NEXTLAB.TECH

Equity in Standardised Testing

How AI Can Level the Playing Field

OECD Conference · Effective Uses of Generative AI in Education · 23 March 2026

Prof. Razvan BOLOGA
CEO Nextlab.tech

NEXTLAB.TECH

Agenda

- AI in Education
- Equity Problem
- Item Generation
- Calibration with AI
- Synthetic Test Takers
- Impact

NEXTLAB.TECH

NEXTLAB.TECH

- Improves education using emerging technologies since 2018
- 5000+ schools, universities and companies from EMEA(Europe, Middle East and Africa)
- Robotics education at scale
- Assessment innovation
- Data insights
- Partners: Microsoft, Oracle, SAP, AWS, Microchip, NXP, TAO Testing, Renault Grouppe
- Scientific advisor for testing: Dr. Alina von Davier

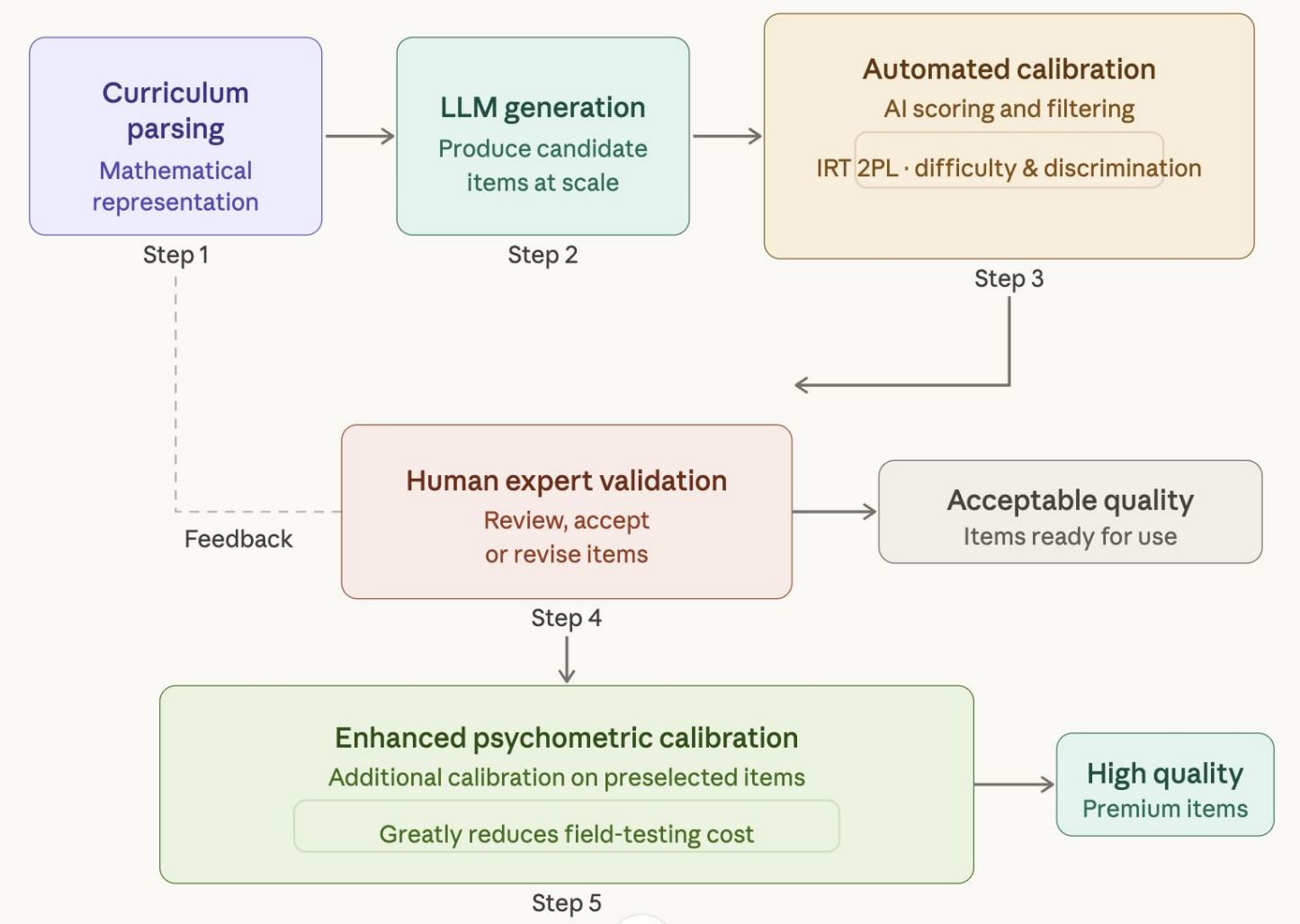
NEXTLAB.TECH

The Problem in Standardised Testing

- High cost per item (up to $6,000)
- Large item banks required
- Expensive pilot samples
- <40% countries with quality testing

NEXTLAB.TECH

AI Item Generation Pipeline



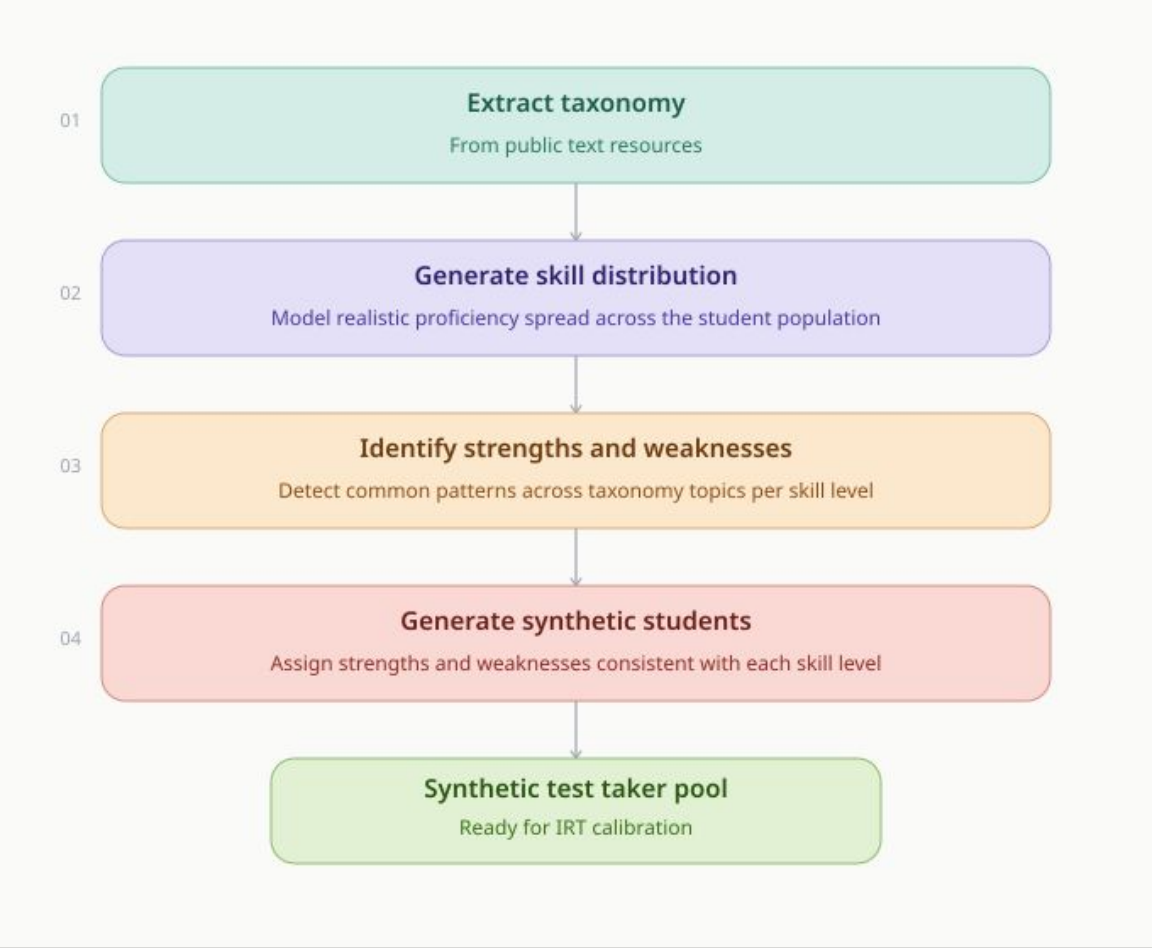
NEXTLAB.TECH

Synthetic Test Takers

- Simulated students
- Full ability distribution
- Thousands of responses
- Models behavior & errors

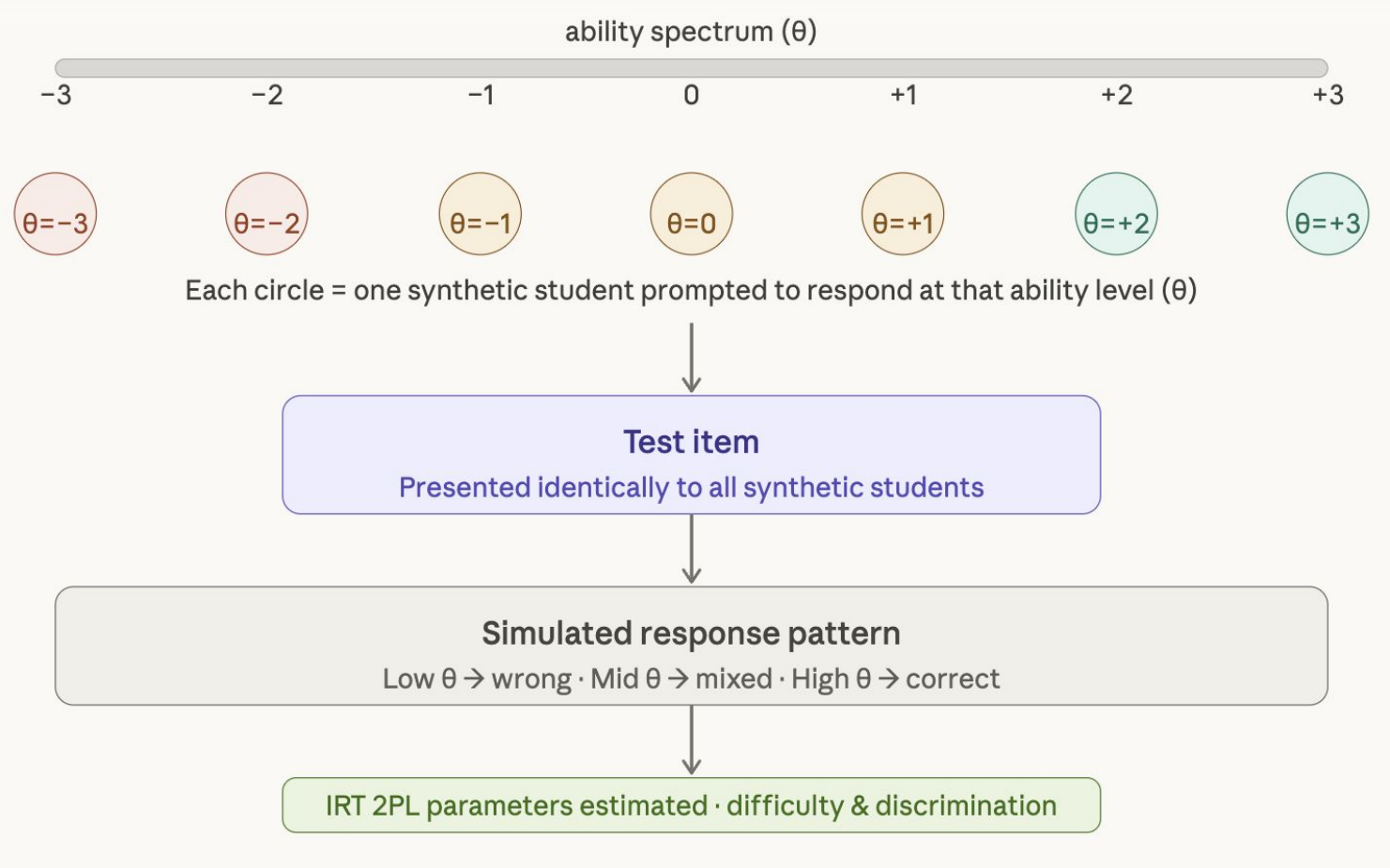
NEXTLAB.TECH

Test taker generation



NEXTLAB.TECH

LLM-based item calibration via synthetic students



NEXTLAB.TECH

Use Cases

- Pre-pilot testing
- Test equating
- Adaptive calibration
- Bias detection
- Cut-score validation

NEXTLAB.TECH

Impact and Trade-offs

Measured impact

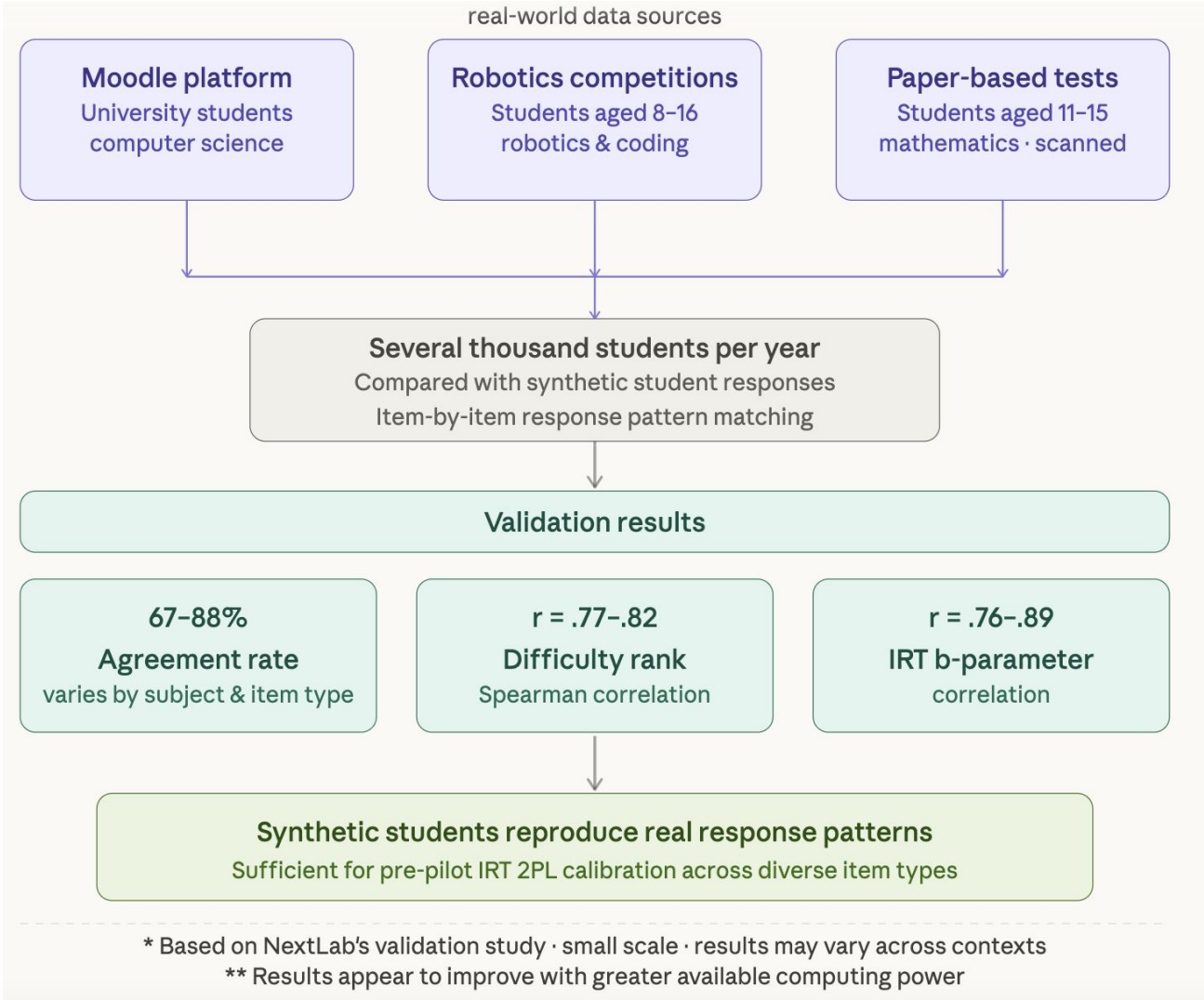
82%	12×	10,000+
cost reduction	faster delivery	synthetic students

Pros	Cons
Democratic access Quality assessment in more countries & institutions	**Large interdisciplinary team** AI engineers, data scientists & SMEs
Multilingual Items generated across languages at scale	**Huge cloud consumption** Up to $50,000 / day in AI cloud services
Efficiency 12× faster · items ready in days, not months	**Large item banks favoured** More cost-effective for larger-scale programmes
Cost savings 82% reduction in per-item production cost	**LLMs still make mistakes** Errors in content, logic & calibration estimates

* Figures based on NextLab's experience · still at small scale · results may vary at full scale

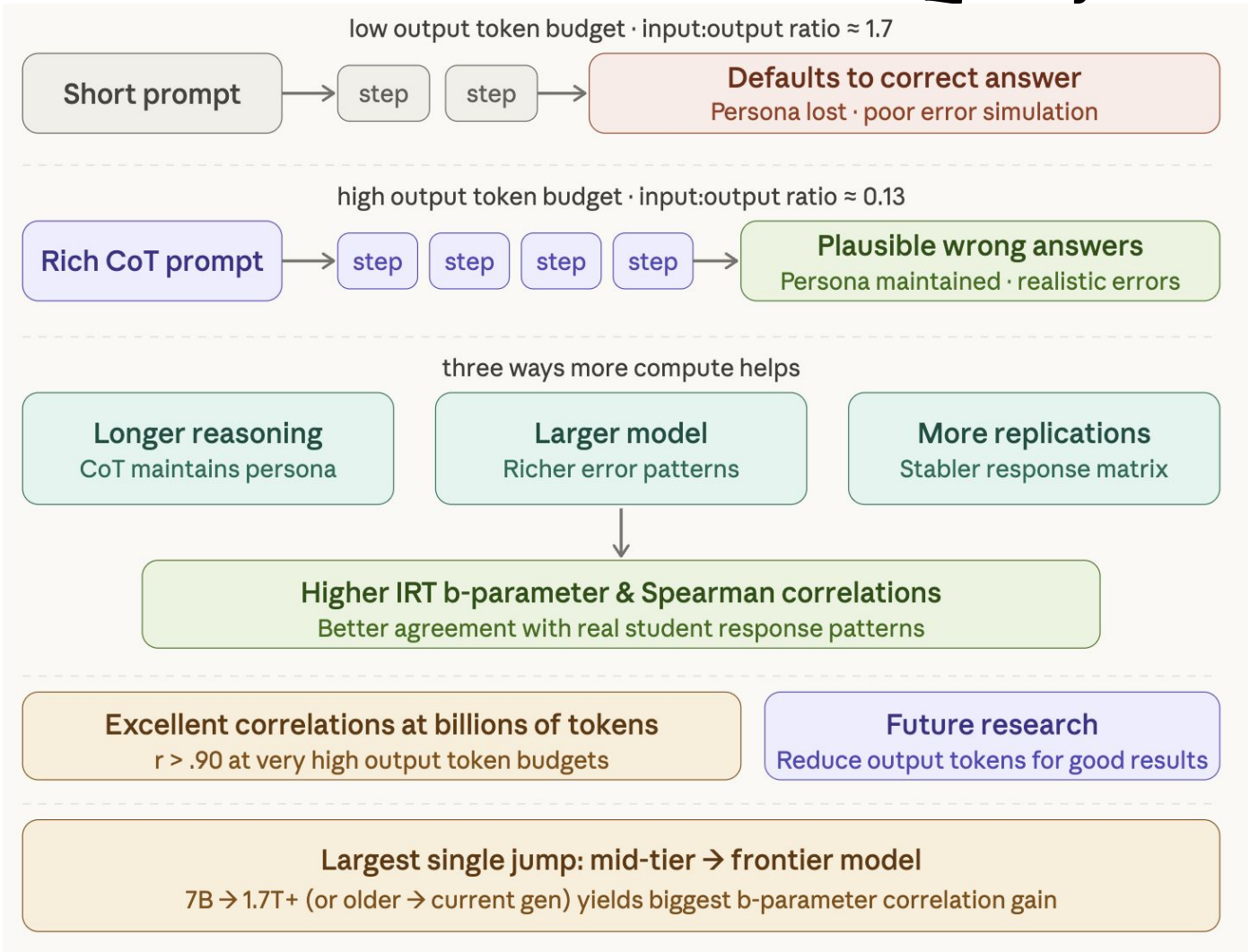
NEXTLAB.TECH

Method validation - synthetic vs real test takers



NEXTLAB.TECH

More Tokens=Better Simulation Quality



NEXTLAB.TECH

Deployments (Eastern Europe)

- 🇷🇴 Romania(4 projects)
  - Romania Tech Nation program ,7 years(2019-2025), 3000-5000 test takers per year, age 8-16, subjects covered, robotics and computer science, 2024-2025 using LLM for test calibration
  - 40+ schools in 3 projects
- 🇷🇸 Serbia(in discussion)
- 🇲🇩 Moldova(in discussion)

NEXTLAB.TECH

Thank You

AI makes it possible.

Humans make it right.

razvan.bologa@nextlab.tech