

India EdTech Mergers and Acquisitions Using AI Cohort Analytics to Test Revenue Quality

A Decision Framework for Retention Outcomes Unit Economics Data Governance and Integration

Authored by Chennakeshav Adya

Independent Researcher

September 2026

Abstract

Background. India education-technology transactions combine subscription, course, test-preparation, institutional, financing and hybrid-delivery economics. Reported growth can conceal changes in acquisition mix, discounting, refunds, deferred service obligations, inactive paid users, weak completion or outcomes that cannot be linked to the acquired product. *Objective.* This paper develops a buyer-side framework for testing retention, outcomes and revenue quality with governed cohort analytics before an acquisition, merger or growth investment. *Approach.* It connects contract and cash records to learner identity, access, engagement, service delivery, renewal, complaints and outcomes. It defines benchmark cohorts, reconciliation controls, artificial-intelligence use cases, data-protection gates, a quality-of-revenue bridge and transaction responses. A hypothetical acquisition case demonstrates how the same evidence can affect valuation, working capital, earn-outs and the first 100 days. *Findings.* Cohort analytics becomes decision-useful when cohort definitions follow the underlying customer promise and reconcile to audited finance records. Activity, completion and outcomes require separate measurement. Artificial intelligence can improve anomaly detection and segmentation, but its output remains provisional until the buyer can reproduce the data lineage, validate performance and assess subgroup effects. *Implications.* Acquirers, founders, lenders and investment committees can use one evidence ledger to connect commercial diligence with accounting, technology, legal and integration work. The worked case is illustrative and does not describe an identified company. Every financial, regulatory and operating assumption requires transaction-specific verification.

JEL Classification: G34, I21, L86, M41, C55

Keywords: India, education technology, mergers and acquisitions, cohort analytics, retention, revenue quality, learning outcomes, data governance, artificial intelligence

1 Define the acquisition question before building the model

An education-technology acquisition rarely turns on a single growth rate. The buyer may seek entry into test preparation, access to a learner community, a recognised faculty network, enterprise distribution, curriculum intellectual property, regional reach or a lower-cost digital delivery platform. Each thesis requires different evidence. A distribution thesis needs channel-level acquisition and renewal. A learning-platform thesis needs active use, completion and outcome evidence. A consolidation thesis needs duplicative cost, product overlap and migration risk. The diligence model should begin with the investment decision that management and the board must make.

India's Competition Commission approved an acquisition of shareholding in Sorting Hat Technologies and its merger into upGrad Education in July 2026.[1] The public combination summary stated that the transaction would enable entry into online test preparation and broaden the range of learners served.[2]

The Commission had also approved an acquisition involving Aakash Educational Services in June 2025.[3] These public records confirm continuing transaction activity and show that strategic scope can cross online, classroom and hybrid education. They do not provide the private operating evidence required to value an individual target.

The diligence perimeter should match the legal entities, brands, products, channels and periods included in the purchase agreement. Buyers often receive dashboards whose definitions changed as the company grew. A paid learner may mean a purchaser, an activated account, a student attached to a family payer, or an instalment plan that later defaulted. A renewal may mean a second transaction, an upgrade, an automatically extended subscription or the movement of the same learner into a different course. The transaction team should freeze a data dictionary before using any reported metric.

The first output is an acquisition-question register. It states the thesis, the metric needed to test it, the source records, the accountable reviewer and the consequence of a failed test. This prevents an impressive analytics exercise from becoming detached from valuation and contract decisions. It also identifies missing evidence early enough for the buyer to narrow the thesis, adjust price, add protection or withdraw.

Table 1. Acquisition questions and decision evidence

Acquisition question	Minimum evidence	Failure signal	Possible transaction response
Does the target retain valuable learners	customer promise, identity, payment, access and renewal records	renewal depends on discounting, channel shift or reclassification	revise forecast and valuation
Does revenue reflect delivered service	contracts, cash, refunds, deferred revenue and delivery logs	recognised revenue exceeds supported delivery or collectability	quality-of-revenue adjustment and indemnity review
Do reported outcomes belong to the product	baseline, assessment, attendance and attribution evidence	results use selected learners or uncontrolled comparisons	exclude outcome premium and require future measurement
Can the platform scale safely	architecture, service levels, support, model and data lineage	growth raises outages, support burden, bias or privacy exposure	price investment needs and add closing conditions
Can integration preserve the asset	learner migration, faculty, curriculum, channel and brand dependencies	value relies on people or systems without retention and migration plans	adjust structure and integration commitments

Each question should have a defined source, test and transaction consequence.

2 Build a learner economic event model

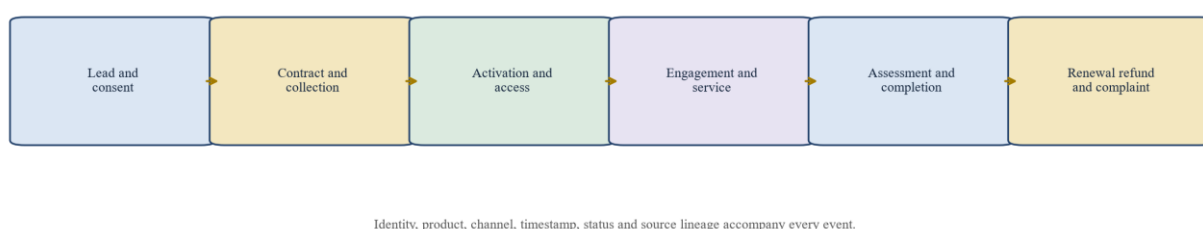
The buyer needs a common event model that connects commercial, learning and finance systems. The chain usually begins with a lead or institutional opportunity and continues through consent, purchase, collection, activation, access, engagement, service delivery, assessment, completion, renewal, refund, complaint and outcome. Each event has a timestamp, source system, identifier, product, channel and status. The model should preserve the original event rather than replacing it with a later dashboard classification.

Identity resolution is the first analytical control. One learner may use a parent's telephone number, a school email, several devices and multiple course accounts. One payer may fund several learners. Enterprise contracts may deliver seats whose end users are created later. The buyer should define household, payer, learner and account relationships separately. Matching rules need confidence scores and exception review because false joins inflate cross-sell and retention, while missed joins understate them.

The event model also needs a customer-promise layer. A recorded course sale may include live instruction, recordings, practice tests, mentorship, physical material, placement support or access over a defined term. The buyer cannot test service delivery from login data alone. It should map each promised component to observable evidence and finance treatment. India's Ind AS 115 material describes revenue recognition around the transfer of promised goods or services and the amount of consideration expected from the customer.[17] Transaction diligence should reconcile operational evidence to the target's accounting policy and audited financial statements with qualified accounting advice.

The raw event ledger remains immutable. Transformations, exclusions and joins are versioned. Every board metric can then be traced to source rows. This discipline also makes artificial-intelligence analysis safer because the model receives a documented population instead of an undocumented dashboard extract.

Figure 1. Learner economic event model for transaction diligence



Proposed evidence chain. The buyer preserves each source event and reconciles commercial, learning and finance records.

3 Reconstruct cohorts around the customer promise

A cohort is a group whose members share a defined starting condition. Calendar-month purchasers are useful for cash and acquisition analysis. Course-start cohorts are better for attendance and completion. Examination-cycle cohorts may suit test preparation. Enterprise learners may need contract and activation cohorts. The buyer should use several cohort views and state which decision each view supports.

The cohort denominator deserves the same attention as the numerator. A renewal rate based only on learners eligible to renew differs from a rate based on all purchasers. A completion rate can exclude cancellations, refunded learners or people who bought access but never activated. Exclusions may be justified, yet they need visibility. The model should show the full starting population, each removal, the remaining eligible group and the measured result.

Public education data helps test whether private operating claims are plausible. The Ministry of Education's UDISE+ 2023-24 report recorded about 248 million enrolments across roughly 1.47 million schools and described its collection and responsibility limitations.[5] The Ministry now lists UDISE+ 2024-25 reports, while NITI Aayog's 2026 school-education study analyses access, transition, retention and learning evidence.[6] These sources provide context for market size and learner progression. They cannot validate a target's proprietary acquisition or renewal claims.

The cohort design should separate observed data from assumptions. Observed fields include purchase date, amount collected, course start, attendance and refund. Derived fields include active learner, expected completion and predicted renewal. Management estimates may influence missing-data treatment, but the diligence report should label them and show sensitivity. A buyer can then distinguish what happened from what the target believes will happen.

Vintage maturity must also be explicit. A cohort that started three months before the diligence cut-off cannot provide a twelve-month renewal observation. The analyst should show the share of each cohort that has reached every measurement point and avoid filling an immature tail with the average of older cohorts without explanation. Where forecasts are necessary, the report should identify the observed period, the extrapolation method and the range created by alternative mature cohorts. This prevents recent fast growth from contributing a full lifetime value before the relevant behaviour has occurred.

Acquired and migrated learners deserve separate flags. A platform can inherit users through a prior acquisition, institutional partnership or database migration. Their first visible event may occur long after the original purchase. Treating them as newly acquired or continuously retained can distort both acquisition cost and survival. The buyer should obtain a migration crosswalk, identify records that lack prior history and present those users as a distinct population until continuity is proven.

Table 2. Cohort definitions for different diligence questions

Cohort view	Starting population	Primary measures	Main limitation
Purchase cohort	completed orders in a defined period	collection, refund, activation and repeat purchase	mixes different service-start dates
Course-start cohort	enrolled learners at scheduled start	attendance, engagement, completion and assessment	excludes sales that never activate unless reconciled
Exam-cycle cohort	learners preparing for the same examination cycle	persistence, test performance and progression	examination changes can limit comparability
Enterprise contract cohort	seats or learners attached to one contract vintage	activation, utilisation, renewal and expansion	payer and learner decisions differ
Channel cohort	learners attributed to one acquisition source	acquisition cost, conversion, payback and retention	attribution rules may change over time

A single cohort definition cannot answer every commercial and educational question.

4 Reconcile revenue cash delivery and learner activity

Revenue quality begins with reconciliation. The buyer should bridge gross billings to cash collected, refunds, taxes, deferred revenue, recognised revenue and receivables. It should then connect the financial bridge to service delivery and learner activity. The resulting view explains whether growth came from durable demand, longer prepaid commitments, aggressive discounting, instalment credit, acquisitions or accounting classification.

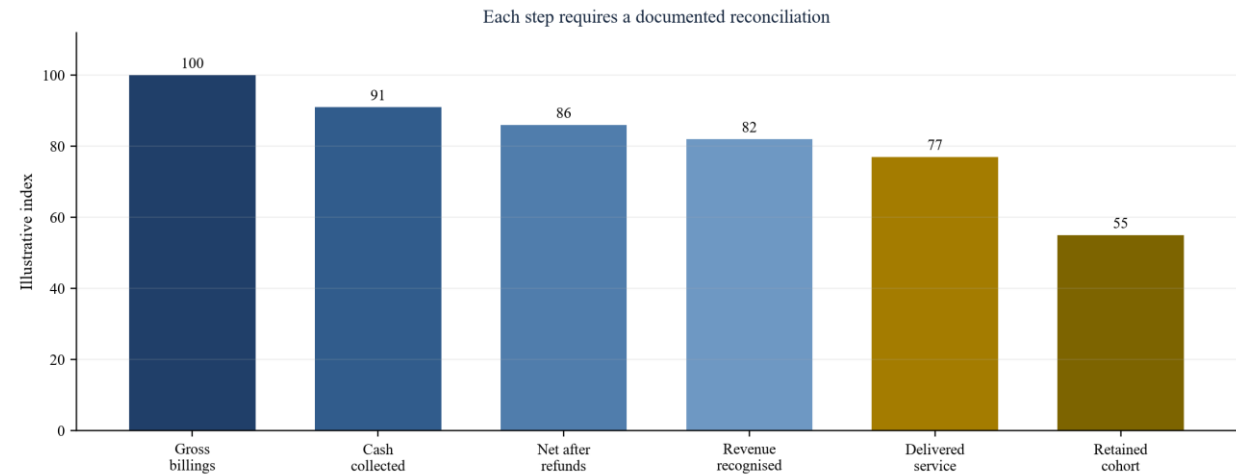
Cash can arrive before service, after service or through a financing provider. A twelve-month course paid upfront creates a different working-capital profile from monthly subscriptions. An instalment plan may record a sale while collection risk remains. A refund policy can shift reported revenue and cash after a learner disengages. Institutional contracts may include implementation, licences, support and outcome-linked consideration. Each model requires its own contract and performance-obligation analysis.

The target's published audited accounts and offer documents provide the starting policy description. PhysicsWallah filed a final prospectus with SEBI in November 2025, following earlier offer documents.[4] The filing is a public source for that issuer's disclosed business and financial information. It should not be treated as a benchmark for a different target without checking product, period and accounting differences.

The reconciliation also tests operational completeness. If finance records show 100,000 paid enrolments and the learning platform contains 82,000 usable learner records, the buyer needs a documented explanation. Offline delivery, enterprise seats, delayed provisioning and migrated systems may explain part of the gap. Unresolved differences affect confidence in retention and outcome analysis. The buyer

should report the unreconciled population separately rather than allocating it through an opaque assumption.

Figure 2. Quality of revenue reconciliation



Proposed bridge from commercial transaction to reported revenue and repeat economics.

5 Measure retention as behaviour and economics

Retention has several meanings. Access retention measures whether a learner remains active during the promised term. Product retention measures continued use of a course or platform. Commercial retention measures renewal or repeat purchase. Revenue retention measures the value retained after contraction, expansion and churn. Outcome persistence measures whether learning or progression continues. The buyer should present these separately.

For consumer products, the cohort curve can show activation in the first week, sustained use over the course, completion and later purchase. For test preparation, seasonality and examination timing dominate. For enterprise learning, seat utilisation and contract renewal may be separated by months. The buyer should select observation windows that reflect the product promise. A thirty-day activity rate may be irrelevant for a six-month examination course if scheduled usage is concentrated around classes and tests.

Discount and acquisition channel can change the apparent curve. A target may improve headline learner growth by adding low-price users with weak persistence. A premium cohort may retain better but require expensive faculty or counselling. The model should therefore stratify cohorts by product, list price, realised price, channel, geography, delivery mode and learner segment. It should also show enough population to judge uncertainty.

Artificial intelligence can help cluster behavioural sequences or identify early disengagement, but the buyer should retain transparent survival curves and simple benchmark models. If a complex model claims higher predictive accuracy, the diligence team should test it on an untouched period and compare calibration across material subgroups. The output should support a concrete decision such as the provision for refunds, the forecast renewal base or the integration retention plan.

Table 3. Retention measures and common interpretation errors

Measure	Calculation basis	Decision supported	Common error
activation	purchasers who complete defined first use	onboarding and provisioning	counting an automated account creation as learner action
period activity	eligible learners with qualifying activity	service use and support planning	changing the qualifying event between cohorts

Measure	Calculation basis	Decision supported	Common error
completion	eligible starters who satisfy a stated requirement	delivery and product effectiveness	excluding inactive or refunded learners without reconciliation
logo renewal	institutional customers that renew	enterprise retention	weighting a small and large contract equally without disclosure
gross revenue retention	opening recurring value retained before expansion	downside durability	adding upsell to conceal churn
net revenue retention	opening recurring value after contraction churn and expansion	account growth	mixing new customers into the opening base

The metric should follow the economic and educational promise made to the customer.

6 Test outcomes without overclaiming causality

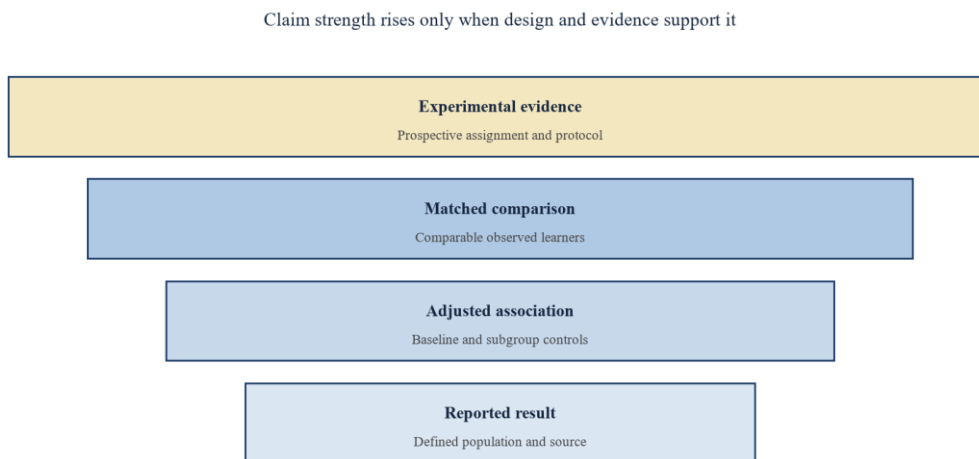
Education outcomes can strengthen an investment thesis when the measurement is credible and tied to the product. The buyer should begin with the claimed outcome: examination success, score improvement, course completion, placement, wage progression, school progression or another defined result. It should then identify the baseline, comparison, observation period, missing data and competing explanations.

ASER 2024 provides independent evidence on enrolment, foundational learning and digital access in rural India.[7] The National Education Policy supports technology use while calling for rigorous and transparent evaluation before scaling interventions.[8] Those sources reinforce the need to measure learning and access carefully. They do not prove the effectiveness of a target's course.

Selection bias is common. Highly motivated learners may purchase a programme and also perform better. A company may report outcomes only for learners who complete all assessments or consent to publication. Examination difficulty and cut-offs change. Schools or tutoring outside the platform contribute to results. A buyer can improve the analysis through matched comparisons, baseline adjustment and sensitivity tests, but causal language requires an appropriate design.

The transaction report should distinguish descriptive, predictive and causal evidence. Descriptive analysis reports what happened within the observed cohort. Predictive analysis estimates which learners are likely to persist or succeed. Causal analysis asks what result arose because of the intervention. Each level can support decisions when labelled correctly. The buyer should avoid paying an outcome premium based on selected testimonials or a model that cannot reproduce its training population.

Figure 3. Evidence ladder for education outcomes



Stronger claims require stronger design, data quality and independent review.

7 Rebuild unit economics by cohort

Aggregate contribution margin can hide deterioration. The buyer should rebuild unit economics for each acquisition and product cohort using realised revenue, cash collection, refunds, content royalties, faculty cost, counselling, support, payment charges, financing subsidy, physical material, platform cost and attributable marketing. Shared cost allocation should be shown separately from direct cohort economics.

Customer acquisition cost requires a controlled denominator. Marketing spend may support brand, lead generation, remarketing and sales conversion across several products. Organic traffic can reflect earlier paid investment. Affiliate or counsellor commissions may be recorded outside marketing. The buyer should reconcile the spend ledger to attributed acquisitions and show unattributed cost. Modelled attribution is useful, but its assumptions should remain visible.

Lifetime value is especially sensitive. A simple model multiplies average contribution by expected retention. A stronger model follows actual cohort cash flows, separates observed and forecast periods, applies refund and default behaviour, and discounts future contribution. It can include expansion or progression only when the learner identity is stable and the next purchase reflects genuine economic value. The model should show how much value arises from unobserved future behaviour.

Faculty and content economics need the same cohort treatment. A recorded class may be reused across a large population at low marginal cost, while live teaching, doubt resolution and counselling scale with attendance. Star faculty may receive fixed compensation, revenue share, guarantees or retention payments. Content rights may sit with an individual, an affiliate or the target. The buyer should map those costs and rights to the cohorts that depend on them. A high-margin digital course can require material future spend when a syllabus, examination pattern or faculty relationship changes.

The enterprise channel introduces concentration and implementation economics. Contract value can include unactivated seats, custom content, integration and support. Expansion within one customer can raise net revenue retention while learner utilisation falls. The buyer should show contracted seats, activated learners, qualifying use, delivered services, invoice milestones and collections for each account vintage. Sales incentives and partner revenue shares should follow the same period and customer population.

The investment committee needs a bridge from reported unit economics to verified economics. Adjustments may cover unsupported attribution, unpaid instalments, refunds after the measurement date, free extensions, service obligations, faculty dependency and recurring platform investment. The buyer can then compare price with evidence-based value rather than a headline ratio.

Table 4. Cohort unit economics build

Component	Preferred basis	Diligence test	Valuation treatment
realised price	transaction value net of tax discount and credit	reconcile order to collection	use collected or collectible value
refunds and reversals	cohort-level actual plus incurred exposure	test policy and post-period events	reduce revenue and cash contribution
delivery cost	faculty content support material and platform usage	map cost to promised service	include remaining obligation
acquisition cost	reconciled campaign sales and commission ledger	show attributed and unattributed spend	stress payback by channel
renewal contribution	actual repeat margin for mature cohorts	confirm identity and eligibility	forecast from comparable cohorts
terminal value	value beyond observed renewal periods	disclose horizon and decay	apply sensitivity and avoid double count

Observed cash and cost should be separated from forecast value.

8 Use artificial intelligence within a controlled diligence process

Artificial intelligence can accelerate record matching, anomaly detection, behavioural sequence analysis, text classification and predictive modelling. It can flag improbable attendance, duplicated identities, unusual refunds, channel discontinuities, support complaints or course-level churn. It can also help reviewers prioritise samples from millions of events. These uses still require a controlled population, tested rules and human review.

The diligence team should begin with deterministic checks. It can reconcile row counts, sums, timestamps, duplicates, impossible sequences and referential integrity before training any model. Simple statistical baselines then establish whether a complex method adds value. A model that predicts renewal should be evaluated on time-separated data, because random splitting can leak patterns from the same campaign or course into training and testing.

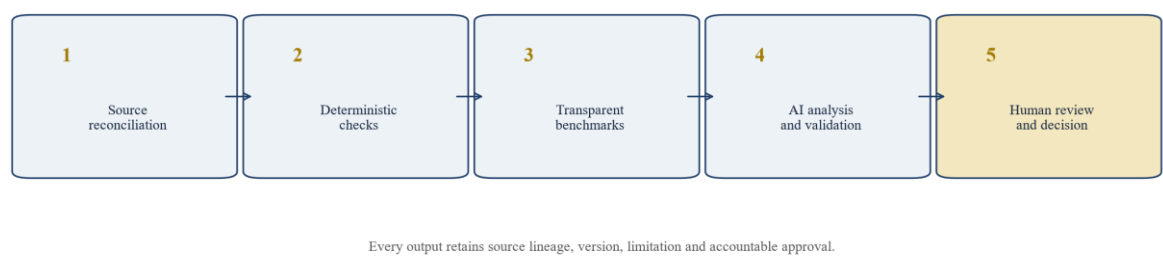
NITI Aayog's responsible-AI papers identify principles including safety, equality, inclusivity, privacy, transparency and accountability, and discuss operational governance.[13][14] UNESCO's guidance on generative AI in education emphasises privacy, age-appropriate use and human-centred governance.[15] These frameworks support a diligence control system in which models have owners, documented purposes, versioned data and routes for human challenge.

The buyer also needs reproducibility. The data room should contain the code or specification, environment, feature definitions, training period, evaluation results, approvals and monitoring history for any model material to revenue, retention, learner treatment or claims. If the target cannot reproduce a dashboard or model, the buyer should treat the output as unsupported until rebuilt from source evidence.

Generative AI use needs an additional boundary. A model may summarise support tickets, classify complaints, produce learning content or assist counsellors. Diligence should identify the provider, model version, prompts or policy, data sent to the service, output review and contractual rights. The buyer should test whether personal or confidential information entered the system, whether generated content was reviewed for accuracy and rights, and whether a model change can alter customer experience without approval. A vendor demonstration does not substitute for evidence from the deployed workflow.

The validation record should include operational cost. A model that improves prediction slightly may require expensive labelling, infrastructure, monitoring and specialist staff. The buyer should compare the incremental decision benefit with those recurring costs and with a simpler rule. This connects model diligence to the acquisition case and prevents technical complexity from being mistaken for commercial advantage.

Figure 4. Controlled analytics sequence for transaction diligence



Automated analysis follows data reconciliation and precedes accountable transaction decisions.

9 Govern learner data and child data through the transaction

Education datasets can contain identity, age, contact, location, assessment, disability, behaviour, payment and family information. The buyer should identify which entity collected each field, for what purpose, under which notice or consent, for how long, and with which processors. It should distinguish personal data required to deliver the service from fields collected for experimentation or marketing.

India's Digital Personal Data Protection Act 2023 creates obligations for lawful processing and addresses withdrawal of consent, data-fiduciary duties and children's data.[11] MeitY notified the Digital Personal Data Protection Rules 2025 in November 2025 with phased commencement described in official materials.[12] Transaction teams need current Indian counsel because commencement dates, exemptions, significant-data-fiduciary obligations and sector-specific rules can affect the target and integration plan.

The buyer's diligence request should minimise new exposure. Aggregate or de-identified extracts may answer many commercial questions. Identifiable records should move through an approved secure environment with access logging and purpose limits. The seller should confirm that disclosure is permitted and that processor contracts, retention schedules and deletion practices align with the legal basis. The buyer should avoid copying a complete learner database merely because it exists.

Model governance adds further questions. Training data may include minors or inferred attributes. Features can proxy for income, geography, language or gender. A retention intervention may direct scarce counselling toward learners predicted to pay rather than those needing educational support. The transaction review should test feature necessity, subgroup performance, intervention policy and complaint routes. These controls protect learners and also reveal liabilities that could impair valuation.

Table 5. Data governance diligence controls

Control area	Evidence	Test	Transaction consequence
purpose and notice	privacy notices consent records and processing register	map each field and use to a stated purpose	remediation plan or restricted use

Control area	Evidence	Test	Transaction consequence
child data	age handling guardian process product controls	test collection marketing and profiling pathways	closing condition or product limitation
processors	contracts sub-processors locations and security terms	reconcile actual data flows to agreements	novation amendment or indemnity review
retention and deletion	schedules logs backups and account closure process	sample expired and deleted records	quantify remediation and migration scope
model features	feature register training population and approvals	identify sensitive fields proxies and unsupported reuse	remove feature retrain or constrain decision use

Legal conclusions require transaction-specific advice and current verification.

10 Detect manipulation leakage and metric drift

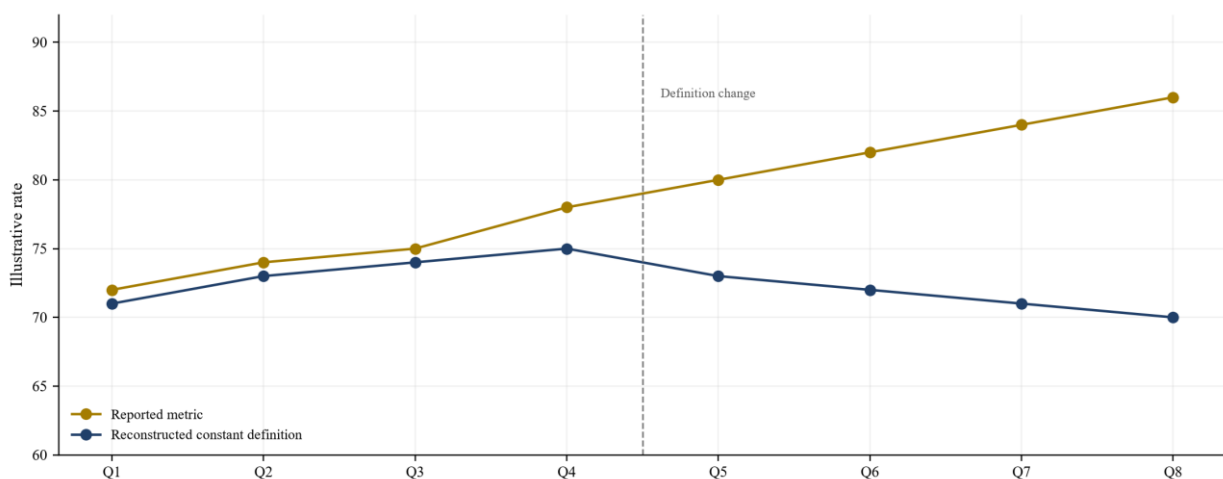
Rapidly growing platforms accumulate metric drift. Product teams change event names. Marketing changes attribution windows. Finance revises revenue policy. Acquired systems use different identities. A dashboard may remain visually stable while its population changes. The buyer should therefore compare metric specifications and source distributions across time.

Anomaly detection can reveal duplicate accounts, synthetic activity, improbable study sessions, burst attendance, repeated assessment patterns or refund clusters. These signals require investigation. A learner may legitimately download content in one session or share a household device. A batch process may create identical timestamps. The report should distinguish the automated flag from the reviewed conclusion.

Leakage is a separate risk. A renewal model may use fields recorded after renewal, directly or through a derived feature. An outcome model may include assessment information unavailable at the decision date. Random train-test splits can put the same learner, household, teacher or course in both sets. The buyer should define the decision timestamp and exclude information that would not have existed then.

Metric drift also affects integration. If the acquirer plans to migrate learners, combine curricula or alter pricing, historic retention may lose relevance. The model should identify stable drivers and transaction-sensitive drivers. It can then stress changes in faculty, brand, product, platform, payment method and support. The integration plan should preserve the features that evidence shows are linked to retention while avoiding claims of causality that the analysis cannot support.

Figure 5. Metric integrity and drift review



The buyer tests definition, data population and model performance across successive periods.

11 Work through a hypothetical acquisition case

Consider a hypothetical Indian test-preparation platform with consumer subscriptions, hybrid centres and a small enterprise channel. Management reports 1.2 million paid learners over three years, strong course completion and increasing repeat purchase. The buyer is evaluating a controlling acquisition. All figures in this section are assumptions created solely to demonstrate the framework.

The transaction team receives order, payment, refund, learning-platform, centre-attendance, assessment, support and general-ledger extracts. Identity resolution links 92 percent of paid orders to a learner with high confidence. Five percent links only to a payer or institutional seat, and three percent remains unresolved. The buyer preserves each category rather than assigning unresolved records to the active population.

The reported repeat-purchase rate is 38 percent. A constant-definition reconstruction finds 31 percent among eligible learners. The difference arises from counting free extensions as repeat purchases, joining some household purchases to the wrong learner and excluding refunded first purchases. Mature organic cohorts show 37 percent, paid-social cohorts 24 percent and centre-originated cohorts 35 percent. The mix has shifted toward paid social, so an aggregate historic rate overstates the forward base.

Management reports 74 percent completion. The reconstructed course-start denominator produces 58 percent, with a further 9 percent classified as materially engaged without meeting the published completion rule. Examination results are available for 41 percent of learners and cannot be linked to a controlled comparison. The buyer treats the result as descriptive evidence and removes the proposed valuation premium for proven outcome superiority.

The finance bridge identifies cash collection below gross billings because of instalment defaults and refunds. Remaining live-class obligations require faculty and support cost after the measurement date. The buyer reduces forecast contribution, adds a closing working-capital mechanism and proposes an earn-out based on cash contribution from defined post-close cohorts. It also requires a data-remediation plan and consent review before migrating learner records.

Table 6. Hypothetical diligence findings and transaction responses

Finding	Reported	Reconstructed	Decision response
learner identity coverage	not stated	92 percent high confidence	exclude unresolved users from retention claims and fund remediation
repeat purchase	38 percent	31 percent constant definition	reduce forecast renewal and test price sensitivity
course completion	74 percent	58 percent under documented rule	revise outcome narrative and integration support plan
examination outcome coverage	selected result set	41 percent linked coverage	treat as descriptive and commission prospective measurement
collection quality	gross billings emphasis	defaults refunds and timing reduce cash conversion	add quality-of-revenue and working-capital adjustments
privacy readiness	general policy available	lineage and child-data gaps require review	make remediation and migration controls contractual

All values are illustrative assumptions and do not describe an identified business.

12 Translate evidence into valuation and deal protection

The cohort model should feed a valuation bridge. Revenue adjustments address unsupported recognition, refund exposure, collection risk and acquired revenue. Margin adjustments address remaining delivery obligations, faculty cost, support and platform investment. Forecast adjustments address acquisition mix, retention, renewal and price. Capital adjustments address working capital, deferred revenue, debt-like items and required remediation.

The buyer should avoid converting every uncertainty into a discount rate. Specific risks often have more direct responses. A shortfall in cash collection can affect net debt or working capital. A data-governance gap can become a closing condition, covenant, indemnity or ring-fenced remediation budget. Uncertain renewal can support an earn-out tied to a defined post-close cohort. Key faculty dependency can require retention arrangements and integration commitments.

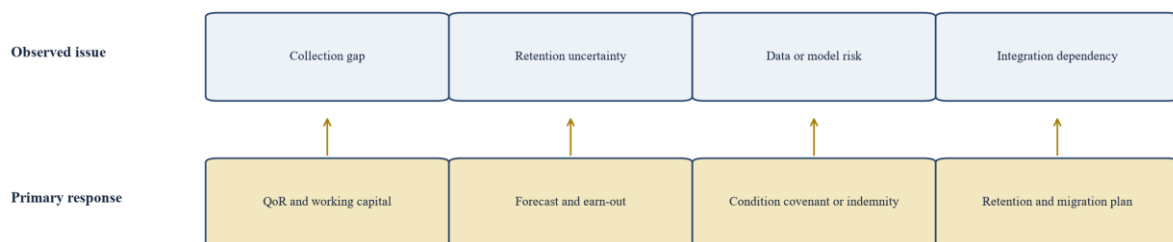
Earn-out metrics require the same discipline as diligence metrics. The agreement should define the population, event, period, exclusions, accounting policy, data source, access rights and dispute process. A revenue earn-out can reward discounting or low-quality sales. A cash-contribution measure can better reflect collection and delivery cost, although it requires careful allocation rules. Outcome metrics may be unsuitable when attribution, observation period or regulatory constraints remain unresolved.

Consideration mechanics should anticipate management control after closing. The buyer may change pricing, marketing, platform, faculty or accounting systems during the earn-out period. The seller may argue that those changes reduced the measured result. The agreement should allocate operating discretion, require consistent measurement and define how approved changes affect the metric. Independent access to source evidence and a short calculation timetable reduce the risk that the earn-out becomes a second transaction dispute.

Representations and indemnities should follow identified facts. A buyer can request statements about ownership of learner data, validity of consents, accuracy of specified metrics, compliance of advertisements and ownership of content. The negotiated wording, knowledge qualifiers, caps and survival periods require legal advice. The diligence ledger helps counsel connect requested protection to evidence instead of relying on a generic technology schedule.

IFRS 3 sets the international accounting framework for recognising and measuring business combinations, while Indian statutory and accounting requirements need qualified advice for the actual transaction.[20] The commercial diligence model informs price and protection. It does not replace purchase-price-allocation, legal, tax or financial-reporting work.

Figure 6. From cohort evidence to transaction terms



Price, structure and integration actions remain traceable to the diligence evidence ledger.

Proposed decision path. Each adjustment cites the evidence and approving forum.

13 Plan integration before signing

An education-technology deal can lose value through rushed migration. Learners may face new logins, altered schedules, lost progress, changed faculty or inconsistent support. Enterprise customers may need consent before assignment or data transfer. Marketing and outcome claims may need reapproval. The buyer should convert diligence findings into a day-one and 100-day plan before signing.

The first priority is service continuity. Product, faculty, centre, support, payments and assessment teams identify the obligations due to active learners. The integration plan records which platform and brand continue, which records move, how users are informed, how consent and preference are handled, and how unresolved accounts are treated. The buyer should preserve an audit trail of the original record and migration result.

The second priority is metric continuity. The combined business needs a common data dictionary and a period of parallel reporting. Historic dashboards should remain accessible. New definitions should carry effective dates and bridges to old definitions. The integration team can then distinguish genuine operating change from measurement change.

The third priority is value capture. The buyer sequences cross-sell, product consolidation, procurement, platform and organisation changes according to evidence and customer risk. A product with strong organic retention may deserve protection while paid acquisition is recalibrated. A course with weak completion may need redesign before broader distribution. Integration milestones should reflect these choices rather than a generic synergy calendar.

Table 7. First 100 day integration control plan

Period	Required control	Evidence	Approval question
before signing	identify active learner obligations and critical dependencies	contracts service calendar faculty and platform map	can service continue through closing
before closing	approve data transfer migration and communication design	legal basis notices processor map and test plan	can records move lawfully and accurately
day one	preserve access payment support and incident routes	readiness tests escalation roster and monitoring	are learners protected from transaction disruption
days 1 to 30	run parallel metrics and reconcile combined populations	old and new dashboards source bridges and exceptions	do reported changes reflect operations or definitions
days 31 to 100	implement evidence-based product and channel actions	cohort contribution completion complaints and renewal	which actions improve value without weakening service

Timing depends on the transaction and should follow legal and operational approval.

14 Establish governance for the combined business

The combined company needs governance that links finance, product, education, data, legal and customer operations. A metric council can own definitions and changes. A model committee can approve material AI uses, validation and monitoring. A learner-protection forum can review complaints, accessibility, child safeguards and advertising claims. The board or investment committee receives a concise pack that connects these controls to value and risk.

Consumer claims require specific attention. India's Central Consumer Protection Authority issued 2024 guidelines addressing misleading advertisements in the coaching sector, including concealed information and false guarantees.[10] The buyer should review historical and planned claims about ranks, selections,

outcomes, faculty, course duration and refund. The cohort model can test whether the supporting population and disclosure match the claim.

Governance should include thresholds and consequences. A material change in refund rate may trigger cash forecasting and customer remediation. Model drift can suspend an automated intervention. A migration reconciliation gap can stop the next batch. A privacy incident follows the approved response process. These are operating controls with owners and evidence, not statements of intent.

The board should also review whether analytics serves educational and commercial objectives responsibly. Engagement optimisation can increase screen time without improving learning. Retention interventions can prioritise high-value customers. Outcome claims can narrow teaching toward measured tests. Management should define acceptable use, monitor unintended effects and keep educators involved in product decisions.

15 Conclusion

India education-technology transactions require an evidence chain that connects the customer promise to cash, service, behaviour, outcomes and renewal. Cohort analytics provides that chain when the buyer fixes definitions, preserves source lineage and reconciles the learner population to finance records. The result is a clearer view of revenue quality and the operating work required after closing.

Artificial intelligence expands the scale and depth of review. It can match records, detect anomalies, segment behaviour and predict risk. Its contribution depends on governed data, transparent benchmarks, time-aware validation, subgroup testing and accountable decisions. Unsupported model outputs should not determine price, learner treatment or public outcome claims.

The practical deliverable is a decision ledger. Each material finding identifies the source, population, method, limitation, owner and transaction consequence. The ledger can support valuation, working capital, earn-outs, protections and the integration plan. It also gives the combined business a disciplined starting point for reporting and model governance.

The hypothetical case shows how reported growth can change after definitions and records are reconciled. The conclusion is not that every difference signals misconduct. System history, product complexity and genuine business-model variation can produce differences. The buyer still needs a complete explanation before paying for the associated value.

Limitations and further research

This framework does not estimate the value of an identified company. Private operating data, contracts, accounting policies, tax positions, legal rights, regulatory status and technology architecture were unavailable. The worked figures are hypothetical. Public sources describe market, policy and transaction context and do not validate a private target's performance.

Future research could test how cohort definitions differ across examination preparation, higher education, enterprise skilling, school software and hybrid centres. Further work could also compare earn-out designs, study privacy-preserving record linkage and evaluate how migration affects learner outcomes and renewal. Such research would require appropriately governed transaction or company data.

Frequently asked questions

Q: What should an EdTech buyer reconcile first? | A: Reconcile contracts, orders, cash, refunds, recognised revenue and learner identities before calculating retention or outcomes. The unresolved population should remain visible.

Q: Which cohort definition is best for an acquisition? | A: The definition depends on the question. Purchase cohorts support cash analysis, course-start cohorts support completion, exam-cycle cohorts support progression and contract cohorts support enterprise renewal.

Q: Can login activity prove that a learner received the service? | A: Login activity is one signal. The buyer should map every promised service component to evidence such as classes, content, assessment, support or material delivery.

Q: How should artificial intelligence be used in diligence? | A: Use it for matching, anomaly detection, segmentation and prediction after source reconciliation. Compare it with transparent benchmarks and require reproducibility, subgroup testing and human review.

Q: Can outcome data justify a valuation premium? | A: Outcome evidence may support a premium when the population, baseline, observation period, missing data and attribution are credible. Selected testimonials or uncontrolled results provide weaker support.

Q: What makes an EdTech earn-out measurable? | A: The agreement should define the cohort, event, period, exclusions, accounting policy, source system, access rights and dispute route. Cash contribution may align quality better than gross sales in some cases.

Q: How should learner data be shared in a transaction? | A: Use the minimum data required, prefer aggregate or de-identified extracts, confirm the legal basis, restrict access, log use and obtain current Indian legal advice for personal and child data.

Q: What should happen during the first 100 days? | A: Preserve learner service, run old and new metrics in parallel, reconcile migrated records and sequence product and channel changes according to verified cohort evidence.

References

- [1] Competition Commission of India. Commission approves acquisition of shareholding in Sorting Hat Technologies and merger with upGrad Education. <https://www.cci.gov.in/antitrust/press-release/details/673>
- [2] Competition Commission of India. Summary of the proposed combination involving upGrad Education and Sorting Hat Technologies. <https://www.cci.gov.in/images/summaryorders/en/summary1778154774.pdf>
- [3] Competition Commission of India. Acquisition involving Aakash Educational Services by Manipal entities. <https://www.cci.gov.in/public/images/pressrelease/en/press-release1750768043.pdf>
- [4] Securities and Exchange Board of India. Physicswallah Limited Prospectus. https://www.sebi.gov.in/filings/public-issues/nov-2025/physicswallah-limited-prospectus_97816.html
- [5] Ministry of Education India. UDISE Plus Report 2023 to 2024 NEP Structure. https://www.education.gov.in/sites/upload_files/mhrd/files/statistics-new/udise_report_nep_23_24.pdf
- [6] NITI Aayog. School Education System in India Temporal Analysis and Policy Roadmap for Quality Enhancement. <https://www.niti.gov.in/node/2288>
- [7] ASER Centre. Annual Status of Education Report 2024. <https://asercentre.org/aser-2024/>
- [8] Ministry of Education India. National Education Policy 2020. https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf
- [9] Ministry of Education India. Promotion of Digital Education. <https://www.education.gov.in/promotion-digital-education>
- [10] Central Consumer Protection Authority. Guidelines for Prevention of Misleading Advertisement in Coaching Sector 2024. <https://consumeraffairs.nic.in/sites/default/files/file-uploads/latestnews/Guidelines%20for%20Prevention%20of%20Misleading%20Advertisement%20in%20Coaching%20Sector%2C%202024.pdf>
- [11] Ministry of Electronics and Information Technology. Digital Personal Data Protection Act 2023. <https://www.meity.gov.in/static/uploads/2024/02/Digital-Personal-Data-Protection-Act-2023.pdf>
- [12] Ministry of Electronics and Information Technology. Digital Personal Data Protection Rules 2025. <https://www.meity.gov.in/documents/act-and-policies/digital-personal-dataprotection-rules-2025gDOxUjMtQWa?pageTitle=Digital-Personal-Data-ProtectionRules-2025>
- [13] NITI Aayog. Approach Document for India Part 1 Principles for Responsible AI. <https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf>
- [14] NITI Aayog. Approach Document for India Part 2 Operationalizing Principles for Responsible AI. <https://www.niti.gov.in/node/463>

- [15] UNESCO. Guidance for Generative AI in Education and Research. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
- [16] UNESCO. Recommendation on the Ethics of Artificial Intelligence. <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>
- [17] Institute of Chartered Accountants of India. Educational Material on Indian Accounting Standard 115 Revenue from Contracts with Customers. <https://www.icaai.org/post/educational-material-on-indian-accounting-standard-ind-as-115-revenue-from-contracts-with-customers-17-08-2018>
- [18] Ministry of Corporate Affairs India. Companies Act 2013. <https://www.mca.gov.in/content/dam/mca/pdf/CompaniesAct2013.pdf>
- [19] Securities and Exchange Board of India. Issue of Capital and Disclosure Requirements Regulations. https://www.sebi.gov.in/legal/regulations/jan-2026/securities-and-exchange-board-of-india-issue-of-capital-and-disclosure-requirements-regulations-2018-last-amended-on-december-5-2025-_98959.html
- [20] IFRS Foundation. IFRS 3 Business Combinations. <https://www.ifrs.org/issued-standards/list-of-standards/ifrs-3-business-combinations/>
- [21] OECD. Digital Education Outlook 2023 Towards an Effective Digital Education Ecosystem. https://www.oecd.org/en/publications/oecd-digital-education-outlook-2023_c74f03de-en.html
- [22] World Bank. Reimagining Human Connections Technology and Innovation in Education. <https://www.worldbank.org/en/topic/edutech/publication/reimagining-human-connections-technology-and-innovation-in-education-at-the-world-bank>

About the Author

Chennakeshav (CK) is a corporate finance and investment banking executive with 25+ years of global experience in deal origination, structuring and execution across M&A, growth capital and corporate strategy. He has led value-creation mandates for founders, corporates and funds — bridging the boardroom view to hands-on execution and closure.

His career spans Morgan Stanley, HSBC, Lloyds Banking Group, EWEC, ADQ portfolio companies and Emirates Growth Fund, across TMT, real estate, fintech, deeptech, cleantech, infrastructure and energy. He has partnered with C-suite leaders, private equity and venture funds, sovereign wealth funds and family offices to finance complex fund raises and scale-up ventures, and has led M&A due diligence, post-merger integration and business-transformation initiatives to create value.

At Matchpoint Partners, he is Managing Partner and leads the firm's corporate finance, M&A and capital-raising practice. He holds an MBA from London Business School, an engineering degree from VTU and a Master of Laws (LLM, in progress) from UCL London.

An active start-up mentor, CK mentors at Techstars, DIFC FinTech Hive, Startup Grind, Founder Institute and IN5, serves as Entrepreneur Mentor in Residence (EMiR) at London Business School, and judges the Entrepreneurship World Cup.

<https://www.linkedin.com/in/ckadya/>

<https://www.matchpoint-partners.com/team/ck-adya.html>

This paper is part of a continuing series on the structure of private and alternative markets. The views expressed are the author's own. The paper is for information only, describes market structure in general terms, and does not constitute investment, legal, tax or regulatory advice or a recommendation in respect of any security, vehicle or counterparty.