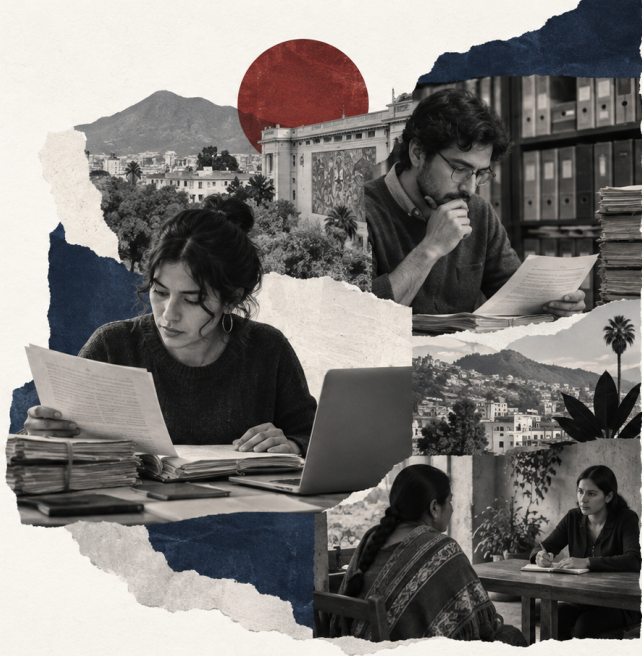

IA e investigación en ciencias sociales

Fallar antes, aprender mejor

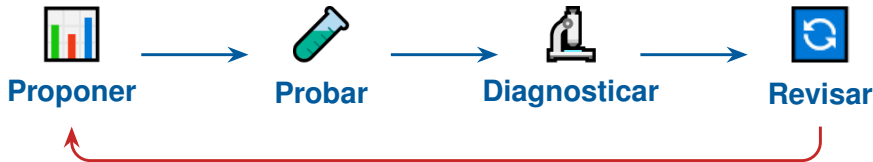
Valentina González-Rostani

University of Southern California



Investigar es un proceso iterativo

Los agentes de IA organizan tareas, usan herramientas y revisan su trabajo con nueva información.



Explorar alternativas y detectar debilidades antes de ampliar el estudio.

Dos riesgos al investigar



Explorar muy poco

Descartar ideas prometedoras antes de probarlas.

Ejemplo: abandonar un proyecto por el costo de reunir datos.

Explorar más exige mejores pruebas.



Cuestionar muy poco

Invertir demasiado en diseños o medidas débiles.

Ejemplo: aplicar a más casos una medida que confunde conceptos.

La crítica debe buscar fallas



Proponer

Diseñar alternativas y explicitar supuestos.



Cuestionar

Buscar errores, contraejemplos y condiciones de fracaso.



Contrastar

Verificar con datos, fuentes y revisión humana.

El acuerdo entre agentes no sustituye evidencia independiente.

Creatividad y crítica en la investigación

Aplicación	Falla de creatividad	Falla de crítica
	Oportunidades perdidas	Inversión en ideas débiles
Medición de organizaciones	Se descartan proyectos por el costo de reunir muchos datos.	Se adoptan medidas defectuosas que afectan los análisis posteriores.
Experimentos y auditorías	Diseños débiles o con baja potencia pueden ocultar efectos reales.	La evidencia contaminada o poco validada parece creíble.
Papel de la IA agéntica	Amplía la investigación al reducir los costos de exploración e iteración.	Mejora la evaluación con validación iterativa y pruebas exigentes.

La IA como herramienta de investigación

Seis usos: validar según el objetivo



Seis usos de la IA en ciencias sociales

Grandes modelos de lenguaje (LLM)



**1. Anotación
y medición**



2. Experimentos



**3. Encuestas
sintéticas**



4. Simulaciones



**5. Recolección
de datos**



6. Auditoría de IA

La revolución de los costos

Ejemplo ilustrativo de 2026: clasificar un millón de tuits.

Configuración intensiva

GPT-5.2.

Alto esfuerzo de razonamiento.

1 tuit por solicitud.

Sin optimizar la ejecución.

US\$149 000

Configuración optimizada

GPT-5-nano.

Sin razonamiento.

100 tuits por solicitud.

Agrupación y API por lotes.

US\$2,90

Costo estimado $\approx$ 50 000 veces menor. Falta validar la calidad.

Seis usos para fallar antes y aprender mejor

Una aplicación del marco de [Fallar antes, aprender mejor](#).

- | | |
|--------------------------------|--|
| 1. Anotación y medición | Probar medidas y corregir errores a tiempo. |
| 2. Experimentos | Pilotear tratamientos antes de ampliar el estudio. |
| 3. Encuestas sintéticas | Detectar problemas en las preguntas antes del piloto con personas. |
| 4. Simulaciones | Poner a prueba mecanismos y supuestos. |
| 5. Recolección de datos | Hacer viables proyectos antes demasiado costosos. |
| 6. Auditoría de IA | Buscar fallas en las respuestas del modelo. |

Simulaciones y auditoría de IA extienden el marco general del capítulo.

1. Anotación y medición

Convertir textos, imágenes o videos en categorías o medidas.

Definir la medida

Criterios de codificación, ejemplos y revisión de desacuerdos.

Evaluar errores por grupo e idioma con casos reservados para la prueba.

Ejemplo hipotético

Una categoría aparece en el 0,5% de los casos.

99,5%

de exactitud al responder siempre «no», omitiendo todos los positivos.

La exactitud global puede ocultar errores en categorías poco frecuentes.

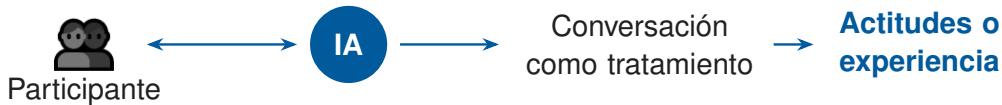
2. Experimentos

El estímulo es lo que recibe o experimenta la persona.

La IA crea el material



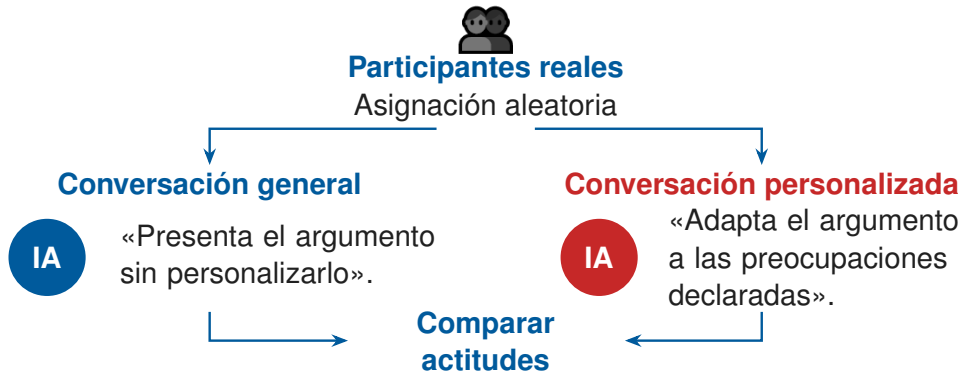
La IA es el interlocutor



Definir qué cambia, registrar lo generado y hacer pruebas piloto con personas.

2. Experimentos: conversar con una IA

¿Personalizar una conversación cambia las actitudes?



Ejemplo ilustrativo. Mantener comparables el tema, la extensión y la información.

3. Encuestas sintéticas

Perfil construido



35 años, Lima
Empleo informal

LLM

Modelo de lenguaje

Respuesta generada

6/10

Ejemplo ilustrativo

«Con este perfil, califica de 0 a 10 tu confianza en la municipalidad».

Validar según el objetivo

Individuos

Errores por persona y subgrupo.

Poblaciones

Medias, dispersión y distribuciones.

Relaciones o efectos

Contraste con evidencia independiente.

Reproducir promedios no garantiza reproducir relaciones ni efectos.

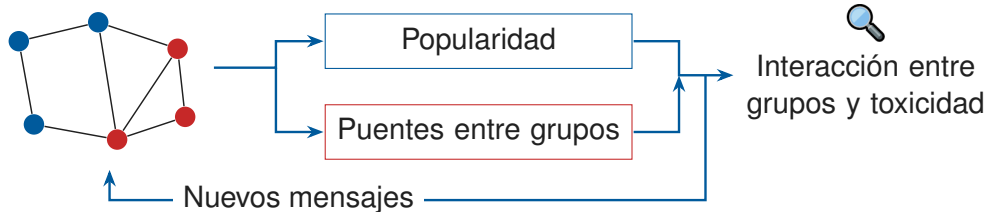
4. Simulaciones

¿Cómo cambia la conversación al recomendar otros contenidos?

Mismos perfiles y red

Cambiar la recomendación

Medir



Ejemplo: «Con tu perfil y memoria, decide qué publicar ante estos mensajes».

Verificar

Probar instrucciones, memoria y orden.

Validar

Contrastar patrones con datos externos.

5. Recolección de datos



Localizar

Definir el universo de documentos y los criterios de inclusión.



Extraer

Guardar cada valor, su fuente y la fecha de consulta.



Verificar

Revisar cobertura, datos faltantes y errores de extracción.

Una base útil permite rastrear cada dato y proteger información sensible.

6. Auditoría de IA

Auditar su desempeño y estudiar las consecuencias de su uso.

Qué estudiar

Capacidades y sesgos en tareas concretas.

Efectos sobre decisiones y conductas de personas u organizaciones.

Auditar respuestas

Controlar contexto, historial y orden de opciones.

Definir repeticiones y alcance: idiomas, versiones y fechas.

Las respuestas del modelo no prueban por sí solas efectos sobre personas.

Aprender de las fallas

Detectar una falla permite revisar el diseño o acotar la conclusión.

Falla detectada



La medida confunde marcos económicos y culturales.



La personalización cambia también el tono.



El patrón desaparece con cambios menores en las instrucciones.



Siguiente paso

Separar criterios y volver a codificar.

Revisar los estímulos y repetir el piloto.

Revisar supuestos.
Validar con datos externos.

Detectar fallas a tiempo permite revisar antes de invertir más.

Evaluación y rigor

Buscar fallas antes de confiar



Desafíos de la investigación con IA



Reproducibilidad

Los modelos cambian o dejan de estar disponibles.



Referencias de validación

Los criterios humanos también necesitan revisión.



Sesgos del modelo

Los errores pueden variar entre grupos y tareas.



Equidad y acceso

Los recursos para investigar difieren entre instituciones.



Opacidad

Los sistemas cerrados limitan qué podemos inspeccionar.



Obsolescencia rápida

Actualizar el modelo exige reevaluar los resultados.

Buenas prácticas para investigar con LLM



1. Comparar costo y calidad antes de ampliar el estudio.



3. Validar según la pregunta, el grupo y el contexto.



5. Revalidar al cambiar el modelo o los datos.



7. Guardar entradas y salidas, protegiendo datos sensibles.



2. Separar la generación de alternativas de su crítica.



4. Registrar versión y fecha, instrucciones y parámetros.



6. Fijar el protocolo y reservar datos para el análisis confirmatorio.



8. Explicitar qué se quiere inferir y con qué límites.

Los registros permiten auditar. Repetir el proceso puede dar resultados distintos.

Ideas clave para investigar con IA



Explorar más alternativas y detectar fallas antes.



Comparar costo y calidad antes de ampliar el estudio.



Validar según la pregunta, el contexto y la población.



Documentar instrucciones, versiones y decisiones.



Contrastar los resultados con evidencia independiente.



El equipo responde por las inferencias, los sesgos y la privacidad.

Gracias

gonzalez.rostani@usc.edu

gonzalez-rostani.com

Referencias

Crabtree, González-Rostani y Kim (2026)

Charles Crabtree, Valentina González-Rostani y Jae Yeon Kim.

Failing Faster, Learning Better: Agentic AI and Empirical Social Science Research.

En Rohan Alexander (ed.), *Quantitative Social Science in the Age of AI*, capítulo 8.

[Leer el capítulo](#)

Barrie et al. (2026)

Christopher Barrie, Lisa Argyle, James Bisbee, Michael Heseltine, Christopher Lucas, Jon Mellon, Alexis Palmer, Margaret Roberts y Arthur Spirling.

AI and Research Methods. APSA Preprints.

[doi:10.33774/apsa-2026-h59kk](https://doi.org/10.33774/apsa-2026-h59kk)