

AI Safety Colombia

¿Qué necesita el campo de AI safety
para crecer?



**AI Safety
Colombia**

El ecosistema de AI safety tiene tres insumos complementarios:

- Plata.
- Ideas/talento investigador.
- Capacidad de conversión (orgs, programas, grantmakers, ops).

**¿Tenemos
suficiente plata?**

OpenAI Foundation ANNOUNCED \$25.0B headline pool · AIS-relevant ≈ \$2.5B–\$12.5B Two pillars: health / curing disease, and “technical AI resilience” (safety & security). source ↗ \$25B initial commitment; AIS-relevant range assumes the “AI resilience” pillar is ~10–50%. Separate ~26% PBC stake ≈ \$130B.	Anthropic IPO philanthropy SPECULATIVE \$37.8B headline pool · AIS-relevant ≈ \$2.0B–\$20.0B 7 cofounders pledged ~80% of wealth; large safety-aligned employee equity, much routed via Coefficient / GWWC. source ↗ Unlocked by IPO/secondary (~\$6B tender; ~\$965B valuation, May 2026). Pledges non-binding; AIS share a guessed 5–50% subset.	Good Ventures / Coefficient (today's anchor) ESTIMATE \$100M headline pool · AIS-relevant ≈ \$80M–\$150M Dustin Moskowitz / Cari Tuna, the largest current AIS funder, plus SFF (Jaan Tallinn) and LTFF. source ↗ Present-day baseline: the whole field runs on ~\$0.1–0.25B/yr today.
---	--	---

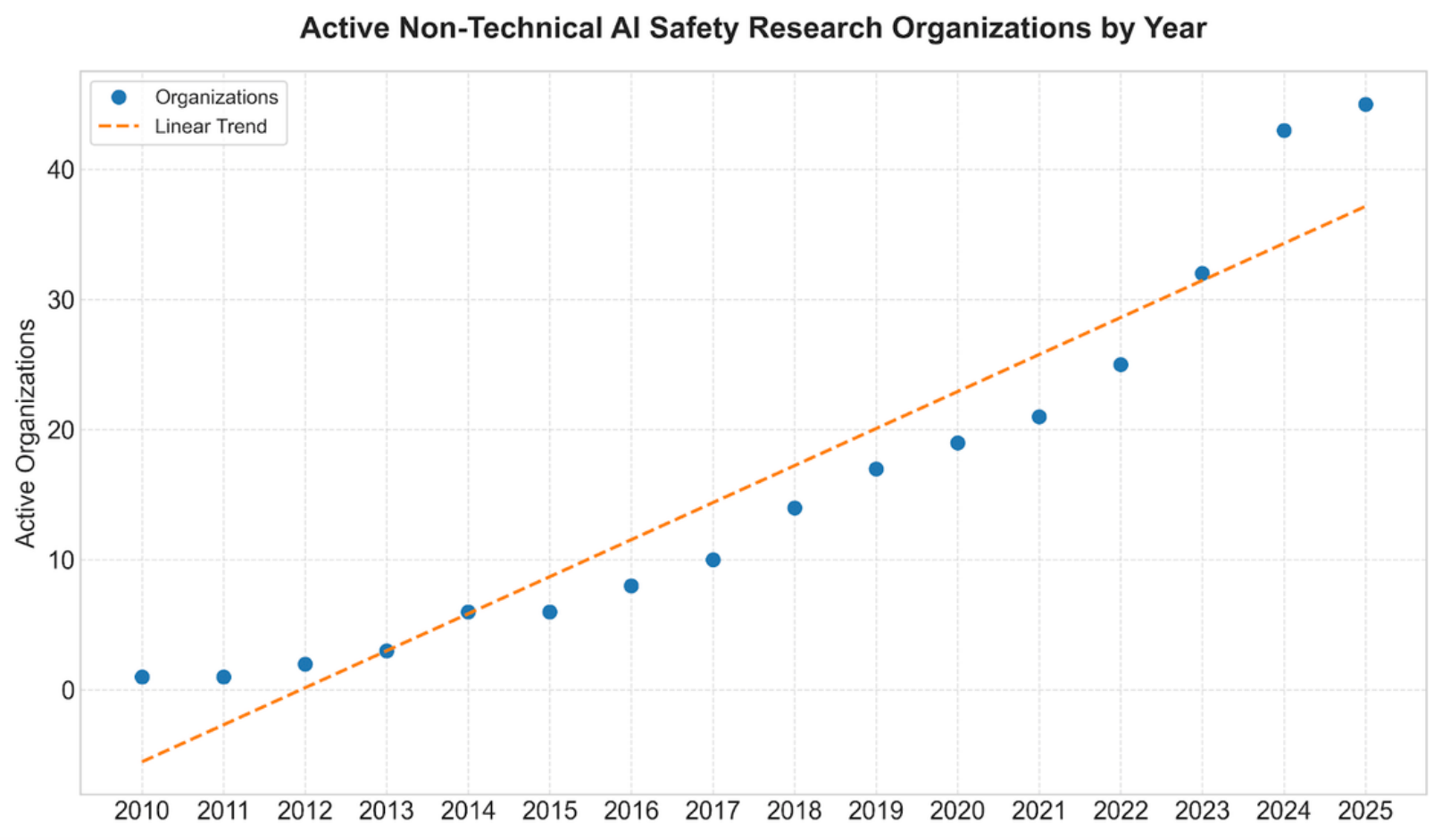
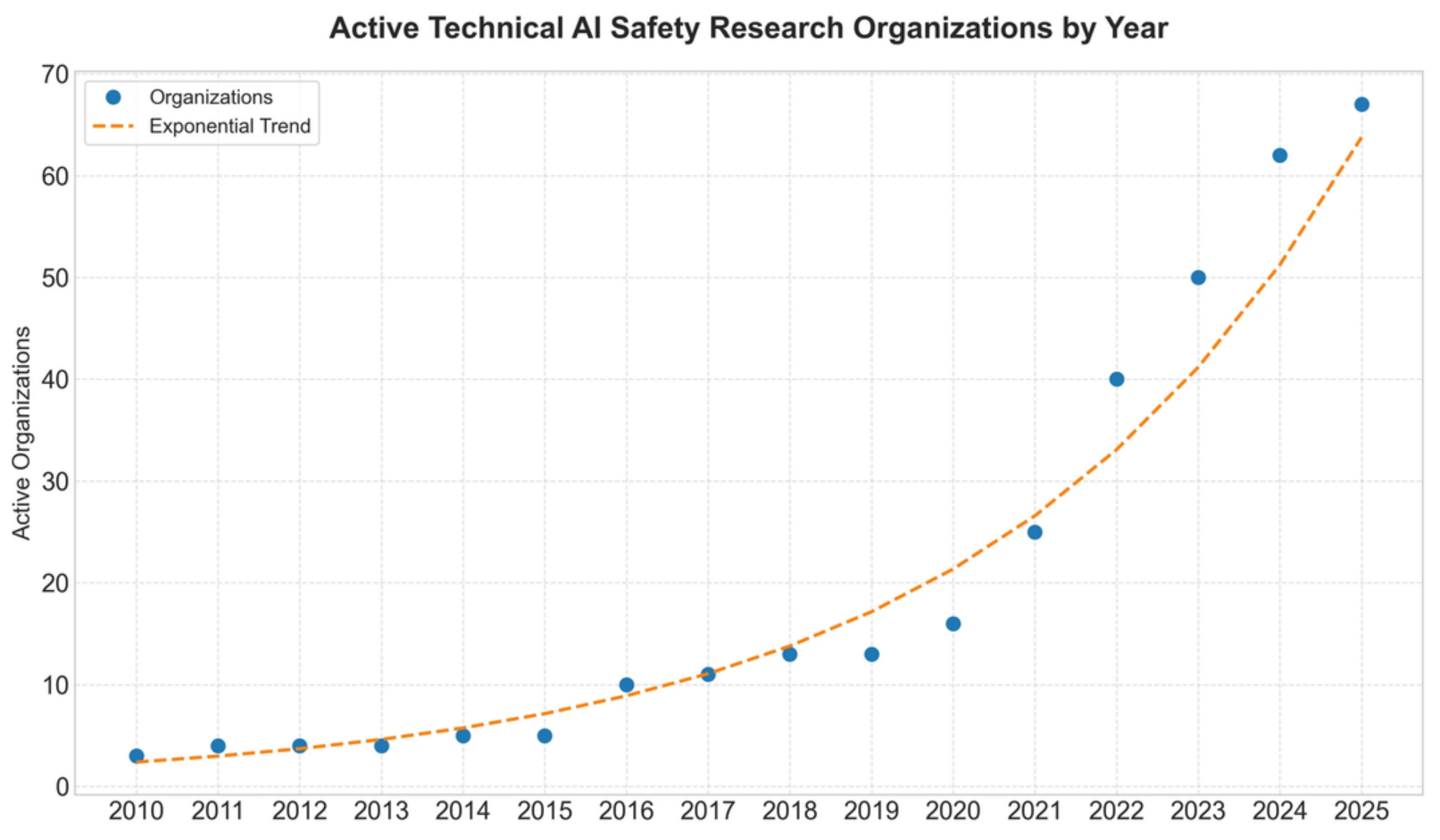
Hoy: USD ~\$0.25 mil millones/año

Plausible: USD ~\$1–10 mil millones/año

Fuente: AI Safety Field Map (Harry Waterman, 2025).

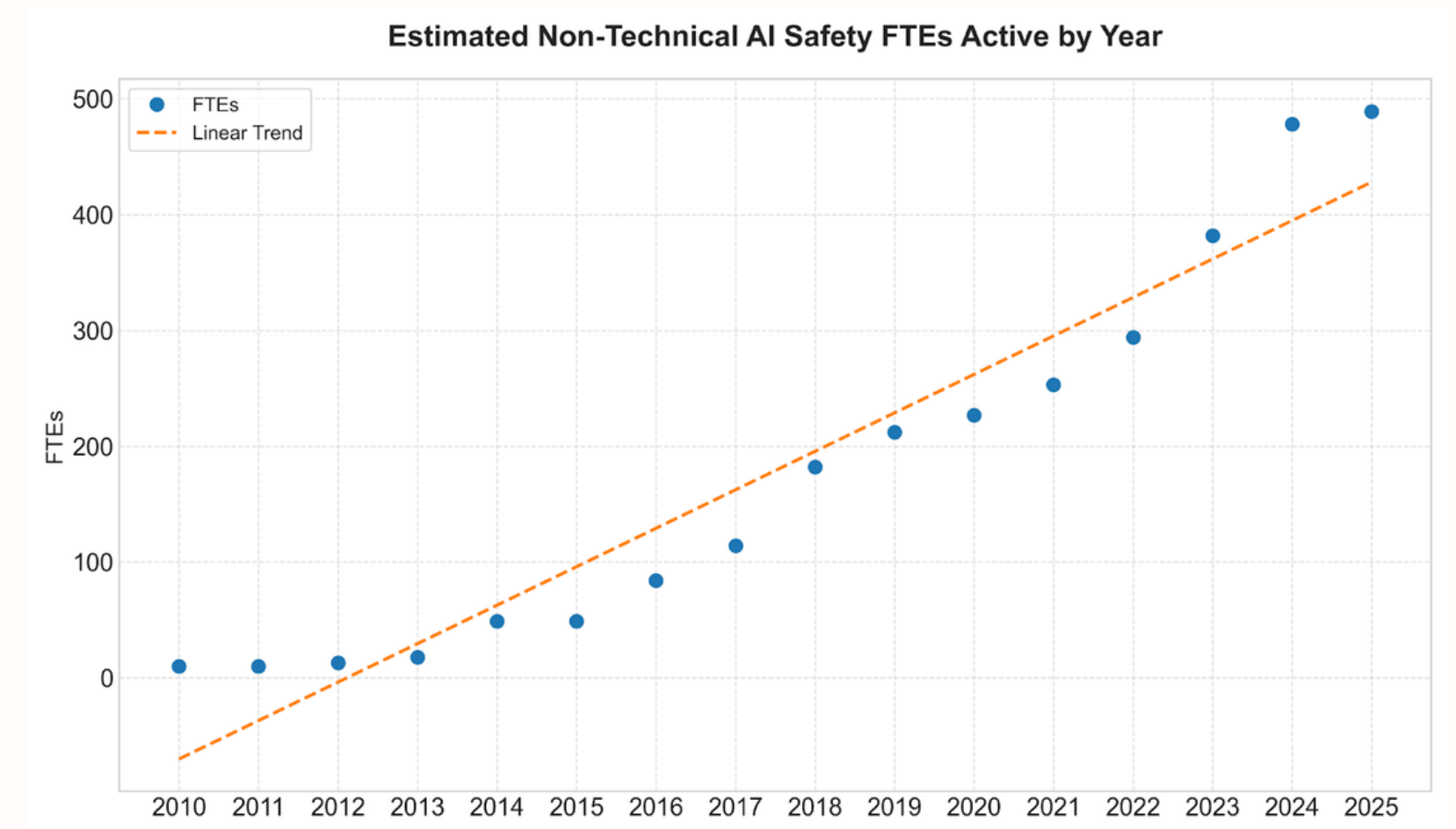
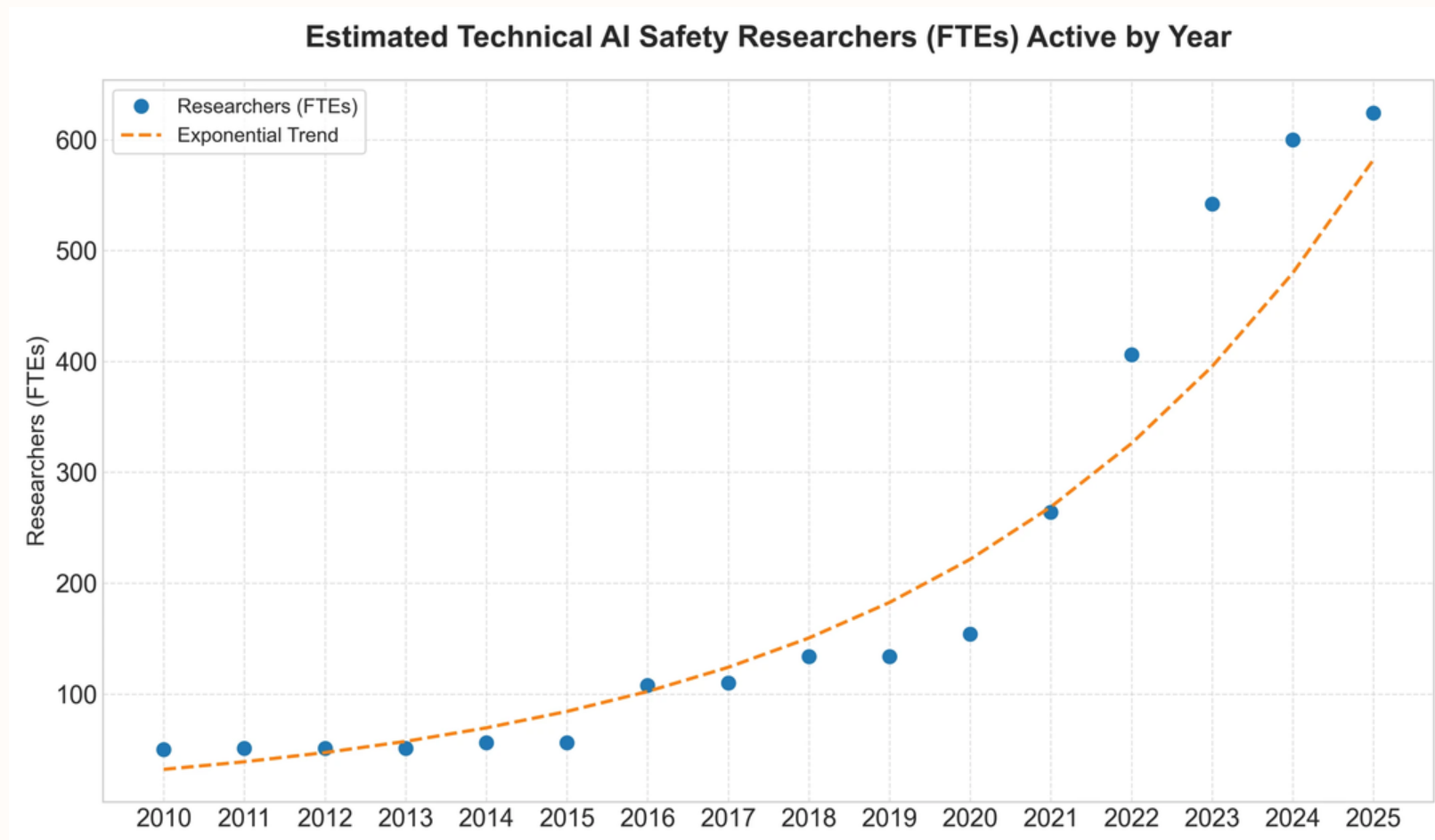
**¿Tenemos
talento/ideas?**

Crecimiento de organizaciones de AI safety



Fuente: AI Safety Field Growth Analysis 2025 (Stephen McAleese, 2025).

Crecimiento de investigadores de AI safety



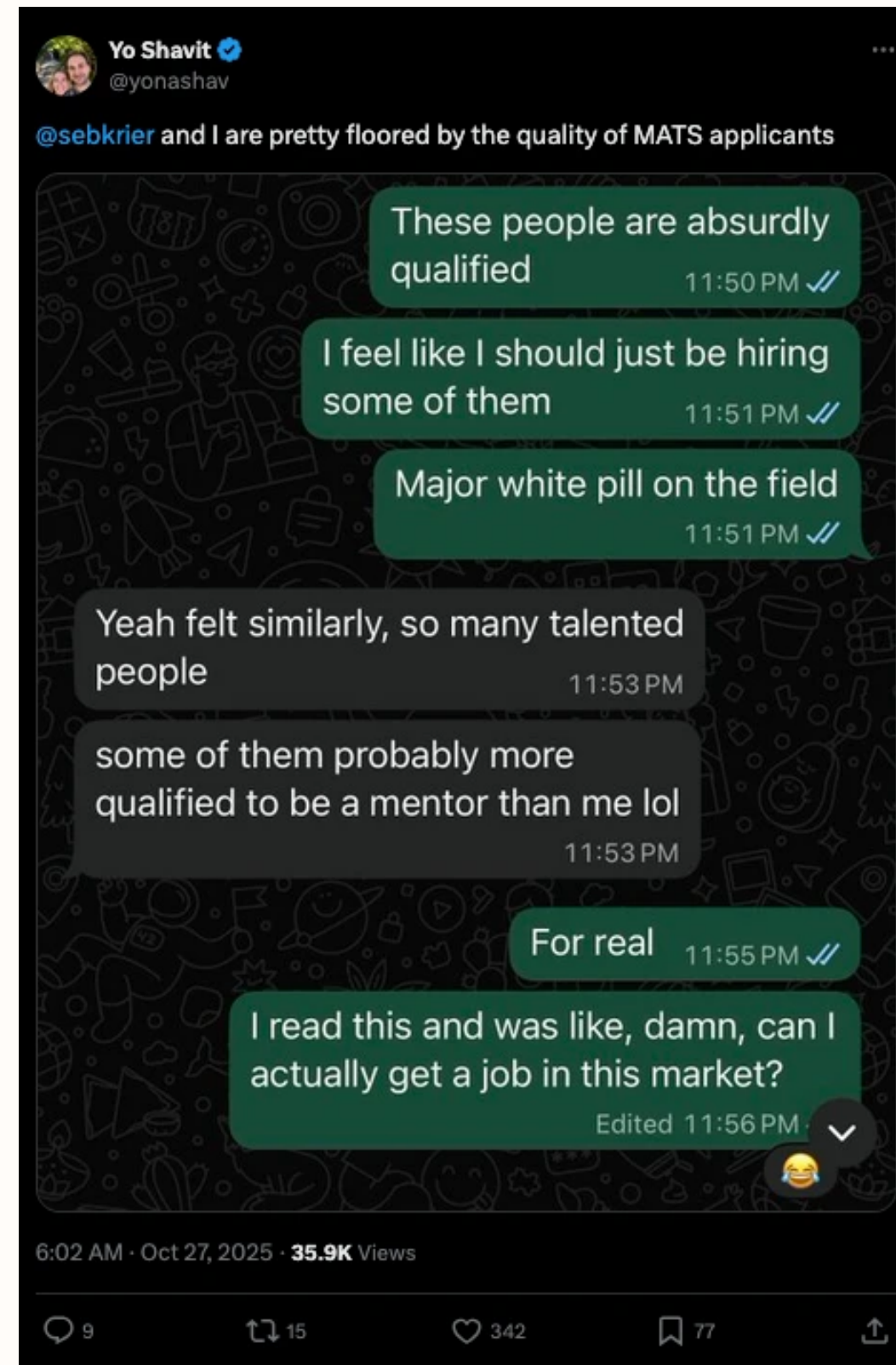
Fuente: AI Safety Field Growth Analysis 2025 (Stephen McAleese, 2025).

**¿Tenemos suficiente
capacidad de conversión?**

“Dado que la capacidad para investigar *grants* es limitada, contratar a un grantmaker menos normalmente significa que esos millones simplemente permanecerán en una cuenta durante otro año, en lugar de destinarse a fines útiles.

...Con frecuencia cerramos procesos de contratación con menos personas contratadas de las que habíamos previsto.” **(traducido del inglés)**

Fuente: AI safety is extremely bottlenecked on grantmakers (Luke Muehlhauser, 2026).



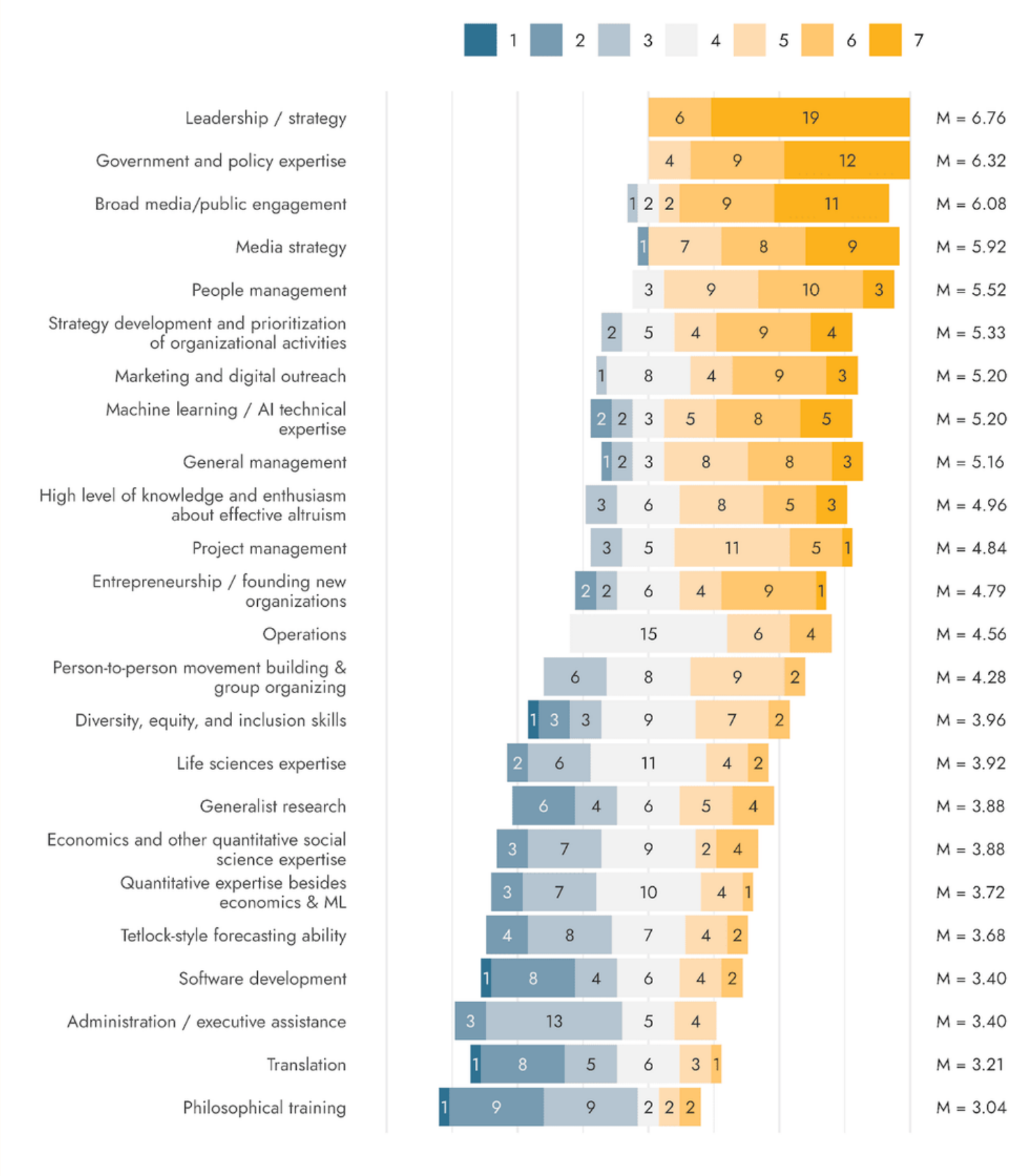
Fuente: AI Safety has a scaling problem (Boyd Kane, 2025).

Survey of 26 AI safety hiring managers and experts · approximate mentions per role type

Operations is the biggest hiring gap in AI safety



Fuente: Scaling organisations making AI go well (Matt Beard, 2026).



Fuente: Meta Coordination Forum (2024) / Talent Need Survey (David Moss & Willem Slegers, 2025).

**¿Qué se está haciendo
para solucionarlo?**

	Technical research	Policy & governance	Generalist / ops / comms
Stage 3 — Frontier (competitive, usually full-time)	MATS (+ Neel Nanda stream) Anthropic Fellows OpenAI Safety Fellowship Astra (Constellation) CHAI Fellowship PIBBSS Iliad Fellowship	Horizon (DC placements) RAND TASP IAPS Institute for Law & AI TechCongress Navigators Incubator	Real jobs / founding / grantmaking Catalyze Impact (incubator) Seldon Lab Constellation Incubator BlueDot Incubator Week
Stage 2 — Mentored / first research (some context; a real project)	SPAR (Kairos) ARENA LASR Labs ERA:AI Pivotal CBAI Summer Global AI Safety Fellowship MARS AI Safety Camp	GovAI Fellowship Talos (EU) FAS Policy Entrepreneurship Arcadia Impact Taskforce ERA:AI (gov track) CAIS AI & Society MATS (policy track) SE-Asia gov fellowship	Generator Residency Tarbell (journalism) Frame (comms) Pathfinder (Kairos — univ organizing) Gen Stream
Stage 1 — Foundations (structured intro, low bar)	BlueDot AISF — Alignment ML4Good (technical) Iliad Intensive BASE (technical track)	BlueDot AISF — Governance AI Safety Collab (gov) ML4Good (governance) BASE (governance track)	BlueDot Context Week Lens Academy (AI Risk / AI Futures) Intro to Cooperative AI Global Challenges Project (Kairos) Lateral Workshop
Stage 0 — Do something now (no gate; start this week)	Apart hackathons Self-study (AI Safety Atlas, ARENA materials) Join a local university group	Reading groups Respond to a real policy consultation Policy sprints	Write in public (Substack) Run a local event 80,000 Hours / Probably Good / Successif / HIP advising

Lateral Workshop

A three-day workshop for experienced professionals moving into AI safety.

Berkeley, California • Sept 11–13

Gen Stream

SUPPORTED BY KAIROS

AI safety programme in London for generalists. Ship a real project and leave connections, and next steps.

ENGLAND

Pathfinder Fellowship

Accelerating campus leaders in AI safety and policy

[Program](#)
[Research](#)
[Mentors](#)

Founding and Field-Building

This track is for prospective founders, field-builders, high-agency generalists, and general managers launch new AI safety initiatives. Over 12 weeks (with a potential 6–12 month extension), fellows build something new with weekly guidance from mentors (founders, field-builders, sitting CEOs, program directors, etc.) who have done it themselves.

Generator Residency

A three-month residency for AI safety generalists. Berkeley, Summer 2026.

KAIROS × CONSTELLATION

Fuente: How to Properly Reject AI Safety Fellowship Applicants: Lessons from Rejection Letters from Top Organizations (Caleb Eli Price, 2026).

Skillset	n	share of all postings
Research	1571	43%
Policy	890	24.4%
Software engineering	758	20.7%
Operations ¹	702 (or ~ 554) ¹	19.2% (or ~15.2%)
Information security	683	18.7%
Management	512	14%
Strategy	298 (after June 1, 2024) ²	12.2% (after June 1, 2024) ²
Outreach	253	6.9%
Data	83	2.3%
Legal	44	1.2%
Finance	43	1.2%
Other	176	4.8%

Fuente: What 3,654 Job Postings Tell Us About Talent Needs in AI Safety. (Weronika Żurek, 2026).

“El canal de contratación más habitual fue el contacto personal: dirigirse directamente a alguien ya conocido, una recomendación de un empleado que dejaba el puesto o una relación establecida en una conferencia.” **(traducido del inglés)**

Fuente: AI safety is maturing. Is its generalist talent pipeline keeping up?
(Uladzislau Linnik, & Enisa Ismaili 2026).

**¿Cómo puedo empezar a
contribuir yo también?**

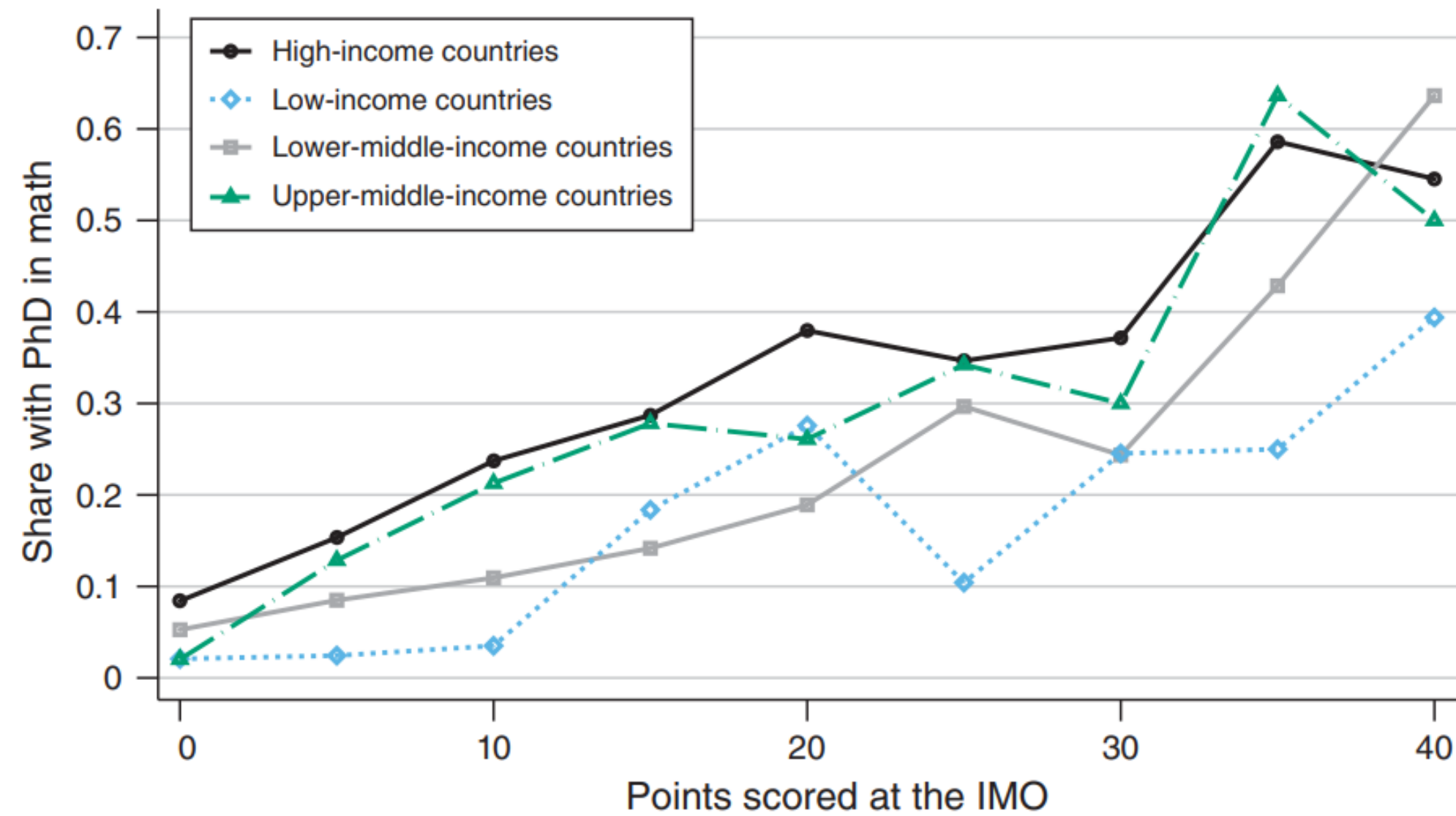


FIGURE 3. SHARE GETTING A PhD IN MATHEMATICS ACROSS COUNTRY INCOME GROUPS

Fuente: Invisible Geniuses: Could the Knowledge Frontier Advance Faster? (Agarwal & Gaule, 2020).

¡Crea una comunidad!

