

JUNIO 2026

Construyendo evidencia sobre el uso de IA en la financiación de I+D

El experimento en la sombra de Neotec

Hugo Cuello

Analista Senior de Políticas · Innovation Growth Lab

El Innovation Growth Lab



Un **laboratorio de políticas** para aumentar el impacto de los programas de innovación y productividad. Somos una **organización sin ánimo de lucro** que une a gobiernos e investigadores para afrontar retos clave por medio de nuevas ideas, experimentación, datos y evidencia.



50+ AGENCIAS DE GOBIERNO

Apoyadas para hacerlas más experimentales y informadas por datos

70+ EXPERIMENTOS

Apoyados en 28 países

300+ INVESTIGADORES

Forman parte de la IGL Research Network

Las políticas de ciencia, innovación y productividad no siempre son muy científicas, innovadoras y productivas. Existimos para cambiar esa realidad.

El programa Neotec

CONTEXTO · 1.1

Instrumento estrella del CDTI para apoyar a empresas de base tecnológica en fase temprana.

≈40M€

Presupuesto en 2025
(la mitad para Seguridad y Defensa, 5M
para proyectos liderados por mujeres)

≈250k€

Subvenciones por proyecto
(325k€ en condiciones específicas)

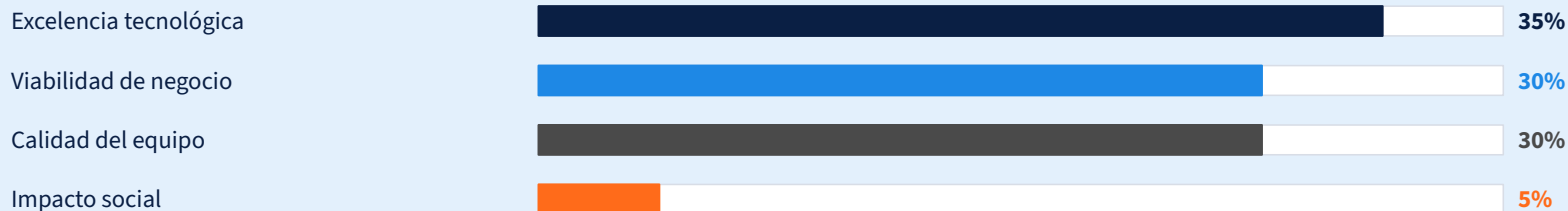
130

Solicitudes aprobadas en 2025

28%

Tasa de éxito de las solicitudes en 2025
(466 presentadas en total)

Ponderación histórica del sistema de evaluación hasta 2024



El punto de partida: Neotec y el impacto social

Hallazgos clave de la evaluación previa del programa Neotec (2018–2021). Las cifras comparan empresas **beneficiarias** (que recibieron la ayuda) frente a **no beneficiarias**.

● IMPACTO NEGATIVO R.29 · ACCESO A LA TECNOLOGÍA

Las tecnologías financiadas son menos accesibles para colectivos vulnerables

La probabilidad de que la tecnología sea accesible **se reduce** si la empresa recibe la ayuda, en comparación con las no financiadas.

PROBABILIDAD DE ACCESIBILIDAD — EFECTO DEL PROGRAMA

Población rural	–12%	Baja cualificación	–11%
Personas con discapacidad	–10%	Bajo nivel de renta	–9%

● SIN EFECTO R.32 · INCLUSIÓN SOCIAL

El programa no incrementa la contribución a colectivos vulnerables

Entre el 50% y el 60% de las empresas declara contribuir a colectivos en riesgo de exclusión, **sin diferencias significativas** entre beneficiarias y no beneficiarias.

EMPRESAS QUE DECLARAN MEJORAR LA CALIDAD DE VIDA DE COLECTIVOS VULNERABLES

BENEFICIARIAS	=	NO BENEFICIARIAS
50–60%		50–60%

● SIN EFECTO R.28 · RESOLUCIÓN DE PROBLEMAS SOCIALES

El programa no orienta la innovación hacia problemas sociales

Las áreas de impacto social abordadas son las mismas en beneficiarias y no beneficiarias. **No hay efecto atribuible al programa.**

PROBLEMAS SOCIALES ABORDADOS POR LOS PROYECTOS NEOTEC

Salud y calidad de vida 60%	Brecha digital 24%
Desigualdades sociales 11%	Educación 9%
	Pobreza 6%
Desigualdad de género 6%	

● IMPACTO NEGATIVO R.27 · TECNOLOGÍAS MEDIOAMBIENTALES

Menor desarrollo de tecnologías para adaptación climática e hídrica

Las beneficiarias son **menos propensas** a desarrollar este tipo de tecnologías que las no financiadas.

EFECTO DEL PROGRAMA SOBRE LA PROBABILIDAD DE INNOVAR

Adaptación al cambio climático	–18%	Uso sostenible de recursos hídricos	–14%
--------------------------------	------	-------------------------------------	------



La evaluación previa concluye que Neotec, en su forma actual, **no genera mejoras en impacto social** y, en varios indicadores clave (accesibilidad, adaptación climática, recursos hídricos), **tiene un efecto negativo** respecto a no recibir la ayuda. Este es el punto de partida que motiva el experimento.

Un nuevo enfoque para abordar el problema

CONTEXTO • 1.3

El CDTI se convierte en miembro de IGL en 2024, que abrió más la puerta a la experimentación y al uso de evidencia. Comenzamos un proceso de ideación y reflexión junto con académicos de primer nivel.

LA PREGUNTA DE INVESTIGACIÓN

¿Puede la IA ayudar al CDTI a identificar mejor los proyectos con mayor impacto social?

Posible solución: Mejorar la evaluación del criterio de impacto social con apoyo de la IA.

Poner a prueba las evaluaciones humanas, con diferentes niveles de experiencia en Neotec y en impacto social, y observar si la IA les ayuda a tomar mejores decisiones.

QUÉ PODRÍA RESPONDER UN EXPERIMENTO

- ✓ Si la IA mejora la consistencia de las evaluaciones
- ✓ Si reduce el ruido en las puntuaciones
- ✓ Si ayuda a identificar señales relevantes
- ✓ Si los evaluadores realmente la usan y responden a ella

Del RCT al experimento en la sombra

CONDICIONES PARA EL EXPERIMENTO · 2.1

La ambición inicial chocó con barreras institucionales y se reorientó hacia un diseño más viable.

PLAN ORIGINAL

Ensayo controlado aleatorio (RCT) sobre las decisiones reales de financiación

- ✗ Conflictos legales con principios de competencia justa
- ✗ Riesgos de impugnación por parte de solicitantes
- ✗ Horizontes temporales largos para medir impacto
- ✗ Coordinación política y administrativa compleja



SOLUCIÓN ADOPTADA

Experimento en la sombra: en paralelo al proceso real, sin afectar las decisiones de financiación

- ✓ Se conserva el rigor experimental sin riesgo legal
- ✓ Permite comparar diferentes enfoques de evaluación
- ✓ Las evaluaciones se hacen tras concluir la oficial
- ✓ Aprendizaje operativo dentro del marco institucional

La arquitectura de la colaboración

CONDICIONES PARA EL EXPERIMENTO · 2.2

Tres actores con roles complementarios hicieron posible el experimento.

RESPONSABLE PÚBLICO

Departamento de Estrategia y Evaluación, CDTI

- Conocimiento institucional
- Acceso a datos y procesos operativos
- Implementación dentro de la organización

LÍDER ACADÉMICO

Humboldt University (IGL Affiliate)

- Diseño metodológico
- Rigor de investigación y protocolos
- Análisis estadístico de datos y resultados

INTERMEDIARIO (KNOWLEDGE BROKER)

IGL

- Traducción entre lógicas académica y de política
- Coordinación entre actores
- Diseño del sistema de IA y gestión de expertos externos

El sistema de IA desarrollado por IGL

EL EXPERIMENTO • 3.1

Entrenado con **283 solicitudes** aprobadas (2021–2023), combinando **textos de solicitudes** y **resultados de impacto social** de las encuestas de fin de proyecto.

1

Modelo predictivo de lenguaje

Analiza el resumen ejecutivo y la sección de impacto social de cada propuesta para generar una puntuación numérica de 0 a 10.

PRODUCE INTERNAMENTE (NO SE COMPARTE CON EL EVALUADOR)

Puntuación: 7.5 / 10

Escala 0-10 de impacto social esperado

2

Modelo generativo (Gemini)

Usa la propuesta y la puntuación predicha para generar una justificación textual con evidencia relevante, protegiendo la confidencialidad.

PRODUCE

Texto explicativo

Probado en dos formatos: neutral y con justificaciones (persuasivo)

Diseño del experimento

EL EXPERIMENTO • 3.2

Los evaluadores revisan el **Resumen Ejecutivo** y la sección de **Impacto Social** de cada solicitud, previamente anonimizada.

Revisión de solicitudes

Los **expertos externos** y los **voluntarios CDTI** evalúan y puntúan propuestas históricas (2021-2023) y algunas recientes (2025)
Los **evaluadores Neotec** revisan principalmente propuestas recientes (2025), con sus puntuaciones originales

Tras puntuar (o revisar) la solicitud, los participantes ven el **texto generado generado por IA**
Se les ofrece la oportunidad de mantener o modificar su puntuación original

1

Análisis Fase I • Capacidad predictiva

¿Cuál es la relación entre las puntuaciones y los resultados de los proyectos (en base a las encuestas de fin de proyecto)? ¿La IA predice mejor los resultados posteriores?

2

Análisis Fase II • Uso de IA

¿Cómo utilizan los evaluadores los aportes de la IA?
¿Modifican sus puntuaciones? Si lo hacen, ¿cómo y cuándo? ¿Mejora la IA la relación entre puntuación y resultado?

Ejecución del experimento

EL EXPERIMENTO • 3.4

La participación de **116 evaluadores** combinó técnicos de Neotec, voluntarios del CDTI y expertos externos en impacto social.

30

Expertos externos

Especializados en impacto social

52

Voluntarios CDTI

Sin experiencia previa en Neotec

34

Evaluadores Neotec

Técnicos del programa

348

Propuestas revisadas

828

Evaluaciones realizadas

83%

Tasa de participación

≈70%

Completaron en la sesión presencial

Resultados principales

Versión resumida

Más resultados y explicaciones en el informe del CDTI y en el futuro informe académico

Todos los evaluadores tienen capacidad predictiva limitada



Cómo medimos la capacidad predictiva

Calculamos la capacidad predictiva en una escala de:



Se basa en cómo los evaluadores ordenan las propuestas de mejor a peor, comparando esto con el ranking de resultados.



Qué captura la escala

Una puntuación de 0 significa que los rankings están completamente invertidos (la peor predicha como la mejor).

Una puntuación de 100 significa alineación perfecta entre los rankings predichos y los resultados reales.

El modelo de IA puede tener mayor capacidad predictiva



Predicción del modelo IA

(N=25)

46



Evaluadores humanos

(N=894)

38

Qué muestra esta comparación

Al calcular la capacidad predictiva de las puntuaciones producidas por el modelo de IA, obtenemos una métrica ligeramente superior a la de los evaluadores humanos.



Limitación: esto se basa en una muestra separada y muy pequeña (N=25).

La IA puede mejorar las puntuaciones de los evaluadores



Qué muestra esta comparación

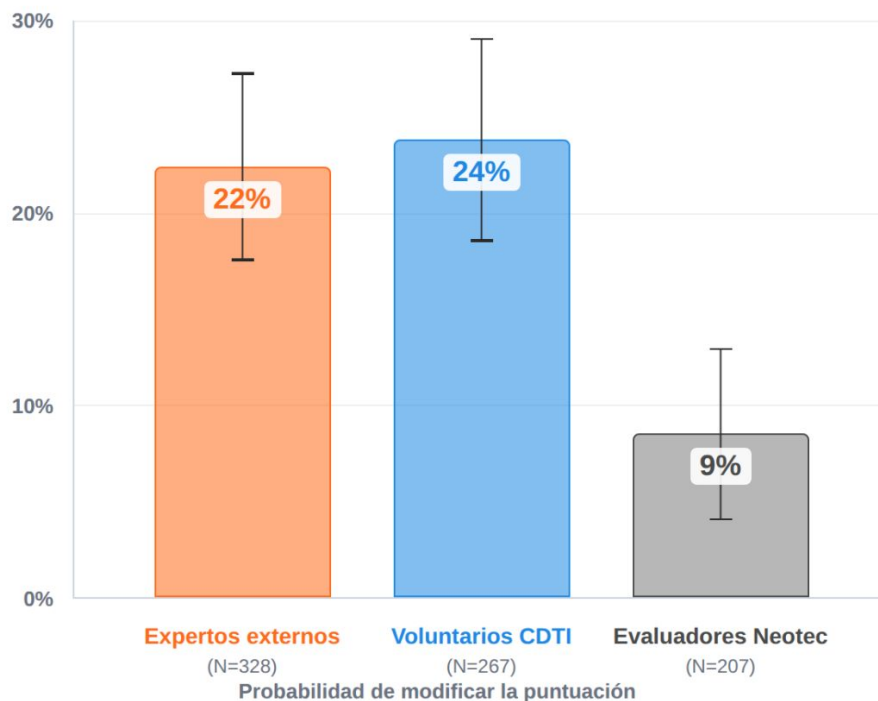
◆ Sin IA ◆ Con asistencia de IA

En todos los grupos de evaluadores, la capacidad predictiva aumenta cuando las puntuaciones se producen con asistencia de IA, observándose la mayor mejora entre los evaluadores Neotec.



Limitación: para los evaluadores Neotec solo disponemos de una muestra de 31 propuestas del periodo 2021–2023. Revisaron principalmente sus propuestas de 2025, para las cuales aún no tenemos datos de resultados.

Los evaluadores usan la IA de forma selectiva

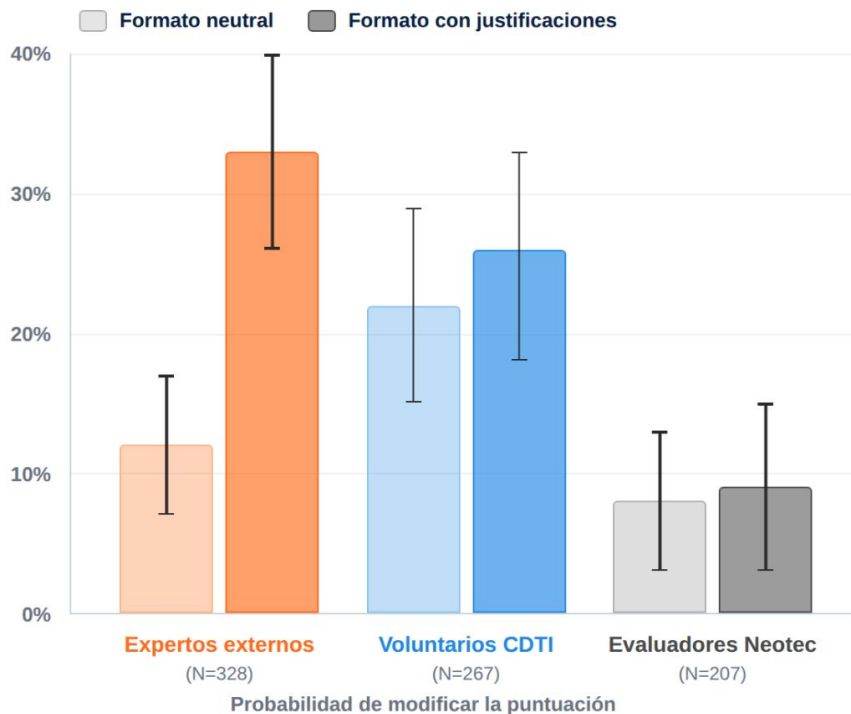


Cuándo modifican su puntuación los evaluadores

Los evaluadores son más propensos a modificar su puntuación cuando la IA:

- a Sugiere **bajar la puntuación** (miedo a los falsos positivos).
- b Sugiere una puntuación **más alejada de la suya** (información novedosa y útil).
- c Proporciona una **sugerencia correcta** (más cercana a los resultados reales).

El diseño importa: el formato de la IA cambia el comportamiento



Dos formatos de presentación de la IA

Presentamos el texto generado por la IA en uno de dos formatos:

1 Formato neutral
resume la evidencia identificada por la IA (solo hechos).

2 Formato con justificaciones
resume la misma evidencia pero con lenguaje persuasivo (juicios de valor).

Lecciones de los resultados

LECCIONES · 4.1

LECCIÓN 1

Predecir impacto social es inherentemente difícil

- Baja capacidad predictiva en todos los grupos.
- Los expertos especializados no superan al resto.
- Refleja un problema estructural de medición, no solo de pericia.
- Vincula con debates internacionales sobre evaluación de procesos de financiación.

LECCIÓN 2

La IA funciona como apoyo cognitivo, no como oráculo

- Mejoras pequeñas pero consistentes en capacidad predictiva.
- Los evaluadores no siguen ciegamente las recomendaciones.
- El contexto institucional condiciona la receptividad.
- El formato de presentación puede influir en la aceptación.

Lecciones del experimento en su conjunto

LECCIONES • 4.2

LECCIÓN 3

Experimentar requiere más que pericia técnica

- La experimentación en lo público es también un reto organizacional y relacional.
- Las preocupaciones institucionales (datos, equidad, riesgos operativos) son legítimas y requieren negociación continua.
- Es clave un rol intermediario (knowledge broker) que conecte lo académico, lo operativo y lo institucional.

LECCIÓN 4

Flexibilidad y confianza son esenciales

- Los planes iniciales casi nunca se mantienen intactos.
- Adaptar el diseño (RCT → sombra) fue una fortaleza, no una debilidad.
- Los "campeones" internos en la institución fueron decisivos para superar resistencias.
- Esto es solo una prueba con un modelo de IA. Se pueden generar otros modelos en el futuro.

Gracias

El experimento en la sombra de Neotec
InnovationGrowthLab.org



*Regístrate para más noticias sobre experimentación en
políticas de innovación y ciencia*

Hugo Cuello

Analista Senior de Políticas · Innovation Growth Lab
hugo.cuello@nesta.org.uk

Junio 2026