

Presentado por



Google Cloud

Informe

# Accelerate State of DevOps 2023



Patrocinadores Premier



Deloitte.

Qarik

# Índice

## Introducción

Resumen ejecutivo .....	03
Conceptos y mediciones .....	06

## Capítulo 1

Comparación con otras empresas .....	10
--------------------------------------	----

## Capítulo 2

Enfoque en los usuarios .....	17
-------------------------------	----

## Capítulo 3

Las capacidades técnicas predicen el rendimiento .....	20
--	----

## Capítulo 4

La documentación es la base .....	27
-----------------------------------	----

## Capítulo 5

La confiabilidad desbloquea el rendimiento .....	31
--	----

## Capítulo 6

La infraestructura flexible es clave para el éxito .....	38
--	----

## Capítulo 7

Nada de esto funciona sin inversión en cultura .....	45
--	----

## Capítulo 8

Cómo, cuándo y por qué importa quién eres .....	51
---	----

## Epílogo

Reflexiones finales .....	57
Agradecimientos .....	58
Autores .....	59
Metodología .....	62
Datos demográficos y organizativos .....	72
Los modelos .....	81
Lecturas adicionales .....	91
Apéndice .....	92

Todas las citas fueron consultadas el 27 de septiembre de 2023

# Resumen ejecutivo

Durante casi una década, el programa de investigación DORA ha estudiado las capacidades y mediciones de las organizaciones de tecnología de alto rendimiento. Recibimos comentarios de más de 36,000 profesionales de organizaciones de diferentes tamaños y sectores. ¡Gracias por compartir tus observaciones!

DORA busca comprender la relación entre las formas de trabajo (es decir, las capacidades) y los resultados: logros considerables que son relevantes en toda la organización y para las personas que la conforman. Este estudio usa evaluaciones estadísticas rigurosas y no depende de ninguna plataforma (consulta [Metodología](#)).

Esperamos que estas observaciones brinden a los líderes y profesionales una idea de dónde pueden generar un impacto.

En el estudio de este año, se exploraron tres resultados clave y las capacidades que contribuyen a lograr esos resultados:

- **Rendimiento organizativo:** La organización debería generar no solo ingresos, sino también valor para los clientes y la comunidad general.
- **Rendimiento de los equipos:** Se refiere a la habilidad de un equipo de servicio o de aplicaciones para crear valor, innovar y colaborar.
- **Bienestar de los empleados:** Las estrategias que adopta una organización o un equipo deberían beneficiar a los empleados; por ejemplo, reducir el agotamiento, fomentar una experiencia laboral satisfactoria y aumentar la capacidad de las personas para producir resultados valiosos (es decir, productividad).

En el estudio, también se exploraron medios y mediciones de rendimiento que, a menudo, definimos como fines en sí mismos:

- **Rendimiento de la entrega de software:** Los equipos pueden cambiar sus sistemas de tecnología de manera segura, rápida y eficiente.
- **Rendimiento operativo:** El servicio proporciona una experiencia confiable a sus usuarios.



# Hallazgos clave



## Establece una cultura saludable

La cultura es la base para desarrollar capacidades técnicas, impulsar el rendimiento técnico, lograr objetivos de rendimiento organizativo y ayudar a los empleados a alcanzar el éxito. Los equipos con culturas generativas tienen un rendimiento organizativo un 30% mayor.



## No te olvides de los usuarios

Un enfoque en el usuario puede fundamentar e impulsar mejoras en todas las capacidades técnicas, culturales y de procesos que analizamos en el estudio. Los equipos pueden realizar implementaciones exitosas tan rápido como quieran, pero, si no tienen al usuario en mente, el esfuerzo puede ser en vano. Los equipos que se enfocan en el usuario tienen un rendimiento organizativo un 40% mayor.



## Optimiza el rendimiento de la entrega de software con revisiones de código más rápidas

Acelerar las revisiones de código es uno de los caminos más eficaces para mejorar el rendimiento de la entrega de software. Los equipos con revisiones de código más rápidas tienen un rendimiento de entrega de software un 50% mayor.



## Amplifica las capacidades técnicas con documentación de calidad

La documentación de alta calidad amplifica el impacto que tienen las capacidades técnicas en el rendimiento organizativo. Por ejemplo, se calcula que el desarrollo basado en troncales tiene un impacto 12.8 veces mayor en el rendimiento organizativo cuando se implementa documentación de alta calidad en comparación con la documentación de baja calidad.



## Aumenta la flexibilidad de la infraestructura con la nube

La computación en la nube es beneficiosa porque crea una infraestructura flexible. Usar una nube pública, por ejemplo, genera un aumento del 22% en la flexibilidad de la infraestructura en comparación con no usar la nube. A su vez, esa flexibilidad lleva a un rendimiento organizativo un 30% mayor en comparación con las infraestructuras no flexibles. Para obtener el máximo valor de la nube, la clave es aprovechar las características y capacidades diferenciadoras que ofrece la nube, en específico, la flexibilidad de infraestructura.



## Equilibra la velocidad de entrega, el rendimiento operativo y el foco en el usuario

Necesitas que el rendimiento de entrega de software y el operativo sean sólidos para que el rendimiento organizativo alcance su máximo potencial. Mantener el equilibrio con el enfoque en el usuario brinda los mejores resultados organizativos, a la vez que mejora el bienestar de los empleados.



## Distribuye el trabajo de manera justa

Las personas que se sienten poco representadas y las mujeres o aquellas personas que deciden autodefinir su género tienen niveles de agotamiento más altos. Lo más probable es que haya varios factores sistémicos y ambientales que provocan ese efecto. Es de esperarse que observemos que los encuestados cuyo trabajo es más repetitivo tengan más probabilidades de experimentar niveles más altos de agotamiento, y los miembros de grupos poco representados tienen más probabilidades de hacer más trabajo repetitivo. En comparación con sus contrapartes, los encuestados poco representados tienen un agotamiento un 24% mayor; por otra parte, hacen un 29% más de trabajo repetitivo. Las mujeres o aquellas personas que autodefinen su género hacen un 40% más de trabajo repetitivo que los hombres.

# Aplica estadísticas de DORA en tu contexto

Es probable que los equipos que adoptan una mentalidad y una práctica de mejora continua tengan más beneficios.<sup>1</sup> DORA puede ayudar a influenciar tus iniciativas de mejora.

Para aprovechar al máximo este estudio, considéralo en el contexto de tu equipo y tus usuarios. Por ejemplo, antes mencionamos que los equipos con las revisiones de código más rápidas tienen un rendimiento un 50% mayor en la entrega de software. Sin embargo, es poco probable que este rendimiento mejore si las revisiones de código ya son rápidas, pero la velocidad está limitada en otro lugar en el sistema. Es posible contextualizar el estudio cuando los profesionales tienen conversaciones sobre cómo se realiza el trabajo en la actualidad. Las conversaciones pueden mejorar la empatía, colaboración y comprensión de las motivaciones de cada participante.

El trabajo de mejora nunca termina. Encuentra un cuello de botella en tu sistema, solúcnalo y repite el proceso. Las comparaciones más importantes se hacen al analizar la misma aplicación con el paso del tiempo y no otras aplicaciones, organizaciones o sectores.

## Métricas y mediciones

Las métricas y los paneles ayudan a los equipos a supervisar su progreso y corregir el rumbo.

Los profesionales y líderes quieren mejorar el rendimiento organizativo y de los equipos, así como su bienestar. Sin embargo, ni la medición ni la entrega de software son el objetivo.

Enfocarse en las métricas de rendimiento puede generar comportamientos ineficaces. Invertir en capacidades y capacitación es una mejor forma de posibilitar el éxito. Los equipos que más aprenden son los que mejoran más.

## No puedes mejorar por tu cuenta

Podemos aprender de las experiencias de otros; un foro excelente para compartir iniciativas de mejora y aprender de ellas es el sitio de la Comunidad de DORA <https://dora.community>.

<sup>1</sup>Informe Accelerate State of DevOps 2022. <https://dora.dev/research/2022/dora-report/2022-dora-accelerate-state-of-devops-report.pdf#page=7>



# Conceptos y mediciones

En esta sección, se incluyen descripciones de los conceptos que DORA intenta medir.<sup>1</sup> Esos elementos forman la base de este informe y nuestros modelos. En estas secciones, se incluyen ingredientes que usamos para elaborar este informe. Por lo tanto, es importante que los autores seamos claros sobre qué son esos conceptos y que hablemos de ellos de manera consistente. El objetivo de las siguientes tablas es proporcionar claridad y una base que brinde a todos (tanto lectores como autores) una terminología compartida.

Debido a que muchos de los conceptos en este informe son multifacéticos, a menudo usamos varios indicadores para captarlos. Una manera de evaluar qué tan bien captamos esos conceptos es usar análisis factorial exploratorio y confirmatorio. Puedes leer más sobre esos procesos en [Metodología](#). Después de evaluar nuestros métodos de medición, hicimos una escala con las puntuaciones de 0 a 10, en la que 0 representa la falta total de un concepto y 10 representa su presencia máxima. Creemos que, de esta manera, se estandariza la forma en la que hablamos sobre cómo funcionan esos conceptos y nos ayuda a comparar datos a lo largo de los años.

Cada uno de los conceptos que analizamos está acompañado de la siguiente información:

- Un ícono para ayudar a comunicar el significado y facilitar la búsqueda cuando se use este capítulo como referencia
- La puntuación promedio de ese concepto en la muestra (la media)
- Los límites del rango intercuartílico (IQR). Al proporcionar los dos números (percentiles 25 y 75) en los que reside el 50% medio de los datos, esos límites deberían ayudar a transmitir la variedad de respuestas
- El valor medio en un conjunto de datos (la mediana). Si es considerablemente diferente de la media, podría indicar que los datos están sesgados
- Una descripción del concepto y cómo lo medimos

<sup>1</sup>Las preguntas de la encuesta que se usaron en el análisis están publicadas en <https://dora.dev>.

# Resultados clave

Los resultados clave son los objetivos que creemos que las personas, los equipos y las organizaciones quieren lograr (rendimiento organizativo, por ejemplo) o evitar (agotamiento, por ejemplo). En consecuencia, creemos que las mediciones son importantes para que las personas se evalúen a sí mismas, a sus equipos y a sus organizaciones.

 Rendimiento organizativo	 Rendimiento del equipo	 Rendimiento de la entrega de software
Media 6.3	IQR 5-8	Mediana 6.3
Las organizaciones de alto rendimiento tienen más clientes, ganancias más altas y una mayor participación de mercado relativa para su producto o servicio principal.		Media 7.6
IQR 6.6-9	Mediana 8	Media 6.3
Los equipos con un rendimiento alto se adaptan al cambio, confían el uno en el otro, trabajan de manera eficiente, innovan y colaboran.		IQR 5.1-7.8
Media 6.4		Mediana 6.4
 Rendimiento operativo	 Objetivos de confiabilidad	Las siguientes cuatro métricas miden la velocidad y estabilidad de la entrega de software: <ul style="list-style-type: none"><li>• Frecuencia de implementación</li><li>• Plazo de entrega de los cambios</li><li>• Tasa de errores de los cambios</li><li>• Tiempo de recuperación para implementaciones con errores</li></ul>
Media 6.2	IQR 5-7.5	Media 7.5
El grado en que un servicio puede cumplir con las expectativas de los usuarios, incluidas mediciones como disponibilidad y rendimiento.		Media 7.5
El grado en que un servicio logra los objetivos establecidos en relación con mediciones como disponibilidad, rendimiento y corrección.		Media 7.5

El **bienestar** es una combinación de agotamiento, productividad y satisfacción laboral.

 Agotamiento	 Productividad	 Satisfacción laboral
Media 4.1	IQR 2-6	Mediana 4
No solo el efecto físico y psicológico del trabajo, sino también el valor y el significado que cada persona le da a su trabajo. El agotamiento provoca cinismo. <sup>2</sup>		Media 7.9
Una persona productiva realiza un trabajo que se alinea con sus habilidades, crea valor y le permite ser eficiente.		IQR 5.7-7.1
Media 7.1		Mediana 7.1

<sup>2</sup>Maslach C., y Leiter MP. "Understanding the burnout experience: recent research and its implications for psychiatry". World Psychiatry. Junio de 2016; 15(2):103-11. doi: 10.1002/wps.20311. PMID: 27265691; PMCID: PMC4911781.

<sup>3</sup>Warr, P., Cook, J. y Wall, T. "Scales for the measurement of some work attitudes and aspects of psychological well-being". Journal of Occupational Psychology, 52(2), 1979. 129-148. <https://doi.org/10.1111/j.2044-8325.1979.tb00448.x>

# Procesos y capacidades técnicas

Estos son actividades, prácticas o estados que pueden surgir en un equipo o una organización. En otras palabras, lo que hacen los equipos o sus formas de ser.

 <b>Contribución de la inteligencia artificial</b>  Media      IQR      Mediana 3.3      0.3-6.3      2.4  La importancia del rol de la inteligencia artificial en la contribución a distintas tareas técnicas.	 <b>Arquitectura con acoplamiento bajo</b>  Media      IQR      Mediana 6.4      4.7-8.3      6.7  Software que puede escribirse, probarse e implementarse de manera independiente.	 <b>Integración continua</b>  Media      IQR      Mediana 6.9      5-8.9      7.8  La práctica de desarrollar y probar cambios de software automáticamente.
 <b>Entrega continua</b>  Media      IQR      Mediana 7.0      5.7-8.7      7.3  La habilidad de llevar a producción cambios de todo tipo, como funciones nuevas, cambios de configuración, correcciones de errores y experimentos, o ponerlos a disposición de los usuarios de manera segura, rápida y sustentable. <sup>4</sup>	 <b>Velocidad de la revisión de código</b>  Media      IQR      Mediana 6.5      6-8      6  Una sola pregunta que evalúa el tiempo necesario desde la solicitud de extracción hasta la revisión del cambio de código.	 <b>Documentación</b>  Media      IQR      Mediana 5.8      3.8-7.9      6.25  La calidad del contenido escrito que crean las personas en la organización y usan en su trabajo diario.
 <b>Prácticas de confiabilidad</b>  Media      IQR      Mediana 5.9      3.9-8.3      6.1  Actividades y prácticas que usan los equipos para mejorar el rendimiento operativo de los servicios.	 <b>Desarrollo basado en troncales</b>  Media      IQR      Mediana 5.6      3.9-7.8      5.6  La práctica de realizar cambios pequeños y frecuentes que se combinen de manera regular en la rama principal de código del sistema de control de versión.	 <b>Infraestructura flexible</b>  Media      IQR      Mediana 6.6      5-8.3      7.3  Infraestructura escalable que es elástica, accesible y medible. <sup>5</sup>

<sup>4</sup>"What is Continuous Delivery" <https://continuousdelivery.com/>

<sup>5</sup>Instituto Nacional de Normas y Tecnología (2018). NIST, The NIST Definition of Cloud Computing. Disponible en <https://csrc.nist.gov/pubs/sp/800-145/final>

# Aspectos culturales

No es sencillo definir el término "cultura", pero podríamos decir que son las normas prevalecientes (como la flexibilidad), la orientación predominante (como el enfoque en el usuario) y la atmósfera (como la estabilidad de la organización) del lugar de trabajo.

 Distribución del trabajo	 Flexibilidad	 Seguridad laboral
<b>Media</b> 5.8	<b>IQR</b> 3.8-7.9	<b>Mediana</b> 5.8
Procesos formales que ayudan a los empleados a distribuir tareas de manera equitativa dentro de un equipo.	Cómo, dónde y cuándo una persona trabaja en sus tareas. <sup>6</sup>	Una medición de un solo elemento con la que se pregunta la frecuencia con la que alguien se preocupa por su seguridad laboral. Las puntuaciones más altas equivalen a menos preocupación.
 Estabilidad organizativa	 Conocimiento compartido	 Enfoque en los usuarios
<b>Media</b> 7.2	<b>IQR</b> 6.7-8.3	<b>Mediana</b> 8.3
Una medición de un solo elemento con la que se pregunta qué tan estable o inestable es el entorno laboral para los empleados.	Cómo se esparran las ideas y la información en la organización. Los miembros de los equipos responden preguntas una vez y ponen la información a disposición de otras personas. Nadie tiene que esperar por las respuestas. <sup>7</sup>	Comprender e incorporar las necesidades y los objetivos de los usuarios para mejorar los productos y servicios. <sup>8</sup>
 Cultura organizativa de Westrum		
<b>Media</b> 7.3	<b>IQR</b> 6.1-8.6	<b>Mediana</b> 7.8
La manera en que una organización suele responder a oportunidades y problemas. Hay tres tipos de cultura: generativa, burocrática y patológica. <sup>9</sup>	<p><sup>6</sup>Shirfin, Nicole V. y Jesse S. Michel. "Flexible work arrangements and employee health: A meta-analytic review". Work &amp; Stress 36(1), 2022, 60-85.</p> <p><sup>7</sup>"2022 Developer Survey", <a href="https://survey.stackoverflow.co/2022#overview">https://survey.stackoverflow.co/2022#overview</a></p> <p><sup>8</sup>Kersten, Mik. "Project to Product: How to survive and thrive in the age of digital disruption with the flow framework". IT Revolution, 2018, 54. <a href="https://itrevolution.com/product/project-to-product/">https://itrevolution.com/product/project-to-product/</a></p> <p><sup>9</sup>Westrum R. "A typology of organisational cultures". BMJ Quality &amp; Safety, 2004. <a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1765804/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1765804/</a></p>	

## Capítulo 1

# Comparación con otras empresas

Ley de Goodhart: cuando una medición se convierte en un objetivo, deja de ser una buena medición.<sup>1</sup>

## Conclusiones

El primer paso para mejorar el rendimiento es definir un punto de referencia para el rendimiento de entrega de software, el enfoque en los usuarios y el rendimiento operativo actuales de una aplicación. Esas mediciones ayudan a los equipos a evaluar su desempeño y brindan un buen indicador de los cambios con el paso del tiempo.

Sin embargo, esas mediciones no son el medio por el que mejorará un equipo. Con ese punto de referencia, es importante evaluar la fortaleza de un equipo en una amplia variedad de personas, procesos y capacidades técnicas para identificar qué podría estar limitando el progreso.<sup>2</sup> Luego, los equipos necesitan tiempo y espacio para alinearse, experimentar y volver a evaluar. Repetir ese proceso ayudará a los equipos a adoptar una mentalidad y práctica de mejora continua.

<sup>1</sup>Strathern, Marilyn (1997). "‘Improving ratings’: audit in the British University system". European Review. John Wiley & Sons. 5(3): 305–321. doi:10.1002/(SICI)1234-981X(199707)5:3<305::AID-EURO184>3.0.CO;2-4. S2CID 145644958.

<sup>2</sup>Este informe y los recursos detallados en “Capability catalog” (<https://dora.dev/devops-capabilities/>) pueden ser útiles.

<sup>3</sup>Forsgren, N., Storey, M.A., et. al. "The SPACE of Developer Productivity: There's more to it than you think", 2021. <https://queue.acm.org/detail.cfm?id=3454124>

Presta atención a estas y otras dificultades a la hora de usar estas comparaciones:

- **Comparaciones no equivalentes.** Es poco probable que comparar aplicaciones únicamente en función de estos grupos sea útil. Al hacerlo, se descarta el contexto de cada aplicación de formas que podrían ser negativas para el objetivo de mejorar.
- **Definir métricas como un objetivo.** Ignorar la ley de Goodhart y hacer afirmaciones generales como "todas las aplicaciones deben tener un rendimiento ‘extraordinario’ para fin de año" aumenta la probabilidad de que los equipos intenten manipular las métricas.
- **Una métrica para gobernarlas a todas.** Intentar medir sistemas complejos con la "única métrica que importa". Usar una combinación de métricas para comprender mejor.<sup>3</sup>
- **Métricas con alcance limitado.** Las personas suelen medir lo que sea más sencillo medir y no lo que sea más significativo.
- **Usar el sector como un escudo contra la mejora.** Por ejemplo, algunos equipos en sectores muy regulados podrían usar las reglamentaciones como una razón para no alterar el statu quo.

Para obtener más detalles sobre los hallazgos y consejos para impulsar una mentalidad y práctica de mejora continua, consulta “How to transform” en: [dora.dev/devops-capabilities/cultural/devops-culture-transform/](https://dora.dev/devops-capabilities/cultural/devops-culture-transform/).

## Introducción

Todos los años realizamos uno o más análisis de grupos para encontrar tendencias comunes en todas las aplicaciones. Recomendamos usar estos análisis para comprender tu situación, pero sin obsesionarte con esas comparaciones. Las mejores comparaciones son aquellas que se realizan con el paso del tiempo en las mismas aplicaciones, en lugar de entre distintas aplicaciones, que siempre tendrán diferentes contextos.

Los equipos desarrollan software para los usuarios, quienes, en última instancia, juzgarán la confiabilidad y utilidad del servicio. Los equipos que se enfocan en las necesidades de los usuarios están mejor preparados para desarrollar el producto adecuado. Combinar el enfoque en los usuarios con el rendimiento operativo y de la entrega de software significa que esos equipos también están preparados para desarrollar correctamente el producto.

**Los equipos que se enfocan en las necesidades de los usuarios desarrollan el producto adecuado Y lo hacen de la manera correcta.**

## Resultados

### Rendimiento de la entrega de software

Usamos las siguientes mediciones para evaluar el rendimiento de la entrega de software:

- **Plazo de entrega del cambio:** cuánto demora un cambio en pasar de confirmado a implementado
- **Frecuencia de la implementación:** con qué frecuencia pasan los cambios a producción
- **Tasa de errores de cambio:** con qué frecuencia se introduce en una implementación un error que requiere intervención inmediata
- **Tiempo de recuperación para implementaciones con errores:** cuánto demora la recuperación después de una implementación con errores

Un enfoque común para mejorar estas cuatro mediciones es reducir el tamaño del lote de cambios de una aplicación.<sup>4</sup> Los cambios más pequeños son más fáciles de racionalizar y hacer avanzar por el proceso de entrega. También es más sencillo recuperarse después de un error cuando los cambios son pequeños. Los equipos deberían hacer cada cambio lo más pequeño posible para que el proceso de entrega sea rápido y estable. Trabajar de esta forma contribuye tanto a la velocidad del cambio como a su estabilidad.

<sup>4</sup>A veces, una función puede desglosarse en muchos cambios que se entregan de manera independiente. Nuestras mediciones del rendimiento de la entrega de software evalúan los cambios hechos a una aplicación o servicio.

Este año, definimos mejor las mediciones del rendimiento de la entrega de software. Lee más sobre esos cambios en “Una mejor definición de cómo medimos el rendimiento de la entrega de software”, en el [Apéndice](#).

Este es un resumen de cómo les está yendo a los encuestados de este año con el rendimiento de la entrega de software:

Nivel de rendimiento	Frecuencia de implementación	Plazo de entrega de los cambios	Tasa de errores de cambio	Tiempo de recuperación para implementaciones con errores	% de encuestados
Élite	Según demanda	Menos de un día	5%	Menos de una hora	18%
Alto	Entre una vez al día y una vez por semana	Entre un día y una semana	10%	Menos de un día	31%
Medio	Entre una vez por semana y una vez por mes	Entre una semana y un mes	15%	Entre un día y una semana	33%
Bajo	Entre una vez por semana y una vez por mes	Entre una semana y un mes	64%	Entre un mes y seis meses	17%



## Rendimiento operativo

Para evaluar el rendimiento operativo, les preguntamos a los encuestados con qué frecuencia su servicio hace lo siguiente:

- Recibe informes sobre usuarios finales que no están satisfechos con la confiabilidad del sistema.
- No está disponible, funciona más lento de lo esperado o tiene errores.

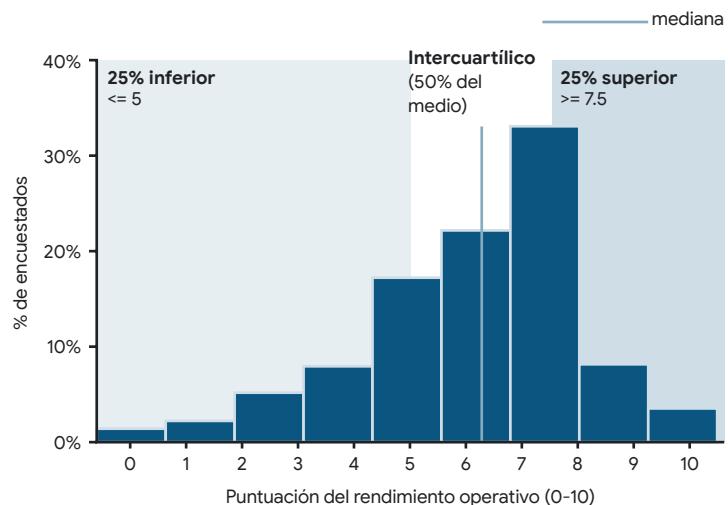
Para explorar de qué manera el rendimiento operativo predice el rendimiento organizativo, consulta el [capítulo 5, La confiabilidad desbloquea el rendimiento](#).

## Enfoque en los usuarios

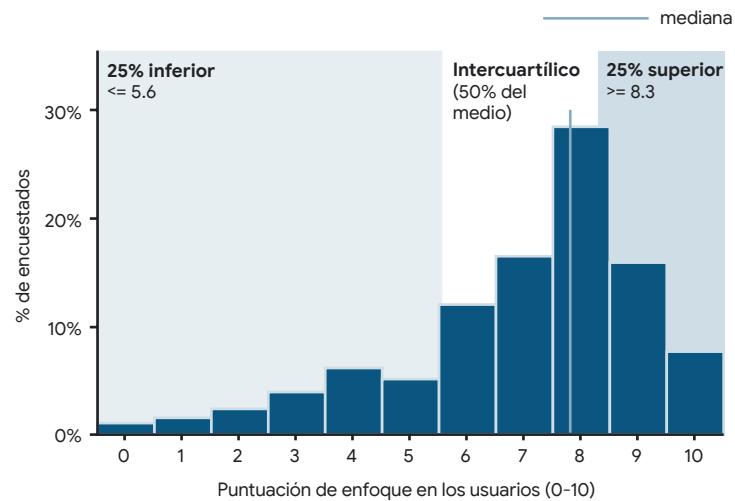
Una aplicación o servicio que se enfoca en los usuarios se desarrolla con el usuario final en mente. Desarrollar un producto como este requiere una buena noción de qué necesitan los usuarios para incorporarlo en la hoja de ruta del producto. Para evaluar qué tan enfocados están los encuestados en los usuarios, preguntamos en qué medida lo siguiente es verdadero:

- Su equipo comprende bien qué quieren lograr los usuarios.
- El éxito de su equipo se evalúa de acuerdo con el valor que proporciona a su organización y a los usuarios de la aplicación.
- Las especificaciones (por ejemplo, la planificación de requisitos) se revisan de manera continua y las prioridades se modifican según los indicadores de usuarios.

Este es un resumen de cómo les está yendo a los encuestados de este año con el rendimiento operativo:



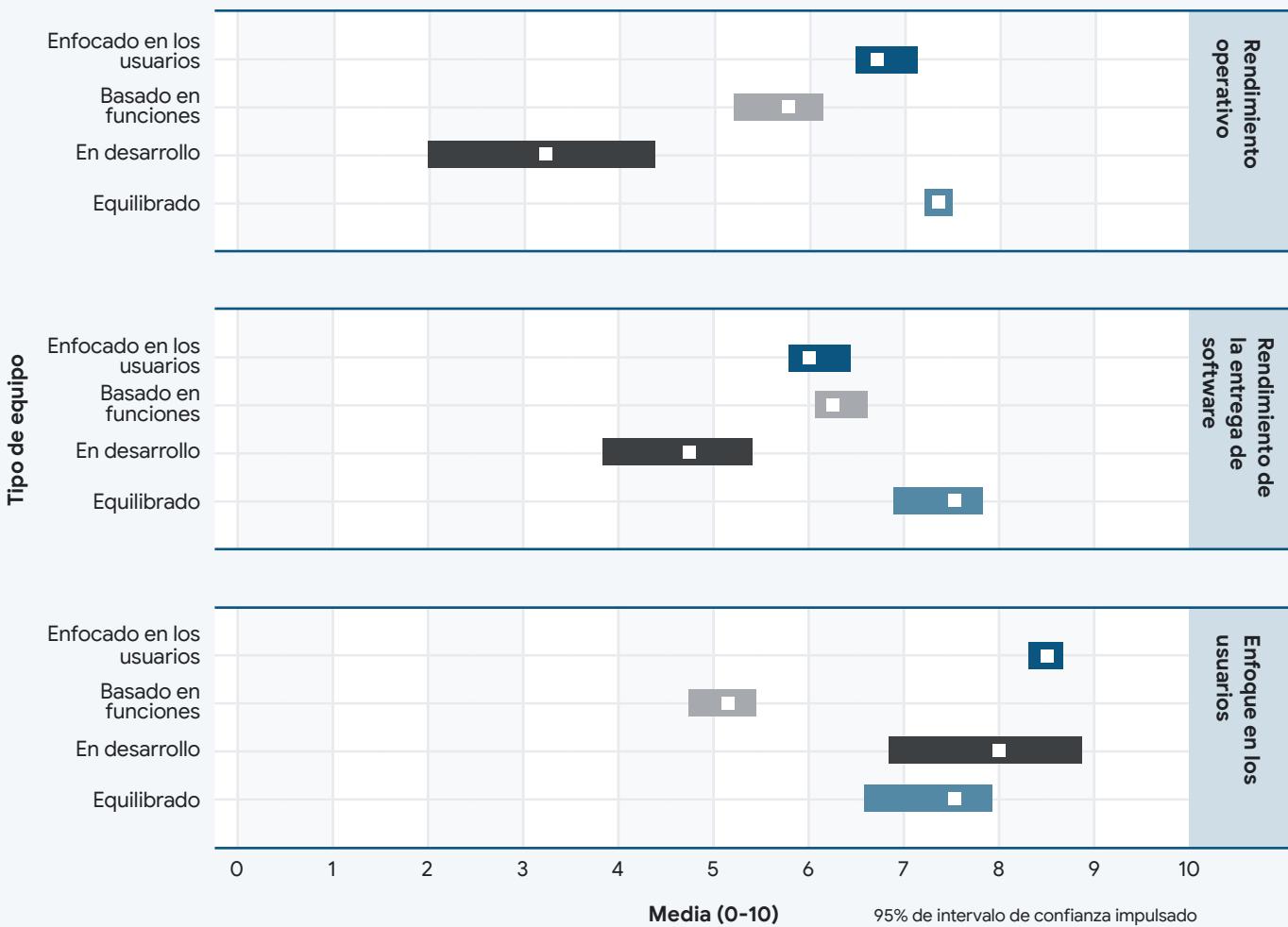
Este es un resumen de qué tan enfocados están los encuestados de este año en los usuarios:



## Agrupar tipos de equipos

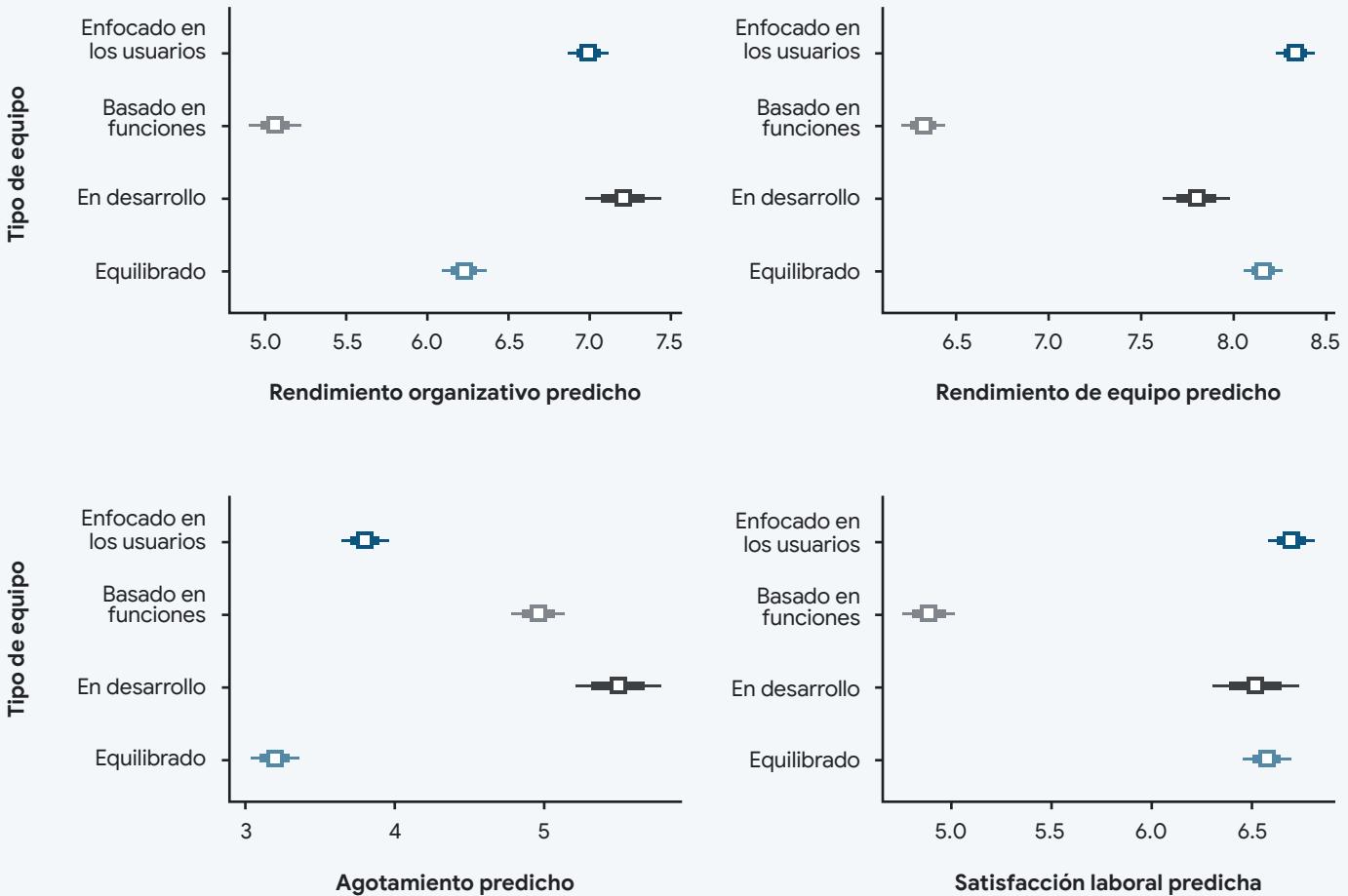
Comparar el rendimiento de la entrega de software, el rendimiento operativo y el enfoque en los usuarios como una unidad revela cuatro tipos de equipos. Como todas las mediciones que se tienen en cuenta para crearlos, estos tipos están al nivel de una aplicación o servicio.

Denominamos a los tipos de equipos "Enfocado en los usuarios", "Basado en funciones", "En desarrollo" y "Equilibrado".



Piensa en las métricas de rendimiento que analizamos como perillas que una organización o equipo puede ajustar para cambiar el rendimiento organizativo y de los equipos, así como el bienestar de las personas en el equipo.

Los siguientes gráficos muestran los resultados de rendimiento que se predijeron para cada tipo de equipo.



\* El punto representa la estimación puntual para la media del tipo de equipo. Intervalo grueso en el que cae el 66% de las simulaciones. Intervalo delgado en el que cae el 89% de las simulaciones.

Cada tipo de equipo tiene características únicas, representa una parte sustancial de nuestros encuestados y tiene resultados diferentes. Es probable que tu equipo no coincida exactamente con uno solo y tampoco esperamos que tu tipo de equipo no cambie con el paso del tiempo.



# ¿Qué significan estos resultados?

## » **Tipo de equipo enfocado en los usuarios**

Este tipo de equipo muestra el mayor enfoque en las necesidades de los usuarios. Aunado a un rendimiento operativo y de la entrega de software sólidos, este enfoque predice los niveles más altos de rendimiento organizativo. Sin embargo, se observa un poco más de agotamiento que en el equipo equilibrado. Mejorar el rendimiento operativo o el de la entrega de software podría ser la mejor forma de reducir el agotamiento en esos equipos.

## » **Tipo de equipo basado en funciones**

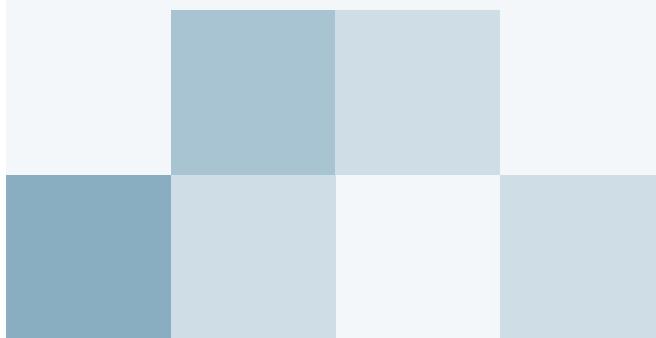
Este tipo de equipo prioriza el envío de funciones. Un enfoque incesante en el envío podría distraer al equipo y evitar que se satisfagan las necesidades de los usuarios, como demuestran las cifras más bajas de enfoque en los usuarios y rendimiento operativo. Esos tipos de equipo informan algunos de los niveles más altos de agotamiento y los niveles más bajos de satisfacción laboral, rendimiento del equipo y rendimiento organizativo. Los empleados valoran la entrega de valor y no solo de funciones. Los equipos basados en funciones podrían beneficiarse de reflexionar sobre las necesidades de sus usuarios como una forma de obtener más valor de las funciones que se envían.

## » **Tipo de equipo en desarrollo**

Este tipo de equipo logra un buen rendimiento organizativo al enfocarse en las necesidades de los usuarios de la aplicación. Sin embargo, es posible que sigan desarrollando su adaptación del producto al mercado o capacidades técnicas. Esos equipos se encuentran con más frecuencia en organizaciones más pequeñas. Su rendimiento de la entrega de software y operativo fue más bajo, y los equipos que trabajan en esas aplicaciones informan niveles de agotamiento más altos que los tipos de equipo equilibrado o enfocado en los usuarios. Es posible que estos equipos tengan procesos muy pesados o tareas laboriosas que podrían automatizarse para mejorar el rendimiento operativo y de la entrega de software.

## » **Tipo de equipo equilibrado**

Este tipo de equipo mostró un enfoque equilibrado y sustentable. Usan la tecnología de forma sustentable para lograr un buen rendimiento organizativo, del equipo y una buena satisfacción laboral. Estos equipos también informan los niveles más bajos de agotamiento. Además, adaptaron sus capacidades de forma tal que permite que haya un buen rendimiento en las tres mediciones. Aumentar el enfoque en los usuarios podría ser el camino a un mejor rendimiento organizativo.



# Enfocarse en los usuarios predice el rendimiento organizativo

## Conclusiones

Las organizaciones pueden obsesionarse con las tendencias de administración y tecnología más recientes cuando buscan mejorar la productividad de los desarrolladores y el rendimiento organizativo. Nuestra investigación muestra que enfocarse en los usuarios para desarrollar aplicaciones y servicios es uno de los indicadores más sólidos del rendimiento organizativo a nivel general. Para mejorar el rendimiento, desarrolla una comprensión profunda de los usuarios; luego, modifica e incorpora de manera iterativa sus comentarios.

## Introducción

El movimiento de DevOps comenzó como una forma de fomentar una mejor colaboración entre equipos de desarrollo y operativos para mejorar el valor del usuario. La alineación generó buenos resultados al principio y permitió que se ampliaran ideas y capacidades de DevOps más allá de esos dos departamentos. En la actualidad, las organizaciones de tecnología de alto rendimiento reconocen la importancia de la alineación entre todos los equipos para alcanzar objetivos organizativos.

Estudiamos tres características fundamentales de enfocarse en los usuarios:

- Qué tan bien los equipos comprenden las necesidades de sus usuarios
- Qué tan bien alineado está el equipo en relación con satisfacer las necesidades de los usuarios
- Cómo se usan los comentarios de los usuarios a la hora de priorizar el trabajo

Los equipos con un sólido enfoque en los usuarios tienen un  
**40%**  
más de rendimiento organizativo

# Resultados

Observamos que enfocarse en los usuarios para desarrollar software genera aumentos considerables en el rendimiento. Las organizaciones pueden experimentar una cascada de beneficios cuando priorizan al usuario. Los comentarios de los usuarios ayudan a los equipos a priorizar proyectos y a crear productos y servicios que satisfagan las necesidades de los usuarios. Este enfoque lleva a una mejor experiencia y mayor satisfacción del usuario, y mayores ingresos.

## El efecto de enfocarse en los usuarios en...

Rendimiento organizativo		Aumento sustancial
Rendimiento del equipo		Aumento sustancial
Rendimiento de la entrega de software		Aumento menor
Rendimiento operativo		Aumento sustancial
Desarrollo basado en troncales		Aumento sustancial
Prácticas de confiabilidad		Aumento sustancial
Integración continua		Aumento sustancial
Entrega continua		Aumento sustancial
Arquitectura con acoplamiento bajo		Aumento sustancial
Agotamiento		Reducción menor*
Satisfacción laboral		Aumento sustancial
Productividad		Aumento sustancial

\* Reducir el agotamiento es bueno.

## ¿Qué significan estos resultados?

Enfocarse en las necesidades de los usuarios es un buen indicador del rendimiento organizativo general. Para lograr un enfoque sólido en los usuarios, se requieren incentivos, alineación y formas de trabajo adecuados. Un enfoque en el usuario puede fundamentar e impulsar mejoras en todas las capacidades técnicas, culturales y de procesos que se analizan en este estudio.

A continuación, exploramos cómo esos resultados podrían afectar a los distintos equipos en tu organización:

### Equipos de desarrollo y entrega de productos

Enfocarse en el usuario ayuda a garantizar que los equipos de desarrollo y entrega de productos desarrollen los productos adecuados para sus usuarios y, con suerte, que lo hagan de manera sustentable. Los equipos equilibrados hacen justamente eso. Muestran solidez en la entrega y el rendimiento operativo y organizativo, con un enfoque robusto en las necesidades de los usuarios. Los miembros de esos equipos se benefician de una comprensión clara de las necesidades de los usuarios y la habilidad de ajustar planes en función de los comentarios de los usuarios.

Los resultados muestran que los equipos basados en funciones no lograron un rendimiento organizativo alto. Al parecer, esos equipos priorizan demás el rendimiento de la entrega en detrimento del rendimiento organizativo y el bienestar de los empleados en la organización.

### Equipos operativos

Es posible que los equipos que se enfocan en el rendimiento operativo trabajen arduamente para optimizar las métricas del sistema, como el uso de la

CPU. Pero si no comprenden qué espera el usuario de un servicio, podrían recibir informes frecuentes de usuarios sobre un rendimiento bajo. Las prácticas de ingeniería de confiabilidad de sitios (SRE), como identificar indicadores de nivel de servicio que les importan a los usuarios y definir objetivos de nivel de servicio para satisfacer a un usuario típico, pueden ayudar a los equipos operativos a tener una mentalidad más enfocada en los usuarios.



### Equipos de ingeniería de plataformas

Los equipos de ingeniería de plataformas podrían adoptar un enfoque del tipo "deséralala y vendrán". Un enfoque más exitoso podría ser tratar a los desarrolladores como usuarios de su plataforma. Para cambiar el enfoque, los equipos de ingeniería de plataformas deben comprender cómo trabajan los desarrolladores para identificar y eliminar de forma correcta las áreas de fricción. Los equipos pueden usar las mediciones de rendimiento operativo y entrega de software como indicadores para supervisar si las iniciativas de la plataforma ayudan a los equipos a lograr mejores resultados.



### Líderes

Mediante la creación de estructuras de incentivos para recompensar a los equipos por entregar valor a los usuarios, los líderes pueden ayudar a crear un entorno en el que prospere un enfoque en los usuarios. Sin estas estructuras, es posible que los equipos sientan que simplemente están midiendo la cantidad de funciones entregadas o una reducción en las interrupciones del servicio. DORA estudió el rol del liderazgo transformacional<sup>1</sup> y tiene consejos para los líderes que están listos para mejorar esa capacidad. Para obtener más información, consulta "DevOps Capabilities: Transformational Leadership" en <https://dora.dev/devops-capabilities/cultural/transformational-leadership>



## Recursos para comenzar

Desarrollar capacidades más enfocadas en los usuarios es un motor importante para el éxito. Nuestros hallazgos de 2023 refuerzan los de 2018,<sup>2</sup> cuando observamos que las capacidades de administración lean de productos predicen el rendimiento organizativo y de la entrega de software.

Mejora el rendimiento de tu equipo mediante la adopción de capacidades enfocadas en los usuarios como comentarios de los clientes,<sup>3</sup> visibilidad del trabajo en la cadena de valor,<sup>4</sup> trabajo en lotes pequeños<sup>5</sup> y experimentación del equipo.<sup>6</sup>

<sup>1</sup>Informe State of DevOps 2017. <https://dora.dev/publications/pdf/state-of-devops-2017.pdf>. 12-19.

<sup>2</sup>Informe Accelerate: State of DevOps 2018: Strategies for a New Economy. <https://dora.dev/publications/pdf/state-of-devops-2018.pdf>. 49-51.

<sup>3</sup>"Customer feedback". <https://dora.dev/devops-capabilities/process/customer-feedback/>

<sup>4</sup>"Visibility of work in the value stream". <https://dora.dev/devops-capabilities/process/work-visibility-in-value-stream/>

<sup>5</sup>"Working in small batches". <https://dora.dev/devops-capabilities/process/working-in-small-batches/>

<sup>6</sup>"Team experimentation". <https://dora.dev/devops-capabilities/process/team-experimentation/>

## Capítulo 3

# Las capacidades técnicas predicen el rendimiento

## Conclusiones

Es probable que invertir recursos y esfuerzos en la integración continua, una arquitectura con acoplamiento bajo y una mayor velocidad de la revisión de código lleve a muchos resultados beneficiosos, como mejoras en el rendimiento organizativo, operativo, de los equipos y de la entrega de software. Al hacerlo, nada se ve perjudicado y, a menudo, se observan beneficios en el bienestar de las personas que trabajan en la aplicación o servicio.

## Introducción

En el resumen ejecutivo, explicamos las capacidades técnicas que estudiamos y cómo afectaron a las distintas mediciones de rendimiento y bienestar. Un componente central de DORA siempre ha sido la exploración y cuantificación del grado en el cual los distintos procesos y las capacidades técnicas predicen el rendimiento.

**Este año, estudiamos la manera en que las siguientes capacidades técnicas predicen el rendimiento:**

- Inteligencia artificial
- Desarrollo basado en troncales
- Arquitectura con acoplamiento bajo
- Integración continua
- Revisión de código rápida

**Analizamos cómo predecían las siguientes mediciones de rendimiento:**

- Rendimiento de los equipos
- Rendimiento organizativo
- Rendimiento de la entrega de software
- Rendimiento operativo

**Además, probamos la conexión entre esas capacidades y un número de indicadores para determinar cómo afectaron a las personas en su trabajo:**

- Agotamiento
- Productividad
- Satisfacción laboral

# Resultados

Las capacidades y los procesos técnicos que analizamos tienen efectos variados, pero en líneas generales positivos en las mediciones de rendimiento clave.

Capacidades y procesos técnicos	Efecto en el rendimiento de equipos	Efecto en el rendimiento organizativo	Efecto en el rendimiento de la entrega de software	Efecto en el rendimiento operativo				
IA		No se observó ningún efecto		Aumento menor		Reducción menor		Reducción sustancial
Integración continua		Aumento menor		Aumento menor		Aumento menor		No se observó ningún efecto
Velocidad de la revisión de código		Aumento menor		No se observó ningún efecto		Aumento sustancial		Aumento sustancial
Arquitectura con acoplamiento bajo		Aumento sustancial		Aumento sustancial		Aumento menor		Aumento sustancial
Desarrollo basado en troncales		Aumento menor		Aumento menor		Aumento menor		Reducción menor

Los equipos con acoplamiento bajo (o quizás los equipos que tienen una arquitectura con acoplamiento bajo) pueden hacer cambios significativos en sus sistemas sin la participación de otros equipos. Esto les permite trabajar más rápido. Cuando los expertos en la materia están más cerca del equipo, pueden revisar el código más rápido porque tienen una mejor comprensión del impacto de los cambios. Un diseño con acoplamiento bajo permite que el equipo pruebe, desarrolle e implemente sin que otros equipos sean un posible cuello de botella.

Incluso con el impacto más pequeño de nuestros cambios en una arquitectura con acoplamiento bajo, debemos garantizar que no se generen conflictos con los otros desarrolladores del equipo. Los equipos que trabajan en lotes pequeños reducen las oportunidades de conflicto, lo que garantiza que, en cada confirmación, se desarrolle el software y se activen pruebas automatizadas, lo que brinda comentarios rápidos a los desarrolladores.

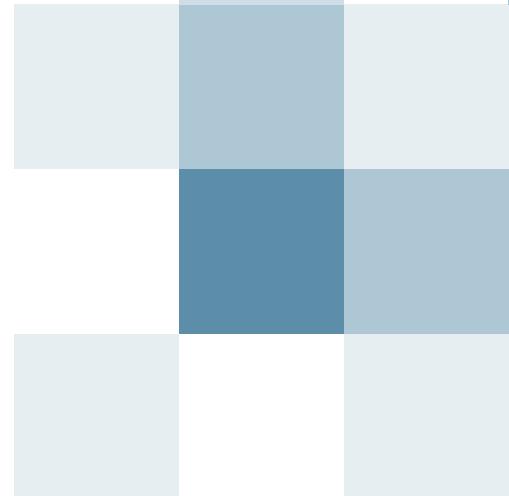
Los equipos con tiempos de revisión de código más cortos tienen un rendimiento un 50% mejor en la entrega de software. Los procesos eficientes de revisión de código llevan a mejoras en el código, transferencia de conocimientos, responsabilidad compartida de código, responsabilidad del equipo y transparencia.

¿Las revisiones de código son un cuello de botella?  
Cuando evalúas el proceso de revisión de código y sus efectos en el plazo de entrega de los cambios, puedes obtener estadísticas sobre oportunidades de mejora. Ten en cuenta las siguientes preguntas:

- ¿Las revisiones de código entre pares están incorporadas en tu proceso?
- ¿Cuánto tiempo hay entre la finalización del código y su revisión?
- ¿Cuál es el tamaño promedio del lote de las revisiones de código?
- ¿Cuántos equipos participan de las revisiones?
- ¿Cuántas ubicaciones geográficas diferentes hay en las revisiones?
- ¿El equipo mejora la automatización de calidad del código en función de las sugerencias de las revisiones de código?

Se observó que, si pasa mucho tiempo entre la finalización del código y su revisión, hay un impacto negativo en la efectividad del desarrollador y la calidad del software entregado. La participación de varios equipos en distintas ubicaciones geográficas lleva a un plazo más prolongado, menos compromiso en el proceso y costos más altos.<sup>1</sup>

En 2022, observamos que las capacidades técnicas dependen una de la otra. Mejorar la velocidad de revisión de código puede contribuir a mejorar distintas capacidades técnicas, incluidos la capacidad de mantenimiento del código, la cultura de aprendizaje (transferencia de conocimientos) y el desarrollo de una cultura generativa.<sup>2</sup>



Las revisiones de código más rápidas son uno de los beneficios de los equipos con acoplamiento bajo, que llevan también a mejoras significativas en el rendimiento operativo y de la entrega de software. Hay distintos caminos para mejorar la eficiencia de las revisiones de código. Cuando el código que se revisa solo afecta el alcance de la arquitectura del equipo, el revisor tiene una mejor comprensión del impacto que tendrá el código en el sistema. Mientras más pequeña sea la revisión, más fácil será para el revisor comprender las implicaciones del cambio. Trabajar en lotes pequeños mejora el ciclo de comentarios, la eficiencia y el enfoque del equipo.<sup>3</sup> La programación en pareja es una práctica que puede reducir el tiempo de revisión de código sin importar la arquitectura o las prácticas actuales de integración.<sup>4</sup>

Además, esas capacidades y procesos no muestran un impacto negativo en el bienestar de las personas que hacen el trabajo. De hecho, la mayoría predice mejoras en el bienestar de las personas.

<sup>1</sup> "Investigating the effectiveness of peer code review in distributed software development based on objective and subjective data". <https://jserd.springeropen.com/articles/10.1186/s40411-018-0058-0>

<sup>2</sup> "Expectations, Outcomes, and Challenges of Modern Code Review". <https://dl.acm.org/doi/10.5555/2486788.2486882>

<sup>3</sup> "Working in small batches". <https://dora.dev/devops-capabilities/process/working-in-small-batches/>

<sup>4</sup> "On Pair Programming". <https://martinfowler.com/articles/on-pair-programming.html>

Capacidades y procesos técnicos	Efecto en el agotamiento*	Efecto en la satisfacción laboral	Efecto en la productividad
IA	Reducción menor	Aumento menor	Aumento menor
Integración continua	Sin efecto	Aumento menor	Sin efecto
Velocidad de la revisión de código	Reducción sustancial	Aumento menor	Aumento menor
Arquitectura con acoplamiento bajo	Reducción sustancial	Aumento sustancial	Aumento sustancial
Desarrollo basado en troncales	Aumento sustancial	Sin efecto	Sin efecto

\* Notarás que el esquema de colores está invertido para el agotamiento. Esto se debe a que la reducción del agotamiento es algo bueno.

Observamos que el uso de arquitectura con acoplamiento bajo, integraciones continuas y revisiones de código eficientes permiten que los equipos mejoren los resultados organizativos y, al mismo tiempo, conserven o incluso mejoren su bienestar.

Cuando los equipos tienen la autonomía para mejorar y mantener un sistema confiable que brinda valor a sus usuarios, observan mejoras en la satisfacción laboral y en el rendimiento de equipos y de la entrega de software. La arquitectura tiene un rol considerable en la habilidad de un equipo para enfocarse en el usuario y mejorar la entrega de software. Al ser pequeños y enfocarse en el usuario, los equipos observaron mejoras significativas en el desarrollo basado en troncales, la arquitectura con acoplamiento bajo, la integración continua, la entrega continua y la SRE. Si quieras mejorar las capacidades técnicas, proporciona oportunidades para que los equipos experimenten y mejoren de manera continua.<sup>5</sup>



<sup>5</sup> "Team experimentation". <https://dora.dev/devops-capabilities/process/team-experimentation>

# Beneficios de la entrega continua

Autor: Dave Farley

El principio fundamental de la entrega continua (CD) es trabajar para que el software esté siempre en un estado que se pueda lanzar. Para lograrlo, debemos trabajar con alta calidad. De esa forma, cuando detectamos un problema, es fácil solucionarlo para recuperar ese estado de manera rápida y sencilla.

Para mantener el software en ese estado ideal listo para lanzar, debemos trabajar para establecer comentarios rápidos y una recuperación de fallas muy veloz.

Como lector del informe de este año, supongo que esas ideas te resultan conocidas. Las métricas de **estabilidad** (tasa de errores de cambio y tiempo de recuperación para implementaciones con errores) están relacionadas principalmente con la calidad, y las métricas de **capacidad de procesamiento** (plazo de entrega para los cambios y frecuencia de implementación) están relacionadas con los comentarios y la facilidad de detección de cualquier problema.

Si haces CD, tendrás una puntuación alta en **estabilidad y capacidad de procesamiento**. Si tienes puntuaciones altas en **estabilidad y capacidad de procesamiento**, es casi imposible que no hagas CD para lograrlos.

El análisis de este año incluye una revisión de cómo las capacidades impulsan el rendimiento mediante la identificación de mediadores de cada capacidad. CD, la capacidad de lanzar cambios de todo tipo a pedido de manera rápida, segura y sustentable, es un mediador considerable de muchas capacidades técnicas. En otras palabras, esas capacidades funcionan porque crean un entorno que posibilita

la CD. La práctica de la CD, a su vez, proporciona el mecanismo por el cual esas capacidades pueden predecir un rendimiento más sólido de la entrega de software.

La capacidad de lanzar es un estándar importante para cumplir en general el desarrollo de software. Por eso, se enfatiza en la CD. Esta característica es relevante porque es una declaración de calidad subjetiva pero definitiva y que se puede aplicar a un contexto. El grado de rigor que define que algo se pueda lanzar puede ser distinto si trabajamos en sistemas fundamentales para la seguridad que si escribimos software para una pastelería. En ambos casos, que algo esté listo para lanzarse indica que hicimos todo lo necesario para establecer que el código está listo y es lo suficientemente bueno y seguro para ponerse a disposición de los usuarios.

Por lo tanto, realizar optimizaciones para que los cambios puedan lanzarse significa optimizar para una definición específica del contexto de calidad mínima aceptable del sistema.

Los equipos que priorizan obtener comentarios rápidos y de alta calidad, y actuar en consecuencia tienen un mejor rendimiento de la entrega de software.

Me sorprende un poco que la integración continua (CI) y el desarrollo basado en troncales no tuvieran un mayor impacto en el rendimiento de la entrega de software. La CI en especial me parece fundamental, por lo que eso desafía mi visión del mundo, pero desenredar esos conceptos es complejo. Por ejemplo, ¿cómo podemos lograr puntuaciones altas en la capacidad de procesamiento si el código no se integra y cómo podemos asegurarnos de que la estabilidad sea alta si no la verificamos? Para mí, la CI nos ayuda a responder esas preguntas, por lo que es un mediador clave del rendimiento de la entrega de software. ¿Será un problema de interpretación o algo más profundo y más importante? ¡Qué misterio!

Capacidades y procesos técnicos	Efecto en el rendimiento de la entrega de software	¿Se mediaron con la entrega continua?*
IA	Reducción menor	<input checked="" type="checkbox"/> No
Integración continua	Aumento menor	<input checked="" type="checkbox"/> Sí, totalmente
Velocidad de la revisión de código	Aumento sustancial	<input checked="" type="checkbox"/> Sí, parcialmente
Arquitectura con acoplamiento bajo	Aumento menor	<input checked="" type="checkbox"/> Sí, parcialmente
Desarrollo basado en troncales	Aumento menor	<input checked="" type="checkbox"/> Sí, totalmente

\* La mediación es una prueba que evalúa posibles mecanismos o rutas subyacentes de un efecto. Puedes decir, por ejemplo: "los datos apoyan la hipótesis de que el efecto del desarrollo basado en troncales en el rendimiento de la entrega de software ocurre por medio de la implementación continua (el mediador)". La mediación está completa cuando todo el efecto se puede explicar con el mediador. La mediación parcial indica que solo una parte del efecto se explica con un mediador.

## Optimizar las capacidades y los procesos organizativos

Sabemos que la cultura impulsa el éxito. Pero ¿qué impulsa la cultura? Esta es una pregunta interesante cuya respuesta es la favorita de todos: depende.

Desde el punto de vista de un profesional, mejorar la forma en que trabaja de manera diaria tiene un impacto positivo en elementos culturales como compartir el riesgo, aumentar la cooperación y establecer la seguridad psicológica. Por ejemplo, integrar cambios con regularidad en la rama principal del sistema de control de versión aumenta la colaboración y el conocimiento compartido. Tener equipos de seguridad que trabajen en conjunto con los desarrolladores y colaboren en la política como código aumenta la confianza entre los equipos y en los cambios que se implementan.

Desde una perspectiva de liderazgo, la cultura comienza con el reconocimiento y la capacitación sobre la importancia de la cultura. El liderazgo transformacional<sup>6</sup> puede ayudar a fomentar un entorno sin culpas en el que se alienten la experimentación y el aprendizaje, además de darles confianza y una voz a los profesionales. La función de los ingenieros es resolver problemas complejos, no solo responder solicitudes de tareas. Para hacerlo, necesitan visibilidad del negocio y la autonomía para tomar medidas. En última instancia, la cultura depende del liderazgo.

Idealmente, los mejores resultados se obtienen de analizar la cultura desde una perspectiva de arriba abajo y de abajo arriba.

<sup>6</sup> "Transformational leadership". <https://dora.dev/devops-capabilities/cultural/transformational-leadership>

# Inteligencia artificial (IA)

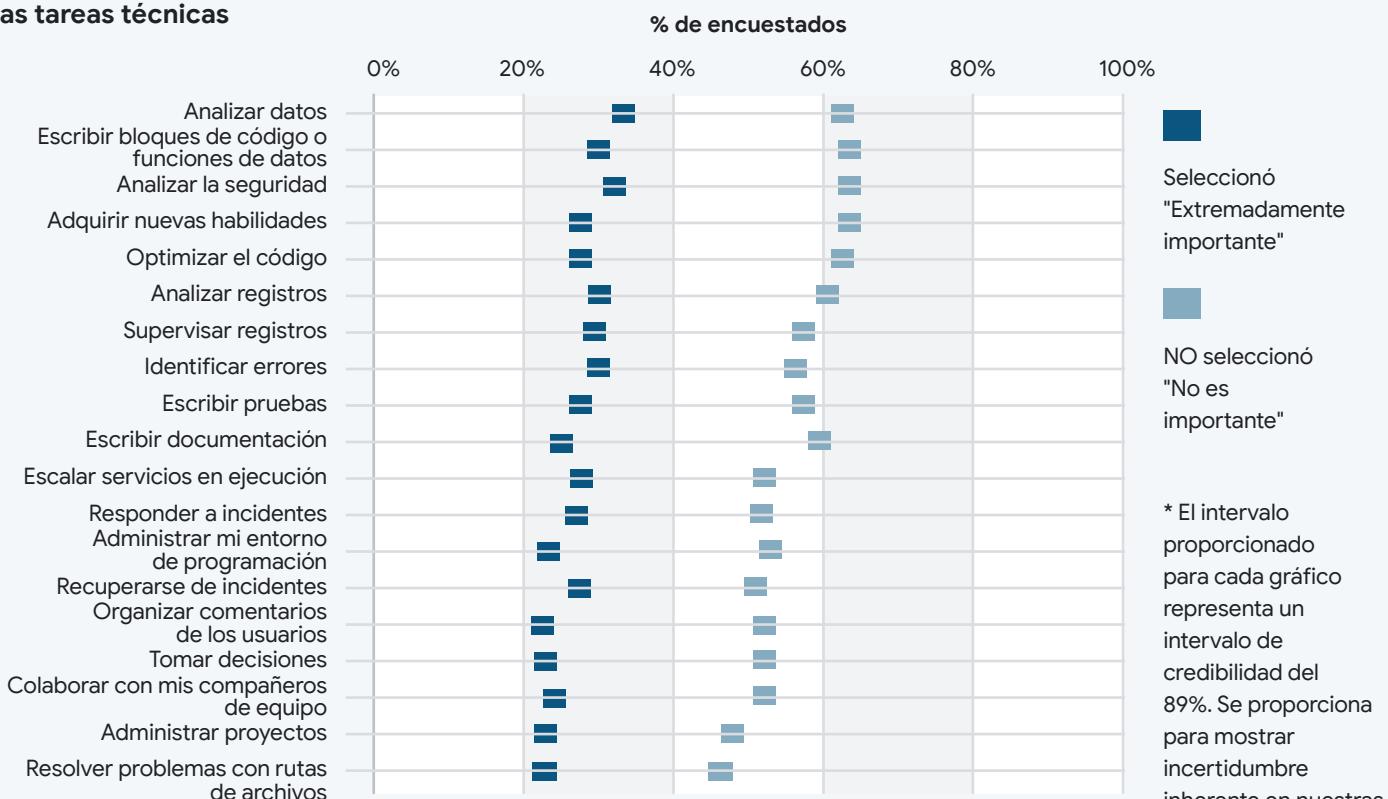
Algunos analistas y tecnólogos tienen la hipótesis de que la IA mejorará el rendimiento de los equipos de software sin afectar de manera negativa el bienestar profesional. Hasta el momento, la evidencia de la encuesta no apoya esa teoría, sino que sugiere que la IA mejora ligeramente las mediciones de bienestar individual (como el agotamiento y la satisfacción laboral), pero tiene un efecto neutro o quizás negativo en los resultados a nivel del grupo (como el rendimiento del equipo y de la entrega de software).

Creemos que la etapa inicial de la adopción de herramientas de IA entre empresas podría ayudar a explicar esas diferencias en la evidencia. Es probable que algunas empresas grandes estén evaluando distintas herramientas potenciadas por IA antes de tomar una decisión sobre si usarlas de manera más generalizada. Hay mucho entusiasmo sobre el potencial de las herramientas de desarrollo de IA, como demostró el hecho de que la mayoría de las personas incorporó al menos un poco de IA en las tareas incluidas en la encuesta. Esto se muestra en el siguiente gráfico. Sin embargo, anticipamos que pasará un tiempo para que las herramientas impulsadas por IA se usen de manera generalizada y coordinada en el sector.

## La importancia de la IA

### La contribución de la IA en las tareas técnicas

En la aplicación o el servicio principal en el que trabajas, ¿qué tan importante es en la actualidad el rol de la inteligencia artificial (IA) a la hora de contribuir en las siguientes tareas?



Estamos muy interesados en ver cómo crece la adopción con el paso del tiempo y el impacto que ese crecimiento tendrá en las mediciones y los resultados de rendimiento que son importantes para las organizaciones.

# La documentación es la base



## Conclusiones

La documentación de calidad es fundamental. Permite la implementación exitosa de las capacidades técnicas y amplifica el impacto que tienen en el rendimiento organizativo. La documentación también tiene un impacto positivo en los resultados, como el rendimiento del equipo, la productividad y la satisfacción laboral. Sin embargo, elevar la calidad de la documentación no genera un mayor bienestar para todos. A medida que mejora la calidad de la documentación, algunos encuestados informan mayores niveles de agotamiento.

## Introducción

Este año, analizamos en más detalle la documentación interna: el conocimiento escrito que usan las personas en la organización todos los días. Estudiamos el impacto de la documentación en las capacidades técnicas y los resultados clave.

Para evaluar la calidad de la documentación, medimos en qué grado la documentación es confiable, fácil de buscar, actualizada y relevante. Después calculamos una puntuación para toda la experiencia de documentación. No evaluamos la documentación página por página, sino como un todo.

# Resultados

## La documentación es la base: mejora y amplifica las capacidades técnicas

Como observamos en 2021<sup>1</sup> y 2022,<sup>2</sup> la calidad de la documentación sigue impulsando la implementación exitosa de las capacidades técnicas que estudiamos.

Como se muestra en la siguiente tabla, la calidad de la documentación también amplifica el impacto de cada capacidad técnica en el rendimiento organizativo, similar a lo que observamos en 2022.<sup>3</sup>

Capacidad técnica	Amplificación del impacto en el rendimiento organizativo*
Integración continua	2.4 veces*
Entrega continua	2.7 veces*
Desarrollo basado en troncales	12.8 veces*
Arquitectura con acoplamiento bajo	1.2 veces*
Prácticas de confiabilidad	1.4 veces*
Contribución de la inteligencia artificial	1.5 veces*

\* Método de cálculo: Impacto de la capacidad técnica con documentación de alta calidad dividido por el impacto de la capacidad técnica con documentación de baja calidad

<sup>1</sup> Informe “Accelerate State of DevOps 2021”. <https://dora.dev/publications/pdf/state-of-devops-2021.pdf>

<sup>2</sup> Informe “Accelerate State of DevOps 2022”. <https://dora.dev/research/2022/dora-report/2022-dora-accelerate-state-of-devops-report.pdf>

<sup>3</sup> Informe de 2022 de profundización de datos “State of DevOps: Documentation is like sunshine”. <https://cloud.google.com/blog/products/devops-sre/deep-dive-into-2022-state-of-devops-report-on-documentation>

## La documentación predice personas y organizaciones productivas y felices

Además de mejorar las capacidades técnicas, observamos que la documentación de calidad tiene un impacto positivo en el bienestar de las personas: reduce el agotamiento, y aumenta la satisfacción laboral y la productividad. Observamos que parte del impacto se debe a que la documentación de calidad aumenta el conocimiento compartido.

Ese efecto no es una gran sorpresa. Es más fácil completar tareas cuando sabes cómo hacerlo, y el trabajo es menos frustrante cuando se comparte el conocimiento.

Aspectos del bienestar	Efecto de la documentación de calidad
Agotamiento	 Reducción sustancial
Satisfacción laboral	 Aumento sustancial
Productividad	 Aumento sustancial

La documentación de calidad también impulsa resultados clave, lo que afecta el rendimiento organizativo, operativo y del equipo.

Resultados clave	Efecto de la documentación de calidad
Rendimiento del equipo	 Aumento sustancial
Rendimiento organizativo	 Aumento sustancial
Rendimiento de la entrega de software	 Sin efecto*
Rendimiento operativo	 Aumento sustancial

\* Seguimos sorprendiéndonos de no encontrar efectos de la documentación de calidad en el rendimiento de la entrega de software. Este es el segundo año en que observamos este comportamiento.

¿A qué se debe este efecto positivo en los tres resultados clave? Como lector, usar documentación clara es benéfico. Es posible que el proceso de escritura también sea un factor. Crear documentación de alta calidad requiere que los equipos primero tomen decisiones sobre los procesos. La documentación puede obligar a los equipos de una organización a discutir de manera explícita qué hacer y cómo hacerlo, y a alinearse.

La documentación de calidad también funciona como un repositorio del conocimiento de los equipos, incluso con el ir y venir de las personas. Ayuda a escalar el conocimiento, tanto en la organización como con el paso del tiempo.

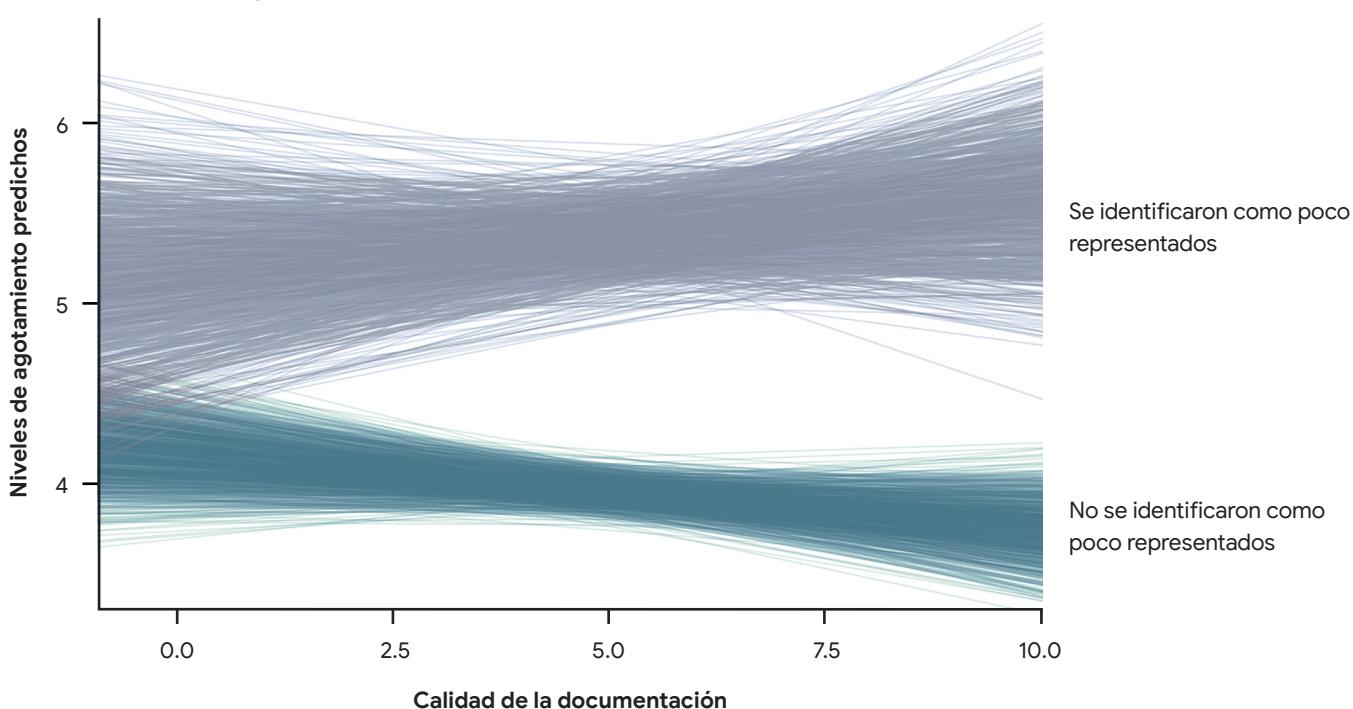
### **¿La documentación está ligada a la reducción del bienestar de algunas personas?**

Observamos una tendencia inesperada cuando prestamos atención a los encuestados que se identificaron como poco representados. Para este grupo, la calidad de la documentación está vinculada a un aumento en el agotamiento.

Para este hallazgo, también analizamos el género y nos sorprendió no encontrar ningún efecto. Todos los encuestados que se identificaron como hombres, mujeres o autodefinieron su género observaron una reducción significativa del agotamiento con documentación de alta calidad. Sin embargo, las personas que se identificaron como poco representadas, sin importar su identidad de género, informaron una tasa de agotamiento mayor en la presencia de documentación de calidad.

En el siguiente gráfico, se muestran predicciones simuladas basadas en nuestros datos. En el conjunto inferior, se observa una reducción del agotamiento para la mayoría de los encuestados a medida que aumenta la calidad de la documentación. Sin embargo, en el conjunto superior, observamos que el agotamiento aumenta significativamente para las personas que se identifican como poco representadas.

En este gráfico se muestran 1,000 líneas simuladas para cada grupo. Las líneas más densas indican que la pendiente es más probable a la luz de nuestros datos.



Este hallazgo es similar para la calidad de la documentación, la cultura generativa y la estabilidad del equipo: a medida que aumentan estos atributos, también lo hace el agotamiento para las personas que se identifican como poco representadas.

#### ¿Qué pasa con la documentación?

Se requiere trabajo para crear y mantener documentación de alta calidad. Es una tarea técnica, con un impacto considerable en las capacidades técnicas, la productividad del equipo y el rendimiento organizativo. Además, es trabajo que puede no ser reconocido de manera constante por su importancia y el impacto que tiene. ¿Las personas que se identifican como poco representadas hacen una cantidad desproporcionada de este trabajo y, si es así, esto ayuda a explicar el efecto en el agotamiento?

¿Es posible que depender del uso de documentación sea problemático? Con una mejor calidad de la documentación, ¿el conocimiento compartido no aumenta para algunos de los encuestados? O, si aumenta, ¿no es suficiente para contrarrestar otros aspectos que generan agotamiento en este grupo?

Es posible que exista un factor completamente diferente que impulse la documentación de calidad, pero también cree o mantenga el agotamiento para encuestados que se identifican como poco representados. Se requieren más estudios.

Al parecer, importa quién eres en el equipo. Hay aspectos del lugar de trabajo, como la documentación de calidad, que tienen beneficios significativos para el equipo y la organización. Pero también podrían estar vinculados con resultados negativos para algunas personas. Exploramos más este tema en el [capítulo 8, Cómo, cuándo y por qué importa quién eres](#).

<sup>4</sup> "Accelerate State of DevOps 2021", 22. <https://dora.dev/publications/pdf/state-of-devops-2021.pdf#page=22>

## Recursos para comenzar

Consulta el informe de 2021 para conocer prácticas que impulsan la documentación de calidad.<sup>4</sup> Este año, también observamos que la distribución del trabajo, incluidos los procesos formales relacionados con la documentación, aumenta considerablemente la calidad de la documentación.

**Existen muchos recursos y capacitaciones sobre la escritura técnica. Puedes obtener más información en los siguientes vínculos:**

- Society for Technical Communications ([stc.org](http://stc.org))
- Cursos de escritura técnica para ingenieros ([developers.google.com/tech-writing](http://developers.google.com/tech-writing))
- Write the docs ([writethedocs.org](http://writethedocs.org))

## Capítulo 5

# La confiabilidad desbloquea el rendimiento

## Conclusiones

Las prácticas de confiabilidad sólidas predicen un mejor rendimiento operativo, organizativo y de los equipos. Los datos muestran que los efectos de mejorar estas prácticas siguen un camino no lineal; es decir, a veces es posible que las mejoras de rendimiento parezcan detenerse mientras las organizaciones desarrollan capacidades más sólidas. Sin embargo, con el paso del tiempo, mantener el compromiso con estas prácticas sigue prediciendo buenos resultados.

## Introducción

“Confiabilidad” es un término ampliamente usado en el espacio de las operaciones de TI. Se define como el grado en que un servicio cumple los objetivos establecidos en relación con mediciones como disponibilidad, rendimiento y corrección. Un enfoque común para lograr resultados de confiabilidad es la SRE, que surgió en Google (<https://sre.google>) y ahora se practica en muchas organizaciones. La SRE prioriza el aprendizaje empírico, la colaboración multifuncional, la amplia dependencia en la automatización y el uso de técnicas de medición, entre ellas los objetivos de nivel de servicio (SLO).

Muchas organizaciones usan prácticas de confiabilidad sin llamarlas SRE. Entre los términos alternativos se incluyen “ingeniería de producción”, “equipos de plataformas”, “equipos de infraestructura” y “TechOps”, entre otros. Para evaluar el alcance de estas prácticas de la forma más objetiva posible, se usa lenguaje neutro y descriptivo en el texto de nuestra encuesta.

También recopilamos datos de los resultados de la ingeniería de confiabilidad: el grado en que los equipos pueden alcanzar sus objetivos de confiabilidad. Tanto las prácticas como los resultados de confiabilidad (que denominamos **rendimiento operativo**) se reflejan en nuestro modelo predictivo junto con otras capacidades.

### Prácticas de confiabilidad

Les pedimos a los encuestados que piensen en la confiabilidad en relación con tres aspectos esenciales de sus operaciones. Primero, ¿tienen planes de mitigación para sus dependencias? Segundo, ¿prueban de manera periódica sus planes de recuperación ante desastres mediante interrupciones simuladas, comutaciones por error prácticas o ejercicios de simulación? Por último, cuando no alcanzan los objetivos de confiabilidad, ¿realizan tareas de mejora o cambian las prioridades y ajustan su trabajo?

Creemos que estas mediciones encapsulan el espíritu de un equipo que sigue principios establecidos de SRE como "aceptar el riesgo" y "medir la felicidad del usuario". Tal equipo establece un objetivo razonable que se alinea con la felicidad del usuario. Luego, realiza pruebas para asegurarse de que puede alcanzar ese objetivo, pero cambia el plan si tiene dificultades. Usamos esto como representación de un equipo que "hace SRE" bien sin vincular evaluaciones de equipos a implementaciones específicas de SRE.

## Resultados

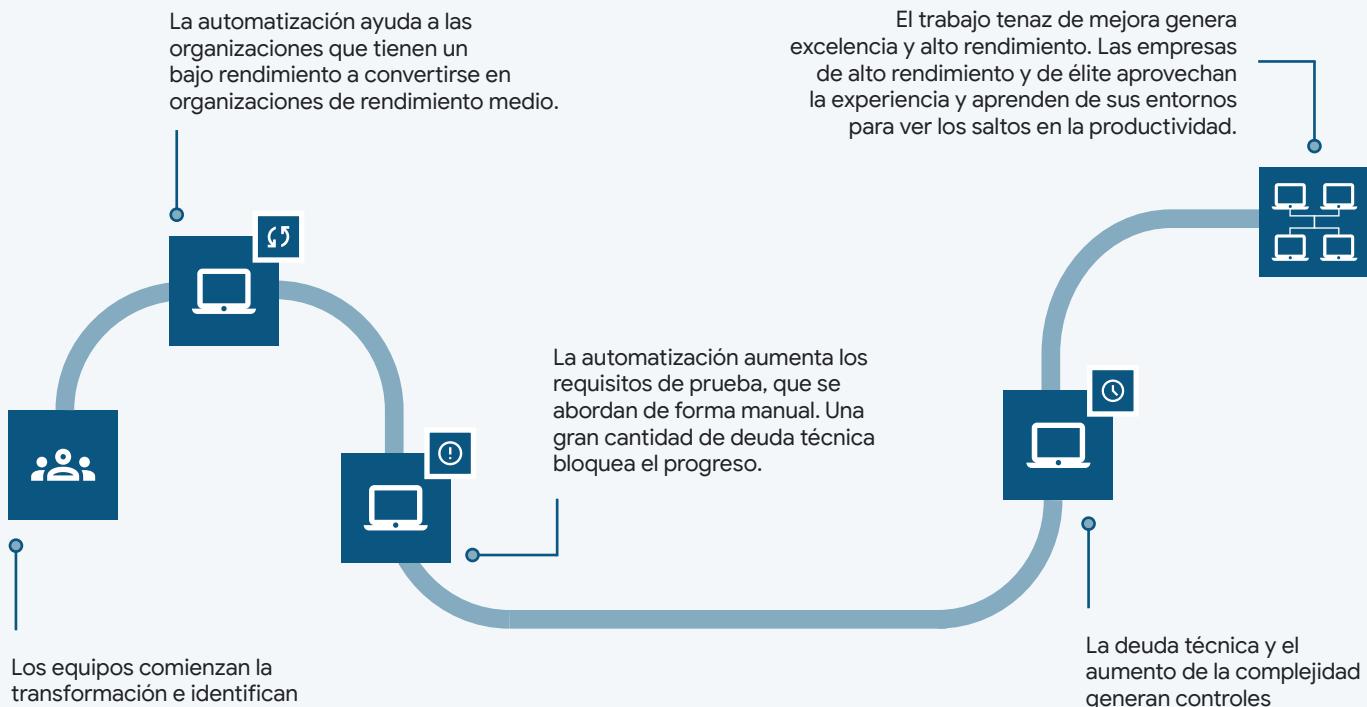
### Confirmar la curva J de las prácticas de confiabilidad

Desde 2018, DORA tiene la teoría de que hay una relación no lineal (figura 1) entre el rendimiento operativo y prácticas como la automatización.

A medida que profundizamos las exploraciones de las prácticas de confiabilidad, observamos evidencia de este patrón en los datos de las encuestas.

En 2022, lo medimos de manera directa. Encuestamos a equipos y notamos que la relación entre las prácticas y los resultados de confiabilidad siguió este tipo de curva no lineal (figura 2). Eso sugirió que los equipos observaron ganancias considerables de confiabilidad únicamente después de adoptar muchas prácticas en la materia.<sup>1</sup> Sin embargo, no creemos que analizar los datos de esta forma muestre el panorama completo. La curva de 2022 dio la impresión de que la SRE es solo para expertos o que no vale la pena invertir en ella, lo que no coincide con la experiencia de muchos equipos de SRE. Se necesitaban más datos.

**Figura 1: Curva J hipotética de 2018**

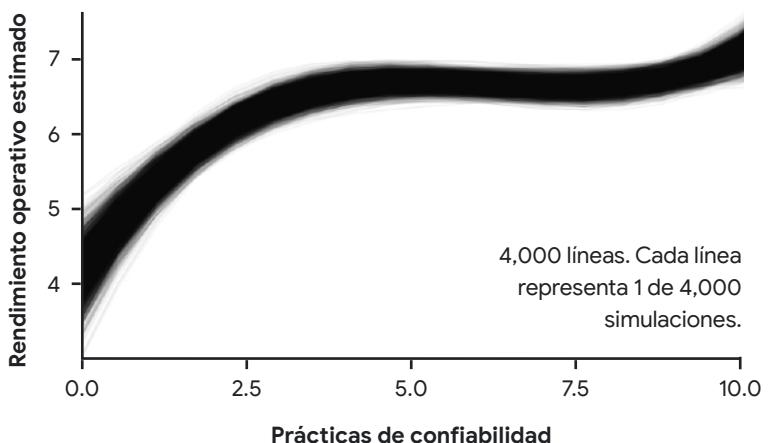


<sup>1</sup> Informe "Accelerate State of DevOps 2022". 27-28. [https://dora.dev/research/2022/dora-report/\\_2022-dora-accelerate-state-of-devops-report.pdf#page=22](https://dora.dev/research/2022/dora-report/_2022-dora-accelerate-state-of-devops-report.pdf#page=22)

**Figura 2: Curva de 2022**



**Figura 3: Curva de 2023**



En 2023, pudimos hacer más preguntas, lo que ayudó a definir mejor una curva que coincide más con las experiencias vividas. La curva nueva se asemeja más a la curva J hipotética de transformación que se describió en el informe de 2018 (consulta la sección [Metodología](#) para obtener más información sobre cómo realizamos el análisis). Esto sugiere que hay ganancias anticipadas de adoptar prácticas de confiabilidad, seguidas de un período de inactividad a medida que la complejidad introduce desafíos nuevos y, finalmente, otro aumento en el rendimiento operativo. Los resultados refuerzan lo que observamos en muchos equipos.

La curva es importante por varias razones:

- Ayuda a las empresas a racionalizar y financiar la adopción inicial de SRE, incluso si no buscan alcanzar niveles extremos de confiabilidad o no esperan realizar grandes inversiones en SRE. Adoptar incluso niveles pequeños de prácticas de confiabilidad puede generar mejoras del rendimiento operativo, lo que tiene aún más efectos beneficiosos en el rendimiento organizativo y de los equipos.

<sup>2</sup> "How to transform". <https://dora.dev/devops-capabilities/cultural/devops-culture-transform/>

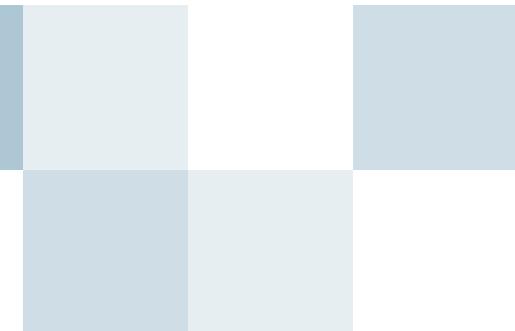
- Prepara a las empresas que *quieren* invertir mucho en confiabilidad para pasar el período de inactividad. Puede ser tentador esperar resultados lineales positivos de inversiones a largo plazo en SRE, pero los datos indican que no es el caso. Cuando los equipos saben de la no linealidad de esta curva con anticipación, pueden tomar una decisión sobre si deben hacer esta inversión y planificar con anticipación para asegurarse de no abandonar la iniciativa antes de aprovechar todos los beneficios.
- Es posible que cambios como estos requieran una transformación cultural.<sup>2</sup> Observamos que el éxito proviene de una combinación del cambio de abajo arriba y arriba abajo. Los equipos pueden adoptar prácticas de confiabilidad y cosechar las recompensas inmediatas, luego el liderazgo puede presumir esos beneficios a otros equipos, reforzarlos y fomentarlos. Esos incentivos y programas estructurados pueden diseñarse con la curva J en mente.

## Prácticas de confiabilidad y bienestar

Las prácticas tradicionales de operaciones son altamente reactivas y a menudo se relacionan más con el estado del sistema técnico que con la felicidad de sus usuarios. Las alertas de guardia para aquello que no tiene un impacto en la experiencia de los usuarios, las tareas manuales repetitivas, el miedo a cometer errores y experiencias similares generan agotamiento y afectan el bienestar de las personas en el equipo.

Observamos lo opuesto en los equipos que aprovechan las prácticas de confiabilidad. Los equipos informan una productividad y satisfacción laboral más altas, y niveles más bajos de agotamiento que sus colegas que no usan estas prácticas. Sospechamos que estas mejoras en el bienestar se deben a algunas prácticas publicadas de SRE:

- Reducir el trabajo repetitivo<sup>3</sup>
- Análisis retrospectivos libres de culpa<sup>4</sup>
- Autonomía del equipo<sup>5</sup>
- Escalamiento sublineal de equipos<sup>6</sup>



<sup>3</sup>Beyer, Betsy et al. Site Reliability Engineering: How Google Runs Production Systems. O'Reilly, 2016, 49–54. <https://sre.google/sre-book/eliminating-toil/>

<sup>4</sup>Ibid. <https://sre.google/sre-book/postmortem-culture/>

<sup>5</sup>Beyer, Betsy et al. The Site Reliability Workbook. O'Reilly, 2018. <https://sre.google/workbook/team-lifecycles/>

<sup>6</sup>Brookbank, James y McGhee, Steve. Enterprise Roadmap to SRE. O'Reilly, 2022, 11. <https://sre.google/resources/practices-and-processes/enterprise-roadmap-to-sre/>

## Rendimiento operativo

También les pedimos a los encuestados que describieran el rendimiento operativo de su servicio. Primero, preguntamos con cuánta frecuencia reciben mensajes directamente de sus usuarios sobre falta de satisfacción relacionada con la confiabilidad del servicio. Después, les preguntamos con cuánta frecuencia su servicio no está disponible, es lento o no opera correctamente.

### Las prácticas de confiabilidad amplifican el rendimiento organizativo y de los equipos por medio del rendimiento operativo

Cuando adoptan prácticas de confiabilidad, los equipos mejoran su rendimiento operativo. Si una organización puede operar su flota de producción de manera eficaz, observamos que eso *amplifica* otros resultados. Si los resultados son altos, las prácticas de confiabilidad los aumentarán. Si los resultados son bajos, las prácticas de confiabilidad no serán de ayuda, sino que se mantendrán iguales.

Los sistemas confiables necesitan contar con capacidades de software adecuadas para tus clientes, las cuales deben entregarse de forma eficaz. Esto tiene sentido porque el propósito de la SRE nunca fue operar en un vacío. Alcanzar los objetivos de confiabilidad es una métrica clave del éxito de los equipos de SRE, y eso se refleja en el rendimiento operativo. Aunque es probable que haya otros beneficios del uso de prácticas de confiabilidad, los datos sugieren que el más importante es el impacto en el rendimiento operativo. Además, el rendimiento operativo aumentado tiene beneficios más allá del estado del servicio. De hecho, observamos evidencia de que el uso de prácticas de confiabilidad predice un mayor bienestar para los profesionales.

### El rendimiento operativo afecta el bienestar

Una percepción común del sector es que los servicios altamente confiables tienen un impacto negativo en el bienestar de los operadores del servicio,

por ejemplo, debido a actividades de guardia o mantenimiento de emergencia fuera del horario laboral. Sin embargo, observamos que el rendimiento operativo alto en realidad genera menos agotamiento, así como mayor productividad y satisfacción laboral. Esto se alinea con el principio de SRE de reducir el trabajo repetitivo;<sup>7</sup> automatizar las partes manuales de la operación es satisfactorio para las personas y también genera una carga continua reducida para el equipo.



### **El rendimiento organizativo y de los equipos amplifica el rendimiento operativo**

Observamos que el rendimiento operativo tiene un efecto positivo considerable en el rendimiento organizativo y de los equipos. Esto no debería ser una sorpresa para los seguidores del movimiento de DevOps. Poder operar la máquina de manera eficaz permite a los equipos lograr más, lo que permite que las organizaciones prosperen.



### **El rendimiento operativo amplifica el rendimiento de la entrega de software**

Si bien el rendimiento de la entrega de software puede mejorar tanto el rendimiento de los equipos como el organizativo, los dos mejoran de manera considerable con el rendimiento operativo. Además, los equipos de entrega de software de alto rendimiento no alcanzarán un rendimiento de equipos y organizativo muy altos sin también lograr un rendimiento operativo alto. Se necesitan los dos. De hecho, los equipos que mejoran su rendimiento de la entrega de software sin los niveles correspondientes de rendimiento operativo terminan teniendo resultados organizativos peores. Por lo tanto, si puedes escribir rápidamente código sorprendente, pero que no puede ejecutarse en producción de forma tal que cumpla con las expectativas del público, no habrá ninguna recompensa del mercado.

## **¿Qué falta y qué sigue?**

Creemos que hay más mediciones que pueden ayudarnos a comprender esas interacciones. Por ejemplo, una pregunta común este año fue el efecto de la administración de costos en estas capacidades y resultados. Algunas organizaciones son más sensibles a los costos que otras y eso tiene implicaciones sobre la manera en que la organización hace planes y toma decisiones. De manera similar, especulamos que las prácticas de confiabilidad podrían surgir de culturas altamente colaborativas,<sup>8</sup> incluso sin que se busquen o planifiquen explícitamente. Queremos comprender mejor de qué manera los equipos modifican sus prácticas actuales de operaciones de TI y cómo esa evolución afecta la confiabilidad del sistema, el rendimiento de los equipos y el bienestar.

Sobre todo, nos gustaría conocer tu opinión. Súmate a nosotros y otros profesionales en [DORA.community](https://dora.community).<sup>9</sup> La SRE todavía es un campo nuevo. Su impacto es diferente en cada organización que adopta prácticas de confiabilidad o en las organizaciones que se dan cuenta de que han estado haciendo SRE todo este tiempo. Los cambios son lentos, y queremos tomar mediciones consistentes para mostrar el progreso con el paso del tiempo. Como comunidad, podemos compartir lo que funciona y ayudarnos a mejorar los unos a los otros a lo largo del camino.

<sup>7</sup> Beyer, Betsy et al. Site Reliability Engineering: How Google Runs Production Systems. O'Reilly, 2016, 49–54. <https://sre.google/sre-book/eliminating-toil/>

<sup>8</sup> Brookbank, James y McGhee, Steve. Enterprise Roadmap to SRE. O'Reilly, 2022, 5. <https://sre.google/resources/practices-and-processes/enterprise-roadmap-to-sre/>

<sup>9</sup> Comunidad DORA, <https://dora.community/>

# Cómo Google hace SRE

En Google, la SRE tienen dos décadas de evolución dentro de una organización en crecimiento y con financiamiento suficiente. El ímpetu por la SRE fue una necesidad para habilitar el hipercrecimiento de la Búsqueda de Google y Google Ads sin invertir mucho dinero. Al comienzo, la naturaleza en tiempo real de esos productos reveló la necesidad de alta confiabilidad, ya que un error temporal en la búsqueda o los anuncios es una pérdida inmediata de clientes, sin segundas oportunidades. Un anuncio dinámico debe calcularse en milisegundos; una búsqueda lenta degrada toda la promesa de la marca de Google.

Al mismo tiempo, la SRE se estaba desarrollando en la concepción de un tipo nuevo de empresa: una empresa basada en ingeniería y con una organización ascendente que eligió desarrollar en lugar de comprar. Google aprovechó un personal de ingenieros de confiabilidad de sitios (SRE) con amplios conocimientos académicos sobre temas de informática, como sistemas distribuidos y diseño de compiladores. Ese ADN cultural proporcionó un entorno enriquecido para que surgiera y prosperara la SRE. Al equipo de SRE se le confiaron no solo las llaves de producción, sino también encontrar formas innovadoras de escalar los sistemas.

Pero ¿cómo escaló verticalmente la SRE con el tiempo? En una palabra: sublinealmente. Es decir, Google no podía duplicar la cantidad de ingenieros de SRE empleados cada vez que Google duplicara su tamaño. Dado el ritmo al cual los primeros productos de Google escalaban verticalmente para satisfacer la demanda global (y con la introducción de productos nuevos como Gmail, Google Maps, Android, YouTube y Google Cloud), no era posible escalar esos nuevos y escasos ingenieros de SRE en la misma proporción a la que crecía la base de clientes.

<sup>10</sup> Benjamin Treynor Sloss, vicepresidente de Ingeniería. Site Reliability Engineering: How Google Runs Production Systems. O'Reilly, 2016. [https://sre.google/sre-book/introduction/#id-2opuzSjFr](https://sre.google/sre-book/part-1-introduction/)

Surgió un dicho:

- La SRE no debe escalar de manera lineal con la cantidad de usuarios.
- La SRE no debe escalar de manera lineal con la cantidad de servidores.
- La SRE no debe escalar de manera lineal con la cantidad de clústeres.
- La SRE no debe escalar de manera lineal con la cantidad de servicios.

Se desarrolló una estructura de administración que permitió mantener ese modelo de crecimiento limitado. Los SRE se alinearon y cooperaron entre sí, y se mantuvieron muy cerca de los equipos de desarrollo de productos con los que trabajaban. Además, rendían cuentas mediante su propia cadena de administración, con Ben Treynor-Sloss<sup>10</sup> a la cabeza. Los equipos se segmentaron en áreas de producto (PA) que se alineaban exactamente con las PA de desarrollo de productos. Los equipos de SRE trabajaban con sus equipos de desarrollo para decidir cuál era la mejor forma de aprovechar a los ingenieros a su disposición. Los equipos de desarrollo podrían financiar SRE nuevos directamente. No todos los equipos de desarrollo crearon productos para los

usuarios: muchos eran equipos de infraestructura compartida, como Bigtable<sup>11</sup> (almacenamiento de datos estructurados), Borg<sup>12</sup> (programación de procesamiento) y Colossus<sup>13</sup> (almacenamiento distribuido). Esos servicios de infraestructura compartida permitieron que los equipos dirigidos a los clientes pudieran escalar sin un equipo de SRE propio.

Mantener a esos equipos en su propia organización permitió que mantuvieran un proceso coherente de contratación y promoción. Los equipos de SRE solían ser diez veces más pequeños (a veces más) en comparación con los equipos de desarrollo. Por eso, era importante garantizar que los equipos de SRE tuvieran autonomía y un objetivo claro que estuviera alineado con los principios de SRE.

Los SRE desarrollaron sus propios productos internos y contaban con sus propios gerentes de producto. Los clientes de los productos de SRE eran otros equipos que querían optimizar la producción. Los equipos de SRE desarrollaron productos relacionados con la respuesta ante incidentes, supervisión y observabilidad, administración de versiones, planificación de capacidad y solución de problemas.

El proceso por el cual los equipos de SRE crecieron siempre consideró la sustentabilidad y la salud del equipo. Por ejemplo, un equipo nuevo de SRE de guardia tenía un tamaño mínimo de 12: dos sitios con 6 miembros. Eso posibilitó la cobertura entre zonas horarias. Además, contar con una cantidad suficiente de personas en un equipo permitió un buen equilibrio entre la vida laboral y personal sin que nadie tuviera que sufrir agotamiento.

La SRE sigue adaptándose en la actualidad. No todos los equipos siguieron con precisión los modelos que se describen arriba, y algunos eligieron caminos alternativos. Organizaciones grandes como la SRE en Google también se ven afectadas por las condiciones del mercado, al igual que cualquier otro grupo grande, por lo que la flexibilidad es importante. Pero sobre todo, los equipos de SRE siguen sus principios: aceptar el riesgo, medir niveles de servicio, eliminar el trabajo repetitivo, adoptar la automatización y simplificar.



<sup>11</sup>Chang, Fay et al. Bigtable: A Distributed Storage System for Structured Data. 7th USENIX Symposium on Operating Systems Design and Implementation (OSDI). USENIX, 2006, 205–218. <https://research.google/pubs/pub27898/>

<sup>12</sup>Verma, Abhishek et al. Large-scale cluster management at Google with Borg. Proceedings of the European Conference on Computer Systems (EuroSys). ACM, Bordeaux, Francia, 2015. <https://research.google/pubs/pub43438/>

<sup>13</sup>Hildebrand, Dean et al. Colossus under the hood: a peek into Google's scalable storage system. 19 de abril de 2021. <https://cloud.google.com/blog/products/storage-data-transfer/a-peek-behind-colossus-googles-file-system>

## Capítulo 6

# La infraestructura flexible es clave para el éxito

## Conclusiones

La infraestructura flexible es un predictor del rendimiento organizativo, operativo, de los equipos y de la entrega de software. La computación en la nube es un habilitador crucial de la infraestructura flexible, pero ese beneficio no se logra automáticamente: nuestros datos muestran que la parte importante es cómo usas la nube.

## Introducción

A lo largo de una gran parte de la investigación de DORA, les preguntamos a los profesionales sobre su infraestructura. Para ello, nos enfocamos en las características esenciales de la computación en la nube, como lo define el Instituto Nacional de Normas y Tecnología (NIST):<sup>1</sup>

- Autoservicio a pedido
- Acceso amplio a la red
- Agrupación de recursos
- Elasticidad rápida
- Servicio medido

Hemos observado constantemente que esas cinco características predicen mejoras en el rendimiento organizativo y de la entrega de software. Este año, queríamos ver si usar computación en la nube predecía infraestructura más flexible.

Las infraestructuras flexibles predicen un  
30%  
más de rendimiento organizativo frente a  
las infraestructuras inflexibles

<sup>1</sup>Publicación especial 800-145 de NIST: "The NIST Definition of Cloud Computing".

	Porcentaje
Múltiples nubes	19.6%
Nube pública	51.5%
Nube híbrida	33.6%
Entorno local	19.8%
De escritorio	3.2%
Otro	2.5%

\* Los encuestados pudieron seleccionar varias respuestas.

## Resultados

Una vez más, confirmamos los hallazgos previos: *la forma en que un equipo usa la nube es un mejor predictor del rendimiento que simplemente indicar que usan la nube*. Si bien el uso de la nube puede ser un habilitador potente, no produce beneficios automáticamente. De hecho, observamos indicadores sólidos de que la nube pública lleva a una *reducción* del rendimiento operativo y de software a menos que los equipos usen infraestructura flexible. Este hallazgo promueve la idea de que solo hacer "lift-and-shift" (el acto de mover cargas de trabajo de un centro de datos a la nube) no es beneficioso y puede generar efectos negativos.

El uso de la computación en la nube se asocia con una reducción considerable del agotamiento y un aumento sustancial en la satisfacción laboral y la productividad.

### Entornos de computación

En la tabla anterior, se detalla el lugar en el que se ejecuta la aplicación o el servicio principal según los encuestados.



Tipo de nube	Rendimiento organizativo	Rendimiento del equipo	Rendimiento de la entrega de software	Rendimiento operativo
Privada		No hay indicios de impacto		No hay indicios de impacto
Pública		Aumentos sustanciales asociados con usar computación en la nube		Reducciones sustanciales asociadas con usar computación en la nube
Híbrida		Aumentos muy sustanciales asociados con usar computación en la nube		No hay indicios de impacto
Múltiples				No hay indicios de impacto

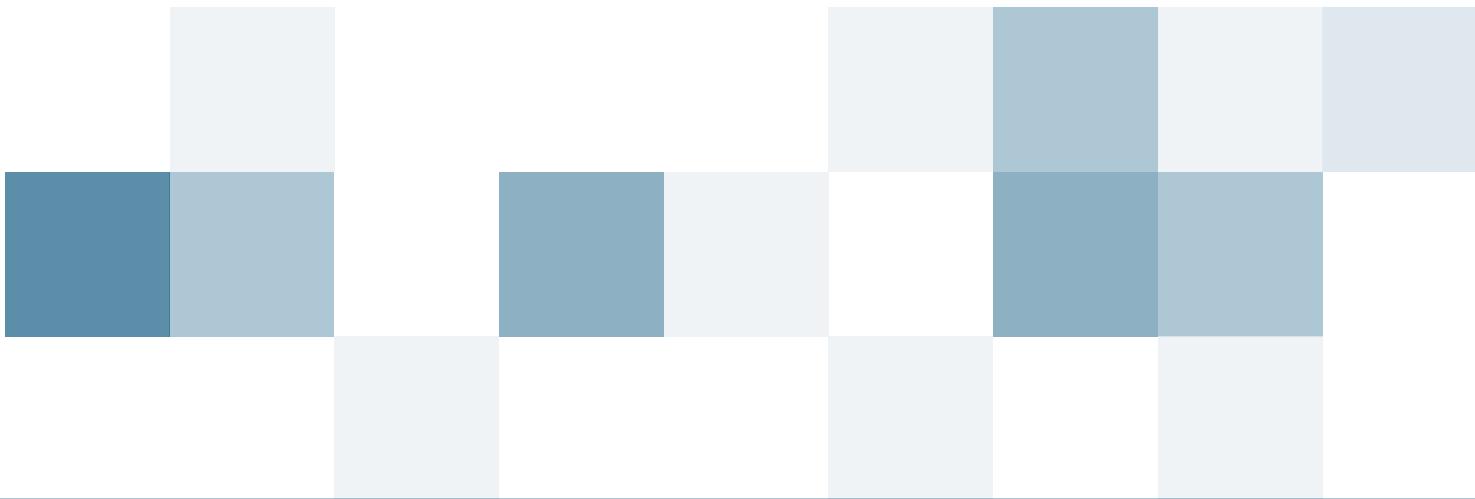
### Simplemente "usar la nube" proporciona resultados mixtos

Como se muestra en la tabla de resultados, simplemente "usar la nube" tiene un impacto neutral o negativo en el rendimiento operativo y de la entrega de software. Es probable que ese impacto sea el resultado de profesionales que han dado el primer paso en su recorrido a la nube y ahora deban trabajar en un entorno nuevo, con herramientas nuevas y algunos procesos diferentes. A menudo, las empresas usan la nube de la misma forma en que lo hicieron en sus propios centros de datos, solo que con las

complejidades y la carga cognitiva adicionales del entorno nuevo. No poder adaptarse a este entorno nuevo no mejora el rendimiento operativo ni el de la entrega de software, sino que los perjudica.

La única excepción de este hallazgo es el rendimiento operativo en el contexto de la nube privada.

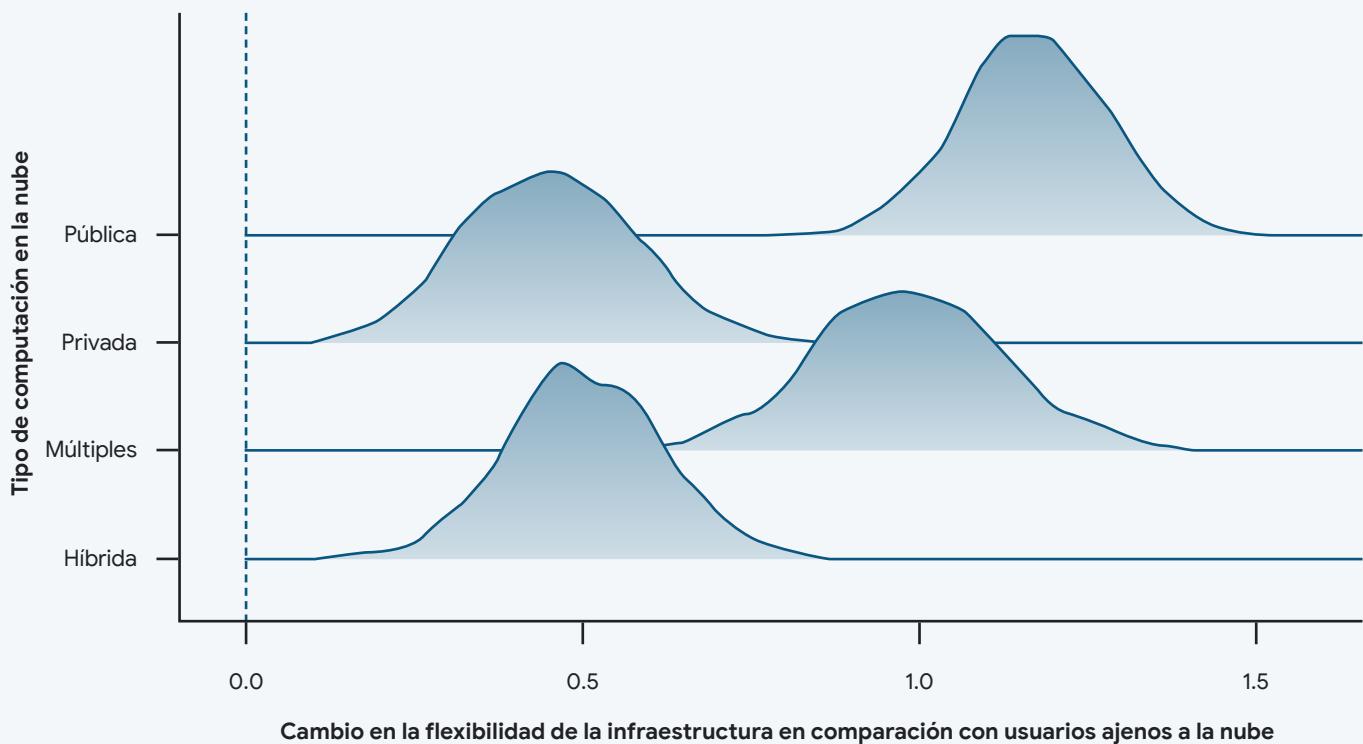
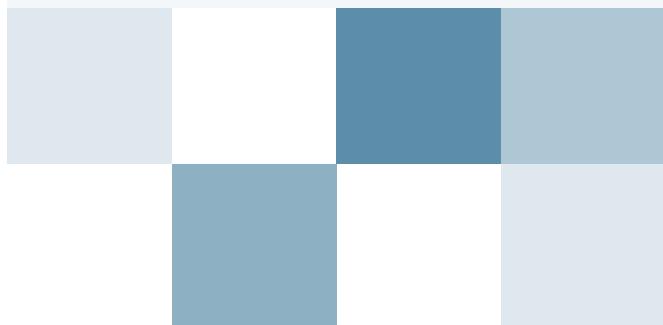
Lo que sí mejora el rendimiento operativo y de la entrega de software es la **infraestructura flexible**, que analizaremos en breve.



### **La infraestructura de nube habilita la flexibilidad**

Usar una nube pública genera un aumento del 22% en la flexibilidad de la infraestructura en comparación con no usar la nube. Usar múltiples nubes también generó un aumento, aunque menos que una sola nube pública; la pregunta obvia es por qué. Nuestros datos muestran que la infraestructura flexible, con frecuencia habilitada por la computación en la nube, genera un mayor impacto que solo usar una plataforma de nube. Para la mayoría, la nube representa una forma nueva de trabajar y se requiere tiempo para dominarla. Como cada plataforma de nube es diferente, eso significa que, a medida que aumentas las plataformas de nube, también incrementa la carga cognitiva requerida para operar en cada plataforma.

**Usar una nube pública genera un aumento en la flexibilidad de la infraestructura en comparación con no usar la nube.**



## Las infraestructuras flexibles predicen un rendimiento más alto en resultados clave

Capacidad	Rendimiento organizativo	Rendimiento del equipo	Rendimiento de la entrega de software	Rendimiento operativo
Infraestructura flexible	Aumentos sustanciales asociados con una infraestructura más flexible	Aumentos sustanciales asociados con una infraestructura más flexible	Aumentos sustanciales asociados con una infraestructura más flexible	Aumentos asociados con una infraestructura más flexible

Es importante reconocer que la infraestructura flexible optimiza el rendimiento organizativo, operativo, de los equipos y de la entrega de software. Muchas organizaciones eligen hacer lift-and-shift para llevar la infraestructura a la nube, lo cual puede ser un gran primer paso, pero es solo el comienzo del recorrido. Si decides hacer lift-and-shift de una parte de tus cargas de trabajo, el próximo paso es modernizarlas mediante una refactorización para aprovechar la infraestructura flexible.

## La computación en la nube tiene un impacto positivo en resultados clave por medio de la infraestructura flexible

Tipo de nube con infraestructura flexible	Rendimiento organizativo	Rendimiento del equipo	Rendimiento de la entrega de software	Rendimiento operativo
Privada	Mediado parcialmente por la infraestructura flexible	Mediado parcialmente por la infraestructura flexible	Mediado parcialmente por la infraestructura flexible	Mediado parcialmente por la infraestructura flexible
Pública	Mediado completamente por la infraestructura flexible	Mediado completamente por la infraestructura flexible	Mediado parcialmente por la infraestructura flexible	Mediado parcialmente por la infraestructura flexible
Híbrida	Mediado parcialmente por la infraestructura flexible	Mediado completamente por la infraestructura flexible	Mediado parcialmente por la infraestructura flexible	Mediado completamente por la infraestructura flexible
Múltiples	Mediado parcialmente por la infraestructura flexible	Mediado parcialmente por la infraestructura flexible	Mediado parcialmente por la infraestructura flexible	Mediado parcialmente por la infraestructura flexible

Cuando se usan las plataformas de computación en la nube de una forma que maximiza las características de la infraestructura flexible, predicen un impacto positivo en el rendimiento operativo y de la entrega de software. La diferencia en el impacto indica algo que ya sabe la mayoría de los profesionales y líderes: simplemente trasladar las cargas de trabajo de un centro de datos a la nube no es la clave del éxito. Lo importante es aprovechar la infraestructura flexible que habilita la nube.

Tipo de infraestructura	Resultados
Nube con una infraestructura flexible	
Nube sin flexibilidad	

Para maximizar el posible beneficio, debes reconsiderar cómo desarrollas, pruebas, implementas y supervisas tus aplicaciones. Una parte importante de ese proceso es aprovechar las cinco características de la computación en la nube: autoservicio a pedido, acceso amplio a la red, agrupación de recursos, elasticidad rápida y servicio medido.



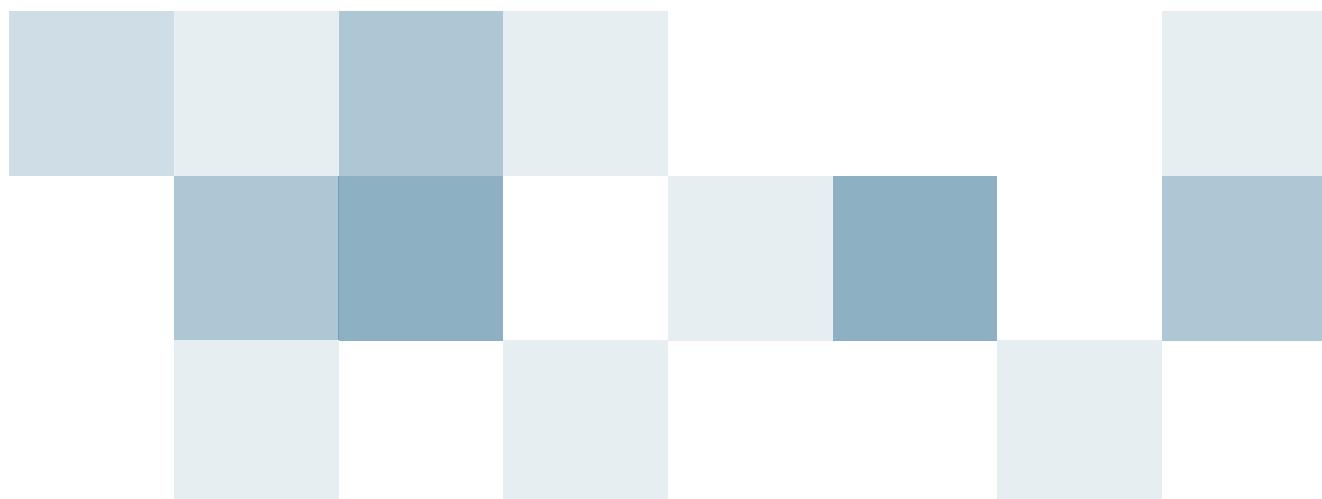
## La computación en la nube mejora el bienestar

Tipo de nube	Agotamiento*	Satisfacción laboral	Productividad
Privada		No hay indicios de impacto	
Pública		Reducciones muy sustanciales asociadas con usar computación en la nube	
Híbrida		No hay indicios de impacto	
Múltiples		No hay indicios de impacto	

\* Notarás que el esquema de colores está invertido para el agotamiento. Esto se debe a que la reducción del agotamiento es algo bueno.

Los datos muestran que la computación en la nube es muy beneficiosa para el bienestar de los empleados. Observamos un aumento sustancial tanto en la satisfacción laboral como en la productividad, y un impacto neutral o positivo en el agotamiento. Dicho de otra forma, la nube no tiene un efecto perjudicial en el bienestar, aunque la computación en la nube conlleva una carga cognitiva adicional, el aprendizaje de herramientas y nuevas formas de trabajar.

En cuanto que profesionales, tenemos algunas hipótesis sobre por qué se observa eso. A los ingenieros les gusta aprender y resolver problemas, y les gusta trabajar en un entorno con características de computación flexible. Aprender sobre tecnologías nuevas es divertido y, además, una excelente forma de desarrollo profesional. Los ingenieros están más felices cuando le va bien a su organización.

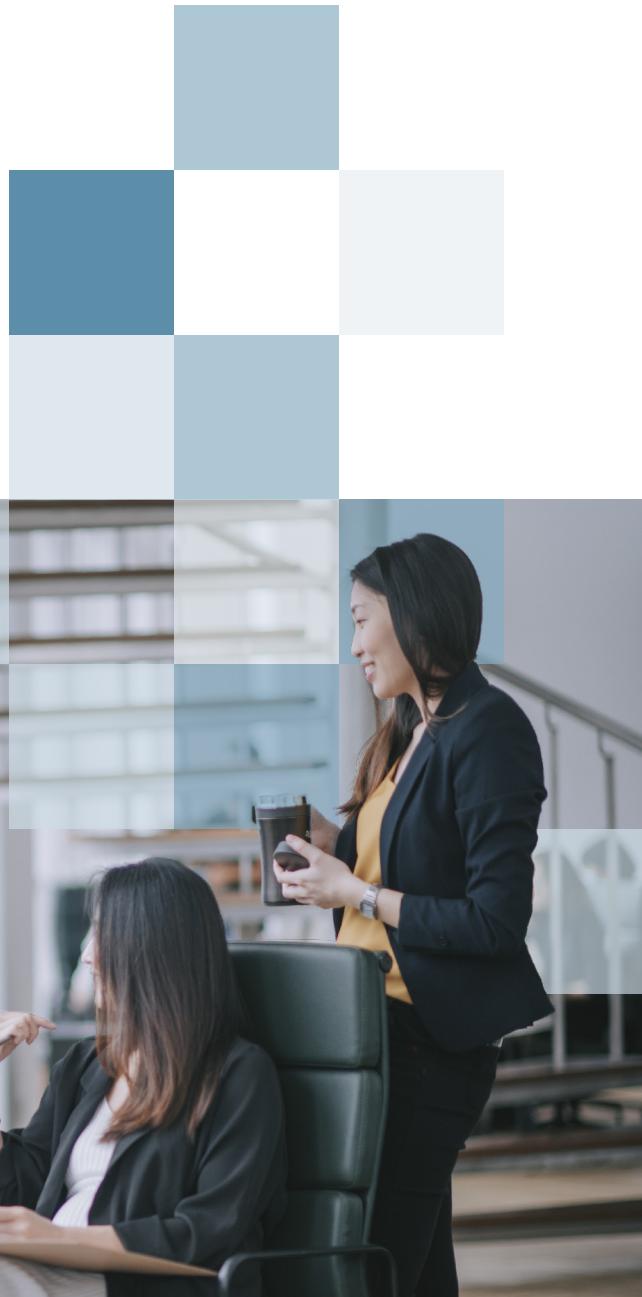


## Capítulo 7

# Nada de esto funciona sin inversión en cultura

## Conclusiones

La cultura es un impulsor clave del bienestar de los empleados y del rendimiento organizativo. Una cultura saludable puede ayudar a reducir el agotamiento, y aumentar la productividad y la satisfacción laboral. También genera aumentos considerables en el rendimiento organizativo, operativo, de la entrega de software y de los equipos. Una cultura organizativa saludable puede ayudar a los equipos a ser más exitosos a la hora de implementar capacidades técnicas asociadas con resultados mejorados.



# Introducción

No es fácil definir la cultura. Nos enfocamos en indicadores que nos digan algo sobre las experiencias de las personas en el trabajo. Usamos la tipología de cultura organizativa de Westrum<sup>1</sup> porque ha sido un buen predictor constante del rendimiento. Este año, observamos que las organizaciones que tienen una cultura generativa, según la define Westrum, siguen teniendo un buen rendimiento.

En la siguiente tabla, se detallan aspectos que creemos que contribuyen a la cultura de los equipos y la organización.

Aspecto	Definición
Cultura organizativa de Westrum	La manera en que una organización suele responder a oportunidades y problemas. Hay tres tipos de cultura: generativa, burocrática y patológica.
Estabilidad de la organización	Qué tan estable o inestable es el entorno para los empleados.
Seguridad laboral	Con cuánta frecuencia los empleados se preocupan por su seguridad laboral.
Flexibilidad	Cómo, dónde y cuándo se trabaja.
Conocimiento compartido	Cómo se esparcen las ideas y la información en la organización. Los miembros de los equipos responden preguntas una vez y la información se pone a disposición de otras personas. Nadie tiene que esperar por las respuestas.
Enfoque en los usuarios	El enfoque está en el usuario final cuando se desarrolla software y hay una comprensión profunda de las necesidades y los objetivos de los usuarios. Se usan indicadores del usuario para mejorar los productos y servicios.
Distribución del trabajo	Procesos formales que ayudan a los equipos a distribuir tareas complejas de manera equitativa entre sus miembros.

Los límites entre aspectos culturales, capacidades de procesos y capacidades técnicas no siempre son claros. Creemos que la cultura emerge de prácticas, y estas emergen de la cultura. Hablaremos sobre esto más adelante cuando analicemos nuestros hallazgos.

<sup>1</sup><http://bmj.co/1BRGh5q>

# ¿Qué descubrimos y qué significa?

## Una cultura saludable mejora los resultados clave

En términos generales, una cultura saludable tiene un impacto positivo en todos los resultados clave. Replicamos los hallazgos de años anteriores relacionados con que una cultura generativa mejora el rendimiento organizativo, operativo y de la entrega de software. También optimiza la nueva métrica de rendimiento de este año: rendimiento de los equipos.

Observamos que un enfoque centrado en los usuarios para desarrollar software genera aumentos considerables en el rendimiento. Vale la pena destacar eso. Las organizaciones pueden experimentar una cascada de beneficios cuando priorizan al usuario. Los comentarios de los usuarios ayudan a los equipos a priorizar proyectos y a crear productos y servicios que satisfagan las necesidades de los usuarios. Esto lleva a una mejor experiencia y mayor satisfacción del usuario, y mayores ingresos.

También evaluamos el estado de la cultura de una organización midiendo la distribución del trabajo entre equipos. Observamos que la distribución equitativa del trabajo beneficia el rendimiento organizativo y de los equipos. Sin embargo, observamos que la distribución equitativa del trabajo estaba asociada con un rendimiento *más bajo* de la entrega de software. Quizás los procesos formales relacionados con la distribución del trabajo ralentizan la finalización de tareas complejas que forman parte de la canalización de la entrega de software. También es posible que los procesos formales tengan un impacto sobre qué miembro del equipo debe encargarse de una tarea determinada.

Otro hallazgo que parece disonante es que la estabilidad de la organización muestra una reducción

Los equipos con culturas generativas tienen un

30%

más de rendimiento organizativo que los equipos que no la tienen

pequeña pero significativa en el rendimiento de la entrega de software. Una posible explicación es que las organizaciones más establecidas (y probablemente más grandes) no resienten la presión de seguirles el ritmo a organizaciones más nuevas y menos establecidas (y más pequeñas). Es posible que las organizaciones más establecidas ya cuenten con un producto conformado, lo que les da flexibilidad en torno a la velocidad de su entrega de software.

Cuando la información fluye fácilmente, las tareas se completan. Descubrimos que los niveles más altos de información compartida se asociaban con aumentos en el rendimiento operativo y de la entrega de software. Cuando es fácil acceder a la información y hay pocos silos de conocimiento, las personas pueden dedicar más tiempo a tareas importantes en lugar de rastrear la información necesaria para realizar esas tareas.

Por último, los arreglos flexibles de trabajo, en los que los empleados pueden determinar cuándo, dónde y cómo trabajan, tienen un impacto benéfico en todas las métricas de rendimiento. Eso es verdadero en especial para el rendimiento de la entrega de software. Incluso a medida que las organizaciones endurecen sus políticas de trabajo remoto, es probable que permitirles a los empleados conservar cierta flexibilidad tenga un beneficio.

Aspecto de la cultura	Efecto en el rendimiento de equipos	Efecto en el rendimiento organizativo	Efecto en el rendimiento de la entrega de software	Efecto en el rendimiento operativo
Cultura organizativa de Westrum	Aumento sustancial	Aumento sustancial	Aumento sustancial	Aumento sustancial
Estabilidad de la organización	Aumento menor	Aumento sustancial	Reducción menor	Sin efecto
Seguridad laboral	Aumento menor	Sin efecto	Aumento menor	Aumento menor
Flexibilidad	Aumento menor	Aumento menor	Aumento sustancial	Aumento menor
Conocimiento compartido	Aumento menor	Reducción menor	Aumento sustancial	Aumento sustancial
Enfoque en los usuarios	Aumento sustancial	Aumento sustancial	Aumento menor	Aumento sustancial
Distribución del trabajo	Aumento sustancial	Aumento sustancial	Reducción sustancial	Sin efecto



## Una cultura saludable mejora las capacidades técnicas

Nuestros hallazgos sugieren que una buena cultura ayuda a mejorar la implementación de capacidades técnicas. Creemos que la relación entre cultura y capacidades técnicas es recíproca: la cultura emerge de las prácticas, y las prácticas emergen de la cultura.

La cultura es amplia y difícil de definir, mientras que las capacidades técnicas suelen tener un alcance y estar bien definidas. Eso tiene implicaciones en la forma en que las personas en una organización pueden ayudar a impulsar el cambio.

Por ejemplo, los líderes pueden crear estructuras de incentivos que promuevan una cultura generativa. Tanto los líderes como los colaboradores individuales pueden enfatizar un enfoque centrado en usuarios

para el desarrollo de software. Los colaboradores individuales pueden ayudar a impulsar la implementación de capacidades técnicas que mejoran el rendimiento, como el desarrollo basado en troncales, la integración continua, las prácticas de confiabilidad y la arquitectura con acoplamiento bajo. Implementar esas capacidades técnicas no es fácil, y su éxito depende de que las personas trabajen juntas, sean abiertas, se apoyen entre sí y aprendan las unas de las otras. Todos esos son elementos de una cultura saludable. Esos equipos pueden convertirse en ejemplos para otros dentro de la organización, los cuales podrían sentirse empoderados para impulsar el cambio usando todo lo que tengan a su alcance.

Los cambios duraderos y significativos en la cultura de una organización ocurren por medio de iniciativas simultáneas ascendentes y descendentes.

Aspecto de la cultura	Efecto en el desarrollo basado en troncales	Efecto en las prácticas de confiabilidad	Efecto en la integración continua	Efecto en la entrega continua	Efecto en la arquitectura con acoplamiento bajo
Cultura organizativa de Westrum	Aumento sustancial	Aumento sustancial	Aumento sustancial	Aumento sustancial	Aumento sustancial
Estabilidad de la organización	Aumento menor	Aumento sustancial	Sin efecto	Sin efecto	Sin efecto
Seguridad laboral	Reducción menor	Reducción menor	Sin efecto	Sin efecto	Sin efecto
Flexibilidad	Sin efecto	Reducción menor	Aumento sustancial	Aumento menor	Aumento sustancial
Conocimiento compartido	Sin efecto	Sin efecto	Sin efecto	Aumento menor	Aumento menor
Enfoque en los usuarios	Aumento sustancial	Aumento sustancial	Aumento sustancial	Aumento sustancial	Aumento sustancial
Distribución del trabajo	Aumento sustancial	Aumento sustancial	Aumento sustancial	Aumento sustancial	Aumento sustancial

## Una cultura saludable mejora el bienestar de los empleados

Una cultura saludable lleva a niveles altos de bienestar de los empleados, ya que reduce el agotamiento y aumenta la satisfacción laboral y la productividad. El bienestar de los empleados no es meramente deseable: es un aspecto fundamental del estado y el éxito generales de una organización.

¿Qué pasa cuando las organizaciones no invierten en una cultura mejor? Aumenta la probabilidad de

agotamiento y se reduce la satisfacción laboral. Los empleados empiezan a ser cínicos y su productividad disminuye. La salud física y psicológica también se ve impactada de manera negativa.<sup>2,3</sup> El agotamiento es persistente; no es algo que las personas superan después de tomarse unas vacaciones. El agotamiento también aumenta la rotación: los empleados renuncian para buscar entornos de trabajo más saludables.<sup>4</sup> Por lo tanto, para aliviar el agotamiento, se requieren cambios organizativos que aborden sus causas.

Aspecto de la cultura	Efecto en el agotamiento*	Efecto en la satisfacción laboral	Efecto en la productividad
Cultura organizativa de Westrum	 Reducción sustancial	 Aumento sustancial	 Aumento sustancial
Estabilidad de la organización	 Reducción sustancial	 Aumento sustancial	 Aumento menor
Seguridad laboral	 Reducción sustancial	 Aumento menor	 Aumento menor
Flexibilidad	 Reducción menor	 Aumento menor	 Aumento menor
Conocimiento compartido	 Reducción sustancial	 Aumento menor	 Aumento menor
Enfoque en los usuarios	 Reducción menor	 Aumento sustancial	 Aumento sustancial
Distribución del trabajo	 Sin efecto	 Aumento menor	 Aumento menor

\* Notarás que el esquema de colores está invertido para el agotamiento. Esto se debe a que la reducción del agotamiento es algo bueno.

<sup>2</sup> Adam Bayes, Gabriela Tavella y Gordon Parker (2021). "The biology of burnout: Causes and consequences". The World Journal of Biological Psychiatry, 22:9, 686–698. DOI: 10.1080/15622975.2021.1907713. <https://doi.org/10.1080/15622975.2021.1907713>

<sup>3</sup> Maslach C, Leiter MP. "Understanding the burnout experience: recent research and its implications for psychiatry". World Psychiatry, junio de 2016, 15(2), 103–11. DOI: 10.1002/wps.20311. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4911781/>. PMID: 27265691; PMCID: PMC4911781.

<sup>4</sup> L.A. Kelly et al. "Impact of nurse burnout on organizational and position turnover". Nursing Outlook, enero-febrero de 2021, 96–102. DOI: [doi.org/10.1016/j.outlook.2020.06.008](https://doi.org/10.1016/j.outlook.2020.06.008)

## Capítulo 8

# Cómo, cuándo y por qué importa quién eres

## Conclusiones

Quién eres importa. Observamos que determinados grupos de encuestados tienen resultados diferentes en comparación con otros; por ejemplo, más agotamiento o menos productividad. También identificamos prácticas específicas que puedes implementar para mitigar algunos de esos resultados negativos.

## Introducción

Un fenómeno general se extendió por el análisis de 2022: la forma en que está configurado el trabajo podría ser propicia para el bienestar de algunas personas, pero no de todas.

En 2022, descubrimos que las personas que se identifican como poco representadas informaron niveles más altos de agotamiento.<sup>1</sup> En este capítulo, veremos que este hallazgo se replica. Además, empezaremos a abordar por qué los grupos poco representados tienen más probabilidades de tener agotamiento y qué factores ayudan a evitarlo.

Asimismo, la inestabilidad que se siente en muchos sectores llevó a preguntas en torno a los empleados nuevos. Las organizaciones están preocupadas de que tome mucho tiempo para que sean productivos, por lo que buscan formas de ayudarlos a ponerse al día más rápido. Aquí también analizaremos eso.

## ¿Qué descubrimos y qué significa?

### Algunas personas están más agotadas que otras

El año pasado, descubrimos que los encuestados que se identifican como mujeres o autodefinen su género, y los que se identificaron como poco representados informaron estar más agotados que los encuestados que se identifican como hombres y no como poco representados. Esos hallazgos son coherentes con un cuerpo de investigación anterior que sugiere que las personas poco representadas sienten un grado más alto de agotamiento<sup>2</sup> y estrés laboral<sup>3</sup> que sus pares representados.

Por esos motivos, nos interesó analizar si esas disparidades en el agotamiento volverían a aparecer en nuestros datos este año, y así lo hicieron. Los encuestados que se identifican como mujeres o autodefinen su género informaron sentir niveles un 6% mayores de agotamiento que los encuestados que se identifican como hombres. Los encuestados que se identifican como poco representados informaron niveles de agotamiento un 24% más altos que los encuestados que no se identifican como poco representados.

<sup>1</sup> Informe Accelerate State of DevOps 2022. <https://dora.dev/research/2022/dora-report/>

<sup>2</sup> Sigalit Ronen y Ayala Malach Pines. "Gender Differences in Engineers' Burnout". Equal Opportunities International, 7 de noviembre de 2008. <https://www.emerald.com/insight/content/doi/10.1108/02610150810916749/full/html>

<sup>3</sup> Dalessandro C., Lovell A., Tanner OC. "Race, Marginalization, and Perceptions of Stress Among Workers Worldwide Post-2020". Sociological Inquiry, 3 de agosto de 2023. <https://onlinelibrary.wiley.com/doi/10.1111/soin.12505>

## Algunos tipos de trabajo predicen más agotamiento

Observamos que ciertos aspectos del lugar de trabajo que pueden parecer neutrales o beneficiosos, como la documentación de calidad o un equipo estable, no reducen el agotamiento para todas las personas (consulta el [capítulo 4, La documentación es la base](#)).

Nos preguntamos si esto podría deberse a tareas que benefician a la organización, pero contribuyen al agotamiento de algunas personas.<sup>4</sup>

Para comprender la experiencia de los encuestados relacionada con el agotamiento, preguntamos sobre el trabajo que hacen, y lo medimos de dos formas analizando lo siguiente:

- Tareas específicas, como programación, reuniones o asistencia a compañeros de equipo
- Características del trabajo, como tareas no planificadas, su visibilidad o cantidad de trabajo repetitivo

Las características del trabajo son importantes porque es posible que la misma tarea genere experiencias distintas para personas diferentes o en distintos momentos. Por ejemplo, algunas revisiones de código pueden ser trabajo repetitivo no planificado y otras pueden ser altamente visibles dentro del equipo, ya que muestran el liderazgo y la experiencia técnica.<sup>5</sup>

Los encuestados que se identifican como poco representados informaron que hacen un 24% más de trabajo repetitivo que sus contrapartes. Los encuestados que se identifican como mujeres o autodefinen su género informaron hacer un 40% más de trabajo repetitivo que los encuestados que se identifican como hombres. Esos dos grupos de encuestados también informaron hacer más trabajo no planificado, además de tareas que no son tan visibles para sus colegas ni se alinean directamente con sus habilidades profesionales. Estos hallazgos explican parcialmente el agotamiento que informaron esos grupos.

<sup>4</sup> Linda Babcock, Brenda Peyser, Lise Vesterlund y Laurie Weingart. *The No Club*. Nueva York: Simon and Schuster, 2022, 17.

<sup>5</sup> Murphy-Hill, E. et al. "Systemic Gender Inequities in Who Reviews Code". Computer Supported Cooperative Work, 2023, publicación pendiente, <https://research.google/pubs/pub52204>

<sup>6</sup> Babcock et al., *The No Club*, 17.

## Tareas que no llevan a un ascenso

En *The No Club*, Babcock et al. describen un tipo específico de trabajo: "Una tarea que no lleva a un ascenso es importante para tu organización, pero no te ayudará a avanzar profesionalmente".<sup>6</sup>

Aquí el término *avanzar* tiene un significado amplio; por ejemplo, una mayor compensación o la posibilidad de promoverse para conseguir otros puestos.

La evidencia muestra que las mujeres hacen más de ese tipo de trabajo. Babcock et al. describen las razones de esa distribución desigual. Por ejemplo, es más probable que se les pida a las mujeres que hagan ese tipo de tareas y que acepten hacerlas cuando rechazarlas conlleva un costo social.

También describen las consecuencias de la distribución desigual de esas tareas; por ejemplo, algunas mujeres declaran lo siguiente:

- Observan un impacto negativo en su carrera o ingresos.
- Toman más horas para tener un volumen de trabajo adecuado que sea relevante para su carrera.

## **Los procesos formales de distribución del trabajo reducen el agotamiento para algunos encuestados**

Les preguntamos a los encuestados si tienen procesos formales para distribuir el trabajo de forma uniforme. Lo denominamos **distribución del trabajo**, y esperábamos observar que mitiga el agotamiento que sienten algunos encuestados.

Descubrimos que la distribución del trabajo sí redujo el agotamiento para los encuestados que se identifican como hombres, mujeres o que autodefinen su género. Con un alto nivel de distribución del trabajo, la diferencia en el agotamiento entre géneros desapareció.

Nos sorprendió descubrir que la distribución del trabajo no tuvo ningún impacto en el nivel de agotamiento que sintieron los encuestados que se identifican como poco representados. Ese hallazgo plantea más preguntas: ¿Los procesos formales para distribuir el trabajo de manera uniforme siguen generando una distribución desigual del trabajo? ¿El "trabajo equitativo" tiene en cuenta las características de las tareas, como interrupciones y visibilidad? ¿Y cómo mitigamos otros factores que contribuyen al agotamiento, además de las tareas laborales, que podrían ser más significativos para este grupo?



# Promover una cultura de pertenencia

**Autor: Dr. Jeffrey Winer,  
psicólogo, Boston Children's Hospital, profesor adjunto en Harvard Medical School**

## Un hallazgo clave y un poco de contexto

Un hallazgo clave de este informe es que las personas que se definen como poco representadas sienten mucho más agotamiento que sus colegas. En el informe, se exploraron algunas razones posibles de eso. En esta sección, queremos conectar esos hallazgos con una investigación más amplia sobre la pertenencia y las estrategias de prácticas organizativas asociadas.

Identificarse como poco representado en un grupo demuestra vulnerabilidad hacia la "incertidumbre de la pertenencia",<sup>7</sup> un fenómeno psicológico bien establecido (consulta "[Understanding and Overcoming Belonging Uncertainty](#)" del Dr. Geoffrey Cohen).<sup>8</sup>

Esta incertidumbre (por ejemplo, "¿pertenezco a este lugar?", "¿las personas como yo pueden tener éxito aquí?") se refuerza o redefine por medio de las experiencias continuas de las personas y sus interpretaciones de esas experiencias. Esos procesos bien establecidos relacionados con la incertidumbre de la pertenencia pueden ayudar a contextualizar el hallazgo de este informe de que las personas que se identifican como poco representadas afirman tener niveles de agotamiento más altos.

## ¿Qué pueden hacer las organizaciones?

Es importante recordar que la diversidad, inclusión, equidad y pertenencia significan cosas distintas y, para lograrlos, se requieren estrategias diferentes, interconectadas y sostenidas. Lograr la pertenencia requiere compromisos verdaderos y sostenidos.

Si alguien tiene dificultades en una organización, la primera pregunta no debería ser "¿Qué le pasa a esa persona?", sino "¿Por qué tendría sentido que alguien se sienta así y qué elementos estructurales de nuestra organización generan ese sentimiento?" (por ejemplo, qué elementos hacen que se mantenga ese sentimiento o lo agravan).

Cuando se identifican problemas, los cambios deberían hacerse a nivel de la organización, a la vez que debería proporcionarse apoyo a nivel individual: un enfoque complementario. Brindar apoyo a las personas para generar un impacto en los sistemas que gobiernan una organización permitirá la aplicación de cambios en el sistema que durarán más que los actores individuales. Tomar esa mentalidad de sistemas y sustentabilidad permitirá que los cambios se incorporen en la institución, de manera que duren más que los actores individuales. Esa cualidad generativa es lo que permite que las organizaciones se esfuerzen para lograr la pertenencia. Esforzarse es clave aquí. La pertenencia se construye por medio de la acción y experiencia sostenidas; nunca se termina y, por eso, es tan esencial para la productividad y el estado del lugar de trabajo.

Hay diversas herramientas para apoyar a las organizaciones en esta área. Por ejemplo, en el informe sobre la soledad de 2023 del Director general de Salud Pública, se identifica que la conexión social y la pertenencia son antídotos clave para la soledad y el agotamiento.<sup>9</sup>

<sup>7</sup> Walton GM; Cohen GL. "A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students". Science, Nueva York, N.Y., último acceso 20 de septiembre de 2023, <https://pubmed.ncbi.nlm.nih.gov/21415354/>

<sup>8</sup> <https://behavioralscientist.org/understanding-and-overcoming-belonging-uncertainty/>

<sup>9</sup> Office of the Assistant Secretary for Health (OASH). "New Surgeon General Advisory Raises Alarm about the Devastating Impact of the Epidemic of Loneliness and Isolation in the United States". HHS.gov, 3 de mayo de 2023. <https://www.hhs.gov/about/news/2023/05/03/new-surgeon-general-advisory-raises-alarm-about-devastating-impact-epidemic-loneliness-isolation-united-states.html>

## **Los empleados nuevos tienen dificultades con la productividad**

Los empleados nuevos (con menos de 1 año de experiencia en el equipo) tienen una puntuación un 8% más baja en productividad que sus compañeros con más de 1 año de experiencia. Quizás eso sea de esperar. Empezar en un equipo nuevo es un desafío y, aunque tengas experiencia en el puesto, la cantidad de conocimiento específico del equipo requerida para comenzar a trabajar puede ser abrumadora. Además, formar parte de un equipo es más que solo un asunto de habilidades y conocimiento. De manera anecdótica, también hay un componente social que es esencial para la productividad. Aspectos como la pertenencia, sentirse como un miembro que contribuye y la seguridad psicológica demoran en desarrollarse.

## **¿Hay algo que pueda ayudar a los empleados nuevos a acelerar el ritmo?**

Tenemos la hipótesis de que las organizaciones podrían ayudar a los empleados nuevos de tres formas:

- Con documentación de alta calidad
- Con la incorporación de inteligencia artificial en los flujos de trabajo, que se ha demostrado en otros estudios que es más útil para los trabajadores sin experiencia
- Trabajando en conjunto de forma presencial, que se ha sugerido que podría ser particularmente beneficioso en la etapa de integración



Los resultados de este año sugieren que la documentación de alta calidad genera impactos considerables en la productividad (consulta el [capítulo 4, La documentación es la base](#)) y la IA tiene beneficios menores en la productividad de una persona (consulta el [capítulo 3, Las capacidades técnicas predicen el rendimiento](#)). No tenemos motivos para pensar que esos efectos no se extenderían a los empleados nuevos. Cuando analizamos los datos, observamos lo siguiente: estas prácticas ayudan a los empleados nuevos, pero no más o menos que al resto de las personas. En otras palabras, los empleados nuevos no obtienen beneficios especiales de estas prácticas.

Si quieras ayudar a los empleados nuevos y al resto de las personas, la documentación de alta calidad es un buen comienzo debido a su efecto claro y sustancial en la productividad. Vale la pena aclarar que la productividad de los empleados nuevos en equipos con documentación bien escrita (1 desviación estándar por encima del promedio) es de un 130% en comparación con equipos con documentación mal escrita (1 desviación estándar por debajo del promedio).

La discusión sobre el retorno de la oficina será breve para evitar echar leña al fuego. Hablamos sobre la importancia de la flexibilidad en el [capítulo 7, Nada de esto funciona sin inversión en cultura](#). Además, nuestros datos no son experimentales y, aunque tratamos de controlar los factores que podrían sesgar los resultados, los beneficios del arreglo laboral personal son un tema complejo y sociológicamente enriquecido, por lo que es difícil sacar conclusiones claras (ten esto en mente cuando leas estudios o reflexiones sobre el tema). Lo que indican claramente los datos es que la flexibilidad tiene un impacto positivo en la productividad. Lo que no está claro es si el lugar donde trabajas tiene algún efecto.

El mismo relato es verdadero para los empleados nuevos. No observamos evidencia de que trabajar juntos de manera presencial tenga un beneficio específico para los empleados nuevos. Si quieras optimizar la productividad y darles a los empleados nuevos flexibilidad en cuanto a cómo, dónde y cuándo trabajan, parece ser una apuesta más segura que obligarlos a estar en la oficina. Sin duda, las organizaciones no optimizan solo para mejorar la productividad (y probablemente no deberían hacerlo). También creemos que la productividad en términos del trabajo genera valor, no solo resultados (no líneas de código), y trabajo que no provoca agotamiento ni es repetitivo.



## Epílogo

# Reflexiones finales

Gracias por participar en la investigación de este año y leer el informe. Siempre estamos buscando maneras mejores de explorar las conexiones entre cómo trabajan los equipos y los resultados que pueden alcanzar.

La conclusión más importante de nuestro programa de investigación multianual es que los equipos que adoptan una mentalidad y práctica de mejora continua pueden alcanzar los mejores resultados.

Las capacidades que exploramos pueden usarse como perillas que impulsan resultados. Algunas de esas perillas están al alcance de las personas, mientras que otras solo son accesibles mediante un esfuerzo coordinado en toda la organización. Identifica qué perillas deben ajustarse en tu organización y, luego, haz inversiones en esos ajustes.

El trabajo para realizar mejoras nunca termina, pero puede crear éxito a largo plazo para las personas, los equipos y las organizaciones. Los líderes y los profesionales comparten la responsabilidad para impulsar las mejoras.

## ¿Cómo pondrás en práctica este estudio?

Explora estos hallazgos en el contexto de tu organización, equipos y los servicios que brindas a tus clientes.

Para compartir tus experiencias, aprender de otros y obtener inspiración de otras personas en el recorrido de la mejora continua, únete a la comunidad de DORA en <https://dora.community>.

# Agradecimientos

Cada año, este informe cuenta con el apoyo de una gran familia de colaboradores apasionados de todo el mundo. Todos los pasos de su producción (diseño de preguntas de la encuesta, localización, análisis, redacción, edición y diagramación) tienen el aporte de colegas que ayudaron a hacer realidad este gran esfuerzo. Los autores desean agradecer a todas esas personas por sus aportes, orientación y camaradería.

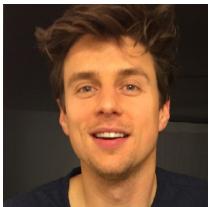
## Colaboradores

Equipo central	Editores	Yuki Iwanari	Rob Edwards
James Brookbank	Mandy Grover	Vincent Jobard	Dave Farley
Kim Castillo	Jay Hauser	Gustavo Lapa	Steve Fenton
Derek DeBellis	Stan McKenzie	Mauricio Meléndez	Dr. Nicole Forsgren
Nathen Harvey	Anna Eames Mikkawi	Jeremie Patonnier	Aaron Gillies
Michelle Irvine	Mike Pope	Miguel Reyes	Denali Lumma
Amanda Lewis	Tabitha Smith	Pedro Sousa	Emerson Murphy-Hill
Eric Maxwell	Olinda Turner	Laurent Tardif	Harini Sampath
Steve McGhee	<b>Localización de la encuesta</b>	Kimmy Wu	Robin Savinar
Dave Stanke		Vinicius Xavier	Dustin Smith
Kevin Storer		Yoshi Yamaguchi	Jess Tsimeris
Daniella Villalba		<b>Asesores y expertos en el campo</b>	Dr. Laurie Weingart
Brenna Washington	Kuma Arakawa		Betsalel (Saul) Williamson
	William Bartlett		Dr. Jeffrey Winer
	Antonio Guzmán		
	Shogo Hamada	Lisa Crispin	

## Patrocinadores



# Autores



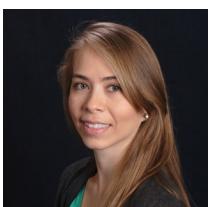
## Derek DeBellis

Derek es un investigador cuantitativo de experiencia del usuario en Google y el investigador principal de DORA. Se enfoca en investigación de encuestas, análisis de registros y buscar formas de medir conceptos que demuestren que un producto o una función brindan verdadero valor a las personas. Derek ha publicado sobre la interacción entre humanos y la IA, el impacto del COVID-19 en el abandono del tabaquismo, el diseño ante errores de PLN, el rol de la experiencia del usuario en discusiones sobre la privacidad, la cultura de equipos y la relación de la IA con la productividad y el bienestar de los empleados. Su investigación extracurricular explora formas de simular la propagación de creencias y poder.



## Amanda Lewis

Es la jefa de desarrollo de DORA.community e ingeniera de relaciones de desarrolladores en el equipo DORA Advocacy de Google Cloud. Ha dedicado su carrera a crear conexiones con desarrolladores, operadores, gerentes de producto, administración de proyectos y liderazgo. Trabajó en equipos que desarrollaron plataformas de comercio electrónico, sistemas de administración de contenido, herramientas de observabilidad y que apoyaban a desarrolladores. Esas conexiones y conversaciones generan clientes felices y mejores resultados para el negocio. Aporta su experiencia y empatía al trabajo que hace para ayudar a los equipos a comprender e implementar prácticas de confiabilidad y entrega de software.



## Daniella Villalba

Es investigadora de la experiencia del usuario en Google. Usa la investigación con encuestas para comprender los factores que hacen felices y productivos a los desarrolladores. Antes de Google, Daniella estudió los beneficios del entrenamiento en meditación y los factores psicosociales que afectan las experiencias de los estudiantes universitarios. Obtuvo un doctorado en psicología experimental de Florida State University.



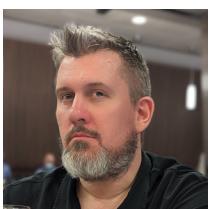
### Dave Farley

Es el director general y fundador de Continuous Delivery Ltd., autor de *Modern Software Engineering* y coautor del exitoso libro *Continuous Delivery*. Es uno de los autores del Manifiesto Reactivo y ganador del Premio Duke por el proyecto de código abierto LMAX Disruptor. Dave es pionero de la entrega continua, líder de pensamiento y profesional experto en CD, DevOps, desarrollo basado en pruebas (TDD) y diseño de software. Tiene una larga trayectoria en la creación de equipos de alto rendimiento, la configuración de organizaciones para el éxito y la creación de software excepcional. Está comprometido con compartir su experiencia y técnicas con desarrolladores de todo el mundo para ayudarlos a mejorar el diseño, la calidad y la confiabilidad de su software. Comparte su experiencia por medio de su consultoría,<sup>1</sup> canal de YouTube,<sup>2</sup> y cursos de capacitación.



### Eric Maxwell

Encabeza la práctica de transformación de DevOps en Google, donde asesora a las mejores empresas del mundo sobre cómo mejorar brindando valor más rápido. Eric se desempeñó la primera mitad de su carrera como ingeniero en el terreno, automatizando todo lo posible y generando empatía con otros profesionales. Fue el cofundador del Programa de modernización de aplicaciones en la nube (CAMP) de Google y es miembro del equipo de DORA. Antes de Google, Eric se dedicó a preparar cosas increíbles con otros colegas de Chef Software.



### James Brookbank

Es arquitecto de soluciones de nube de Google. Los arquitectos de soluciones ayudan a los clientes de Google Cloud resolviendo problemas técnicos complejos y proporcionando orientación experta relacionada con la arquitectura. Antes de formar parte de Google, James trabajó en varias empresas grandes con un enfoque en la infraestructura de TI y servicios financieros.



### Dr. Jeffrey Winer

El Dr. Winer es psicólogo, consultor de sistemas de salud conductual y desarrollador de tratamientos psicológicos en Trauma and Community Resilience Center (TCRC) del Boston Children's Hospital y profesor adjunto de Harvard Medical School. Con sus colegas del TCRC, su trabajo se enfoca principalmente en desarrollar, probar, difundir e implementar intervenciones psicosociales culturalmente responsivas e informadas sobre traumas para jóvenes y familias de inmigrantes y refugiados. Es coautor del libro *Mental Health Practice with Immigrant and Refugee Youth: A Socioecological Framework*.<sup>3</sup> Ha sido asesor en programas en Estados Unidos y Canadá. Las herramientas de intervención y prevención psicológica que ayudó a desarrollar o adaptar se usan actualmente en todo el mundo. Para obtener más información, consulta <http://www.drjeffwiner.com>.

<sup>1</sup> <https://continuous-delivery.co.uk/engineering-for-software>

<sup>2</sup> <https://www.youtube.com/@ContinuousDelivery>

<sup>3</sup> <https://www.apa.org/pubs/books/4317536>



### **Kevin Storer**

Kevin es investigador de la experiencia del usuario en Google, donde lleva a cabo estudios dirigidos a comprender cómo los equipos de desarrollo de software interactúan con herramientas de DevOps y por medio de ellas. Antes de unirse a Google, Kevin obtuvo un doctorado en informática de la University of California, Irvine. En los sectores público y privado, es autor de publicaciones de alto impacto sobre temas como programación centrada en los humanos, experiencia de los desarrolladores, comportamiento de la información, accesibilidad y computación ubicua.



### **Kim Castillo**

Es administradora de programas de experiencia del usuario en Google. Kim dirige la iniciativa multifuncional detrás de DORA, desde supervisar las operaciones de investigación hasta la publicación de este informe. También trabaja en investigación sobre UX para Duet AI en Google Cloud. Antes de Google, Kim hizo carrera en entrega de software, trabajando en administración de programas técnicos y asesoramiento ágil. Sus raíces están en la investigación psicosocial, con un enfoque en temas como ejecuciones extrajudiciales, desarrollo urbano pobre y trauma y resiliencia comunitarios en Filipinas, su país natal. En DORA, Kim combina sus distintas pasiones en prácticas de entrega de software, seguridad psicológica y organizaciones con el modelo Teal.



### **Michelle Irvine**

Es redactora técnica en Google y lleva a cabo investigación relacionada con el impacto y la producción de documentación técnica. Antes de Google, trabajó en edición educativa y como redactora técnica de software de simulación de física. Michelle tiene una licenciatura en física y una maestría en diseño de comunicación y retórica de la University of Waterloo.



### **Nathen Harvey**

Lidera el equipo DORA Advocacy como gerente de ingeniería de relaciones de desarrolladores en Google Cloud. Nathen tuvo el privilegio de trabajar con algunos de los mejores equipos y comunidades de código abierto para ayudarlos a aplicar los principios y las prácticas de DevOps y SRE. Ha sido el coautor del informe *Accelerate State of DevOps* durante los últimos tres años. Nathen también coeditó y contribuyó a la publicación de *97 Things Every Cloud Engineer Should Know*.



### **Steve McGhee**

Es defensor de confiabilidad y ayuda a equipos a comprender cuál es la mejor forma de desarrollar y operar servicios confiables de primer nivel. Antes de eso, trabajó por más de 10 años como ingeniero de confiabilidad de sitios en Google y aprendió cómo escalar sistemas globales en Búsqueda, YouTube, Android y Google Cloud. Administró varios equipos de ingeniería en California, Japón y el Reino Unido. Steve también pasó un tiempo con una empresa de California para ayudarlos a realizar la transición a la nube.

# Metodología

En este capítulo, se detalla el recorrido de un conjunto inicial de ideas al informe que tienes ahora (con ideas un poco más completas). Esperamos que responda muchas de tus preguntas sobre cómo se genera este informe y que funcione como esquema para ayudarte a investigar por tu cuenta.

## Paso 1: Generar un conjunto de resultados que nos parezcan importantes para las organizaciones de tecnología de alto rendimiento

Esto es fundamental. Nuestro programa se basa en ayudar a orientar a las personas hacia los fines que les interesan. Si no sabemos adónde quieren llegar las personas, las organizaciones o los equipos, empezamos con el pie izquierdo. ¿Cómo podemos averiguarlo? Usamos una mezcla de investigación cualitativa (es decir, preguntarles a las personas qué quieren lograr ellos, sus equipos y organizaciones), encuestas, interacción con la comunidad general y muchos talleres. Una y otra vez, descubrimos resultados como los siguientes:

- **Rendimiento organizativo.** La organización debería generar no solo ingresos, sino también valor para los clientes y la comunidad en general.
- **Rendimiento del equipo.** Se refiere a la habilidad de un equipo de aplicaciones o servicio para crear valor, innovar y colaborar.
- **Bienestar de los empleados.** Las estrategias que adopta una organización o equipo deberían beneficiar a los empleados; por ejemplo, reducir el agotamiento, fomentar una experiencia laboral satisfactoria y aumentar la capacidad de las personas para producir resultados valiosos (es decir, productividad).

También escuchamos que se mencionan objetivos como los siguientes:

- **Rendimiento de la entrega de software.** Los equipos pueden implementar software de manera rápida y exitosa.
- **Rendimiento operativo.** El software que se envía proporciona una experiencia confiable al usuario.

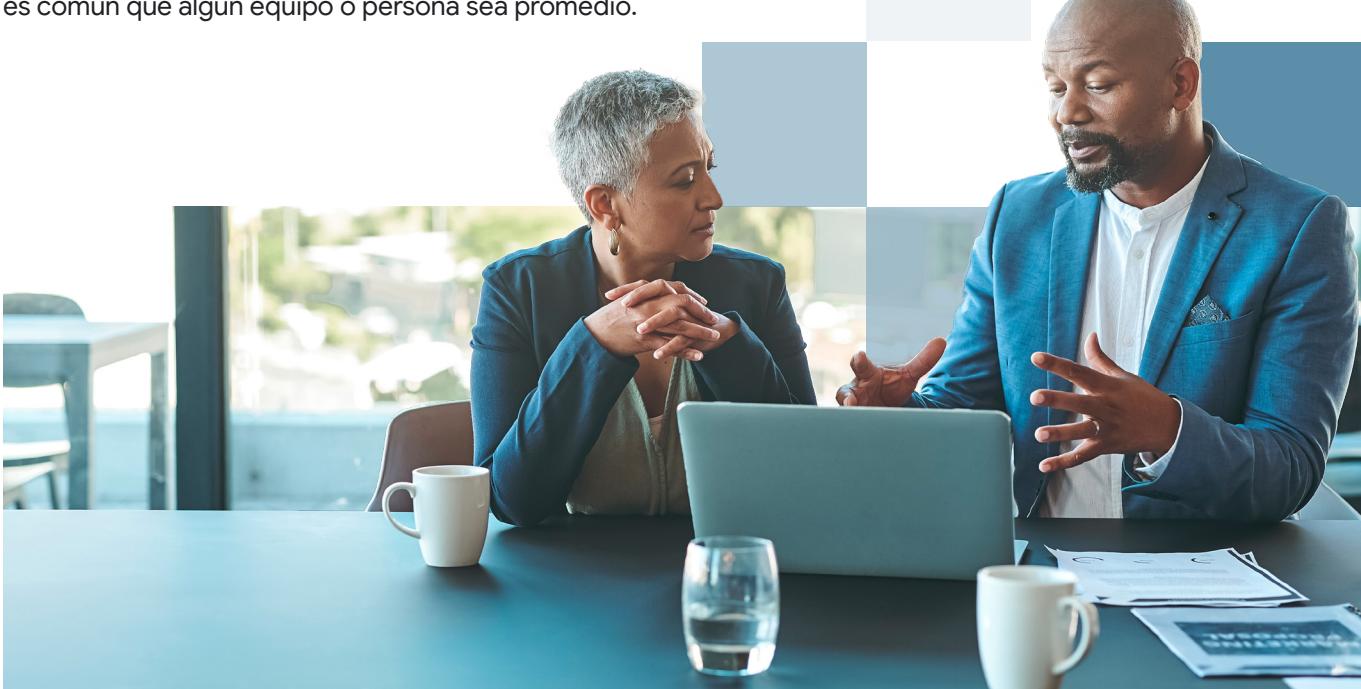


## Paso 2: Crear hipótesis sobre cómo, cuándo y por qué se logran esos resultados

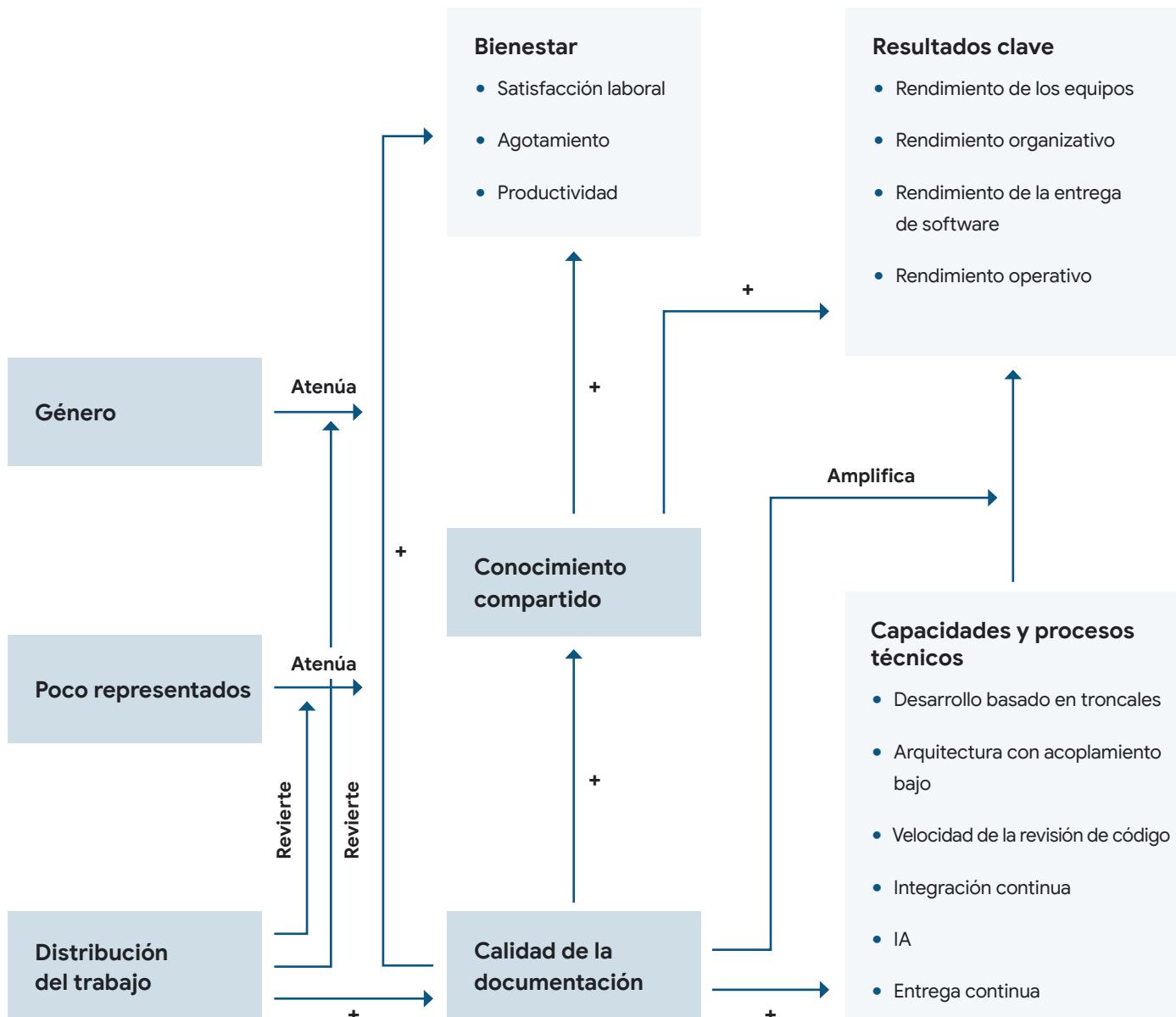
Con los resultados del paso 1 en mente, debemos crear hipótesis sobre rutas que podrían usar los equipos para alcanzarlos. Para eso, es necesario buscar factores que parezcan tener un impacto confiable en un resultado. Queremos decir algo como que, si todo se mantiene igual,  $x$  tiene un efecto en  $y$ . Esa información puede ayudar a los profesionales a tomar decisiones fundamentadas en datos sobre qué tipo de cambios probar.

No solo queremos entender qué rutas tienen un impacto, sino que también queremos explorar bajo qué *condiciones* podrían tener un impacto mayor o menor. Eso equivale a preguntar "cuándo" y "para quién". Por ejemplo, se ha comprobado que la calidad de la documentación reduce drásticamente el agotamiento... en promedio. Sin embargo, cuando prestamos atención a los encuestados que informan que son poco representados, lo opuesto es cierto: la calidad de la documentación *aumenta* el agotamiento. Comprender las condiciones de esos efectos es esencial porque no es común que algún equipo o persona sea promedio.

Además, proponemos hipótesis sobre *mecanismos* que explican *por qué* o *cómo*. Este año, planteamos la hipótesis de que las personas que se identifican como poco representadas sienten más agotamiento, en función de resultados que observamos el año pasado y en una amplia bibliografía sobre el tema. La primera pregunta que se hacen las personas es: "*¿Por qué ocurre esto?*". Para tratar de responder, planteamos hipótesis sobre posibles mecanismos a evaluar. Por ejemplo, las personas que se identifican como poco representadas podrían sentir más agotamiento porque deben hacer o se les asigna más trabajo repetitivo.



Luego, esas hipótesis se ordenan para que podamos diseñar nuestra encuesta y ejecutar los análisis. El siguiente es un ejemplo de un modelo hipotético para el capítulo de documentación, con efectos, condicionalidades y mecanismos:

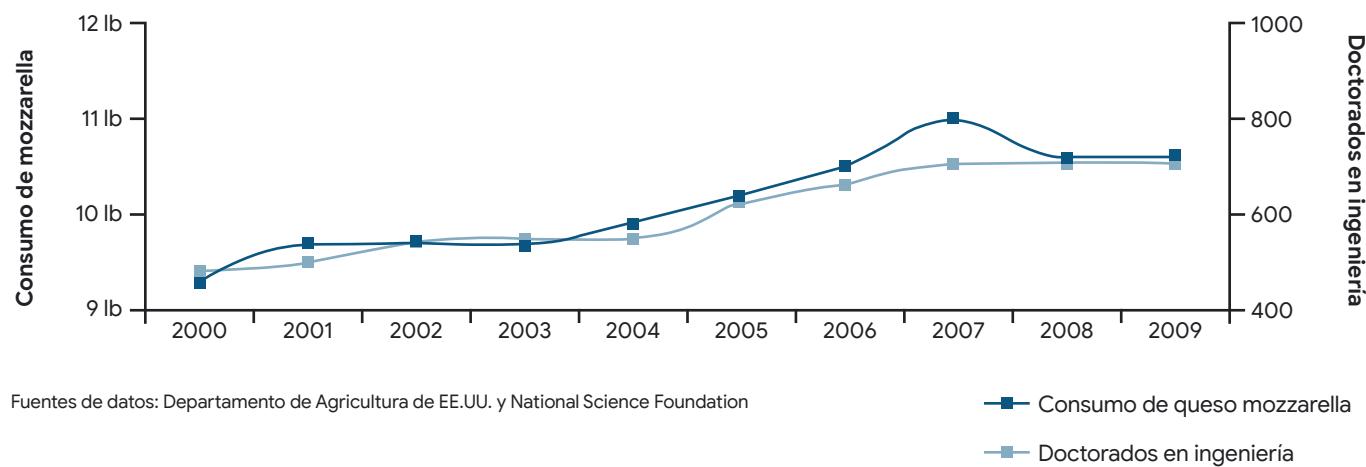


## Paso 3: Plantear hipótesis sobre posibles confusiones

Si alguna vez analizaste datos, es probable que hayas encontrado correlaciones espurias. Quizá conozcas un sitio web que muestra muchas correlaciones espurias,<sup>1</sup> como el siguiente ejemplo:

**El consumo de queso mozzarella per cápita se correlaciona con los doctorados de ingeniería civil otorgados**

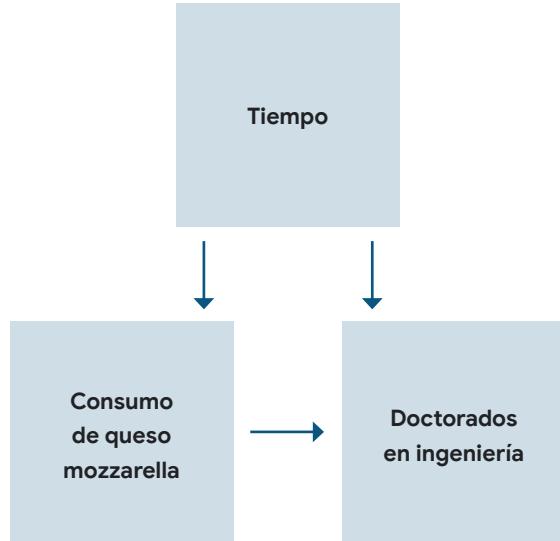
Correlación: 95.86% ( $r=0.958648$ )



Es poco probable que haya una conexión causal entre los doctorados de ingeniería y el consumo de queso mozzarella.

No obstante, hay un elemento de confusión que se asoma detrás de esa relación: el tiempo. Si el consumo de queso mozzarella y los doctorados en ingeniería tienen tendencia positiva en el mismo periodo, es probable que tengan una correlación positiva.

Es probable que incluir el tiempo en un modelo o quitar la tendencia de los datos anule la relación. Podemos ilustrar el modelo de la siguiente manera:



<sup>1</sup> <https://www.tylervigen.com/spurious-correlations>

Si no tomamos en cuenta el tiempo (una tercera variable), es posible que los datos muestren una relación espuria entre el consumo de queso mozzarella y los doctorados de ingeniería.

Hay herramientas que ayudan a los investigadores con esto, como Dagitty (<https://dagitty.net/dags.html>), que nos permite especificar el modelo causal. Para ayudarnos a calcular de manera adecuada el efecto de X en Y, nos indica las implicaciones del modelo, lo que debemos tener en cuenta y lo que *no* debemos considerar. Las herramientas como Dagitty pueden llevar a la conclusión de que la correlación podría no implicar causación, pero sin duda indica la manera en que alguien piensa sobre ello.

Es imposible capturar todos los elementos que sesgan las estimaciones de los investigadores; piensa en encuestas de 50 horas y en la omnisciencia. De cualquier manera, nos esforzamos por registrar las rutas de sesgo para brindar estimaciones precisas sobre los efectos de distintas actividades, tecnologías y estructuras en los resultados que son importantes para ti. En última instancia, muchos profesionales desean saber cuáles son los factores que tendrán un impacto en esos resultados clave. Los modelos que no toman en consideración el sesgo no podrán proporcionar a los profesionales la orientación que necesitan. No queremos decirle a alguien que el queso mozzarella aumenta el rendimiento de la entrega de software, y cometer ese error sería mucho más fácil de lo que te imaginas.



## Paso 4: Desarrollar la encuesta

Hay tres aspectos para desarrollar la encuesta: operacionalización, experiencia y localización.

### Operacionalización

Queremos mediciones que capturen de manera adecuada los conceptos que nos interesan y que lo hagan de forma confiable. Traducir un concepto abstracto en algo medible es el arte de la operacionalización. Esas mediciones son los ingredientes que forman la base de todo el análisis. Si nuestras mediciones no brindan indicadores claros, ¿cómo podemos confiar en el resto del análisis? ¿Cómo medimos un concepto tan elusivo como, por ejemplo, la productividad? ¿Qué pasa con el agotamiento o el rendimiento operativo?

Primero, consultamos la bibliografía para ver si hay mediciones exitosas que ya existan. Si podemos usar mediciones validadas previamente en nuestra encuesta, obtenemos un puente de la encuesta a toda la bibliografía que se generó en torno a ese tema. El uso continuo de la *tipología de culturas organizativas* de Westrum es un ejemplo de nuestra reutilización de mediciones validadas previamente.

Sin embargo, muchos conceptos no fueron validados previamente para los campos que investigamos. En ese caso, llevamos a cabo una investigación cualitativa para desentrañar cómo se entiende el concepto y analizamos la bibliografía filosófica sobre sus particularidades.

### Experiencia de la encuesta

Queremos que la encuesta sea comprensible, sencilla, no más larga de lo necesario y ampliamente accesible. Esos objetivos son difíciles, debido a todas las preguntas que queremos hacer, la comprensión técnica que se necesita para responder y la variación en la nomenclatura para determinadas prácticas. Hacemos evaluaciones remotas y sin moderadores para asegurarnos de que la encuesta tenga un rendimiento superior a ciertos umbrales. Para eso, es necesario hacer múltiples iteraciones.

### Localización

Personas de todo el mundo responden la encuesta todos los años. Este año, nos esforzamos para que la encuesta fuera más accesible para un público más amplio mediante su localización al inglés, español, francés, portugués y japonés. Esta fue una iniciativa comunitaria, liderada por algunos miembros increíbles de la comunidad de DORA. Googlers de todo el mundo contribuyeron a esta iniciativa, además de un socio en el campo: muchas gracias a Zenika (<https://www.zenika.com>) por la localización al francés. Esperamos ampliar esas iniciativas y lograr que la encuesta sea verdaderamente intercultural.

## Paso 5: Recopilar respuestas de la encuesta

Usamos varios canales para el reclutamiento, que se clasifican en dos categorías: *orgánico* y *de paneles*.

El *enfoque orgánico* es usar todos los medios sociales a nuestra disposición para hacer saber que hay una encuesta que queremos que completen las personas. Publicamos entradas de blog. Usamos campañas por correo electrónico. Publicamos contenido en redes sociales y les pedimos a los miembros de la comunidad que hagan lo mismo (es decir, muestreo de bola de nieve).

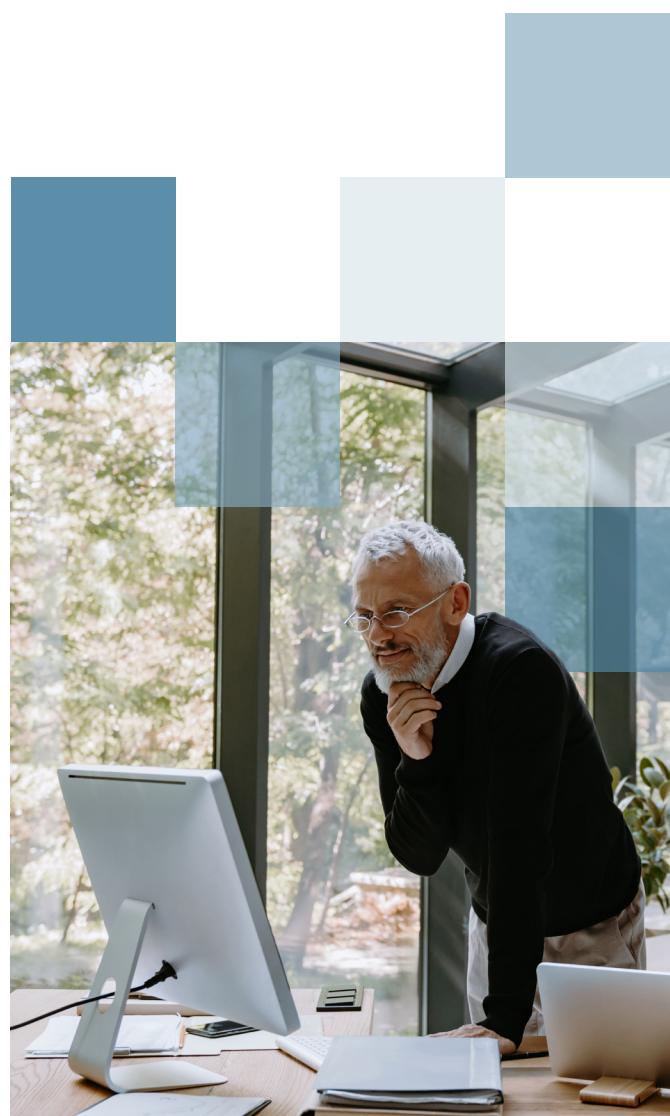
Usamos el *enfoque de paneles* para complementar el canal orgánico. Intentamos reclutar a personas tradicionalmente poco representadas en la comunidad técnica en general y buscamos respuestas adecuadas de determinados sectores y tipos de organización. En pocas palabras, aquí es donde tomamos algo de control sobre el reclutamiento, el cual no tenemos con el enfoque orgánico. El enfoque de paneles también nos permite asegurarnos de tener suficientes encuestados porque nunca sabemos si el enfoque orgánico arrojará las respuestas necesarias para hacer los tipos de análisis que hacemos.

## Paso 6. Analizar los datos

Hay tres pasos centrales del análisis: limpieza de datos, validación de medición y evaluación de modelos.

### Limpieza de datos

El objetivo de la limpieza de datos es aumentar la relación entre señal y ruido. Por distintas razones, algunas respuestas son ruido. Algunos patrones de respuesta pueden indicar que la persona está distraída, haciendo la encuesta sin leer con atención o no responde de buena fe. Esto es inusual en nuestros datos, pero ocurre. Las respuestas con indicios de tergiversación de las experiencias reales de una persona se excluyen del análisis porque son ruido. El desafío a la hora de descartar el ruido es asegurarse de no desechar la señal, en especial de manera sesgada o de una forma que valide nuestras hipótesis. Por ejemplo, si concluimos que nadie podría tener una puntuación alta en un valor determinado y baja en otro, podríamos excluir a ese encuestado, lo que haría que los datos estén más alineados con nuestras creencias y aumentaría las probabilidades de que se confirmen nuestras hipótesis.



## Validación de mediciones

Al comienzo de este informe, hablamos sobre los conceptos que tratamos de medir. Podríamos participar de muchos juegos del lenguaje distintos, pero, según un punto de vista, esta medición de un concepto se llama variable.<sup>2</sup> Estas variables son los ingredientes de los modelos, que son los elementos que se incluyen en nuestra investigación. Hay dos formas generales de analizar la validez de estas mediciones: internamente y externamente.

Para comprender la validez interna de la medición, analizamos lo que creemos que indica la presencia de un concepto. Por ejemplo, el hecho de que las personas usen la documentación para resolver problemas podría indicar documentación de calidad.

Una mayoría de nuestras variables consta de varios indicadores porque los constructos que nos interesan parecen ser multifacéticos. Para comprender la naturaleza multifacética de una variable, probamos qué tan bien combinan los elementos que usamos para representar ese constructo. Si lo hacen bien (es decir, comparten un nivel alto de varianza comunitaria), suponemos que hay algo subyacente, como el concepto de interés.

Piensa en la felicidad, por ejemplo: es multifacética. Cuando alguien está feliz, esperamos que se sienta, actúe y piense de cierta manera. Suponemos que la felicidad es subyacente a determinado patrón de sentimientos, pensamientos y acciones. Por lo tanto, esperamos que ciertos tipos de sentimientos, pensamientos y acciones surjan cuando la felicidad está presente. Luego, haríamos preguntas sobre esos sentimientos, pensamientos y acciones. Usaríamos análisis factorial confirmatorio para evaluar si en verdad aparecen al mismo tiempo.

Este año usamos el paquete de R lavaan para hacer el análisis.<sup>3</sup> Lavaan devuelve una variedad de estadísticas de bondad de ajuste que nos ayudan a comprender si un constructo de verdad representa la forma en que las personas responden las preguntas. Si los indicadores de un concepto no se combinan bien, es posible que se deba revisar el concepto o descartarlo porque está claro que no encontramos una forma confiable de medirlo.

Para determinar la validez externa de un constructo, hay que analizar cómo encaja en el mundo. Podemos esperar que un constructo tenga determinadas relaciones con otros. A veces podemos esperar que dos constructos tengan una relación negativa, como la felicidad y la tristeza. Si la medición de la felicidad tiene una correlación positiva con la tristeza, podríamos poner en duda la medición o la teoría. Del mismo modo, podríamos esperar que dos constructos tengan relaciones positivas, pero no muy sólidas. Es probable que la productividad y la satisfacción laboral tengan una correlación positiva, pero no creemos que sean idénticas. Si la correlación es demasiado alta, podríamos decir que parece que estamos midiendo lo mismo. Esto, a su vez, significa que nuestras mediciones no están calibradas lo suficiente como para detectar las diferencias entre los dos conceptos o la diferencia que planteamos no existe.

<sup>2</sup> Moore, Will H. y David A. Siegel. A mathematics course for political and social research. Princeton University Press, 2013.

<sup>3</sup> Rosseel, Y. "lavaan: An R Package for Structural Equation Modeling". Journal of Statistical Software, 48(2), 2012. 1–36. <https://doi.org/10.18637/jss.v048.i02>

## Evaluación del modelo

En los pasos 2 y 3, desarrollamos modelos hipotéticos. Después de obtener datos limpios en el paso 6, vemos qué tan bien se adaptan esos modelos a los datos. Este año adoptamos un enfoque bayesiano para comprender qué tan plausibles son las distintas hipótesis según los datos, en lugar de qué tan probables son los datos dada una hipótesis nula (es decir, no hay ningún efecto). Las herramientas principales que usamos en R son blavaan<sup>4</sup> y rstanarm.<sup>5</sup> Podemos evaluar la probabilidad de que un efecto sea considerable o dramático y no simplemente menor. Para evaluar un modelo, optamos por la parsimonia. Eso implica empezar con un modelo muy simplista y agregar complejidad hasta que no se justifique más. Por ejemplo, predecimos que el rendimiento organizativo es el producto de la interacción entre el rendimiento de la entrega de software y el operativo.

Nuestro modelo simplista no incluye la interacción:

Rendimiento organizativo ~ Rendimiento de la entrega de software + Rendimiento operativo

En nuestro segundo modelo, se agrega la interacción:

Rendimiento organizativo ~ Rendimiento de la entrega de software + Rendimiento operativo + Rendimiento de la entrega de software × Rendimiento operativo

Según las recomendaciones en Regression and other stories<sup>6</sup> y Statistical Rethinking,<sup>7</sup> usamos la validación cruzada dejando uno fuera (LOOCV) y el criterio de información de aplicación amplia Watanabe–Akaike<sup>8</sup> para determinar si es necesaria la complejidad adicional.

<sup>4</sup> Merkle, Yves Rosseel y Edgar C. "blavaan: Bayesian structural equation models via parameter expansion". arXiv, versión preliminar, 2015. <https://doi.org/10.48550/arXiv.1511.05604>

<sup>5</sup> Goodrich, Ben, Jonah Gabry, Imdad Ali y Sam Brilleman. "rstanarm: Bayesian applied regression modeling via Stan". Paquete R versión 2, no. 1 (2020).

<sup>6</sup> Gelman, Andrew, Jennifer Hill y Aki Vehtari. Regression and Other Stories. Cambridge University Press, 2020.

<sup>7</sup> McElreath, Richard. Statistical rethinking: A Bayesian course with examples in R and Stan. Chapman y Hall/CRC, 2018.

<sup>8</sup> Vehtari, Aki, Andrew Gelman y Jonah Gabry. "Practical Bayesian model evaluation using leave-one-out cross-Validation and WAIC". Statistics and Computing, 27, 2017. 1413-1432. <https://doi.org/10.1007/s11222-016-9696-4>

## Paso 7: Hallazgos del informe

Luego, revisamos esos resultados como equipo. Este año, pasamos algunos días juntos en Boulder, Colorado, sintetizando los datos con las experiencias de los expertos en la materia. Hicimos esto para todos los capítulos del informe, una hipótesis a la vez. La interpretación de datos siempre tiene riesgos de tener sesgos, especulaciones, anécdotas y suposiciones. Esos riesgos se mitigaron teniendo muchas personas con antecedentes diversos en una sala en la que se alentaron las preguntas, la divergencia, los puntos de vista únicos y la curiosidad.<sup>9</sup>

Con los resultados a mano, los autores del informe volvieron a sus lugares de origen y escribieron. Durante el proceso de redacción, se consultó con editores y expertos en la materia. Contar con esos puntos de vista fue fundamental para ayudarnos a comunicar nuestras ideas. La persona responsable de analizar esos datos también debió asegurarse de que nada de lo que dijéramos se desviara de los datos.

Esos capítulos se agruparon en un diseño cohesivo de nuestros talentosos socios de diseño, BrightCarbon.<sup>10</sup>

## Paso 8. Sintetizar los hallazgos con la comunidad

Nos apoyamos en la participación de la comunidad para identificar formas de aprovechar e interpretar estos hallazgos. Intentamos ser específicos en nuestras recomendaciones, pero, en última instancia, hay innumerables implementaciones que un equipo podría probar en función de los resultados que descubrimos. Por ejemplo, la arquitectura con acoplamiento bajo parece ser una práctica beneficiosa según los resultados que medimos. Pero, sin duda, no hay una única forma de establecer una arquitectura con acoplamiento bajo. Generar y compartir enfoques como comunidad es la única manera de mejorar continuamente. Nuestro mapa del mundo es una interpretación y abstracción del territorio y contexto en el que tú, tu equipo y tu organización operan.

Para participar en la comunidad global de práctica de DORA, visita el sitio de DORA Community (<https://dora.community>).

<sup>9</sup> Stasser, G., y Titus, W. (1985). "Pooling of unshared information in group decision making: Biased information sampling during discussion". *Journal of Personality and Social Psychology*, 48(6), 1985. 1467–1478. <https://doi.org/10.1037/0022-3514.48.6.1467>

<sup>10</sup> <https://www.brightcarbon.com/>

# Datos demográficos y organizativos

## Quién respondió la encuesta

Durante casi una década, el programa de investigación de DORA ha estudiado las capacidades, prácticas y mediciones de las organizaciones de tecnología de alto rendimiento. Recibimos opiniones de más 36,000 profesionales que trabajan en organizaciones de todos los tamaños y en muchos sectores diferentes. ¡Gracias por compartir tus observaciones! Este año, casi 3,000 profesionales de distintos sectores de todo el mundo compartieron sus experiencias para ayudar a aumentar nuestra comprensión de los factores que impulsan a las organizaciones de tecnología de alto rendimiento.

Este año, se aprovechó la investigación de Stack Overflow en las preguntas de datos demográficos y organizativos. Participaron más de 70,000 encuestados en la Stack Overflow Developer Survey de 2022.<sup>1</sup> Esa encuesta no llegó a todos los profesionales técnicos por distintos motivos, pero es lo más cercano a un censo de desarrolladores. Con la idea de la población que nos proporcionó la encuesta, identificamos sesgos en las respuestas en nuestros datos y comprendimos hasta dónde podríamos generalizar nuestros hallazgos. Esta información y las

preguntas sobre datos demográficos y organizativos incluidas en la Stack Overflow Developer Survey están bien diseñadas y vale la pena tomarlas prestadas. En cuanto a la encuesta de Stack Overflow, nuestro conjunto de muestra incluye una proporción más alta de mujeres, participantes con discapacidad y participantes que trabajan en organizaciones más grandes. Nuestro conjunto de muestra es similar al de Stack Overflow en cuanto a origen étnico.

La encuesta de este año registró 3.6 veces más encuestados orgánicos que la de 2022.

La encuesta de este año registró  
3.6 veces  
más encuestados orgánicos que en 2022

<sup>1</sup> <https://survey.stackoverflow.co/2022#overview>

# Datos demográficos

## Género

En comparación con 2022, la muestra de este año tuvo una proporción menor de mujeres encuestadas (12% frente a 18%).

Género	% de encuestados
Prefiero no decirlo	3%
O, en tus propias palabras	2%
Mujer	12%
Hombre	81%

## Discapacidad

Identificamos la discapacidad en función de seis dimensiones que siguen los lineamientos de la Lista breve de preguntas sobre la discapacidad del Grupo de Washington.<sup>2</sup> Este es el quinto año que hacemos preguntas sobre discapacidad. El porcentaje de personas con discapacidad se redujo de 11% en 2022 a 6% en 2023.

Discapacidad	% de encuestados
Ninguna de las discapacidades aplica	87%
Sí	6%
Prefiere no decirlo / no respondió	7%

## Poco representados

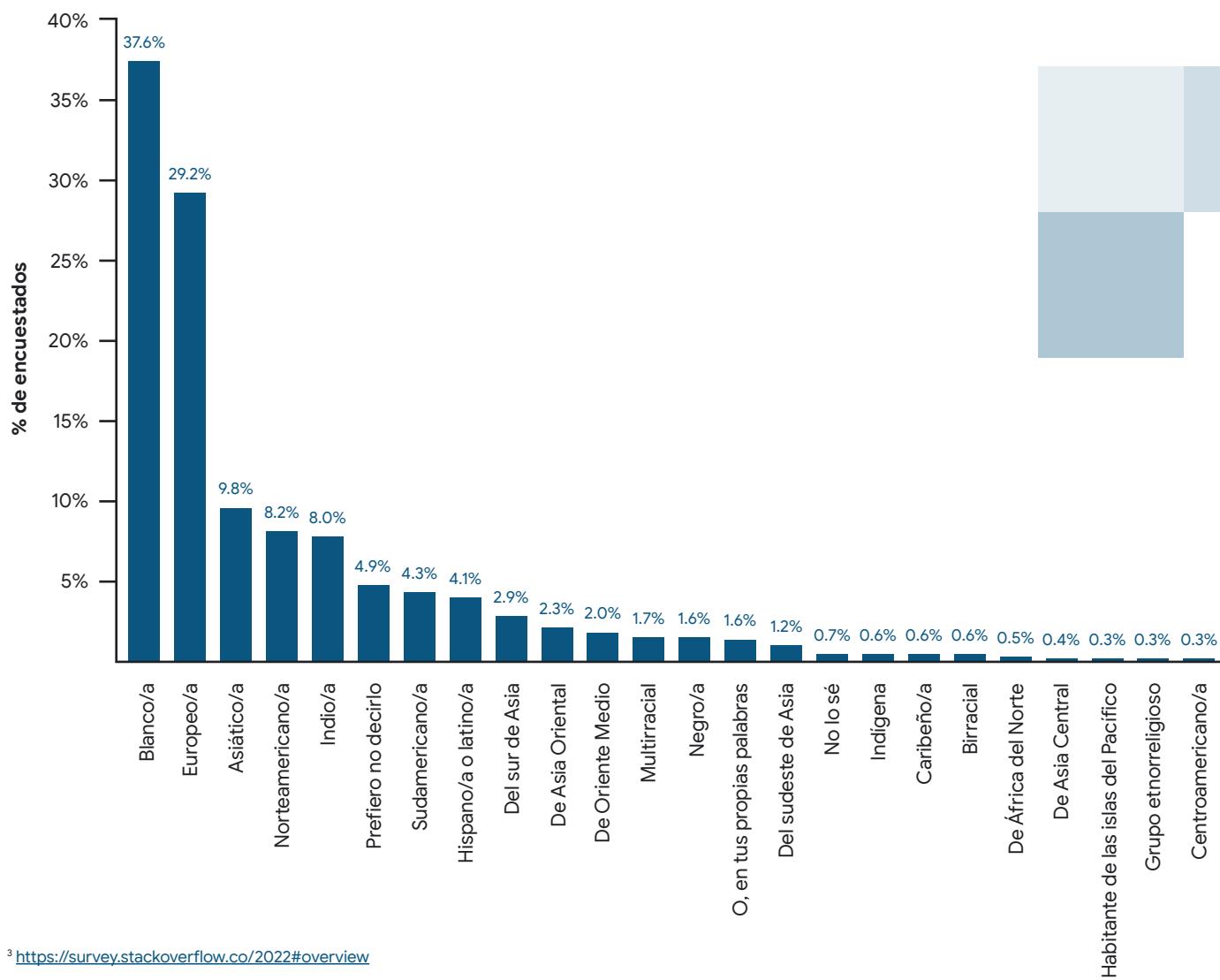
La identificación como miembro de un grupo poco representado puede referirse al origen étnico, el género o alguna otra característica. Este es el sexto año que preguntamos sobre la poca representación. El porcentaje de personas que se identifican como poco representadas se redujo un poco de 19% en 2022 a 15% en 2023.

Poco representados	% de encuestados
No	77%
Sí	15%
Prefiere no responder	7%

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

## Origen étnico

Tomamos la pregunta de la encuesta Stack Overflow Developer de 2022.<sup>3</sup> Como se mencionó antes, nuestro conjunto de muestra es similar, con una desviación notable: tenemos una proporción más baja de europeos.

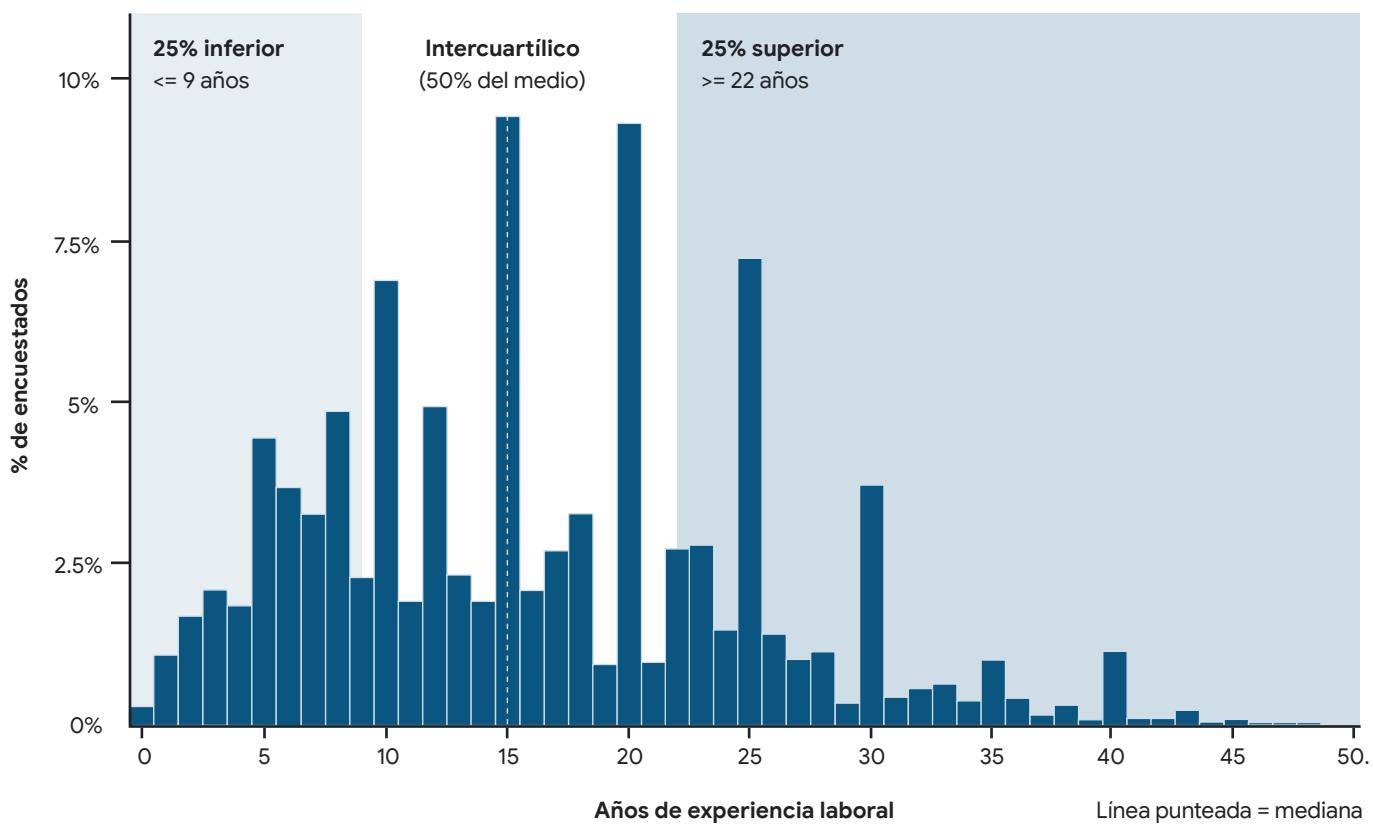


<sup>3</sup> <https://survey.stackoverflow.co/2022#overview>

# Años de experiencia

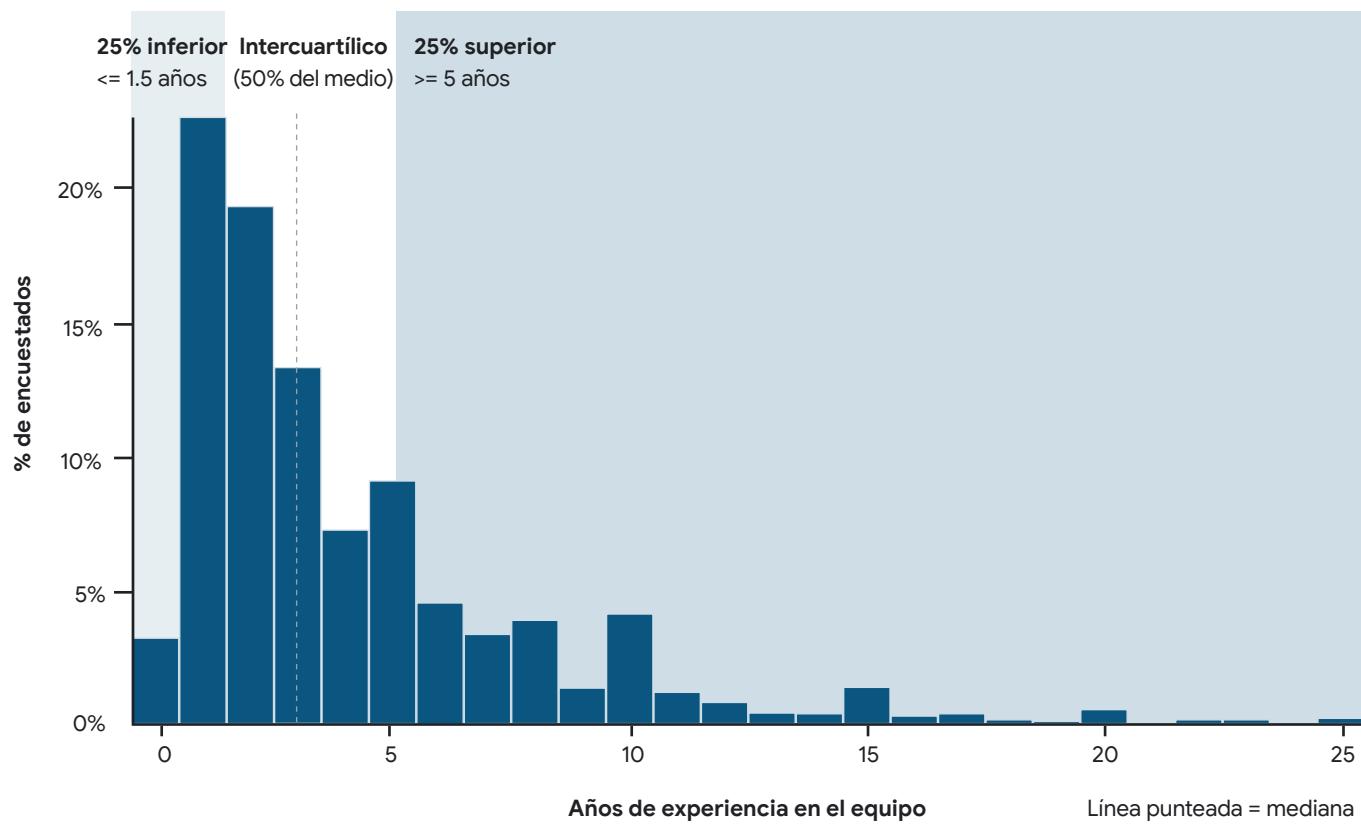
## Experiencia laboral

Queríamos saber cuánto tiempo ha trabajado una persona. Aquí hicimos la siguiente pregunta: "¿Cuántos años de experiencia laboral tienes?". En pocas palabras, recopilamos datos de un grupo de profesionales con bastante experiencia. Es decir, el 50% de los encuestados tenía 15 años de experiencia o más. El 25% de los encuestados tenía más de 22 años de experiencia. El 25% de los encuestados tenía menos de 9 años de experiencia. En retrospectiva, no es obvio qué entiende una persona por "trabajo".



## Experiencia en el equipo

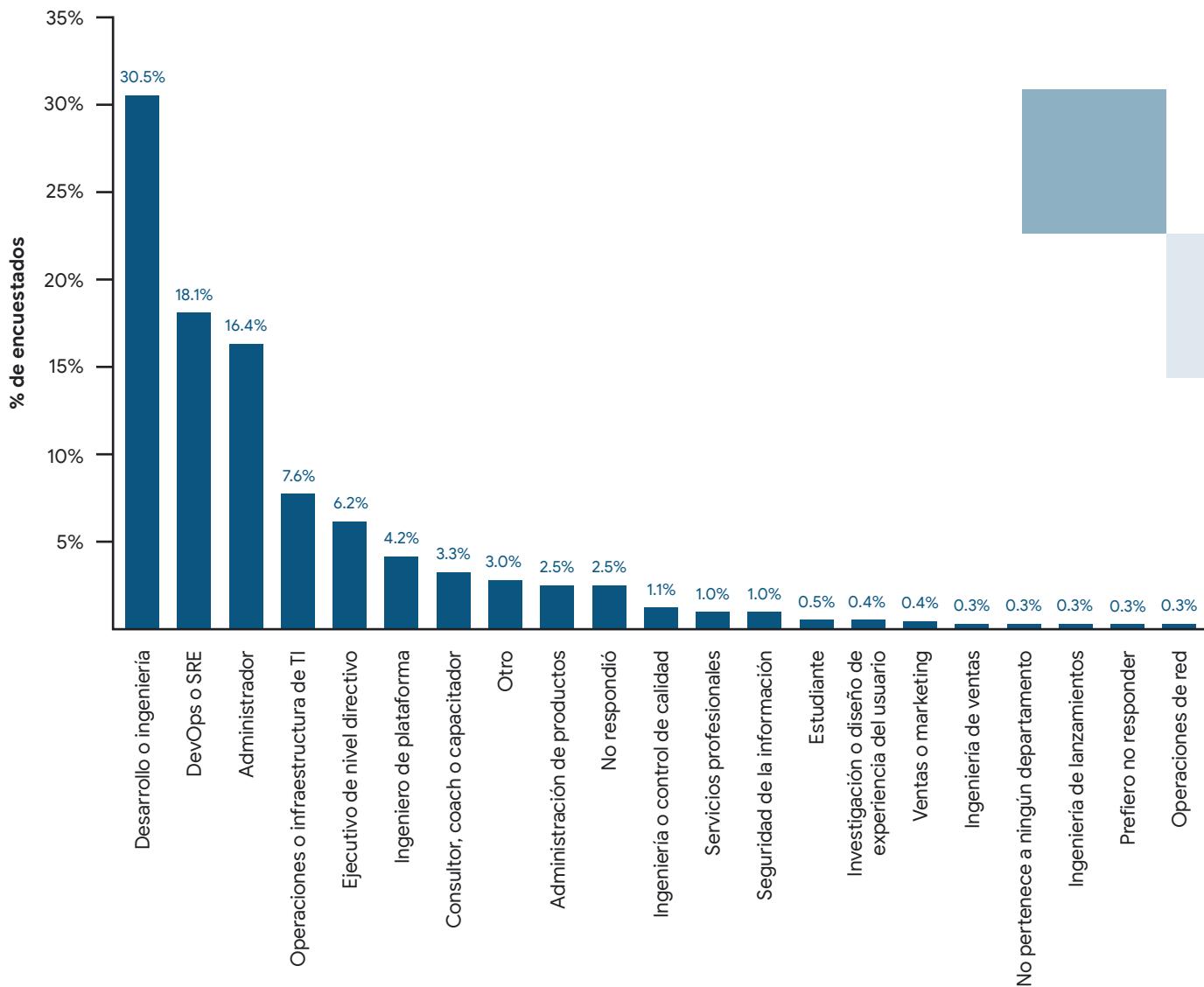
A veces, trabajar en un equipo nuevo se siente como volver a empezar. Para analizar ese aspecto, preguntamos lo siguiente: "¿Hace cuántos años trabajas en el equipo en el que estás actualmente?". Aunque la experiencia laboral parecía alta, muchas de las personas encuestadas eran nuevas en sus equipos. El 50% de los encuestados ha estado en su equipo nuevo hace menos de 3 años. El 25% ha formado parte de su equipo por un período igual o inferior a 1.5 años. Solo el 25% de los encuestados ha estado en su equipo desde hace 5 años o más. ¿Esto refleja una mentalidad de mejora continua en las personas encuestadas? ¿O refleja el movimiento y la inestabilidad de la economía?



## Función

El 72% de los encuestados son personas que trabajan en equipos de desarrollo o ingeniería (30%), de DevOps o SRE (18%), de operaciones o infraestructura de TI (8%), o son gerentes (16%). En 2022, las personas en esos puestos representaban el 85% de los encuestados.

La disminución en la cantidad de encuestados de esas cuatro funciones sugiere que pudimos llegar a más personas en puestos diferentes. La proporción de equipos de operaciones o infraestructura de TI (8%) volvió a niveles de 2021 (9%) después de un cambio de rumbo en 2022 (19%).



# Datos organizativos

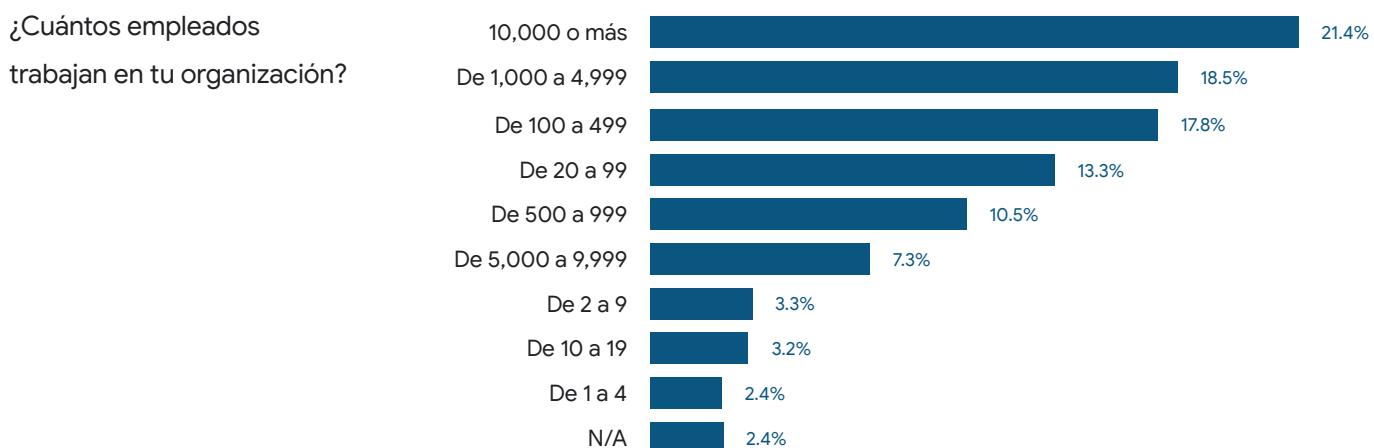
## Sector

## % de encuestados



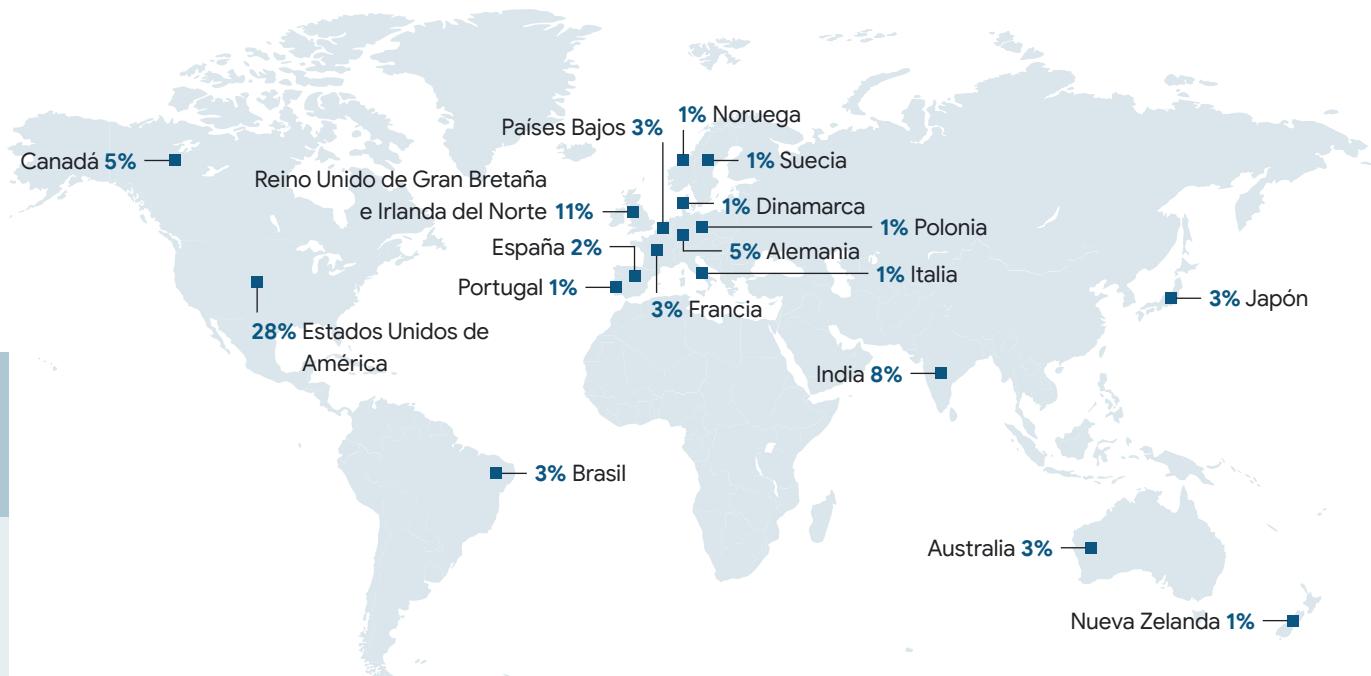
## Cantidad de empleados

## % de encuestados



## País

Siempre nos emociona ver que personas de todo el mundo participan en la encuesta. ¡Muchas gracias a todos!



País	Dinamarca	Lituania	Túnez	Bangladés	Guatemala
EE.UU.					
Reino Unido	Suiza	Tailandia	Uruguay	República Dominicana	Honduras
India	Austria	Hungría	Afganistán	Ghana	Letonia
Canadá	Kenia	Israel	Argelia	Región Administrativa Especial de Hong Kong	Libano
Alemania	Sudáfrica	Vietnam	Egipto	Kazajistán	Luxemburgo
Australia	Argentina	EAU	Estonia	Birmania	Maldivas
Brasil	República Checa	Bulgaria	Islandia	Arabia Saudita	Malta
No aplicable	Bélgica	Croacia	Irán	Somalia	Mauricio
Países Bajos	Colombia	Ecuador	Nigeria	Sudán	Mongolia
Japón	Finlandia	Indonesia	Perú	Uganda	Marruecos
Francia	Irlanda	Filipinas	Eslovaquia	Albania	Nepal
España	China	Armenia	Eslovenia	Bahamas	Catar
Suecia	Rumania	Georgia	Corea del Sur	Bielorrusia	Antigua República Yugoslava de Macedonia
Italia	Singapur	Grecia	Sri Lanka	Bolivia	
Nueva Zelanda	México	Malasia	Andorra	Camboya	Trinidad y Tobago
Polonia	Türkiye	Pakistán	Angola	Costa Rica	República Unida de Tanzania
Noruega	Ucrania	Federación Rusa	Antigua y Barbuda	Yibuti	
Portugal	Chile	Serbia	Baréin	El Salvador	Zimbabue

# Arreglo laboral

## Estado laboral

El 88% de los encuestados son empleados de tiempo completo. El 10% de los encuestados son contratistas. Algunos contratistas informan experiencias ampliamente diferentes a las de los empleados de tiempo completo.

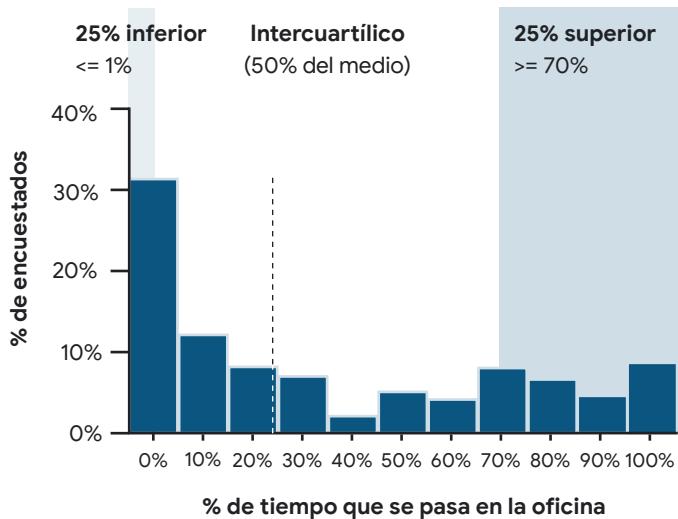
Contrato	% de encuestados
Empleado de tiempo completo	88%
Contratistas de tiempo completo	8%
Empleado a tiempo parcial	2%
Contratista a tiempo parcial	2%

\* Con respecto a la aplicación o el servicio principal en el que trabajas, ¿cuál es la mejor descripción de tu estado laboral en la organización propietaria de la aplicación o el servicio?

La diferencia en la experiencia podría deberse al lugar que ocupan en el equipo. Algunos contratistas informan que están integrados en el equipo con el que trabajan. Esto quiere decir que trabajan estrechamente con los miembros del equipo todos los días y consideran que no hay diferencias entre ellos y un empleado de tiempo completo. El 70% de los encuestados contratistas está muy de acuerdo o de acuerdo con la afirmación de que están integrados en el equipo.

## Ubicación

El patrón de respuestas de este año indica que, a pesar de los intentos para volver a la oficina, el trabajo desde casa sigue siendo una realidad para muchos trabajadores. Cerca del 33% de los encuestados trabaja casi exclusivamente desde casa (pasa menos del 5% del tiempo en la oficina). El 63% de los encuestados trabaja desde casa más de lo que trabaja en la oficina. Para el resto de los encuestados, el trabajo híbrido podría ser el arreglo más común, ya que el 75% pasa menos del 70% de su tiempo en la oficina. No hay muchas personas que tengan un vínculo estrecho con la oficina. Solo el 9% de los encuestados está en la oficina más del 95% del tiempo.



# Los modelos

## Introducción

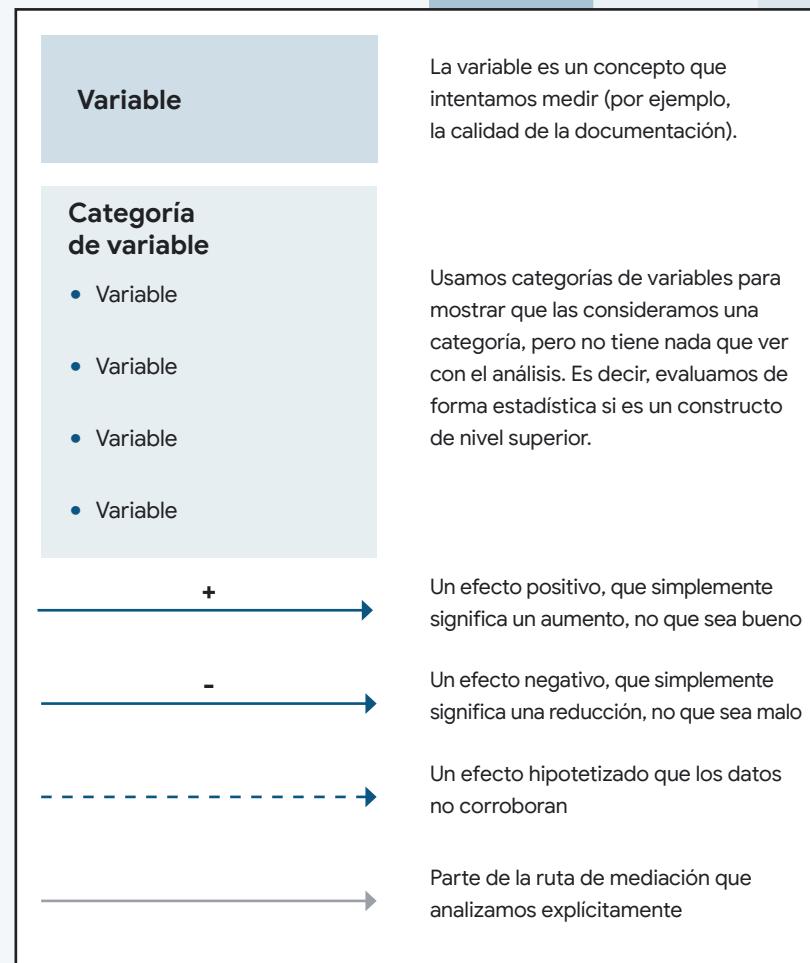
Tradicionalmente, creábamos un modelo gigante. Este año decidimos desglosarlo en varios modelos por las siguientes razones:

- Los modelos grandes pueden descontrolarse rápidamente. Cada una de las variables que se agregan cambia la forma en la que funciona el modelo. Eso puede generar estimaciones imprecisas y dificulta encontrar la razón de un cambio.
- Este año, creamos nuestras hipótesis una sección a la vez. Por lo tanto, tiene sentido crear un modelo para cada sección.
- No es obvio cuál es el beneficio de un modelo gigante para estimar el efecto de X en Y. Para comprender el impacto de X en Y, usamos grafos acíclicos dirigidos para ayudar a entender qué variables de control deberíamos incluir o no en el modelo.
- Sería muy difícil para el lector darle sentido a la cantidad de hipótesis que abordamos este año en un modelo gigante. Imagina combinar todas las visualizaciones incluidas a continuación en una sola.



# ¿Cómo leo estos diagramas?

Una vez que aprendes cómo leer estos diagramas, verás que son herramientas eficientes para comunicar mucha información.



**Ten en cuenta que los modelos son sumas generales.** Categorizamos algunas variables juntas para facilitar la lectura. Gracias a esta estrategia, la flecha que va a una categoría de variable, desde una categoría de variable, o ambas, puede ser el patrón general de resultados, aunque quizás no sea verdadero para todas las variables en la categoría. Por ejemplo, compartir conocimiento tiene un impacto positivo en la mayoría de los resultados. Por lo tanto, trazaríamos una flecha con el símbolo más (+) que iría de compartir conocimiento a la categoría de variable de los resultados clave. Sin embargo, compartir conocimiento no tiene un impacto positivo en el rendimiento de la entrega de software. Para conocer los detalles, consulta los capítulos relevantes.

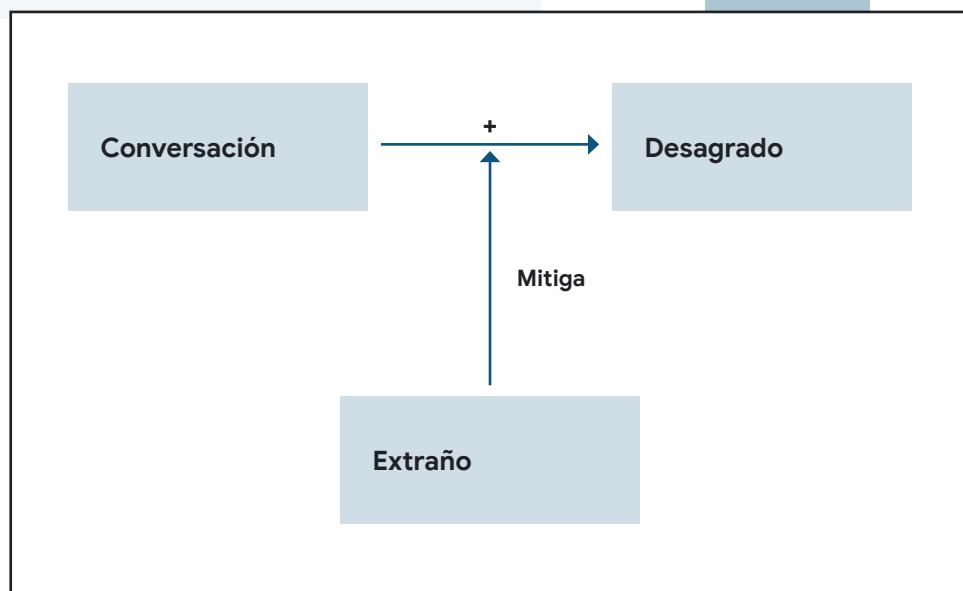
## Ejemplo de moderación

La moderación es un concepto difícil de entender en estadística, pero, en el mundo real, equivale a decir "depende". Analicemos un ejemplo rápido para aclarar el concepto de moderación en el contexto de este informe.

En la tercera temporada de *Curb Your Enthusiasm*, Larry David dice: "No me gusta hablar con personas que conozco, pero no tengo problema para hablar con extraños". Esto nos ofrece un diagrama rápido para analizar:

**"No me gusta hablar con personas que conozco, pero no tengo problema para hablar con extraños".**

Larry David



En este diagrama, se muestra que, para Larry David, la conversación tiene un impacto positivo en el desagrado. Aquí, positivo significa tan solo un aumento, no que sea necesariamente bueno. Esto se demuestra mediante la línea negra sólida entre "conversación" y "desagrado" con una flecha que apunta a "desagrado". La flecha sugiere que creemos que el flujo causal va de "conversación" a "desagrado". De lo que podemos observar, las conversaciones le desagradan a Larry.

El segundo aspecto a destacar es que "extraño" (aquí para representar el extraño booleano sí / no) no apunta a otra variable. En su lugar, apunta a un efecto, una flecha. Eso significa que pensamos que "extraño" no modifica una variable, sino un efecto. Por eso,

dibujamos la flecha de "extraño" a otra flecha, no a otra variable. Lo que decimos es que lo que impacta el efecto de "conversación" en "desagrado" es si Larry habla o no con un extraño. En otras palabras, el efecto de "conversación" en "desagrado" depende de si la persona con la que conversa Larry es un extraño o no. Cuando la persona es un extraño, Larry "no tiene ningún problema" con el efecto de la conversación. Podríamos decir que los extraños mitigan el desagrado que siente Larry mientras conversa.

Existen diferentes maneras en las que algo podría moderar otra cosa:

- **Amplifica**: hace que los efectos positivos sean más positivos y que los negativos sean más negativos.
- **Atenúa**: debilita el efecto.
- **Mitiga**: hace que los efectos positivos sean menos positivos y que los negativos sean menos negativos.
- **Revierte**: hace que los efectos positivos se vuelvan negativos y que los negativos se vuelvan positivos.
- **Modifica**: a veces el efecto cambia, pero el patrón no puede plasmarse de manera adecuada en una sola palabra. Eso sucede frecuentemente con las variables categóricas como causas. Por ejemplo, algunos sectores podrían comportarse de manera diferente en condiciones distintas, como fluctuaciones económicas.

#### Ejemplo de mediación

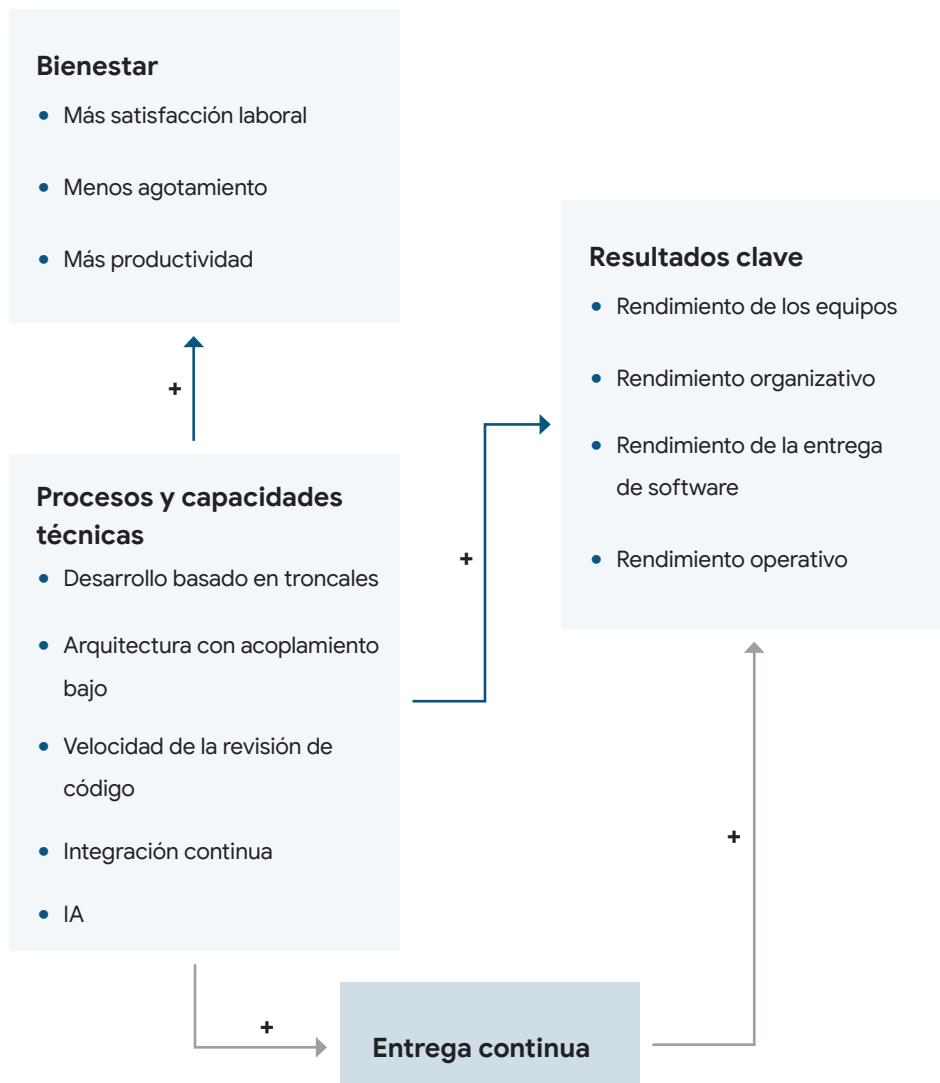
Al igual que con la moderación, pensamos mucho en la mediación. En la raíz de la mediación, se encuentran *por qué* o *cómo*. ¿Por qué el sol hace que una planta crezca más? ¿Por qué, cuando como fresas, percibo un sabor dulce? ¿Cómo es que presionar el pedal del acelerador hace que mi automóvil se mueva? ¿Cómo reducen el malestar los analgésicos? Podemos probar la mediación en estadística,<sup>1</sup> lo que significa que podemos decir: "Parece que el efecto de *X* en *Y* se explica de manera total o parcial con *M*". Por ejemplo, el efecto del sol en el tamaño de una planta se explica con la fotosíntesis.



<sup>1</sup><https://lavaan.ugent.be/tutorial/mediation.html>

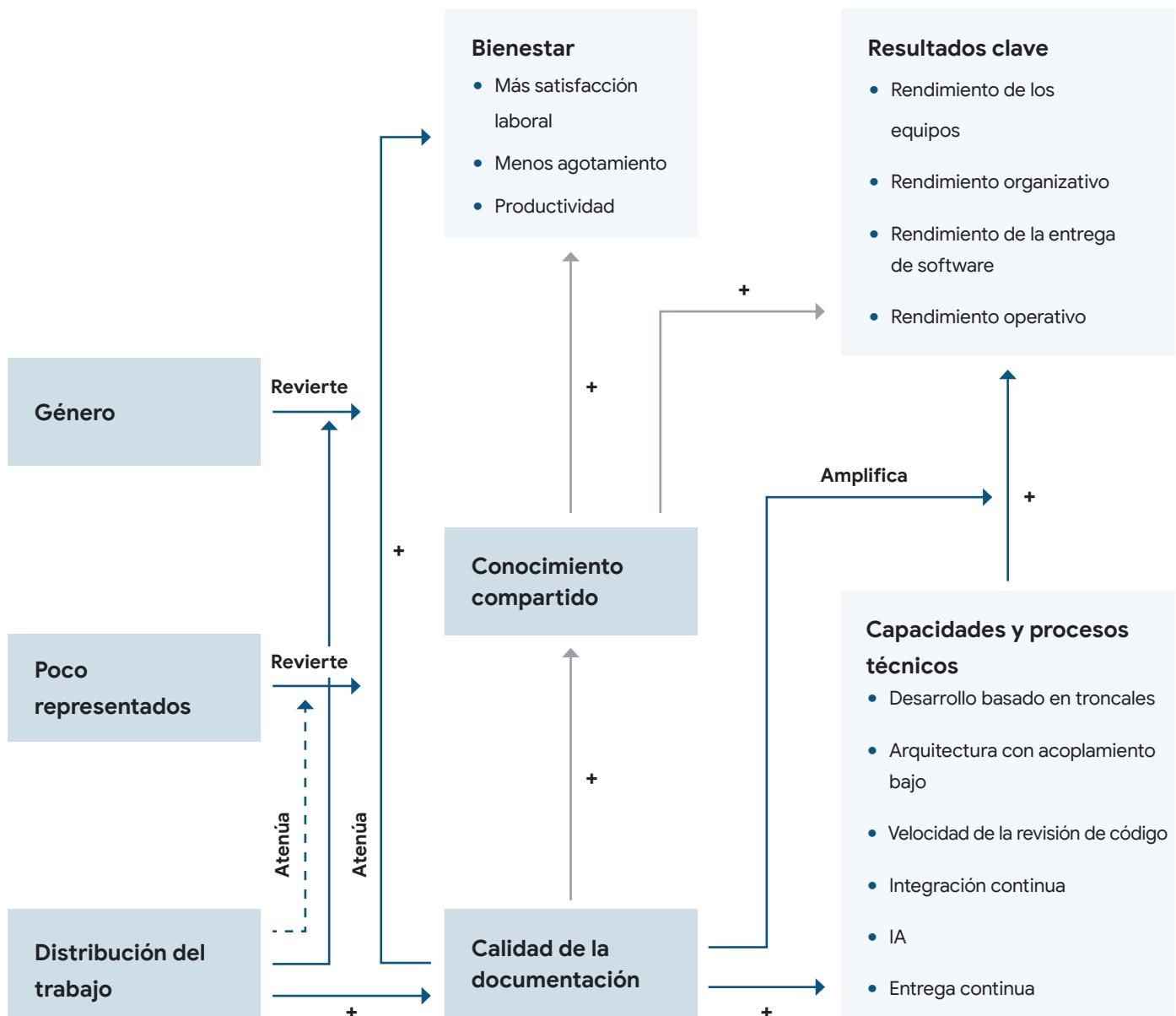
# Modelo del capítulo 3 Las capacidades técnicas predicen el rendimiento

Los procesos y las capacidades técnicas tienen un impacto positivo en el bienestar y los resultados clave. El efecto de las capacidades técnicas en los resultados clave está mediado parcialmente por la entrega continua, de forma tal que esas capacidades crean un entorno de entrega continua que tiene un impacto posterior en esos resultados clave.



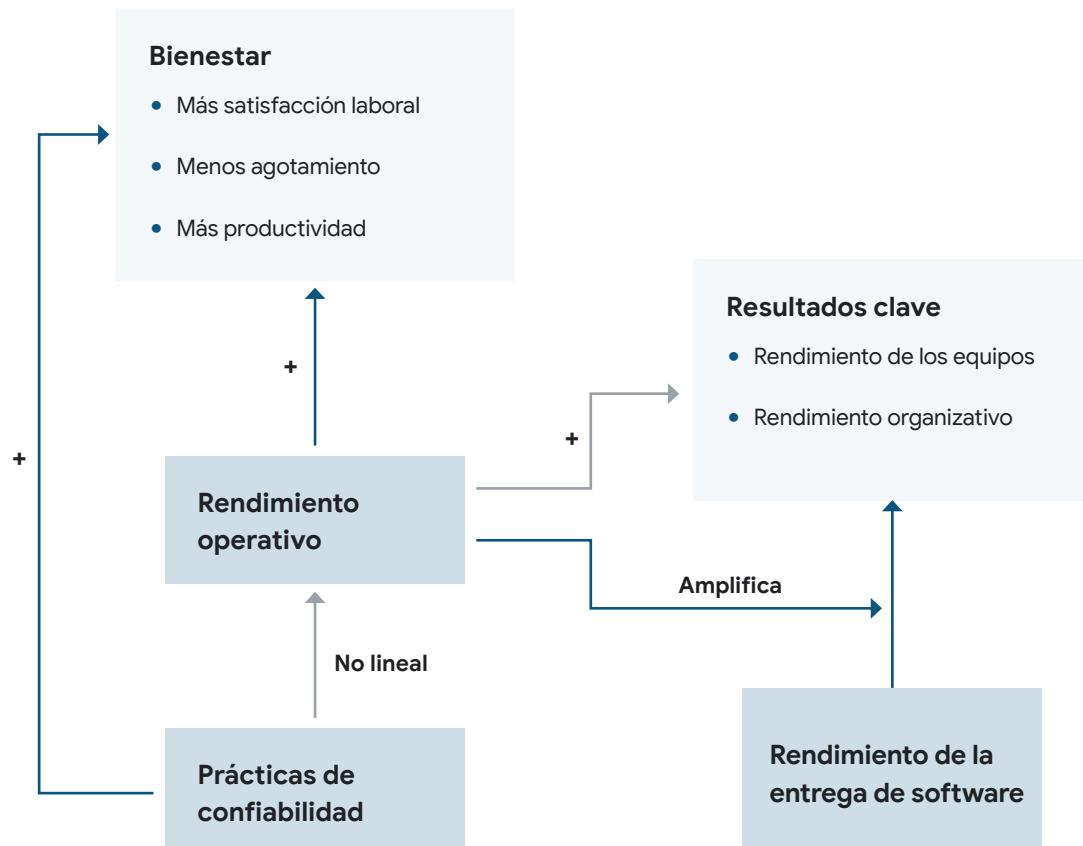
# Modelo del capítulo 4 La documentación es la base

En resumen, la documentación de alta calidad tiene efectos positivos en el bienestar y los resultados clave. Esto es verdad en parte porque se crea un entorno en el que se puede compartir conocimiento. La documentación de alta calidad también ayuda a los equipos a establecer procesos y capacidades técnicas; además, ayuda a que esos procesos y capacidades técnicas tengan un impacto aún mayor en los resultados clave. Por último, la calidad de la documentación no genera un mejor bienestar para todos. Recomendamos leer la sección para conocer el desglose de este hallazgo complejo.



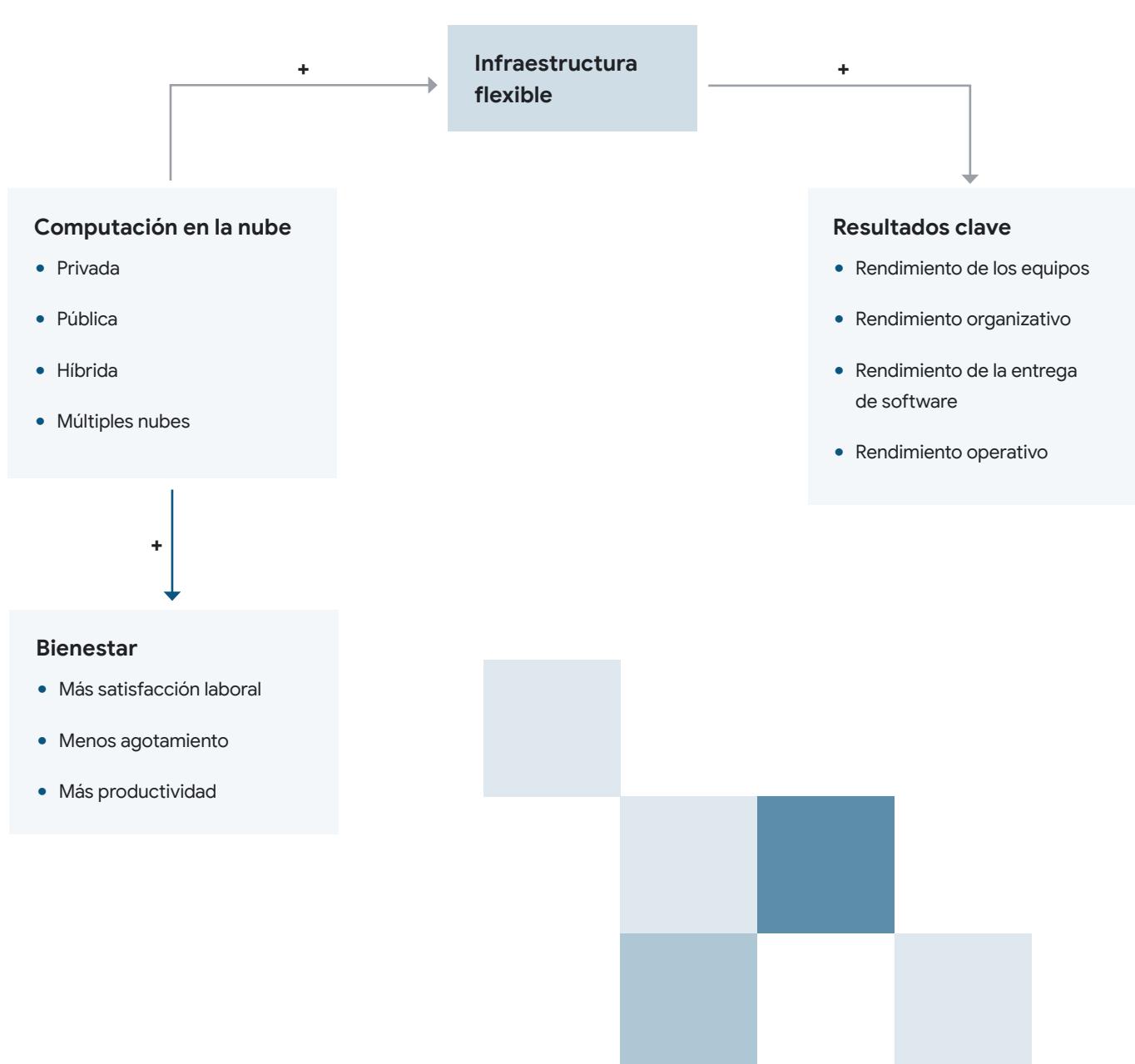
# Modelo del capítulo 5 La confiabilidad desbloquea el rendimiento

Aquí exploramos el rol central del rendimiento operativo. No solo mejora el bienestar y los resultados clave, sino que también amplifica el efecto del rendimiento de la entrega de software. Las prácticas de confiabilidad tienen una relación no lineal con el rendimiento operativo. Recomendamos consultar este capítulo para comprender esos y otros detalles.



# Modelo del capítulo 6 La infraestructura flexible es clave para el éxito

La computación en la nube genera un impacto en los resultados clave porque proporciona una infraestructura más flexible. La computación en la nube también mejora el bienestar.



# Modelo del capítulo 7 Nada de esto funciona sin inversión en cultura

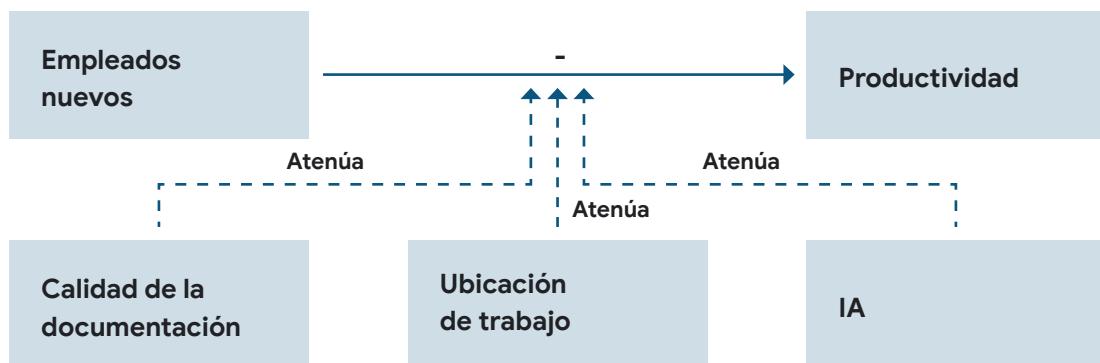
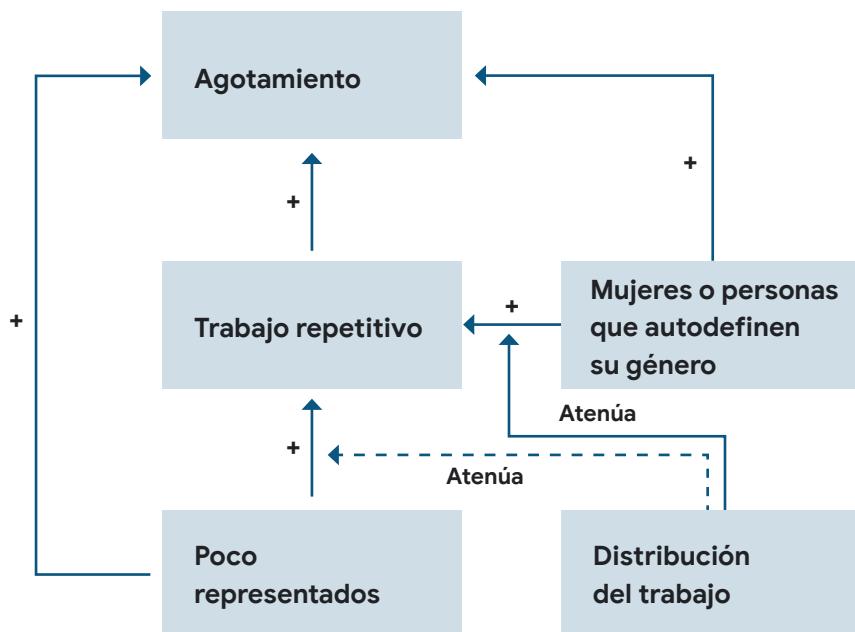
Podemos observar que la cultura está en el centro de muchas cosas en este diagrama.

Descubrimos que la cultura tiene una relación positiva con las capacidades técnicas, los resultados clave y el bienestar.



# Modelos del capítulo 8 Cómo, cuándo y por qué importa quién eres

En esta sección, hay dos modelos. En uno se explora por qué y cuándo las personas que se identifican como poco representadas y las que no se identifican como hombres tienen niveles de agotamiento más altos. En el otro modelo, se explora si la calidad de la documentación, la ubicación de trabajo o la IA pueden ayudar a los empleados nuevos a ser más productivos.



# Lecturas adicionales

Únete a DORA Community para debatir, aprender y colaborar en torno a la mejora de la entrega de software y el rendimiento operativo. [DORA.community](https://dora.community)

Realiza la revisión rápida sobre DevOps de DORA. <https://dora.dev/quickcheck>

Explora las capacidades técnicas, de procesos y culturales que impulsan un mayor rendimiento organizativo y de entrega de software. <https://dora.dev/devops-capabilities/>

Encuentra recursos sobre SRE <https://sre.google https://goo.gle/enterprise-roadmap-sre>

Lee el libro: *Acelera: La ciencia del desarrollo Lean y DevOps: Construcción y escalado de organizaciones tecnológicas de alto rendimiento IT Revolution*. <https://itrevolution.com/product/accelerate/>

Descubre una constelación adecuada de métricas para tu equipo usando SPACE Framework. "The SPACE of Developer Productivity: There's more to it than you think". <https://queue.acm.org/detail.cfm?id=3454124>

Existen varios estudios de investigación sobre revisiones de código modernas. A continuación, se incluyen algunos informes para explorar:

- "Expectations, Outcomes, and Challenges of Modern Code Review". <https://dl.acm.org/doi/10.5555/2486788.2486882>

- "Code Reviews - From bottlenecks to Superpowers". <https://learning.acm.org/techtalks/codereviews>
- "Modern Code Review- A Case Study at Google". <https://dl.acm.org/doi/10.1145/3183519.3183525>
- "The Choice of Code Review Process: A Survey on the State of the Practice". [https://link.springer.com/chapter/10.1007/978-3-319-69926-4\\_9](https://link.springer.com/chapter/10.1007/978-3-319-69926-4_9)
- "Investigating the effectiveness of peer code review in distributed software development based on objective and subjective data". <https://jserd.springeropen.com/articles/10.1186/s40411-018-0058-0>

Lee el libro: *The No Club: Putting a Stop to Women's Dead-End Work*. Simon & Schuster. <https://www.simonandschuster.com/books/The-No-Club/Linda-Babcock/9781982152338>

Publicaciones del programa de investigación de DORA, incluidos informes anteriores de Accelerate State of DevOps. <https://dora.dev/publications/>

Preguntas frecuentes sobre el estudio y los informes. <http://dora.dev/faq>

Erratas: Lee el informe y envía cambios, correcciones y aclaraciones a <https://dora.dev/publications/errata>

# Apéndice

## Una mejor definición de cómo medimos el rendimiento de la entrega de software

Este año, hicimos cambios en la forma en que evaluamos los errores de cambio y la recuperación de una falla.

Cambiamos la forma en que los encuestados informaron su **tasa de errores de cambio**. En años anteriores, los encuestados tenían seis opciones (0-15%, 16-30%, etcétera). Este año les brindamos a los encuestados un control deslizante para que pudieran seleccionar cualquier valor entre 0% y 100%.

Hicimos este cambio por dos razones:

- La tasa de errores de cambio siempre tuvo un comportamiento un poco diferente que las otras tres mediciones del rendimiento de la entrega de software.<sup>1</sup> Tenemos la teoría de que esto podría deberse en parte al tamaño de los buckets. Proporcionar más precisión en la respuesta podría generar un mejor rendimiento estadístico. Pudimos validar esta teoría.
- Planteamos la hipótesis de que los equipos tienen una mejor comprensión de su tasa de errores de cambio en la actualidad que cuando comenzó la investigación hace casi una década. No tenemos forma de validar esta hipótesis.

Hicimos preguntas sobre la recuperación ante fallas de dos formas diferentes este año:

- Antes, hacíamos esta pregunta: "En la aplicación o el servicio principal en el que trabajas, ¿cuánto tiempo se necesita por lo general para restablecer el servicio cuando ocurre un incidente o un defecto que afecta a los usuarios (por ejemplo, una interrupción no planificada o una inhabilitación del servicio)?".
- Este año, agregamos los siguientes calificadores a la pregunta (las diferencias están en negrita aquí, pero no fue así en la encuesta): "En la aplicación o el servicio principal en el que trabajas, ¿cuánto tiempo se necesita por lo general para restablecer el servicio **después de que un cambio en producción o en el lanzamiento a los usuarios provoca un servicio degradado (por ejemplo, provoca una interrupción en el servicio o lo inhabilita) y, en consecuencia, necesita una corrección (por ejemplo, se requiere aplicar un parche rápido, realizar reversión, implementar una corrección o aplicar un parche)**?".

<sup>1</sup> Forsgren, N., Humble, J., y Kim, G. Acelera: *La ciencia del desarrollo Lean y DevOps: Construcción y escalado de organizaciones tecnológicas de alto rendimiento*. IT Revolution Press, 2018, 37-38.

La forma anterior de hacer preguntas sobre tiempos de recuperación no permitía una distinción entre una falla iniciada por un cambio de software y una falla iniciada por algo como un terremoto que interrumpe un servicio en un centro de datos. Teníamos la hipótesis de que una redacción más precisa nos permitiría comparar tipos de falla similares entre sí y que los términos estarían más alineados estadísticamente con las otras tres mediciones de rendimiento de entrega de software.

Ahora usamos el término "tiempo de recuperación para implementaciones con errores" para distinguir nuestra medición de la frase más genérica "tiempo de restablecimiento" que usamos en el pasado y, a veces, abreviamos como "MTTR". MTTR generó algo de confusión en la comunidad: ¿la "M" es de media o mediana? Además, los profesionales que querían saber más sobre fallas, como los del campo de resiliencia de la ingeniería, están dejando atrás MTTR como una medición confiable para orientar el aprendizaje y la mejora.<sup>2</sup>

La pregunta que se agregó recientemente y una métrica nueva, tiempo de recuperación para implementaciones con errores, están más alineadas con el espíritu de medir el rendimiento de la entrega de software.



<sup>2</sup> "Moving Past Simple Incident Metrics: Courtney Nash on the VOID". <https://www.infoq.com/articles/incident-metrics-void/>

# Los cálculos detrás de las comparaciones

A lo largo de este informe, hay estadísticas que indican que tener una medición más alta para una variable determinada lleva a puntuaciones N veces mayores en otra característica. ¿Qué es esto? ¿A qué se debe? Esta es nuestra receta para crear estas fórmulas:

1. Crea una fórmula a partir del modelo evaluado usando técnicas de regresión que toman en consideración posibles rutas de sesgo:

$$\text{Felicidad} \sim 5.64 + 0.19 * \text{luz solar} + 0.14 * \text{temperatura}$$

2. Para la variable de interés, como la luz solar, busca valores altos y bajos para comparar:

media = 6.3; desviación estándar = 1.4

Luz solar alta = 1 desviación estándar por sobre la media = media + desviación estándar =  $6.3 + 1.4 = 7.7$

Luz solar baja = 1 desviación estándar debajo de la media = media - desviación estándar =  $6.3 - 1.4 = 4.9$

3. Calcula la media para las covariables. En este ejemplo, eso sería la temperatura, que es igual a 0.24 (temperatura estandarizada).
4. Completa la fórmula del paso uno para luz solar alta y luz solar baja. Notarás que solo un número es diferente en la fórmula. Así es como mantenemos todo el resto igual y aislamos solo la única diferencia que nos interesa.

Felicidad predicha de la luz solar alta =

$$5.64 + 0.19 * 7.7 + 0.14 * 0.24 = 7.1$$

Felicidad predicha de la luz solar baja =

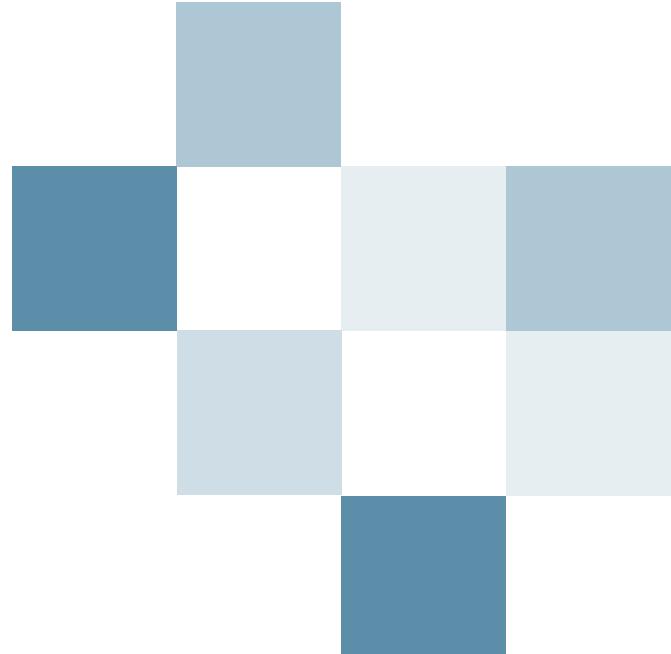
$$5.64 + 0.19 * 4.9 + 0.14 * 0.24 = 6.6$$

5. Calcula la proporción:

$$\frac{\text{felicidad predicha}}{\text{de la luz solar alta}} = \frac{7.1}{6.6} = 1.1 \text{ veces}$$

$\text{felicidad predicha}$   
 $\text{de la luz solar alta}$   
\_\_\_\_\_  
 $\text{felicidad predicha}$   
 $\text{de la luz solar baja}$

6. La proporción sugiere que los niveles altos de luz solar llevan a niveles de felicidad un 10% más altos en comparación con los niveles bajos.



## ¿Qué es una simulación?

No significa que inventamos los datos. Usamos estadísticas bayesianas para calcular una **posterior**, que intenta capturar "la frecuencia esperada en la que aparecerán diferentes valores del parámetro".<sup>3</sup> La "simulación" toma de esa posterior más de 1,000 veces para explorar los valores que son más creíbles para un parámetro (media, ponderación beta, sigma, intersección, etc.) dados nuestros datos. "Imagina que la posterior es un bucket lleno de valores del parámetro, cifras como 0.1, 0.7, 0.5, 1, etc. Dentro del bucket, cada valor existe en proporción a su probabilidad posterior, tal que los valores cercanos al pico son mucho más comunes que aquellos en las colas".<sup>4</sup>

Esto equivale a nuestro uso de simulaciones para explorar interpretaciones posibles de los datos y tener una idea de cuánta incertidumbre hay.

Puedes considerar cada simulación como una pequeña IA que conoce solo nuestros datos y un par de reglas que intentan completar un espacio en blanco (parámetro) con una suposición fundamentada. Haces eso 4,000 veces y obtienes 4,000 suposiciones de pequeñas IA para un parámetro determinado. Puedes aprender mucho de esas suposiciones. Puedes conocer la suposición promedio, entre qué valores cae el 89% de estas, cuántas están arriba de un nivel determinado, cuánta variación hay, etc. Incluso puedes divertirte combinando suposiciones (simulaciones) en muchos modelos.

Cuando mostramos un gráfico con varias líneas o una distribución de valores potenciales, intentamos mostrarte lo que es más razonable dados nuestros datos y cuánta incertidumbre hay.

<sup>3</sup> McElreath, Richard. Statistical rethinking: A Bayesian course with examples in R and Stan. Chapman and Hall/CRC, 2018, p. 50.

<sup>4</sup> McElreath, Richard. Statistical rethinking: A Bayesian course with examples in R and Stan. Chapman and Hall/CRC, 2018, p. 52.

