instrument from 44% to 46%, and reports not stating a time horizon from 60% to 58%. The double weighting risk is therefore bounded at 4% of units and does not alter any count on which the synthesis rests.

Table 4 sets out each included report. The evidence base is recent: 20 of 25 reports (80%) were published from 2023 onwards and 15 (60%) in 2025 or 2026, with the earliest two appearing in 2018. Geographically it is uneven. East Asia and Pacific (8 reports) and Europe and Central Asia (7) together account for 15 of 25, while Sub-Saharan Africa contributes 2 and South Asia, North America and the Middle East and North Africa contribute 1 each; a further 5 reports are multi-region or global in coverage. By World Bank income group, upper-middle-income settings dominate (11 reports), followed by mixed-income coverage (6), lower-middle-income (4) and high-income (4). Eighteen reports (72%) examine a single country.

Table 4. Characteristics of the 25 included reports.

No.	Study	Setting (income)	Programme or policy studied	Scale	Design	MMAT criteria met
1	Shi et al. (2023) [39]	China (UM)	National Digital Competence Training Program for teachers	National	Mixed methods	3 of 4
2	Luo et al. (2026) [40]	China (UM)	Teacher Professional Development at Scale (TPD@Scale)	Subnational	Quantitative	4 of 4
3	Bansal & Choudhary (2024) [41]	India (LM)	Pradhan Mantri Gramin Digital 4Saksharta Abhiyan (PMGDISHA)	National	Qualitative	4 of 4
4	Shava & Muringa (2025) [42]	South Africa (UM)	National ICT policy framework and digital inclusion programmes	National	Review	3 of 4†
5	Kerrick & Cumberland (2026) [43]	United States (H)	Digital Literacy Train-the-Trainer Program (DLT3)	Subnational	Case study	2 of 4
6	Choudhary & Bansal (2022a) [44]	Multi-country (Mx)	Corpus of 86 digital literacy training programmes	Multi-country	Review	4 of 4†
7	Ou et al. (2025) [45]	China (UM)	National digital literacy programme for K-12 teachers	National	Quantitative	4 of 4
8	Sabdenov et al. (2025) [46]	Kazakhstan (UM)	Civil service digital recruitment and training system	National	Mixed methods	4 of 4
9	Tjokro et al. (2025) [47]	Indonesia, EU, USA (Mx)	National legal framework for digital human capital	Multi-country	Mixed methods	3 of 4
10	Chetty, Aneja, et al. (2018) [48]	G20 (Mx)	G20 digital skills strategies; National Qualifications Framework	Multi-country	Mixed methods	2 of 4
11	Choudhary & Bansal	Multi-country	Corpus of 86 digital literacy	Multi-country	Review	4 of 4†

No.	Study	Setting (income)	Programme or policy studied	Scale	Design	MMAT criteria met
	(2022b) [37]	(Mx)	training programmes (companion)			
12	Maiyapan et al. (2026) [49]	Thailand (UM)	Ministry of Public Health Digital Health Strategy	Subnational	Quantitative	4 of 4
13	Miltner et al. (2026) [50]	United Kingdom (H)	Government-backed coding and data science bootcamps	National	Quantitative	3 of 4
14	Widaningsih et al. (2025) [51]	Indonesia (UM)	BBPPMPV technical competency training for TVET teachers	Subnational	Quantitative	3 of 4
15	Bilan et al. (2023) [52]	Ukraine (LM)	Advanced civil service training, National Agency for Civil Service	National	Quantitative	3 of 4
16	Yooyativong & Kamyod (2023) [53]	Thailand (UM)	Digital Farmer Development Framework (NBTC fund)	Subnational	Case study	2 of 4
17	Assanova et al. (2025) [54]	Kazakhstan (UM)	Digital Kazakhstan; Astana Hub; National AI Platform	National	Mixed methods	4 of 4
18	Burns et al. (2020) [55]	Serbia, Russia (UM)	Pro-VET (Erasmus+ CBHE); VET Educator Trainer Network	Multi-country	Mixed methods	3 of 4
19	Adewojo (2026) [56]	Africa, multi-country (Mx)	Library digital literacy capacity-building initiatives	Multi-country	Review	4 of 4†
20	Hackl & Najdi (2024) [57]	Lebanon (LM)	Digital Skills Training programme (World Food Programme)	National	Qualitative	4 of 4
21	Dunan et al. (2025) [58]	Indonesia (UM)	Digital Talent Scholarship and	National	Qualitative	4 of 4

No.	Study	Setting (income)	Programme or policy studied	Scale	Design	MMAT criteria met
22	Chetty, Liu, et al. (2018a) [59]	G20 (Mx)	ICT Jamboree (Kominfo) G20 digital training management and evaluation mechanisms	Multi-country	Review	2 of 4†
23	Hachem et al. (2025) [60]	Sweden (H)	Accreditation of an AI mass-market MOOC	National	Case study	4 of 4
24	Tu et al. (2025) [61]	Vietnam (LM)	Digital skills training in higher education institutions	National	Mixed methods	3 of 4
25	Esposito et al. (2025) [62]	Italy (H)	Rome Technopole Spoke 4 AI ethics pilot module	Subnational	Mixed methods	3 of 4

Note. Income group: H = high; UM = upper-middle; LM = lower-middle; Mx = mixed coverage. MMAT ratings (Hong et al., 2018) are descriptive and were not used to exclude reports. The MMAT column should be read only as a count of criteria met, not as a quality score. Five of the 25 reports use review designs and fall outside the tool's stated scope; they are marked with † accordingly and are excluded from the appraisal-based sensitivity analysis reported in Section 3.3. † Review design; the count is shown for transparency only, and these five reports are also excluded from Fig.3, Panel B.

Table 5 aggregates the study and programme characteristics. Governments initiate most of the programmes studied (15 of 25), and government appears as a partner in 24 of 25 (96%), but the partnership base beyond the state is thin: industry appears in 11 reports (44%), training providers in 10 (40%), and universities and multilateral bodies in 9 each (36%). Public-only funding covers 12 programmes, while donor funding, co-funding and mixed models cover 11 between them. Methodologically the corpus is dominated by mixed-methods (8) and quantitative (6) designs, and appraisal quality is generally sound: 21 of 25 reports (84%) attained three or four MMAT stars. Sixteen reports (64%) state an explicit theoretical frame. National-scale programmes (12) outnumber multi-country (7) and subnational (6) ones. Education is the single largest sector (8 reports) and the largest contributing discipline (11 reports), which means that the literature on national digital talent programmes is being written mainly from within education research rather than from public administration or management.

Table 5. Aggregate characteristics of the included reports (n = 25).

Characteristic	Categories (n)
Publication year	2025 (10); 2026 (5); 2023 (3); 2018 (2); 2022 (2); 2024 (2); 2020 (1)
World region	East Asia & Pacific (8); Europe & Central Asia (7); global or multi-region (5); Sub-Saharan Africa (2); Middle East & North Africa (1); North America (1); South Asia (1)
Income group	Upper-middle (11); mixed (6); lower-middle (4); high (4)
Programme scale	National (12); multi-country (7); subnational (6)
Initiating actor	Government (15); public-private partnership (6); multilateral (3); industry-led under public mandate (1)
Funding model	Public only (12); mixed (4); donor-funded (4); co-funded (3); not reported (2)
Partner types present	Government (24); industry (11); training provider (10); multilateral (9); university (9); non-governmental organisation (7)
Economic sector	Cross-sector (10); education (8); public administration (3); ICT (2); health (1); agriculture (1)
Study design	Mixed methods (8); quantitative (6); review (5); qualitative (3); case study (3)
Contributing discipline	Education (11); public administration (4); economics (3); interdisciplinary (3); sociology (2); information systems (1); management (1)
MMAT rating	**** (12); *** (9); ** (4)
Explicit theory stated	Yes (16); no (9)

Note. Partner types are not mutually exclusive; a report may name several

Two further features of this evidence base are worth recording because they constrain what can be asked of it. First, the unit of analysis is fragmented: 8 reports analyse individuals, 7 the programme itself, 4 whole countries, 4 several levels at once, and 1 each an organisation and a sector. Comparison across reports is therefore rarely comparison of like with like. Second, programme maturity is almost never stated. Only 3 of the 25 reports (12%) record how long the programme had been running at the time of study, so in 22 cases the reader cannot tell whether a reported outcome reflects a mature system or a pilot. Sample sizes vary by four orders of magnitude, from 8 interviews in a single university [60] to 45,648 participating teachers in one province [39], which reflects the breadth of designs rather than any weakness in individual reports.

3.2 Conceptualisation and Measurement of Programme Performance

Performance is more often assumed than defined. Fourteen reports (56%) state an explicit performance construct; the remaining 11 (44%) report indicators without defining what performance means. Even among the 14, the construct is usually stated as an aim rather than specified as a measurable object, as in the assessment of “the implementation and effectiveness of a national plan for a digital competence training program” [39] or the question whether bootcamps “make good on their key claims,

particularly around employment and salaries” [50]. Chetty, Aneja, et al. [48] are unusual in treating the measurement of programme performance as itself the object of inquiry. Across the corpus, 105 distinct performance indicators were charted, a mean of 4.2 per report.

Where a theoretical anchor is used, it is more often a competence taxonomy than a theory of programme performance. Of the 16 reports (64%) stating an explicit frame, 6 build on an established digital competence taxonomy such as DigComp, the UNESCO Digital Literacy Global Framework or a national equivalent. The remaining 10 draw on general social science theory. Human capital theory is the most frequent of these and appears in only 3 reports, empowerment theory in 2, and the rest once each, so there is no shared theoretical vocabulary in this field. Nine reports (36%) state no theoretical frame. In two of these the absence is itself part of the argument: Choudhary & Bansal [37] report that the training programmes they reviewed lacked any theoretical grounding, and Adewojo [56] calls for an Africa-centred digital literacy framework in place of imported ones.

Coverage of the CIPP dimensions is broad but front-loaded. Input is addressed by 23 reports (92%) and Context by 22 (88%), whereas Process is addressed by 18 (72%) and Product by 17 (68%). Twelve reports (48%) touch all four dimensions and only one addresses a single dimension. The imbalance is sharper once outcome level is

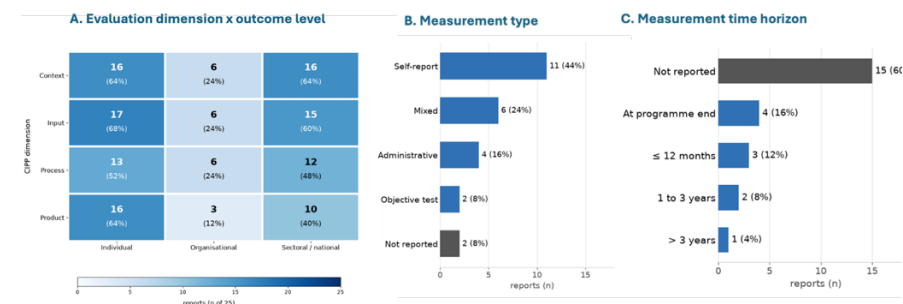


Fig. 2. Measurement practice across the 25 included reports. Panel A: CIPP evaluation dimension [10] by outcome level, with cell values showing the number of reports. Panels B and C: measurement type and measurement time horizon

added. Fig.2 maps the two together. Individual-level outcomes are reported by 19 reports (76%) and sectoral or national outcomes by 16 (64%), but organisational outcomes by only 6 (24%). The weakest cell in the matrix is Product at organisational level (3 reports, 12%), followed by Product at sectoral or national level (10 reports, 40%). In other words, the literature describes what is put into these programmes and the systems around them far more thoroughly than it describes what comes out of them above the level of the individual learner.

Measurement practice itself is thin (Fig. 2, panels B and C). Self-report is the dominant instrument, used by 11 reports (44%), with a further 6 (24%) combining self-report and other sources; administrative data are used by 4 (16%) and objective

tests by only 2 (8%). Time horizon is not reported at all by 15 reports (60%). Where it is reported, it is short: 4 reports (16%) measure at programme completion and 3 (12%) within 12 months, while 2 (8%) reach one to three years and 1 (4%) exceeds three years. A comparison group appears in a single report (4%), and even there the comparison is between publicly and privately funded training institutes rather than between programme participants and non-participants [39]. No report in the corpus estimates a programme counterfactual.

Table 6 groups the 105 indicators into 11 empirically derived themes. Two contrasts stand out. First, indicators capturing participant perception or self-assessment appear in 8 reports (32%), twice as many as the 4 reports (16%) carrying labour-market or economic outcome indicators. Second, indicators of behavioural adoption or transfer into practice appear in just 2 reports (8%), so the link between competence acquired and competence used is almost entirely unexamined. Reach, access and participation is the most frequent theme (18 indicators across 12 reports), yet as Section 3.4 shows, reach is measured in a way that permits interpretation in only a small minority of cases.

Table 6. Taxonomy of the 105 charted performance indicators.

Code	Indicator Theme	CIPP dimension	Indicators	Reports (%)
IT4	Reach, access and participation	Process	18	12 (48)
IT8	Equity and disparity	Cross-cutting	12	11 (44)
IT9	Participant perception and self-assessment	Product, self-reported	15	8 (32)
IT5	Delivery inputs and provider quality	Input	10	7 (28)
IT6	Institutional and governance arrangements	Context / Input	10	7 (28)
IT1	Competence attainment	Product, individual	9	7 (28)
IT11	Measurement infrastructure	Context	9	6 (24)
IT3	Labour market and economic outcomes	Product, individual and national	6	4 (16)
IT10	Macro-system indicators	Product, sectoral and national	8	3 (12)
IT7	Financial sustainability	Input / Process	3	3 (12)
IT2	Behavioural adoption and transfer to practice	Product, individual	5	2 (8)

3.3 Governance Mechanisms and Their Reported Valence

Ninety-eight verbatim governance mechanisms were charted, a mean of 3.9 per report. Coverage of the collaborative governance framework is complete at the structural end and sparse at the relational end (Fig. 3, panel A). System context and drivers are present in all 25 reports (100%) and capacity for joint action in 24 (96%),

whereas principled engagement appears in 13 (52%) and shared motivation in only 11 (44%). Sixteen reports (64%) address six or more of the eight components, so the gap is not one of shallow coverage overall but of a consistent silence about the relational dynamics through which collaboration is sustained.

Valence is predominantly negative. At report level, 11 reports (44%) carry a barrier-dominant reading of the governance arrangements they describe, 9 (36%) a mixed reading, 3 (12%) an enabling reading and 2 (8%) an ambiguous one. Panel B of Fig. 3 tests whether this negative tilt is an artefact of weaker evidence, and the result runs the other way. One check on this reading deserves separate statement, and excluded the five review-design reports that fall outside the MMAT’s stated scope. Among the 20 empirical reports, no report carrying an enabler-dominant reading meets all four applicable MMAT criteria, while 3 of the 8 barrier-dominant empirical reports do. The negative reading therefore rests on the methodologically stronger reports rather than the weaker ones, which is the opposite of what a problem-reporting bias would produce. Before excluding the review-design reports the corresponding figures were 0 of the enabler-dominant and 5 of 11 barrier-dominant reports; the direction is unchanged.

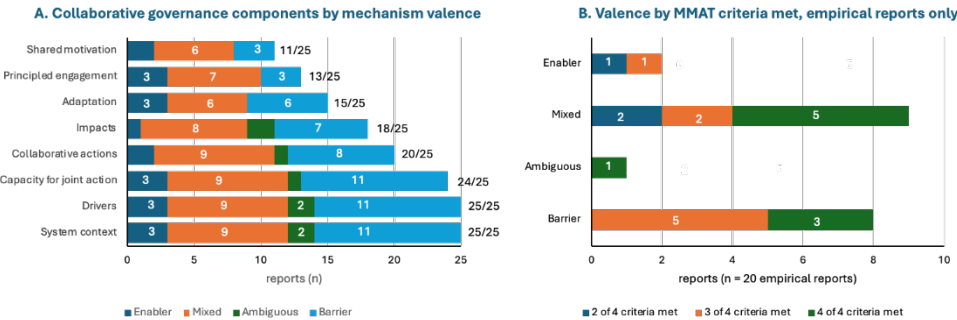


Fig. 3. Governance mechanisms across the 25 included reports. Panel A: the eight components of the integrative framework for collaborative governance Emerson et al. [4] by mechanism valence, ordered by coverage. Panel B: report-level mechanism valence by the number of MMAT criteria met, restricted to the 20 empirical reports; the five review-design reports are excluded because the tool cannot be used to appraise reviews [63].

Twenty-four of the 25 reports (96%) also yielded at least one mechanism that could not be mapped onto the framework. These inductive codes cluster into recognisable groups: standardisation instruments such as national qualifications frameworks [48], [51]; measurement infrastructure as a precondition for evidence-based governance [46], [59]; financial sustainability tied to donor cycles [56]; post-programme labour-market filters [50]; and access preconditions such as assistive technology and language [54], [58].

Line-by-line coding of the 98 mechanisms produced 98 descriptive codes and 14 descriptive themes, a mean of 3.4 themes per report (range 2 to 4). Fig. 2 shows the resulting structure and Table 6 reports the themes with their coverage and the share of

units coded as barriers. The last new descriptive theme appeared at the twentieth of 25 reports, so saturation was approached but is not claimed.

The two most widely shared themes describe partnership architecture. DT4, the most frequent, records formal multi-actor vehicles, as in the account of technical academies operating “under a Foundation of Participation model, which brings together public and private stakeholders, including companies, universities, research centers, local authorities” [62]. DT2 records standardisation instruments, typically qualification frameworks used to align provision with employer demand [48]. Against these, the themes concentrated at the delivery end read very differently. DT6 records capacity that thins out below the centre: “Despite progressive strategies such as the National Integrated ICT Policy White Paper and the Digital Economy Masterplan, provincial and municipal levels of implementation are partial” [42].

Table 7. Fourteen descriptive themes derived from 98 governance-mechanism units.

Code	Descriptive theme	Units	Reports (%)	Units coded barrier (%)	Stage
DT4	Multi-actor partnership as delivery vehicle	16	13 (52)	5 (31)	Design
DT5	Structured engagement processes	11	11 (44)	2 (18)	Design
DT1	Policy mandate as initiating force	10	10 (40)	4 (40)	Design
DT2	Standards and qualification infrastructure as coordination device	11	9 (36)	5 (45)	Design
DT3	Financing arrangements govern continuity	11	8 (32)	6 (55)	Delivery
DT6	Uneven implementation capacity across levels and regions	9	7 (28)	8 (89)	Delivery
DT9	Thin monitoring and evaluation infrastructure	6	5 (20)	2 (33)	Not stage-specific
DT10	Exclusion mechanisms operating beyond programme control	5	4 (16)	3 (60)	Not stage-specific
DT8	Weak control of provider quality and delivery design	4	4 (16)	2 (50)	Delivery
DT14	Policy support alone yields no return	4	4 (16)	2 (50)	Not stage-specific
DT7	Low and rarely measured reach	3	3 (12)	3 (100)	Delivery
DT12	Adaptive workarounds and successor policies	3	3 (12)	1 (33)	Not stage-specific
DT11	Legitimacy withheld or denied by the state	3	2 (8)	1 (33)	Not stage-specific
DT13	Endogenous capacity created by the programme	2	2 (8)	0 (0)	Not stage-specific

3.4 Analytical Themes

Six analytical themes were generated from the descriptive themes (Fig.4). The first concerns where fragmentation occurs. Grouping the descriptive themes by locus, the

four themes describing programme design and formation (DT1, DT2, DT4, DT5) contain 48 units of which 16 (33%) are coded as barriers, whereas the four themes describing delivery and implementation (DT3, DT6, DT7, DT8) contain 27 units of which 19 (70%) are coded as barriers. This literature documents collaboration at the point where programmes are designed and fragmentation at the point where they are delivered. DT6, uneven implementation capacity, is the clearest single case, with 8 of its 9 units coded as barriers across 7 reports. 23 units or 23,5% from corpus (DT9, DT10, DT11, DT12, DT13, DT14) were excluded since they are conceptually implementation. Two qualifications belong with this contrast. It is computed on the 75 of 98 mechanism units that the stage rule in Section 2.6 assigns to one stage or the other, so it describes 77% of the corpus rather than all of it. Because the contrast carries much of the article's argument, we subjected it to 17 alternative specifications, reported in Table 8: each of the eight stage-assigned themes was removed in turn, each was reassigned to the opposing stage in turn, and the companion report was excluded. The gap is 37 percentage points at baseline and ranges from 17 to 48 points across the 17 alternatives. It does not reverse or close in any of them. The narrowest result, 17 points, arises when DT5 is reassigned from design to delivery; the widest, 48 points, when DT3 is removed from the delivery group. Sensitivity of the design and delivery contrast are shown in Table 8.

Table 8. Sensitivity of the design and delivery contrast to 17 alternative specifications.

Specification	Design barrier % (n units)	Delivery barrier % (n units)	Gap (pp)
Baseline: design DT1, DT2, DT4,DT5;delivery DT3,DT6, DT7,DT8	33(48)	70(27)	37
Leave out DT1	32(38)	70(27)	38
Leave out DT2	30(37)	70(27)	40
Leave out DT4	34(32)	70(27)	32
Leave out DT5	38(37)	70(27)	32
Leave out DT3	33(48)	81(16)	48
Leave out DT6	33(48)	61(18)	28
Leave out DT7	33(48)	67(24)	34
Leave out DT8	33(48)	74(23)	41
Move DT1 to delivery	32(48)	62(37)	30
Move DT2 to delivery	30(37)	63(38)	33
Move DT4 to delivery	34(32)	56(43)	22
Move DT5 to delivery	38(37)	55(38)	17
Move DT3 to delivery	37(59)	81(16)	44
Move DT6 to delivery	42(57)	61(18)	19
Move DT7 to delivery	37(51)	67(24)	30
Move DT8 to delivery	35(52)	74(23)	39
Companion report excluded(94 units)	32(47)	68(25)	36
Specifications in which the gap reverse or closes			0 of 18

Note: Percentages are of coded mechanism units at each stage carrying a barrier valence. Computed from the coded mechanism file.

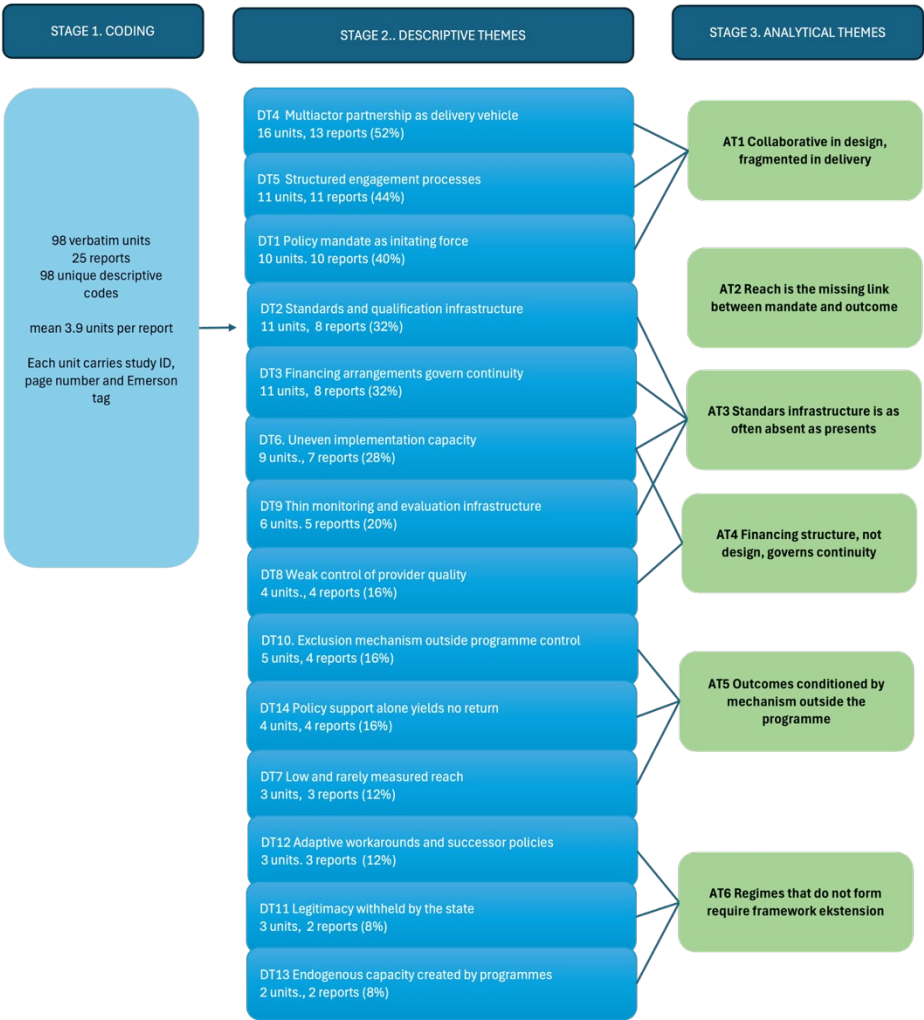
The second concerns reach. Only 3 of the 25 reports (12%) measure who actually received training, and all three find that the majority did not. In Kazakhstan, 56.5% of civil servants had never received formal digital skills training after recruitment [46]; in Thailand, 74.19% of primary healthcare personnel had never attended digital health training [49]; in West Java, government-run training had reached only 19.4% of vocational teachers [51]. Three countries, three sectors and three designs converge without exception on the same result.

The third concerns standards. Of the 11 units in DT2, 5 (45%) describe a competence or qualification framework as absent, not yet mandated or still to be developed rather than as an instrument in operation [46], [47], [52], [61]. DT2 is one of five themes appearing across all four income groups, so the absence of standards infrastructure is not confined to lower-income settings.

The fourth concerns continuity. DT3 appears in 8 reports and is barrier-dominant in 5 of them. What is reported is structural rather than a general shortage of funds: 92% of donor-funded library projects ceased once grants expired, with consequent staff turnover and loss of institutional memory [56]; allocations were withheld or diverted, producing delays [42]; and absent funds blocked technology implementation in otherwise well-designed programmes [55].

The fifth concerns determinants outside the programme boundary. DT10 and DT14 together describe regulatory exclusion of refugees from digital platforms and financial systems [57], labour-market screening through narrow notions of cultural fit that operate against bootcamp graduates [50], geographic concentration and language barriers [54], and assistive-technology availability as a precondition for using acquired skills [58]. The associated outcome evidence is consistent: only 13% of participants in the Digital Skills Training programme were employed within 12 months of completion [57].

The sixth concerns regimes that do not form. Three configurations in the corpus fall outside a framework that presumes a collaborative regime has been established. In the first, legitimacy is sought and refused: efforts by the funding agency “to obtain governmental endorsement of digital skills training and online work for displaced Syrians in Lebanon were unsuccessful” [57]. In the second, the state declines to build a regime at all, having “chose not to systematise MOOC-related activities on a national level” [60]. In the third, institutional actors diverge from national system rules in order to make delivery possible, altering admission parameters to work around system limitations [60]. These configurations appear in only 2 reports, but in each the absence or refusal of a regime, rather than its internal dynamics, determined the outcome. The evidential basis for this theme should be stated plainly rather than left to be inferred. Three configurations is a small number, and they come from two reports. We therefore treat this theme as a hypothesis-generating observation about a boundary condition of the framework, not as an established finding, and Section 4.6 develops it on than footing.



Saturation: the last new descriptive theme appeared at report 20 of 25. Mean 3.4 themes per report (range 2-4)

Fig. 4. Structure of the thematic synthesis, from 98 verbatim mechanism units through 14 descriptive themes to 6 analytical themes, following Three Stage Synthesis by Thomas & Harden [36].

Applying the pre-specified decision rule to the cross-tabulations produces a clear separation between conceptual and geographic coverage. The CIPP by outcome level matrix contains no empty cells: all 12 cells are evidence-present, so gaps in the measurement of performance are a matter of depth rather than absence. The governance matrix is more uneven. Of its 32 cells, 21 are evidence-present, 8 evidence-scarce and 3 evidence-absent, the last being ambiguous mechanisms within

principled engagement, shared motivation and adaptation. Enabling mechanisms are scarce in three components: shared motivation, collaborative actions and impacts.

The geographic map is the sparsest of the three. Table 9 shows that of 21 region by scale cells, only 4 are evidence-present, 7 are evidence-scarce and 10 are evidence-absent. National-scale evidence exists in quantity for East Asia and Pacific and for Europe and Central Asia alone. There is no multi-country evidence for East Asia and Pacific, the Middle East and North Africa, North America or South Asia, and no subnational evidence outside East Asia and Pacific, Europe and Central Asia and North America.

Table 9. Evidence gap map by world region and programme scale.

World region	Multi-country	National	Subnational
East Asia & Pacific	Absent	Present	Present
Europe & Central Asia	Scarce	Present	Scarce
Global or multi-region	Present	Absent	Absent
Middle East & North Africa	Absent	Scarce	Absent
North America	Absent	Absent	Scarce
South Asia	Absent	Scarce	Absent
Sub-Saharan Africa	Scarce	Scarce	Absent

Note. Classification followed the rule fixed in advance and reported in Section 2.6: no studies = evidence-absent; one or two = evidence-scarce; three or more = evidence-present.

4 Discussion

4.1 Collaboration is Theorised Where It Holds and Left Unexamined Where It Breaks

The most robust pattern in this corpus is positional rather than quantitative. Governance mechanisms coded to programme design and formation carry a barrier valence in 33% of units, while those coded to delivery and implementation do so in 70%. The literature therefore documents collaboration at the point where programmes are constituted and fragmentation at the point where they are run. The component distribution in Fig. 4 reinforces this. System context and drivers appear in all 25 reports and capacity for joint action in 24, yet principled engagement appears in 13 and shared motivation in 11. What is consistently present is the structural scaffolding of collaboration; what is consistently missing is the relational work that sustains it.

This should not be read only as a gap in the literature. It is also a limit in how the integrative framework [4] has been operationalised. The near-universal coding of capacity for joint action alongside a barrier valence in 8 of the 9 units on uneven implementation capacity shows that the component is being recorded as an intention rather than as an attainment. Recent work makes the same distinction from other directions: performance management inside collaborative arrangements remains under-studied and rarely connected to how those arrangements were designed [8] jointly agreed institutional settlements can turn into standoffs precisely at the implementation stage [64]; and the informational capacity of the implementing

bureaucracy, not the design of the policy, determines whether citizens can actually be served [15]. The problem is generic to multi-level systems, where policy capacity is unevenly distributed across tiers [65], [66] and where design and implementation are routinely governed by different actors [9], [67]. Our contribution is to show that in the digital talent domain the split is not incidental. It is where the barrier language concentrates.

4.2 Reach, Not Effect Size, is the Binding Constraint on This Evidence Base

Only 3 of 25 reports measure who actually received training, and all three find that most of the target population did not: 56.5% of Kazakh civil servants had never received formal digital skills training [46], 74.19% of Thai primary healthcare personnel had never attended digital health training [49], and provincial training had reached 19.4% of Indonesian vocational teachers [51]. Three countries, three sectors and three designs converge without exception.

The critical implication is that this field is debating effectiveness before it has established coverage. An effect estimated on participants says nothing about a programme whose participants are a small and self-selected fraction of its mandate. What is striking is that the corpus never frames non-participation in the vocabulary that exists for it. Public administration research has established that awareness and learning costs, not preference, drive much of the gap between eligibility and participation [68], that reducing those costs measurably raises take-up [14], and that burden is better conceptualised through onerous experience than through simple compliance cost [69]. The evidence that vulnerable citizens perceive higher burden is itself contested and needs careful testing [70], which makes the absence of any burden measurement in our corpus more consequential rather than less. A literature that could explain the field's central empirical finding is simply not being drawn on.

Measurement practice compounds the problem. Self-report is the dominant instrument (44%), yet self-appraised and demonstrated digital competence diverge systematically and unevenly across groups [11], and measuring digital skill in Global South settings requires instruments validated for those settings rather than imported ones [3]. Time horizon is unreported in 60% of reports, and where reported it is short. This is disqualifying given that the labour market effects of training in a middle-income setting emerge only over multi-year horizons and can reverse sign in the interim [12]. Comparable programme evaluations that do follow beneficiaries find modest and conditional effects [71], and outcomes depend heavily on how individualised delivery is [72]. Against this, a corpus in which one report has a comparison group and none estimates a counterfactual cannot support the effectiveness claims made of national digital talent programmes.

4.3 The Remedy the Field Requests is the One Specialists Question

Standards infrastructure is the most frequently proposed solution in this corpus and one of the least frequently realised: 5 of the 11 units in this theme describe a competence or qualification framework as absent, not yet mandated or still to be built, and the theme appears across all four income groups. The reflex is to recommend a

unified framework. That recommendation deserves more scrutiny than the corpus gives it. Allais (2023) shows that occupational standards and competency-based qualifications are extensively adopted in African vocational systems yet deliver limited value, and argues that the skills anticipation apparatus itself needs rethinking. Even the European reference taxonomy carries substantive gaps, with soft skills only implicitly represented in ESCO [73]. The critical reading is therefore not that these systems lack frameworks and should acquire them, but that frameworks are being treated as a solved instrument when their effectiveness is contested. Adewojo's [56] call for an Africa-centred framework, rather than a further import, is closer to the specialist position than the generic calls elsewhere in the corpus.

4.4 Continuity is a Property of Financing Structure, Not of Programme Design

Financing arrangements appear in 8 reports and are barrier-dominant in 5. What is reported is structural: 92% of donor-funded library projects ceased when grants expired, with staff turnover and loss of institutional memory [56] and allocations were withheld or diverted elsewhere [42]. This matches evidence from adjacent fields that donor dependency produces organisational precarity rather than merely funding shortfalls [74], and it supports the argument that externally financed capacity building tends to build capacity that does not stay [75]. The framework consequence is specific. Adaptation, as specified by Emerson et al. [4], captures adjustment within a continuing regime; it does not capture termination. Continuity therefore has to be modelled as an outcome with its own determinants rather than assumed as a background condition.

4.5 Outcomes are Conditioned Beyond the Programme Boundary, Against a Moving Target

Two themes in this corpus locate the determinants of outcome outside the programme: regulatory exclusion from digital platforms and financial systems, labour market screening, and access preconditions such as assistive technology and language. The associated outcome evidence is stark, with 13% of participants employed within 12 months of completion in the strongest documented case [57], a pattern also reported for digital livelihood schemes in the same region [76]. The mechanism our corpus describes qualitatively as cultural fit screening has a precise counterpart in the economics literature: culture used as a hiring criterion produces systemic discrimination even inside procedurally fair processes [77]. Programme evaluation that ignores these filters risks attributing to the programme a failure generated after it ends.

The target is also moving. Large-scale analyses of job postings show artificial intelligence reshaping the structure of skill demand [1] and altering hiring behaviour directly [2]. A field that measures competence by self-report at programme completion is therefore calibrating against a specification that changes faster than its own measurement cycle. Even where competence is acquired, transfer into practice is neither automatic nor frequent, and depends on workplace support conditions that this

corpus almost never charts [78], [79]. That behavioural adoption and transfer appear in only 2 of 25 reports is consistent with, and partly explained by, that omission.

4.6 Regime Status as A Candidate Prior Variable: A Proposition for Testing

The integrative framework presumes that a collaborative governance regime has formed and then analyses its dynamics. Three configurations in this corpus fall outside that premise: legitimacy sought and refused [57], a state declining to systematise activity nationally, and institutional actors diverging from national rules to make delivery possible [60]. In each, the absence or refusal of a regime, not its internal dynamics, determined the outcome. We therefore propose regime status as a prior variable with four states, formed, refused, declined and circumvented, together with three additions arising from Sections 4.2 to 4.5. Fig. 5 sets out the proposed extension against the coverage observed here. The proposal is consistent with work showing that policy legitimacy is actively constructed and contested through institutional work rather than given [80]. We advance it as a hypothesis generated by 2 reports, not as a validated construct. The strength of this claim requires an explicit statement of its evidential base, which is narrow. The three non-forming configurations that motivate it come from two of the 25 reports. Two reports cannot establish a boundary condition of a framework as widely used as this one. We therefore advance the extension as a proposition to be tested rather than a validated modification, and we state what would test it: a study design that samples intended collaborations rather than realised ones. So that in cases in which a regime was mandated but never formed remain observable rather than dropping out of the sampling frame before analysis begins. The value of the proposition, if it has any, lies in identifying that sampling problem, not in the three cases themselves. Figure 5 should accordingly be read as a proposed research agenda.

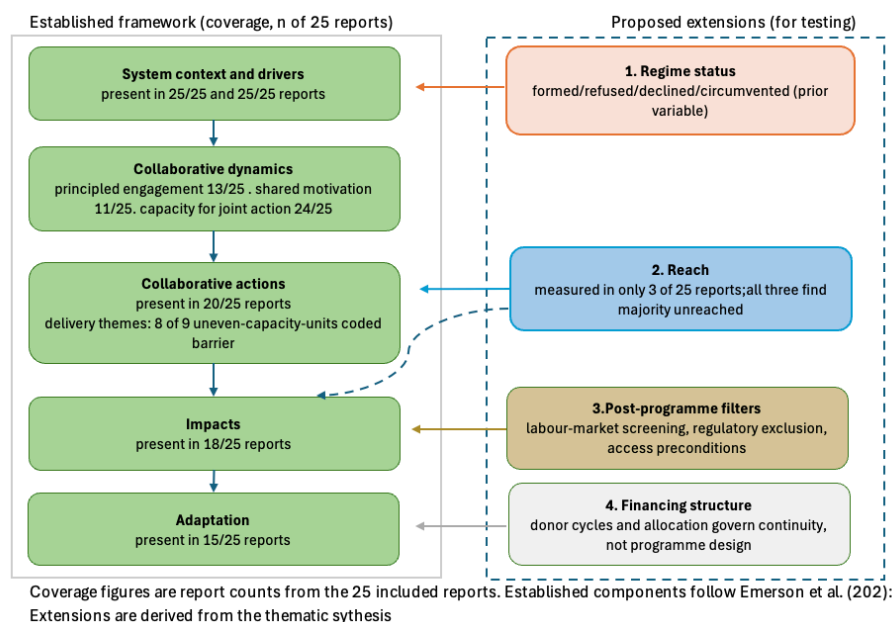


Fig. 5. Proposed extension of the integrative framework for collaborative governance [4]. Left: established components with their coverage in this review. Right: four extensions derived from the thematic synthesis. Proposed rather than established: the additions shown here are derived from three configurations in two reports and are advanced as propositions for testing.

4.7 Limitations

Four limitations bound these conclusions. First, coverage is restricted to Scopus-indexed, English-language journal articles, so grey literature and non-English evidence are absent. Several large national programmes that are prominent in practice do not appear in the corpus at all. Second, 25 reports is a small base, and descriptive theme saturation was approached but not reached, with the last new theme appearing at the twentieth report. Third, journal quartile could not be verified for any report, so no claim about publication venue quality is made anywhere in this article. The MMAT results should be read as descriptive appraisal, not as a quality filter. One counterargument deserves direct treatment. The corpus reads pessimistically, and a reader may reasonably suspect publication bias toward problems. The evidence runs the other way. A fourth limitation concerns the coding itself. The classification of the 105 indicators and 98 mechanisms was performed by one researcher and verified by a second against verbatim extracts, rather than independently duplicated, so no interrater coefficient is available for these classifications and the possibility of systematic single-coder drift cannot be excluded statistically. It is bounded instead by the requirement that every classification be traceable to a quoted extract with a page reference and all of which are available for inspection. On the first limitation, the direction of the likely bias can be named even though its size cannot: Scopus indexes journals less completely than ERIC, so education-specific evidence is the most

probable omission, and the geographic map in Section 3.5 is the result most exposed to it.

4.8 Implications

For research, three priorities follow. Reach should be reported as a standard indicator, since without it no effect estimate is interpretable. Outcome measurement should extend to at least 12 months and use administrative or assessed data rather than self-report. Evaluations should model post-programme filters explicitly instead of treating them as noise. For policy, the practical implication is that partnership architecture at the design stage is not evidence of delivery capacity, and that financing structure should be treated as a design variable with consequences for continuity rather than as an administrative detail.

5 Conclusion

This scoping review charted 25 peer-reviewed reports on publicly mandated digital talent development programmes and synthesised the 98 governance mechanisms and 105 performance indicators they report. The evidence base is recent, geographically concentrated and thinnest at the point that matters most. Performance is defined explicitly in just over half the reports, measured predominantly by self-report, and seldom followed beyond programme completion. Only three establish how much of the intended population was reached, and all three find that most was not.

Two findings hold beyond description. Governance is documented as collaborative where these programmes are designed and as fragmented where they are delivered, with barrier-coded mechanisms rising from 33% to 70% between the two. And the pessimistic reading of this literature cannot be dismissed as a bias toward reporting problems, because it rests on the methodologically stronger reports, not the weaker ones. From these we argue that regime status, financing structure, reach and post-programme filters should be treated as variables, not background assumptions. The framework extension we draw from this rests on three configurations in two reports and is offered as a proposition for testing, not as a validated modification.

The practical conclusion is narrow and firm. Until reach is reported as standard and outcomes are measured beyond twelve months using data that is not self-reported, claims about the effectiveness of national digital talent programmes cannot be evaluated. That is a fixable problem, and fixing it is the precondition for everything else this field wants to know.

Acknowledgment

The authors would like to express sincere gratitude to Ministry of Communications and Digital of Indonesia for financial support and Universitas Pendidikan Indonesia for providing the facilities necessary to complete this study.

Declaration of competing interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of generative AI and AI-assisted technologies. During the preparation of this work, the author(s) used Claude to assist with literatures search and QuillBot to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication. The authors take full responsibility for the content of the published article.

Data availability. The data supporting this study are available from the corresponding author on reasonable request. No supplementary files accompany this article.

Ethics. This review analyzed published literature only. No human participants were involved, and no ethical approval was required.

References

- [1] L. Zhang and C. Zhang, "Toward sustainable workforce development: How AI reshapes skill demand structure—Evidence from 67 million job postings in China," *Sustainability*, vol. 18, no. 10, p. 4905, 2026, doi: 10.3390/su18104905.
- [2] E. Engberg *et al.*, "Artificial intelligence, hiring and employment: Job postings evidence from Sweden," *Appl. Econ. Lett.*, 2025, doi: 10.1080/13504851.2025.2497431.
- [3] D. I. Yasmine, F. Colombijn, A. J. A. M. van Deursen, and E. van Ingen, "Internet skills, digital engagement, and outcomes among urban youth in Jakarta, Indonesia: Incorporating a Global South context in measuring digital inequality," *Inf. Commun. Soc.*, vol. 28, no. 15, pp. 2774–2794, 2025, doi: 10.1080/1369118X.2025.2479778.
- [4] K. Emerson, T. Nabatchi, and S. Balogh, "An integrative framework for collaborative governance," *Journal of Public Administration Research and Theory*, vol. 22, no. 1, pp. 1–29, 2012, doi: 10.1093/jopart/mur011.
- [5] D. Sagarik, "How digital government capabilities strengthen public sector workforce resilience," *International Review of Public Administration*, vol. 29, no. 4, pp. 353–362, 2024, doi: 10.1080/12294659.2024.2432096.
- [6] J. Crusoe, J. Magnusson, and J. Eklund, "Digital transformation decoupling: The impact of willful ignorance on public sector digital transformation," *Gov. Inf. Q.*, vol. 41, no. 3, p. 101958, 2024, doi: 10.1016/j.giq.2024.101958.
- [7] K. Norling, "Digital transformation or digital standstill? Status quo bias in Swedish public sector strategies," *Transforming Government: People, Process and Policy*, vol. 19, no. 1, pp. 91–107, 2025, doi: 10.1108/TG-04-2024-0078.
- [8] M. Waardenburg, M. Groenleer, and J. de Jong, "Performance management in collaborative governance: A review of the literature and synthesis of the challenges," *Public Performance & Management Review*, vol. 48, no. 4, pp. 735–767, 2025, doi: 10.1080/15309576.2025.2510973.

- [9] S. Thabit, A. Sancino, and L. Mora, "Strategic public value(s) governance: A systematic literature review and framework for analysis," *Public Adm. Rev.*, vol. 85, no. 3, pp. 885–906, 2025, doi: 10.1111/puar.13877.
- [10] D. L. Stufflebeam and G. Zhang, *The CIPP evaluation model: How to evaluate for improvement and accountability*. Guilford Press, 2017.
- [11] S. Hadad, A. Watted, and I. Blau, "Cultural background in digital literacy of elementary and middle school students: Self-appraisal versus actual performance," *J. Comput. Assist. Learn.*, vol. 39, no. 5, pp. 1591–1606, 2023, doi: 10.1111/jcal.12820.
- [12] A. Doerr and R. Novella, "The long-term effects of job training on labor market and skills outcomes in Chile," *Labour Econ.*, vol. 91, p. 102619, 2024, doi: 10.1016/j.labeco.2024.102619.
- [13] W. Wei, X. Chen, and X. Ni, "Awareness, costs, and take-up of social benefits: Evidence from a field experiment in China," *Journal of Public Administration Research and Theory*, 2026, doi: 10.1093/jopart/muag013.
- [14] P. Cantarelli, N. Belle, P. Herd, and D. P. Moynihan, "Reducing administrative burdens to increase the take-up of public services: The case of vaccination intentions," *Public Management Review*, vol. 26, no. 12, pp. 3556–3576, 2024, doi: 10.1080/14719037.2024.2345203.
- [15] R. Peeters, C. Rentería, and G. M. Cejudo, "How information capacity shapes policy implementation: A comparison of administrative burdens in COVID-19 vaccination programs in the United States, Mexico, and the Netherlands," *Gov. Inf. Q.*, vol. 40, no. 4, p. 101871, 2023, doi: 10.1016/j.giq.2023.101871.
- [16] S. Allais, "Why skills anticipation in African VET systems needs to be decolonized: The wide-spread use and limited value of occupational standards and competency-based qualifications," *Int. J. Educ. Dev.*, vol. 102, p. 102873, 2023, doi: 10.1016/j.ijedudev.2023.102873.
- [17] S. Livingstone, G. Mascheroni, and M. Stoilova, "The outcomes of gaining digital skills for young people's lives and wellbeing: A systematic evidence review," *New Media Soc.*, vol. 25, no. 5, pp. 1176–1202, 2023, doi: 10.1177/14614448211043189.
- [18] L. Zhang, C. Yang, and Y. Zheng, "Digital competence for sustainable education of pre-service teachers: A systematic literature review (2014–2024)," *Front. Psychol.*, vol. 16, p. 1710983, 2026, doi: 10.3389/fpsyg.2025.1710983.
- [19] S. León-Herrera *et al.*, "Postgraduate digital health training programmes for primary care physicians: A scoping review protocol," *BMJ Open*, vol. 16, no. 1, p. e109715, 2026, doi: 10.1136/bmjopen-2025-109715.
- [20] M. Nagai, R. Mufwaya-Nsene, M. M. Thandar, S. Matsuoka, S. Okawa, and N. Fujita, "Digitalized human resources for health information systems in low- and middle-income countries: A scoping review," *Hum. Resour. Health*, vol. 24, no. 1, p. 3, 2025, doi: 10.1186/s12960-025-01043-x.
- [21] A. Segal, N. W. Doyle, and Y. Yaffe, "Digital skills assessments in occupational therapy: A scoping review," *The American Journal of Occupational Therapy*, vol. 80, no. 2, p. 8002205160, 2026, doi: 10.5014/ajot.2026.051419.
- [22] P. M. Caligiuri, D. G. Collings, H. De Cieri, and M. B. Lazarova, "Global talent management: A critical review and research agenda for the new organizational reality," *An-*

- nual Review of Organizational Psychology and Organizational Behavior*, vol. 11, pp. 393–421, 2024, doi: 10.1146/annurev-orgpsych-111821-033121.
- [23] O. Smagina and J.-F. Stich, “Talent development in higher education institutions: Systematic literature review and research agenda,” *Human Resource Development Review*, vol. 25, no. 3, pp. 231–260, 2026, doi: 10.1177/15344843251381389.
 - [24] Z. Munn, M. D. J. Peters, C. Stern, C. Tufanaru, A. McArthur, and E. Aromataris, “Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach,” *BMC Med. Res. Methodol.*, vol. 18, p. 143, 2018, doi: 10.1186/s12874-018-0611-x.
 - [25] H. Arksey and L. O’Malley, “Scoping studies: Towards a methodological framework,” *Int. J. Soc. Res. Methodol.*, vol. 8, no. 1, pp. 19–32, 2005, doi: 10.1080/1364557032000119616.
 - [26] D. Levac, H. Colquhoun, and K. K. O’Brien, “Scoping studies: Advancing the methodology,” *Implementation Science*, vol. 5, p. 69, 2010, doi: 10.1186/1748-5908-5-69.
 - [27] M. D. J. Peters *et al.*, “Updated methodological guidance for the conduct of scoping reviews,” *JBIEvid. Synth.*, vol. 18, no. 10, pp. 2119–2126, 2020, doi: 10.11124/JBIES-20-00167.
 - [28] A. C. Tricco *et al.*, “PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation,” *Ann. Intern. Med.*, vol. 169, no. 7, pp. 467–473, 2018, doi: 10.7326/M18-0850.
 - [29] M. D. J. Peters *et al.*, “Updated methodological guidance for the conduct of scoping reviews,” *JBIEvid. Synth.*, vol. 18, no. 10, pp. 2119–2126, 2020, doi: 10.11124/JBIES-20-00167.
 - [30] M. D. J. Peters *et al.*, “Best practice guidance and reporting items for the development of scoping review protocols,” *JBIEvid. Synth.*, vol. 20, no. 4, pp. 953–968, 2022, doi: 10.11124/JBIES-21-00242.
 - [31] T. Byrt, J. Bishop, and J. B. Carlin, “Bias, prevalence and kappa,” *J. Clin. Epidemiol.*, vol. 46, no. 5, pp. 423–429, 1993, doi: 10.1016/0895-4356(93)90018-V.
 - [32] J. R. Landis and G. G. Koch, “The measurement of observer agreement for categorical data,” *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977, doi: 10.2307/2529310.
 - [33] T. Byrt, J. Bishop, and J. B. Carlin, “Bias, prevalence and kappa,” *J. Clin. Epidemiol.*, vol. 46, no. 5, pp. 423–429, 1993, doi: 10.1016/0895-4356(93)90018-V.
 - [34] M. J. Page *et al.*, “The PRISMA 2020 statement: An updated guideline for reporting systematic reviews,” *BMJ*, vol. 372, p. n71, 2021, doi: 10.1136/bmj.n71.
 - [35] Q. N. Hong *et al.*, “The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers,” *Education for Information*, vol. 34, no. 4, pp. 285–291, 2018, doi: 10.3233/EFI-180221.
 - [36] J. Thomas and A. Harden, “Methods for the thematic synthesis of qualitative research in systematic reviews,” *BMC Med. Res. Methodol.*, vol. 8, p. 45, 2008, doi: 10.1186/1471-2288-8-45.
 - [37] H. Choudhary and N. Bansal, “Barriers affecting the effectiveness of digital literacy training programs (DLTPs) for marginalised populations: A systematic literature review,” *Journal of Technical Education and Training*, vol. 14, no. 1, pp. 110–127, 2022, doi: 10.30880/jtet.2022.14.01.010.

- [38] H. Choudhary and N. Bansal, "Addressing digital divide through digital literacy training programs: A systematic literature review," *Digital Education Review*, pp. 224–248, 2022, doi: 10.1344/DER.2022.41.224-248.
- [39] Q. Shi, M. Lan, and X. Wan, "The implementation mechanism and effectiveness of a national plan of a digital competence training program for Chinese primary and secondary school teachers," *Sustainability*, vol. 15, no. 24, p. 16944, 2023, doi: 10.3390/su152416944.
- [40] J. Luo, H. Han, and M. Sun, "Teachers' digital literacy within the Teacher Professional Development at Scale program: Measurement, profiles and predictors," *The Asia-Pacific Education Researcher*, vol. 35, no. 3, pp. 643–653, 2026, doi: 10.1007/s40299-025-01061-2.
- [41] N. Bansal and H. Choudhary, "Fostering digital equity: Evaluating impact of digital literacy training on internet outcomes in rural marginalised communities in India," *International Journal of Lifelong Education*, vol. 43, no. 5, pp. 473–493, 2024, doi: 10.1080/02601370.2024.2347327.
- [42] E. Shava and T. P. Muringa, "Digital empowerment of women in the South African public sector," *Frontiers in Sociology*, vol. 10, p. 1604857, 2025, doi: 10.3389/fsoc.2025.1604857.
- [43] S. Kerrick and D. Cumberland, "Academic innovation for small business empowerment: Micro-credentials as strategic tools," *Journal of Small Business Strategy*, vol. 36, no. 2, pp. 38–46, 2026, doi: 10.53703/001c.157803.
- [44] H. Choudhary and N. Bansal, "Addressing digital divide through digital literacy training programs: A systematic literature review," *Digital Education Review*, pp. 224–248, 2022, doi: 10.1344/DER.2022.41.224-248.
- [45] Q. Ou, S. Wu, Y. Jiang, Z. He, and H. Luo, "From training to practice: A multi-group analysis of factors influencing K-12 teachers' integration of digital educational resources (DERs)," *PLoS One*, vol. 20, no. 12, p. e0338543, 2025, doi: 10.1371/journal.pone.0338543.
- [46] R. Sabdenov, G. Abdrakhmanova, and N. Pussyrmanov, "Assessing the integration of digital competencies into civil service recruitment and training in Kazakhstan," *Problems and Perspectives in Management*, vol. 23, no. 2, pp. 561–571, 2025, doi: 10.21511/ppm.23(2).2025.40.
- [47] G. Tjokro, Z. Zaenuri, H. Yanto, F. Rokhman, A. Rajagukguk, and T. Sanni, "Reforming legal frameworks for human capital: Digital strategies driving Industry 5.0 and sustainability," *Journal of Law and Legal Reform*, vol. 6, no. 3, pp. 1123–1174, 2025, doi: 10.15294/jllr.v6i3.22645.
- [48] K. Chetty, U. Aneja, V. Mishra, N. Gcora, and J. Josie, "Bridging the digital divide in the G20: Skills for the new age," *Economics*, vol. 12, no. 1, pp. 2018–24, 2018, doi: 10.5018/economics-ejournal.ja.2018-24.
- [49] T. Maiyapan, N. Prasit, N. Nilnate, J. Yarasee, W. Phordee, and A. Nonthamat, "Factors associated with perceived need for digital health competency development among primary healthcare personnel in Kalasin Province Thailand," *Discover Public Health*, vol. 23, no. 1, p. 1127, 2026, doi: 10.1186/s12982-026-02515-8.

- [50] K. M. Miltner, G. Kadar-Satat, E. Ashton, and L. Scorer, "Who's coding? An analysis of UK bootcamp experiences and outcomes," *Big Data Soc.*, vol. 13, no. 1, p. 20539517251412372, 2026, doi: 10.1177/20539517251412373.
- [51] L. Widaningsih, S. Rahayu, V. Dwiyantri, W. Widaningsih, and Muktiarni, "Analysis and mapping of technical competency needs for TVET teachers in the era of Industry 4.0," *Journal of Technical Education and Training*, vol. 17, no. 2, pp. 151–164, 2025, doi: 10.30880/jtet.2025.17.02.011.
- [52] Y. Bilan, H. Mishchuk, and N. Samoliuk, "Digital skills of civil servants: Assessing readiness for successful interaction in e-society," *Acta Polytechnica Hungarica*, vol. 20, no. 3, pp. 155–174, 2023, doi: 10.12700/APH.20.3.2023.3.10.
- [53] T. Yooyativong and C. Kamyod, "IoT technology and digital upskilling framework for farmers in the northern rural area of Thailand," *Journal of Mobile Multimedia*, vol. 19, no. 5, pp. 1129–1152, 2023, doi: 10.13052/jmm1550-4646.1952.
- [54] A. Assanova, A. Issaeva, Z. Dzhubalieva, R. Arzikulova, and A. Uchkampirova, "Financial-digital investments in human capital as a factor of sustainable economic growth: Assessment, dynamics and impact of artificial intelligence (a Kazakhstan case study)," *Economic Annals-XXI*, vol. 213, no. 1–2, pp. 18–29, 2025, doi: 10.21003/ea.V213-02.
- [55] E. Burns, E. Silvennoinen, V. A. Kopnov, D. E. Shchipanova, N. Papić-Blagojević, and S. Tomašević, "Supporting the development of digitally competent VET teachers in Serbia and Russia," *The Education and Science Journal*, vol. 22, no. 9, pp. 174–203, 2020, doi: 10.17853/1994-5639-2020-9-174-203.
- [56] A. A. Adewojo, "Digital literacy in African public and academic libraries: A systematic review of capacity-building initiatives and competency frameworks (2010–2025)," *Digit. Libr. Perspect.*, 2026, doi: 10.1108/DLP-09-2025-0136.
- [57] A. Hackl and W. Najdi, "Online work as humanitarian relief? The promise and limitations of digital livelihoods for Syrian refugees and Lebanese youth during times of crisis," *Environment and Planning A: Economy and Space*, vol. 56, no. 1, pp. 100–116, 2024, doi: 10.1177/0308518X231184470.
- [58] A. Dunan, B. Mudjiyanto, Karman, and D. Walujo, "Unveiling disability empowerment: Evaluating ICT skill enhancement initiatives in Indonesia," *Media Commun.*, vol. 13, p. 9078, 2025, doi: 10.17645/mac.9078.
- [59] K. Chetty, Q. Liu, N. Gcora, J. Josie, W. Li, and F. Chen, "Bridging the digital divide: Measuring digital literacy," *Economics*, vol. 12, no. 1, pp. 2018–23, 2018, doi: 10.5018/economics-ejournal.ja.2018-23.
- [60] H.-H. Hachem, M. Wiggberg, T. Osborne, J. Gulliksen, and F. Heintz, "Give MOOCs some credit: A system-divergent innovation accrediting an AI mass-market MOOC at a Swedish university," *Cogent Education*, vol. 12, no. 1, p. 2529421, 2025, doi: 10.1080/2331186X.2025.2529421.
- [61] D. T. Tu, B. T. H. Linh, N. H. X. Truong, T. T. M. Nguyet, and H. M. Linh, "Digital workforce training: An empirical study of higher education institutions in Vietnam," *Multidisciplinary Reviews*, vol. 8, no. 11, p. e2025347, 2025, doi: 10.31893/multirev.2025347.
- [62] G. Esposito *et al.*, "Embedding AI ethics in technical training: A multi-stakeholder pilot module emphasizing co-design and interdisciplinary collaboration at Rome Technopole," *Educ. Sci. (Basel)*, vol. 15, no. 10, p. 1416, 2025, doi: 10.3390/educsci15101416.

- [63] J. Thomas and A. Harden, "Methods for the thematic synthesis of qualitative research in systematic reviews," *BMC Med. Res. Methodol.*, vol. 8, p. 45, 2008, doi: 10.1186/1471-2288-8-45.
- [64] A. Akbik, A. Bobic, C. Freudlsperger, and M. Migliorati, "When compromise becomes standoff: Joint implementation and the case of the European Public Prosecutor's Office," *Governance*, vol. 39, no. 3, p. e70138, 2026, doi: 10.1111/gove.70138.
- [65] N. C. Bandelow, J. Hornung, F. Sager, and I. Schröder, "Multilevel interdependencies and policy capacity in Europe," *European Policy Analysis*, vol. 10, no. 1, pp. 6–9, 2024, doi: 10.1002/epa2.1202.
- [66] A. Shaheen, F. Hussain, and M. Mumtaz, "Diagnosing individual policy capacity for climate governance: Evidence from national and subnational institutions in Pakistan," *Review of Policy Research*, vol. 43, no. 2, p. e70086, 2026, doi: 10.1111/ropr.70086.
- [67] A. S. Muhammad, L. Isiaku, and A. S. Shehu, "Bridging the policy design and implementation gap in Nigeria's startup ecosystem," *Transforming Government: People, Process and Policy*, vol. 20, no. 3, pp. 899–924, 2026, doi: 10.1108/TG-09-2025-0295.
- [68] W. Wei, X. Chen, and X. Ni, "Awareness, costs, and take-up of social benefits: Evidence from a field experiment in China," *Journal of Public Administration Research and Theory*, 2026, doi: 10.1093/jopart/muag013.
- [69] P.-M. Daigneault, "Reconceptualizing administrative burden around onerous experiences," *Perspect. Public Manag. Gov.*, 2024, doi: 10.1093/ppmgov/gvae010.
- [70] R. Vogel, A. Dahlweg, and F. Hattke, "Do vulnerable citizens (really) perceive higher bureaucracy costs? Testing a key claim of the administrative burden framework," *Public Adm. Rev.*, vol. 85, no. 4, pp. 1098–1114, 2025, doi: 10.1111/puar.13932.
- [71] A. H. Ansari and A. Amegah, "Assessing the labour market outcomes of beneficiaries of the Sierra Leone Skills Development Project," *Dev. Pract.*, vol. 35, no. 7, pp. 1102–1115, 2025, doi: 10.1080/09614524.2025.2505023.
- [72] H. M. Gjersøe, A. I. Alecu, M. Bråthen, and E. Dahl, "Labour market programmes, service individualisation and employment outcomes," *Soc. Policy Adm.*, vol. 60, no. 1, pp. 1–11, 2026, doi: 10.1111/spol.13132.
- [73] C. Panzaru and A. Grama, "Towards explicit soft skills labelling in ESCO through semantic NLP analysis," *J. Labour Mark. Res.*, vol. 59, no. 1, p. 18, 2025, doi: 10.1186/s12651-025-00409-x.
- [74] M. D. Pak Güre, C. Karakaya, and M. Başçillar, "Precarity under donor dependency: Social workers' experiences of funding cuts in migration NGOs in Türkiye," *The British Journal of Social Work*, 2026, doi: 10.1093/bjsw/bcag113.
- [75] S. Lezak, "From capacity building to capacity sharing," *Nat. Sustain.*, vol. 7, no. 1, pp. 1–3, 2023, doi: 10.1038/s41893-023-01250-1.
- [76] R. Shibli and S. Kouzi, "Digital livelihoods undone," *Journal of Humanitarian Affairs*, vol. 4, no. 3, pp. 42–47, 2023, doi: 10.7227/jha.095.
- [77] D. Meurs and P. A. Puhani, "Culture as a hiring criterion: Systemic discrimination in a procedurally fair hiring process," *Labour Econ.*, vol. 87, p. 102482, 2024, doi: 10.1016/j.labeco.2023.102482.
- [78] H. A. Hamzah, A. J. Marcinko, B. Stephens, and M. Weick, "Making soft skills 'stick': A systematic scoping review and integrated training transfer framework grounded in behav-

- ournal science,” *European Journal of Work and Organizational Psychology*, vol. 34, no. 2, pp. 237–250, 2025, doi: 10.1080/1359432X.2024.2376909.
- [79] L. Mehner, S. Rothenbusch, and S. Kauffeld, “How to maximize the impact of workplace training: A mixed-method analysis of social support, training transfer and knowledge sharing,” *European Journal of Work and Organizational Psychology*, vol. 34, no. 2, pp. 201–217, 2025, doi: 10.1080/1359432X.2024.2319082.
- [80] Q. Cheng, Y. Sun, J. Song, and L. Song, “Unpacking the role of institutional work in the dynamics of policy legitimacy: The case of demolition policy in China,” *Chinese Public Administration Review*, 2026, doi: 10.1177/15396754251413103.