

DORA

2024

Google Cloud

Accelerate State of DevOps

Patrocinadores Gold

 catchpoint  chronosphere  DATADOG

Deloitte.

 Excella

 Gearset

 liatrío

 Middleware

 OPSERA

10

Diez años de DORA

Contenido

Resumen ejecutivo	3	Reflexiones finales	83
Rendimiento del envío de software	9	Agradecimientos	85
Inteligencia artificial: adopción y actitudes	17	Autores	87
Análisis del impacto secundario de la IA	27	Datos demográficos y firmográficos	91
Ingeniería de plataformas	47	Metodología	99
Experiencia de los desarrolladores	57	Modelos	113
Transformación desde el liderazgo	69	Lecturas recomendadas	117
Diez años de DORA	77		

Resumen ejecutivo

DORA lleva más de diez años investigando las competencias, prácticas y métricas de empresas y equipos tecnológicos con un alto rendimiento. Este es nuestro décimo informe de DORA. Hemos recabado información de más de 39.000 profesionales de todo el mundo que trabajan en empresas de distintos tamaños y sectores. Muchas gracias por participar con una contribución tan importante para nuestra investigación.

El equipo de DORA recopila datos a través de una encuesta anual mundial entre profesionales con roles técnicos y complementarios. En la encuesta se incluyen preguntas relacionadas con las formas de trabajar y los logros pertinentes para las empresas y las personas que trabajan en ellas.

Usamos una rigurosa metodología de evaluación estadística para entender las relaciones entre estos factores y averiguar cómo contribuye cada uno de ellos al éxito de los equipos y empresas.

Este año, hemos ampliado la encuesta con entrevistas en profundidad a profesionales para obtener información más exhaustiva, triangular los datos y contextualizar mejor nuestros hallazgos. Para obtener más información consulta el capítulo [Metodología](#).

Estos son los principales logros y resultados que hemos investigado este año:

Reducción del desgaste

El desgaste es un estado de cansancio emocional, físico y mental provocado por un estrés prolongado o excesivo. Normalmente, se caracteriza por sentimientos de escepticismo, desapego y ausencia de logros personales.

Flujo

El flujo mide el nivel de concentración que tiende a alcanzar una persona durante las tareas de desarrollo.

Satisfacción laboral

La satisfacción laboral mide las sensaciones generales de una persona respecto a su trabajo.

Rendimiento empresarial

Mide el rendimiento de una empresa en aspectos como la rentabilidad, la cuota de mercado, el número total de clientes, la eficiencia operativa, la satisfacción del cliente, la calidad de los productos y servicios y la capacidad para cumplir objetivos.

Rendimiento del producto

Mide la usabilidad, la funcionalidad, el valor, la disponibilidad, el rendimiento (por ejemplo, la latencia) y la seguridad de un producto.

Productividad

La productividad mide el grado en el que una persona se siente eficaz y eficiente en su trabajo para crear valor y cumplir sus tareas.

Rendimiento del equipo

Este parámetro mide la capacidad de un equipo para colaborar, innovar, trabajar de manera eficiente, confiar los unos en los otros y adaptarse.

Principales conclusiones

La IA está teniendo un amplio impacto

La IA está generando un cambio de paradigma en el campo del desarrollo de software. Los resultados están siendo muy prometedores en lo que respecta al número de primeros usuarios, aunque deben tomarse con cautela.

Ventajas de la adopción de la IA:

- Flujo
- Productividad
- Satisfacción laboral
- Calidad del código
- Documentación interna
- Procesos de revisión
- Rendimiento del equipo
- Rendimiento empresarial

Sin embargo, la adopción de la IA también trae consigo algunos efectos negativos. Hemos visto cierta reducción del rendimiento del envío de software, mientras que su efecto en el rendimiento del producto aún es incierto. Además, los usuarios afirman que el aumento de la adopción de la IA está reduciendo el tiempo que dedican a tareas que generan valor, un hallazgo curioso en el que profundizaremos más adelante.

Los equipos deben seguir experimentando y aprendiendo sobre el impacto de basarse cada vez más en la IA.

La adopción de la IA aumenta conforme lo hace la confianza en ella

El uso de inteligencia artificial generativa (IA generativa) hace que los desarrolladores se sientan más productivos. Además, los que confían más en ella también la usan más. En este sentido, hay margen de mejora, ya que el 39,2 % de los encuestados afirmaron tener poca o ninguna confianza en la IA.

El enfoque orientado al usuario potencia el rendimiento

Las empresas que priorizan la experiencia del usuario final crean productos de mayor calidad, y sus desarrolladores, además de tender menos al desgaste, se sienten más productivos y satisfechos.

El liderazgo en la transformación es importante

El liderazgo en el ámbito de la transformación mejora la productividad de los empleados, la satisfacción laboral y el rendimiento empresarial, del equipo y del producto. Además, ayuda a reducir el desgaste de la plantilla.

Las prioridades estables aumentan la productividad y el bienestar

Si las prioridades de la empresa son inestables, la productividad se reduce de manera significativa y el desgaste aumenta en gran medida, incluso en aquellas que cuentan con buenos líderes, documentos internos de calidad y una estrategia de desarrollo de software centrada en los usuarios.

La ingeniería de plataformas puede impulsar la productividad

La ingeniería de plataformas tiene un impacto positivo en la productividad y el rendimiento empresarial, pero el rendimiento del envío de software invita a la cautela.

La nube posibilita la flexibilidad de la infraestructura

Una infraestructura flexible puede aumentar el rendimiento empresarial, pero migrar a la nube sin aprovechar su flexibilidad puede ser más dañino que quedarse en el centro de datos. Para que la migración salga bien, es necesario transformar los enfoques, los procesos y las tecnologías.

Se puede alcanzar un alto rendimiento del envío de software

Los equipos con mayor rendimiento destacan en las cuatro métricas de rendimiento del envío de software (plazo de entrega de los cambios, frecuencia de despliegue, tasa de errores al aplicar cambios y tiempo de recuperación tras un despliegue fallido), mientras que los que presentan el rendimiento más bajo ofrecen resultados deficientes en todas ellas. Hemos estudiado a equipos de cada vertical del sector en cada grupo de rendimiento.

Aplicar las estadísticas de DORA

Para que DORA propicie avances en un equipo y una empresa, es necesario evaluar el rendimiento actual, identificar las áreas en las que invertir y hacer mejoras, y contar con bucles de retroalimentación para determinar el progreso. Los equipos que adoptan una mentalidad y una práctica de mejora continua tienen más probabilidades de lograr mejores resultados. Invierte en reforzar la estructura de tu empresa de forma que puedas volver a hacerlo cada cierto tiempo.

Los hallazgos de nuestra investigación pueden proporcionarte información útil para tus propios experimentos e hipótesis. Es importante probar y medir el impacto de los cambios para ver qué funciona mejor en tu equipo y empresa. Así, te resultará más fácil validar nuestros hallazgos. Ten en cuenta que es probable que tus resultados difieran. No dudes en compartir tu progreso con nosotros para que aprendamos de tu experiencia.

Para la estrategia de mejora, recomendamos adoptar un enfoque experimental.

1. Identifica un área o resultado que quieras mejorar.
2. Mide tu referencia o estado actual.
3. Desarrolla un conjunto de hipótesis sobre qué podría acercarte al estado que buscas.
4. Acuerda un plan de mejora y comprométete a cumplirlo.
5. Haz el trabajo necesario.
6. Mide el progreso alcanzado.
7. Repite el proceso. El trabajo de mejora se consigue de manera incremental con repeticiones.

DORA

COMMUNITY



No puedes mejorar en solitario

Aprendemos de la experiencia de los demás. La comunidad DORA es un foro estupendo para compartir y conocer iniciativas de mejora: <https://dora.community>.

Rendimiento del envío de software

Los equipos tecnológicos siempre necesitan formas de medir el rendimiento para poder evaluar su situación actual, saber qué mejoras priorizar y validar su progreso. DORA ha validado de manera repetida cuatro métricas de envío de software (las cuatro claves) para medir de forma eficaz los resultados del proceso de envío de software.



Las cuatro claves

Las cuatro claves de DORA se han usado para medir la velocidad y la estabilidad de los cambios de software. Esto incluye todo tipo de cambios, incluidos aquellos relacionados con la configuración y los cambios de código.



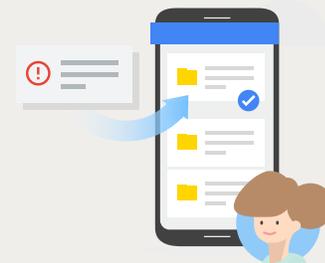
Plazo de entrega de los cambios: el tiempo que tarda cada confirmación de código o cambio en desplegarse correctamente en la producción.



Frecuencia de despliegue: asiduidad con la que se despliegan los cambios de aplicación en la producción.



Tasa de errores al aplicar cambios: el porcentaje de despliegues que provocan errores de producción¹ y requieren paquetes de revisiones o restauraciones.



Tiempo de recuperación tras un despliegue fallido: cuánto tarda en producirse la recuperación después de un despliegue fallido.

Hemos visto que estas métricas suelen estar relacionadas entre sí: los equipos con mejor rendimiento presentan buenos resultados en las cuatro, mientras que los que tienen un rendimiento bajo, presentan malos resultados.

Evolución de las métricas de rendimiento del envío de software

El análisis de las cuatro métricas clave ofrece un valor atípico desde hace tiempo: la tasa de errores al aplicar cambios.² La tasa de errores al aplicar cambios está estrechamente correlacionada con otras tres métricas, pero las pruebas y los métodos estadísticos nos impiden combinar las cuatro en un solo factor. La conexión mejoró al modificar la forma de respuesta a la pregunta sobre la tasa de errores al aplicar cambios, pero nos parecía que podría haber algo más.

Desde hace tiempo, tenemos la hipótesis de que la tasa de errores al aplicar cambios funciona como indicador de la cantidad de veces que un equipo tiene que reelaborar un trabajo. Cuando el envío falla, el equipo tiene que arreglarlo, lo que suele suponer otro cambio.

Para demostrar esta teoría, añadimos otra pregunta este año sobre la tasa de reelaboración de una aplicación: "Respecto a la aplicación o servicio principal en el que trabajas, ¿cuántos despliegues no planeados se realizaron aproximadamente para solucionar un error para los usuarios de la aplicación en los últimos seis meses?".

Nuestro análisis de datos confirmó nuestra hipótesis de que la tasa de reelaboración y la tasa de errores al aplicar cambios están relacionadas. Juntas, las dos métricas crean un factor fiable de estabilidad del envío de software.

Esto queda patente también en el análisis de los niveles de rendimiento de software. Este año, más de la mitad de los equipos de nuestro estudio presentan diferencias en la velocidad y la estabilidad del software. Estas diferencias nos llevaron a investigar el rendimiento del envío de software según dos factores:



A lo largo de este informe, utilizamos varias veces el concepto de rendimiento del envío de software y ambos factores en nuestro análisis. Para describir el rendimiento del envío de software se tienen en cuenta las cinco métricas.

Al describir la velocidad de envío de software, se usan el plazo de entrega de los cambios, la frecuencia de despliegue y el tiempo de recuperación tras un despliegue fallido. Este factor mide la velocidad al hacer modificaciones de cualquier tipo, tanto cambios normales como en respuesta a un fallo.

Al describir la estabilidad del envío de software, se usan la tasa de errores al aplicar cambios y la tasa de reelaboración. Este factor mide la posibilidad de que los despliegues conlleven trabajo adicional inmediato sin querer.



Niveles de rendimiento

Cada año, preguntamos a los encuestados acerca del rendimiento del envío de software de la aplicación o servicio principal en el que trabajan. Estudiamos sus respuestas con un análisis de clústeres, un método estadístico que identifica respuestas similares entre sí, pero distintas a las de otros grupos de respuestas.

Realizamos un análisis de clústeres en las cuatro métricas de envío de software originales para mantener la coherencia con los análisis de clústeres de años anteriores.

En nuestro análisis del rendimiento del envío de software, aparecieron cuatro clústeres de respuestas. No establecemos estos niveles con antelación, sino que surgen a partir de las respuestas de la encuesta. Este proceso nos permite conocer la situación del rendimiento del envío de software de todos los encuestados cada año.

Los datos de este año nos dieron cuatro clústeres distintos, como se muestra más abajo.

Nivel de rendimiento	Plazo de entrega de los cambios	Frecuencia de despliegue	Tasa de errores al aplicar cambios	Tiempo de recuperación tras un despliegue fallido	Porcentaje de encuestados*
Élite	Menos de un día	Bajo demanda (varios despliegues al día)	5 %	Menos de una hora	19 % (18-20 %)
Alto	Entre un día y una semana	Entre una vez al día y una vez a la semana	20 %	Menos de un día	22 % (21-23 %)
Medio	Entre una semana y un mes	Entre una vez a la semana y una vez al mes	10 %	Menos de un día	35 % (33-36 %)
Bajo	Entre uno y seis meses	Entre una vez al mes y una vez cada seis meses	40 %	Entre una semana y un mes	25 % (23-26 %)

*Intervalo de incertidumbre del 89 %

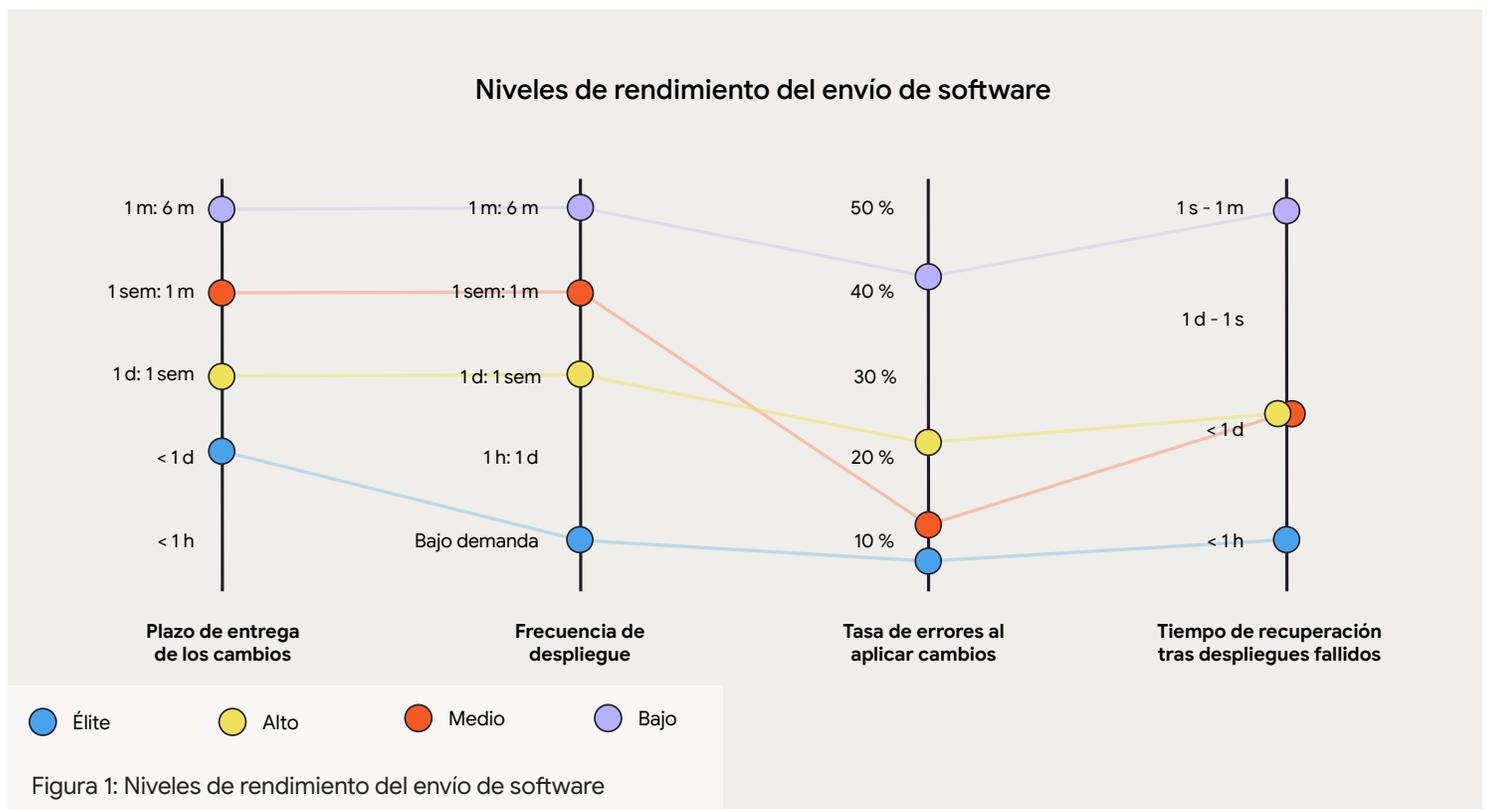
¿Velocidad o estabilidad?

En los cuatro clústeres, la velocidad y la estabilidad están correlacionadas. Esta correlación se mantiene incluso en el clúster de rendimiento medio (naranja), donde la velocidad es más baja y la estabilidad es más alta que en el clúster de alto rendimiento (amarillo). Esto sugiere que hay otros factores que influyen en el rendimiento del clúster, además de la velocidad y la estabilidad. Por ejemplo, en el clúster de rendimiento medio se pueden ver cambios más frecuentes en el envío.

¿Qué es mejor: despliegues más frecuentes o con menos fallos?

Quizás no haya una respuesta universal a esta pregunta, ya que depende de la aplicación o el servicio que se analice, de los objetivos del equipo que trabaja en esa aplicación y, sobre todo, de las expectativas de los usuarios.

Hemos tomado la decisión de llamar "equipos de alto rendimiento" a los más rápidos y "equipos de rendimiento medio" a los que son más lentos, pero más estables. Esta decisión destaca uno de los posibles inconvenientes de usar estos niveles de rendimiento: para un equipo, debería ser más importante mejorar que alcanzar un determinado nivel de rendimiento. Los mejores equipos son los que consiguen una **mejora** de élite, pero no necesariamente un **rendimiento** de élite.



Si los comparamos con los equipos de bajo rendimiento, los equipos de rendimiento de élite presentan las siguientes cifras:

127
veces más
rápidos

en el plazo
de entrega

182
veces más

despliegues
al año

8 veces
menos

errores al aplicar
cambios

2293
veces más
rapidez

de recuperación
tras un
despliegue
fallido

Cómo usar los clústeres de rendimiento

Los clústeres de rendimiento proporcionan datos de referencia que muestran el rendimiento del envío de software de los encuestados de este año. Se han diseñado como inspiración para demostrar que el rendimiento de élite es factible.

Creemos que es más importante que los equipos se centren en mejorar su rendimiento general, más que en alcanzar un nivel de rendimiento concreto. Los mejores equipos son los que consiguen una **mejora** de élite, pero no necesariamente un **rendimiento** de élite.

El sector no afecta a los niveles de rendimiento de manera significativa

Según nuestra investigación, el sector casi nunca³ permite predecir el rendimiento del envío de software, y hay equipos de alto rendimiento en cada vertical. Esto no significa que cada sector no se tope con obstáculos específicos, sino que ningún sector en particular parece ser mejor o peor en cuanto al rendimiento del envío de software.

Uso de las métricas de rendimiento del envío de software

Cada aplicación o servicio tiene su propio contexto único. Esta complejidad hace que sea muy difícil predecir cómo puede afectar un cambio al rendimiento general del sistema. Además, es prácticamente imposible hacer cambios en una empresa de uno en uno. Si tenemos en cuenta esta complejidad, ¿cómo se pueden usar las métricas de rendimiento del envío de software como guía para nuestras iniciativas de mejora?

Para empezar, identifica la aplicación o servicio principal que quieras medir y mejorar. A continuación, recomendamos que el equipo multifuncional responsable de la aplicación se reúna para medir y establecer su rendimiento del envío de software en ese momento. La encuesta breve de DORA (<https://dora.dev/quickcheck>) puede servir como guía en la reunión y ayudar a fijar ese valor de referencia. Tu equipo también tendrá que entender qué está impidiendo que el rendimiento sea mayor.

Una manera eficaz de detectar estos impedimentos es realizar un ejercicio de asignación de flujo de valor⁴ en equipo.

Después, será el momento de identificar y acordar un plan de mejora, que puede centrarse en mejorar alguna de las competencias identificadas por DORA⁵ o cualquier otro aspecto específico de tu aplicación o empresa.

Ya que tenemos este plan a mano, es hora de ponerse a trabajar. Dedicar los recursos necesarios a esta estrategia de mejora y tener en cuenta todo lo que vayas aprendiendo durante el proceso.

Cuando el cambio ya se haya implementado y afianzado, será el momento de volver a evaluar las cuatro claves. ¿Cómo han evolucionado desde que el equipo implementó el cambio? ¿Qué has aprendido en este tiempo?

Repetir este proceso permite integrar en el equipo una práctica de mejora continua.

Recuerda que el cambio no se produce de la noche a la mañana. Hay que adoptar un enfoque iterativo que fomente la creación de un entorno propicio para el aprendizaje, la fluidez y las sugerencias rápidas⁶.

-
1. Solo consideramos que un despliegue constituye un error de aplicación de cambios si provoca algún fallo tras llegar a la fase de producción, donde puede afectar a los usuarios finales. Por el contrario, que un cambio se detenga antes de llegar a producción demuestra la capacidad del proceso de despliegue para detectar errores.
 2. Forsgren, Nicole, Jez Humble y Gene Kim. 2018. *Accelerate: The Science Behind DevOps: Building and Scaling High Performing Technology Organizations*. IT Revolution Press. págs. 37-38
 3. Según el informe *Accelerate State of DevOps del 2019* (pág. 32), el comercio minorista tuvo un rendimiento del envío de software significativamente superior: <https://dora.dev/research/2019/dora-report/2019-dora-accelerate-state-of-devops-report.pdf#page=32>
 4. <https://dora.dev/guides/value-stream-management/>
 5. <https://dora.dev/capabilities>
 6. <https://dora.dev/research>

Inteligencia artificial: adopción y actitudes



Datos clave

La gran mayoría de las empresas e de todos los sectores que participaron en la encuesta están cambiando sus prioridades para incorporar mejor la IA en sus aplicaciones y servicios. Del mismo modo, la gran mayoría de los profesionales del desarrollo confían en la IA para cumplir las responsabilidades básicas de su puesto y afirman haber experimentado mejoras en la productividad como resultado. Hay una percepción generalizada entre los desarrolladores de que es necesario usar la IA para seguir siendo competitivos en el mercado actual, y parece ser un factor importante a la hora de adoptar la IA, tanto para las empresas e como para los profesionales del desarrollo de manera individual.

Introducción

No es fácil ignorar el gran impacto que ha tenido la IA en el trabajo de desarrollo, ya que no paran de publicarse artículos periodísticos en medios populares que describen sus efectos, tanto buenos¹ como malos² o peores.³ En nuestro informe Accelerate State of DevOps del 2023⁴ solo hablamos de la IA como una de las muchas competencias técnicas que afectan al rendimiento, así que este año hemos querido indagar en el tema de una forma más exhaustiva.

Como el uso de la IA en el trabajo de desarrollo profesional está cambiando rápidamente desde ser algo marginal a estar omnipresente, creemos que nuestro informe del 2024 es una importante oportunidad para evaluar la adopción, el uso y las actitudes de los profesionales del desarrollo en un punto de inflexión fundamental para el sector.

Resultados

Adopción de la inteligencia artificial

Las respuestas sobre la adopción de la IA parecen indicar que cada vez se reconoce más que la IA ya no es algo que se vislumbra en el horizonte, sino que ya está aquí y, casi con total seguridad, ha venido para quedarse.

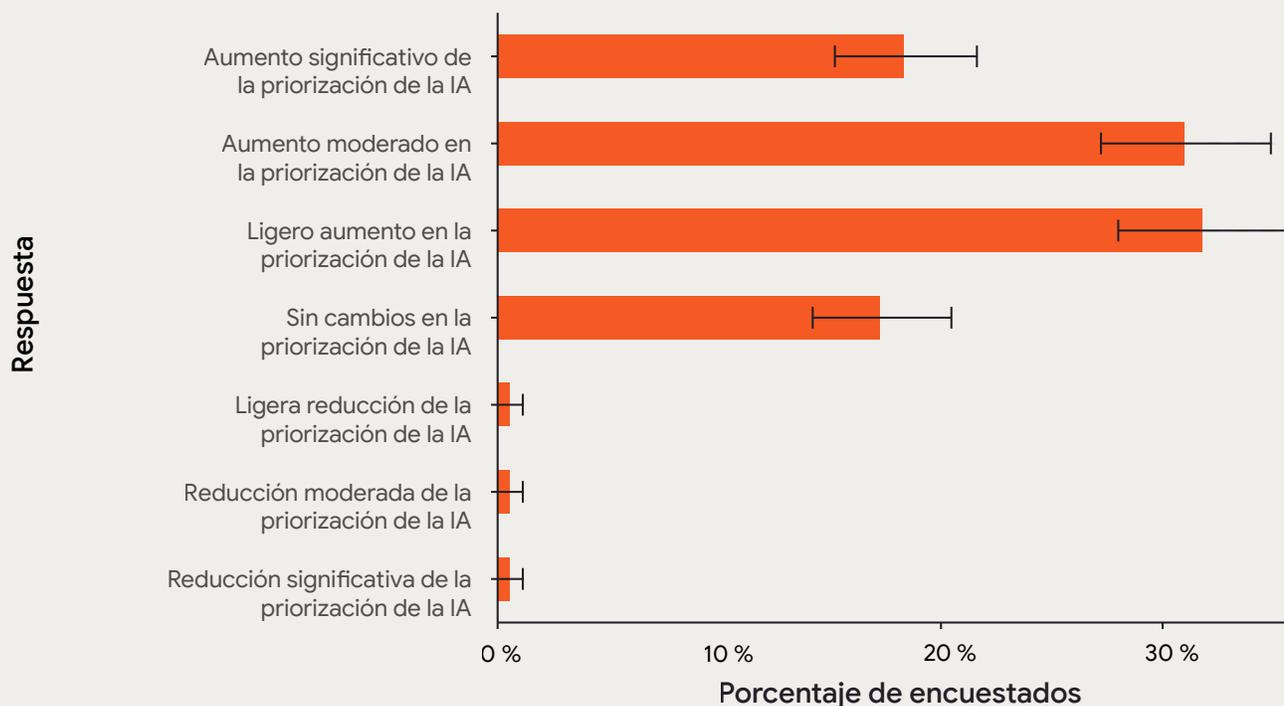
Adopción empresarial de la inteligencia artificial

La gran mayoría de los encuestados (81 %) afirmaron que sus empresas habían cambiado sus prioridades para incrementar

la incorporación de la IA en sus aplicaciones y servicios. El 49,2 % de los encuestados incluso describieron la magnitud del cambio como algo "moderado" o "significativo".

Cabe destacar que el 3 % de los encuestados indicaron que sus empresas ya no se estaban centrando tanto en la IA (dentro del margen de error de nuestra encuesta). El 78 % de los encuestados señalaron que confiaban en la transparencia de sus empresas respecto a cómo planean usar la IA como resultado de esos cambios en las prioridades. Estos datos se muestran en la figura 2.

Cambios en las prioridades empresariales relacionados con la IA



La barra de error representa un intervalo de incertidumbre del 89 %

Figura 2: Percepciones de los encuestados sobre los cambios de prioridades en sus empresas en cuanto al grado de incorporación de la IA en sus aplicaciones o servicios.

Participantes de todos los sectores representados en la encuesta indicaron niveles de confianza de la IA en su trabajo diario estadísticamente idénticos, lo que sugiere que esta rápida adopción de la IA se está produciendo de manera uniforme en todos los sectores, lo cual nos sorprendió un poco. Cada sector puede presentar amplias diferencias según la rigidez de las restricciones normativas y el ritmo histórico de innovación, factores que pueden afectar a las tasas de adopción de tecnología.

Sin embargo, vimos que los encuestados que trabajan en grandes empresas indicaron una menor confianza en la IA en su trabajo diario que los que pertenecen a empresas más pequeñas, algo que concuerda con estudios anteriores que indicaban que las empresas más grandes se estaban adaptando más lentamente al cambio tecnológico debido a que las complejidades organizativas y los costes de coordinación eran mayores.⁵



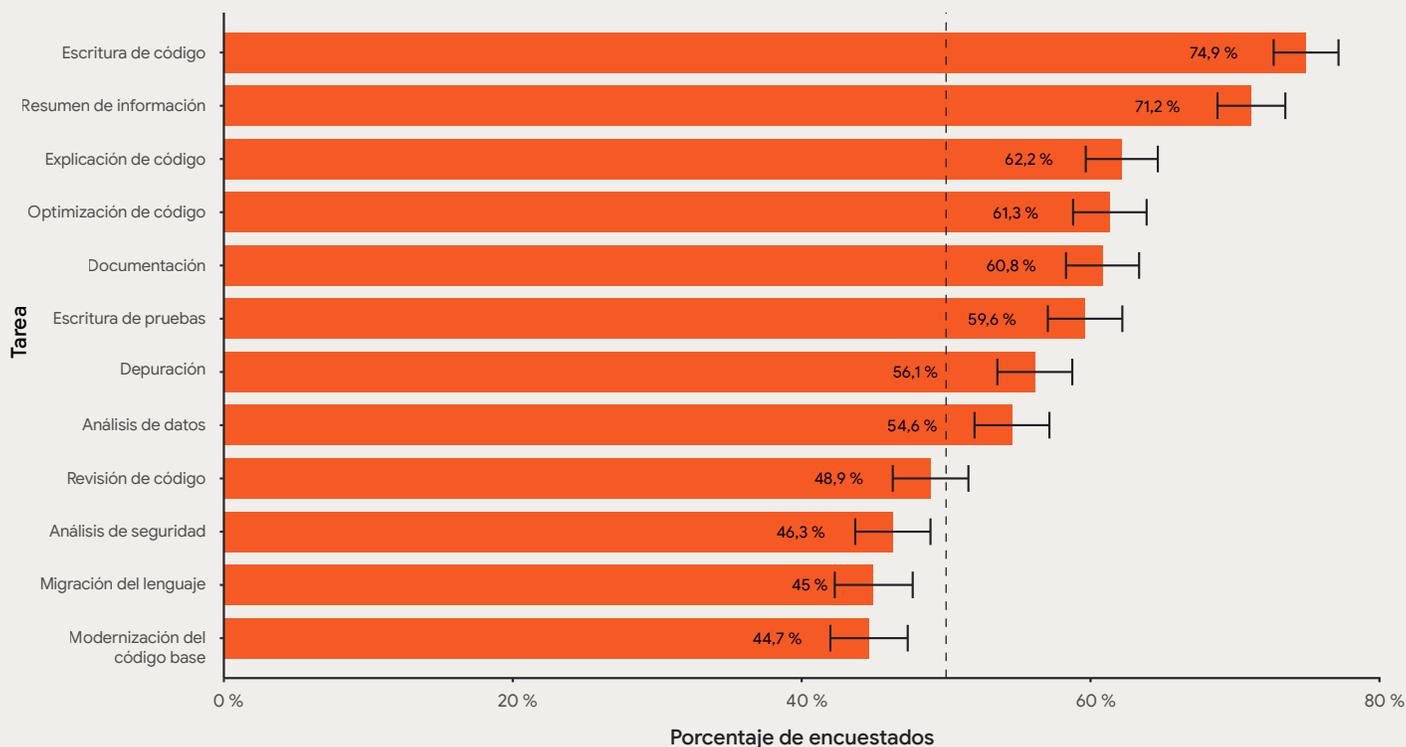
Adopción individual de la inteligencia artificial

A nivel individual, hemos visto que el 75,9 % de los encuestados confían en la IA, al menos en parte, para cumplir una o varias de sus responsabilidades profesionales diarias. Entre estas responsabilidades laborales se incluyen las siguientes tareas, para las que la mayoría de los encuestados afirmaron confiar en la IA:

1. Escribir código
2. Resumir información
3. Explicar código desconocido
4. Optimizar código
5. Documentar código
6. Escribir pruebas
7. Depurar código
8. Analizar datos

De todas las tareas incluidas en las respuestas de nuestra encuesta, los casos prácticos más habituales en los que se usaba la IA en desarrollo de software fueron escribir código y resumir información: un 74,9 % y un 71,2 de los encuestados cuyas responsabilidades incluían estas tareas afirmaron confiar en la IA para hacerlas (al menos, en parte). Estos datos se muestran en la figura 3.

Confianza en la IA para hacer tareas



La barra de error representa un intervalo de credibilidad del 89 %

Figura 3: Porcentaje de los encuestados que confían en la IA, al menos en parte, para realizar doce tareas de desarrollo habituales

Los bots de chat fueron la interfaz más común a través de la que los encuestados interactuaron con la IA en su trabajo diario (78,2 %), seguidos de las interfaces web externas (73,9 %) y las herramientas de IA incrustadas en sus IDEs (72,9 %). Era menos habitual que los encuestados usaran la IA a través de interfaces web internas (58,1 %) y como parte de flujos de procesamiento de CI/CD automatizados (50,2 %).

Sin embargo, somos conscientes de que conocer la IA que se usa en los flujos de procesamiento de CI/CD y en las plataformas internas de los encuestados probablemente dependa de la frecuencia

con la que interactúan con estas tecnologías. Por ello, es probable que las cifras sean un poco más bajas que las reales.

Hemos visto que los científicos de datos y los especialistas en aprendizaje automático, por lo general, confiaban más en la IA que los encuestados con otros puestos. Sin embargo, el uso de la IA de los ingenieros de hardware suele ser menor que el del resto de los encuestados, algo que quizás se deba a que sus responsabilidades no suelen incluir las tareas para las que se suele usar IA que hemos mencionado arriba.



Factores que impulsan la adopción de la inteligencia artificial

Los entrevistados a menudo vinculaban la decisión de adoptar la IA a las presiones de la competencia y a la necesidad, tanto de las empresas como de los desarrolladores, de mantenerse al día con los estándares del sector, en los que cada vez se incluyen más las competencias en IA.

Para varias empresas en las que trabajaban los participantes, usar la IA se consideraba "una gran cuestión de marketing" (P3)⁶ que servía como rasgo diferenciador frente a la competencia. Saber que sus competidores estaban empezando a adoptar la IA en sus propios procesos hizo que incluso una empresa renunciara a aplicar la "enorme burocracia" normalmente

relacionada con la adopción de nuevas tecnologías por la urgencia que sentían por adoptar la IA, preguntándose "¿qué pasaría si nuestros competidores se nos adelantaran?" (P11).

A nivel individual, muchos participantes relacionaron la adopción de la IA con la sensación de que saber usar la IA para el desarrollo de software es "prácticamente el nuevo requisito básico para acceder a puestos de ingeniero" (P9). Varios participantes sugirieron que otros desarrolladores deberían adoptar rápidamente la IA en su flujo de trabajo de desarrollo porque "están pasando tantas cosas en esta área que es difícil llevar el ritmo... Creo que, si no la usas, te quedas atrás muy rápido" (P4).

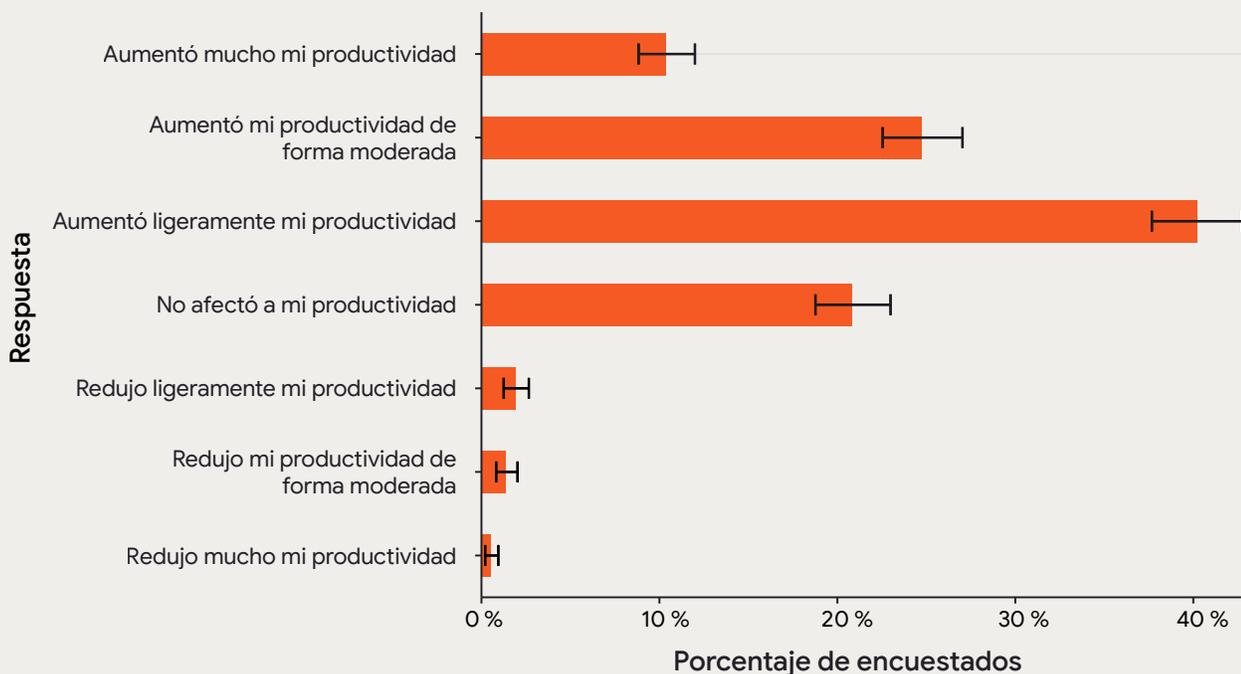
Percepciones sobre la inteligencia artificial

Mejoras de rendimiento gracias a la inteligencia artificial

Para la mayor parte de las empresas y los desarrolladores que están adoptando la IA, usarla en el trabajo de desarrollo parece tener muchas ventajas. El 75 % de los encuestados comunicaron importantes mejoras en la productividad derivadas del uso de la IA en los tres meses anteriores a nuestra encuesta, que se realizó a principios del 2024.

En particular, más de un tercio de los encuestados describieron las mejoras experimentadas en la productividad de magnitud moderada (25 %) o extrema (10 %). Menos del 10 % de los encuestados afirmaron haber experimentado impactos negativos en su productividad (ni siquiera leves) debido a la IA. Estos datos se muestran en la figura 4.

Percepciones de los cambios en la productividad por la IA



La barra de error representa un intervalo de incertidumbre del 89 %

Figura 4: Percepciones de los encuestados respecto al impacto de la IA en su productividad

Si nos centramos en puestos concretos, los encuestados que indicaron las mayores mejoras en la productividad debido a la IA fueron los profesionales de la seguridad, los administradores de sistemas y los desarrolladores de full stack. Aunque los desarrolladores de aplicaciones móviles, los ingenieros de Site Reliability Engineering y los gestores de proyectos también señalaron mejoras en la productividad, las cuantificaron en magnitudes más bajas que los encuestados en los otros puestos mencionados.

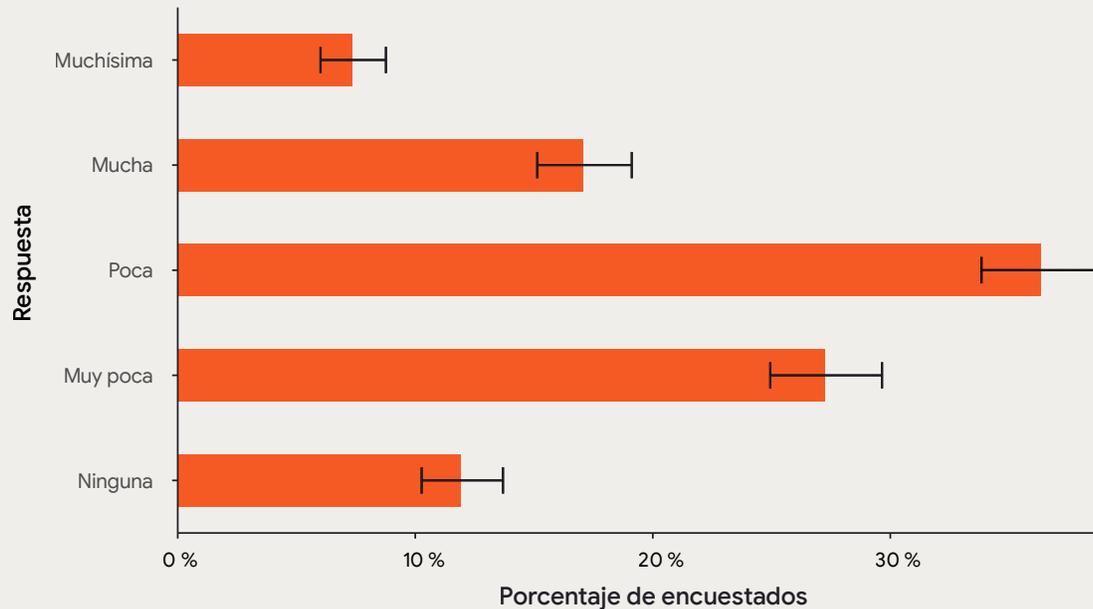
Aunque sospechábamos que el carácter novedoso de la IA en el trabajo de desarrollo, y su correspondiente curva de aprendizaje, podría afectar negativamente a la capacidad de los desarrolladores para escribir código, nuestros hallazgos no confirmaron esa hipótesis. Solo el 5 % de los encuestados señalaron que la IA había obstaculizado en cierto grado su capacidad para escribir código. De hecho, el 67 % de los encuestados indicaron al menos alguna mejora en su capacidad para escribir código como resultado de las herramientas de programación asistidas por IA, y alrededor del 10 % vieron mejoras "extremas" gracias a la IA.

Confianza en el código generado por IA

Las percepciones de los participantes respecto a la fiabilidad del código generado por IA usado en el trabajo de desarrollo son complejas. Aunque la gran mayoría de los participantes (87,9 %) señalaron cierto nivel de confianza en la calidad del código generado por IA, el grado de confianza en su calidad era por lo general bajo, ya que el 39,2 % indicó tener poca (27,3 %) o ninguna confianza (11,9 %). Estos datos se muestran en la figura 5.



Confianza en la calidad del código generado por IA



La barra de error representa un intervalo de incertidumbre del 89 %

Figura 5: Confianza indicada por los encuestados en la calidad del código generado por IA.

Dado que la encuesta demuestra que los desarrolladores están adoptando rápidamente la IA, que la usan y que la perciben como un elemento que contribuye de manera positiva al rendimiento, nos ha sorprendido la falta de confianza general en ella. Merece la pena señalar que, durante nuestras entrevistas, muchos participantes explicaron que estaban dispuestos a modificar los resultados del código generado por IA usado en su trabajo, o que esperaban tener que hacerlo.

Un participante incluso comparó la necesidad de evaluar y modificar los resultados del código generado por IA con "los comienzos de StackOverflow, [cuando] siempre dabas por hecho que los usuarios de StackOverflow tenían mucha experiencia y sabían exactamente qué hacer, y entonces solo copiabas y pegabas, y explotaba todo" (P2).

Al no tratarse de un problema nuevo, es posible que participantes como P3 piensen que las empresas no están "preocupadas porque alguien haga cosas como copiar y pegar código de Copilot o ChatGPT [porque] hay muchas capas para comprobarlo" con los procesos de control de calidad del código que tienen.

Nuestra hipótesis es que los desarrolladores no esperan necesariamente tener una confianza absoluta en la precisión del código generado por IA ni parece que esta confianza absoluta sea fundamental para que les resulte útil usarlo. Más bien parece que el código generado por IA, que casi siempre es correcto y puede perfeccionarse con un par de cambios, resulta aceptable, suficientemente valioso para motivar su adopción y uso generalizados, y compatible con los procesos de control de calidad actuales.

Expectativas sobre el futuro de la IA

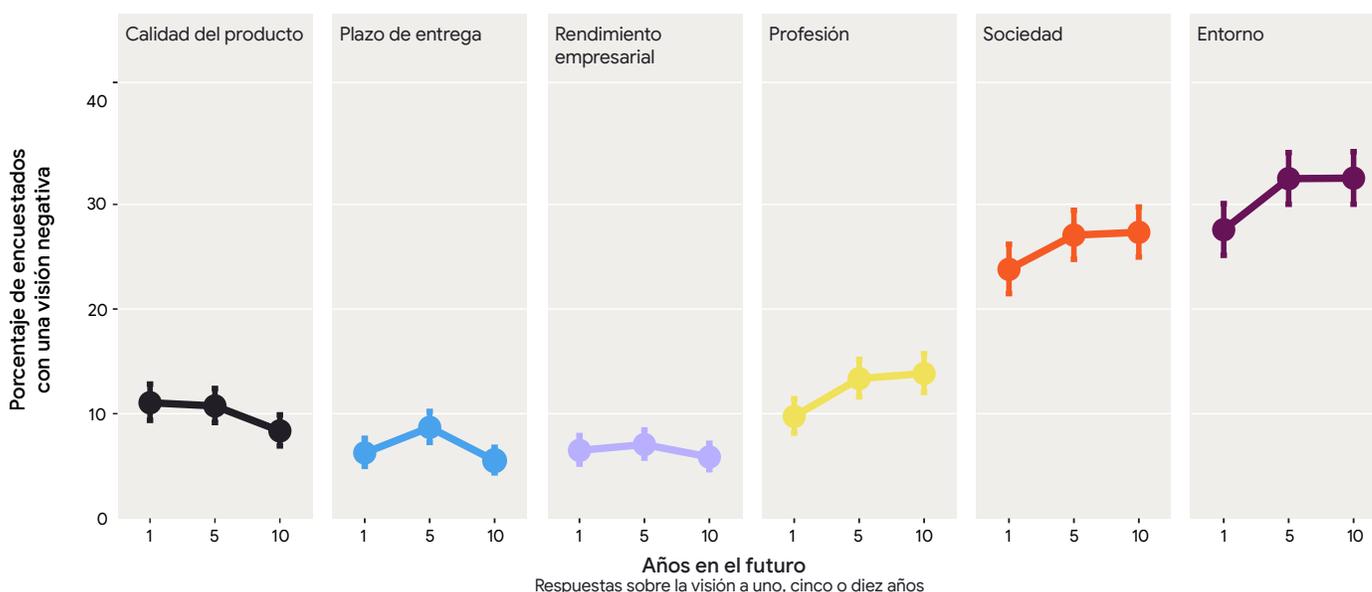
En general, nuestros hallazgos indican que la IA ya ha tenido un gran impacto en el trabajo de los profesionales del desarrollo, una tendencia que creemos que seguirá creciendo. Aunque sería imposible predecir exactamente cómo afectará la IA al desarrollo (y al mundo) en el futuro, pedimos a los encuestados que especularan y compartieran sus expectativas sobre el impacto de la IA en un año, en cinco y en diez.

Aunque los encuestados indicaron un impacto bastante positivo de la IA en su trabajo de desarrollo reciente, no tenían tantas esperanzas en relación con sus efectos en el futuro.

De manera optimista, y coherente con nuestros hallazgos sobre la influencia positiva de la IA en el rendimiento de los profesionales, los encuestados dijeron que esperaban que la calidad de sus productos siguiera mejorando como resultado del uso de la IA en uno, cinco y diez años.

Sin embargo, también señalaron que esperaban que la IA tuviera efectos negativos en su carrera profesional, el medioambiente y la sociedad en su conjunto, y que creían que esos efectos negativos se materializarían por completo en unos cinco años. Estos datos se muestran en la figura 6.

Impactos negativos esperados de la IA



La barra de error representa un intervalo de credibilidad del 89 %

Figura 6: Expectativas de los encuestados sobre los futuros efectos negativos de la IA en uno, cinco y diez años.

Los entrevistados tenían los mismos sentimientos encontrados que los encuestados en cuanto al impacto de la IA en el futuro. Algunos se preguntaban sobre las acciones legales venideras en un panorama normativo aún incierto, y les preocupaba estar "en el lado malo, si se toman decisiones" (P3).

Otros mostraron ansiedad y se preguntaban "¿Va a sustituir a las personas? ¿Quién sabe? Quizás" (P2), mientras que sus compañeros acallaron sus miedos al establecer paralelismos con el pasado, cuando "decían que el efecto 2000 iba a provocar el caos, etc., porque era lo nuevo en aquel momento,

[pero] no se sustituyó a nadie. De hecho, se crearon más puestos de trabajo. Creo que pasará lo mismo con la IA" (P1).

Los efectos que tendrá la IA en el mundo en el futuro siguen sin estar claros, pero, este año, nuestra encuesta indica con firmeza que la IA ha provocado un cambio de paradigma en el desarrollo de software que no puede pasarse por alto. Hasta ahora, los profesionales del desarrollo han recibido bien estos cambios.



1. <https://www.sciencedaily.com/releases/2024/03/240306144729.htm>

2. <https://tech.co/news/list-ai-failures-mistakes-errors>

3. <https://klyker.com/absurd-yoga-poses-generated-by-ai/>

4. <https://dora.dev/dora-report-2023>

5. Rogers, Everett M., Arvind Singhal y Margaret M. Quinlan. "Diffusion of innovations". An integrated approach to communication theory and research. Routledge, 2014. 432-44, Tornatzky, L. G. y Fleischer, M. (1990). The processes of technological innovation. Lexington, MA: Lexington Books

6. (P[N]), por ejemplo (P1), es el seudónimo de cada entrevistado.

Análisis del impacto secundario de la IA



Datos clave

En este capítulo, investigamos el impacto de la adopción de la IA en general, desde los desarrolladores individuales hasta empresas completas. Los hallazgos revelan una imagen compleja con ventajas claras e inconvenientes inesperados. Aunque la adopción de la IA refuerza la productividad individual, el flujo y la satisfacción laboral, también puede reducir el tiempo dedicado al trabajo de más valor.

De igual forma, la IA afecta de manera positiva a la calidad, la documentación y los procesos de revisión del código, pero, sorprendentemente, estos beneficios no se traducen en una mejora en el rendimiento del envío de software. De hecho, la adopción de la IA parece perjudicial en este ámbito, mientras que su efecto en el rendimiento de los productos sigue siendo insignificante.

A pesar de estos obstáculos, la adopción de la IA se relaciona con la mejora en el rendimiento de los equipos y empresas. Para cerrar este capítulo, hacemos un llamamiento a evaluar de forma crítica el papel de la IA en el desarrollo de software y adaptar de manera proactiva su aplicación para aprovechar al máximo las ventajas y mitigar las consecuencias inesperadas.

La situación de la IA y DORA

Se estima que los principales gigantes tecnológicos invertirán aproximadamente un billón de USD en el desarrollo de la IA en los próximos cinco años.¹ Esta estimación está en sintonía con la estadística que presentamos en el capítulo "Inteligencia artificial: adopción y actitudes" que muestra que el 81 % de los encuestados afirman que su empresa ha cambiado sus recursos para dedicarlos al desarrollo de la IA.

Los impactos de la IA en el medioambiente representan el resto de los costes. También hay estimaciones que sugieren que, para el 2030, la IA supondrá un incremento del 160 % en la demanda de energía de los centros de datos.² El entrenamiento de un modelo de IA puede alcanzar prácticamente "el consumo de electricidad anual de más de 1000 hogares de Estados Unidos".³ Por tanto, no es de extrañar que más del 30 % de los encuestados piensen que la IA va a afectar negativamente al medioambiente.

Más allá de los costes de desarrollo y medioambientales, también hay que considerar los posibles costes de adopción.

Podrían generarse por muchos motivos, desde el descenso de la productividad hasta la contratación de especialistas. Estos costes de adopción también se producirán en la sociedad, ya que más de un tercio de los encuestados creen que la IA tendrá perjuicios para la sociedad en la próxima década. Teniendo en cuenta estos costes, parece natural que exista una profunda curiosidad por saber qué se obtendrá a cambio.

Esta curiosidad se ha manifestado en una gran variedad de medios de comunicación, artículos y estudios en los que se mezclan opiniones y datos, al menos en cierta medida.



Algunos creen que la IA ha mejorado significativamente las capacidades de la humanidad,⁴ otros sugieren que es poco más que una herramienta útil para hacer tareas⁵ y otros tantos piensan que será la perdición para la humanidad.⁶

Los indicios sobre próximos resultados, como la capacidad de realizar por completo y correctamente una tarea concreta, son muy positivos.⁷ Sin embargo, cuando se trata de resultados más lejanos, como el código base de un equipo, no son tan claros ni tan positivos. Por ejemplo, algunas investigaciones sugieren que la tasa de modificación de código se ha duplicado en comparación con el valor de referencia de antes del 2021.⁸

No nos sorprende lo difícil que resulta entender estos efectos secundarios. Cuanto más alejado está el efecto de la causa, menos pronunciada y evidente es la conexión.

Evaluar los efectos secundarios de la IA es similar a cuantificar el efecto de una roca que se tira a un lago. Podemos atribuir fácilmente casi todas las ondas cercanas al punto de impacto de la roca en el agua, pero, cuanto más nos alejamos de él, menos marcado es el efecto de la roca y más difícil resulta relacionar las olas con su impacto.

La IA es, básicamente, una roca lanzada a un mar embravecido, lleno de otros procesos y dinámicas. Entender hasta qué punto las olas están causadas por la IA (o cualquier tecnología o práctica) es todo un reto. Quizás esto explique, en parte, por qué ha costado tanto que el sector adopte un conjunto de marcos de análisis y métricas basado en principios para entender el impacto de la IA.⁹

Nuestro enfoque está creado específicamente para resultar útil en este tipo de desafíos. DORA se ha diseñado para entender la utilidad o inutilidad de una práctica. Durante los últimos 10 años, hemos investigado los efectos secundarios de numerosas prácticas, como prácticas de seguridad, liderazgo transformacional, culturas generativas, documentación, integración y entrega continuas, y enfoques orientados al usuario.¹⁰

Creemos que el enfoque de DORA¹¹ puede ayudarnos a saber más sobre el impacto de la IA, especialmente al investigar sus efectos en diferentes resultados.



Medir la adopción de la IA

El primer reto para conocer el impacto de la adopción de la IA es medirla. Determinamos que medir la frecuencia de uso probablemente no sea tan significativo como medir la confianza para entender la centralización de la IA en los flujos de trabajo de desarrollo. Una persona puede revisar código o redactar documentos varias veces al mes o cada dos meses y aun así considerar esas tareas como una parte esencial de su trabajo.

Al mismo tiempo, que alguien use la IA con frecuencia no significa que lo haga para trabajos que considere importantes o fundamentales para su puesto.

Teniendo todo esto en cuenta, preguntamos a los encuestados por su uso de la IA tanto en general como para hacer tareas concretas. En el [capítulo anterior](#) se detallan los resultados de la encuesta y su interpretación.

Mediante el análisis factorial, descubrimos que el elemento de la encuesta de confianza "general" en la IA se solapaba en gran medida con la confianza en la IA indicada para las siguientes tareas:

- Escritura de código
- Resumen de información
- Explicación de código
- Optimización de código
- Documentación
- Escritura de pruebas

La alta coincidencia y covarianza entre estos siete elementos sugiere que existe un factor subyacente, al que llamamos "adopción de IA".

El impacto de la IA en las personas conforma una historia de claras ventajas (y algunos posibles perjuicios)

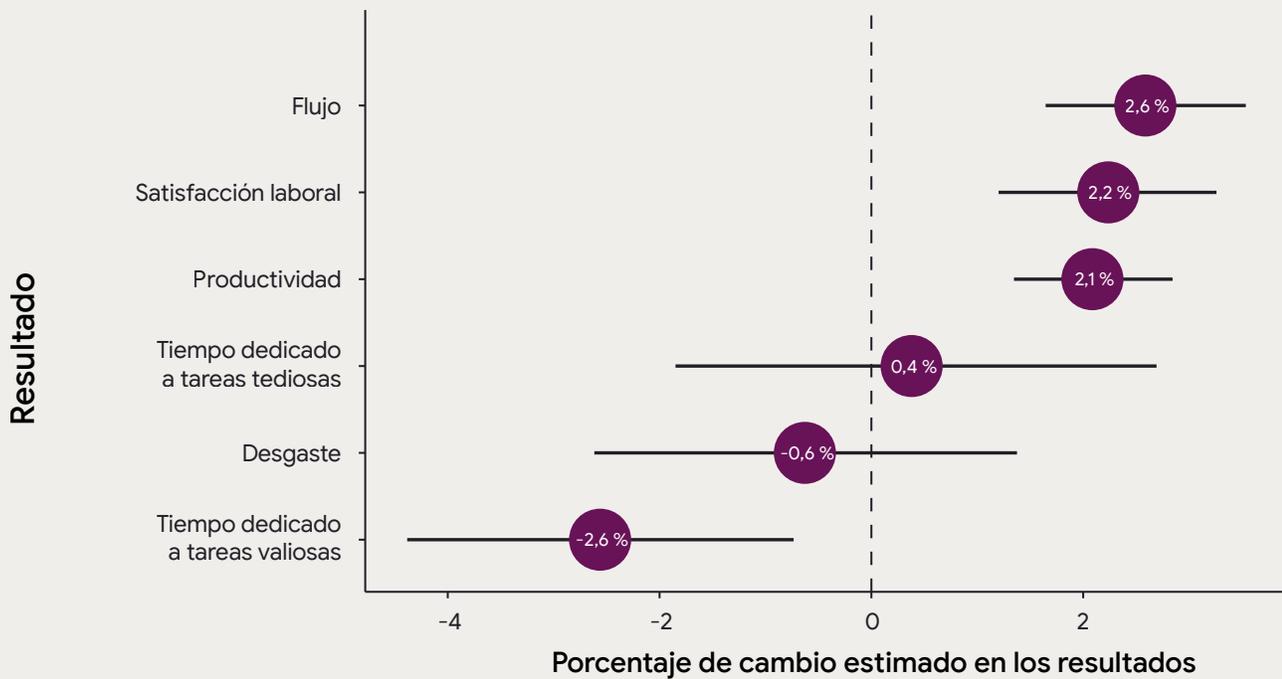
Como hacemos cada año, medimos una gran variedad de constructos relacionados con el éxito y el bienestar de una persona:

Satisfacción laboral	Elemento único diseñado para establecer la opinión general de alguien sobre su trabajo.
Desgaste	Factor que engloba la naturaleza multifacética del desgaste, donde se incluyen las dimensiones física, emocional y psicológica, así como su impacto en la vida personal.
Flujo	Elemento único diseñado para establecer el grado de concentración que suele tener una persona durante las tareas de desarrollo.
Productividad	Puntuación factorial diseñada para medir el grado en el que una persona se siente eficaz y eficiente en su trabajo para crear valor y cumplir sus tareas.
Tiempo dedicado a tareas tediosas	Elemento único que mide el porcentaje de tiempo que dedica una persona a tareas manuales repetitivas que ofrecen muy poco valor a largo plazo.
Tiempo dedicado a tareas valiosas	Elemento único que mide el porcentaje de tiempo que dedica una persona a tareas que considera que tienen valor.

Queríamos averiguar si la forma de responder a estas preguntas cambiaba en relación con la adopción de la IA, y los resultados sugieren que es probable que sea así.

En la figura 7 se muestran nuestras mejores estimaciones sobre el impacto en el éxito y bienestar de una persona que tendría adoptar la IA.

Si una persona aumenta la adopción de la IA en un 25 %...



Punto = valor estimado

Barra de error = intervalo de incertidumbre del 89 %

Figura 7: Impactos de la adopción de la IA en el éxito y el bienestar individuales

Las ventajas claras

La historia de las ventajas de adoptar la IA para los usuarios individuales es bastante positiva, pero, como cualquier buena historia, no está totalmente exenta de contratiempos.

Lo que parece evidente es que la IA tiene un impacto significativo y beneficioso en el flujo, la productividad y la satisfacción laboral (consulta la figura 7).

Por ejemplo, es probable que la productividad aumente aproximadamente un 2,1% cuando la adopción de la IA de un usuario individual se incrementa en un 25% (consulta la figura 7). Puede parecer poco, pero hay que tener en cuenta que estamos hablando a nivel individual e imaginar este patrón en cientos o incluso miles de desarrolladores.

Este patrón es el que esperábamos. Creemos que se basa, en parte, en la capacidad de la IA para sintetizar fuentes de información dispares y dar una respuesta altamente personalizada en una sola ubicación. Hacer algo así sin ayuda conlleva tiempo y muchos cambios de contexto, además de ser menos probable que fomente el flujo.

Si tenemos en cuenta la alta conexión que tienen la productividad y el flujo con la satisfacción laboral, no debería sorprendernos ver que la adopción de la IA redunde en una mayor satisfacción laboral.

Los posibles contras

Aquí es donde se complica un poco la historia. Una propuesta de valor para adoptar la IA es que ayudará a los usuarios a pasar más tiempo haciendo tareas valiosas. Esto significa que, al automatizar las tareas tediosas, manuales y repetitivas, esperamos que los encuestados tengan más tiempo libre para encargarse de "algo mejor". Sin embargo, nuestros datos sugieren que la adopción de la IA puede tener el efecto contrario (reducir el tiempo dedicado a hacer tareas valiosas), y parece que el tiempo dedicado al trabajo tedioso no se ve afectado.

Los marcadores correspondientes al bienestar de los encuestados, como el flujo, la satisfacción laboral y la productividad, se han asociado históricamente al tiempo dedicado a hacer tareas valiosas. Por eso, es sorprendente ver los aumentos en estos indicadores a pesar de haber reducido el tiempo dedicado a esas tareas.

Si queremos explicar bien estos patrones, habrá que abordar esta aparente incongruencia. Una buena explicación de una película no puede obviar una escena que la contradice, ni una buena explicación de un libro puede pasar por alto un capítulo que no concuerda del todo. De la misma forma, una buena explicación de estos patrones no puede centrarse únicamente en el subconjunto de patrones que nos permite contar una historia sencilla.

Hay innumerables hipótesis que podrían explicar los datos, pero se nos ocurrió una que parece disculpar que el flujo, la productividad y la satisfacción laboral se beneficien del uso de la IA mientras que el tiempo dedicado a tareas valiosas disminuye y las tareas tediosas no presentan cambios.

A esta hipótesis la llamamos "la hipótesis del vacío". Al aumentar la productividad y el flujo, la IA ayuda a los usuarios a trabajar de manera más eficiente, y esa eficiencia les permite acabar más rápido el trabajo que consideran más valioso.

Entonces, se crea el vacío; hay tiempo extra. La IA no roba valor del trabajo de los encuestados, sino que lo agiliza.



Pero, ¿qué es el trabajo valioso?

Para darles sentido a estos hallazgos contraintuitivos, investigamos detenidamente qué tipos de tareas consideran valiosas o tediosas los encuestados.

Las ideas tradicionales, nuestros informes antiguos y los datos cualitativos de nuestras entrevistas sugieren que los encuestados consideran que las tareas relacionadas con el desarrollo, como la programación, son valiosas, mientras que las que tienen menos valor, o incluso resultan tediosas, son aquellas relacionadas con la coordinación de la empresa, como la asistencia a reuniones. En este marco de categorización, la IA está en una mejor posición para contribuir al trabajo "valioso" que al "tedioso", según las definiciones de los propios encuestados.

Al recurrir a los datos cualitativos de nuestras entrevistas, descubrimos que, al responder a la pregunta del moderador sobre si consideraban que su trabajo es "significativo", los participantes solían medir el valor de su trabajo en relación con su impacto en el trabajo de otros.

Esto lo confirman las pruebas de DORA de los dos años anteriores relativas al impacto, muy beneficioso, del enfoque orientado al usuario sobre la satisfacción laboral.

Por ejemplo, al describir un reciente cambio de puesto, P10¹² indicó lo siguiente al explicar por qué había tomado la decisión: "Me permite influir en más personas. Me permite influir en más cosas". De la misma forma, P11 explicó: "Si creas algo desde cero y ves que llega al consumidor o al cliente, tienes una sensación de éxito, sabes que has hecho algo que la gente usa".

Entender que la "importancia" del trabajo de desarrollo deriva del impacto de la solución creada (y no directamente de la escritura de código) ayuda a explicar por qué hemos visto que los encuestados dedican menos tiempo a tareas valiosas, pero se sienten más satisfechos con su trabajo.

Aunque la IA está facilitando y agilizando las tareas que consideran valiosas, en realidad no sirve de ayuda para las tareas que no disfrutan. Que esté sucediendo esto mientras el trabajo manual y el desgaste se mantienen sin cambios a pesar de la adopción de la IA significa que no ha conseguido ayudarnos a evitar la monotonía de las reuniones, la burocracia y muchas otras tareas tediosas (figura 8).

La buena noticia es que la IA no ha empeorado las cosas ni ha afectado negativamente al bienestar de los encuestados.



Figura 8: No son datos, pero sí una representación de nuestra hipótesis de que la IA está ayudando a hacer las tareas valiosas, pero no las tediosas.

El prometedor impacto de la IA en los flujos de trabajo de desarrollo

En la última sección exploramos los resultados centrándonos en los usuarios individuales. Para el siguiente conjunto de resultados cambiamos el enfoque e indagamos en los procesos, los códigos base y la coordinación de los equipos. Aquí presentamos una lista con los resultados que hemos medido:

Complejidad del código	Hasta qué punto la complejidad y sofisticación del código minan la productividad.
Deuda técnica	Hasta qué punto la deuda técnica actual en la aplicación o servicio principal ha minado la productividad en los últimos seis meses.
Velocidad de revisión del código	El tiempo medio necesario para efectuar una revisión de código completa para la aplicación o servicio principal.
Velocidad de aprobación	La duración típica desde que se propone un cambio de código hasta que se recibe la aprobación para usarlo en la producción de la aplicación o servicio principal.
Coordinación de un equipo multidisciplinar	El grado de conformidad con esta afirmación: "En los últimos tres meses, he podido colaborar con los miembros de un equipo multidisciplinar de manera eficaz".
Calidad del código	El nivel de satisfacción o insatisfacción con la calidad del código subyacente al servicio o aplicación principal en los últimos seis meses.
Calidad de la documentación	La percepción de la documentación interna (manuales, información, comentarios de código) en términos de fiabilidad, buscabilidad, actualización y capacidad de ayuda.

Como antes, nuestro objetivo en este caso también es entender si estos aspectos parecen variar con la adopción de la IA. En la figura 9 se muestran nuestras mejores estimaciones respecto al cambio en estos resultados en relación con un aumento del 25 % en la adopción de la IA.

En general, estos patrones apuntan a una historia muy convincente en favor de la IA. Aquí presentamos los resultados más importantes de esta sección.

Un aumento del 25 % en la adopción de la IA se asocia con...

Un incremento del 7,5 % en la calidad de la documentación

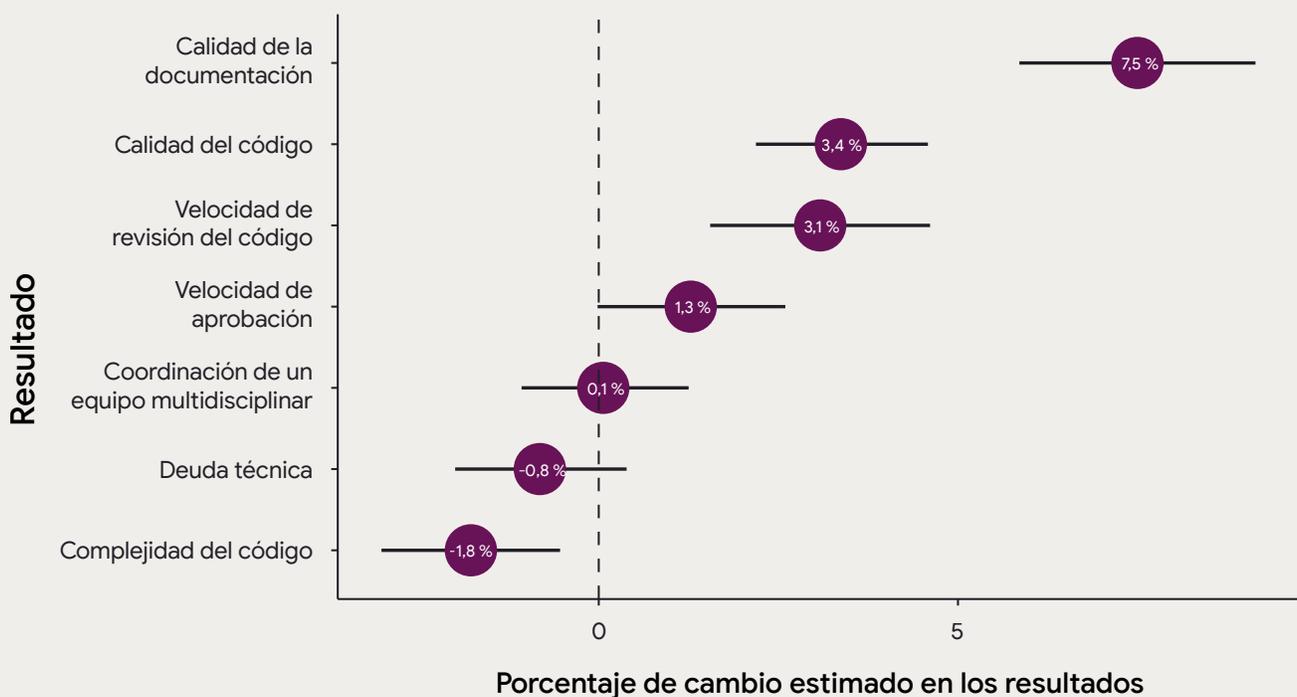
Un incremento del 3,4 % en la calidad del código

Un incremento del 3,1 % en la velocidad de revisión del código

Un incremento del 1,3 % en la velocidad de aprobación

Una disminución del 1,8 % en la complejidad del código

Si la adopción de la IA aumenta en un 25 %...



Punto = valor estimado

Barra de error = intervalo de incertidumbre del 89 %

Figura 9: Impactos de la adopción de la IA en las empresas.

Los datos presentados en el capítulo "Inteligencia artificial: adopción y actitudes" muestran que el uso más común de la IA es la escritura de código. El 67 % de los encuestados sostienen que la IA les está ayudando a mejorar su código. Aquí tenemos otra confirmación de esa idea. La IA parece mejorar la calidad del código y reducir su complejidad (figura 9). Cuando se combina con la posible refactorización del código antiguo, el código de alta calidad generado por IA podría mejorar el código base en general. Este código base podría mejorarse aún más con un acceso optimizado a documentación de calidad, que ahora mismo los usuarios están generando mediante IA (consulta [Inteligencia artificial: adopción y actitudes](#)).

Un código de mejor calidad es más fácil de revisar y aprobar. Si se combina con revisiones de código asistidas por IA, podemos hacer las revisiones y aprobaciones más rápido. Este patrón se refleja claramente en los datos (figura 9).

Ahora bien, que las revisiones y aprobaciones del código sean más rápidas no significa que los procesos de revisión y aprobación del código sean mejores ni más exhaustivos. Es posible que estemos ganando velocidad por una confianza excesiva en la ayuda de la IA en el proceso o en el código que ha generado. Este hallazgo no es incompatible con los patrones de la figura 9, pero tampoco ofrece una conclusión obvia.

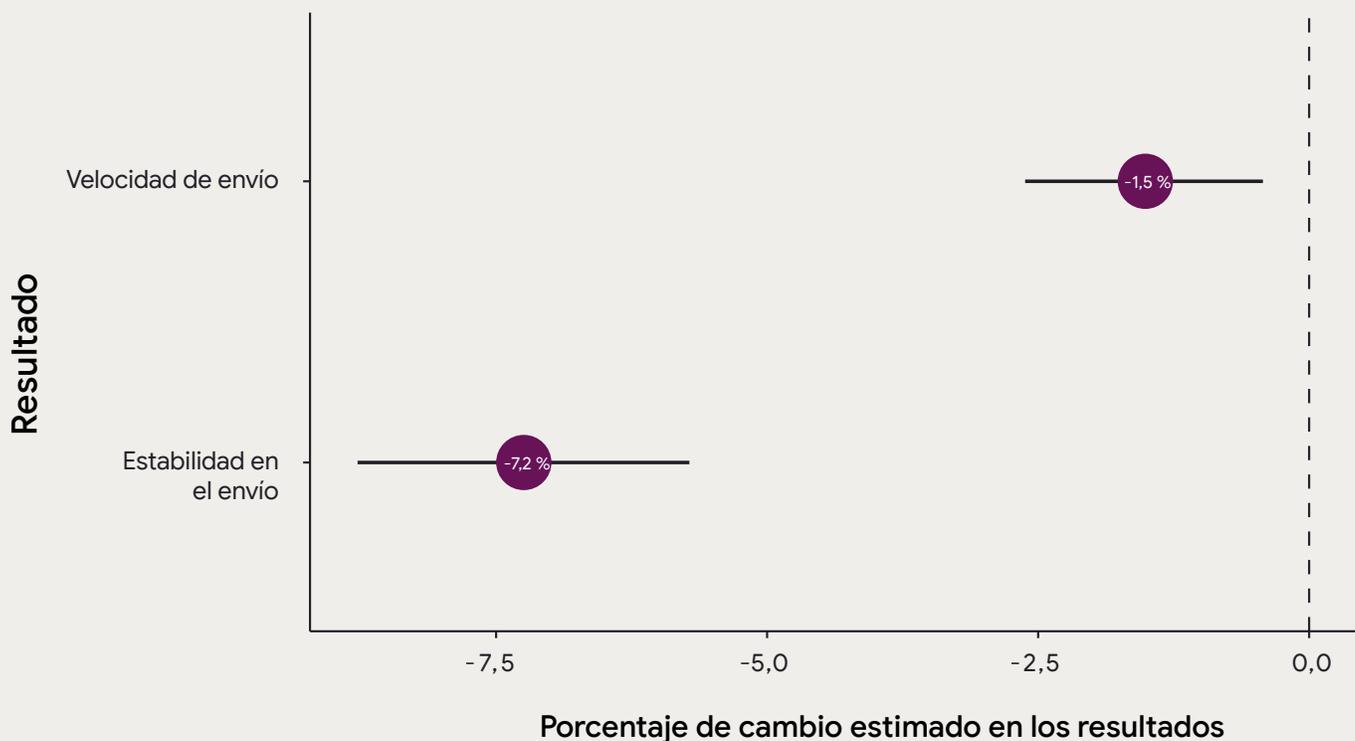
Tampoco es obvio que la calidad del código y de la documentación esté mejorando porque se estén generando por IA ni que la IA haya mejorado nuestra capacidad para sacar valor de un código y una documentación que, de otro modo, se habrían considerado de baja calidad. ¿Y si estamos bajando un poco el umbral de lo que consideramos código y documentación de calidad cuando usamos IA porque la propia IA tiene capacidad suficiente para darle sentido? Estas dos formas de entender estos patrones no son interpretaciones que se excluyan entre sí: ambas podrían estar contribuyendo a ellos.

Lo que parece claro en estos patrones es que la IA ayuda a los usuarios a sacar más partido de los documentos que necesitan y de los códigos base con los que trabajan. Además, ayuda a reducir los cuellos de botella del proceso de revisión y aprobación del código, y, por tanto, los costes que implican. Lo que no es tan obvio es cómo lo está haciendo exactamente y si estas ventajas conllevan beneficios secundarios, como mejoras en el envío de software.

La IA afecta negativamente al rendimiento del envío

Durante los últimos años, hemos visto que los indicadores de velocidad y estabilidad del envío de software han empezado a mostrar cierta independencia entre sí. Aunque se mantiene la asociación tradicional entre velocidad y estabilidad, las pruebas recientes sugieren que estos factores operan con suficiente independencia como para justificar su estudio por separado.

Si la adopción de la IA aumenta en un 25 %...



Punto = valor estimado

Barra de error = intervalo de incertidumbre del 89 %

Figura 10: Impactos de la adopción de la IA en la velocidad y la estabilidad del envío.

Al contrario de lo que esperábamos, nuestros hallazgos indican que la adopción de la IA está afectando negativamente al rendimiento del envío de software. Hemos visto que el efecto en la velocidad de envío es pequeño, pero suele ser negativo (se estima una reducción del 1,5 % por cada aumento del 25 % en la adopción de la IA). El impacto negativo en la estabilidad del envío es mayor (se estima una reducción del 7,2 % por cada aumento del 25 % en la adopción de la IA). Estos datos se muestran en la figura 10.

Históricamente, nuestra investigación ha mostrado que las mejoras en el proceso de desarrollo de software, donde se incluyen las mejoras de la calidad de la documentación y del código, la velocidad de revisión y aprobación del código, y la reducción de la complejidad del código, conllevan mejoras en el envío de software. Por ello, nos sorprendió ver que la IA mejora todas estas métricas del proceso, pero parece afectar negativamente a nuestras métricas de rendimiento relacionadas con la velocidad y la estabilidad en el envío.

A partir de nuestros hallazgos de años anteriores, nuestra hipótesis es que el cambio de paradigma fundamental que ha supuesto la IA en términos de velocidad de generación de código y productividad de los encuestados ha hecho que se olvide uno de los principios básicos de DORA: la importancia de los tamaños de lote pequeños. Esto quiere decir que, como la IA permite que los

usuarios produzcan una cantidad de código mucho mayor en el mismo tiempo, es posible, incluso probable, que el tamaño de las listas de cambios también esté creciendo. DORA ha demostrado de manera coherente que los cambios de mayor envergadura son más lentos y propensos a crear inestabilidad.

Si lo observamos todo en conjunto, nuestros datos sugieren que mejorar el proceso de desarrollo no mejora automáticamente el envío de software o, al menos, no con el cumplimiento adecuado de los conceptos de envío eficaz de software, como que haya tamaños de lote pequeños y mecanismos de prueba fiables.

Los beneficios de la IA sobre muchos factores individuales y organizativos importantes que fomentan las condiciones necesarias para un alto rendimiento del envío de software nos invitan a ser optimistas. Sin embargo, parece que la IA no es la panacea.



Los equipos y empresas de alto rendimiento usan IA, pero no parece que eso beneficie a los productos.

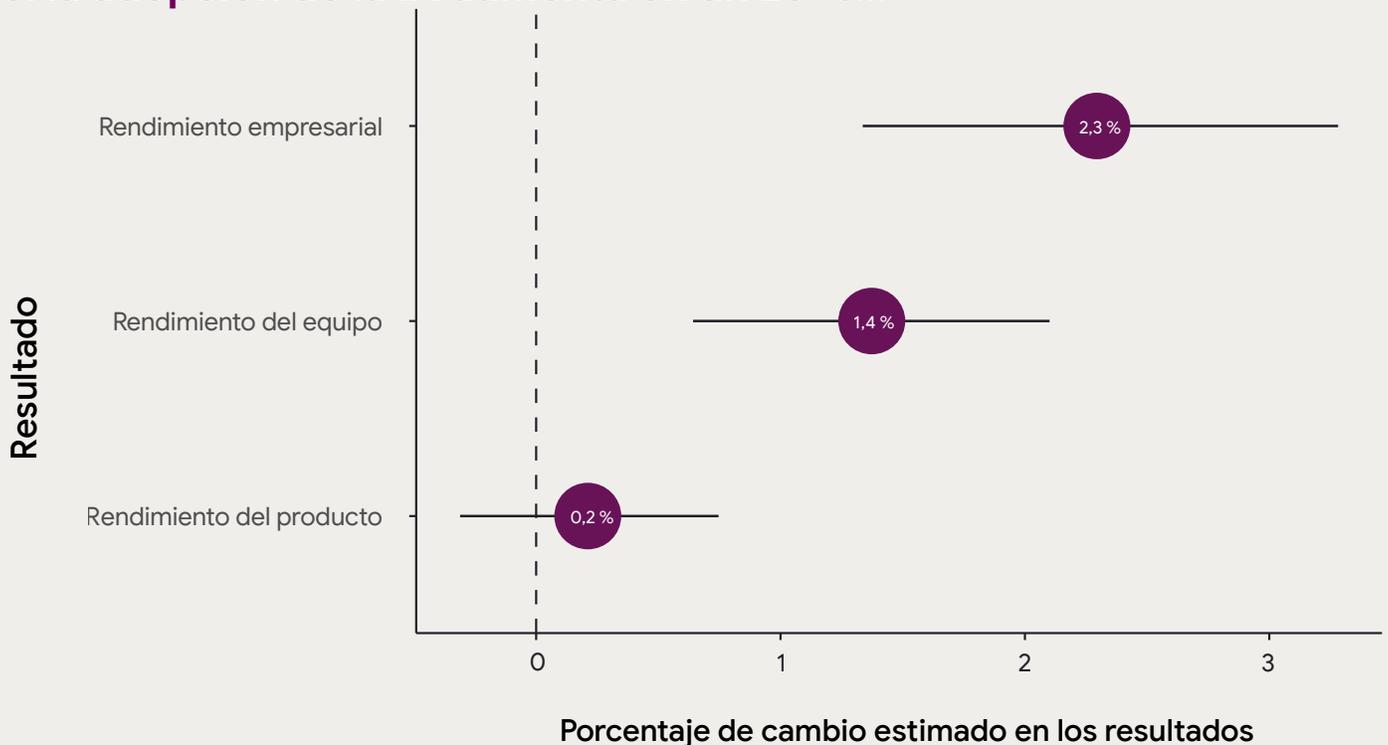
Ahora vamos a observar la relación de la IA con nuestros resultados menos directos:

Rendimiento empresarial	Puntuación factorial que representa el rendimiento, la rentabilidad, la cuota de mercado, los clientes totales, la eficiencia operativa, la satisfacción del cliente, la calidad de los productos/servicios y la capacidad para lograr los objetivos de la empresa en general.
Rendimiento del equipo	Puntuación factorial que representa la capacidad de un equipo para colaborar, innovar, trabajar de manera eficiente, confiar los unos en los otros y adaptarse.
Rendimiento del producto	Puntuación factorial que representa la usabilidad, la funcionalidad, el valor, la disponibilidad, el rendimiento (por ejemplo, la latencia) y la seguridad de un producto.

Establecer una conexión entre estos resultados y un usuario que adopta la IA es un proceso difícil y lleno de interferencias. A veces, parece que estamos intentando analizar el impacto de lo que comiste ayer en el buen rendimiento de tu empresa este año.

Existe una lógica que explica el salto entre el nivel micro (por ejemplo, una persona) y el macro (por ejemplo, una empresa). En el capítulo [Metodología hablamos sobre ese salto inferencial](#). Por ahora, veamos solo las asociaciones:

Si la adopción de la IA aumenta en un 25 %...



Punto = valor estimado

Barra de error = intervalo de incertidumbre del 89 %

Figura 11: Impactos de la adopción de la IA en el rendimiento de empresas, equipos y productos.

El rendimiento a nivel de empresa (un aumento estimado del 2,3 % por cada aumento del 25 % en la adopción de la IA) y de equipo (un aumento estimado del 1,4 % por cada aumento del 25 % en la adopción de la IA) parece inclinar la balanza en favor de la adopción de la IA (figura 11). Sin embargo, el rendimiento del producto sigue sin mostrar una asociación obvia con la adopción de la IA. Ahora, podemos pasar a intentar entender qué hay detrás de estos efectos.

Nuestra hipótesis es que los factores que contribuyen a un gran rendimiento empresarial y de los equipos son diferentes a los que influyen en el rendimiento de los productos.

Los equipos y las empresas confían mucho en la comunicación, el intercambio de conocimientos, la toma de decisiones y la cultura saludable. La IA podría reducir algunos cuellos de botella en esas áreas, lo que afectaría positivamente a los equipos y las empresas.

Sin embargo, el éxito de los productos podría depender de otros factores. Seguramente los buenos productos tengan causas subyacentes similares a las de los equipos y empresas de alto rendimiento, pero es probable que exista una conexión más cercana y directa con el flujo de trabajo de desarrollo y el envío de software, dos procesos que quizás puedan seguir estabilizándose tras la introducción de la IA.

La importancia única de los aspectos técnicos que subyacen a un buen producto podría explicarlo en parte, pero también hay cierto grado de arte y empatía tras un gran producto. Quizás sea algo difícil de creer para aquellos que piensan que cualquier problema se puede

resolver mediante la computación, pero determinados elementos del desarrollo de productos, como la creatividad o el diseño de la experiencia de usuario, pueden basarse enormemente aún (o para siempre) en la intuición y los conocimientos humanos.

Sin lugar a dudas, el rendimiento de las empresas, equipos y productos sigue estando interconectado. Si observamos las correlaciones bivariadas (Pearson), vemos que el rendimiento del producto tiene una correlación positiva moderada tanto con el rendimiento del equipo ($r = 0,56$, intervalo de confianza del 95 % = de 0,51 a 0,60) como con el rendimiento empresarial ($r = 0,47$, intervalo de confianza del 95 % = de 0,41 a 0,53).

Estos resultados influyen unos en otros de manera recíproca y crean interdependencias claras. Los equipos de alto rendimiento suelen desarrollar mejores productos, pero heredar un producto mediocre puede socavar su éxito. De la misma forma, las empresas de alto rendimiento fomentan el alto rendimiento de sus equipos a través de los recursos y procesos, pero las dificultades organizativas pueden suponer un impedimento. Por tanto, si la adopción de la IA beneficia tanto a los equipos y las empresas, es razonable esperar que también tenga ventajas para los productos.

La adopción de la IA está en sus primeras fases. Algunas ventajas e inconvenientes pueden tardar un tiempo en materializarse, ya sea por la naturaleza inherente del impacto de la IA o por la curva de aprendizaje asociada con la eficacia de su uso.

Quizás sea tan sencillo como que aún estamos averiguando cómo puede ayudar la IA a las empresas y a los equipos antes de ser totalmente conscientes de su potencial para la innovación y el desarrollo de los productos. En la figura 12 intentamos representar cómo podría desarrollarse la historia.

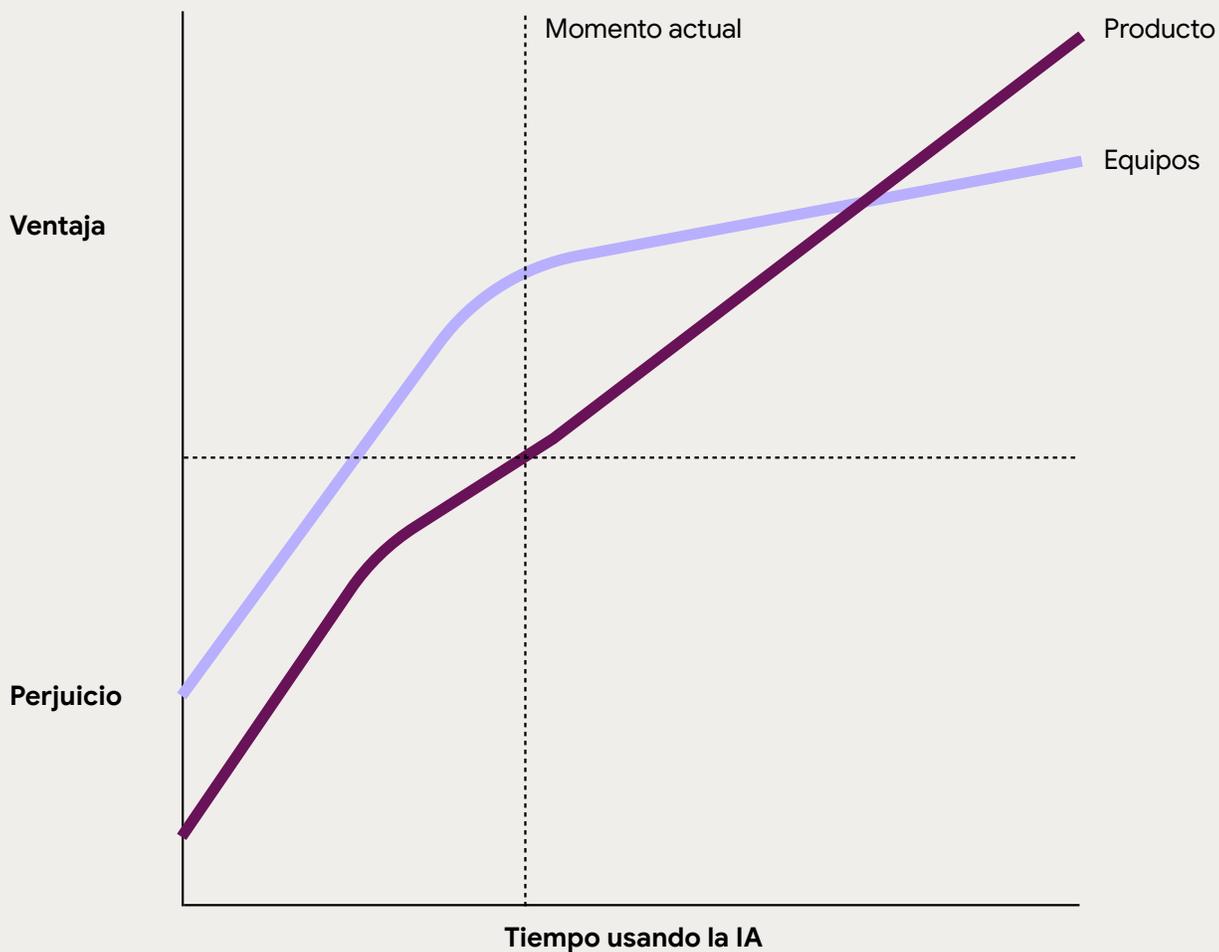


Figura 12: Representaciones de diferentes curvas de aprendizaje. Abstracción para fines demostrativos que no se basa en datos reales.

¿Y ahora qué?

Queríamos entender el potencial de la IA en el momento actual como ayuda para los usuarios, equipos y empresas. Los patrones que están apareciendo ponen de relieve que no es todo palabrería, sino que verdaderamente se está materializando algo.

Hay pruebas claras a favor de la adopción de la IA. Dicho esto, también es bastante evidente que hay muchos posibles obstáculos, problemas y formas en que la IA podría provocar efectos perjudiciales.

Adoptar la IA a escala quizás no sea tan fácil como pulsar un botón. Una estrategia medida, transparente y adaptable puede traer consigo grandes ventajas, pero tendrá que ser una estrategia codesarrollada por líderes, equipos, empresas e investigadores, así como por los propios desarrolladores de la IA.

Los líderes y las empresas tendrán que encontrar formas de priorizar la adopción en aquellas áreas en las que los empleados se vean más beneficiados.

Estas son algunas ideas sobre cómo orientar tu estrategia de adopción de la IA:

Define una misión clara para la IA y políticas beneficiosas para tu empresa y equipo.

Ofrece información transparente a los empleados sobre la misión de la IA, los objetivos y el plan de adopción. Al articular tanto la visión general como las políticas específicas (abordando problemas relacionados con los procedimientos, como la ubicación permitida del código y las herramientas disponibles), podrás disipar los celos y convertir la IA en una herramienta con la que todos puedan centrarse en tareas más valiosas, satisfactorias y creativas.

Crea una cultura de aprendizaje y experimentación continuos con la IA.

Promueve un entorno que fomente la investigación continua de las herramientas de IA reservando tiempo para que los usuarios y los equipos puedan descubrir casos prácticos útiles y dándoles la autonomía necesaria para decidir cómo aprovecharla. Fomenta la confianza en las tecnologías de la IA con experiencias prácticas en entornos aislados o de bajo riesgo. Para mitigar aún más los riesgos, puedes dedicarte a desarrollar una automatización de pruebas sólida. Implementa un marco de medición para evaluar la IA, no solo en cuanto a la adopción en sí, sino también a los principales impactos secundarios: cómo capacita a tus empleados, desarrolla el potencial de tus equipos y ofrece ventajas para quienes confían en tus productos.

Reconoce y equilibra los pros y contras de la IA para conseguir una ventaja competitiva.

Al reconocer los posibles inconvenientes, como dedicar menos tiempo a tareas valiosas, el exceso de dependencia de la IA, la posibilidad de que las ventajas obtenidas en un área provoquen dificultades en otra, y los impactos en la velocidad y la estabilidad del envío de software, podrás identificar las oportunidades para evitar escollos y dar forma a la trayectoria de la IA en tu empresa y equipo de forma positiva. Entender no solo cómo la IA puede ser beneficiosa, sino también cómo puede ser perjudicial, te permitirá acelerar las curvas de aprendizaje e investigar y convertir lo aprendido en acciones y en una ventaja competitiva real.

No cabe duda de que hay muchas razones para sentirnos entusiasmados, y aún más para seguir aprendiendo. El equipo de DORA seguirá investigando e intentando ofrecer perspectivas honestas, precisas y útiles, al igual que hemos hecho durante estos diez últimos años.

1. <https://www.goldmansachs.com/insights/top-of-mind/gen-ai-too-much-spend-too-little-benefit>

2. <https://www.goldmansachs.com/insights/articles/AI-poised-to-drive-160-increase-in-power-demand>

3. <https://www.washington.edu/news/2023/07/27/how-much-energy-does-chatgpt-use/>

4. <https://www.gatesnotes.com/The-Age-of-AI-Has-Begun>

5. <https://www.businessinsider.com/ai-chatgpt-homework-cheating-machine-sam-altman-openai-2024-8>

6. <https://www.safe.ai/work/statement-on-ai-risk>

7. <https://github.blog/news-insights/research/research-quantifying-github-copilots-impact-on-developer-productivity-and-happiness/>

8. https://www.gitclear.com/coding_on_copilot_data_shows_ais_downward_pressure_on_code_quality

9. <https://www.nytimes.com/2024/04/15/technology/ai-models-measurement.html>

10. <https://dora.dev/capabilities>

11. Es importante aclarar que no es un enfoque único, pero de alguna forma lo es para este ámbito

12. (P[N]), por ejemplo (P1), es el seudónimo de cada entrevistado.

Ingeniería de plataformas



Introducción

La ingeniería de plataformas es una disciplina de ingeniería emergente que cada vez acapara más interés y tiene más presencia en el sector. Los líderes del sector, como Spotify y Netflix, junto con libros como *Team Topologies*¹, han ayudado a despertar el entusiasmo en el público.

La ingeniería de plataformas es una disciplina sociotécnica en la que los ingenieros se centran tanto en las interacciones sociales entre diferentes equipos como en los aspectos técnicos de la automatización, el autoservicio y la repetibilidad de los procesos. Los conceptos que subyacen de la ingeniería de plataformas lleva años estudiándolos, entre otros, el equipo de DORA.

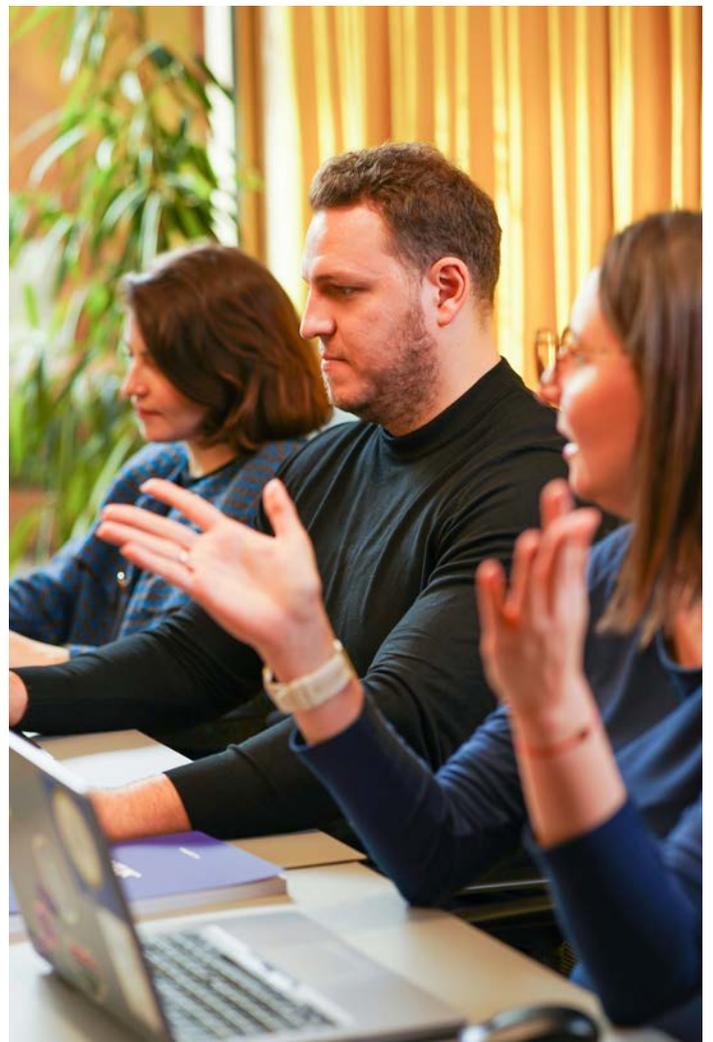
Por lo general, nuestra investigación aborda la forma de enviar software a usuarios externos, mientras que los resultados de los equipos de plataformas suelen ser un conjunto de APIs, herramientas y servicios internos diseñados para respaldar el desarrollo de software y el ciclo de vida de las operaciones.

En la ingeniería de plataformas, se dedica mucha atención y energía a mejorar la experiencia de los desarrolladores. Para ello, se crean procedimientos o flujos de trabajo de autoservicio altamente automatizados que los usuarios de la plataforma utilizan al interactuar con los recursos necesarios para enviar y usar las aplicaciones. El objetivo es abstraer las complejidades de crear y enviar software para que el desarrollador solo tenga que preocuparse del código.

Algunos ejemplos de las tareas automatizadas con estos procedimientos son el aprovisionamiento de aplicaciones y bases de datos, la gestión de esquemas, la ejecución de pruebas, la creación e implementación del aprovisionamiento de infraestructuras, y la gestión de DNS.

Hay conceptos de la ingeniería de plataformas, como la reducción de competencias² en un sistema compartido, que pueden parecer contrarios a enfoques como el de "quien lo crea lo ejecuta". Sin embargo, consideramos la ingeniería de plataformas como un método para escalar la adopción de estas prácticas en una empresa, ya que, cuando una competencia está en la plataforma, los equipos cuentan con ella de forma gratuita básicamente por usar la plataforma.

Por ejemplo, si la plataforma puede ejecutar pruebas unitarias y comunicar los resultados directamente a los equipos de desarrollo (sin que estos tengan que crear ni gestionar el entorno de ejecución de las pruebas), la función de plataforma de integración continua les permite centrarse en escribir pruebas de alta calidad. En este ejemplo, la función de integración continua puede escalarse en toda la empresa y ayudar a distintos equipos a mejorar su competencia al realizar pruebas continuas³ y automatizarlas.⁴

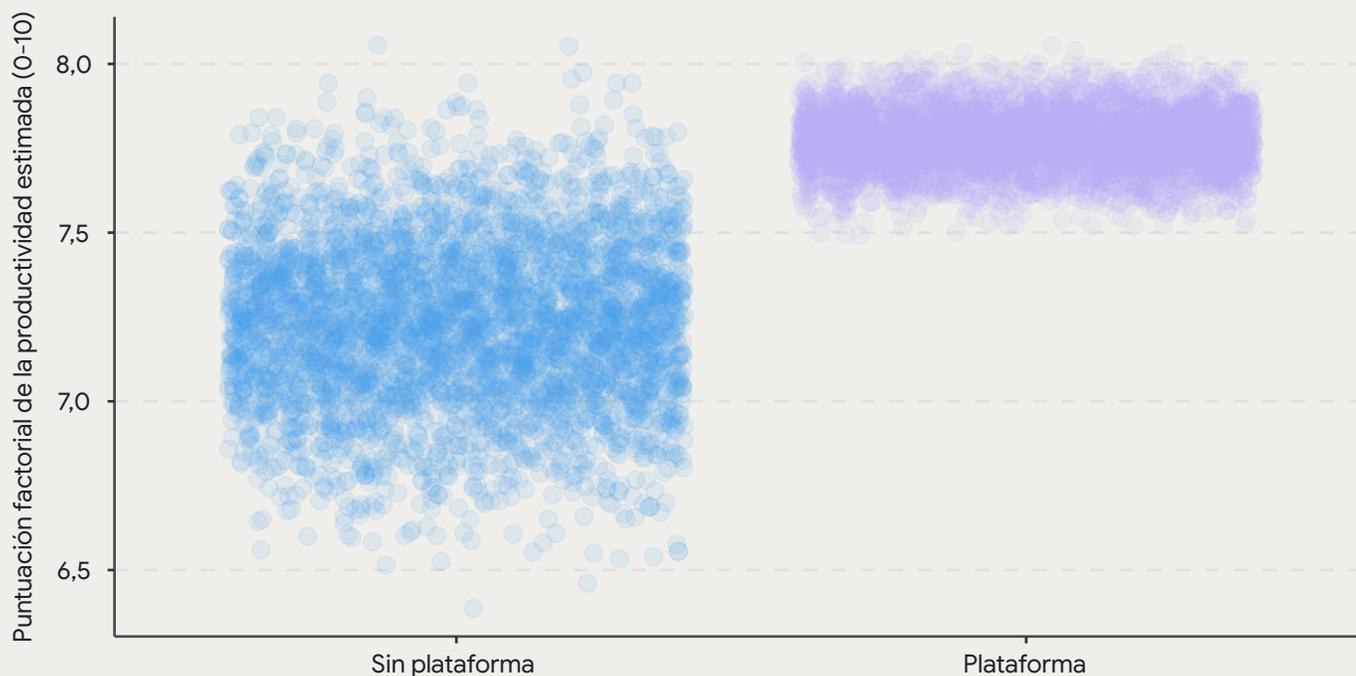


Un factor clave para que salga bien es abordar la ingeniería de plataformas con un enfoque orientado al usuario (en el contexto de una plataforma interna para desarrolladores, los usuarios son los desarrolladores), la independencia de los desarrolladores y la mentalidad centrada en el producto. Tampoco nos sorprende mucho si tenemos en cuenta que el enfoque orientado al usuario se identificó como un factor clave para mejorar el rendimiento empresarial, tanto este año como en años anteriores.⁵ Sin este enfoque, más que una ayuda, la plataforma sería un impedimento.

En el informe de este año, quisimos comprobar la relación entre las plataformas y el rendimiento operativo y del envío de software. Algunos de los resultados fueron positivos. Los usuarios de plataformas internas para desarrolladores

tuvieron niveles de productividad individual un 8 % más altos, una cifra que aumenta hasta el 10 % si hablamos del rendimiento como equipo. Además, el rendimiento operativo y del envío de software de una empresa crece un 6 % al usar una plataforma. No obstante, estas mejoras no están exentas de inconvenientes. La velocidad y la estabilidad de los cambios disminuyeron un 8 % y un 14 % respectivamente, un resultado inesperado.

En las siguientes secciones, profundizaremos en las cifras y los detalles, y desvelaremos algunos datos sorprendentes de esta encuesta. Aplicar nuestros hallazgos clave puede potenciar el éxito de una plataforma, tanto si la iniciativa de ingeniería de plataformas se acaba de poner en marcha como si ya hace años que está funcionando.



Cada punto es una de las 8000 estimaciones de la puntuación de productividad media más plausible.

Figura 13: Factor de productividad para los usuarios cuando usan o no usan una plataforma interna para desarrolladores.

La promesa de la ingeniería de plataformas

Dadas las posibles ventajas en materia de eficiencia y productividad que se pueden conseguir con la práctica, las plataformas internas para desarrolladores están despertando cada vez más interés en gran parte del sector de TI y de los desarrolladores de software. Para la encuesta de este año, definimos la "plataforma interna para desarrolladores" de una manera bastante amplia⁶ y descubrimos que el 89 % de los participantes estaban usando alguna. Los modelos de interacción son muy diversos entre ellos.

Los puntos de datos van en consonancia con el amplio nivel de interés del sector por la ingeniería de plataformas y el carácter emergente de este campo.

En general, el impacto de las plataformas es positivo: los usuarios eran un 8 % más productivos y los equipos mostraron un desempeño un 10 % mayor al usar una plataforma interna para desarrolladores.

Y más allá de la productividad, también vimos ventajas por usar una plataforma en el rendimiento general de las empresas, con un aumento del 6 %. En general, las empresas pueden enviar software, satisfacer las necesidades de los usuarios e impulsar el valor empresarial más rápido gracias al uso de una plataforma.

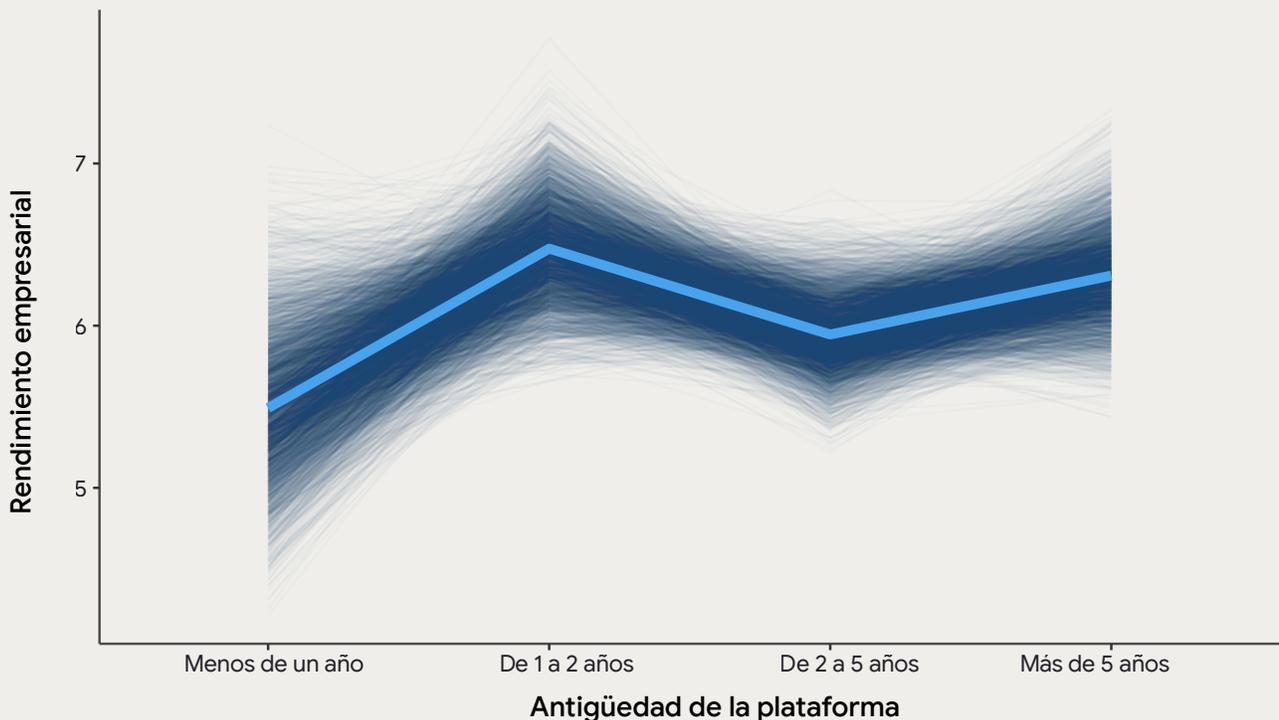


Figura 14: Cambio en el rendimiento de una empresa al usar una plataforma interna para desarrolladores en comparación con la antigüedad de la plataforma.

Si analizamos la productividad teniendo en cuenta la antigüedad de la plataforma, vemos un ligero aumento del rendimiento inicial al poner en marcha la iniciativa de ingeniería de plataformas, seguido de periodos de descenso y recuperación conforme la plataforma se va consolidando y gana madurez. Se trata del patrón típico de las iniciativas de transformación, que muestran ventajas al principio, pero en las que surgen dificultades con el paso del tiempo.

A largo plazo, las mejoras en la productividad se mantienen, lo que demuestra el potencial general de la plataforma interna para desarrolladores en el envío de software y los procesos operativos.



Resultado clave: impacto de la independencia de los desarrolladores

La independencia de los desarrolladores tuvo un importante impacto en el nivel de productividad de los usuarios individuales y equipos al enviar software con una plataforma interna para desarrolladores. La independencia de los desarrolladores se define como su "capacidad para realizar sus tareas durante todo el ciclo de vida de la aplicación sin recurrir a un equipo de capacitación".

Tanto a nivel de equipo como de usuarios individuales, se aprecia una mejora del 5 % en la productividad cuando los usuarios de la plataforma pueden llevar a cabo sus tareas de principio a fin sin equipos de capacitación. Este hallazgo se relaciona con uno de los principios fundamentales de la ingeniería de plataformas: priorizar los flujos de autoservicio.

Es fundamental para los equipos de plataformas porque pone el foco en la recopilación de los comentarios de los usuarios, que constituye una parte muy importante del proceso de ingeniería de plataformas. Las respuestas a la encuesta no indicaron qué tipo de comentarios son los más eficaces, pero normalmente son las conversaciones informales, las herramientas de seguimiento de incidencias y, en menor medida, el desarrollo conjunto de manera continuada, encuestas, telemetría y entrevistas.

Todos estos métodos pueden ayudar a entender si los usuarios son capaces de completar sus tareas de manera independiente. Los datos de la encuesta también destacan el impacto negativo de no recopilar comentarios sobre la plataforma.

Resultado secundario: impacto de un equipo de plataforma específico

Curiosamente, tener un equipo de plataforma específico apenas tuvo impacto alguno en la productividad de los usuarios individuales. Sin embargo, a nivel de equipo, supuso una mejora de la productividad del 6 %. Esto resulta sorprendente porque se trata de un impacto desigual y sugiere que tener un equipo de plataforma específico es útil para los usuarios individuales, pero tiene más impacto en los equipos en conjunto.

Como los equipos se componen de varios desarrolladores con distintas responsabilidades y habilidades, es natural que tengan un conjunto de tareas más diverso que un ingeniero concreto. Es posible que contar con un equipo de ingeniería de plataforma específico permita que la plataforma respalde mejor la diversidad de las tareas de un equipo.

En general, el impacto de tener una plataforma interna para desarrolladores es positivo en términos de productividad.

Estos son los factores clave:

Un enfoque orientado al usuario que favorece la independencia de los desarrolladores a través del autoservicio y los flujos de trabajo que se pueden llevar a cabo de principio a fin de manera autónoma. En el contexto de la plataforma, los usuarios son equipos de desarrollo e ingeniería interna.

Igual que sucede con otras transformaciones, en la ingeniería de plataformas también se produce la "curva en J", por lo que las ventajas en la productividad se estabilizarán con la mejora continua.

El inconveniente inesperado

Aunque la ingeniería de plataformas presenta algunos beneficios claros, como una mayor sensación de productividad entre los equipos y usuarios individuales, y mejoras en el rendimiento de la empresa, también presenta un inconveniente inesperado: una reducción de la velocidad y de la estabilidad de los cambios.

De forma fortuita, descubrimos una conexión muy interesante entre la inestabilidad de los cambios y el desgaste.

Velocidad

En el caso de la velocidad, vimos que se redujo aproximadamente un 8 % en comparación con la de quienes no usan ninguna plataforma. Tenemos algunas hipótesis sobre cuál podría ser la causa subyacente.

Primero, toda la maquinaria adicional por la que tienen que pasar los cambios hasta desplegarse en la producción reduce su velocidad general. Cuando se usa una plataforma interna para desarrolladores para crear y enviar software, lo normal es que suela haber un mayor número de "transferencias" entre los sistemas e, implícitamente, entre los equipos.

Por ejemplo, cuando se confirma el código para llevar un control del código fuente, hay diferentes sistemas que lo seleccionan

automáticamente para hacer pruebas, comprobaciones de seguridad, despliegues y tareas de supervisión.

Cada una de estas transferencias suma algo de tiempo al proceso general, lo que reduce la velocidad, a pesar de que se produce un incremento neto de la capacidad para completar el trabajo.

En segundo lugar, en el caso de los encuestados que indicaron que están obligados a "usar únicamente la plataforma para realizar tareas durante todo el ciclo de vida de la aplicación", se produjo un descenso del 6 % en la velocidad. Aunque no se trata de una conexión definitiva, también podría relacionarse con la primera hipótesis.

Si el número de sistemas y herramientas que forman parte del desarrollo y lanzamiento de software aumenta debido a la presencia de una plataforma, la relación entre la exclusividad y el descenso de la productividad podría explicarse por el hecho de que se obligue a usar la plataforma aunque no sea lo más adecuado para un fin concreto o aumente de manera intrínseca la latencia del proceso.

Para remediarlo, es importante que en tus iniciativas de ingeniería de plataformas mantengas un enfoque orientado al usuario y a su independencia.

Inestabilidad de los cambios y desgaste

Al fijarnos en la estabilidad de los cambios en las aplicaciones desarrolladas y operadas con una plataforma interna para desarrolladores, vimos que, curiosamente, la estabilidad de los cambios se redujo en un 14 %. Esto indica que la tasa de errores al aplicar cambios y la tasa de reelaboración aumentan significativamente al usar una plataforma.

Aunque vimos algo aún más interesante en los resultados: la combinación de inestabilidad y uso de una plataforma se relaciona con niveles más altos de desgaste. Esto no quiere decir que las plataformas provoquen desgaste, sino que la combinación de inestabilidad y plataformas puede ser especialmente problemática en este sentido. Algo similar sucede con el descenso en la velocidad: tampoco sabemos a ciencia cierta por qué se produce el cambio en el desgaste, pero tenemos algunas hipótesis.

En primer lugar, la plataforma permite que los desarrolladores y equipos envíen cambios con más confianza en que, si son incorrectos, se solucionarán rápidamente. En este sentido, el mayor nivel de inestabilidad no tiene por qué ser necesariamente negativo, ya que la plataforma ayuda a los equipos a experimentar y realizar cambios, lo que a su vez aumenta la tasa de errores al aplicar cambios y la reelaboración.

Otra idea es que la plataforma no es eficaz para garantizar la calidad de los cambios o despliegues en la producción.

También puede ser que la plataforma ofrezca una función de pruebas automáticas que se ejecute siempre que se incluyan pruebas en la aplicación, pero que los equipos de aplicaciones no la aprovechen al máximo y prioricen la velocidad frente a la calidad, sin mejorar sus pruebas. En cualquier caso, los cambios incorrectos van avanzando por el proceso, lo que se traduce en tareas de reelaboración.

Una tercera posibilidad es que los equipos con un alto nivel de inestabilidad en los cambios y desgaste tiendan a crear plataformas para mejorar ambas cuestiones. Esto tiene sentido, ya que la ingeniería de plataformas se suele ver como una práctica que reduce el desgaste y aumenta la capacidad de enviar cambios más pequeños de manera coherente. Con esta hipótesis, la ingeniería de plataformas sería el reflejo de una empresa en la que hay desgaste e inestabilidad en los cambios.

En los dos primeros escenarios, las tareas de reelaboración que posibilita la plataforma pueden verse como un obstáculo, lo que también podría incrementar el desgaste. El segundo escenario en particular, en el que la plataforma permite que se realicen cambios inadecuados, contribuiría más al desgaste, pero en ambos casos el equipo o usuario seguiría sintiéndose productivo debido a su capacidad para enviar cambios y funciones. En el tercer escenario, la iniciativa de ingeniería de plataformas se implementa como solución para abordar la inestabilidad en los cambios y el desgaste.

Equilibrar los pros y contras

Aunque la ingeniería de plataformas no es la panacea, puede ser una importante disciplina para el proceso general de desarrollo de software y operaciones relacionadas. Como con cualquier disciplina, la ingeniería de plataformas tiene ventajas y desventajas.

Basándonos en nuestra investigación, se pueden tomar un par de medidas para equilibrarlas al poner en marcha una iniciativa de ingeniería de plataformas. Eso ayudará a tu empresa a disfrutar de las ventajas de la ingeniería de plataformas al tiempo que se supervisan y gestionan los posibles contras.

Primero, prioriza las competencias de la plataforma que fomenten la independencia de los desarrolladores y el autoservicio. Al hacerlo, es importante encontrar un equilibrio que no obligue a usar la plataforma para todos los aspectos del ciclo de vida de la aplicación, algo que podría repercutir negativamente en la independencia de los desarrolladores.

Recomendamos que la plataforma permita a los usuarios usar las herramientas y funciones de automatización disponibles por separado para fomentar la independencia, aunque eso suponga un aumento de la complejidad. Para compensarlo, puede designarse un equipo de plataformas específico que colabore activamente con los usuarios de la plataforma y recopile sus comentarios.

La colaboración y los comentarios refuerzan el enfoque orientado a los usuarios de la iniciativa y contribuirán al éxito de la plataforma a largo plazo. Como los datos demuestran, se pueden usar muchos métodos distintos para recopilar comentarios, así que emplea más de un enfoque para contar con tantas opiniones como sea posible.

En segundo lugar, supervisa detenidamente la inestabilidad de los cambios en tus aplicaciones e intenta entender si es intencional o no. Las plataformas pueden fomentar la experimentación en términos de inestabilidad, aumento de la productividad y mejora del rendimiento a escala.

Sin embargo, esa misma inestabilidad puede generar desgaste e inestabilidad general, por lo que hay que supervisarla detenidamente y tenerla en cuenta a lo largo del proceso de ingeniería de plataformas. Al mismo tiempo, es importante que seas consciente del nivel de inestabilidad que puedes soportar. Usar objetivos de nivel de servicio (SLOs) y presupuestos de errores de Site Reliability Engineering (SRE) puede ayudarte a medir tu tolerancia al riesgo y la eficacia de la plataforma a la hora de posibilitar la experimentación de manera segura.

Las plataformas internas para desarrolladores se centran en gran medida en la experiencia de los desarrolladores, pero hay muchos otros equipos (como administradores de bases de datos y profesionales del ámbito de la seguridad y las operaciones) que tienen que poder enviar y usar software de manera eficaz.

En tus iniciativas de ingeniería de plataformas, fomenta una cultura de orientación al usuario y mejora continua en todos los equipos que se ajuste a los objetivos de la empresa.

De esta forma, las funciones, servicios y APIs de la plataforma tendrán la misma finalidad: ayudar a los usuarios individuales y equipos a enviar software y generar valor empresarial.



1. Skelton, Matthew y Pais, Manuel. 2019. Team Topologies: Organizing Business and Technology Teams for Fast Flow. IT Revolution Press. <https://teamtopologies.com/>
2. <https://cloud.google.com/blog/products/application-development/richard-seroter-on-shifting-down-vs-shifting-left>
3. <https://dora.dev/capabilities/continuous-integration/>
4. <https://dora.dev/capabilities/test-automation/>
5. <https://dora.dev/research/2023/>, <https://dora.dev/research/2016/>
6. <https://dora.dev/research/2024/questions/#platform-engineering>

Experiencia de los desarrolladores



Datos clave

El software no se desarrolla por sí solo. Aun con ayuda de la IA, el software lo crean los usuarios, y sus experiencias prácticas son fundamentales para el éxito de las empresas.

En el informe de este año, volvimos a ver que, para que las empresas y sus empleados prosperen, es necesario que lo que los desarrolladores crean concuerde con lo que los usuarios necesitan. Los desarrolladores son más productivos, menos propensos a sufrir desgaste y más capaces de crear productos de alta calidad si desarrollan software con un enfoque orientado al usuario.

En última instancia, el software se crea para las personas, así que es responsabilidad de la empresa fomentar entornos que ayuden a los desarrolladores a centrarse en crear software que mejore la experiencia de usuario. También hemos descubierto que los entornos estables, en los que no se están cambiando constantemente las prioridades, generan mejoras pequeñas pero significativas en la productividad, además de disminuir bastante el desgaste de los empleados.

Los factores ambientales tienen importantes consecuencias en la calidad de los productos desarrollados y en la experiencia general de los desarrolladores encargados de crear esos productos.

Pon al usuario primero y se estabilizará casi todo

Creemos que los desarrolladores tienen un trabajo muy interesante. Están a la vanguardia de los avances tecnológicos y ayudan a conformar nuestras formas de vivir, trabajar e interactuar con el mundo.

Su trabajo está ligado de manera intrínseca a las personas, es decir, a los usuarios del software y las aplicaciones que crean. Aun así, los desarrolladores suelen trabajar en entornos en los que se priorizan las funciones y la innovación y ponen menos énfasis en averiguar de qué forma aportan valor esas funciones a las personas que usan los productos que ellos crean.

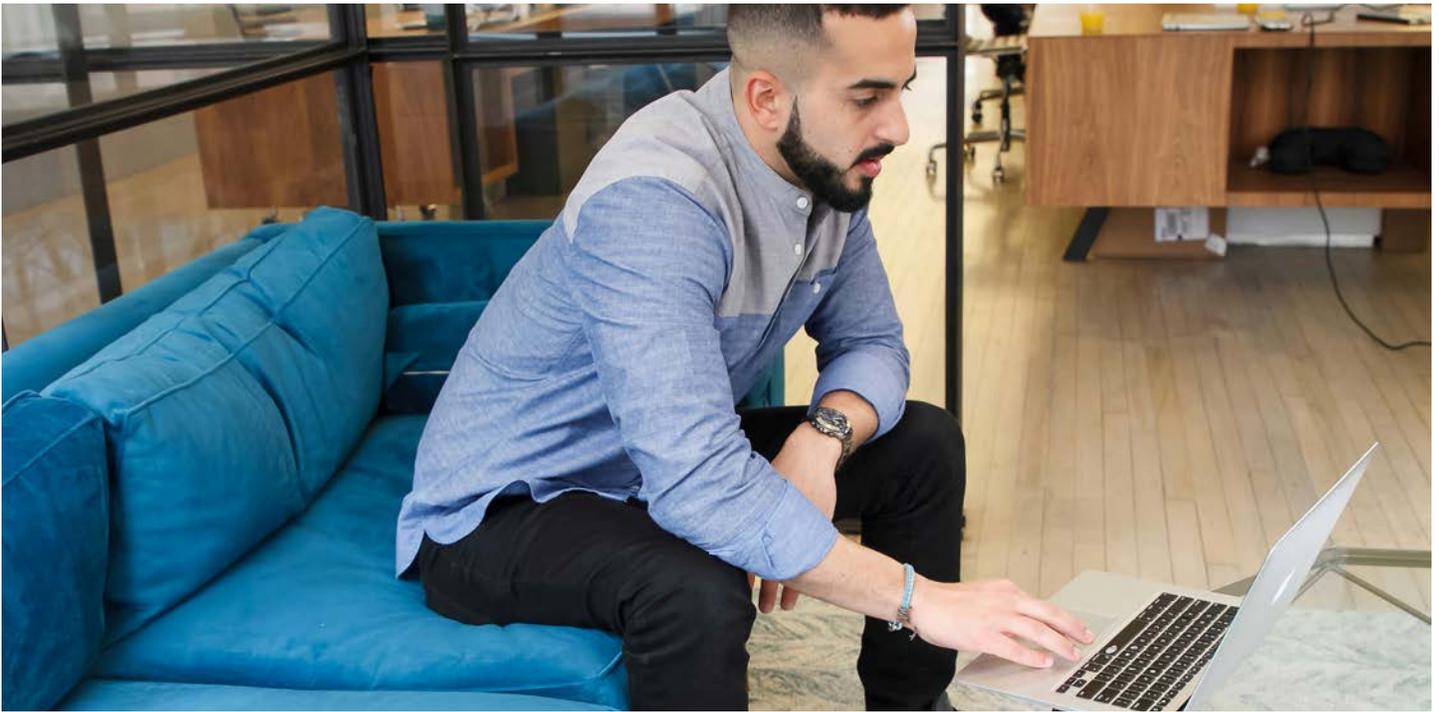
Aquí recopilamos pruebas claras que demuestran que adoptar un proceso de desarrollo de software en el que se

priorice al usuario final beneficia tanto a los empleados como a la empresa.

Este año, hemos hecho preguntas centradas en entender si los desarrolladores:

1. Usan los comentarios de los usuarios para modificar las funciones y reorganizar sus prioridades
2. Saben qué quieren conseguir los usuarios con una aplicación o servicio específico
3. Creen que centrarse en el usuario es fundamental para que la empresa tenga éxito
4. Creen que la experiencia del usuario es una de las principales prioridades empresariales





Nuestros resultados y lo que implican

Nuestros datos sugieren claramente que las empresas que se guían por las necesidades y los problemas de los usuarios crean mejores productos.

Hemos visto que centrarse en el usuario mejora la productividad y la satisfacción laboral, al tiempo que reduce el riesgo de desgaste.

Es importante destacar que estas ventajas van más allá de los empleados a nivel individual y abarcan toda la empresa. En años anteriores, destacamos que las empresas de alto rendimiento envían software de manera rápida y fiable. Esto significa que el rendimiento del envío de software es uno de los requisitos para el éxito.

Sin embargo, nuestros datos indican que hay otra vía para lograr el éxito:

Los desarrolladores, sus empleadores y, a fin de cuentas, las empresas en general pueden adoptar un enfoque de desarrollo de software orientado al usuario.

Descubrimos que, cuando las empresas conocen y entienden las necesidades de los usuarios, no es necesario que haya velocidad ni estabilidad del envío de software para garantizar la calidad de un producto. El producto será de alta calidad siempre que se priorice la experiencia de usuario.

Cuando las empresas no se centran en el usuario, no incorporan sus opiniones en el proceso de desarrollo, y la única vía para asegurar la calidad del producto es apostar todo a un envío estable y rápido (consulta la figura 15).

Sabemos que algunas empresas tienden a centrarse en crear funciones y tecnologías innovadoras. Si nos fijamos en el valor, tiene sentido, porque, después de todo, los desarrolladores conocen los pros y contras de la tecnología mucho mejor que el usuario medio.

Sin embargo, desarrollar software basándose en suposiciones sobre la experiencia de usuario aumenta la probabilidad de que se creen funciones que quizás sean muy llamativas, pero que apenas se usen en la práctica.¹

Cuando las empresas y los empleados entienden cómo interaccionan sus usuarios con el mundo, es más probable que creen funciones que satisfagan sus necesidades reales, lo que aumenta las posibilidades de que de verdad se usen.

Centrarte en crear productos para los usuarios es garantía de calidad.

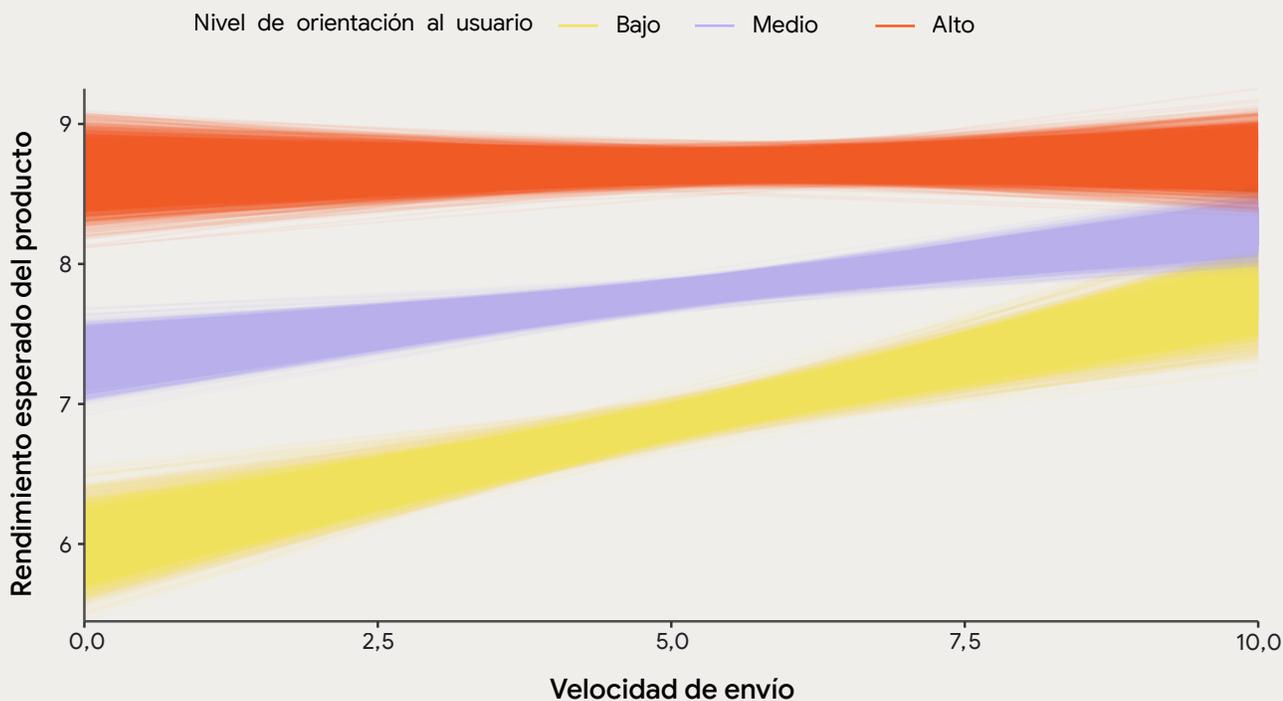


Figura 15: Rendimiento del producto y velocidad en el envío según 3 niveles de orientación hacia el usuario

¿Por qué desarrollar software con un enfoque orientado al usuario es una filosofía y práctica tan determinante?

Según la investigación académica, conocer el propósito del trabajo beneficia a los empleados y a las empresas.^{2,3}

Por ejemplo, en una encuesta reciente, el 93 % de los trabajadores indicaron que es importante tener un trabajo en el que sientas que haces algo importante.⁴ De manera similar, otra encuesta descubrió que, de media, los encuestados serían capaces de renunciar al 23 % de sus futuras ganancias si eso implicara tener un trabajo que siempre fuera importante.⁵

Resulta llamativo que los empleados estén dispuestos a hacer ese sacrificio. Dice mucho sobre lo que los motiva e indica que quieren dedicar su tiempo a hacer algo que importe.

"Sería magnífico que todos pudiéramos trabajar en una empresa que repercutiera en personas ajenas a ella, o [en] la comunidad local, de manera positiva. No siempre es así. No siempre es posible. En gran medida, el objetivo de la conducción autónoma es que las personas que saben conducir puedan dormirse mientras van por la autopista. Pero yo no estoy aquí para eso. Yo quiero ayudar a aquellos que no conducen a ir a donde quieran, a tener la libertad de hacer lo que quieran". (P2)⁶

Traza una dirección clara:

Adoptar un enfoque de desarrollo de software orientado al usuario puede cambiar por completo la forma en la que los desarrolladores ven su trabajo. En lugar de enviar funciones arbitrarias y presuponer que los usuarios podrían usarlas, los desarrolladores pueden aprovechar los comentarios de los usuarios para saber qué priorizar durante la creación.

Este enfoque les asegura que las funciones en las que trabajan tienen una razón de ser. De repente, su trabajo tiene sentido: asegurarse de que los usuarios disfruten de una excelente experiencia al usar sus productos y servicios. Ya no hay una desconexión entre el software que se desarrolla y el mundo en el que se usa.

Los desarrolladores pueden ver el impacto directo de su trabajo a través del software que crean.

"Como empresa, tenemos presión por obtener resultados. Hay tantas cosas para llamar la atención y debates sobre cómo quieres mejorar, que parece que los últimos cambios en nuestra estructura provocan que nos centremos más en los resultados que en la calidad. Para mí, personalmente, es una auténtica pesadilla". (P9)

Aumenta las colaboraciones interfuncionales:

Por mucho talento que tenga, ningún desarrollador puede crear software solo. Crear productos de alta calidad requiere la colaboración de muchas personas, normalmente con talentos diferentes pero complementarios.

Un enfoque de desarrollo centrado en el usuario permite que los desarrolladores colaboren en equipos interfuncionales dentro de la empresa. De esta forma, sus responsabilidades van más allá del mero envío de software: ahora forman parte de un equipo diseñado para crear increíbles experiencias para los usuarios.

Adoptar este enfoque en el desarrollo de software puede ayudar a los desarrolladores a eliminar los silos, buscar puntos en común, fomentar el trabajo en equipo y crear oportunidades para aprender más los unos de los otros. Solucionar problemas adquiere otra dimensión. No se trata solo de cómo solucionar problemas técnicos, sino de cómo hacerlo de la forma más útil para el usuario.

Este enfoque puede potenciar la implicación de los empleados y crear un entorno más estimulante a nivel intelectual para diluir la sensación de estancamiento que se asocia con el desgaste.

¿Qué pueden hacer las empresas?

Basándonos en nuestros hallazgos, recomendamos a las empresas invertir tiempo y recursos en conocer a sus usuarios, así como centrarse en entender para quiénes están creando las soluciones y cuáles son sus problemas. No nos cabe duda de que es una inversión que vale la pena.

Resiste la tentación de hacer suposiciones sobre los usuarios. Obsérvalos en sus entornos, hazles preguntas y sé lo suficientemente humilde para hacer los cambios necesarios en función de lo que te digan. Así, los desarrolladores serán más productivos y menos propensos a experimentar desgaste, al tiempo que crearán productos de mayor calidad.

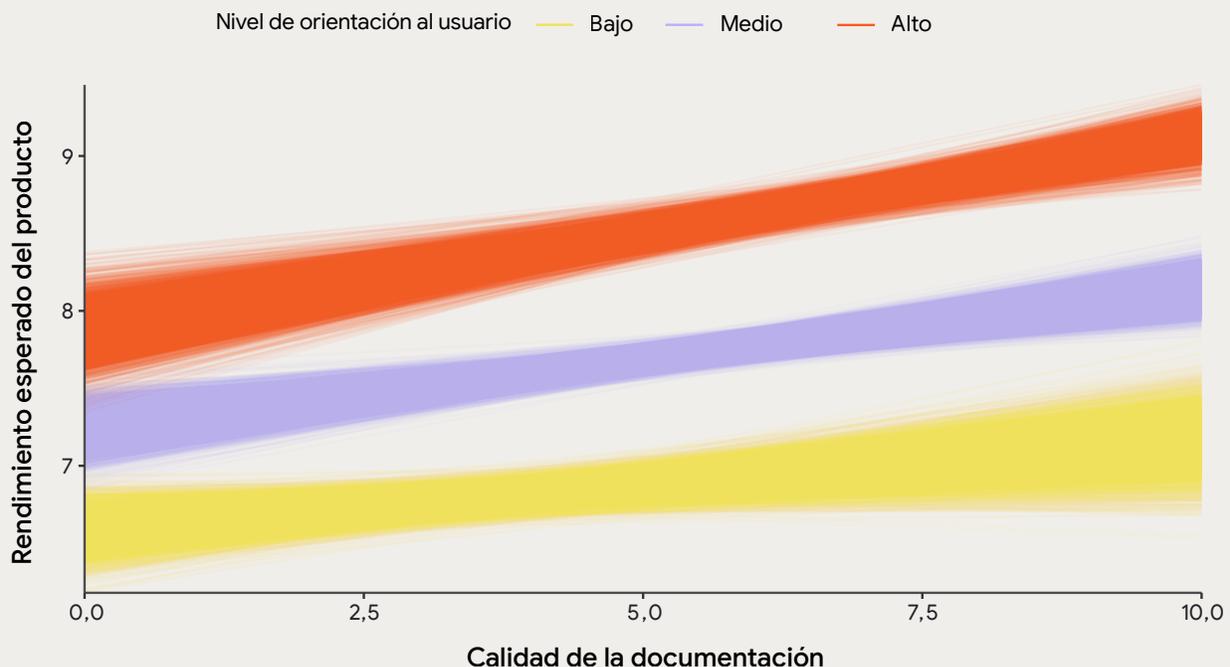
La combinación de una buena documentación y un enfoque de desarrollo de software orientado al usuario da muy buenos resultados.

Los equipos que se centran en los usuarios ven aumentar el rendimiento del producto. Al combinar este enfoque en el usuario con un entorno donde prima la documentación interna de calidad, el aumento del rendimiento del producto se multiplica (consulta la figura 16). Este hallazgo es similar al comportamiento que vemos cuando la documentación amplía el impacto de una competencia técnica en el rendimiento empresarial.⁷

La documentación ayuda a propagar las indicaciones y los comentarios de los usuarios por todo el equipo, y los lleva hasta el propio producto.

Hemos visto que la documentación interna no afecta significativamente al rendimiento esperado del producto sin las indicaciones de los usuarios. Sin embargo, si un equipo tiene una documentación interna de alta calidad, las indicaciones de los usuarios incluidas en ella tendrán un mayor impacto en el rendimiento del producto.

Empezamos a analizar la documentación en 2021 y todos los años vemos el amplio impacto de la documentación de calidad. Este año, nuestros hallazgos suman el impacto de la documentación interna a las predicciones de rendimiento del producto.



El gráfico incluye 12.000 líneas de simulaciones con las que intentamos estimar el patrón más plausible
Figura 16: Rendimiento del producto y calidad de la documentación según 3 niveles de orientación hacia el usuario

Cultura de la documentación

El Manifiesto Ágil defiende que "es más importante que el software funcione que tener una documentación exhaustiva"⁷ No obstante, seguimos viendo que la documentación de calidad es un componente fundamental para que el software funcione.

Es posible que la expresión "documentación exhaustiva" haga referencia a prácticas inadecuadas en las que puede incluirse la documentación. La documentación problemática puede ser, entre otras, aquella creada solo con fines burocráticos o para disimular la falta de confianza entre directivos y empleados. En este tipo de cultura también se incluyen prácticas como escribir documentos, pero nunca sobre cómo mantenerlos ni afianzarlos.

En estos casos, seguramente consideraríamos que se trata de documentación de baja calidad. Este tipo de contenido se escribe para la audiencia incorrecta, por lo que no tiene el mismo efecto al intentar usarlo durante el trabajo. Además, el exceso de documentación puede suponer un problema de la misma magnitud que su escasez.

Para evaluar la calidad de la documentación nos basamos en atributos como su buscabilidad y fiabilidad. Es importante recordar que, en la documentación interna, la audiencia principal son los compañeros e incluso quien la redacta cuando necesite llevar a cabo tareas específicas en el futuro.⁸ Los equipos con una buena cultura de la documentación se centran en las necesidades de estos lectores. Es otra forma de centrarse en lo que importa a los usuarios.

Puedes crear tu propia cultura de la documentación siguiendo las prácticas que hemos visto para crear documentación de calidad, como:

Documentar casos prácticos fundamentales.

Recibir formación en escritura técnica.

Definir las responsabilidades y procesos relacionados con la actualización de la documentación.

Repartir el trabajo de documentación entre distintos miembros del equipo.

Mantener la documentación como parte del ciclo de vida del desarrollo de software.

Eliminar la información obsoleta o redundante.

Reconocer el trabajo de documentación en las revisiones de rendimiento y los ascensos.

Los peligros de cambiar constantemente de prioridades

Sabemos lo que se siente. Te has pasado los últimos meses trabajando en una función nueva. Eres consciente de lo que necesitan tus usuarios, y eso te motiva y te ayuda a centrarte en la tarea. De repente (o eso parece), el equipo directivo decide cambiar las prioridades de la empresa. Ahora no está claro si paralizarán tu proyecto, lo fragmentarán, le sumarán capas o lo convertirán en otra cosa.

Esta experiencia tan habitual puede tener grandes repercusiones para los empleados y las empresas. Aquí echamos un vistazo a lo que pasa cuando las empresas están constantemente cambiando de prioridades.

Nuestros resultados y lo que implican

Nuestros hallazgos muestran que, cuando las empresas tienen prioridades inestables, generalmente la productividad se reduce de forma leve pero significativa, mientras que el desgaste aumenta bastante.

Nuestros datos apuntan a que es difícil mitigar ese mayor desgaste, así que estudiamos si tener buenos líderes, documentos internos de calidad y un enfoque de desarrollo de software orientado al usuario podría ayudar a contrarrestar el efecto que tiene el cambio de prioridades en el desgaste.

La respuesta es que no. Aunque una empresa presente todos esos aspectos positivos, si las prioridades son inestables, todavía existirá el riesgo de que los empleados sufran desgaste.

¿Por qué la inestabilidad de prioridades de la empresa afecta negativamente al bienestar de sus empleados?

Tenemos la hipótesis de que la inestabilidad de las prioridades de una empresa incrementa el desgaste de sus empleados porque crea expectativas confusas, disminuye su sensación de control y aumenta sus cargas de trabajo.

La verdad es que no creemos que el problema sea que las prioridades cambien, ya que los objetivos de negocio y la dirección de los productos están cambiando constantemente y la capacidad de adaptar las prioridades puede ser algo bueno para una empresa.

Creemos que lo que afecta negativamente al bienestar de los empleados es la frecuencia con la que cambian las prioridades. La incertidumbre que acompaña a esa inestabilidad de las prioridades implica un problema crónico relativo a la frecuencia con la que cambian las prioridades.

Décadas de investigación académica han demostrado los efectos perjudiciales del estrés crónico en la salud y el bienestar,⁹ y hemos visto paralelismos entre la investigación sobre el estrés crónico y nuestros hallazgos. La inestabilidad crónica provoca incertidumbre y reduce la percepción de control: la combinación perfecta para el desgaste.

¿Qué ocurre cuando se estabilizan las prioridades?

Nuestros hallazgos son algo enrevesados. Hemos descubierto que, cuando se estabilizan las prioridades, disminuye el rendimiento del envío de software. Se vuelve más lento y menos estable.

Nuestra hipótesis es que puede deberse a que, como es más probable que las empresas con prioridades estables tengan productos y servicios que funcionen bien, los cambios se hacen con menos frecuencia. También puede ser que la estabilidad de las prioridades disminuya los envíos y derive en que haya lotes más grandes de lo recomendado.

Sea cual sea la razón, nos parece sorprendente. ¿Por qué crees que estabilizar las prioridades de la empresa reduce la velocidad y la estabilidad del envío de software?

Crear IA para usuarios finales aporta estabilidad a las prioridades, pero no al envío.

Incorporar experiencias basadas en IA para los usuarios finales estabiliza las prioridades de la empresa. Parece una forma rimbombante de recomendar la IA, pero no interpretamos que este hallazgo nos diga algo significativo sobre la propia IA.

Por el contrario, creemos que cambiar las iniciativas hacia la adopción de la IA aporta claridad y sirve de guía para las empresas. Esta claridad, y no la IA, es lo que estabiliza las prioridades de la empresa.

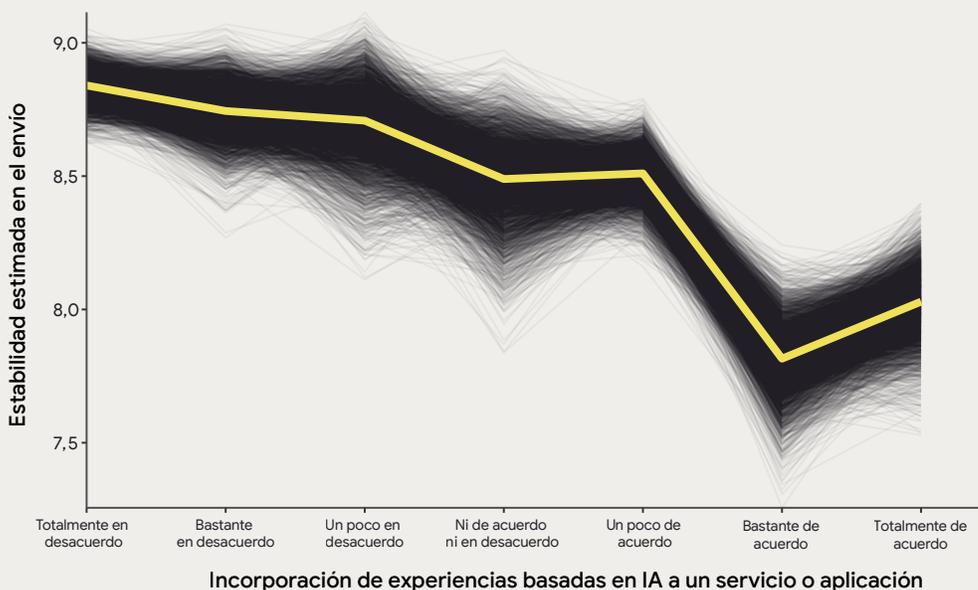
Merece la pena subrayarlo porque explica un poco lo que pasa en las empresas cuando aparecen nuevas tecnologías. Las nuevas tecnologías traen consigo cambios, y las empresas necesitan tiempo para adaptarse.

Es probable que ese periodo desestabilice las prioridades, ya que los líderes tienen que averiguar cuál es la mejor estrategia para la empresa. Cuando las aguas vuelven a su cauce, y las empresas tienen claros sus próximos pasos, las prioridades empiezan a estabilizarse.

No obstante, estabilizar las prioridades no se traduce inmediatamente en la estabilización del proceso de envío de software.

Nuestros análisis muestran que incorporar experiencias basadas en IA en un servicio o aplicación trae consigo dificultades y quebraderos de cabeza.

Hemos visto que los equipos que han dado el paso han experimentado una importante reducción del 10 % en la estabilidad del envío de software respecto a quienes no lo han hecho. Aquí representamos el reto.



*Cada línea es una de las 4000 simulaciones que hemos realizado para intentar estimar el patrón más plausible

Figura 17: La estabilidad en el envío de software en relación con la incorporación de experiencias basadas en IA en un servicio o aplicación

¿Qué pueden hacer las empresas?

La respuesta, aunque es sencilla, quizás no sea tan simple. Basándonos en nuestros hallazgos, recomendamos a las empresas que se centren en estabilizar sus prioridades, ya que se trata de una forma segura de contrarrestar los efectos negativos de las prioridades inestables, que contribuyen al desgaste de los empleados.

Según nuestros hallazgos, esos efectos negativos no desaparecen aunque haya buenos líderes, documentación de calidad y un enfoque de desarrollo de software orientado al usuario. Esto nos hace pensar que, además de crear estabilidad, no hay mucho más que las empresas pueden hacer para evitar el desgaste, más allá de encontrar formas de 1) estabilizar las prioridades y 2) asegurarse de que el día a día de sus empleados no se ve afectado por el cambio constante de prioridades.

-
1. <https://www.nngroup.com/articles/bridging-the-designer-user-gap/>
 2. <https://executiveeducation.wharton.upenn.edu/thought-leadership/wharton-at-work/2024/03/creating-meaning-at-work/>
 3. <https://www.apa.org/pubs/reports/work-in-america/2023-workplace-health-well-being>
 4. <https://bigthink.com/the-present/harvard-business-review-americans-meaningful-work/>
 5. <https://hbr.org/2018/11/9-out-of-10-people-are-willing-to-earn-less-money-to-do-more-meaningful-work>
 6. (P[N]), por ejemplo (P1), es el seudónimo de cada entrevistado.
 7. <https://cloud.google.com/blog/products/devops-sre/deep-dive-into-2022-state-of-devops-report-on-documentation> y el informe Accelerate State of DevOps del 2023. <https://dora.dev/research/2023/dora-report>
 8. <https://agilemanifesto.org/>
 9. Hay otras audiencias, como la dirección, los organismos supervisores o los auditores.
 10. Cohen S, Janicki-Deverts D, Miller GE. Psychological Stress and Disease. JAMA. 2007;298(14):1685–1687.doi:10.1001/jama.298.14.1685

Transformación desde el liderazgo

El éxito de la transformación depende de muchos factores. Este año, hemos visto que los equipos de alto rendimiento son aquellos que priorizan la estabilidad, se centran en sus usuarios, tienen buenos líderes y elaboran documentación de calidad. Nuestra investigación señala algunas rutas útiles para ayudarte a trazar el mejor rumbo hacia la transformación.

Hemos visto que la clave del éxito está en abordar la transformación desde el punto de vista de la mejora continua. Los equipos de alto rendimiento de nuestro estudio entienden las variables que les impiden avanzar y mejoran de manera metódica y continua usando como referencia las métricas de DORA. Aunque el éxito a largo plazo exige excelencia en todos los pilares, una década de investigación apunta a cuatro formas específicas y eficaces de empezar a transformar tu empresa.



Liderazgo en el ámbito de la transformación

El liderazgo en el ámbito de la transformación es un modelo en el que los líderes apelan a los valores y la idea de propósito de los empleados, facilitando un cambio organizativo a gran escala, para inspirarlos y motivarlos a mejorar su rendimiento.

Estos líderes animan a sus equipos a trabajar para lograr un objetivo común a través de las siguientes dimensiones:¹

Visión	Tienen una visión clara de hacia dónde van tanto su equipo como la empresa.
Comunicación inspiradora	Dicen cosas positivas sobre su equipo, hacen que los empleados se sientan orgullosos de formar parte de su empresa y los animan a ver los cambios como oportunidades.
Estimulación intelectual	Animan a los miembros de su equipo a pensar de formas diferentes sobre problemas antiguos y a replantearse algunas de sus suposiciones acerca de su trabajo.
Liderazgo alentador	Tienen en cuenta los sentimientos de los demás antes de actuar y son considerados con las necesidades personales ajenas.
Reconocimiento personal	Elogian a los miembros de su equipo cuando trabajan mejor que la media y reconocen las mejoras de calidad en el trabajo de sus compañeros.

Este año, hemos descubierto que el liderazgo transformador mejora la productividad de los empleados. Hemos descubierto que aumentar el liderazgo transformador en un 25 % desencadena un aumento del 9 % en la productividad de los empleados.

Pero el liderazgo transformador no solo ayuda a mejorar la productividad, sino que también puede traer consigo:

- Un menor desgaste de los empleados
- Una mayor satisfacción laboral
- Un mayor rendimiento de los equipos
- Un mayor rendimiento de los productos
- Un mayor rendimiento de la empresa

Según nuestra investigación, en el 2017 hubo una relación estadísticamente significativa entre las cualidades de liderazgo mencionadas anteriormente y el rendimiento de TI. Los equipos de alto rendimiento tenían líderes con altas puntuaciones en las cinco características, mientras que los de bajo rendimiento tenían las puntuaciones más bajas. Además, vimos una estrecha correlación entre el liderazgo transformador y los resultados del Employee Net Promoter Score (eNPS), que mide la probabilidad de que alguien recomiende trabajar en una empresa.

Dicho esto, el liderazgo transformador por sí solo no garantiza un alto rendimiento, pero sí lo posibilita.

El liderazgo transformador resulta crucial para favorecer la adopción de competencias y prácticas técnicas y de gestión de productos. Para conseguirlo, hay que 1) delegar la autoridad y autonomía a los equipos, 2) proporcionarles las métricas e inteligencia empresarial necesarias para solucionar problemas, y 3) crear estructuras de incentivos por la materialización del valor, en vez de por el lanzamiento de funciones.

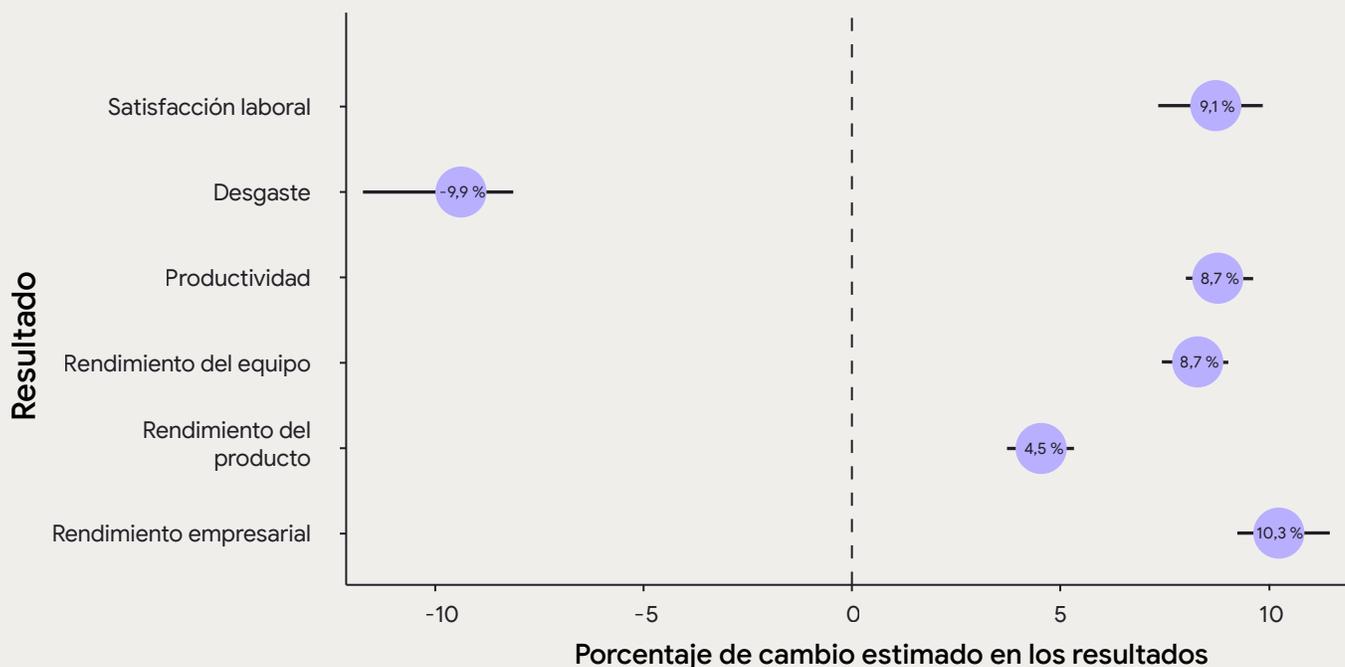
La transformación requiere tiempo y herramientas. Los líderes deben asignar los recursos pensando específicamente en la mejora. Los buenos líderes tienen un papel fundamental a la hora de dar a sus equipos el tiempo y la financiación que necesitan para avanzar. No se debe esperar que los ingenieros aprendan cosas nuevas y se formen en automatización en su tiempo libre, sino que esas tareas deben formar parte de su horario laboral.

Nuestra investigación ha contribuido a cambiar la idea de que el departamento de TI es un centro de costes: se trata de una inversión que potencia el éxito de la empresa. En el 2020, escribimos un informe sobre el retorno de la inversión de la transformación de DevOps,² en el que se incluyen cálculos que pueden ayudarte a estimar el valor potencial de la inversión en mejoras de TI.

El retorno monetario es solo una de las muchas cosas que puedes esperar obtener de esta inversión. Nuestra investigación del 2015 demostró que "la inversión de una empresa en DevOps

está estrechamente correlacionada con la cultura empresarial; la capacidad de los equipos de desarrollo, operaciones y seguridad de la información para obtener resultados satisfactorios; tasas bajas de desgaste; un liderazgo más eficaz; y la implementación adecuada de prácticas de entrega continua y gestión optimizada".³ Recomendamos dedicar específicamente una parte de la capacidad a las mejoras.

Si el liderazgo transformador aumenta en un 25 %...



Punto = valor estimado

Barra de error = intervalo de incertidumbre del 89 %

Figura 18: Impacto del liderazgo transformador en diversos resultados.

Apuesta sin miedo por el enfoque orientado al usuario

La investigación de este año demuestra que las empresas con buenos líderes y centradas en crear software que aborde las necesidades de los usuarios desarrollan mejores productos: es una combinación ganadora. Cuando el usuario está en el centro del desarrollo de software, los líderes pueden articular una visión clara.

El objetivo final es crear productos que los usuarios adoren. Como vimos en el capítulo [Experiencia de los desarrolladores](#), centrarse en el usuario da una razón de ser a las funciones del producto. Así, los desarrolladores las crean con la confianza de saber que mejorarán la experiencia de usuario.

Vemos que los equipos que muestran un fuerte deseo de entender y satisfacer las necesidades de los usuarios, y que cuentan con los mecanismos necesarios para recopilar, monitorizar y responder a sus comentarios, presentan los niveles más altos de rendimiento empresarial. De hecho, las empresas pueden tener éxito incluso sin altos niveles de velocidad y estabilidad del software, siempre que se centren en el usuario. En el 2023, descubrimos que los equipos orientados al usuario tienen un rendimiento empresarial un 40 % más alto en comparación con los que no.⁴ También vimos que se producía una mejora en este sentido en el 2016.

La investigación de este año repite los hallazgos anteriores. Los equipos que se centran en el usuario hacen mejores productos.

Además de que los productos mejoran, los empleados están más satisfechos con sus puestos y es menos probable que sufran desgaste.

El envío de software rápido y estable proporciona más oportunidades de experimentación y aprendizaje a las empresas, y lo ideal sería que esos experimentos e iteraciones se basaran en los comentarios de los usuarios. El envío de software rápido y estable permite experimentar, entender mejor las necesidades de los usuarios y responder rápidamente si esas necesidades no se están cumpliendo.

Incorporar la velocidad y la estabilidad en el envío también permite adaptarse con más facilidad a los cambios en el mercado o a la competencia.

Es importante recordar que tus desarrolladores internos son también usuarios. Las plataformas internas para desarrolladores (IDPs) permiten que las empresas aporten valor a los desarrolladores, quienes, a su vez, aportan valor a los usuarios externos o a otros usuarios internos.

Nuestra investigación demuestra que las IDPs verdaderamente útiles son las que se desarrollan como productos y están orientadas al usuario para ofrecer una experiencia que permita a los desarrolladores trabajar con independencia. Una IDP desplegada de esta forma fomenta el aumento de la productividad de las personas, los equipos y la empresa.

Básate en los datos para las decisiones de tu empresa

Es fundamental tener la capacidad de visualizar los progresos para llegar al éxito. En los últimos 10 años, hemos defendido que las empresas deben basarse en datos para tomar decisiones. Las cuatro métricas clave de DORA⁵ se han convertido en un estándar internacional para medir el rendimiento del envío de software, pero eso solo es una parte de la historia. Hemos identificado más de 30 competencias y procesos⁶ que pueden usarse para impulsar mejoras en las empresas.

El valor de las métricas está en su capacidad para indicarte si estás mejorando. Las cuatro métricas clave se deben usar a nivel de aplicación y servicio, no de empresa ni de línea de negocio. Utiliza las métricas para identificar tus esfuerzos de mejora continua, no para comparar unos equipos con otros ni a unos empleados con otros.

Las métricas tampoco deben usarse como modelo de madurez para equipos de aplicaciones o servicios. Tener un rendimiento bajo, medio, alto o de élite es interesante, pero es un indicador que debe tomarse con cautela, ya que se trata de simples denominaciones con poco valor en el contexto de tu proceso de transformación.

A medida que nuestra investigación va avanzando y evolucionando, te animamos a ir más allá de las cuatro claves. Ha quedado claro que las métricas relacionadas con los comentarios de los usuarios son tan

importantes como estas cuatro métricas clave, y creemos que esto se debe a que la mayoría de los equipos han ideado soluciones factibles para mejorar la velocidad y la estabilidad. Como resultado, las ventajas obtenidas con la velocidad y la estabilidad disminuyen, puesto que el alto rendimiento se ha propagado notablemente.

Viendo la transformación de manera holística, recomendamos crear paneles y representaciones gráficas que combinen métricas técnicas (como nuestras cuatro métricas clave y las métricas de fiabilidad) y empresariales, ya que eso permite reducir las distancias entre las iniciativas de transformación descendentes y ascendentes. También ayuda a conectar la guía central, los objetivos y resultados clave, y los objetivos de los empleados con las inversiones en TI, y te puede servir para cuantificar el retorno de inversión.

Creemos que las métricas son esenciales para alcanzar la excelencia. Las métricas facilitan la toma de decisiones. Cuantas más métricas recopiles, cuantitativas y cualitativas, mejores y más fundamentadas serán las decisiones que tomes. Cada persona tendrá su criterio sobre el valor o el sentido de los datos, pero suele ser preferible usarlos como base para tomar una decisión que confiar en una opinión o intuición.



Migra por completo a la nube o quédate en el centro de datos

Llevamos desde el 2018 investigando la relación entre las 5 características del cloud computing definidas por el NIST⁷ (autoservicio bajo demanda, amplio acceso a la red, agrupamiento de recursos, elasticidad rápida y servicio medido o infraestructura flexible) y el rendimiento empresarial. Hemos visto que los equipos de mayor éxito tienen más probabilidades de aprovechar las ventajas de la infraestructura flexible que los que van más rezagados.

El año pasado, nuestra investigación nos dejó el resultado más llamativo hasta la fecha sobre esta cuestión: usar la nube sin aprovechar las ventajas de las cinco características puede ser perjudicial

y reducir el rendimiento empresarial. Si las empresas no están dispuestas a transformar por completo su aplicación o servicio, es mejor que se queden en el centro de datos. Por supuesto, para lograrlo, no basta con adoptar herramientas o tecnologías sin más. A menudo se trata de un cambio total de paradigma a la hora de diseñar, crear, desplegar y ejecutar aplicaciones. Hacer cambios a gran escala es más fácil si se empieza con un pequeño número de servicios, por lo que recomendamos un enfoque iterativo que ayude a los equipos y a las empresas a aprender y mejorar conforme vayan avanzando.

Resumen

Llevamos más de 10 años viendo sistemáticamente que la transformación es esencial para el éxito. Y lo que muchas empresas no entienden es que la transformación no es un destino, sino un proceso de mejora continua.⁸ Nuestra investigación lo deja claro: las empresas que no están continuamente mejorando se acaban quedando atrás. Del mismo modo, las que adoptan una mentalidad de mejora continua muestran los mayores niveles de éxito.

Es importante ser conscientes de que habrá algunos obstáculos y cierta incomodidad a lo largo del proceso. Nuestra investigación muestra una pequeña caída del rendimiento seguida de muchas ganancias (lo que también se conoce como "curva en J") con DevOps,⁹ SRE¹⁰ y, este año, la ingeniería de plataformas. Es algo normal y, si sigues mejorando de manera continua, la situación se enderezará y llegarás al otro extremo del proceso en mejor forma que cuando empezaste.

La idea de un proceso que nunca acaba puede parecer abrumadora. Es fácil atascarse en la planificación o en diseñar la transformación perfecta. La clave del éxito está en remangarse y ponerse a trabajar. El objetivo de la empresa y de tus equipos debería ser, sencillamente, ser un poco mejores que el día anterior. El objetivo de nuestros últimos 10 años de investigación es el mismo que tenemos para el futuro: ayudarte a mejorar mejor.

1. Dimensions of transformational leadership: Conceptual and empirical extensions - Rafferty, A. E., y Griffin, M. A.

2. The ROI of DevOps Transformation. <https://dora.dev/research/2020/>

3. Informe State of DevOps del 2015. <https://dora.dev/research/2015/2015-state-of-devops-report.pdf#page=25>

4. Informe Accelerate State of DevOps del 2023. <https://dora.dev/research/2023/dora-report/2023-dora-accelerate-state-of-devops-report.pdf#page=17>

5. Las cuatro métricas clave de DORA. <https://dora.dev/guides/dora-metrics-four-keys/>

6. Competencias y procesos de DORA. <https://dora.dev/capabilities/>

7. Las 5 características del cloud computing definidas por el NIST. <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf>

8. Proceso de mejora continua. <https://cloud.google.com/transform/moving-shields-into-position-organizing-security-for-digital-transformation>

9. Informe Accelerate State of DevOps del 2018. <https://dora.dev/research/2018/dora-report/>

10. Informe State of DevOps del 2022. <https://dora.dev/research/2022/dora-report/>

Diez años de DORA



Historia

El movimiento DevOps surgió en el 2009 a partir de dos eventos de la misma temática, pero sin ninguna otra relación. En junio de ese año, John Allspaw y Paul Hammond dieron una charla en la Velocity Conference titulada "10+ deploys per day: Dev & ops cooperation at Flickr" (Más de 10 despliegues al día: la cooperación entre departamentos de desarrollo y operaciones en Flickr).¹ Unos meses después, Patrick Debois retomó el tema al dirigir un equipo de voluntarios para organizar el primer evento DevOpsDays en Gante (Bélgica).²

No pasó mucho tiempo hasta que la comunidad de DevOps quiso saber cómo estaba evolucionando el concepto. Alana Brown, que trabajaba en Puppet Labs, realizó una encuesta en el 2011 para saber más sobre DevOps. Esta encuesta sirvió para confirmar que "trabajar siguiendo prácticas de DevOps se está convirtiendo en una nueva forma de hacer negocios de TI".

Conforme el movimiento siguió creciendo hasta abarcar nuevos sectores y empresas, Alana aprovechó su éxito y colaboró con IT Revolution Press para realizar otra encuesta en el año 2012, cuyos resultados se publicaron en el informe State of DevOps del 2013.³

Al año siguiente, la Dra. Nicole Forsgren se unió al equipo de investigación, aportando más rigor científico al programa. En el informe State of DevOps del 2014⁴ ya se estableció la conexión entre el rendimiento del envío de software y el rendimiento empresarial, y se descubrió que "la

capitalización de mercado de las empresas que cotizaban en bolsa con equipos de TI de alto rendimiento creció un 50 % más en tres años que la de aquellas con equipos de TI de bajo rendimiento".

La tendencia de hacer informes anuales se consolidó en el 2016, y Forsgren, Jez Humble y Gene Kim fundaron DevOps Research and Assessment (DORA). En el informe State of DevOps de ese año se incluyeron cálculos para ayudar a cuantificar las inversiones hechas por los equipos que adoptaron prácticas de DevOps. Este trabajo se amplió en el informe The ROI of DevOps Transformation⁵, publicado en el 2020.

*En el 2017, IT Revolution Press publicó Accelerate: The science behind devops: Building and scaling high performing technology organizations,*⁶ escrito por Forsgren, Humble y Kim. En este libro se resumen los primeros años del programa de investigación y se hace hincapié en las competencias que fomentan las mejoras.

DORA publicó como empresa un informe independiente en el año 2018 titulado "Accelerate State of DevOps: Strategies for a New Economy".⁷ El equipo de Puppet siguió publicando una serie de informes,⁸ sin relación con DORA, desde ese mismo año.

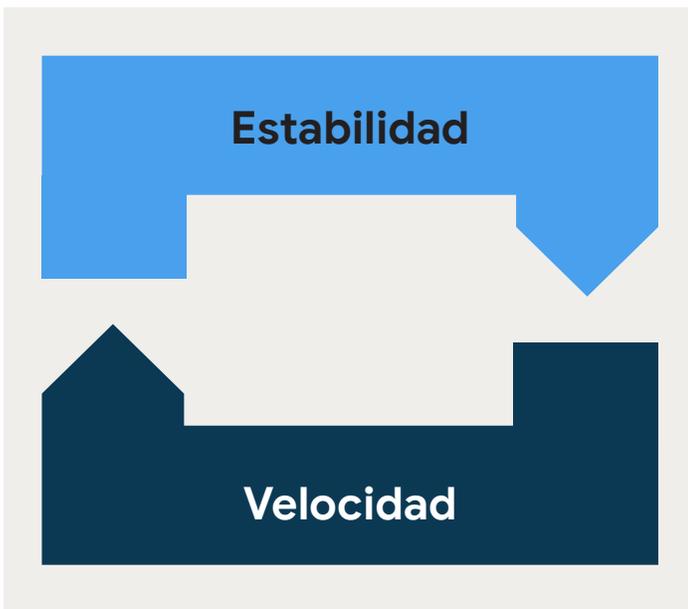
A finales del 2018, Google Cloud adquirió DORA⁹ para continuar la investigación científica independiente de la plataforma. Este año celebramos el décimo informe de DORA,¹⁰ y nos encanta compartir nuestros hallazgos con el mundo. ¡Gracias por leerlo!

Datos clave de DORA

Los equipos no tienen que sacrificar la velocidad en favor de la estabilidad

Los equipos tecnológicos siempre necesitan formas de medir el rendimiento para poder evaluar su situación actual, saber qué mejoras priorizar y validar su progreso. DORA ha identificado y validado cuatro métricas de envío de software (las cuatro claves) para medir de forma eficaz los resultados del proceso de envío de software. Estas métricas del rendimiento del envío de software se han convertido en un estándar del sector.

La investigación ha demostrado que la velocidad y la estabilidad de los cambios suelen ir de la mano; y hemos visto equipos de todas las verticales del sector que logran altos niveles en ambos.



Los equipos pueden medir las cuatro claves de muchas formas, incluidas las siguientes:

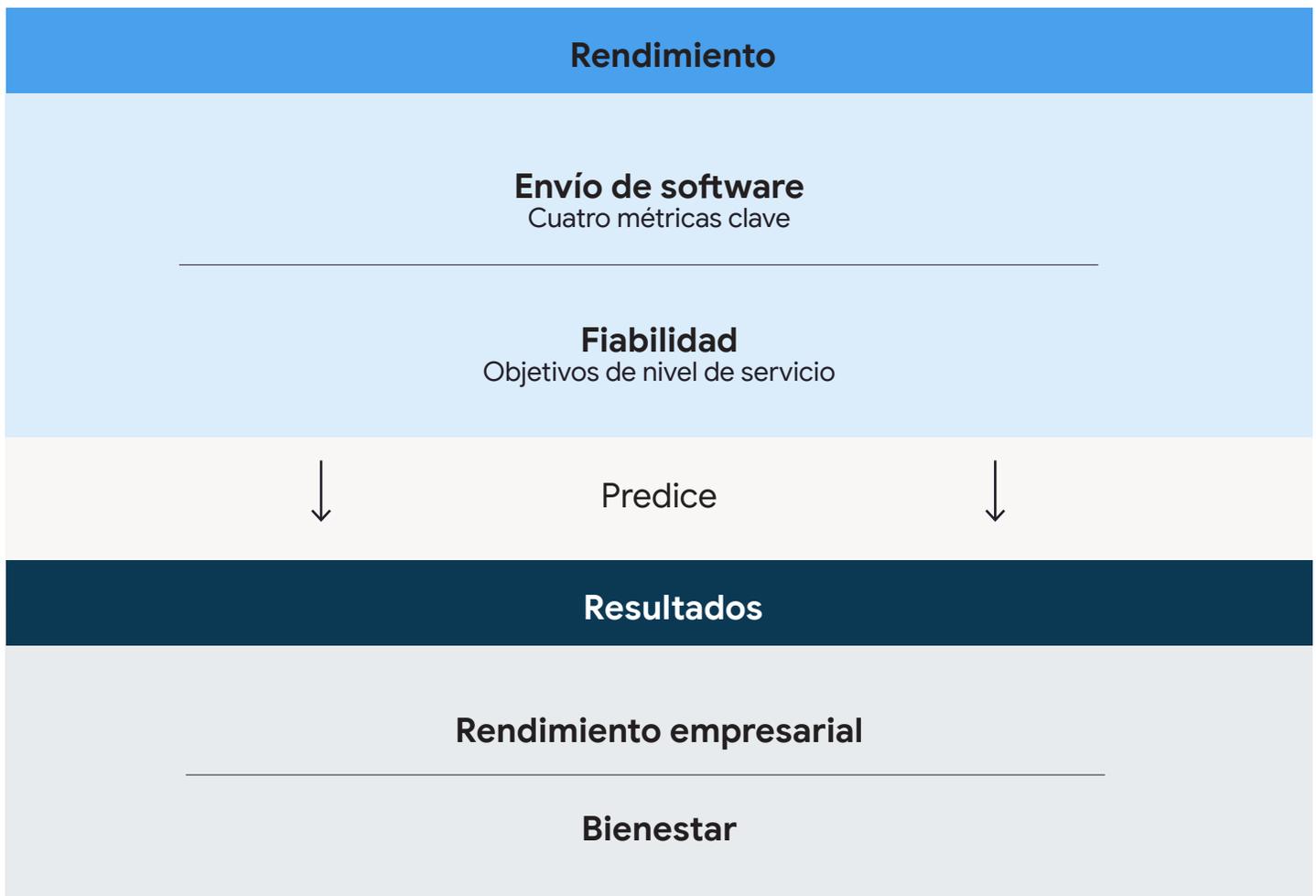
- Conversaciones y reflexiones durante las reuniones de equipo
- La encuesta breve de DORA (<https://dora.dev/quickcheck>)
- Herramientas comerciales y de código disponibles¹¹ en la categoría de inteligencia de ingeniería de software (SEI)
- Integraciones creadas a medida para herramientas específicas que usa un equipo

El envío de software y el rendimiento operativo fomentan el rendimiento empresarial

DORA usa las cuatro claves para medir el rendimiento del envío de software. El rendimiento operativo se estudió por primera vez en DORA en el 2018 y mide la capacidad para hacer y mantener promesas y afirmaciones sobre un producto o servicio de software.

Los mejores resultados se obtienen cuando se combinan el envío de software y el rendimiento operativo para fomentar el rendimiento empresarial y el bienestar de los empleados.

Los profesionales que trabajan en equipos tecnológicos conocen la importancia de reducir la fricción en el proceso de envío al tiempo que satisfacen las expectativas de fiabilidad de los usuarios de la aplicación.





La cultura es esencial para el éxito

Uno de los indicadores más claros del rendimiento es la cultura de la empresa, y hemos visto de forma continuada el potencial de una cultura de alta confianza en la que se fomenta un entorno de aprendizaje y colaboración. Por ejemplo, nuestra investigación del 2022 demostró que la cultura era el mayor indicador de las prácticas de seguridad en el desarrollo de aplicaciones de una empresa.¹²

La cultura afecta a cada aspecto de nuestra investigación, es multifacética y está en constante evolución. A lo largo de los años, hemos usado muchas métricas diferentes inspirándonos en investigaciones como la tipología de la cultura organizativa de Westrum.¹³ En nuestras métricas de bienestar hemos incluido el desgaste, la productividad y la satisfacción laboral.

Mejorar mejor

Animamos a los equipos a marcarse el objetivo de mejorar mejor. Hacer progresos requiere adoptar una mentalidad y una práctica de mejora continua, para lo cual es necesario contar con algún método que permita evaluar el rendimiento en ese momento, priorizar las mejoras e incorporar mecanismos que ayuden a medir los avances.

Abordar la mejora desde un enfoque experimental implica que habrá una combinación de victorias y fracasos, pero, en ambos casos, los equipos pueden adoptar medidas significativas como resultado de las lecciones aprendidas.

La próxima década

En conjunto, hemos aprendido mucho los unos de los otros durante la última década. Gracias por participar en nuestras encuestas anuales y en la [comunidad de prácticas de DORA](#),¹⁴ y por trabajar con DORA en tu propio contexto.

A medida que el panorama tecnológico sigue cambiando, el equipo de DORA continuará investigando las competencias y prácticas que ayudan a triunfar a las empresas y equipos tecnológicos. Continuaremos priorizando la parte humana de la tecnología y nos comprometemos a seguir publicando investigaciones que sean independientes de la plataforma para que las uses como guía en tu propio proceso.

Muchas de nuestras estadísticas antiguas siguen siendo lo suficientemente relevantes como para fundamentar tu enfoque sobre las tecnologías y prácticas emergentes, pero también nos entusiasma descubrir información nueva junto a nuestros lectores.

Tenemos un compromiso con los principios fundamentales que siempre han formado parte del movimiento DevOps: cultura, colaboración, automatización, aprendizaje y uso de la tecnología para lograr objetivos de negocio. Al incluir los puntos de vista de personas con diferentes roles, incluidas aquellas que quizás no parezcan tan relacionadas con la etiqueta "DevOps", nuestra comunidad y nuestra investigación se ven beneficiadas. También verás que el término "DevOps" va perdiendo protagonismo.

En el informe de este año nos centramos mucho en el uso y los impactos de la inteligencia artificial (IA). Como has visto, la adopción es cada vez mayor y hay mucho margen para la experimentación en este sentido. En el futuro, seguiremos investigando la IA, además de otras tecnologías y prácticas emergentes. Usa nuestras antiguas investigaciones junto con los hallazgos recientes para impulsar la adopción y mejorar la experiencia de todos los miembros de tu equipo.

1. Diapositivas: <https://www.slideshare.net/jallspaw/10-deploys-per-day-dev-and-ops-cooperation-at-flickr>; grabación: <https://www.youtube.com/watch?v=LdOe18KhtT4>

2. <https://legacy.devopsdays.org/events/2009-ghent/>

3. <https://www.puppet.com/resources/history-of-devops-reports#2013>

4. Informe State of DevOps del 2014. <https://dora.dev/research/2014/>

5. The ROI of DevOps Transformation. <https://dora.dev/research/2020/>

6. Forsgren, Nicole, Jez Humble y Gene Kim. 2018. Accelerate: The Science Behind DevOps: Building and Scaling High Performing Technology Organizations. IT Revolution Press.

7. Accelerate State of DevOps: Strategies for a New Economy. <https://dora.dev/research/2018/dora-report/>

8. <https://www.puppet.com/resources/history-of-devops-reports#2018>

9. <https://dora.dev/news/dora-joins-google-cloud>

10. Consideramos que el primer informe de DORA se publicó en el 2014, el año en el que la Dra. Forsgren se unió al programa, a pesar de que DORA en sí se fundó varios años después. En el 2020, el informe no se publicó, por lo que el del año 2024 es el décimo.

11. <https://dora.dev/resources/#source-available-tools>

12. Informe Accelerate State of DevOps del 2022. <https://dora.dev/research/2022/dora-report/>

13. Ron Westrum, "A typology of organisation culture", BMJ Quality & Safety 13, número 2(2004), doi:10.1136/qshc.2003.009522

14. <https://dora.community>

Reflexiones finales

En la última década, DORA se ha consolidado como una fuente fiable de investigación, estadísticas e información. Mientras se sigan adoptando nuevas prácticas y tecnologías en el sector, como la ingeniería de plataformas o la inteligencia artificial, DORA continuará investigando nuevas formas de ayudar a los equipos a mejorar. Gracias por haber confiado en DORA durante todo este tiempo.

Replicar nuestra investigación

El área de investigación y los hallazgos del informe de este año son complejos y, en ocasiones, poco claros o incluso contradictorios. Te animamos a replicar nuestra investigación y a centrarte en un solo equipo u empresa para profundizar más.

Realiza experimentos dentro de tu empresa

Los resultados de DORA se pueden convertir en las hipótesis de tus próximos experimentos. Obtén más información sobre el rendimiento de tu equipo e identifica áreas de mejora, que podrían inspirarse en los hallazgos del programa de investigación DORA.

Haz encuestas dentro de tu empresa

Inspírate en este informe y en las preguntas de la encuesta de este año¹ para diseñar tu propia encuesta interna. Puedes incorporar preguntas más detalladas a tu encuesta si las consideras pertinentes para tu audiencia.² Lee el capítulo [Metodología](#) para obtener más información sobre cómo hacemos nuestra investigación. Céntrate en poner en práctica tus hallazgos.

Comparte lo que aprendas

Difunde lo que vayas aprendiendo por toda la empresa. Para compartir lo aprendido, puedes usar desde informes formales para grandes audiencias hasta métodos más informales, como comunidades de práctica o charlas entre compañeros. Prueba diferentes enfoques y descubre cuál funciona mejor en tu contexto y cultura. Esto también es un proceso experimental.



¿Cómo estás aplicando lo aprendido en esta investigación?

Únete a la comunidad DORA para compartir tus experiencias, aprender de los demás e inspirarte en los compañeros con los que recorres el camino hacia la mejora continua: <https://dora.community>.



¹ Encuesta del 2024. <https://dora.dev/research/2024/questions/>

² Experiencias de hacer las encuestas de DORA de manera interna en empresas de software. <https://www.infoq.com/news/2024/08/dora-surveys-software-company/>

Agradecimientos

Este año se cumple un hito especial: el décimo informe de DORA. Queremos darles las gracias a todos los investigadores, expertos, profesionales, líderes y agentes transformadores que han contribuido a dar forma a este corpus de trabajo y han ido evolucionando a la par que nosotros.

Han pasado muchas cosas desde que IT Revolution Press y Puppet Labs publicaron el primer informe State of DevOps y nos sentimos profundamente agradecidos con los fundadores de DORA por allanarnos el camino. Nos parece importante reflexionar sobre todo lo que ha cambiado desde entonces y cuánto hemos aprendido a lo largo de estos años.

Estamos muy agradecidos con todas las personas que han participado en la publicación de este año. Guiar e influir en las prácticas del sector es una enorme responsabilidad, y estas contribuciones tienen un valor incalculable.

Gracias a todos los que han formado parte de esta experiencia, desde el principio hasta la emocionante era de la IA. Ese apoyo y la valiosa información aportada han sido decisivos. ¡A por la próxima década de descubrimientos y colaboración!

Equipo de elaboración del informe de DORA

James Brookbank

Kim Castillo

Derek DeBellis

Benjamin Good

Nathen Harvey

Michelle Irvine

Amanda Lewis

Eric Maxwell

Steve McGhee

Allison Park

Dave Stanke

Kevin Storer

Daniella Villalba

Editor

Seth Rosenblatt

Voluntarios de localización

Andrew Anolasco

Mauricio Meléndez

Marie-Blanche Panthou

Miguel Reyes

Yoshi Yamaguchi

Jinhong Yu

Guías de DORA

Lisa Crispin

Steve Fenton

Denali Lumma

Betsalel (Saul) Williamson

Asesores y expertos en la materia

John Allspaw

Birgitta Böckeler

Sander Bogdan

Michele Chubirka

Thomas De Meo

Jessica DeVita

Rob Edwards

Dra. Nicole Forsgren

Gene Kim y
IT Revolution

Dra. Laura Maguire

James Pashutinski

Ryan J. Salva

Majed Samad

Harini Sampath

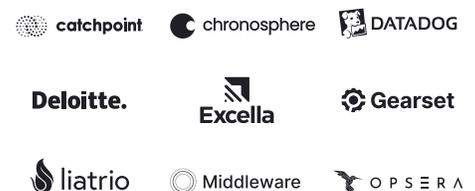
Robin Savinar

Sean Sedlock

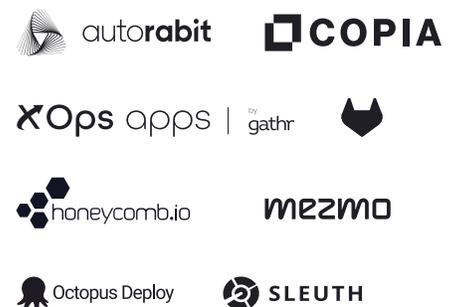
Dustin Smith

Finn Toner

Patrocinadores Gold



Patrocinadores Silver



Autores



Derek DeBellis

Derek es investigador de experiencia de usuario cuantitativa en Google y jefe de la investigación de DORA. Lleva a cabo investigaciones por medio de encuestas, analiza registros y estudia formas de medir conceptos que demuestren que un producto o una función aporta un gran valor a las personas. También ha publicado artículos de investigación sobre la interacción entre los humanos y la inteligencia artificial, el impacto del COVID-19 en el abandono del tabaquismo, el diseño para abordar los errores de PLN, el papel de la experiencia de usuario en los debates sobre la privacidad, la cultura de equipo y la relación de la IA con el bienestar y la productividad de los empleados. En su actual investigación extracurricular analiza formas de simular la propagación de las creencias y el poder.



Kevin M. Storer

El Dr. Kevin M. Storer es investigador de la experiencia de los desarrolladores en Google, donde lidera la investigación cualitativa del equipo de DORA. Su experiencia profesional en ingeniería de software y su formación multidisciplinar de posgrado en ciencias sociales y humanidades le han servido para llevar desde el 2015 liderando estudios sobre desarrolladores de software en los que se centra en el ser humano y abarca un conjunto de contextos problemáticos, perfiles de participantes y métodos de investigación muy diverso. Diversas revistas científicas de prestigio han publicado sus investigaciones sobre temas como inteligencia artificial, extracción de información, sistemas incrustados, lenguajes de programación, computación ubicua o diseño de la interacción.



Amanda Lewis

Amanda Lewis es la jefa de desarrollo de **DORA.community** e ingeniera de relaciones con desarrolladores de Google Cloud y ha dedicado su carrera profesional a crear conexiones entre desarrolladores, operadores, responsables de producto, gestores de proyectos y líderes. Ha trabajado en equipos que desarrollaban plataformas de comercio electrónico, sistemas de gestión de contenido y herramientas de observabilidad, y ha prestado apoyo a desarrolladores. Esas conexiones y conversaciones mejoran la satisfacción de los clientes y los resultados de la empresa. Amanda aporta su experiencia y empatía a la hora de ayudar a los equipos a entender e implementar el envío de software y las prácticas de inteligencia artificial.



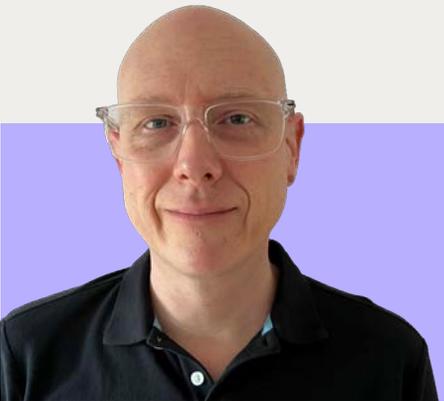
Benjamin Good

Ben Good es arquitecto de soluciones en la nube de Google. Le apasiona mejorar las prácticas de envío de software a través de tecnologías de nube y automatización. Como arquitecto de soluciones, ofrece asesoramiento sobre arquitecturas, publica guías técnicas y hace contribuciones al software libre para ayudar a los clientes de Google Cloud a solucionar problemas. Antes de unirse a Google, Ben dirigió las operaciones de nube de diversas empresas en la zona de Denver/Boulder, donde implementó prácticas de DevOps.



Daniella Villalba

Daniella Villalba es investigadora de la experiencia del usuario en Google, donde utiliza la investigación por medio de encuestas para comprender los factores que mejoran la satisfacción y la productividad de los desarrolladores. Antes de trabajar en Google, Daniella estudió las ventajas de aprender meditación y los factores psicosociales que afectan a las experiencias de los estudiantes universitarios. Tiene un doctorado en Psicología Experimental por la Universidad Internacional de Florida.



Eric Maxwell

Eric Maxwell dirige las prácticas de transformación de DevOps en Google, donde asesora a las empresas más importantes del mundo sobre cómo mejorar al ofrecer valor de forma más rápida. Durante la primera mitad de su carrera, trabajó como ingeniero, automatizando de todo y abogando por la empatía hacia otros profesionales. Colaboró en la creación del Cloud Application Modernization Program (CAMP) de Google y es miembro del equipo de DORA. Antes de trabajar en Google, Eric estuvo cocinando maravillas con otros tíos ingeniosos en Chef Software.



Kim Castillo

Kim Castillo es directora de programas de experiencia de usuario en Google y, desde el 2022, dirige los esfuerzos multifuncionales que hay detrás de DORA, como supervisar sus operaciones de investigación y publicar este informe. Kim también trabaja en la investigación de la experiencia de usuario de Gemini en Google Cloud. Antes de entrar en Google, ocupaba puestos relacionados con la tecnología, donde trabajaba en gestión de programas técnicos y coaching hacia la agilidad. Los orígenes de Kim están en la investigación psicosocial, que dedicó a temas como las ejecuciones extrajudiciales, el desarrollo urbano de zonas pobres y la resiliencia de la comunidad en su país de origen, Filipinas.



Michelle Irvine

Michelle Irvine es redactora técnica en Google y su investigación se centra en la documentación y en otros tipos de comunicaciones técnicas. Antes de llegar a Google, trabajó en el ámbito de la edición para centros educativos y como redactora técnica de software de simulación física. Michelle se licenció en Física y tiene un máster en Retórica y Diseño de Comunicación por la Universidad de Waterloo.



Nathen Harvey

Nathen Harvey dirige el equipo de DORA de Google Cloud. Ha trabajado con excelentes empresas, equipos y comunidades de software libre, de los que ha aprendido mucho. También es coautor de numerosos informes de DORA, así como colaborador y editor de *97 Things Every Cloud Engineer Should Know*, publicado por O'Reilly en el 2020.

Datos demográficos y firmográficos

¿Quién ha respondido la encuesta?

El programa de investigación de DORA lleva más de una década estudiando las competencias, prácticas y métricas de empresas tecnológicas con un alto rendimiento. En él han participado unos 39.000 profesionales que trabajan en empresas de diferentes tamaños y sectores. Gracias por estas valiosas aportaciones. Este año, cerca de 3000 profesionales de diversos sectores de todo el mundo han compartido sus experiencias para ayudarnos a conocer mejor los factores que impulsan a las empresas tecnológicas con un alto rendimiento.

Para las preguntas de carácter demográfico y firmográfico de este año se ha tenido en cuenta la investigación realizada por Stack Overflow.

En la encuesta a desarrolladores de Stack Overflow del 2023, participaron más de 90.000 personas.¹ Aunque no se incluyó a todos los profesionales técnicos, es lo más parecido a un censo mundial de desarrolladores.

Nos hemos hecho una idea de la población a partir de esa encuesta, así que podemos identificar el sesgo de respuesta en nuestros datos y determinar cuánto queremos generalizar nuestras conclusiones. Además, las preguntas de carácter demográfico y firmográfico que se formularon en la encuesta a desarrolladores de Stack Overflow estaban bien elaboradas, por lo que merece la pena tomarlas en consideración.

En resumen, no hay grandes discrepancias entre nuestro conjunto de muestras y el de Stack Overflow. Esto nos lleva a pensar que nuestro conjunto de muestras refleja bien a la población.

Sector

Pedimos a los encuestados que identificaran entre 12 categorías el sector en el que opera principalmente su empresa. Los sectores más comunes en los que trabajaban los encuestados fueron la tecnología (35,69 %), los servicios financieros (15,66 %) y el comercio minorista, consumo o comercio electrónico (9,49 %).

Sector	Porcentaje de encuestados
Tecnología	35,69 %
Servicios financieros	15,66 %
Comercio minorista, consumo o comercio electrónico	9,49 %
Otro	5,94 %
Industria y fabricación	5,49 %
Sanidad y productos farmacéuticos	4,60 %
Medios de comunicación y entretenimiento	4,26 %
Administración pública	3,89 %
Formación	3,66 %
Energía	3,03 %
Seguros	2,39 %
Organizaciones sin ánimo de lucro	1 %

Número de empleados

Pedimos a los encuestados que indicaran el número de empleados de su empresa entre nueve segmentos posibles. La mayoría de los encuestados trabajaba en empresas con 10.000 empleados o más (24,10 %), seguidos de empresas con entre 100 y 499 empleados (18,50 %) y entre 1000 y 9999 empleados (15,60 %).

Tamaño de la empresa	Porcentaje
1	2,0 %
De 2 a 9	3,2 %
De 10 a 19	4,3 %
De 20 a 99	14,5 %
De 100 a 499	18,5 %
De 500 a 999	11,2 %
De 1000 a 4999	15,6 %
De 5000 a 9999	6,7 %
10.000 o más	24,1 %

Discapacidad

Identificamos la discapacidad en seis dimensiones que siguen las pautas del Washington Group Short Set.² Este es el quinto año que hemos preguntado sobre discapacidad. El porcentaje de encuestados que han indicado alguna discapacidad se ha reducido del 11 % en el 2022 al 6 % en el 2023 y al 4 % en el 2024.

Discapacidad	Porcentaje de encuestados
Ninguna de las discapacidades	92 %
Al menos una de las discapacidades	4 %
Prefieren no decirlo	4 %

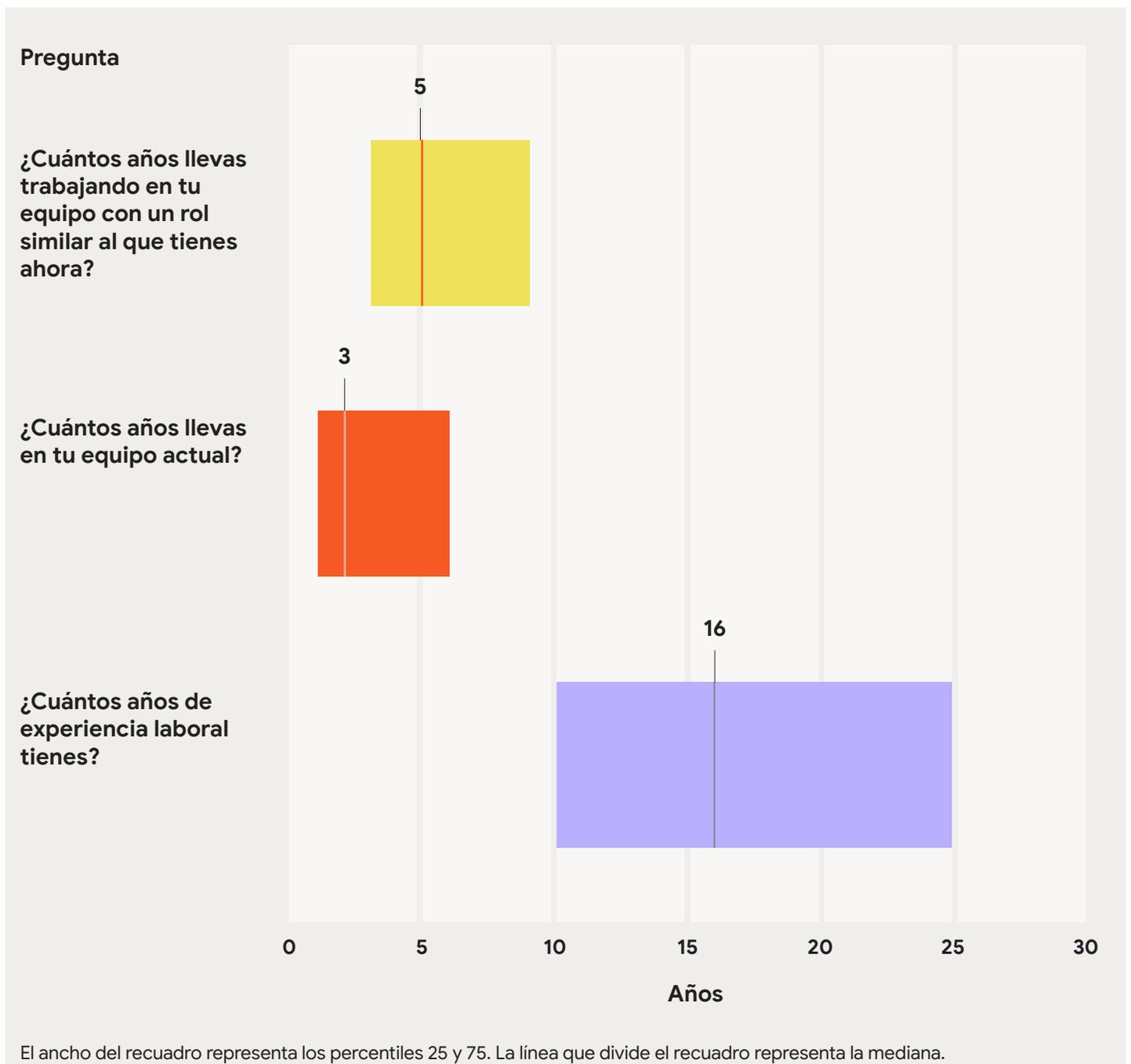
Sexo

Preguntamos a los encuestados por su sexo. El 83 % de ellos se identificaron como hombres, el 12 % como mujeres, el 1 % prefirió usar sus propias palabras y el 4 % decidieron no contestar.

Sexo	Porcentaje
Hombre	83 %
Mujer	12 %
Terminología propia	1 %
Prefieren no contestar	4 %

Experiencia

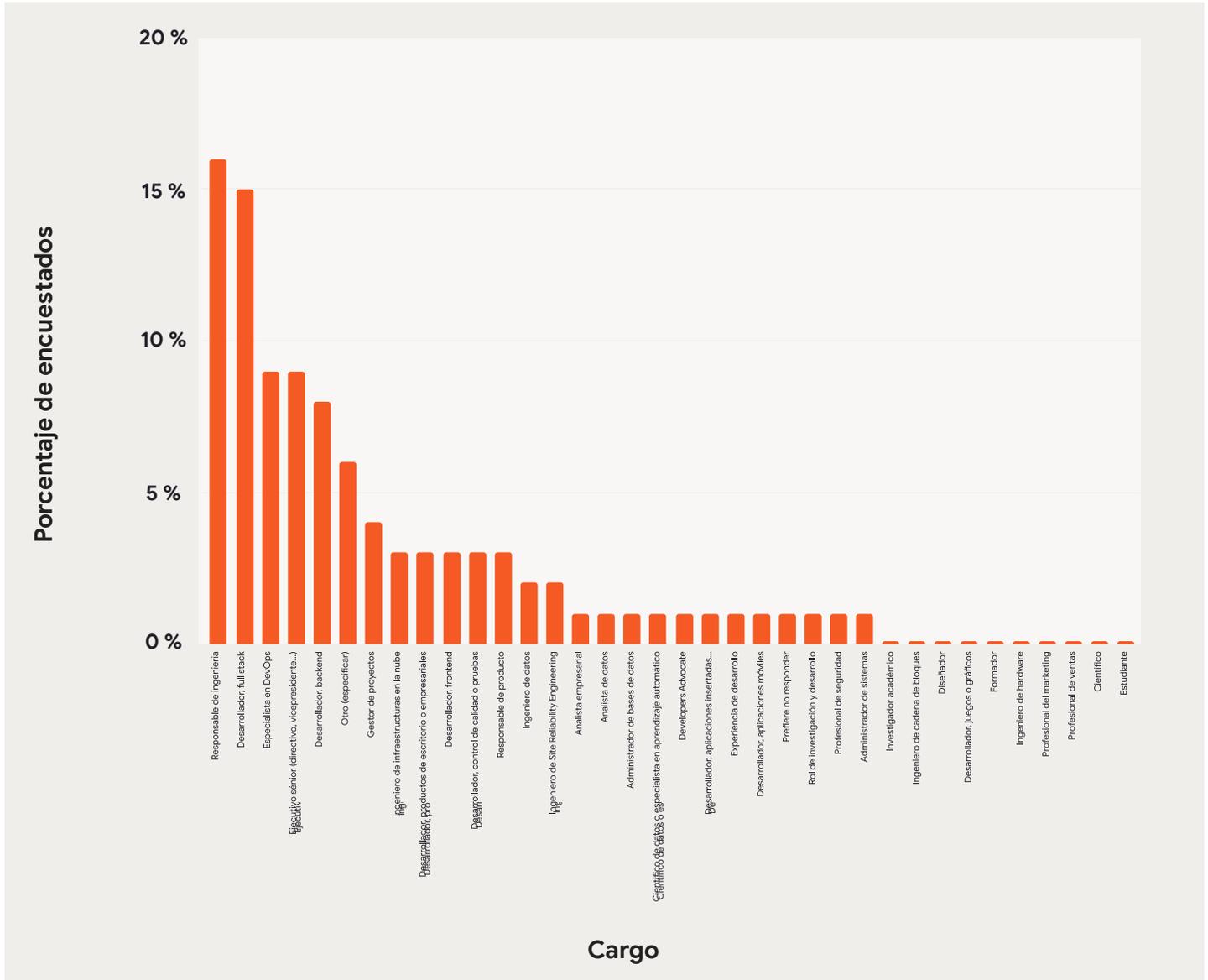
Preguntamos a los encuestados cuántos años de experiencia tenían en su rol y equipo. De media, tenían 16 años de experiencia laboral, cinco años de experiencia en su rol actual y tres años de experiencia en su equipo actual.



Rol

En los análisis, agrupamos algunos roles específicos para poder incluir con relevancia aquellos que representaban a una pequeña proporción de los encuestados. Otras categorías, como las siguientes, tenían mucha representación en nuestros datos.

- Desarrolladores: el 29 % de los encuestados.
- Supervisores: el 23 % de los encuestados.
- Altos directivos: el 9 % de los encuestados (aumento del 33 % desde el 2023).
- Analistas: en torno al 5 % de los encuestados.



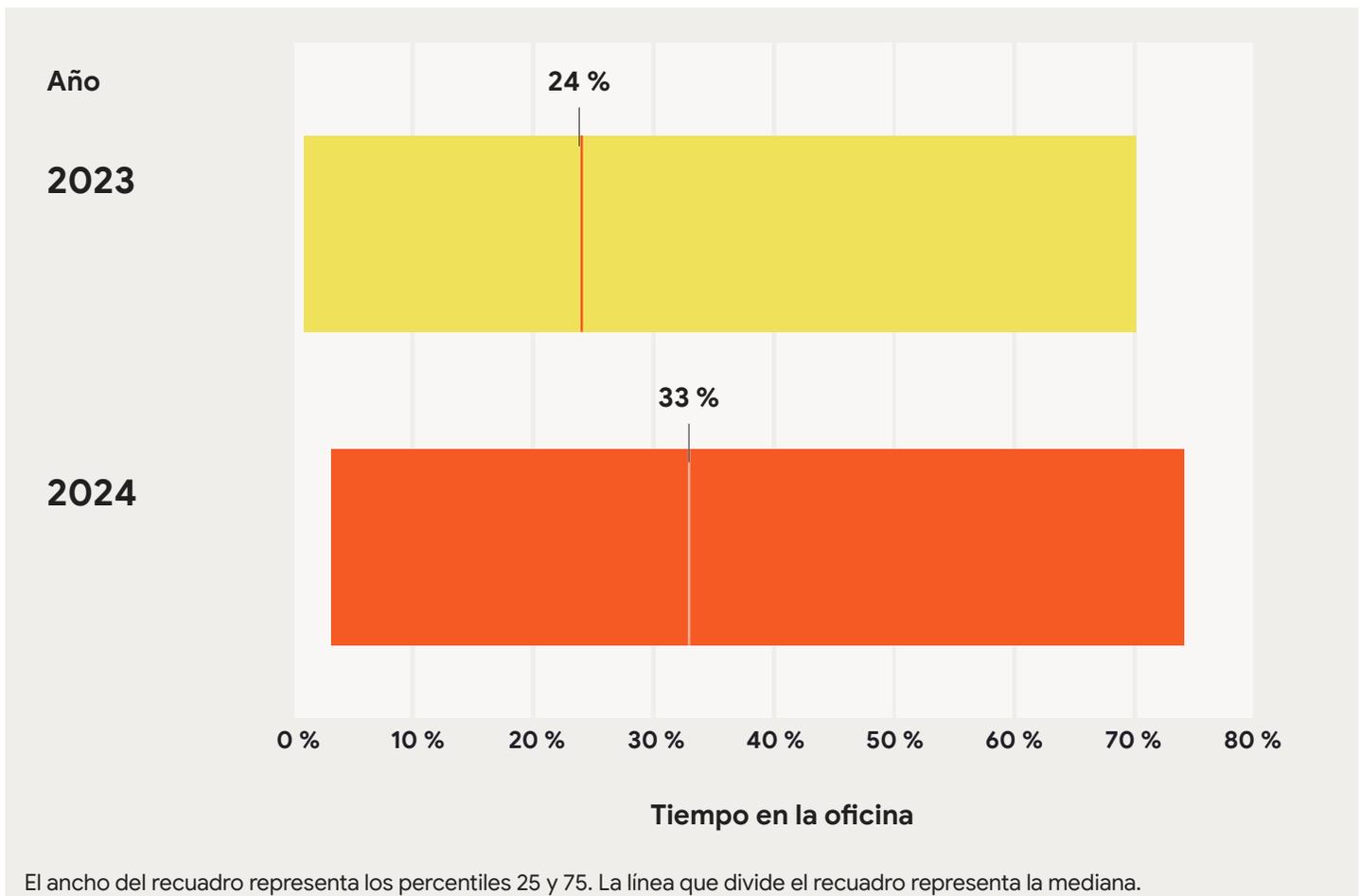
Situación laboral

Preguntamos a los encuestados cuál era su situación laboral en ese momento. La gran mayoría (90 %) eran empleados a tiempo completo en una empresa.

Tipo de empleo	Porcentaje
Contratista a tiempo completo	6 %
Empleado a jornada completa	90 %
Contratista a tiempo parcial	1 %
Empleado a tiempo parcial	2 %

Lugar de trabajo

A pesar de tratarse de otro año con presiones por volver a la oficina, en general se sigue manteniendo el patrón del año anterior, especialmente en los extremos de la distribución. El aumento del 37,5 % en los valores medios sugiere que el trabajo híbrido o, al menos, con algunas visitas presenciales, se está convirtiendo en lo más habitual.



País

Nuestros encuestados procedían de 104 países diferentes. Nos encanta ver que en la encuesta participan personas de todo el mundo. ¡Gracias a todos!



País					
Estados Unidos	Italia	Singapur	Islandia	Luxemburgo	Guatemala
Reino Unido	Suiza	Albania	Irán	Nicaragua	Región Administrativa Especial de Hong Kong
Canadá	Argentina	Georgia	Jordania	Pakistán	Malta
Alemania	México	Grecia	Kenia	Perú	Mauricio
Japón	Portugal	Filipinas	Arabia Saudí	Corea del Sur	Marruecos
India	Austria	Hungría	Eslovaquia	Sri Lanka	Nepal
Francia	Rumanía	Serbia	Eslovenia	Túnez	Paraguay
Brasil	Finlandia	Afganistán	Tailandia	Andorra	Suazilandia
España	Turquía	Argelia	Uzbekistán	Barbados	Siria
Australia	Bulgaria	Egipto	Angola	Belice	Taiwán
Países Bajos	Irlanda	Indonesia	Armenia	Benín	Macedonia del Norte
China	Israel	Federación de Rusia	Bosnia y Herzegovina	Bolivia	Trinidad y Tobago
Suecia	Bélgica	Ucrania	República Dominicana	Burkina Faso	Uruguay
Noruega	Chile	Vietnam	Ecuador	Comoras	Venezuela
Nueva Zelanda	Colombia	Bangladés	Estonia	Costa de Marfil	
Polonia	Chequia	Bielorrusia	Kazajistán	El Salvador	
Sudáfrica	Malasia	Costa Rica	Letonia	Etiopía	
Dinamarca	Nigeria	Croacia	Lituania	Gambia	

Raza y etnia

Preguntamos a los encuestados por su raza y etnia. Nuestro grupo más numeroso de encuestados eran blancos (32,4 %) y/o europeos (22,7 %).

Raza o etnia	Porcentaje
Blanca	32,4
Europea	22,7
Asiática	9,9
Norteamericana	4,6
India	4,1
Prefieren no decirlo	4,1
Hispana o latina	3,5
Sudamericana	3,2
De Asia oriental	2,5
Africana	1,8
Del sur de Asia	1,7
Multirracial	1,5
Terminología propia	1,5
Del sudeste asiático	1,4
Negra	1,3

Raza o etnia	Porcentaje
De Oriente Medio	1,3
Birracial	0,4
Centroamericana	0,4
No lo sé	0,4
Norteafricana	0,4
Caribeña	0,2
De Asia central	0,2
Del sur de Asia	1,7
Grupo etnorreligioso	0,2
De las islas del Pacífico	0,2
Indígena (por ejemplo, indio americano o indígena australiano)	0,1

¹ <https://survey.stackoverflow.co/2023/>

² <https://www.washingtongroup-disability.com/question-sets/wg-short-set-on-functioning-wg-ss/>

Metodología

Se supone que la metodología debe ser una receta que te ayude a replicar nuestro trabajo y a determinar si nuestra forma de generar y analizar los datos puede proporcionarte información valiosa. Aunque no tenemos espacio suficiente para entrar en muchos detalles, esperamos que te sirva como punto de partida.

Desarrollo de la encuesta

Selección de preguntas

Para decidir si incluir o no una pregunta en la encuesta, nos planteamos lo siguiente.

¿Cumple la pregunta estos requisitos?

- **Está hecha de manera que podemos conectar nuestro trabajo con iniciativas anteriores.**
- **Hace referencia a un resultado que el sector quiere conseguir (por ejemplo, alto rendimiento del equipo).**
- **Hace referencia a una competencia a la que el sector se está planteando dedicar sus recursos (por ejemplo, la IA).**
- **Hace referencia a una competencia que creemos que ayudará a los usuarios a lograr sus objetivos (por ejemplo, documentación de calidad).**
- **Nos ayuda a evaluar la representatividad de la muestra (por ejemplo, el rol o el sexo).**
- **Nos ayuda a bloquear los posibles sesgos (por ejemplo, lenguaje de programación o rol).**
- **Está hecha de manera que la gran mayoría de los encuestados pueden responder con cierto grado de precisión.**

Investigamos la bibliografía, interactuamos con la [comunidad DORA](#), realizamos entrevistas cognitivas, llevamos a cabo investigaciones cualitativas paralelas, trabajamos con expertos en la materia y organizamos talleres de equipo para comunicar nuestra decisión sobre si incluir una pregunta en la encuesta.

Experiencia de la encuesta

Nos preocupamos mucho por mejorar la usabilidad de la encuesta. Realizamos entrevistas cognitivas y pruebas de usabilidad para asegurarnos de que cumpla determinadas especificaciones:

- **De media, debería tardarse poco en responderla**
- **El cuestionario debería comprenderse bastante bien**
- **El nivel de esfuerzo debería ser razonablemente bajo, lo que supone un enorme reto dada la naturaleza técnica de los conceptos**

Recogida de datos

Localizaciones

Todos los años responden a nuestra encuesta personas de todo el mundo. Este año hemos trabajado en que la encuesta sea más accesible para una audiencia mayor, y la hemos localizado a inglés, español, francés, portugués, japonés y chino simplificado.



Recopilación de las respuestas de la encuesta

Utilizamos múltiples canales para buscar encuestados. Estos canales se engloban en dos categorías: enfoque orgánico y enfoque de panel.

El enfoque orgánico sirve para usar todos los medios sociales a nuestra disposición para que las personas se enteren de que hemos hecho una encuesta en la que queremos que participen. Creamos entradas de blog, lanzamos campañas por correo electrónico, publicamos en redes sociales y les pedimos a los miembros de la comunidad que también lo hagan (es decir, muestreo por bola de nieve).

El enfoque de panel lo usamos para complementar el canal orgánico. Intentamos buscar a personas tradicionalmente infrarrepresentadas en la comunidad técnica general y tratamos de obtener suficientes respuestas de determinados sectores y tipos de empresa.

En resumen, este enfoque nos permite controlar un poco la captación, algo que contrasta con el enfoque orgánico, donde el control es inexistente. El enfoque de panel también nos permite asegurarnos de que tenemos suficientes encuestados, ya que nunca sabemos si el enfoque orgánico va a conseguir las respuestas necesarias para el tipo de análisis que hacemos. Este año hemos recibido suficientes respuestas orgánicas para realizar nuestro análisis, y el panel ayudó a terminar de seleccionar a nuestro grupo de participantes.

Flujo de la encuesta

Este año, queríamos formular muchas preguntas, pero no teníamos mucho tiempo para hacerlas. Teníamos estas opciones:

- **Hacer una encuesta extremadamente larga**
- **Elegir un subconjunto de áreas en las que centrarnos**
- **Asignar temas diferentes a los encuestados de manera aleatoria**

Como no queríamos renunciar a ninguno de nuestros intereses, designamos tres flujos independientes y asignamos cada uno de ellos a los participantes de manera aleatoria. Los diferentes flujos se solapaban bastante, pero cada uno de ellos profundizaba en un área diferente.

Estas fueron las tres rutas:

- **IA**
- **Lugar de trabajo**
- **Ingeniería de plataformas**

Análisis de la encuesta

Validación de las mediciones

En esta encuesta, intentamos recoger una gran variedad de conceptos. Podríamos sumarnos a numerosos y diferentes juegos lingüísticos, pero uno de ellos denomina "variable" a la medida de un concepto. Las variables son los ingredientes de los modelos, que a su vez son los elementos incluidos en nuestra investigación. Hay dos formas amplias de analizar la validez de estas medidas: interna y externamente.

Para entender la validez interna de la medida, hay que analizar lo que creemos que indica la presencia de un concepto. Por ejemplo, las personas que mencionan la documentación de calidad podrían ser las que usan la documentación para solucionar sus problemas.

La mayoría de nuestras variables se componen de varios indicadores, porque las ideas que nos interesan parecen ser polifacéticas.

Para entender la naturaleza polifacética de una variable, probamos qué tal se combinan los elementos que usamos para representar ese concepto. Si se combinan bien (es decir, comparten un nivel alto de varianza comunal), suponemos que hay algo subyacente en ellos, como el concepto de interés.

Por ejemplo, la felicidad: es polifacética. Cuando alguien es feliz, esperamos que se sienta de una manera determinada, actúe de una manera determinada y piense de una manera

determinada. Damos por hecho que detrás de la felicidad subyace un patrón determinado de sentimientos, pensamientos y acciones.

Por eso, cuando la felicidad está presente, esperamos que surjan a la vez determinados tipos de sentimientos, pensamientos y acciones. Entonces plantearíamos preguntas sobre esos sentimientos, pensamientos y acciones, y usaríamos el análisis factorial confirmatorio para probar si realmente aparecen al mismo tiempo.

Este año hemos utilizado el paquete de R de lavaan¹ para hacer el análisis. lavaan ofrece una variedad de estadísticas ajustadas que nos ayudan a entender si los conceptos realmente representan la forma en que las personas responden a las preguntas.

Si los indicadores de un concepto no combinan bien, podría ser necesario revisar el concepto o abandonarlo, ya que es evidente que no hemos encontrado una manera fiable de medirlo.

La validez externa de una idea consiste en fijarse en cómo encaja en el mundo. Podríamos esperar que una idea estuviera relacionada de alguna manera con otras ideas. Hay veces que podemos esperar que dos ideas tengan una relación negativa, como la felicidad y la tristeza.

Si nuestra medida de felicidad se correlaciona positivamente con la tristeza, quizá debamos cuestionar nuestra medida o nuestra teoría.

De forma similar, podríamos esperar que dos ideas tuvieran relaciones positivas, pero no muy sólidas. Es posible que la productividad y la satisfacción laboral tengan una correlación positiva, pero no pensamos que sean idénticas. Si la correlación es demasiado alta, parecería que estuviéramos midiendo lo mismo. Por tanto, llegamos a la conclusión de que nuestras medidas no están lo suficientemente calibradas para recoger las diferencias entre los dos conceptos o la diferencia sobre la que hemos formulado la hipótesis no existe en realidad.

Evaluación de modelos

Tomamos un conjunto de hipótesis como principio fundamental para crear modelos hipotéticos, pequeños juguetes que intentan identificar algún aspecto sobre cómo funciona el mundo, y analizamos hasta qué punto coincidían estos modelos con los datos recopilados. Para evaluar un modelo nos decantamos por la parsimonia, que supone comenzar con un modelo muy simplista² e ir añadiendo complejidad hasta que deje de estar justificada.

Por ejemplo, predecimos que el rendimiento empresarial es el resultado de la interacción entre el rendimiento del envío del software y el rendimiento operativo. Nuestro modelo simplista no incluye la interacción:

Rendimiento empresarial \sim Rendimiento del envío de software + Rendimiento operativo

Nuestro segundo modelo añade la interacción:

Rendimiento empresarial \sim Rendimiento del envío de software + Rendimiento operativo + Rendimiento del envío de software \times Rendimiento operativo

A partir de las recomendaciones que se hacen en "Regression and other stories"³ y "Statistical Rethinking",⁴ utilizamos la validación cruzada dejando uno fuera (LOOCV)⁵ y el criterio de información de Watanabe-Akaike⁶ para determinar si es necesaria la complejidad adicional.

Grafos acíclicos dirigidos para la inferencia causal

Un modelo validado nos indica lo que necesitamos saber para adoptar un pensamiento causal. Más adelante hablaremos sobre los retos del pensamiento causal.

Estas son algunas de las razones por las que estamos *intentando* adoptar una perspectiva causal:

Pensamos que la cuestión que se plantean nuestros lectores es esencialmente causal: quieren saber si hacer algo provocará algo. Nadie va a invertir en hacer algo si cree que la relación no es causal.

Los resultados de nuestros análisis dependen de nuestra comprensión causal del mundo. Las cifras reales que obtenemos de la regresión cambian según lo que se incluya, y lo que incluyamos en la regresión dependerá de cómo pensemos que se generan los datos, lo que también es causal. Por ello, debemos ser claros.

El pensamiento causal es hacia donde nos llevará nuestra curiosidad y donde queremos pasar mucho tiempo. A menudo nos preguntamos cómo se relacionan distintos aspectos del mundo y por qué. No tenemos que llevar a cabo experimentos en cada ámbito de nuestra vida para adoptar una mentalidad causal.

El pensamiento causal es la base de la acción, así que esperamos que este informe te ayude a tomar decisiones para actuar.

Podemos usar el modelo validado para saber lo que debemos tener en cuenta para entender un efecto. En resumen, nos permite intentar recopilar datos como un experimento A/B donde se intentan crear dos mundos idénticos a partir de uno solo, con una sola diferencia entre ellos. La lógica sugiere que, al hacerlo, cualquier diferencia que aparezca entre estos dos mundos se podría atribuir a esa diferencia inicial.

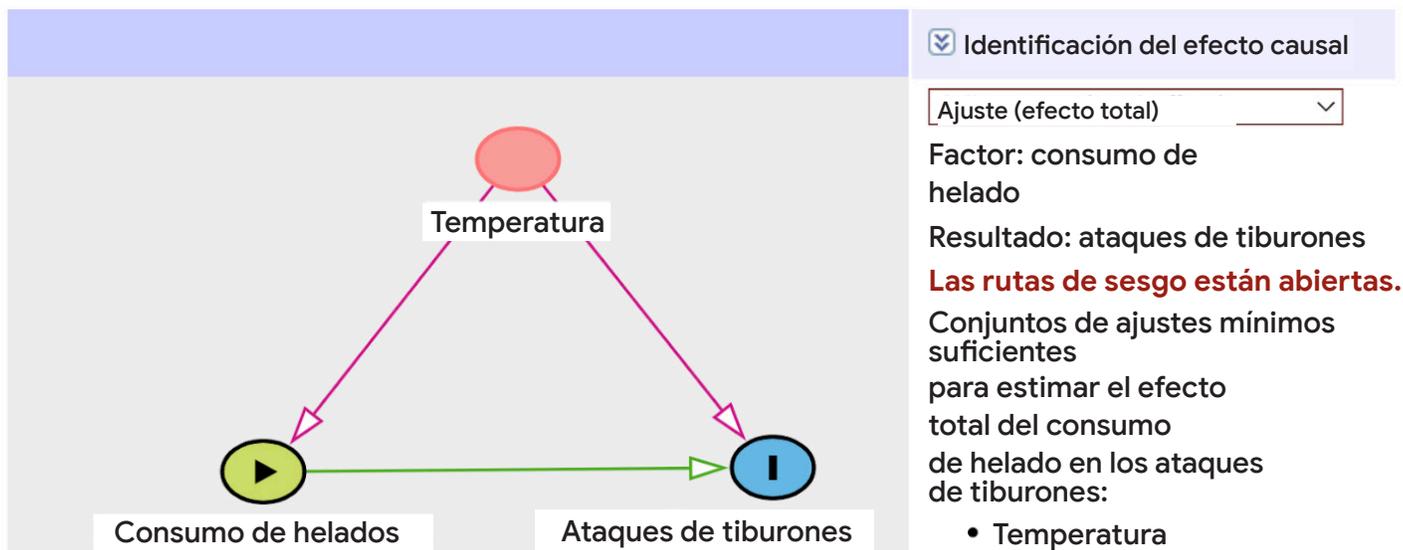
En los datos de observación y de la encuesta, la división no es tan clara: hay muchas diferencias entre los participantes, lo que lleva a confusiones. Nuestro método de inferencia causal pretende tener en cuenta esas diferencias para replicar un experimento, es decir, mantenerlo todo constante excepto una cosa (por ejemplo, la adopción de la IA).

Vamos a analizarlo con el clásico ejemplo de "los helados causan ataques de tiburones". Hay un problema con esa observación: solemos tomar helados los días calurosos, que también es cuando vamos a la playa. La situación en la que las personas comen helado y van a la playa no es igual que la situación en la que no hacen ninguna de las dos cosas. Los datos no siguen la lógica de un experimento, porque tenemos una variable que induce a confusión: la temperatura.

Los grafos acíclicos dirigidos (DAGs) ayudan a identificar las diferencias en el mundo y ofrece enfoques para paliarlas e intentar replicar un experimento manteniendo constante todo lo que forma parte de ese mundo salvo una cosa. Veamos cómo funcionan los DAGs en el ejemplo de los helados y los ataques de tiburones: queremos cuantificar el impacto del consumo de helados en los ataques de tiburones.

Dibujó mi modelo, le digo a la herramienta el efecto que quiero entender y la herramienta me dice qué va a sesgar mi estimación del efecto. En este caso, la herramienta me dice que no puedo estimar el efecto del consumo de helado en los ataques de tiburones sin tener en cuenta la temperatura, lo que supone un enfoque estadístico para intentar igualarlo todo excepto el consumo de helados y ver si los ataques de tiburones siguen fluctuando según este factor.

Presentamos nuestros modelos, como no podía ser de otro modo, en el capítulo [Modelos](#).



Fuente de la imagen: <https://www.dagitty.net/dags.html>.

El grafo acíclico dirigido nos dice qué tener en cuenta en nuestros análisis de efectos particulares.

Por ejemplo, qué debemos tener en cuenta en nuestros análisis del impacto de la adopción de la IA en la productividad.

Estadísticas bayesianas

Hemos hecho este análisis usando estadísticas bayesianas. Este tipo de estadística ofrece muchas ventajas:

- Dejamos de pensar en términos de significativo o insignificante (si les pides a 10 personas que te expliquen los valores p frecuentistas obtendrás 10 respuestas diferentes)
- Queremos saber la probabilidad de la hipótesis según los datos, no la probabilidad de los datos según la hipótesis
- Queremos incorporar lo que ya sabíamos en nuestros modelos o, al menos, ser explícitos respecto a lo que no sabemos⁷
- Nos vemos obligados a confrontar las suposiciones subyacentes al proceso de creación de modelos
- Podemos indagar en las distribuciones posteriores para hacernos una idea de la magnitud y la incertidumbre, así como determinar, en general, cómo y hasta qué punto cobran sentido los datos con el modelo, lo que en última instancia nos permite identificar de una forma bastante precisa lo que sabemos y lo que no sabemos sobre nuestros datos
- Contamos con un marco flexible que aborda muchos problemas estadísticos de una forma bastante unificada

¿Qué quiere decir "simulación"?

No es que nos inventemos los datos. Utilizamos estadísticas bayesianas para calcular una distribución posterior que trata de captar "la frecuencia esperada con la que aparecerán diferentes valores de parámetros".⁸ La parte de simulación extrae datos más de 1000 veces de esta distribución para observar los valores más creíbles de un parámetro (la media, el peso beta, sigma, intercepción, etc.) a partir de nuestros datos.

"Imagina que la distribución posterior es un cubo lleno de valores de parámetros, números como 0,1, 0,7, 0,5, 1. Dentro del cubo, cada valor existe en proporción a la probabilidad de su distribución a posteriori, de tal manera que los valores cercanos al pico son mucho más habituales que los que están en los extremos".⁹

Todo esto equivale a usar simulaciones para estudiar posibles interpretaciones de los datos y hacernos una idea de cuánta incertidumbre

hay. Piensa que cada simulación es como una pequeña IA que no conoce nada más que nuestros datos y algunas reglas, e intenta rellenar los huecos en blanco (parámetros) con una suposición fundamentada. Haz esto 4000 veces y obtendrás las suposiciones de 4000 pequeñas IAs para un parámetro determinado.

Se puede aprender mucho de estas suposiciones. Puedes conocer la suposición media, entre qué valores se encuentra el 89 %¹⁰ de esas suposiciones, cuántas suposiciones hay por encima de un determinado nivel, cuánta variación se da en las suposiciones, etc. Incluso puedes divertirte combinando suposiciones (simulaciones) entre numerosos modelos.

Al incluir un gráfico con un montón de líneas o una distribución de posibles valores, estamos intentando enseñarte lo que resulta más plausible según nuestros datos y el nivel de incertidumbre que hay.

Síntesis de conclusiones con la ayuda de la comunidad

Nuestros hallazgos ofrecen perspectivas muy valiosas para las empresas y los equipos tecnológicos, pero se entienden mejor a través del diálogo y el intercambio de impresiones. La colaboración con la comunidad DORA nos aporta información muy diversa, nos invita a poner en duda nuestras suposiciones y nos ayuda a descubrir nuevas formas de interpretar y aplicar los resultados.

Te animamos a unirse a la comunidad DORA (<https://dora.community>), donde podrás compartir tus experiencias, aprender de sus miembros y descubrir distintos enfoques para implementar estas recomendaciones. Juntos, podemos encontrar la mejor forma de aprovechar estas estadísticas y conseguir un cambio significativo en tu empresa.

Entrevistas

Este año, complementamos nuestra encuesta anual con entrevistas exhaustivas semiestructuradas para triangular, contextualizar y aclarar los resultados cuantitativos. La guía de la entrevista iba en paralelo con los temas de nuestra encuesta y se diseñó para sesiones de aproximadamente 75 minutos cada una, realizadas a distancia a través de Google Meet.

En total, entrevistamos a 11 personas, cuyos perfiles coincidían con los criterios de inclusión de la encuesta. Tanto el vídeo como el audio de la entrevista se grabaron. Las sesiones duraron entre 57 y 85 minutos, lo que supuso un total de 14 horas y 15 minutos de datos recopilados entre todos los participantes. Los datos de los participantes se seudonimizaron con los identificadores P(N), donde N corresponde al orden en el que se entrevistó a los participantes.

Todas las entrevistas se transcribieron con software automatizado. Las transcripciones se codificaron manualmente usando los temas de la entrevista como códigos previos. Las citas que aparecen en la publicación final del informe se revisaron y transcribieron manualmente antes de incluirlas. Las palabras que los autores del informe añadieron a las citas de los participantes aparecen entre paréntesis ([]), las palabras eliminadas se indican mediante puntos suspensivos (...), y se realizaron modificaciones únicamente cuando fue necesario para aumentar la claridad.

Saltos inferenciales en los resultados

Nuestro objetivo es crear una representación pragmática del mundo, algo que todos podamos aprovechar para mejorar nuestra forma de trabajar. Sabemos que estamos simplificando cuestiones complejas, pero ese es a grandes rasgos el objetivo del modelo. Jorge Luis Borges escribió un cuento corto, titulado "Del rigor en la ciencia", en el que habla de un imperio que hace mapas de su territorio a escala 1:1.¹¹ Lo absurdo es que eso hace que el mapa sea totalmente inútil (al menos, esa es mi interpretación). Nuestras simplificaciones pretenden ser útiles.

Dicho eso, hay algunos saltos inferenciales que nos gustaría aclarar.

Causalidad

Según John Stuart Mill, hay que marcar tres casillas para decir que X causa Y:¹²

- **Correlación:** ¿X tiene que covariar con Y?
- **Prioridad temporal:** ¿X tiene que pasar antes que Y?
- ¿Se tienen en cuenta los sesgos (como se ha descrito anteriormente en la sección sobre DAG)?

Estamos seguros de que entendemos la correlación, que a menudo es un procedimiento estadístico estándar. Dado que nuestra encuesta recoge un momento en el tiempo, la prioridad temporal es teórica, pero no forma parte de nuestros datos.

En lo que respecta a los sesgos, como ya hemos mencionado, los tenemos en cuenta en lo que respecta a los modelos

de ecuaciones estructurales y grafos acíclicos dirigidos, pero en este caso se trata de un ejercicio altamente teórico que, al contrario de lo que sucede con la prioridad temporal, tiene implicaciones que podemos investigar en los datos.

Lo que queremos decir es que no hicimos estudios longitudinales ni un experimento propiamente dicho. A pesar de ello, creemos que el pensamiento causal hace referencia a nuestra forma de entender el mundo y hacemos todo lo que podemos para usar técnicas emergentes de inferencia causal para dar buenas estimaciones. La correlación no implica causalidad, pero sí implica cómo entendemos la causalidad.

Fenómenos a nivel micro -> Fenómenos a nivel macro

A menudo, analizamos la relación de las competencias individuales con niveles superiores. Por ejemplo, vinculamos la adopción individual de la IA a una aplicación o servicio y al rendimiento de un equipo. No es algo muy intuitivo de primeras. La historia de cómo un fenómeno macro provoca un fenómeno individual suele ser más sencilla. Es más fácil entender que la inflación (macro) afecta a que yo compre huevos (micro) que pensar que el hecho de que yo no compre huevos causa inflación.

Lo mismo sucede si hablamos de cómo afecta el rendimiento de una empresa (macro) al bienestar de una persona (micro). Desde el punto de vista heurístico, es probable que la empresa influya más en la persona que la persona en la empresa.

Siendo así, ¿por qué decimos que una acción individual afecta a algo como el rendimiento de un equipo o empresa? Hacemos un salto inferencial que pensamos que no es del todo ilógico. En concreto, damos por hecho que, a gran escala, la siguiente afirmación suele ser verdad:

$$p(\text{individuo hace } X \mid \text{empresa hace } X) > p(\text{individuo hace } X \mid \text{empresa no hace } X).$$

Creemos que la probabilidad de que una persona haga algo (X) es mayor si está en una empresa o un equipo que también hace X. De esta forma, las personas que hacen algo representan a equipos y empresas que también tienden a hacer X. Por supuesto, aquí hay bastante ruido, pero debería aparecer un patrón con el que esta suposición nos diera bastantes competencias importantes.

Veamos un ejemplo sin relación con DORA: imagina dos países en los que la altura media difiere. La altura media de los habitantes de uno de los países es de 168 cm (5'6"), mientras que en el otro miden 188 cm (6'2") de media. La desviación típica es idéntica. Si eliges a una persona aleatoria de cada uno de los dos países, ¿de cuál crees que hay más posibilidades que sea la persona más alta? Si lo haces miles de veces, el país con la altura media mayor estará representado por las personas más altas. La altura de los individuos se acercaría a las alturas medias de los países.

No es que sea necesario, pero hicimos una simulación rápida para validar que es cierto:

```
#código R
#fijamos semilla para reproducibilidad
set.seed(10)

#6'2 y 5'6
height_means = c(6 + 1/6, 5.5)

#desviación típica constante en 1/4 de
pie
std_dev =0.25

#selecciones aleatorias
draws = 1000

#selecciones aleatorias del país A
country_a <-rnorm(draws, mean = height_
means[1], sd = std_dev)

#selecciones aleatorias del país B
country_b <-rnorm(draws, mean = height_
means[2], sd = std_dev)

#cómo se representa la diferencia
correcta en las selecciones aleatorias
represented_difference = sum(country_a >
country_b) / 1000

#muestra los resultados como un
porcentaje
represented_difference * 100
```

Los resultados no nos sorprenden. El 97,2 % de las 1000 selecciones aleatorias apuntan en la dirección correcta. Por supuesto, sería muy fácil caer en la trampa de las selecciones no aleatorias, las diferencias más pequeñas entre países y las muestras reducidas. Aun así, la cuestión sigue siendo la misma: las diferencias en el nivel macro tienden a representarse en el nivel micro.

1. Rosseel Y (2012). "lavaan: An R Package for Structural Equation Modeling". Journal of Statistical Software, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
2. Esto también implicaría examinar las posibles fuentes de confusión.
3. Gelman, Andrew, Jennifer Hill y Aki Vehtari. 2021. Regression and Other Stories. S.I. Cambridge University Press.
4. McElreath, Richard. 2016. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. S.I. CRC Press/Taylor & Francis Group.
5. Gelman, Andrew, Jennifer Hill y Aki Vehtari. 2021. Regression and Other Stories. S.I. Cambridge University Press
6. McElreath, Richard. 2016. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. S.I. CRC Press/Taylor & Francis Group.
7. La información anterior es más bien débil (escéptica, neutra y poco reveladora) y hemos comprobado que los resultados no se ven condicionados por ello.
8. McElreath, Richard. Statistical rethinking: A Bayesian course with examples in R and Stan. Chapman and Hall/CRC, 2018, p. 50
9. McElreath, Richard. Statistical rethinking: A Bayesian course with examples in R and Stan. Chapman and Hall/CRC, 2018, p. 52
10. Seguimos el razonamiento de McElreath en la página 56 de Statistical Rethinking para elegir el 89 %. "¿Por qué estos valores? No hay ninguna razón... Y estos valores evitan el típico 95 %, ya que los intervalos habituales del 95 % animan a muchos lectores a probar las hipótesis de manera inconsciente". El intervalo que proporcionamos intenta mostrar un "rango de valores de parámetros compatible con el modelo y los datos" plausible.
11. Borges, J. L. (1999). Collected fictions. Penguin.
12. Duckworth, Angela Lee, Eli Tsukayama y Henry May. "Establishing causality using longitudinal hierarchical linear modeling: An illustration predicting achievement from self-control". Social psychological and personality science 1, número 4 (2010): 311-317.

Modelos

Antes, creábamos un modelo gigante que validábamos usando varias técnicas de modelización de ecuaciones estructurales (mínimos cuadrados parciales, con base en la covarianza, de tipo bayesiano). Para el informe del 2023, cambiamos el enfoque para centrarnos en muchos modelos más pequeños con el objetivo de entender procesos específicos.

Por ejemplo, hicimos un modelo detallado para entender la física de la documentación de calidad. Crear modelos más pequeños¹ y personalizados para entender efectos específicos tiene muchas ventajas:

- **Ayuda a identificar áreas en las que el modelo no encaja bien.**
- **Todo lo que se añade a un modelo ejerce una fuerza, una cierta gravedad. Cuando el modelo se hace más grande, es muy difícil entender todas las variables diferentes que ejercen fuerzas entre sí.**
- **Evita condicionamientos que muestren relaciones espurias.²**

¿Cómo usamos los modelos?

Todos nos preguntamos muchas cosas, pero un gran número de preguntas cruciales tienen la misma estructura:

¿si hacemos X, qué pasa con Y?

La X suele ser una práctica, como crear documentación de calidad, adoptar la IA o invertir en cultura.

La Y suele ser algo que queremos lograr o evitar, algo que podría ocurrir desde el nivel individual (por ejemplo, la productividad) hasta el nivel de la empresa (por ejemplo, la cuota de mercado).

Construimos, evaluamos y usamos los modelos³ para abordar preguntas con esta estructura. Nuestro objetivo es ofrecer una estimación precisa de qué sucede con resultados importantes como consecuencia de haber hecho X.⁴ Cuando registramos efectos, expresamos dos características fundamentales:

1. ¿Qué nivel de certeza tenemos de la **dirección** del efecto, es decir, cómo de claro está que esa práctica será beneficiosa o perjudicial?
2. ¿Qué grado de certeza tenemos de la **magnitud** del efecto? Daremos una estimación relativa de qué nivel de impacto tendrán determinadas prácticas y qué grado de incertidumbre rodea a estas estimaciones.

Estas son algunas de las competencias en las que nos hemos centrado este año:

- **Adopción de la IA**
- **Uso de la plataforma**
- **Antigüedad de la plataforma**
- **Liderazgo en el ámbito de la transformación**
- **Estabilidad de las prioridades**
- **Enfoque orientado al usuario**

Estas son algunos de los resultados y grupos de resultados de este año:

- **Rendimiento y bienestar individuales (por ejemplo, desgaste)**
- **Rendimiento del equipo**
- **Rendimiento del producto**
- **Flujo de trabajo de desarrollo (por ejemplo, complejidad del código base y calidad de los documentos)**
- **Rendimiento del envío de software**
- **Rendimiento empresarial**

Nos centramos en estos resultados porque creemos que son finales. Por supuesto, esto es más cierto para unos resultados que para otros. Si descubrieras que el rendimiento empresarial y de los equipos no tiene nada que ver con el rendimiento del envío de software, seguramente te daría igual que este último fuera bajo.

Sin embargo, esperamos que incluso aunque el rendimiento empresarial no dependiera del bienestar individual de los empleados, aún quisieras priorizarlo.

Un modelo repetido

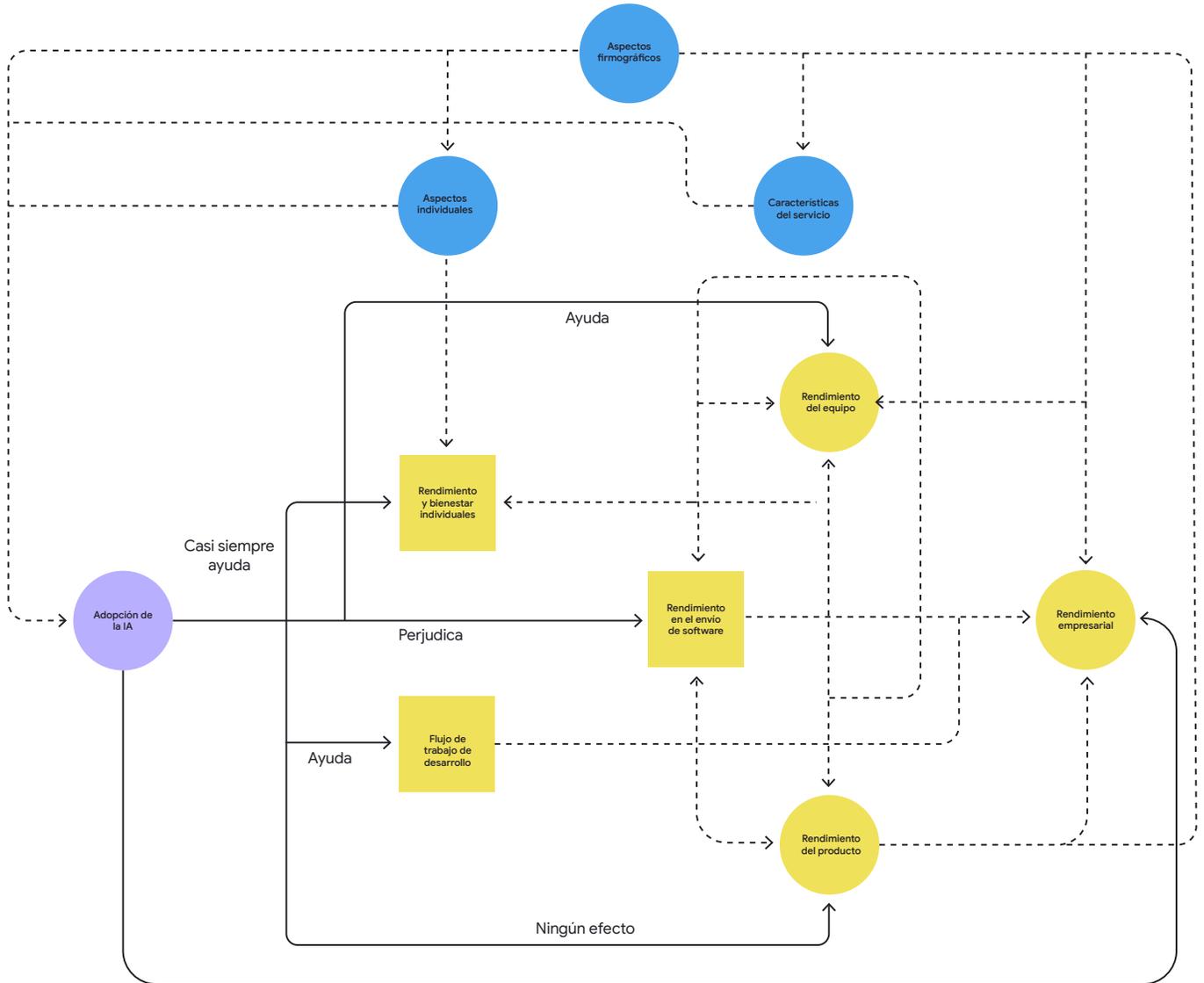
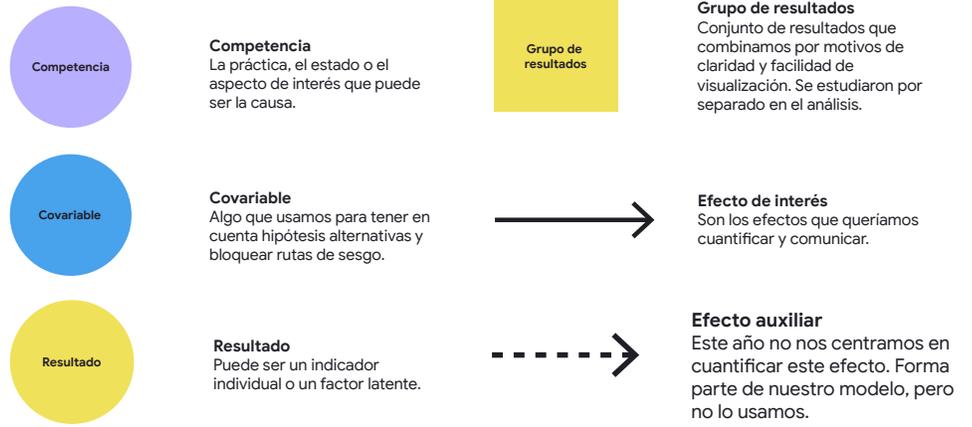
En los tres últimos años, desarrollamos y estudiamos muchas hipótesis sofisticadas, especialmente en lo que respecta a la moderación y a la mediación.

Este año, nos hemos centrado menos en este tipo de hipótesis y hemos dedicado más tiempo a intentar estimar el efecto de una competencia en un resultado. Esto significa que el modelo de cada capacidad es prácticamente el mismo.

Por tanto, el diseño del modelo de los efectos de la adopción de la IA es muy similar al del modelo usado para los efectos del enfoque orientado al usuario. Podríamos haber copiado el modelo y cambiado el nombre de la competencia, pero eso seguramente no habría sido tan útil para los lectores.

En su lugar, nos vamos a limitar a mostrar el modelo de IA, pero es importante saber que es el esquema o formato que está detrás de cada uno de nuestros modelos. Si quieres hacer tu propio análisis, al construir este modelo en una herramienta como [DAGitty](#), deberías replicar aproximadamente las regresiones que usamos en nuestro análisis. Dicho eso, la información que presentamos se ha simplificado un poco para hacerla más legible. Además, aunque los modelos son muy similares para cada competencia, los efectos son diferentes. Por ejemplo, más abajo puede verse que la adopción de la IA suele afectar negativamente al rendimiento del envío de software, pero resulta positiva en aspectos como la documentación interna y el enfoque orientado al usuario. En cada capítulo encontrarás información adicional al respecto.

La clave



1. En "Regression and other stories" (Gelman et. al) se incluyen algunos consejos que arrojan luz, en concreto en las páginas 495 y 496: B.6 Usa varios modelos y B.9 Haz la inferencia causal de manera dirigida, no como consecuencia de una gran regresión
2. En el capítulo 6 de Statistical Rethinking se habla mucho sobre esto. Nos interesan especialmente los sesgos de colisión.
3. En el capítulo de metodología se analiza la relación de estos modelos con los grafos acíclicos dirigidos.
4. También hablamos brevemente sobre causalidad en el capítulo de metodología.

Lecturas recomendadas

Únete a la comunidad DORA para debatir, aprender y colaborar con el objetivo de mejorar el rendimiento del envío de software y de las operaciones. <https://dora.community>

Responde a la encuesta breve de DORA. <https://dora.dev/quickcheck>

Descubre las competencias que fomentan un entorno caracterizado por el aprendizaje, la fluidez y las sugerencias rápidas. <https://dora.dev/capabilities>

Fostering developers' trust in generative artificial intelligence. <https://dora.dev/research/2024/trust-in-ai/>

Lee el libro *Accelerate: The science behind devops: Building and scaling high performing technology organizations*. IT Revolution. <https://itrevolution.com/product/accelerate>

Lee el libro *Team Topologies: Organizing Business and Technology Teams for Fast Flow*. IT Revolution Press. <https://teamtopologies.com/>

Publicaciones del programa de investigación de DORA, donde se incluyen informes de DORA anteriores. <https://dora.dev/publications>

Preguntas frecuentes sobre la investigación y los informes. <http://dora.dev/faq>

Fe de erratas: lee y envía cambios, correcciones y aclaraciones del informe en <https://dora.dev/publications/errata>

Comprueba si esta es la versión más reciente del informe de DORA del 2024: <https://dora.dev/vc/?v=2024.3>

"Accelerate State of DevOps 2024"
de Google LLC está sujeto a la licencia
[CC BY-NC-SA 4.0](#)

